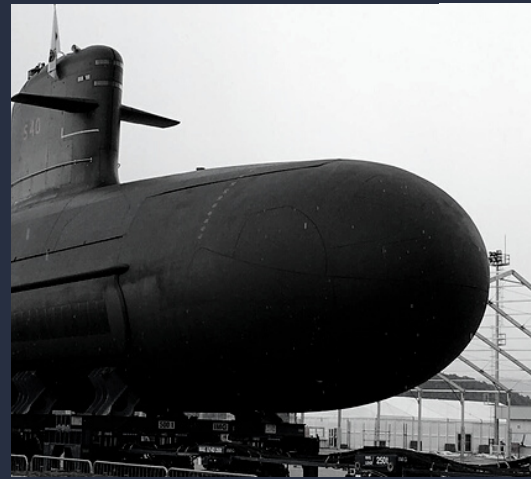


VOLUME III



UNCERTAINTY MODELING: FUNDAMENTAL CONCEPTS AND MODELS

EDITORS: ARIOSTO B. JORGE, CARLA T.M. ANFLOR,
GUILHERME F. GOMES, and SERGIO H.S. CARNEIRO





University of Brasilia (UnB)
Post-Graduate Program - Integrity of Engineering Materials

**Book Series in Discrete Models, Inverse Methods, &
Uncertainty Modeling in Structural Integrity**

Volume III

**Uncertainty Modeling:
Fundamental Concepts and Models**

Book Series Editors

Ariosto Bretanha Jorge (leading editor)

Carla Tatiana Mota Anflor

Guilherme Ferreira Gomes

Sergio Henrique da Silva Carneiro

This Book Series is an initiative of the Post-Graduate Program - Integrity of Engineering Materials (PPG-Integridade) of University of Brasilia (UnB)

With the incentive, encouragement and support from

- Latin American Journal of Solids and Structures (LAJSS) (www.lajss.org)
- Brazilian Society of Mechanical Sciences and Engineering (ABCM) (www.abcm.org.br)
- Brazilian Association of Computational Methods in Engineering (ABMEC) (www.abmec.org.br).

Brasília – DF

Open access digital book published by
University of Brasilia Central Library (BCE/UnB)

2022

Cover page pictures, clockwise from top left: a SAAB Gripen (Brazilian version, monoplace); an EMBRAER's KC-390 transport aircraft, a Brazilian-manufactured Riachuelo-class submarine, a nuclear powerplant at Angra dos Reis, Brazil; an offshore platform operated by PETROBRAS; and a generator in a wind power plant in the northeast region of Brazil.

Several engineering fields of application, such as aerospace, naval, nuclear, energy, mechanical, civil, etc, may take advantage of the technology, innovation and research in the areas discussed in this book.

All pictures in the cover page are public domain.

© 2022



[Creative commons license]

Jorge, Ariosto B.; Anflor, Carla T. M.; Gomes, Guilherme F.; Carneiro, Sergio H. S. (Eds.): Uncertainty Modeling: Fundamental Concepts and Models [1st Edition], UnB, Brasília, DF, Brazil, 2022. Volume III in Discrete Models, Inverse Methods, & Uncertainty Modeling in Structural Integrity.

ISBN 978-65-86503-88-3

DOI: <https://doi.org/10.4322/978-65-86503-88-3>

Edited as an initiative of: Universidade de Brasília
Faculdade do Gama (FGA) / Faculdade de Tecnologia (FT)
Programa de Pós-Graduação em Integridade de Materiais da Engenharia

Address:

Universidade de Brasília, Campus do Gama
Área Especial de Indústria e Projeção – Setor Leste
Prédio UED (Unidade de Ensino e Docência) - 1º Andar
CEP: 72.444-240, Gama - DF, Brasil

Contact: Phone: (+55) (61) 3107-8903 WebSite: <http://www.pgintegridade.unb.br/>
E-mail pgintegr@unb.br; ctanflor@gmail.com; ariosto.b.jorge@gmail.com

Cataloging-in-Publication (CIP) Data:

Prepared by the librarian at University of Brasilia Central Library (BCE/UnB).
Heloiza Faustino dos Santos - CRB 1/1913

U54 Uncertainty modeling [recurso eletrônico] : fundamental concepts and models / editors, Ariosto B. Jorge ... [et al.]. – Brasília : Universidade de Brasília, 2022.
743 p. : il. – (Book series in discrete models, inverse methods, & uncertainty modeling in structural integrity ; v. 3)

Inclui bibliografia.

Modo de acesso: World Wide Web:

<www.pgintegridade.unb.br>.

ISBN 978-65-86503-88-3.

1. Mecânica computacional. 2. Integridade de materiais. 3. Modelagem de incertezas. I. Jorge, Ariosto B. (ed.). II. Série.

CDU 620



Foreword from FGA / UnB



The engineering sector drives and enables the development of a country. The formation of an engineer allows a technical capacity to evaluate, plan, design, suggest and apply all possible techniques in search of the best construction of a technological equipment. Currently, the engineer must be more and more prepared to solve existing problems in various sectors of society. It is through it that societies grow in search of progress.

The recognition of engineering and the training of new professionals increases every year in Brazil. In the 2000s, the University of Brasília (UnB) went through an expansion process, resulting in the implementation of the new UnB Engineering Campus in the city of Gama (UnB-Gama, FGA). Five new undergraduate courses were created: Aerospace Engineering, Automotive Engineering, Electronic Engineering, Energy Engineering and Software Engineering. The UnB Gama Campus project converges to increase the education level of the Brazilian population, especially in the five areas of engineering activity, all in line with current national public policies, aimed at expanding the population's access to quality higher education in the country.

Following the high quality teaching line, the UnB-Gama campus has the Graduate Program in Integrity of Engineering Materials (PPG-Integridade). The program has the following lines of research: Dynamics and Vibrations, Fatigue, Structural Materials, Biomaterials, Structure Fluid Interaction and Numerical Simulation of the Mechanical Behavior of Materials. This book series is an initiative of PPG-Integridade - UnB, organized as a collaborative work involving researchers, engineers, scholars, from several institutions, universities, industry, recognized both nationally and internationally.

Beside the high technical quality and relevance of the topics covered in the books, this series will enable an essential internationalization of the research currently developed within the University of Brasília. Several authors from different countries also contributed to these books, enabling greater interaction between national and international research groups. This internationalization raises the level of academic education for new professionals in the field of engineering, in addition to more advanced scientific research and technological development.

Additionally, this book series features a strong contribution from the industrial sector. Several professionals from different companies collaborated with the writing of some chapters in the three volumes that make up this series. These initiatives are of great strategic importance, as they allow the grouping of different technical capabilities. On the part of companies in the sector, with knowledge of market demands, and on the part of universities, by adding the technical-scientific knowledge of their team of researchers to the improvement of innovative products and services.

This book should be appreciated by anyone in need of knowledge of Materials Integrity. The completeness of Discrete Modeling and Inverse Methods theory combined with the Uncertainty Modeling in Structural Integrity makes these books mandatory for everybody aiming at Direct and Inverse Problems, including model-based and signal-based inverse problems.

Prof. Dr. Sandro A.P. Haddad, Director
UnB-Gama campus (<https://fga.unb.br/>)

Foreword from LAJSS

Book Series in Discrete Models, Inverse Methods, & Uncertainty Modeling in Structural Integrity

Book Series editors: Ariosto B. Jorge, Carla T.M. Anflor, Guilherme F. Gomes, and Sergio H.S. Carneiro

This book series represents a commendable effort in compiling the latest developments on three important Engineering subjects: discrete modeling, inverse methods, and uncertainty structural integrity. Although academic publications on these subjects are plenty, this book series may be the first time that these modern topics are compiled together, grouped in volumes, and made available for the community.

The application of numerical or analytical techniques to model complex Engineering problems, fed by experimental data, usually translated in the form of stochastic information collected from the problem in hand, is much closer to real-world situations than the conventional solution of PDEs. Moreover, inverse problems are becoming almost as common as direct problems, given the need in the industry to maintain current processes working efficiently, as well as to create new solutions based on the immense amount of information available digitally these days. On top of all this, deterministic analysis is slowly giving space to statistically driven structural analysis, delivering upper and lower bound solutions which help immensely the analyst in the decision-making process.

All these trends have been topics of investigation for decades, and in recent years the application of these methods in the industry proves that they have achieved the necessary maturity to be definitely incorporated into the roster of modern Engineering tools. The present book series fulfills its role by collecting and organizing these topics, found otherwise scattered in the literature and not always accessible to industry.

Moreover, many of the chapters compiled in these books present ongoing research topics conducted by capable fellows from academia and research institutes. They contain novel contributions to several investigation fields and constitute therefore a useful source of bibliographical reference and results repository.

The Latin American Journal of Solids and Structures (LAJSS) is honored in supporting the publication of this book series, for it contributes academically and carries technologically significant content in the field of structural mechanics.

On behalf of LAJSS,

Prof. Dr. Marcílio Alves (USP), Editor-in-Chief
Prof. Dr. Rogério J. Marczak (UFRGS), Assoc. Editor
Prof. Dr. Pablo A. Muñoz-Rojas (UDESC), Assoc. Editor
Prof. Dr. Marco L. Bittencourt (Unicamp), Assoc. Editor



Latin American Journal of Solids and Structures (LAJSS)

(www.lajss.org)

Foreword from ABCM

The Brazilian Society of Mechanical Sciences and Engineering – ABCM welcomes enthusiastically the publication of the Book Series in Models, Inverse Methods & Uncertainty Modeling in Structural Integrity.

The initiative, undertaken by Prof. Ariosto B. Jorge, Dr. Carla T.M. Anflor, Dr. Guilherme F. Gomes and Dr. Sergio H. S. Carneiro, with the support of the University of Brasília, is received by the scientific community as a valuable contribution to the dissemination of knowledge encompassing the large number of topics covered in the three volumes of the series.

These topics have been judiciously selected to encompass comprehensively the theoretical aspects, modeling techniques and numerical methods related to Structural Integrity, and are presented in a large collection of chapters authored by renowned experts, from both academia and industry. We gladly realize that many members of ABCM have contributed as authors.

Besides the comprehensive and well-articulated content, one distinguishing characteristic of this book series is that it has been conceived to serve both for educational purposes at graduate level and as an information source for researchers and engineering practitioners, which amplifies, to a large extent, its utility. Another relevant feature is that the material is intended to be available to the public in electronic format at no cost, which highlights the generosity of the authors and editors and their commitment to the most fundamental academic principles.

On behalf of the scientific community of the field of Mechanical Sciences and Engineering, ABCM acknowledges the editors and authors of the present book series for their contribution to the progress of Engineering research and education.

Prof. Dr. Domingos Alves Rade
President of ABCM

On behalf of



Brazilian Society of Mechanical Sciences and
Engineering (ABCM)
(www.abcm.org.br)

Foreword from ABMEC

The whole range of topics related to Direct & Inverse Problems and Modeling of Uncertainties is substantially associated with the needs of the mechanical, civil, aeronautical/aerospace, nuclear, and naval/oceanic industries. Indeed, they play a core role in industrial renewal, contributing to productivity and competitiveness. Especially taking Brazil into account, this book series, conceived as a comprehensive one that covers these important topics, is very welcome.

These themes are also among the main interests of the Brazilian Association of Computational Methods in Engineering, ABMEC. ABMEC is concerned with the application of numerical methods and digital computers to the solution of engineering problems. Its mission is to promote, foster, and organize activities encompassing the development and use of such computational methods in Brazil.

We are fortunate to have the opportunity to support this book series as a collaborative work that intends to involve scholars from different institutions and researchers from industry, with national and international relevance. We sincerely believe that this work will provide a common forum for discussion, education, and research information transfer between the several subjects concerning computational methods in engineering.

Our congratulations to the editors, professors Ariosto Bretanha Jorge, Carla Tatiana Mota Anflor, Sergio Henrique da Silva Carneiro, Guilherme Ferreira Gomes for this important contribution to the Brazilian engineering.

Prof. Dr. Felício Bruzzi Barros
President of ABMEC

On behalf of



abmec
Associação Brasileira de Métodos
Computacionais em Engenharia

Brazilian Association of Computational
Methods in Engineering (ABMEC)
(www.abmec.org.br)

Acknowledgements

Acknowledgements from the Book Series editors

This book series is an initiative of the Graduate Program in Integrity of Engineering Materials (PPG-Integridade) at the University of Brasilia (UnB), Brazil (www.pgintegridade.unb.br).

The editors would like to thank PPG-Integridade and UnB for the initiative, incentive and support for this Book Series project.

The book series is organized as a collaborative work involving researchers, engineers, scholars, engaged in research, development and applications in the related areas, affiliated to several institutions, universities, industry, and recognized both nationally and internationally.

The editors are grateful and would like to show their appreciation to all the co-authors of the book chapters, for their participation, dedication, and support.

The book series is published as a digital version, with ISBN provided by UnB, and DOI for each chapter, provided by the Latin American Journal of Solids and Structures (LAJSS) (www.lajss.org).

The scope of the Book series is in the broad areas of interest of LAJSS, and also of the Brazilian Society of Mechanical Sciences and Engineering (ABCM) (www.abcm.org.br) and the Brazilian Association of Computational Methods in Engineering (ABMEC) (www.abmec.org.br).

For increased visibility, these three institutions are encouraging the divulgation of the Book Series project in their websites.

The editors would like to express their appreciation to LAJSS, ABCM and ABMEC, for their incentive, encouragement and support for this Book Series project.

Ariosto, Carla, Guilherme, Sergio

Brasilia, September 15th, 2022.

Personal dedications from the Book Series editors

A personal dedication from Ariosto Bretanha Jorge, leading editor

In my point of view, this Book Series project represents the culmination of a dedicated academic career, in many aspects intrinsically related to the different research topics and areas covered along the three Volumes of the Book Series.

I would like to thank all my co-editors of the Book Series, all the co-authors of the book chapters, and also all the researchers, scholars, students with whom I had the opportunity to share collaborative work throughout my many years along this academic career. I've enjoyed learning a lot from you all!

To you, my sincere thank you!

I would like also to dedicate this project to my wife Daisy, for her love, understanding, and unconditional support, throughout my entire academic career, and to my daughter Elisa and my son Luís Paulo, for their love and support.

To you, my true love and deepest appreciation!

Ariosto

A personal dedication from Carla Tatiana Mota Anflor

I would like to thank my parents, Maria Emilia and Miro (in memoriam), for their love and support during this infinite academic journey.

Special thanks to Professor Ariosto B. Jorge for inviting me to join the Book Series editors team and for being a great advisor until nowadays. Also, sincere thanks to my colleagues Guilherme Gomes and Sergio Carneiro, for their friendship during our editorial working times.

Carla

A personal dedication from Guilherme Ferreira Gomes

To my beloved father Expedito Ferreira Gomes (*in memoriam*).

"Those we love never die, they just leave before us."

Guilherme

A personal dedication from Sergio Henrique da Silva Carneiro

To my wife Lúcia and my children Pedro, Letícia and Matheus, for their love and support.

To my parents Neli and Joaquim (in memoriam) and my brother Luiz, my true roots, my foundation.

To all my students, past and present, graduate and undergraduate, from whom I have learned so much.

To my colleagues in Academia and research institutes, for all the knowledge and wisdom they have shared with me throughout my career.

To the Brazilian Air Force, for the privilege of serving my Country in the company of dedicated and honored men and women.

Sergio

Brasilia, September 15th, 2022.

Table of Contents

Forewords (FGA/UnB, LAJSS, ABCM, ABMEC); Acknowledgements (Editors)	iii
Chapter 1 Introduction to Probabilistic Methods & Uncertainty Modeling in Direct & Inverse Problems	1
Chapter 2 Application of Data Science in the Sustainable Energy Industry	19
Chapter 3 Overview of Classic Surrogates and Metamodeling Strategies in Structural Analysis	54
Chapter 4 Uncertainty Quantification in Inverse Kinematics Computation of Space Robots by Neural Networks	89
Chapter 5 Metamodel-Based Simulation Optimization: A Systematic Literature Review	119
Chapter 6 Principal Component Analysis: An Overview and Applications in Multivariate Engineering Problems	172
Chapter 7 Gaussian Mixture Models and Their Applications to Fault Diagnosis in Electric Motors	195
Chapter 8 Sensor Placement Optimization of a Helicopter Main Rotor Blade	221
Chapter 9 Challenges for Structural Health Monitoring: Nonlinearities and Uncertainties	258
Chapter 10 Applications of Numerical and Statistical Methods for Monitoring and Control Engineering Variables in a Concrete Dam	279

Chapter 11	
Application of Statistical Processing Techniques to the Impedance-based SHM for the Oil & Gas Industry	337
Chapter 12	
Neuro-Fuzzy Data Normalization Applied for Impedance-based SHM to the Oil & Gas Industry	384
Chapter 13	
On Kriging Techniques & Impedance-based SHM as Applied to Damage Detection in 2D Structures	427
Chapter 14	
Time Series Models and Expansion Optimal Linear Estimation for Time-Dependent Reliability Analysis	459
Chapter 15	
Bayesian Optimal Experimental Design	479
Chapter 16	
Airplane Optimal Robust Design Combined with an Entropy-Statistics Unsupervised Learning Classification Algorithm	502
Chapter 17	
Random Fatigue Under Uniaxial and Multiaxial Loading Conditions	552
Chapter 18	
Random Field Discretization in Stochastic Finite Elements: Element Size Effects and Error Indicators in Reliability Structural Analysis	613
Chapter 19	
On Error Estimators and Bayesian Approaches in Computational Model Validation	640
Book Series: Editors - Bio & Back Cover	744

Chapter 1

Introduction to Probabilistic Methods & Uncertainty Modeling in Direct & Inverse Problems

Chapter details

Chapter DOI:

<https://doi.org/10.4322/978-65-86503-88-3.c01>

Chapter suggested citation / reference style:

Jorge, Ariosto B., et al. (2022). “Introduction to Probabilistic Methods & Uncertainty Modeling in Direct & Inverse Problems”. In Jorge, Ariosto B., et al. (Eds.) *Uncertainty Modeling: Fundamental Concepts and Models*, Vol. III, UnB, Brasilia, DF, Brazil, pp. 1–18. Book series in Discrete Models, Inverse Methods, & Uncertainty Modeling in Structural Integrity.

P.S.: DOI may be included at the end of citation, for completeness.

Book details

Book: Uncertainty Modeling: Fundamental Concepts and Models

Edited by: Jorge, Ariosto B., Anflor, Carla T. M., Gomes, Guilherme F., & Carneiro, Sergio H. S.

Volume III of Book Series in:

Discrete Models, Inverse Methods, & Uncertainty Modeling in Structural Integrity

Published by: UnB City: Brasilia, DF, Brazil Year: 2022

DOI: <https://doi.org/10.4322/978-65-86503-88-3>

Introduction to Probabilistic Methods & Uncertainty Modeling in Direct & Inverse Problems

Ariosto Bretanha Jorge^{1a*}, Carla Tatiana Mota Anflor^{1b},
Guilherme Ferreira Gomes² and Sérgio Henrique da Silva Carneiro^{1c}

¹Post-Graduate Program - Integrity of Engineering Materials, University of Brasilia, Brazil. Book series editors. E-mail: ariosto.b.jorge@gmail.com, anflor@unb.br, shscarneiro@unb.br

²Mechanical Engineering Institute, Federal University of Itajubá, Itajubá, Brazil. Book series editor. E-mail: guilhermefergom@unifei.edu.br

*Corresponding author. Book series leading editor.

Abstract

This chapter presents an overview of the Book Series in Direct Methods, Inverse Methods and Uncertainty Modeling, of which the first two volumes have already been published (Volume I: Model-Based and Signal-Based Inverse Methods, DOI: <https://doi.org/10.4322/978-65-86503-71-5>, and Volume II: Fundamental Concepts and Models for the Direct Problem, DOI: <https://doi.org/10.4322/978-65-86503-83-8>), and a presentation of the current Volume III: Uncertainty Modeling: Fundamental Concepts and Models. The chapter includes an introduction to the different topics in Optimization and Identification Techniques detailed in Volume I, an introduction to topics related to materials, technologies, and discrete modeling detailed in Volume II, and an introduction to topics related to uncertainty modeling and probabilistic methods, comprising the several chapters included in this Volume III of the Book Series.

1 Book Series in Discrete Models, Inverse Methods, & Uncertainty Modeling in Structural Integrity: overview

The Book series in “Discrete Models, Inverse Methods, & Uncertainty Modeling in Structural Integrity” is an initiative of the Post-Graduate Program - Integrity of Engineering Materials (PPG-Integridade) of University of Brasilia (UnB), organized as a collaborative work involving researchers, engineers, scholars, from several institutions, universities, industry, recognized both nationally and internationally.

This book series is an activity related to the Research, Development & Innovation (R,D&I) Project at UnB, titled: “Technological Demonstration Platform for Inverse Methods and Uncertainty Modeling Integrity of Structures and Components”, available at the UnB Central Library (in Portuguese) (Jorge, 2020).

The Book Series project is comprised by three Volumes:

- Volume I – Model-based and Signal-Based Inverse Methods (Jorge, Anflor, et al., 2022d)

- Volume II – Fundamental Concepts and Models for the Direct Problem (Jorge, Anflor, et al., 2022a)
 - Part I – Technologies and Materials Modeling
 - Part II - Discrete Modeling
- Volume III – Uncertainty Modeling: Fundamental Concepts and Models

The different book chapters were elaborated encompassing the relevant project topics, including chapters covering:

- Fundamentals, including topics such as: basic principles, concepts & foundations, for the Direct & Inverse Problems (including model-based and signal-based inverse methods), and for the Modeling of Uncertainties;
- Special Topics, Applications, and Research Review, including topics such as: research review, state-of-the-art & future trend topics, for the Direct & Inverse Problems (including model-based and signal-based inverse methods), and for the Modeling of Uncertainties.

The book chapters were prepared as a collaborative work by researchers, engineers, scholars, involved in research, development and applications in the related areas.

The research areas of interest throughout the book chapters include:

- Modeling of the inverse problem, monitoring & diagnosis / prognosis: models and methods for inverse problems, optimization methods (including techniques such as; multi-objective optimization, topology optimization, evolutionary optimization), Wavelets, Kalman Filter (KF), Particle Filter (PF), Machine Learning (ML), Artificial Intelligence (AI), Data Science (DS), for applications such as Structural Health Monitoring (SHM) (including impedance-based and Lamb Wave-based techniques), Health & Usage Monitoring Systems (HUMS);
- Modeling of the direct problem: mechanics of materials (including metallic materials, composites), structures (including civil, mechanical naval, aeronautical structures) machinery design and mechanical components, fracture mechanics, impact, fatigue, damage tolerance, integrity, mechanical vibrations, dynamics of structures, computational mechanics, including mathematical methods and numerical methods for discrete modeling for continuum mechanics (such as Finite Element Methods (FEM), Boundary Element Methods (BEM), Mesh-Free Methods (MFM));
- Probabilistic methods and modeling of uncertainties: probabilistic methods in engineering, Design of Experiments (DOE), Response Surface Methods (RSM), Risk & Reliability (including structural and system reliability), Uncertainty Modeling (UM) & Uncertainty Quantification (UQ), Bayesian Approaches (BA), stochastic Finite Element approaches (Stochastic FEM, Spectral FEM, Polynomial Chaos), stochastic optimization, meta-modeling (including techniques such as Surrogate Models (SM), Reduced Order Models (ROM)), model Verification & Validation (V&V).

The different models, methods and approaches presented throughout the several chapters in the three Volumes of this Book Series are intended as an introductory presentation of some possibilities of methods that could be used in problems related to integrity of structures and components, and maybe even extended to other engineering areas, as appropriate. The list of models is not unique, and is neither comprehensive nor

exhaustive, and the reader is encouraged to look for different possibilities of methods that may be applicable to the particular engineering problem at hand.

A common aphorism, often presented as "*All models are wrong, but some are useful*", is usually considered to be applicable to scientific models in general, and to statistical models in particular. The aphorism recognizes that statistical or scientific models always fall short of the complexities of reality but can still be of use. The aphorism is generally attributed to the statistician George E. P. Box, although the underlying concept predates Box's writings.

Section 2 presents an overview of Volume I - Model-Based and Signal-Based Inverse Methods, which includes chapters on several Optimization Models and Identification Techniques. An introduction and context is presented, as well as a listing of the chapters included in Volume I.

Section 3 presents an overview of Volume II - Fundamental Concepts and Models for the Direct Problem, which includes chapters distributed in Part I – Technologies and Materials Modeling, and in Part II - Discrete Modeling. An introduction and context is presented, as well as a listing of the chapters included in Volume II.

Section 4 presents an overview of Volume III - Uncertainty Modeling: Fundamental Concepts and Models, which includes chapters on Probabilistic Methods and Uncertainty Modeling. An introduction and context is presented, as well as a listing of the chapters included in Volume III.

Section 5 presents some final remarks and acknowledgements.

2 Overview of Volume I

2.1 Model-Based and Signal-Based Inverse Methods – context

The detection, localization, classification and identification of parameters and/or material properties, related to the integrity of structures and components, with and without defects or damages, involves the modeling of inverse problems, as well as an adequate modeling and quantification of the uncertainties involved in the problem.

The formulation of the direct problem, of the inverse problem, and the related uncertainties modeling, needed for an adequate description of the structure and/or the mechanical component, and of its potential defects or damages, involves multidisciplinary modeling techniques, whose understanding and proper application transcends the field of integrity and damage tolerance, being able to serve as a basis for applications, in other contexts or fields.

Among the application problems of interest for inverse methods, one can cite Structural health monitoring (SHM) and Health and Usage Monitoring Systems (HUMS).

The monitoring of structural integrity (SHM) is a competitive technique for damage detection and identification, wherein information is collected online, and compared with an existing database for an undamaged (“healthy”) structure. From this comparison, real-time information on the presence of damages can be obtained, including their localization, size, propagation speed, and, ultimately, the remaining operational life of the structural component.

The monitoring of mechanical components (HUMS) is a technique which is being used to follow / accompany the state of the integrity of mechanical systems / components (Health) and to monitor the appearance of indicators of the presence of damage (usage) in dynamic systems, such as rotating components (in engines, for example) and in gearboxes (in mechanical transmission systems, for example). In this case, the comparison is made between vibration signals from the healthy components (accumulated historical data) and from the components being monitored, to identify significant discrepancies in the vibration signals, which could be correlated to specific / particular / known damages.

The scientific challenge of the modeling of inverse problems, as well as of the adequate modeling and quantification of the related uncertainties, in a problem of integrity of structures and components, involves several aspects:

- The modeling of the problems (direct problem, inverse problem, uncertainties) needs to be done, whenever possible, by using more than one technique, for each case being described, in order to implement, for the particular problem at hand, model techniques which are independent, complementary, and/or redundant. Whenever possible, more than one model should be used, for redundancy and/or comparative analysis, especially in the case of unavailability of prior data for the healthy structure and/or component.
- The techniques used for inverse methods may involve:
 - i. Optimization techniques, based on multi-objective optimization models, using classical optimization techniques (such as Sequential Quadratic Programming (SQP), BFGS, etc), or evolutionary optimization techniques (such as Genetic Algorithms (GA), Differential Evolution (DE), Sunflower Optimization (SFO), Lichtenberg Algorithm (LA), etc.);
 - ii. Identification techniques based on Artificial Intelligence, Machine Learning, Pattern Recognition, Data Science, etc., models (such as identification models based on Artificial Neural Networks (ANN), for example;
 - iii. Models based on the Wavelet Transform (continuous Wavelet Transform (CWT), discrete Wavelet transform (DWT), with different types and sizes of the Wavelet window, for example);
 - iv. Stochastic models (such as Kalman Filter (KF), Extended Kalman Filter (EKF), Extended Information Filter (EIF), Particle Filter (PF), Least Squares (LS), etc).
- In several situations, the direct models to be implemented may involve different problem physics (Multi-physics Modeling), and multiple scales (Multi-scale Modeling). In such cases, the description for the direct problem may involve coarser global models, and more detailed local models;
- The computational simulations and the experimental / laboratory tests must take into account the additional challenge of properly simulating / representing the local behavior of a complex structure, in the regions of interest, where the defect of damage is expected to be, or is expected to appear. For example, Fracture Mechanics (FM) problems and Damage Tolerance (DT) problems cannot be properly represented by reduced-scale models, as the damaged region must be represented using full-scale models. In these cases, the computational simulations (and also the experimental / laboratory tests) are required to reproduce the situation in the region of the damage using high fidelity local models. Thus, the region of the damage must be modeled in full scale, with the model also representing properly the

geometry, the mechanical properties, and the real loading in that local region (loading that is coming from the external loads that were applied in the structure or in the component as a whole).

- In many cases, inverse problems may belong to the category of ill-posed problems, which represents an additional challenge in the modeling of the problem at hand. In these cases, the approach for the inverse method may require additional hypothesis to be made (for problem regularization, for example), or that a meta-modeling approach is adopted (such as surrogate models, reduced order models, etc), replacing the original model by the proper meta-model, and then solving this approximate model for the problem.
- The modeling of inverse problems, such as in the case of SHM and/or HUMS, must take into account the proper modeling of the sensor behavior, and also the uncertainties associated to these sensors, as well as the simultaneous use of multiple, independent, techniques for monitoring, with different sensors. The optimal sensor positioning, to maximize the Probability of Detection (PoD), may be seen as a topological optimization problem and/or as a stochastic optimization problem.
- The modeling of inverse problems may involve the detection, localization, and identification of parameters and/or material properties (for example, properties such as elasticity modulus, Poisson's ratio, etc), which may vary through time (for example, material degradation through time) and also along the length of the structure or mechanical component (for example, local changes which may occur in the material properties and / or mechanical properties of a composite plate, due to the debonding between the layers of the composite material).

2.2 Chapter topics in Volume I: presentation

Along Volume I of the Book Series, several topics related to model-based and signal-based methods for inverse problems were presented in the several book chapters, representing the collaborative work from researchers, engineers, scholars, engaged in research, development and applications in the related areas, affiliated to several institutions, universities, industry, and recognized both nationally and internationally.

The book chapters in Volume I of the Book Series were distributed as follows:

- Chapter 1: Introduction to Optimization and Identification Techniques for Model-Based and Signal-Based Inverse Problems (Jorge, Anflor, et al., 2022b)
- Chapter 2: Overview of Some Optimization and Identification Techniques for Inverse Problems of Detection, Localization and Parameter Estimation (Sousa et al., 2022)
- Chapter 3: An overview of Linear and Non-linear Programming methods for Structural Optimization (Choze, Santos, Jorge, et al., 2022)
- Chapter 4: Overview of Traditional and Recent Heuristic Optimization Methods (Choze, Santos, & Gomes, 2022)
- Chapter 5: Application of Machine Learning and Multi-Disciplinary/Multi-Objective Optimization Techniques for Conceptual Aircraft Design (Mattos et al., 2022)
- Chapter 6: On a Bio-Inspired Method for Topology Optimization via Map L-Systems and Fractone Modeling (Kobayashi, 2022)

- Chapter 7: Fundamentals on the Topological Derivative concept and its classical applications (Carvalho et al., 2022)
- Chapter 8: Ultrasound Obstacle Identification using the Boundary Element and Topological Derivative Methods (Cisilino & Anflor, 2022)
- Chapter 9: Fundamental Concepts on Wavelet Transforms (Palechor, Silva, Morais, et al., 2022)
- Chapter 10: Application of Wavelet Transforms to Structural Damage Monitoring and Detection (Palechor, Silva, Gomes, et al., 2022)
- Chapter 11: Inverse Methods using KF, EKF, EIF, PF, and LS Techniques for Detection, Localization, and Parameter Estimation (Myers & Jorge, 2022)
- Chapter 12: Fundamental Concepts for Impedance-based Structural Health Monitoring (Finzi Neto & Moura, 2022)
- Chapter 13: Fundamental Concepts for Guided Lamb Wave-based Structural Health Monitoring (Finzi Neto et al., 2022)
- Chapter 14: Machine Learning and Pattern Recognition: Methods and Applications for Integrity Monitoring of Civil Engineering Structures (Alves et al., 2022)

3 Overview of Volume II

3.1 Fundamental Concepts and Models for the Direct Problem – context

The modeling of the Direct Problem includes disciplines related to:

- Materials Science, Mechanics of Materials (applications to engineering materials, such as metallic materials, composites, biomaterials, etc.), Material Characterization Techniques (Digital Image Correlation (DIC), etc.);
- Modeling & Design of Structures (applications to several Engineering branches such as Mechanical, Aeronautical, Civil, Naval, Ocean & Offshore, Nuclear, Energy, Oil & Gas, etc.);
- Design of Machines and Mechanical Components;
- Manufacturing Technologies and Processes (Additive Manufacturing, etc.)
- Structural & Component Integrity (Fracture Mechanics, Fatigue, Impact, Contact, Damage Tolerance, Inspection, Non-Destructive Testing (NDE), Maintenance, Structural Health Monitoring (SHM), Health and Usage Monitoring Systems (HUMS), etc.);
- Mechanical Vibrations, Acoustics, Structural Dynamics (Data Acquisition, Modal Testing & Modal Analysis Techniques, etc.);
- Computational Mechanics, Mathematical Methods, Numerical Methods (such as Finite Element Methods (FEM), Boundary Element Methods (BEM), Meshless / Meshfree Methods, etc.).

Regarding structural integrity applications, the problem of detecting damage in a structure involves modeling the direct problem, that is, the behavior of that structure in the presence of one or more pre-established damages, of assumed shape and size, and at given locations. This modeling can be done, for example, involving the analysis of the distribution of stresses and strains in parts with cracks, done through a FEM, BEM or Meshless modeling, or by analyzing the distribution of sound waves in this cracked part, emitted by a pre-established source, by means of a FEM, BEM or Meshless modeling.

Regarding component integrity applications, for the problem of detecting faults in dynamic components, experimental vibration data can be acquired from accelerometers conveniently distributed in the vicinity of the components of interest. In this case, there would be no direct problem modeling (the direct model is implicit), as the vibration data would be acquired for the healthy components to be later compared with the vibration data of the defective (with failures) components.

With respect to Direct Problem Modeling, several problems of interest in structural integrity can be modeled as initial value problems and/or boundary value problems. In the case of boundary value problems, not varying in time, there is a valid partial differential equation in the region that represents the problem domain, and valid boundary conditions at the boundary of this region. In the case of time-dependent initial value problems, there are also valid initial conditions in this domain.

A boundary value problem can be represented by a general equation of type $Au = p$, valid in the domain D , with boundary conditions valid in the boundary ∂D . The complexity of the problem, the geometry of the boundary, the actions (external loads, restrictions on supports) limit the possibilities of analytical solution for this problem, which generates a demand for approximate solutions, such as, for example, through:

- replacement of the differential operator A by an approximate operator A^* . This is the case of the finite difference method, which adopts an algebraic operator A^* , where the derivative at a given point is expressed through the values of the function at points located in a neighborhood of that point, allowing the construction of a system of linear equations, to be solved with the proper application of the initial and boundary conditions;
- replacement of the variable u by an approximate variable u^* , expressed in the form of a linear combination of a set of basic functions. These functions can be polynomial, trigonometric, etc. The coefficients of this linear combination are the components of this approximate solution in the coordinate system represented by these basic functions. When the number of components tends to infinity, the approximate solution u^* tends to the solution u . Among the approximate solutions based on this method, many correspond to methods called “projective”, where the equation valid in the domain is projected in an auxiliary subspace of functions, called support functions, and integrated in the domain, with its boundary conditions, using integration by parts. This category of methods includes, among others:
 - The Finite Element Method (FEM), using an integration by parts and, in each element, the same functions, both for support and for approximation of the variable u , in order to obtain the solution in points of the domain, called nodes, and, by interpolation using these functions, the solution at all other points in the domain;
 - The Boundary Element Method (BEM), using two integrations by parts and using support functions called fundamental solutions, in order to eliminate the domain integral, leaving only the boundary integrals (also known as boundary integrals). Thus, the solution is initially obtained only at the nodes of the discretized contour. Then, by interpolation, the solution is found at the other points of the contour. Finally, by post-processing, the solution is obtained at any desired point in the domain.

For problems of interest in structural integrity, the finite element method can be efficiently applied to problems in which the domain is inhomogeneous, where the discretization of the elements would already consider this local non-homogeneity.

On the other hand, problems involving domain singularities, such as stress distribution in the vicinity of the crack tip, in fracture mechanics, can be well represented using the boundary element method.

For initial value problems, with variable (in the domain and/or in the frontier) dependent on time, it may need a combination of methods that include the finite difference method, to model the march in time of the variable of interest.

Other numerical methods are also available, such as the Meshless / Meshfree Methods, in which the global system of equilibrium equations is constructed using a node-by-node process, performed in the local domain of each node, and thus there is no need to define *a priori* a mesh of finite elements or a mesh of boundary elements.

For the application of inverse methods in dynamic components, the methodology is based on the comparison of vibration data acquired and stored throughout the life of the components still undamaged, with the most recent vibration data, for the components currently in use, from so that significant changes in the spectrum of acquired signals can be used as a starting point for the detection and identification of damage to these components. In this case, instead of a numerical code for the direct problem, vibration and acoustics data acquisition systems are used, together with structural dynamics, vibration analysis, modal testing & modal analysis techniques.

The direct problem modeling can also include the study of the damage evolution with time (such as, for example, the crack propagation rate), in order to estimate the remaining useful life (safe life) of the structure.

The science and technology behind analyzing structures is not yet mature, and further advancements and research are unnecessary for the basic understanding of structural mechanics phenomena and behavior (design/analysis/test) (Hamm et al., 2021). Some remaining challenges in the structures discipline include many areas still in need of rigorous research to arrive at physics-based methodologies, including: Fracture Mechanics, Composite Materials, New Materials, Manufacturing and Processing Methods, New Applications, Lack of Physics-Based Theories, New Test Methods, Incorporation of Research Advancements, Computational Advances, Multi-Discipline Interactions.

One final note; in the context of philosophy of science, two fundamental issues were pointed out in (Da Costa & French, 2006). One issue concerns the appropriate attitude we should take towards scientific theories - whether we should regard them as true or merely empirically adequate, for example. The other issue concerns the nature of scientific theories and models and how these might best be represented. These two issues together lead us to arguing that theories and models should be regarded as partially rather than wholly true, with issues in terms of belief, theory acceptance, and the realism-antirealism debate.

3.2 Chapter topics in Volume II: presentation

Along Volume II of the Book Series, several topics related to technologies, materials and discrete modeling for direct problems are presented in the several book chapters, representing the collaborative work from researchers, engineers, scholars, engaged in research, development and applications in the related areas, affiliated to several institutions, universities, industry, and recognized both nationally and internationally.

The book chapters in Volume II of the Book Series are distributed as follows:

- Chapter 1: Introduction to the Direct Problem: Materials, Technologies, and Discrete Modeling (Jorge, Anflor, et al., 2022c)

Part I - Technologies, Structural Integrity & Material Modeling

- Chapter 2: Technologies and Materials Issues Related to Naval and Ocean Engineering (Marques, 2022)
- Chapter 3: New Advances in Thermoplastic Commingled Composites: Processing and Molecular Dynamics (Benedetto et al., 2022)
- Chapter 4: The Influence of Weaving Patterns on the Effective Mechanical Response of Reinforced Composites - A Study Through Homogenization (Rocha et al., 2022)
- Chapter 5: Auxetic Structures: Parametric Optimization, Additive Manufacturing, and Applications (Francisco et al., 2022)
- Chapter 6: Additive Manufacturing of Ti-Alloys for Aerospace, Naval and Offshore Applications (Mendoza et al., 2022)
- Chapter 7: Review of Low Cycle Fatigue: Issues in Naval and Offshore Engineering (Delgado-Morales et al., 2022)
- Chapter 8: Modeling of Damage in Turbine Blades for Large Deformations (Aslan, 2022)
- Chapter 9: Stresses on Expansion Joints Used in Penetrations (Suanno & Maneschy, 2022)
- Chapter 10: Leak-Before-Break Technology for Piping in Nuclear Industry (Maneschy, 2022)
- Chapter 11: Fatigue Control in Angra Nuclear Power Plants (Cisternas, 2022)
- Chapter 12: On the use of Modal Test Data in Inverse Problems: Fundamentals and Applications (Gomes et al., 2022)
- Chapter 13: Application of Deep Learning Techniques for the Impedance-based SHM to the Oil & Gas Industry (Rezende et al., 2022)
- Chapter 14: Noise, Vibration, and Health and Usage Monitoring Systems (HUMS) of Aircraft Dynamic Components (Cruz et al., 2022)
- Chapter 15: On Vibration Analysis and Health and Usage Monitoring Systems (HUMS) of Dynamic Components of Helicopters (Jorge et al., 2022)
- Chapter 16: Model-Based Parameter Identification for Helicopter Main Rotor Balancing and Tracking Using Once-per-Revolution Vibration Data (Jorge et al., 2022)

Part II - Discrete Modeling: Finite Element, Boundary Element, Meshfree Methods

- Chapter 17: Finite Element Method for Structural Integrity Problems (Sales et al., 2022)
- Chapter 18: An Application of Finite Element Method and Sensitivity Analysis in Structural Dynamics (Avila & Prado, 2022)
- Chapter 19: The Boundary Element Method for Potential Problems (Albuquerque, Jorge, et al., 2022)
- Chapter 20: The Boundary Element Method for Structural Problems (Albuquerque, Sollero, et al., 2022)
- Chapter 21: The Boundary Element Method for Damage Modeling (Peixoto & Chaves, 2022)
- Chapter 22: Application of Roving Mass Technique Associated with Wavelet Transform to Structural Damage Detection and Localization (Palechor, Santos, et al., 2022)
- Chapter 23: Local Mesh Free Methods in Linear Elasticity and Fracture Mechanics (Oliveira et al., 2022)
- Chapter 24: Meshless smoothed point interpolation methods for damage modeling (Gori, 2022)
- Chapter 25: A New Non-Iterative Reconstruction Method for Solving a Class of Inverse Problems (Novotny, 2022)

4 Overview of Volume III

4.1 Uncertainty Modeling: Fundamental Concepts and Models – context

The modeling and quantification of uncertainties includes disciplines related to:

- the stochasticity of the problem variables, which may be modeled as:
 - random variables (univariate or multivariate, Gaussian Mixture Models, etc.) with parameters (for example: mean; standard deviation) constant in space and time (stationary processes); or
 - stochastic processes (when these parameters are variable over time), described by time series; or
 - random fields (when there is a spatial distribution of the parameters of the different random variables along the structure or the mechanical component). In stochastic finite element methods (SFEM), the discretization of the random fields (related to the distributed material properties, or to the distributed loading, for example) may be different than the finite element discretization itself (in general, one element of the random field may be related to a greater number of finite elements);
- probabilistic methods in engineering, including techniques such as Design of Experiments (DoE), Response Surface Methodology (RSM), Bayesian Approaches, Metamodels (Surrogate Models, Reduced Order Models), sensitivity analysis, stochastic optimization (robust optimization / multi-objective optimization) techniques, reliability methods, etc.

- the replacement of the real problem with a model to be simulated and / or tested, leading to several types of errors, such as: model errors (lack of fidelity of the model to the original problem); random errors (improper or incorrect account of the randomness of the problem variables); discretization errors (use of coarse meshes versus mesh refinement), numerical errors (such as rounding errors, etc.). The modeling and quantification of uncertainties may require the use of concepts related to Model Validation & Verification.

4.2 Chapter topics in Volume III: presentation

Along this Volume III of the Book Series, several topics related to uncertainty modeling and probabilistic methods for the direct and inverse problems are presented in the several book chapters, representing the collaborative work from researchers, engineers, scholars, engaged in research, development and applications in the related areas, affiliated to several institutions, universities, industry, and recognized both nationally and internationally.

The book chapters in this Volume III of the Book Series are distributed as follows:

- Chapter 1: Introduction to Probabilistic Methods & Uncertainty Modeling in Direct & Inverse Problems
- Chapter 2: Application of Data Science in the Sustainable Energy Industry
- Chapter 3: Overview of Classic Surrogates and Metamodeling Strategies in Structural Analysis
- Chapter 4: Uncertainty Quantification in Inverse Kinematics Computation of Space Robots by Neural Networks
- Chapter 5: Metamodel-Based Simulation Optimization: A Systematic Literature Review
- Chapter 6: Principal Component Analysis: An Overview and Applications in Multivariate Engineering Problems
- Chapter 7: Gaussian Mixture Models and Their Applications to Fault Diagnosis in Electric Motors
- Chapter 8: Sensor Placement Optimization of a Helicopter Main Rotor Blade
- Chapter 9: Challenges for Structural Health Monitoring: Nonlinearities and Uncertainties
- Chapter 10: Applications of Numerical and Statistical Methods for Monitoring and Control Engineering Variables in a Concrete Dam
- Chapter 11: Application of Statistical Processing Techniques to the Impedance-based SHM for the Oil & Gas Industry
- Chapter 12: Neuro-Fuzzy Data Normalization Applied for Impedance-based SHM to the Oil & Gas Industry
- Chapter 13: On Kriging Techniques & Impedance-based SHM as Applied to Damage Detection in 2D Structures
- Chapter 14: Time Series Models and Expansion Optimal Linear Estimation for Time-Dependent Reliability Analysis
- Chapter 15: Bayesian Optimal Experimental Design

- Chapter 16: Airplane Optimal Robust Design Combined with an Entropy-Statistics Unsupervised Learning Classification Algorithm
- Chapter 17: Random Fatigue Under Uniaxial and Multiaxial Loading Conditions
- Chapter 18: Random Field Discretization in Stochastic Finite Elements: Element Size Effects and Error Indicators in Reliability Structural Analysis
- Chapter 19: On Error Estimators and Bayesian Approaches in Computational Model Validation

5 Final remarks and acknowledgements

This Chapter presents an overview of the Book Series in Direct Methods, Inverse Methods and Uncertainty Modeling, with focus on:

- The previously published Volume I: Model-Based and Signal-Based Inverse Methods (DOI: <https://doi.org/10.4322/978-65-86503-71-5>), including an introduction to the different topics in Optimization and Identification Techniques detailed in Volume I of the Book Series;
- The previously published Volume II: Fundamental Concepts and Models for the Direct Problem (DOI: <https://doi.org/10.4322/978-65-86503-83-8>), including an introduction to the topics in Materials, Technologies and Discrete Models, detailed in Volume II of the Book Series;
- The current Volume III: Uncertainty Modeling: Fundamental Concepts and Models, including an introduction to the topics in Uncertainty Modeling and Probabilistic Methods, presented in this Volume III of the Book Series.

This Book Series is an initiative of the Graduate Program in Integrity of Engineering Materials (PPG-Integridade) at the University of Brasilia (UnB), Brazil (www.pgintegridade.unb.br). The editors would like to thank PPG-Integridade and UnB for the initiative, incentive and support for this Book Series project.

The book series is organized as a collaborative work involving researchers, engineers, scholars, engaged in research, development and applications in the related areas, affiliated to several institutions, universities, industry, and recognized both nationally and internationally. The editors are grateful and would like to show their appreciation to all the co-authors of the book chapters, for their participation, dedication, and support.

The book series is published as a digital version, with ISBN provided by UnB, and DOI for each chapter, provided by the Latin American Journal of Solids and Structures (LAJSS) (www.lajss.org).

The scope of the Book series is in the broad areas of interest of LAJSS, and also of the Brazilian Society of Mechanical Sciences and Engineering (ABCM) (www.abcm.org.br) and the Brazilian Association of Computational Methods in Engineering (ABMEC) (www.abmec.org.br). For increased visibility, these three institutions are encouraging the divulgation of the Book Series project in their websites. The editors would like to express their appreciation to LAJSS, ABCM and ABMEC, for their incentive, encouragement and support for this Book Series project.

References

- Albuquerque, E. L., Jorge, A. B., & Anflor, C. T. M. (2022). The Boundary Element Method for Potential Problems. In A. B. Jorge, C. T. M. Anflor, G. F. Gomes, & S. H. S. Carneiro (Eds.), *Fundamental Concepts and Models for the Direct Problem* (pp. 688–732). Biblioteca Central da Universidade de Brasilia. <https://doi.org/10.4322/978-65-86503-83-8.c19>
- Albuquerque, E. L., Sollero, P., & Aliabadi, M. H. (2022). The Boundary Element Method for Structural Problems. In A. B. Jorge, C. T. M. Anflor, G. F. Gomes, & S. H. S. Carneiro (Eds.), *Fundamental Concepts and Models for the Direct Problem* (pp. 733–808). Biblioteca Central da Universidade de Brasilia. <https://doi.org/10.4322/978-65-86503-83-8.c20>
- Alves, V. N., Cury, A. A., Roitman, N., & Magluta, C. (2022). Machine Learning and Pattern Recognition: Methods and Applications for Integrity Monitoring of Civil Engineering Structures. In A. B. Jorge, C. T. M. Anflor, G. F. Gomes, & S. H. S. Carneiro (Eds.), *Model-based and Signal-Based Inverse Methods* (pp. 502–535). <https://doi.org/10.4322/978-65-86503-71-5.c14>
- Aslan, O. (2022). Modeling of Damage in Turbine Blades for Large Deformations. In A. B. Jorge, C. T. M. Anflor, G. F. Gomes, & S. H. S. Carneiro (Eds.), *Fundamental Concepts and Models for the Direct Problem* (pp. 209–230). Biblioteca Central da Universidade de Brasilia. <https://doi.org/10.4322/978-65-86503-83-8.c08>
- Avila, S. M., & Prado, Z. J. G. (2022). An Application of Finite Element Method and Sensitivity Analysis in Structural Dynamics. In A. B. Jorge, C. T. M. Anflor, G. F. Gomes, & S. H. S. Carneiro (Eds.), *Fundamental Concepts and Models for the Direct Problem* (pp. 666–687). Biblioteca Central da Universidade de Brasilia. <https://doi.org/10.4322/978-65-86503-83-8.c18>
- Benedetto, R. M., Janotti, A., Gomes, G. F., Ancelotti, A. C., & Botelho, E. C. (2022). New Advances in Thermoplastic Commingled Composites: Processing and Molecular Dynamics. In A. B. Jorge, C. T. M. Anflor, G. F. Gomes, & S. H. S. Carneiro (Eds.), *Fundamental Concepts and Models for the Direct Problem* (pp. 51–75). Biblioteca Central da Universidade de Brasilia. <https://doi.org/10.4322/978-65-86503-83-8.c03>
- Carvalho, F. S., Anflor, C. T. M., Jorge, A. B., Cisilino, A. P., & Marczak, R. J. (2022). Fundamentals on the Topological Derivative concept and its classical applications. In A. B. Jorge, C. T. M. Anflor, G. F. Gomes, & S. H. S. Carneiro (Eds.), *Model-based and Signal-Based Inverse Methods* (pp. 269–300). <https://doi.org/10.4322/978-65-86503-71-5.c07>
- Choze, S. B., Santos, R. R., & Gomes, G. F. (2022). Overview of Traditional and Recent Heuristic Optimization Methods. In A. B. Jorge, C. T. M. Anflor, G. F. Gomes, & S. H. S. Carneiro (Eds.), *Model-based and Signal-Based Inverse Methods* (pp. 107–142). <https://doi.org/10.4322/978-65-86503-71-5.c04>
- Choze, S. B., Santos, R. R., Jorge, A. B., & Gomes, G. F. (2022). An overview of Linear and Non-linear Programming methods for Structural Optimization. In A. B. Jorge, C. T. M. Anflor, G. F. Gomes, & S. H. S. Carneiro (Eds.), *Model-based and Signal-Based Inverse Methods* (pp. 65–106). <https://doi.org/10.4322/978-65-86503-71-5.c03>

- Cisilino, A. P., & Anflor, C. T. M. (2022). Ultrasound Obstacle Identification using the Boundary Element and Topological Derivative Methods. In A. B. Jorge, C. T. M. Anflor, G. F. Gomes, & S. H. S. Carneiro (Eds.), *Model-based and Signal-Based Inverse Methods* (pp. 301–330). <https://doi.org/10.4322/978-65-86503-71-5.c08>
- Cisternas, M. (2022). Fatigue Control in Angra Nuclear Power Plants. In A. B. Jorge, C. T. M. Anflor, G. F. Gomes, & S. H. S. Carneiro (Eds.), *Fundamental Concepts and Models for the Direct Problem* (pp. 277–310). Biblioteca Central da Universidade de Brasilia. <https://doi.org/10.4322/978-65-86503-83-8.c11>
- Cruz, A. R., Jorge, A. B., Mattos, B. S., & Cruz, R. V. (2022). Noise, Vibration, and Health and Usage Monitoring Systems (HUMS) of Aircraft Dynamic Components. In A. B. Jorge, C. T. M. Anflor, G. F. Gomes, & S. H. S. Carneiro (Eds.), *Fundamental Concepts and Models for the Direct Problem* (pp. 386–450). Biblioteca Central da Universidade de Brasilia. <https://doi.org/10.4322/978-65-86503-83-8.c14>
- Da Costa, N. C. A., & French, S. (2006). Science and Partial Truth: A Unitary Approach to Models and Scientific Reasoning. In *Science and Partial Truth: A Unitary Approach to Models and Scientific Reasoning*. <https://doi.org/10.1093/019515651X.001.0001>
- Delgado-Morales, L. L., Palma, V. C., Gu, J., & Molinas, E. (2022). Review of Low Cycle Fatigue: Issues in Naval and Offshore Engineering. In A. B. Jorge, C. T. M. Anflor, G. F. Gomes, & S. H. S. Carneiro (Eds.), *Fundamental Concepts and Models for the Direct Problem* (pp. 184–208). Biblioteca Central da Universidade de Brasilia. <https://doi.org/10.4322/978-65-86503-83-8.c07>
- Finzi Neto, R. M., & Moura, J. R. V. (2022). Fundamental Concepts for Impedance-based Structural Health Monitoring. In A. B. Jorge, C. T. M. Anflor, G. F. Gomes, & S. H. S. Carneiro (Eds.), *Model-based and Signal-Based Inverse Methods* (pp. 443–471). <https://doi.org/10.4322/978-65-86503-71-5.c12>
- Finzi Neto, R. M., Rezende, S. W. F., & Moura, J. R. V. (2022). Fundamental Concepts for Guided Lamb Wave-based Structural Health Monitoring. In A. B. Jorge, C. T. M. Anflor, G. F. Gomes, & S. H. S. Carneiro (Eds.), *Model-based and Signal-Based Inverse Methods* (pp. 472–501). <https://doi.org/10.4322/978-65-86503-71-5.c13>
- Francisco, M. B., Pereira, J. L. J., & Gomes, G. F. (2022). Auxetic Structures: Parametric Optimization, Additive Manufacturing, and Applications. In A. B. Jorge, C. T. M. Anflor, G. F. Gomes, & S. H. S. Carneiro (Eds.), *Fundamental Concepts and Models for the Direct Problem* (pp. 120–153). Biblioteca Central da Universidade de Brasilia. <https://doi.org/10.4322/978-65-86503-83-8.c05>
- Gomes, G. F., Carneiro, S. H. S., Cesnik, C. E. S., & Jorge, A. B. (2022). On the use of Modal Test Data in Inverse Problems: Fundamentals and Applications. In A. B. Jorge, C. T. M. Anflor, G. F. Gomes, & S. H. S. Carneiro (Eds.), *Fundamental Concepts and Models for the Direct Problem* (pp. 311–348). Biblioteca Central da Universidade de Brasilia. <https://doi.org/10.4322/978-65-86503-83-8.c12>
- Gori, L. (2022). Meshless smoothed point interpolation methods for damage modelling. In A. B. Jorge, C. T. M. Anflor, G. F. Gomes, & S. H. S. Carneiro (Eds.), *Fundamental Concepts and Models for the Direct Problem* (pp. 959–1006). Biblioteca Central da Universidade de Brasilia. <https://doi.org/10.4322/978-65-86503-83-8.c24>

- Hamm, K. R., Imtiaz, K. S., & Raju, I. S. (2021). Some remaining challenges in aerospace structures. *AIAA Scitech 2021 Forum*. <https://doi.org/10.2514/6.2021-0574>
- Jorge, A. B. (2020). *Technological Demonstration Platform for Inverse Methods and Uncertainty Modeling Integrity of Structures and Components (Plataforma demonstradora tecnológica para métodos inversos e modelagem de incertezas em integridade de estruturas e componentes)*. University of Brasilia. <https://repositorio.unb.br/handle/10482/39570>
- Jorge, A. B., Anflor, C. T. M., Gomes, G. F., & Carneiro, S. H. S. (Eds.). (2022a). *Fundamental Concepts and Models for the Direct Problem: Vol. II* (1st ed.). Biblioteca Central da Universidade de Brasilia. <https://doi.org/10.4322/978-65-86503-83-8>
- Jorge, A. B., Anflor, C. T. M., Gomes, G. F., & Carneiro, S. H. S. (2022b). Introduction to Optimization and Identification Techniques for Model-Based and Signal-Based Inverse Problems. In A. B. Jorge, C. T. M. Anflor, G. F. Gomes, & S. H. S. Carneiro (Eds.), *Model-based and Signal-Based Inverse Methods* (pp. 1–7). <https://doi.org/10.4322/978-65-86503-71-5.c01>
- Jorge, A. B., Anflor, C. T. M., Gomes, G. F., & Carneiro, S. H. S. (2022c). Introduction to the Direct Problem: Materials, Technologies, and Discrete Modeling. In A. B. Jorge, C. T. M. Anflor, G. F. Gomes, & S. H. S. Carneiro (Eds.), *Fundamental Concepts and Models for the Direct Problem* (pp. 1–13). Biblioteca Central da Universidade de Brasilia. <https://doi.org/10.4322/978-65-86503-83-8.c01>
- Jorge, A. B., Anflor, C. T. M., Gomes, G. F., & Carneiro, S. H. S. (Eds.). (2022d). *Model-based and Signal-Based Inverse Methods*. <https://doi.org/10.4322/978-65-86503-71-5>
- Jorge, A. B., Cruz, R. V., Mattos, B. S., & González, R. A. (2022). On Vibration Analysis and Health and Usage Monitoring Systems (HUMS) of Dynamic Components of Helicopters. In A. B. Jorge, C. T. M. Anflor, G. F. Gomes, & S. H. S. Carneiro (Eds.), *Fundamental Concepts and Models for the Direct Problem* (pp. 451–518). Biblioteca Central da Universidade de Brasilia. <https://doi.org/10.4322/978-65-86503-83-8.c15>
- Jorge, A. B., González, R. A., Mattos, B. S., & Cruz, R. V. (2022). Model-Based Parameter Identification for Helicopter Main Rotor Balancing and Tracking Using Once-per-Revolution Vibration Data. In A. B. Jorge, C. T. M. Anflor, G. F. Gomes, & S. H. S. Carneiro (Eds.), *Fundamental Concepts and Models for the Direct Problem* (pp. 519–616). Biblioteca Central da Universidade de Brasilia. <https://doi.org/10.4322/978-65-86503-83-8.c16>
- Kobayashi, M. H. (2022). On a Bio-Inspired Method for Topology Optimization via Map L-Systems and Fractone Modeling. In A. B. Jorge, C. T. M. Anflor, G. F. Gomes, & S. H. S. Carneiro (Eds.), *Model-based and Signal-Based Inverse Methods* (pp. 237–268). <https://doi.org/10.4322/978-65-86503-71-5.c06>
- Maneschy, J. E. (2022). Leak-Before-Break Technology for Piping in Nuclear Industry. In A. B. Jorge, C. T. M. Anflor, G. F. Gomes, & S. H. S. Carneiro (Eds.), *Fundamental Concepts and Models for the Direct Problem* (pp. 260–276). Biblioteca Central da Universidade de Brasilia. <https://doi.org/10.4322/978-65-86503-83-8.c10>

- Marques, A. L. F. (2022). Technologies and Materials Issues Related to Naval and Ocean Engineering. In A. B. Jorge, C. T. M. Anflor, G. F. Gomes, & S. H. S. Carneiro (Eds.), *Fundamental Concepts and Models for the Direct Problem* (pp. 15–50). Biblioteca Central da Universidade de Brasilia. <https://doi.org/10.4322/978-65-86503-83-8.c02>
- Mattos, B. S., Bortolete, F. A., Fregnani, J. A. T. G., Jorge, A. B., Alves, W. M., & Cruz, R. V. (2022). Application of Machine Learning and Multi-Disciplinary/Multi-Objective Optimization Techniques for Conceptual Aircraft Design. In A. B. Jorge, C. T. M. Anflor, G. F. Gomes, & S. H. S. Carneiro (Eds.), *Model-based and Signal-Based Inverse Methods* (pp. 143–236). <https://doi.org/10.4322/978-65-86503-71-5.c05>
- Mendoza, M. Y., Delgado-Morales, L. L., & Palma, V. C. (2022). Additive Manufacturing of Ti-Alloys for Aerospace, Naval and Offshore Applications. In A. B. Jorge, C. T. M. Anflor, G. F. Gomes, & S. H. S. Carneiro (Eds.), *Fundamental Concepts and Models for the Direct Problem* (pp. 154–183). Biblioteca Central da Universidade de Brasilia. <https://doi.org/10.4322/978-65-86503-83-8.c06>
- Myers, M. R., & Jorge, A. B. (2022). Inverse Methods using KF, EKF, EIF, PF, and LS Techniques for Detection, Localization, and Parameter Estimation. In A. B. Jorge, C. T. M. Anflor, G. F. Gomes, & S. H. S. Carneiro (Eds.), *Model-based and Signal-Based Inverse Methods* (pp. 382–442). <https://doi.org/10.4322/978-65-86503-71-5.c11>
- Novotny, A. A. (2022). A New Non-Iterative Reconstruction Method for Solving a Class of Inverse Problems. In A. B. Jorge, C. T. M. Anflor, G. F. Gomes, & S. H. S. Carneiro (Eds.), *Fundamental Concepts and Models for the Direct Problem* (pp. 1007–1023). Biblioteca Central da Universidade de Brasilia. <https://doi.org/10.4322/978-65-86503-83-8.c25>
- Oliveira, T., Vélez, W., & Portela, A. (2022). Local Mesh Free Methods in Linear Elasticity and Fracture Mechanics. In A. B. Jorge, C. T. M. Anflor, G. F. Gomes, & S. H. S. Carneiro (Eds.), *Fundamental Concepts and Models for the Direct Problem* (pp. 899–958). Biblioteca Central da Universidade de Brasilia. <https://doi.org/10.4322/978-65-86503-83-8.c23>
- Palechor, E. U. L., Santos, J. C., Silva, R. S. Y. C., Machado, M. R., Morais, M. V. G., Bezerra, L. M., & Jorge, A. B. (2022). Application of Roving Mass Technique Associated with Wavelet Transform to Structural Damage Detection and Localization. In A. B. Jorge, C. T. M. Anflor, G. F. Gomes, & S. H. S. Carneiro (Eds.), *Fundamental Concepts and Models for the Direct Problem* (pp. 881–898). Biblioteca Central da Universidade de Brasilia. <https://doi.org/10.4322/978-65-86503-83-8.c22>
- Palechor, E. U. L., Silva, R. S. Y. C., Gomes, G., Morais, M. V. G., Bezerra, L. M., & Jorge, A. B. (2022). Application of Wavelet Transforms to Structural Damage Monitoring and Detection. In A. B. Jorge, C. T. M. Anflor, G. F. Gomes, & S. H. S. Carneiro (Eds.), *Model-based and Signal-Based Inverse Methods* (pp. 357–381). <https://doi.org/10.4322/978-65-86503-71-5.c10>
- Palechor, E. U. L., Silva, R. S. Y. C., Morais, M. V. G., Bezerra, L. M., & Jorge, A. B. (2022). Fundamental Concepts on Wavelet Transforms. In A. B. Jorge, C. T. M. Anflor, G. F. Gomes, & S. H. S. Carneiro (Eds.), *Model-based and Signal-Based Inverse Methods* (pp. 331–356). <https://doi.org/10.4322/978-65-86503-71-5.c09>

- Peixoto, R. G., & Chaves, A. P. (2022). The Boundary Element Method for Damage Modeling. In A. B. Jorge, C. T. M. Anflor, G. F. Gomes, & S. H. S. Carneiro (Eds.), *Fundamental Concepts and Models for the Direct Problem* (pp. 809–880). Biblioteca Central da Universidade de Brasilia. <https://doi.org/10.4322/978-65-86503-83-8.c21>
- Rezende, S. W. F., Moura, J. R. V., Silva, J. W., Rabelo, D. S., Nomelini, Q. S. S., Finzi, R. M., Gallo, C. A., & Ramos, J. E. (2022). Application of Deep Learning Techniques for the Impedance-based SHM to the Oil & Gas Industry. In A. B. Jorge, C. T. M. Anflor, G. F. Gomes, & S. H. S. Carneiro (Eds.), *Fundamental Concepts and Models for the Direct Problem* (pp. 349–385). Biblioteca Central da Universidade de Brasilia. <https://doi.org/10.4322/978-65-86503-83-8.c13>
- Rocha, J. T., Lisboa, T. V., & Marczak, R. J. (2022). The Influence of Weaving Patterns on the Effective Mechanical Response of Reinforced Composites - A Study Through Homogenization. In A. B. Jorge, C. T. M. Anflor, G. F. Gomes, & S. H. S. Carneiro (Eds.), *Fundamental Concepts and Models for the Direct Problem* (pp. 76–119). Biblioteca Central da Universidade de Brasilia. <https://doi.org/10.4322/978-65-86503-83-8.c04>
- Sales, T. P., Martins, P. C. O., & Rade, D. A. (2022). Finite Element Method for Structural Integrity Problems. In A. B. Jorge, C. T. M. Anflor, G. F. Gomes, & S. H. S. Carneiro (Eds.), *Fundamental Concepts and Models for the Direct Problem* (pp. 618–665). Biblioteca Central da Universidade de Brasilia. <https://doi.org/10.4322/978-65-86503-83-8.c17>
- Sousa, B. S., Gomes, G. F., Alexandrino, P. S. L., Cunha, S. S., & Jorge, A. B. (2022). Overview of Some Optimization and Identification Techniques for Inverse Problems of Detection, Localization and Parameter Estimation. In A. B. Jorge, C. T. M. Anflor, G. F. Gomes, & S. H. S. Carneiro (Eds.), *Model-based and Signal-Based Inverse Methods* (pp. 8–64). <https://doi.org/10.4322/978-65-86503-71-5.c02>
- Suanno, R., & Maneschy, J. E. (2022). Stresses on Expansion Joints Used in Penetrations. In A. B. Jorge, C. T. M. Anflor, G. F. Gomes, & S. H. S. Carneiro (Eds.), *Fundamental Concepts and Models for the Direct Problem* (pp. 231–259). Biblioteca Central da Universidade de Brasilia. <https://doi.org/10.4322/978-65-86503-83-8.c09>

Chapter 2

Application of Data Science in the Sustainable Energy Industry

Chapter details

Chapter DOI:

<https://doi.org/10.4322/978-65-86503-88-3.c02>

Chapter suggested citation / reference style:

Marques, André L. F. (2022). “Application of Data Science in the Sustainable Energy Industry”. In Jorge, Ariosto B., et al. (Eds.) *Uncertainty Modeling: Fundamental Concepts and Models*, Vol. III, UnB, Brasilia, DF, Brazil, pp. 19–53. Book series in Discrete Models, Inverse Methods, & Uncertainty Modeling in Structural Integrity.

P.S.: DOI may be included at the end of citation, for completeness.

Book details

Book: Uncertainty Modeling: Fundamental Concepts and Models

Edited by: Jorge, Ariosto B., Anflor, Carla T. M., Gomes, Guilherme F., & Carneiro, Sergio H. S.

Volume III of Book Series in:

Discrete Models, Inverse Methods, & Uncertainty Modeling in Structural Integrity

Published by: UnB City: Brasilia, DF, Brazil Year: 2022

DOI: <https://doi.org/10.4322/978-65-86503-88-3>

Application of Data Science in the Sustainable Energy Industry

André Luís Ferreira Marques

Naval & Nuclear Engineer - Rear Admiral (Naval Constructor) (retired) Brazilian Navy. Affiliations: University of São Paulo (USP), Electrical & Computer Engineering - PhD candidate. Paulista University (UNIP), Electrical Engineering (ICET) - Professor. Green Yellow of Brazil - Data Scientist. Brazil. E-mail: aferreiramarques@uol.com.br

Abstract

The application of Data Science (DS) in engineering fields is presented from a broad approach, taking the basics and historical motivation to apply statistical tools into practical needs, from Health Sciences to structural monitoring, the latter linked to the sustainable energy context, like solar, wind, geothermal and nuclear. Moreover, brief bullets about the DS techniques and models are also covered, most in the Machine Learning branch, but with some indications of Artificial Intelligence applications. Math techniques like Kringing, Support Vector Machines, Cross-Validation, Data Regularization, Surrogate Models, for instance, also make part of the text, where the practical issues are evaluated in a brief technical manner. The chapter covers the structural reliability and its DS application too.

Keywords: Data Science, Machine Learning, sustainable energy, uncertainty modeling, regularization techniques.

1 Introduction

The analysis of data has improved in several dimensions along the evolution of the digital systems in the last three decades, with more data volume been processed faster and faster every year. The Big Data environment has used tools in the analysis of phenomena and scenarios regarding the day routine of people, services, and countries. Data analysis has been a key instrument on the evolution of expansion campaigns of enterprises, marketing operations, public planning among others [1].

One example of how complex a data analysis can be, take the examples from war battles or campaigns, where data from different subjects come together, such as logistics and systems operational performance, to build up complex scenarios, to ease decision on the maximization of the combat capability. For instance, some data have origins from different time zones and production rates, and must be integrated into databanks for further processing, which normally considers the extraction, treatment and launching tasks. Nowadays, the military decision makers have data sources such as satellites, fixed

radar stations, mobile communication units, drones and even from cell phones, piling up amounts of data with different standards, languages, types, and volume. In the end, the analysis of data shall produce a valuable information to support the decision processes and culture in a short and long term as well [2, 3]

Earth environment sciences have been another technical subject with a high profit from the application of DS, manipulating data from satellites, fixed and mobile instruments aboard of planes, ships and voitures [4]. Equally important, people and animals can also provide data to be collected by research centers and universities linked to the environment surveys. The digital hardware has increased its reliability and portability with the development of electric materials and associated software. The data gathered go to databanks, data lakes etc., where can be managed to feed up analytical methods to compose data analysis about air pollution, water quality, land contamination over time, as few examples.

This chapter deals with some applications of Data Science (DS) into engineering fields, focusing current renewable energies options. The scope covers to show and understand the application of DS, throughout some examples, to address technical needs and outcomes, centered in data exploration, data classification and prediction, within dynamic space and time boundaries [5].

2 Development

2.1 Data Science: what's for?

In a few words, DS can be outlined by the combination of mathematics and computing disciplines and tools, including numeric calculus, analytic geometry, statistics, probability, to mention some of them, to handle a set of data, or a dataset, to provide some understanding, or even a better recognition, of characteristics and behaviors that might be inside the dataset. The whole pack of mathematical techniques have been integrated into computer schemes, accelerating the outcomes towards a more data centered conclusions. Figure 1 shows an overall view of the DS scope [6].

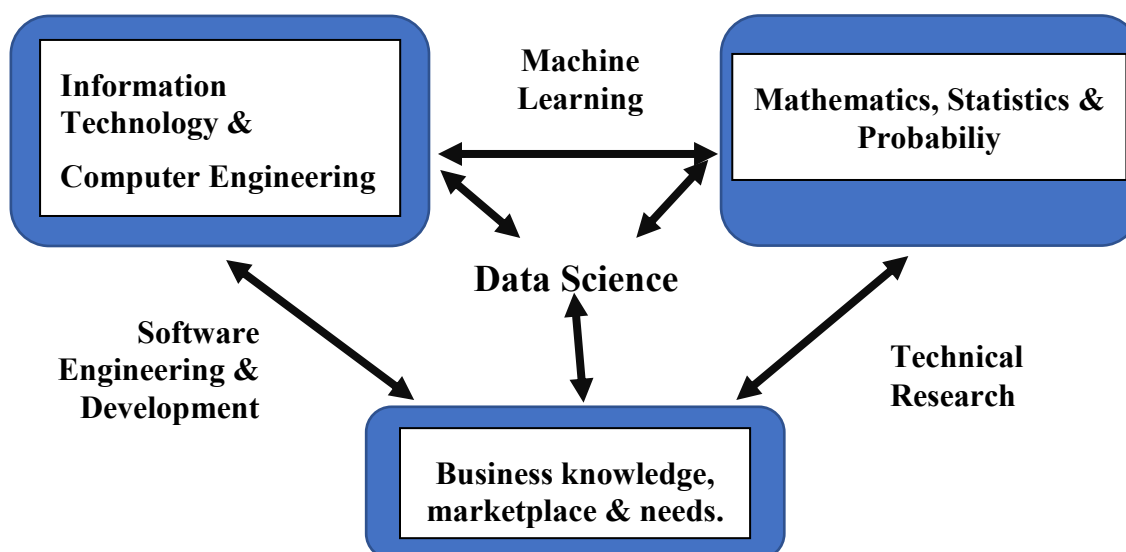


Figure 1: Scope of Data Science and other IT knowledge fields [6].

Imagine an enterprise received a large amount of data, from different sources, like sensors, human observations and so on. The enterprise has some points of interest to check and evaluate whether the data collected may present some insights or ideas never devised. Thus, the enterprise needs a set of criteria and procedures to go over the data and try to grab the outputs foreseen. Nonetheless, the insights may not be present in the way earlier planned and the data may require a cleanup or organization not so easy to detect throughout a simple visual inspection. That is the scenario where DS applies with different approaches.

Just organizing the data in a certain format can provide conclusions or deeper research directions, saving money and work hours from analysis not needed in the end, or more wisely to aim the effort into what really matters for the purpose of the enterprise or investigation. Take the example of the issues dealing with diseases, by checking the data about the occurrences in space and time. Mapping the cases and their characteristics can lead to know about a specific pattern and, thus, to preview future events or trends, which allow the people involved to get ready to dim the intensity of the consequences or even to stop the whole cause or the disease. As a remark, “mapping” means to organize the data into matrices, with dimensions related to the problem, which demands a deep knowledge of the details. That is the contribution of Florence Nightingale, during the Crimea War, which devised a way to organize data, from soldier casualties and deaths, using a bidimensional map, linking the data in terms of time [7]. This tool helped to know more about on how to decrease the harsh time of patient’s recovery and to avoid worse cases as well. Figure 2 shows her drawings, forecasting what it has been known today as “Data Analytics”.

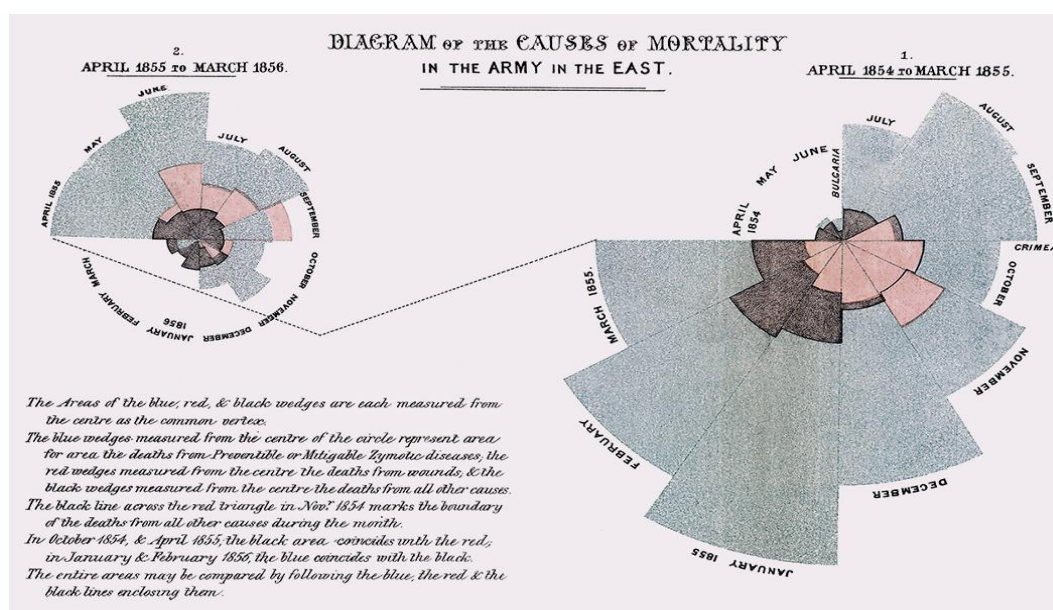


Figure 2: Florence Nightingale maps over soldier casualties and time [7]

With the development of the digital systems and the internet, the uses of math algorithms have spread to almost all areas of knowledge, improving the power and quality of observations, testing, and the accumulated experience [8]. Weather forecast, traffic control and disease control are one of the examples, managing large amounts of data, with different spectra and range of variables, only possible with digital computers and the algorithms tailored tuned. To interpret this huge set of data and build up a conclusion from it, the statistical algorithms have been used in large scale and faster, providing chances for insights not foreseen before. The higher the computing capability, the larger can be the dataset to be handle, and therefore more chances for better information to be checked and transmitted, and to allow more focused decisions to be made.

One fair example of the data handling linking odd parameters like cholera outbreaks and geographic locations, comes from medical doctor Snow, in London (1854), he was not a cartographer though [9]. In the end, but the association of the disease people data and where they used to live and the housing conditions, Dr. Snow was able to take conclusions and to control the situation. In fact, he used something closer today what is known as Data Analytics, well done. To dive into something closer to DS, the team of Dr. Snow could check a math model to classify the people according to the money income, age, water and food supply, education level, the intensity of the fever or other symptoms, for how long they had survived and others. Thus, the output of the model would be an index or number, grading the cholera case in terms of time and position, allowing some guessing how the disease would spread to other parts of London.

This text deals with the DS subject in a broad or horizontal view, with some applications, not covering coding examples, nor detailed or specific math procedures, steps, or models, by their equations. Rather, straight explanations and references will be used.

2.1.1 Brief history of DS

The DS history can be outlined as a progressive crescent timeline, with different increasing rates though. Taking the developments centered to handle information, the set up to data into tables, graphics and maps may be considered one of the first moves of DS, in parallel with description texts about phenomena and human actions. The main reason

for that also may come from religion and political issues, as the first motion springs to act in the societies. This feature may also be made without the scientific reason known as today and, therefore, reviewing historic pieces of plans, rule registers and description texts it is possible to identify the application basics of DS, when the use of rational steps to make conclusions based on collected data, mainly after the XVII and XVIII centuries in Europe.

It is worth noting the data processing period at that time can be taken as a couple of months or weeks. Hour and minutes could not be achieved due to the technology obstacles when dealing with the energy processing. The spread of electric hardware changed this frame after the middle of the XIX Century.

A half century ago almost, the paper “A Mathematical Theory of Communications” (1948) written by Claude Shannon, presented the term ‘bit, coined by John W. Tukey in the prior year, starting the information process path of today, at the time the digital computer evolved. Almost 30 year later, the book “Concise Survey of Computer Methods” [9], written by Peter Naur, synthesized the data processing methods used in several practical subjects, and the definition of DS appeared as: “...the science of dealing with data, once they have been established, while the relation of the data to what they represent is delegated to other fields and sciences”. In a brief review, it resumes the main issues in DS. In 1989, the workshop “Knowledge Discovery in Databases (KDD)” was organized by Gregory P-Shapiro and can be seen as definitive move towards a stronger link between data analysis and statistics.

As private companies have handled more data related to marketing campaigns, the notion of a more rational use of math, statistics, and computer science to check what messages and relations may be present inside a data file, now collected more extensively as the communications webs enlarged exponentially. Thus, in 1996 two marks happened to clarify the DS scope: the book “From Data Mining to Knowledge Discovery in Databases” written by Usama Fayyad, Gregory P-Shapiro and Padhraic Smyth, and the conference “Data Science, classification, and related methods”, held by the International Federation of Classification Societies (IFCS), both presented the first definitions and applications of all practices so far, focusing the classification problems mostly, which can be seen as the first natural field of DS. Leon Breiman pushed the envelope in 2001 with the book “Statistical Modelling: The Two Cultures”, enhancing the importance of statistics application when analyzing data: ‘If our goal as a field is to use data to solve problems, then we need to move away from exclusive dependence on data models and adopt a more diverse set of tools’. This was said because he realized the use of statistics approaches had led to questionable conclusions and that the algorithmic modeling had been used outside statistics, dealing with different sizes of datasets [10].

The evolution of DS has all to do with the development of the digital technology because, now than ever, the capacity to perform hard calculations has been feasible in shorter time intervals (order of milliseconds on average, connecting billions of machines. Electronics and materials have increased their performance in an exponential scale, making data analysis more available to a broader set of researchers, enterprises, and governments. One example of this development is the ‘smartphone’, which is no longer only a ‘mobile phone’, but also a versatile ‘data source’ and ‘data processor’.

Looking ahead, Figure 3 presents a distribution of subjects in terms of time and areas, regarding potential expansion applications, migrating from ‘things’ to thoughts. DS

works among all the sectors taking different signals transformed into binary numbers and checking what information can be draw from them. Currently, optical character processing has gained focus due to applications related to asset security and document automatic processing, saving money and time. A potential field of interest is the natural language processing (NLP) and the ‘wearable hardware’, coming closer and closer to human ordinary tasks.

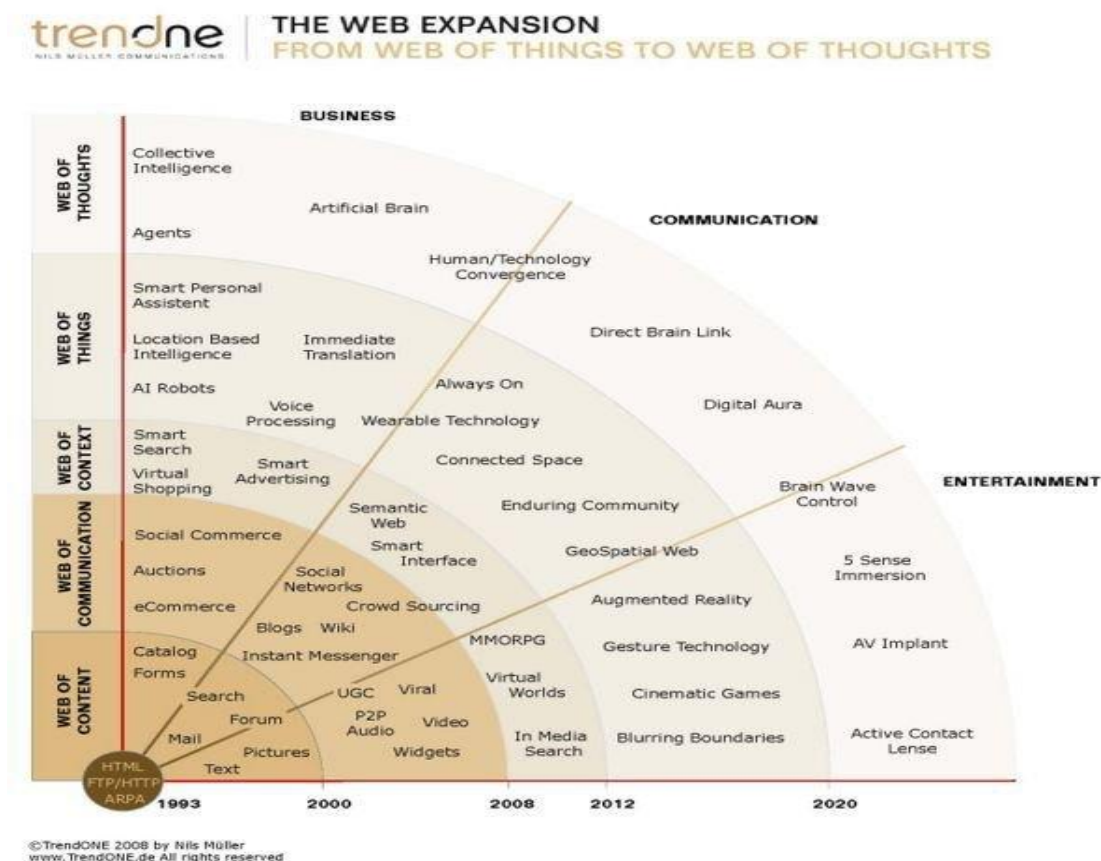


Figure 3: A short view of the web expansion and distribution [11]

In the last twenty years, journals, workshops, and thematic congress have been organized and the application of DS has focused the education alternatives and the market need fulfillment in many countries. As today, the DS has key presence in our routines, some not easy to be detected, and “Data Scientists” belong to the professionals needed for the XXI Century and beyond.

2.2 Sustainable & renewable energies: the main types and frameworks

In the last fifty years, the atmospheric phenomena have taken the world attention, due to the higher intensity of the climatic events, such as flooding, forest fires, ocean level increase, draughts span and others. One of the causes of this complex scenario is the emission of greenhouse gases, where Carbon Dioxide (CO₂) and Methane (CH₄) represents the most important ones for instance [12].

Several solutions have been under evaluation to dim the greenhouse gas emission, such as to reduce the use of fossil fuels, using less environment aggressive energy sources, like

the solar, wind and nuclear options. Each one has a trade-off, and their selection depends on economics, geopolitical and logistics issues. The recent evaluations using data grabbed from satellites, fixed instrument stations and even from people, lead to conclusions data driven, thinking on the balance of the Carbon generation and absorption, in terms of its presence in the atmosphere.

The planning of the ways to go to handle the Greenhouse gases require data management, because the wide source of gases sources and sinks, which may vary in different clusters of time, say long term (more than 10 years) or short term (less than 1 year), depending on the geographic region and human activity present. Among the data sources, the satellites deserve a special attention once they are not land intrusive and nor exposed to the atmospheric harsh events more frequent so far [13]. As a result, the satellites present data in different profiles in time, position, and substance of interest, with the development of new materials and data processing techniques.

With the whole set of data, the evaluation of the gas emission in a specific region can be checked against a potential source of sustainable energy, while dealing with the regional energy matrix. The consensus goes towards a more diverse matrix, combining manifold sources of energy, which may consider the overall Carbon emission reduction within a certain time frame. In this more diverse energy matrix, the sustainable options may be optimized taking their best outputs and operational profile. For instance, based on the data from satellite, the government planners can have elements to plan a smooth energy source transition, migrating from the sources burning oil towards options with low carbon signature, such as wind and solar units.

Figure 4 shows the expansion trend of sustainable energy sources, to provide more flexibility to the energy matrices. It is an example of Data Analytics, with an extract of many energy sources, linking colors to its concept. The first four are represented in grey tones, probably related to not renewable explanations. The other sources have colors such as light green, yellow and blue, in a psychological approach, which makes sense with the message to be sent. It is also clear their increasing rates, mainly after 2025, regardless geopolitical boundary conditions, which can vary in a different manner as considered when collecting and analyzing the data. The option biomass appears as a natural candidate with smaller challenges to be overcome. On the other hand, the option ‘wind-offshore’ can increase but the technical burdens will demand more resources than foreseen today, which may be checked later [14].

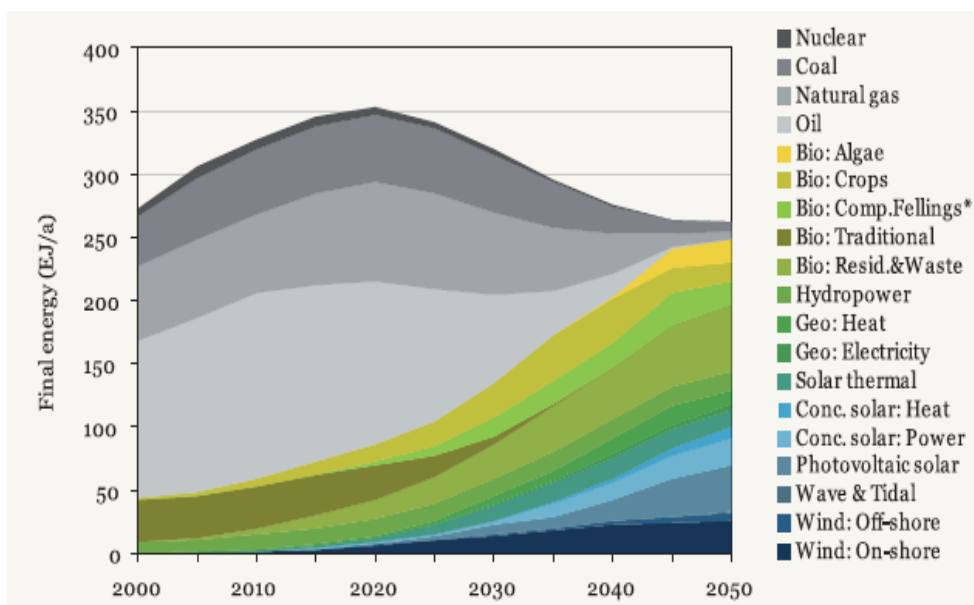


Figure 4: Sustainable energy trends

2.3 The Data Science cycle

In a DS project or experiment, there are steps that can be seen in Figure 5, in a loop cycle, with some internal feedback as well.

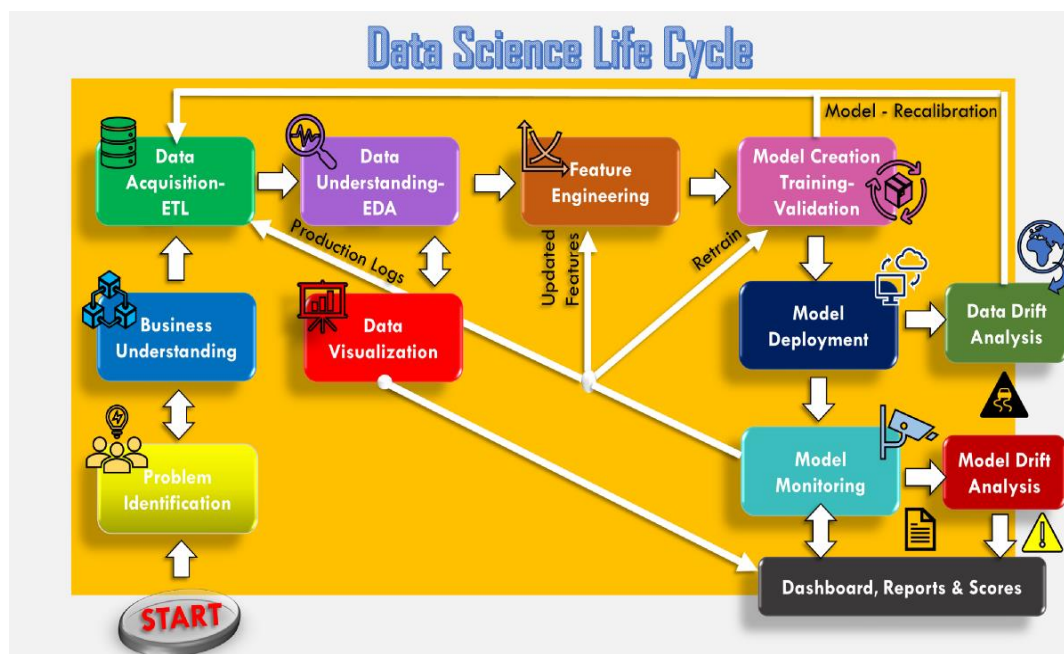


Figure 5: The general Data Science cycle [15]

As seen in other knowledge fields, the definition of the problem is the beginning of the cycle, after all it defines the needs to be fulfilled and the results to be reached. Make note the problem definition represents the hardest part of the cycle, in most of the cases, due to the unknown variables and details not considered initially. This first step has a strong

link with the business understanding because the effort will be done to enlarge or sustain business boundaries, orbiting the need to improve profit or deal conditions in long terms.

The second step into the data field is the data acquisition, with ‘ETL’ operations: extraction, treatment and launching, which will be explained as follows. Extractions deals with to identify the data sources: sensors, collectors, documents, transactions logs and others. Each source may have a proper standard to handle the data: word length, number format, sampling frequency, time scale and so on. Inside the jungle of details and fashion, the data treatment will take care of the standard needed for the data project, cleaning and even reshaping the data set: math transformations, null or incoherent data drop, and data scale adjustment, to mention some of them. In the end, the data treated will represent the useful available set to be worked, and once so, the data can be transmitted or launched to the processing area. The documentation of this treated data receives a key attention because it has details and definitions, which will be key along the data project. The documentation receives the title of ‘Metadata’, combining explanations such as: data origins; how they were gathered; references; time span; accuracy; analytical methods used to validate them; name of persons involved and how to contact them; among others.

Figure 6 shows a short view of the ETL process, and almost not further word is needed to understand its message.

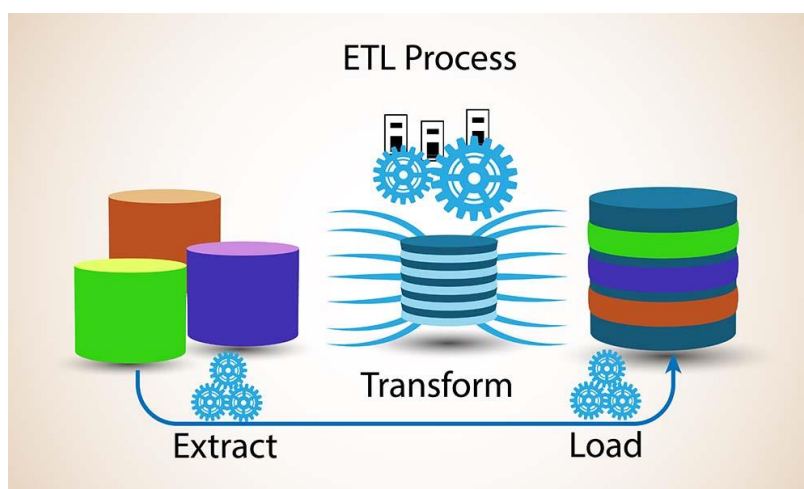


Figure 6: the ETL process in pictures [16]

The next step of the DS cycle is the exploratory data analysis, where the overall boundaries will be assessed, or in other words: how the data is spread, according to the main parameters or variables; which are the maximum and minimum values found; what kind of correlation can be detect at a first glance, although there is no guarantee the correlation makes physical or practical sense; what kind of initial clusters can be spotted; what area the statistics found: mean, standard deviation, mode, median, etc. This exploratory data analysis helps to evaluate whether the data gotten will make part of the solution of the problem to be solved, or whether the data can provide some initial insights. Several tools contribute for this exploratory analysis, and most of them deal with graphics and tables, touching what is also called ‘Data Analytics’ nowadays.

Following, the feature engineering will provide an overview of what parameters matter most, which saves times and resources in the next steps. Feature engineering also employs a set of math tools to understand how significant a certain parameter can be to explain the

data related. Equally important, the feature engineering presents the order of priority of the parameters in that explanation, which indicates a scaling of value linked to each parameter. Sometimes, to save computing resources, taking the feature engineering order of priority will ease any further work. For instance, in a certain job, the feature engineering check every one of the twenty parameters involved, and issued that only four can explain the data with more than 70% of importance. Therefore, the project may consider just these four parameters and drop the rest, saving time and money in the end.

Figure 7 shows some the overall wrap up of the feature engineering, when dealing with a set of variables or parameters of a job analysis problem, where the data about the candidate matters.

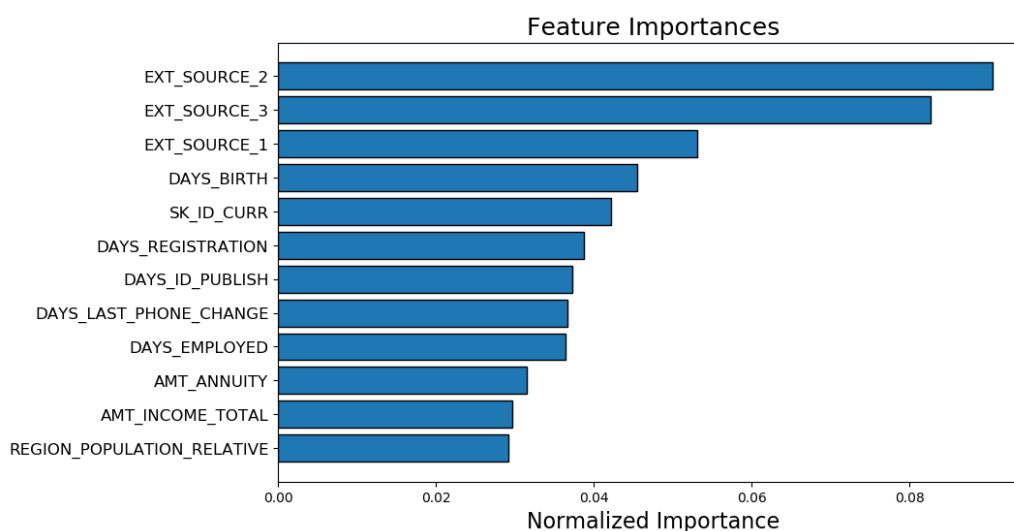


Figure 7: feature engineering relative numbers [17]

The model creation can be seen at the core of the DS cycle because it is the step where most of the math models are choose and set to aid the solution of the initial problem. Using the prior steps, the first idea of which methods may apply normally appears easily, or the ones not to be considered either. If the available data shows a linear correlation, this might be an indication that the Linear Regression approach can make some sense. On the other hand, if the exploratory data analysis does not show this kind of alignment, other models fit better, such as the Logistic Regression or the methods based on Decision Trees, or even a combination of other methos, normally called Ensemble.

Figure 4 shows the suggestion of nine important models for this step, where SVM stands for supported vector machines, KNN for k-nearest neighbors.

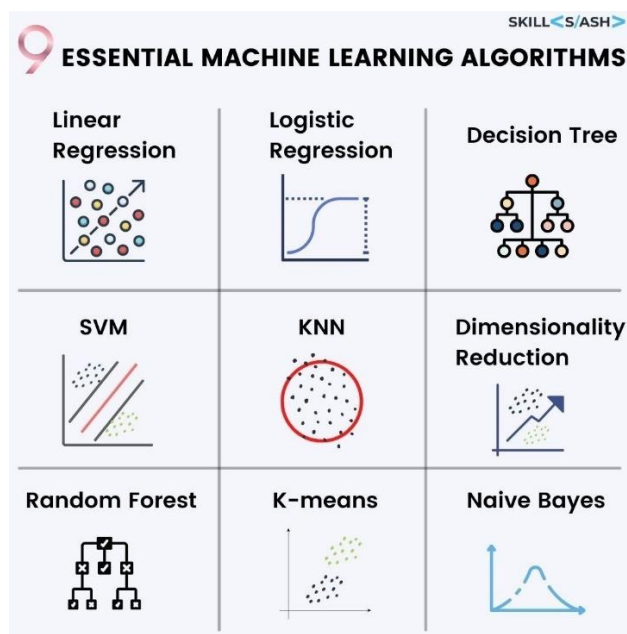


Figure 8: nine first models to be considered in DS [18]

These models are the simplest and the first choice in terms of cost and computing time, although the accuracy and performance may vary. Surprisingly, in some problems, they have a better performance when compared to more complex options as the ensemble ones. Here, the text indicates to get into them for those not so familiar with the DS subjects. Equally important, data verification and validation can be carried out during this phase, and they have methods and strategies that can be seen in Reference [19].

The term ‘model deployment’ in DS means the use of the developed model (classification, prediction etc.) with a new set of data to produce the insights, conclusions and decisions foreseen. In synthesis, it is the reason DS is for. A model can be defined as an equation and its coefficients, or a set of rules that produce the result with the input data. The deployment may consider an interactive process, handling the data pipeline in a continuous way, which demands to reshape the model based on the data time variations. In a first interaction, the best model could be the Random Forest, but as data is input sequentially, in a second round the best model could be then the Adaptive Boost (AdaBoost). Thus, the deploy will require a dynamic tuning, which quite frequent in the credit card and bank operations. Reference [20] has details of the use of the Predict Model Markup Language (PMML), which allows to build up the predict model to be deployed to compliant applications, considering ‘cloud’ options as well.

Monitoring makes part of all quality or action control step and deserves a special attention to achieve what is planned. After all, the use of real data into the models will pay back the investment, if the goals fulfill or overfill the expectancies. The only way to reach this conclusion comes from the model monitoring. The classical way of model monitoring is the performance metrics surveillance, during the software production phase. In some cases, it will be necessary to re-run the model, tuning it again, due to the new input data profile. If the performance decays more than a certain tolerance, then a whole process of re-training, re-defining and re-tuning will be carried out, and will be sent back to the DS cycle.

Reference [21] makes a clear resume of the model monitoring, which is near to the end of the DS cycle, but it can also be set anywhere in the middle, because it deals with the data presentation in two or more dimensions, easing to foresee any relation between data. It has a close relation with the prior step, model monitoring, also sharing some of the technical features: graph or trend line tools, classification plots etc. It may be part of the output package of application with DS.

The data collection starts the whole work and followed by the data treatment, when most of the mishaps and errors present in the dataset can be identified and corrected.

2.3.1 Uncertainty modeling in DS

In logistic chains, with several connections, it seems quite hard to predict all the parameters with a stunning precision and with no errors along the time. For instance, production plans considering a market expansion forecast demand to know or to grab data from different stages. This action will provide data to check each logistic step seems coherent. For instance, linear production rates do not stay for a long period, because outside factors change the logistic parameters in different frequencies and intensities. Therefore, the production rate needs to vary in time to avoid material shortage or overproduction, affecting the business plan.

Many sources indicate two benefits while attacking the data uncertainty: room for improvement and transparency. The more knowledge about the input data, the better can be the entire DS cycle, leaving space to drop data wisely, saving computing resources. Thus, treating the uncertainty remains in the field of ‘additional improvement’ way. More related to the current corporate practices, transparency is an advantage to the client, who will have a clear idea of the boundaries of his data and what may come about from the DS cycle and models. The use of DS models in the Health and Medicine applications demands transparency to aid doctors and researchers to run into acceptable conclusions.

Some authors define the uncertainty as the work with incomplete or incorrect data or even information. The uncertainty management can be done with math tools, such as probability and the Fourier Transforms, catching the time variance more accurately than visual observation or linear techniques. The three main sources of uncertainties can be nominated as:

- Imperfect or incomplete models, due to lack of experience or business knowledge.
- Noise, because the associated data without a reasonable explanation.
- incomplete data coverage, restricting the work field and bounding it wrongly.

There are two classical ways to cluster the uncertainty: aleatory or epistemic. The first one deals with the inherent noise associated with the data or environment where the data is captured. The second one aims the uncertainty of the model parameters, or how unable a model can the data relations. As seen in human activities, further training tasks can dim this uncertainty, according to the move towards the test phase. In the opposite, no additional training will help to reduce the aleatory uncertainty, only better measuring clearance will have a chance to reduce it. Suppose a system has cranky sensors but the model can explain more than 60% of the data. This situation will have more aleatory uncertainty than epistemic, and, by chance, has allowed this 60% score. While changing with better sensors, the aleatory branch can be reduced and now, the epistemic will be suitable to try the more training tasks, for instance, to reach a fairer result.

In a broad view, the application of probability and statistics aid to quantify the expected value and its variability, based on the data domain. Although both math fields were not devised for data variation strictly, they can be adjusted to detect data variance and its implications. Naïve-Bayes models consider the use of probability and account for the possible variations in a set of data, learnt from a train-test split. Reference [22] shows some of the math techniques bound for this issue, within the scope of Big Data.

Probability theories take the randomness and generally handle the statistical characteristics of the input data. Within this field, the Bayesian theory considers the interpretation of the probability, based on prior knowledge, which depends on the user background, and taken as an expression of a rational degrees of belief about uncertain values. Other tool is the “Belief function theory”, which aggregates imperfect data through an information fusion process. In addition, when dealing with classification tasks, the Classification entropy measures ambiguity among classes throughout the index of confidence, which varies between 0 and 1, and its values closer to 0 point out complete classification in only a single class. In the opposite, values closer to one indicate links among several classes.

The concepts related to the Fuzzy theory are also used, to measure uncertainty in classes, notably in natural language examples. Fuzzy logic then handles the uncertainty associated with human perception by creating a reasoning mechanism, taking a more human or subjective approach. Other tool focuses the Shannon’s entropy, which quantifies the amount of information in a variable, coming from the application of theory of communication and transmission of information. It provides a method of information quantification when it is not feasible to quantify weights. To model inaccurate or incomplete data, the Probabilistic theory and Shannon’s entropy are often employed. For ambiguous data, due its nature, the Fuzzy theory is more recommended.

Another example is the Monte Carlo dropout technique, which is linked to Artificial Intelligence networks. In short words, this method drops some intermediate layers/neurons to produce a model away from overfitting, in a model regularization fashion. The dropout layers are taken away randomly after the training phase. For the test phase, the dropout layers will be available and, to reduce the uncertainty, this procedure shall be used more than three times, dropping out the layers randomly at each time, while predicting the results. Thus, the prediction will be tuned with less uncertainty as much as the process is repeated.

Now, think other way around, instead repeating the same procedure, even with a randomly process in the middle, or aleatory steps, take multiple process in parallel, doing the quite same tasks, at the same time. When taking multiple choices, the process has a label of ‘ensemble’, but each choice taking some layers randomly. This feature also produces different model parameters trying to reduce the uncertainty. As a general recommendation, take more than 3 different choices to work in parallel, which is equivalent of the dropout of at least 3 different intermediate layers of the Monte Carlo method above.

2.3.2 Regularization techniques

In many applications of DS, a common trouble has the designation of ‘overfitting’, which means the classification or prediction model tries to fit all the data, no matter it is an outlier, or a mistake not dropped out after the exploratory data analysis (EDA). The

aftermath of the overfitting is a model unable to predict or classify extra intake data, showing poor efficiency. This situation happens with a mathematical model that cannot capture the underlying data structure, taken more parameters that can be explained by the data itself. On the other hand, underfitting means a math model lacking parameters to explain all the data. It is quite easy to be detected when using a lower order model trying to fit a data distribution with a higher order of complexity, such as a second order model (i.e., parabolic) trying to shape a third or higher order data profile (i.e. hyperbolic). The prediction performance with both overfitting or underfitting is poor in the end, and these situations need to be skipped.

It is not advisable to operate with an overfitted model. The Figure 9 presents the comparison of the training and validation phases of a prediction model. It comes from the development of a neural network where the number of epochs is represented in the horizontal axis. The math accuracy is shown in the vertical axel.

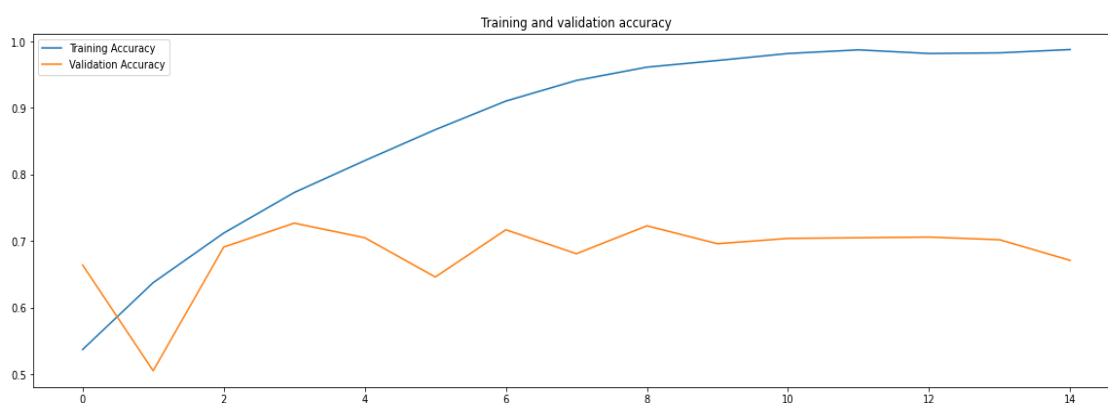


Figure 9: Comparison of model accuracy for the training and validation phases - I

When the accuracy of the training phase is higher than in the validation stage, it indicates the overfitting occurrence. The more trained is the model, the higher its accuracy. Nonetheless, the validation accuracy does not increase as the number of epochs increase, allowing the conclusion the model is not recommended. Depending on the input data and the split of training and validation, the suitable prediction model will show a different pattern. The best predict model shall have a better validation accuracy than in the training phase, because working like that shows to be prepared to produce fair prediction with data not known yet or fresh data. The expected shape said above is presented by figure 10, and regard the upward rate of the validation phase, as the accuracy increases.



Figure 10: comparison of model accuracy for the training and validation phases -- II

Regularization is a set of techniques to avoid the overfitting, using a compensation function that penalizes the parameters responsible for the fluctuations to fit all the data, without any key explanation. Among these techniques, the ‘least absolute shrinkage and selection operator’, or Lasso – L1, the Ridge – L2 and a combination of both or Elastic net are the most famous used, three options in the end. The Lasso model operates into the sum of the module of the parameters and a penalty coefficient to take the sum towards 0. The parameters without significance, or which are redundant with others, will receive weights equal to 0. In the Ridge approach, it works with the sum of each square of the parameters and a penalty coefficient that reduces the sum, but without assigning it to 0. The L1 tends to provide a sparse solution mostly, being indicated to select prediction models or to carry out the feature importance determination, but it is biased towards the non-null parameters. The L1 works well when there many parameters, in comparison to the number of data or observations. The L2 approach leads to a quite complete result model, although it is not recommended to select models, because some least germane parameter will stay in the model.

2.3.3 Missing data fill in with Bootstrap

Incomplete data reveals to be more common than expected, in the normal DS life, requiring an evaluation and decision on how to handle them. The first shoot could be to drop out these data or positions, but it brings burdens not easy to cope with sometimes. Thus, a classical way to treat the missing data is to fill in them with reasonable guessing or estimation. One at hand is to use the nearest data, or its mean, trying to maintain such a coherence, although some anomaly may be present at that point, and it will be missed unfortunately. Another quick act is to fill in with the mode, solving some statistics issue in advance, but again introducing some external bias by choice. In all of them, it normal to consider filling in with a single value, no matter its origin.

Filling the missing data may consider the following:

- a) Deleting the rows or columns where the missing data is detected, in a tradeoff between to have a better algorithm accuracy or performance against to drop other key data;
- b) Imputing values, which may be continuous or categorical, in a tradeoff between to make something to work or to input some bias;

In the first option, which may be seen as the easiest, the burdens remain with the Data Scientist, which shall take the overall boundary and initial conditions to solve the stated problem. The Exploratory Data Analysis can provide a first assessment about the data quality and whether to drop the missing or weird values will represent an odd procedure.

The second choice to fill in the missing data can choose from some options as below:

- a) the median or mean of the data of the same parameter, which harms the covariance of the data.
- b) the last or closest observation, which can also be used the very right next one. Either way, some bias will be shown, small though.
- c) the use of other algorithms, as a predictor tool, from Machine Learning or even Deep Learning. In a first glance, linear interpolation appears a reasonable candidate method, but this decision requires a better assessment.

Reference [23] presents another way to consider this data placement, with a combination of prediction methods and, in the end, many values will be available to fill in, depending on fair judgments, because the missing value may be at random, partially, or even not at all.

For this text, the meaning of bootstrap refers to a method of inferring results, for a population, from outcomes taken from a set of smaller random samples of that population, with the replacement during the sampling process, to get advantage of the aleatory process to enhance the quality of the result, demanding some computer power to carry on. Bootstrap is advised for problems with a large data population and tight schedules to reach a solution. The overall scheme is presented in Figure 11. Make note the way the sampling is carried out in many different combinations, but with the same pattern, say the number of original data at each sampling. The order does not matter at all.

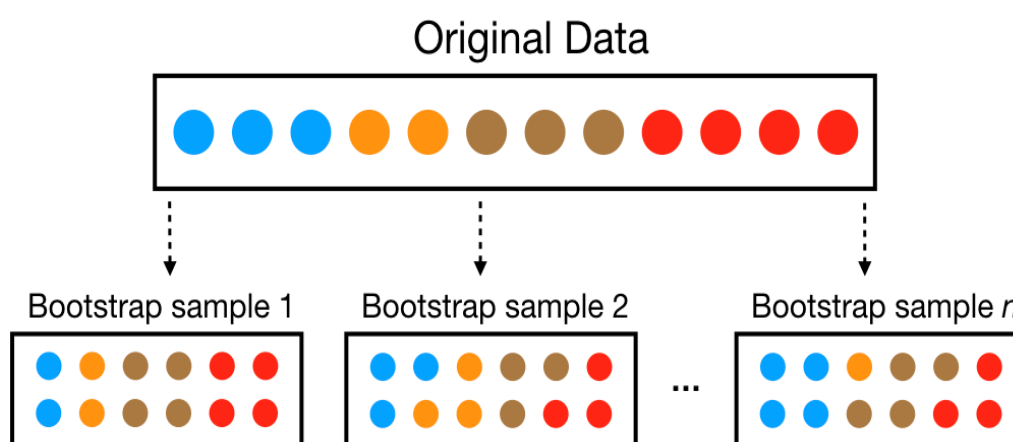


Figure 11: an overview of the bootstrap approach [23]

As a clear picture of the combined technique to fill in the missing value, using more than one single attempt, Figure 12 shows how the bootstrap can propose an answer for this problem. The simplest and common procedure would be to take out the bootstrap phase and proceed in a single line. Nonetheless, the combined sampling technique reveals a more complex process, thus demanding more computer resources and the use of a mean guessed value in the end. The term ‘multiple imputation’ can be the use of different models to predict the value, such as linear regression, decision trees and others. In the end, each option will pose a value, and the one to be chosen can come from the mean between all multiple imputation trials. The common start is the incomplete raw dataset. The ‘decision tree’ following could be any other math model to handle the problem definition and so on.

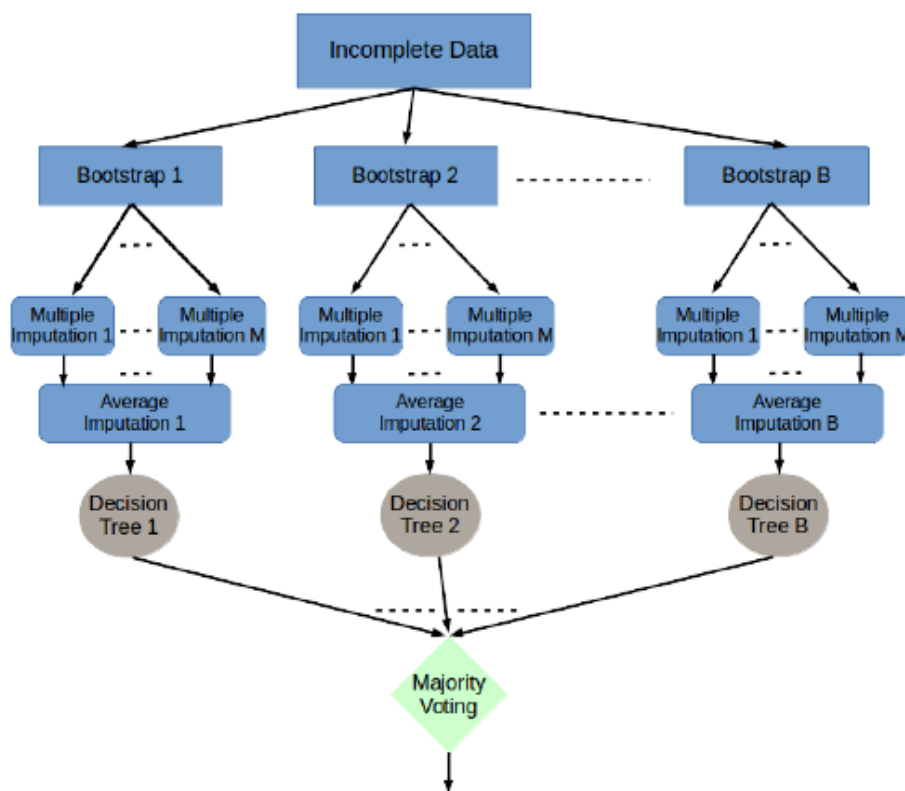


Figure 12: Bootstrap operations to full in a missing value [23]

It is recommended to note the ‘average imputation’ phases help to work on the uncertainties of the raw data, because it manages the variations randomly, preventing any strong bias to prevail. The use of decision trees also represents a pro factor because it can explain the reasons for a certain indication, even taking scarce or too diverse fields of data distribution. The majority voting stage may introduce some decision criteria based on the experience of the researchers or any specialist team. For instance, in the nuclear industry, normally the choice of a value will require a voting process 4 to 1, or take 4 options to select 1, based on how dissimilar, or not, the choice is shown against the other 3.

2.3.4 Surrogate models

The meaning of the word ‘surrogate’ is representative, substitute or equivalent for the scope of this text. Taking the physics of fluid aerodynamics, the concept of drag or pressure drop is expressed in terms of the quotient between force and area, derived from phenomena observation, modelling and the natural simplification, coming from dimension analysis, because it is simple and straight. Nonetheless, there is another way to comprehend the same fluid dynamics, now accepting a ‘surrogate model’, simpler and less accurate either, in other words is an approximation model.

In DS, for instance, the Machine Learning tools are a kind of surrogate models, because they are based on observations & measurements with plentiful of data, normally, belonging to complex phenomena, hard to see any correlation or direct explanation of the output at the beginning. While using surrogate models, the computing costs can be managed more suitably and the whole work can operate with more protective means to avoid leakage of germane data or information.

The spline of the surrogate models lays on the data input, which can come from natural observations or aleatory simulation. The more diverse and numerous the data, the better the surrogate model can provide the output needed. The initial Data Exploratory Analysis has a key value to decipher any quick correlation inside the input data, allowing a better tuning of the surrogate model as expected. After that, the next step is to construct the parameter features that will make part of the model, also known as feature engineering. Some parameters come with the input data, but other may be lined up to increase the chance of a fair result. These featured engineering parameters may be a combination of the natural parameters, or an implementation of other parameters based on the study and knowledge related to the data, requiring previous skills to carry on wisely. For instance, one input data transformation needed sometime is the “log transformation”, to reshape the input numbers in a better way to understand the phenomena or natural correlation, caring to make the inverse transform when checking the output for the answer of the problem.

The surrogate will be a set of equations or business rules that take the input data and can provide an answer, based on the method reasoning to do so. The equations have coefficients which after combination with the input features provide the suggested answer (prediction or classification).

As a practical example of surrogate models, reference [24] works with surrogate models to predict the net present value from different oil extraction options, using data from oil wells already in use. They try to maximize the NPV by choosing the number of oil wells and their spatial coordinates according to the maximum oil production. The input data is complex because it takes manifold parameters from Geology, Chemistry, Logistics, Environmental rules and so on, a situation well suited to apply the Machine Learning algorithms. The reference considered the following regression (prediction) methods: support vector machine (SVM), gradient tree boosting (GTB), genetic algorithm (GA), multi-layer perceptron (MLP), elastic net (EN) which takes care of the order regularizations L1 and L2 altogether, and the K nearest neighbor (KNN), which was recommend in the end to predict the NPV.

2.3.5 Lower order models

As a general approach in any scientific field, it is better to work with models or problems with the lowest orders of complexity or magnitude. One benefit is the lower processing or computing costs, besides a faster answer and agility in the decision phase. In DS, the complexity of a math model deals with the number of parameters, for instance, to be handled. Nevertheless, the order reduction shall provide the quite same result as the high order models.

The common procedure is to take the measurements or data and derive the prediction models, which will have a high order for instance. Secondly, these models will be transformed into the low order models, by math transforms, such as projection, optimal Hankel norm approximation, or Krylov methods. More recently, a new suggestion has emerged as ‘data driven order reduction’, which does not need to go over a high dimension model firstly, as it can be seen with more details in reference [25].

Specifically in the DS field, the Principal Component Analysis (PCA) is a key technique to evaluate the dimensions that can be scrapped to reduce the order of the problem and related math model. The concept is to produce components by a combination of the

original parameters, in a manner that the components made will not covary among themselves, as they were orthogonal as possible or kind of. In terms of analytical geometry, the principal component will be a direction that explains the maximum amount of variance of the parameters. Therefore, the new direction or axis takes most of the information from the parameters and will be like axels where the difference between parameters will be detected easily. After the PCA, the variables or parameters will not have the original one, but a wise combination of them. The key trade-off to manage is some decrease of accuracy when reducing the problem order.

PCA requires a set of operations, normally:

- a) Standardization – the initial variables will have their values into a standard range, normally between 0 and 1, or -1 and 1, taking care with the diversity of range sets, using the standard deviation and the mean value of the variables or parameters.
- b) Matrix of covariance – this step is needed because variables may a high covariance and have redundant information as well. The covariance matrix is a numerical indication to check this dual condition.
- c) Eigenvectors and eigenvalues of the covariance matrix – these concepts are useful to determine the principal components, which will be not correlated among themselves, based on the eigenvectors and eigenvalues. The first principal component will retain most of the information from the data, and the second the most important information also remaining and so on. The principal components shall behave like orthogonal ones. The higher the eigenvalue, the more principal the component is.
- d) Determination of the features of the vectors – the dimensions of a matrix made of the vectors will be the principal components to be chosen, a decision that depends on the order of magnitude to accept. The columns of the matrix will be the vectors.
- e) Redistribution of the original data into the new axis or components – this step is needed to shape the data into the new dimension axis or references, with a smaller order of magnitude and most of the original information retained.

In a simple graph explanation, Figure 13 shows the concepts above in a 2D approach. The initial variables were draw in black, referring to the original parameters. Searching for new orientations axis, following the steps right above, the new dimensions will be the red and blue options, where the red has retained more information than the blue, in a way to reduce the order of the model. In this example, the order remains the same, but it is easier to follow the math steps.

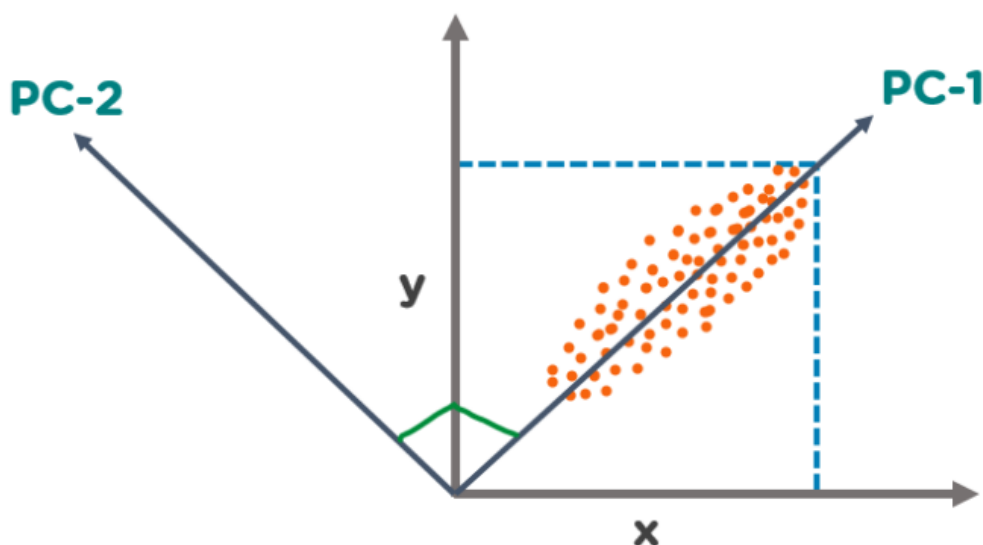


Figure 13: 2D example of Principal Component Analysis [26]

Note that the new variables orientation (red and blue) will be a combination of the black ones (x and y), requiring attention when analyzing the final prediction results, for instance. It may be needed a recovery transformation to stick to the original problem definition.

2.3.6 Kriging techniques for interpolation

As an example of a practical use of DS interpolation or data fitting, the term ‘Kriging’ refers to a specific math tool related to geostatistical procedure to adjust a surface to a 3D pattern: two localization variables (x , y , latitude, longitude etc.) and a third variable related to an interest value (z , outcome of function etc.). The technique is also known as Wiener-Kolgomorov prediction. It takes a limited set of observed data to make inference of a large data space. It condenses complex math operations to fit data into a functional pattern, with direct applications in Environmental and Health Sciences, Chemical Engineering and others, which deals with spatial distributions, particle dispersion inside variant media, among others, demanding a high computer capacity to work with data varying in time and space. The original motivation for the development of this technique was the gold dispersion prediction in South Africa as an oddity.

The main concept remains to predict a value based on average-data in the surroundings of the region of interest. An interesting way to see this technique is like a subset of the Bayesian inference. It commences with a distribution over functions. In geostatistical models, data comes from random events, although it may not be in this way really. This basic assumption refers to a spatial inference for locations without measured data, computing also the uncertainty associated to the prediction. The first task comprises to build up a random process capable to describe the measured data. The solution in the geostatistical framework focuses the assumption of diverse degrees of stationarity related to the random function, to infer a set of statistic values.

While comparing to Linear Regression or Gaussian decays, the Kriging technique uses spatial correlation between sampled data to predict the values in the completely spatial field, considering now the empirical measurements, rather than on a preconceived spatial

distribution. As a bonus, the Kriging technique produces uncertainty estimates of the neighborhoods of each predicted value. The closer the position to the interpolation point, the greater the weight for the spatial distribution. The farther positions receive smaller weights consequently. To prevent bias in the prediction phase, cluster points also receive smaller weights, because they do not add much information as single points. The Kriging technique tries to minimize the prediction error for a specific point, making its prediction at the measured points equals to the measured values.

The Kriging technique uses variograms extensively, which are a visual taking of the covariance between sample points. For each pair of sample points, the semi-variance, or the half mean-squared difference of their values, is associated to the distance between them. The actual measured points make part of a variogram labeled as ‘experimental’, or the ‘reality’. The ‘model variogram’ is the distributional model that best adjust to the data, considering a set of math functions allowed previously, such as linear, spherical or power models, which includes the exponential too. Figures 14 and 15 show variograms and the concepts above. Note the vertical axel represents the statistical parameter (variance, semi-variance), as the horizontal axis is the distance between the points.

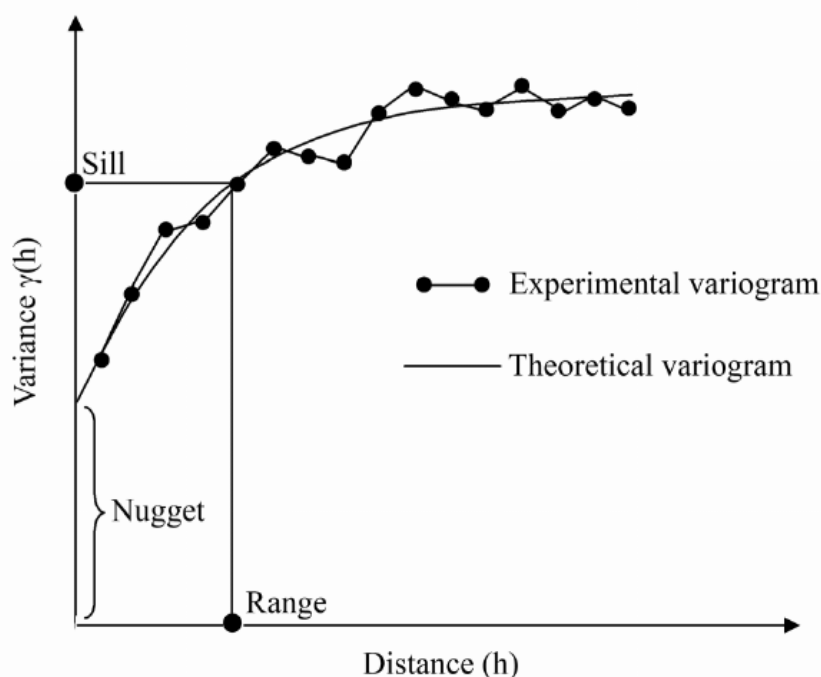


Figure 14: 2D variogram linked to the Kriging technique [27]

As shown right above, the experimental variogram deals with the measured points and are linked by a straight line, showing a discrete profile, and the theoretical variogram is a continuous line, generated by math functions prior selected. The term ‘Nugget’ means the initial predicted value for the distance equals to 0 or the ‘small-scale’ variability of the data. If the variance is low, the correlation between the points is high, and the opposite is true, checked at the upper part of the variogram, where the continuous line becomes almost horizontal. The Sill means the maximum variability between pairs of points. Conceptually, the Kriging model has the values of the measured points, although it has worked with a sample set of these points.

As seen below, different math models produce different theoretical models, which tries to reproduce the experimental model in the best way. The spherical and circular equations tend to occupy the upper left part of the graph, while the linear and exponential stay closer to the natural diagonal, or inclination of 45° . The Gaussian model has an odd shape, with an inflection point somewhere the middle of the Distance range, depending on the problem and data available.

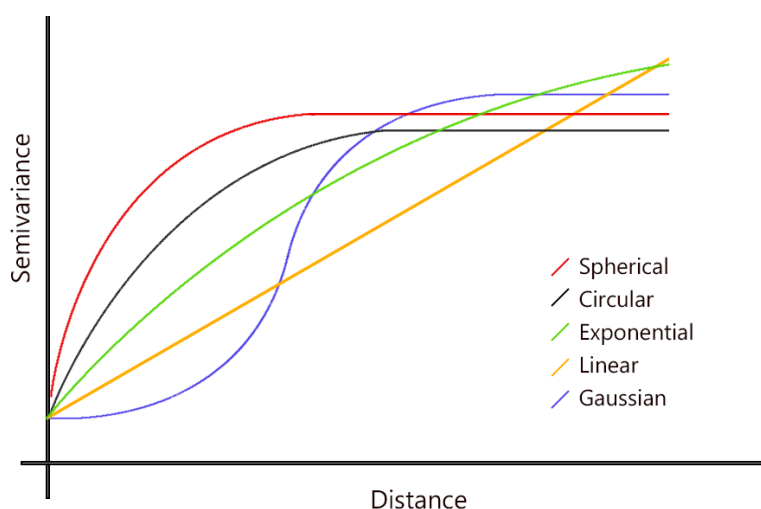


Figure 15: Different math models allowed in the Kriging model [27]

2.3.7 Bayesian techniques for data pipelines

As already mentioned in along this text, a prediction or classification model may require a constant update when dealing with a data pipeline, which is quite common with logistic and energy efficiency applications. Data comes from several sources or measuring instruments, following the sample strategies, and stored into databanks or data lakes to fill in DS tools. This update can be done by several techniques and schemes, depending on the experience of the DS team and their prior experiences, recent data recorded, and the information derived from them. Thus, the data background has a strong influence on the work to carry out, throughout the learning involved from different directions.

The first natural choice to a data update focuses the frequentist approach, which in a short essence deals with the question: Which model, or set of parameter values, are we most certain is the true model? Frequentist statistics is based on hypothetically resampling the data from an underlying true model, using the concept of likelihood, straightforward [28]. Nonetheless, this approach does not consider any previous experience or wise observation of hidden details eventually. Thus, another way to update the DS model may take concepts from the Bayes Theorem and its aftermaths. In this sense, our belief is taken into consideration, which sticks to the most common behavior, although it represents a hard technical step, due to the variability involved. For instance, referring to past data, the rate of sick people in a town may vary with the local temperature and season, with figures of 100 sick people for every 5°C decrease in the winter. These numbers may not be so precise as stated, requiring a distribution approach to be more consistent. Let take a Gaussian

distribution (bell shape) to set the prior knowledge, setting an average and standard deviation. It is a straight application of the probability distribution functions or 'pdf'. Using a high mean and small standard deviation means the prior knowledge reflects a strong certainty about the data or situation in case. Being humbler, however, knowing that recent quick changes in the application of vaccines, let take a smaller average and larger standard deviation, reshaping the bell outline of the Gaussian distribution assigned to this case or problem. After all, the parameters can be seen as random variables and the pdf concept applies. The propose above to use Gaussian distribution on the parameter's estimation represents a way to regularize the input data, taking care of outliers as an example. Thus, it is the application of the Bayes technique, setting something based on the previous knowledge or evidence of an occurrence related to problem. Equally important, it is a iterative process, allowing permanent reviews of guessing and conclusions, based on the outcomes, in a closed or feedback loop.

Hierarchical models are a standard Bayes approach for data updating, using a scheme of data clustering following a knowledge path. The advantage of this way is to consider different data sources, with high or low data quantities, and dealing wisely with the outliers, because the Gaussian distributions shapes will take that math.

Another type of subject inside the data update is the measurement of the accuracy. Sometimes a set of conditions arise [28]:

- a) There is no data enough to carry on the training phase, typical in problems of shortage of data. This scenario gets worse with the need to use cross validation.
- b) The time involved in the training phase may be too long.
- c) When using an unsupervised technique, like k-means, in classification for instance, the accuracy may not be easy to compute, because there is no labeled data to detect the learning improvement.

As a result, the application of the Bayesian method in the data update will be made to the point where the model accuracy does not change significantly. Also, the Bayesian method handles the errors linked to the measurement of the data with a probabilistic way, more suitable to be treated mathematically.

The Bayesian parameters are reviewed based on the new data input, changing the model set previously, in a continuous way. The update of the model aims to reduce the distance between the forecast results and the new measured data, and it is worth noting the model update time interval must be shorter than the data input process, otherwise the model will be always not fitted to handle the problem. This gap time will require a piece of money in the whole DS cycle.

2.4 The application of Data Science in engineering

Many DS applications have been closer to the common understanding recently, with the diffusion of algorithms linked to the prediction of human behavior, sales campaigns, and logistic planning. In this chapter, the use of DS will focus the renewable energy field. For instance, the application of the math algorithms and digital computing into the analysis of materials behavior provide more insights about the interactions between particles, in different size scales, which affects the capacity of the material to transfer heat, to conduct electricity or to recover from strain loads.

The two large application areas of DS in engineering are classification and prediction. The first area deals with to recognize in what kind of groups, patterns, or clusters a set of data can be assigned. In this task, the group identification comes from the alignment of several parameters, as close as possible, of common characteristics among the data sampled.

2.4.1 DS and solar and wind power sources

In this section, some practical applications of DS covering types of sustainable energy will be addressed. With a top-down approach, Reference [29] presents a scheme on how DS can be applied to propose or indicate a set of sustainable energy sources can be on the radar of energy planners or policy makers. It starts with data from the sources, towards a set of parameters, engineered or not, in a big data scale. Then, depending to the class of the problem, clustering, classification, or regression, or both at the same time, can be done using the vast DS toolbox, providing more than straight presentation from Data Analytics.

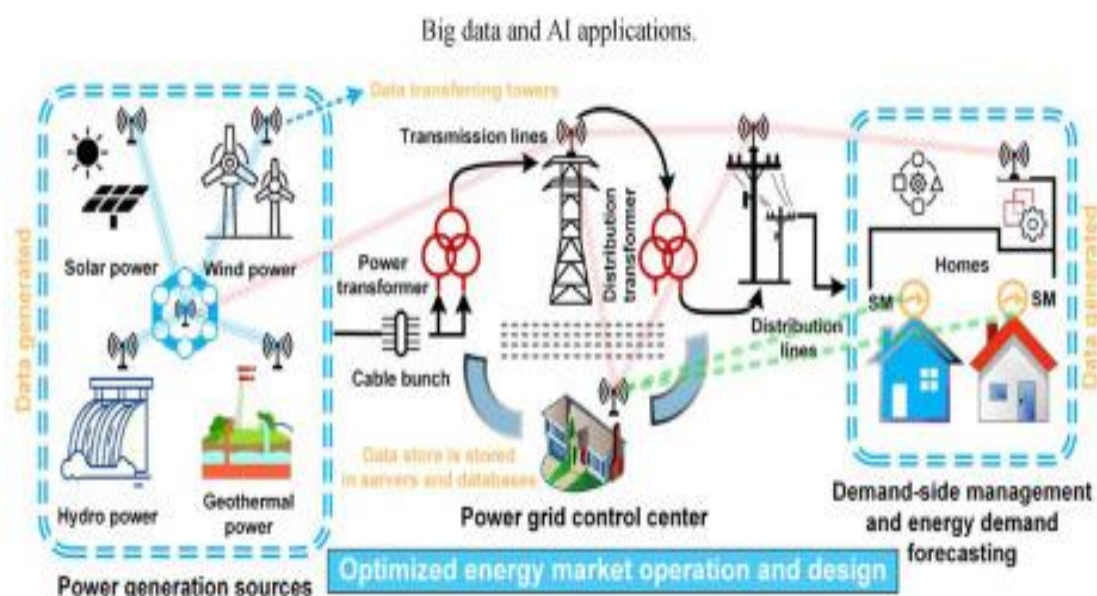


Figure 16: General scheme for DS application into renewable/sustainable energy field. [29]

As the sustainable energies make more common to our routines nowadays, the combined use of different sources makes all the sense to optimize the investment. This is not new in the energy market, but the difference calls because the decision process comes closer to the users than it used to be, with large generator and distributor groups taking the wheel for most of the time.

Reference [30] worked on a combined model with solar and wind energy generators, quite easy to have in most of the houses and buildings near the shorelines. Both exhibits energy production transients based on the local and climate boundary conditions, such as landscape details, latitude, and others. For instance, solar radiance varies along time and latitude, due to astronomical known reasons, but the efficiency of the solar panels depends also on the local temperature. Above 40 °C, the photovoltaic efficiency of the current solar panels drops drastically, in factors 5:1 or worse than that, depending on installation

and maintenance practices. Figure 17 shows the technical hardware involved, where the solar option is labeled by salmon color and the wind option by green tone.

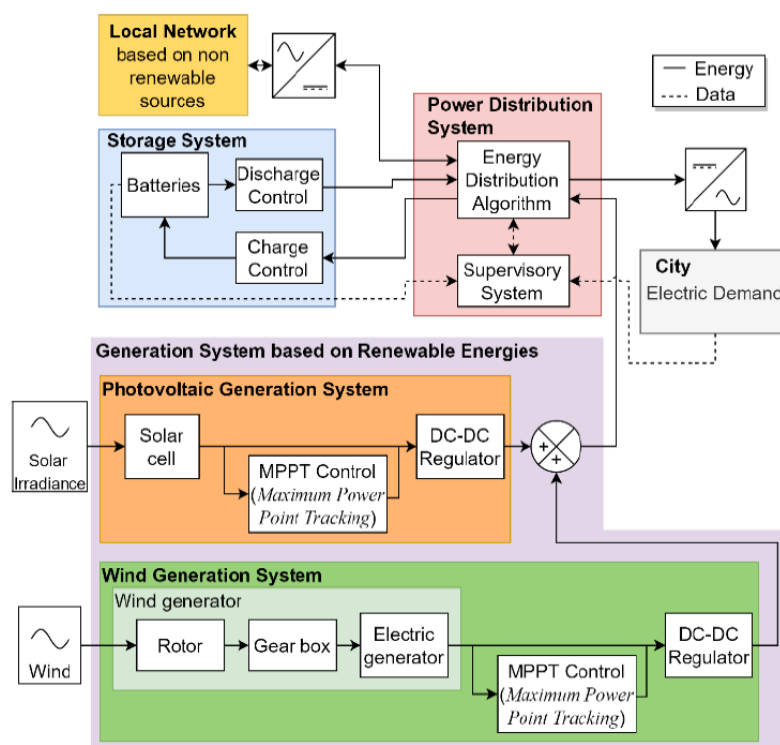


Figure 17: Wind and solar power sources and DS [30]

Both energy inputs are variable, signed by the wave form on the left bottom. The thinking core of the proposed system is the Energy Distribution Algorithm (purple box), which will compute a set of data to decide how to do with the energy stored, taking at least two possible decisions: supply or not the local network, with has equivalent sustainable energy sources, or supply or not the city, which is a bigger energy sink than the first one. Data comes along the energy to be managed, and that is the key difference from the conventional energy grid available in most cities. With DS, it is possible to use its tools to bound the decision along different time intervals, to maximize profits or cut unreasonable costs, based on the electric efficiency of each option or the combination of both [31].

2.4.2 DS and geothermal energy

Geothermal energy has gained the screens more recently, although it is well known for centuries, because it uses the natural heat source from the Earth's core and can be linked to the factors bound with the occurrence of earthquakes. This energy source does not require sophisticated technology to be used, long pipelines, heat exchangers and industrial instrumentation, available from the oil industry mostly. Depending on the use, it can be set anywhere, but the economics play a key decision factor as usual. That is the point where DS comes in.

Reference [32] refers to a practical case to identify the best locations to build up geothermal stations, based on earth big data, from different sources in time and location. The general result of the research is presented in Figure 18.

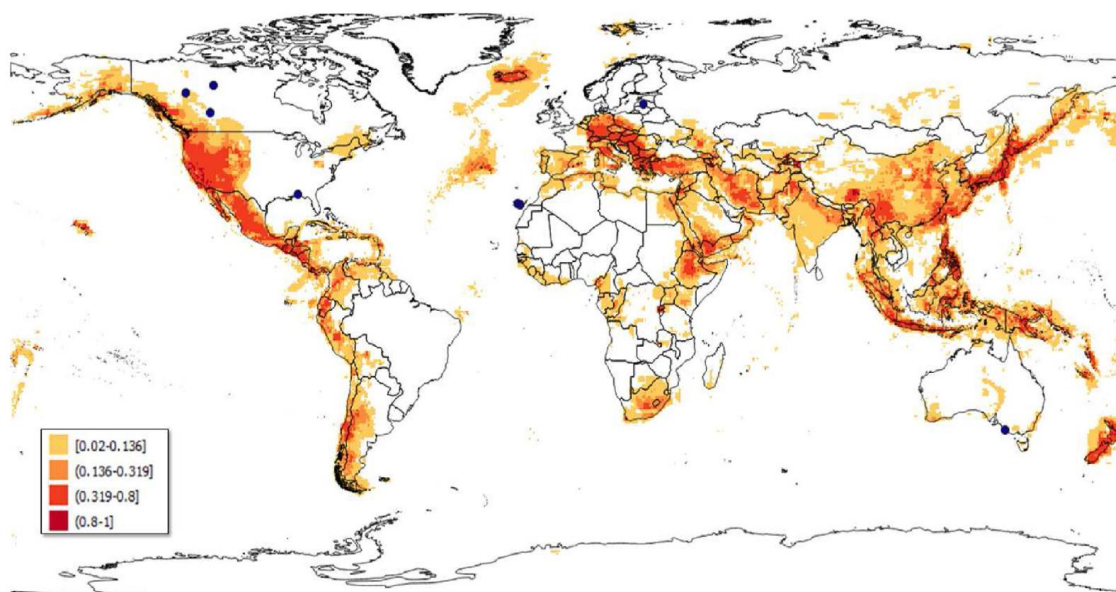


Figure 18: Optimal geothermal distribution for energy management [32]

The higher the number or reddish, the better potential to establish a geothermal energy station. This survey came from the application of DS, taking data from geological sources, time and geographic coordinates, as well as data from satellites of CO₂ distribution, earthquakes density, depth, magnitudes, global heat flow, groundwater sources, surfaced air temperatures, distance from land lines, among others. The model used to process the data sources, after engineered features, was the Maximum Entropy, which states the best way to explain a dataset will be the one with the highest entropy according to constraints with a prior knowledge. For instance, the current geothermal plants are a well-known constraint, also the new geological anomalies produced by recent volcano eruptions etc.

Taking the data pipelines from survey installations or instrumented field systems, or even mobile, the evaluation models can be updated constantly, refreshing the model parameters as discussed in the items above.

2.4.3 DS and nuclear energy

In the nuclear industry, the application of safeguards has taken more importance to prevent proliferation of nuclear materials and technology. It is necessary to avoid the use of weapons of mass destruction, such as thermonuclear or even dirty bombs, which can leave long time consequences for people and their generations.

One of the safeguards instruments is the frontier inspection of goods being transported. Some industrial nuclear hardware has its profile very close to ordinary ones, such as pumps, valves, or heat exchangers, although constructed under restricted materials and refined techniques. Thus, the customs agents must have a special training to be able to recognize whether an industrial good may be controlled by the safeguard regimes adopted. One way to help those agents considers the developing of an intelligent system, trained with certified data, and based on AI, to provide a first guidance whether the good should be controlled or not. In brief words, it is the application of NPL, in this case digital images.

Reference [33] presents a simple option focusing the most important nuclear hardware associated both to the fuel cycle and for nuclear power reactors. The more the AI systems is trained, the better will be the chances for a fair alarm to the customs agents, paying attention for the overfitting indicated the items above.

Another application of DS in the nuclear field deals with data filling regarding the operational performance, as pointed out by Reference [34], helping to detect problems in the operation of a nuclear system or plant. The target here are the operational parameters, which may have missing values in their records, due to a faulty instrumentation, lack of communications connections or even human errors, and all combined, as it was detected in the Chernobyl's accident in 1986. Unfortunately, at that time, the DS was a kind of starting and would help to prevent the horrible outcomes. In their studies, the authors considered the options of the least squares support vector machines, the K-means clustering algorithm based on a noise algorithm. In the end, the reference uses wavelet analysis and the Mahalanobis distance, for data work, when dealing with analog sources for instance. The algorithm proposed by the authors works better than the mean procedure for instance, because it treats the oddities with more detailed scrutiny.

2.4.4 DS and wind power: structural monitoring application

The occurrence of winds on Earth comes from the heat transfer in the atmosphere, influenced by the incident heat flux from the Sun mostly. Thus, the prediction of wind can be based on physical models of heat transfer of large masses of gases, assuming the air as a classical gas in most of the situations.

However, the number of variables and their crossing influences make this problem complex to be solved, even with numeric procedures and computers, although feasible with some accuracy and many hypotheses. The other way to predict the wind occurrence is data driven, based on large quantities of accumulated data from real observations and following the DS cycle presented at the beginning of this text.

One of the applications of the data drive approach aims the structural reliability of wind towers used in the wind energy sector. In a few words, the structural tower is a vertical beam, with a concentrated load on the top and with a connection on the ground without any freedom degree. The dynamic load on the structure comes from the wind itself on the propeller blades, on the tower and from the ground to the tower, in a first definition, together with the static loads, such as weight and ground reactions. The loads will stress the structure material and strain will be detected, providing a data source on how the dislocations and other geometry changes occur in the material lattice.

This strain field is the data input for DS applications. Figure 19 shows the material distortion or singularity will be digitally treated to be a set of numbers, as an example of NLP application. Particularly, the arrangement has a standard AI framework to produce the dataset [35].

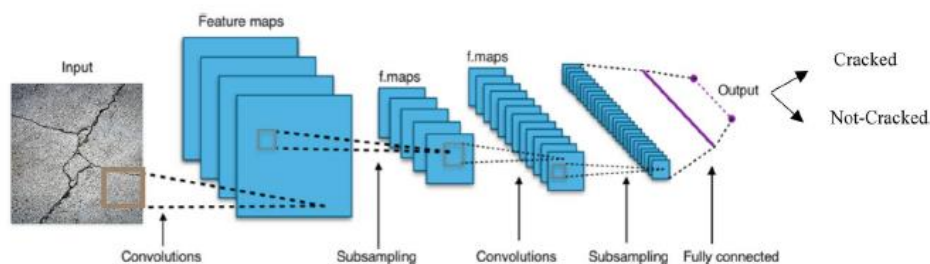


Figure 19: digital matrix data treatment for DS applications [35]

With the structural sensors, checking the strain regime of the tower, mainly at positions with stress concentration (distortion or change of geometries etc.), another set of sensors work to measure the wind characteristics, such as speed, direction, and variation rates of them, among others associated with the electric power production and transmission. Combining the data, the tower has a dataset associated to structure behavior, which must be within a safety standard envelope, associated to the concept of structural reliability.

Reference [36] presents an example of the use of AI to handle the situation when a blade of a wind turbine turns to be unbalanced, indicating the points with faults, suitable to be also useful in predictive maintenance programs. It is the same kind of application found in the aerospace industry more recently, improving the quality assurance in most cases.

Before the WWII, the term ‘reliability’ was associated to any system that had produced the same output for a time interval, close to the concept ‘repeatability’. Nowadays, with the statistics advances, the concept of reliability has incorporated probability distribution functions (pdf) and can be framed as the probability of a system to perform its functions during a defined time interval. Connecting DS into this field, the prediction of structural status, according to dynamic loads, materials degradation, wind variations, can use manifold ML methods, starting with the clustering of the engineering factors and without using any specific instructions. ML looks like a surrogate model of the wind tower structural response, submitted to load conditions (stochastic or deterministic).

Classification can be set to indicate whether a set of data will associate to a safe state. On the other hand, ML prediction can provide a parameter future value, or index, of a structural component under a certain function analysis. Both cases can manage the ML techniques to do so, using different definitions of “distance”, as the Manhattan distance for clustering.

As outlined above, a data drive method using ML demands a dataset with input parameters and one output at least, and with this arrangement, algorithms can be set to make predictions, which can be supervised or non-supervised. Also relevant, DL can also be set to perform the same tasks, with perceptrons and learning layers for instance. Structurally speaking, the prediction aims to forecast a limit state or boundary, throughout the concept of Limit State Functions (LSF), which may consider physical models or not. As already known from other structure applications, the use of ML in the wind power structures can be challenging, due to the need to have data from reliable measuring systems and from lab tests. Summing up, the use of ML may face a rare event, with few available data, nearing the situation of high dimensionality, or more parameters than occurrences, requiring special math techniques. The selection of the parameters demands a closer analysis in the problem definition as well, which will deal with a time-series in most of the cases.

The ML models used for the application focused here are shown in Figure 20. Noteworthy that many can be used together, in different phases of the DS cycle, say in classification and prediction, in this order, to provide an insight about the need. References and summarize them for the structural applications highlighted.

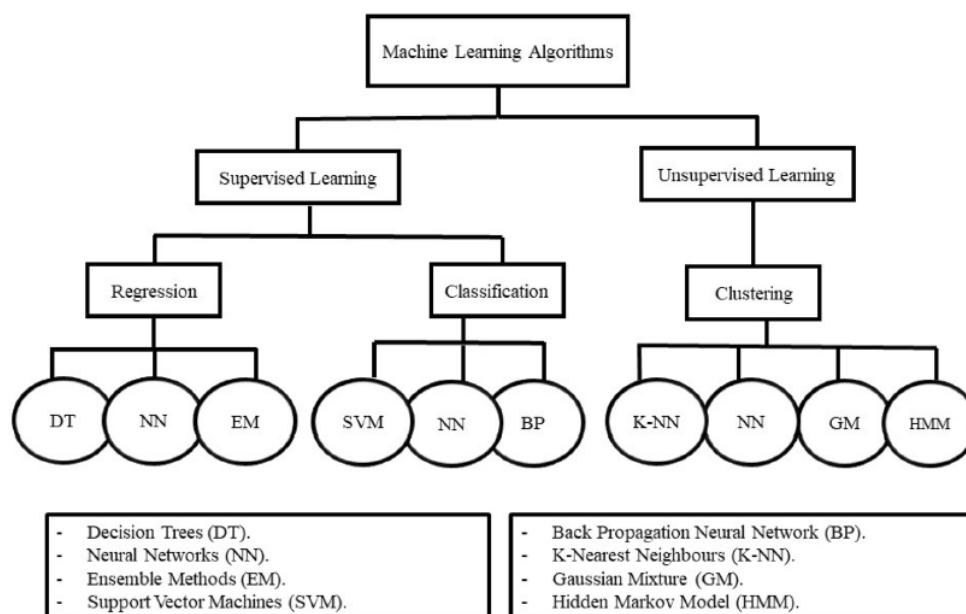


Figure 20: ML models aimed for structural monitoring application [36]

In terms of AI, the feedforward Multi-Layer Perceptron is commonly used in structural reliability studies, with four layers initially, which may vary based on the problem of interest, but as a matter of fact, the number of layers does not need to be large indeed. The Adaptive Sampling Methods can be used with MLP, tackling the data training phase, refining the AI model till a specific accuracy is achieved. The selection of the training points is based on the training selected previously and the associated performance index, leading to a lower computing cost. Reference [37] has a fair literature review of the application of AI in the structural reliability.

Looking into the wind power practical applications, Figures 21 and 22 show the overall instrumentation used with DS to monitor the performance of a wind power electric generator in an offshore application, that is more complex than on land, and the parallel with a structural bridge. Those systems have similarities on providing field data to check the structural reliability of their own cases [38].



Figure 21: wind power monitoring system linked to DS [38]

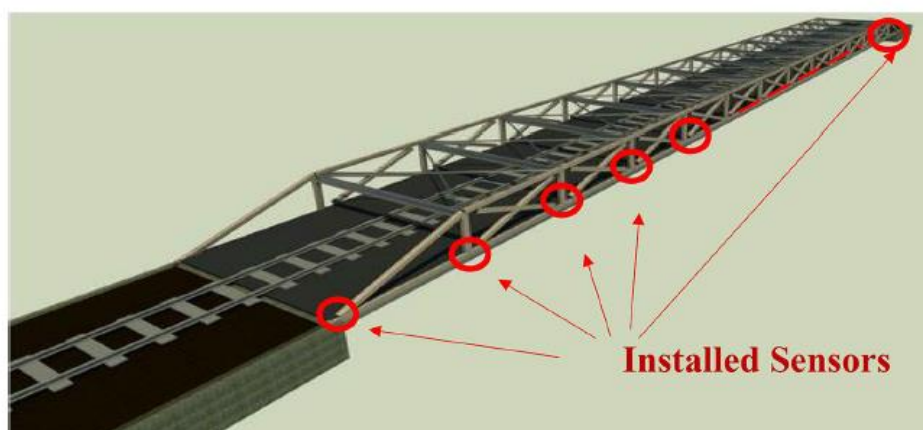


Figure 22: Structural bridge monitoring system linked to DS [38]

In both cases, data is gathered from a set of sensors, with a sampling frequency, into a data bus or pipeline, building up a timeline or time series, to be treated with DS algorithms. As seen in other engineering applications, this scheme will help other assessments and to feed up the industry with feedback from field applications, associated to the design, testing and improvement of new advanced materials and construction techniques, managing fuzzy phenomena as delamination of carbon fiber structures under dynamic loads and solar radiation together. The DS cycle will provide insights for a data driven decision on the choice of the resins, for instance, or the way the carbon fiber thread will be designed and carried out, nearer to the statistical control and structure health monitoring application, as shown in Figure 23.

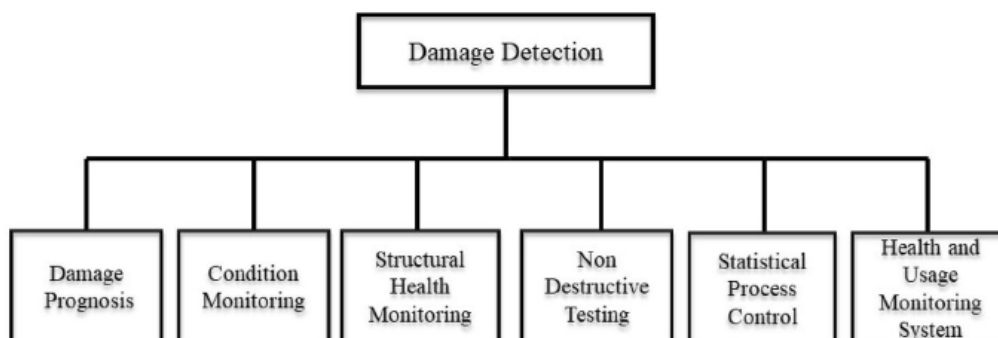


Figure 23: Damage detection fields related to advanced structures [38]

The DS application will require the data confirmation, data reliability and other data quality features, at a high level, because the data driven decision may have impact in the cost and performance evaluation of the wind power structure and, moreover, also have links with disputes among constructors and operators in case of incidents or even accidents and damages.

3 Concluding Remarks

In this chapter, a general view of the ship design was presented, covering its phases and relevant technical issues. The characterization of the ship dimensions and systems, along design drawings and materials application, was described in a progressive way. As the sections of the chapter were presented, the engineering materials field had been expanded in terms of technical figures and details. The focus was to provide a spark to the reader about the variety of topics around the ship materials selection and its evolution, without the pretense explore all the fields in deep. The data presented came from open sources, intending to provide an order of magnitude of the reasoning discussed along the text, and they may be seen as mean values. All topics here shown are under continuous development, mainly the nuclear examples, and more recently the application of Machine Learning and Artificial Intelligence. As a high potential subject for technical improvement, the environmental impact reduction with new fuels and propulsion systems has its place to catch the general attention.

Acknowledgements

The author thanks University of São Paulo (USP) and Paulista University (UNIP) for all their academic and technical support.

References

- [1] <https://www.analyticsinsight.net/top-10-big-data-disruption-coming-in-in-2022/>, accessed on June, 30th 2022.

- [2] Desclaux, Gilles. (2018). Big Data & Artificial Intelligence for Military Decision Making.
- [3] <https://www.iai.it/sites/default/files/978195445000.pdf>, accessed on June, 30th 2022.
- [4] Ferster, Colin & Coops, Nicholas. (2015). Integrating volunteered smartphone data with multispectral remote sensing to estimate forest fuels. *International Journal of Digital Earth*. 9. 1-26. 10.1080/17538947.2014.100286.
- [5] <https://people.smp.uq.edu.au/DirkKroese/DSML/DSML.pdf>, accessed on June, 20th 2022.
- [6] Chen, Jiangping & Ayala, Brenda & Alsmadi, Duha. (2018). Fundamentals of Data Science for Future Data Scientists. 10.1201/9781315209555-6.
- [7] Mcdonald, Lynn. (2013). Florence Nightingale, statistics and the Crimean War. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*. 177. 10.1111/rssa.12026.
- [8] https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1063041/ddat-playbook-march-2022.pdf, accessed on June 20th, 2022.
- [9] <https://doi.org/10.1002/asi.4630270213>, accessed on June, 20th 2022.
- [10] <http://www2.math.uu.se/~thulin/mm/breiman.pdf>, accessed on June, 17th 2022.
- [11] Dineva, Snejana & Nedeva, Veselina. (2018). Expanding Web and Innovation Skills for 21st Century.
- [12] Kosse, Pascal & Blomberg, Kati & Mikola, Anna & Heinonen, Mari & Kuokkanen, Anna & Lange, Ruben-Laurids & Lübken, Manfred & Wichern, Marc. (2016). Climate change and green- house gas emissions within the context of urban waste water management. *WaterSolutions*. 89-94.
- [13] https://www.ipcc.ch/site/assets/uploads/2018/05/ipcc_wg_I_1992_suppl_report_section_a1.pdf, accessed on June, 16th 2022.
- [14] <https://www.eia.gov/todayinenergy/detail.php?id=41533>, accessed on June, 17th 2022.
- [15] <https://morioh.com/p/9ba3ed05fad1>, accessed on June, 15th 2022.
- [16] <https://www.dreamstime.com/illustration/etl-process.html>, accessed on June, 15th 2022.
- [17] Zheng, Huan, and Yanghui Wu. 2019. "A XGBoost Model with Weather Similarity Analysis and Feature Engineering for Short-Term Wind Power Forecasting" *Applied Sciences* 9, no. 15: 3019. <https://doi.org/10.3390/app9153019>.
- [18] <https://twitter.com/datasciencedojo/status/1404167692173692931>, accessed on June, 30th 2022.
- [19] <https://www.analyticsvidhya.com/blog/2021/03/data-validation-and-data-verification-from-dictionary-to-machine-learning/>, accessed on June, 15th 2022.
- [20] <https://www.ibm.com/docs/en/db2/11.1?topic=analytics-pmml-markup-language-data-mining>, accessed on June, 20th 2022.

- [21] <https://www.oracle.com/a/ocom/docs/data-science-lifecycle-ebook.pdf>, accessed on June, 15th 2022.
- [22] <https://www.webpages.uidaho.edu/~stevel/517/The%20Data%20Science%20Design%20Manual.pdf>, accessed on June 30th 2022.
- [23] Aljuaid, Tahani & Sasi, Sreela. (2016). Proper imputation techniques for missing values in data sets. 1-5. 10.1109/ICDSE.2016.7823957.
- [24] Yetunde M. Aladeitan, Akeem O. Arinkoola, Okhiria D. Udebhulu & David O. Ogbe | Wenjun Xu (Reviewing editor) (2019) Surrogate modelling approach: A solution to oil rim production optimization, Cogent Engineering, 6:1, DOI: 10.1080/23311916.2019.1631009
- [25] Brunton, Steven & Kutz, J.. (2020). 7 Data-driven methods for reduced-order modeling. 10.1515/9783110671490-007.
- [26] <https://www.simplilearn.com/tutorials/machine-learning-tutorial/principal-component-analysis>, accessed on June 30th 2022.
- [27] <https://doi.org/10.1016/B978-0-12-410437-2.00008-4>, accessed on June, 20th 2022.
- [28] Redman, Thomas. (2005). Measuring data accuracy: A framework and review. Information Quality. 21-36.
- [29] <https://doi.org/10.1016/j.jclepro.2021.125834>, accessed on June, 20th 2022.
- [30] C. F. Camilo, J. M. Rosário, M. M. Jefry and D. Dumur, "Data Science-based Sizing Approach for Renewable Energy Systems," 2020 IX International Congress of Mechatronics Engineering and Automation (CIIMA), 2020, pp. 1-6, doi: 10.1109/CIIMA50553.2020.9290303.
- [31] P. Li, R. Dargaville, F. Liu, J. Xia and Y. -D. Song, "Data-Based Statistical Property Analyzing and Storage Sizing for Hybrid Renewable Energy Systems," in IEEE Transactions on Industrial Electronics, vol. 62, no. 11, pp. 6996-7008, Nov. 2015, doi: 10.1109/TIE.2015.2438052.
- [32] <https://doi.org/10.1016/j.jclepro.2020.121874>, accessed on June, 20th 2022.
- [33] Marques, André Luis Ferreira (2021). Image captions generation for non-proliferation control with Artificial Intelligence. Brazil: ABEN. INAC-2021.
- [34] <https://www.hindawi.com/journals/stni/2022/4172622/>, accessed on June, 30th 2022.
- [35] Garrett, Joseph & Mei, Hanfei & Giurgiutiu, Victor. (2022). An Artificial Intelligence Approach to Fatigue Crack Length Estimation from Acoustic Emission Waves in Thin Metallic Plates. Applied Sciences. 12. 1372. 10.3390/app12031372.
- [36] Saraygord Afshari, S., Enayatollahi, F., Xu, X., & Liang, X. (2022). Machine learning-based methods in structural reliability analysis: A review. In Reliability Engineering & System Safety (Vol. 219, p. 108223). Elsevier BV. <https://doi.org/10.1016/j.res.2021.108223>.
- [37] <https://doi.org/10.1016/j.res.2021.108223>, accessed on June, 30th 2022.

- [38] Flah, Majdi & Vargas, Itzel & Ben Chaabene, Wassim & Nehdi, Moncef. (2020). Machine Learning Algorithms in Civil Structural Health Monitoring: A Systematic Review. Archives of Computational Methods in Engineering. 28. 10.1007/s11831-020-09471-9.

Suggested Reading

- A - Energies 2022, 15(2), 578; <https://doi.org/10.3390/en15020578>
B - <https://hdsr.mitpress.mit.edu/pub/d9j96ne4/release/3>
C - <https://www.hindawi.com/journals/sp/2020/8616039/>
D - <https://www.sciencedirect.com/science/article/pii/S0925753519308835>

Chapter 3

Overview of Classic Surrogates and Metamodeling Strategies in Structural Analysis

Chapter details

Chapter DOI:

<https://doi.org/10.4322/978-65-86503-88-3.c03>

Chapter suggested citation / reference style:

Santos, Rogerio R., et al. (2022). “Overview of Classic Surrogates and Metamodeling Strategies in Structural Analysis”. In Jorge, Ariosto B., et al. (Eds.) *Uncertainty Modeling: Fundamental Concepts and Models*, Vol. III, UnB, Brasilia, DF, Brazil, pp. 54–88. Book series in Discrete Models, Inverse Methods, & Uncertainty Modeling in Structural Integrity.

P.S.: DOI may be included at the end of citation, for completeness.

Book details

Book: Uncertainty Modeling: Fundamental Concepts and Models

Edited by: Jorge, Ariosto B., Anflor, Carla T. M., Gomes, Guilherme F., & Carneiro, Sergio H. S.

Volume III of Book Series in:

Discrete Models, Inverse Methods, & Uncertainty Modeling in Structural Integrity

Published by: UnB City: Brasilia, DF, Brazil Year: 2022

DOI: <https://doi.org/10.4322/978-65-86503-88-3>

Overview of Classic Surrogates and Metamodeling Strategies in Structural Analysis

Rogério R. Santos^{1*}, Guilherme F. Gomes²,
and Ariosto B. Jorge³

¹Division of Mechanical Engineering, Technological Institute of Aeronautics - ITA, Brazil. E-mail: rsantos9@gmail.com

²Mechanical Engineering Institute, Federal University of Itajuba - UNIFEI, Brazil. E-mail: guilhermefergom@unifei.edu.br

³Post-Graduate Program - Integrity of Engineering Materials, University of Brasilia, Brazil. E-mail: ariosto.b.jorge@gmail.com

*Corresponding author

Abstract

In some situations, the need for multiple evaluations of a computational model is not feasible due to the large cost in terms of budget or time required. The dominant practice consists in the creation of statistical approximations for the response of the model and performing additional analysis on the meta-model. The present contribution discusses some classic alternatives for the establishment of experimental designs. Next, the Least-Squares and Kernel methods are outlined as alternatives to approximate models. Subsequently, characteristics of the Parallel, Sequential, and Multi-level ensembling methods are presented. The current work is concluded with considerations on the goodness-of-fit assessment.

Keywords: Surrogate modeling, Metamodeling, Design of Experiments, Machine Learning.

1 Introduction

Contemporary projects in the field of structural analysis have to deal with a lot of data. The need to translate raw numbers into intelligence that drives real-world projects requires the ability to classify and organize information that arises from multiple and sometimes conflicting requirements.

In this context, how to break down and structure complex problems and data sets is of paramount importance. Design of Experiments (DoE) is concerned with designing tasks capable of describing and explaining the variation of information under specific hypothesized conditions. It is associated with experiments in which the design introduces conditions that directly affect the variation/response.

Inferring the system's behavior from a reduced set of data can contribute to reducing the computational budget of a project and increase the probability of identifying unforeseen scenarios. In the present contribution, classical and contemporary techniques are presented and their

suitability for structural analysis is discussed. It starts with the concepts of Design of Experiments, Metamodeling and Surrogate alternatives for the establishment of experimental designs. Next, the Least-Squares and Kernel methods are outlined as alternatives of approximate models. Subsequently, characteristics of the Parallel, Sequential and Multi-level ensembling methods are presented. Considerations on the goodness-of-fit appraisal will concludes the present contribution.

2 Design of Experiments, Metamodeling and Surrogates

In certain circumstances, it is not feasible or practical to calculate exact values for the functions $F(\{X\})$, $G_j(\{X\})$ and $H_k(\{X\})$ pertaining to the optimization problem formulation seen in chapters “An overview of Linear and Non-linear Programming methods for Structural Optimization” (Choze et al. [2022b]) and “Overview of Traditional and Recent Heuristic Optimization Methods” (Choze et al. [2022a]). One such possibility is whenever the computational cost is too burdensome, and even the specialized structural synthesis approximations (Choze et al. [2022b]) cannot provide an adequate solution. Another situation, in itself or combined with excessive computational cost, is when the calculation engines responsible for providing the values of the many quantities involved belong to computational domains that are disparate enough to jeopardize a direct integration, and hence a common frame of reference has to be obtained to reconcile them.

In either case, it is mainstream practice to create statistical approximations of these quantities, henceforth denoted as $\hat{F}(\{X\})$, $\hat{G}_j(\{X\})$ and $\hat{H}_k(\{X\})$, with the “ $\hat{}$ ” notation used to highlight their non-exact nature. As approximated entities, there are mechanisms to drive the computational burden down and, both as closed-form equations or represented through standardized database functions, their computational compatibility can be assured to the extent that they are operated together within the same optimization procedure. The overall procedure for obtaining these approximations is explained after its introduction in Table 1 (Chihara and Hesterberg [2018]).

From left to right, the first step consists of observing enough values of decision variables combinations ($\{X\}$) and the corresponding objective/constraint values that they yield. Explicitly, the yielded values, which are collectively denoted as $\{Y\}$, are dependent upon $\{X\}$ which, in statistical nomenclature, are the independent variables (within reason, freely chosen to map the decision space). Depending on the specific purposes of each study, one or more suitable design of experiments (first column in Table 1) will determine specifically which setups of independent variables are observed to form a matrix where, in a given layout, each row corresponds to an observation (as determined by the applicable/available experimental design) and each column maps into the corresponding element of the vector of decision variables $\{X\}$. A statistical surrogate will map an approximation to describe the underlying relationship between independent and dependent variables, pairing each row in $[X]$ (now in plain matrix notation) with the corresponding one in $\{Y\}$ (still in vector notation, focusing on a single response at a time). One or more such methods are chosen from the second column in Table 1. Although the number of choice possibilities is potentially high, reflecting a myriad of different approaches, the fundamental operating principles underpinning all of them are common, relating to the optimization of a goodness-of-fit metric. Finally, and most importantly, even though a suitable goodness of fit may be reached once executing of steps 2 and 3 within Table 1 (with strictly training data, as more usual, or with a mix of training and validation data), additional appraisal, corresponding to step 4, is strongly advised for a final and complete performance assessment of the surrogate’s capabilities.

Specific techniques and additional comments relative to the steps outlined above are presented in the subsequent sections. Several of these instances are supplemented with their corresponding computer codes implemented with the R statistical programming platform (R Core Team [2020]), indicated as “Code”. For this purpose, the command `install.packages` is the main tool to add new packages to the R software. It takes a vector of names and a destination library, downloads

Table 1: Standardized steps sequence for building statistical approximations, also known as meta-models or surrogates.

Experimental Design	Model Choice	Model Fitting		Quality Verification
(Fractional) Factorial	Polynomial	Least Squares Regression	Response Surface	Analysis of Residuals
Central Composite		Weighted Least Squares Regression		ANOVA
Box-Behnken				:
Alphabet-Optimal	Splines	Best Linear Unbiased Predictor	Bayesian Models	
Orthogonal	Frequency Domain	Best Linear Predictor		
Plackett-Burman	Kernel Smoothing			
Hexagon	Radial Basis Functions			
Hybrid				
Latin Hypercube				
Enumerative	Neural Networks Rule based or Decision Tree	Back-Propagation Entropy	Neural Networks Inductive Learning	Residual Error
Random				

the packages from the repositories (from a local folder or remote web server) and installs them in the R environment. To install the packages needed to run the codes in this chapter, for example, enter the following:

```
install.packages('FrF2', 'rsm', 'lhs', 'glmnet', 'DiceKriging',  
'RSNNS', 'randomForest')
```

Help is available at the command prompt via `?install.packages` command.

3 Experimental Design

From a statistical theory perspective, a more exhaustive list of possible experimental design approaches is richly documented in the literature, such as in references [Myers et al. \[2016\]](#), [Box et al. \[2005\]](#), and [Khuri and Cornell \[2019\]](#). The following sections are extracts of this broader picture that more closely correspond to the state-of-the-art and practice for computer analysis and optimization of structural systems. Nevertheless, active research is ongoing in the direction of broadening the scope and possibilities within the field, and the interested reader is referred to [Garud and Kraft \[2017a\]](#), [Garud and Kraft \[2017b\]](#), [Bhosekar and Ierapetritou \[2018\]](#), [Boukouvala and Floudas \[2017\]](#), and [Pronzato and Müller \[2012\]](#).

3.1 Full and fractional factorial designs

Initial exploration and understanding of decision spaces in m variables tends to be undertaken by screening experimental designs, whose purpose is to filter out variables that do not influence the outcome(s). Among the various options for screening designs, the simplest ones are factorial designs at two levels, which serve to identify linear relationships between inputs and outputs. For $m = 5$ and variable ranges codified between -1 for lower bound and $+1$ for upper bound, the resulting combination would be as appears on Table 2.

The 32 total combinations arise from 2^m , for two levels (at the lower and upper bounds for each variable) over $m = 5$ variables, and define an hypercube (each edge corresponding to one dimension or variable) with the experimental points at the vertices. The exponential growth with respect to m indicates that, in practice, this may even be a simple approach, but is only feasible for screening purposes with moderate values of m and/or low cost for performing each individual run. This limitation may be better perceived if one considers $m = 10$ variables yielding 1024 total runs just for screening on a linear relationship basis. For completeness of this discussion, consider also the inclusion of simple curvature with base 3 would result in $3^{10} = 59049$ runs, proving the point that straight factorial combination may only work in very particular cases and may lead to experimental inefficiencies in exchange for its simplicity. For the purposes of computational coverage, creation of full factorial designs in R ([R Core Team \[2020\]](#)) is illustrated in Code 1 for 3 factors (A, B and C).

As an initial step to trade-off complexity and efficiency, the introduction of fractional factorial designs leverages the idea of incurring into the partial cost of factorial designs while reducing the clarity of the conclusions to be obtained. The multiple trade-off points within this balancing act are measured by different resolutions, which become an attribute of the experimental design similarly to its base and number of variables. In formal terms, resolution is defined as the minimum word length in the defining relation, excluding 1. Whereas [Myers et al. \[2016\]](#) and [Box et al. \[2005\]](#) present complete theoretical discussions about the defining relations of different experimental designs and, consequently, their resolutions, the practical aspect to be emphasized is that higher resolutions are desired, with the following recommendations: a minimum of 3 (typically noted as III) for variable screening, a minimum of 4 (typically noted as IV) for model fitting and, finally, a minimum of 5 (typically noted as V) for optimization.

To illustrate these concepts, consider alternatives for fractional designs that can be derived from Table 2, utilizing only part of the 5 total factors and confounding the remaining k ones, that

Table 2: Full factorial experimental design at two-levels with $m = 5$ independent / input variables, whose ranges are codified between -1 (lower bound) and $+1$ (upper bound).

Run	Variable 1	Variable 2	Variable 3	Variable 4	Variable 5
1	-1	-1	-1	-1	-1
2	-1	-1	-1	-1	1
3	-1	-1	-1	1	-1
4	-1	-1	-1	1	1
5	-1	-1	1	-1	-1
6	-1	-1	1	-1	1
7	-1	-1	1	1	-1
8	-1	-1	1	1	1
9	-1	1	-1	-1	-1
10	-1	1	-1	-1	1
11	-1	1	-1	1	-1
12	-1	1	-1	1	1
13	-1	1	1	-1	-1
14	-1	1	1	-1	1
15	-1	1	1	1	-1
16	-1	1	1	1	1
17	1	-1	-1	-1	-1
18	1	-1	-1	-1	1
19	1	-1	-1	1	-1
20	1	-1	-1	1	1
21	1	-1	1	-1	-1
22	1	-1	1	-1	1
23	1	-1	1	1	-1
24	1	-1	1	1	1
25	1	1	-1	-1	-1
26	1	1	-1	-1	1
27	1	1	-1	1	-1
28	1	1	-1	1	1
29	1	1	1	-1	-1
30	1	1	1	-1	1
31	1	1	1	1	-1
32	1	1	1	1	1

is, 2^{5-k} . The resulting possibilities are detailed in Table 3, which is generalized for a large number of variables/fractions/resolutions of practical applicability in NIS.

Code 1: Full Factorial Design for 3 factors (A B and C)

```

expand.grid(A = c("-1", "+1"), B = c("-1", "+1"),
C = c("-1", "+1"))
A B C
1 -1 -1 -1
2 1 -1 -1
3 -1 1 -1
4 1 1 -1
5 -1 -1 1
6 1 -1 1
7 -1 1 1
8 1 1 1

```

Code 2 brings the R counterpart of the $k = 1, 2_{R=V}^{5-1}$ design also listed in Table 3. Please note the difference in notation, where the factors are alluded to as capital letters A, B, C, D and E instead of numbers, both versions being prevalent (and conceptually equivalent) in the literature.

Code 2: Regular Fractional Factorial 2-level design

```

FrF2(16, 5, generators='ABCD')
  A B C D E
1  1 -1 1 -1 1
2 -1 1 -1 -1 -1
3 -1 -1 1 1 1
4  1 1 1 -1 -1
5  1 -1 1 1 -1
6  1 1 1 1 1
7  1 -1 -1 1 1
8  1 1 -1 -1 1
9 -1 1 1 1 -1
10 1 1 -1 1 -1
11 -1 1 -1 1 1
12 -1 -1 -1 1 -1
13 -1 -1 1 -1 -1
14 -1 1 1 -1 1
15 -1 -1 -1 -1 1
16  1 -1 -1 -1 -1
class=design, type= FrF2.generators

```

Even with the introduction of the fractional strategies and the consequent reduction in the number of runs, some situations result in non-affordable experimental effort. This is often the case when non-linearity prevail and/or the experimental investigation actually targets some level of prediction, beyond pure screening.

Such situations require alternative approaches, and the most popular ones for computational experiments, especially when dealing with structural mechanics, are described in the next two subsections.

Table 3: Fractional factorial experimental design options with various resolutions out of the $m = 5$ independent/input variables listed in Table 2. * confounded variable/Factor.

Fraction (k) and naming	Runs						Resolution	Confounding Structure
$k = 1,$ $2_{R=V}^{5-1}$	Run	$F1$	$F2$	$F3$	$F4$	$F5^*$	V	DEFINING RELATION: $I = 12345$
	1	-1	-1	-1	-1	1		
	2	-1	-1	-1	1	-1		CONFOUNDED
	3	-1	-1	1	-1	-1		VARIABLE(S): $F5 = F1 \times F2 \times F3 \times F4$
	4	-1	-1	1	1	1		
	5	-1	1	-1	-1	-1		CONFOUNDED
	6	-1	1	-1	1	1		STRUCTURE:
	7	-1	1	1	-1	1		$F1 = F1 + F2345$
	8	-1	1	1	1	-1		$F2 = F2 + F1345$
	9	1	-1	-1	-1	-1		$F3 = F3 + F1245$
	10	1	-1	-1	1	1		$F4 = F4 + F1235$
	11	1	-1	1	-1	1		$F5 = F5 + F1234$
	12	1	-1	1	1	-1		Note: each variable/Factor con-
	13	1	1	-1	-1	1		founded with interactions of or-
	14	1	1	-1	1	-1		der 4, which are negligible.
	15	1	1	1	-1	-1		
16	1	1	1	1	1			
$k = 2,$ $2_{R=III}^{5-2}$	Run	$F1$	$F2$	$F3$	$F4^*$	$F5^*$	III	DEFINING RELATION: $I = 124 = 135 = 2345$
	1	-1	-1	-1	1	1		
	2	-1	-1	1	1	1		CONFOUNDED
	3	-1	1	-1	-1	1		VARIABLE(S):
	4	-1	1	1	-1	-1		$F4 = F1 \times F2$
	5	1	-1	-1	-1	-1		$F5 = F1 \times F3$
	6	1	-1	1	-1	1		CONFOUNDED
	7	1	1	-1	1	-1		STRUCTURE:
8	1	1	1	1	1	$F1 = F1 + F24 + F35 + F12345$		
						$F2 = F2 + F14 + F345 + F1235$		
						$F3 = F3 + F15 + F245 + F1234$		
						$F4 = F4 + F12 + F235 + F1345$		
						$F5 = F5 + F13 + F234 + F1245$		
						Note: each variable/Factor con-		
						founded with interactions of or-		
						der 4, which are negligible.		

3.2 Central composite designs

Central Composite designs are suitable for situations of moderate dimensionality (when some form of fractional or even full factorial design fits the experimental budget) but it is important to consider non-linear effects that will certainly not be captured via two-level combinations. In this context, consider the initial scenario with 3 independent variables and the fractional form 2^{3-1} resulting in 4 total linear experimental runs, in the portion denoted as “cube” (signifying edges with equal lengths). Adding a total of 5 additional runs (one at the “center”, and the remaining at the far edges of a “star”) would cover the non-linear effects as represented schematically in Figure 1, which is mirrored in Table 4.

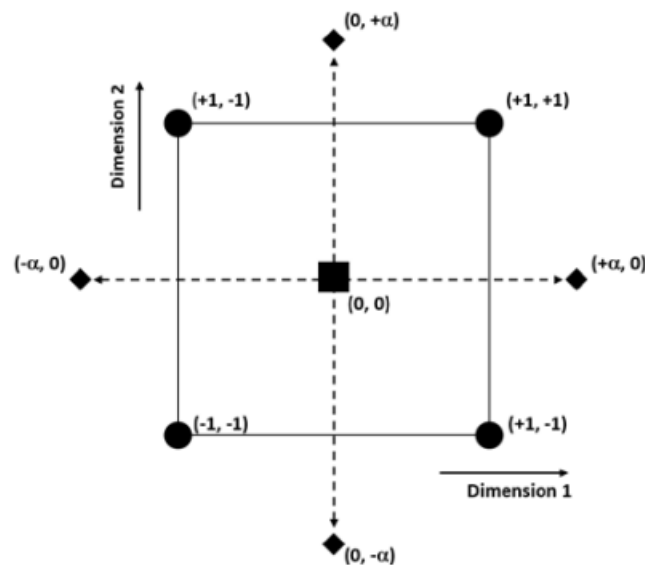


Figure 1: Diagram of a Central Composite Design with linear and non-linear portions represented in two-dimensions.

Table 4: Central Composite Design corresponding to Figure 1, with 2 factors covered at 5 total levels.

Run #	F1	F2
1 (“cube”)	-1	-1
2 (“cube”)	-1	1
3 (“cube”)	1	-1
4 (“cube”)	1	1
5 (“center”)	0	0
6 (“star”)	$-\alpha$	0
7 (“star”)	$+\alpha$	0
8 (“star”)	0	$-\alpha$
9 (“star”)	0	$+\alpha$

The square vertices are the factorial points spanning the range between the linear lower bound (-1) and the linear upper bound (+1), and the qualifier “linear” is deliberate to indicate that the overall ranges have to be examined and equalized so that no experimental points are configured outside of any feasible experimental range. The reason, which also appears on Figure 1, is that the

non-linear effects are captured by the rhombuses α units afar from the center (square icon), which usually assume absolute values larger than unity. Determination of the values of α depend upon a number of factors, both formal/mathematical and also practical:

- α for rotation: ensures that the resulting design yields predictive variances (i.e., uncertainties) that, at any point x , depend only on the distance of x from the design center point. A design with this desirable property can be rotated around its center point without changing the prediction variance at x , hence its nomenclature;
- α for orthogonality: guarantees that the effect of one factor or interaction can be estimated separately from the effect of any other factor or interaction in the model, which counters the factor confusion arising from fractional factorial designs and their associated (lower) resolution;
- Special cases and arbitrary values of α : practical constraints may limit the choice of values for α in such a way that neither rotation or orthogonality is obtained, requiring additional and more careful assessment of goodness-of-fit prior to utilizing the model resulting from the experiment. One possibility is that, due to lack of experimental bandwidth, the values of α are not allowed to go beyond the linear range $[-1, +1]$, in which case the resulting designs are called face centered composite designs.

Once again, the R code counterpart is offered in Code 3, highlighting configuration parameters necessary for the creation of a Central Composite Design.

Code 3: Central Composite Design example

```

library(rsm)
n_var = 3
n_center_points_cube = 4
n_center_points_star = 6
ccd(n_var, n0 = c(n_center_points_cube,
  n_center_points_star),
  alpha = "orthogonal")

  run.order std.order x1.as.is x2.as.is x3.as.is Block
1 1 5 -1 -1 1 1
2 2 9 0 0 0 1
3 3 2 1 -1 -1 1
4 4 3 -1 1 -1 1
5 5 11 0 0 0 1
6 6 4 1 1 -1 1
7 7 8 1 1 1 1
8 8 1 -1 -1 -1 1
9 9 6 1 -1 1 1
10 10 12 0 0 0 1
11 11 10 0 0 0 1
12 12 7 -1 1 1 1
13 1 10 0 0 0 2
14 2 7 0 0 0 2
15 3 9 0 0 0 2
16 4 2 2 0 0 2
17 5 8 0 0 0 2
18 6 3 0 -2 0 2
19 7 4 0 2 0 2
20 8 5 0 0 -2 2
21 9 1 -2 0 0 2
22 10 12 0 0 0 2
23 11 11 0 0 0 2
24 12 6 0 0 2 2

```

3.3 Latin Hypercube (computational) designs

Since their inception in the seminal paper by Sacks et al, 1989, computer experimental designs have gained popularity due to their inherent suitability regarding phenomena that are immune to experimental variation. In this sense, any of the experimental setups discussed in sections 3.1 and 3.2 above would in principle require replications at each point whenever feasible, if they are undertaken in the actual physical world. In computers however, barring minor differences due to platforms (chipset, operating systems, etc.), such replicates would yield the same results at any given experimental run.

At any measure, with or without replications, application scenarios with relatively high dimensionality and any degree of non-linearity may push beyond experimental budget feasibility even for Central Composite Designs. In such scenarios, the ideal combination of attributes for the experimental design would encompass:

1. Low, and ideally fixed/predetermined experimental cost (total number of runs);
2. Ability to explore various levels and capture non-linearity;
3. Ability to execute item 2 minimizing redundancy of information aggregated by each individual experimental run.

Whereas a random sample with N points over m dimensions would comply with these 3 requirements, it would offer low to no control over coverage of the space, with the enduring possibility that some areas are over-sampled at the expense of voids in other locations (Figure 2a). On the other hand, a true space filling experimental design would look like Figure 2b and, as implied by the name, experimental samples would be distributed as to satisfy the uniform filling of the available space, such as minimizing the maximum distance or, conversely, maximizing the minimum distance between any 2 given points in m dimensional space.

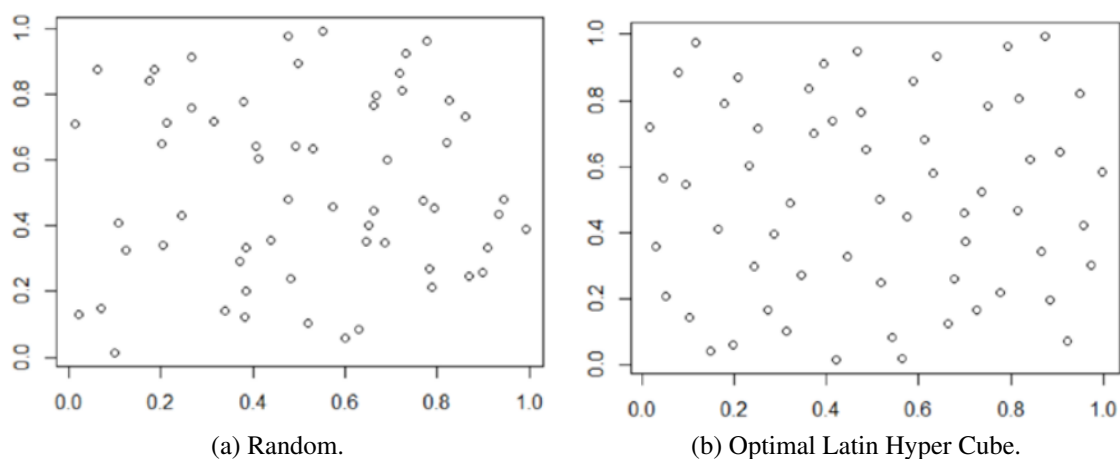


Figure 2: Random vs. Space filling Design with 60 points over 2 dimensions between 0 and 1. Generated in R with random seed = 1 for Figure (2a) and function optimumLHS for Figure (2b).

Code 4 contains the necessary R syntax to create designs such as that of Figure 2b.

Code 4: Optimum Latin Hypercube sample

```

library(lhs)
n_var = 3
n_samples = 16
optimumLHS(n = n_samples, k = n_var)
      [,1]      [,2]      [,3]
[1,] 0.77986305 0.25962784 0.98113240
[2,] 0.73374286 0.89983958 0.65652518
[3,] 0.08106826 0.38917024 0.80616234
[4,] 0.67524880 0.03049537 0.36286886
[5,] 0.34750631 0.78054941 0.84859144
[6,] 0.12511944 0.17081607 0.10528484
[7,] 0.52469292 0.60694614 0.04280036
[8,] 0.41165871 0.54098857 0.50797994
[9,] 0.98482138 0.36344077 0.73544520
[10,] 0.84308190 0.18906118 0.20958842
[11,] 0.02456373 0.64927328 0.25217138
[12,] 0.61763298 0.46786939 0.89379079
[13,] 0.47930504 0.98500063 0.15556912
[14,] 0.22676730 0.86768002 0.48733852
[15,] 0.26289577 0.06276729 0.58312047
[16,] 0.92962543 0.69242589 0.43707276

```

The ability to cover multidimensional spaces more uniformly than pure random numbers and the amount of control relative to overall experimental cost have converted Latin Hypercube Sampling (LHS) designs into the *de facto* standard of computational design and analysis of experiments.

4 Choice and Fitting of Approximate Models

4.1 Least-Square based regression methods

The most fundamental method of empirical surrogates building is the Ordinary Least Squares (OLS for brevity) regression, an average filter that creates an approximation linear in the coefficients that assumes a form along the lines of the one depicted in Equation 1:

$$\hat{y} = a_0 + a_1x_1 + a_2x_2 + \cdots + a_nx_n + b_1x_1^2 + b_2x_2^2 + \cdots + b_nx_n^2 + c_1x_1x_2 + \cdots + c_mx_{n-1}x_n \quad (1)$$

which can also be represented through the compact linear combination/dot product form in Equation 2, of coefficients β over t model terms x , each one bearing its respective degree of non-linearity at the exponent:

$$\hat{y} = \sum_{i=1}^t \beta_i x_i \quad (2)$$

Specific for the historically usual case of a quadratic response surface with interactions (Barthelemy and Haftka [1993], Kaufman et al. [1996], Sobieszczanski-Sobieski and Haftka [1997]), the true response y (the function that yields all of the observations $\{Y\}$) is approximated by \hat{y} accounting

for all of the independent variables belonging to the n -dimensional vector $\{X\}$ in their pure linear forms (weighted by the coefficients $a_i, i = 1, \dots, n$), their quadratic forms, as the simplest possible representation of potential curvature/non-linearity (pondered by $b_i, i = 1, \dots, n$) and also the interactions, in which all pairwise multiplications of the terms in $\{X\}$ are exhausted and pondered by $c_i, i = 1, \dots, m$ with $m = C_{n,2}$ (combination of the n terms within $\{X\}$, two at-a-time). Optionally, the model also contains a free term or intercept, which approximates the average of the data generating process responsible for creating all possible values of y .

Collecting all the $2n+m+1$ coefficients into the vector of coefficients $\{\beta\}$, the OLS regression method applies the deterministic linear algebra set of operations described in Equation 3, which satisfy the least-squares criterion in Equation 4, in which the sum of residuals (between true and estimated response values) squared is minimized:

$$\{\beta\} = ([X]^T \cdot [X])^{-1} \cdot [X]^T \cdot \{Y\} \quad (3)$$

$$\{\beta\} = \operatorname{argmin} L(\{\beta\}) = \operatorname{argmin} \frac{1}{2N} \sum_N (\{\beta\} \cdot [X] - \{Y\})^2 \quad (4)$$

Noticing that the OLS criterion is indeed an optimization problem in itself (minimizing the loss function $L(\{\beta\})$, with the objective function being a slightly modified version of the sum of squares of residuals, exhibiting the following convenient features:

- Its quadratic nature is amenable to optimization methods;
- Once derived, as part of the optimization procedure, the number 2 in both the exponent and the denominator are mutually canceled;
- The whole expression is normalized by the number of available samples N , avoiding artificial inflation of the sum of squares of residuals for larger sample sizes.

The computational counterpart of Equations 3 and 4 in the R environment is shown in Code 5.

Code 5: Least-square regression

```
library(stats)

x = data.frame(matrix(rnorm(200), nrow=100, ncol=2))
colnames(x) = c('x1', 'x2')
noise = rnorm(100)
y = 2*x[,1] + (0.01*x[,2])^2 + 0.5*noise

modell = lm(y ~ x1 + x2, data = data.frame(x,y))

z = data.frame(matrix(c(1,1), nrow=1, ncol=2))
colnames(z) = c('x1', 'x2')
fc1 = predict(modell, z)

cat('R2: ', summary(modell)$r.squared,
    '\nPrediction: ', fc1)

R2: 0.942392
Prediction: 1.896089
```

Equation 4, which makes the underlying optimization procedure explicit, offers important levers for improving the quality of the model, specifically with respect to model/term selection. Based on the principle of parsimony, the effective number of terms in the model will be narrowed down to only the most significance, which improve the bias-variance trade-off in the fashion schematically represented in Figure 3.

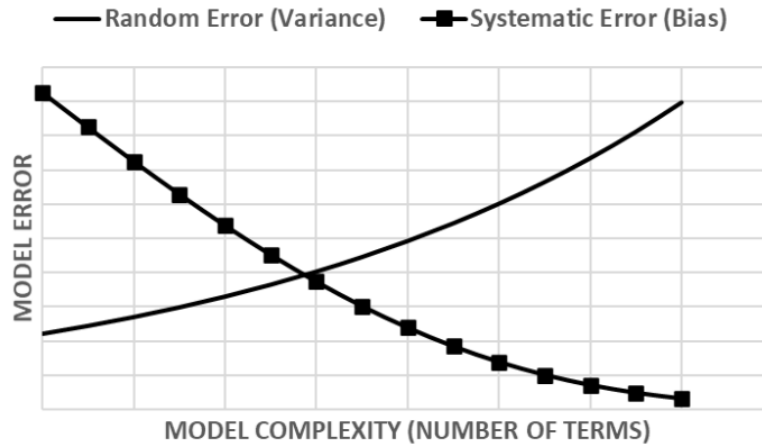


Figure 3: Schematic representation of the bias-variance trade-off for model selection.

Mathematically, the optimal trade-off illustrated in Figure 3 comes out of an extension of Equation 4, also known as regularized regression

$$\min\{L(\beta) + \lambda [\alpha \|\{\beta\}\|_1 + (1 - \alpha) \|\{\beta\}\|_2]\} \quad (5)$$

where the minimization of the loss-function in $\{\beta\}$ becomes a constrained procedure via penalization, which is mediated by the regularization parameter λ and, in its most general form, is applied onto two components, the first $\|\cdot\|_1$ and second $\|\cdot\|_2$ norms of the vector of coefficients.

Both $\|\cdot\|_1$ and $\|\cdot\|_2$ are minimized when the coefficients with smaller values vanish, which inherently leads to model/term selection. When α equals 1, the penalization is maximized, since the first norm summing up all the terms of $\{\beta\}$ (both the significant ones and those tending to zero) is added to the objective function. When α equals 0, on the other hand, the penalization is not as severe, with the overall summation of coefficients yielding a smaller number, since those with values tending to zero become even smaller once they are squared. These two particular cases of Equation 5 correspond to LASSO and Ridge Regression, respectively, with the acronym LASSO (Least Absolute Shrinkage and Selection Operator) reflecting the most severe penalization case. In practice, it is most usual that $0 \leq \alpha \leq 1$, and the flexibility to pick the best possible value within the unit interval is reflected by the model taking the name of “Elastic Net”. The presence λ and α introduce the possibility of regularization (i.e., inherent model selection by dropping the less significant terms) at the expense of increased complexity for the loss function to be minimized, now dependent on these two additional parameters.

Reflecting the enhancement/generalization of Equation 6 over the formulation in Equations 3-4, Code 6 contains the R implementation of the regularized regression approach, with the family of Generalized Linear Models.

Code 6: Generalized linear model with “elasticnet” regularization

```

library(glmnet)

x = matrix(rnorm(200), nrow=100, ncol=2)
noise = rnorm(100)
y = 2*x[,1] + (0.01*x[,2])^2 + 0.5*noise

model2 = glmnet(x, y, beta='elnet')

z = matrix(c(1,1), nrow=1, ncol=2)
fc2 = predict(model2, z)
cat('R2: ', tail(model2\%dev.ratio, 1),
    '\nPrediction: ', tail(fc2[1,], 1))

R2: 0.9405139
Prediction: 1.989429

```

Up to this point, improvements in the model performance is sought quantitatively, that is, deciding how many terms remain in the model. For models of polynomial form, such as those considered far, each term is then a monomial, whose limited flexibility resides mostly on its degree of non-linearity (i.e., exponent) and, even so, typically at the expense of demanding significant additional data. The fundamental limitations of monomials, however, is their stiffness, as illustrated by the Runge phenomenon in Figure 4.

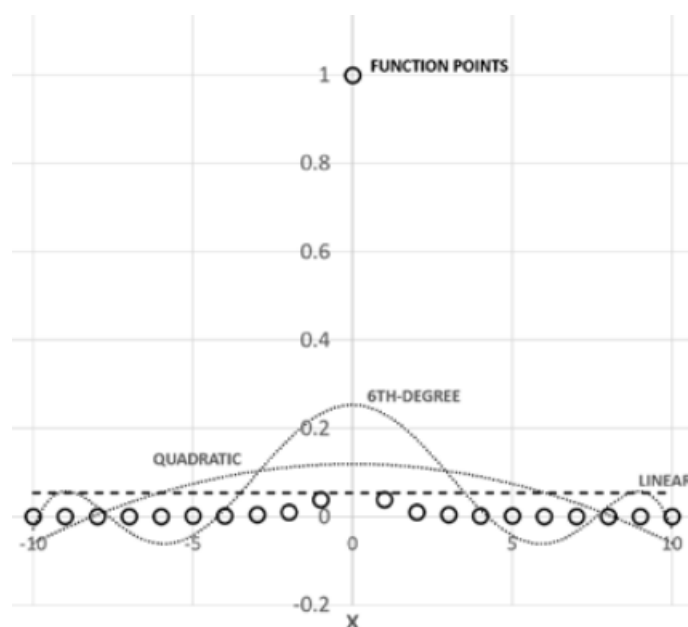


Figure 4: Illustration of the Runge phenomenon, whereby monomial stiffness leads to oscillations when model order is increased in an attempt to capture underlying non-linearity driven by the outlier at 0.

The Runge phenomenon is illustrated by the Runge function described in Equation 6, with the outlier in 0 causing local inflation of the sum of squares of residuals and peripheral oscillations, hence severely limiting the accuracy of a monomial based model which operates as an average

filter (recall that the outlier strongly deviates from that average), not prone to be improved by regularization or, at least, by regularization alone.

$$R(x) = \frac{1}{1 + 25x^2} \quad (6)$$

Several efforts of applied and theoretical nature have been implemented to circumvent these limitations associated with monomial stiffness within the realm of fundamental regression, and the interested reader is encouraged to pursue non-linear regression (Frank [1995]), robust regression (in the sense of immunity to outliers, as detailed in Andersen [2008]) and Generalized Additive models (Hastie [2017]).

However, setting the stage of methods that would become collectively known (and quite successful) as Machine Learning approximations (Hastie et al. [2009]), there is an operating principle that replaces monomials by kernels. Therefore, they are also known as kernel methods, which display the following salient features:

- Parametrization, which result in flexibility (through configuration/optimization of the parameters values) that counters the monomial's inherent stiffness;
- Combination of linear approach (dot product) and non-linear representation (the kernel function);
- In the extreme case, monomials are the simplest possible configuration, where the kernels are just the variables themselves.

4.2 Kernel methods

In its simplest, yet widely used forms, the kernels belong to the family of Radial Basis functions, and is applied for modeling purposes as follows:

1. Collect a sample of size N , comprised of independent variables $[X]$ and dependent variables $\{Y\}$, using the same notation applied for regression methods in Equations 1 to 6;
2. As in Equation 7, calculate the D , which is symmetric and square of order N , containing the Euclidean distances (second norms) between all possible pairwise combinations of independent variables within the sample matrix $[X]$ collected in Step 1. It is important to highlight that the terms in the main diagonal all default to 0 (distances between each point in the sample and themselves) and that, by way of this step, the term-based logic for the model is replaced by a location-based logic, with locations mapped in terms of the distances and, moreover, the distances themselves default to simple scalars irrespective of the original dimensionality of the problem (number of input or independent variables);

$$D = \begin{bmatrix} \|[x]_{1,N}, [X]_{1,N}\| & \|[x]_{2,N}, [X]_{1,N}\| & \cdots & \|[x]_{N,1}, [X]_{1,N}\| \\ \|[x]_{1,N}, [X]_{2,N}\| & \|[x]_{2,N}, [X]_{2,N}\| & \cdots & \|[x]_{N,2}, [X]_{2,N}\| \\ \vdots & \vdots & \ddots & \vdots \\ \|[x]_{1,N}, [X]_{N,N}\| & \|[x]_{2,N}, [X]_{N,N}\| & \cdots & \|[x]_{N,N}, [X]_{N,N}\| \end{bmatrix} \quad (7)$$

3. Transform the distances matrix D by calculating the radial-basis function value for each of its terms $D_{i,j}$. These radial-basis functions, as suggested by their name, assign weight/penalizations to each distance as they depart radially from the points where samples are

available. A number of usual radial-basis functions is listed in Equations 8-13:

$$f_{rb}(D_{i,j}) = D_{i,j} \quad (8)$$

$$f_{rb}(D_{i,j}) = (D_{i,j})^3 \quad (9)$$

$$f_{rb}(D_{i,j}) = (D_{i,j})^2 \cdot \ln(D_{i,j}) \quad (10)$$

$$f_{rb}(\sigma, D_{i,j}) = \exp\left[\frac{-(D_{i,j})^2}{2\sigma^2}\right] \quad (11)$$

$$f_{rb}(\sigma, D_{i,j}) = \sqrt{r^2 + \sigma^2} \quad (12)$$

$$f_{rb}(\sigma, D_{i,j}) = 1/\sqrt{r^2 + \sigma^2} \quad (13)$$

In the order in which they are listed, the radial basis functions in Equations 8 to 13 are the kernels with linear, cubic, spline, gaussian, multi-quadric and inverse multi-quadric shapes, corresponding to different penalization strategies as a function of location distance. For nomenclature purposes, this step is also known as the “kernelization” of the distances in the input space;

4. This step is involved with estimation of the response value at a new point $\{X_{new}\}$ not belonging to the sample $[X]$ collected in Step 1. Since the method is distance based, a total of N Euclidean distances are to be calculated, each of them turned into a scalar that captures a distance which is represented within a space with the same dimensionality of the functional relationship between inputs and output. Equation 14 represents this idea mathematically:

$$D_{X_{new}} = \left\{ \begin{array}{l} \|\{X\}_1, \{X\}_{new}\| \\ \|\{X\}_2, \{X\}_{new}\| \\ \vdots \\ \|\{X\}_N, \{X\}_{new}\| \end{array} \right\} \quad (14)$$

5. “Kernelize” the distances vector determined in Equation 14, analogously to the application of Step 2 over the matrix in Equation 7;
6. Pre-multiply the “kernelized” distances matrix emerging out of Step 2 by the vector of sampled outputs (dependent variable) $\{Y\}$, observing that they both possess the same number of rows N . This operation yields the so-called weights of the model;
7. Perform the dot product (or linear combination) between the vector of weights obtained in Step 6 and the “kernelized” vector of distances between the inference and the sample points, obtained in Step 5. At this point, because of the linear combination operator applied over “kernelized” entities (in this case, distances within the input space), the philosophies of linear combination (proper of the regression methods discussed between Equations 1 and 6 and kernels are merged together, resulting in a model with a general form that is still linear in its parameters, but flexible to any arbitrary non-linear kernel within each component, as opposed to the stiff nature of the monomials. This combination of advantages has resulted in wide and successful adoption of these models within the realm of Machine Learning, and is summarized by Equation 15 below:

$$Y = \sum_N w_i \cdot f_{rb}(D) \quad (15)$$

Please note that despite Equation 15 representing an inference (i.e., an estimated value), the notation Y is used instead of \hat{Y} to indicate that, in its original form, the family of radial basis

function methods is interpolative, with perfect recall of the response values Y at the N points used training (a condition to be relaxed at advanced variations of the approach to be discussed in the sequence).

Specifically, Equations 12 to 13 correspond to functional forms that depend on additional parameters (namely, σ), whose value(s) are determined either by deterministic linear algebra type of operations (Steps 1 to 7 detailed above), or through an optimization method aimed at minimizing a loss function associated with model accuracy. The latter is the logic underpinning the two methods to be discussed next (Kriging and Gaussian Processes), that derive as enhancements of the fundamental kernel idea.

Hence, introducing a more sophisticated version of the gaussian kernel of Equation 12, a full covariance map is established as a function of a “nugget” (or “jitter”) C_0 , a “sill” $C_0 + C_1$ and a “range” a . The terms between double quotes reflect the nomenclature of the Variogram based Kriging (also known as Variographic method), in which the smoothness of the input space is mapped according to:

- A discontinuity around each known point (C_0), to account for the possibility of noisy data records. Even in the case of computer experiments, which are deterministic by nature, the introduction of a minor noise level may be beneficial (and in extreme cases, necessary) to reduce the stiffness of matrix inversion operations to be introduced in the sequence;
- A limit for the variance at infinity ($C_0 + C_1$), whereby prediction bounds that measure estimate uncertainties will not grow indefinitely even in the case of extrapolation beyond the input ranges used to train the model. This is a fundamental difference with respect to least-squares based regression and, whereas extrapolation should always be practiced with caution, there is a level of robustness inherent to Kriging;
- A distance (a) beyond which the variance stabilizes as described in the item above

$$C(D) = \begin{cases} C_0 + C_1 & \text{if } |D| = 0 \\ C_1 \cdot \exp\left(\frac{-3D}{a}\right) & \text{if } |D| > 0 \end{cases} \quad (16)$$

Procedurally, the Variographic Kriging method applies as follows:

1. The variographic kernel of Equation 16 is applied over the distance matrix D , generating its “kernelized” version analogously to Step 2 of the Radial Basis approach discussed earlier. The “kernelized” distance matrix is the covariance matrix itself, in which geometrical distances are converted into statistical distances, speaking for the smoothness of the input space. Smoothness is an important consideration because, among the fundamental assumptions of the method, it is considered that points that are statistically close in the input space give rise to points that will also be statistically close in the output space;
2. Involving prediction, this step requires calculation of the input space distance vector between a point whose output is to be predicted, with respect to all known points (the sample used to train the model). This is equivalent to Step 4 of the Radial Basis procedure;
3. Analogously to Step 5 in the Radial Basis discussion, “kernelize” the distance vector obtained in the previous step, converting it into the covariance vector between training and prediction points;
4. Invert the covariance matrix obtained in Step 1. Numerical difficulties in this inversion step may be circumvented by adding a small random noise, which is the “nugget” or “jitter” term of the variographic kernel introduced in Equation 16;

5. Calculate a weights vector by pre-multiplying the inverse covariance matrix obtained in Step 4 by the covariance vector between training and prediction points, calculated in Step 3;
6. Obtain the mean (in the average sense) predictions by multiplying the weights vector by the vector of known outputs from the training sample and summing over the components. This means that the mean predicted value is some form of a linear combination (dot product) of the known outputs, mediated by their respective covariance-driven weights;
7. Repeat Step 6, replacing the vector of known outputs by the covariance vector between training and prediction points obtained in Step 3, to obtain the predictive variance. This is a distinctive feature of the Kriging method in terms of uncertainty quantification. Uncertainty will be higher in regions of the input space with sparse sampling, but will be bounded to a finite value as per the discussion of the “sill” and “range” terms in Equation 16.

Application of Steps 1 to 7 as described indicates two distinctive features that are worth highlighting:

- a) while the final estimation step (for both predictive mean and predictive variance) is a linear combination, the kernel is flexible enough to capture a wide gamut of non-linear behaviors, similarly to the discussed via Equation 14 for the Radial Basis approach;
- b) a linear algebra based procedure but, given the additional parameters C_0, C_1 and a , their values should be simultaneously optimized to satisfy criteria pertinent to the context, as for example (and very usually) reducing predictive variance.

As far as the optimization of aspect “b” above goes, re-framing the Kriging formulation in terms of its counterpart method known as Gaussian Process (GP) offers additional options in terms of flexibility and performance. Regardless, given the similarity between Kriging and GP, they are often used and named interchangeably, and the following formulation may allude to both:

$$C(D) = \sigma_f^2 \cdot \exp\left(-\frac{D}{2l^2}\right) + \sigma_n^2 \cdot \delta(D) \quad (17)$$

where the covariance structure depends on a functional variance σ_f^2 and a noise term σ_n^2 , the variance parameters of two multi-variate gaussian distributions (hence the name of the method) that represent the estimates and the error, respectively. The additional parameter l is the length scale and, overall, the presence of the negative exponent that defaults to 1.0 for distances equal to zero carries over from Equation 12. Lastly, the δ operator introduces the noise term σ_n^2 only outside of the main diagonal of the resulting covariance matrix, obtained with the use of Equation 17 to “kernelize” the geometric distance matrix D , indicating that the method interpolates exactly over the points known from the training sample.

With the above considerations, the task of fitting the model (i.e., identifying the values of the parameters σ_f^2, σ_n^2 and l) morphs into finding the multivariate gaussian distributions whose variances should be equal to σ_f^2 and σ_n^2 . In classical statistical practice, the determination of probability distributions that best fit to data is known as Maximum Likelihood Estimation (MLE for short), expressing the procedure of finding the probability distribution(s) that most likely give rise to the observed data (Millar [2011]). For this purpose, a new procedural combination of matrix algebra and parameter optimization ensues as follows:

1. Calculate the covariance matrix by “kernelizing” the geometrical distance matrix D according to Equation 17;

2. Invert the covariance matrix obtained in Step 1;
3. Again involving prediction, use Equation 17 to “kernelize” the vector(s) of distances between known training points and unknown one(s), converting it into the covariance vector between training and prediction points;
4. To obtain the weights, pre-multiply the covariance vector between training and prediction points (obtained in Step 3) by the inverse covariance matrix from Step 2;
5. To obtain the predictive mean, perform the dot product between the weights arising from Step 4 and the output training vector;
6. Multiply the covariance vector between training and prediction points obtained in Step 3 by its transpose (equivalent to the dot product of this vector by itself, hence resulting in a scalar);
7. Subtract the value obtained in Step 6 from any of the equal values that belong to the main diagonal of the covariance matrix obtained in Step 1. The variances go along this main diagonal, and this subtraction is an adjustment that leads into the predictive variance at the inference point.

Once again, matrix algebra operations provide both the predictive mean and variance, that is, an estimate and the associated uncertainty. However, since the entire procedure is influenced by the values of σ_f^2 , σ_n^2 and l , they need to be estimated in the MLE sense via Equation 18:

$$\max(\log(P(y|x, \{\theta\}))) = -\frac{1}{2}\{y\}^T[C]^{-1}\{y\} - \frac{1}{2}\log(|C|) - \frac{N}{2}\log(2\pi) \quad (18)$$

which indicates that the negative log-likelihood of observing outputs y given inputs x and model parameters $\theta = \{\sigma_f^2, \sigma_n^2, l^2\}$ is maximum for values within θ that influence the 3 terms in the right-hand side of the equation, which depend only on the known training outputs y , the covariance matrix C (where the parameters $\theta = \{\sigma_f^2, \sigma_n^2, l^2\}$ matter) and the sample size N .

Code 7 implements the rather equivalent Kriging/Gaussian Process approaches in R.

Code 7: Kriging model

```
library('DiceKriging')

x = data.frame(matrix(rnorm(200), nrow=100, ncol=2))
colnames(x) = c('x1', 'x2')
noise = rnorm(100)
y = 2*x[,1] + (0.01*x[,2])^2 + 0.5*noise

model3 = km(design = x, response = y)

z = data.frame(matrix(c(1,1), nrow=1, ncol=2))
colnames(z) = c('x1', 'x2')
fc = predict.km(object = model3, newdata = z, type = 'SK')
cat('Prediction: ', fc$mean)

Prediction: 1.118517
```

Further exploration of the idea of kernels, with model fitting procedures that combine linear algebra matrix manipulations with parameter optimization, allows the opportunity to consider yet another method: neural networks. Although there are numerous variations of the fundamental concept, and significant recent expansion within the specialized application field of Deep Learning (Goodfellow et al. [2016]), the focus on the aforementioned operating principles pivots towards many, and typically simpler kernels, combined with minimization of the predictive error via a mechanism known as back-propagation. Providing materiality to these concepts, consider the (shallow) neural network topology in Figure 5.

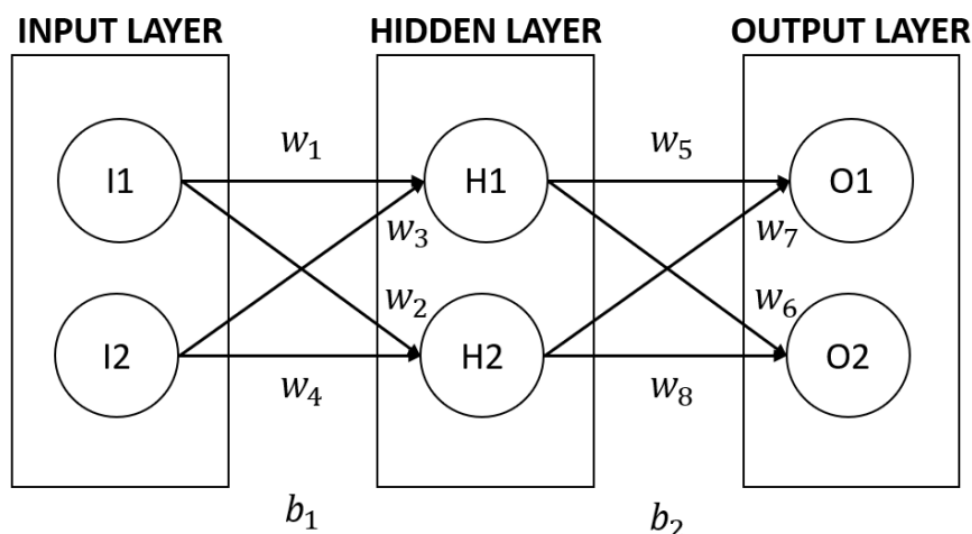


Figure 5: Shallow neural network topology, with input and output layers surrounding one, and only one hidden layer (hence, shallow), all of them containing 2 neurons (kernels) each.

The two input kernels/neurons $I1$ and $I2$ receive the values from the input variables, so that there are as many neurons in the input variables as the dimensionality of the predictive problem itself. The inputs are then transmitted into the hidden layer(s), whereas a single one is capable to map relationships between inputs and outputs at any arbitrary degree of non-linearity.

This transmission step occurs over weighted channels w_1, w_2, w_3 and w_4 such that, in the general case, the actual number of channels corresponds to the number of neurons in the input layer (i.e., the number of input variables) multiplied by the number of neurons in the hidden layer (arbitrarily chosen by the modeler, aiming at balancing bias and variance type of errors).

Over the course of this transmission, an optional bias term b_1 can be added to offset the linear combination of the input values and the transmission weights, as indicated in Equation 19:

$$\begin{aligned} H1_{input} &= w_1 I1 + w_3 I2 + b_1 \\ H2_{input} &= w_2 I1 + w_4 I2 + b_1 \end{aligned} \quad (19)$$

Differently than in the input layer, the neurons/kernels in the hidden layer process their inputs before transmitting them forward, utilizing some non-linear activation function mostly intended to saturate a potentially wide range of variation of the input values into a narrow and normalized (between 0 and 1, or -1 and 1) band. A prevalent choice for this role is the sigmoid function recalled by Equation 20. At this point, the kernel principle whereby a linear combination undergoes

a non-linear transformation is repeated.

$$H1_{output} = K(H1_{input}) = \frac{1}{1 + \exp(-H1_{input})} \quad (20)$$

$$H2_{output} = K(H2_{input}) = \frac{1}{1 + \exp(-H2_{input})}$$

After this non-linear transformation via the kernel K , the same process taking place between the input and the hidden layer takes place between the hidden layer and the output one, with generation of the corresponding input via Equation 19, but utilizing weights and bias w_5, w_6, w_7, w_8 and bias b_2 as indicated in Figure 5, and output via Equation 20.

The final output thus produced is compared to the true output values y_1 and y_2 at the 2 arbitrarily chosen neurons/kernels at the output layer, generating an error signal that triggers the core of the neural networks model building: the backpropagation method. The naming is owed to the fact that the error is back-propagated into the previous layers, in reverse order, and their respective weights and biases are re-adjusted so that in the next forward pass (again involving Equations 19 and 20 across all layers, as described above), the final error is reduced. The process repeats until convergence, and the following detailed backpropagation steps apply for a network with the topology as in Figure 5 and a neuron/kernel activation function as in Equation 20:

1. Calculate the total error in the differentiable form of the quadratic loss function in Equation 21, accumulating over the i neurons in the output layer:

$$E_{total} = \frac{1}{2} \sum_i (y_i - O_{output,i})^2 \quad (21)$$

2. Express the sensitivity of the total error calculated in Equation 21 with respect to the weights of the output layer, applying the chain rule of derivation as in Equation 22,

$$\frac{\partial E_{total}}{\partial w} = \frac{\partial E_{total}}{\partial O_{output}} \cdot \frac{\partial O_{output}}{\partial K} \cdot \frac{\partial K}{\partial w} \quad (22)$$

where, in this specific sequence: the total error depends on the outcome of the output layer, which depends on the kernel transformation at the same layer, which depends on the input it received based on the application of the weights from the previous (hidden) layer.

3. Calculate the 3 derivative terms of the chain rule in Equation 22 which, for a sigmoid activation/kernel function as in Equation 20, results in Equations 23-25 below:

$$\frac{\partial E_{total}}{\partial O_{output}} = -(y_i - O_{output,i}) \quad (23)$$

$$\frac{\partial O_{output}}{\partial K} = O_{output,i} \cdot (1 - O_{output,i}) \quad (24)$$

$$\frac{\partial K}{\partial w} = H_{output,i} \quad (25)$$

4. Multiply the 3 terms of the chain derivative expressed in Equations 23-25 to calculate the expression stated in Step 1;
5. Update the weights by using any of the optimization methods discussed in chapter "An overview of Linear and Non-linear Programming methods for Structural Optimization" (Choze et al. [2022b]), given the availability of its total amount (Equation 21, as calculated in Step 1) and its sensitivities in analytical form (Equations 23-25, corresponding to

Step 3). Although any optimization method can be used, the simplest line search in the direction of the gradient with step size/learning rate α , as expressed in Equation 26,

$$\Delta w = -\alpha \frac{\partial E_{total}}{\partial w} \quad (26)$$

is the most usual, given the convenience of the derivatives availability (some implementations will make the value of α adaptive for efficient convergence). On the other end of the spectrum, Boltzmann Machines is the usual nomenclature for neural network models fitted when genetic algorithms are used to perform this step.

6. Run the backpropagation of errors one further layer backwards, that is, repeat the procedure between Equations 21 and 26 for the connection between the input and the hidden layer. Because the total error actually carries over from the output layer, an extra implicit derivative must be accounted for, and the term corresponding to Equation 23 is further unpacked as shown in Equation 27, for every neuron j in the hidden layer and i in the output layer:

$$\frac{\partial E_{total}}{\partial H_{output,j}} = \sum_j \frac{\partial E_{output,i}}{\partial H_{output,j}} \quad (27)$$

and each term in the summation is a repetition of Equation 20, implicit within its first term (Equation 23). The two remaining terms are calculated as before, per Equations 24 and 25, giving rise to the sensitivity of the total error with respect to the weights;

7. Repeat Step 5 with the sensitivity of error versus weights obtained in Step 6, finalizing the backward pass and getting ready to initiate the next forward pass (Equations 19 and 20) if the method is to proceed for an additional iteration.

Requirements for the amount of training data and number of iterations until successful convergence vary on a case-dependent basis, but tend to be high as compared to other methods. Nevertheless, the wall clock time for training neural networks is reduced due to the simplicity of each of the individual operations.

As a concluding comment about neural networks, the arbitrary number of neurons in the hidden layer is equivalent to the number of terms in polynomial least-squares regression models (Equations 1 to 5), and hence should be chosen to balance between over-fitting (excessive number of terms leading to increased variance driven predictive errors) and under-fitting (insufficient number of terms leading to bias driven predictive errors). The practitioner is especially cautioned relative to a natural tendency for over-fitting, barring extreme low number of hidden neurons, and appropriate metrics to gauge this phenomenon are offered in the next section about goodness-of-fit. Many variations of the basic Neural Network method are available in the R statistical computing platform, with one of them illustrated in Code 8.

Code 8: Radial Basis Function (RBF) network

```
library(RSNNS)

x = matrix(rnorm(200), nrow=100, ncol=2)
noise = rnorm(100)
y = 2*x[,1] + (0.01*x[,2])^2 + 0.5*noise

model4 = rbf(x, y, size=40, maxit=1000)

z = matrix(c(1,1), nrow=1, ncol=2)
predict(model4, z)

1.864962
```

5 Ensemble Methods

Whereas each one of the modeling approaches described in section 4.1-4.2 individually possess their benefits and drawbacks, the practice of building model ensembles assumes that any individual models are “weak predictors”, which can be enhanced by some combination strategy capable of collectively downplay the shortcomings and amplify the strengths.

Operationalization of models ensembling strongly relies upon the theory of statistical resampling (Chihara and Hesterberg [2018]) whose primary techniques, the jackknife (deterministic sub setting, performed without replacement) and the bootstrap (random sub setting, performed with replacement) both leverage the most fundamental topic of sampling distributions (Devire [2015]). In particular, bootstrapping is the preferred method for mainstream model ensembling due to its ability to balance sample size (recalling that shortage of data is more likely than not), introduction of variability for model diversity (otherwise the very purpose of ensembling is defeated) and, depending upon specific circumstances, independence of the sub-samples. Within this construct, ensemble modeling techniques fall within the following categories.

5.1 Parallel ensembling

The main steps are:

1. Perform the bootstrapping, extracting S sub-samples out of N total available points;
2. Fit S independent (weak) learners, utilizing each of the individual subsets obtained in Step 1;
3. Summarize the S independent (weak) learners into a combination, which is typically an average for predictive models whose output is a continuous random variable and either an average (soft combination) or a majority vote (hard combination) if the output is categorical.

When performing Step 1, the goal is to generate as many subsets as possible, provided that each one is still large enough to allow for reasonable individual models to be produced, and at the same time minimizing the correlation (i.e., repetition of data that gets resampled with replacement) across samples. Certainly, satisfactory fulfillment of all these conflicting requirements depends upon a large enough N to start with.

For Step 2, it is usual to choose each independent learner to be of the same type before they are “bootstrap aggregated” or “bagged” in Step 3. This homogeneity does not mean that the intended diversity of the ensembling is compromised, not only because each individual learner arises from a different (and ideally independent) data set, but also due to parameter choices that can add to achievement of this goal. For example, it is common to balance individual learners that overfit (residing at the right hand end of the spectrum in Figure 3) and underfit (residing at the left hand end of the spectrum in Figure 3) so that the resulting ensemble best balances the bias-variance trade-off. Moreover, it is also common practice to fit the individual weak learners with different feature sets (i.e., different versions of the matrix of independent variables $[X]$) so that additional diversity is obtained.

Last but not least, the most important aspect of Step 3 is finding the optimal weighting strategy for each individual weak learner, which is possible by applying variations of Equations 4 and 5 when regularization is or is not considered, respectively. In either case, the weights are decision variables for minimizing a suitable loss function, ideally out of an independent data set used exclusively for testing purposes (additional details on candidate metrics presented in section 6 about goodness-of-fit appraisal).

5.2 Sequential ensembling

Steps 1 and 2 for sequential ensembling occur as in the parallel methods, and the fundamental difference in Step 3 is that, instead of solving a single instance of weights optimization, many of them are resolved in sequence as in Equation 28:

$$EM_i = EM_{i-1} + w_i \cdot WM_i \quad (28)$$

where the ensemble model EM at the i -th iteration is updated from the previous one by adding a weighted contribution of one of the possible weak learners WM , such that the loss function is minimized. Because the addition of each further learner “boosts” the previous one(s), and this is done adaptively at each iteration, this approach is often referred to as “adaptive boosting” or “adaboost” for short, and can be generalized as in Equation 29 for a loss function L :

$$\operatorname{argmin}_{w, WM} \left(\sum_{i=1}^N L \left[\hat{Y}, EM_{i-1}(X) \right] + w_i \cdot WM_i \right) \quad (29)$$

When the steepest descent (Vanderplaats [1998]) optimization method is applied to solving Equation 29, relying solely on information about the gradient of the loss function from any iteration to the next, the so-called gradient boosting method (GBM, for short) is configured, and Equation 30 represents it accordingly:

$$EM_i = EM_{i-1} - w_i \cdot \nabla E_{i-1} \quad (30)$$

where the weights w_i function as the steps to be taken along the line search following the direction of the gradient of the error E in the previous ensemble configuration. It is assumed that the loss function is continuously differentiable and, when it incorporates regularization terms, along the lines of Equation 5, the nomenclature xgboost (for “extreme” gradient boosting) is applied.

5.3 Multi-level ensembling

A common thread between parallel and sequential ensembling strategies is that 1) both processes are guided through minimization of a loss function and 2) it is typical, although not mandatory,

that models of the same type are combined together. Conversely, multi-level ensembling techniques will 1) fit a meta-model, at another level (hence the nomenclature), combining all existing learners to match given response values and 2) typically, although not mandatory, combine models of different types in the process. Equation 31 represents the idea in matrix notation, with the solution of the system of k simultaneous equations yields the coefficients/weights w that combine alternative weak learners to best match the ground truth $\{Y\}$.

$$\begin{bmatrix} w_1 \cdot \hat{Y}_1^{WM_1} & w_2 \cdot \hat{Y}_1^{WM_2} & \dots & w_k \cdot \hat{Y}_1^{WM_k} \\ w_1 \cdot \hat{Y}_2^{WM_1} & w_2 \cdot \hat{Y}_2^{WM_2} & \dots & w_k \cdot \hat{Y}_2^{WM_k} \\ \vdots & \vdots & \ddots & \vdots \\ w_1 \cdot \hat{Y}_k^{WM_1} & w_2 \cdot \hat{Y}_k^{WM_2} & \dots & w_k \cdot \hat{Y}_k^{WM_k} \end{bmatrix} = \begin{Bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_k \end{Bmatrix} \quad (31)$$

Reflecting the increased popularity of ensemble surrogates, the R statistical platform incorporates many of the methods, with Code 9 containing the Random Forest in particular:

Code 9: Random Forest algorithm

```
library(randomForest)

x = matrix(rnorm(200), nrow=100, ncol=2)
noise = rnorm(100)
y = 2*x[,1] + (0.01*x[,2])^2 + 0.5*noise

model5 = randomForest(x, y, ntree = 500)

z = matrix(c(1,1), nrow=1, ncol=2)
predict(model5, z)

1.905261
```

6 Goodness-of-Fit Appraisal

Complete appraisal of goodness-of-fit depends on two complementary angles: model recall (its ability to accurately estimate the outputs in training data) and generalization (its ability to accurately estimate the outputs in independent test data, not utilized during training). Whereas generalization is evidently more important for a final determination of the model qualities, recall analysis should be also considered as a tool for building understanding, helpful to improve the model overall. For this purpose, the foremost recall metric is the coefficient of multiple determination, shown in Equation 32:

$$R^2 = \frac{SS_R}{SS_Y} = 1 - \frac{SS_E}{SS_Y} \quad (32)$$

It defines the percent ratio between the explained variance (by the model) and the total variance, and is calculated as follows:

1. Extract the difference between all known training output Y and their average. Square these differences and add them together to obtain SS_Y ;
2. Repeat Step 1, but for the differences between actual and estimated (by the model) Y . These model errors (E) are the residuals, and this step obtains their sum of squares;

3. Using the principle of variance additivity, obtain the sum of squares of regression (R) by subtracting the term in Step 2 from that in Step 1;
4. Possessing all terms in Equation 32, apply it to calculate the coefficient of multiple determination.

The adjusted coefficient of multiple determination, on its hand, places effort on recognizing the principle of parsimony, and penalizes the metric in Equation 32 as the model acquires additional terms (i.e., operates on higher dimension input space). Denoting the number of model terms as p , it is possible to use this adjusted form of the coefficient of multiple determination to understand how it balances with respect to the available number of training samples N :

$$R_a^2 = 1 - \left(\frac{N-1}{N-p} \right) \cdot (1 - R^2) \quad (33)$$

The structure of Equation 33 is such that the need to fit additional parameters from a fixed training sample size is penalized, and the effect is similar to that from Equation 5 which regularizes the regression by eliminating less or non-essential terms.

Supplementing Equations 32 and 33, the Analysis of Variance (ANOVA for short) offers an additional way into the sources of error from random variation and also terms in the model/dimensionality of input space, as shown in Table 5.

Table 5: Analysis of Variance (ANOVA) for measuring goodness-of-fit.

Source of Variability	Statistical degrees-of-freedom (dof)	Sums of Squares (SS)	Mean Squares (MS)	F-Statistic
Model	$p - 1$	Step 3, R^2 calculation		
Residuals	$N - p$	Step 1, R^2 calculation	$\frac{SS}{dof}$	$\frac{MS_{model}}{MS_{residual}}$
Total	$N - 1$	Step 2, R^2 calculation		

It should be noted that Mean Squares are actually variances, and the F-Statistic, as a ratio between variances, follows the Snedecor probability distribution and, in this context, measures the probability that the model is only able to explain an amount of the total variance which is statistically equivalent to the residuals variance. In other words, the F-Statistic measures the probability that the model can only explain as much of the data as the error itself, and is hence useless. With this interpretation, adequate goodness-of-fit requires this number to be as low as possible, contrary to the coefficient of multiple determination and its adjusted form.

The possibility of calculating model generalization metrics assumes that there are independent observations, not consumed for the purposes of model training, that can be spared. This tends to be a strong assumption, considering the considerations made in section 3 on experimental design and the challenges of balancing data availability and model quality. For this reason, traditional partitions such as sparing some percentage (typically 20% - 30%) of the overall data for testing will only perform when there is abundance of data, usually not the case with high fidelity simulations of structures. The solution then resides in Cross-validation, which operates according to the following algorithm:

1. Create a fold, by sparing a small percentage of the available data for testing, the rest being dedicated to training;
2. Train the model with the training data portion (majority) of the fold determined in Step 1;

3. Test the model with the testing data portion (minority) of the fold determined in Step 1 and calculate performance metrics;
4. Repeat Steps 2 and 3 until all possible folds are exhausted;
5. Analyze the distribution of the performance metrics calculated in Step 3 over the many folds.

This approach assumes that there will be no replacement of the data across the folds and, at the limit, the LOOCV (Leave-One-Out-Cross-Validation) will happen when only one data point is retained for testing within each fold. This is the most time-consuming option, but the one that yields the most complete possible version for the distribution of whatever performance metric is chosen in Step 3. As for the metrics themselves, the RMSE (Root Mean Squared Error) is a typical choice, and it is defined in Equation 34 as the square-root of the average quadratic error between true values (Y) and their model estimates (\hat{Y}):

$$RMSE = \sqrt{\frac{\sum_N (\hat{Y} - Y)^2}{N}} \quad (34)$$

It should be noted that RMSE is not the only metric applicable to appraise model generalization, albeit it is by and large the most used one, more so in the context of cross-validation described herein. Other alternatives consist of calculating variance explained type of metrics, similar to those in Equations 32 and 33, but using data that was not directly used to build the model in the first place. Hence, it is conceivable to calculate the sum of squares of test residuals (i.e., those arising from independent data) as inputs to such equations for estimating the portion of the overall variance not explained by the model. Caution should be exercised though with the basis for comparison, which should remain standardized so that such metrics are meaningful. One of the approaches to address this concern is to use the “out-of-sample” (OOS) version of Equation 35, which is defined as follows:

$$R_{OOS}^2 = 1 - \frac{RMSE_m}{RMSE_b} \quad (35)$$

and depends upon $RMSE_m$, calculated for the model using Equation 34, and an ideal (possibly arbitrarily set) benchmark $RMSE_b$ for a real or hypothetical model. It should be noted that, contrary to the $RMSE$ defined in Equation 34, $-\infty \leq R_{OOS}^2 \leq 1$ and a value of 0 represents a tie between the model being appraised and the benchmark. Conversely, positive/negative values signify that the model outperforms/underperforms the benchmark.

A totally different perspective, which is more solid but difficult to automate, counts on understanding the error structure of the models in detail rather than relying solely upon summary metrics such as any of those offered in Equations 32 to 35 and Table 5. Models that have exhausted all possible information contained in the data yield errors that behave as random numbers, ideally centered around zero (indicating absence of bias) and with small spread (indicating absence of variance). These two features allude to the error components depicted in Figure 3 and it expected that one can only be improved at the expense of the other, forcing a trade-off. Nevertheless, the two characteristics can be combined to some extent when the errors are Normally distributed (i.e., behave as random numbers that follow the Normal probability distribution with mean 0 and some variance), and then the exercise of appraising goodness of fit defaults to measuring deviation/adherence of the modeling errors with respect to the Normal reference.

Following this rationale, specific techniques may be more visual, such as those illustrated in Figures 6, 7 and 8, or rather parametric (and easier to automate), such as the tests of Normality summarized in Table 6.

Figure 6 shows distinct scenarios in which the residuals distributions vary from closely adhering to random (Figure 6a) to exhibiting multiple symptoms of Normality violation (Figure 6b). In the former, uniformly scattered points representing the residuals spread symmetrically around zero, without revealing any relevant pattern indicative of non-randomness. In the latter, on the contrary, multiple patterns arise by way of repeated ascending/descending patterns, as well as lack of symmetry and outliers, indicating that multiple processes are manifesting themselves in tandem, rather than a single random distribution representative of a good model fit. The particular case explored in Figure 6c represent the case in which the model lacks complexity to represent a linear trend, and is hence under-fitted. Generalizations of this concept could show very visible and identifiable patterns such as parabolas or other shapes, which would be associated with even more severe under-fitting other than just missing a linear trend. Finally, the scenario depicted in Figure 6d is the most subtle and potentially damaging one, because an apparently random residual distribution is actually strongly biased towards positive. Realistic applications will pose the additional challenging of combining all these types of archetypal behaviors into a single residual distribution, making it very complex to issue an appraisal about goodness of fit.

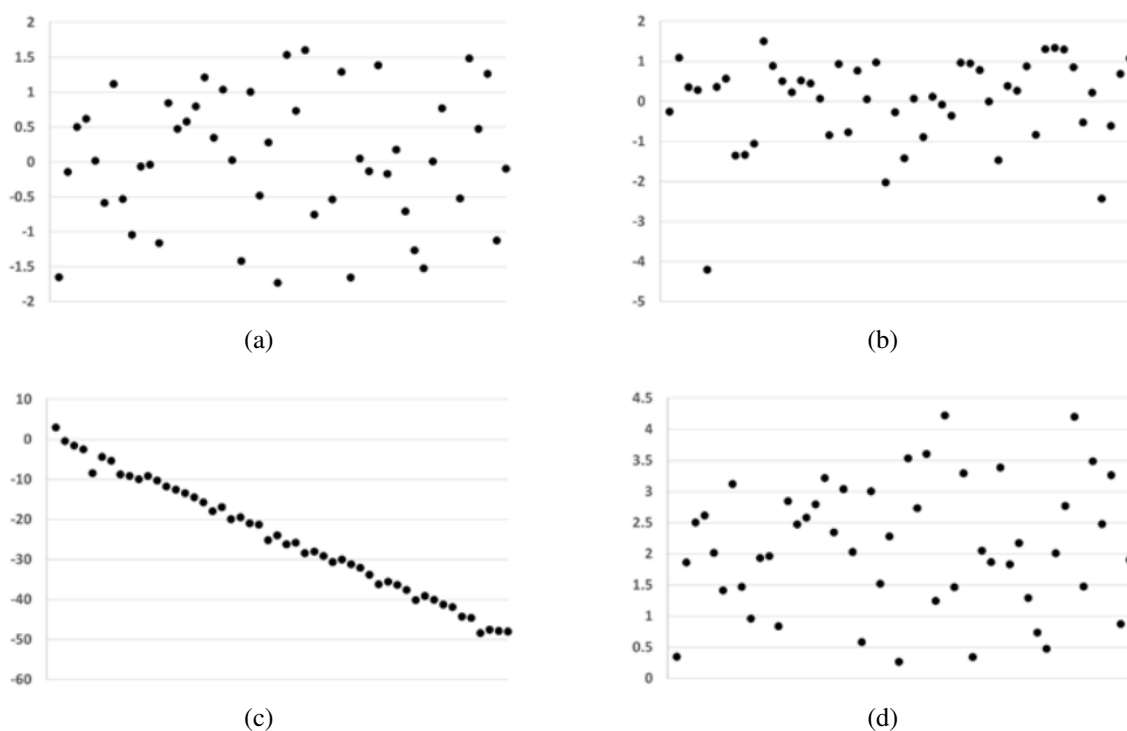


Figure 6: Different patterns of residuals for goodness of fit analysis.

Panels (6a) and (6b) in Figure 6 are in opposite ends of the spectrum with respect to residuals that closely adhere to adequate versus poor goodness of fit, respectively. Since such verdicts are seldom this binary, Figure 7 offers an additional appraisal instrument by plotting the residuals versus lagged versions of themselves. With $lag = 1$ in these particular plots, it is possible to confirm the absence of randomness in panel (7b), which corresponds to the same panel in Figure 6 and, at the same time, certain patterns arise in panel (7a), again homologous to the same in Figure 6. For Figure 7, the lagged plot is powerful enough to reveal that the data from Figure 7a is actually a random draw from a Normal distribution with mean 0 and variance 1 ($N(0, 1)$) distribution with only 50 points, which are still not sufficient to display full blown randomness as it would have been the case with a much larger sample size, but still satisfactorily random/Normal on a relative

scale.

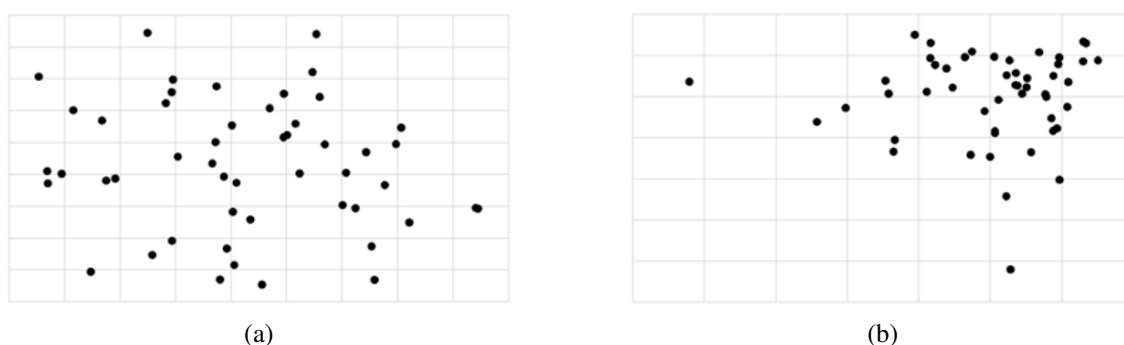


Figure 7: Plot of lagged residuals for goodness of fit analysis.

On a related note, the lagging period can be generalized to other values and, at the limit, their collection at several values gives rise to auto-correlation and partial auto-correlation plots, which are paramount in time-series analysis (Shumway and Stoffer [2017]) and, in the current context, would reveal additional details about the residuals structure for the purposes of model appraisal.

Figure 8 again looks at the comparison between panels (a) and (b) from Figures 6 and 7, however analyzing the very same data through a different instrument: the (Normal) probability plot, which are generated as follows:

1. Sort the residuals in ascending order;
2. Estimate the probability of each residual value based on its rank order by dividing the position in the sorted sequence by the total number of residuals N . Per this method, the estimated rank order probability for the n -th residual in the sequence is n/N , and the last value in the series results in 1 (100% of the residuals covered, producing an empirical cumulative probability distribution);
3. Calculate the z-scores (distances from the mean normalized by the standard deviation) using the probabilities obtained in Step 2 and the inverse cumulative version of the distribution of reference. If the inverse Normal cumulative is chosen, the empirical Normal z-scores are obtained;
4. Calculate the adherence of the cloud of z-scores obtained in Step 3 to the theoretical distribution (Normal, in the current context) by fitting a straight line and measuring the goodness of fit through the R^2 (as in Equation 32).

Specifically, in Figure 8, panel (a) confirms a highly satisfactory adherence to normality (R^2 nearing 100%), whereas the various violations of Normality observed in panel (b) in Figures 6 and 7 are confirmed by a very low R^2 .

The way of measuring adherence to Normality through a metric, like the R^2 in Figure 8, is extending the concept through formal Hypothesis testing, whereby a statistic is calculated and compared to a critical value to determine if the “null hypothesis” (residual data follows the reference distribution, to be assumed as Normal) can be accepted and at which confidence level. Several versions of this principle are summarized in Table 6, which lists the most commonly used tests of Normality, applicable to the specific purpose of measuring goodness-of-fit through the $i = 1, \dots, N$ available values of residuals denoted as ε .

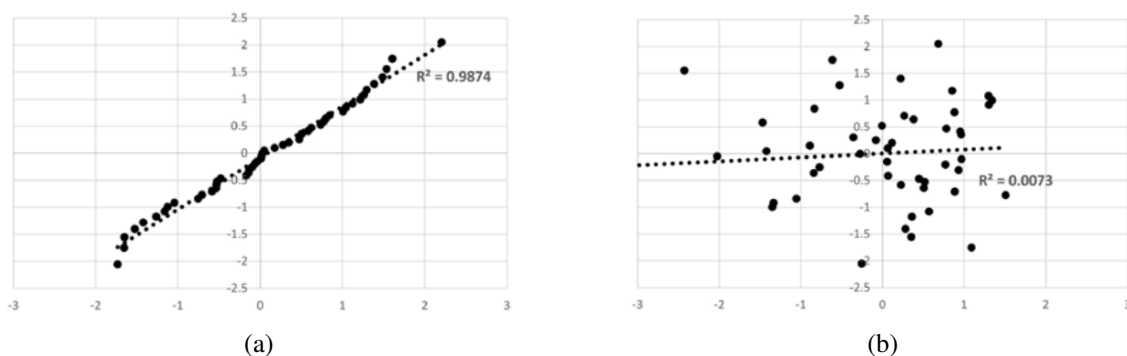


Figure 8: Normal Probability Plot, also known as quantiles plot, for goodness of fit analysis.

In every case, F and C are the reference (i.e., standard Normal with zero mean and unit variance) and empirical (i.e., from the residual data itself) cumulative probability distribution functions (again, their comparison relates to the procedure used to produce Figure 8).

To conclude this section on goodness-of-fit, residuals analysis is presented as the ultimate ground truth in appraising if a model is or is not suited for a particular application. Even for model types such as Kriging and Gaussian Processes, which possess inherent uncertainty quantification features as displayed at their respective descriptions (Equations 16 to 18 and associated content), the application of the metrics shown in this section, and more emphatically cross-validation and residual analysis, are thoroughly recommended to the practitioner.

Table 6: Collection of usual tests of hypothesis to assess goodness of fit through Normality.

Anderson-Darling Can be specified to reference distributions other than Normal by generalizing F :

$$A^2 = -N - S$$

$$S = \sum_{i=1}^N \frac{(2i-1)}{N} \cdot [\ln F(\varepsilon_i) + \ln(1 - F(\varepsilon_N + 1 - i))]$$

Liliefors Utilizes the supremum (sup) operator, which finds the lowest upper bound of a set of values:

$$L = \sup_{\varepsilon} |F(\varepsilon) - C(\varepsilon)|$$

Kolmogorov-Smirnov Can be specified to reference distributions other than Normal by generalizing F :

$$D = \max_{1 \leq i \leq N} \left(F(\varepsilon_i) - \frac{i-1}{N}, \frac{i}{N} - F(\varepsilon_i) \right)$$

Kulbach-Leibler Divergence In this formula, already particularized for the Normal case, but also possible to be generalized depending on the density function chosen for pre-multiplying the natural logarithm and appearing in its argument's denominator. The lower case $c(\varepsilon)$ in the natural logarithm argument's numerator is the empirical density:

$$D_k(P|\mathbb{N}) = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{\varepsilon-\mu}{\sigma}\right)^2} \cdot \ln \left[c(\varepsilon) \cdot \left\{ \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{\varepsilon-\mu}{\sigma}\right)^2} \right\}^{-1} \right] d\varepsilon$$

Jarque-Bera Strictly for Normality test, does not depend on a reference distribution, but rather on the residuals skewness S and kurtosis k :

$$JB = \frac{n}{6} \left[S^2 + \frac{1}{4}(k-3)^2 \right]$$

Chi-Squared The observed (O_i) and expected (E_i) terms correspond to the number of residuals counted within a certain bin or range i (expected are per the Normal distribution for using residuals to appraise goodness of fit, but may also be generalized to other distributions). The χ^2 itself is a random variable arising from the sums of squares of Normals, as implied by the formula:

$$\chi^2 = \sum_{i=1}^N \frac{(O_i - E_i)^2}{E_i}$$

References

- NIST/SEMATECH *e-Handbook of Statistical Methods*. URL <https://www.itl.nist.gov/div898/handbook/pri/section3/pri3347.htm>.
- R. Andersen. *Modern methods for robust regression*. Number 152. Sage, 2008.
- J.-F. Barthelemy and R. T. Haftka. Approximation concepts for optimum structural design - a review. *Structural optimization*, 5(3):129–144, 1993.
- A. Bhosekar and M. Ierapetritou. Advances in surrogate based modeling, feasibility analysis, and optimization: A review. *Computers and Chemical Engineering*, 108:250–267, 2018.
- F. Boukouvala and C. A. Floudas. ARGONAUT: AlgoRithms for Global Optimization of coN-strAined grey-box compUTational problems. *Optimization Letters*, 11(5):895–913, 2017.
- G. E. P. Box, J. S. Hunter, and W. G. Hunter. *Statistics for Experimenters: Design, Innovation, and Discovery*. Wiley-Interscience, 2nd edition, 2005.
- L. M. Chihara and T. C. Hesterberg. *Mathematical Statistics with Resampling and R*. Wiley, 2nd edition, 2018.
- S. B. Choze, R. R. Santos, and G. F. Gomes. Overview of traditional and recent heuristic optimization methods. In *Model-Based and Signal-Based Inverse Methods*, volume 1 of *Discrete Models, Inverse Methods & Uncertainty Modeling in Structural Integrity*, chapter 4, pages 107–142. University of Brasilia, 2022a. URL <https://doi.org/10.4322/978-65-86503-71-5.c04>.
- S. B. Choze, R. R. Santos, A. B. Jorge, and G. F. Gomes. An overview of linear and non-linear programming methods for structural optimization. In *Model-Based and Signal-Based Inverse Methods*, volume 1 of *Discrete Models, Inverse Methods & Uncertainty Modeling in Structural Integrity*, chapter 3, pages 65–106. University of Brasilia, 2022b. URL <https://doi.org/10.4322/978-65-86503-71-5.c03>.
- J. L. Devire. *Probability and Statistics for Engineering and the Sciences*. Cengage Learning, 9th edition, 2015.
- I. E. Frank. Modern nonlinear regression methods. *Chemometrics and intelligent laboratory systems*, 27(1):1–19, 1995.
- K. I. A. Garud, S. S. and M. Kraft. Design of computer experiments: A review. *Computers and Chemical Engineering*, 106:71–95, 2017a.
- K. I. A. Garud, S. S. and M. Kraft. Smart sampling algorithm for surrogate model development. *Computers and Chemical Engineering*, 96:103–114, 2017b.
- I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. Adaptive Computation and Machine Learning series. The MIT Press, illustrated edition, 2016.
- T. Hastie, R. Tibshirani, J. H. Friedman, and J. H. Friedman. *The elements of statistical learning: data mining, inference, and prediction*, volume 2. Springer, 2009.
- T. J. Hastie. Generalized additive models. In *Statistical models in S*, pages 249–307. Routledge, 2017.

- M. Kaufman, V. Balabanov, A. A. Giunta, B. Grossman, W. H. Mason, S. L. Burgee, R. T. Haftka, and L. T. Watson. Variable-complexity response surface approximations for wing structural weight in hsct design. *Computational Mechanics*, 18(2):112–126, 1996.
- A. I. Khuri and J. A. Cornell. *Response Surfaces: Designs and Analyses*. CRC Press, 2nd edition, 2019.
- R. B. Millar. *Maximum Likelihood Estimation and Inference: With Examples in R, SAS and ADMB*. Wiley, 1st edition, 2011.
- R. H. Myers, D. C. Montgomery, and C. M. Anderson-Cook. *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*. Wiley Series in Probability and Statistics, 4th edition, 2016.
- L. Pronzato and W. G. Müller. Design of computer experiments: space filling and beyond. *Statistics and Computing*, 22(3):681–701, 2012.
- R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2020. URL <https://www.R-project.org/>.
- R. H. Shumway and D. S. Stoffer. *Time Series Analysis and Its Applications: With R Examples*. Springer International Publishing, 4th edition, 2017. URL <https://doi.org/10.1007/978-3-319-52452-8>.
- J. Sobieszczanski-Sobieski and R. T. Haftka. Multidisciplinary aerospace design optimization: survey of recent developments. *Structural optimization*, 14(1):1–23, 1997.
- G. N. Vanderplaats. *Numerical Optimization Techniques for Engineering Design*. Vanderpaats Research and Development, 1998.

Chapter 4

Uncertainty Quantification in Inverse Kinematics Computation of Space Robots by Neural Networks

Chapter details

Chapter DOI:

<https://doi.org/10.4322/978-65-86503-88-3.c04>

Chapter suggested citation / reference style:

Santos, Rogerio R., et al. (2022). “Uncertainty Quantification in Inverse Kinematics Computation of Space Robots by Neural Networks”. In Jorge, Ariosto B., et al. (Eds.) *Uncertainty Modeling: Fundamental Concepts and Models*, Vol. III, UnB, Brasilia, DF, Brazil, pp. 89–118. Book series in Discrete Models, Inverse Methods, & Uncertainty Modeling in Structural Integrity.

P.S.: DOI may be included at the end of citation, for completeness.

Book details

Book: Uncertainty Modeling: Fundamental Concepts and Models

Edited by: Jorge, Ariosto B., Anflor, Carla T. M., Gomes, Guilherme F., & Carneiro, Sergio H. S.

Volume III of Book Series in:

Discrete Models, Inverse Methods, & Uncertainty Modeling in Structural Integrity

Published by: UnB City: Brasilia, DF, Brazil Year: 2022

DOI: <https://doi.org/10.4322/978-65-86503-88-3>

Uncertainty Quantification in Inverse Kinematics Computation of Space Robots by Neural Networks

Rogério R. Santos^{1*}, Ijar M. da Fonseca², and Domingos A. Rade³

¹Division of Mechanical Engineering, Technological Institute of Aeronautics - ITA, Brazil. E-mail: rsantos9@gmail.com

²Division of Aeronautical and Aerospace Engineering, Technological Institute of Aeronautics - ITA, Brazil. E-mail: ijar@ita.br

³Division of Mechanical Engineering, Technological Institute of Aeronautics - ITA, Brazil. E-mail: rade@ita.br

*Corresponding author

Abstract

Service satellite are used in a variety of applications, justified by scientific, economic, strategic and social benefits. The perspective of working with space equipment is largely based on the assumption that there are robotic mechanisms capable of handling multiple payloads. Normally, robot path planning involves determining the inverse kinematics configuration, since the workspace is expressed in Cartesian coordinates and the robot control is performed through joint angles. The present study discusses strategies for calculating the inverse kinematics of a space serial manipulator. The results of a classical industrial kinematics model and a multibody dynamics formulation are compared in the context of space operation. A Neural Network model is considered to quantify the uncertainty about the end-effector positioning and joint angle values. It was found that the Neural Network is a robust predictor that, together with a Nonlinear Programming strategy, represents an effective methodology for inverse kinematics calculations. This result contributes to the establishment of algorithms that will increase the operational autonomy of space equipment in different scenarios of orbital operation.

Keywords: Robot Path Planning, Berthing, Satellite Dynamics, Machine Learning, PyTorch, OpenModelica

1 Introduction

The movement autonomy of space robots is a contemporary challenge of great relevance. It consists in the ability of fulfilling goals without the need of external control operators. There is a wide range of applications for service satellites, and its relevance is justified by scientific, economic and strategic benefits.

Long-distance space travel requires the existence of space deposits along the way. The service satellites are enabling agents of such infrastructure. Another activity destined for the service satellite is the manipulation of objects in orbit, enabling the removal of space debris and maintenance of space vehicles in orbit. Additionally, as the positioning of satellites in the Geostationary Earth Orbit (GEO) depends on the availability of a free and viable position in orbit, there is commercial and security interest in removing failed or unused satellites for reuse from its orbital position. Additionally, the maintenance of geostationary satellites increases their lifespan and leads to better use of space.

The repair of space equipment is another important application of service satellites, for example, the Hubble Space Telescope's maintenance service mission (1993). The lifespan of valuable resources can be extended through maintenance and upgrade actions. From this perspective, we see the need to develop the equivalent of a "technical assistance network" in orbit to solve problems of critical failures in space equipment, leading to the interruption or complete loss of the space mission.

For the activities of providing services in orbit, in addition to the strategic importance regarding the domain of space exploration technology, it also involves the economic aspect of great relevance already identified by companies around the world. We live in a new phase of the space race characterized by the participation of private companies, focused on the economic exploitation of space resources.

NASA's Artemis program (Smith et al. [2020]) will send astronauts to the Moon through a global initiative in the engineering, technology development, and process improvement necessary to safely conduct human exploration to the Moon. This initiative includes the contribution of companies and international partners to the exploration and development of the Moon.

Companies such as SpaceX (Mróz et al. [2022]) and Blue Origin have been involved in the development of space transport equipment, focusing on economic viability through the reuse of launch vehicles. Boeing provides a series of resources oriented towards space exploration. Among the resources offered, there is the provision of a global positioning system infrastructure (Global Positioning System IIF) (Kuang et al. [2017]) composed of a large number of satellites. Boeing Commercial Satellites designs and markets communication and cargo satellites for commercial purposes, such as telecommunication infrastructure, broadband, and scientific applications (Skinner et al. [2013]).

Airbus presents on its website the possibility of space equipment maintenance services to be offered in the coming years. The "Airbus O. CUBED Services" will be operated by a category of space vehicle called "SpaceTug". The SpaceTug family of service satellites contains advanced technologies for robotics, electric propulsion, computer vision-based navigation, as well as approaching and intercepting strategies. On-orbit services (OOS) will be offered in three main categories:

- Inspection and maintenance service for telecommunications satellites;
- On-orbit servicing for geostationary orbit (GEO) and low-Earth orbit (LEO);

- Debris Removal Service from the Space vehicles orbits.

On-orbit vehicle maintenance and orbital debris removal services should operate autonomously (Aglietti et al. [2020]) based on intelligent strategies for moving and manipulating objects. The feasibility of this vision is largely based on the existence of robotic mechanisms that are able to act effectively in the manipulation of different objects. The view on the commercial importance of space exploration is confirmed by the large number of satellites in operation today. According to ESPI [2020a], the number of satellites operating in Earth orbit in March 2019 was 2062, distributed as follows:

- Communication: 773;
- Earth observation: 768;
- Scientific research: 105;
- Navigation: 138;
- Technology: 265;
- Other missions: 13.

In ESPI [2020b] the policies of different European countries for the development of space programs are presented, together with justifications on the strategic importance of such an initiative. According to the information in the report, two concerns common to the countries involved are:

- Development of technologies that lead to the autonomy of space operations, in order to establish situational awareness for action at the international level, that is, the ability to recognize any possible issues once you arrive at the scene and act proactively to avoid a negative impact;
- Expansion of knowledge and training of people involved in space activities as well as maintenance of the competence of specialists in the area of space security.

In this context, from a strategic point of view, the development of space exploration technologies is fundamental, in order to contribute to the maintenance of national sovereignty, as well as to contribute to the development of cutting-edge technologies with great industrial and commercial impact.

The paper is organized as follows. The Section 2 presents the main characteristics for robot kinematics representation. Inverse kinematics of space robots is discussed in Section 3. The Section 4 presents a popular interpolation scheme for path planning. The Section 5 outlines the principles involved in torque calculation, while the feature-rich approach of multibody dynamics is discussed in Section 6. Numerical experiments that quantify uncertainty are detailed in Section 7. The Section 8 summarizes the conclusions of the present study.

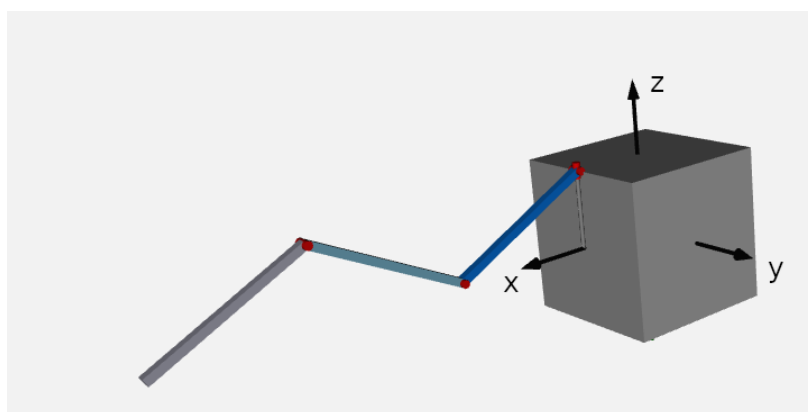


Figure 1: Space robot composed by a satellite and a serial manipulator.

2 Space Robots

The concept of space robot considered in the present paper consists of a space satellite with a serial manipulator attached to it. The schematic representation of the spaceship is given in Figure 1.

The kinematic model of the industrial serial manipulator describes the influence of joint angles in the placement of the end-effector. Usually, an n -link manipulator is mathematically represented by a set of homogeneous transformations $A_i, i = 1, \dots, n$ that describe the Cartesian positioning from the base $P_{base}(x, y, z)$ to the end-effector $P_{end}(x, y, z)$. Following this convention, $A_1 = T_0^1, A_2 = T_1^2, \dots, A_n = T_{n-1}^n$ and the serial manipulator can be represented by the matrix

$$T_0^n(q) = T_0^1(q_1)T_1^2(q_2) \dots T_{n-1}^n(q_n) \quad (1)$$

where q_i is the i -th joint coordinate, that is, $q_i = \theta_i$ for rotational joint or $q_i = d_i$ for translation joint (Craig [2005]).

In the context of applications, the paper by da Fonseca [2021] presents an overview of the space age and the first robotic probes preceding the landing of the astronauts on the Moon. It discusses space robotics fundamental concepts, classification, and safety-critical aspects for space robotics applications. The space robotics state of the art presented focuses on the International Space Station (ISS) Extravehicular Activities (EVA), the planetary explorations (like the current Mars exploration), JAXA's Hayabusa mission that rendezvoused and landed on the 25143 Itokawa asteroid, and in the ESA/DLR's Rosetta spacecraft carrying its Philae robotic module which landed on the 67P / Churyumov-Gerasimenko comet. The article includes a discussion of on-orbit servicing and rendezvous & docking/berthing (RVD/B).

2.1 Robot Kinematics

To determine the inverse kinematics of a serial robot manipulator involves the computation of the joint angles from a given Cartesian position and orientation of the end-effector.

This computation is of primary importance in the programming and control of robot manipulators. The corresponding procedure has to do with nonlinear algebraic equations,

considering that there is no analytical closed-form solution for a general robot structure. Consequently, it is necessary to establish a way to compute the inverse kinematics for a wide class of robots and related tasks that are assigned to them.

When the end-effector moves along a continuous trajectory, it is not possible to switch from one kind of solution to another, arbitrarily. Uchiyama [1979] pointed out that the trajectories of multiple solutions intersect at singular points and solutions can be switched only at those points. The singular points of a robotic mechanism are defined as the singular points of the implicit function, namely, the joint angles that make the Jacobian matrix not to have full rank. Besides, singular points can bring control difficulties and poor positioning accuracy.

Moreover, when working into complex environment, the manipulator can be subjected to work into restricted workspaces, both for sided constraints (like walls and floor), and for internal constraints (as the geometry of manipulated object). Effective optimization strategies to overcome such difficulties are proposed by dos Santos et al. [2008b], dos Santos et al. [2010], and Santos et al. [2022], among others.

Adopting the Denavit-Hartenberg convention (Spong and Vidyasagar [1989], Craig [2005]) to define A_i , any serial robot can be represented by Equation(1). Hence, given the base P_{base} and joint $q_i(d, \theta)$ values, the Cartesian position of the end-effector P_{end} is given by

$$P_{end} = T_0^n(q).P_{base}. \quad (2)$$

2.2 Inverse Kinematics

Adopting the Denavit-Hartenberg convention (Craig [2005], Spong and Vidyasagar [1989]) to define A_i , any serial robot can be represented by Equation 1. Hence, given the base P_{base} and joint $q_i(d, \theta)$ values, it is possible to determine the end-effector cartesian position P_{end} through the equality

$$P_{end} = T_0^n(q).P_{base}. \quad (3)$$

To solve the inverse kinematics problem means to solve the inverse problem, i.e., given P_{base} and P_{end} find $q_i(d, \theta)$ for which Equation 3 holds.

One approach to solve the inverse kinematics problem is to find a closed-form solution by using algebra and geometry. Although this approach is generally more desirable in applying the solution to real-time control of robots, it is not always possible to obtain closed-form solutions for manipulators with an arbitrary mechanism due to the complexity of Equation 3.

Rather, the class of manipulators mechanisms for which the closed-form solutions are guaranteed is very limited. The algebraic approach to closed-form solutions is related to finding q through various algebraic transformations of Equation 3; the geometric approach means finding q by using geometrical heuristics to take advantage of the special structure of the manipulator.

The paper by da Fonseca et al. [2019] considers the forward and inverse kinematics for a space robot. The main focus is to compute the position and orientation of manipulators' end-effectors relative to their platform. The set of coordinate frames follows the convention for frames that appears in the literature for ground robot manipulators. The kinematics related to the spacecraft attitude is added in the formulation, providing

an overview of the kinematics equations for spacecraft-type manipulators. The inertial, orbital, and body-fixed coordinate frames are included in the kinematics study.

2.3 Inverse kinematics as an optimization problem

Considering the manipulator kinematics described by Equation 3 and an end-effector point defined as P_{end} , then it is desired to get $q(d, \theta)$ values that obey Equation 3. However,

$$\begin{aligned} T_0^n(q).P_{base} = P_{end} &\Rightarrow T_0^n(q).P_{base} - P_{end} = \vec{0} \Rightarrow \\ |T_0^n(q).P_{base} - P_{end}| &= 0 \end{aligned} \quad (4)$$

Therefore, the optimal solution q^* (i.e. minimal) of the objective function given by

$$f_1(d, \theta) = |T_0^n(q(d, \theta)).P_{base} - P_{end}|^2 = 0 \quad (5)$$

is also solution of Equation 3, because $q(d, \theta)$ satisfies Equation 5 if, and only if, it satisfies Equation 3. It is an expression for the misalignment between the base and tool frames, that is, the transformation matrix align both systems of coordinates, the symbol $|\cdot|$ is a vector norm, and the squared refers to the quadratic error (dos Santos et al. [2005], dos Santos et al. [2010]). Then, the inverse kinematics parameter $q(d, \theta)$ can be achieved through an optimization procedure by using Equation 5 as the objective function.

This approach has the main advantage of representing a generic procedure that can be developed for different architectures of serial robot manipulators. As will be shown, this approach leads to efficient computation of the inverse kinematics (positioning and orientation) by using global and local optimization methods.

Along the investigation of the optimization problem, there are two kinds of solution points (Luenberger and Ye [2008]): *local minimum points*, and *global minimum points*. The definitions below are related to these solution points.

Definition A point $q^* \in \Omega$ is said to be a *relative minimum point* or a *local minimum point* of f over Ω if there is an $\epsilon > 0$ such that $f(q) \geq f(q^*)$ for all $q \in \Omega$ and $|q - q^*| < \epsilon$. If $f(q) > f(q^*)$ for all $q \in \Omega$, $q \neq q^*$, $|q - q^*| < \epsilon$, then q^* is said to be a *strict relative minimum point* of f over Ω .

Definition A point $q^* \in \Omega$ is said to be a *global minimum point* of f over Ω if $f(q) \geq f(q^*)$ for all $q \in \Omega$. If $f(q) > f(q^*)$ for all $q \in \Omega$, $q \neq q^*$, then q^* is said to be a *strict global minimum point* of f over Ω .

As the inverse kinematics solution $q(d, \theta)$ of Equation 5 is also a global minimum, it is possible to obtain the inverse kinematics by means of a global optimization procedure.

The local minimum of Equation 5 can be easily obtained by classical nonlinear optimization methods, such as the Gradient method, the Conjugate Gradient method, and the Quasi-Newton based methods, among others. However, it can result a minimum which does not satisfy $f_1(q) = 0$. Then, Equation 5 does not hold. To avoid this situation, the initial point is changed and the optimization is performed until the global minimum is achieved.

To improve the robustness of the process, it is possible to use a global minimizer. There are many methods devoted to find a global minimum of a nonlinear objective function, despite that to find a global minimum is not an easy task. Well known methods such as genetic algorithms (Le Grand and Merz Jr. [1993]), differential evolution algorithm (Storn and Price [1997]), simulated annealing (Romeijn and Smith [1994]), or other heuristic (Choze et al. [2022a]) could be used in this case. The main characteristic of these methods is that the global (or near global) optimum is obtained through a high number of functional evaluations.

3 Inverse Kinematics of Space Robots

To solve the inverse kinematics problem means to solve the inverse problem, i.e., given P_{base} and P_{end} , find $q_i(d, \theta)$ for which Equation 2 holds. It will require the solution of the inverse problem

$$\min_{\theta_i} |T_0^n(q(\theta_i))P_{base} - P_{end}| \quad (6)$$

The Denavit-Hartenberg parameters of the manipulator, as seen in Figure 1, is shown in Table 1. The parameter a refers to a fixed translation while α represents a fixed rotation

Table 1: Denavit-Hartenberg parameters.

Joint	a (m)	α (rad)	d (m)	θ (rad)
1	0	1.57	0	θ_1
2	1	0	0	θ_2
3	1	0	0	θ_3
4	1	0	0	θ_4

of the coordinate system. The joint value d express a translation joint while θ represents the angle of a rotational joint (Craig [2005], Spong and Vidyasagar [1989]).

The classical formulation of the industrial device, Equation 2, contains the hypothesis that the base of the manipulator is fixed. Since the satellite is traveling in orbit and may floats under the reaction forces and torque arising from the joints' motion control, the system dynamics (satellite plus the attached manipulator) is more complex than that of manipulators dynamics on ground. Therefore, to contemplate the dynamic effects along the movement, a higher fidelity model is considered.

The high fidelity model is build using the OpenModelica software (Fritzson et al. [2005]), which provides a multibody dynamics library and will compute the coupled effects between the robot and the satellite while floating in the space. The parameters are shown in Table 2.

Given the displacement Δ of the satellite, the position $P_{effector}$ of the end-effector and the position P_{target} of the target, the inverse problem for obtaining the inverse kinematics is given by

$$\min_{\theta_i} |\{P_{effector}(\theta_i) + \delta(\theta_i)\} - P_{target}| \quad (7)$$

which minimum value corresponds to the Euclidean distance between the end-effector and the target. The disturbance due to multiple dynamic effects and coupling between satellite and robot is accounted for by the δ parameter (Fonseca et al. [2021]).

Table 2: Satellite and robot parameters for dynamics computation.

Device	Dimension _{<i>x,y,z</i>} (m)	Mass (kg)
Satellite	$1 \times 1 \times 1$	90
Link 1	$1 \times 0.05 \times 0.05$	0.3
Link 2	$1 \times 0.05 \times 0.05$	0.3
Link 3	$1 \times 0.05 \times 0.05$	0.3

4 Path Planning

An important aspect of the robotic device is the ability to compute a trajectory that describes the desired motion of a manipulator in multidimensional space. It consists in a time history of position, velocity, and acceleration for each degree of freedom of the manipulator.

Multiple strategies are available and in the simplest case the mission will specify the desired goal position and orientation of the end-effector and the exact shape of the path to get there, the duration, the velocity profile, and other details will be automatically established by an algorithm.

The desired outcome of a plan will be represented in terms of the state and actions that are executed. The main goals of a trajectory planning system are:

1. Feasibility: find a plan that causes arrival at a final state, regardless of its efficiency;
2. Optimality: find a feasible plan that optimizes performance following one or more performance indices, in addition to arriving in a goal state.

Furthermore, in order to guarantee smooth paths, some sort of constraints may be included on the spatial and temporal aspects of the path through the usage of via points, which consists in the establishment of intermediate points between the initial and the final configurations.

The initial position of the manipulator is known according to a Cartesian reference and also in the form of a set of joint angles. A planning strategy shall provide a set of joint angles whose value at t_0 is the initial position of the joint and the set of joint angles at t_f is the desired goal position of that joint. In making a smooth motion through a polynomial function, at least four constraints will arise. Two constraints are due to the selection of initial and final values:

$$\begin{aligned}\theta(0) &= \theta_0 \\ \theta(t_f) &= \theta_f\end{aligned}$$

Two additional constraints are due to the continuity in velocity, which in this case means that the initial and final velocities are zero:

$$\begin{aligned}\dot{\theta}(0) &= 0 \\ \dot{\theta}(t_f) &= 0\end{aligned}$$

These four constraints can be satisfied by a polynomial of third degree. Satisfying such constraints uniquely specify a particular cubic polynomial. A cubic has the form

$$\theta(t) = a_0 + a_1t + a_2t^2 + a_3t^3$$

and the joint velocity and acceleration are:

$$\begin{aligned}\dot{\theta}(t) &= a_1 + 2a_2t + 3a_3t^2 \\ \ddot{\theta}(t) &= 2a_2 + 6a_3t\end{aligned}$$

Combining the equations of velocity and acceleration with the four desired constraints yields four equations in four unknowns:

$$\begin{aligned}\theta_0 &= a_0 \\ \theta_f &= a_0 + a_1t_f + a_2t_f^2 + a_3t_f^3 \\ 0 &= a_1 \\ 0 &= a_1 + 2a_2t_f + 3a_3t_f^2\end{aligned}$$

Solving the equations for a_i , it follows that the explicit values of constants are

$$\begin{aligned}a_0 &= \theta_0 \\ a_1 &= 0 \\ a_2 &= \frac{3}{t_f^2}(\theta_f - \theta_0) \\ a_3 &= -\frac{2}{t_f^3}(\theta_f - \theta_0)\end{aligned}$$

and uniquely defines a cubic polynomial.

In the case $t_0 = 0$ s, $t_f = 30$ s, $\theta_0 = 0$ rad and $\theta_f = \frac{\pi}{2}$ rad, the a_i parameters are $a_0 = 0$, $a_1 = 0$, $a_2 = 0.0052$ and $a_3 = -0.0001$. The polynomial $\theta(t)$ is shown in Figure 2.

This concept can be generalized through an interpolation scheme. For this, the intermediate reference r_1 and the final reference point r_2 are added to the analysis. By moving the position of the reference points, a spline interpolation scheme is applied to the references r_0 , r_1 and r_2 . The concept is shown in Figure 3. Different interpolation data will lead to a different physical movement.

5 Torque and mechanical power

The Lagrangian $L = K - P$ is defined by the difference between the kinetic energy K and the potential energy P of the system. The dynamics of the system can be described by the Lagrange equation as given by:

$$u_i = \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) - \frac{\partial L}{\partial q_i} \quad (8)$$

where, q_i are the generalized coordinates (θ_i for rotational joints and d_i for prismatic joints); \dot{q}_i are the generalized velocities (angular velocity $\dot{\theta}_i$ for rotational joints and \dot{q}_i for prismatic joints); and u_i are the generalized forces.

By using Equation 8, the generalized forces u_i can be written as:

$$u_i = \sum_{j=1}^n D_{ij} \ddot{q}_j + I_{ai} \ddot{q}_i + \sum_{i=1}^n \sum_{k=1}^j C_{ijk} \dot{q}_j \dot{q}_k + G_i \quad (9)$$

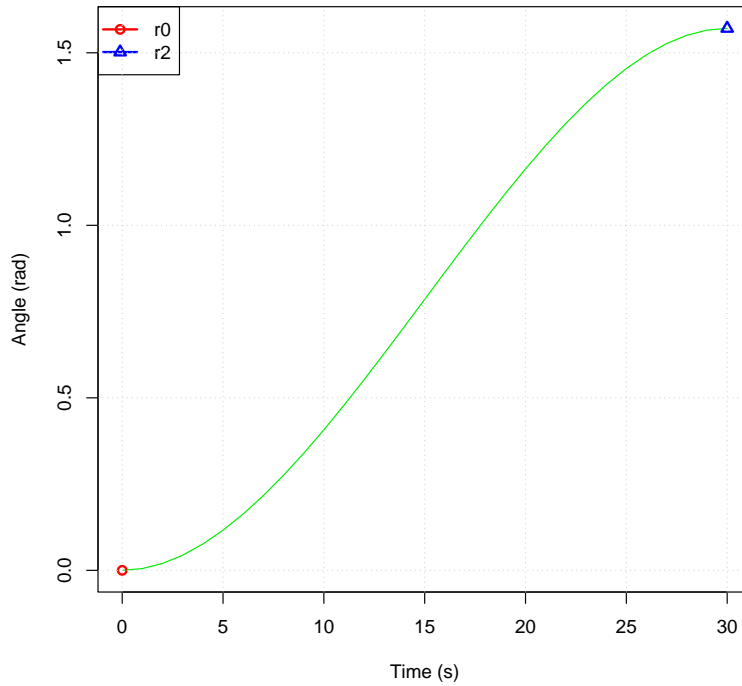


Figure 2: Cubic interpolation of joint coordinates.

where, D_{ij} is the inertia matrix of the system; C_{ijk} is the matrix that takes into account Coriolis and centrifugal effects; $I_{\alpha i}$ represent the inertia of the actuators, and G_i is the vector of gravitational forces.

Due to the relationship that exists between energy and force, the minimal energy can be estimated from the generalized force $u_i(t)$ that is associated to each joint i at the time instant t . The mechanical power can be used for design purposes, as defined by:

$$f = \int_{t_0}^{t_f} \sum_{i=1}^n (u_i^T(t) \dot{q}_i(t))^2 dt \quad (10)$$

This expression is representative of the phenomenon under study because it considers both the kinematic and dynamical aspects of the trajectory, simultaneously (Steffen Jr. and Saramago [1997], Saramago and Steffen Jr. [1998]).

Should be pointed out that by using a discrete set with limited number of intermediate reference points, the interpolation schema provides a trajectory with relevant information in terms of mechanical energy (dos Santos et al. [2008a]).

In the scenario of space missions, the joint forces and torques may incur in disturbance over the satellite and end-effector movements. The paper by da Fonseca et al. [2018] deals with the dynamic analysis of the close approach phase of a space robot to a target spacecraft. The computer simulations of the equations of the dynamics (relative orbital and attitude equations of motion) yields the grasping conditions and the results are compared with the inverse kinematics grasping operation implemented in a lab experiment.

An alternative to describe the spacecraft dynamics is to use software that requires only a geometric description of the system, as described in the next section.

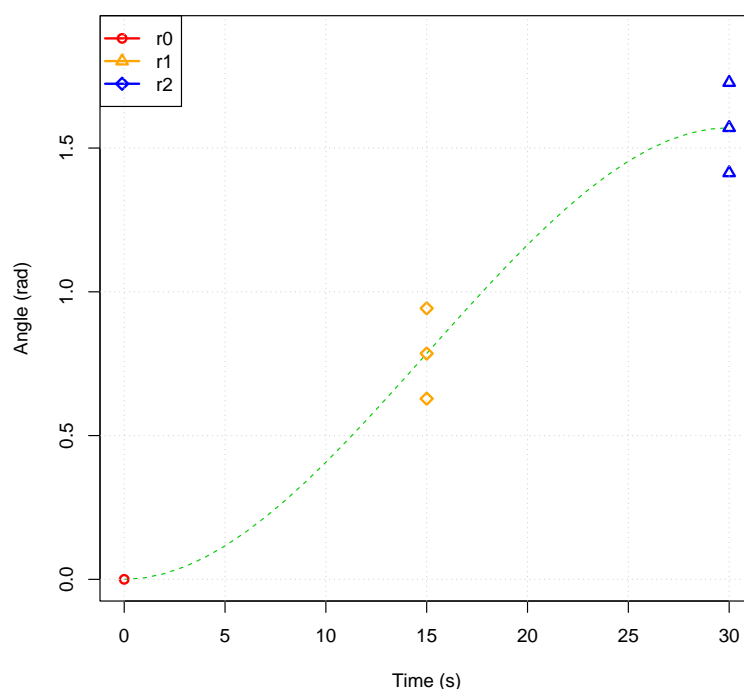


Figure 3: Spline interpolation of joint coordinates through reference points r_1 and r_2 .

6 Multibody modeling approach with OpenModelica software package

The paper Pulecchi and Lovera [2006] discusses the suitability of Modelica language for modeling Attitude and Orbit Control Systems (AOCS). The ability to use multiple coordinate frames through visual connectors, the existence of rich MultiBody Library to describe spacecraft dynamics, and the ability to describe multi-physics domains are some key aspect cited by the authors.

The Modelica language allows defining models in a declarative way, in hierarchical modules, combining several formalisms: ordinary differential equations (ODE), differential-algebraic equations (DAE), bond graphs, and finite state automata.

The multiple-domain capability of the language allows combining electrical, mechanical, hydraulic, and thermodynamic model components on the same application Fritzson et al. [2005]. By default, OpenModelica transforms a Modelica model into an ODE representation to perform a simulation by using numerical integration methods. DASSL is the default solver in OpenModelica. It is an implicit, high-order, multi-step solver with a step-size control (Petzold [1982]).

Elementary objects such as mass, spring, damper, and joints can be represented by visual elements, and do not require any explicit knowledge of the mathematical equations. An example of a 4-joint, 3-link robotic manipulator is shown in Figure 4.

As a result, the satellite body, joints, links and other mechanical and electrical elements are represented by building blocks. In this context, there is no need to write the dynamics equations explicitly. Code 1 shows a fragment of the automatically generated

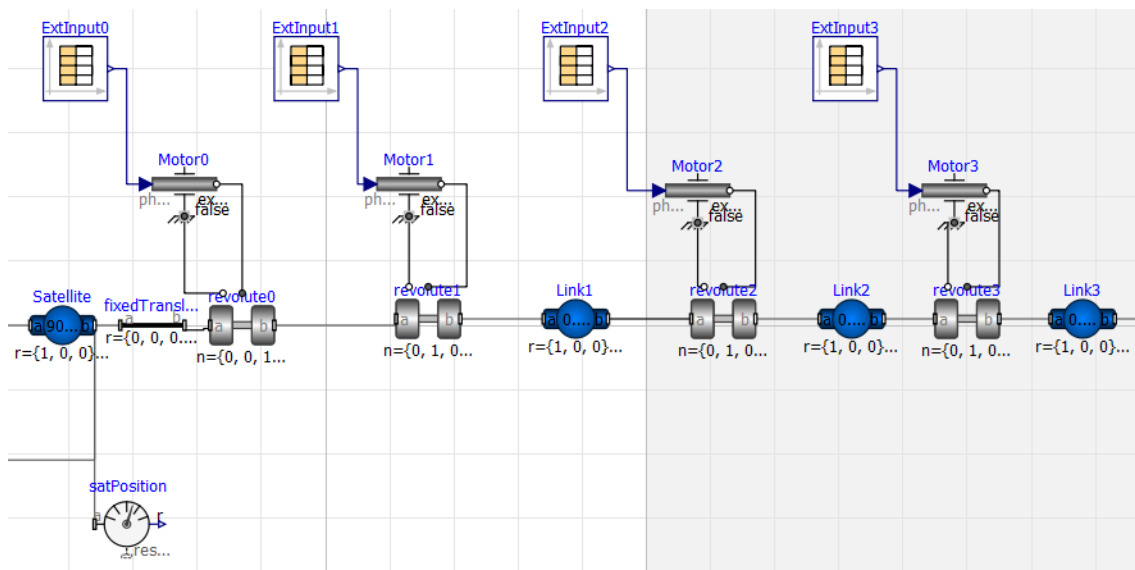


Figure 4: Fragment of a Space Robot model in OpenModelica.

equations. The corresponding Denavit-Hartenberg parameters are given in Table 1.

The satellite and robotic arm of the present work rely on a computational model built in OpenModelica software. A more in-depth discussion of software features and aerospace applications can be found in the works of Fritzson et al. [2006], Elmqvist et al. [1998], Lovera and Pulecchi [2006], Looye [2008], Briese et al. [2017], Moorman and Looye [2002], and Chen et al. [2017].

Some commercial front-ends for Modelica include AMESim (Siemens Software), Dymola (Dassault Systemes), Wolfram SystemModeler (Wolfram Research), SimulationX (ESI ITI GmbH), MapleSim (Maplesoft) and CATIA Systems Dynamic Behavior (Dassault Systemes).

Code 1: Code fragment of a Space Robot model in OpenModelica.

```

model SpaceRobot
inner Modelica.Mechanics.MultiBody.World world(axisLength =
  ↪ 1, gravityType = Modelica.Mechanics.MultiBody.Types.
  ↪ GravityTypes.UniformGravity, n = {0, 0, -1});
Modelica.Mechanics.MultiBody.Joints.Revolute revolute1(n =
  ↪ {0, 1, 0}, phi(displayUnit = "rad", fixed = false,
  ↪ start = 0), useAxisFlange = true);
Modelica.Mechanics.MultiBody.Joints.Revolute revolute2(n =
  ↪ {0, 1, 0}, phi(displayUnit = "rad"), useAxisFlange =
  ↪ true);
(...)
Modelica.Mechanics.MultiBody.Parts.BodyShape Link1(m = 0.3,
  ↪ r = {1, 0, 0}, r_CM = {0.5, 0, 0}, shapeType = "box
  ↪ ", sphereDiameter = 0.01);
Modelica.Mechanics.MultiBody.Parts.BodyShape Link2(color =
  ↪ {135, 206, 235}, m = 0.3, r = {1, 0, 0}, r_CM = {0.5,
  ↪ 0, 0}, shapeType = "box", sphereDiameter = 0.01);
Modelica.Mechanics.MultiBody.Sensors.AbsolutePosition
  ↪ effectorPosition(resolveInFrame = Modelica.Mechanics.
  ↪ MultiBody.Types.ResolveInFrameA.world);
Modelica.Mechanics.MultiBody.Parts.BodyShape Satellite(
  ↪ color = {200, 200, 200}, height = 0.99, length = 0.99,
  ↪ m = 90, r = {1, 0, 0}, r_CM = {0.5, 0, 0}, shapeType
  ↪ = "box", sphereDiameter = 0.1, width = 0.99,
  ↪ widthDirection = {0, 0, 1});
(...)
equation
connect(world.frame_b, traZ.frame_a);
connect(Satellite.frame_b, satAngles.frame_a);
connect(Satellite.frame_b, satPosition.frame_a);
connect(revolute3.frame_b, link3.frame_a)
(...)
connect(Link1.frame_b, revolute2.frame_a);
connect(Link2.frame_b, revolute3.frame_a);
connect(revolute0.frame_b, revolute1.frame_a);
connect(Satellite.frame_b, fixedTranslation.frame_a);
annotation(uses(Modelica(version = "3.2.2")), Diagram);
end SpaceRobot;

```

7 Uncertainty Quantification

The satellite and the robotic serial manipulator are coupled as shown in Figure 5. Four joint coordinates and tree links are considered. The Denavit-Hartenberg parameters (Table 1) that describe this robot are associated with eight reference points r_1, \dots, r_8 , as shown in Table 3.

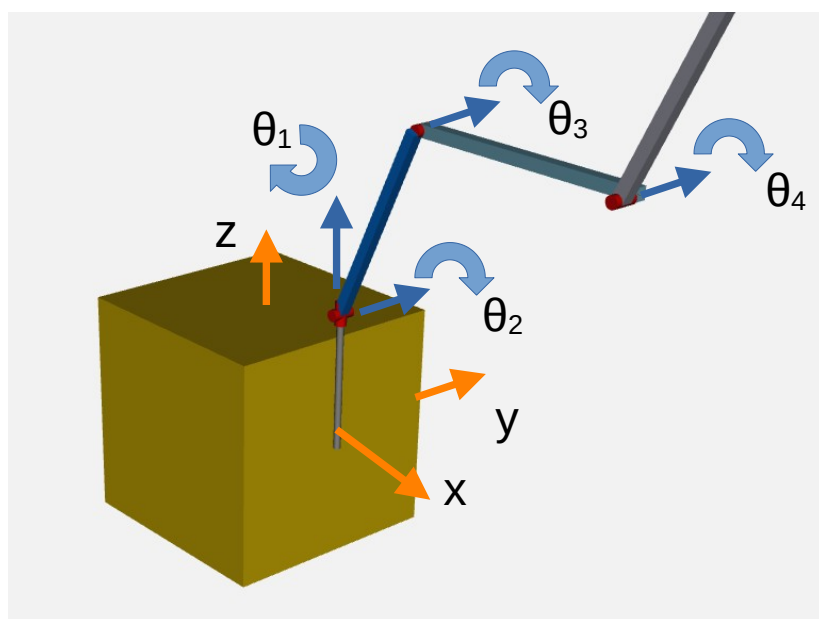


Figure 5: Space robot manipulator.

Table 3: Joint coordinate nomenclature.

Joint	Design variables	Joint coordinate
1	r_1, r_2	$\theta_1(r_1, r_2, t)$
2	r_3, r_4	$\theta_2(r_3, r_4, t)$
3	r_5, r_6	$\theta_3(r_5, r_6, t)$
4	r_7, r_8	$\theta_4(r_7, r_8, t)$

The berthing operation involves rendezvous between two orbiting space vehicles followed by berthing (grasping or capture operation). The relative attitude implies synchronizing the attitude motion of both spacecraft so that the relative attitude is zero. The same is valid for the relative position. When the relative velocity is zero, the position of the center of mass of both spacecraft stays at the same distance from each other.

In the present work, the berthing box (Fehse [2003]) is centered at the position $(x, y, z) = (2.50, 0.00, 0.70)$ m in relation to the satellite origin and encompasses the uncertainty of $(\Delta x, \Delta y, \Delta z) = (\pm 0.21, \pm 0.21, \pm 0.21)$ m.

The numerical computations were performed using the R software (R Core Team [2020]) and considered the following step:

- Nonlinear programming algorithm (Johnson [2020]) solving the inverse problem (Equation 6);
- Nonlinear programming algorithm solving the inverse problem (Equation 7);
- Neural Network algorithm (Paszke et al. [2019]) learning the inverse kinematics solution from historical data.

These steps are evaluated as detailed in Algorithm 1.

Algorithm 1: Inverse kinematics computation.

```

 $t_0 \leftarrow 0$ 
 $t_f \leftarrow$  final time
 $P_{end} \leftarrow$  Cartesian coordinates of target
 $(\theta_1, \theta_2, \theta_3, \theta_4)_0 = (0, 0, 0, 0)$ 
/* Generate estimate from low fidelity model */
 $\tilde{\theta}_i \leftarrow \min_{\theta_i} |T_0^n(q(\theta_i))P_{base} - P_{end}|, i = 1, \dots, 4$ 
/* Optimize high fidelity model */
Initial guess  $\leftarrow \tilde{\theta}_i$ 
 $\vec{\theta}_f \leftarrow \min_{\theta_i} |T_0^n(q(\theta_i)) \{P_{base}(\theta_i) + \Delta(\theta_i)\} - P_{end}|, i = 1, \dots, 4$ 
/* Save historical designs */
 $n \leftarrow$  number of optimization iterations
 $\vec{\theta}_{hist} = (\vec{\theta}_f)_k, k = 1, \dots, n$ 
 $\vec{P}_{hist} = (\vec{P}_{end})_k, k = 1, \dots, n$ 
 $\vec{\Delta}_{hist} = (\vec{\Delta}(\theta_f))_k, k = 1, \dots, n$ 
/* Neural network training */
 $RNN \leftarrow$  training from historical data  $\{\vec{\theta}_{hist}, \vec{P}_{hist}, \vec{\Delta}_{hist}\}$ .
```

In the following, 36 cases with distinct target points P_{end} were analyzed. The results provided by the DH algorithm are summarized in Table 4. The error corresponds to the distance between the end-effector and the target, computed through the high fidelity model (Equation 7). The measurement of the distance between the end-effector and the target confirmed that the stationary point was an optimal design.

The error dispersion, summarizing the distance between the end-effector and the target for all cases, is shown in Figure 6. After the optimization process, the final result (Figure 6, on the right side) was below the threshold $\varepsilon = 0.004$ m. The choice of a Nonlinear Programming strategy has a direct influence on the convergence rate and on the accuracy of the result. A comprehensive exposition of techniques and features can be found in the chapter “An overview of Linear and Non-linear Programming methods for Structural Optimization” (Choze et al. [2022b]) of this book series.

Table 4: Number of iterations and final error obtained from DH strategy.

Case	Target (m)	DH Iter.	DH Error (m)
1	(2.500, 0.000, 0.700)	141	0.004
2	(2.710, 0.210, 0.910)	121	0.003
3	(2.710, 0.210, 0.490)	174	0.003
4	(2.710, -0.210, 0.910)	111	0.003
5	(2.710, -0.210, 0.490)	153	0.003
6	(2.290, 0.210, 0.910)	117	0.003
7	(2.290, 0.210, 0.490)	151	0.002
8	(2.290, -0.210, 0.910)	127	0.003
9	(2.290, -0.210, 0.490)	139	0.003
10	(2.471, 0.050, 0.600)	145	0.003
11	(2.600, -0.023, 0.708)	136	0.003
12	(2.576, -0.051, 0.607)	150	0.003
13	(2.486, 0.103, 0.688)	144	0.003
14	(2.428, 0.007, 0.653)	137	0.003
15	(2.478, -0.102, 0.616)	139	0.003
16	(2.559, -0.011, 0.670)	145	0.003
17	(2.568, 0.090, 0.675)	136	0.004
18	(2.583, 0.056, 0.721)	137	0.003
19	(2.405, 0.075, 0.741)	166	0.003
20	(2.535, -0.078, 0.777)	134	0.004
21	(2.501, -0.061, 0.635)	148	0.002
22	(2.596, 0.035, 0.803)	117	0.002
23	(2.518, 0.023, 0.658)	141	0.003
24	(2.417, -0.070, 0.757)	141	0.003
25	(2.461, -0.088, 0.693)	139	0.002
26	(2.398, -0.037, 0.623)	148	0.003
27	(2.456, 0.070, 0.745)	132	0.003
28	(2.493, -0.015, 0.731)	130	0.002
29	(2.525, -0.096, 0.700)	150	0.003
30	(2.507, 0.013, 0.783)	136	0.003
31	(2.545, 0.040, 0.630)	137	0.002
32	(2.442, -0.031, 0.767)	137	0.003
33	(2.423, -0.003, 0.714)	141	0.003
34	(2.552, -0.050, 0.760)	141	0.002
35	(2.443, 0.087, 0.644)	126	0.002
36	(2.531, 0.066, 0.796)	132	0.003

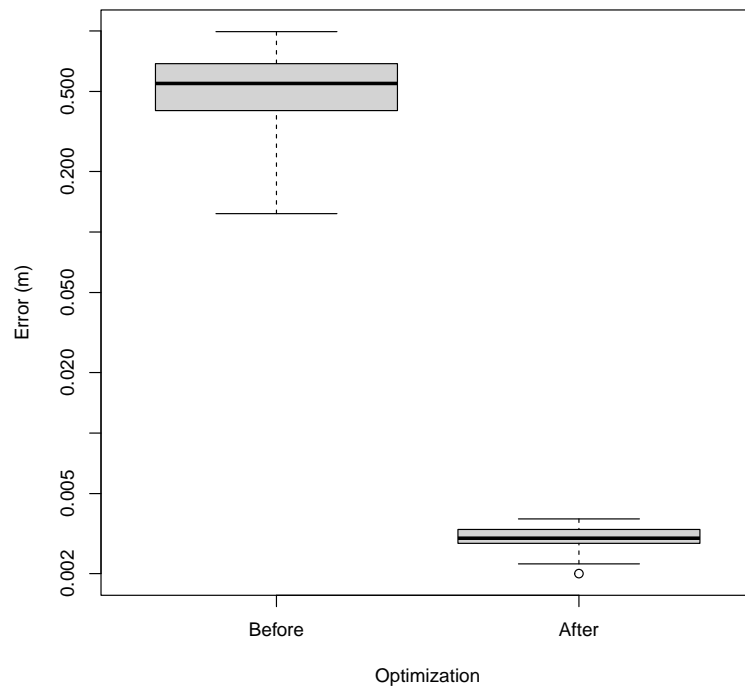


Figure 6: Error dispersion before and after the optimization process.

Code 2: Code of a Neural Network architecture to estimate the Cartesian position error.

```
class Net(T.nn.Module):
    def __init__(self):
        super(Net, self).__init__()
        self.hid1 = T.nn.Linear(8, 2400)
        self.outp = T.nn.Linear(2400, 1)

    def forward(self, x):
        x = T.relu(self.hid1(x))
        x = self.outp(x)
        return x
```

7.1 Error due to model simplification

Once the satellite is traveling in space, it may present oscillations influenced by the robotic movement, and is subject to the influence of external forces from the environment. As a result, the hypothesis of fixed robotic base (Equation 1) is not feasible in some situations. The Neural Network model, as shown in Code 2, is considered to estimate the discrepancy between the model with a fixed base and the model with a floating base.

The uncertainty that arises from this approach is evaluated by the computation of the same joint design $\vec{\theta}(r, t)$ by using the high fidelity model (Figure 4), and the difference between results of Equations 6 and 7 is shown in Figure 7. There is a similar pattern in terms of error dispersion along x , y and z dimensions, as shown in Figure 8.

The error is negligible in some situations but relevant in another. As a result, the optimization problem 6 given an useful estimate of joint angles at the expense of low computational power requirements. Notice that the median of the error (horizontal heavy line on the plot) is similar in all three dimensions.

The most important reference points are those representing the final joint angles. Joint angles at the final time $\theta_i(t_f)$ corresponds to a Cartesian position of the end-effector. The cumulative error between the target and the end-effector is shown in Figure 9.

The z -axis has large deviation than other dimensions. According to the present convention (Figure 5) there is a higher displacement in the vertical axis while evaluating the target set (Table 4). The procedure is further detailed in Algorithm 2.

Algorithm 2: Inverse kinematics computation through the mathematical model.

```
 $t_0 \leftarrow 0$ 
 $t_f \leftarrow$  final time
 $P_{end} \leftarrow$  Cartesian coordinates of target
 $(\theta_1, \theta_2, \theta_3, \theta_4)_0 = (0, 0, 0, 0)$ 
/* Generate estimate from low fidelity model */
 $\vec{\theta}_f \leftarrow \min_{\theta_i} |T_0^n(q(\theta_i))P_{base} - P_{end}|, i = 1, \dots, 4$ 
```

The main advantage of this procedure is the low computational cost of the CPU, which allows the evaluation of a multitude of scenarios under a low computational budget.

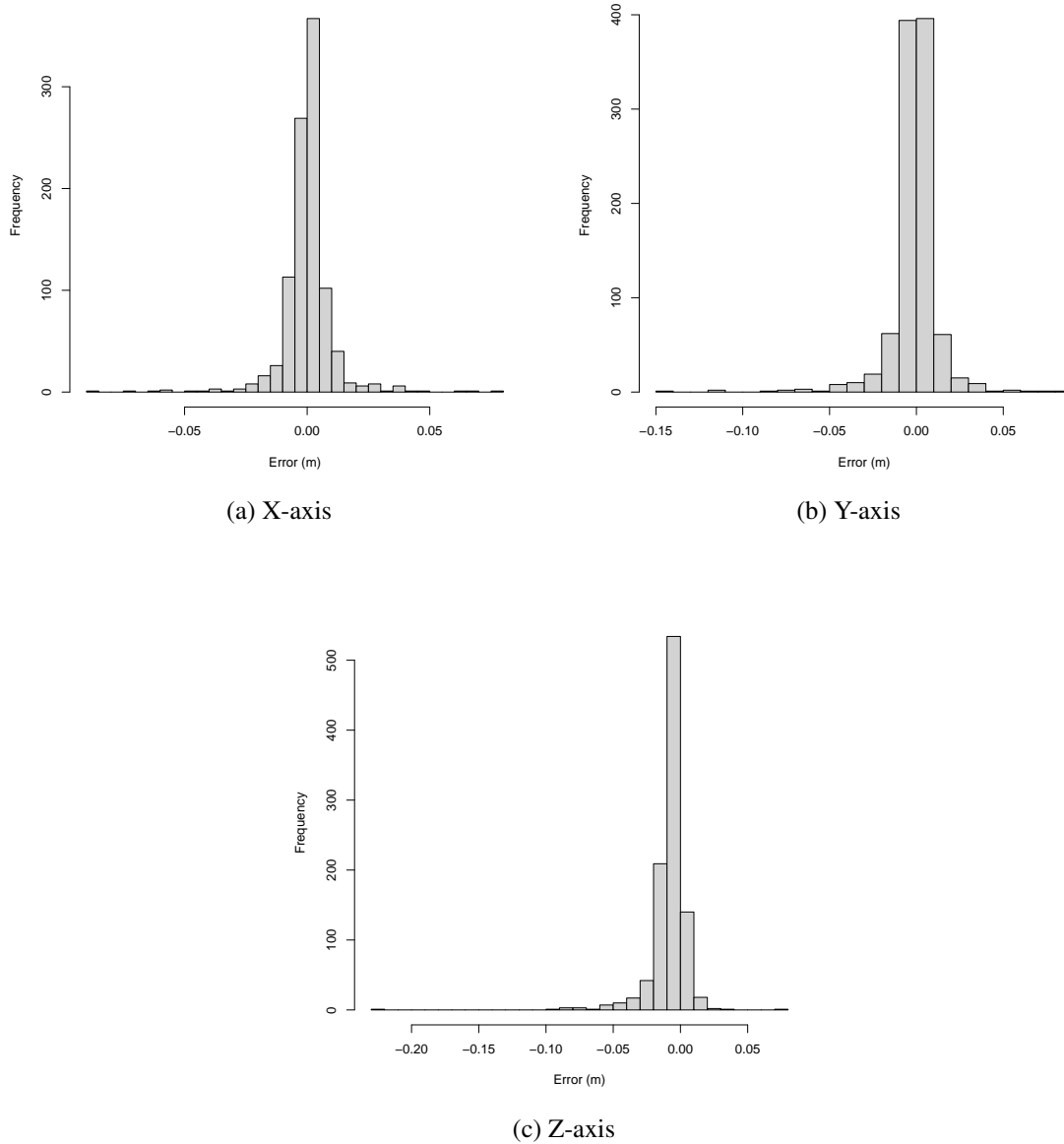


Figure 7: Historical error histogram along the x, y and z axes.

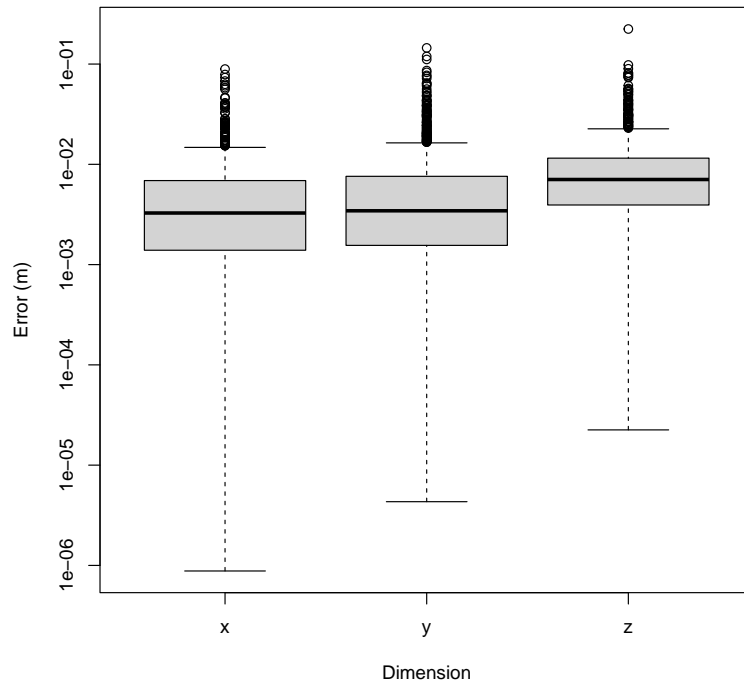


Figure 8: Comparison of error dispersion along the x, y and z axes.

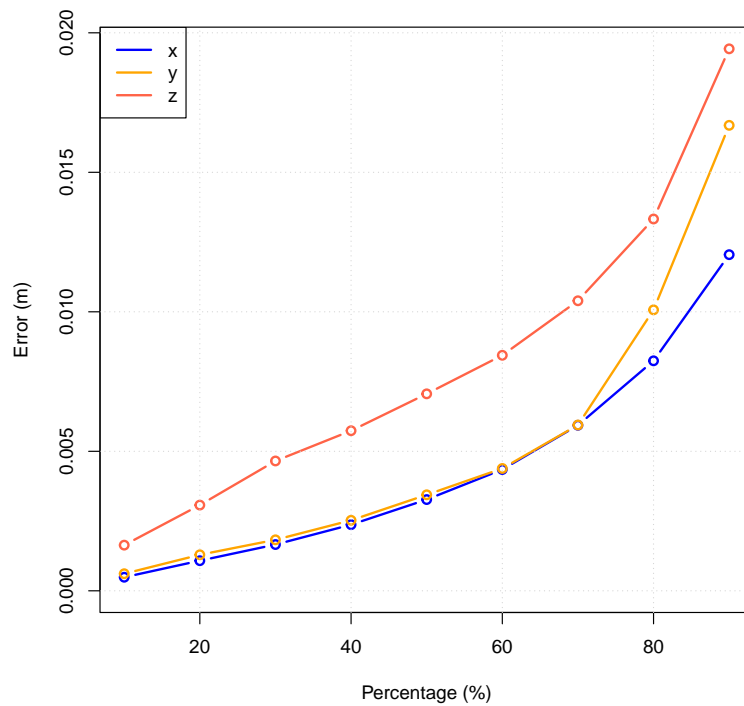


Figure 9: Cumulative error along the x, y and z axes.

Code 3: Code of a Neural Network architecture to estimate the joint angle value.

```

class Net(T.nn.Module):
    def __init__(self):
        super(Net, self).__init__()
        self.hid1 = T.nn.Linear(10, 2700)
        self.hid2 = T.nn.Linear(2700, 1031)
        self.hid3 = T.nn.Linear(1031, 394)
        self.hid4 = T.nn.Linear(394, 150)
        self.hid5 = T.nn.Linear(150, 57)
        self.hid6 = T.nn.Linear(57, 22)
        self.outp = T.nn.Linear(22, 1)

    def forward(self, x):
        x = T.relu(self.hid1(x))
        x = T.relu(self.hid2(x))
        x = T.relu(self.hid3(x))
        x = T.relu(self.hid4(x))
        x = T.relu(self.hid5(x))
        x = T.relu(self.hid6(x))
        x = self.outp(x)
        return x

```

7.2 Error on meta-modeling estimate

The calculation of the inverse kinematics through Equation 7 generated a history of 4949 joint coordinates θ and positions of end-effector P_{end} .

This section will discuss the existence of a meta-model that can learn from this data and predict the inverse kinematics configuration of other target points. The training data will be based exclusively on the historical data set acquired from the optimization procedure done in the previous section.

The values used to train a Neural Network consist of a subset of 3959 sample points, which corresponds to 80% of the data available. The meta-model will predict the optimal value of each reference point $\vec{r} = (r_1, \dots, r_8)$. These points are used to calculate the value of the joint coordinates $\theta(\vec{r}, t)$ in the time interval $t \in [t_0, t_f]$. The model is tested against a subset of 990 points not used in the training phase, which corresponds to 20% of the data. The general architecture of the Neural Network is presented in the Code 3, following the syntax of PyTorch.

The difference between reference points $\vec{e} = \vec{r}_{opt} - \vec{r}_{rnn}$ computed through optimization and the Neural Network is summarized in Figure 10. The majority of the results are near zero. The mean value and standard deviation are $(\bar{e}_1, \sigma_1) = (0.0000, 0.0421)$ and $(\bar{e}_2, \sigma_2) = (-0.0007, 0.0050)$ for r_1 and r_2 , respectively.

Continuing with the analysis of the second joint, the mean value and the standard deviation are $(\bar{e}_3, \sigma_3) = (-0.0152, 0.1937)$ and $(\bar{e}_4, \sigma_4) = (-0.0039, 0.0085)$ for r_3 and r_4 , respectively. The histogram is shown in Figure 11.

The analysis of the third joint led to the mean value and standard deviation of $(\bar{e}_5, \sigma_5) =$

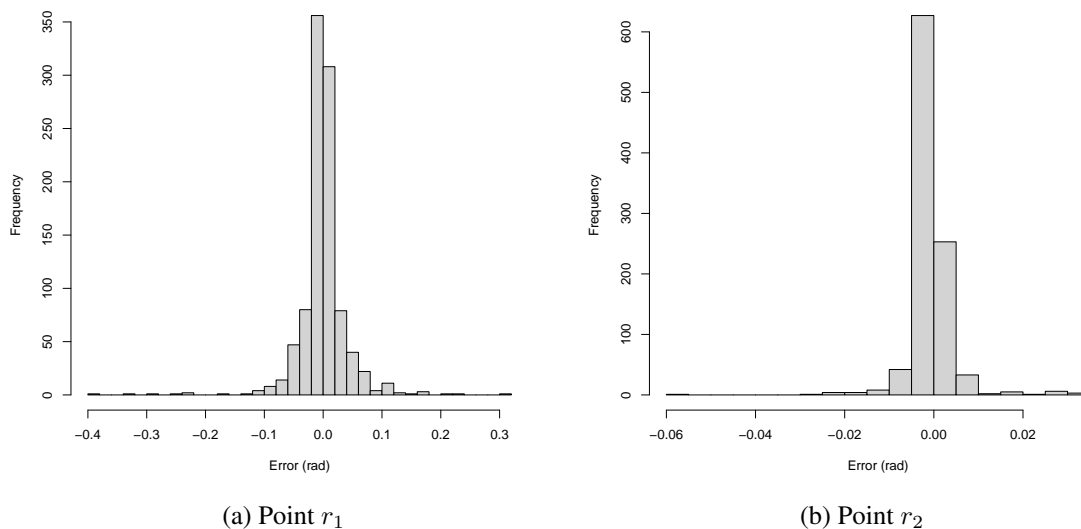


Figure 10: Historical error histogram of joint 1.

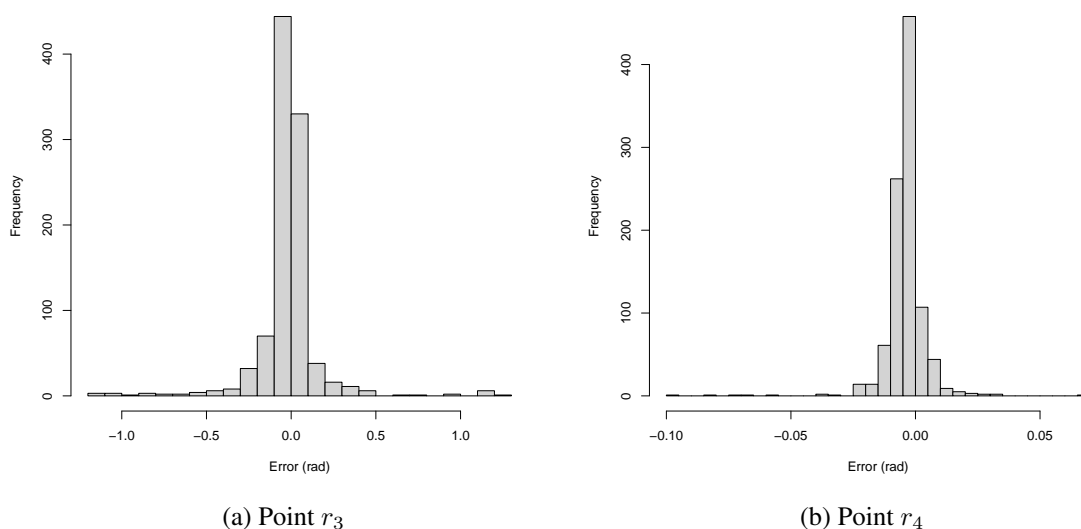


Figure 11: Historical error histogram of joint 2.

$(0.0126, 0.2007)$ and $(\bar{e}_6, \sigma_6) = (-0.0011, 0.0102)$ for the r_5 and r_6 reference points, respectively. The histogram is shown in Figure 12.

Finally, the analysis of the fourth joint resulted in a mean value and standard deviation of $(\bar{e}_7, \sigma_7) = (-0.0106, 0.1944)$ and $(\bar{e}_8, \sigma_8) = (0.0046, 0.0132)$ for the reference points r_7 and r_8 , respectively. The histogram is shown in Figure 13.

The standard deviations at the end times $\sigma_2, \sigma_4, \sigma_6$ and σ_8 are smaller than their intermediate time counterparts ($\sigma_1, \sigma_3, \sigma_5$ and σ_7). It means that the calculation of the reference point at the final time (r_2, r_4, r_6 and r_8) is more accurate than the estimated value at the

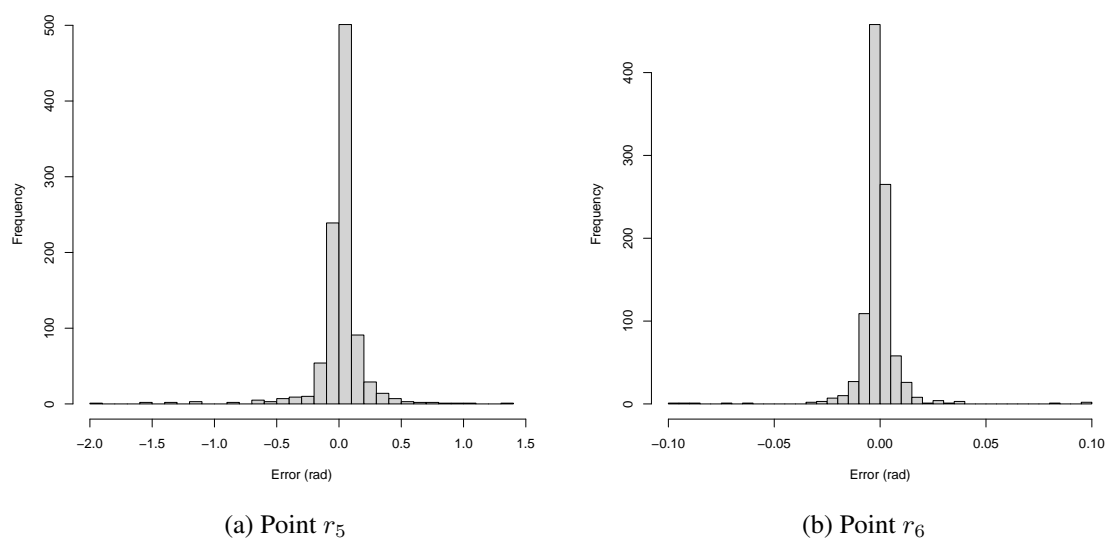


Figure 12: Historical error histogram of joint 3.

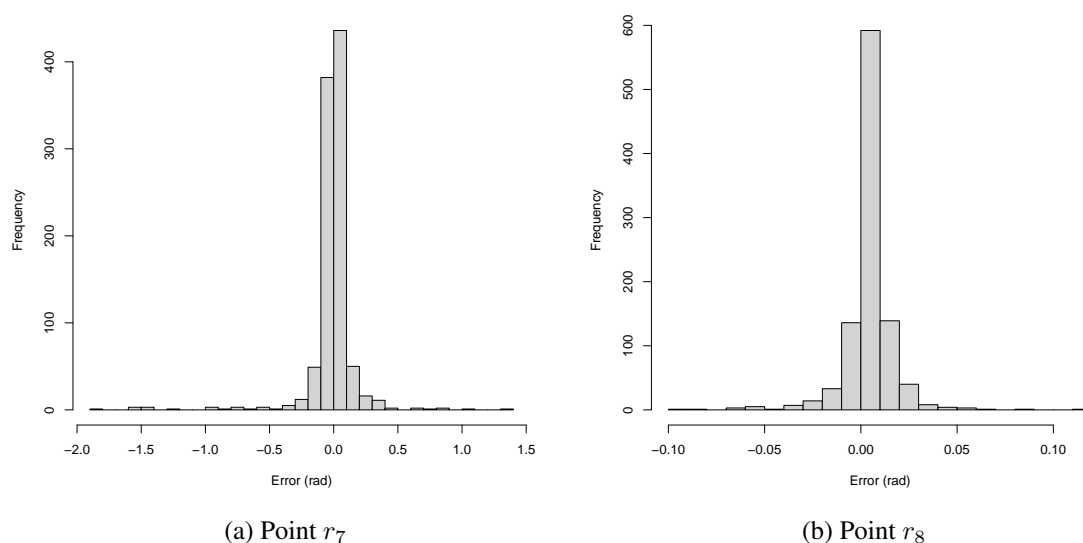


Figure 13: Historical error histogram of joint 4.

intermediate time. It is important from a practical point of view because the joint angle at the end time is critical to the accuracy of the end-effector towards the target.

A comparison of the error dispersion between all reference points is presented in Figure 14. Note that the median and quartiles follow a similar trend as discussed in the previous paragraph.

The accumulated error of the final time reference points (r_2, r_4, r_6 and r_8) are shown in Figure 15. The biggest error was found in the last reference point r_8 . A valid approach to mitigating this problem is to investigate the existence of a custom neural network archi-

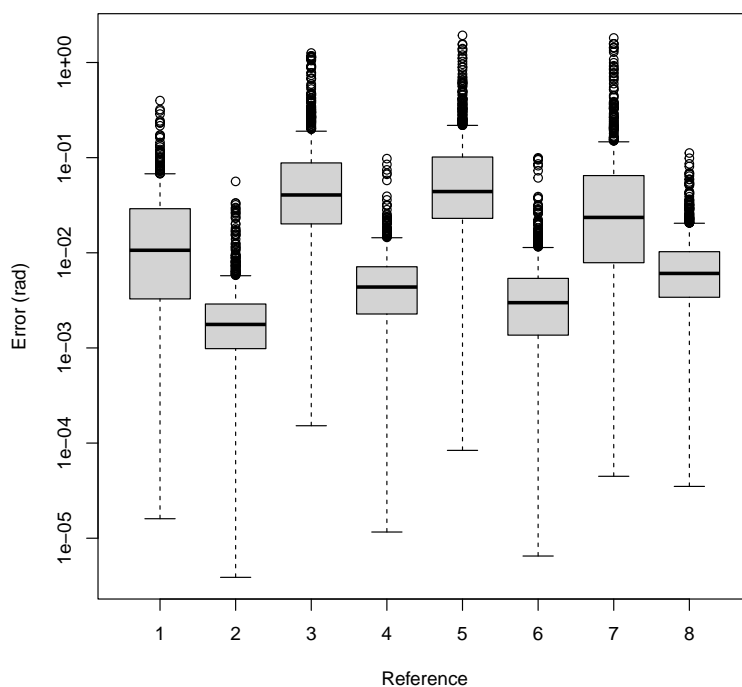


Figure 14: Comparison of error dispersion along the reference points.

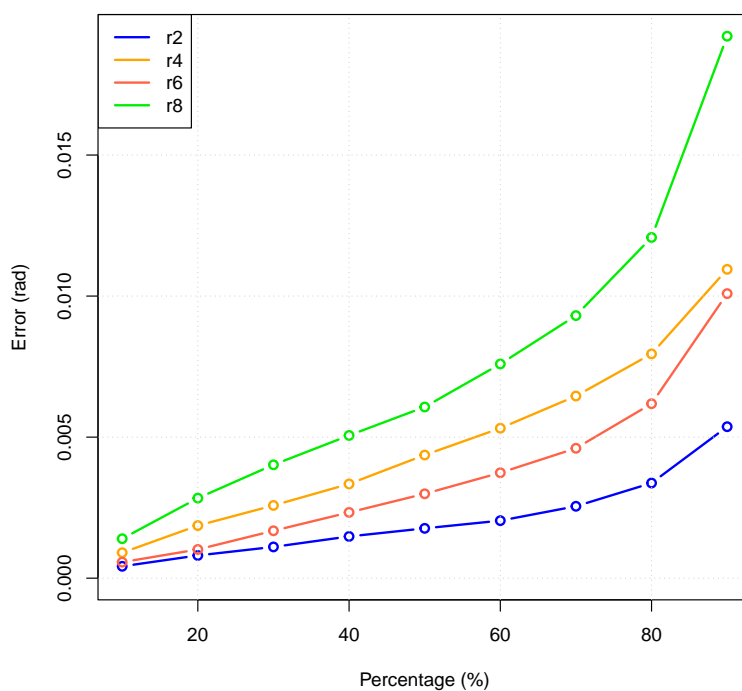


Figure 15: Cumulative error in multiple reference points.

ture that improves the quality of results. In the present study, the same neural network architecture was used as a meta-model for all reference points. Although an independent training procedure was performed to estimate the weights of the network used in the calculations of each reference r_i , it was found that the quality of the model is different in each case.

As a final observation, the results of a neural network proved to be useful to estimate the value of reference points that will meet the inverse kinematics criteria. Despite the fact that in specific circumstances a large error can occur, this methodology is really efficient in providing a good estimate for the optimization procedure. As a result, the combined use of a Neural Network (for a robust estimation) and the technique of Nonlinear Programming (for an accurate solution) is recommended as a powerful tool in the path planning of space robots.

8 Conclusion

The present study analyzed the uncertainty that arises in the movement of a space robot composed of a satellite and a serial manipulator. The first model described the robot kinematics using rotation and translation matrices. No force effects are considered.

The second model used a multibody formulation written in OpenModelica software. In this case, forces and torques are transmitted between the robot and the satellite. The end-effector positioning error in both cases was compared.

A Neural Network was trained on the PyTorch platform to identify disparities between models and learn from historical data. The trained model was found to be a robust predictor of joint angle values. In most cases, it gives a result with sufficient accuracy.

The Nonlinear Programming methodology benefited from this estimate as initial guess for the optimization process. The procedure is efficient in providing accurate calculations for inverse kinematics computation.

It was confirmed that the combined use of Neural Network and Nonlinear Programming strategies is an efficient approach for the calculations of inverse kinematics, which increases robustness and speeds up the process.

Acknowledgments

D.A. Rade acknowledges the Brazilian research agencies Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) (grant #2018 / 15894-0) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (grant #312068 / 2020-4) for the financial support to his research work. Ijar M. da Fonseca acknowledges the Brazilian Space Agency (AEB-UNIESPACO, grant #104006129).

References

- G. S. Aglietti, B. Taylor, S. Fellowes, T. Salmon, I. Retat, A. Hall, T. Chabot, A. Pisseloup, C. Cox, A. Mafficini, et al. The active space debris removal mission removedebris. part 2: in orbit operations. *Acta Astronautica*, 168:310–322, 2020.
- L. E. Briese, A. Klöckner, and M. Reiner. The dlr environment library for multi-disciplinary aerospace applications. In *Proceedings of the 12th International Modelica Conference, Prague, Czech Republic, May 15-17, 2017*, number 132, pages 929–938. Linköping University Electronic Press, 2017.
- F. Chen, G. Zhao, Y. Yu, J. Wang, and P. Cao. Modelica-based modeling and simulation of satellite on-orbit deployment and attitude control. In *4th International Conference on Machinery, Materials and Information Technology Applications*, pages 196–204. Atlantis Press, 2017.
- S. B. Choze, R. R. Santos, and G. F. Gomes. Overview of traditional and recent heuristic optimization methods. In *Model-Based and Signal-Based Inverse Methods*, volume 1 of *Discrete Models, Inverse Methods & Uncertainty Modeling in Structural Integrity*, chapter 4, pages 107–142. University of Brasilia, 2022a. URL <https://doi.org/10.4322/978-65-86503-71-5.c04>.
- S. B. Choze, R. R. Santos, A. B. Jorge, and G. F. Gomes. An overview of linear and non-linear programming methods for structural optimization. In *Model-Based and Signal-Based Inverse Methods*, volume 1 of *Discrete Models, Inverse Methods & Uncertainty Modeling in Structural Integrity*, chapter 3, pages 65–106. University of Brasilia, 2022b. URL <https://doi.org/10.4322/978-65-86503-71-5.c03>.
- J. J. Craig. *Introduction to robotics: Mechanics and control, 3rd Edition*. Pearson Education, 2005.
- I. M. da Fonseca. Space robotics and associated space applications. In *Vibration Engineering and Technology of Machinery*, pages 151–170. Springer, 2021.
- I. M. da Fonseca, L. M. Unfried, G. L. B. Lima, M. R. da Silva, O. Satome, and E. J. de Oliveira. Space robot dynamic analysis of the relative orbital and attitude motion in the close range rendezvous phase and grasping of a target space vehicle. In *69th International Astronautical Congress (IAC)*. 2018.
- I. M. da Fonseca, M. N. Pontuschka, and G. L. B. Lima. Kinematics for spacecraft-type robotic manipulators. In *Kinematics-Analysis and Applications*. IntechOpen, 2019.
- R. R. dos Santos, V. Steffen, Jr., and S. de Fatima Pereira Saramago. Solving the inverse kinematics problem through performance index optimization. In *Proceedings of the XXVI Iberian Latin-American Congress on Computational Methods in Engineering CILAMCE 2005*. Brazilian Assoc. For Comp. Mechanics & Latin American Assoc. of Comp. Methods in Engineering, 2005.
- R. R. dos Santos, V. Steffen Jr., and S. d. F. Saramago. Robot path planning in a constrained workspace by using optimal control techniques. *Multibody System Dynamics*, 19(1):159–177, 2008a.

- R. R. dos Santos, V. Steffen Jr., and S. d. F. P. Saramago. Optimal path planning and task adjustment for cooperative flexible manipulators. In *ABCMS Symposium Series in Mechatronics*, volume 3, pages 236–245. 2008b.
- R. R. dos Santos, V. Steffen, and S. d. F. P. Saramago. Optimal task placement of a serial robot manipulator for manipulability and mechanical power optimization. *Intelligent Information Management*, 2:512–525, 2010. doi: 10.4236/iim.2010.29061.
- H. Elmqvist, S. E. Mattsson, and M. Otter. Modelica: The new object-oriented modeling language. In *12th European Simulation Multiconference, Manchester, UK*, volume 5, 1998.
- ESPI. Espi report 71 - towards a european approach to space traffic management - full report, 2020a.
- ESPI. Espi report 72 - europe, space and defence - full report, 2020b.
- W. Fehse. *Automated rendezvous and docking of spacecraft*, volume 16. Cambridge university press, 2003.
- I. M. D. Fonseca, P. Gasbarri, E. F. Moura, and R. R. Santos. Trajectory generation for an on-orbit robot manipulator. In *Proceedings of the 72nd International Astronautical Congress (IAC)*. IAC, 2021.
- P. Fritzson, P. Aronsson, H. Lundvall, K. Nyström, A. Pop, L. Saldamli, and D. Broman. The OpenModelica Modeling, Simulation, and Software Development Environment. *Simulation News Europe*, 15(44/45):8–16, Dec. 2005.
- P. Fritzson, P. Aronsson, A. Pop, H. Lundvall, K. Nyström, L. Saldamli, D. Broman, and A. Sandholm. Openmodelica - a free open-source environment for system modeling, simulation, and teaching. In *2006 IEEE Conference on Computer Aided Control System Design, 2006 IEEE International Conference on Control Applications, 2006 IEEE International Symposium on Intelligent Control*, pages 1588–1595, 2006. doi: 10.1109/CACSD-CCA-ISIC.2006.4776878.
- S. G. Johnson. The NLOpt nonlinear-optimization package. 2020. URL <https://cran.r-project.org/package=nloptr>.
- D. Kuang, S. Desai, and A. Sibois. Observed features of gps block iiif satellite yaw maneuvers and corresponding modeling. *GPS solutions*, 21(2):739–745, 2017.
- S. M. Le Grand and K. M. Merz Jr. The application of the genetic algorithm to the minimization of potential energy functions. *Journal of Global Optimization*, 3(1):49 – 66, 1993.
- G. Looye. The new dlr flight dynamics library. In *Proceedings of the 6th International Modelica Conference*, volume 1, pages 193–202, 2008.
- M. Lovera and T. Pulecchi. Object-oriented modelling for spacecraft dynamics: A case study. In *2006 IEEE Conference on Computer Aided Control System Design*,

- 2006 *IEEE International Conference on Control Applications*, 2006 *IEEE International Symposium on Intelligent Control*, pages 1898–1903, 2006. doi: 10.1109/CACSD-CCA-ISIC.2006.4776930.
- D. G. Luenberger and Y. Ye. *Linear and Nonlinear Programming*. Springer, 3rd edition, 2008.
- D. Moorman and G. Looye. The modelica flight dynamics library. In *Proceedings of the 2nd International Modelica Conference, Oberpfaffenhofen, Germany, 2002*.
- P. Mróz, A. Otarola, T. A. Prince, R. Dekany, D. A. Duev, M. J. Graham, S. L. Groom, F. J. Masci, and M. S. Medford. Impact of the spacex starlink satellites on the zwicky transient facility survey observations. *The Astrophysical Journal Letters*, 924(2):L30, 2022.
- A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, et al. PyTorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems*, 32, 2019.
- L. Petzold. Dassel. a differential/algebraic system solver. Technical report, Lawrence Livermore National Lab., CA (United States), 1982.
- T. Pulecchi and M. Lovera. A modelica library for space flight dynamics. In *In Proceedings of the 5th International Modelica Conference*. Citeseer, 2006.
- R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2020. URL <https://www.R-project.org/>.
- H. E. Romeijn and R. L. Smith. Simulated annealing for constrained global optimization. *Journal of Global Optimization*, 5(2):101 – 126, 1994.
- R. R. Santos, D. A. Rade, and I. M. da Fonseca. A machine learning strategy for optimal path planning of space robotic manipulator in on-orbit servicing. *Acta Astronautica*, 191:41–54, 2022.
- S. F. P. Saramago and V. Steffen Jr. Optimization of the trajectory planning of robot manipulators taking into-account the dynamics of the system. *Mechanism and Machine Theory*, 33(7):883–894, 1998.
- M. A. Skinner, T. M. Kelecyc, S. A. Gregory, J. P. Toth, D. Liang, D. Yamanaka, S. Kent, R. Tjoelker, D. Margineantu, A. L. Allison, et al. Commercial space situational awareness: an investigation of ground-based ssa concepts to support commercial geo satellite operators. *Proc. AMOS*, 2013.
- M. Smith, D. Craig, N. Herrmann, E. Mahoney, J. Krezel, N. McIntyre, and K. Goodliff. The artemis program: an overview of nasa’s activities to return humans to the moon. In *2020 IEEE Aerospace Conference*, pages 1–10. IEEE, 2020.
- M. W. Spong and M. Vidyasagar. *Robot Dynamics and Control*. John Wiley and Sons, 1989.

- V. Steffen Jr. and S. F. P. Saramago. Optimization techniques for off-line trajectories planning of robot manipulators. *Non-linear Dynamics, Chaos, Control and their Applications in Engineering Sciences*, 1:363–368, 1997.
- R. Storn and K. Price. Differential evolution - a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization*, 11:341 – 359, 1997.
- M. Uchiyama. A study of computer control of a mechanical arm. *Bulletin of the Japan Society of Mechanical Engineers*, 22:1640–1647, 1979.

Chapter 5

Metamodel-Based Simulation Optimization: A Systematic Literature Review

Chapter details

Chapter DOI:

<https://doi.org/10.4322/978-65-86503-88-3.c05>

Chapter suggested citation / reference style:

Amaral, João V. S., et al. (2022). “Metamodel-Based Simulation Optimization: A Systematic Literature Review”. In Jorge, Ariosto B., et al. (Eds.) *Uncertainty Modeling: Fundamental Concepts and Models*, Vol. III, UnB, Brasilia, DF, Brazil, pp. 119–171. Book series in Discrete Models, Inverse Methods, & Uncertainty Modeling in Structural Integrity.

P.S.: DOI may be included at the end of citation, for completeness.

Book details

Book: Uncertainty Modeling: Fundamental Concepts and Models

Edited by: Jorge, Ariosto B., Anflor, Carla T. M., Gomes, Guilherme F., & Carneiro, Sergio H. S.

Volume III of Book Series in:

Discrete Models, Inverse Methods, & Uncertainty Modeling in Structural Integrity

Published by: UnB City: Brasilia, DF, Brazil Year: 2022

DOI: <https://doi.org/10.4322/978-65-86503-88-3>

Metamodel-Based Simulation Optimization: A Systematic Literature Review

João Victor Soares do Amaral^{1*}, José Arnaldo Barra Montevechi¹, and Rafael de Carvalho Miranda¹

¹Industrial Engineering and Management Institute, Federal University of Itajubá, Brazil.

E-mail addresses: joao.amaral@unifei.edu.br (J.V.S. Amaral), montevechi@unifei.com.br (J.A.B. Montevechi), rafael.miranda@unifei.edu.br (R.C. Miranda),

*Corresponding author

Abstract

Over the past few decades, metamodeling tools have received attention for their ability to represent and improve complex systems. The use of metamodeling techniques in optimization problems via simulation has grown considerably in recent years to promote more robust and agile decision-making, determining the best scenario among the solution space to the problem. In this sense, this Chapter discusses the state of the art of metamodeling-based simulation optimization, presenting the gaps, opportunities, and future perspectives found in literature.

Keywords: Metamodeling; Simulation; Optimization; Machine Learning.

1. Introduction

Production systems need to include increasingly effective decision-making tools to ensure organizational competitiveness, customer satisfaction, and cost reductions (Salam and Khan, 2016). According to Law (2013), analytical solutions are effective when systems are relatively simple. However, the author states that most real systems are too complex for analytical analyses. Simulations were developed to predict and evaluate complex, stochastic, and often analytically intractable system performance (Xu et al., 2016).

One advantage of using simulations is reduced decision-making risk, since simulations allow different scenarios to be evaluated without physically interfering in production systems (Helleno et al., 2015). Since not all possible scenarios can be practically evaluated, Simulation Optimization (SO) is used to achieve these goals (Barton, 2009).

The SO field has progressed significantly over the past decade with newly developed algorithms, implementations, and applications. There have been significant contributions to this field of research given the work of researchers and practitioners, mainly in the fields of engineering, industrial operations, mathematical programming, statistics, machine learning, and computer science (Amaran et al., 2016).

SO finds the input values of a simulation model that gives the best outputs (Belgin, 2019). According to Fu (2002), SO can be divided into two phases: 1) generating possible candidates for solving the problem; and 2) evaluating solutions using simulation model estimates. Figure 1 shows this process.

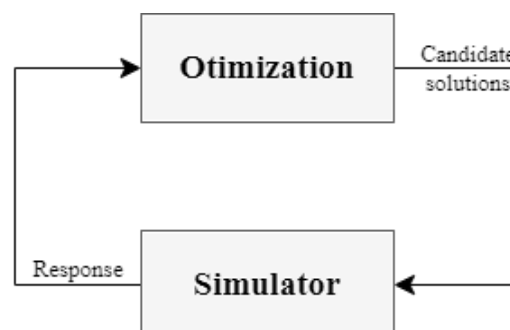


Figure 1: General SO process.
Source: Adapted from Fu (2002).

According to Miranda et al. (2017), several SO techniques have been developed in recent years, e.g., heuristics and metaheuristics, that have still been little explored by scientific literature.

SO methods can be classified into two categories, model-based and metamodel-based methods. Moghaddam and Mahlooji (2017), explain that the optimization and simulation modules are directly integrated with each other in model-based methods, where optimization occurs according to the solutions that need to be evaluated, then the simulation evaluates them. A third component, called a metamodel, exists in metamodel-based approaches, which identifies and estimates the relationship between simulation model inputs and outputs as a deterministic mathematical function that is used to evaluate possible simulation and optimization solutions (Barton and Meckesheimer, 2006; Moghaddam and Mahlooji, 2017; Parnianifard et al., 2020a).

The model-based approach is the most popular and widespread method according to literature. However, when a simulation model is very complex, and when a problem solution space is large, a SO process using these techniques is too slow, and may require hours or months to converge to an acceptable local optimum value, making use unfeasible for several applications (Cai et al., 2019; Miranda et al., 2014). On the other hand, metamodeling techniques have been increasingly used over the last decade to reduce computational optimization times (Parnianifard et al., 2019).

Metamodeling techniques have been applied in several engineering and management branches, e.g., optimally allocating operators and equipment on factory floors (Amaral et al., 2022), planning and controlling traffic to minimize travel times (Xiao Chen, Osorio, and Santos, 2019), structurally designing vehicles to reduce weight (Fu and Sahin, 2004), finding optimal sizes for concrete barrier designs (Yin et al., 2016), and geometrically optimizing laminated composite wind turbine blades to reduce weight (Albanesi et al., 2018), etc.

Currently, several algorithms are used to build metamodels. Some examples include Response Surface Methodology (RSM) (Hassrnaayebi et al., 2019; Shirazi; Mahdavi; Amiri, 2011), the *Kriging* algorithm (Baquela and Olivera, 2019; Pedrielli et al., 2020), the Radial Basis Function (RBF) (Christelis and Mantoglou, 2016; Yin et al., 2016), and Artificial Neural Networks (ANN) (Raei et al., 2019; Storti et al., 2019).

Given these various methods, and the growing importance of metamodeling for complex engineering problems, this study seeks to provide readers with a review of the main metamodeling concepts as applied to SO problems. Furthermore, this chapter will conduct a Systematic Literature Review (SLR), to provide a holistic view of how this topic is addressed in literature.

2. Background

According to Lin and Chen (2015), optimization via simulation is a powerful decision-making tool, since this technique can understand various relationships in complex real-world systems, to identify the best solution to a given problem. According to Souza Junior et al. (2019), SO helps decision-makers invest or allocate resources in new or existing production systems.

Amaran et al. (2016), define SO as an optimized Objective Function (OF), that is subject to constraints, which is then evaluated using stochastic simulations. Carson and Maria (1997), state that SO is a process of identifying the best input values for simulated system variables, and evaluating solutions using a loop created between the optimization algorithm and the simulation model.

The SO field of study has evolved significantly over the last decade, in light of newly developed algorithms, software programs, and applications. Amaran et al. (2016), state that the general formulation of an SO problem can be expressed as in Eq. (1).

$$\begin{aligned}
 & \text{Min } \mathbb{E}_\omega [f(x, y, \omega)] \\
 & \text{s. a. } \mathbb{E}_\omega [g(x, y, \omega)] \leq 0 \\
 & \quad h(x, y) \leq 0 \\
 & \quad x_l \leq x \leq x_u \\
 & \quad x \in \mathbb{R}^n, y \in \mathbb{D}^m
 \end{aligned} \tag{1}$$

Where \mathbb{E}_ω is the expected value of function f , which is evaluated by simulating continuous x inputs, or discrete y inputs, subjected to a random number vector ω . Likewise, restrictions are defined by function g values that are evaluated in each simulation. Problems may also contain other constraints (represented by h) that do not involve random variables, or contain constraints linked to decision variables.

By contrast, Miranda et al. (2017), states that long computational times are required for the optimization algorithm to converge to a good result for SO problems addressing complex systems with very large solution spaces. Barton (2009), states that this has led researchers to develop specialized SO methods, which include:

- ranking and selection;
- heuristics and metaheuristics;
- random searches; and
- metamodeling.

According to De La Fuente and Smith (2017), the more complex the studied system, the more time required for performing optimizations. According to the authors, to obtain good results in a timely manner, one would need to elevate problems to another level of abstraction, i.e., a metamodel.

Metamodeling is commonly used for developing representations of simulation models (De La Fuente et al., 2019). A metamodel can understand relationships between decision variable values and simulation outputs, providing OF approximations faster than the original simulation model (Jalali and Nieuwenhuyse, 2015). In layman's terms, a simulation model is an abstraction of a real system, and similarly, a metamodel is an abstraction of a simulation model (a model of another model), making evaluations simpler and faster, albeit while less accurately representing real systems. Figure 2 shows this relationship.

Once a metamodel has been trained and validated it can replace the original simulation model in SO processes, thereby considerably reducing computational times for optimization convergence, as shown in Figure 3. This process is called Metamodel Based Simulation Optimization (MBSO).

Several metamodeling methods and applications have emerged in literature in recent years to create more reliable metamodels. Barton (2009), stands out, especially for comparing the main methods used for these applications, i.e., RSM, RBF, Kriging, RNA, and spline. Hannah, Powell, and Dunson (2014), proposed an MBSO method based on a semi-convex regression. Yousefi et al. (2018), used ensemble methods in an agent-based simulation model to reduce average wait times in the healthcare sector,

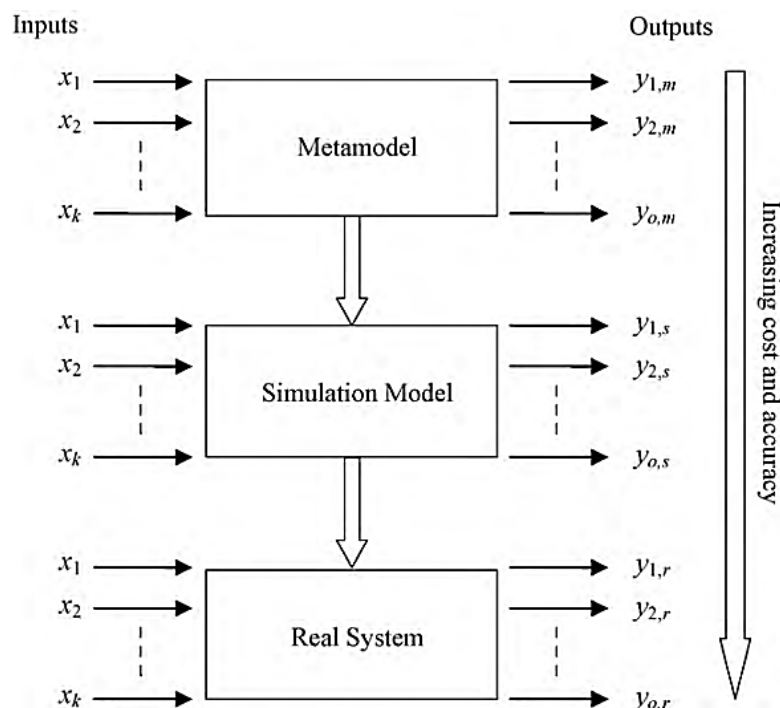


Figure 2: Relationship between real-world scenarios, simulation models, and metamodels.
 Source: (Li et al., 2010).

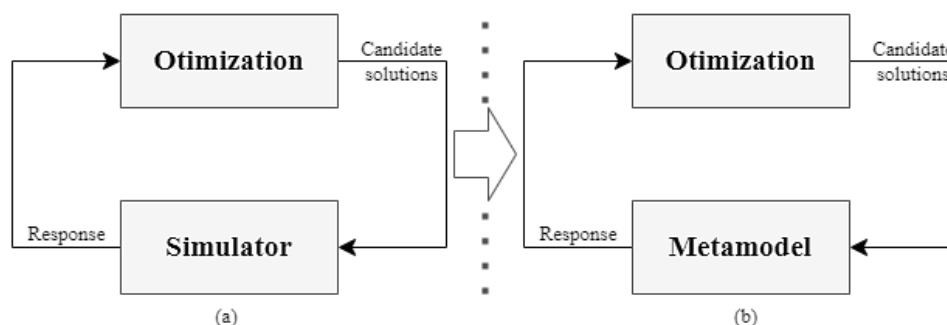


Figure 3: SO process (a) and MBSO process (b).

Souza Junior et al. (2019), studied the integrations between metamodeling, parallel processing and metaheuristics applied to production engineering problems, and Woldemariam and Lemu (2019), used ANN and Genetic Algorithm (GA) metamodeling to optimize fluid dynamics analysis models. Hullén et al. (2020), proposed strategies for dealing with SO uncertainty by combining ML with stochastic programming, robust optimizations, and discrepancy modeling.

Sections 2.1 – 2.3 address the concepts involved in the main machine learning algorithms used to build metamodels, performance metrics, hyperparameter optimization techniques used in these algorithms, and a brief explanation of one of the most widely-used MBSO designs, the Latin Hypercube Design (LHD), all to aid in better understanding the algorithms and techniques common to metamodeling studies.

2.1 Machine Learning Algorithms

2.1.1 Gaussian Process

Also known as the Kriging algorithm, the Gaussian Process (GP) is a non-parametric Bayesian approach using variables following Gaussian distribution (Kleijnen, 2014). The main goal is to define a probability for infinite functions, by adding finite points that result in a function passing directly over these points (Ostergard et al., 2018). Nguyen-Tuong et al. (2009), state that given a set of n training data $\{x_i, y_i\}_{i=1}^n$, the function $f(x_i)$ seeks to learn from input data (x_i) transformed into response variables (y_i), converging according to Eq. (2), where ϵ_i is the Gaussian noise function, with zero mean, and variance σ_n^2 .

$$y_i = f(x_i) + \epsilon_i \quad (2)$$

The GP is a stochastic process, whereby any finite group of model entry points are incorporated into Gaussian distribution, expressed as $f(x) \sim GP[\mu(x), K(X, X)]$, where $\mu(x)$ and $K(X, X)$ are the mean and the covariance functions, respectively (Wan and Ren, 2015). Here, the covariance functions are normally given via a Gaussian kernel, as per Eq. (3) (Nguyen-Tuong et al., 2009).

$$k(x_p, x_q) = \sigma_s^2 \exp\left(-\frac{1}{2}(x_p - x_q)^T W(x_p - x_q)\right) \quad (3)$$

Where σ_s^2 and W are the variance and amplitude of the Gaussian kernel. Furthermore, Eq. (4) shows the joint distribution for the estimated ($f(x_*)$) and observed (y) values, given a given input x_* . By conditioning the joint distribution, one can extract the predicted mean value ($f(x_*)$), and the associated variance ($V(x_*)$), as per Eq. (5) and Eq. (6) (Nguyen-Tuong et al., 2009).

$$\begin{bmatrix} y \\ f(x_*) \end{bmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} K(X, X) + \sigma_*^2 & k(X, x_*) \\ k(x_*, X) & k(x_*, x_*) \end{bmatrix}\right) \quad (4)$$

$$f(x_*) = k_*^T (K + \sigma_*^2 I)^{-1} y = k_*^T \alpha \quad (5)$$

$$V(x_*) = k(x_*, x_*) - k_*^T (K + \sigma_*^2 I)^{-1} k_* \quad (6)$$

Where $k_* = k(X, x_*)$, $K = K(X, X)$, and α are the SO prediction vectors.

2.1.2 Polynomial Regression

PR is a collection of statistical and mathematical techniques that can be applied to develop, improve, and optimize processes. It has been successfully used in recent decades to represent stochastic systems (Parnianifard et al., 2020a). PR was developed over 50 years ago for exploration and extrapolation applications (Barton and Meckesheimer, 2006), and according to Parnianifard et al. (2019), literature cites PR as being used to:

1. approximate relationships between dependent and independent variables;
2. investigate and determine the best process operating conditions; and
3. implement more robust responses.

According to Myers, Montgomery, and Anderson-Cook (2016), PR for a given response is based on Taylor series expansion theory to a data set. The general formulation for linear regression (first-order PR) is given by Eq. (7) (Parnianifard et al., 2019).

$$y = f(X) = \hat{\beta}_0 + \sum_{i=1}^k \hat{\beta}_i x_i + \varepsilon \quad (7)$$

Where k is the number of independent variables. If more complex response surfaces cannot be represented by linear PR, quadratic (second-order) PR can be used. Its general formulation is given in Eq. (8) (Dengiz and Belgin, 2015; Osorio and Bierlaire, 2013; Parnianifard et al., 2019; Parnianifard et al., 2020a).

$$y = f(X) = \hat{\beta}_0 + \sum_{i=1}^k \hat{\beta}_i x_i + \sum_{i=1}^k \hat{\beta}_{ii} x_i^2 + \sum_{i=1}^k \sum_{i' < i''=2}^k \hat{\beta}_{ii'} x_i x_{i'} + \varepsilon \quad (8)$$

Where $\hat{\beta}_0$, $\hat{\beta}_i$, $\hat{\beta}_{ii}$ and $\hat{\beta}_{ii'}$ are the unknown coefficients, and ε is the model noise. The number of expressions p is $p = k + 1$ for linear regression, $p = \frac{1}{2}(k + 1)(k + 2)$ for quadratic regression, and $p = \frac{1}{6}(k + 1)(k + 2)(k + 3)$ for cubic regression (Parnianifard et al., 2019). One of the main advantages of PR is the model interpretability and adaptability to different data types. However, Ostergard et al. (2018), state that these can be computationally inefficient compared to more sophisticated high dimension expansions, e.g., GP and SVM.

2.1.3 Artificial Neural Networks

ANNs have been widely applied to classification, regression, and time-series problems. They consist of an input layer, one or more hidden layers, and an output layer. (Hassoun, 1995). The general ANN and neuron structures are shown in Figures 4 and 5, respectively.

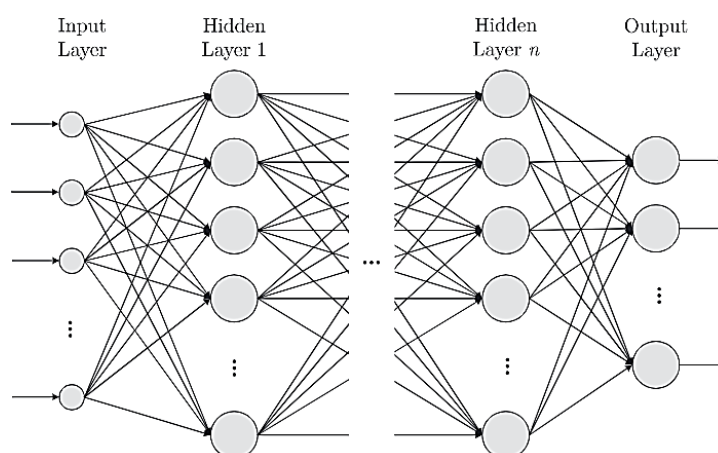


Figure 4 - Structure of an Artificial Neural network
Source: Adapted from Storti et al. (2019)

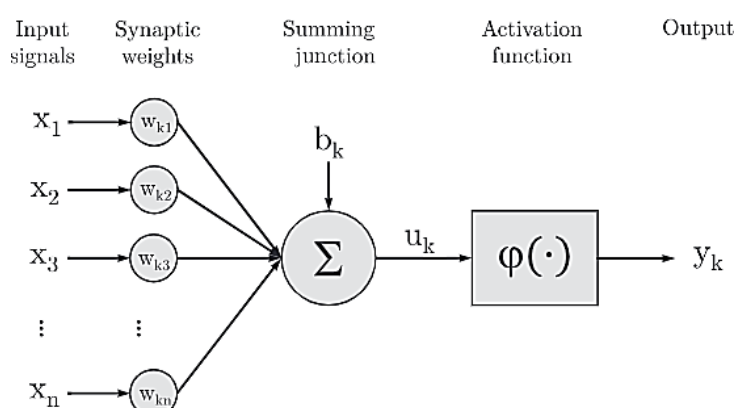


Figure 5 - General structure of a Neuron
Source: Adapted from Storti et al. (2019)

According to Li et al. (2010), information transmission occurs via neuron connections through layers. Also according to the authors, a typical ANN can be represented with three layers and one neuron in the output layer, as per the following formula:

$$\hat{y} = \hat{f}(X) = \sum_{j=1}^j w_j f \left(\sum_{i=1}^k v_{ij} f(x_i) + a_j \right) + \beta + \varepsilon \quad (9)$$

Where:

- X is a k dimension vector;
- $f(\cdot)$ is the transfer function;
- ε the random error with mean 0;
- v_{ij} represents the connection weights between the i neurons in the input layer and the j neurons in the hidden layer;
- a_j is the bias of the j^{th} hidden layer neuron;
- w_j is the weight of connections between the j^{th} hidden neuron and the output neuron;
- j is the total number of neurons in the hidden layer; and
- β is the output neuron bias (Li et al., 2010).

According to Oliveira et al. (2010), weights and biases are adjusted using supervised training algorithms. The back-propagation algorithm is most widely used, where the model is fed with standard inputs, and the signal is propagated to the output layer to generate predictions. This error then adjusts the output neuron, hidden neuron, and input neuron connection weights.

2.1.4 Random Forest

Breiman (2001), introduced the Random Forest (RF) method, which combines many randomly uncorrelated decision trees to obtain predictions based on average votes among all individuals (Li et al., 2020). RF is one of the most efficient ML algorithms in terms of accuracy and execution speed for large databases. It is also a simple method, and its variables and results can be easily interpreted (Crisci et al., 2012).

Dai et al. (2018), state that RF is a combination of decision trees $\{h(x, \theta_k), k = 1, 2, \dots\}$, where each tree receives a random vector $\{\theta_k\}$. Outputs are numerical, and training data are independently and identically distributed in regression problems (Breiman, 2001). The root mean square error associated with each tree $h(x)$ is defined by Eq. (10), for the X vector and Y output (Dai et al., 2018).

$$E_{X,Y}(Y - h(X))^2 \quad (10)$$

The procedure for adjusting the RF is simple. The first step involves adjusting all m independent trees. Then a sample size N is obtained at each interaction, and k attributes are randomly selected to generate new uncorrelated trees (De La Fuente and Smith, 2017). The individual predictions of each m^{th} tree are guided by a set of *if-then-else* rules, characterized by Θ (Hastie et al., 2008). According to Breiman (2001), an FR forecast is calculated using the average from the k trees $h(x)$, belonging to the forest, as per Eq. (11).

$$f(X) = \frac{1}{M} \sum_{m=1}^M T(X: \theta_m) \quad (11)$$

Where M is the number of trees, $T(X: \theta_m)$ is the model for a single tree, X is the training data vector, and θ_m is the set of rules for the trees.

2.1.5 Gradient Boosted Trees

The Boosting method was initially proposed by Schapire (1990), to build highly accurate models based on low learning converging algorithms, or decision trees for Gradient Boosted Trees (GBT). Friedman (2002), proposed a variation of this method called Stochastic Gradient Boosting, introducing randomness to the procedure proposed by Schapire (1990). Friedman (2002), states that the trees are trained using a random sample \tilde{N} of the training data at each iteration of the algorithm without replacements, so $\tilde{N} < N$, where N is the size of the training data. The GBT algorithm differs from bagging methods, since the first seeks to sequentially build models (trees), while the second is adjusted based on models built in parallel (Xia et al., 2017).

This algorithm is based on sets of decision trees. A new tree is generated at each interaction based on the residues of previous trees (De La Fuente and Smith, 2017). Given

a dataset $\{x_i; y_i\}$ size N , input attributes are represented by x_i , while y_i is the response variable. The objective of the algorithm is to determine the function $F^*(x)$ that expresses the relationship between x and y to minimize the loss function as per Eq. (12). This relationship is additively approximated in Eq. (13) (Xia et al., 2017).

$$F^*(x) = \arg \min_f \Psi(y, F(x)) \quad (12)$$

$$F_m^*(x) = \sum_{m=0}^M \beta_m h(x; \Theta_m) \quad (13)$$

The ideal base learners ($h(x; \Theta_m)$) are represented by $\beta_m h(x; \Theta_m)$, where β_m and Θ_m are optimal coefficients and parameters for the base learner, respectively. Friedman (2002), presents a pseudocode for GBT as described in Algorithm 1.

Algorithm 1: Stochastic Gradient Boosted Trees

```

1   $F_0(x) = \arg \min_{\gamma} \Psi(y_i, \gamma)$ 
2  For  $m = 1$  to  $M$  do:
3     $\{\pi(i)\}_1^N = \text{rand}_{\text{perm}} \{i\}_1^N$ 
4     $\tilde{y}_{\pi(i)m} = - \left[ \frac{\partial \Psi(y_{\pi(i)}, F(x_{\pi(i)}))}{\partial F(x_{\pi(i)})} \right]_{F(x)=F_{m-1}(x)}$ ,  $i = 1, \tilde{N}$ 
5     $\{R_{lm}\}_1^N = L - \text{terminal node tree}(\{\tilde{y}_{\pi(i)m}, x_{\pi(i)}\}_1^{\tilde{N}})$ 
6     $\gamma_{lm} = \arg \min_{\gamma} \sum_{x_{\pi(i)} \in R_{lm}} \Psi(y_{\pi(i)}, F_{m-1}(x_{\pi(i)}) + \gamma)$ 
7     $F_m(x) = F_{m-1}(x) + \nu \cdot \gamma_{lm} l(x \in R_{lm})$ 
8  End For.
```

Source: Friedman (2002).

Where $\Psi(y_i, \gamma)$ represents the loss function, R_{lm} represents each region formed by the l nodes of the m^{th} trees, and $\pi(i)\}_1^N$ represents integer (1, ..., N) permutations for training data $\{y_i, x_i\}_1^N$.

Line (4) in the algorithm gives the gradient of the loss function, or the pseudo-residues of the last iteration, and γ_{lm} is the set of optimal expansion coefficients corresponding to each tree node in the m^{th} iteration.

2.1.6 Support Vector Machine

A Support Vector Machine (SVM) from Vapnik (1999), is one of the most popular ML methods proposed in recent years (Cao et al., 2020). It was initially built to classify problems but has been widely used for regression problems (Cao et al., 2020, Hastie et al., 2008).

For classification problems, the method seeks boundaries to separate training data into respective classes with the least amount of error. A hyperplane is defined based on vectors built from observations located at boundaries, so data on different sides of the hyperplane are allocated in different classes (Ostergard et al., 2018).

Jeng (2006), explains that SVM tries to find a maximum margin between two classes using linear vector combinations for these types of problems, and this is solved

using quadratic programming with linear constraints. According to the authors, this algorithm can be extended to solve nonlinear regression problems by introducing the \mathcal{E} parameter, which represents the loss function. Li et al. (2010), gave a mathematical formulation of SVM as per Eq. (14):

$$\hat{y} = \hat{f}(X) = w\phi(X) + b \quad (14)$$

Where $X \in \mathfrak{R}^k$ is the k -dimensional input, ϕ is a nonlinear transformation from \mathfrak{R}^k to \mathfrak{R}^h , with $h > k$, and $\hat{f}(X)$ describes the hyperplane in \mathfrak{R}^h . The authors emphasize that coefficients w and b can be obtained by minimizing the risk function, as per Eq. (15).

$$R_{svm}(C) = C \sum_{i=1}^n C(\hat{f}(X_i), y_i) + \frac{1}{2} \|w\|^2 \quad (15)$$

Where:

$$C(\hat{f}(X_i), y_i) = \begin{cases} |\hat{f}(X) - y| - \mathcal{E} & \text{se } |\hat{f}(X) - y| \geq \mathcal{E} \\ 0 & \text{Contrary Case} \end{cases} \quad (16)$$

The final SVM method prediction function is obtained after solving the minimization problem via Eq. (17) (Vapnik, 1999).

$$f(X, a_i, a_i^*) = \sum_{i=1}^n (a_i - a_i^*)K(X, X_i) + b \quad (17)$$

Where X is the new model input, $K(X, X_i)$ is the Kernian function representing the distance between X and X_i , and a_i, a_i^* and b are coefficients obtained by minimizing the risk function (Li et al., 2010).

Oliveira et al. (2010), state that SVM efficiency depends on correctly choosing the hyper-parameters C , epsilon (\mathcal{E}) and Kernel (K). According to Cherkassky and Ma (2004), the ideal C value of can be calculate via Eq. (18):

$$C = \max(|\bar{y} + 3\sigma_y|, |\bar{y} - 3\sigma_y|) \quad (18)$$

Where \bar{y} and σ_y are the mean and standard deviations of the y values in the training data. To calculate \mathcal{E} , the authors recommend making a first pass with the algorithm using the default parameters to calculate the prediction error δ associated with the database, where $\delta = \hat{y} - y_i, (i = 1, \dots, n)$.

$$\mathcal{E} = 3\sigma \sqrt{\frac{\ln n}{n}} \quad (19)$$

Where σ is the standard deviation of error δ , and n is the number of database examples.

2.2 Performance Metrics

An important step in building metamodels is evaluating them to choose the model that best fits the data, and aid in selecting and/or optimizing the parameters. Bergmeir and Benítez (2012), stated that the Mean Square Error (MSE) or Root MSE (RMSE), the Mean Absolute Error (MAE), and the Mean Absolute Percentage Error (MAPE), were the main error metrics cited in literature. According to Hu et al. (2006), the coefficient of correlation (R^2) was another important regression metric.

Taking y_i as the observed (actual) value at point i , \hat{y}_i as the predicted value, and n as the size of the test data, RMSE, MAE and MAPE can be calculated as per Eq. (20), Eq. (21), and Eq. (22), respectively (Bergmeir and Benitez, 2012; Chai and Draxler, 2014).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (20)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (21)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| 100 \frac{y_i - \hat{y}_i}{y_i} \right| \quad (22)$$

There is no consensus as to the model accuracy values for metamodel evaluation metrics. However, one could state that a metamodel with $R^2 > 0.90$ is reasonable, $R^2 > 0.95$ is accurate, and $R^2 > 0.99$ is almost perfect (Ostergard et al., 2018).

R^2 is an accuracy metric with an amplitude that varies from 0 to 1. As shown in Eq. (23), it is calculated by dividing the sum of the squares of the residuals $(y_i - \hat{y}_i)^2$ by the total sum of the squared residuals $(y_i - \bar{y}_i)^2$. High R^2 values correspond to good model and training data fits, but do not guarantee reduced errors (Hu et al., 2006).

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad (23)$$

Where:

$$\bar{y}_i = \frac{1}{n} \sum_{i=1}^n y_i \quad (24)$$

2.3 Hyper-Parameter Optimization

Choosing adequate hyper-parameters for ML algorithms is a fundamental step for increasing the learning capacity of these algorithms. Given that these parameters cannot be manually defined, one must use optimization techniques to find the best performing configuration extraction from the algorithm (Xia et al., 2017).

Currently, there are several optimization methods for selecting hyper-parameters, like AG, ant colony, grid search, Bayesian optimization, etc. (Hullen et al., 2020; Kim; Boukouvala, 2020; Li et al., 2019). In our study, we chose to use GA given its execution speed, its ability to adapt to different variable types, and given that the initial values do not need to be defined to perform optimization. Table 1 gives a brief comparison of the main hyper-parameter optimization methods.

Table 1 - Comparison of common methods for hyper-parameter optimization.

Optimization Methods	Strengths	Weaknesses	Complexity (time)
<i>Grid Search</i>	Simple	Takes a lot of time and is efficient only for categorical variables	n^k
<i>Randon Search</i>	Is more efficient than the Grid Search	Does not consider previous results and is not efficient with conditional variables	n
<i>Gradient-based</i>	Allows for parallelism	Supports only continuous variables and detects only large locations	n^k
<i>Bayesian GP</i>	Has fast convergence for continuous variables	Low capacity for parallelism and is not efficient for conditional variables	n^3
<i>Bayesian TPE</i>	Has fast convergence for continuous variables	Low capacity for parallelism	$n \log n$
<i>Sequential model-based algorithm configuration</i>	Is efficient for all variable types	Low capacity for parallelism	$n \log n$
<i>Hyperband</i>	Maintains conditional dependencies	Not efficient for conditional variables and requires small subsets to be relevant	$n \log n$
<i>Bayesian hyperBand</i>	Is efficient for all variable types	Requires small subsets to be representative	$n \log n$
<i>Genetic algorithm</i>	Allows for parallelism	Low capacity for parallelism	n^2
<i>Particle swarm</i>	Is efficient for all variables types	Requires proper start up	$n \log n$
	Allows for parallelism		

k is the number of hyperparameters, and n is the number of hyperparameter values.

Source: Adapted from Yang and Shami (2020).

A Genetic Algorithm (GA) is a heuristic technique that was inspired by the theory of natural selection. It was initially proposed by De Jong (1975), and Holland (1975), and is based on Darwin's theory of the survival of the fittest. It can be successfully applied to linear and nonlinear programming problems (Bashiri and Geranmayeh 2011; Lessmann et al., 2006). The search space in GA comprises candidate solutions, each represented by a chromosome (Alabas et al., 2002). Each chromosome comprises genes, representing

encoded parameters for a candidate solution, and aptitude, or OF, to reach the best solution for the problem. The objective of GA is to combine genes to obtain new chromosomes with better OF values (Li et al., 2010).

The GA comprises three basic operators, i.e., reproduction, crossover, and mutation. Reproduction is a process by which individual parameters, or genes, are copied based on their OF values. Individuals with higher OF values are more likely to generate one or more offspring in the next generation. When a parameter is reproduced it creates an exact replica of itself. Then, the crossover operator randomly combines newly produced parameters. Finally, mutations generate changes to parameter values, playing a secondary role, because mutation rates are generally low (Alabas et al., 2002; Storti et al., 2019).

GA applied to hyper-parameter optimization was studied by Hou et al. (2019), and Hullen et al. (2020). This method evaluates several possible parameter combinations $\{a_1, \dots, a_i\}$ of the ML algorithm. The cross-validation technique is used to calculate the error associated with each combination. k -fold cross-validation is one of the most important techniques for evaluating regression and classification methods. It divides training data both equally and randomly into k parts, using $k-1$ parts for training the algorithm, and the remainder for testing. The error of each array is calculated k times, and a different part k_i is selected for testing at each iteration, with $i = \{1, \dots, k\}$. Then, performance is measured based on averaged errors in k validation interactions, which is a more robust metric than if it were measured using just one observation (Bergmeir and Benítez, 2012). Roy and Datta (2019), suggested using 10-fold cross-validation.

2.4 Design of Experiments

According to Montevechi et al. (2012), experiments are a series of tests with purposeful changes made to input variables to identify and analyze the effects of these changes to response variables within the system. According to the authors, the Design of Experiments (DOE), is a set of methods for designing and conducting experiments. Several DOE techniques have been cited in literature, e.g., orthogonal arrays (Miranda et al., 2017), Central Composite Designs (CCD), (Villarreal-Marroquín et al., 2013), factorial designs (Dengiz et al., 2006), etc.

According to the Structured Literature Review (SLR) in Section 4, the most common DOE method for metamodeling projects is the Latin Hypercube Design (LHD). According to Parnianifard et al. (2020), LHD is a computational DOE technique based on space-filling. For n input variables, m sample points are randomly obtained within m equally probable intervals. Figure 6 gives an example of an LHD array with $n=2$ decision variables, and a $m=7$ sample size.

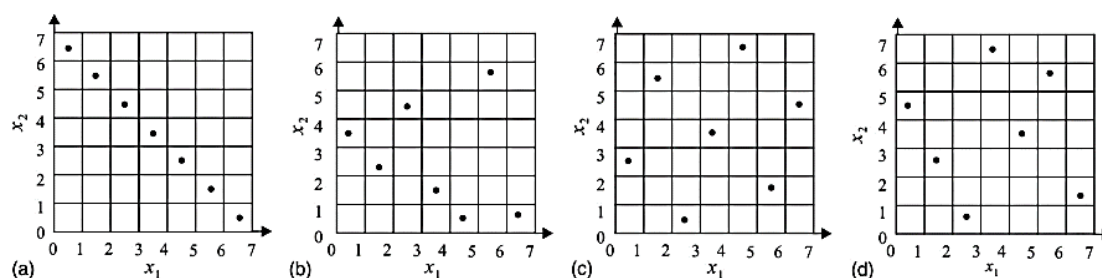


Figure 6: Examples of LHD arrays with 2 input variables and 7 sample points Source: Adapted from Luo, Ji and Lu (2019)

According to Parnianifard et al. (2020), the LHD procedures are:

- The value ranges for each variable are divided into m subranges of equally probable magnitudes. The number of sampling points m is defined by the researcher. In general, it should be greater than the total points produced by the CCD.
- The second step is allocating the m points between each lower and upper interval limit, with each point (1, 2, ..., m) in one row and column of the sample space.

3. Systematic Literature Review

As shown in the previous section, there is a wide range of methodologies and applications surrounding MBSO. Sections 3 and 4 provide a Systematic Literature Review on the topic to provide the reader with an overview of how metamodeling has been approached both in academia and in practice.

3.1 Literature Reviews on Metamodeling Simulation Optimization

We conducted a search in the Scopus and Web of Science scientific databases in July 2020 to look for studies including Literature Reviews (LR) on MBSO, or other related topics. In this step we searched for terms related to metamodeling (metamodel, metamodeling, meta-model, or machine learning), optimization via simulation (Simulation optimization, simulation-based optimization, or optimization via simulation) using Boolean logic (AND/OR), and Literature Reviews (literature reviews, the state-of-the-arts, overviews, or surveys), with a filter for documents in English.

12 LRs on topics related to MBSO were returned, however, these focused on a certain stages of metamodeling projects, one or a few metamodeling algorithms or types of simulations, and none contained complete discussions on MBSO knowledge that should be used by academics and practitioners for developing projects. This section contributes by offering a comprehensive view of MBSO using the scientific rigor of SLR (Y. Xiao and Watson 2019), and CIMO (Context, Interventions, Mechanisms, Outcomes) (Denyer et al., 2008).

The first study returned was published from the Winter Simulation Conference in 1986, addressing the frequency domain experimentation technique and RSM for dealing with discrete variable problems (Schruben, 1986). At the same conference, twenty-three years after the publication of the study by Schruben (1986), Barton (2009) gave an overview of SED metamodeling techniques, describing the main algorithms used in metamodeling according to literature i.e., RSM, RNA, Kriging, spline, RBF, and provided a framework for conducting metamodeling projects using local or global searches. That same year, Kleijnen (2009), published a review of Kriging metamodeling and its association with techniques like bootstrapping, DOE, sensitivity analysis, and optimization.

In 2010, a study systematically compared five SO metamodeling algorithms (i.e., RNA, RBF, Kriging, SVM, and multivariate regression splines) for decision support systems (LI et al., 2010). In 2014, a study was carried out on Kriging metamodeling with parametric and distribution-free bootstrapping applied to optimization problems via

deterministic and/or stochastic simulations (Kleijnen 2014). That same year, Figueira and Almada-Lobo (2014), studied hybrid SO method classifications and taxonomies, explaining questions on procedures and methods related to metamodeling.

In 2015, Dellino et al. (2015), described robust optimization concepts using Kriging metamodeling in a chapter of their book, applying these to an economic order quantity inventory model. A survey was carried out again in 2015 on the practical use of search experimentation techniques, i.e., SO, DOE, and metamodeling applied in SED (Hoad et al., 2015). Two years later, an LR was conducted presenting a general review of several metamodeling algorithms (i.e., polynomial regression, Kriging, RBF, SVM, and mixed metamodels), DOE techniques (i.e., geometrical, statistical, and optimal designs), and mechanisms (i.e., expected improvements, efficient global optimizations, bumpiness, etc.), for problems containing continuous variables (Vu et al. 2017). That same year, two other LRs addressed metamodeling using Kriging and polynomial regression (Kleijnen 2017a, 2017b).

Finally, only one study used a SLR to study metamodeling for robust black-box SO subjected to uncertainties (Parnianifard et al. 2019). These studies served as the basis for conducting our SLR on MBSO. We should mention that our SLR did not seek to conduct an exhaustive review of literature on MBSO, but rather to provide academics and practitioners in the area with a broad view of the subject by consolidating and explaining existing MBSO projects, while also defining the frontier for knowledge in this area, by identifying and presenting current opportunities and challenges related to this topic.

3.2 The Systematic Literature Review Method

A SLR is a methodology for outlining the theoretical-scientific basis necessary for understanding a topic. This is done by collecting, understanding, synthesizing, and evaluating a set of scientific studies (Levy and Ellis, 2006; Xiao and Watson, 2019). According to Sousa Junior et al. (2019), SLRs assist researchers in performing methodological steps, e.g., defining problems, selecting methods, collecting data, and performing analysis, to mitigate flaws in the conclusions of studies, thereby offering greater clarity on employed research procedures.

SLR is not the same as a traditional LR, but rather seeks to clearly explore a specific issue by consulting existing literature. An SLR differs from other LR methodologies given its exact principle, whereby researchers must pre-specify and clarify their inclusion/exclusion criteria to readers (Denyer and Tranfield, 2009). Oliveira et al. (2016), point out that one of the advantages of SLR relative to traditional RL methods, is that SLRs avoid data bias and distortions. A SLR mainly seeks to promote advancements and to develop new studies, given the fact that a new frontier of knowledge can only be established after defining the current frontier.

3.3 Employing the Systematic Literature Review Method

Our SLR adopted the same procedures as Oliveira et al. (2016), and Sousa Junior et al. (2019), to systematically investigate MBSO knowledge, by identifying borders and gaps in existing literature, thereby providing new guidelines for future studies. This was done in four main stages, i.e., planning, researching/sorting, analyzing/synthesizing, and presenting the results.

3.3.1 Planning

This step was developed by a panel of experts in the area to better investigate this field of research. The panel included 3 Ph.Ds with academic and practical experience in MBSO projects, in addition to a Ph.D. and M.Sc. researching the same topic.

The study was inspired given the expressive growth in MBSO publications, the differing applications to knowledge areas, and given the lack of studies comprehensively and systematically synthesizing and discussing knowledge on this subject. We expect that our results can aid all project stages within this area, and present the state-of-the-arts for MBSO by compiling existing studies, and mainly proposing future opportunities for additional research.

The main SLR objectives are:

- identify and explore the main MBSO aspects in studies;
- analyze and summarize the results;
- discuss the results, define the state-of-the-arts, and any research gaps and trends that could be further explored.

To do this, Research Questions (RQ) need to be defined to extract and analyze intrinsic knowledge on the topic. Several frameworks and methods for structuring RQs have been addressed in literature e.g., PICO, PICOS or PICOC (Population, Interventions, Comparisons, Outcomes, and Context) (Kloda et al. 2020), SPICE (Settings, Perspectives, Interventions, Comparisons, and Evaluations) (Booth, 2006), SPIDER (Samples, Phenomenon of Interest, Designs, Evaluations, and Research Type) (Methley et al. 2014), ECLIPSE (Expectations, Clients, Location, Impacts, Professionals, and Services) (Wildridge and Bell 2002), etc.

We used the CIMO-logic framework as proposed by Denyer et al. (2008), which combines problem Contexts requiring Interventions, generated using a Mechanism, in order to produce Outcomes. This approach has been used by several authors to study engineering problems (Aouadni et al., 2019; Balan, 2018; Makhashen et al., 2020; Costa et al., 2020; Ilk et al., 2020; Kochan and Nowicki, 2018; Tanila et al., 2020). The RQs in our study were divided into the four elements outlined in the CIMO-logic method:

- context: real or theoretical problems...
- interventions: ... using metamodeling...
- mechanisms: ...combined with optimization techniques via simulations...
- outcomes: ...to find the best performing solution within the project scope and constraints.

The RQs were formulated in compliance with these objectives, within the scope of this study, and to identify existing gaps in literature. The RQs are:

- RQ1) What is the role of metamodeling in SO projects, and what are its main applications? (Context)
- RQ2) What are the main simulation and optimization techniques used with metamodeling? (Mechanisms)
- RQ3) What are the most widely-used metamodeling algorithms in MBSO? (Interventions)

- RQ4) What sampling strategies are normally used? (Interventions)
- RQ5) How are results measured/validated? (Outcomes)

The information extracted from the articles was divided into two categories to answer the RQs, as shown in Figure 7.

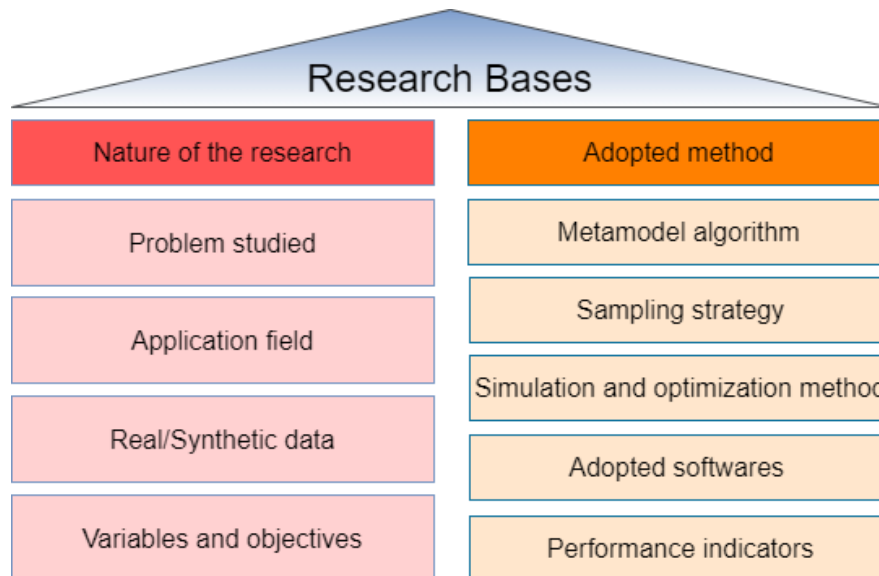


Figure 7: Research pillars and article extracted data

The first category investigates the nature of the work relative to the area of industry that the studies focus on, what types of problems are addressed, whether the problem is real or theoretical (based on data from literature), and how decision variables and problem objectives are framed. The second category deals with method-related issues e.g., simulation type, the metamodeling algorithms used, the search engines used in the optimizations, sampling strategies, software programs, and how results were measured. Finally, the main MBSO frameworks from literature will be presented, along with two new frameworks that complement the main points found in the SLR.

3.3.2 Researching and Sorting the Articles

We conducted exploratory research in the Scopus and Web of Science (WoS) scientific databases on the topic, since these databases are the main data citation sources (Mongeon and Paul-Hus 2016), and are representative of publications on the subject (AGHAEI et al., 2013). Thus, terms related to metamodeling (metamodel, metamodeling, meta-model, and machine learning), and SO (Simulation optimization, simulation-based optimization, optimization via simulation, and SO) were searched for in both databases using Boolean Logic (AND/OR) in the process. The meta-search returned 188 articles from the WoS and 293 articles from Scopus, resulting in 481 studies.

The initial article selection phase for downloading was made by reading the titles/abstracts of the articles that met the following inclusion criteria:

- a) contained SO metamodeling search terms in titles/abstracts/keywords;
- b) published in peer-reviewed journals;
- c) written in English; and
- d) had full versions available for download.

209 studies were excluded from the search after applying the inclusion criteria for failing to meet at least one of the previously defined criteria. 3 (0.6%) did not have the search terms in the titles, abstracts, or keywords, 203 (42, 2%) were not published in peer-reviewed journals, 5 (0.8%) were not written in English, and 1 was not available for download. 269 articles were downloaded in total for quality assessment, 144 from Scopus, and 125 from the WoS. Then 104 articles were removed due to duplications.

Quality assessments consist of reading the titles, abstracts, objectives, methods, results, and discussions. 21 articles were excluded from analysis at this stage for the following reasons, 11 (52.3%) mentioned “machine learning” but did not use ML or metamodeling, 8 (38.1%) used metamodeling/ML for purposes other than SO, and 2 (9.5%) did not use simulations. 143 articles were selected for full reading after the quality assessment for information extraction and data analysis. Figure 8 shows the search/screening procedures.

Searching/screening procediments			
	Scopus	WoS	Total
Databases selection			
Search the literature	n = 293	n = 188	n = 481
Screen for inclusion	n = 144	n = 125	n = 165
Assess quality	n = 126	n = 121	n = 143

Figure 8: Searching/Sorting

Figure 8 shows the results of the search/sort phase. 293 articles were downloaded from Scopus, and 188 from WoS after the search phase. 144 articles from Scopus and 21 from WoS were included for screening. 104 eligible articles from WoS were excluded because they were duplicates with Scopus. Finally, 126 studies from Scopus and 17 from WoS were selected for analysis and data extraction in the quality assessment phase.

3.3.3 Analysis and Synthesis

The articles were evaluated in 9 items to answer the RQs, as shown in Figure 7, corresponding to:

1. define the problem;
2. industrial sector;
3. data source;
4. variables and objective types;
5. metamodeling algorithms;
6. sampling strategies;
7. simulation and optimization methods;
8. software used;
9. performance indicators and validation methods.

Similar nomenclature standardizations were performed after data extraction, e.g., the gaussian process and Kriging were understood as being the same algorithm, and the Latin Hypercube Design and Latin Hypercube Sampling refer to the same DOE technique.

Descriptive data, table, and graph statistics were made after the synthesis using the Software R program. The results were analyzed in line with the RQs related to the three aforementioned categories, serving as a basis for the results and conclusions that aided us in developing a body of evidence on the main MBSO perspectives discussed in this SLR.

3.3.4 Results Presentation

The results were summarized in graphs and tables to discuss the topic, revealing compelling SLR insights. We reflected on and identified the gaps and perceptions on the development of this theme within the scientific community based on the analysis of the interactions between SO and metamodeling.

The main results are summarized in Tables 2 - 6 below following the analyses and syntheses, wherein all RQs will be answered, and where we discuss the state-of-the-arts for MBSO, presenting practices observed in literature, and possibilities for future research on this topic. The 143 articles were divided into two categories according to the problem origin, e.g., real or theoretical, to elucidate the relationship between academia and practical problems facing industry/service providers.

4. Results and SLR Discussion

This section presents the data obtained from the articles and answers the RQs, in the order following the pillar arrangements, i.e., nature of research, methods, and state-of-the-arts.

4.1 Nature of research

This section discusses the “nature of research” pillar, which is directly linked to RQ1 (“What is the role of MBSO, and what are its main applications?”). The main problem-related aspects and industrial sectors, economic sectors, and problem specificities like objectives and sizes studied in literature are presented. Table 2 shows the problem type, and industrial and economic sectors.

The objects of study in the articles were classified and grouped into more than 18 problem classes. Given the wide variety of problems, only the six most frequent classes are shown in Table 2 i.e., Process Control, Stock Control, Product Design, Resource Allocation, Mathematical Models, Traffic Models, together constituting more than 65% of the problems found. The most studied issue was using MBSO for control processes to optimize production parameters for certain goods. Most works (74.1%), addressed this problem type relative to managing water resources to find well-water injection and extraction pumping rates for supplying demand while reducing energy costs, wear and tear, and watershed contamination e.g., (Ataie-Ashtiani et al. 2014; Candelieri et al., 2018; Hussain et al., 2015; Roy and Datta 2019a; Timani and Peralta 2017). Other applications were determining optimal parameters for plastic injection processes (Dang 2014; Villarreal-Marroquín et al., 2011, 2013), cutting parameters for semiconductor wires (Monostori and Viharos 2001), and sheet metal pressing (WANG et al., 2018).

Table 2 - Nature of the research.

Problem Type	Real	Theoretical	Total	Cum. %
Process Control	16 (10.4%)	10 (07.2%)	26 (17.2%)	17.2
Stock Control	01 (00.6%)	20 (13.1%)	21 (13.7%)	31.2
Product Design	08 (05.2%)	08 (05.2%)	16 (10.5%)	41.6
Resource Allocations	11 (07.2%)	05 (03.3%)	16 (10.5%)	52.1
Mathematical Models	00 (00.0%)	11 (07.2%)	11 (07.2%)	59.3
Traffic Models	08 (05.2%)	01 (00.6%)	09 (5.88%)	65.1
Industrial Sector				
Not Specified	03 (02.3%)	36 (27.9%)	39 (30.2%)	30.2
Water Management	19 (14.7%)	10 (07.7%)	29 (22.5%)	52.7
Urban Traffic	18 (13.9%)	02 (01.5%)	20 (15.5%)	68.2
Energy Generations / Distribution	06 (04.6%)	01 (00.8%)	07 (05.4%)	73.6
<i>Health care</i>	06 (04.6%)	00 (00.0%)	06 (03.9%)	78.3
Pharmaceutical / Chemical Industry	03 (02.3%)	02 (1.5%)	05 (03.9%)	82.2
Economic Sector				
Not Specified	03 (02.3%)	36 (27.9%)	39 (30.2%)	30.2
Primary	19 (14.7%)	10 (07.7%)	29 (22.4%)	52.7
Secondary	15 (11.6%)	07 (05.4%)	22 (17.0%)	69.8
Tertiary	33 (25.5%)	06 (04.6%)	39 (30.2%)	100.0

The second problem class was related to inventory controls and management activities, e.g., sizing the number of kanbans for shop floors e.g., (Alabas et al., 2002; Dengiz; Alabas-Uslu and Dengiz, 2009; Hurrion, 1997; Huyet and Paris, 2001), and lot size (Sousa Junior et al., 2019; Greenwood et al., 1998; Hachicha, 2011; Horng and Lin, 2017; Jackson et al., 2019; Moghaddam and Mahlooji, 2016, 2017; Nezhad and Mahlooji, 2014; Parnianifard et al., 2020b; Zakerifar et al., 2011). Both dealt with theoretical cases, where data sources came from other articles, or from classic problems in specialized literature. In our SLR, only one article addressed the stock control problem by studying a real case using metamodeling via Kriging in discrete SO to optimize emergency fuel stocks in ships (Quan et al. 2013).

Many authors have also studied MBSO techniques to optimize product designs. Examples include applications related to the automotive (Fu and Sahin, 2004; Ivanova and Kuhnt, 2014; Stork et al., 2008), energy (Sharif and Hammad 2019; Storti et al., 2019), hydropower (Bananmah et al., 2020; Mooselu et al., 2019), and transit infrastructure (Yin et al., 2016) sectors. Many authors have also studied MBSO applied to resource allocation problems (e.g., (Coelho and Pinto 2018; Song et al., 2005; Yousefi and Yousefi 2019; Zeinali et al., 2015), to synthetic mathematical functions (e.g., Baquela and Olivera, 2019; Kim and Boukouvala, 2020; Gonzalez et al., 2020; Wang et al., 2020), and to urban traffic issues (Chen et al., 2019; Chong and Osorio 2018; Osorio and Bierlaire 2013; Osorio and Chong 2015).

Another important factor is that theoretical studies and real cases are balanced in both of these problem classes, and in most of other classes found in the SLR, indicating the interest and development in the area, in academia, and in practice. However, it is worth mentioning that studies focused on theoretical cases for some problem classes, e.g., inventory controls, and we recommend exploring this problem more in real contexts, since inventory management is a strategic activity for most organizations.

The second aspect studied was the industrial sector of focus in the studies. 52.0% of the articles focused on water resource management, energy generation/distribution, urban traffic planning, healthcare, and the chemical/pharmaceutical industry. This shows that most MBSO efforts are focused on essential goods and services production, providing cost reductions and improvements that benefit the greatest number of people. About 30.2% of all studies did not specify the industrial sector. It is noteworthy that among the real cases, only 4.3% did not specify the industrial sector.

This is positive because more precise problem indications help future researchers and practitioners select the methodologies that best fit their industrial sector. Regarding the economic sector of the studies, primary, secondary, and tertiary sectors were equally represented, characterizing the breadth and versatility of MBSO techniques. Table 3 shows the variables and objectives of the studies.

Table 3 - Problem variables and objectives.

Variables	Real	Theoretical	Total	Cum. %
Nº of Variables (median)	10	5	6	-
Continuous	34 (26.7%)	37 (29.1%)	71 (55.9%)	55.9
Discreet	12 (09.4%)	23 (19.1%)	35 (27.6%)	83.5
Hybrid	07 (05.5%)	07 (05.5%)	14 (11.0%)	94.5
Binary	02 (01.6%)	01 (00.8%)	03 (02.4%)	96.9
Not Specified	00 (00.0%)	03 (02.4%)	03 (02.4%)	99.2
Categorical	01 (00.8%)	00 (00.0%)	01 (00.8%)	100.0
Optimization Objectives				
Single Objective	53 (35.6%)	54 (36.2%)	107 (71.8%)	71.8
Multi Objective	18 (12.1%)	22 (14.8%)	40 (25.8%)	98.7
Not Specified	01 (00.7%)	01 (00.7%)	02 (01.3%)	100.0

Regarding the problem sizes studied, real case applications had a median of 10 variables, compared to 5 variables for theoretical cases, as shown in Table 3. Furthermore, many authors have studied problems with more than 100 variables (e.g., (Chen et al., 2019; Chong and Osorio 2018; Jackson et al., 2019; Osorio 2019b; Osorio and Punzo 2019; Timani and Peralta 2017), extending up to 16,200 variables for a mesoscopic simulation model calibration problem for controlling traffic (Osorio 2019a). Regarding the type of variable, 83.5% of the problems had continuous or discrete variables. The first type had infinite solution spaces, and the second had delimited solution spaces, with an average of 203,963 possible scenarios, ranging from 36 to 1×10^{62} scenarios, with a 90% percentile at 2.5×10^{20} .

Although most studies addressed single-objective problems, 25.8% studied metamodeling for multi-objective optimization problems, mainly dealing with identifying Pareto Frontiers. We can observe the importance of using metamodeling, since recursive SO methods are practically unfeasible for dimensionality and solution space problems found in this SLR, in addition to proving to be good choices for multi objective optimizations, which are a class of NP-hard problems (Baquela and Olivera 2019).

We also analyzed metamodeling use versus model-based techniques. Although 61.3% of all studies did not specify the reason for choosing metamodeling over other methods, 33.9% of all articles used metamodeling to reduce project optimization times, while other methods could not converge with adequate computation time given restricted simulation budgets (e.g., Parnianifard et al., 2018; Song et al., 2005; Villarreal-Marroquín

et al., 2013). Furthermore, 3.0% used metamodeling to provide quick answers to operational problems requiring short-term decision-making time frames (e.g., Dunke and Nickel, 2020; Calahorrano et al., 2016). Metamodeling was also used to perform joint optimizations taking data from several models (e.g., Bartz-Beielstein et al., 2018), to quickly generate entire solution spaces (e.g., Stork et al., 2008), and better understand optimization processes (e.g., Huyet and Paris, 2001).

We can answer RQ1 (“What is the role of metamodeling in SO projects, and what are its main applications?”), according to the information in Tables 2 and 3, and in the discussion. Metamodeling is used to reduce computational times for problems where solution spaces and model complexity make more traditional recursive methods too expensive or unfeasible. Furthermore, recursive methods are unfeasible depending on the decision-making time frame, in light of the advent of industry 4.0, and the use of simulations as decision support tools in real or near real time, meaning that metamodels need to be used. Once trained and validated, they can generate good results in mere seconds.

Several authors have used this strategy for operational support systems (e.g., Dunke and Nickel, 2020; Calahorrano et al., 2016; Pedrielli et al., 2020; Shirazi et al., 2011; Steer et al., 2012). Regarding the main MBSO applications in real cases, process control problems, resource allocation and sizing, urban traffic controls, and product designs are particularly worth mentioning. Furthermore, there was fair distribution among different economic sectors, allowing us to infer that MBSO is a broad method that can generate good results when facing various problems, including NP-hard problems.

4.2 Methods used

This section presents the most widely-used methods, according to literature, for metamodeling algorithms, search methods (optimization), simulation types, sampling strategies, software programs, and programming languages. Tables 4 and 5 summarize the results, showing the six most frequent technologies/methods for each studied topic.

We noticed that the number of occurrences exceeds the total number of articles analyzed for certain topics due to the fact that some studies used more than one method, either by comparing or combining them. For example, Li et al. (2010), compared metamodels using RNA, RBF, SVM, Kriging, and MARS for a resource allocation problem, and Bartz-Beielstein et al. (2018), combined metamodels using linear regression, RF, and Kriging via an ensemble model, etc. (e.g., Lal and Datta, 2020; Parnianifard et al., 2020a; Ranjbar et al., 2020).

The first aspect in Table 4 was the type of simulation used by the researchers. 25% of the studies did not state the type of simulation used, and this only appeared as a recommendation for future studies so researchers can identify the type of simulation applicable to their specific problems in literature, and identify which tools can be used and associated with it. Of all the articles, 26.3% used SED. According to Law (2013), SED is defined as a system modeling and simulation technique that evolves over time, where events change state at discrete points in time.

Several authors have used SED for problems related to the secondary and tertiary economic sectors, like inventory controls (e.g., Dengiz et al., 2009; Quan et al., 2015), shop floor planning (e.g., Dunke and Nickel, 2020; Greinacher et al., 2020), allocating and sizing resources (Aydin et al., 2018; Coelho and Pinto 2018; Cuckler et al., 2017), and production scheduling (Chang, 2016; Hassannayebi et al., 2019). No SED

metamodeling application was conducted for the primary economic sector, perhaps because this type of simulation is commonly taught in specific engineering courses, limiting this tool's dissemination. We can therefore state that there is a possible literature gap that future researchers could explore, i.e., using metamodeling and SED to optimize processes in the primary sector.

Table 4 - Method used.

Type of Simulation	Real	Theoretical	Total	Cum. %
SED	14 (09.7%)	24 (16.6%)	38 (26.3%)	26.4
Not Specified	12 (08.3%)	24 (16.6%)	36 (25.0%)	51.4
Finite Element Method	13 (09.0%)	13 (09.0%)	26 (18.0%)	69.4
Simulating Traffic	16 (11.1%)	1 (00.7%)	17 (11.8%)	81.3
Monte Carlo Simulations	01 (00.7%)	5 (03.5%)	06 (04.2%)	85.4
Finite Volume Method	04 (02.8%)	00 (00.0%)	04 (02.8%)	88.2
Optimization Method				
Genetic Algorithm	12 (13.1%)	10 (10.9%)	22 (24.1%)	24.2
Trust Region	11 (12.0%)	01 (01.1%)	12 (13.1%)	37.4
NSGAI	08 (08.8%)	02 (02.2%)	10 (10.9%)	48.4
Particle swarm	06 (06.6%)	03 (03.3%)	09 (09.9%)	58.2
Simulated annealing	03 (03.3%)	02 (02.2%)	05 (05.5%)	63.7
Metamodeling Algorithm				
Kriging	22 (09.0%)	31 (12.7%)	53 (21.7%)	21.7
Polynomial Regression	20 (08.2%)	19 (07.8%)	39 (16.0%)	37.7
RNA	26 (10.7%)	13 (05.3%)	39 (16.0%)	53.7
Ensembles	06 (02.5%)	15 (05.1%)	21 (08.6%)	62.3
RBF	06 (02.5%)	09 (03.7%)	15 (06.1%)	68.4
SVM	02 (00.8%)	08 (03.3%)	10 (04.1%)	71.5
Sample Strategy				
Latin hypercube design	23 (15.2%)	38 (25.1%)	61 (40.3%)	40.4
Random	10 (06.6%)	10 (06.6%)	20 (13.2%)	53.6
Not Specified	10 (06.6%)	03 (02.0%)	13 (08.6%)	62.3
Factorial design	04 (02.6%)	06 (04.0%)	10 (06.6%)	68.9
Central composite design	03 (02.0%)	03 (02.0%)	06 (04.0%)	72.8
Orthogonal array	01 (00.7%)	04 (02.6%)	05 (03.3%)	76.2

By contrast, the second most widely-used simulation type was numerical simulation using the finite element method (18.0%), and most applications were focused on the primary economic sector (50.0%), for optimizing pumping rates for water extractions (e.g., Dhar and Datta, 2009; Hussain et al., 2015; Roy and Datta, 2018a, 2019b). Traffic simulations were conducted for 11.8% of all articles, corresponding to microscopic models (Guo and Zhang 2014; Osorio and Bierlaire, 2013; Osorio and Nanduri, 2015), mesoscopic models (Chen et al., 2019; Chen et al., 2016; Osorio, 2019a), and macroscopic models (Chen et al., 2014; Zhang et al., 2017). Metamodeling acts as a means for minimizing transit model error in this type of simulation by optimizing the model's input parameters (e.g., origin-destination matrix), minimizing the average travel time by optimizing traffic signal cycle times, defining public transportation and/or toll fees, etc.

Several other types of simulations were also identified, but were less-popular than those mentioned before. These include Monte Carlo simulations (e.g., Song et al., 2005; Wade, 2019), the finite volume method (e.g., Mooselu et al., 2019; Storti et al., 2019), the differential elements method (e.g., Alizadeh et al., 2017; Li et al., 2020), and agent-based simulations (e.g., Yousefi and Yousefi, 2019; Zeinali et al., 2015), etc.

The second item shown in Table 4 was the optimization method used in the studies. It is worth mentioning that metamodels are not optimization methods, but rather simplified models of original simulations. Researchers need to apply appropriate search methods to metamodels to perform optimizations. Many optimization algorithms were found in the studies. Genetic Algorithms were most-commonly used, especially for single-objective network planning problems (Shariatinasab et al., 2008; Vahidi et al., 2008; Ye and You, 2015), for resource allocation (Sousa Junior et al., 2020; Yousefi et al., 2018), process control (Ali et al., 2018; Luo et al., 2019; Ouyang et al., 2017), and product designs (Albanesi et al., 2018; Storti et al., 2019).

The second most frequently used search method was the trust region method, which is mostly used with quadratic polynomial metamodels for urban traffic control problems (Chen et al., 2019; Chong and Osorio, 2018; Osorio and Bierlaire, 2013; Osorio and Chong, 2015; Osorio and Nanduri, 2015), traffic model calibrations (Li et al., 2017; Osorio 2019b), emergency evacuations (Lv et al. 2015), and network planning (Ye and You 2015). The third most commonly applied search method was the Non-dominated Sorting Genetic Algorithm II (NSGA-II), an evolutionary multi-objective optimization algorithm, which was used mostly for optimizations, especially optimizations for process control problems. (Alizadeh et al., 2017; Dhar and Datta, 2009; Hussain et al., 2015; Ranjbar et al., 2020; Sreekanth and Datta 2011), and for product designs (Mooselu et al., 2019; Sharif and Hammad 2019)

Other commonly used methods in literature were particle swarm optimizations (e.g., Li et al., 2020; Yadav et al., 2016), and simulated annealing (e.g., Christelis and Mantoglou, 2016; Ciuffo and Punzo, 2014). By stratifying the main types of simulations, we see that Genetic Algorithms were used most in SED, the finite element method, and the finite volume method. The trust-region was most used for traffic simulations, and Monte Carlo simulations used mostly Gold Region, Ranking, and Selection methods.

With respect to RQ2 ("What are the main simulation and optimization techniques used in conjunction with metamodeling?"), we can conclude that SED is the most popular simulation technique considering the total number of publications, and the most popular technique for secondary economic sector applications, given that most industrial processes have independent and identically distributed variables, and can be modeled as discrete systems. The finite element method was most common for the primary economic sector, led by studies on process controls for water extraction activities. Traffic simulations, in particular microscopic simulations, for urban traffic planning and management, were most frequently used as simulation techniques in the tertiary economic sector. Furthermore, researchers used Genetic Algorithms as a search engine in single-objective optimization processes more frequently, and the NSGA II algorithm more frequently for multi-objective optimizations.

The third analyzed item was the type of metamodeling algorithm used. Table 4 shows that the most frequent algorithm was Kriging, appearing in 53 articles (21.7%). Kriging was initially proposed by Daniel G. Krige (1951), for geostatistical applications. It is a highly flexible interpolation method, given its ability to employ a wide variety of correlation functions (Parnianifard et al., 2018). Kriging has been applied to a wide

variety of problems because of this flexibility, e.g., for controlling inventories (Azizi, Seifi, and Moghadam 2019; Quan et al., 2013; Smew, Young, and Geraghty 2013), product designs (Bartz-Beielstein et al., 2018; Ivanova and Kuhnt, 2014; Yaohui 2017), process controls (Boukouvala and Ierapetritou, 2013; Candelieri et al., 2018; Roy and Datta 2018b), etc.

PR is another class of classic algorithms found in literature (16%). PR is an explicit mathematical expression, that is either linear or nonlinear, that aims to approximate relationships between independent and dependent variables (Parnianifard et al., 2019). PR is commonly found in literature as RSM (Aydin et al., 2018; Cuckler et al. 2017; Hassannayebi et al., 2019), as quadratic polynomials (Xiao et al., 2015; Zlatinov and Laskowski, 2015), and as first-orders (Hong et al., 2013; Villarreal-Marroquin et al., 2013). Many authors have also used polynomials to approximate true functions in local searches via the trust region method (e.g., Chen et al., 2019; Chen et al., 2014; Chong and Osorio, 2018; Osorio and Chong, 2015; Osorio and Nanduri, 2015).

ANNs were the third most frequently used algorithm. These algorithms are some of the most important artificial intelligence techniques, given their ability in solving extremely difficult problems that do not have any known analytical solutions. ANNs comprise both interconnected and adaptive elements, which respond to stimuli, similar to the human nervous system (Hachicha, 2011). Choosing the appropriate architecture is a crucial success factor for ANN applications, however, standard formula cannot be determined, and each case must be analyzed individually.

Hyperparameter selection techniques used in metamodels are important for these types of problems, and trial and error methods (Ataie-Ashtiani et al., 2014; Sreekanth and Datta, 2011), or more efficient optimization methods, like Genetic Algorithms for selecting hyper-parameters, can be used (Hullen et al., 2020; Kim et al., 2016).

Several authors (4th position) also used ensemble methods, or algorithms that combined several metamodels, be them similar or not, to formulate metamodels with greater learning capacities, like bagging (Chen et al., 2017; Yousefi et al., 2018), stacking (Bartz-Beielstein et al., 2018), AdaBoost (Yousefi et al., 2018; Yousefi and Yousefi, 2019), etc.

Algorithms using decision tree-based ensembles were the most frequent, and RF use was common (e.g., Bartz-Beielstein et al., 2018; Candelieri et al., 2018; Sousa Junior et al., 2019; Sousa Junior et al., 2020; Steer et al., 2012), along with GBT (e.g., (Sousa Junior et al., 2019; Sousa Junior et al., 2020). Many applications deal with using RBF algorithms (Sousa Junior et al., 2019; Sousa Junior et al., 2020). e.g., Ali et al., 2018; Christelis and Mantoglou, 2016; Horng and Lin, 2018; Krityakierne et al., 2016; Stork et al., 2008), and SVM (e.g., Hou et al., 2019; Lal and Datta, 2020; Li et al., 2019; Mirenderesgi and Mousavi, 2016).

Based on the above data, we can answer RQ3 (“What are the most commonly used metamodeling algorithms in MBSO?”). The main algorithms are Kriging, Polynomial Regression (especially RSM), ANNs, ensemble methods (in special RF and GBT), RBF, and SVM. The most common algorithms stratified by type of simulation in SED are Polynomial Regressions, Kriging, ANNs, and decision trees. For the finite element method, they were Kriging, ANNs, Polynomial Regressions, MARS. and RBF. 80% of all traffic simulations used either Polynomial Regressions or Kriging. Monte Carlo simulations and the finite volume method mostly used ANNs. Table 5 shows the main hyper-parameters of these algorithms and their frequencies .

The last item in Table 4 was the sampling strategy used by the researchers to build the metamodel, addressing RQ4 (“Which sampling strategies were used?”). One of the critical success factors in metamodel use is the training database. So researchers must choose the correct sampling method that can reduce the number of simulations needed for modeling, while also representing the search space well (Vu et al. 2017). 61 articles (40.3%) used the DOE technique, known as LHD, followed by random sampling (13.2%), Factorial Design (6.6%), CCD (4.0%), the Orthogonal Array (3.1%), etc., in addition to another 13 articles (8.6%), that either did not use, or did not specify, the sampling technique.

LHD was first proposed by McKay et al. (1979), and is a very popular stratified sampling method for metamodeling studies. Given n samples and s variables, LHD divides each variable’s region of into n disjointed intervals of equal probability. Then the sample space is selected by building a $n \times s$ matrix. The columns are randomly selected from a permutation of $[1, \dots, n]$, and each row corresponds to a hyper-rectangle cell. After building the matrix, a point from each cell is sampled (Luo et al., 2019).

In addition to good representative samples in the solution space, the sample size in LHD can be defined by the user. This can lead to success in building metamodels that, unlike techniques like the orthogonal array or CCD, have greater decision-making power, in function of trade-offs between experimental performance time and metamodel accuracy. Although the sample size can be defined by the user, the academic consensus is that it should be at least 10 times greater than the number of problem variables (Kim and Boukouvala, 2020; Parnianifard et al., 2018; Parnianifard and Azfanizam, 2020; Smew et al., 2013). We should note that LHD was the most used simulation type, except for in traffic simulations, which used random sampling more frequently.

Furthermore, 80 articles (61.5% of the studies using some sampling technique) used the fixed sampling strategy, i.e., only a set of samples size w is taken at the beginning of the project, and the metamodel is trained with this set alone. For the incremental strategy (38.5%), an initial data sample size ξ is taken, and the metamodel is trained and optimized with the precision and region direction results containing the optimum level, adding a new sample set size δ_i to the database. Then the metamodel is retrained with base size $\xi_{i+1} = \xi_i + \delta_i$, where $i \in \mathbb{Z}$ is the number of algorithm iterations.

This algorithm, which is responsible for recursive and adaptive learning, is referred to as Efficient Global Optimization (EGO) in most studies. The choice for the point(s) that should be added to each interaction is usually made using the Expected Improvement (EI) algorithm (Coelho and Pinto 2018; Pedrielli et al., 2020; Quan et al., 2013; Rojas-Gonzalez et al., 2020; Wang et al. 2018). Table 6 shows the main software programs used, and performance metrics found in literature.

Regarding the type of technology used in the studies, we observed from the first topic in Table 6 that most authors (19.7%), did not specify the modeling and simulation software they used, while 14.9% used programming languages like C++ (e.g., Song et al., 2011), Python (e.g., (Sousa Junior et al., 2019; Kim and Boukouvala, 2020), or Matlab (e.g., Parnianifard and Azfanizam, 2020; Rojas-Gonzalez et al., 2020). Regarding proprietary software used for simulations, we can cite the use of Arena®, one of the first software programs developed with 2D animation for modeling discrete systems. FEMWATER and MoldFlow were the most common software programs for the finite element method, along with Aimsun for traffic simulations.

Table 5 – The main metamodeling algorithm hyper-parameters.

Algorithm	Amplitude	Most Frequent
Kriging		
Correlation Function	{Gaussian, polynomial, exponential}	Gaussian
Lengthscale (θ)	$[10^{-10}; 10^{10}]$	10
Epsilon ¹	$[10^{-6}; 10^{-5}]$	-
Geometric Tolerance ¹	[0; 0.3]	-
Max N°. of vectors ¹	[38; 47]	-
RP		
Degree	{1°, 2°, 3°, 4°, 5°}	2°
RNA		
Transfer function	{linear, sigmoid, quadratic error, hyperbolic tangent, rectified linear unit, RBF}	Linear
Training function	{backpropagation, levenberg-marquardt, feedforward}	Backpropagation
Learning rate	[0.01; 0.4]	0.1
Epochs	[200; 1500]	1500
N° of layers	[1; 4]	1
N° of neurons	[1; 66]	4
Momentum	0.1	0.1
RF		
N° of trees	[10; 50]	10
Maximum Depth	2	2
N° of variables per node ³	$\text{int}(p/3) - 1, \text{int}(p/3), \text{int}(p/3) + 1$	-
GBT		
N° of trees	[1; 100]	100
Learning rate ²	[0.05; 0.30]	-
Min. Obs. per leaf ²	[0.12; 0.50]	-
RBF		
Function	{gaussian, splines, multiquadratics, cubic}	Gaussian
Scale (sigma)	[1; 20]	1
Regularization	$[10^{-10}; 10^{10}]$	-
SVM		
C ⁴	$[\bar{y} - 3\sigma ; \bar{y} + 3\sigma]$	20.73
Epsilon ⁴	$[0; 3\sigma\sqrt{\ln n/n}]$	0.0779
Kernel	{RBF, Gaussian}	Gaussian
Gama	[0.0224; 0.1294]	0.0673

¹Hyper-parameter cited but not quantified in the SLR, so the hyper-parameter as suggested by Rostami, Khaksar, and Manshad (2014) was used.

²Hyper-parameter cited but not quantified in the SLR, therefore, the hyper-parameter as suggested by Ganjisaffar, Caruana, and Lopes (2011) was used.

³Hyper-parameter cited but not quantified in SLR articles, however, this was soon adopted as suggested by Dai et al. (2018), where p is the number of decision variables in the problem.

⁴The C and epsilon values should be calculated from an algorithm first run using the default parameters, where \bar{y} and σ_y are the mean and standard deviations of the y values in the training data, n is the sample size, and σ is the variance of the difference between the estimated and observed values. For further details see Oliveira et al. (2010).

Table 6 - Software and metrics.

Simulation Software	Real	Theoretical	Total	Cum. %
Not Specified	05 (03.4%)	24 (16.3%)	29 (19.7%)	19.7
Programming Language	07 (04.7%)	15 (10.2%)	22 (15.0%)	34.7
Arena	04 (02.7%)	08 (05.4%)	12 (08.1%)	42.9
FEMWATER	04 (02.7%)	04 (02.7%)	08 (05.4%)	48.3
Aimsun	07 (04.7%)	00 (00.0%)	07 (04.7%)	53.1
MoldFlow	02 (01.3%)	02 (01.4%)	04 (02.7%)	55.8
Metamodeling Software				
Matlab	26 (18.5%)	26 (18.5%)	52 (37.1%)	37.1
Not Specified	24 (17.1%)	24 (17.1%)	48 (34.2%)	71.4
Python	03 (02.1%)	03 (02.1%)	06 (04.2%)	75.7
R	03 (02.1%)	03 (02.1%)	06 (04.2%)	80.0
GAMS	03 (02.1%)	03 (02.1%)	06 (04.2%)	84.3
Minitab	02 (01.4%)	02 (01.4%)	04 (02.8%)	87.1
Performance Metrics				
Mean square error (MSE)/Root MSE	32 (12.6%)	36 (14.1%)	68 (26.8%)	26.8
Coefficient of Correlation (R , R^2 , R_{adj})	24 (09.4%)	26 (10.2%)	50 (19.7%)	46.5
Not Specified	24 (09.4%)	13 (05.1%)	37 (14.6%)	61.0
Mean absolute percentage error (MAPE)	14 (05.5%)	12 (04.7%)	26 (10.2%)	71.3
Mean absolute error (MAE)	05 (01.9%)	07 (02.7%)	12 (04.7%)	76.0
Standardized residuals error (SRE)	02 (00.7%)	05 (02.0%)	07 (02.7%)	78.7
Validation Technique				
Not Specified	28 (22.6%)	20 (16.1%)	48 (38.7%)	38.7
Train/test split	17 (13.7%)	11 (08.9%)	28 (22.6%)	61.3
Leave one out cross-validation	09 (07.3%)	11 (08.9%)	20 (16.1%)	77.4
k-fold cross-validation	06 (04.8%)	11 (08.9%)	17 (13.7%)	91.1
Does not use validation	00 (00.0%)	04 (03.2%)	04 (03.2%)	94.4
New test data	02 (01.6%)	02 (01.6%)	04 (03.2%)	97.6

Regarding the software or programming language used in developing the metamodels and the optimization structure, Matlab was used by 37.1% of all researchers. However, Python and R, which are open-source languages, were widely used in the field of Data Science, and have gained traction in recent years. All 6 works (4.3%), using R were published from 2014 onwards, and studies using Python were published between 2019 and 2020. Therefore, we can conclude that Matlab is widely used and popular for MBSO projects. However, the recent and successful use of R and Python makes them good alternatives, mainly given the growth in communities focused on ML studies (Pedregosa et al., 2011), and newly created and improved algorithms, that may increase metamodel efficiency and fit.

The last item analyzed in this section was the performance metrics used to evaluate the metamodels, to answer RQ5 (“How are the results measured/validated?”). The three most widely used metrics were MSE (or root MSE), the coefficient of determination R , R^2 , R_{adj}^2 , MAPE, and MAE. More details on these metrics are given in Section 2.2.

The metamodel validation process is a fundamental step, since researchers can evaluate and compare the effectiveness of different metamodels based on simulated system representativeness. Of all the articles analyzed, 38.3% did not specify the type of

validation process used, and we recommend that future research pay more attention to this stage, since validation is an important part of the MBSO methodological process, and provides support and reliability to metamodels. 22.9% of all articles (e.g., Sousa Junior et al., 2019; Hussain et al., 2015; Luo et al., 2019; Raei et al., 2019), used the database division strategy in two parts, the first for training the metamodel (around 70%-80% of the data), and the second for testing/validating (20%-30% of data).

However, cross-validation techniques were used in most studies, and 20 (16.4%), of all authors used leave-one-out cross-validation (e.g., Azizi et al., 2019; Xiqun et al., 2019), and 15 (12.3%), used k-fold cross-validation, (e.g., Coelho and Pinto, 2018; Hou et al., 2019; Wade, 2019). For a more detailed explanation of cross-validation techniques see Bergmeir and Benitez (2012).

4.3 MBSO Frameworks

Given the relevance and complexity of this topic, and the wide variety of joint techniques and applications for metamodeling, some MBSO frameworks have emerged to help researchers and practitioners with their projects. Existing frameworks, as proposed in literature, must be analyzed to propose a MBSO in SED framework to obtain insights that result in better assertiveness when designing suggested steps. Thus, the same keywords and search criteria/article screening used in this SLR were again applied to journal articles alone.

Osorio and Bierlaire (2013), proposed a framework optimizing urban traffic simulations, integrating metamodeling with quadratic polynomials and the derivative-free trust region optimization algorithm. Pedrielli et al. (2020), proposed a framework for real-time route optimization for ship collision detection systems. The method used metamodeling and agent-based simulations to generate contingency plans for possible collisions, minimizing risks and negative effects from route changes.

Other frameworks have also addressed specific engineering problems, e.g., Dang (2014), who used metamodeling, DOE, and Computer Aided Engineering to optimize plastic injection processes. Ranjbar et al. (2020), developed a multi-objective optimization procedure with metamodeling and NSGA II to manage water intrusion into coastal aquifers.

Parnianifard et al. (2020b), proposed a robust optimization framework to optimize OF while minimizing response variance using metamodeling, Taguchi orthogonal arrays, Pareto frontiers, and Bootstrapping sensitivity analysis.

Barton (2009) proposed a general optimization framework using metamodeling in SED models. The SO process was bifurcated into two possible strategies, as either local or global metamodeling, as shown in Figure 4.8. The global metamodeling strategy, which is the focus of this paper, adjusts the metamodel to represent the entire search space, and this global approximation is used in iterative optimization processes. The local strategy alternates the metamodeling and optimization steps, since as optimization moves in the search space, new local regions of θ space are explored, and new metamodel approximations are adjusted.

The global metamodeling strategy proposed by Barton (2009), was designed with seven main steps. In the first step, the search space region must be identified by defining the problem constraints. The type of metamodel is chosen in the second step. The author suggests using RNA, RBF, splines, or Kriging to do this. The following three steps deal

with choosing the type of experiment - which can be performed via fractional or complete factorial arrangements, CCD, LHD, etc. - conducting the experiments in the simulation model, and adjusting the metamodel to the experimental arrangement. If the error metrics are unacceptable, one should go back to the third step and change the metamodel type or experimental design. The last two steps address the optimization phase, where the optimization algorithm must be applied, and the optimal simulation model response must be evaluated.

It is worth mentioning that there are several methodologies for metamodeling, however, there are two main identifiable aspects, i.e., fixed sampling and incremental sampling. In fixed sampling, a sample set with N experiments is generated at the beginning of the project and the metamodel is trained using this set. Figure 9 shows a fixed sampling metamodeling framework.

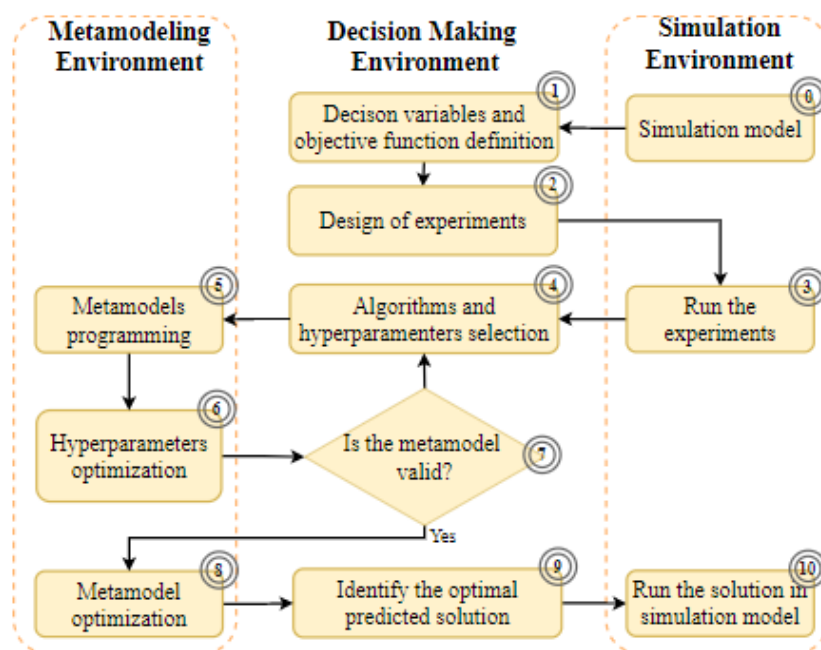


Figure 9: Fixed sampling metamodeling framework

The framework steps are as follows:

1. Determine the decision variables (x_1, \dots, x_n), and their variation amplitudes, along with the OF that will be optimized (y);
2. Define the experimental arrangement and select a sample of m scenarios that will comprise the metamodel training base. The LHD array is usually suggested for this step;
3. Run experiments on the simulation model, and store the simulated data;
4. Select the ML algorithms and the hyper-parameters that will be optimized, along with the variation range;
5. Build the metamodel structure in the software or programming language;
6. Perform hyperparameter optimization by validating and testing using the cross-validation technique. The GA is usually suggested as an optimization method;
7. Evaluate whether the error and adjustment metrics are acceptable according to the decision-maker criteria. If yes, then go to step 8. If no, then go back to step 4;

8. Run the metamodel optimization to find the decision variable values that optimize the predicted OF value defined in Step 1. The decision-maker can use any optimization techniques they see fit, e.g., AG, Ranking and Selection, NSGA II, etc.;
9. Identify the solution that optimizes the OF predicted by the metamodel;
10. Evaluate the solution from the simulation model.

By contrast, the incremental sampling approach is adaptive, and performs experiments preferably in more promising solution space regions. A small sample set is used to train the metamodels, and these are used to find new promising trial points in the simulation. These new points are added to the first set, and the loop continues until meeting the stop criterion. Figure 10 shows the incremental sampling framework.

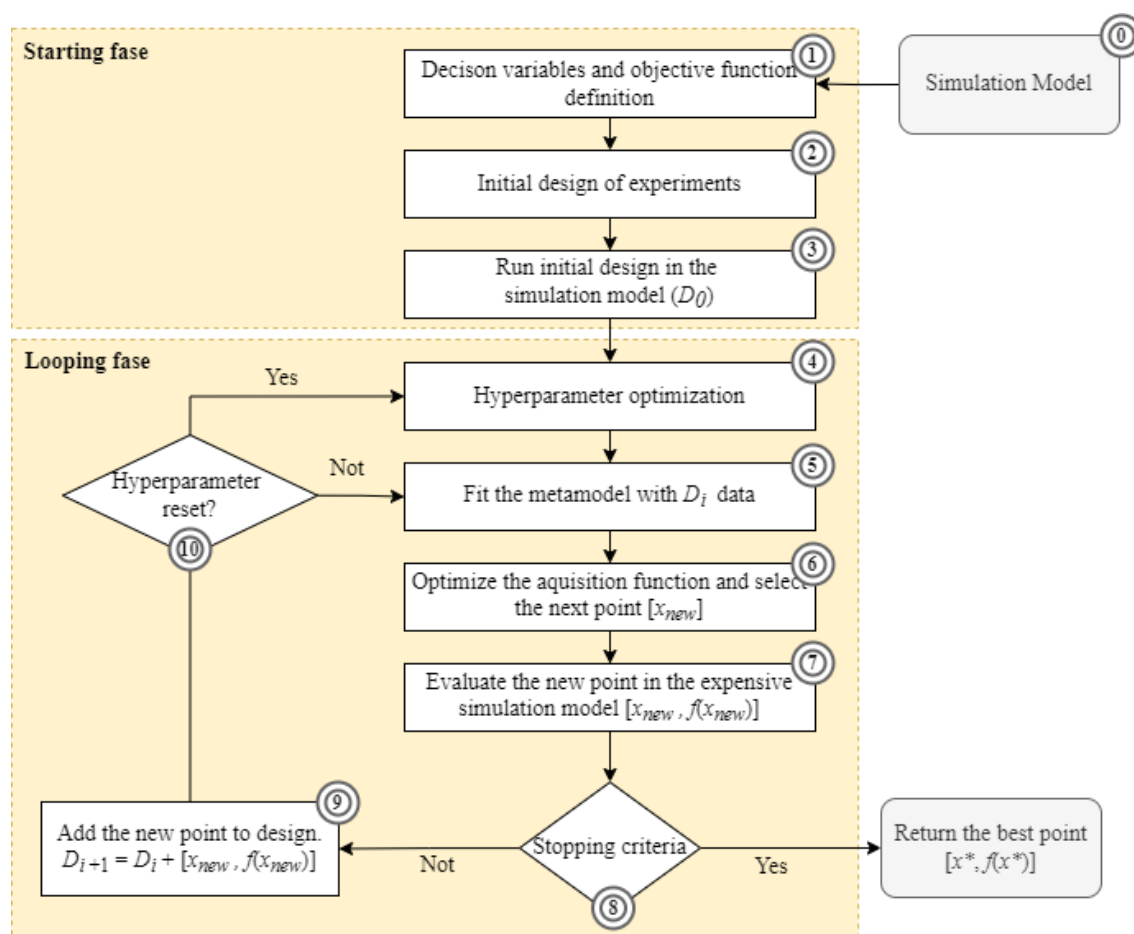


Figure 10: Incremental sampling metamodeling framework

The framework steps are as follows:

- 1) Determine the decision variables (x_1, \dots, x_n), and their variation amplitudes, along with the OF that will be optimized (y);
- 2) Define the experimental setup and select a sample of m scenarios that will comprise the initial metamodels training base. The LHD array is usually suggested for this step;
- 3) Run experiments on the simulation model, and store the simulated data;

- 4) Optimize the hyper-parameters via validations and testing using the cross-validation technique on the selected metamodel and the initial generated experiment base. The GA is usually suggested as an optimization method;
- 5) Train the metamodel using available experiments;
- 6) Use the metamodel to predict the point that optimizes the acquisition function to find the point (or points) that should be tested in the simulation. The acquisition function can be used from simpler functions, e.g., the highest/smallest $\hat{f}(x)$ value, or more complex functions, e.g., Expected Improvement or Probability of improvement, which seek the point that best meets trade-offs between regions with higher global optimum ($\hat{f}(x)$ high) probabilities, and regions with high uncertainty ($\hat{\sigma}(x)$ high);
- 7) Evaluate this new point (or set of points) in the original simulation model;
- 8) Check if the stop criteria have been met. If yes, return the optimal point. If no, proceed to step 9. The maximum number of iterations or simulations, the metamodel reaching certain precision, or the maximum number of iterations without improving the optimal response, can all be used as stop criteria;
- 9) Add the new test point to the current test base;
- 10) Evaluate whether the metamodel's hyper-parameters need to be adjusted. Comparing the metamodel error of the current iteration with the previous one can be used as a criterion. If further hyper-parameter adjustments are needed, proceed to step 4. If not, proceed to step 5.

These frameworks were developed based on analyzing good practices observed in the articles from the SLR.

4.4 Future directions for MBSO research

Based on the articles studied in this SLR, we identified that trends and gaps in literature can constitute opportunities involving adopted MBSO contexts and methods. Regarding the methods used, growing development in the area of computational science has greatly contributed to improved learning algorithms, which can consequently be studied by creating more accurate metamodels. One algorithm that has been gaining traction in the ML field, given its flexibility and learning ability, is Deep Learning (Higham and Higham, 2019; Xu et al., 2020). However, it has not been used to create SO metamodels, meaning that future studies could explore this research opportunity.

Beyond Deep Learning, there are still many opportunities for using ML algorithms already addressed in literature, but which are still being developed. Classical algorithms like Kriging and ANNs have shown considerable growth in recent years, especially when varying how these algorithms are used, e.g., (Rojas Gonzalez, Jalali, and Van Nieuwenhuyse 2020; Yousefi et al., 2018), or by using them with other techniques to build innovative solutions, e.g., (Ranjbar et al., 2020; Yaohui, 2017). Extreme learning machines e.g., (Li et al., 2020), SVM (Li et al., 2019), and ensemble methods, e.g., (Lal and Datta 2020), are other algorithms that have gained recent attention in MBSO, have high learning capabilities, and are under active development in the field of artificial intelligence.

Hyper-parameter optimization techniques increase metamodel effectiveness in correctly representing simulation models. Optimization methods like AG, particle swarms, Bayesian optimizations, and search grids for hyper-parameter optimization have

led to positive results in several studies, e.g. (Amiri et al., 2019; Hou et al., 2019; Hullen et al., 2020; Kim and Boukouvala, 2020; Li et al., 2019), but research is still incipient.

The global optimization method is another technique that has shown good results for building metamodels, which uses EGO and EI techniques to build adaptive and interactive metamodels. Kriging was mostly used as a metamodeling algorithm for applications involving this strategy, e.g., (Coelho and Pinto, 2018; Ivanova and Kuhnt, 2014; Wang et al., 2020). We recommend that future research explore this adaptive strategy for building metamodels based on other algorithms, like ANNs, SVM, Extreme Learning Machines, decision trees, etc.

Metamodeling has been used to address more complex and large-scale problems over the years. Feature selection techniques have been used more and more to deal with increasingly large databases in the ML and data mining communities. Feature selection can be interpreted as a set of techniques used to reduce problem size, removing irrelevant or redundant variables from the database (Bommert et al. 2020). However, these techniques were only addressed by Kim and Boukouvala (2020), and Sharif and Hammad (2019), and we suggest that future research address these techniques, like Principal Component Analysis, Wrappers, Filters, etc., to treat data in the MBSO process.

Regarding the context in which the studies were conducted, the highest growth trend was observed in the tertiary economic sector, driven mainly by metamodeling use in logistical problems, like cargo transportation and storage, and calibrating and optimizing of urban traffic models. However, it is worth mentioning that no MBSO application-optimized operations at distribution centers, a potential field for application, especially given e-commerce expansion and reverse logistics challenges. Furthermore, the COVID-19 pandemic intensified the pre-existing trends of using modeling and simulation techniques in healthcare (Currie et al., 2020), giving metamodeling an opportunity to reduce inherent healthcare model complexity and provide faster more consistent decision making.

Finally, with advancements in industry & service 4.0, many organizations from different sectors have demanded Digital Twin studies and applications. According to Mueller et al. (2017), Digital Twins are physical and virtual system junctions (simulators), connected in full or in partial autonomy, with real-time controls, which can work with human decision-making interfaces. This limits model-based SO technique use, since decision making is done in real time, or in short time intervals, leading to metamodeling replacing optimization simulation models via Digital Twin simulations. Dunke and Nickel (2020), Calahorrano et al. (2016), Pedrielli et al. (2020), Shirazi et al. (2011), and Steer et al. (2012), addressed operationally applying metamodels for decision making in near real time.

5 Final Considerations

This paper sought to (a) carry out an SLR on MBSO to (b) identify and explore the main aspects related to the topic, (c) analyze and summarize the results to answer the RQs, and (d) define the state-of-the arts, gaps, and trends that could be explored in future research.

SLR proved to be satisfactory in exploring and analyzing existing scientific literature. The guidelines proposed by Denyer and Tranfield (2009), and Denyer et al. (2008), greatly aided in better conducting this study, preventing possible distortions and author partialities in the results, thereby allowing us to meet our first objective. It is worth

mentioning that this study did not seek to exhaustively analyze all articles on the subject, but to systematically analyze a representative sample that would allow us to consistently answer the RQs, generating insights into MBSO practices, to assist future researchers and practitioners in understanding and identifying the frontier of existing knowledge in the area.

According to our second and third objectives, Sections 4.1 and 4.2 presented the main aspects related to the MBSO contexts of the studies, along with the simulation, optimization, sampling, and metamodeling methods adopted in the studies. Finally, Section 4.3 addressed the main frameworks and best MBSO project practices. Section 4.4 discussed the gaps and trends identified in this study, thereby responding to our final proposed objective.

Regarding the RQs, we can start by concluding that metamodeling has reduced computational time spent on SO problems, where recursive models would be expensive, or unfeasible. Metamodeling has been applied to many problems, with emphasis on MBSO optimizations for process control parameters, sizing, resource allocation, urban traffic controls, and product designs.

MBSO studies were initiated predominantly by industry. However, they have tended to expand into primary sectors, especially for water extraction and treatment, and into the tertiary economic sector, focusing on smart city concepts and better healthcare management.

Our second conclusion deals with the adopted methods. We were unable to reach a global consensus on which method or technique is most suitable for each problem, since this decision is often made based on available resources and the prior knowledge of researchers. The most frequently used simulation techniques were SED, the finite element method, and microscopic simulations of urban traffic. Regarding the most frequently used optimization methods, Genetic Algorithms were used for dealing with single-objective problems, the NSGA II algorithm was used for multi-objective problems in SED models, the finite elements, finite volumes, and trust-region methods were used for traffic simulations.

Lastly, the most used algorithms for building metamodels were Kriging, PR, RNAs, RBF, and SVM, along with ensemble methods (Bagging, AdaBoost, etc.), which combine one or more metamodeling algorithm to formulate a metamodel that can learn better, e.g., RF and GBT. Also, LHD and Random design were the main sampling strategies for search spaces.

Acknowledgements

The authors are grateful to CAPES, CNPq, FAPEMIG, and the UNIFEI Postgraduate Program in Production Engineering for supporting this research.

References

Aghaei Chadegani, A., Salehi, H., Md Yunus, M. M., Farhadi, H., Fooladi, M., Farhadi, M., & Ale Ebrahim, N. (2013). A comparison between two main academic literature collections: Web of science and scopus databases. *Asian Social Science*, 9(5), 18–26. <https://doi.org/10.5539/ass.v9n5p18>

Alabas, C., Altiparmak, F., & Dengiz, B. (2002). A comparison of the performance of artificial intelligence techniques for optimizing the number of kanbans. *Journal of the Operational Research Society*, 53(8), 907–914. <https://doi.org/10.1057/palgrave.jors.2601395>

Albanesi, A., Roman, N., Bre, F., & Fachinotti, V. (2018). A metamodel-based optimization approach to reduce the weight of composite laminated wind turbine blades. *Composite Structures*, 194(February), 345–356. <https://doi.org/10.1016/j.compstruct.2018.04.015>

Ali, W., Khan, M. S., Qyyum, M. A., & Lee, M. (2018). Surrogate-assisted modeling and optimization of a natural-gas liquefaction plant. *Computers and Chemical Engineering*, 118, 132–142. <https://doi.org/10.1016/j.compchemeng.2018.08.003>

Alizadeh, M. R., Nikoo, M. R., & Rakhshandehroo, G. R. (2017). Developing a Multi-Objective Conflict-Resolution Model for Optimal Groundwater Management Based on Fallback Bargaining Models and Social Choice Rules: a Case Study. *Water Resources Management*, 31(5), 1457–1472. <https://doi.org/10.1007/s11269-017-1588-7>

Amaral, J. V. S. do, de Carvalho Miranda, R., Montevechi, J. A. B., dos Santos, C. H., & Gabriel, G. T. (2022). Metamodeling-based simulation optimization in manufacturing problems: a comparative study. *The International Journal of Advanced Manufacturing Technology*, 20. <https://doi.org/10.1007/s00170-022-09072-9>

Amaran, S., Sahinidis, N. V., Sharda, B., & Bury, S. J. (2016). Simulation optimization: a review of algorithms and applications. *Annals of Operations Research*, 240(1), 351–380. <https://doi.org/10.1007/s10479-015-2019-x>

Amiri, F., Shirazi, B., & Tajdin, A. (2019). Multi-objective simulation optimization for uncertain resource assignment and job sequence in automated flexible job shop. *Applied Soft Computing Journal*, 75, 190–202. <https://doi.org/10.1016/j.asoc.2018.11.015>

Aouadni, S., Aouadni, I., & Rebaï, A. (2019). A systematic review on supplier selection and order allocation problems. *Journal of Industrial Engineering International*, 15(2009), 267–289. <https://doi.org/10.1007/s40092-019-00334-y>

Ataie-Ashtiani, B., Ketabchi, H., & Rajabi, M. M. (2014). Optimal Management of a Freshwater Lens in a Small Island Using Surrogate Models and Evolutionary Algorithms. *Journal of Hydrologic Engineering*, 19(2), 339–354. [https://doi.org/10.1061/\(asce\)he.1943-5584.0000809](https://doi.org/10.1061/(asce)he.1943-5584.0000809)

Aydin, C. S., Ozgurler, S., Durmusoglu, M. B., & Ozgurler, M. (2018). Response surface approach to robust design of assembly cells through simulation. *Assembly Automation*, 38(4), 450–464. <https://doi.org/10.1108/AA-08-2017-093>

Azizi, M. J., Seifi, F., & Moghadam, S. (2019). A robust simulation optimization algorithm using kriging and particle swarm optimization: Application to surgery room optimization. *Communications in Statistics: Simulation and Computation*, 0(0), 1–17. <https://doi.org/10.1080/03610918.2019.1593452>

Bălan, C. (2018). The disruptive impact of future advanced ICTs on maritime transport: a systematic review. *Supply Chain Management*, 25(2), 157–175. <https://doi.org/10.1108/SCM-03-2018-0133>

Bananmah, M., Reza Nikoo, M., Nematollahi, B., & Sadegh, M. (2020). Optimizing chute-flip bucket system based on meta-modelling approach. *Canadian Journal of Civil Engineering*, 47(5), 584–595. <https://doi.org/10.1139/cjce-2018-0534>

Baquela, E. G., & Olivera, A. C. (2019). A novel hybrid multi-objective metamodel-based evolutionary optimization algorithm. *Operations Research Perspectives*, 6(November 2018), 100098. <https://doi.org/10.1016/j.orp.2019.100098>

Barton, R. R. (2009a). Simulation Optimization Using Metamodel. *Winter Simulation Conference*, 2, 230–238.

Barton, R. R. (2009b). Simulation Optimization Using Metamodel. In *Winter Simulation Conference* (Issue 2).

Barton, R. R., & Meckesheimer, M. (2006). Chapter 18 Metamodel-Based Simulation Optimization. In *Handbooks in Operations Research and Management Science* (Vol. 13, Issue C, pp. 535–574). [https://doi.org/10.1016/S0927-0507\(06\)13018-2](https://doi.org/10.1016/S0927-0507(06)13018-2)

Bartz-Beielstein, T., Zaeferrer, M., & Pham, Q. C. (2018). Optimization via multimodel simulation: A new approach to optimization of cyclone separator geometries. *Structural and Multidisciplinary Optimization*, 58(3), 919–933. <https://doi.org/10.1007/s00158-018-1934-2>

Bashiri, M., & Geranmayeh, A. F. (2011). Tuning the parameters of an artificial neural network using central composite design and genetic algorithm. *Scientia Iranica*, 18(6), 1600–1608. <https://doi.org/10.1016/j.scient.2011.08.031>

Belgin, O. (2019). Data Envelopment Analysis Based Metamodeling for Multi Objective Simulation Optimization in a Manufacturing Line. *Sigma Journal of Engineering and Natural Sciences*, 37(4), 1435–1449.

Bergmeir, C., & Benítez, J. M. (2012). On the use of cross-validation for time series predictor evaluation. *Information Sciences*, 191, 192–213. <https://doi.org/10.1016/j.ins.2011.12.028>

Bin Makhshen, Y., Rafi-ul-Shan, P. M., Bashiri, M., Hasan, R., Amar, H., & Khan, M. N. (2020). Exploring the role of ambidexterity and cooperation in designing resilient fashion supply chains: a multi-evidence-based approach. *Journal of Enterprise Information Management*. <https://doi.org/10.1108/JEIM-08-2019-0213>

Bommert, A., Sun, X., Bischl, B., Rahnenführer, J., & Lang, M. (2020). Benchmark for filter methods for feature selection in high-dimensional classification data. *Computational Statistics and Data Analysis*, 143, 106839. <https://doi.org/10.1016/j.csda.2019.106839>

- Booth, A. (2006). Clear and present questions: formulating questions for evidence based practice. *Library Hi Tech*, 24(3), 355–368. <https://doi.org/10.1108/07378831111138242>
- Boukouvala, F., & Ierapetritou, M. G. (2013). Surrogate-based optimization of expensive flowsheet modeling for continuous pharmaceutical manufacturing. *Journal of Pharmaceutical Innovation*, 8(2), 131–145. <https://doi.org/10.1007/s12247-013-9154-1>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5–32. <https://doi.org/10.1201/9780429469275-8>
- Cai, X., Qiu, H., Gao, L., Jiang, C., & Shao, X. (2019). An efficient surrogate-assisted particle swarm optimization algorithm for high-dimensional expensive problems. *Knowledge-Based Systems*, 184, 104901. <https://doi.org/10.1016/j.knosys.2019.104901>
- Candelieri, A., Perego, R., & Archetti, F. (2018). Bayesian optimization of pump operations in water distribution systems. *Journal of Global Optimization*, 71(1), 213–235. <https://doi.org/10.1007/s10898-018-0641-2>
- Cao, Y., Xu, Y., & Du, J. (2020). Multi-variable estimation-based safe screening rule for small sphere and large margin support vector machine. *Knowledge-Based Systems*, 191(xxxx), 105223. <https://doi.org/10.1016/j.knosys.2019.105223>
- Carson, Y., & Maria, A. (1997). Simulation optimization: Methods and applications. *Winter Simulation Conference Proceedings*, 118–126.
- Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? -Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, 7(3), 1247–1250. <https://doi.org/10.5194/gmd-7-1247-2014>
- Chang, K. H., & Hsieh, L. Y. (2016). Determination of Wafer Start Mix in Semiconductor Manufacturing during New Technology Ramp-Up: Model, Solution Method, and an Empirical Study. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 46(2), 294–302. <https://doi.org/10.1109/TSMC.2015.2426174>
- Chen, M., Izady, A., & Abdalla, O. A. (2017). An efficient surrogate-based simulation-optimization method for calibrating a regional MODFLOW model. *Journal of Hydrology*, 544, 591–603. <https://doi.org/10.1016/j.jhydrol.2016.12.011>
- Chen, X., He, X., Xiong, C., Zhu, Z., & Zhang, L. (2019). A Bayesian stochastic kriging optimization model dealing with heteroscedastic simulation noise for freeway traffic management. *Transportation Science*, 53(2), 545–565. <https://doi.org/10.1287/trsc.2018.0819>
- Chen, X. M., Xiong, C., He, X., Zhu, Z., & Zhang, L. (2016). Time-of-day vehicle mileage fees for congestion mitigation and revenue generation: A simulation-based optimization method and its real-world application. *Transportation Research Part C: Emerging Technologies*, 63, 71–95. <https://doi.org/10.1016/j.trc.2015.12.001>
- Chen, X., Osorio, C., & Santos, B. F. (2019). Simulation-based travel time reliable signal control. *Transportation Science*, 53(2), 523–544. <https://doi.org/10.1287/trsc.2017.0812>

Chen, X., Yin, M., Song, M., Zhang, L., & Li, M. (2014). Social welfare maximization of multimodal transportation: Theory, metamodel, and application to Tianjin Ecocity, China. *Transportation Research Record*, 2451(2451), 36–49. <https://doi.org/10.3141/2451-05>

Cherkassky, V., & Ma, Y. (2004). Practical selection of SVM parameters and noise estimation for SVM regression. *Neural Networks*, 17(1), 113–126. [https://doi.org/10.1016/S0893-6080\(03\)00169-2](https://doi.org/10.1016/S0893-6080(03)00169-2)

Chong, L., & Osorio, C. (2018). A simulation-based optimization algorithm for dynamic large-scale urban transportation problems. *Transportation Science*, 52(3), 637–656. <https://doi.org/10.1287/trsc.2016.0717>

Christelis, V., & Mantoglou, A. (2016). Pumping Optimization of Coastal Aquifers Assisted by Adaptive Metamodelling Methods and Radial Basis Functions. *Water Resources Management*, 30(15), 5845–5859. <https://doi.org/10.1007/s11269-016-1337-3>

Ciuffo, B., & Punzo, V. (2014). “No free lunch” Theorems applied to the calibration of traffic simulation models. *IEEE Transactions on Intelligent Transportation Systems*, 15(2), 553–562. <https://doi.org/10.1109/TITS.2013.2287720>

Coelho, G. F., & Pinto, L. R. (2018). Kriging-based simulation optimization: An emergency medical system application. *Journal of the Operational Research Society*, 69(12), 2006–2020. <https://doi.org/10.1080/01605682.2017.1418149>

Costa, E., Soares, A. L., & de Sousa, J. P. (2020). Industrial business associations improving the internationalisation of SMEs with digital platforms: A design science research approach. *International Journal of Information Management*, 53(January), 102070. <https://doi.org/10.1016/j.ijinfomgt.2020.102070>

Crisci, C., Ghattas, B., & Perera, G. (2012). A review of supervised machine learning algorithms and their applications to ecological data. *Ecological Modelling*, 240, 113–122. <https://doi.org/10.1016/j.ecolmodel.2012.03.001>

Cuckler, R., Chang, K. H., & Hsieh, L. Y. (2017). Optimal Parallel Machine Allocation Problem in IC Packaging Using IC-PSO: An Empirical Study. *Asia-Pacific Journal of Operational Research*, 34(6), 1–20. <https://doi.org/10.1142/S0217595917500348>

Currie, C. S. M., Fowler, J. W., Kotiadis, K., Monks, T., Onggo, B. S., Robertson, D. A., & Tako, A. A. (2020). How simulation modelling can help reduce the impact of COVID-19. *Journal of Simulation*, 14(2), 83–97. <https://doi.org/10.1080/17477778.2020.1751570>

Dai, B., Gu, C., Zhao, E., & Qin, X. (2018). Statistical model optimized random forest regression model for concrete dam deformation monitoring. *Structural Control and Health Monitoring*, 25(6), 2170. <https://doi.org/10.1002/stc.2170>

Dang, X. P. (2014). General frameworks for optimization of plastic injection molding process parameters. *Simulation Modelling Practice and Theory*, 41, 15–27. <https://doi.org/10.1016/j.simpat.2013.11.003>

De Jong, K. A. (1975). *Analysis of the Behavior of a Class of Genetic Adaptive Systems*.

De La Fuente, R., Erazo, I., & Smith, R. L. (2019). Enabling intelligent processes in simulation utilizing the tensorflow deep learning resources. *Proceedings - Winter Simulation Conference, 2018-Decem*, 1108–1119. <https://doi.org/10.1109/WSC.2018.8632539>

De La Fuente, R., & Smith, R. (2017). Metamodeling a system dynamics model: a contemporary comparison of methods. *Winter Simulation Conference* *Winter Simulation Conference*, 53(9), 1689–1699. <https://doi.org/10.1017/CBO9781107415324.004>

Dellino, G., Kleijnen, J. P. C., & Meloni, C. (2015). Metamodel-based robust simulation-optimization: An overview. *Operations Research/ Computer Science Interfaces Series*, 59, 27–54. https://doi.org/10.1007/978-1-4899-7547-8_2

Dengiz, B., Alabas-Uslu, C., & Dengiz, O. (2009). Optimization of manufacturing systems using a neural network metamodel with a new training approach. *Journal of the Operational Research Society*, 60(9), 1191–1197. <https://doi.org/10.1057/palgrave.jors.2602620>

Dengiz, B., Bektas, T., & Ultanir, A. E. (2006). Simulation optimization based DSS application: A diamond tool production line in industry. *Simulation Modelling Practice and Theory*, 14(3), 296–312. <https://doi.org/10.1016/j.simpat.2005.07.001>

Dengiz, B., İç, Y. T., & Belgin, O. (2015). A meta-model based simulation optimization using hybrid simulation-analytical modeling to increase the productivity in automotive industry. *Mathematics and Computers in Simulation*, 120, 120–128. <https://doi.org/10.1016/j.matcom.2015.07.005>

Denyer, D., & Tranfield, D. (2009). Producing a systematic review. In *In D. A. Buchanan & A. Bryman (Eds.), The Sage handbook of organizational research methods* (pp. 671–689).

Denyer, D., Tranfield, D., & Van Aken, J. E. (2008). Developing design propositions through research synthesis. *Organization Studies*, 29(3), 393–413. <https://doi.org/10.1177/0170840607088020>

Dhar, A., & Datta, B. (2009). Saltwater Intrusion Management of Coastal Aquifers. I: Linked Simulation-Optimization. *Journal of Hydrologic Engineering*, 14(12), 1263–1272. [https://doi.org/10.1061/\(asce\)he.1943-5584.0000097](https://doi.org/10.1061/(asce)he.1943-5584.0000097)

Dunke, F., & Nickel, S. (2020). Neural networks for the metamodeling of simulation models with online decision making. *Simulation Modelling Practice and Theory*, 99(September 2019), 102016. <https://doi.org/10.1016/j.simpat.2019.102016>

Figueira, G., & Almada-Lobo, B. (2014). Hybrid simulation-optimization methods: A taxonomy and discussion. *Simulation Modelling Practice and Theory*, 46, 118–134. <https://doi.org/10.1016/j.simpat.2014.03.007>

Friedman, J. H. (2002). Stochastic gradient boosting. *Computational Statistics and Data Analysis*, 38(4), 367–378. [https://doi.org/10.1016/S0167-9473\(01\)00065-2](https://doi.org/10.1016/S0167-9473(01)00065-2)

Fu, M. C. (2002). Feature Article: Optimization for simulation: Theory vs. Practice. *INFORMS Journal on Computing*, 14(3), 192–215. <https://doi.org/10.1287/ijoc.14.3.192.113>

Fu, Y., & Sahin, K. H. (2004). Better Optimization of Nonlinear Uncertain Systems (BONUS) for vehicle structural design. *Annals of Operations Research*, 132(1–4), 69–84. <https://doi.org/10.1023/B:ANOR.0000045277.22566.5f>

Ganjisaffar, Y., Caruana, R., & Lopes, C. V. (2011). Bagging gradient-boosted trees for high precision, low variance ranking models. *SIGIR'11 - Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval*, c, 85–94. <https://doi.org/10.1145/2009916.2009932>

Ghorbani Mooselu, M., Nikoo, M. R., Bakhtiari Rayani, N., & Izady, A. (2019). Fuzzy Multi-Objective Simulation-Optimization of Stepped Spillways Considering Flood Uncertainty. *Water Resources Management*, 33(7), 2261–2275. <https://doi.org/10.1007/s11269-019-02263-2>

Greenwood, A. G., Rees, L. P., & Siochi, F. C. (1998). An investigation of the behavior of simulation response surfaces. *European Journal of Operational Research*, 110(2), 282–313. [https://doi.org/10.1016/S0377-2217\(97\)00255-5](https://doi.org/10.1016/S0377-2217(97)00255-5)

Greinacher, S., Overbeck, L., Kuhnle, A., Krahe, C., & Lanza, G. (2020). Multi-objective optimization of lean and resource efficient manufacturing systems. *Production Engineering*, 14(2), 165–176. <https://doi.org/10.1007/s11740-019-00945-9>

Guo, R., & Zhang, Y. (2014). Exploration of correlation between environmental factors and mobility at signalized intersections. *Transportation Research Part D: Transport and Environment*, 32, 24–34. <https://doi.org/10.1016/j.trd.2014.05.011>

Hachicha, W. (2011). A simulation metamodeling based neural networks for lot-sizing problem in MTO sector. *International Journal of Simulation Modelling*, 10(4), 191–203. [https://doi.org/10.2507/IJSIMM10\(4\)3.188](https://doi.org/10.2507/IJSIMM10(4)3.188)

Hannah, L. A., Powell, W. B., & Ducson, D. B. (2014). SEMICONVEX REGRESSION FOR METAMODELING-BASED OPTIMIZATION*. *SIAM J. OPTIM.*, 24(2), 573–597.

Hassannayebi, E., Boroun, M., Alaei Jordehi, S., & Kor, H. (2019). Train schedule optimization in a high-speed railway system using a hybrid simulation and meta-model approach. *Computers and Industrial Engineering*, 138(July), 106110. <https://doi.org/10.1016/j.cie.2019.106110>

Hassoun, M. H. (1995). Fundamentals of Artificial Neural Networks. In *MIT Press*. <https://doi.org/10.1177/0261018311403863>

Hastie, T., Tibshirani, R., & Friedman, J. (2008). The Elements of Statistical Learning. In *Springer* (Vol. 2).

Helleno, A. L., Pimentel, C. A., Ferro, R., Santos, P. F., Oliveira, M. C., & Simon, A. T. (2015). Integrating value stream mapping and discrete events simulation as decision making tools in operation management. *International Journal of Advanced Manufacturing Technology*, 80(5–8), 1059–1066. <https://doi.org/10.1007/s00170-015-7087-1>

Higham, C. F., & Higham, D. J. (2019). Deep learning: An introduction for applied mathematicians. *SIAM Review*, 61(4), 860–891. <https://doi.org/10.1137/18M1165748>

Hoad, K., Monks, T., & O'Brien, F. (2015). The use of search experimentation in discrete-event simulation practice. *Journal of the Operational Research Society*, 66(7), 1155–1168. <https://doi.org/10.1057/jors.2014.79>

Holland, J. H. (1975). *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence* (1^o). MIT Press.

Hong, J. H., Seo, K. M., & Kim, T. G. (2013). Simulation-based optimization for design parameter exploration in hybrid system: A defense system example. *Simulation*, 89(3), 362–380. <https://doi.org/10.1177/0037549712466707>

Hornig, S. C., & Lin, S. S. (2017). Ordinal optimization based metaheuristic algorithm for optimal inventory policy of assemble-to-order systems. *Applied Mathematical Modelling*, 42, 43–57. <https://doi.org/10.1016/j.apm.2016.10.002>

Hornig, S. C., & Lin, S. S. (2018). Embedding Advanced Harmony Search in Ordinal Optimization to Maximize Throughput Rate of Flow Line. *Arabian Journal for Science and Engineering*, 43(2), 1015–1031. <https://doi.org/10.1007/s13369-017-2864-9>

Hou, Z., Dai, Z., Lao, W., Wang, Y., & Lu, W. (2019). Application of Mixed-Integer Nonlinear Optimization Programming Based on Ensemble Surrogate Model for Dense Nonaqueous Phase Liquid Source Identification in Groundwater. *Environmental Engineering Science*, 36(6), 699–709. <https://doi.org/10.1089/ees.2018.0366>

Hu, B., Palta, M., & Shao, J. (2006). Properties of R² statistics for logistic regression. *Statistics in Medicine*, 25(8), 1383–1395. <https://doi.org/10.1002/sim.2300>

Hüllen, G., Zhai, J., Kim, S. H., Sinha, A., Realff, M. J., & Boukouvala, F. (2020). Managing uncertainty in data-driven simulation-based optimization. *Computers and Chemical Engineering*, 136, 106519. <https://doi.org/10.1016/j.compchemeng.2019.106519>

Hurrion, R. D. (1997). An example of simulation optimisation using a neural network metamodel: Finding the optimum number of kanbans in a manufacturing system. *Journal of the Operational Research Society*, 48(11), 1105–1112. <https://doi.org/10.1057/palgrave.jors.2600468>

- Hussain, M. S., Javadi, A. A., Ahangar-Asr, A., & Farmani, R. (2015). A surrogate model for simulation-optimization of aquifer systems subjected to seawater intrusion. *Journal of Hydrology*, 523, 542–554. <https://doi.org/10.1016/j.jhydrol.2015.01.079>
- Huyet, A. L., & Paris, J. L. (2001). Optimizing simulated manufacturing systems using machine learning coupled to evolutionary algorithms. *IEEE International Conference on Emerging Technologies and Factory Automation, ETFA*, 1, 17–21. <https://doi.org/10.1109/ETFA.2001.996349>
- Ilk, N., Shang, G., & Goes, P. (2020). Improving customer routing in contact centers: An automated triage design based on text analytics. *Journal of Operations Management*, 66(5), 553–577. <https://doi.org/10.1002/joom.1084>
- Ivanova, M., & Kuhnt, S. (2014). A parallel optimization algorithm based on FANOVA decomposition. *Quality and Reliability Engineering International*, 30(7), 961–974. <https://doi.org/10.1002/qre.1710>
- Jackson, I., Tolujevs, J., Lang, S., & Kegenbekov, Z. (2019). Metamodelling of Inventory-Control Simulations Based on a Multilayer Perceptron. *Transport and Telecommunication*, 20(3), 251–259. <https://doi.org/10.2478/ttj-2019-0021>
- Jagannatham, A. K., & Kumar, V. (2016). Introduction to game theory. In *Decision Sciences: Theory and Practice* (pp. 79–144). <https://doi.org/10.1201/9781315183176>
- Jalali, H., & Nieuwenhuysse, I. Van. (2015). Simulation optimization in inventory replenishment: A classification. *IIE Transactions*, 47(11), 1217–1235. <https://doi.org/10.1080/0740817X.2015.1019162>
- Jeng, J. T. (2006). Hybrid approach of selecting hyperparameters of support vector machine for regression. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 36(3), 699–709. <https://doi.org/10.1109/TSMCB.2005.861067>
- Kim, S. H., & Boukouvala, F. (2020). Machine learning-based surrogate modeling for data-driven optimization: a comparison of subset selection for regression techniques. *Optimization Letters*, 14(4), 989–1010. <https://doi.org/10.1007/s11590-019-01428-7>
- Kim, W., Jeon, Y., & Kim, Y. (2016). Simulation-based optimization of an integrated daylighting and HVAC system using the design of experiments method. *Applied Energy*, 162, 666–674. <https://doi.org/10.1016/j.apenergy.2015.10.153>
- Kleijnen, J. P. C. (2014). Simulation-optimization via Kriging and bootstrapping: a survey. *Journal of Simulation*, 8(4), 241–250. <https://doi.org/10.1057/jos.2014.4>
- Kleijnen, J. P. C. (2017a). Regression and Kriging metamodels with their experimental designs in simulation: A review. *European Journal of Operational Research*, 256(1), 1–16. <https://doi.org/10.1016/j.ejor.2016.06.041>
- Kleijnen, J. P. C. (2017b). Simulation Optimization Through Regression or Kriging Metamodels. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2969730>

Kloda, L. A., Boruff, J. T., & Cavalcante, A. S. (2020). A comparison of patient, intervention, comparison, outcome (PICO) to a new, alternative clinical question framework for search skills, search results, and self-efficacy: A randomized controlled trial. *Journal of the Medical Library Association*, 108(2), 185–194. <https://doi.org/10.5195/jmla.2020.739>

Kochan, C. G., & Nowicki, D. R. (2018). Supply chain resilience: a systematic literature review and typological framework. *International Journal of Physical Distribution and Logistics Management*, 48(8), 842–865. <https://doi.org/10.1108/IJPDLM-02-2017-0099>

Krige, D. G. (1951). A Statistical Approach to Some Basic Mine Valuation Problems on the Witwatersrand. *Journal of the Chemical Metallurgical & Society of South Mining Africa*, 52(6), 119–139.

Krityakierne, T., Akhtar, T., & Shoemaker, C. A. (2016). SOP: parallel surrogate global optimization with Pareto center selection for computationally expensive single objective problems. *Journal of Global Optimization*, 66(3), 417–437. <https://doi.org/10.1007/s10898-016-0407-7>

Lal, A., & Datta, B. (2020). Performance Evaluation of Homogeneous and Heterogeneous Ensemble Models for Groundwater Salinity Predictions: a Regional-Scale Comparison Study. *Water, Air, and Soil Pollution*, 231(6). <https://doi.org/10.1007/s11270-020-04693-w>

Law, A. M. (2013). Simulation Modeling and Analysis, FIFTH EDITION. In *Simulation Modeling and Analysis*.

Lessmann, S., Stahlbock, R., & Crone, S. F. (2006). Genetic algorithms for support vector machine model selection. *IEEE International Conference on Neural Networks - Conference Proceedings*, 3063–3069. <https://doi.org/10.1109/ijcnn.2006.247266>

Levy, Y., & Ellis, T. J. (2006). A systems approach to conduct an effective literature review in support of information systems research. *Informing Science*, 9(May 2014), 181–211. <https://doi.org/10.28945/479>

Li, G., Hwai-yong Tan, M., & Hui Ng, S. (2019). Metamodel-based optimization of stochastic computer models for engineering design under uncertain objective function. *IIE Transactions*, 51(5), 517–530. <https://doi.org/10.1080/24725854.2018.1504355>

Li, J., Lu, W., Wang, H., Fan, Y., & Chang, Z. (2020). Groundwater contamination source identification based on a hybrid particle swarm optimization-extreme learning machine. *Journal of Hydrology*, 584, 124657. <https://doi.org/10.1016/j.jhydrol.2020.124657>

Li, W., Wolinski, D., & Lin, M. C. (2017). City-scale traffic animation using statistical learning and metamodel-based optimization. *ACM Transactions on Graphics*, 36(6), 1–12. <https://doi.org/10.1145/3130800.3130847>

Li, Y. F., Ng, S. H., Xie, M., & Goh, T. N. (2010). A systematic comparison of metamodeling techniques for simulation optimization in Decision Support Systems.

Applied Soft Computing Journal, 10(4), 1257–1273.
<https://doi.org/10.1016/j.asoc.2009.11.034>

Li, Y. S., Chi, H., Shao, X. Y., Qi, M. L., & Xu, B. G. (2020). A novel random forest approach for imbalance problem in crime linkage. *Knowledge-Based Systems*, 195(xxxx), 105738. <https://doi.org/10.1016/j.knosys.2020.105738>

Lin, J. T., & Chen, C. M. (2015). Simulation optimization approach for hybrid flow shop scheduling problem in semiconductor back-end manufacturing. *Simulation Modelling Practice and Theory*, 51, 100–114. <https://doi.org/10.1016/j.simpat.2014.10.008>

Luo, J., Ji, Y., & Lu, W. (2019). Comparison of Surrogate Models Based on Different Sampling Methods for Groundwater Remediation. *Journal of Water Resources Planning and Management*, 145(5), 04019015. [https://doi.org/10.1061/\(asce\)wr.1943-5452.0001062](https://doi.org/10.1061/(asce)wr.1943-5452.0001062)

Lupera Calahorrano, G. J., Shokry, A., Campaña, G., & Espuña, A. (2016). Application of the Meta-Multiparametric methodology to the control of emissions in the industry under continuous and discrete uncertain parameters. *Chemical Engineering Research and Design*, 115, 365–373. <https://doi.org/10.1016/j.cherd.2016.09.006>

Lv, Y., Zhang, X., Kang, W., & Duan, Y. (2015). Managing Emergency Traffic Evacuation with a Partially Random Destination Allocation Strategy: A Computational-Experiment-Based Optimization Approach. *IEEE Transactions on Intelligent Transportation Systems*, 16(4), 2182–2191. <https://doi.org/10.1109/TITS.2015.2399852>

McKay, M. D., Beckman, R. J., & Conover, W. J. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21(2), 239–245. <https://doi.org/10.1080/00401706.1979.10489755>

Methley, A. M., Campbell, S., Chew-Graham, C., McNally, R., & Cheraghi-Sohi, S. (2014). PICO, PICOS and SPIDER: A comparison study of specificity and sensitivity in three search tools for qualitative systematic reviews. *BMC Health Services Research*, 14(1). <https://doi.org/10.1186/s12913-014-0579-0>

Miranda, R. D. C., Montevechi, J. A. B., Silva, A. F. Da, & Marins, F. A. S. (2014). A new approach to reducing search space and increasing efficiency in simulation optimization problems via the fuzzy-DEA-BCC. *Mathematical Problems in Engineering*, 2014. <https://doi.org/10.1155/2014/450367>

Miranda, R. de C., Montevechi, J. A. B., da Silva, A. F., & Marins, F. A. S. (2017). Increasing the efficiency in integer simulation optimization: Reducing the search space through data envelopment analysis and orthogonal arrays. *European Journal of Operational Research*, 262(2), 673–681. <https://doi.org/10.1016/j.ejor.2017.04.016>

Mirfenderesgi, G., & Mousavi, S. J. (2016). Adaptive meta-modeling-based simulation optimization in basin-scale optimum water allocation: A comparative analysis of meta-models. *Journal of Hydroinformatics*, 18(3), 446–465. <https://doi.org/10.2166/hydro.2015.157>

- Moghaddam, S., & Mahlooji, H. (2016). Robust simulation optimization using ϕ -divergence. *International Journal of Industrial Engineering Computations*, 7(4), 517–534. <https://doi.org/10.5267/j.ijiec.2016.5.003>
- Moghaddam, S., & Mahlooji, H. (2017). A new metamodel-based method for solving semi-expensive simulation optimization problems. *Communications in Statistics: Simulation and Computation*, 46(6), 4795–4811. <https://doi.org/10.1080/03610918.2015.1134567>
- Mohammad Nezhad, A., & Mahlooji, H. (2014). An artificial neural network meta-model for constrained simulation optimization. *Journal of the Operational Research Society*, 65(8), 1232–1244. <https://doi.org/10.1057/jors.2013.73>
- Mongeon, P., & Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: a comparative analysis. *Scientometrics*, 106(1), 213–228. <https://doi.org/10.1007/s11192-015-1765-5>
- Monostori, L., & Viharos, Z. J. (2001). Hybrid, al- and simulation-supported optimisation of process chains and production plants. *CIRP Annals - Manufacturing Technology*, 50(1), 353–356. [https://doi.org/10.1016/S0007-8506\(07\)62138-6](https://doi.org/10.1016/S0007-8506(07)62138-6)
- Mueller, E., Chen, X. L., & Riedel, R. (2017). Challenges and Requirements for the Application of Industry 4.0: A Special Insight with the Usage of Cyber-Physical System. *Chinese Journal of Mechanical Engineering (English Edition)*, 30(5), 1050–1057. <https://doi.org/10.1007/s10033-017-0164-7>
- Myers, R. H., Montgomery, D. C., & Anderson-Cook, C. M. (2016). *Response surface methodology: process and product optimization using designed experiments* (4th). John Wiley & Sons.
- Nguyen-Tuong, D., Seeger, M., & Peters, J. (2009). Model learning with local Gaussian process regression. *Advanced Robotics*, 23(15), 2015–2034. <https://doi.org/10.1163/016918609X12529286896877>
- Oliveira, A. L. I., Braga, P. L., Lima, R. M. F., & Cornélio, M. L. (2010). GA-based method for feature selection and parameters optimization for machine learning regression applied to software effort estimation. *Information and Software Technology*, 52(11), 1155–1166. <https://doi.org/10.1016/j.infsof.2010.05.009>
- Oliveira, J. B., Lima, R. S., & Montevechi, J. A. B. (2016). Perspectives and relationships in Supply Chain Simulation: A systematic literature review. *Simulation Modelling Practice and Theory*, 62, 166–191. <https://doi.org/10.1016/j.simpat.2016.02.001>
- Osorio, C. (2019a). Dynamic origin-destination matrix calibration for large-scale network simulators. *Transportation Research Part C: Emerging Technologies*, 98(September 2018), 186–206. <https://doi.org/10.1016/j.trc.2018.09.023>
- Osorio, C. (2019b). High-dimensional offline origin-destination (OD) demand calibration for stochastic traffic simulators of large-scale road networks. *Transportation Research Part B: Methodological*, 124, 18–43. <https://doi.org/10.1016/j.trb.2019.01.005>

Osorio, C., & Bierlaire, M. (2013). A simulation-based optimization framework for urban transportation problems. *Operations Research*, 61(6), 1333–1345. <https://doi.org/10.1287/opre.2013.1226>

Osorio, C., & Chong, L. (2015). A computationally efficient simulation-based optimization algorithm for large-scale urban transportation problems. *Transportation Science*, 49(3), 623–636. <https://doi.org/10.1287/trsc.2014.0550>

Osorio, C., & Nanduri, K. (2015). Urban transportation emissions mitigation: Coupling high-resolution vehicular emissions and traffic models for traffic signal optimization. *Transportation Research Part B: Methodological*, 81, 520–538. <https://doi.org/10.1016/j.trb.2014.12.007>

Osorio, C., & Punzo, V. (2019). Efficient calibration of microscopic car-following models for large-scale stochastic network simulators. *Transportation Research Part B: Methodological*, 119, 156–173. <https://doi.org/10.1016/j.trb.2018.09.005>

Østergård, T., Jensen, R. L., & Maagaard, S. E. (2018). A comparison of six metamodeling techniques applied to building performance simulations. *Applied Energy*, 211(October 2017), 89–103. <https://doi.org/10.1016/j.apenergy.2017.10.102>

Ouyang, Q., Lu, W., Lin, J., Deng, W., & Cheng, W. (2017). Conservative strategy-based ensemble surrogate model for optimal groundwater remediation design at DNAPLs-contaminated sites. *Journal of Contaminant Hydrology*, 203(May), 1–8. <https://doi.org/10.1016/j.jconhyd.2017.05.007>

Parnianifard, A., & Azfanizam, A. S. (2020). Metamodel-based robust simulation-optimization assisted optimal design of multiloop integer and fractional-order PID controller. *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields*, 33(1), 1–20. <https://doi.org/10.1002/jnm.2679>

Parnianifard, A., Azfanizam, A. S., Ariffin, M. K. A., & Ismail, M. I. S. (2020a). Comparative study of metamodeling and sampling design for expensive and semi-expensive simulation models under uncertainty. *Simulation*, 96(1), 89–110. <https://doi.org/10.1177/0037549719846988>

Parnianifard, A., Azfanizam, A. S., Ariffin, M. K. A., & Ismail, M. I. S. (2020b). Crossing weighted uncertainty scenarios assisted distribution-free metamodel-based robust simulation optimization. *Engineering with Computers*, 36(1), 139–150. <https://doi.org/10.1007/s00366-018-00690-0>

Parnianifard, A., Azfanizam, A. S., Ariffin, M. K. A., Ismail, M. I. S., & Ale Ebrahim, N. (2019). Recent developments in metamodel based robust black-box simulation optimization: An overview. *Decision Science Letters*, 8(1), 17–44. <https://doi.org/10.5267/j.dsl.2018.5.004>

Parnianifard, A., Azfanizam, A. S., Ariffin, M. K. A., Ismail, M. I. S., Maghami, M. R., & Gomes, C. (2018). Kriging and Latin hypercube sampling assisted simulation optimization in optimal design of PID controller for speed control of DC motor. *Journal*

of Computational and Theoretical Nanoscience, 15(5), 1471–1479.
<https://doi.org/10.1166/jctn.2018.7379>

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., & Cournapeau, D. (2011). Scikit-learn: Machine Learning in Python. *Journal Of Machine Learning Research*, 12(1), 2825–2830. <https://doi.org/10.1145/2786984.2786995>

Pedrielli, G., Xing, Y., Peh, J. H., Koh, K. W., & Ng, S. H. (2020). A real time simulation optimization framework for vessel collision avoidance and the case of singapore strait. *IEEE Transactions on Intelligent Transportation Systems*, 21(3), 1204–1215. <https://doi.org/10.1109/TITS.2019.2903824>

Quan, N., Yin, J., Ng, S. H., & Lee, L. H. (2013). Simulation optimization via kriging: A sequential search using expected improvement with computing budget constraints. *IIE Transactions (Institute of Industrial Engineers)*, 45(7), 763–780. <https://doi.org/10.1080/0740817X.2012.706377>

Raei, E., Reza Alizadeh, M., Reza Nikoo, M., & Adamowski, J. (2019). Multi-objective decision-making for green infrastructure planning (LID-BMPs) in urban storm water management under uncertainty. *Journal of Hydrology*, 579(January), 124091. <https://doi.org/10.1016/j.jhydrol.2019.124091>

Ranjbar, A., Mahjouri, N., & Cherubini, C. (2020). Development of an efficient conjunctive meta-model-based decision-making framework for saltwater intrusion management in coastal aquifers. *Journal of Hydro-Environment Research*, 29(November 2019), 45–58. <https://doi.org/10.1016/j.jher.2019.11.005>

Rojas Gonzalez, S., Jalali, H., & Van Nieuwenhuysse, I. (2020). A multiobjective stochastic simulation optimization algorithm. *European Journal of Operational Research*, 284(1), 212–226. <https://doi.org/10.1016/j.ejor.2019.12.014>

Rostami, H., & Khaksar Manshad, A. (2014). Application of evolutionary Gaussian processes regression by particle swarm optimization for prediction of dew point pressure in gas condensate reservoirs. *Neural Computing and Applications*, 24(3–4), 705–713. <https://doi.org/10.1007/s00521-012-1275-9>

Roy, D. K., & Datta, B. (2018a). Influence of sea level rise on multiobjective management of saltwater intrusion in coastal aquifers. *Journal of Hydrologic Engineering*, 23(8), 1–17. [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0001680](https://doi.org/10.1061/(ASCE)HE.1943-5584.0001680)

Roy, D. K., & Datta, B. (2018b). Trained meta-models and evolutionary algorithm based multi-objective management of coastal aquifers under parameter uncertainty. *Journal of Hydroinformatics*, 20(6), 1247–1267. <https://doi.org/10.2166/hydro.2018.087>

Roy, D. K., & Datta, B. (2019a). Adaptive Management of Coastal Aquifers Using Entropy-Set Pair Analysis–Based Three-Dimensional Sequential Monitoring Network Design. *Journal of Hydrologic Engineering*, 24(3), 04018072. [https://doi.org/10.1061/\(asce\)he.1943-5584.0001765](https://doi.org/10.1061/(asce)he.1943-5584.0001765)

Roy, D. K., & Datta, B. (2019b). An Ensemble Meta-Modelling Approach Using the Dempster-Shafer Theory of Evidence for Developing Saltwater Intrusion Management Strategies in Coastal Aquifers. *Water Resources Management*, 33(2), 775–795. <https://doi.org/10.1007/s11269-018-2142-y>

Salam, M. A., & Khan, S. A. (2016). Simulation based decision support system for optimization: A case of Thai logistics service provider. *Industrial Management and Data Systems*, 116(2), 236–254. <https://doi.org/10.1108/IMDS-05-2015-0192>

Schapire, R. E. (1990). The Strength of Weak Learnability. *Machine Learning*, 5(2), 197–227. <https://doi.org/10.1023/A:1022648800760>

Schruben, L. (1986). Simulation optimization using frequenct domain methods. *Winter Simulation Conference*, 366–369. <https://doi.org/10.1017/CBO9781107415324.004>

Shariatinasab, R., Vahidi, B., Hosseinian, S. H., & Ametani, A. (2008). Optimization of surge arrester's location on EHV and UHV power networks using simulation optimization method. *IEEE Transactions on Power and Energy*, 128(12), 1465–1472. <https://doi.org/10.1541/ieejpes.128.1465>

Sharif, S. A., & Hammad, A. (2019). Developing surrogate ANN for selecting near-optimal building energy renovation methods considering energy consumption, LCC and LCA. *Journal of Building Engineering*, 25(April), 100790. <https://doi.org/10.1016/j.jobe.2019.100790>

Shirazi, B., Mahdavi, I., & Mahdavi-Amiri, N. (2011). ICoSim-FMS: An intelligent co-simulator for the adaptive control of complex flexible manufacturing systems. *Simulation Modelling Practice and Theory*, 19(7), 1668–1688. <https://doi.org/10.1016/j.simpat.2011.04.003>

Smew, W., Young, P., & Geraghty, J. (2013). Supply chain analysis using simulation, gaussian process modelling and optimisation. *International Journal of Simulation Modelling*, 12(3), 178–189. [https://doi.org/10.2507/IJSIMM12\(3\)4.239](https://doi.org/10.2507/IJSIMM12(3)4.239)

Song, C., Guan, X., Zhao, Q., & Ho, Y. C. (2005). Machine learning approach for determining feasible plans of a remanufacturing system. *IEEE Transactions on Automation Science and Engineering*, 2(3), 262–275. <https://doi.org/10.1109/TASE.2005.849090>

Song, C., Guan, X., Zhao, Q., & Jia, Q. S. (2011). Remanufacturing planning based on constrained ordinal optimization. *Frontiers of Electrical and Electronic Engineering in China*, 6(3), 443–452. <https://doi.org/10.1007/s11460-011-0162-y>

Sousa Junior, W. T., Montevechi, J. A. B., Miranda, R. de C., & Campos, A. T. (2019). Discrete simulation-based optimization methods for industrial engineering problems: A systematic literature review. *Computers and Industrial Engineering*, 128(December 2018), 526–540. <https://doi.org/10.1016/j.cie.2018.12.073>

Sousa Junior, W. T., Montevechi, J. A. B., Miranda, R. de C., Oliveira, M. L. M. de, & Campos, A. T. (2020). Shop floor simulation optimization using machine learning to

improve parallel metaheuristics. *Expert Systems with Applications*.
<https://doi.org/10.1016/j.eswa.2020.113272>

Sousa Junior, W. T., Montevechi, J. A. B., Miranda, R. de C., Rocha, F., & Vilela, F. F. (2019). Economic lot-size using machine learning, parallelism, metaheuristic and simulation. *International Journal of Simulation Modelling*, 18(2), 205–216.
[https://doi.org/10.2507/IJSIMM18\(2\)461](https://doi.org/10.2507/IJSIMM18(2)461)

Sreekanth, J., & Datta, B. (2011). Optimal combined operation of production and barrier wells for the control of saltwater intrusion in coastal groundwater well fields. *Desalination and Water Treatment*, 32(1–3), 72–78.
<https://doi.org/10.5004/dwt.2011.2680>

Steer, K. C. B., Wirth, A., & Halgamuge, S. K. (2012). Decision tree ensembles for online operation of large smart grids. *Energy Conversion and Management*, 59, 9–18.
<https://doi.org/10.1016/j.enconman.2012.01.010>

Stork, A., Thole, C. A., Klimenko, S., Nikitin, I., Nikitina, L., & Astakhov, Y. (2008). Towards interactive simulation in automotive design. *Visual Computer*, 24(11), 947–953.
<https://doi.org/10.1007/s00371-008-0274-4>

Storti, B. A., Dorella, J. J., Roman, N. D., Peralta, I., & Albanesi, A. E. (2019). Improving the efficiency of a Savonius wind turbine by designing a set of deflector plates with a metamodel-based optimization approach. *Energy*, 186, 115814.
<https://doi.org/10.1016/j.energy.2019.07.144>

Tanila, T., Tenhunen, H., & Hirvonen, P. (2020). Value mechanisms in the implementation of intelligent patient flow management system - A multiple case study. *Studies in Health Technology and Informatics*, 270, 708–712.
<https://doi.org/10.3233/SHTI200252>

Timani, B., & Peralta, R. (2017). Aggregated surrogate simulator for groundwater-surface water management via simulation-optimization modeling: Theory, development and tests. *Environmental Modelling and Software*, 96, 210–231.
<https://doi.org/10.1016/j.envsoft.2017.06.014>

Vahidi, B., Tavakoli, M. R. B., & Hosseinian, S. H. (2008). Determining arresters best positions in power system for lightning shielding failure protection using simulation optimization approach. *EUROPEAN TRANSACTIONS ON ELECTRICAL POWER*, 20(3), 255–276. <https://doi.org/10.1002/etep>

Vapnik, V. N. (1999). *The Nature of Statistical Learning Theory*.

Villarreal-Marroquín, M. G., Mulyana, R., Castro, J. M., & Cabrera-Ríos, M. (2011). Selecting process parameter in injection molding via simulation optimization. *Journal of Polymer Engineering*, 31(5), 387–395. <https://doi.org/10.1515/POLYENG.2011.086>

Villarreal-Marroquín, M. G., Svenson, J. D., Sun, F., Santner, T. J., Dean, A., & Castro, J. M. (2013). A comparison of two metamodel-based methodologies for multiple criteria

simulation optimization using an injection molding case study. *Journal of Polymer Engineering*, 33(3), 193–209. <https://doi.org/10.1515/polyeng-2013-0022>

Vu, K. K., D'Ambrosio, C., Hamadi, Y., & Liberti, L. (2017). Surrogate-based methods for black-box optimization. *International Transactions in Operational Research*, 24(3), 393–424. <https://doi.org/10.1111/itor.12292>

Wade, B. M. (2019). Creating surrogate models for an air and missile defense simulation using design of experiments and neural networks. *Journal of Defense Modeling and Simulation*. <https://doi.org/10.1177/1548512919877987>

Wan, H. P., & Ren, W. X. (2015). Parameter selection in finite-element-model updating by global sensitivity analysis using gaussian process metamodel. *Journal of Structural Engineering*, 141(6), 1–11. [https://doi.org/10.1061/\(ASCE\)ST.1943-541X.0001108](https://doi.org/10.1061/(ASCE)ST.1943-541X.0001108)

Wang, H., Chen, L., & Li, E. (2018). Time dependent sheet metal forming optimization by using Gaussian process assisted firefly algorithm. *International Journal of Material Forming*, 11(2), 279–295. <https://doi.org/10.1007/s12289-017-1352-9>

Wang, H., Yuan, J., & Ng, S. H. (2020). Gaussian process based optimization algorithms with input uncertainty. *IISE Transactions*, 52(4), 377–393. <https://doi.org/10.1080/24725854.2019.1639859>

Wildridge, V., & Bell, L. (2002). How CLIP became ECLIPSE: a mnemonic to assist in searching for health policy/ management information. *Health Information and Libraries Journal*, 19(1), 113–115. <https://doi.org/10.1002/ajpa.21406>

Woldemariam, E. T., & Lemu, H. G. (2019). A machine learning based framework for model approximation followed by design optimization for expensive numerical simulation-based optimization problems. *Proceedings of the International Offshore and Polar Engineering Conference*, 1, 943–950.

Xia, Y., Liu, C., Li, Y. Y., & Liu, N. (2017). A boosted decision tree approach using Bayesian hyper-parameter optimization for credit scoring. *Expert Systems with Applications*, 78, 225–241. <https://doi.org/10.1016/j.eswa.2017.02.017>

Xiao, H., Lee, L. H., & Chen, C. H. (2015). Optimal budget allocation rule for simulation optimization using quadratic regression in partitioned domains. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 45(7), 1047–1062. <https://doi.org/10.1109/TSMC.2014.2383997>

Xiao, Y., & Watson, M. (2019). Guidance on Conducting a Systematic Literature Review. *Journal of Planning Education and Research*, 39(1), 93–112. <https://doi.org/10.1177/0739456X17723971>

Xu, J., Huang, E., Hsieh, L., Lee, L. H., Jia, Q. S., & Chen, C. H. (2016). Simulation optimization in the era of Industrial 4.0 and the Industrial Internet. *Journal of Simulation*, 10(4), 310–320. <https://doi.org/10.1057/s41273-016-0037-6>

Xu, L., Zhao, J., Li, C., Li, C., Wang, X., & Xie, Z. (2020). Simulation and prediction of hydrological processes based on firefly algorithm with deep learning and support vector for regression. *International Journal of Parallel, Emergent and Distributed Systems*, 35(3), 288–296. <https://doi.org/10.1080/17445760.2019.1593409>

Yadav, B., Ch, S., Mathur, S., & Adamowski, J. (2016). Estimation of in-situ bioremediation system cost using a hybrid Extreme Learning Machine (ELM)-particle swarm optimization approach. *Journal of Hydrology*, 543, 373–385. <https://doi.org/10.1016/j.jhydrol.2016.10.013>

Yang, L., & Shami, A. (2020). On hyperparameter optimization of machine learning algorithms: Theory and practice. *Neurocomputing*, 415, 295–316. <https://doi.org/10.1016/j.neucom.2020.07.061>

Yaohui, L. (2017). A Kriging-based global optimization method using multi-points infill search criterion. *Journal of Algorithms and Computational Technology*, 11(4), 366–377. <https://doi.org/10.1177/1748301817725307>

Ye, W., & You, F. (2015). A simulation-based optimization method for solving the integrated supply chain network design and inventory control problem under uncertainty. *Chemical Engineering Transactions*, 45, 499–504. <https://doi.org/10.3303/CET1545084>

Yin, H., Fang, H., Wang, Q., & Wen, G. (2016). Design optimization of a MASH TL-3 concrete barrier using RBF-based metamodels and nonlinear finite element simulations. *Engineering Structures*, 114, 122–134. <https://doi.org/10.1016/j.engstruct.2016.02.009>

Yousefi, M., & Yousefi, M. (2019). Human resource allocation in an emergency department: A metamodel-based simulation optimization. *Kybernetes*, 49(3), 779–796. <https://doi.org/10.1108/K-12-2018-0675>

Yousefi, M., Yousefi, M., Ferreira, R. P. M., Kim, J. H., & Fogliatto, F. S. (2018). Chaotic genetic algorithm and Adaboost ensemble metamodeling approach for optimum resource planning in emergency departments. *Artificial Intelligence in Medicine*, 84, 23–33. <https://doi.org/10.1016/j.artmed.2017.10.002>

Zakerifar, M., Biles, W. E., & Evans, G. W. (2011). Kriging metamodeling in multiple-objective simulation optimization. *Simulation*, 87(10), 843–856. <https://doi.org/10.1177/0037549711411964>

Zeinali, F., Mahootchi, M., & Sepehri, M. M. (2015). Resource planning in the emergency departments: A simulation-based metamodeling approach. *Simulation Modelling Practice and Theory*, 53(1), 123–138. <https://doi.org/10.1016/j.simpat.2015.02.002>

Zhang, C., Osorio, C., & Flötteröd, G. (2017). Efficient calibration techniques for large-scale traffic simulators. *Transportation Research Part B: Methodological*, 97, 214–239. <https://doi.org/10.1016/j.trb.2016.12.005>

Zlatinov, M., & Laskowski, G. (2015). Hybrid large-eddy simulation optimization of a fundamental turbine blade turbulated cooling passage. *Journal of Guidance, Control, and Dynamics*, 38(6), 1292–1297. <https://doi.org/10.2514/1.B35713>

Chapter 6

Principal Component Analysis: An Overview and Applications in Multivariate Engineering Problems

Chapter details

Chapter DOI:

<https://doi.org/10.4322/978-65-86503-88-3.c06>

Chapter suggested citation / reference style:

Almeida, Fabricio A., et al. (2022). “Principal Component Analysis: An Overview and Applications in Multivariate Engineering Problems”. In Jorge, Ariosto B., et al. (Eds.) *Uncertainty Modeling: Fundamental Concepts and Models*, Vol. III, UnB, Brasilia, DF, Brazil, pp. 172–194. Book series in Discrete Models, Inverse Methods, & Uncertainty Modeling in Structural Integrity.

P.S.: DOI may be included at the end of citation, for completeness.

Book details

Book: Uncertainty Modeling: Fundamental Concepts and Models

Edited by: Jorge, Ariosto B., Anflor, Carla T. M., Gomes, Guilherme F., & Carneiro, Sergio H. S.

Volume III of Book Series in:

Discrete Models, Inverse Methods, & Uncertainty Modeling in Structural Integrity

Published by: UnB City: Brasilia, DF, Brazil Year: 2022

DOI: <https://doi.org/10.4322/978-65-86503-88-3>

Principal Component Analysis: An Overview and Applications in Multivariate Engineering Problems

Fabricio A. Almeida^{1*}, Guilherme F. Gomes², Pedro P. Balestrassi³,
Gabriela Belinato⁴, and Pedro A. R. C. Rosa⁵

¹Institute of Electrical Systems and Energy, Federal University of Itajubá (UNIFEI), Itajubá, Brazil – fabricio-almeida@unifei.edu.br

²Mechanical Engineering Institute, Federal University of Itajubá (UNIFEI), Itajubá, Brazil – guilhermefergom@unifei.edu.br

³Institute of Electrical Systems and Energy, Federal University of Itajubá (UNIFEI), Itajubá, Brazil – pedro@unifei.edu.br

⁴Federal Institute of South Minas Gerais (IFSULDEMINAS), Brazil – gabrielabelinato@ifsuldeminas.edu.br

⁵Department of Mechanical Engineering, Technician Superior Institute of Lisbon, University of Lisbon (ULisboa), Portugal – pedro.rosa@tecnico.ulisboa.pt

*Corresponding author

Abstract

This chapter presents an overview of principal component analysis (PCA), introducing and presenting the steps for using this powerful technique for data processing. In addition, three different examples are described, applying PCA to different multivariate engineering problems: in the manufacturing process using laser beam machining; power quality indices and in turbofan engine degradation data.

Keywords: *principal component analysis; multivariate statistics; laser beam machining; power quality indices; turbofan engine degradation.*

1 Contextualization

Most datasets usually have data with multiple characteristics that can be analyzed. For example, in a standard machining process, which features characteristics such as average roughness (R_a), cutting tool wear and material removal rate (MRR). Such characteristics present a relationship with each other, which can be statistically verified through their variance-covariance structure. The need to understand the relationships

between several variables of a correlated nature makes multivariate analysis an intrinsically complex subject (Johnson, R.A., Wichern, 2007).

When analyzing a set of characteristics, using univariate strategies (which deal with only one variable at a time) can bring unsatisfactory or even inadequate results. This can happen, because the multicollinearity existing in the set would be neglected (Almeida et al., 2020). Thus, it must be necessary to verify the relationship between the characteristics (which usually present correlation between them), requiring the use of multivariate strategies, promoting more informative and robust evaluations (Ferreira, 2018).

Among the commonly used strategies, principal component analysis (PCA) stands out, which was introduced by Pearson (Pearson, 1901) and later attributed differently by Hotelling (Hotelling, 1933). PCA is characterized as an exploratory multivariate technique that models correlated data from the variance-covariance structure (Ferreira, 2018). In addition, this technique allows the reduction of the dimensionality of the dataset (Gaudêncio et al., 2019; Jolliffe, 2010), finding a linear combination of uncorrelated variables that adequately explains the original variables, with the least possible loss of information (Mardia et al., 1995). In this way, the principal components can be obtained through a diagonalization, specifically, of defined semipositive symmetric matrices (Ferreira, 2018). The use of this technique can be found in many studies with different applications, such as: (Bounoua & Bakdi, 2021; Mahmoudi et al., 2021; Nhu et al., 2020; Song & Li, 2021; Yu et al., 2020).

Based on the previous discussion, this chapter will present an overview of the PCA strategy, indicating how the application should be carried out and interpreted, from the previous analysis of the data (before the application of the PCA). In addition, the steps to be considered by using this technique in the applied in datasets with multiple characteristics will be discussed. Finally, three different examples will be explored using problems in several areas of engineering, such as the manufacturing process, power quality indices and health monitoring of aeronautical engine.

2 Principal Component Analysis

As previously inferred, PCA is characterized by being a multivariate technique widely used to interpret and reduce extensive and correlated data (Wang & Chien, 2010). Thus, the first step to consider before using the PCA is to verify the significance of the data correlation structure. This analysis can be done through correlation tests such as Pearson's, in which it is possible to verify (through the p-value of the test) if the data present a significant correlation. In addition, it is possible to verify whether the characteristics are directly or inversely correlated. Correlation between the characteristics can be verified according to Equation (1):

$$Cor_{y_i y_j} = \frac{CoVar_{y_i y_j}}{\sqrt{Var_{y_i} Var_{y_j}}} \quad \forall i = 1, 2, \dots, q; \quad j = 1, 2, \dots, q \quad (1)$$

Where:

Var_{y_i} and Var_{y_j} are i^{th} and j^{th} variance;

$CoVar_{y_i, y_j}$ represents the covariance between the characteristics.

Datasets with a significant level of correlation usually present an ellipsoidal geometric structure, while variables without correlation, that is, independent, present a spherical structure. Figure 1 illustrates both behaviors: no significant correlation and significant correlation level.

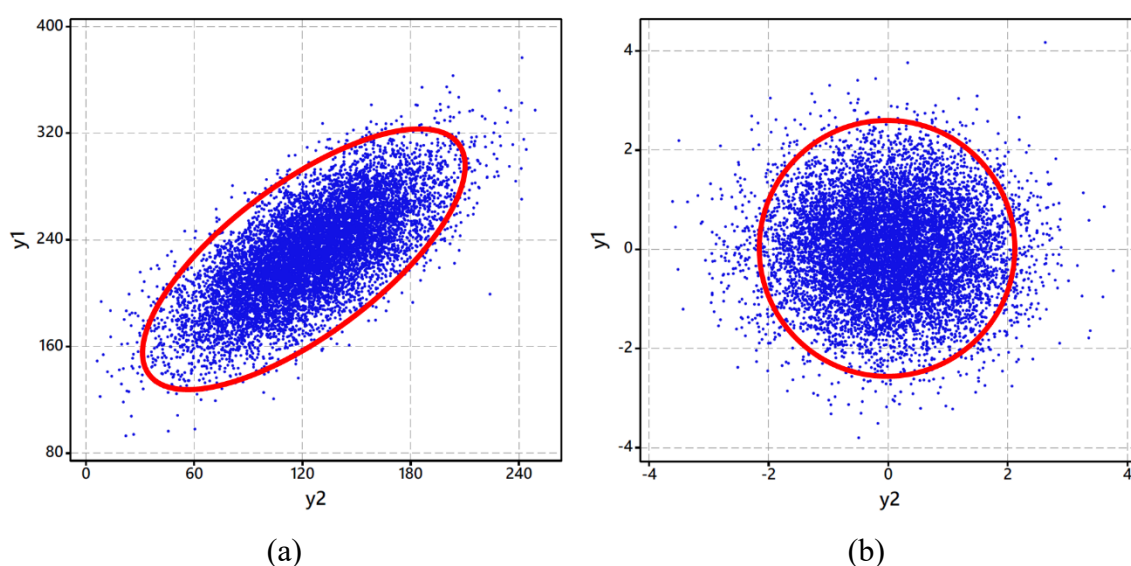


Figure 1: Data behavior with correlation: (a) $r = +0.705$; (b) $r = 0$ (Almeida, 2021).

If the data set does not show significant correlation, there is no need to use a multivariate strategy. However, when dealing with variables from the same source, there is usually a significant correlation between the analyzed characteristics. In this way, one can proceed with the application of the PCA strategy.

It is known that PCA minimizes the dimensionality of the original variables, in order to absorb significant elements in the principal axis, while maintaining the error variation in the secondary axes. Consequently, PCA stands out for being very widespread in the literature to reduce the computational effort in analyzes involving large and correlated data sets. This strategy makes use of an orthogonal conversion to transform the observations into a set of variables that are not correlated with each other (Almeida, 2021). Thus, one of the necessary parameterizations for the application of the PCA is to define the ideal number of principal components (PC_i) that will be considered in the model. This choice is not arbitrary and can be defined through specific criteria.

One of the guidelines used when defining the amount of PC_i is the Kaiser criterion (Johnson, R.A., Wichern, 2007). This criterion indicates that: the more correlation there is between the variables, the smaller the number of components needed to represent the observations. Thus, the components need to explain at least 80% of the accumulated

variance. For example, on a suitable dataset with seven correlated characteristics: if the first principal component (PC_1) has an explanation of 63.4% and the second principal component (PC_2) has a 19.9% explanation, it is known that two components are enough to adequately explain the original data set. Thus, there is a dimensionality reduction of 71.43% of the original dataset.

In addition to the percentage of explanation, another criterion commonly used to define the number of principal components is based on their eigenvalue. Thus, if the eigenvalue associated with the component is greater than or equal to 1 ($\lambda \geq 1$), this principal component will be associated in the model. Figure 2 shows the “Scree Plot” graph, which exemplifies the behavior of the eigenvalue in data from a welding process (adapted from (Almeida, 2017)).

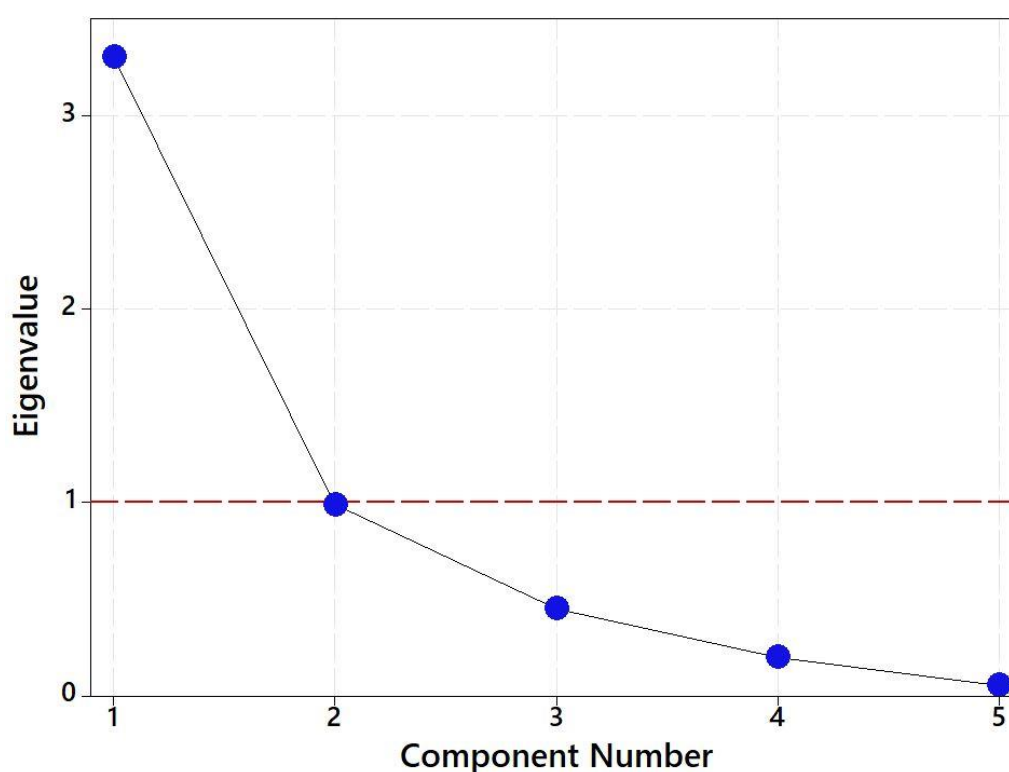


Figure 2: Scree plot

An alternative way to represent the behavior of both criteria (to choose the number of principal components) is through the Pareto chart (Figure 3). This graph illustrates the study's eigenvalue along with the explanation percentage of each component. In addition, the percentage of cumulative explanation is also presented, favoring decision making.

(Intentionally left blank)

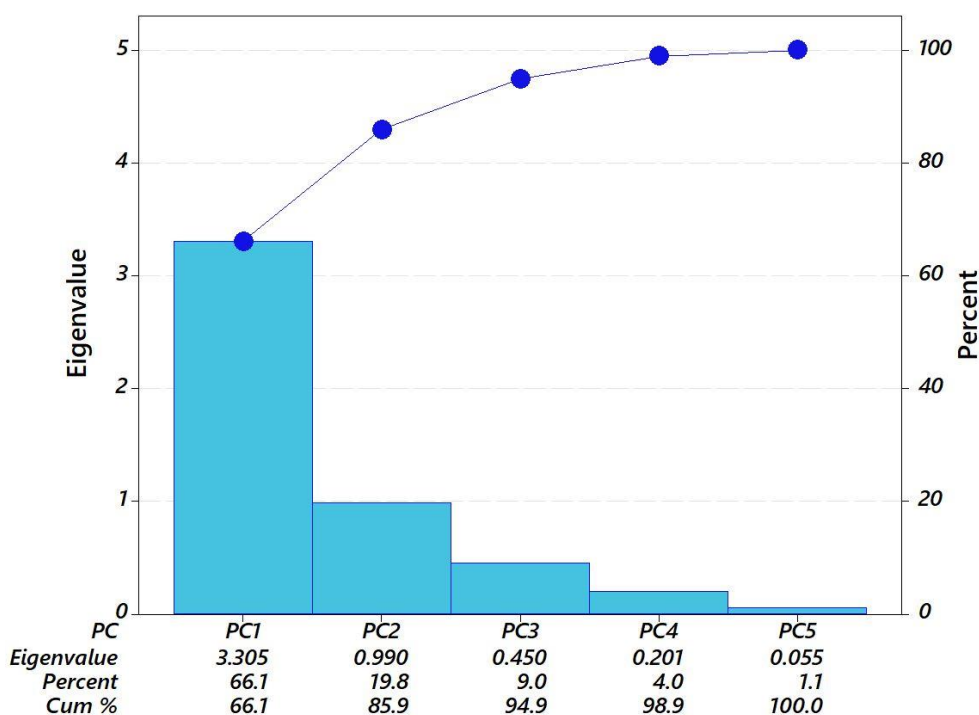


Figure 3: Pareto chart with eigenvalues and percentage of explanation

Since PCA is the combination of a linear set for q random variables Y_1, Y_2, \dots, Y_q , it can be inferred that the coordinate system represents a new set of coordinates before its original rotation, where the new axes hold the greatest data variability (Johnson, R.A., Wichern, 2007). Figure 4 shows the geometric interpretation of the axes.

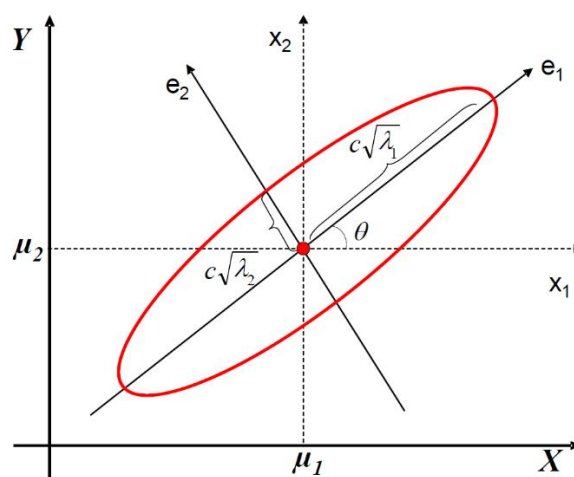


Figure 4. Constant density ellipsoid (Almeida et al., 2020).

PCA aims to find a combination of uncorrelated variables that adequately explains the original variables (Velasco et al., 2020). For this, we consider the random vector $\mathbf{X}^T = [X_1, X_2, \dots, X_p]$ which has the covariance matrix Σ with eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \dots \geq \lambda_p \geq 0$. Then, the linear combinations can be described as in Equation (2).

$$\begin{aligned}
Y_1 &= a_1^T X = a_{11}X_1 + a_{21}X_2 + \dots + a_{p1}X_p \\
Y_2 &= a_2^T X = a_{12}X_1 + a_{22}X_2 + \dots + a_{p2}X_p \\
&\dots \\
Y_p &= a_p^T X = a_{1p}X_1 + a_{2p}X_2 + \dots + a_{pp}X_p
\end{aligned} \tag{2}$$

Considering that Y_i for the i^{th} principal component, then, in Equation (3) and Equation (4):

$$\text{Var}(Y_i) = a_i^T \sum a_i = e_i^T \sum e_i; \quad \forall i = 1, 2, \dots, p \tag{3}$$

$$\text{CoVar}(Y_i, Y_k) = a_i^T \sum a_k = e_i^T \sum e_k; \quad \forall i, k = 1, 2, \dots, p \tag{4}$$

Thus, the principal components represent the uncorrelated linear combinations Y_1, Y_2, \dots, Y_p , in which the variances described in Equation (3) are the largest possible. That is, the i^{th} component can be defined from Equation (5), which was previously obtained through the formulation written in Equation (6) (Almeida, 2017).

$$PC_i = e_i^T Y = e_{1i}Y_1 + e_{2i}Y_2 + \dots + e_{qi}Y_q \quad i = 1, 2, \dots, q \tag{5}$$

$$\begin{aligned}
&\text{Max} \quad \text{Var} \left[e_i^T Y \right] \\
&\text{subject to:} \\
&\quad e_i^T e_i = 1 \\
&\quad \text{Cov} \left[e_i^T Y, e_k^T Y \right] = 0 \\
&\quad \quad k < i
\end{aligned} \tag{6}$$

Therefore, the original variables can be replaced by an uncorrelated linear set, i.e., the scores of the principal components. According to Johnson, R.A., Wichern, (2007), in possession of the standardized data matrix \mathbf{Z} and the eigenvectors matrix \mathbf{E} (from a multivariate set), the scores of the principal components can be obtained by Equation (7).

$$PC_{\text{score}} = \mathbf{Z}^T \mathbf{E} = \begin{bmatrix} \left(\frac{y_{11} - \bar{y}_1}{\sqrt{s_{11}}} \right) & \left(\frac{y_{12} - \bar{y}_2}{\sqrt{s_{22}}} \right) & \dots & \left(\frac{y_{1q} - \bar{y}_q}{\sqrt{s_{qq}}} \right) \\ \left(\frac{y_{21} - \bar{y}_1}{\sqrt{s_{11}}} \right) & \left(\frac{y_{22} - \bar{y}_2}{\sqrt{s_{22}}} \right) & \dots & \left(\frac{y_{2q} - \bar{y}_q}{\sqrt{s_{qq}}} \right) \\ \vdots & \vdots & \ddots & \vdots \\ \left(\frac{y_{n1} - \bar{y}_1}{\sqrt{s_{11}}} \right) & \left(\frac{y_{n2} - \bar{y}_2}{\sqrt{s_{22}}} \right) & \dots & \left(\frac{y_{nq} - \bar{y}_q}{\sqrt{s_{qq}}} \right) \end{bmatrix}^T \times \begin{bmatrix} e_{11} & e_{12} & \dots & e_{1q} \\ e_{21} & e_{22} & \dots & e_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ e_{q1} & e_{q2} & \dots & e_{qq} \end{bmatrix} \tag{7}$$

Principal component scores are dimensionless, independent representations of an entire original dataset. These scores can be used in the most varied applications: from techniques such as artificial neural networks, cluster analysis to optimization techniques.

In the following topics, three real examples will be detailed using PCA on data from different multivariate processes in engineering. In such manner, applications of PCA will be addressed in data from a laser machining process (Belinato et al., 2019); in power quality sector (Almeida et al., 2021) and on data from NASA's Turbofan engine degradation analysis (Saxena & Goebel, 2008).

3 On the Use of PCA in Engineering Applications

3.1 PCA-based LBM process

The first example to be addressed will be based on the article by Belinato et al., (2019). In this study, the authors used the PCA technique with other statistical methods for experimentation and optimization in the machining process through laser beam machining (LBM). In general, this machining process has multiple objectives due to its quality characteristics. The authors performed a design of experiments (DOE) using the Response Surface Methodology (RSM) for the parameters of laser frequency (f), cut speed (S), laser power (I). The quality characteristics investigated were: material removal rate (MRR) and different roughness metrics (R_a , R_q , R_z , R_p and R_t). Table 1 and Table 2 present the parameter levels and the experimental matrix with the multiple quality characteristics analyzed, respectively.

The experiments were performed on a Deckel Maho Lasertec® machine model DML40SI (Figure 5a) and data collection used a Mahr® rugosimeter model M300 with an RD18 measuring device (Figure 5b). The workpieces used can be seen in Figure 6. All planning, experiment and data collection were carried out at Instituto Superior Técnico, University of Lisbon.

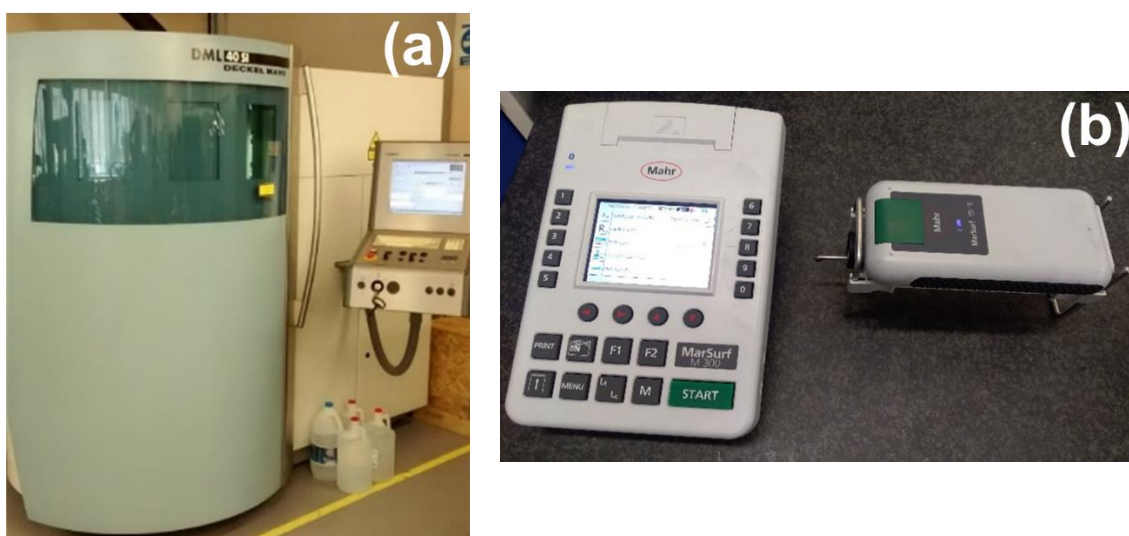


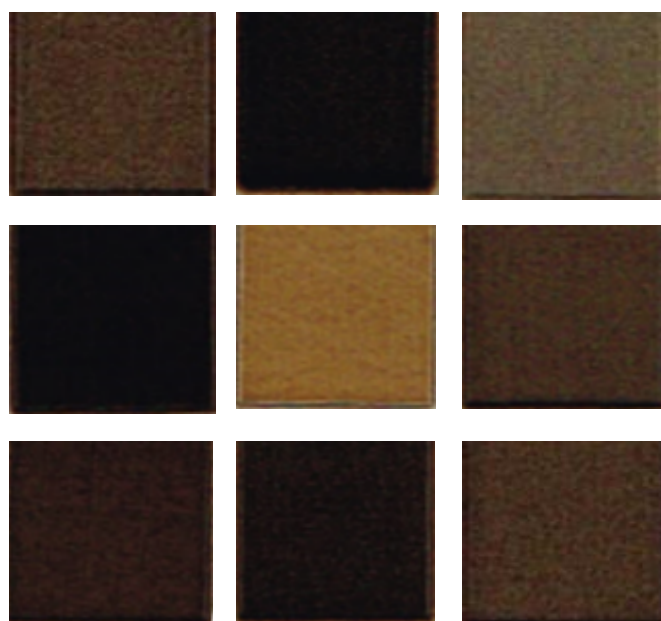
Figure 5: (a) DML40SI LBM machine and (b) Mahr® M300 rugosimeter (Belinato et al., 2019).

Table 1: Input parameters and levels (Belinato et al., 2019).

Input parameters	Level				
	-1.682	-1	0	+1	+1.682
f [kHz]	11.2	15	20.5	26	29.7
S [mm/min]	29.5	200	450	700	870.4
I [%]	26.3	40	60	80	93.6

Table 2: Experimental matrix (Belinato et al., 2019).

N	Setup			Responses					
	f [kHz]	S [mm/min]	I [%]	R_a [μ m]	R_q [μ m]	R_z [μ m]	R_p [μ m]	R_t [μ m]	MRR [cm ³ /s]
1	15.0	200.0	40.0	4.54	5.85	31.30	14.46	39.58	5.95 x10 ⁻⁴
2	26.0	200.0	40.0	2.12	2.65	12.77	6.70	14.00	4.20 x10 ⁻⁴
3	15.0	700.0	40.0	7.27	9.03	42.06	22.59	47.90	7.81 x10 ⁻⁴
4	26.0	700.0	40.0	3.68	4.53	17.47	8.82	18.86	4.21 x10 ⁻⁴
5	15.0	200.0	80.0	12.38	15.07	66.30	33.68	78.49	1.84 x10 ⁻³
6	26.0	200.0	80.0	5.28	6.91	35.33	20.30	45.70	6.57 x10 ⁻⁴
7	15.0	700.0	80.0	11.82	14.66	63.70	31.40	79.66	2.05 x10 ⁻³
8	26.0	700.0	80.0	6.08	7.25	36.26	17.15	52.88	2.33 x10 ⁻³
9	11.2	450.0	60.0	11.93	14.93	61.52	33.55	80.09	1.77 x10 ⁻³
10	29.7	450.0	60.0	3.14	4.71	14.96	8.04	17.93	1.21 x10 ⁻³
11	20.5	29.55	60.0	11.91	15.91	69.30	33.67	95.14	2.24 x10 ⁻⁴
12	20.5	870.45	60.0	3.07	4.14	26.24	12.64	30.49	1.24 x10 ⁻³
13	20.5	450.0	26.36	3.64	4.36	17.57	8.68	22.42	3.07 x10 ⁻⁴
14	20.5	450.0	93.64	10.73	12.92	71.64	32.47	89.21	2.70 x10 ⁻³
15	20.5	450.0	60.0	6.22	7.62	36.61	17.18	47.00	1.54 x10 ⁻³
16	20.5	450.0	60.0	5.90	7.38	36.71	17.81	42.50	1.57 x10 ⁻³
17	20.5	450.0	60.0	6.17	7.51	36.35	16.91	38.37	1.56 x10 ⁻³
18	20.5	450.0	60.0	6.23	7.71	36.14	18.10	41.29	1.56 x10 ⁻³
19	20.5	450.0	60.0	5.88	7.55	36.52	17.95	44.76	1.58 x10 ⁻³
20	20.5	450.0	60.0	6.17	7.60	35.60	17.33	38.25	1.60 x10 ⁻³

**Figure 6: Machined surface workpieces (Belinato et al., 2019).**

To emphasize principal components analysis, initially a correlation analysis is performed to verify the behavior of the quality characteristics. From Table 3, it appears that there is a significant correlation between the characteristics. This analysis can be verified using Pearson's correlation test and the relationship between the quality characteristics can also be verified through the correlation matrix illustrated in Figure 7. Then, information on eigenvalue and percentage of explanation is showed to define the number of components to be used in the analysis. Figure 8 presents the Pareto chart for the eigenvalues and cumulative percentage of explanation of the components. Through this analysis, it can be seen that the first component (PC₁) is the only one to present an eigenvalue greater than 1, in addition to explaining 86% of the original data. Thus, only 1 component is sufficient to represent all the quality characteristics analyzed.

Knowing the number of components needed (only 1), in this case, it is enough to extract the component scores, which represent the original dataset in a dimensionless way. From this result, it is possible to verify that the multivariate PCA technique provides a data dimensionality reduction by 83.33%. This result favors a less complex analysis with less computational effort for the next steps (such as optimization, forecasting etc). For more details about the process and the approach performed by the authors, see the study by Belinato et al., (2019).

Table 3: Correlation analysis for the LBM process characteristics (Belinato et al., 2019).

	Ra				
Rq	0.996 ⁽¹⁾ 0.000 ⁽²⁾	Rq			
Rz	0.973 ⁽¹⁾ 0.000 ⁽²⁾	0.969 ⁽¹⁾ 0.000 ⁽²⁾	Rz		
Rp	0.979 ⁽¹⁾ 0.000 ⁽²⁾	0.979 ⁽¹⁾ 0.000 ⁽²⁾	0.989 ⁽¹⁾ 0.000 ⁽²⁾	Rp	
Rt	0.960 ⁽¹⁾ 0.000 ⁽²⁾	0.962 ⁽¹⁾ 0.000 ⁽²⁾	0.987 ⁽¹⁾ 0.000 ⁽²⁾	0.976 ⁽¹⁾ 0.000 ⁽²⁾	Rt
MRR	0.460 ⁽¹⁾ 0.041 ⁽²⁾	0.415 ⁽¹⁾ 0.069 ⁽²⁾	0.500 ⁽¹⁾ 0.025 ⁽²⁾	0.454 ⁽¹⁾ 0.044 ⁽²⁾	0.462 ⁽¹⁾ 0.040 ⁽²⁾

(1) Pearson correlation

(2) P-Value

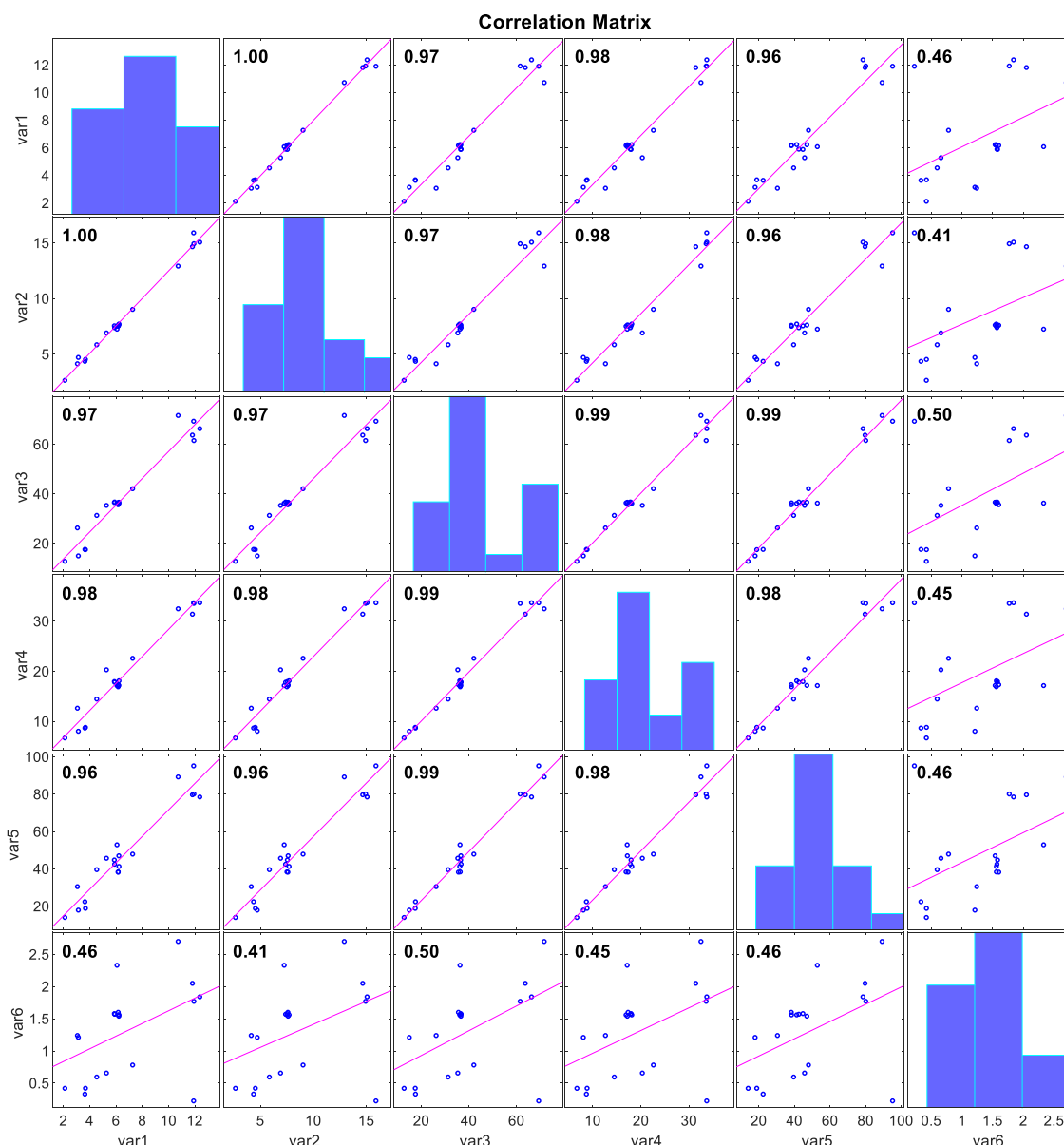


Figure 7: Correlation matrix from LBM process.

3.2 On the use of PCA in electric power substation quality indices

The second example to be addressed is derived from an analysis of power quality indices of electric power substations addressed by Almeida et al., (2021). In this study, the authors used data from 17 substations with 31 power quality characteristics. The substations are located in southeastern Brazil, corresponding to a total area of 41,241 km², about 90% of the state of Espírito Santo. Figure 9 illustrates, geographically, the location of the analyzed substations.

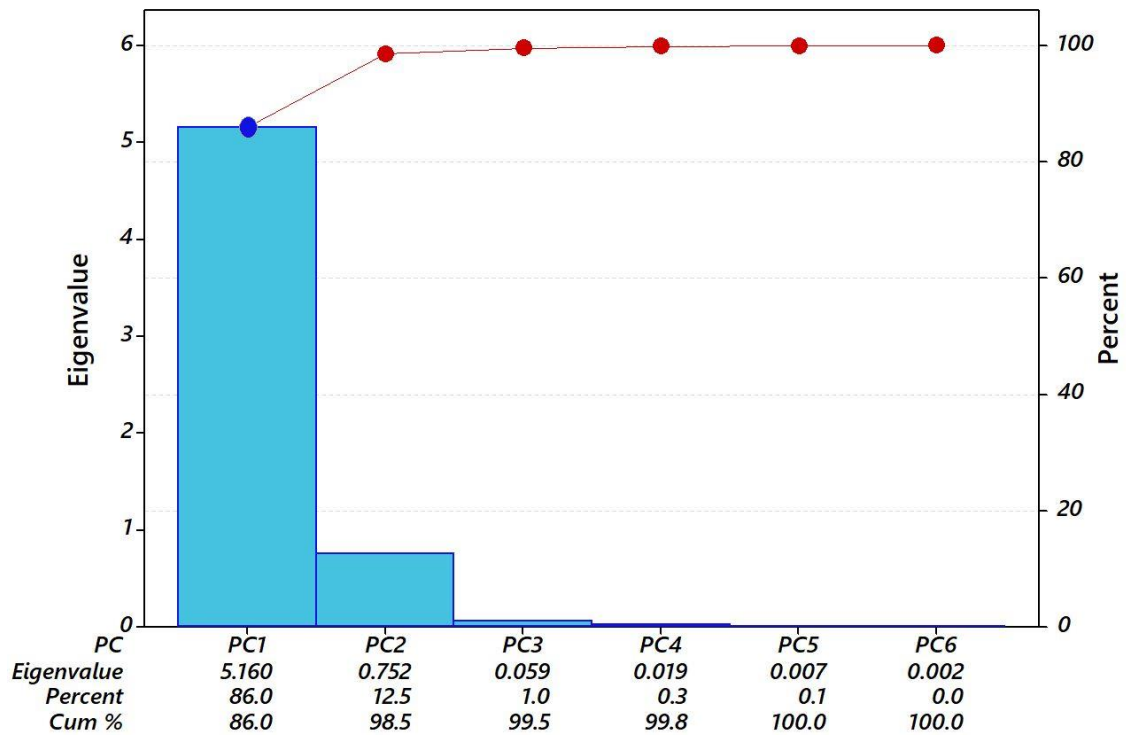


Figure 8: Pareto chart with eigenvalues and percentage of explanation (Belinato et al., 2019).

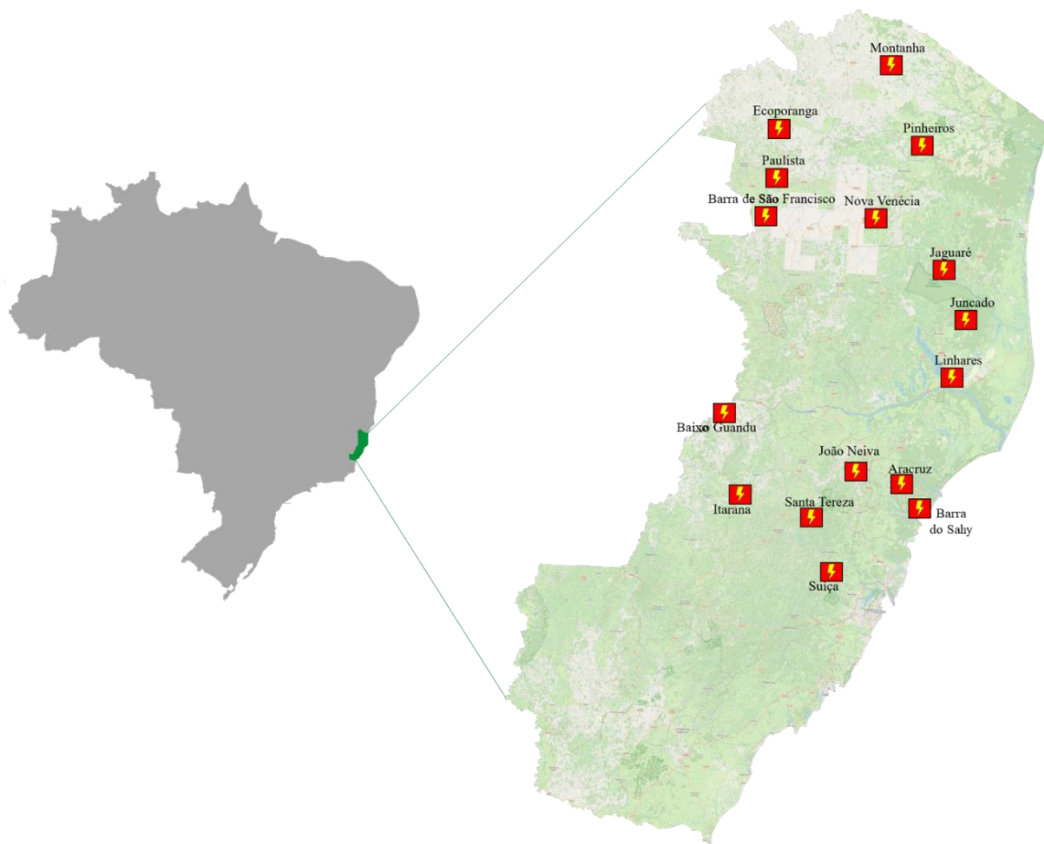


Figure 9: Location of the investigated substations in the State of Espírito Santo (Almeida et al., 2022).

Power quality measurements were collected over a year, in order to cover different seasonalities that influence the performance of the electricity distribution network (such as rain, winds, among other phenomena). In addition, 30 power quality monitors from Schweitzer Engineering Laboratories, model SEL 734, were used to acquire this data. The behavior of the dataset towards the substations is illustrated in Figure 10. The set were applied in a method to find the best clustering technique for this set and, consequently, classify the substations based on the power quality. Due to the correlated structure of the data (originally available in Miranda et al., (2016), the use of exploratory techniques such as PCA is recommended.

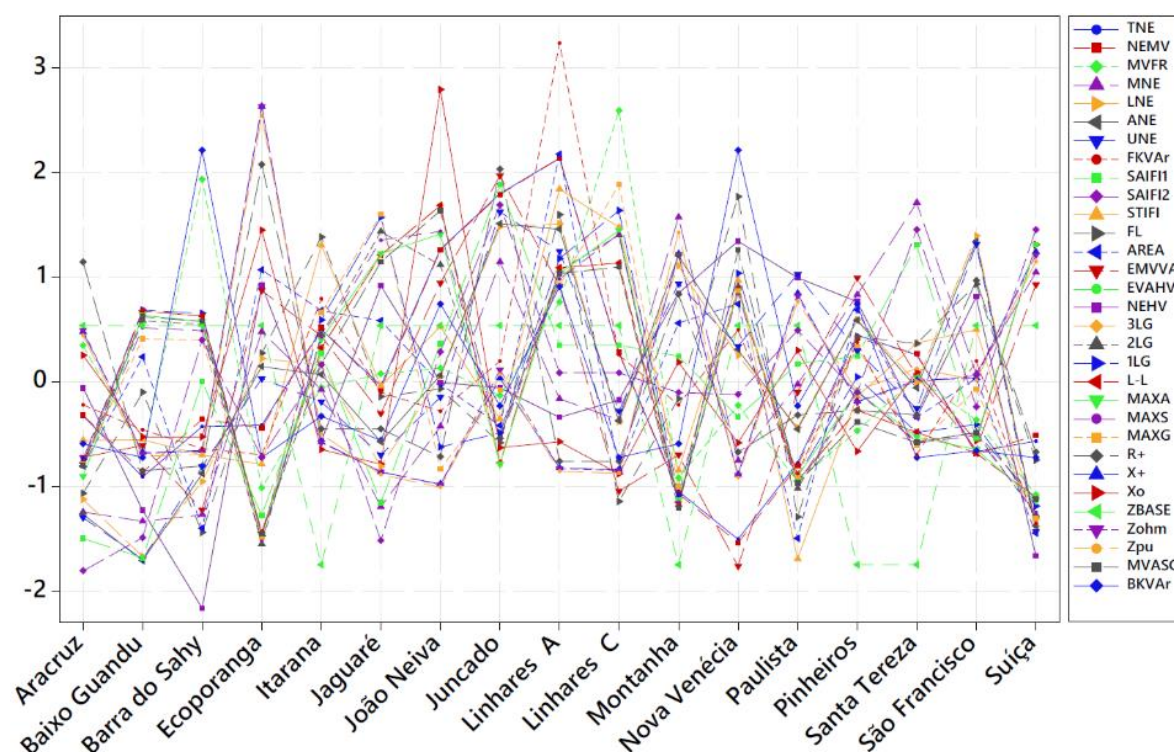


Figure 10: Relationship between power quality indices and substations (Almeida, 2021).

After test and confirm that data structure presents a significant correlation, the eigenvalues and the percentage of explanation of each component are calculated. This procedure is intended to determine the number of principal components needed for the study. As an alternative view of the first example (applied to the process by LBM), the behavior of the eigenvalues can also be visualized through the Scree plot (Figure 11). In this graph it is possible to verify that the first six components present eigenvalues greater than 1 ($\lambda \geq 1$). Figure 12 illustrates the behavior of eigenvalues 1 and 2 of their respective principal components.

(Intentionally left blank)

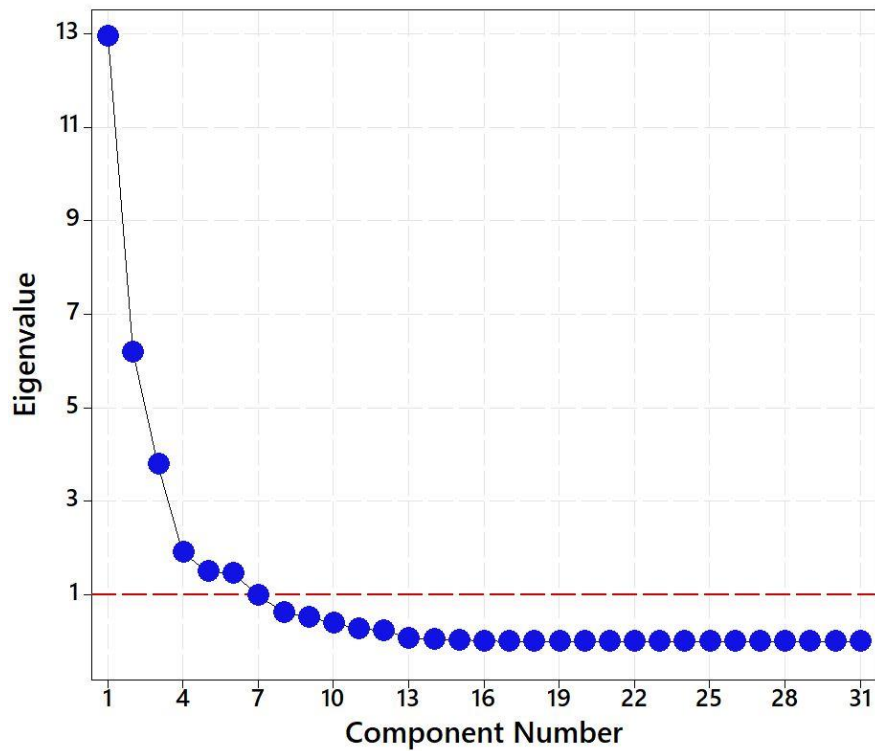


Figure 11: Scree plot of eigenvalues from power quality indices (adapted from Almeida, 2021).

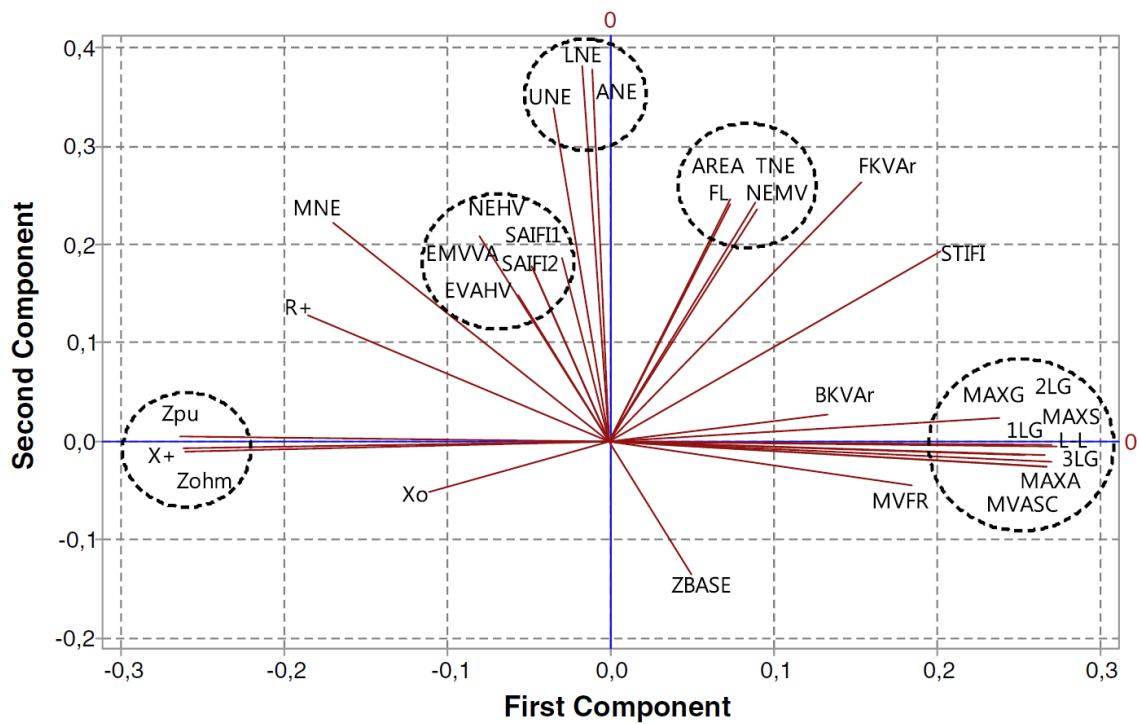


Figure 12: Plot of eigenvectors 1 and 2 (Miranda et al., 2016).

Complementarily, the Pareto chart (Figure 13) also provides information on the explanation percentage of each principal component. In this graph it is possible to verify that the fourth component presents a cumulative percentage of 80%. In this case, it is advisable to use the number of components necessary to respect the eigenvalue criterion. By using 6 components instead of 4, there data explanation contribution increases from 9.53%, in addition to properly respecting the Kaiser criterion discussed above.

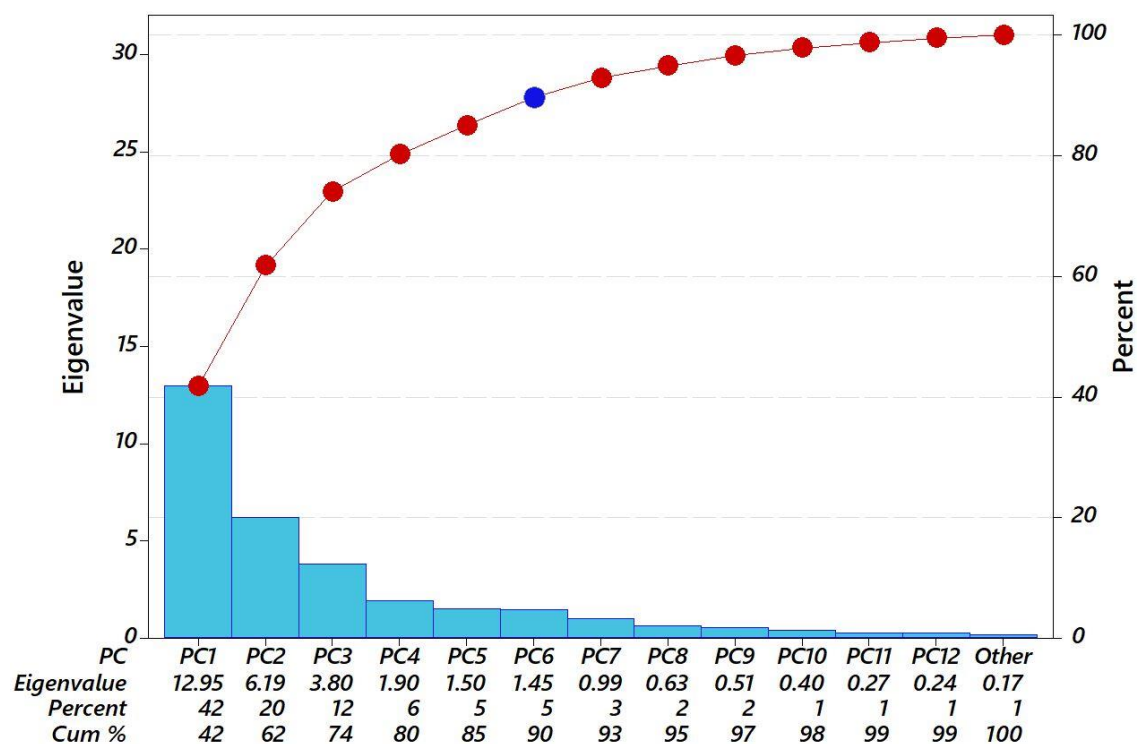


Figure 13: PCA Pareto chart for substation data.

There is a discussion in literature about the number of components to be used in exploratory analysis studies, both for principal components and factor analysis (another widely used exploratory technique). Some of these discussions can be verified in studies such as Visinescu & Evangelopoulos (2014) and Almeida (2021).

Considering the six principal components, the extraction of scores can be performed, thus generating 6 vectors of dimensionless and independent scores. These values adequately represent and explain the 31 original variables of the study. From this application, there is an 80.5% reduction in the data dimensionality. This result favors further analyzes that involve computational effort, as in the application of cluster algorithms used in the aforementioned power quality studies. The following studies bring more details and information about the object of study: (Almeida et al., 2022); (Almeida et al., 2021) and (Miranda et al., 2016).

3.3 On the use of PCA in Turbofan engine degradation data

The analysis and investigation of aeronautical engine data is widely investigated by industries and researchers (Chatterjee & Litt, 2003; Deng et al., 2020; Goebel et al.,

2007; Kurosaki et al., 2004; Listou Ellefsen et al., 2019; Saxena et al., 2008; Xu et al., 2020), comprising the collection of numerous information generated by several different sensors in specific positions. The diagnosis of this magnitude generates a large amount of data that usually present a multivariate characteristic. Thus, the PCA strategy can also be explored in this context. Data referring to turbofan engine degradation predictions will be analyzed. This set can be found in the public archives of the National Aeronautics and Space Administration (NASA), referring to the study by Saxena et al., (2008) and available at Saxena and Goebel, (2008). The data were generated using C-MAPSS – Commercial Modular AeroPropulsion System Simulation software, based on the behavior of turbofan engines, shown in the diagram of Figure 14.

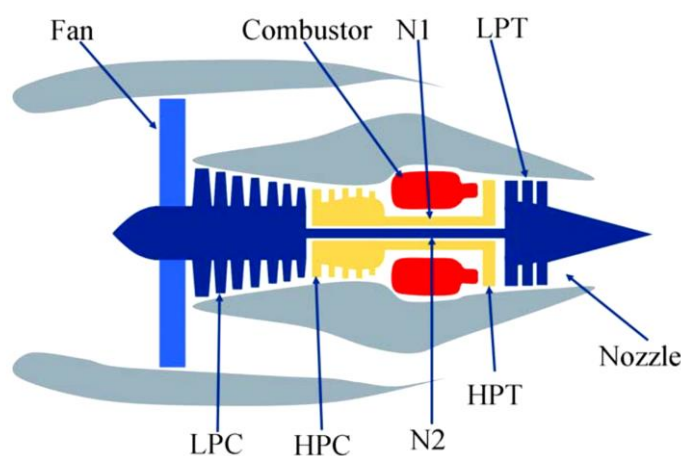


Figure 14: Diagram of engine simulated (Saxena et al., 2008).

For this example, a set will be selected that describes the simulation of 218 motors in constant execution until the moment of failure (between 127 and 356 cycles). The set presents information for analysis, using 21 C-MAPSS outputs with 33,991 collections for each characteristic (totaling 713,811 data). Table 4 describes the characteristics of the C-MAPSS outputs. Figure 15 illustrates the behavior of the sensors as a function of the engine's remaining useful life (RUL). The complete dataset and more details about the collection are available in the articles and database mentioned above.

Through correlation analysis it is possible to assume that there is a significant variance-covariance structure for the whole set, considering a confidence interval of 95%. Given the large amount of data, the correlation analysis will not be presented here, but it can be easily replicated with the help of any statistical software.

From the correlation level, the data are able to be used in the PCA strategy. Thus, the eigenvalues and contribution percentage of each component are initially verified to estimate the ideal amount for this data set. Through this initial analysis, it is possible to verify that only two principal components present $\lambda \geq 1$ (16,908 and 3,578, respectively). The behavior of the eigenvalues can be verified through Figure 16. In addition, the second component presents an accumulated explanation percentage of 97.55% of the data. Such results infer that only two components are sufficient to adequately represent all 21 quality characteristics collected through the sensors. The

Pareto chart illustrates the behavior of both results mentioned above, as shown in Figure 17.

Table 4: C-MAPSS outputs (Adapted from (Saxena et al., 2008)).

Symbol	Description	Units
T2	Total temperature at fan inlet	°R
T24	Total temperature at LPC outlet	°R
T30	Total temperature at HPC outlet	°R
T50	Total temperature at LPT outlet	°R
P2	Pressure at fan inlet	psia
P15	Total pressure in bypass-duct	psia
P30	Total pressure at HPC outlet	psia
Nf	Physical fan speed	rpm
Nc	Physical core speed	rpm
epr	Engine pressure ratio (P50/P2)	--
Ps30	Static pressure at HPC outlet	psia
phi	Ratio of fuel flow to Ps30	pps/psi
NRf	Corrected fan speed	rpm
NRc	Corrected core speed	rpm
BPR	Bypass Ratio	--
farB	Burner fuel-air ratio	--
htBleed	Bleed Enthalpy	--
Nf_dmd	Demanded fan speed	rpm
PCNfR_dmd	Demanded corrected fan speed	rpm
W31	HPT coolant bleed	lbm/s
W32	LPT coolant bleed lbm/s	lbm/s

(Intentionally left blank)

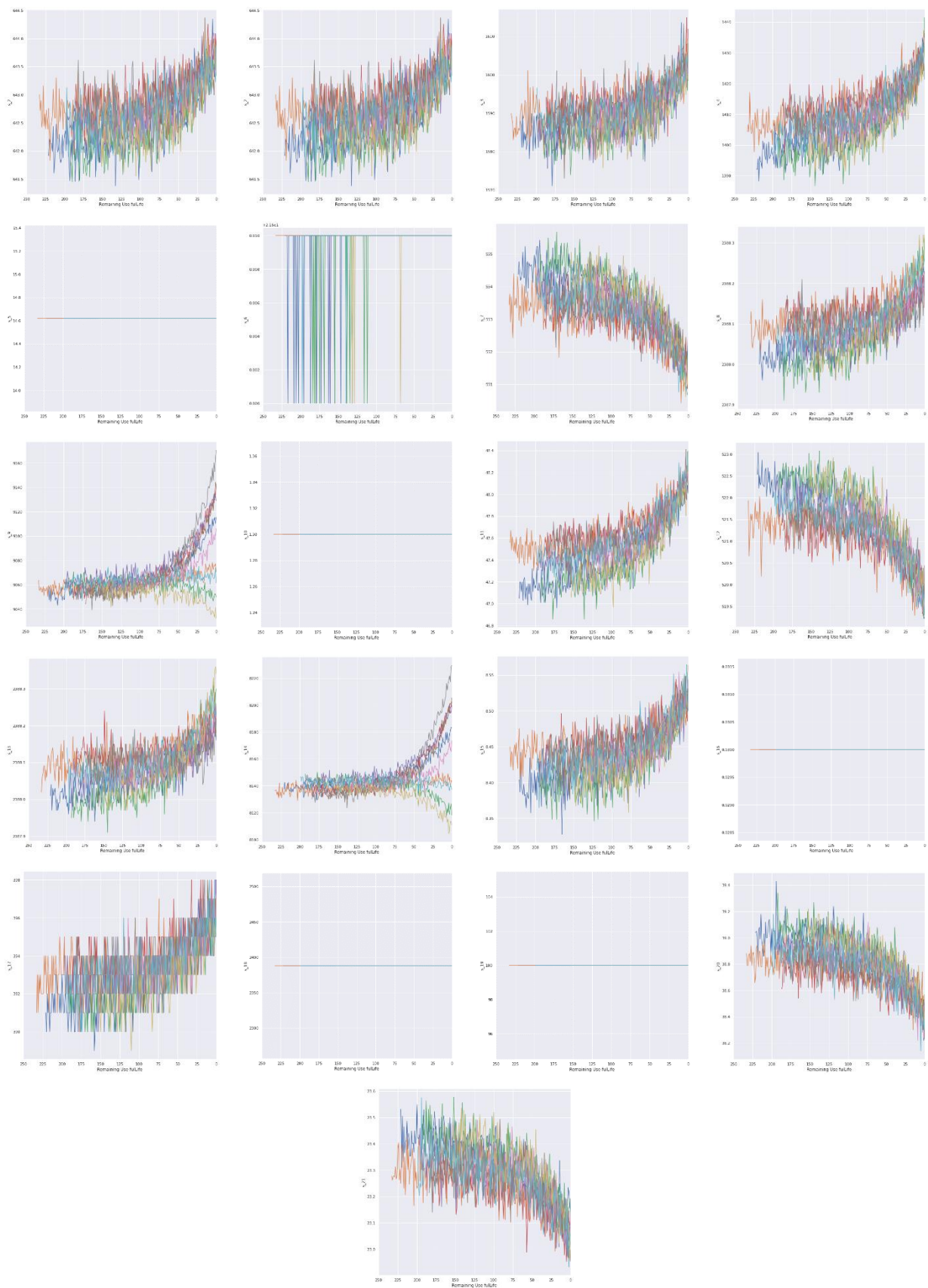


Figure 15: Sensors signal (#1 to #21) in function of the engine's RUL.

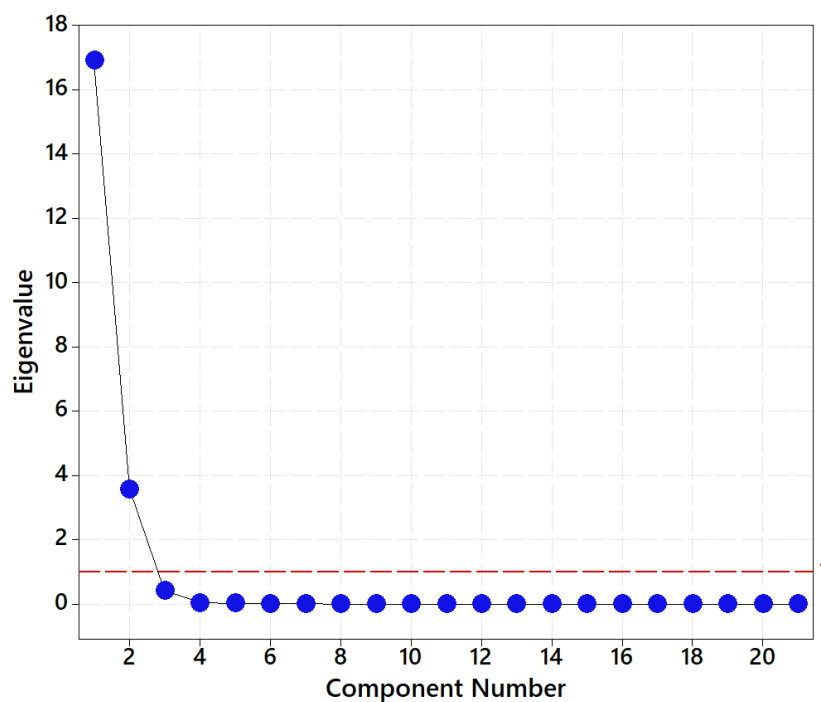


Figure 16: Scree plot of eigenvalues from C-MAPSS outputs.

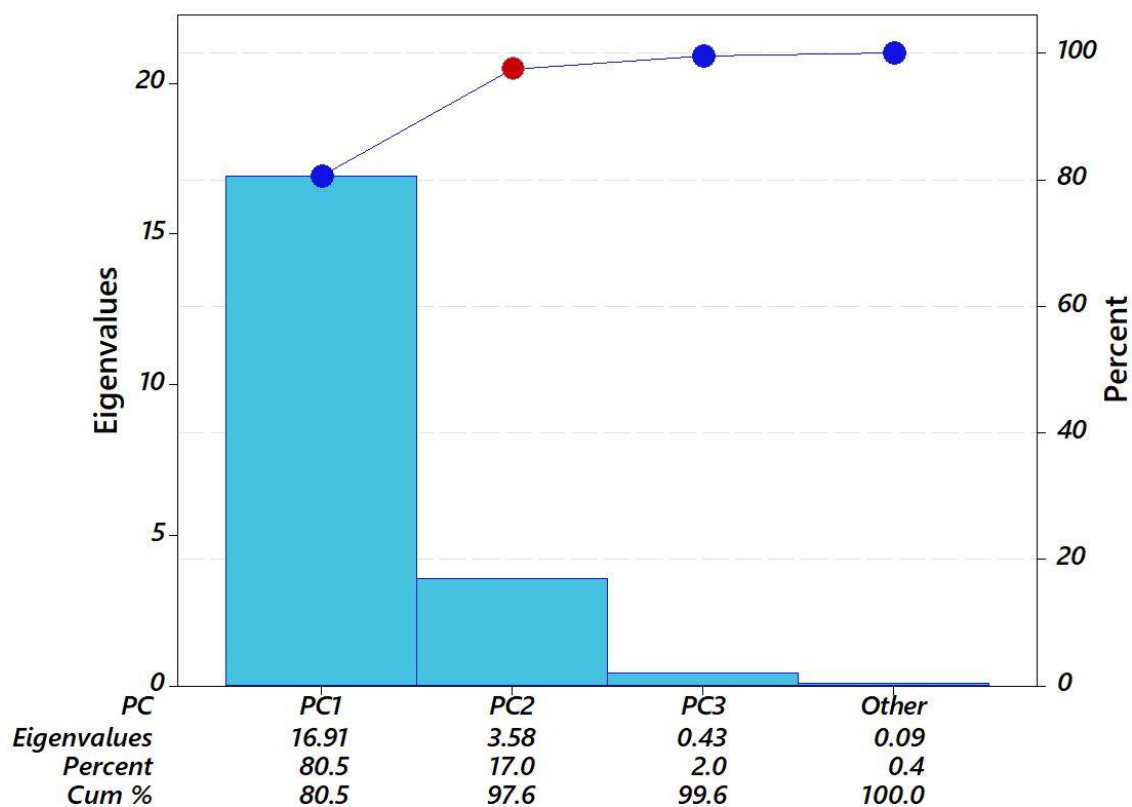


Figure 17: PCA Pareto chart for C-MAPSS outputs.

In a complementary way, it is possible to verify which variables have the greatest effect on these components. This relationship can be seen through a loading plot, as described in Figure 18. In this graph, the variables closer to 0 have a weak influence and, antagonistically, those close to 1 or -1 have a strong influence on the component.

From the ideal number of components (in this case, only 2), it is possible to extract the scores that will represent all the original variables through independent and dimensionless vectors. From the application of the analysis, it is possible to reach a reduction of the data dimensionality of 95.25%, i.e., the original set that had 713,811 data points can be properly represented by 67,822 data points. This significant reduction favors the computational and analytical performance of the data, reducing the time and helping for more accurate evaluations, since the PCA considers the correlated structure of the data. The following studies bring more details and information about the object of study: Saxena et al., (2008); Saxena and Goebel, (2008).

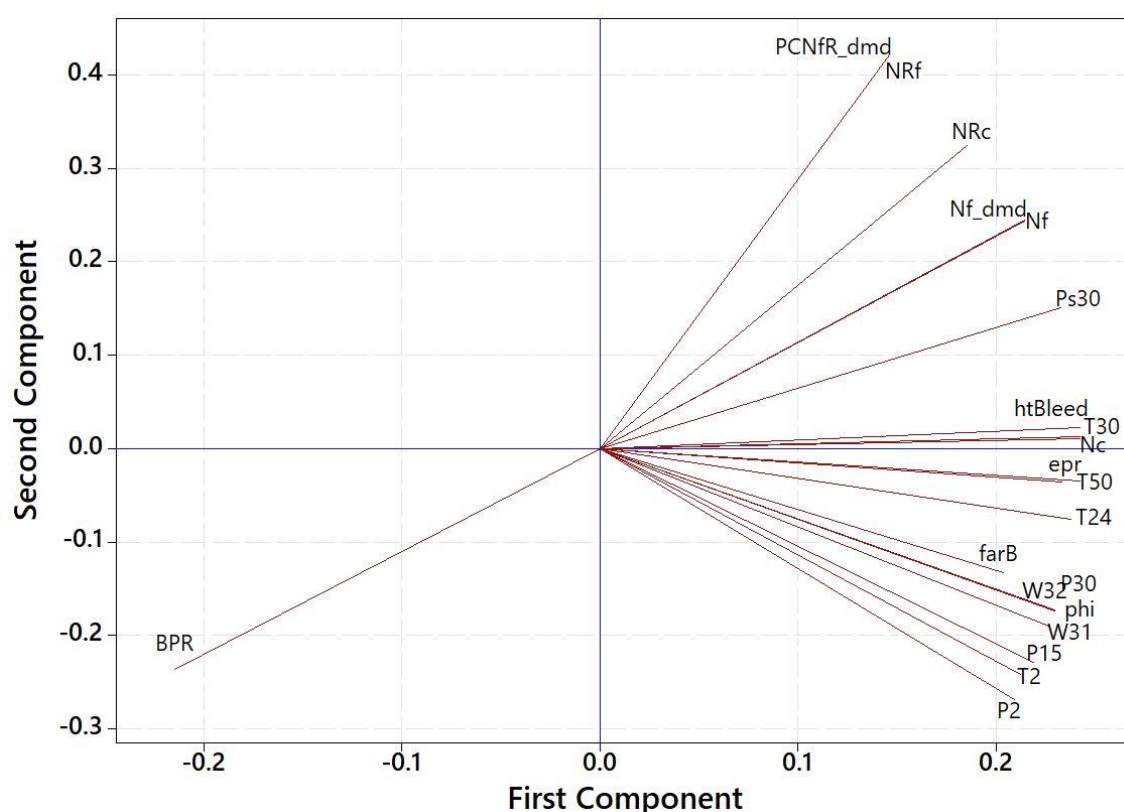


Figure 18: Loading plot from the C-MAPSS outputs

4 Conclusion

In this chapter we present a general and brief overview of the principal component analysis technique. This technique is characterized as a powerful tool to reduce the data dimensionality and create uncorrelated response vectors, being widely used in several segments. In addition to an explanation of this strategy, this chapter described 3 examples that cover different areas of engineering, showing the possibilities of using PCA. Finally, it is hoped that this chapter will help readers understand PCA, as well as how to apply this strategy to other multivariate data sets.

Acknowledgements

The authors would like to acknowledge the financial support from the Brazilian agency CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) under

process number 150117/2021–3, CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), and FAPEMIG (Fundação de Amparo à Pesquisa do Estado de Minas Gerais - APQ-00385-18).

References

- Almeida, F. A. de. (2021). Aprimoramento do poder discriminatório de funções elipsoidais modificadas por cargas fatoriais rotacionadas na formação otimizada de agrupamentos. Universidade Federal de Itajubá (UNIFEI).
- Almeida, F. A. De, Mello, L. G. D., Romão, E. L., Gomes, G. F., Gomes, J. H. D. F., Paiva, A. P. De, Filho, J. M., & Balestrassi, P. P. (2021). A PCA-Based Consistency and Sensitivity Approach for Assessing Linkage Methods in Voltage Sag Studies. *IEEE Access*, 9, 84871–84885. <https://doi.org/10.1109/ACCESS.2021.3088436>
- Almeida, F. A. de, Romão, E. L., Gomes, G. F., Gomes, J. H. de F., Paiva, A. P. de, Miranda Filho, J., & Balestrassi, P. P. (2022). Combining machine learning techniques with Kappa–Kendall indexes for robust hard-cluster assessment in substation pattern recognition. *Electric Power Systems Research*, 206, 107778. <https://doi.org/https://doi.org/10.1016/j.epsr.2022.107778>
- Almeida, F. A. (2017). *Análise Multivariada do Sistema de Medição de um Processo de Solda a Ponto por Resistência Elétrica utilizando Componentes Principais Ponderados*. Universidade Federal de Itajubá.
- Almeida, F. A., Leite, R. R., Gomes, G. F., Gomes, J. H. de F., & de Paiva, A. P. (2020). Multivariate data quality assessment based on rotated factor scores and confidence ellipsoids. *Decision Support Systems*, 129, 113173. <https://doi.org/10.1016/j.dss.2019.113173>
- Belinato, G., de Almeida, F. A., de Paiva, A. P., de Freitas Gomes, J. H., Balestrassi, P. P., & Rosa, P. A. R. C. (2019). A multivariate normal boundary intersection PCA-based approach to reduce dimensionality in optimization problems for LBM process. *Engineering with Computers*, 35(4). <https://doi.org/10.1007/s00366-018-0678-3>
- Bounoua, W., & Bakdi, A. (2021). Fault detection and diagnosis of nonlinear dynamical processes through correlation dimension and fractal analysis based dynamic kernel PCA. *Chemical Engineering Science*, 229, 116099. <https://doi.org/https://doi.org/10.1016/j.ces.2020.116099>
- Chatterjee, S., & Litt, J. (2003). Online Model Parameter Estimation of Jet Engine Degradation for Autonomous Propulsion Control. In *AIAA Guidance, Navigation, and Control Conference and Exhibit*. American Institute of Aeronautics and Astronautics. <https://doi.org/doi:10.2514/6.2003-5425>
- Deng, Y., Bucchianico, A. Di, & Pechenizkiy, M. (2020). Controlling the accuracy and uncertainty trade-off in RUL prediction with a surrogate Wiener propagation model. *Reliability Engineering & System Safety*, 196, 106727. <https://doi.org/https://doi.org/10.1016/j.ress.2019.106727>
- Ferreira, D. F. (Federal U. of L. (2018). *Estatística Multivariada* (3th ed.). UFLA.

- Gaudêncio, J. H. D., Almeida, F. A., Turrioni, J. B., Quinino, R. C., Balestrassi, P. P., & Paiva, A. P. (2019). A multiobjective optimization model for machining quality in the AISI 12L14 steel turning process using fuzzy multivariate mean square error. *Precision Engineering*, 56(December 2018), 303–320. <https://doi.org/10.1016/j.precisioneng.2019.01.001>
- Goebel, K., Qiu, H., Eklund, N., & Yan, W. (2007). Modeling Propagation of Gas Path Damage. *2007 IEEE Aerospace Conference*, 1–8. <https://doi.org/10.1109/AERO.2007.352835>
- Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24(7), 498–520. <https://doi.org/10.1037/h0070888>
- Johnson, R.A., Wichern, D. (2007). *Applied Multivariate Statistical Analysis* (6th ed.). Prentice-Hall.
- Jolliffe, I. T. (2010). *Principal Component Analysis* (2nd ed.). Springer Series in Statistics.
- Kurosaki, M., Morioka, T., Ebina, K., Maruyama, M., Yasuda, T., & Endoh, M. (2004). Fault Detection and Identification in an IM270 Gas Turbine Using Measurements for Engine Control. *Journal of Engineering for Gas Turbines and Power*, 126(4), 726–732. <https://doi.org/10.1115/1.1787515>
- Listou Ellefsen, A., Bjørlykhaug, E., Æsøy, V., Ushakov, S., & Zhang, H. (2019). Remaining useful life predictions for turbofan engine degradation using semi-supervised deep architecture. *Reliability Engineering & System Safety*, 183, 240–251. <https://doi.org/https://doi.org/10.1016/j.ress.2018.11.027>
- Mahmoudi, M. R., Heydari, M. H., Qasem, S. N., Mosavi, A., & Band, S. S. (2021). Principal component analysis to study the relations between the spread rates of COVID-19 in high risks countries. *Alexandria Engineering Journal*, 60(1), 457–464. <https://doi.org/https://doi.org/10.1016/j.aej.2020.09.013>
- Mardia, K. V., Kent, J. T., & Bibby, J. M. (1995). *Multivariate Analysis* (5th ed.). Academic Press Limited.
- Miranda, J., Maria, J., Filho, D. C., Paulo, A., Vitor, P., Souza, G. De, & Tomasin, S. (2016). A PCA-based approach for substation clustering for voltage sag studies in the Brazilian new energy context. *Electric Power Systems Research*, 136, 31–42. <https://doi.org/10.1016/j.epsr.2016.02.012>
- Nhu, V.-H., Samui, P., Kumar, D., Singh, A., Hoang, N.-D., & Tien Bui, D. (2020). Advanced soft computing techniques for predicting soil compression coefficient in engineering project: a comparative study. *Engineering with Computers*, 36(4), 1405–1416. <https://doi.org/10.1007/s00366-019-00772-7>
- Pearson, K. (1901). *On Lines and Planes of Closest Fit to Systems of Points in Space*. University College. https://books.google.com.br/books?id=uGt_YgEACAAJ
- Saxena, A., & Goebel, K. (2008). *Turbofan Engine Degradation Simulation Data Set*. NASA Ames Prognostics Data Repository, NASA Ames Research Center, Moffett Field, CA. <http://ti.arc.nasa.gov/project/prognostic-data-repository>
- Saxena, A., Goebel, K., Simon, D., & Eklund, N. (2008). Damage propagation

- modeling for aircraft engine run-to-failure simulation. *2008 International Conference on Prognostics and Health Management*, 1–9. <https://doi.org/10.1109/PHM.2008.4711414>
- Song, J., & Li, B. (2021). Nonlinear and additive principal component analysis for functional data. *Journal of Multivariate Analysis*, 181, 104675. <https://doi.org/https://doi.org/10.1016/j.jmva.2020.104675>
- Velasco, J. A., Amaris, H., & Alonso, M. (2020). Deep Learning loss model for large-scale low voltage smart grids. *International Journal of Electrical Power & Energy Systems*, 121, 106054. <https://doi.org/https://doi.org/10.1016/j.ijepes.2020.106054>
- Visinescu, L. L., & Evangelopoulos, N. (2014). Orthogonal rotations in latent semantic analysis: An empirical study. *Decision Support Systems*, 62, 131–143. <https://doi.org/10.1016/J.DSS.2014.03.010>
- Wang, F.-K., & Chien, T.-W. (2010). Process-oriented basis representation for a multivariate gauge study. *Computers & Industrial Engineering*, 58(1), 143–150. <https://doi.org/https://doi.org/10.1016/j.cie.2009.10.001>
- Xu, H., Fard, N., & Fang, Y. (2020). Time series chain graph for modeling reliability covariates in degradation process. *Reliability Engineering & System Safety*, 204, 107207. <https://doi.org/https://doi.org/10.1016/j.ress.2020.107207>
- Yu, Y., Peng, M., Wang, H., Ma, Z., & Li, W. (2020). Improved PCA model for multiple fault detection, isolation and reconstruction of sensors in nuclear power plant. *Annals of Nuclear Energy*, 148, 107662. <https://doi.org/https://doi.org/10.1016/j.anucene.2020.107662>

Chapter 7

Gaussian Mixture Models and Their Applications to Fault Diagnosis in Electric Motors

Chapter details

Chapter DOI:

<https://doi.org/10.4322/978-65-86503-88-3.c07>

Chapter suggested citation / reference style:

Ribeiro Junior, Ronny F., et al. (2022). “Gaussian Mixture Models and Their Applications to Fault Diagnosis in Electric Motors”. In Jorge, Ariosto B., et al. (Eds.) *Uncertainty Modeling: Fundamental Concepts and Models*, Vol. III, UnB, Brasilia, DF, Brazil, pp. 195–220. Book series in Discrete Models, Inverse Methods, & Uncertainty Modeling in Structural Integrity.

P.S.: DOI may be included at the end of citation, for completeness.

Book details

Book: Uncertainty Modeling: Fundamental Concepts and Models

Edited by: Jorge, Ariosto B., Anflor, Carla T. M., Gomes, Guilherme F., & Carneiro, Sergio H. S.

Volume III of Book Series in:

Discrete Models, Inverse Methods, & Uncertainty Modeling in Structural Integrity

Published by: UnB City: Brasilia, DF, Brazil Year: 2022

DOI: <https://doi.org/10.4322/978-65-86503-88-3>

Gaussian Mixture Models and Their Applications to Fault Diagnosis in Electric Motors

Ronny Francis Ribeiro Junior^{1*}, Fabricio Alves de Almeida²,
Ariosto Bretanha Jorge³ and Guilherme Ferreira Gomes⁴

¹Mechanical Engineering Institute, Federal University of Itajubá (UNIFEI), Itajubá, Brazil – ronnyjr@unifei.edu.br

²Institute of Electrical Systems and Energy, Federal University of Itajubá (UNIFEI), Itajubá, Brazil – fabricio-almeida@unifei.edu.br

³Post-Graduate Program - Integrity of Engineering Materials, University of Brasilia, Brazil. E-mail: ariosto.b.jorge@gmail.com

⁴Mechanical Engineering Institute, Federal University of Itajubá (UNIFEI), Itajubá, Brazil – guilhermefergom@unifei.edu.br;

*Corresponding author.

Abstract

This chapter presents the theoretical and practical aspects of Gaussian mixture models method. The introduction section describes the motivation for using the Gaussian mixture model. Next, the historical development, concepts, parameters and equations about Gaussian mixture model are presented. Following, a case study on fault classification in electric motors using Gaussian mixture model is described. Finally, the conclusion section ends the chapter.

Keywords: Gaussian mixture model; fault; electric motors; artificial intelligence; vibration; structural health monitoring.

1 Introduction

Clustering is a useful technique for finding structure in a dataset. From the statistical point of view, clustering methods may be divided into probability model-based approaches and nonparametric approaches. The probability model-based approach assumes that the data set follows a mixture model of probability (Yang *et al.*, 2012).

So, any model that represents a dataset with a mix of two or more probability distributions is known as a mixture model. Mixture models aid in cluster identification by fitting a finite number of probability distributions to the data and repeatedly adjusting the parameters of those distributions until they best match the underlying data. Mixture models also aid in describing the presence of subpopulations within an aggregate population that do not require an observed data set to identify the subpopulation to which an individual observation belongs (known as

unsupervised learning in machine learning). In SHM, mixture models help to identify whether there is more than one failure mode in the data.

In real applications, many datasets may be described by Gaussian distributions. As a result, it is reasonable and straightforward to infer that the clusters are derived from various Gaussian distributions. In other words, the dataset is attempted to be modeled as a blend of various Gaussian distributions. Therefore, Gaussian mixture modeling, which fits Gaussian (or normal) distributions to data, is the most frequent type of mixture modeling.

By the end of this chapter, we hope you'll have an understanding of how the Gaussian mixture model works. We will also present an example of how to apply the GMM to a real case to help the reader to understand how to use mixture model for clustering problems.

2 Gaussian Mixture Model

Gaussian mixture models (GMMs) are statistical models that fit data to a set of stochastic models. These models, as GMMs, fit Gaussian distributions to a collection of data in order to split the data into discrete clusters. One of the benefits of mixed model clustering is that it is a soft clustering strategy (Rhys, 2020). Because it is a soft method, the likelihood of each example belonging to a cluster may be quantified. As a result, more information may be obtained to make better judgments. For example, if certain data has 49 percent of belonging to cluster A and 51 percent of belonging to cluster B, how confident can we declare that it belongs to cluster A or B? As a result, each Gaussian in the mixture model represents a possible cluster. Once our Gaussian mixture best matches the data, we can compute the likelihood of each case belonging to each cluster and allocate instances to the most likely cluster.

2.1 Historical Development

The first instance of Gaussian mixture models being used is usually attributed to Karl Pearson (1894). Pearson noted non-normal patterns of forehead to body length ratios in female shore crab populations. Therefore, the author used not only one normal to model the data, but two. This decision was motivated by the suspicion that asymmetries in the histograms of these length ratio could indicate an evolutionary divergence. By adjusting the mixture's parameters, Pearson could make a model that is able to identify two distinct species of shore crab. Despite being successful, the mathematical formulation for the period was an extremely difficult challenge, making the method not widely used.

Subsequently, Wolfe (1963,1965,1970) put a lot of effort into developing Gaussian mixture models for clustering. In his work, Wolfe developed an interactive scheme to estimate the parameters of Gaussians, something innovative for the period. As the main challenge of mixing models is to estimate the parameters of the Gaussians, this step was extremely important for the development of the method.

Finally, the mixture model methods were popularized by Duda and Hard (1973). The authors in their work introduced the maximum likelihood estimation (MLE) technique. The MLE uses an observed data to estimate the parameters of a probability distribution (Myung, 2003). With this technique allied to the beginning of modern computing, it was possible to develop the mixture models.

2.2 Parameters of Gaussian Mixture Models

To define a one-dimensional Gaussian distribution, two parameters (mean and variance) are required. Before beginning the process, it is required to define the number of clusters that are expected to be identified within the dataset; if this number is unknown, it is feasible to approximate it using various approaches that will be detailed in section 2.5.

For now, let's suppose the dataset contains only two distinct clusters (k and j). The GMM begins by picking random values for the mean and variance of the two Gaussian distributions. The likelihood of belonging to one of the two clusters is then estimated. The Bayes' theorem is the most often used method for performing this computation (Rhys, 2020). Using Bayes' theorem, we get (Kolmogorov & Bharucha-Reid, 2018):

$$p(k|x) = \frac{p(x|k)p(k)}{p(x)} \quad (1)$$

In this case, $p(k|x)$ represents the probability of case x belonging to Gaussian k ; $p(x|k)$ represents the probability of observing case x if you sampled from Gaussian k ; $p(k)$ represents the probability of a randomly selected case belonging to Gaussian k ; and $p(x)$ represents the probability of drawing case x if you sampled from the entire mixture model as a whole. The evidence, $p(x)$, is thus the probability of pulling case x from either Gaussian distribution.

However, in the case study we need to calculate the probability of belonging to either cluster. therefore, we need to add to the formula the probability of each event occurring independently. thus, Equation 1 can be rewritten for the case of two cluster (k and j), as shown in Equation 2 and 3 (Rhys, 2020):

$$p(k|x_i) = \frac{p(x_i|k)p(k)}{p(x_i|k)p(k) + p(x_i|j)p(j)} \quad (2)$$

$$p(j|x_i) = \frac{p(x_i|j)p(j)}{p(x_i|k)p(k) + p(x_i|j)p(j)} \quad (3)$$

All that remains now is to compute the likelihood and prior. The Likelihood function of the Gaussian distribution shows us the relative likelihood of drawing a case with a certain value from a Gaussian distribution with a specific mean and variance combination. Equation 4 depicts the probability density function for the Gaussian distribution k :

$$p(x_i|k) = \frac{1}{\sqrt{2\Pi\sigma_k^2}} e^{-\frac{(x_i-\mu_k)^2}{2\sigma_k^2}} \quad (4)$$

Where, μ_k and σ_k^2 are the mean and variance of the k Gaussian.

At the start of the process, the prior probabilities, like the Gaussian means and variances, are generated at random. However, for subsequent calculations, the iterative procedure of this

method must be followed. Expectation Maximization is used for this. Figure 1 shows the convergence process of a one-dimensional GMM.

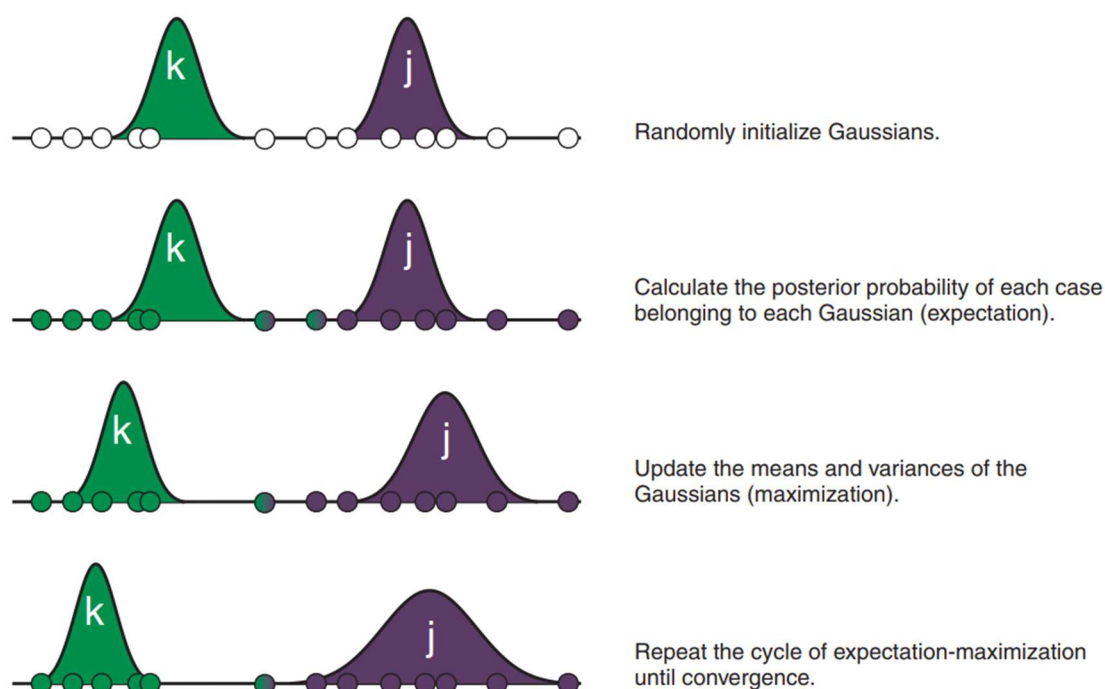


Figure 1 – Convergence process of a one-dimensional Gaussian mixture model (adapted from Rhys, 2020).

2.3 Expectation-maximization

Maximum likelihood estimation is a method of estimating a dataset's density by exploring across probability distributions and their parameters. When there are variables that interact with those in the dataset but are hidden or not observed, maximum likelihood becomes intractable (Murphy, 2012). These variables are also called latent variables.

The expectation-maximization algorithm is a method for calculating maximum likelihood when latent variables are present. This is accomplished by first estimating the values of the latent variables (E-step), then optimizing the model (M-step), and finally repeating these two stages until convergence is achieved. It is a powerful and versatile method that is often used in density estimation with missing data, such as clustering algorithms like the Gaussian mixture model (Hastie *et al.*, 2009).

The EM method may be characterized by the following equations given a set of observable data (U), a set of latent data (J), and a vector of unknown parameters (θ), along with likelihood function $L(\theta, U, J) = p(U, J | \theta)$ (Moon, 1996; Dellaert, 2002).

Expectation step (E step)

$$Q(\Theta | \Theta^{(t)}) = E_{J|U, \Theta^{(t)}}[\log L(\Theta; U, J)] \quad (5)$$

Maximization step (M step)

$$\Theta^{(t+1)} = \underset{\Theta}{\operatorname{argmax}} Q(\Theta | \Theta^{(t)}) \quad (6)$$

The algorithm's iterative procedure begins with a random value of θ . The probability of each potential value of J is then computed, given θ . The new θ parameters are estimated using these values. The procedure is repeated until the algorithm achieves convergence. A maximum number of iterations or a stopping condition is typically employed to determine an algorithm's convergence. There are various alternatives for the halting criterion. The most common is to halt the algorithm due to a lack of progress in the likelihood function (Mc Nicholas, 2016), which means to stop the process when:

$$l^{k+1} - l^k < \varepsilon \quad (7)$$

Where l is the likelihood function value, and k denotes the number of the iteration. The value of ε is defined for the sake of the problem under consideration.

2.4 Multivariate clustering problems

All the development in the previous sections has been with a one-dimensional GMM in mind. However, most of the time, we will come across datasets with multiple variables. Therefore, there is a version of GMM for multiple variables. One-dimensional GMMs are defined by two parameters: mean and variance. Multivariate GMM is defined by two new parameters: the centroid and the covariance matrix.

In a nutshell, the centroid is a vector of means that represents an average for each dimension in the dataset. The covariance matrix is a square matrix whose elements are the covariance between the dataset's variables. The covariance matrix is often a non-standardized number. However, you may normalize this matrix by dividing the value by the product of the standard deviations of the variables.

Therefore, the equations change to the multivariate case. In general, what happens is a generalization of the one-dimensional case to more dimensions. Equations 8 to 10 describe the calculation of the centroid, covariance matrix and likelihood for the multivariate case (Rhys, 2020). Figure 2 shows the GMM algorithm for a bivariate case.

$$\mu_{k,a} = \sum_{i=1}^n \left(\frac{p(k|x_i)}{n \times p(k)} \right) x_{i,a} \quad (8)$$

$$(\sigma_k)_{a,b} = \sum_{i=1}^n \left(\frac{p(k|x_i)}{n \times p(k)} \right) (x_{i,a} - \mu_{k,a})(x_{i,b} - \mu_{k,b}) \quad (9)$$

$$p(x_i|k) = \frac{1}{\sqrt{(2\pi)^p |\Sigma_k|}} e^{-0.5(\vec{x}_i - \vec{\mu}_k)^T \Sigma_k^{-1} (\vec{x}_i - \vec{\mu}_k)} \quad (10)$$

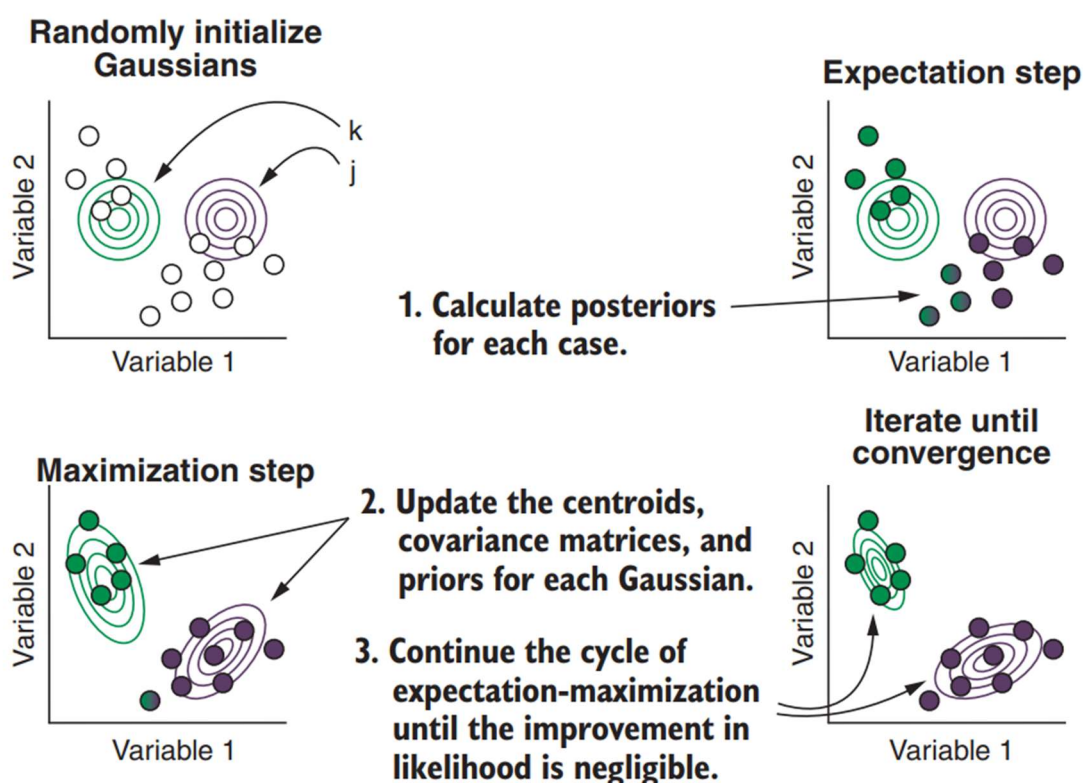


Figure 2 – Gaussian mixture model algorithm for a bivariate case (adapted from Rhys, 2020).

2.5 Number of components in Gaussian Mixture Model

Choosing the number of components simply selects the number of Gaussians for your model. However, when you have no prior knowledge of your dataset, it is necessary to use techniques to assist in this choice, i.e., model selection. In short, a fitting process is performed with several models on a given dataset, and one is chosen among all the others. It is worth mentioning that this is an alternative when there is no information about the dataset. If you know that there are three different types of clusters and the model selection indicates a different number, it is recommended that you use the three and not the one given by the model. There are several ways to choose the number of components in a GMM; even the trial-and-error method can be used. However, there are more optimized methods for this task. In this section, the main methods used for model selection will be discussed.

2.5.1 Akaike Information Criterion

Hirotsugu Akaike created the Akaike information criterion (AIC) in 1971 (Akaike, 1974). AIC is a statistical estimator that helps in the selection of the optimal model for a given dataset. The lower the AIC score for a dataset, the better the fit for the data (Profillidis & Botzoris, 2018). Because it is a comparative method, it can only compare the score value for the same dataset. The AIC formula is defined by Equation 11 (Hastie et al., 2009).

$$AIC = -\frac{2}{N} \cdot LL + 2\frac{k}{N} \quad (11)$$

Where N is the number of examples in the training dataset, LL is the log-likelihood of the model on the training dataset, and k is the number of parameters in the model.

2.5.2 Bayesian Information Criterion

The Bayesian information criterion (BIC) was developed by Gideon Schwarz in 1978 (Schwarz, 1978). Like the AIC, the BIC is a statistical estimator for selecting a model that best fits the dataset. The main difference between BIC and AIC is that BIC penalizes more complex models more sharply. This makes simpler models tend to have a lower score, making them more likely to be chosen. BIC is mathematically defined by Equation 12 (Hastie *et al.*, 2009).

$$BIC = -\frac{2}{N} \cdot LL + \log(N) \cdot k \quad (12)$$

Where $\log()$ has the base- e called the natural logarithm, LL is the log-likelihood of the model, N is the number of examples in the training dataset, and k is the number of parameters in the model.

2.5.3 Minimum description length

The Minimum description length (MDL) concept is a strong inductive inference approach that serves as the foundation for statistical modeling, pattern recognition, and machine learning. The MDL idea was first suggested in 1978 by Rissanen (Rissanen, 1978). The notion is that with the assistance of a model, every data set can be suitably encoded, and that the more we compress the data by eliminating redundancy from it, the more we find underlying regularities in the data. As a result, code length is directly connected to the model's generalization capabilities, and the model that gives the shortest description of the data should be chosen (Myung, 2001). Like the other two models, the lowest score is chosen. MDL is calculated as Equation 13 (Hastie *et al.*, 2009).

$$MDL = -\log(P(\theta)) - \log(P(y|\theta)) \quad (13)$$

Where $\log()$ has the base- e called the natural logarithm, θ is the model parameters, y is the target values.

2.5.4 Silhouette Coefficient

Silhouette coefficient is a method to estimate distance between cluster, proposed by Peter Rousseeuw in 1987 (Rousseeuw, 1987). The silhouette coefficient compares the similarity of an object to its own cluster (cohesion) to other clusters (separation). The silhouette has a value between -1 and +1, with a high value indicating that the object is well fitted to its own cluster and a low value indicating that the object is poorly adapted to surrounding clusters. The

clustering setup is useful if the majority of the items have a high value. If a large number of points have a low or negative value, the clustering configuration may have too many or too few clusters. The coefficient is defined for each sample and is defined by Equation 14.

$$s = \frac{b - a}{\max(a, b)} \quad (14)$$

where a is the mean distance between a sample and all other points in the same class and b is the mean distance between a sample and all other points in the next nearest cluster.

3 Case study: Motor Fault Classification

In this section we will present a practical case where the Gaussian mixture model was used to cluster, identify, and diagnose six different failures in electric motor. We hope that this example will help to better understand the usefulness of the GMM.

3.1 Experimental Setup

All tests were carried out at PS Soluções in Itajuba (Brazil). The test bench (Figure 1) employed a 0.5 HP induction motor that was directly linked to the SpectraQuest machine failure simulator (Ribeiro Junior et al., 2020). Accelerometer 1 detects vibration in the y direction, whereas accelerometer 2 detects vibration in the x direction. Table 1 also illustrates the motor's basic characteristics at 60 and 50 Hz.

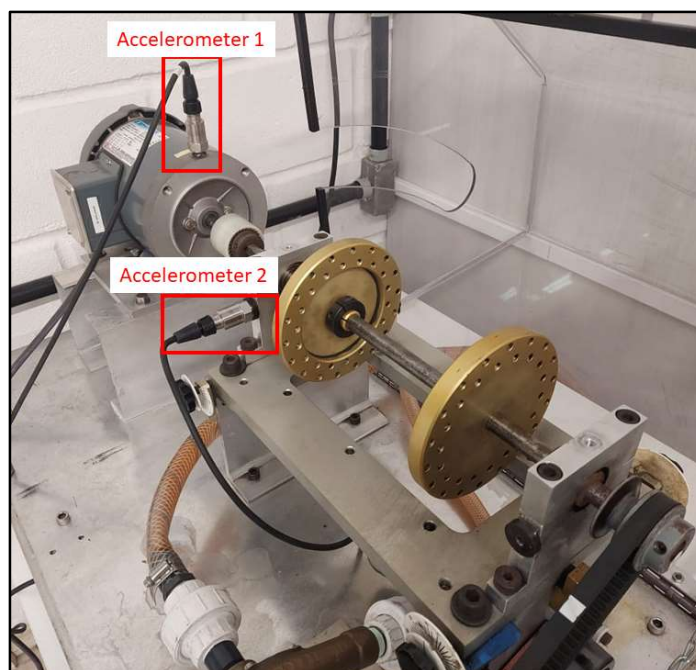


Figure 3. Experimental setup (adapted from Ribeiro Junior et al., 2022).

Table 1. Marathon motor nameplate (adapted from Ribeiro Junior et al., 2022).

Parameter	60 Hz	50 Hz
Power (HP)	0.50	0.33
Rotation (RPM)	3450	2850
Voltage (V)	208-230/460	190/380
Amperage (A)	2.1-2.2/1.1	2.0/1.0
Service Factor	1.15	1.15

The experimental bench is made up of five electric motors that may test up to seven operational conditions. These conditions include normal, bent shaft, broken bar, misalignment, mechanical looseness, bearing problem, and imbalanced (Figure 4).

- **Normal** condition is the electric motor in a healthy state, without any fault.
- **Bent shaft** refers to the motor whose shaft is slightly bent (approximately 4 mm), inducing a scratch between the rotor and the stator as shown in Figure 4(a). This fault mainly occurs due to overloads and coupling conditions.
- **Broken bar** refers to the motor whose squirrel cage rotor bar is broken. This fault is more common on bigger electric motors. As shown in Figure 4(b) this fault is simulated by drilling 6 holes with 2 mm diameter in the rotor.
- **Misalignment** refers to the motor whose shaft is misaligned. This fault occurs due to bad installation or overloads. Figure 4(c) shows how this defect is induced. On the bench, it is possible to generate up to 3 mm of misalignment
- **Mechanical looseness** refers to the motor whose screws are loosening. This fault commonly occurs due to bad installation but high vibration may cause this fault. Figure 4(d) shows which screws are loosened.
- **Bearing fault** refers to the motor whose bearing is damaged. This is the most common fault in electric motors. Figure 4(e) shows the bearing fault in the outer race. The bearing used is SKF-6203.
- **Unbalanced** refers to the motor whose load is unbalanced. This fault is very common in electric motors, being the reason for other faults. Figure 4(f) shows how this fault is induced.

These conditions cover the most frequent electric motor failures. Several investigations confirm that the selected failures are the most frequently defects in electric motors (Randal, 2011; Plante et al., 2016; Zarei et al, 2014; Choudhary et al., 2019; ABRAMAN, 2011).

In all, four rounds of tests with all operational conditions were done, gathering 28 signals from each accelerometer. To properly analyze the faults, the defects were removed and installed in different motors for each run.

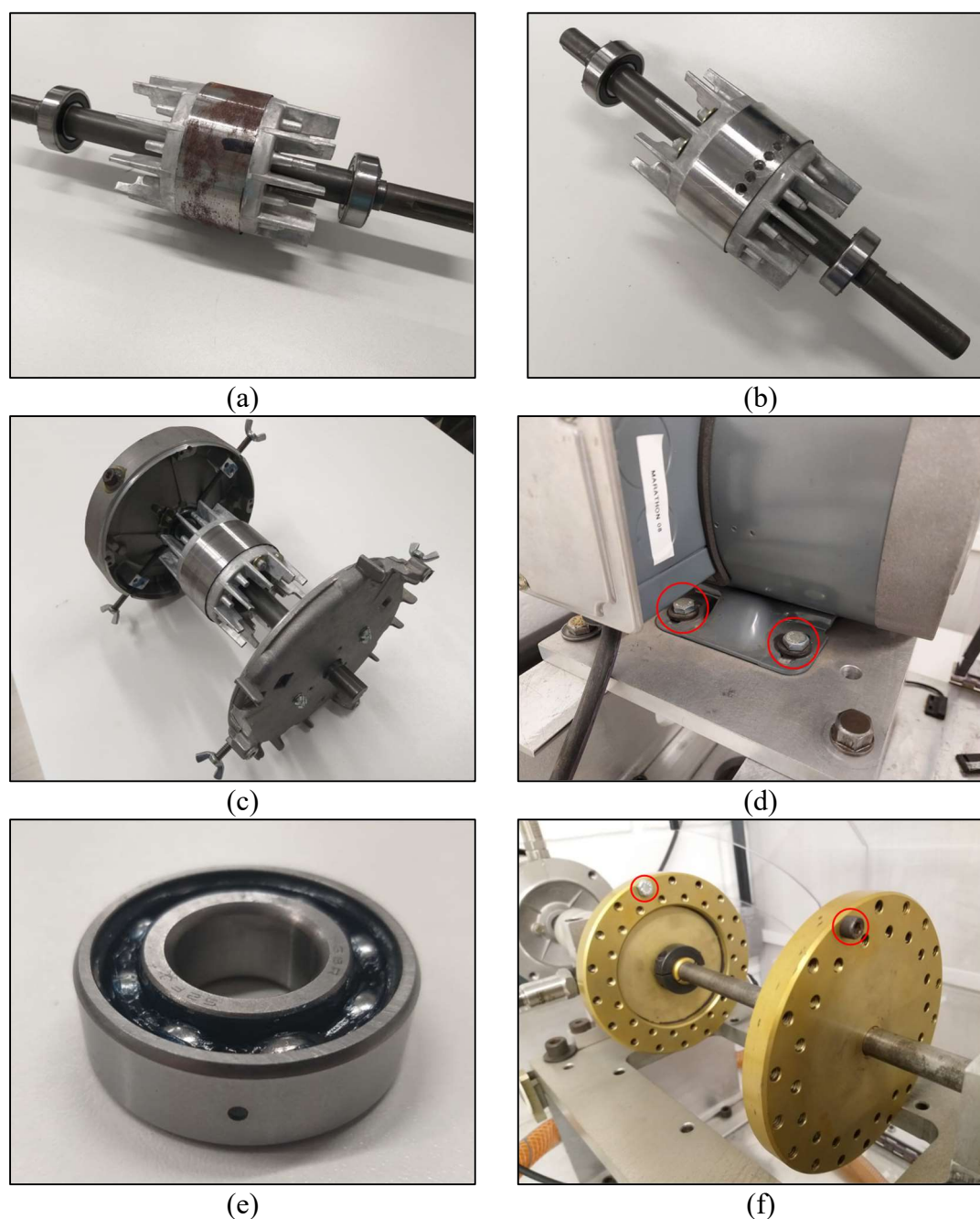


Figure 4. Fault types: (a) Bent shaft, (b) Broken bar, (c) Misalignment, (d) Mechanical looseness, (e) Bearing fault, (f) Unbalanced (adapted from Ribeiro Junior et al., 2022).

3.2 Vibration signal acquisition

The IMI uniaxial accelerometer records the two vibration signals. The accelerometer has a frequency range of ± 3 dB, a measuring range of ± 50 g, a sensitivity of 100 mV/g, and resonance at a frequency of 25 kHz. Because the frequency of the recorded signals is low and the engine mass is significantly more than the mass of the accelerometer, these facts make the sensor suitable for vibration measurement. A threaded base links the first accelerometer to the motor. The second is also kept in place by a threaded base that is put into the bearing housing. Araldite was used to bind all threaded bases. The acquisition system (Preditor®) gathers 262144 sample points at a sampling frequency of 11718.5 Hz and an acquisition window of 22.37 seconds.

3.3 Procedures of the Proposed Method

The primary step is to gather all vibration signals from various operational circumstances. For each operational state, we gathered four signals, 75 percent for training and 25 percent for testing. Following that, the signals are divided into smaller sizes, each with 512 sample points. Finally, we employ RMS and Skewness statistics to extract features from vibration signals in order to improve classification results. The RSM and Skewness measures were chosen because of their significance and capacity to aggregate information, as proved by Ribeiro Junior et al (2020). Other statistics employed were Kurtosis, SRA, Peak-to-Peak, Crest factor, and Shape factor. However, none of them performed as well as the duo of RMS and Skewness. The main parameter of GMM is the covariance matrix, being possible to choose 4 different types:

- Full: each component has its own general covariance matrix
- Tied: all components share the same general covariance matrix
- Diag: each component has its own diagonal covariance matrix
- Spherical: each component has its own single variance

Because each data set can have a best-fitting parameter, the optimal form of covariance will be picked. The average of two random variables' degrees of dependency or linear numerical interrelationship is their covariance or variance. Choosing the best covariance type for your dataset can help the GMM model perform better.

Another method is to begin the process supervised, that is, with the mean of each distribution as the initial weight. This reduces the number of iterations required for optimizers to converge while also avoiding incorrect convergence. Because the model was begun supervised, the number of mixing components will be limited to two (healthy and a faulty motor). The number of iterations was limited to 10,000 with a 10^{-6} convergence criterion.

The Mahalanobis distance is then computed. The Mahalanobis distance is the distance between two points in multivariate space. Variables in a normal Euclidean space are represented by axes drawn at right angles to each other. To measure the distance between two points, use a ruler. For uncorrelated data, the Euclidean distance equals the Mahalanobis distance. When two or more variables are connected, the axes are no longer at right angles, prohibiting ruler measurements. Figure 5 shows a comparison between the Mahalanobis distance and the Euclidean distance in a two-dimensional space. The elongated cluster's Euclidean distance to its centroid is greater than that of the round cluster. The Mahalanobis distance to the centroid of an elongated cluster, on the other hand, is lower than that of a spherical one. Furthermore, if you have more than three variables, you will be unable to plot them at all in normal 3D space. The Mahalanobis distance, which quantifies distances between points, including correlated points, solves this measurement difficulty. The Mahalanobis distance is a measure of distance relative to the centroid, which may be thought of as the overall mean for the Mahalanobis distance is a measure of distance relative to the centroid, which is the general mean for multivariate data. The centroid in multivariate space is the point at which all meanings from all variables coincide. The bigger the Mahalanobis distance, the greater the distance between the data point and the centroid. The formula version utilized is described by Equation 15 (Hamil et.al, 2016):

$$d_i = [(x_i - \bar{x})^T C^{-1} (x_i - \bar{x})]^{0.5} \tag{15}$$

Where x_i is the observation point, \bar{x} is the mean of the reference cluster (normal condition) and C^{-1} is the inverse covarian matrix of independent variables.

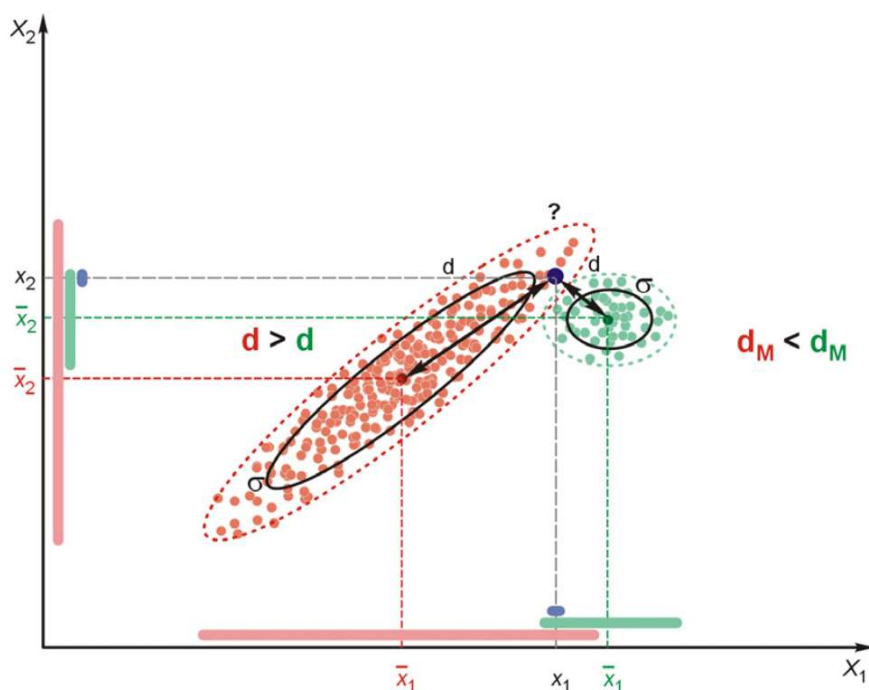


Figure 5. Comparison between Mahalanobis distance and Euclidian distance (adapted from Woźniak et al., 2019).

All of the features in the proposed method were created in Python using the Scikit-learn library. Figure 6 depicts the overall flowchart of the approach followed in this study.

(Intentionally left blank)

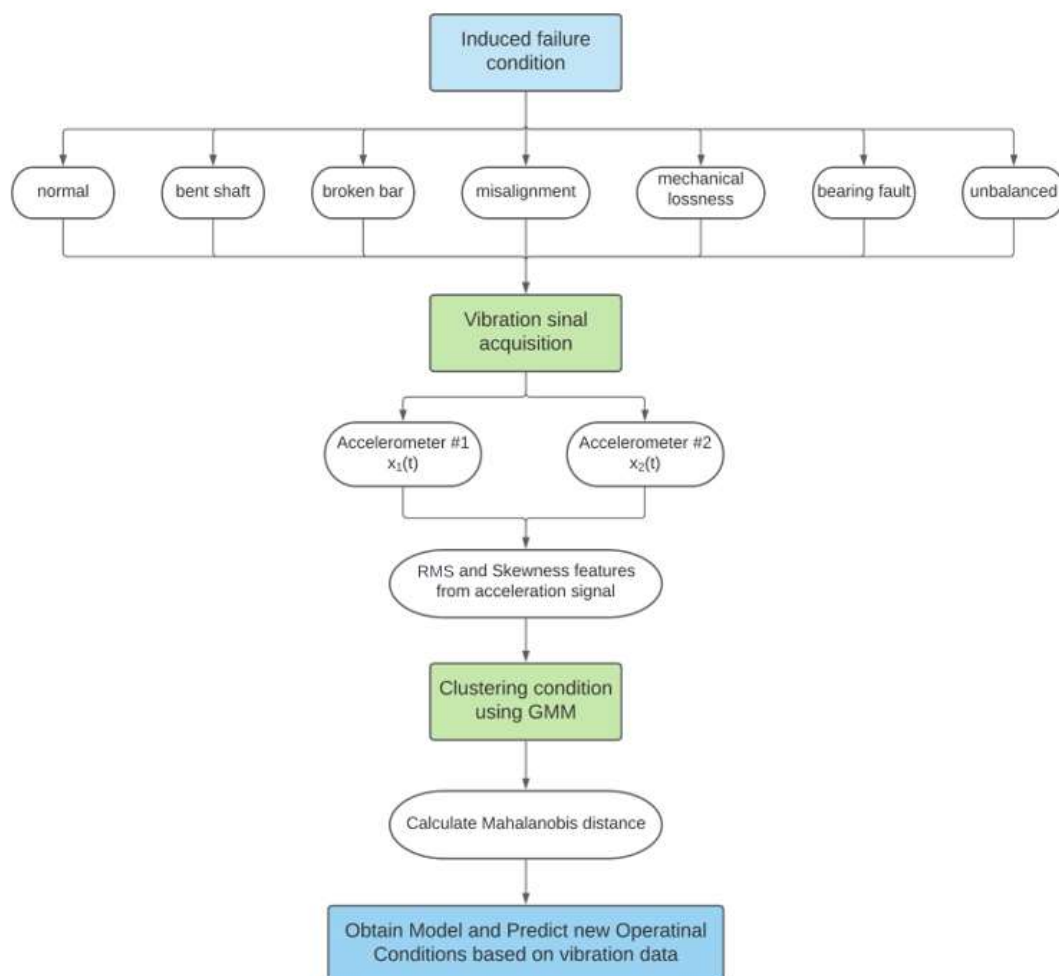


Figure 6. General flowchart of the motor operation condition clustering using GMM (adapted from Ribeiro Junior, 2022).

3.4 Results and Discussion

In this section, we will discuss and exhibit the primary and most important findings from the vibration signal analysis in seven distinct operating conditions (six faults) in electric motors. Figure 7 compares the time domain vibration signals for the six fault scenarios to the accelerometer 1 normal signal. Figure 8 shows the same findings when accelerometer 2 is used. In certain ways, it is obvious that the signal varies under fault circumstances, each with its own distinct pattern. This is especially true for accelerometer 1 and operating circumstances characterized by bearing failure and a lack of mechanical looseness. As a result, only accelerometer signal 1 will be used as a GMM input. Figure 9 also provides the scatter plot data for the accelerometer signal. Except for the mechanical loss working state (Figure 9(d)), when probable nonlinearities are present, a dot pattern is found in all conditions.

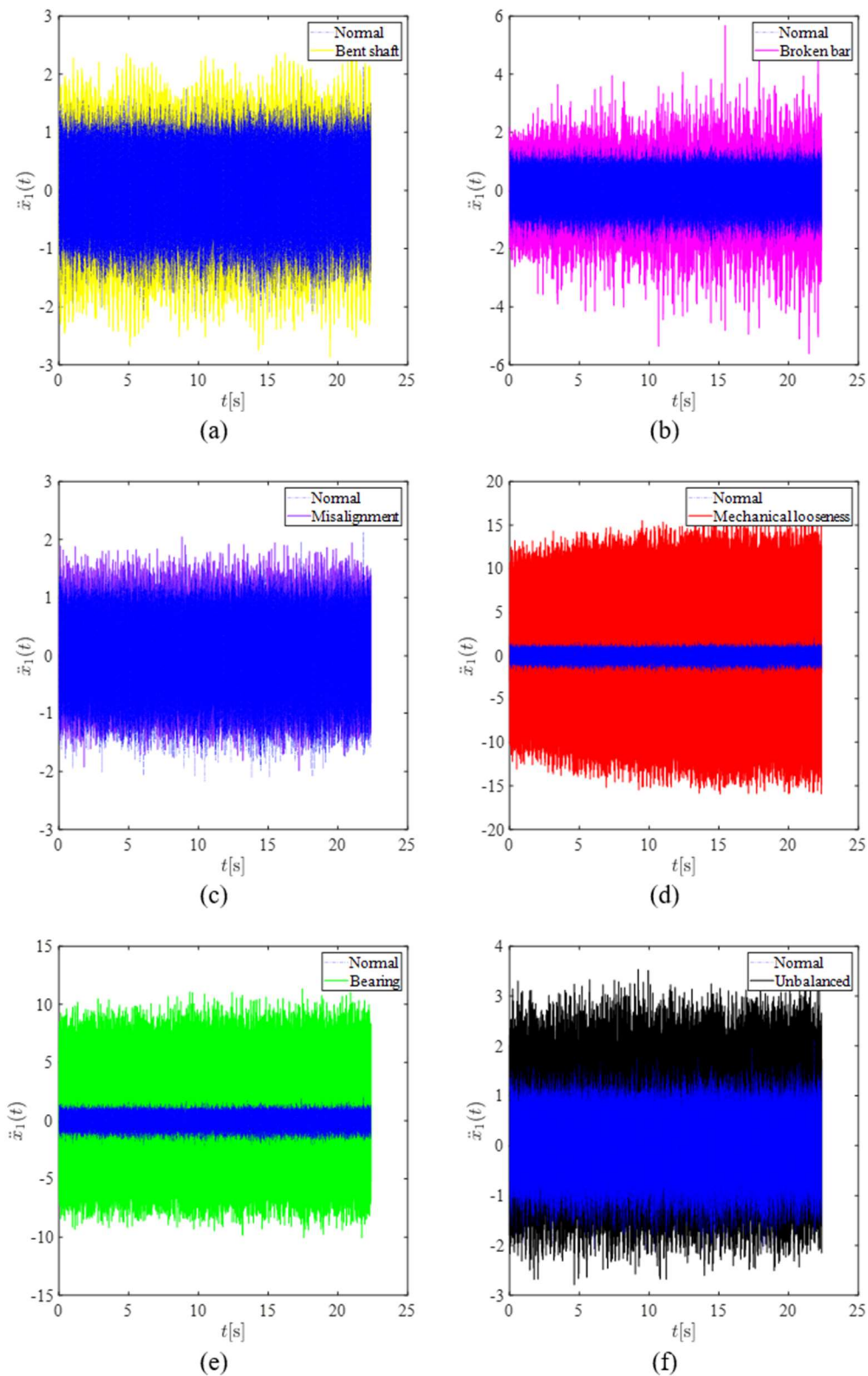


Figure 7. Vibrations signal for the six different fault conditions considering the vibration accelerometer #1 (adapted from Ribeiro Junior, 2022).

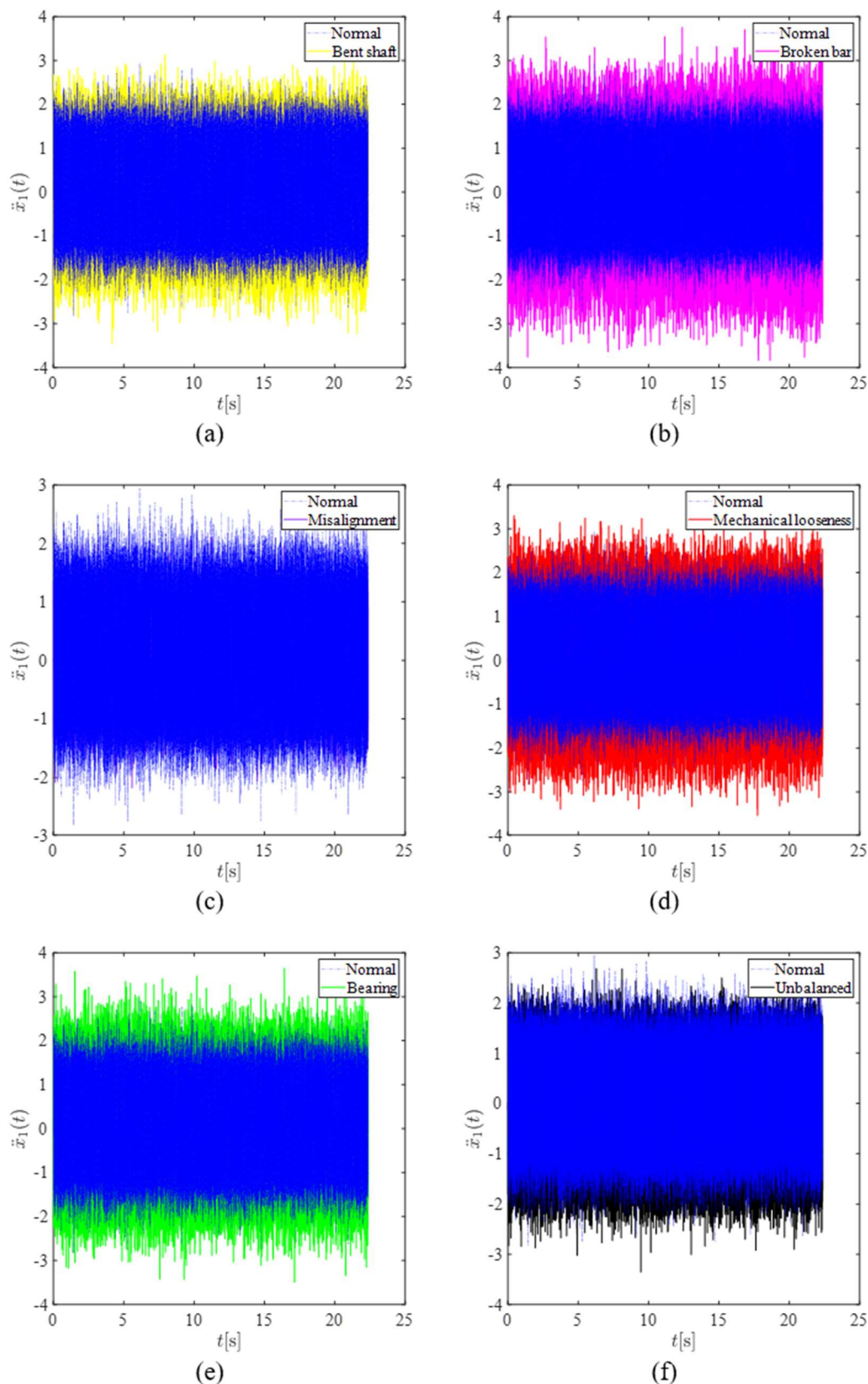


Figure 8. Vibrations signal for the six different fault conditions considering the vibration accelerometer #2 (adapted from Ribeiro Junior, 2022).

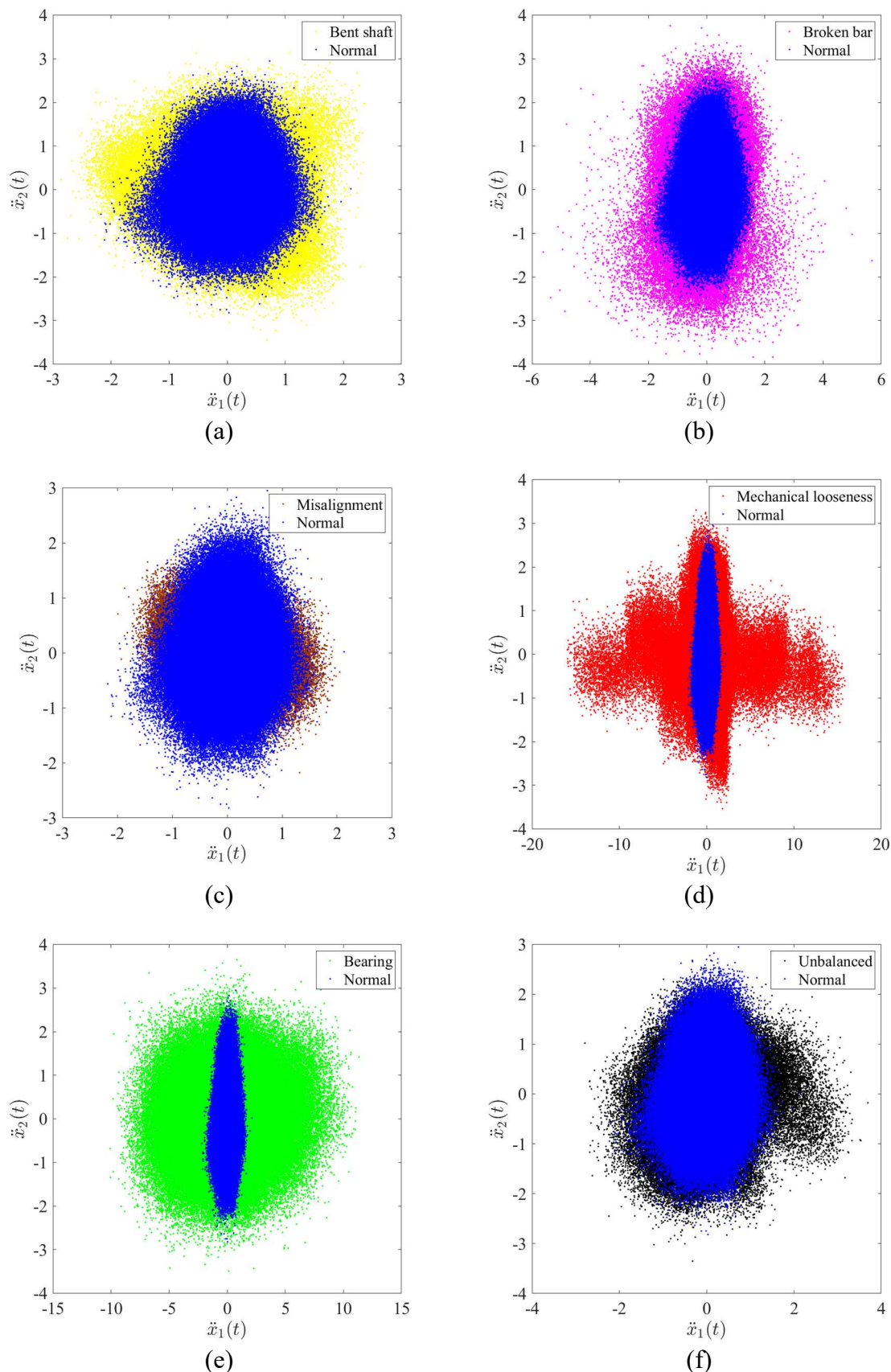


Figure 9. Scatter vibration signal for the six different fault conditions considering the vibration considering both the accelerometers #1 and #2 (adapted from Ribeiro Junior, 2022).

The second step of machine learning was carried out after an exploration phase of vibration data in various operating situations (caused failures). An unsupervised learning method was required during this step. Because it is a flexible approach characterized as "soft clustering," a Gaussian mixing model (GMM) was employed. In other terms, a GMM is a probabilistic model in which all data points are created by a finite number of Gaussian distributions with unknown parameters.

Figures 10–15 illustrate the training results for the GMM algorithm under all operational circumstances. All covariance matrix parameters are displayed in the findings (full, tied, diag, and spherical). For all figures, the \cdot represents training data and the X represents test data.

Some clustering results are more segregated (distant) than others. The clusters are closer together in the case of misalignment. The groups of bent and broken bars are far enough apart yet share a border. In the other cases, the generated groupings are more separated. Because GMM is a soft algorithm, its usage is justified at this stage. When the various forms of covariance matrix are analyzed, the full or diagonal covariance matrix settings are the best for the dataset, with the full option slightly better. This setup has an accuracy of more than 90% for all defects assessed. The findings are comparable to earlier researches (Wang et al., 2009; Yan et al., 2017; Hong et al., 2019; Pani et al., 2020).

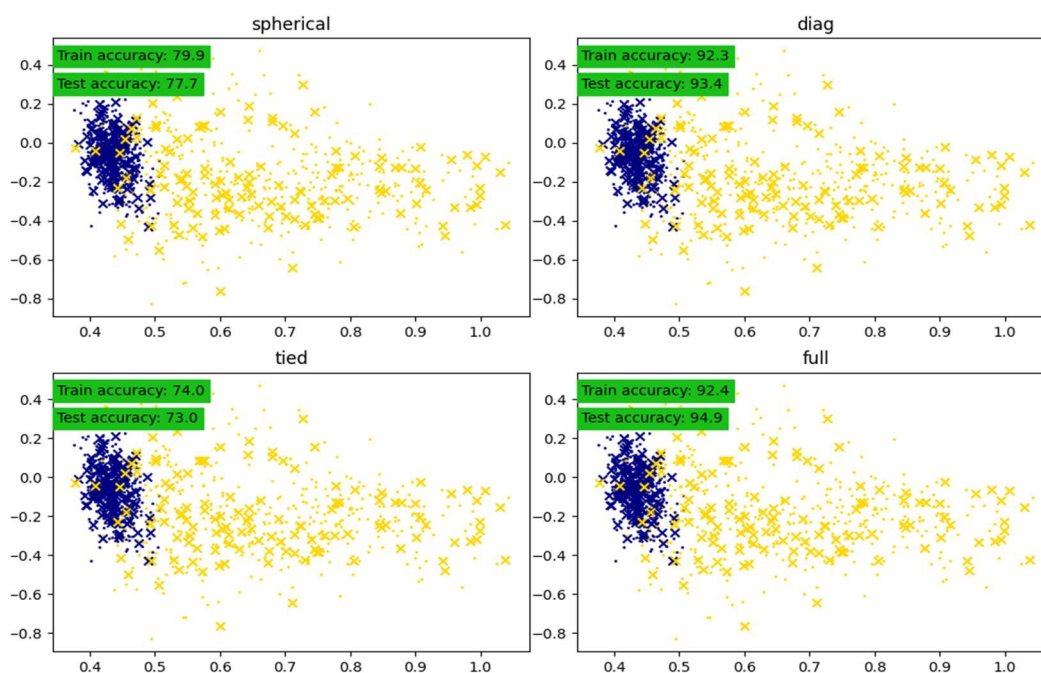


Figure 10. Gaussian mixture model results (RSM vs. Skewness) for the Bent and Normal conditions (adapted from Ribeiro Junior, 2022).

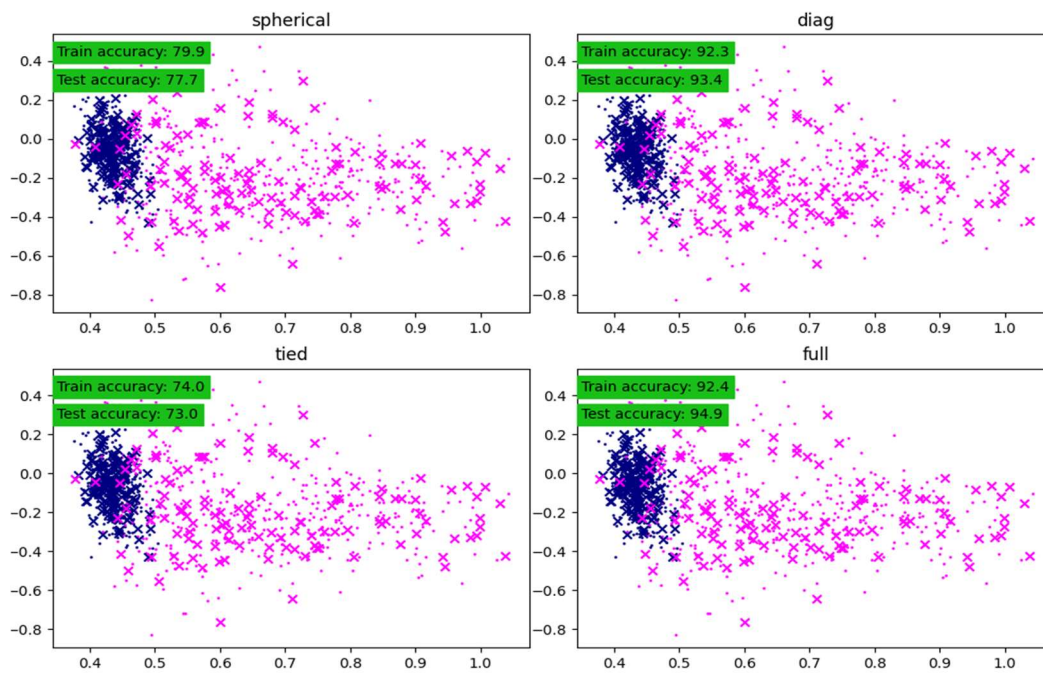


Figure 11. Gaussian mixture model results (RSM vs. Skewness) for the Broken bar and Normal conditions (adapted from Ribeiro Junior, 2022).

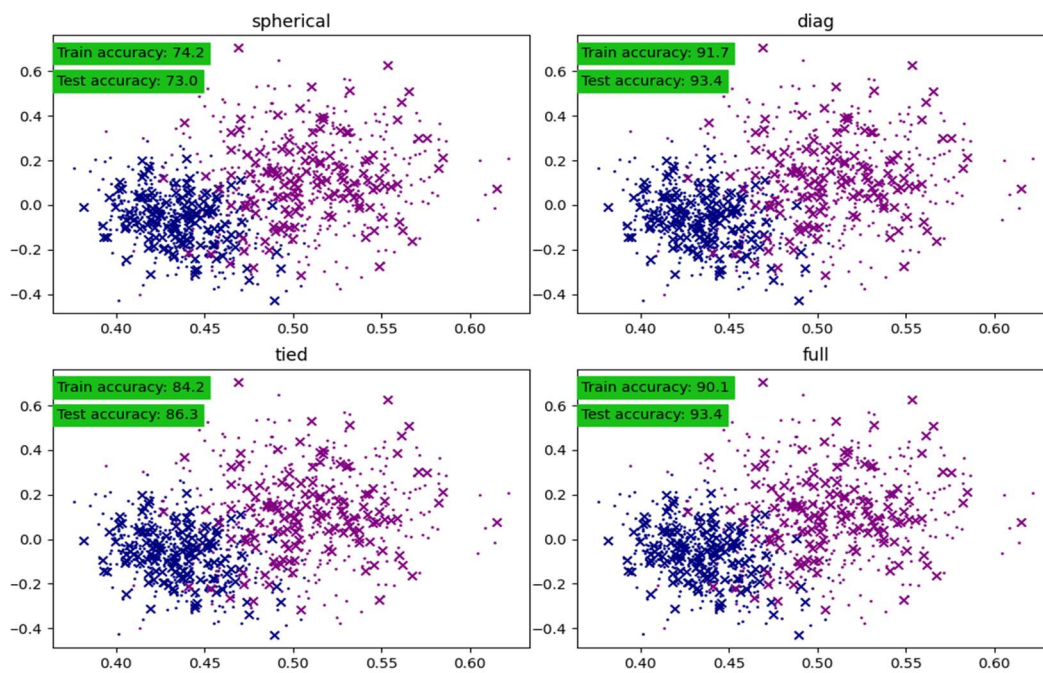


Figure 12. Gaussian mixture model results (RSM vs. Skewness) for the Misalignment and Normal conditions (adapted from Ribeiro Junior, 2022).

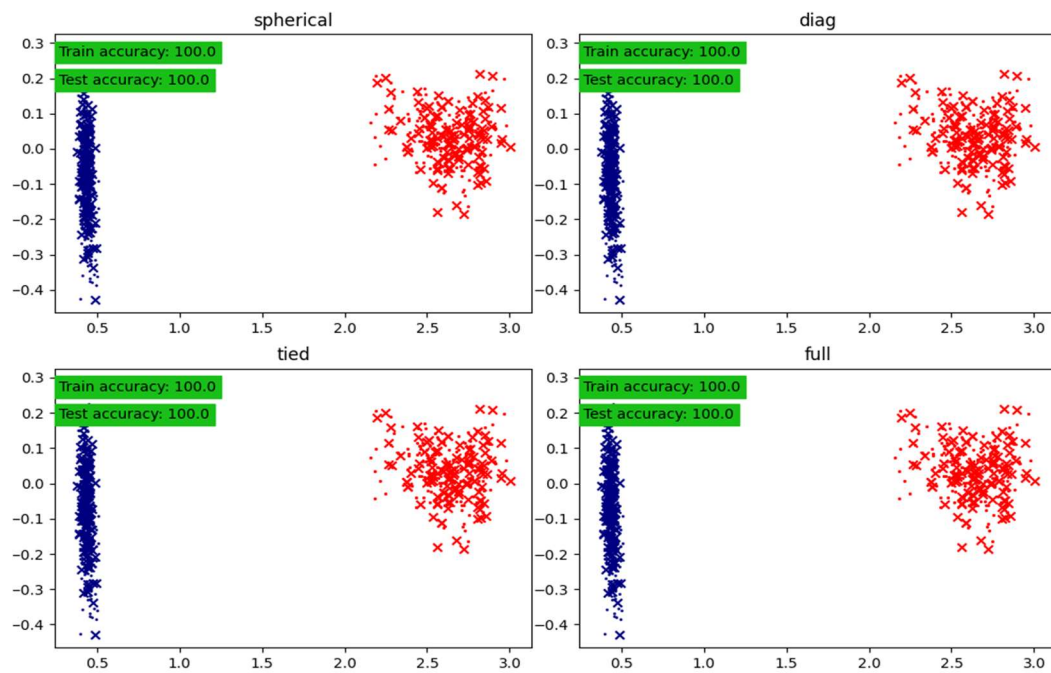


Figure 13. Gaussian mixture model results (RSM vs. Skewness) for the Mechanical looseness and Normal conditions (adapted from Ribeiro Junior, 2022).

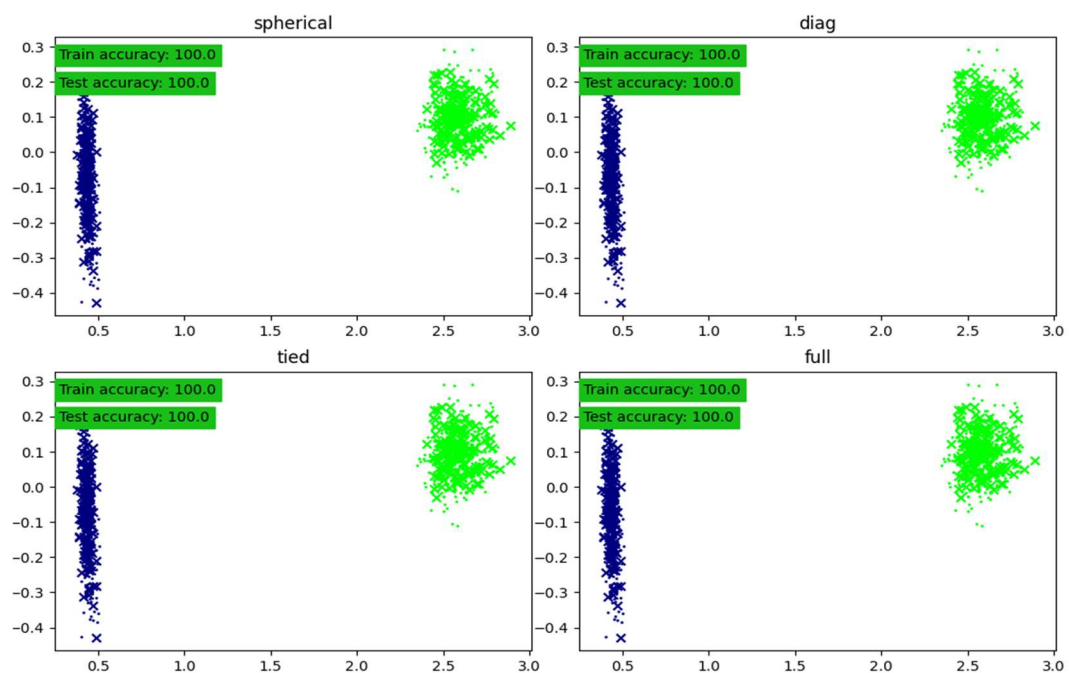


Figure 14. Gaussian mixture model results (RSM vs. Skewness) for the Bearing and Normal conditions (adapted from Ribeiro Junior, 2022).

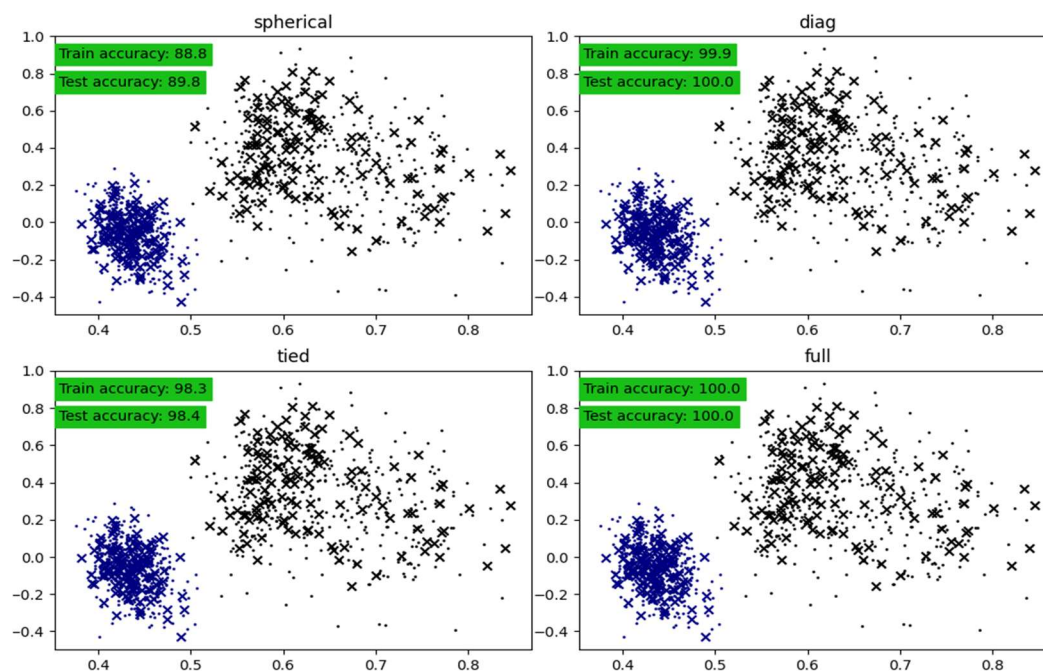
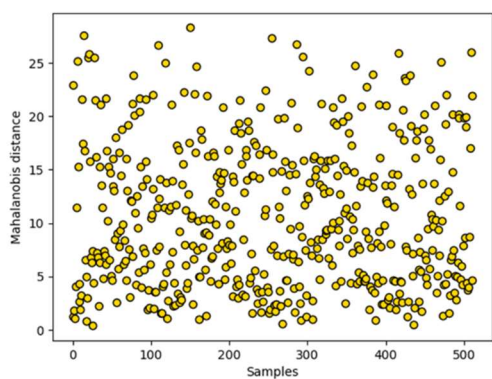


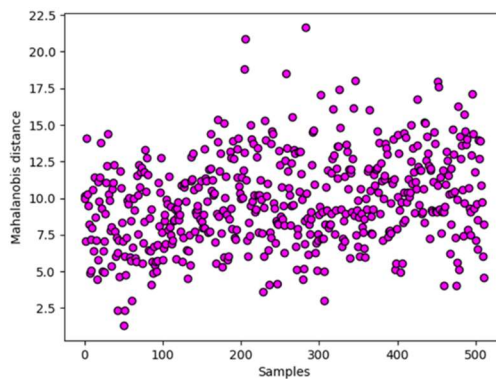
Figure 15. Gaussian mixture model results (RSM vs. Skewness) for the Unbalanced and Normal conditions (adapted from Ribeiro Junior, 2022).

The Mahalanobis distance measure was used to determine how far apart the clusters are. The Mahalanobis distances of the Gaussian distributions adapted to the usual operating state are shown in Figure 16. It was discovered that the mechanical looseness situation produced a certain pattern of distances. Figure 17 organizes all of the distances. It should be noted that mechanical looseness and bearing faults are the most extreme operation condition. This suggests that these are the most easily identified defects. Furthermore, the Mahalanobis distance pattern may be used to distinguish one fault from another, i.e., identify which fault it is. This is not achievable with GMM alone since it simply detects the presence of a fault. Figure 17 shows that each defect has a pattern and various distance value ranges. As a result of integrating the GMM with the Mahalanobis distance, it is able to not only identify but also discriminate between faults.

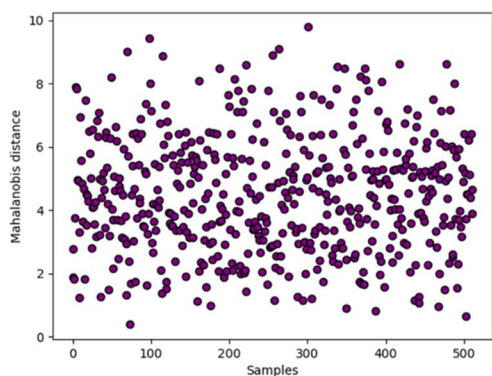
(Intentionally left blank)



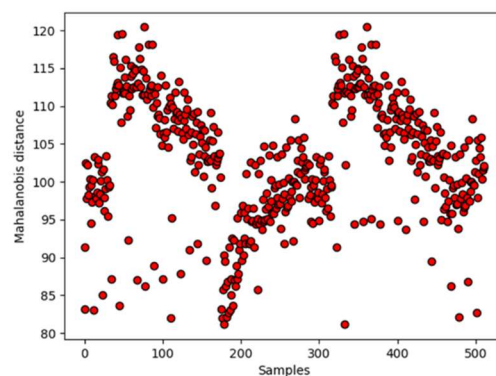
(a) bent



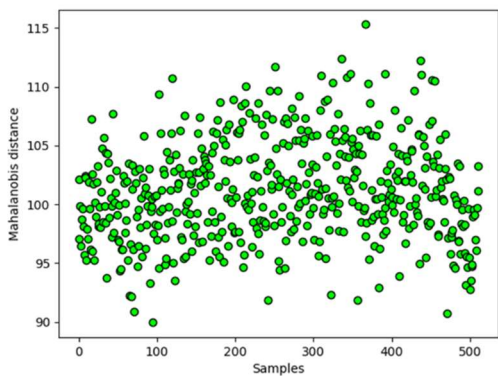
(b) broken bar



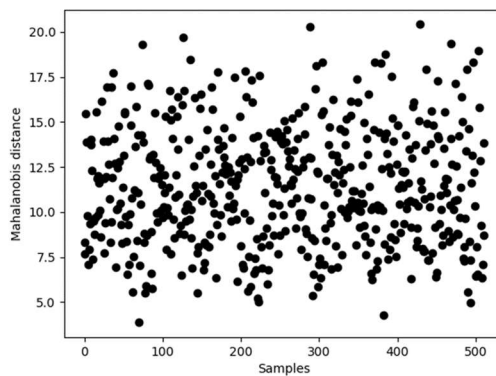
(c) misalignment



(d) mechanical looseness



(e) bearing fault



(f) unbalanced

Figure 16. Mahalanobis distance for the different induced operational conditions (adapted from Ribeiro Junior, 2022).

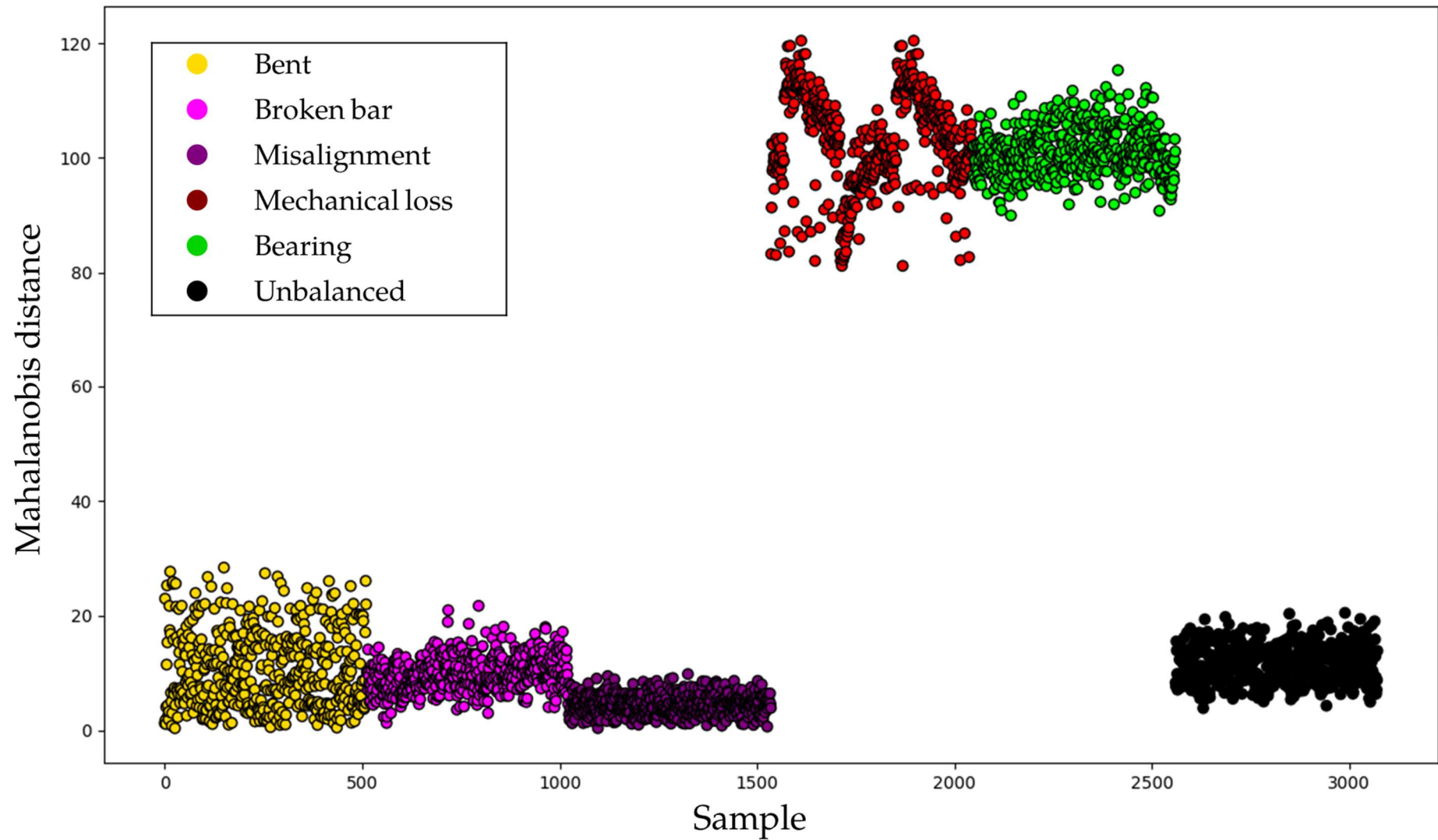


Figure 17. Mahalanobis distance for all the evaluated and induced operational condition (adapted from Ribeiro Junior, 2022).

4 Chapter Conclusion

In this chapter we present the theory of the Gaussian mixture model as well as its effectiveness in clustering and classifying data. In addition, the GMM theory was applied to a real problem. By applying the GMM to a real case, it was possible to demonstrate and discuss the strength of the applied method, as well as the challenges in using it. Also, as it is an unsupervised learning algorithm, this method does not need human support to carry out the identification of faults. Finally, we believe that with this chapter the readers have understood how the GMM works as well as ways to apply the method to other engineering problems.

Acknowledgements

The authors would like to acknowledge the financial support from the Brazilian agency CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – 150117/2021-3), and FAPEMIG (Fundação de Amparo à Pesquisa do Estado de Minas Gerais - APQ-00385-18). Also, the authors would like to acknowledge the support from PS Soluções, that provided the experimental laboratory to carry out this Chapter.

References

- Abraman, A. situação da Manutenção no Brasil. In: Anais do 26º Congresso Brasileiro de Manutenção. Curitiba: Abraman, 2011.
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6), 716-723.
- Choudhary, A., Goyal, D., Shimi, S. L., & Akula, A. (2019). Condition monitoring and fault diagnosis of induction motors: A review. *Archives of Computational Methods in Engineering*, 26(4), 1221-1238.
- Dellaert, F. (2002). The expectation maximization algorithm. Georgia Institute of Technology.
- Duda, R. O., & Hart, P. E. (1973). *Pattern classification and scene analysis* (Vol. 3, pp. 731-739). New York: Wiley.
- Hamill, P., Giordano, M., Ward, C., Giles, D., & Holben, B. (2016). An AERONET-based aerosol classification using the Mahalanobis distance. *Atmospheric Environment*, 140, 213-233.
- Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). *The elements of statistical learning: data mining, inference, and prediction* (Vol. 2, pp. 1-758). New York: springer.
- Hong, Y., Kim, M., Lee, H., Park, J. J., & Lee, D. (2019). Early fault diagnosis and classification of ball bearing using enhanced kurtogram and Gaussian mixture model. *IEEE Transactions on Instrumentation and Measurement*, 68(12), 4746-4755.
- Kolmogorov, A. N., & Bharucha-Reid, A. T. (2018). *Foundations of the theory of probability: Second English Edition*. Courier Dover Publications.

- McNicholas, P. D. (2016). Mixture model-based classification. Chapman and Hall/CRC.
- Moon, T. K. (1996). The expectation-maximization algorithm. *IEEE Signal processing magazine*, 13(6), 47-60.
- Murphy, K. P. (2012). *Machine learning: a probabilistic perspective*. MIT press.
- Myung, I. J. (2001). *Computational approaches to model evaluation*.
- Myung, I. J. (2003). Tutorial on maximum likelihood estimation. *Journal of mathematical Psychology*, 47(1), 90-100.
- Panić, B., Klemenc, J., & Nagode, M. (2020). Gaussian Mixture Model Based Classification Revisited: Application to the Bearing Fault Classification. *Strojniski Vestnik/Journal of Mechanical Engineering*, 66(4).
- Pearson, K. (1894). Contributions to the mathematical theory of evolution. *Philosophical Transactions of the Royal Society of London. A*, 185, 71-110.
- Plante, T., Stanley, L., Nejadpak, A., & Yang, C. X. (2016, September). Rotating machine fault detection using principal component analysis of vibration signal. In *2016 IEEE Autotestcon* (pp. 1-7).
- Profillidis, V. A., & Botzoris, G. N. (2018). *Modeling of transport demand: Analyzing, calculating, and forecasting transport demand*. Elsevier.
- Randal, R.B. *Vibration-based Condition Monitoring: Industrial, aerospace and automotive applications*. John Wiley & Sons, 2011.
- Rhys, H. (2020). *Machine Learning with R, the tidyverse, and mlr*. Simon and Schuster.
- Ribeiro Junior, R. F., dos Santos Areias, I. A., Campos, M. M., Teixeira, C. E., da Silva, L. E. B., & Gomes, G. F. (2022). Fault Detection and Diagnosis in Electric Motors Using Convolution Neural Network and Short-Time Fourier Transform. *Journal of Vibration Engineering & Technologies*, 1-12.
- Ribeiro Junior, R.F., de Almeida, F.A. & Gomes, G.F. Fault classification in three-phase motors based on vibration signal analysis and artificial neural networks. *Neural Comput & Applic* 32, 15171–15189 (2020).
- Ribeiro Junior, R. F., Machine learning-based fault detection and diagnosis in electric motors (2022).
- Rissanen, J. (1978). Modeling by shortest data description. *Automatica*, 14(5), 465-471.
- Rousseeuw, P. J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics*, 20, 53-65.
- Schwarz, G. (1978). Estimating the dimension of a model. *The annals of statistics*, 461-464.
- Wang, G. F., Li, Y. B., & Luo, Z. G. (2009). Fault classification of rolling bearing based on reconstructed phase space and Gaussian mixture model. *Journal of Sound and Vibration*, 323(3-5), 1077-1089.

Wolfe, J. H. (1963). Object cluster analysis of social areas (Doctoral dissertation, University of California).

Wolfe, J. H. (1965). A computer program for the maximum likelihood analysis of types. Technical Bulletin 65-15, U.S. Naval Personnel Research Activity.

Wolfe, J. H. (1970). Pattern clustering by multivariate mixture analysis. *Multivariate Behavioral Research* 5, 329–350

Woźniak, M., Gałazka-Friedman, J., Duda, P., Jakubowska, M., Rzepecka, P., & Karwowski, Ł. (2019). Application of Mössbauer spectroscopy, multidimensional discriminant analysis, and Mahalanobis distance for classification of equilibrated ordinary chondrites. *Meteoritics & Planetary Science*, 54(8), 1828-1839.

Yan, H. C., Zhou, J. H., & Pang, C. K. (2017). Gaussian mixture model using semisupervised learning for probabilistic fault diagnosis under new data categories. *IEEE Transactions on Instrumentation and Measurement*, 66(4), 723-733.

Yang, M. S., Lai, C. Y., & Lin, C. Y. (2012). A robust EM clustering algorithm for Gaussian mixture models. *Pattern Recognition*, 45(11), 3950-3961.

Chapter 8

Sensor Placement Optimization of a Helicopter Main Rotor Blade

Chapter details

Chapter DOI:

<https://doi.org/10.4322/978-65-86503-88-3.c08>

Chapter suggested citation / reference style:

Pereira, João L. J., et al. (2022). “Sensor Placement Optimization of a Helicopter Main Rotor Blade”. In Jorge, Ariosto B., et al. (Eds.) *Uncertainty Modeling: Fundamental Concepts and Models*, Vol. III, UnB, Brasilia, DF, Brazil, pp. 221–257. Book series in Discrete Models, Inverse Methods, & Uncertainty Modeling in Structural Integrity.

P.S.: DOI may be included at the end of citation, for completeness.

Book details

Book: Uncertainty Modeling: Fundamental Concepts and Models

Edited by: Jorge, Ariosto B., Anflor, Carla T. M., Gomes, Guilherme F., & Carneiro, Sergio H. S.

Volume III of Book Series in:

Discrete Models, Inverse Methods, & Uncertainty Modeling in Structural Integrity

Published by: UnB City: Brasilia, DF, Brazil Year: 2022

DOI: <https://doi.org/10.4322/978-65-86503-88-3>

Sensor Placement Optimization of a Helicopter Main Rotor Blade

João Luiz Junho Pereira^{1*}, Matheus Brendon Francisco², Ariosto Bretanha Jorge³, and Guilherme Ferreira Gomes⁴

¹Mechanical Engineering Institute, Federal University of Itajubá (UNIFEI), Itajubá, Brazil – joaoluizjp@unifei.edu.br <mailto:joaoluizjp@unifei.edu.br>

² Management and Production Engineering Institute, Federal University of Itajubá (UNIFEI), Itajubá, Brazil – matheus_brendon@unifei.edu.br

³Post-Graduate Program - Integrity of Engineering Materials, University of Brasilia, Brazil. E-mail: ariosto.b.jorge@gmail.com

⁴Mechanical Engineering Institute, Federal University of Itajubá (UNIFEI), Itajubá, Brazil – guilhermefergom@unifei.edu.br;

*Corresponding author

Abstract

This chapter describes the importance of one of the major areas within Structural Health Monitoring and Health and Usage Monitoring Systems: Sensor Placement Optimization (SPO). In this sense, the most important SPO metrics are presented and discussed with the reader. The methodological procedure of the optimization problem statement is also considered and discussed. The significance of rotary-wing aircraft, damages that can occur in the blade, works related to damage identification, modern methods and algorithms that have been used in this context, and a comprehensive SPO study for a helicopter main rotor blade will be presented. Finally, a new technique that considers the number of sensors as an objective associated with some of the main metrics used in SPO. The technique uses the Multi-objective Lichtenberg Algorithm and Feature Selection, an important area used in Data Mining.

Keywords: Structural Health Monitoring, Sensor Placement Optimization, Helicopter rotor blade.

1 Introduction

Helicopters have great importance due to its capabilities of travel speed (200-300 km/h) and maneuverability, which promotes movement along the shortest track and vertical takeoff (Khabarov & Komshin [2021]). These unique qualities make this aircraft efficient for urban transport of people, rescue in hard places, use on offshore oil rigs and military forces, among other applications.

Being a rotary-wing aircraft, or rotorcraft, its most important element is the main rotor blade (MRB), which is usually made of composite material. Due to the intense in-service conditions of aerodynamics, temperature variation, and accelerations, damage can be induced in these structures (Ahmad *et al.* [2020]). Some examples are: fracture of matrix, delamination (most important), fiber breakage, fiber waviness, creep deformation, etc. (VOICU *et al.* [2020]).

As a matter of fact, composite structures don't give any clue till the collapse. It is necessary to have methodologies for the assessment, identification and prediction of damages in order to avoid catastrophic failures. Efficient systems, such as Structural Health Monitoring (SHM) or Health and Usage Monitoring Systems (HUMS), allow early identification of damage through non-destructive inspection and integrated sensors exploring vibration/modal measures. The method is based on the principle that a damaged in a structure changes natural frequencies, mode shapes, modal strain energy and damping ratios (Gomes *et al.* [2020]; Pereira *et al.* [2021]).

Then, it is possible to identify the damage using inverse modeling and computational intelligence analyzing the output vibration parameters of the system. The main data processing techniques used, nowadays, are meta-heuristics, neural networks, non-probabilistic methodologies, and time-series analyses. However, there is no methodology that is very effective for all problems. New technologies have been proposed and there is space for new ones (Burgos [2020]). There are not many works in the literature that have been dedicated to studying damage in MRB and most modeled the study using some direct method, that is, the problem is elaborated and their answers are directly analyzed.

In case of experimental studies, Cattarius & Inman [2000] induced punctual increase and decrease of mass and observed significant variations of natural frequencies in a real MRB. A fact also proved by Santos *et al.* [2013]. Ghoshal *et al.* [2001] tested different types of sensors and concludes that those linked to modal properties are the most reliable in damage identification. Kiddy & Pines [2001] used an experimental methodology to conclude that localized damage in low modal energy locations is difficult to detect. Santos *et al.* [2016] compared coordinate modal assurance criterion and modal strain energy method to detect damage on a full-size composite MRB. The authors concluded that the last method was more sensitive to the presence of damage.

Equally important, other authors used numerical modeling to carry out their related studies. Pawar & Ganguli [2005] studied the effect of matrix cracks in composite MRB using strain measurement and concludes that is possible to find the localization and density of a matrix cracks using this metric. The same authors extend the study to delaminations and fiber breakage, analyzing stiffness, natural frequencies, deflections, root forces, root moments and strains in forward flights (Pawar & Gangulli, [2007c]). The authors also concluded that modal data are appropriate for damage identification. In

like manner, Ahmad *et al.* [2020] created a faithful MRB with contour conditions and forces identical to a flight situation to study the progressive failure. The authors found that failure starts from the root chord and tend toward the tip chord. Shahani & Mohammadi [2015] analyzed the fatigue life of a numerical blade using safe-life and damage tolerance approaches and concludes that the skin is the component that has the shortest lifetime.

The first works using SHM and inverse methods associated with finite element method (FEM) updating on MRB appeared at the beginning of the century. Pawar & Ganguli [2003] created a genetic fuzzy system to find the damage location and extent: a fuzzy system qualifies the changes in natural frequencies in four damage levels and the Genetic Algorithm (GA) optimizes its rule-base and membership functions. Despite being a qualitative damage identification and the authors have used a composite beam as a MRB, the methodology proved to be efficient. The same authors used the same strategy in two other studies: *i*) using displacements instead natural frequencies and residual life classes instead damage levels (Pawar & Gangulli, [2007]) and *ii*) changing the genetic fuzzy system by the Support Vector Machine (SVM) classifier, still using natural frequencies to find damage levels, having success even in noisy situations (Pawar & Gangulli, [2007b]). Reddy & Ganguli [2003] used artificial neural networks (ANN) with modal data generated by the FEM to identify MRB damage. The authors have success and concluded that is enough use the first 5 mode shapes.

Up to the present time, Gomes *et al.* [2020b] was the first to propose a methodology for identifying MRB damage that locate and quantify the damage severity. The authors proposed an inverse method using the FEM updating associated with the Bat Optimization Algorithm (BA) and uniformly distributed 10 sensors on the blade to measure displacements due to mode shapes.

As can be seen, none of them proposed a MRB Sensor Placement Optimization (SPO) methodology. Just like none of them either proposed a methodology using multi-objective meta-heuristics. At the same time, modern methodologies have been proposed for damage identification in other structures using: *i*) FEM updating, *ii*) frequency response function, *iii*) ground excitation, *iv*) signal processing, *v*) new machine learning algorithms, among others. The FEM updating is the most precise and have two ways: *i*) direct problem modeling, where a structural numerical model is made, and *ii*) inverse problem modeling, where the model is constantly evaluated by an optimization algorithm that minimizes an objective function composed by structural characteristics (Milad *et al.* [2019]; Assis & Gomes [2021]).

The SPO problem in structures can be complex, multi-modal, with multiple variables and even multiple objective functions. The complexity becomes even greater if there is interest in sensors number reduction. Therefore, the efficiency of the methodology and the quality of the answers strongly depend on the optimization algorithm used and meta-heuristics have proven to be the best.

Some important works using mono-objective meta-heuristics in SPO are: *i*) Genetic Algorithm (GA) (Gomes *et al.* [2019]), *ii*) Ant Colony Optimization (ACO) (Braun *et al.* [2015]; Mishra *et al.* [2019]), *iii*) Particle Swarm Optimization (PSO) (Chen & Yu [2018]; Kaveh & Maniat [2015]), *iv*) Firefly Algorithm (FA) (Pan *et al.* [2016]), *v*) Sunflower Optimization (SFO) (Gomes *et al.* [2019]), *vi*) BA (Zenzen *et al.* [2018]; Gomes *et al.* [2020b]), among others.

However, recent studies indicate that the use of multi-objective meta-heuristics, given their ability to evaluate several metrics at the same time, has found better results both in SPO and in damage identification (Alkayem *et al.* [2017]; Gomes *et al.* [2018], Zhou *et al.* [2021]). Furthermore, the fundamental principle of SHM is select the smallest number possible of measurement locations from a structure and represent the system with the highest possible accuracy, which are obviously conflicting objectives (Barthorpe & Worden [2009]; Pereira *et al.* [2021b]).

Even so, the studies found in SPO aims cost or energy minimization while the efficiency of data reading is improved: *i)* Ferentinos & Tsiligiridis [2007]: wireless sensors, *ii)* Lin *et al.* [2018]: three-dimensional truss structure, *iii)* Gomes *et al.* [2018]: The first using FIM and MSI at the same time, finding excellent results, *iv)* Cao *et al.* [2020]: proposed a method named distance coefficient-multi-objective information fusion algorithm (D-MOIF) and two objective functions based on other metrics. The author was the first to study SPO in a more complex structure, *v)* Zhou *et al.* [2021]: showed that the Multi-objective FA obtained better results than the NSGA-II, and *vi)* Colombo *et al.* [2022]: proposed a statistical methodology that evaluated the positioning of sensors together with the damage identification.

To the best author's knowledge, no works are found that consider the number of sensors as one of the objectives. In larger and more complex structures, as in a MRB, the need to develop multi-objective methodologies that select the smallest number of sensors that bring the maximum information about the structure is essential. The first study to do so was recently published and is the result of the Research Group in Computational Mechanics (GEMEC) at UNIFEI (Pereira *et al.* [2022b]). The technique uses the Multi-objective Lichtenberg Algorithm and techniques employed in Feature Selection, an important area in Data Mining. The methodology applied as well as unpublished information will be presented in this chapter.

So, in this chapter will be introduced the importance of studying SPO in an MRB, discussing in detail the appearance of damage in it, how these damages influence its mechanical properties, how the problem of SPO is treated, including your main metrics and makes a case study of SPO considering a single objective, maximizing one of the main metrics, and considering these same metrics, but having the number of sensors as an objective in a multi-objective SPO.

2 Backgrounds

The concepts necessary for the development of this chapter are: *i)* types and characteristics of damages in MRB, *ii)* how modal data can help identify damage, *iii)* the main metrics used to interpret and analyze modal data, and *iv)* what is the Lichtenberg Algorithm meta-heuristic and its multi-objective version, not yet shown in this book series.

2.1 Damage in main helicopter rotor blade

Composite materials have distinctive qualities, such as low weight, high strength, remarkable stiffness related to its specific mass, and high capacity to withstand fatigue and corrosion. These reasons justify using them in the manufacture of MRB, which have in their constructions multiple constituents and a pronounced anisotropic behavior. However, they may have failure mechanisms. In addition to possible manufacturing

defects, the helicopter in flight can present severe conditions of aerodynamics, temperature variation, and accelerations loads that contribute to emergence of damage (Ahmad *et al.* [2020]).

Small defects in blade fabrication, such as autoclave molding, resin transfer mold or resin film infusion, can progress severely when in operation. The main damages resulting from this are: non-uniformly positioned laminae, uneven distributed resin, missing or cut fibers, incomplete resin maturation, porosity or trapped air bubbles, signs of wear or scratches on the material, and fiber or weave undulation (Huo & Zhang [2012]).

In operation, overloads, environmental factors or impact with foreign objects can aggravate previous manufacturing defects or give rise to new damages. Some examples are: fracture of matrix, fiber breakage, debonding, delamination, fiber waviness, and creep deformation (Voicu *et al.* [2020]). A matrix crack develops between two or more parallel laminae and can propagate beyond them. Despite being a type of damage that does not present an immediate risk, it can be a starting mechanism for more serious ones. However, fiber breakage considerably reduces the Young's modulus and the load bearing capacity of the material, what makes it a damage that requires more attention (Amabilia *et al.* [2017]).

Debonding occurs when the adhesive cannot bond the surfaces which it is applied, which happens when external or environmental loads breakage chemical, physical, or mechanical links between the surface and adhesive. According to Voicu *et al.* [2020], material ageing, low speed impact or flaws in the development process can lead to this damage. Another damage that considerably reduces the strength of a composite material is the fiber waviness. This damage appears due to the local buckling of prepreg or wet hoop-wound filaments strand (Vinet & Gamby [2008]).

The most critical and important damage is delamination, which appears as the result of impact of foreign objects. This damage can lead the entire structure to failure, as the material strength is proportional with the degree of delamination (Latifi *et al.* [2015]). This damage is the loss of stiffness in the material due to spaces formed between the adjacent layers of a laminate and is the main source of cracks in composite structures (Pantano [2019]).

2.2 Damage monitoring using modal data

These damages cause structural degradation and change natural frequencies, mode shapes, modal strain energy and damping ratios in very particular ways. Then, it is possible to identify the damage using inverse modeling and computational intelligence analyzing the output modal parameters of the system.

The most used modal properties in damage identification are natural frequencies and mode shapes (Gomes *et al.*, [2020]). Both have their advantages and drawbacks. According to Gopalakrishnan *et al.* [2011], natural frequencies have low sensitivity in determining damage location when compared to mode shapes. However, according to Worden & Friswell [2009], natural frequencies can be estimated very accurately (1% error) without a complete modal test. A single random excitation test can determine it through a basic spectral analysis. Still, according to Guan *et al.* [2017], it can be obtained using a single and random structural point.

According to Hearn & Testa [1991], the change of the i -th natural frequency after a change in the global stiffness matrix ΔK_N in a damaged structure is represented by Equation (1).

$$\Delta\omega_i^2 = \frac{(\epsilon_N(\Phi_i))^T \Delta K_N (\epsilon_N(\Phi_i))}{\Phi_i^T M \Phi_i} \quad (1)$$

where Φ_i defines the i -th mode shape, M is the mass matrix, and $\epsilon_N(\Phi_i)$ is the strain vector calculated by the mode shape.

However, while the natural frequency is a property that considers the entire structure, mode shapes are able to provide local modal properties. For this reason, it has been the most used to identify localized damage. It has greater sensitivity to structural damage and provides detailed information about the extent and location of it (Pereira *et al.*, [2021]; Gomes *et al.* [2018]).

The mode shapes are the eigenvectors Φ and natural frequencies are the eigenvalues λ of the problem expressed in Equation (2). Both can be obtained solving this Equation.

$$[K - \lambda_i M] \Phi_i = 0 \quad (2)$$

where Φ_i is the eigenvector associated with the i -th mode shape, λ_i is the eigenvalue, M is the mass matrix, and K is the stiffness matrix, which is the most affected matrix when damage comes.

In delamination case, for example, the stiffness matrix is the only one affected and according to Pereira *et al.* [2021], the variation in the i -th mode shape is in Equation (3):

$$\Delta\Phi_i = \sum_{r=1, r \neq i}^p \frac{-\Phi_r^T \Delta K \Phi_i}{\lambda_r - \lambda_i} \quad (3)$$

where Φ_i is the eigenvector and λ_i the eigenvalue associated with the i -th mode shape, ΔK is the variation stiffness matrix, and λ_r the eigenvalue and Φ_r the eigenvector associated with the damaged structure.

Solving these equations for complex structures is not an easy task. However, advances in computing and FEM make it possible to obtain, both natural frequencies and mode shapes, for any type of structure and for any given excitation range in fast numerical simulations. Therefore, it is possible through the mode shapes to calculate the displacement of any point of this structure for each of the natural frequencies, due to the particular ways that each structure vibrates when excited near its natural frequencies. Figure 1 shows the mode shapes of a helicopter main rotor blade (MRB).

Using FEM, the variation of the global stiffness matrix is simplified as shown Equation (4). Many authors have used this equation to represent delaminations (PEREIRA *et al.* [2021]; GOMES *et al.*, [2020b]).

$$\Delta K = \sum_{k=1}^N \alpha_k K_k \quad (4)$$

where α is a scalar multiplier (between 0 and 1) that modifies the original K stiffness in the k element.

FEM is a powerful tool that allows finding the displacement at each node in each mode shape, however, it is unable to answer the following question: where and how many sensors should I place on a specific structure to obtain as much information as possible about how it vibrates? Note that the answer to this question not only allows identifying the sensing that provides the highest level of information about the structure, but also allows for a more accurate and efficient damage identification.

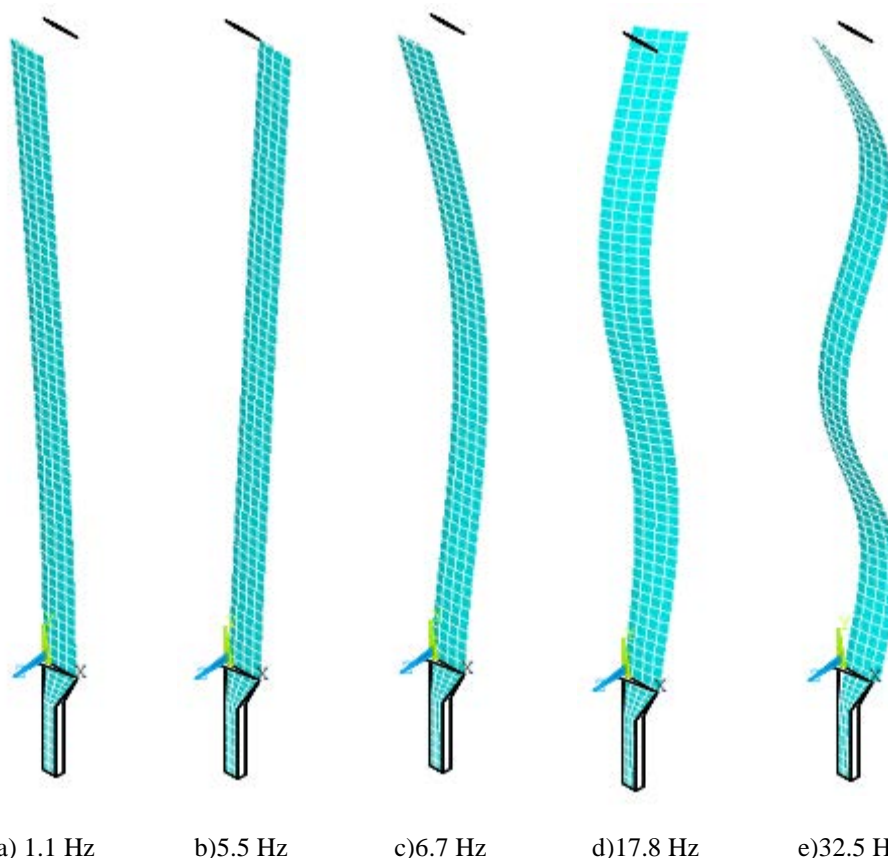


Figure 1: First five MRB natural frequencies (adapted from Pereira *et al.* [2022b])

It is intuitive for the reader to imagine that the more sensors, the more information will be acquired and that is one of the biggest challenges of sensor positioning: a greater number of sensors means more weight, cost, processed data size, and changes to the modal properties of the structure as more mass, wiring, and other items are added to it. Therefore, it is shown that the acquisition of information and the number of sensors are conflicting objectives and from here arises an area within SHM called Sensor Placement Optimization (SPO).

Knowing that each point of a structure in a given mode shape is a point where a displacement can be measured and therefore a sensor can be installed, it is intended to identify which combinations of these sensors provide the greatest displacements and still use as few sensors as possible. To answer this question, several SPO metrics have been proposed, which although they approach these displacement measures in different ways, they all aim for the same answer.

2.3 Modal metrics in Structural Health Monitoring

This section presents the main SPO metrics used in recent years. From them, an objective function can be proposed so that, when optimized by an optimization algorithm, the set of sensors that delivers the most information is obtained. Some best known metrics are: Kinetic Energy (KE) (Heo *et al.* [1997]), Effective Independence (Efi) (Yang & Lu [2017]), Average Driving-Point Residue (ADPR) (Barthorpe & Worden [2009]), Eigenvalue Vector Product (EVP) (Barthorpe & Worden [2009]), Information Entropy (IE) (Yuen & Kuok [2015]), Fisher Information Matrix (FIM) (Benner *et al.* [2017]), and Modal Assurance Criterion (MAC) (Jin *et al.* [2015]).

The main component of all these metrics is the mode shape matrix (Φ). To understand it, see Figure 1 again. It is possible to see that the blade is separated into small squares that are finite elements. Each vertex of each element is a node and it is at this point that FEM calculates the displacement due to the mode shape. As can be seen, there are many nodes for the blade and each node can present different displacements for different mode shapes. The mode shape matrix is the matrix whose number of rows i is equal to the number of selected sensors and the number of columns j is equal to the number of mode shapes to be considered. Therefore, element a_{ij} is the displacement of node i in mode shape j .

Therefore, one of the main metrics is the KE. It gives a measure of the dynamic contribution of each element of FEM to each of the mode shapes and can be calculated through Equation (5).

$$KE_{ij} = \Phi_{ij} \sum_j M_{ij} \Phi_{ij} \omega_j^2 \quad (5)$$

where Φ is the mode shape matrix, M is the mass matrix, i and j refers to the degrees of freedom and n , the n -th mode (Heo *et al.* [1997]). Note that the calculated kinetic energy grows with the square of the natural frequency of the considered mode shape, so the method tends to value higher frequency nodal displacements.

Efi is one of the most used metric in large structures, which is an efficient unbiased estimator. Its objective is to select positions of measurements linearly independent providing high information. The covariance matrix of the estimated error can be expressed in Equation (6). Note that unlike KE, mass and natural frequency are not considered.

$$Efi = \text{diag}(\Phi [\Phi^T \Phi]^{-1} \Phi^T) = \text{diag}(QQ^T) \quad (6)$$

where Q is an orthonormal matrix and Φ is the mode shape matrix. This metric has a few drawbacks: *i*) it might choose sensor locations with less energy content, just like KE, what can leave sensing incomplete and bring vulnerability in noisy conditions (Cao *et al.* [2020]), *ii*) it presented poor results under conditions of uncertainty (Kammer [1992]).

The ADPR provides a measure of the contribution of any point to the global response. For all N mode shapes, the ADPR in the i -th node can be calculated using Equation (7).

$$ADPR_i = \sum_{j=1}^N \frac{\Phi_{ij}^2}{\omega_j} \quad (7)$$

where ω_j is the j -th natural frequency and Φ is the mode shape matrix.

The EVP calculates the product of the mode shapes for its locations for N modes to be measured: a maximum for this product is a point with optimum measurement candidate. In the i -th, it can be calculated by Equation (8):

$$EVP_i = \prod_{j=1}^N |\Phi_{ij}| \quad (8)$$

IE is an efficient metric to finding the best combination of structural tests that can minimize the negative consequence of uncertainty. The metric is based on a Bayesian statistical method, where a probability density function $p(\theta|D)$ (Equation 9) is used to quantify the uncertainties in the parameters θ . The calculation is defined in Equation (10).

$$p(\theta|D) = c [J(\theta; D)]^{-(NN_0-1)/2} \pi_\theta(\theta) \quad (9)$$

$$H(D) = E_\theta [-\ln p(\theta|D)] = -\int p(\theta|D) \ln p(\theta|D) d\theta \quad (10)$$

where E_θ is the mathematical expectation to θ , D is the dynamic test data, $J(\theta|D)$ is the fit measure between the measured and the response time histories, N_0 is the number of DOF of the structure and N is the number of sampled data. Note that, while EfI gives the relative importance of the chosen DOF, IE is related to the total maximum limit of entropy.

The FIM is the main metric used in SPO, where the optimal sensor configuration is taken as the one that maximizes the norm of the FIM and minimizes the expected Bayesian loss function involving the trace of the inverse of the FIM (Gomes *et al.* [2018]). The array of sensors can be given in the form of an estimation problem with a corresponding FIM given in Equation (11) (Kammer [1991]). The modal response is estimated based on the data measured by the sensors.

$$Q = \Phi_S^T W \Phi_S \quad (11)$$

where W is a weighting matrix and Q , when maximized, results in the minimization covariance matrix error. This results in the maximization of the signal intensity and the independence of the main directions (Rao *et al.* [2015]). This metric showed good performance in noise and multi-axial data (Kammer [1992]; Kammer & Tinker [2004]).

MAC is a metric proposed by Carne and Dohrmann [1994] that aims to find sensor locations to ensure that all inner products between distinguishable shape vectors have relatively small cosines. It allows comparison of the different mode shapes and is defined by Equation (12).

$$MAC = \frac{(\Phi_i^T \Phi_j)^2}{(\Phi_i^T \Phi_i)(\Phi_j^T \Phi_j)} \quad (12)$$

where Φ_i and Φ_j are the i -th and j -th column vectors in matrix Φ . The matrix MAC values range between 0 and 1, where 1 indicates a high similarity between the modal vectors. In an ideal condition, the MAC matrix has its null and all elements of main diagonal are 1. So whether it's an OSP or damage identification problem, it is desired to maximize the sum of diagonal values. Note that Efi is MAC^{-1} (Tan & Zhang [2019]).

Many studies show significant correlation and statistical equality between: *i*) Efi, KE, and FIM, and *ii*) MAC and IE (Li *et al.* [2007]; Yi & Li [2012]). Gomes *et al.* [2018b] also found a strong correlation between the metrics FIM, KE, EI, EVP and ADPR when studying them for OSP in a laminate composite plate. The authors also conclude that Efi presented the worst result among the metrics compared. Table 1 summarizes and shows the objective functions derived from these metrics to be used in SPO or Damage Identification.

Table 1: Sensor placement objective functions used in Structural Health Monitoring

Metric	Objective function
KE	$J = \max(KE)$
Efi	$J = \max(\text{diag}(QQ^T))$
ADPR	$J = \max(ADPR)$
EVP	$J = \max(EVP)$
IE	$J = \min(H(D))$
FIM	$J = \max \det(Q) $
MAC	$J = \max(\text{diag}(MAC))$

2.4 Lichtenberg and Multi-objective Lichtenberg algorithms

With the objective function defined, meta-heuristics such as GA, PSO, DE, FA, or SFO can be applied. As shown, they are the best and most used optimization algorithms for complex problems. In the MRB SPO problem of this chapter, the Lichtenberg Algorithm (LA) (Pereira *et al.* [2021a]) will be applied, when mono-objective. The LA is a powerful meta-heuristic inspired by the physical phenomenon of intra-cloud lightning and Lichtenberg figures and is the first hybrid algorithm in the literature that combines population and trajectory optimization routines at the same time.

The meta-heuristic, also developed at GEMEC, has been successful in identifying cracks (Pereira *et al.* [2020]) and delaminations (Pereira *et al.* [2021]), isogrid tube optimization (Francisco *et al.* [2020]; Pereira *et al.* [2022a]), hydrogen production optimization (Souza *et al.* [2022]), wind forecast optimization (Tian & Wang [2022]), parameter optimization in manufacturing processes (Challan *et al.* [2022]; Mohanty *et al.* [2022]), and multi-objective optimization (Pereira *et al.* [2022]), among other applications. The LA algorithm in the mono-objective version has already been well explained in previous chapters. In this chapter, the Multi-objective Lichtenberg Algorithm (MOLA) will be presented and used in the multi-objective SPO. It is the first multi-objective meta-heuristic created and registered in Brazil.

The algorithm optimizes by creating Lichtenberg Figures (LF) that are thrown into the search space (of design variables) with random sizes and rotations at each iteration. In the mono-objective version, the LF was always centered on the best point of the

previous iteration, or on the best point of all iterations so far. In the multi-objective version, these figures are triggered from some point of the variables that make up the current Pareto front. As the LF is made up of many points, only a Pop (population) number is randomly selected in this structure to be evaluated in the objective functions.

As the LF has different sizes in each iteration, sometimes it is small and favors exploitation, sometimes it is large and favors exploration. At each iteration, the new and old solutions are compared using the Pareto dominance relationship, thus constituting the current Pareto front. The algorithm stops when all N_{iter} iterations are complete. See Figure 2. Figure two shows the relationship between the search space (of variables) and the objective space (where the Pareto front is). Different colors show the LF in different iterations.

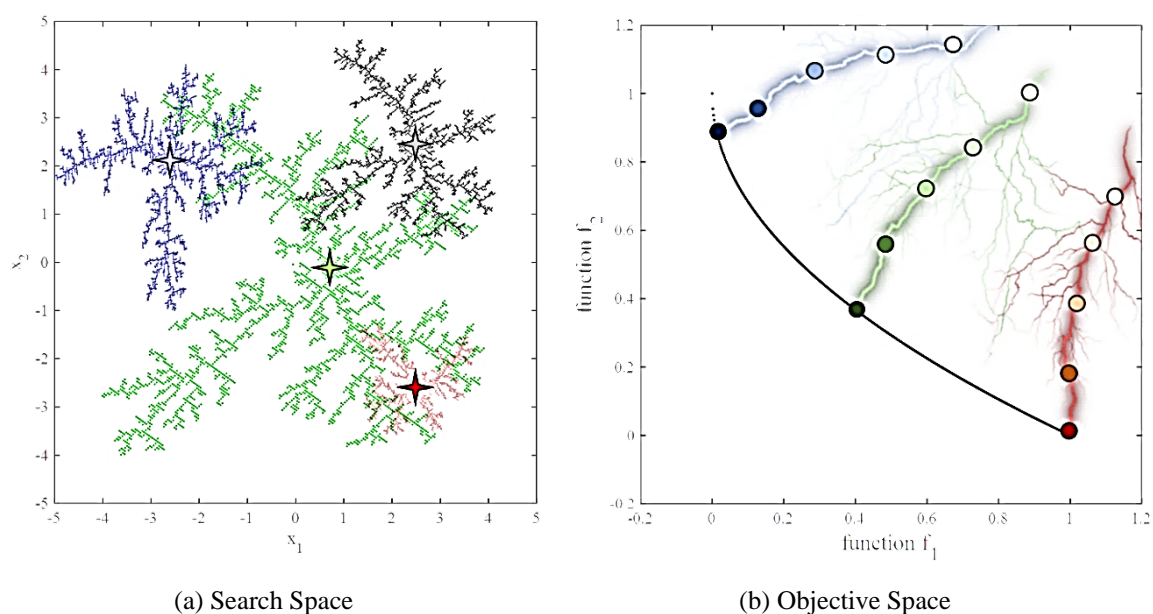


Figure 2: MOLA in a problem of two design variables and two objectives (PEREIRA *et al.*, 2022)

This is the general functioning of MOLA, however, it has other parameters besides Pop and N_{iter} . Three parameters control the LF creation: creation radius (R_c), number of particles (N_p), and sickness coefficient (S). The first two are associated with the size of the LF, which grows in a matrix space whose number of rows and columns is equal to twice R_c . N_p defines the number of particles or unitary elements that can be in this matrix space, remembering that it can be less than the set if the cluster of unitary values reaches R_c . S controls the density of the LF, being close to zero very dense (reminiscent of a cloud) or not, when close to 1, causing a faster cluster growing.

The LF creation process can take about 2 minutes, which would make the optimization process slow. Then another parameter called M was created that controls the creation of the LF and brings three options: *i*) The LF can be created only once and used in all iterations ($M = 1$), *ii*) an LF can be created at each iteration, which is not recommended ($M = 2$), or *iii*) an LF with optimized R_c , N_p and S values can be loaded and in this case the entire optimization process can take less than 1s ($M = 0$). Another additional parameter is the Ref, which creates a local LF at a ref size of the global one, always

having half of all Pop, which favors exploitation. See Figure 3 of the actual LF in the search space.

Finally, it has two last parameters of MOLA related to the control of the algorithm in the construction of the Pareto front and they greatly increases its execution: the number of grids (N_{grid}) and the number of admissible solutions in the Pareto front (N_s). The first parameter divides the objective space into N_{grid} hyper cubes, decreasing the regions of comparison and analysis of the Pareto dominance relationship (Coello *et al.*, 2004). N_s limits the number of solutions in the PF excluding those that are close. Therefore, MOLA has 8 parameters: Pop , N_{iter} , R_c , N_p , S , ref , N_{grid} , and N_s . Figure 4 summarizes MOLA.

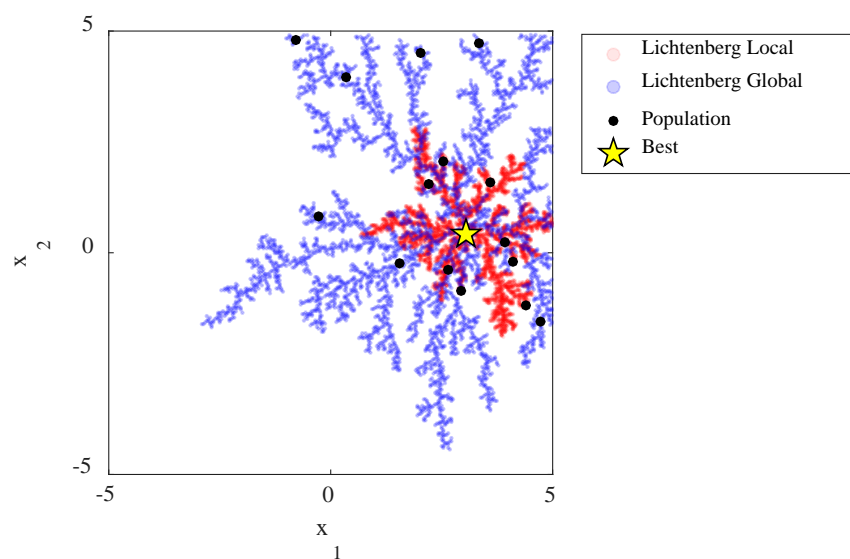


Figure 3: Lichtenberg Figures at one random iteration ($pop = 15$ and $ref = 0.4$) (PEREIRA *et al.*, 2022a)

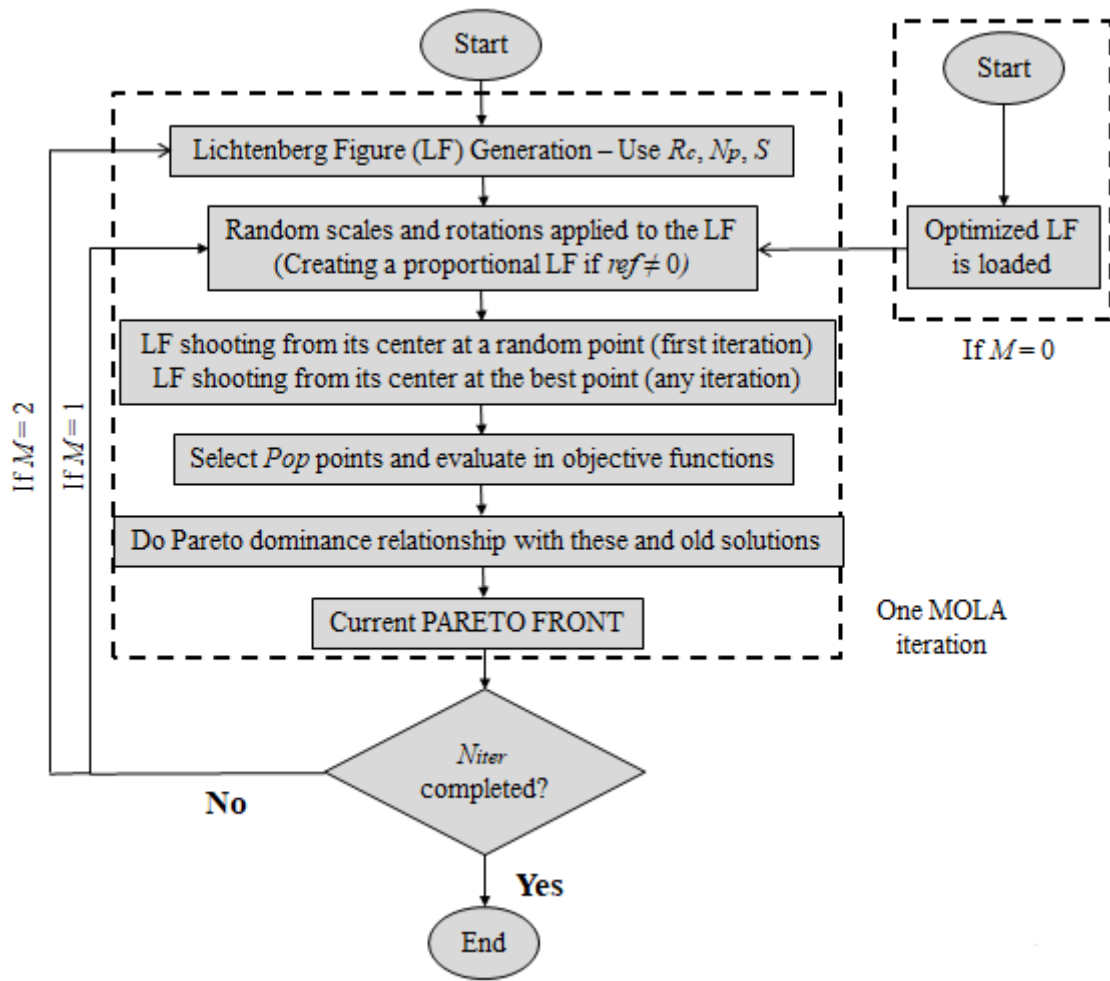


Figure 4: MOLA flowchart (PEREIRA *et al.*, 2022a)

3 Helicopter Main Rotor Blade Case Studies

The metrics discussed in Section 2.3 associated with meta-heuristics in the form of objective functions have been the most accurate and efficient approach used in SPO. As discussed in the introduction to this chapter, the vast majority of workers use a mono-objective meta-heuristic to solve one of the objective functions in Table 1 for any type of structure. However, a new technique was developed by GEMEC using feature selection for multi-objective SPO in which the number of sensors is placed as one of the objectives. In this section, studies will be carried out using both approaches, mono and multi-objective SPO, in a helicopter MRB. For this, a blade identical to an AS-350 MRB was created using FEM. See the real model in Figure 5.

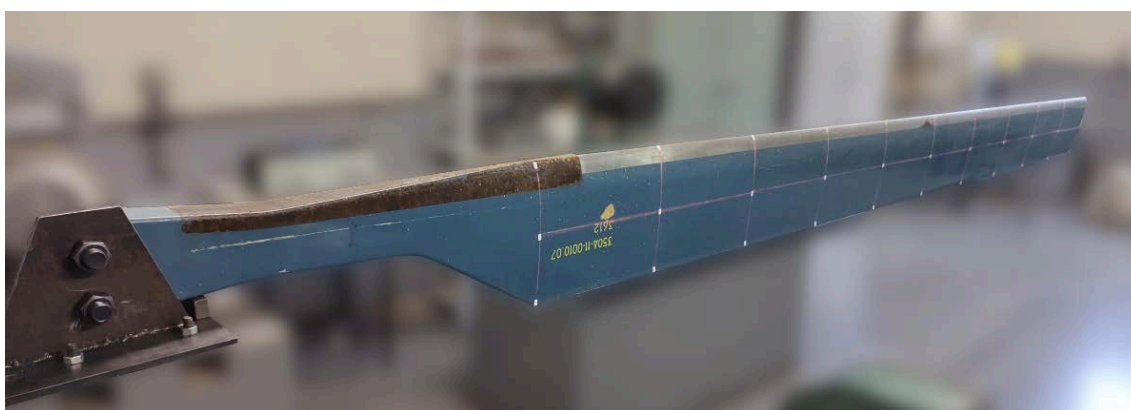


Figure 5: AS-350 MRB (Pereira *et al.* [2022b])

For this, a blade identical to an AS-350 MRB was created using FEM. The model was created in the ANSYS® APDL software and has a total length of 4665mm, being 3880mm with a NACA0012 aerodynamic profile. The element used was the SHELL281 with 8 nodes and 6 degree of freedom per node. Figure 6 shows the solid and all elements and nodes, being each node a candidate position for sensor positioning. Figure 1 showed the first 5 blade mode shapes for this numerical model.

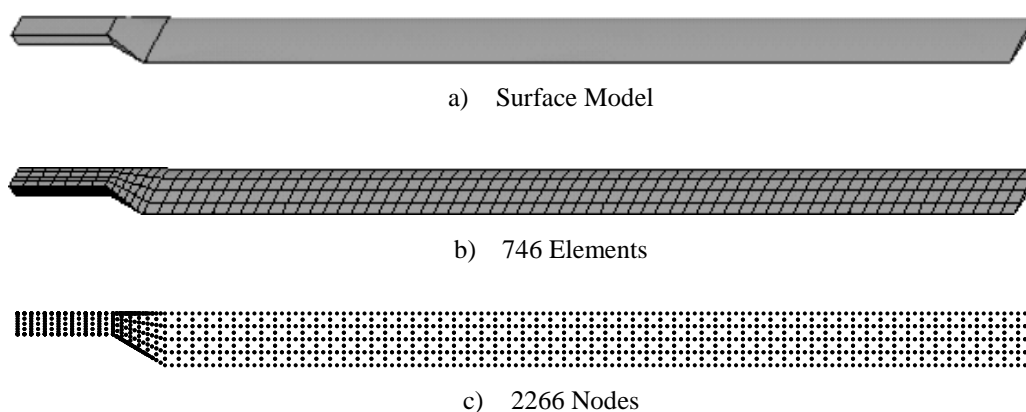


Figure 6: Numerical AS-350 MRB (adapted from Pereira *et al.* [2022b])

3.1 Mono-objective sensor placement optimization

The problem of mono-objective optimization using meta-heuristics consists of some basic steps: *i)* definition of the meta-heuristic to be used and configure its parameters, mainly population and number of iterations; *ii)* definition of the objective function that will guide the optimizer in the search for optimal solutions, *iii)* definition of upper and lower bounds (LB and UB), which define the search space in which the meta-heuristic will look for the best solutions. In this way, each variable in the problem will have a search range; and *iv)* definition of the constraints of the problem, if any.

In SPO, the objective function F can be any of those in Table 1, the variables are the number of sensors n to be allocated, $\vec{X} = \{ S_1, S_2, \dots, S_n \}$. The lower and upper bounds are all the numbers of nodes available in the structure, in the MRB case (Figure 2c), from 1 to 2266. Then each of the variables will have an LB of 1 and a UB of 2266, whose vector length is equal n to be positioned. In this case, it is also convenient to add a constraint g_l that prevents sensors from overlapping, that is, so that two sensors are not positioned in the same place.

Another decision that the SPO problem operator must address is how many mode shapes he will consider. It could be just the first, it could be a multi-objective optimization considering more than one if they are conflicting, or composing an objective function that is the sum of the m mode shapes to be considered, as will be adopted here. The criterion adopted to consider the 5 modes of Figure 1 is that an MRB in the field has excitations ranging from 0 to 25 Hz (Tamer *et al.* [2021]). The SPO problem considering mono-objective optimization and KE metric is expressed in Equation 13.

$$\begin{aligned} \max F(\vec{X}) &= \sum_{i=1}^m KE_i \\ \text{subject to:} \\ 1 \leq S_n &\leq 2266, \text{ for } n = 1, 2, \dots, n \\ g_l(\vec{X}) &= S_n - S_{n+1} > 1 \end{aligned} \quad (13)$$

The LA parameters used in this chapter are those in Table 2, they are optimal parameters found by the authors. Table 1 consists of 7 objective functions. Equation 10 brings the SPO problem to the first one (KE). Note that the number of sensors that will be allocated must be defined before starting the optimization process. In this chapter, 5 sensors will be considered, as it is equal to the mode shapes number being used, a minimum recommendation in the literature (Gomes *et al.* [2018]). Applying the LA with the Parameters in Table 2 for each of these metrics and considering 5 sensors, we have after seven simulations the results in Figure 3. The reader is invited to compare the mode shapes of Figure 1 with the results of Figure 7, remembering that all metrics seek to measure displacements in their own way.

Table 2: Tuned LA parameters

R_c	N_p	S	ref	Pop	N_{iter}
200	10^6	1	0.4	200	200

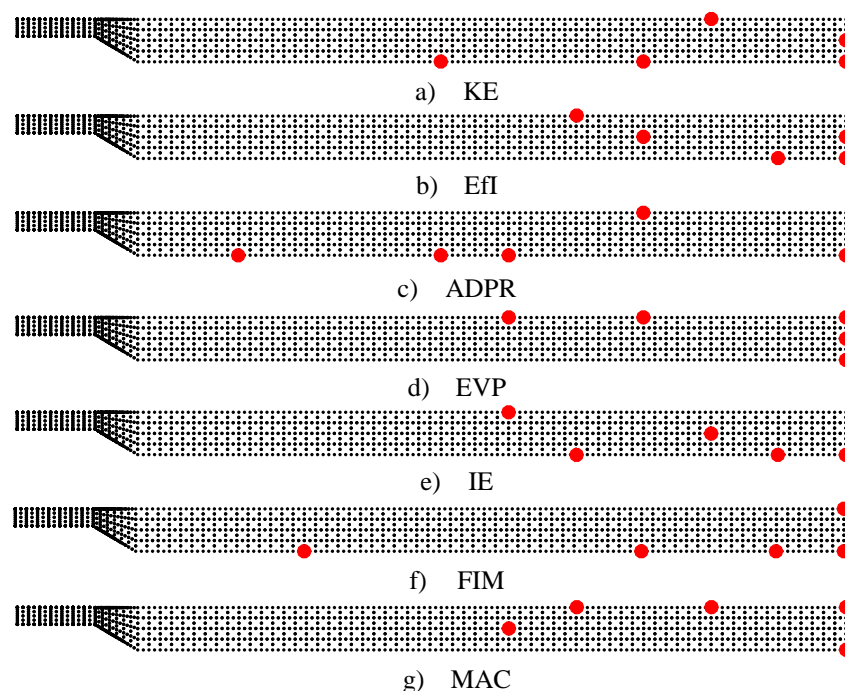


Figure 7: Sensor configurations considering 5 sensors after SPO for each metric

Since the blade is divided into three sections: root, center and tip, with the root located to the left of the images in Figure 3, it can be seen in Figure 1 that the largest nodal displacements for all mode shapes are at its tip. Then comes the central section. Considering only 5 sensors to be distributed, it can be seen that all metrics have at least one sensor at the tip, being the EVP with the highest number exactly at it. Also, all metrics selected the blade tip node that is furthest from the root support, that is, the node with the most movement flexibility of the entire MRB. All metrics have most sensors between the central region and the tip, being that only ADPR and FIM that placed a single sensor in the root section. EfI and FIM neglected the center section in relation to the other metrics. Also, as already observed in other studies in the literature, it is possible to see a sensor distribution relationship between the MAC and IE metrics and between EfI and KE, although in this case KE has considered the central section.

3.2 Multi-objective sensor placement optimization

Most of the works found in the literature perform the SPO as shown in Section 3.1. Note that for each metric or objective function, the algorithm had to be run once and the number of sensors was fixed. However, it is known that the basic principle and biggest challenge in SHM in relation to SPO is to maximize the level of information acquired from the structure by minimizing the number of sensors, which are clearly conflicting objectives.

Developing a multi-objective optimization method capable of finding a Pareto front (PF) that associates each of these metrics with the number of sensors not only allows finding a larger number of sensor configurations, since for each sensor number there will be a sensor configuration and the respective metric value, as well as allowing the

problem operator to evaluate how the information level of that structure varies with the addition of sensors for that metric.

Based on this, it is possible for the problem operator for a given structure to know, for example, if the increase in the number of sensors causes the information level to grow linearly, exponentially, or parabolically. This would imply linear, convex or concave Pareto fronts. It also allows determining whether the addition of sensors allows the level of information to increase significantly up to a certain point and, from then on, it changes insignificantly. Unlike the single-objective approach, with just one execution of the SPO problem, a PF will be found that will have the best possible metric value for each integer number of sensors and each one of these solutions will have the sensors positions.

Despite being a very innovative methodology, this has not yet been done not only due to the greater programming difficulty, but also to the lack of methods. However, new techniques have been developed for data mining and filtering before machine learning algorithms training and prediction. This large area, named Feature Selection (FS), aims to minimize the number of features and maximize a model accuracy, having as its most important approach the wrapper-based method, which using meta-heuristics, it selects different subsets within the dataset and seeks to select those that present the best accuracy values (Sharma & Kaur [2021]).

In the FS problem, there is a vector of length equal to the number of original problem features, whose the selected feature worth one and the rejected, zero. Or, in this case, the sensor will be selected with a value of one and excluded with a value of zero. Therefore is a binary (and discrete) optimization problem. Here begins the first challenges, since basically all meta-heuristics are programmed to deal with continuous problems. The reader may fall into the error of thinking that a simple round could solve the problem, but this simple strategy leads to poor results.

In literature there are two main groups that can convert a optimization algorithm in binary: i) general approach, where the algorithm operators are not modified, such as transfer function and angle modulation, and ii) specific approaches, were the algorithm structure is modified, such as Boolean or quantum binary techniques (Crawford *et al.*, [2017]). The easiest and most effective is to adopt a transfer function and there are two families: *s-shaped* and *v-shaped*. See Figure 8. There are many studies comparing both and most agree that *v-shaped* transfer function is better for FS. The two main reasons for this are that this transfer function does not force continuous values to assume the values of 0 or 1 and that using it, there is a greater chance that a given feature will not be selected (Mirjalili & Lewis [2013] ; Ghosh *et al.* [2020]).

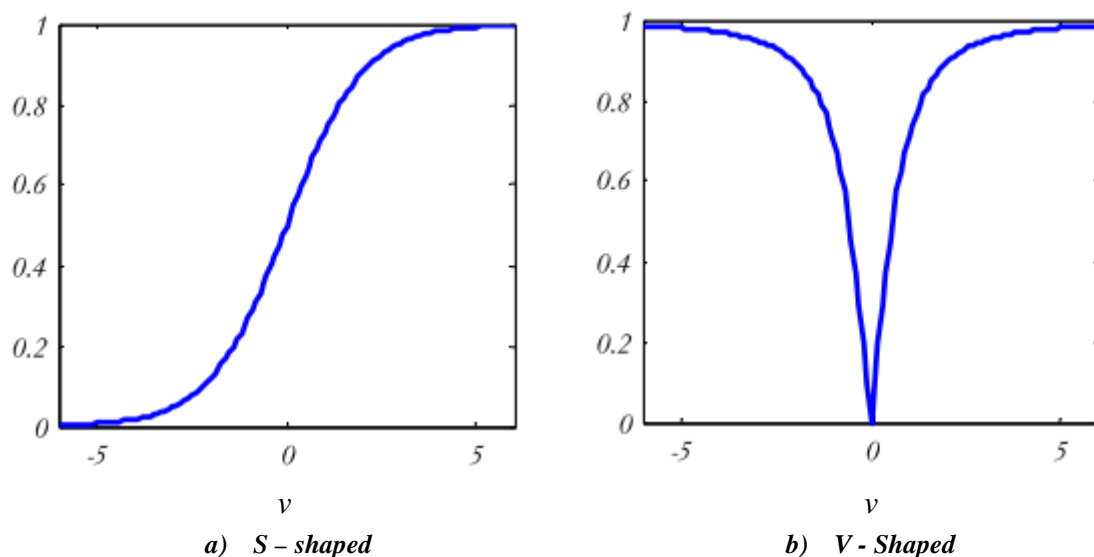


Figure 8: main groups of transfer functions

Thus, MOLA will be equipped with the *v-shaped* transfer function. The binarization process uses Equations 14 and 15 (Mirjalili & Lewis [2013]). Equation 1 represents one of the forms of this transfer function. In it, each continuous element of the vector of decision variables generated by the LA is an input.

$$T(\Delta x) = \left| \frac{\Delta x}{\sqrt{\Delta x^2 + 1}} \right| \tag{14}$$

Then, the $T(\Delta x)$ value is used in Equation 15 to calculate the probability of changing each element to 0 or 1. The *rand* is a randomly generated number between 0 and 1. When $T(\Delta x)$ has a small value, the chance of inverting the element's value is also small.

$$X_{t+1} = \begin{cases} -X_t, & \text{rand} < T(\Delta x_{t+1}) \\ X_t, & \text{rand} \geq T(\Delta x_{t+1}) \end{cases} \tag{15}$$

Implementing the *v-shaped* transfer function, MOLA becomes appropriate for the SPO problem. The node position in FEM has a integer number assigned to it, which generates a candidate sensors vector equal to the nodes number n , i.e., $\overline{S}_c = \{1, \dots, n\}$. In the same way, MOLA generates a binary vector of the same length (\overline{B}). Through Equation 16 can be found the selected sensors vector (\overline{S}_s) (Pereira *et al.*, [2022b]):

$$\overline{S}_s = \overline{S}_c * \overline{B}^T \tag{16}$$

For each vector \overline{S}_s generated are calculated the number of sensors the metric M value, which can be any one of Table 1. The solutions generated in the objective space are compared through the Pareto dominance relationship and dominated solutions are excluded. That is, for a fixed number of sensors, any solution generated with worse metric values is excluded. Likewise, considering a fixed metric value, any generated

solution that has a greater number of sensors for the same metric value is eliminated. The optimization problem is expressed in Equation 17 (Pereira *et al.*, [2022b]). Note that there is no way for sensors to overlap. The a_i is obtained using Equations 14 and 15 through the continuous x values between 0 and 1 delivered by MOLA.

$$F(\vec{S}_s) = \left\{ \min(\text{length}(\vec{S}_s)) ; \max(M(\vec{S}_s)) \right\}$$

where:

$$\begin{aligned} \vec{S}_s &= \vec{S}_c * \vec{B}^T \\ \vec{S}_c &= \{1, \dots, n\} \\ \vec{B} &= \{a_1, \dots, a_n\} \end{aligned} \quad (17)$$

With the optimization problem defined in Equation 17, the binary MOLA could already be applied. However, if all 2266 nodes were considered, the problem would have an optimization problem with 2266 variables. The greater the dimensionality, regardless of the meta-heuristic to be used, more complex is the optimization problem. Looking at Figure 6c, it is possible to notice the number of extremely close nodes and obviously, not all of them need to be evaluated. Therefore, some simplifications can be made to improve the optimization process. 34 well-spaced and distributed sensors will be considered. They composed the vector of candidate sensors \vec{S}_c and are represented in Figure 9.



Figure 9: Candidate nodes for sensor allocation (Pereira *et al.*, [2022b])

Another important consideration and simplification of this work is that, as can be seen in Figure 1, the MRB has vibration modes with spatial displacements, that is, in all directions. Knowing that the FEM delivers the nodal displacements in the x , y and z directions, the total displacement will be considered in the optimization problem, calculated using Euclidian distance.

Having defined the optimization problem of Equation 17 and adopted some simplifications strategies, the binary MOLA can be applied with the parameters of Table 3 for all the objective functions of Table 1, considering the first five mode shapes of the MRB, resulting in 7 multi-objective optimization problems. In Pereira *et al.* [2022b] was considered the first six mode shapes and this methodology was called Multi-objective Sensor Selection and Placement Optimization based on Lichtenberg Algorithm (MOSSPOLA). Just for the reader to have an idea, for each of these problems, MOSSPOLA tested $200 \times 200 = 40000$ sensor configurations in an average time of 18s. And just in this single run it will try to find at least one sensor configuration for each number of sensors, unlike the traditional SPO methodology that the number of sensors is fixed and therefore each program run is limited to only this number of sensors. The Pareto fronts found are shown in Figure 10.

Table 3: Tuned MOLA parameters

R_c	N_p	S	ref	Pop	N_{iter}	N_{grid}	N_s
200	10^6	1	0.4	200	200	30	100

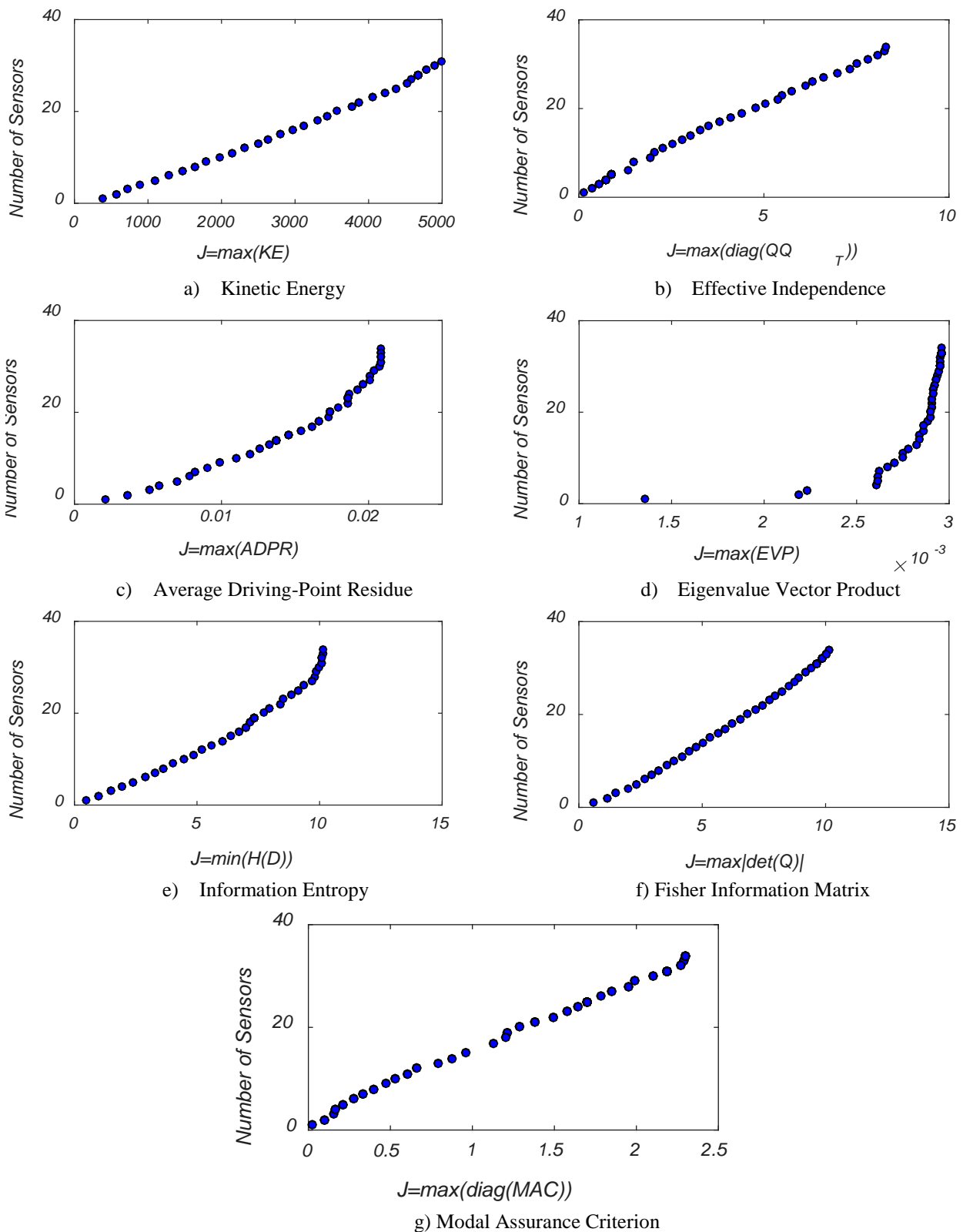


Figure 10: Answering how MRB information varies with the number of sensors using Pareto front for seven SPO metrics and the first five mode shapes

Interesting discussions can be made from Figure 10, remembering that these Pareto fronts refer exclusively to the MRB and its first 5 mode shapes. The first is that the applied methodology was able to successfully find at least one sensor configuration for each number of sensors in all metrics. It was also noticed that for all metrics, the increase in the number of sensors will always lead to an increase in the level of information acquired from the MRB.

For the KE, Efl, and FIM metrics, the MRB information level grows linearly with the increase in the number of sensors up to the number 34, the interval used in this study. Already the ADPR, IE, and MAC metrics have a smoothly convex behavior and in both the information level does not significantly improves from 25 sensors. The EVP metric was the one that showed the most information gain for the first 4 sensors, and from this the information level does not grow significantly anymore too. Still, it is possible to confirm with this methodology the relationship between some metrics, already found in other studies: *i)* Efl, KE, and FIM (Gomes *et al.* [2018b]; Li *et al.* [2007]), and *ii)* MAC and IE (Yi & Li [2012]).

Each solution of these Pareto fronts has a sensor configuration with its number and positions. It would be exhausting and tiring to present every sensor configuration found for every metric. Considering that the number of mode shapes used is 5 and it is recommended that the number of sensors used is at least equal to this number, Figure 11 shows all solutions for 5 sensors. Figures 12 and 13 bring the solutions for 7 and 9 sensors respectively. Remembering that good damage identification (delamination) results have already been found using only 10 well-distributed sensors without any SPO study (Gomes *et al.* [2020b]). The scale that shows the distances of the sensors on the blade is in meters.

It is possible to see that in Figures 11, 12 and 13 that again all metrics allocate at least one sensor at the end of the MRB, with the IE for five sensors (Figure 11e) being the only one that not selected the node at the lower tip, furthest from the root base and point of greater flexibility. Can be also seen that the metrics still prefer to select nodes in the central and tip regions, with the ADPR metric being the only one to select at least one point in the root section in all cases. And this pattern remains even increasing the number of sensors (from 5 to 7 and 9).

Figure 11 represents the multi-objective SPO solutions for 5 sensors, which can be compared with the 5-sensor solutions of the single-objective SPO of Figure 7. Can be seen that the sensor configurations for the metrics KE, Efl, ADPR, EVP, and IE are very similar, though not identical. As for the FIM metric, although both allocated more than 40% of the sensors in the tip, in the multi-objective version it is clear that the metric allocated more sensors in the central region. The opposite happened with the MAC, which in the multi-objective version allocated even more sensors in the tip than in the central section. These are not significant differences, but they did happen.

Therefore, still discussing Figure 11, KE, ADPR, EVP, and FIM were the only metrics that had better sensor distribution, as they divided the available sensors well in the most energetic regions of the blade, at the extreme tip and exactly in the middle. Therefore, still discussing Figure 11, KE, EVP and FIM were the only metrics to distribute the sensors well in the regions of greater displacement, center and tip. Again, the APR metric was the only one to allocate sensors at the root and the MAC metric was the one that put the most sensors at the tip.

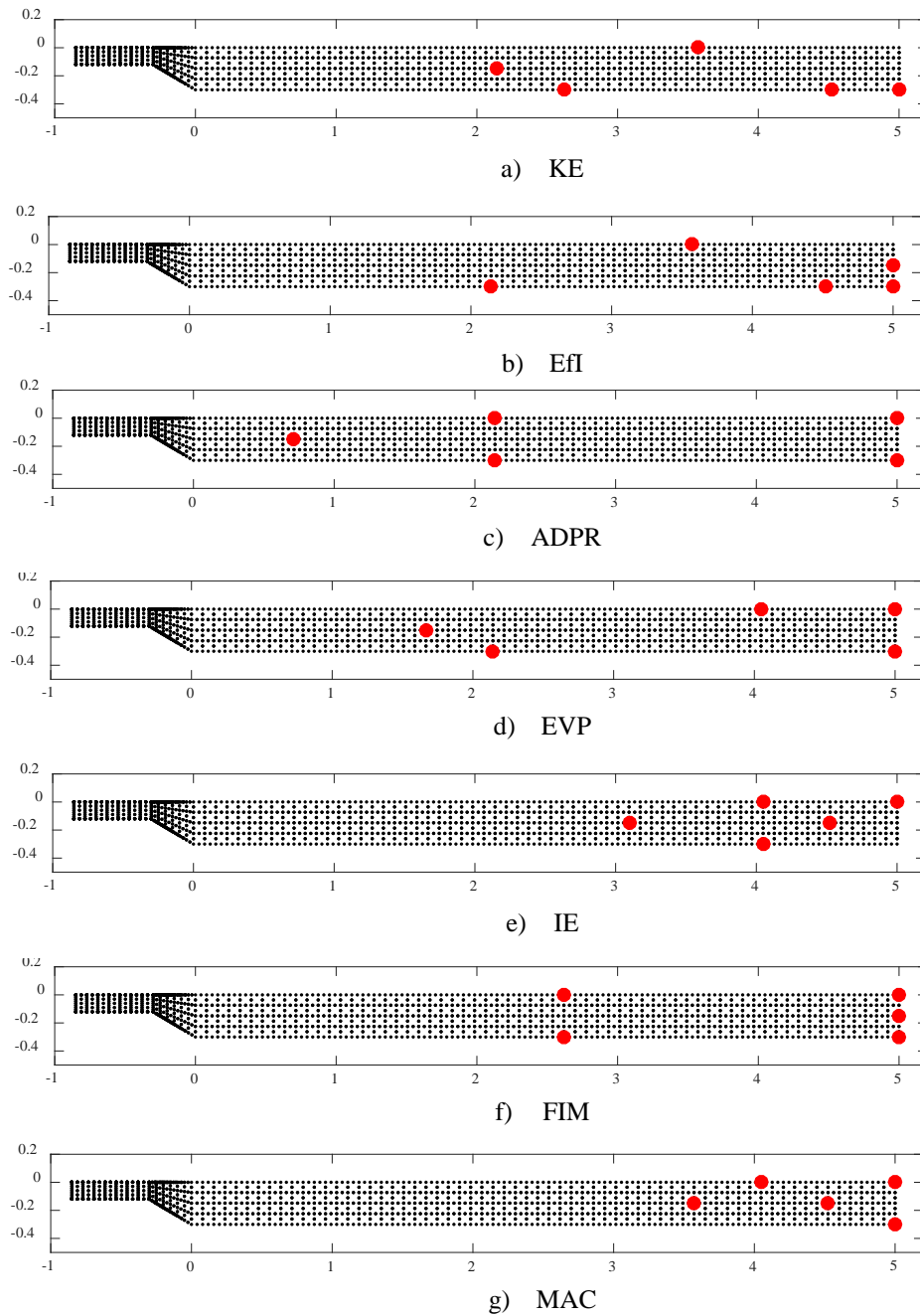


Figure 11: Best sensor configuration for five sensors in different SPO metrics.

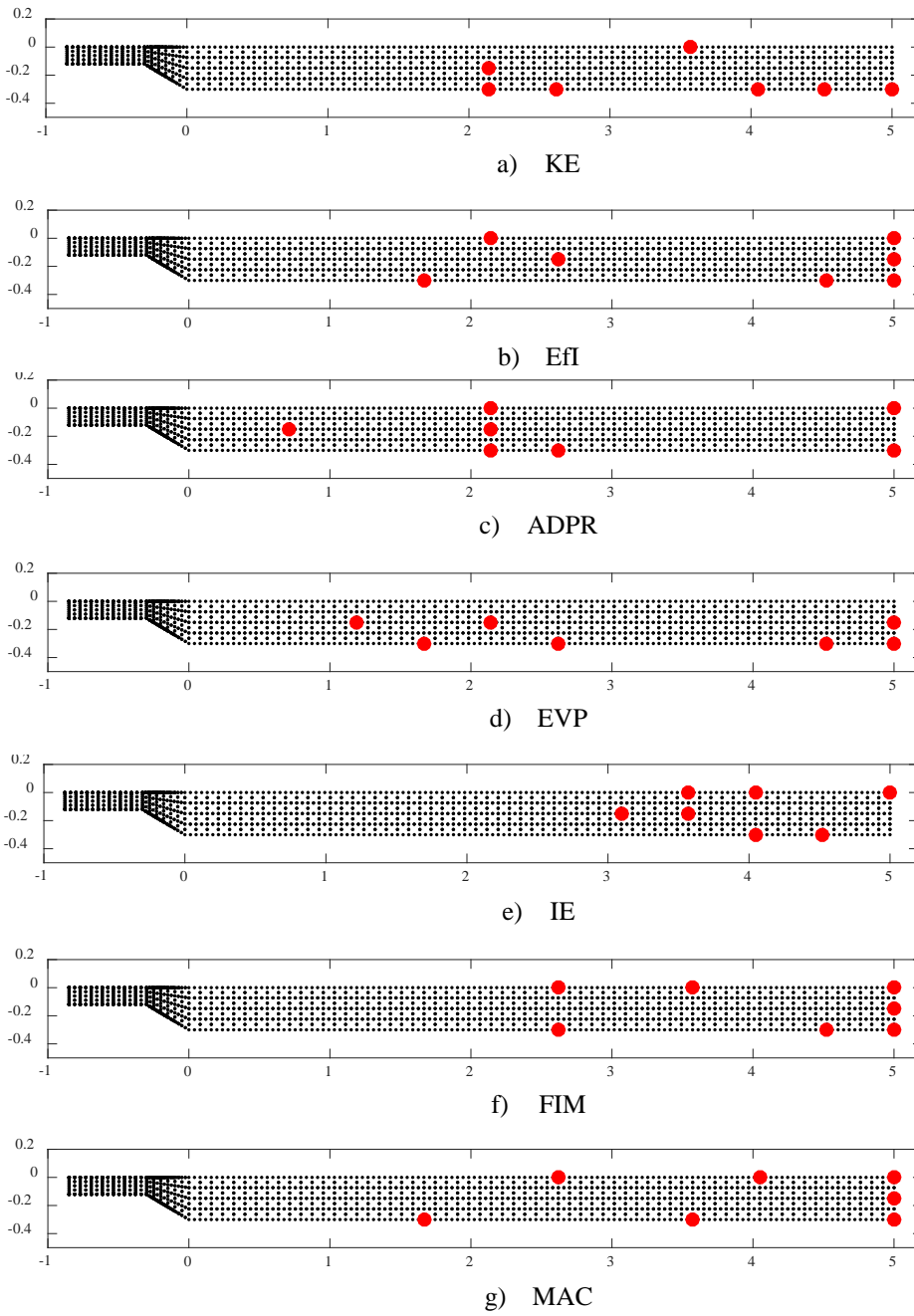


Figure 12: Best sensor configuration for seven sensors in different SPO metrics.

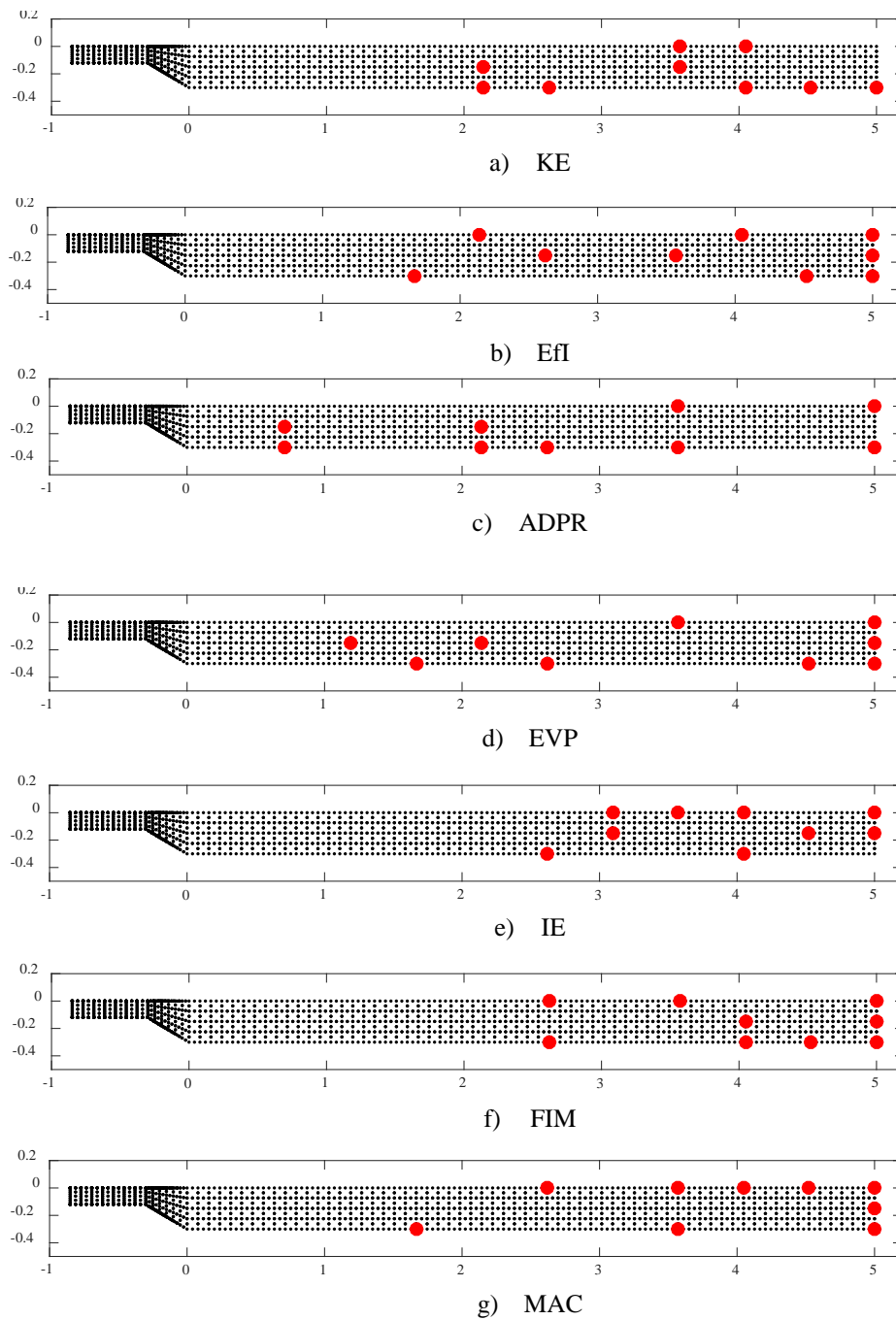


Figure 13: Best sensor configuration for nine sensors in different SPO metrics.

Increasing the number of sensors to 7 (Figure 12), all metrics maintain the same sensor distribution pattern, with the exception of MAC and EfI, which better distributes sensors between the central and root sections. Note that as shown in the Pareto fronts of Figure 10, these metrics have increased sensing in the central region, raising the information level of the entire blade. So, these join the metrics KE, EVP and FIM, which continue to have the best sensor distributions. IE continues to place all sensors in the tip section and ADPR still has two sensors at the tip and the majority between the root and beginning of the central section. The EVP metric now has sensors closer to the root. Considering 9 sensors, the offer of sensors is considerable and all metrics are distributed more evenly between the higher displacement sections of the blade. However, the IE and FIM metrics continue to allocate most of the sensors between the exact middle of the blade and the tip, smoothly accompanied by the MAC.

Being damage identification the main purpose of sensing, it is known that in general, the closer the damage is to the sensor, the easier it is to diagnose. Therefore, a good distribution of sensors implies in a more efficient HUMS for all blade sections. The only metrics that had a good sensors distribution for the number of sensors analyzed and at the same time always had a greater allocation of sensors in regions with greater displacements due to mode shapes were KE and EVP.

Figure 14 shows how the sensor distribution is modified for different numbers of sensors for these two metrics, starting from 1 to 30 sensors. As discussed before, the node with the greatest flexibility and therefore the greatest range of motion is at the tip, at the vertex opposite to the blade attachment by the root, and both metrics consider this point as the point of maximum information for a single sensor. From then on, the EVP metric allocates more sensors exactly at the tip and at the same time it is the one that has sensors closer to the root. Therefore, because it is more distributed, the EVP metric has the lowest sensor density. However, both highly value the tip and center sections.

The EVP and KE metrics seem to be the most suitable for the MRB. The sensor configurations for 5 sensors considering mono and multi-objective optimization and for 7 and 9 sensors considering multi-objective optimization were presented for these metrics and for others 5 important metrics in literature.

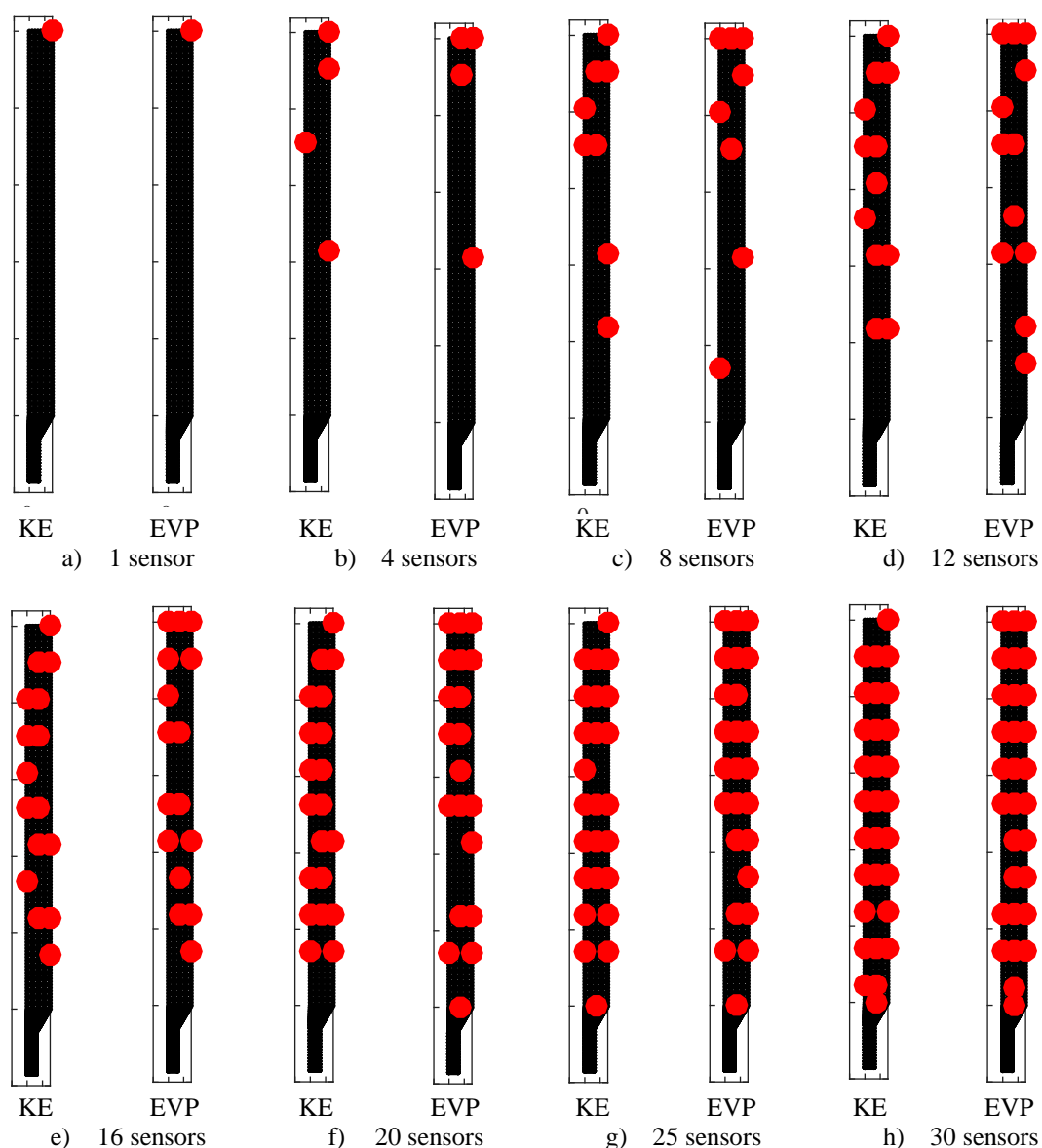


Figure 14: Sensor distribution variation as increases the sensors number

4 Concluding Remarks

This chapter showed the relevance of using rotary-wing aircraft and focused on the study of damage to its main component: the main rotor blade. A discussion of the types of damage that ranged from blade manufacture to damage that can arise in use was made. It also detailed how modal data can help engineers analyze the presence of damage in structures and showed how the study of sensor positioning is crucial for a good Health and Usage Monitoring System. This chapter presented the main sensor placement optimization metrics most used in the literature, this chapter used mono and multi-objective optimizations to find the best sensor configurations for the blade. The techniques used were developed in the Research Group in Computational Mechanics (GEMEC) at UNIFEI.

5 Perspectives for Future Research Work

The future research suggestions in this chapter relate mainly to two fronts: continuation of the SPO theme and adjustment of the numerical model in the FEM to better fit the reality of the MRB.

5.1 Sensor placement optimization continuation

The study of SPO in MRB carried out in this chapter follows the line of more recent works for different structures in the literature. A meta-heuristic is employed and the objective is to find the sensor positions that have the largest displacements for the mode shapes to be considered. Obviously, these positions will also represent the most significant changes in a possible damage presence. Furthermore, in the optimization process, only the mode shapes and natural frequencies of a single FEM simulation are considered: the healthy and intact blade. This brings one of the main advantages of this methodology: finding sensor configurations in complete optimization simulations in less than 3 minutes.

However, this methodology is unable to answer, for example, which of all sensor configurations in Figures 7, 11, 12, and 13 is more accurate in finding damage in any blade section. Could the accuracy of a specific sensor configuration change for damage in different regions? Could sensor configurations that are more present in the tip section be better at diagnosing damage anywhere on the blade? The answers to these questions can only be confirmed after damage identification studies using the sensor configurations found.

In this case, meta-heuristics have also been widely applied, but the objective functions may be different from those in Table 1, although they can also be used. In this case, Equation 4 can be used in the FEM to induce damage with well-defined intensity (α) and location (n_e) in any structures with N_e elements, with the damage rate varying between 0 and 100% (structure intact). An example of objective function widely used to guide the meta-heuristic in a damage identification problem is represented in Equation 18, where this optimization problem is also defined (Pereira *et al.* [2021]).

$$\min J_{\Phi} = \sum_{i=1}^n \sqrt{\left(1 - \frac{\Phi_{i,s}^{MH}}{\Phi_{i,s}^{real}}\right)^2}$$

(18)

subject to:

$$1 \leq n_e \leq N_e$$

$$0 \leq \alpha \leq 1$$

where $\Phi_{i,s}^{MH}$ are the nodal displacements in mode shape i in the sensors found in the SPO problem for random damages suggested by some meta-heuristic and $\Phi_{i,s}^{real}$ is the reading of the nodal displacements in the real structure in the same node positions, which could be the MRB. Or, if it's a numerical simulation problem, $\Phi_{i,s}^{real}$ could be the displacement data of the mode shapes using the same FEM structure with a known and induced damage. The meta-heuristic optimizes suggesting damages, calculating their

modal data, and comparing them with the real one. When the difference is zero, the damage will be identical. Note that for each point in the population in each iteration, the FEM is activated, which makes the optimization process to identify damages computationally expensive. A single simulation can take just over a day.

Nonetheless, this is the only way to validate the accuracy of the sensor configurations found. In normal way and as has been discussed in the literature, each of the sensor configurations of Figures 7, 11, 12, and 13 can have three damages induced in the root, tip and center sections. At least S simulations can be run for each damage and each sensor configuration and can be calculated which of these had the smallest error in relation to α and n_e . A similar way frequently used referring mainly to n_e is the Probability of Detection (PoD) (Diacenco [2016]). This is a measure that is the rate of times that the optimization problem of Equation 18 was executed and found the exact location of the damage n_e in relation to all trials, which can also be used as an objective function to be maximized.

Therefore, two interesting directions can be suggested in relation to the SPO study in this chapter, in order of increasing computational cost: *i*) Do a general damage identification study with the sensor configurations found in Figures 7, 11, 12, and 13 to see which has the smallest error or the largest PoD and *ii*) Use the methodology presented with MOLA and Feature Selection for multi-objective SPO and for each sensor configuration proposed, run the optimization problem of Equation 18 S times and calculate the PoD or error, which will be the second objective function instead of the main SPO metrics. Note that it may have an exhausting computational cost, but to the best knowledge of the authors it has never been done before and it can bring promising results, since the first objective is the number of sensors.

Some suggestions regarding the algorithms can also be presented: *i*) Test other meta-heuristics in place of MOLA and compare the results of Figure 10 using robust metrics for comparison of Pareto fronts when there is no true Pareto front as a reference, such as HyperVolume (Kumar *et al.*, [2021]), *ii*) Use chaos theory or lévy flights in MOLA to try to improve the Pareto fronts results, *iii*) Test other methods of binarizing the continuous values of MOLA and also compare Pareto fronts to try to improve the results, *iv*) apply binary MOLA on different/new SPO metrics, *v*) apply the SPO methodology presented in different structures, since both the method and the application in MRB are new and have not yet been applied in any other structure. This includes repeating all previous items, among others.

5.2 Improvement of the finite element method and its alignment with reality

Still on the continuations of the SPO problem presented in this chapter, future suggestions can be added regarding modeling in FEM. Although a numerical model of the MRB with NACA0012 profile and geometric characteristics similar to the real AS-350 blade was used, to the detriment of many works in the literature that used simply a beam, simplifications were made and may influence the SPO result.

As can be seen in Figure 6a, the MRB is simply fixed by the root surface and fixation by screws in the head rotor, which usually have a metallic jacket, was not modeled. Still, in fact, the MRB is connected to the hub via three rings (and Degrees of Freedom):

flapping, drag, and pitch-change, which allows movement in these same directions in flight. See Figure 4.

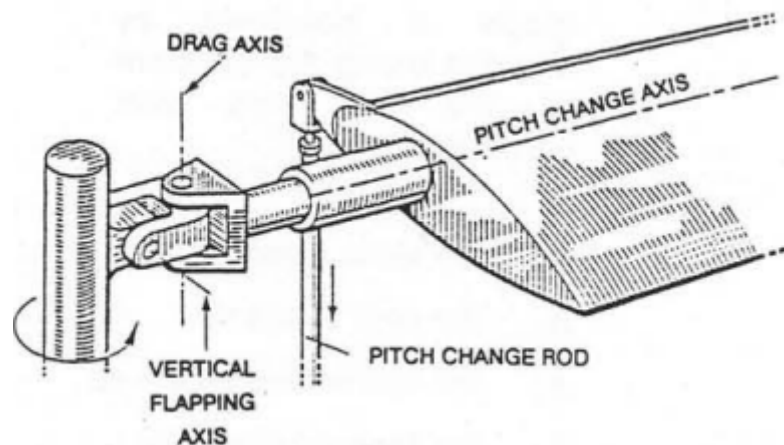


Figure 15: Axes of freedom for the MRB (Bramwell et al. [2001])

Still comparing Figures 6a and 5, it is also possible to observe that some sections have a spline for corner smoothing, which is also not included in the numerical model. The same occurred with the leading edge that has a protective steel frame. These are geometric and boundary conditions details.

The following simplification can significantly affect the SPO results and is about the dynamic modeling. As seen and discussed, the natural frequencies and mode shapes were calculated considering that the blade is simply static and supported by the root at the hub. However, there are several sources of forces and vibrations in a helicopter in flight and can change the MRB rigidity. The most important is the centrifugal force due to the rotation of the rotor. With it, a traction stress is added to each element of the blade, getting bigger from the tip to the root section of the MRB.

The common and unavoidable sources of vibrations in the main blade that would be welcome at the FEM are: *i*) variable aerodynamic forces resulting from the cyclic movement of the blades in translational flight, *ii*) centrifugal forces from the center of gravity variation position due the flapping motion of the blades, and *iii*) aerodynamic forces from the air flow in rotors, among others. These concepts have been better discussed in earlier chapters.

Other studies can also be done in relation to the type of induced damage. As a suggestion of section 5.1, a case study of damage identification degrading the elastic property of a single element was shown, that is widely used in literature to simulate delamination. However, other types of damage can be modeled, such as circular holes (which can represent ballistic damage) and cracks, to verify if the sensor configurations found are capable of diagnosing them.

Acknowledgements

The authors would like to acknowledge the financial support from the Brazilian agency CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), and FAPEMIG (Fundação de Amparo à Pesquisa do Estado de Minas Gerais - APQ-00385-18).

References

- Ahmad, K., Baig, Y., Rahman, H., & Hasham, H. J. (2020). Progressive failure analysis of helicopter rotor blade under aeroelastic loading. *Aviation*, 24(1), 33-41.
- Alkayem, N. F., Cao, M., Zhang, Y., Bayat, M., & Su, Z. (2017). Structural damage detection using finite element model updating with evolutionary algorithms: a survey. *Neural Computing and Applications*, 30(2), 389–411.
- Amafabia, D. Montalvão, O. David-West, G. Haritos, A Review of Structural Health Monitoring Techniques as Applied to Composite Structures, in *SDHM Structural Durability and Health Monitoring*, Vol. 11, No. 2, pp. 91-147, (2017).
- Assis, F. M., & Gomes, G. F. (2021). Crack identification in laminated composites based on modal responses using metaheuristics, artificial neural networks and response surface method: a comparative study. *Archive of Applied Mechanics*.
- Barthorpe RJ, Worden K (2009) Sensor placement optimization. *Encyclopedia of Structural Health Monitoring*.
- Benner P, Herzog R, Lang N, Riedel I, Saak J (2017) Comparison of model order reduction methods for optimal sensor placement for thermo-elastic models.
- Bramwell, A.R.S., Done, J., Balmford, D., “Bramwell’s helicopter dynamics”, Reed Educational and Professional Publ. Ltd, 2001, 373 p
- Braun, C. E., L.D. Chiwiacowsky, A.T. Gomez, Variations of Ant Colony Optimization for the solution of the structural damage identification problem, *Procedia Comput. Sci.* 51 (2015) 875–884
- Burgos, D. A. T., Vargas, R. C. G., Pedraza, C., Agis, D., & Pozo, F. (2020). Damage Identification in Structural Health Monitoring: A Brief Review from its Implementation to the Use of Data-Driven Applications. *Sensors*, 20(3), 733.
- Cao, X., Chen, J., Xu, Q., & Li, J. (2020). A distance coefficient-multi objective information fusion algorithm for optimal sensor placement in structural health monitoring. *Advances in Structural Engineering*, 136943322096437.
- Carne TG and Dohrmann CR. A modal test design strategy for model correlation. Albuquerque, NM: Sandia National Labs, 1994.

- Cattarius, J., & Inman, D. J. (2000). Experimental verification of intelligent fault detection in rotor blades. *International Journal of Systems Science*, 31(11), 1375–1379.
- Challan, M., Jeet, S., Bagal, D. K., Mishra, L., Pattanaik, A. K., & Barua, A. (2022). Fabrication and mechanical characterization of red mud based Al2025-T6 MMC using Lichtenberg optimization algorithm and Whale optimization algorithm. *Materials Today: Proceedings*, 50, 1346-1353.
- Chen, L. Yu, A new structural damage detection strategy of hybrid PSO with Monte Carlo simulations and experimental verifications, *Measurement* 122 (2018) 658–669.
- Coello, C. A. C., Lamont, G. B. & Lechuga, M. S. (2004). Handling multiple objectives with particle swarm optimization. *Evolutionary Computation, IEEE Transactions on*, 8(3), 256-279.
- Colombo, M.D. Todd, C. Sbarufatti, M. Giglio. (2022). On statistical Multi-Objective optimization of sensor networks and optimal detector derivation for structural health monitoring. *Mechanical Systems and Signal Processing*. Volume 167, Part A , 108528. ISSN 0888-3270.
- Crawford, R. Soto, G. Astorga, J. García, C. Castro, F. Paredes, Putting continuous metaheuristics to work in binary search spaces, *Complexity* 2 (2017) 1–19.
- Diacenco, A. A. Optimization Modeling Applied to the Sensor Topology for Identification and Localization of Damages in Composite Plate (In Portuguese: Modelagem de Otimização Aplicada à Topologia de Sensores para Identificação e Localização de Danos em Placa Compósita). 2016. Doctoral Thesis, Federal University of Itajubá, MG. <https://repositorio.unifei.edu.br/jspui/handle/123456789/582>
- Ferentinos, K. P. & Tsiligiridis, T. A. (2007). Adaptive design optimization of wireless sensor networks using genetic algorithms. *Computer Networks*, 51(4), 1031–1051.
- Francisco MB, Junqueira DM, Oliver GA, Pereira JLJ, da Cunha SS, Gomes GF. Design optimizations of carbon fibre reinforced polymer isogrid lower limb prosthesis using particle swarm optimization and Lichtenberg algorithm. *Eng Optim* 2020.
- Ghosh, K. K., Ritam Guha, R., Bera S. K. S-Shaped versus V-Shaped Transfer Functions for Binary Manta Ray Foraging Optimization in Feature Selection Problem. *Research Square*, 2020. DOI 10.21203/rs.3.rs-28684/v1
- Ghoshal, A., Harrison, J., Sundaresan, M. J., Hughes, D., & Schulz, M. J. (2001). Damage detection testing on a helicopter flexbeam. *Journal of Intelligent Material Systems and Structures*, 12(5), 315–330
- Gomes, G. F., Almeida, F. A., Silva Lopes Alexandrino, P., Cunha, S. S., Sousa, B. S., & Ancelotti, A. C. (2018). A multi-objective sensor placement optimization for SHM

systems considering Fisher information matrix and mode shape interpolation. *Engineering with Computers*.

Gomes, G. F., S.S. da Cunha, P.D.S.L. (2018b). Alexandrino, B.S. de Sousa, A.C. Ancelotti, Sensor placement optimization applied to laminated composite plates under vibration, *Struct. Multidiscipl. Optimiz.* 58 (5) 2099–2118.

Gomes G. F., Sebastiao Simões da Cunha Jr., Antonio Carlos Ancelotti Jr., A sunflower optimization (SFO) algorithm applied to damage identification on laminated composite plates, *Eng. Comput.* 35 (2) (2019) 619–626

Gomes, G. F., & Giovani, R. S. (2020). An efficient two-step damage identification method using sunflower optimization algorithm and mode shape curvature (MSDBI–SFO). *Engineering with Computers*.

Gomes, G. F., Souza Chaves, J. A., & de Almeida, F. A. (2020b). An inverse damage location problem applied to AS-350 rotor blades using bat optimization algorithm and multi-axial vibration data. *Mechanical Systems and Signal Processing*, 145, 106932.

Gopalakrishnan *et al.* (2011) Srinivasan Gopalakrishnan, Massimo Ruzzene, e Sathyanarayana Hanagud. Computational techniques for damage detection, classification and quantification. Em *Computational Techniques for Structural Health Monitoring*, pages 407–461. Springer.

Guan et al. (2017) Xuefei Guan, Yongxiang Wang, e Jingjing He. A probabilistic damage identification method for shear structure components based on cross-entropy optimizations. *Entropy*, 19(1):27.

Hearn & Testa (1991) George Hearn e Rene B Testa. Modal analysis for damage detection in structures. *Journal of Structural Engineering*, 117(10):3042–3063.

Heo G, Wang M and Satpathi D. Optimal transducer placement for health monitoring of long span bridge. *Soil Dyn Earthq Eng* 1997; 16(7–8): 495–502

Hou, W., & Zhang, W. Advanced Composite Materials defects/damages and health monitoring, in *Proceedings of the IEEE 2012 Prognostics and System Health Management Conference (PHM-2012 Beijing)*, Beijing, pp. 1-5, (2012).

Jin H, Xia J and Wang YQ. Optimal sensor placement for space modal identification of crane structures based on an improved harmony search algorithm. *J Zhejiang Univ: Sc A* 2015; 16(6): 464–477.

Kammer DC. Sensor placement for on-orbit modal identification and correlation of large space structures. *J Guid Control Dynam* 1991; 14(2): 251–259.

Kammer DC. Effects of noise on sensor placement for on-orbit modal identification of large space structures. *J Dyn Syst* 1992; 114(3): 436–443.

- Kammer DC and Tinker ML. Optimal placement of triaxial accelerometers for modal vibration tests. *Mech Syst Signal Pr* 2004; 18(1): 29–41.
- Khabarov, S. S., & Komshin, A. S. (2021). Fiber-Optic Measurement Technology and the Phase-Chronometric Method for Controlling and Monitoring the Technical Condition of Aircraft Structures. *Measurement Techniques*, 64(2), 131–138.
- Kiddy, J., & Pines, D. (2001). Experimental validation of a damage detection technique for helicopter main rotor blades. *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, 215(3), 209–220.
- Latifi, F.P. van der Meer, L.J. Sluys, A level set model for simulating fatigue-driven delamination in composites, *International Journal of Fatigue*, Vol. 80, pp. 434–442, (2015)
- Li D, Li H and Fritzen C. The connection between effective independence and modal kinetic energy methods for sensor placement. *J Sound Vib* 2007; 305(4–5): 945–955.
- Lin, J.-F., Xu, Y.-L., & Law, S.-S. (2018). Structural damage detection-oriented multi-type sensor placement with multi-objective optimization. *Journal of Sound and Vibration*, 422, 568–589.
- Milad Jahangiri, M.A. Najafgholipour, S.M. Dehghan, M.A. Hadianfard. (2019) The efficiency of a novel identification method for structural damage assessment using the first vibration mode data, *J. Sound Vibr.* 458. 1–16.
- Mirjalili S, Lewis A (2013) S-shaped versus V-shaped transfer functions for binary particle swarm optimization. *Swarm Evolut Comput* 9:1–14
- Mishra, S.K. Barman, D. Maity, D.K. Maiti, Ant lion optimisation algorithm for structural damage detection using vibration data, *J. Civil Struct. Health Monitor.* 9 (1) (2019) 117–136.
- Mohanty, A., Nag, K. S., Bagal, D. K., Barua, A., Jeet, S., Mahapatra, S. S., & Cherkia, H. (2022). Parametric optimization of parameters affecting dimension precision of FDM printed part using hybrid Taguchi-MARCOS-nature inspired heuristic optimization technique. *Materials Today: Proceedings*, 50, 893-903.
- Pan, L. Yu, Z.P. Chen, W.F. Luo, H.L. Liu, A hybrid self-adaptive Firefly-Nelder-Mead algorithm for structural damage detection, *Smart Struct. Syst.* 17 (6) (2016) 957–980.
- Pantano, Cohesive Model for the Simulation of Crack Initiation and Propagation in Mixed-Mode I/II in Composite Materials, *Applied Composite Materials*, Vol. 26, No. 4, pp. 1207-1225, (2019).
- Pawar, P. M., & Ganguli, R. (2003). Genetic fuzzy system for damage detection in beams and helicopter rotor blades. *Computer Methods in Applied Mechanics and Engineering*, 192(16-18), 2031–2057.

- Pawar, P., & Ganguli, R. (2005). On the effect of matrix cracks in composite helicopter rotor blade. *Composites Science and Technology*, 65(3-4), 581–594.
- Pawar, P., & Ganguli, R. (2007). Fuzzy-Logic-Based Health Monitoring and Residual-Life Prediction for Composite Helicopter Rotor. *Journal of Aircraft*, 44(3), 981–995.
- Pawar, P. M., & Sung Nam Jung. (2007b). Support Vector Machine based Online Composite Helicopter Rotor Blade Damage Detection System. *Journal of Intelligent Material Systems and Structures*, 19(10), 1217–1228.
- Pawar, P. M., & Ganguli, R. (2007c). On the effect of progressive damage on composite helicopter rotor system behavior. *Composite Structures*, 78(3), 410–423.
- Pereira JLJ, Chuman M, Cunha SS, Gomes GF. Lichtenberg optimization algorithm applied to crack tip identification in thin plate-like structures. *Eng Comput (Swansea, Wales)* 2020;38:151–66. <https://doi.org/10.1108/EC-12-2019-0564>.
- Pereira JLJ, Francisco MB, Cunha SS da, Gomes GF. (2021). A powerful Lichtenberg Optimization Algorithm: A damage identification case study. *Eng Appl Artif Intell*; 97:104055.
- Pereira JLJ, Francisco MB, Diniz CA, Antônio Oliver G, Cunha SS, Gomes GF. Lichtenberg algorithm: A novel hybrid physics-based meta-heuristic for global optimization. *Expert Syst Appl* 2021a;170.
- Pereira, J.L.J., Oliver, G.A., Francisco, M.B. *et al.* (2021b). A Review of Multi-objective Optimization: Methods and Algorithms in Mechanical Engineering Problems. *Arch Computat Methods Eng*.
- Pereira JLJ, Guilherme Antônio Oliver, Matheus Brendon Francisco, Sebastião Simões Cunha Jr, Guilherme Ferreira Gomes. (2022). Multi-objective lichtenberg algorithm: A hybrid physics-based meta-heuristic for solving engineering problems. *Expert Systems with Applications*. Volume 187, 115939, ISSN 0957-4174.
- Pereira, J.L.J., Francisco, M.B., Ribeiro, R.F. et al. Deep multiobjective design optimization of CFRP isogrid tubes using lichtenberg algorithm. *Soft Comput* (2022a)
- Pereira, J.L.J., Francisco, M.B., Souza Chaves, J. A., Sebastião Simões Cunha Jr & Gomes, G. F. (2022b). Multi-objective sensor placement optimization of helicopter rotor blade based on feature selection. *Mechanical Systems and Signal Processing*. Volume 180. 109466. ISSN 0888-3270. <https://doi.org/10.1016/j.ymssp.2022.109466>.
- Shahani, A. R., & Mohammadi, S. (2015). Damage tolerance and classic fatigue life prediction of a helicopter main rotor blade. *Meccanica*, 51(8), 1869–1886.
- Tian, Z. & Wang, J. Variable frequency wind speed trend prediction system based on combined neural network and improved multi-objective optimization algorithm. *Energy*. Volume 254. Part A. 2022.

- Kaveh, M. Maniat, Damage detection based on MCSS and PSO using modal data, *Smart Struct. Syst.* 15 (5) (2015) 1253–1270
- Rao ARM, Lakshmi K, Kumar SK (2015) Detection of delamination in laminated composites with limited measurements combining pca and dynamic qpso. *Adv Eng Softw* 86:85–106.
- Reddy, R. R. K., & Ganguli, R. (2003). Structural damage detection in a helicopter rotor blade using radial basis function neural networks. *Smart Materials and Structures*, 12(2), 232–241.
- Santos, F. B., Peeters, H. Van der Auweraer, and L. C. S. Góes, “Experimental damage detection of a helicopter main rotor blade based on modal properties,” in *Proceedings of the International Conference on Structural Engineering Dynamics (ICEDyn 2013)*, Sesimbra, Portugal, 17–19 June 2013.
- Santos, F. L. M., Peeters, B., Van der Auweraer, H., Góes, L. C. S., & Desmet, W. (2016). Vibration-based damage detection for a composite helicopter main rotor blade. *Case Studies in Mechanical Systems and Signal Processing*, 3, 22–27.
- Sharma, Manik; KAUR, Prableen. A Comprehensive Analysis of Nature-Inspired Meta-Heuristic Techniques for Feature Selection Problem. *Archives of Computational Methods in Engineering*, v. 28, n. 3, 2021.
- Souza, T. A. Z., J. L. J. Pereira, M. B. Francisco, C. A. R. Sotomonte, B. Jun Ma, G. F. Gomes & C. J. R. Coronado (2022) Multi-objective optimization for methane, glycerol, and ethanol steam reforming using lichtenberg algorithm, *International Journal of Green Energy*
- Tamer, A., Zanoni, A., Cocco, A. et al. A numerical study of vibration-induced instrument reading capability degradation in helicopter pilots. *CEAS Aeronaut J* 12, 427–440 (2021).
- Tan, Y., & Zhang, L. (2019). Computational methodologies for optimal sensor placement in structural health monitoring: A review. *Structural Health Monitoring*, 147592171987757.
- Vinet, D. Gamby, Prediction of long-term mechanical behaviour of fibre composites from the observation of micro-buckling appearing during creep compression tests, *Composites Science and Technology*, Vol. 68, No. 2, pp. 526–536, (2008).
- Voicu, Andrei-Daniel, Hadăr, Anton, Vlăsceanu, Daniel and Pastramă, Ștefan-Dan. (2020). SHM Monitoring Methods and Sensors with Applications to Composite Helicopter Blades: A Review. *Acta Universitatis Cibiniensis. Technical Series*, vol.72, no.1, pp.1-11.
- Worden & Friswell (2009) Keith Worden e Michael I Friswell. Modal–vibration-based damage identification. *Encyclopedia of Structural Health Monitoring*.

- Yang C, Lu Z (2017) An interval effective independence method for optimal sensor placement based on non-probabilistic approach. *Sci China Technol Sci* 60(2):186–198.
- Yi TH and Li HN. Methodology developments in sensor placement for health monitoring of civil infrastructures. *Int J Distrib Sens N* 2012; 8(8): 612726.
- Yuen K-V, Kuok S-C (2015) Efficient bayesian sensor placement algorithm for structural identification: a general approach for multi-type sensory systems. *Earthq Eng Struct Dyn* 44(5):757–774
- Zenzen, I. Belaidi, S. Khatir, M.A. Wahab, A damage identification technique for beam-like and truss structures based on FRF and Bat Algorithm, *Comptes Rendus Mécanique* 346 (12) (2018) 1253–1266.
- Zhou, G. D., Yi, T.-H., Xie, M.-X., Li, H.-N., & Xu, J.-H. (2021). Optimal Wireless Sensor Placement in Structural Health Monitoring Emphasizing Information Effectiveness and Network Performance. *Journal of Aerospace Engineering*, 34(2), 04020112.

Chapter 9

Challenges for Structural Health Monitoring: Nonlinearities and Uncertainties

Chapter details

Chapter DOI:

<https://doi.org/10.4322/978-65-86503-88-3.c09>

Chapter suggested citation / reference style:

Villani, Luis G. G., et al. (2022). “Challenges for Structural Health Monitoring: Nonlinearities and Uncertainties”. In Jorge, Ariosto B., et al. (Eds.) *Uncertainty Modeling: Fundamental Concepts and Models*, Vol. III, UnB, Brasilia, DF, Brazil, pp. 258–278. Book series in Discrete Models, Inverse Methods, & Uncertainty Modeling in Structural Integrity.

P.S.: DOI may be included at the end of citation, for completeness.

Book details

Book: Uncertainty Modeling: Fundamental Concepts and Models

Edited by: Jorge, Ariosto B., Anflor, Carla T. M., Gomes, Guilherme F., & Carneiro, Sergio H. S.

Volume III of Book Series in:

Discrete Models, Inverse Methods, & Uncertainty Modeling in Structural Integrity

Published by: UnB City: Brasilia, DF, Brazil Year: 2022

DOI: <https://doi.org/10.4322/978-65-86503-88-3>

Challenges for Structural Health Monitoring: Nonlinearities and Uncertainties

Luis G. G. Villani^{1*}, Sidney B. Shiki², Americo Cunha Jr³,
Samuel da Silva^{4*} and Michael D. Todd⁵

¹Universidade Federal do Espírito Santo (UFES), Brazil. e-mail: luis.villani@ufes.br

²Universidade Federal de São Carlos (UFSCar), Brazil, bruce@ufscar.br

³Universidade do Estado do Rio de Janeiro (UERJ), Brazil, americo@ime.uerj.br

^{4*}Universidade Estadual Paulista (UNESP), Brazil, samuel.silva13@unesp.br

⁵University California San Diego, EUA, mdtodd@eng.ucsd.edu

*samuel.silva13@unesp.br

Abstract

Implementing structural health monitoring (SHM) techniques with a high Technology Readiness Level (TRL) is still challenging due to several practical requirements and assumptions to apply the fundamental methods. Between them, two issues earn some special attention: the linearity hypothesis and the robustness to the natural variability of data. The first point to overcome is that many structural engineering systems inherently behave nonlinearly during operation, even in a healthy state. Here, the assumption of linearity is typically inaccurate, eliminating large classes of feature extraction techniques. This issue is more complicated when the Damage also induces additional nonlinearities, e.g., cracking. The second aspect is the need to quantify the parameters' variation and uncertainties and signal data to interrogate the structural state. This chapter proposes introducing these challenges and some examples of addressing them in this context.

Keywords: nonlinearities; uncertainty quantification; damage detection

1 Introduction

Implementing a Structural Health Monitoring (SHM) strategy is essential to avoid and prevent catastrophic failures in civil, mechanical, and aeronautical systems. Different paradigms utilize physics-based methods (numerical models coupled with model updating, for example) and data-driven methods (where models or trends are inferred from the data alone). Regardless of approach, SHM requires sensors to be deployed on the structure to measure raw information from which it is necessary to extract features sensitive to whatever target Damage mechanisms are of interest. These features are then subject to some sort of classification or discrimination process to make informed decisions about the state of the Damage. Damage is a mechanism that modifies the properties of mass, damping, stiffness, connectivity, or boundary conditions, individually or mutually, that prevent the structure from performing its intended design functions. The primary issue

is that the variation observed in these properties can be caused by several different sources, not necessarily only by the presence of Damage. This is a lack of *specificity*. Furthermore, data itself contains many sources of noise that corrupt the measurements and propagate the error into the features and the decision space; this leads to a lack of *sensitivity*. These combined challenges of specificity and sensitivity are problematic for practical, long-term SHM applications, reducing its potential performance.

Among all practical limitations, this chapter focuses on two issues that may decrease the performance of an SHM method, which we have been investigating for the last ten years. The first is the inherent nonlinear behavior of the vibration motion of flexible and light structures. A nonlinear behavior that leads to the breaking of the superposition principle can be induced when a linear system is subjected to the presence of some types of Damage. That is the case when a crack is presented in a simple clamped-free beam when the level of force excitation is low. When the Damage is propagated, a breathing crack causes a bilinear effect that violates the system's linearity and manifests in the data vibration. Modern engineering systems are manufactured with advanced materials, which demands constitutive equations that are more complex than traditional ones breaking the assumptions of linearity even in a healthy state. Consequently, an algorithm of SHM is required to separate the nonlinear effects induced by damaged from the inherent nonlinearities.

The second challenge to overcome is the presence of uncertainties in structure monitoring. The measured data may include heavy noise and systemic uncertainties caused by boundary condition variations, property change not caused by Damage, variation in the parameters caused by temperature changes, etc. Consequently, the features extracted from measured signals can be corrupted to compute the damage indices. On the other hand, the models built to monitor the structures are also full of uncertainties in the definition of their parameters and the assumptions and simplifications adopted throughout the modeling. Therefore, to ensure statistical reliability in the results obtained, such uncertainties must be considered in the construction and testing stage of the methodologies.

Both limitations severely harm damage detection because they can increase the level of false alarms and reduce the reliability of the diagnosis. It is essential to simultaneously develop and apply an SHM methodology to deal with these concerns. The key idea to get around these problems is to choose features that are insensitive to these variations, use projections and transformations, or use extensive data for learning purposes, increasing the amount of available measured data. Unfortunately, both options can be problematic and non-trivial to be quickly implemented. This chapter introduces these challenges by exploiting two benchmarks examples in simple beam structures with nonlinear behavior and uncertainties. Data of both setups are shared to allow the reader to deal with methods to observe the effects of nonlinearities and uncertainties in the performance degradation of the SHM algorithms.

Section 2 introduces general concepts of uncertainty quantification (UQ) and demonstrates how these natural variabilities can harm the damage-sensitive features used in traditional SHM algorithms. Section 3 illustrates a general approach for integrating the assumptions and concepts of uncertainty quantification and nonlinear system identification in damage detection for SHM purposes in nonlinear systems. First, the section presents an overview of data-driven modeling and system identification (ID) to detect and understand, based on data, the systems' nonlinear behavior. Many different ID models could be chosen, but a nonparametric approach based on the Volterra series is selected because this series is a straightforward generalization of the convolution concept of linear systems and strongly correlates with the frequency response functions for higher-order components. Section 4 provides two examples of applying the described methodology using experimental data. Firstly, under conditions where a deterministic model is enough to guarantee

the detection of structural variations, where the variability of the data is small. And then a more interesting experiment, involving a more significant data variability, where the stochastic approach is needed to ensure good performance (closer to the real world). Finally, the Section 5 provides the main conclusions and propositions for future work.

2 Overview in Uncertainty Quantification (UQ) and its relevance to SHM

2.1 General definitions

The uncertainties can be classified into two main groups [Soize, 2012, 2017]:

- **Data uncertainties:** also known as *aleatory*, these uncertainties are intrinsic to scenarios with variabilities, such as noise in the measured data (experimental data) and variations concerning the nominal configuration of the structures, due to geometric imperfections, manufacturing irregularities, environmental conditions, etc. These uncertainties can not be eliminated, only better characterized;
- **Model uncertainties:** also known as *epistemic*, these uncertainties result from the limited knowledge about the model structure to be used, i.e., ignorance about the system's physics. These uncertainties can be mitigated or even eliminated by increasing knowledge about the behavior of the system in analysis.

Dealing with model uncertainty is no simple task. Although they can be reduced as the models are improved, it is practically impossible to build models that are totally faithful to the physical phenomena involved. In this case, it is interesting that the model can provide some level of information about the uncertainties involved in its own prediction estimation process. Physical models encounter great difficulty in performing this kind of estimation, which is much more common when considering machine learning-based models. In the context of this work, we assume that model uncertainties can be neglected, although we know that in reality, they will be present. This simplification is made by assuming that the chosen baseline nonlinear model (the Volterra series) is efficient enough to describe the phenomena under study, so those model uncertainties will not have much influence on the final result. The results obtained experimentally support this belief.

On the other hand, as mentioned, data uncertainties cannot be eliminated, but they can be better characterized. A better characterization of this type of uncertainty can give the model the ability to predict the behavior of the dynamic systems under study with a certain level of statistical reliability. When dealing with experiments where data variability can be significant (almost all real-world applications), this capability can improve the performance of the monitoring metrics adopted. The most usual way to consider data uncertainties in estimating a numerical model is to apply a parametric probabilistic approach [Soize, 2017], that considers the model parameters sensitive to the presence of uncertainties as random objects (such as random variables, random vectors or random processes). After characterizing the joint-distribution of these parameters, the underlying uncertainties can be propagated through the model using the well-known Monte Carlo (MC) method, and statistics of the model output can be obtained. Experimental applications will make clearer the advantage of adopting this kind of approach when estimating stochastic models in the context of damage detection.

2.2 The problem of data variation in the context of SHM

As mentioned before, in the context of SHM, the presence of data variation can confuse the deterministic metrics, as shown in Villani et al. [2019b], being the use of probabilistic models necessary

to improve the results. Figure 1 highlights how the data variation can complicate the implementation of classical SHM methodologies. Modal parameters are standard features to consider in the damage detection problem because Damage induces variations in structural properties, which are reflected in the vibration modes of the structures. It is interesting to observe that depending on the level of data variation in the measured signals and extension of the Damage, it is difficult to observe the presence of Damage when the modal parameters are used as the damage-sensitive feature. There is a superposition between the features calculated on reference condition and under damage I condition. Therefore, if we use the modal parameters directly as damage indicators in a deterministic approach or even using powerful machine learning algorithms, we will not be able to detect damage condition I, under these circumstances. The details about the data can be found in Villani et al. [2019c].

When dealing with data analysis directly (data-based monitoring), it is possible to project these features in dimensions where the variability of the measured data is reduced or try to compensate the variability using a reference parameter, the temperature, for example. It is also possible to extract features from the signals that may be less influenced by data variability and implement more powerful classification algorithms. However, when dealing with SHM based on prediction models, as is the case where nonlinear behavior is involved, some statistical certification for the model is always necessary because the presence of uncertainties also influences the parameters of the models. Therefore, these uncertainties must be characterized and incorporated into the model identification step to build stochastic models capable of performing the prediction with statistical reliability.

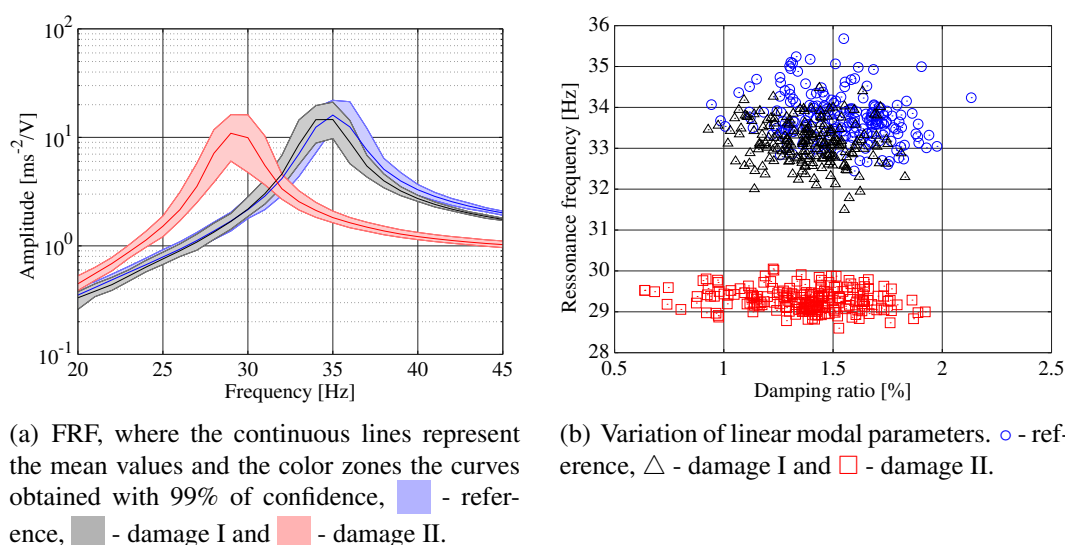


Figure 1: Uncertainties induced by variations in the measured data.

3 An integrated SHM approach using Nonlinearities and UQ

3.1 Nonlinear system identification

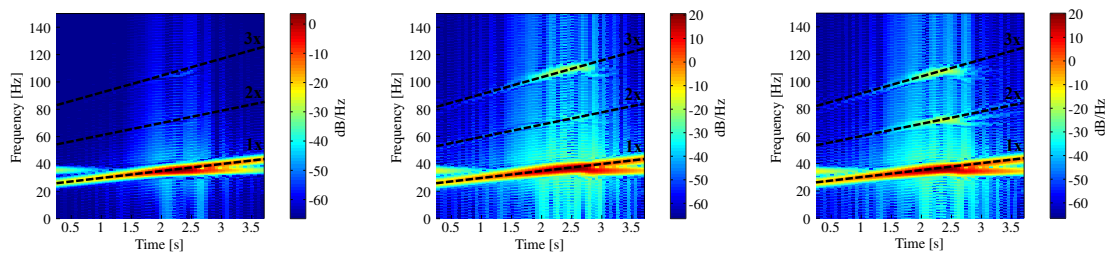
There are two main ways to handle the presence of nonlinear phenomena in SHM problems [Nichols and Todd, 2009]:

- **The healthy structures present linear behavior:** Damage induces the linear structure to present nonlinear phenomena in its behavior. The Damage can be detected by observing

nonlinear behavior in the measured systems' output signals. When we want to achieve higher steps of SHM, usually, a nonlinear model is necessary to predict the structure's output signals in damaged conditions. Many works were developed in this sense, aiming to detect delaminations [Ghrib et al., 2018], breathing cracks [Rébillat et al., 2014], unbalance in rotor systems [Xia et al., 2016], and others;

- **The structures present nonlinear behavior even in healthy conditions:** the structure presents nonlinear phenomena in its output signal in the reference condition, and this behavior can be confused with damages when the approaches described previously are applied [Bornn et al., 2010]. This situation makes necessary the use of more sophisticated methods that can differentiate the effects caused by the damages and the ones caused by the inherent nonlinearities.

The second scenario is more critical. Figure 2 shows an example of the situation described; the data used was measured considering a magneto-elastic nonlinear structure exposed to the presence of a structural change that emulates the behavior of a breathing crack. The presence of multiple harmonics in the responses suggests nonlinear oscillatory behavior. The structure presents linear behavior when a low energy level is applied; however, when high amplitudes are achieved, it manifests nonlinear phenomena. Moreover, the structure is exposed to the presence of a breathing crack that induces nonlinear behavior. When dealing with the second scenario, the construction of a model capable of predicting the intrinsically nonlinear behavior of the system and providing information that allows differentiating the intrinsic phenomena of the system from that caused by Damage is of great relevance. Many nonlinear models could be proposed to solve this issue, and the reader is invited to see Kerschen et al. [2007], Worden et al. [2007] for exploring this vast universe.



(a) Reference - low amplitude. (b) Reference - high amplitude. (c) Damaged - high amplitude.

Figure 2: Time-frequency diagram of the system's output exemplifying the presence of nonlinear phenomena.

The authors decided to use the Volterra series nonlinear data-driven model for dealing with the intrinsically nonlinear behavior of such systems. The choice is based on the capability of the Volterra series to separate linear and nonlinear contributions to the predicted response since this model represents a generalization of the linear convolution concept.

Consider a discrete-time causal nonlinear system with a single output $k \in \mathbb{Z}_+ \mapsto y(k)$ caused by a single input $k \in \mathbb{Z}_+ \mapsto u(k)$, with \mathbb{Z}_+ representing the set of nonnegative integers. Through the discrete-time Volterra series, nonlinear system's output can be written in the form

$$y(k) = \sum_{\eta=1}^{\infty} \sum_{n_1=1}^{N_1} \dots \sum_{n_{\eta}=1}^{N_{\eta}} \mathcal{H}_{\eta}(n_1, \dots, n_{\eta}) \prod_{i=1}^{\eta} u(k - n_i) = \underbrace{y_1(k)}_{linear} + \underbrace{y_2(k) + \dots + y_{\eta}(k)}_{nonlinear}, \quad (1)$$

where $(n_1, \dots, n_\eta) \in \mathbb{Z}_+^\eta \mapsto \mathcal{H}_\eta(n_1, \dots, n_\eta)$ represents the η -order Volterra kernel, N the number of input lags considered in each kernel, and $k \in \mathbb{Z}_+ \mapsto \{y_1(k), y_2(k), \dots, y_\eta(k)\}$ represent, respectively, the first kernel (linear) contribution, the second kernel (quadratic) contribution, and the η -order kernel contribution [Schetzen, 1980].

The main disadvantage of the approach is the challenge in achieving the convergence when a high number of terms is used [Shiki et al., 2017, Villani et al., 2019a]. In order to reduce the number of terms necessary to obtain a good approximation, the Volterra kernels \mathcal{H}_η can be expanded into an orthonormal basis, such as the Kautz functions [Kautz, 1954]. In this way, the series can be rewritten according to the approximation

$$y(k) \approx \sum_{\eta=1}^{\infty} \sum_{i_1=1}^{J_1} \dots \sum_{i_\eta=1}^{J_\eta} \mathcal{B}_\eta(i_1, \dots, i_\eta) \prod_{j=1}^{\eta} l_{\eta, i_j}(k), \quad (2)$$

where J represents the number of Kautz functions used in each orthonormal projections of the Volterra kernels, $(i_1, \dots, i_\eta) \in \mathbb{Z}_+^\eta \mapsto \mathcal{B}_\eta(i_1, \dots, i_\eta)$ represents the η -order Volterra kernel expanded in the orthonormal basis, $k \in \mathbb{Z}_+ \mapsto l_{\eta, i_j}(k)$ is a simple filtering of the input signal $u(k)$ by the Kautz function ψ_{η, i_j} related to each kernel. Information about the Kautz functions can be found in da Silva et al. [2010]. The kernels' coefficients can be estimated using linear regression algorithms, such as the least-squares method [Shiki et al., 2017].

The classical version of the Volterra series expanded using Kautz filters is a useful deterministic nonlinear model. However, as mentioned before, practical applications involving real-world data deal with the presence of uncertainties. If the experimental data present small variability, related to a low level of noise in the measurements, the deterministic version of the model is sufficient to monitor the system, on the other hand, if the data present high variability, related to noise and operational conditions, the model can be rewritten to take into account *data uncertainties*. The *model uncertainties* are not considered here, which means that we assume that the nonlinear model is able to describe the nonlinearities considered in the analysis performed. A parametric probabilistic approach is employed, meaning that the model parameters subjected to uncertainties are described as random variables or random processes [Soize, 2017]. The stochastic version of the Volterra-Kautz model can be described as

$$y(\theta, k) \approx \sum_{\eta=1}^{\infty} \sum_{i_1=1}^{J_1} \dots \sum_{i_\eta=1}^{J_\eta} \mathbb{B}_\eta(\theta, i_1, \dots, i_\eta) \prod_{j=1}^{\eta} \mathbb{l}_{\eta, i_j}(\theta, k), \quad (3)$$

where, in this new version, the random process $(\theta, k) \in \Theta \times \mathbb{Z}_+ \mapsto y(\theta, k)$ represents the stochastic nonlinear system's output, $(\theta, n_1, \dots, n_\eta) \in \Theta \times \mathbb{Z}_+^\eta \mapsto \mathbb{B}_\eta(\theta, i_1, \dots, i_\eta)$ represents the random version of the η -order Volterra kernel expanded in the Kautz basis, and the random process $(\theta, k) \in \Theta \times \mathbb{Z}_+ \mapsto \mathbb{l}_{i_j}(\theta, k)$ is a simple filtering of the deterministic input signal $u(k)$ by the random Kautz function Ψ_{η, i_j} , i.e.,

$$\mathbb{l}_{\eta, i_j}(\theta, k) = \sum_{n_i=0}^{J_\eta} \Psi_{\eta, i_j}(\theta, n_i) u(k - n_i). \quad (4)$$

Estimating the random kernels (the random coefficients) can be done based on the deterministic method described before and Monte Carlo simulations [Kroese et al., 2011, Cunha Jr et al., 2014]. Since we have the random kernels computed when a new input signal is applied, the output signal can be obtained using the uncertainty propagation through the model (again implementing

Monte Carlo simulations). The stochastic model is versatile and can be used to describe a variety of polynomial nonlinearities, considering data variation. Additionally, the component filtering characteristic of the model enables the comparison between the use of linear and nonlinear metrics and the extraction of nonlinear features sensitive to the presence of nonlinear behavior induced by Damage. More details can be obtained in Villani et al. [2019c].

3.2 Extraction of features sensitive to Damage

With the method described in the previous section, it is possible to identify a model representing a structural system's linear and nonlinear dynamics. One can use the discrepancies detected in a measured output y_{exp} concerning the Volterra series considering η terms in the expansion described in Eq. 1. The η -th order prediction error can then be defined by

$$e_\eta = y_{exp} - \sum_{m=1}^{\eta} y_m. \quad (5)$$

It is expected that the prediction error presents an important change in its statistical properties when the system is subjected to damage. To summarize the deviation of a structure in an unknown structural state, a λ_η damage index is applied

$$\lambda_\eta = \frac{\sigma(e_{\eta,unk})}{\sigma(e_{\eta,ref})}, \quad (6)$$

where σ denotes the sample standard deviation, $e_{\eta,unk}$ and $e_{\eta,ref}$ are the prediction errors measured in an unknown structural condition and the baseline state, respectively. This index is similar to the one used in the article of Sohn and Farrar [2001] used to detect nonlinear Damage in an initially linear system. A pure linear dynamics can represent the λ_η index for $\eta = 1$ in the case where the Volterra expansion matches a simple linear convolution [Peng et al., 2021], or it can take into account higher-order contributions of the Volterra series expansion for $\eta > 1$. This property is used in the present chapter to evaluate possible issues that might appear when using linear models to represent and monitor inherently nonlinear structures.

It is worth mentioning that the features described above make use of the deterministic version of the Volterra series model and are very effective for applications where the uncertainties have low influence, as will be demonstrated in the experimental example. When dealing with a higher level of variability, such features can fail in detecting the variations only related to the presence of Damage, and new features should be determined, considering the nonlinear stochastic model.

Let's assume that the first three kernels are enough to describe the nonlinearity of the monitored system. The stochastic model represented by Eq. 3 can be estimated with the system operating under healthy conditions. This model provides a family of models, through MC simulations, that represent the mechanical system in the reference state. This family of models can be used to formulate a new damage index, stochastic in this case, as follows

$$\mathbb{I}_{lin} = [\Lambda_{lin}(\theta, i_1) \quad \mathbb{C}_{lin}(\theta, n_{pca})]_{(N_s \times (J_1 + n_{pca}))}, \quad (7)$$

for the linear components, and

$$\mathbb{I}_{nlin} = [\Lambda_{qua}(\theta, i_1 = i_2) \quad \Lambda_{cub}(\theta, i_1 = i_2 = i_3) \quad \dots \\ \dots \quad \mathbb{C}_{qua}(\theta, n_{pca}) \quad \mathbb{C}_{cub}(\theta, n_{pca})]_{(N_s \times (J_2 + J_3 + 2n_{pca}))}, \quad (8)$$

for the nonlinear ones, where N_s represents the number of MC simulations used. These indices take into account the identified coefficients of each kernel $\{\Lambda_{lin}(\theta, i_1), \Lambda_{qua}(\theta, i_1 = i_2)$ and $\Lambda_{cub}(\theta, i_1 = i_2 = i_3)\}$ and the principal components of the contributions of each kernel to the total output calculated using the stochastic model

$$\begin{aligned} y_{lin}(\theta, k) &\gg \text{PCA} \gg \mathbb{C}_{lin}(\theta, 1), \dots, \mathbb{C}_{lin}(\theta, n_{pca}), \\ y_{qua}(\theta, k) &\gg \text{PCA} \gg \mathbb{C}_{qua}(\theta, 1), \dots, \mathbb{C}_{qua}(\theta, n_{pca}), \\ y_{cub}(\theta, k) &\gg \text{PCA} \gg \mathbb{C}_{cub}(\theta, 1), \dots, \mathbb{C}_{cub}(\theta, n_{pca}), \end{aligned} \quad (9)$$

where n_{pca} represents the number of principal components to be used. This is a brief description and more information about the mathematical formulation can be found in Villani et al. [2019c]. The main advantage here is that the damage index in the reference condition takes into account the natural variability of the experimental data, reducing the probability of false positives in the analysis while separating out the effect of nonlinearities in the system dynamics. When the mechanical system operates in an unknown condition a new index can be calculated and compared with this reference stochastic model using supervised methods, described below.

3.3 Classification of structural states by using supervised methods

For the index calculated using Eq. 6, a simple outlier analysis using thresholds is enough to differentiate the damage states from the reference condition because the index is uni-dimensional. For doing so, simply define a theoretical distribution to represent the index in the reference condition and use the probability measures of false alarms to define the threshold value.

However, for the stochastic indices (Eqs. 7 and 8), this direct procedure is not possible. There are several ways to approach multivariate data classification. One simple and effective way is to use metrics based on distance measures to generate a uni-dimensional index that can also be represented by a theoretical distribution. In this sense, the Mahalanobis distance can be used

$$\mathcal{D}_m^2 = [\mathcal{I}_m - \mu_{\mathbb{I}_m}]^T \Sigma_{\mathbb{I}_m}^{-1} [\mathcal{I}_m - \mu_{\mathbb{I}_m}], \quad (10)$$

where the sub-index m represents both linear and nonlinear features, $\Sigma_{\mathbb{I}_m}$ and $\mu_{\mathbb{I}_m}$ are the covariance matrix and the mean vector of the feature vector completed considering the reference condition (using Eqs. 7 and 8), and \mathcal{I}_m represents the index vector calculated with the mechanical system in an unknown condition. This metric calculates the distance between \mathcal{I}_m and the mean of \mathbb{I}_m taken into account the dispersion and correlation of \mathbb{I}_m . Now, a theoretical probability distribution can be proposed for \mathcal{D}_m^2 and an outlier analysis can be implemented through threshold definition.

It is worth mentioning that different machine learning methods could be used to monitor the structural condition through the proposed damage indices. The choice depends on the ability to separate the data for the features. Since the features calculated here can remarkably separate the different conditions, a simple outlier analysis using one-dimensional metrics is enough to detect the occurrence of Damage. For states considering more incipient Damage, more robust methods may be needed.

3.4 Definition of thresholds: statistical approaches for hypothesis tests

By considering the deterministic approach with the λ_η index (Eq. 6), it is possible to define a statistical methodology assuming that the prediction errors have a Gaussian distribution [Shiki et al., 2017]. With this, the λ_η indicator can be related to a F distribution which describes the

ratio between the variances of two random variables [Bendat and Piersol, 2011]. In this sense, the expected probability density function $P(\lambda)$ is given by

$$P(\lambda) = \frac{2 \left[\frac{v_1}{v_2} \right]^{\frac{v_1}{2}} \lambda^{v_1-1}}{\beta \left(\frac{v_1}{2}, \frac{v_2}{2} \right) \left[1 + \frac{v_1}{v_2} \lambda^2 \right]^{\left(\frac{v_1+v_2}{2} \right)}}, \quad (11)$$

where v_1 and v_2 are the number of degrees of freedom for the prediction errors in the unknown and reference conditions respectively and β is the function:

$$\beta(a_1, a_2) = \frac{(a_1 - 1)! (a_2 - 1)!}{(a_1 + a_2 - 1)!}, \quad (12)$$

where a_1 and a_2 are the input arguments of the β function. In this sense, using the deterministic version of the Volterra series, the damage detection process can be summarized by a statistical hypothesis test to detect a significant increase in the prediction error under a significance level α :

$$\begin{cases} H_0 : \sigma(e_{\eta,unk}) = \sigma(e_{\eta,ref}), \\ H_1 : \sigma(e_{\eta,unk}) > \sigma(e_{\eta,ref}), \end{cases} \quad (13)$$

where H_0 is the null hypothesis which represents the condition where it is likely that the system is still in the reference condition, and H_1 is the alternative hypothesis that represents the case where the standard deviation presented a significant increase due to some modification in the system behavior. This means that in this formulation, it is considered that no other effects are changing the system's response (e.g., environmental variations or uncertainties effects).

When using the Mahalanobis distance calculated (Eq. 10) it is necessary to propose a new hypothesis test. In this regard, we can assume that the Mahalanobis distance can be described by a chi-square distribution when calculated in the reference condition. Within that assumption, we need to ensure the premise of independence and normality in the underlying multivariate features used to calculate the distance measure [Yeager et al., 2019]. Therefore, by integrating the chi-square PDF, we can calculate the probability of a new distance value to belong to the theoretical distribution [Grimmett and Welsh, 2014]

$$p_m = F(\mathcal{D}_m^2 | \nu) = \int_0^{\mathcal{D}_m^2} \frac{t^{(\nu-2)/2} e^{-t/2}}{2^{\nu/2} \Gamma(\nu/2)} dt, \quad (14)$$

where p_m is the probability of the value \mathcal{D}_m^2 belonging to the chi-square distribution, ν is the number of degrees-of-freedom, and $\Gamma(\cdot)$ is the Gamma function.

At last, we can propose a probability threshold value to decide if the sample belongs or not to the theoretical distribution. This sensitivity value depends on the probability of false alarms tolerated in the specific application. At this point, the hypothesis test can be rewritten

$$\begin{cases} H_0 : p_m \geq \epsilon, \\ H_1 : p_m < \epsilon, \end{cases} \quad (15)$$

where ϵ represents the value of sensitivity chosen. The definition of α and ϵ depends on the practical application and for laboratory studies, various values can be examined to test the performance of the methods.

3.5 Probability of false alarm and ROC curve

With the statistical procedure presented in the subsection 3.4, it is possible to accept or reject the null hypothesis, which can give an insight into the structural state of the dynamic system. The threshold for this binary classification is governed by the significance level of the hypothesis test.

The performance of this classification can be carried out by evaluating the false alarm rate, which represents the number of false positives indicated by a damage index, against the true detection rate, which reflects the sensitivity of the proposed methodology. These two metrics are mapped for different threshold values in a receiver-operating characteristic (ROC) curve, which illustrates the false alarm rates against true detection rates [Farrar and Worden, 2012].

4 Experimental applications

This section presents two case studies to highlight various aspects of damage detection involving nonlinear and uncertain systems. The first one in subsection 4.1 illustrates a nonlinear magneto-elastic system subjected to linear structural changes and a lower level of data variation related to uncertainties. In this situation, deterministic metrics can be applied. In subsection 4.2, the second one presents an uncertainty-robust damage detection scheme in a magneto-elastic system with nonlinear damage scenarios and a higher level of data variation imposed during the experiments.

4.1 Linear damage detection in a nonlinear magneto-elastic system

The first case study was performed in the MAGNetO-eLastIc beAm (MAGNOLIA¹) by the SHM Group from UNESP [Shiki et al., 2017]. It was mainly composed of a cantilever aluminum beam with $300 \times 19 \times 3.2$ [mm] and a steel mass attached to its free end. This mass interacted with a permanent neodymium magnet distant 2 [mm] for the steel end in a mono-stable configuration. Similar systems were already investigated in the literature exploring their rich nonlinear behavior in the bi-stable format with multiple magnets [Barton et al., 2010, Erturk and Inman, 2011].

A Modal Shop 2004E electrodynamic shaker was attached 50 [mm] away from the clamped end with a Dytran load cell model 1022V. A Polytec OFV-525/-5000-S laser vibrometer was used to measure the velocity in the beam's free end. A picture of the experimental setup is presented in Figure 3 (a). The nonlinearity in this system comes from the magnetic interactions between the permanent magnet and the steel mass placed in the beam's free end. The magnet is positioned to attract the beam in the axial direction and during the bending movement. Considering the gap of 2 [mm] in this system, a hardening behavior was observed during the experimental tests in this structure. A frequency response curve to a stepped sine input with a shaker-controlled input of 0.01 V and 0.15 V highlights the hardening behavior of the first mode of this system.

Also, a single bolt with 4 nuts (with 2 grams each) was placed in the middle of the beam to simulate mass variations. In this sense, this case study fits the damage detection problems where the structure is nonlinear in the reference state. Still, the Damage was simulated as a simple mass variation that mainly affects linear parameters of this kind of system.

To identify the Volterra model representing the baseline condition of the nonlinear system, a chirp input with two levels: 0.01 and 0.15 V was configured in the shaker to sweep the frequency range from 10 to 50 Hz with 4096 samples acquired with a sampling frequency of 1024 Hz. Figure 4 shows the short Fourier transform of these signals. It is possible to observe the appearance of odd and even harmonics, especially with higher levels of force.

¹The complete dataset can be obtained on the website of the UNESP SHM Research Group: <https://github.com/shm-unesp/UNESP-MAGNOLIA>

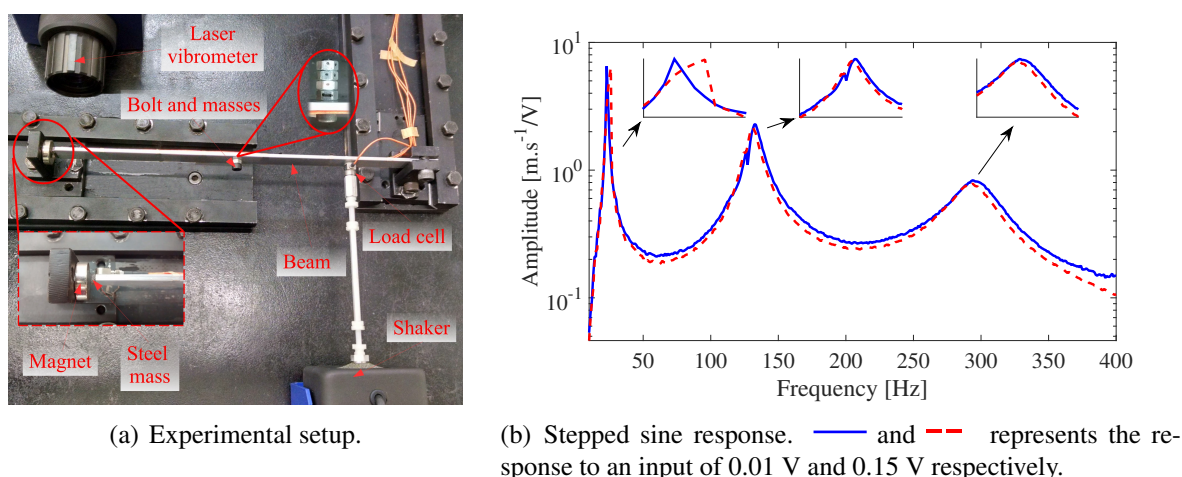


Figure 3: Setup and nonlinear behavior of the MAGNOLIA system.

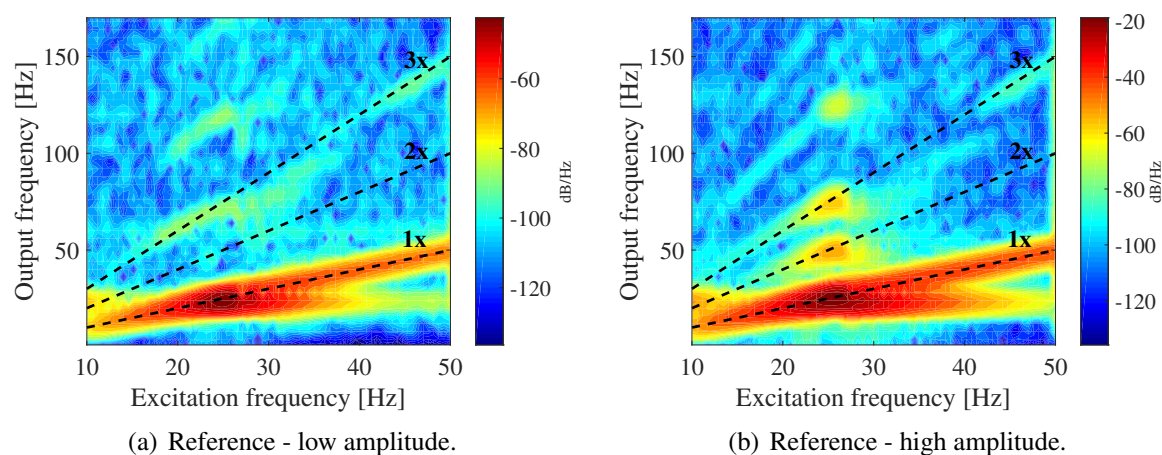


Figure 4: Time-frequency representation of the vibration of the MAGNOLIA system under low and high levels of excitation.

With these signals, we identified the Volterra kernels using the procedure described in section 3.1. Since both second and third-order harmonics were detected, the Volterra model was considered to be presented by the first three kernels. A convergence simulation was performed to set the order of the Kautz filters by calculating the prediction error for different model orders. This resulted in the choice of $J_1 = 2$, $J_2 = 2$ and $J_3 = 8$. Figure 5 illustrates the convergence study and the main diagonals of the first three kernels identified to represent the MAGNOLIA system.

The identified model is a nonparametric representation of the MAGNOLIA system in the baseline condition with 4 masses in the bolt in the middle of the structure. To simulate a linear kind of Damage, these masses were removed one by one and placed back in the system simulating the 8 structural states presented in Table 1. The tests for each structural state were repeated 40 times to check the damage indices' repeatability.

Figures 6 and 7 shows the Volterra-based linear and nonlinear damage indices for all the 8 structural conditions using low and high level chirp input. As shown in Figure 4, higher levels of displacement tend to cause stronger nonlinear effects shown by the presence of harmonics in the response. In this sense, the λ_1 linear index shows to be sensitive to the Damage only in the linear

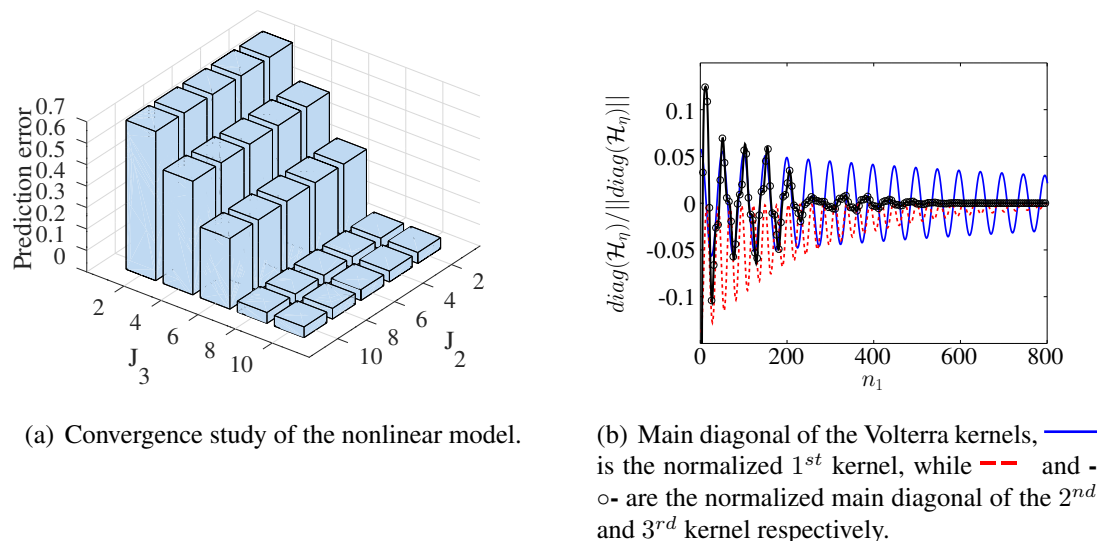


Figure 5: Volterra model representing the MAGNOLIA system.

Table 1: Structural states simulated in the MAGNOLIA system.

State	Condition
1	4 masses (baseline)
2	3 masses (damaged)
3	2 masses (damaged)
4	1 mass (damaged)
5	1 mass (repair)
6	2 masses (repair)
7	3 masses (repair)
8	4 masses (repair)

regime of the MAGNOLIA system. Meanwhile, the nonlinear index λ_3 shows to be sensitive to mass variations in both regimes. This simple example clarifies the possible issues when adopting linear damage indicators for monitoring structures with inherent nonlinear behavior.

Tables 2 and 3 illustrate the results of the Volterra-based damage detection for the low and high-level input amplitude for different values of the significance level of the hypothesis test. By the Table 2 it is visible that both λ_1 and λ_3 indexes were able to accurately detect structural variations without any false alarms. However, Table 3 illustrating the results of the nonlinear regime shows that the linear index fails to detect variations while presenting no false alarms. The nonlinear index presents a true high detection for every case, while false alarms can be minimized by using lower values of significance levels.

The behavior of the hypothesis test for different values of significance levels α is illustrated through the ROC curve for both the linear and nonlinear indices. From this analysis, it is clear that the linear indicator can only increase the detection rate through an increase in the false alarm rate during the nonlinear regime of the system. Meanwhile, the nonlinear index presents high values of detection rate without being affected by the nonlinear structural regime.

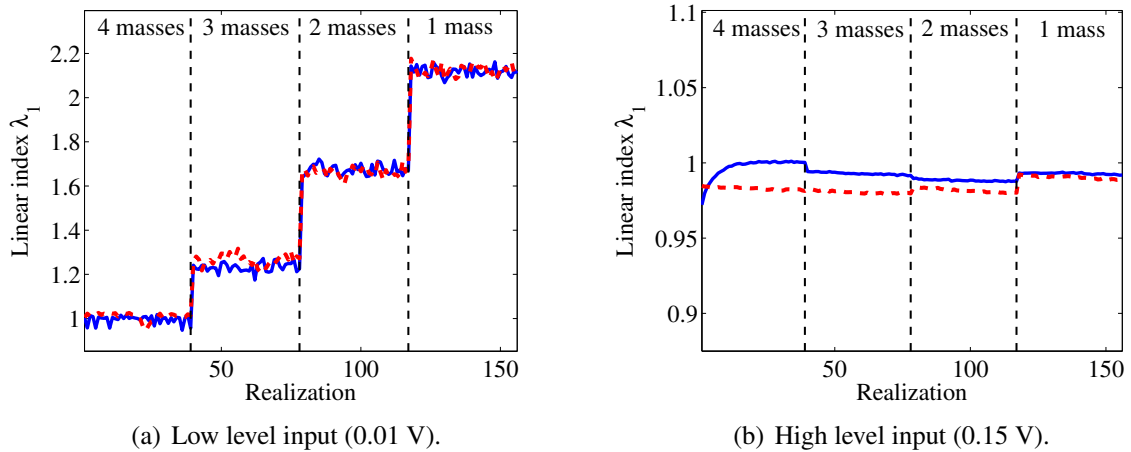


Figure 6: Linear damage index under two different input levels for the MAGNOLIA system. The continuous line — represents the indexes during the damage applications (states 1 to 4) while --- represents the indexes during the repair (states 5 to 8).

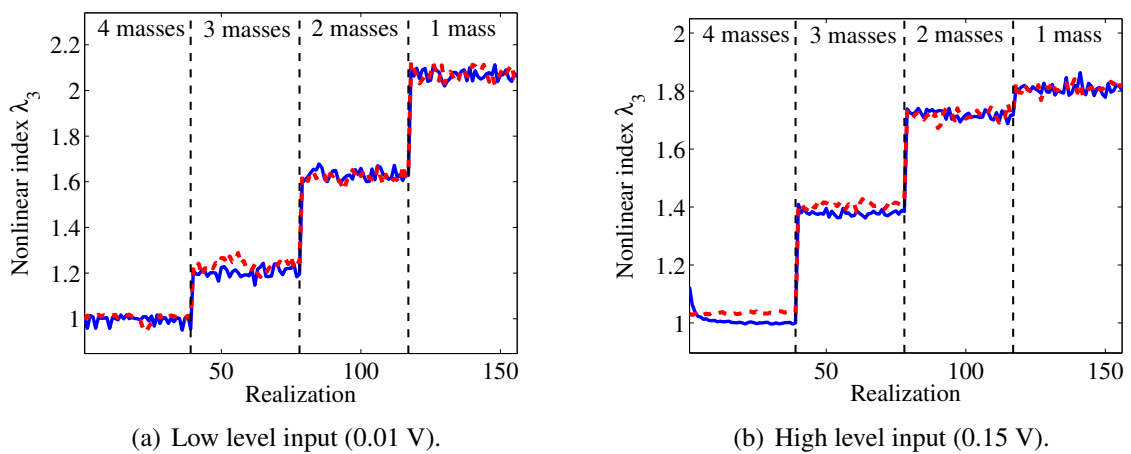


Figure 7: Nonlinear damage index under two different input levels for the MAGNOLIA system. The continuous line — represents the indexes during the damage applications (states 1 to 4) while --- represents the indexes during the repair (states 5 to 8).

Table 2: Results of the hypothesis tests for different significance levels under low level input (0.01 V) on the MAGNOLIA system.

α [%]	Linear index (λ_1)		Nonlinear index (λ_3)	
	False alarm [%]	True detection [%]	False alarm [%]	True detection [%]
2	0	100	0	100
1	0	100	0	100
0.5	0	100	0	100

Table 3: Results of the hypothesis tests for different significance levels under high level input (0.15 V) on the MAGNOLIA system.

α [%]	Linear index (λ_1)		Nonlinear index (λ_3)	
	False alarm [%]	True detection [%]	False alarm [%]	True detection [%]
2	0	0	34.6	100
1	0	0	17.9	100
0.5	0	0	3.85	100

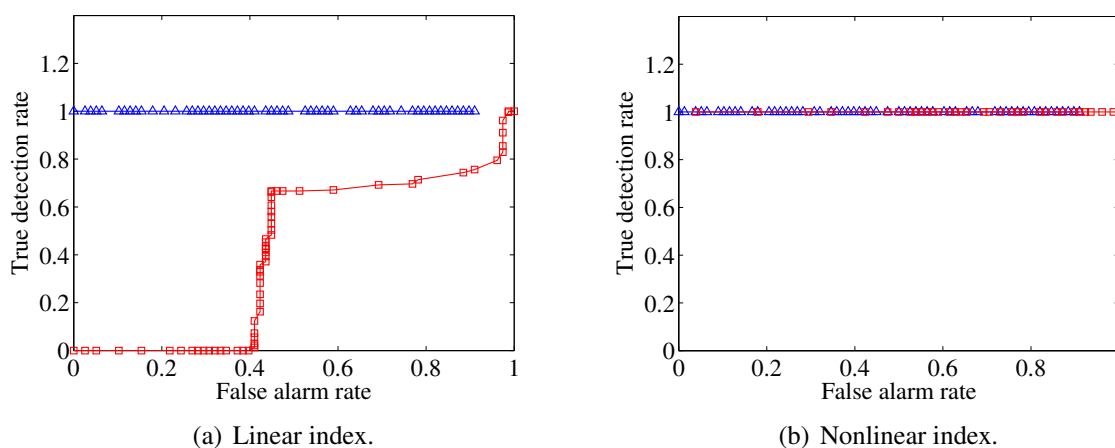


Figure 8: ROC curves of the damage indexes for the MAGNOLIA system. - \triangle - is the curve for the low level input (0.01 V) and - \square - is the curve for the high level input (0.15 V).

4.2 Nonlinear damage in a magneto-elastic beam

The second case study was performed in the nonlinear Damage in a magnEto-eLastic bEam (ADELE²) by the SHM Group from UNESP [Villani et al., 2019c]. This is a more challenging benchmark since data variation was induced throughout the experimental tests and the emulated Damage having nonlinear behavior characteristics. The stochastic formulation becomes necessary in this situation to deal with the uncertainties.

The main characteristics of the experimental setup used is presented in Fig. 9, with the structure in the reference condition (without Damage). Figure 9 (a) shows the structure composed of a cantilever beam, constructed by gluing four beams of Lexan together, $2.4 \times 24 \times 240$ [mm³] each one. At the free boundary, two steel masses are affixed and interact with a magnet, generating a nonlinear behavior in the system response that third-order harmonics can verify in the Time-frequency diagram presented in Fig. 9 (b). The experimental setup uses an Electrodynamic Transducer Labworks Inc. (ET-132) to apply forces to the structure (close to the clamp), an Accelerometer PCB PIEZOTRONICS (352C22) to measure the structure output (close to the free end), and a National Instruments acquisition system to convert and save data. Once again, two levels of input were considered during the experimental tests. A low input amplitude level (1 V RMS) ensures the system oscillates in a linear vibration regime, and a high input amplitude level (6 V RMS) ensures the excitation of the system's nonlinearities.

²The complete dataset can be obtained on the website of the UNESP SHM Research Group: <https://github.com/shm-unesp/ADELE>

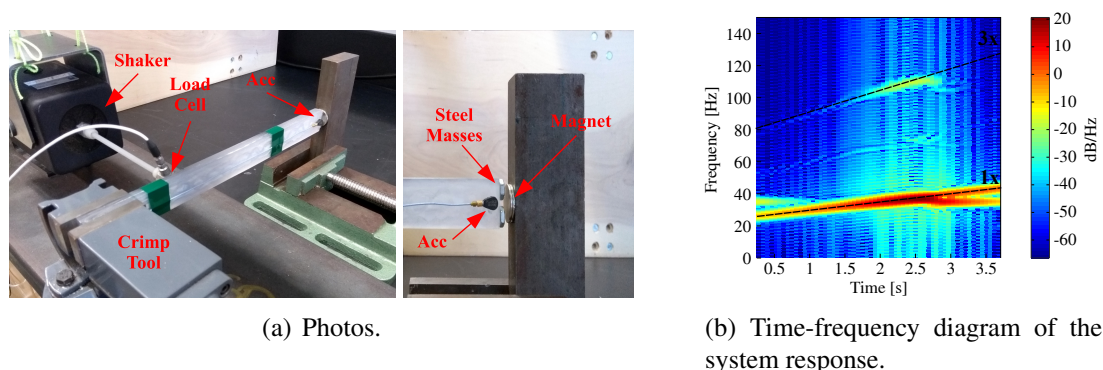


Figure 9: Setup and nonlinear behavior of the ADELE system.

The Damage introduced to the mechanical system aims to simulate a breathing crack behavior. Therefore, four different configurations were constructed:

1. Training beam: 4 intact Lexan beams glued, used to train the model (see Fig. 10 (a));
2. Test beam: 4 intact Lexan beams glued, used to test the model (see Fig. 10 (a));
3. Damage I: 3 intact and 1 cut beams glued (see Fig. 10 (a) and (b));
4. Damage II: 2 intact and 2 cut beams glued (see Fig. 10 (a) and (b));

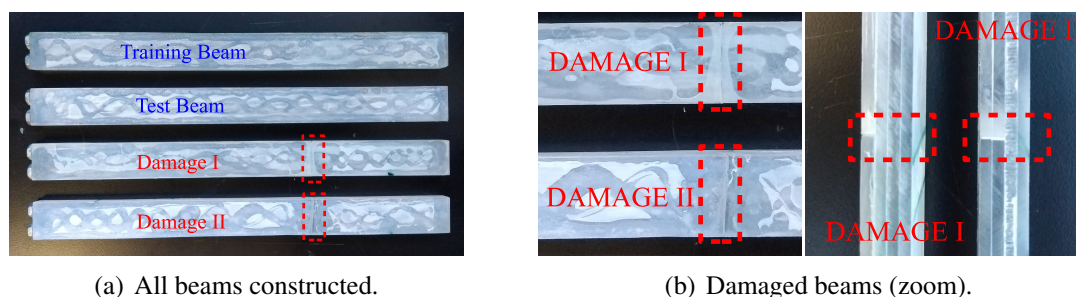


Figure 10: Structural conditions examined.

This kind of structural change can emulate a breathing crack behavior, inducing a nonlinear quadratic behavior in the system that can be verified by observing the results of Fig. 11. The results show that a quadratic harmonic was introduced for the Damage I condition compared with Fig. 9 (b). In contrast, the first and third harmonics are not visually affected. For Damage condition II, all harmonics are altered because of the large extent of Damage.

Furthermore, the data variation was induced by varying the distance between the magnet and the structure. This variation generates great variability in the modal parameters of the structure and of the nonlinear components too. The effect of such variation can be seen in Fig. 1. In this situation, we need to estimate a stochastic model that can represent such variability under reference conditions and separate the nonlinear contributions generated by the presence of the magnet and those generated by the presence of Damage.

For doing so, we can identify a stochastic model using the formulation of Eq. 3, and use this family of models to estimate the indices of Eqs. 7 and 8 that represents the structure in the healthy

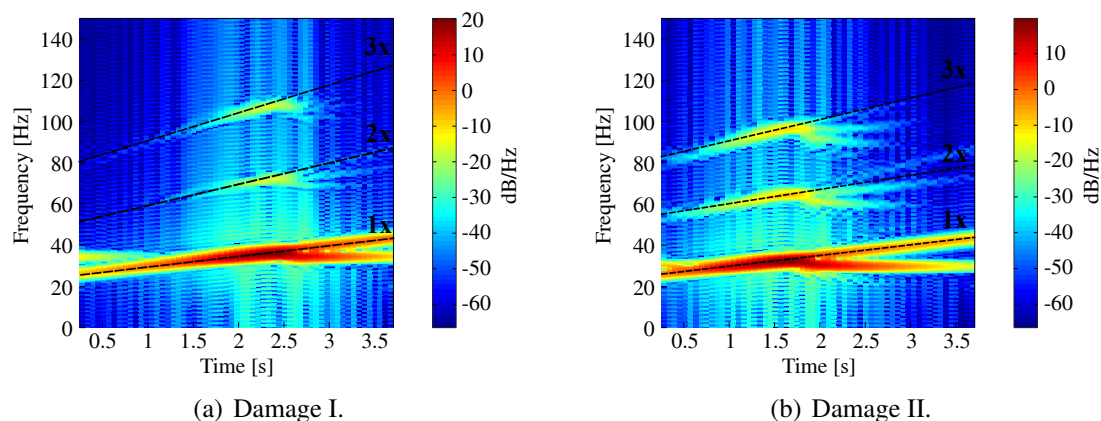


Figure 11: Time-frequency diagram obtained with the structure damaged.

condition. In this step we considered: a chirp input signal varying the excitation frequency from 25 to 40 Hz (first mode shape region) with two levels of amplitude, as mentioned before, $J_1 = 2$, $J_2 = 4$, $J_3 = 6$, and 200 experimental realizations for each structural condition. More information about the step-by-step model estimation procedure can be found in Villani et al. [2019c].

Once the stochastic model is estimated, it can be used to monitor the system. \mathbb{I}_m can be estimated, and in an unknown condition, a new model can be identified to calculate \mathcal{I}_m . Finally, the Mahalanobis distance (Eq. 10) is calculated and plotted in Fig. 12 for different structural conditions. It can be noted that the linear index can not detect the nonlinear behavior induced by Damage at the beginning of the propagation (Damage condition I), as expected. When the Damage also changes the linear components (Damage condition II), the linear index values grow. Since the characteristic of the Damage is to induce nonlinear behavior in the structure, the nonlinear index shows an interesting ability to separate the nonlinearities generated by the presence of the magnet from those induced by the presence of the Damage, even at the beginning of the propagation (Damage condition I). Therefore, it is expected that when applying the proposed hypothesis test (Eq. 15) the nonlinear index will show a more robust behavior, even considering the variability of the data.

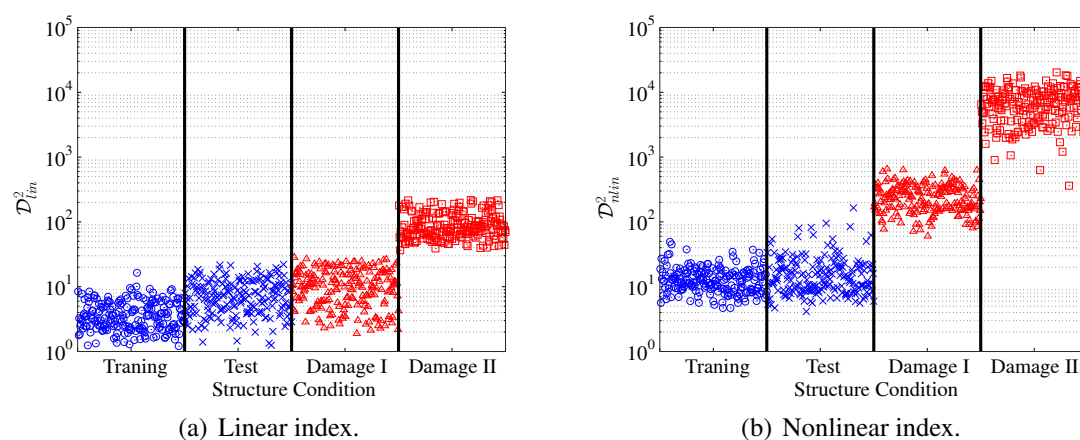


Figure 12: Evolution of the Mahalanobis distance.

Table 4 shows the results obtained for both linear and nonlinear indices when applying the

Table 4: Hypothesis test results for ADELE system.

	ϵ	False detection [%]		True detection [%]	
		Training Beam	Test Beam	Damage I	Damage II
Linear Index	10^{-2}	0.5	15.5	34.5	100
	10^{-4}	0	0	5	100
	10^{-6}	0	0	0	100
	10^{-12}	0	0	0	75
Nonlinear Index	10^{-2}	4.5	12.5	100	100
	10^{-4}	1	7	100	100
	10^{-6}	0	4.5	100	100
	10^{-12}	0	1.0	96.5	100

hypothesis test. The test is applied using various values for the sensitivity (ϵ). Despite presenting a higher level of false positives because of its high sensitivity, the nonlinear index has a considerably better performance in detecting this type of Damage, even considering the great variability of the data. It is worth remembering that a false positive error can generate less catastrophic consequences than a false negative error, as is seen in the case of the linear index. This ratio between false positives and false negatives could be controlled by choosing the value of ϵ optimally, depending on the level of reliability and cost that the error may cause in each practical application.

Finally, we can study the degree of separability of the data that these indices generate by varying the sensitivity of the hypothesis test and constructing the ROC curve. Figure 13 shows the results obtained. The results revealed that monitoring the nonlinear components of the structure output signal, depicted by the second and third-order kernel coefficients and contributions used to fill the feature vector, is an interesting methodology to be deployed in damage detection problems, mainly when the Damage exhibits nonlinear characteristics. Additionally, when data variation represents an issue, the use of the stochastic reference model can improve the capability of the method to distinguish the changes related to the presence of Damage from the ones related to the presence of uncertainties. Knowing that the closer to the point (0,1), the better the index's performance, we can conclude that the nonlinear index performed better.

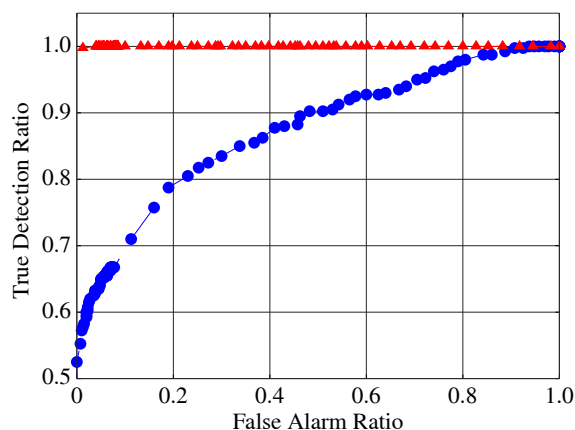


Figure 13: ROC curves of the damage indexes for ADELE system. \circ represents the linear index and \triangle represents the nonlinear index.

5 Final Remarks

This chapter has demonstrated by some simple illustrative examples how nonlinear vibration and operation regimes with parameter uncertainty can compromise the correct diagnosis of the structural state of a system. Some methods of dealing with these difficulties were presented based on systems identification techniques that seek to filter linear and nonlinear components in a stochastic way and group the effects according to their causes and effects. The results confirmed that the procedure's robustness is increased when both effects are mitigated.

It is evident that when not assumed, the possibility of propagating uncertainties and identifying and separating the nonlinear behavior from damage and operation confounded the detection. On the other hand, these techniques demand more knowledge from the SHM engineer to identify these sources and the requirement of implementing some strategy to allow the correct classification.

Acknowledgements

The authors gratefully acknowledge, for the financial support given to this research, the following Brazilian agencies: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) - Finance Code 001; São Paulo Research Foundation (FAPESP), grant numbers 12/09135-3, 13/25148-0, 15/25676-2, 17/24977-4 and 19/19684-3; Brazilian National Council for Scientific and Technological Development (CNPq) grant number 306526/2019-0 ; and Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ) grants 210.167/2019, 211.037/2019 and 201.294/2021.

References

- D. Barton, S. Burrow, and L. Clare. Energy harvesting from vibrations with a nonlinear oscillator. *ASME Journal of Vibrations and Acoustics*, 132(2):021009, 2010. doi: 10.1115/1.4000809.
- J. Bendat and A. Piersol. *Random Data: Analysis and Measurement Procedures*. Wiley Series in Probability and Statistics. Wiley, 2011. ISBN 9781118210826. URL <https://books.google.com.br/books?id=qYSViFRNmlwC>.
- L. Bornn, C. R. Farrar, and G. Park. Damage detection in initially nonlinear systems. *International Journal of Engineering Science*, 48(10):909–920, 2010. ISSN 0020-7225. doi: 10.1016/j.ijengsci.2010.05.011. URL <http://www.sciencedirect.com/science/article/pii/S0020722510001059>. Structural Health Monitoring in the Light of Inverse Problems of Mechanics.
- A. Cunha Jr, R. Nasser, R. Sampaio, H. Lopes, and K. Breitman. Uncertainty quantification through Monte Carlo method in a cloud computing setting. *Computer Physics Communications*, 185:1355–1363, 2014. doi: <https://doi.org/10.1016/j.cpc.2014.01.006>.
- S. da Silva, S. Cogan, and E. Foltête. Nonlinear identification in structural dynamics based on Wiener series and Kautz filters. *Mechanical Systems and Signal Processing*, 24(1):52–58, Jan. 2010. ISSN 0888-3270. URL <http://www.sciencedirect.com/science/article/pii/S0888327009001897>.
- A. Erturk and D. Inman. Broadband piezoelectric power generation on high-energy orbits of the bistable duffing oscillator with electromechanical coupling. *Journal of Sound and Vibration*, 330(10):2339 – 2353, 2011. ISSN 0022-460X. doi: <http://dx.doi.org/10.1016/j.jsv>.

- 2010.11.018. URL <http://www.sciencedirect.com/science/article/pii/S0022460X10007807>. Dynamics of Vibro-Impact Systems.
- C. R. Farrar and K. Worden. *Machine Learning and Statistical Pattern Recognition*, pages 295–320. John Wiley & Sons, Ltd, 2012. ISBN 9781118443118. doi: 10.1002/9781118443118.ch9. URL <http://dx.doi.org/10.1002/9781118443118.ch9>.
- M. Ghrib, M. Rébillat, G. V. des Roches, and N. Mechbal. Automatic damage type classification and severity quantification using signal based and nonlinear model based damage sensitive features. *Journal of Process Control*, 2018. ISSN 0959-1524. doi: <https://doi.org/10.1016/j.jprocont.2018.08.002>. URL <http://www.sciencedirect.com/science/article/pii/S0959152418301975>.
- G. Grimmett and D. Welsh. *Probability: an introduction*. Oxford University Press, 2014.
- W. H. Kautz. Transient synthesis in the time domain. *Circuit Theory, Transactions of the IRE Professional Group on*, CT-1(3):29–39, September 1954. ISSN 0197-6389. doi: 10.1109/TCT.1954.1083588.
- G. Kerschen, K. Worden, A. F. Vakakis, and J.-C. Golinval. Nonlinear system identification in structural dynamics: current status and future directions. In *25th International Modal Analysis Conference, Orlando, 2007*, 2007.
- D. P. Kroese, T. Taimre, and Z. I. Botev. *Handbook of Monte Carlo Methods*. Wiley, 2011.
- J. M. Nichols and M. D. Todd. *Nonlinear Features for SHM Applications*. American Cancer Society, 2009. ISBN 9780470061626. doi: 10.1002/9780470061626.shm049. URL <http://onlinelibrary.wiley.com/doi/abs/10.1002/9780470061626.shm049>.
- Z. Peng, J. Li, H. Hao, and C. Li. Nonlinear structural damage detection using output-only volterra series model. *Structural Control and Health Monitoring*, 28(9):e2802, 2021. doi: <https://doi.org/10.1002/stc.2802>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/stc.2802>.
- M. Rébillat, R. Hajrya, and N. Mechbal. Nonlinear structural damage detection based on cascade of Hammerstein models. *Mechanical Systems and Signal Processing*, 48(1):247–259, 2014. ISSN 0888-3270. doi: <http://dx.doi.org/10.1016/j.ymsp.2014.03.009>. URL <http://www.sciencedirect.com/science/article/pii/S0888327014000843>.
- M. Schetzen. *The Volterra and Wiener Theories of Nonlinear Systems*. Wiley, 1980.
- S. B. Shiki, S. da Silva, and M. D. Todd. On the application of discrete-time Volterra series for the damage detection problem in initially nonlinear systems. *Structural Health Monitoring*, 16(1): 62–78, 2017. doi: 10.1177/1475921716662142. URL <http://dx.doi.org/10.1177/1475921716662142>.
- H. Sohn and C. R. Farrar. Damage diagnosis using time series analysis of vibration signals. *Smart Materials and Structures*, 10(3):446–451, jun 2001. doi: 10.1088/0964-1726/10/3/304. URL <https://doi.org/10.1088/0964-1726/10/3/304>.
- C. Soize. *Stochastic models of uncertainties in computational mechanics*. Amer Society of Civil Engineers, 2012.

- C. Soize. *Uncertainty Quantification: An Accelerated Course with Advanced Applications in Computational Engineering*. Springer, 2017.
- L. G. G. Villani, S. da Silva, and A. Cunha Jr. Damage detection in uncertain nonlinear systems based on stochastic Volterra series. *Mechanical Systems and Signal Processing*, 125:288–310, 2019a. ISSN 0888-3270. doi: <https://doi.org/10.1016/j.ymssp.2018.07.028>. URL <http://www.sciencedirect.com/science/article/pii/S0888327018304291>. Exploring nonlinear benefits in engineering.
- L. G. G. Villani, S. da Silva, A. Cunha Jr, and M. D. Todd. Damage detection in an uncertain nonlinear beam based on stochastic Volterra series: An experimental application. *Mechanical Systems and Signal Processing*, 128:463–478, 2019b. ISSN 0888-3270. doi: <https://doi.org/10.1016/j.ymssp.2019.03.045>. URL <http://www.sciencedirect.com/science/article/pii/S0888327019302250>.
- L. G. G. Villani, S. da Silva, A. Cunha Jr, and M. D. Todd. On the detection of a nonlinear damage in an uncertain nonlinear beam using stochastic Volterra series. *Structural Health Monitoring*, 2019c. doi: 10.1177/1475921719876086. URL <http://doi.org/10.1177/1475921719876086>.
- K. Worden, G. Kerschen, A. F. Vakakis, and J.-C. Golinval. Nonlinear System Identification in Structural Dynamics: A Short (and Biased) History. In *25th International Modal Analysis Conference, Orlando, 2007*, pages 1996–2017, 2007.
- X. Xia, J. Zhou, J. Xiao, and H. Xiao. A novel identification method of Volterra series in rotor-bearing system for fault diagnosis. *Mechanical Systems and Signal Processing*, 66-67:557–567, 2016. ISSN 0888-3270. doi: <https://doi.org/10.1016/j.ymssp.2015.05.006>. URL <http://www.sciencedirect.com/science/article/pii/S0888327015002277>.
- M. Yeager, B. Gregory, C. Key, and M. Todd. On using robust Mahalanobis distance estimations for feature discrimination in a damage detection scenario. *Structural Health Monitoring*, 18(1): 245–253, 2019. doi: 10.1177/1475921717748878. URL <https://doi.org/10.1177/1475921717748878>.

Chapter 10

Applications of Numerical and Statistical Methods for Monitoring and Control Engineering Variables in a Concrete Dam

Chapter details

Chapter DOI:

<https://doi.org/10.4322/978-65-86503-88-3.c10>

Chapter suggested citation / reference style:

Lazzarotto, Emerson, et al. (2022). “Applications of Numerical and Statistical Methods for Monitoring and Control Engineering Variables in a Concrete Dam”. In Jorge, Ariosto B., et al. (Eds.) *Uncertainty Modeling: Fundamental Concepts and Models*, Vol. III, UnB, Brasilia, DF, Brazil, pp. 279–336. Book series in Discrete Models, Inverse Methods, & Uncertainty Modeling in Structural Integrity.

P.S.: DOI may be included at the end of citation, for completeness.

Book details

Book: Uncertainty Modeling: Fundamental Concepts and Models

Edited by: Jorge, Ariosto B., Anflor, Carla T. M., Gomes, Guilherme F., & Carneiro, Sergio H. S.

Volume III of Book Series in:

Discrete Models, Inverse Methods, & Uncertainty Modeling in Structural Integrity

Published by: UnB City: Brasilia, DF, Brazil Year: 2022

DOI: <https://doi.org/10.4322/978-65-86503-88-3>

Applications of Numerical and Statistical Methods for Monitoring and Control Engineering Variables in a Concrete Dam

Emerson Lazzarotto^{1*}, Sheila R. Oro², Francisco A. F. Reinaldo³,
Tásia Hickmann⁴, Samuel B. Rodrigues⁴, Jairo M. Corrêa⁴
and Lucas S. Ribeiro⁴

¹Post-Graduate Program – Management, Technology and Sustainability, West Parana State University, Brazil. E-mail: emerson.lazzarotto@unioeste.br

²Department of Physics, Statistics and Mathematics, Federal University of Technology – Francisco Beltrão – PR, Brazil. E-mail: sheilaro@utfpr.edu.br

³Department of Informatics, Federal University of Technology – Francisco Beltrão – PR, Brazil. E-mail: reinaldo@utfpr.edu.br

⁴Department of Mathematics and Statistics, Federal University of Technology – Medianeira – PR, Brazil. E-mail: hickmann@utfpr.edu.br; samuelb@utfpr.edu.br; jairocorrea@utfpr.edu.br; lribeiro@utfpr.edu.br

*Corresponding author

Abstract

This chapter presents a compilation of techniques and research results published by the authors in journal articles concerning the use of numerical methods and univariate and multivariate statistical methods applied to evaluate, predict and control the behavior of important engineering variables that are monitored by instruments in the dam of the Itaipu hydroelectric plant. The introductory section contextualizes Itaipu and the importance of the methods used. Next, we present a brief theoretical description of univariate forecasting methods, multivariate statistical techniques, and the finite element method. After describing the set of techniques, the reader can always find specific applications in Itaipu of one or more of the preceding techniques. The results, in general, show that the combination of statistical and numerical methods helps to improve the performance of forecasts and, in general, can collaborate in the decision-making of Itaipu engineers and are an additional risk analysis tool that can be applied to any dam.

Keywords: Dam safety and monitoring; Forecasting methods; Multivariate analysis; Finite element method; reduction methods

1 Introduction

In this chapter, we present some numerical and statistical methods applied in the context of dam safety, resulting from research developed in 2012 by the authors during and after their doctorate in the Postgraduate Program of Numerical Methods in Engineering at the Federal University of Paraná (PPGMNE – UFPR), carried out in partnership with the Center for Advanced Studies in Dam Safety of the Itaipu Technological Park (CEASB – PTI), with the support of Itaipu Binacional.

Dams are barriers or structures that cross streams, rivers, or canals to confine and control water flow and are generally used for specific purposes such as water supply, hydroelectric power generation, and irrigation. Among other administrative requirements, dams must have adequate instrumentation and performance monitoring. The maintenance of safe dams is a guarantee of sustainable development and management of water resources ICOLD-CIGB [2008].

Controlling the structural behavior of significant engineering work, such as a hydroelectric power plant dam, is usually done through visual inspections and instrumentation. Visual inspections are most often used to detect accidents and incidents. However, this does not diminish the importance of instrumentation, which is either insufficient or not adequately exploited. In particular, the best instrumentation can minimize accidents and incidents, establish operating rules, propose conservation measures and allow for the deepening of existing knowledge, improving the quality and lifespan of the work Silveira [2003].

The literature from international conferences on large dams emphasizes the need to maintain a safety culture and effective safety management systems. The Itaipu Binacional Hydroelectric Plant is located on the Paraná River, on the stretch of border between Brazil and Paraguay, 14 km north of Friendship Bridge, in the Foz do Iguaçu, in Brazil, and Hernandarias, in Paraguay Itaipu [2022a]. The purpose of the plant is to take advantage of the hydraulic potential of the Paraná River and began to generate energy in 1984 Itaipu [2022a]. The Itaipu dam is made of concrete, rockfill (rocks), and earth and has a length of 7,919 meters and a maximum height of 196 meters, the equivalent of a 65-story building.

Itaipu's installed capacity (power) is 14 gigawatts (GW), through 20 generating units of 700 megawatts (MW) each. Itaipu supplies around 10.8% of the energy consumed in Brazil and 88.5% of Paraguay's consumption. In 2020, the Itaipu plant produced 76,382 GWh. Itaipu Binacional is the world leader in delivering clean and renewable energy, having made more than 2.7 billion MWh since the beginning of its operation Itaipu [2022b].

Figure 1 presents an illustrative image of the Itaipu Dam with its main parts and dimensions, with emphasis on the main dam of relieved gravity (U) where the generating units are located and the spillway (A), which is a concrete structure intended to drain excess water from the reservoir when it is at its maximum level.

The dam's superlative dimensions and strategic importance to Brazil and Paraguay have made dam safety one of Itaipu's priorities, from the design phase to the completion and maintenance of the mega-works. Monitoring the dam's safety began even before the plant's first generating unit was installed and remains highly relevant – not only to maintain and extend the project's useful life but above all to protect life and property in the vicinity of the hydroelectric plant Itaipu [2022c].

To maintain these essential economic numbers from the point of view of environmental issues and the impact on the entire community that is close not only to the dam but also to the whole area of the reservoir, it is necessary to ensure that the behavior of the structures is under control.

Considering that each dam work is unique, Itaipu and CEASB understood that, in addition to

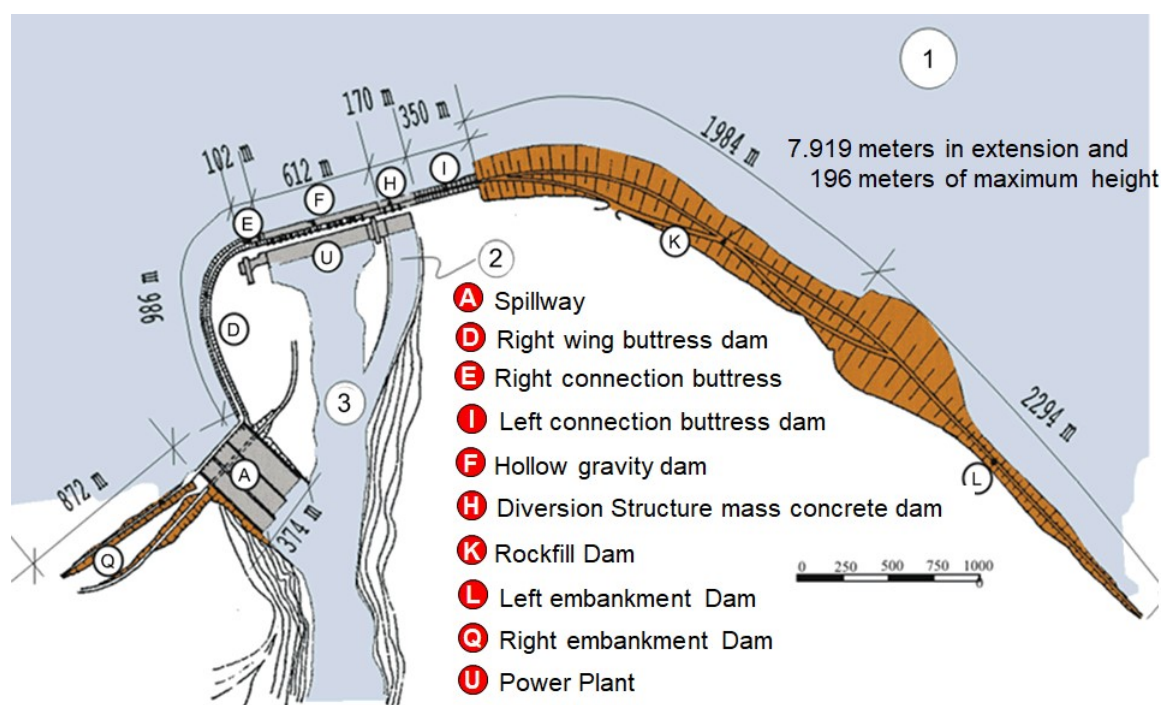


Figure 1: General arrangement of the Itaipu Hydroelectric power plant dam.

Source: ENCC [2019]

visual inspections and the control already performed by the dam's safety engineering team about monitoring instruments, additional investigation through numerical and probabilistic models is essential. The proposal is that such models, in some way, could let the data speak for itself.

The probabilistic models can also be used and combined with other techniques to produce the most accurate predictions possible. They serve as a reliable alert for predicting eventual abnormalities in the dam's behavior, alerting, if necessary, to the early realization of intervention works and actions.

The reader interested only in the applications can find them in the subsections of sections 2 and 3 and subsection 4.3 of section 4.

2 Forecasting Methods

One question to be answered when dealing with dam monitoring data is how can the historical behavior of data readings help predict the future? The prior indication of a possible anomaly is crucial to mitigate financial, environmental, and social effects.

The availability of monitoring data has spread the application of statistical models to predict the dam's behavior. These have complemented visual inspections and numerical models to support decision-making (Salazar et al. [2017]).

This section briefly describes several statistical or artificial intelligence techniques used by the authors and some applications in time series of instrument monitoring data from the Itaipu dam.

2.1 The ARIMAX-GARCH model

Let y_t ($t = 1, \dots, T$) be a stationary time series (or a non-stationary time series that can be transformed into a stationary one) that exhibits linear auto-dependence. Also, consider a list

of $r + 1$ vectors of realizations from $r + 1$ stationary exogenous variables of y_t ($t = 1, \dots, T$), denoted by $\left((x_{1,t})_{t=1}^T, \dots, (x_{(r+1),t})_{t=1}^T \right)$. Based on Box and Tiao [1975] and Pankratz [2012], each realization y_t can be represented by an ARIMAX (p, d, q) model, with the mathematical formulation:

$$\nabla^d y_t = \theta_0 + \sum_{j=1}^p \phi_j y_{t-j} + \sum_{i=0}^q \theta_i e_{t-i} + \sum_{i=1}^{r+1} \sum_{l_i=0}^{L_i} \alpha_{il_i} x_{i,t-l_i} + e_t \quad (1)$$

where B is the backward operator defined by $B^k y_t := y_{t-k}$, with k belonging to \mathbb{Z} ; $\nabla^d := (1 - B)^d$ is the difference operator, with d representing its order; $(\phi_i)_{i=1}^p$ and $(\theta_j)_{j=0}^q$ are the ordered lists of model complex parameters, with $\phi_p \neq 0$ and $\theta_p \neq 0$, and $(\alpha_{il_i})_{l_i=0}^{L_i}$ is the i th ordered list of complex model parameters associated with the exogenous component (these three lists of parameters need to satisfy both the invertibility and the stationarity conditions – see e.g. Hamilton [1994] and Lütkepohl [2005]); e_t is an innovation consisting of a state of the random variable e_t from an uncorrelated stochastic process with zero mean; p and q are, respectively, the orders of the *Auto-Regressive* (AR) part $\sum_{j=1}^p \phi_j y_{t-j}$ (AR(p)) and of the *Moving Average* (MA) part $\sum_{i=0}^q \theta_i e_{t-i}$ (MA(q)); and, L_i is the maximum lag order in the sum $\sum_{l_i=0}^{L_i} \alpha_{il_i} x_{i,t-l_i}$ of the exogenous variable $x_{i,t}$, with $i = 1, \dots, (r + 1)$. Particularly, if $\alpha_{il_i} = 0$, for all $i = 0, \dots, (r + 1)$ and all $l_i = 0, \dots, L_i$, then the model in (1) above becomes a conventional ARIMA (p, d, q). Note that, a SARIMAX (p, d, q) x (P, D, Q)_S model – also known as a multiplicative ARIMAX (p, d, q) model – can be used for modeling a seasonal time series (as in Hamilton [1994]). It generically consists of an ARIMAX model with seasonal components.

In equation (1) above, it is assumed that the innovation term e_t is a realization of an uncorrelated random variable e_t , with zero mean and constant conditional variance, that is, $s_t^2 = s^2$, for all t . However, many time series do not satisfy this stationarity condition (called homoscedasticity). In some cases, the changing conditional variance (volatility) may depend on past squared innovations of the time series or past values of the variance itself such that an ARMA structure, as well as an extension of it, can be adopted for temporally projecting the volatility. This way, unconditionally, the variance is constant but conditional on past values; it is allowed to change in time. According to Bollerslev [1986], the general GARCH (R, S) model for the conditional variance s_t^2 of the innovation e_t is given by

$$\sigma_t^2 = \gamma_0 + \sum_{j=1}^R \gamma_j \sigma_{t-j}^2 + \sum_{i=1}^V \delta_i e_t^2 \quad (2)$$

where the following constraints must hold: $\sum_{j=1}^R \gamma_j + \sum_{i=1}^V \delta_i < 1$; $\gamma_0 > 0$; $\gamma_j = 0$ ($j = 1, \dots, R$) and $\delta_i \geq 0$ ($i = 1, \dots, V$). Equation (2) is used by the WARIMAX-GARCH method to generate in-sample and out-of-sample forecasts of the conditional variance of the time series y_t ($t = 1, \dots, T$). In addition, s_t^2 multiplied by a parameter f can be used in (1) to account for a non-linear effect to construct the forecasting of its level (conditional mean). Notice that the generating mechanism for a GARCH innovation process e_t ($t = 1, \dots, T$) is defined by $e_t := \sigma_t z_t$, where: z_t is a realization of a standardized, independent and identically distributed random variable Z_t ; and σ_t is the conditional standard deviation.

A model compounded by (1) and (2), which accounts for the above mentioned constraints, is called an ARIMAX-GARCH model. In order to obtain the best possible ARIMAX-GARCH

model, three basic steps should be carried out: (i) test the plausible values for the parameters p , d , q and L_i , in (1), as well as the parameters R and V in (2) (which can be obtained through the profile analysis of the plots of simple and partial auto-correlation functions of the ordinary standard innovations z_t ($t = 1, \dots, T$) and the squared standard innovations z_t^2 ($t = 1, \dots, T$), as described by Bollerslev [1986] and Engle [1982]; (ii) define the method to be used to estimate the ARIMAX-GARCH parameters – the most common is the *Maximum Likelihood Estimation* (MLE) method (as in Hamilton [1994]); and, (iii) make a diagnostic check to choose the most parsimonious and adequate model to be used for generating both the in-sample and the out-of-sample forecasts of y_t ($t = 1, \dots, T$) and their volatilities.

It is worth pointing out that similar to the way that the h -steps ahead out-of-sample forecasts of conditional means are produced (for instance, through an estimated ARIMAX model), the h -steps ahead forecasts of conditional variances (volatilities) are generated using an estimated GARCH model (or an extension of it). For more details, please see [18], wherein a *GARCH-in-mean model* produces out-of-sample forecasts of volatilities and employs them in constructing the estimates for the corresponding conditional means. This approach aims to aggregate non-linear information (coming from the squared auto-dependence exhibited by a given time series) from the forecasts of the dependent means to increase its predictive accuracy. Note that the WARIMAX-GARCH method proposed here can adopt any GARCH approach to forecast volatilities.

For a non-stationary time series, stationarity can be achieved by differentiating order d using the difference operator (∇). Also, suppose the time series presents seasonality. In that case, the model can be said to be a multiplicative SARIMA(p,d,q)x(P,D,Q) $_s$, where p is the auto-regressive parameter of the non-seasonal part (using earlier p values to set the forecast); q is the parameter of the moving averages of the non-seasonal position (averaging the errors of p auto-regressive values); d is the differentiation parameter of the series until it becomes stationary; P is the auto-regressive order of the seasonal part; D is the seasonal difference order if the series does not present stationarity; Q is the moving averages parameter of the seasonal period, and s is the seasonal period ($s=52$ in this paper, indicating 52 weeks/year). The SARIMA model in the simplified form of the time series Z_t is defined by the equation:

$$\nabla_s^D \nabla^d Z_t = \phi_0 + \sum_{j=1}^p \phi_j z_{t-j} + \sum_{j=i}^q \theta_j e_{t-i} + e_t \quad (3)$$

where $\nabla_s^D := (1 - B^s)^D$ is the difference operator of the seasonal part; $\nabla^d := (1 - B)^d$ is the difference operator, with d representing its order; B is the backward operator defined by $B^k z_t := z_{t-k}$, with k belonging to Z ; $(\phi_i)_{i=0}^p$ and $(\theta_j)_{j=1}^q$ are the ordered lists of complex model parameters, with $\phi_p \neq 0$ and $\theta_q \neq 0$; e_t is a residue; p and q are, respectively, the orders of the AR(p) part ($\sum_{j=1}^p \phi_j z_{t-j}$) and of the MA(q) part ($\sum_{j=i}^q \theta_j e_{t-i}$).

The univariate Box & Jenkins methodology also allows the use of exogenous variables (or explanatory variables) in the composition of the adjusted model, allowing the study of the impact of these variables in the prediction model. More information about a certain physical phenomenon is added to the SARIMA model. The inclusion of an exogenous variable X in the model described by the equation (3) generates the model SARIMAX(p,d,q)x(P,D,Q) $_s$ defined by:

$$\nabla_s^D \nabla^d Z_t = \phi_0 + \sum_{j=1}^p \phi_j z_{t-j} + \sum_{j=i}^q \theta_j e_{t-i} + \alpha X + e_t \quad (4)$$

Where α is the coefficient of the exogenous variable X .

The pre-modeling phase of the SARIMA and the SARIMAX methods consists of identifying the lagged correlation between the time series of long-term concrete deformations (dependent variable series) and ambient temperature (exogenous variable series).

2.2 Artificial Neural Network

Artificial Neural Networks (ANNs) are well known to be flexible computing frameworks for modeling and forecasting a broad range of stochastic time series exhibiting either linear or non-linear auto-dependence structures. Contrary to many linear statistical forecasting models, stationarity is not required by ANN methods (see, e.g., Teixeira Júnior et al. [2015]). Another critical aspect of ANNs is that they are universal approximates of compact (i.e., closed and bounded) support functions, as shown in Cybenko [1989]. In effect, since observations from a time series y_t ($t= 1, \dots, T$) that exhibit dependency on past values may be seen as points of the domain of an unknown compact support function, it follows that the ANNs are capable of approximating them (for modeling or forecasting) with a high degree of accuracy. According to Zhang [2003], the predictive power of ANNs comes from the parallel processing of the information exhibited by the data. In addition, ANN models are primarily determined by the stochastic characteristics inherent in the time series.

In this context, the feed-forward multi-layer perceptron ANNs (see e.g. Haykin [2001]) are the most widely used neural prediction models for time series forecasting. Particularly, a *single hidden layer ANN* (hereafter, for simplicity, referred to as ANN) is characterized by an artificial network composed of three layers (namely, input, hidden, and output layers) of simple processing units numerically connected by acyclic links. The relationship between the output y_t and the L -lagged inputs, y_{t-k} ($k = 1, \dots, L$), has the following mathematical representation

$$y_t = \alpha_0 + \sum_{j=1}^q \alpha_j g \left(\beta_{0j} + \sum_{i=1}^L \beta_{ij} y_{t-i} \right) + \varepsilon_t, \quad (5)$$

where α_j ($j = 0, 1, \dots, q$) and β_{ij} ($i = 0, 1, \dots, L; j = 0, 1, \dots, q$) are the ANN parameters, called connection weights; L is the number of input nodes; q is the number of hidden nodes; ε_t is the approximation error at time t ; and $g(\cdot)$ is the transfer function, here, a logistic function – although it is possible to adopt other functions (see e.g Haykin [2001]). The logistic function is widely employed as the hidden layer transfer function in neural network forecasting and is mathematically defined by

$$g(x_t) = \frac{1}{1 + \exp(-x_t)}, \quad (6)$$

where $x_t := \beta_{0j} + \sum_{i=1}^L \beta_{ij} y_{t-i}$ and $\exp(\cdot)$ is the exponential function with Euler's basis (as in Zhang [2003]). Due to $g(\cdot)$ being a non-linear transfer function, the ANN model in (5), in fact performs a non-linear mapping of the past observations y_{t-k} ($k = 1, \dots, L$) to produce a forecast for y_t . In general, the model in (5) can be rewritten, as

$$y_t = f(y_{t-1}, y_{t-2}, \dots, y_{t-L}, w) + \varepsilon_t, \quad (7)$$

where w denotes a vector of all ANN parameters and $f(y_{t-1}, y_{t-2}, \dots, y_{t-L}, w)$ is the model determined by the network structure and the connection weights in (5). Note that the neural network as defined above is equivalent to a non-linear auto-regressive model.

In practice, w is unknown and needs to be numerically determined. So, to find its estimated value, \hat{w} , that accounts for some criteria (an objective function), an optimization algorithm is applied

to training data. Although several methodologies are available in the literature, we adopt, in line with Haykin [2001], Levenberg-Marquardt's algorithm. The minimization of the in-sample sum of squared errors (i.e., $\min_w \sum_{t=1}^T \varepsilon_t^2$) is the numerical criteria of this ANN. We have that once \hat{w} is determined.

$$y_t = f(y_{t-1}, y_{t-2}, \dots, y_{t-L}, \hat{w}) + \hat{\varepsilon}_t \quad (8)$$

with $f(y_{t-1}, y_{t-2}, \dots, y_{t-L}, \hat{w}) = \hat{y}_t$ consisting of the optimized ANN outcome at time t , which is the forecast of the state y_t , and $\hat{\varepsilon}_t$ being the forecasting error of \hat{y}_t .

The Recurrent Neural Network (RNN) Nerrand et al. [1993], which uses the architecture Long Short-Term Memory (LSTM) in the hidden layer for a series of sequential data in the training process, allows for better precision in the adjustment in the training data as compared to MLP and RBF networks, due to its property of storing short and long term memory in the training process that produces better accuracy in the forecast Bengio et al. [2012].

2.3 Support Vector Regression

The Support Vector Machines (SVMs) form a system based on the theory of statistical learning or VC theory (Vapnik-Chervonenkis), which has been developed since the 1960s as reports Vapnik [1995]. It was presented for the first time in the Annual Workshop on Computation Learning Theory, COLT 1992 by Boser et al. [1992]. The SVM algorithm was extended for the case of regression by Vapnik [1999] making thus possible the estimation of functions of absolute values. The so-called Supor Vector Regression Machine (SVRM) or Support Vector Regression (SVR) emerged.

The following descriptions of SVR are based on tutorials Smola and Schölkopf [2004]. The basic idea of SVR consists of imagining a margin around the outline of the transformation function and finding a function having at most a previously fixed error in all examples, thus trying to obtain a margin as narrow as possible.

Let it be a sample of training data $(y(t), y(t+1))_{t=1}^{T-1} \subset X \times R$ where X indicates the space of entry patterns (for instance $X = R^{T-1}$), with $t = 1, 2, \dots, T-1$. In Vapnik's ε -SV regression, the objective is to find a function $f(x(t))$, (where $x(t) = (y(t))_{t=1}^{T-1}$), which presents at most a deviation from the targets $d(t)$, with $d(t) = (y(t+1))_{t=1}^{T-1}$, obtained for all the targets in the training. In other words, the method is not concerned with errors as long as they are smaller than, but does not accept any deviation higher than that. Therefore the linear approximation function is described as follows: $f(x) = \langle w, x \rangle + b$ com $w \in X$, $b \in R$, where $\langle \cdot, \cdot \rangle$ denotes the internal product at X . This notwithstanding, it is not always possible to guarantee the viability of the problem, since there are points that violate the restrictions. On account of that, the formulation proposed Vapnik [1995, 1999] and presented in Problem (9) utilizes clearance variables $\xi(t)$ and $\xi^*(t)$ for each t , where $t = 1, 2, \dots, T-1$, in order to deal with the unfeasible restrictions of the problem of optimization, that is, penalizing data that lie outside the margin.

$$\begin{aligned} \text{Minimize} \quad & \frac{1}{2} \|w\|^2 + C \left(\sum_{t=1}^{T-1} \xi(t) + \xi^*(t) \right) \\ \text{subject to :} \quad & \begin{cases} d(t) - \langle w, x(t) \rangle - b \leq \varepsilon + \xi(t) \\ \langle w, x(t) \rangle + b - d(t) \leq \varepsilon + \xi^*(t) \\ \xi(t), \xi^*(t) \geq 0, \quad \forall t = 1, \dots, T-1 \end{cases} \end{aligned} \quad (9)$$

Where $C > 0$ is the regularization constant, that is, of penalization of errors, since it weighs on the terms of the minimization function.

Solving the Problem (9) is not always easy due to the inequality restrictions; the model now occupies the dual space, which provokes the softening of the algorithm.

For the dual formulation, non-negative Lagrange multipliers are introduced as reported by Vapnik [1995, 1999] and Smola and Schölkopf [2004]. Therefore, Problem (10), now by maximization, in its dual form, is given by

$$\begin{aligned} & \text{Maximize } \left(-\frac{1}{2}\right) \sum_{t,j=1}^{T-1} (\alpha(t) - \alpha^*(t)) (\alpha(j) - \alpha^*(j)) \langle x(t), x(j) \rangle + \\ & \quad -\varepsilon \sum_{i=1}^l (\alpha(t) + \alpha^*(t)) + d(t) \sum_{t=1}^{T-1} (\alpha(t) - \alpha^*(t)) \\ & \text{subject to } \sum_{t=1}^{T-1} (\alpha(t) - \alpha^*(t)) = 0 \\ & \quad \alpha^*(t) \in [0, C], \forall t = 1, \dots, T-1 \end{aligned} \quad (10)$$

In this way, the variable w of the primal form comes to be defined as $w = \sum_{t=1}^{T-1} (\alpha(t) - \alpha^*(t)) x(t)$ and the linear approximation function is now represented by equation (11):

$$f(x) = \sum_{t=1}^{T-1} (\alpha(t) - \alpha^*(t)) \langle x(t), x \rangle + b \quad (11)$$

The dual formulation of the SVR problem provides the option of working in a higher-dimensional space. Thus, one may carry out a non-linear mapping of the entry data toward an area of a large dimension, where the linear regression becomes possible. For this purpose the approach based on Kernel functions Vapnik [1995, 1999], $K(x, x') := \langle \phi(x), \phi'(x') \rangle$ is used, that is, there is a function K that receives two points (x, x') and calculates its scale product in a characteristic space, whose introduction in the optimization problem changes its description as Problem (12):

$$\begin{aligned} & \text{Maximize } -\frac{1}{2} \sum_{t,j=1}^{T-1} (\alpha(t) - \alpha^*(t)) (\alpha(j) - \alpha^*(j)) K(x(t), x(j)) + \\ & \quad -\varepsilon \sum_{t=1}^{T-1} (\alpha(t) + \alpha^*(t)) + d(t) \sum_{t=1}^{T-1} (\alpha(t) - \alpha^*(t)) \\ & \text{subject to } \sum_{t=1}^{T-1} (\alpha(t) - \alpha^*(t)) = 0 \\ & \quad \alpha(t), \alpha^*(t) \in [0, C], \quad \forall t = 1, \dots, T-1 \end{aligned} \quad (12)$$

And the approximation function of non-linear SVR comes to be given by Equation (13):

$$f(x) = \sum_{t=1}^{T-1} (\alpha(t) - \alpha^*(t)) K(x(t), x) + b \quad (13)$$

According to Yu et al. [2006], the most utilized Kernel functions are linear Kernel: $K(x_i, x_j) := \langle x_i, x_j \rangle$; sigmoidal Kernel: $K(x_i, x_j) := \tanh[\gamma \langle x_i, x_j \rangle + c]$; polynomial kernel: $K(x_i, x_j) := [\gamma \langle x_i, x_j \rangle + c]^d$ and gaussian kernel: $K(x_i, x_j) := \exp(-\gamma \|x_i - x_j\|^2)$. With the inclusion of

Kernel, the SVMs acquire the characteristic described by Cortes and Vapnik [1995] which must respect the Mercer conditions Vapnik [1995, 1999] and Smola and Schölkopf [2004] to ensure that the Kernel functions realize the dot product in the space of characteristics.

As Beltrami et al. [2011], the results obtained by SVR depend significantly on the values of the regularization constant C , margin ε , the type of Kernel function, and their respective parameters.

2.4 Wavelet decomposition of level r

Let l^2 be the collection of all scalar-valued complex infinite sequences $y_t (t \in \mathbb{Z})$ in $\mathbb{C}^{\mathbb{Z}}$ such that $\sum_{t \in \mathbb{Z}} |y_t|^2 < \infty$, where \mathbb{C} and \mathbb{Z} denote, respectively, the sets of all complex and integer numbers; and assume that the function $\langle ; \rangle$ of l^2 into \mathbb{C} is the usual inner product (as in Kubrusly [2011]). An element $\omega(\cdot) \in l^2$ – with $\langle ; \rangle$ – is a l^2 – wavelet function, if the sequence $\omega_{m,n}(\cdot) := 2^{\frac{m}{2}} \omega(2^m(\cdot) - n)$, where $(n, m) \in \mathbb{Z} \times \mathbb{Z}$, forms an orthonormal basis for l^2 ; and a member $\phi(\cdot) \in l^2$ – with $\langle ; \rangle$ – consist of a l^2 -scaling function, if the sequence $\phi_{m,n}(\cdot) := 2^{\frac{m}{2}} \phi(2^m(\cdot) - n)$, where $(m, n) \in \mathbb{Z} \times \mathbb{Z}$, holds: $\langle \phi_{m',n'}(\cdot); \phi_{j,k}(\cdot) \rangle \neq 0$, whenever $m' = j$ and $n' \neq k$; and, if otherwise, $\langle \phi_{m',n'}(\cdot); \phi_{j,k}(\cdot) \rangle = 0$, where $m', n', j, k \in \mathbb{Z}$ (see e.g. Levan and Kubrusly [2003] and Mallat [2008]). Based on Kubrusly and Levan [2006], it can be seen that the subset $\left\{ \omega_{m,n}(\cdot) \right\}_{\substack{cm,n \in \mathbb{Z} \\ m \geq m_0}} \cup \left\{ \phi_{m_0,n}(\cdot) \right\}_{n \in \mathbb{Z}}$ of l^2 -wavelet and l^2 -scaling functions, where m_0 is a fixed integer value, is, in fact, an orthonormal basis for l^2 . Accordingly, any sequence $y_t (t \in \mathbb{Z})$ in l^2 can be orthogonally decomposed, in terms of an orthonormal basis $\left\{ \omega_{m,n}(\cdot) \right\}_{\substack{cm,n \in \mathbb{Z} \\ m \geq m_0}} \cup \left\{ \phi_{m_0,n}(\cdot) \right\}_{n \in \mathbb{Z}}$, as represented by

$$y_t = y_{A_{m_0},t} + \sum_{m=m_0}^{+\infty} y_{D_m,t}, \quad (14)$$

where $y_{A_{m_0},t} := \sum_{n \in \mathbb{Z}} a_{m_0,n} \phi_{m_0,n}(\cdot)$ is the WC of approximation at level m_0 of the state y_t , with $a_{m_0,n} := \langle y_t; \phi_{m_0,n}(\cdot) \rangle$; and $y_{D_m,t} := \sum_{n \in \mathbb{Z}} d_{m,n} \omega_{m,n}(\cdot)$ is the WC of detail at level m of y_t , with $d_{m,n} := \langle y_t; \omega_{m,n}(\cdot) \rangle$. The orthogonal decomposition in (14) is usually called a *wavelet decomposition*.

Tautologically, any finite (scalar-valued complex) time series $y_t (t = 1, \dots, T)$ can be interpreted as an infinite sequence $y_t (t \in \mathbb{Z})$ in l^2 , defined as $y_t = y_t$, if $t \in \{1, 2, \dots, T\}$; and $y_t = 0$, if $t \in \mathbb{Z} - \{1, 2, \dots, T\}$. Therefore, any finite time series $y_t (t = 1, \dots, T)$ can be orthogonally decomposed by the wavelet decomposition in (14).

In practical terms, once it is impossible to model separately all WCs generated by the expansion (14), an adaptation is required to obtain a finite number of components. Thus, according to [16], a good alternative may occur thoroughly by means of the *wavelet decomposition of level r* of $y_t (t = 1, \dots, T)$, where $r > 0$ and $r \in \mathbb{Z}$, is given by

$$y_t = \tilde{y}_{A_{m_0},t} + \sum_{m=m_0}^{m_0+(r-1)} \tilde{y}_{D_m,t} + \varepsilon_t, \quad (15)$$

where m_0 is the level parameter (which is often assumed to be equal to r); ε_t is the approximation error term, that is, the difference between the state y_t and its (wavelet) approximation $\tilde{y}_{A_{m_0},t} + \sum_{m=m_0}^{m_0+(r-1)} \tilde{y}_{D_m,t}$ (nevertheless, in practice, it is usually assumed that ε_t is equals zero); $\tilde{y}_{A_{m_0},t} := \sum_{n=0}^{2M-m_0-1} a_{m_0,n} \phi_{m_0,n}(t)$ and $\tilde{y}_{D_m,t} := \sum_{n=0}^{2^{(M-m_0)}-1} d_{m,n} \omega_{m,n}(t)$, which are, respectively, WCs of

approximation at level m_0 and of detail at level m and consist, respectively, of the approximations to $y_{A_{m_0,t}}$ and $y_{D_{m,t}}$, in (14); and M is a parameter that takes an integer value such that $T = 2^M$. If T is not an integer power of 2, the sequence y_t ($t = 1, \dots, T$) is usually completed with zeros such that its length T is increased up to the next integer power of 2. This procedure may be carried out because the zeros added up do not affect the calculation of the WCs $\tilde{y}_{A_{m_0,t}}$ and $\tilde{y}_{D_{m,t}}$ generate in (15) (see e.g. Haven et al. [2012]), preserving the auto-correlation y_t and its components, in (15), for all t , where $t = 1, \dots, T$.

2.5 Bootstrap

Introduced by Efron [1992] the Bootstrap method is a non-parametric statistical procedure whose main idea is the re-sampling with repositioning of the original data to create new sets of data that allow the estimation of a measure of interest. The Bootstrap method is proper when no probabilistic model can be applied to the collection of data under analysis and when the amount of data is insufficient for using the Central Limit Theorem Davison and Hinkley [1997]. Tibshirani [1996] suggests, in the context of regressions, two kinds of Bootstrap sampling that permit the determination of the trust interval: Bootstrap pairs and Bootstrap residual. Bootstrap pairs consider the set of original data sampled and replaced in the original set, each with the same probability of being selected. Bootstrap residual, in its turn, requires a regression model adjusted over the original collection of data and the calculation of the residues. This way, the residues are re-sampled, and new series are obtained. This paper considers the Bootstrap residual technique to achieve the trust interval. Following is a summary of the Bootstrap residual algorithm Tibshirani [1996]:

In the time series $(y(t))_{t=1}^T$, with $t = 1, 2, \dots, T', T' + 1, \dots, T$ values of the original sample, the algorithm is given by the following steps:

- i To obtain the adjusted model $(\hat{y}(t))_{t=T'}^T$ regarding the original data $(y(t))_{t=1}^T$ where T' indicates the non-considered degrees of liberty;
- ii To determine the residues of the adjusted model, $r = (y(t))_{t=T'}^T - (\hat{y}(t))_{t=T'}^T$;
- iii To generate Bootstrap samples, each of size $T - T'$ to be obtained through the re-sampling of the residues, obtaining $(r^b(t))_{t=T'}^T$, which will be the B Bootstrap samples; To generate new synthetic series $(y^b(t))_{t=T'}^T = (\hat{y}(t) + r^b(t))_{t=T'}^T$;
- iv To create new synthetic series $(y^b(t))_{t=T'}^T = (\hat{y}(t) + r^b(t))_{t=T'}^T$;
- v For each synthetic series, to obtain the adjusted model $(\hat{y}^b(t))_{t=T'}^{T+h}$, $h \in N$ indicate the values to be forecast outside the training sample.

- vi To estimate the standard error of the t^{th} forecasted value using $\left[\frac{1}{B-1} \sum_{b=1}^B (\hat{y}^b(t) - \bar{\hat{y}}^b(t))^2 \right]^{1/2}$
 where $\bar{\hat{y}}^b(t) = \frac{1}{B} \sum_{b=1}^B \hat{y}^b(t)$.

For further details about Bootstrap, the work of Efron and Tibshirani [1994] and Davison and Hinkley [1997] may be consulted.

2.6 Trust intervals

The application of regression models involves the assessment of uncertainties that surround those who estimate parameters. The trust interval is a valuable statistic and performs this vital function. They are used to estimate the parameters of the regression model Pan and Politis [2016].

Consider patterns $(y(t), \mathbf{x}(t))_{t=1}^T$ and the proposed regression model as $y(\mathbf{x}) = f(\mathbf{x}, \mathbf{w}) + e(\mathbf{x})$ where $e(\mathbf{x})$ are the targets, \mathbf{w} are the true parameters of the unknown model $f(\mathbf{x}, \mathbf{w})$ and $e(\mathbf{x})$ the noise (aleatory variable) with zero average. The estimators of \mathbf{w} are $\hat{\mathbf{w}}$ and $g(\mathbf{x}, \hat{\mathbf{w}})$ is the approximation of $f(\mathbf{x}, \mathbf{w})$, where $f(\mathbf{x}, \mathbf{w})$ can be interpreted as the average of the y targets given the \mathbf{x} entry. According to Heskes [1997], for regression problems the accuracy of the true regression estimate leads to the construction of the trust interval. And according to Khosravi et al. [2014] the uncertainty of the adjusted regression model is given by the variance $\sigma_g^2(x)$ associated to the approximation $g(\mathbf{x}, \hat{\mathbf{w}})$ of $f(\mathbf{x}, \mathbf{w})$. Khosravi et al. [2014] worked with Bootstrap and neural networks and suggested a trust interval expressed in Equation (16):

$$m(\mathbf{x}) - t_{1-\frac{\alpha}{2}, gl} \sigma_g(\mathbf{x}) \leq f(\mathbf{x}, \mathbf{w}) \leq m(\mathbf{x}) + t_{1-\frac{\alpha}{2}, gl} \sigma_g(\mathbf{x}) \quad (16)$$

where $t_{1-\frac{\alpha}{2}, gl}$ comes from the distribution t of Student with a α level of significance, gl is the number of degrees of liberty, defined as the difference between the number of training samples and the number of parameters of the model and $m(\mathbf{x}_i) = \frac{1}{B} \sum_{b=1}^B g(\mathbf{x}_i, \hat{\mathbf{w}})$.

According to Haykin [2009], Support Vector Machines is a category of neural networks with alterations in the re-alimentation of the layers. Thus, this paper used the trust interval as a describer in Equation (16).

2.7 Forecast combination

The first papers that dealt with forecast combinations were developed by the researchers Reid [1968], and Bates and Granger [1969], as Clemen [1989] reports. According to Clemen [1989] combination is an attractive method to obtain forecasts, since instead of choosing the best technique, the problem becomes deciding which techniques may be of greater help for improved accuracy, since the performance of combination is, in general, assessed through accuracy. In Wallis [2011] and Mancuso and Werner [2013], one can find historical reviews in which many of the main articles on the combination of individual forecast methods are mentioned. These reports bring up more than forty-two years of research in this area, proving utilization and efficiency, besides showing a variety of applications and approaches.

Faria and Mubwandarinkwa [2008] report the existence of two possibilities to determine a forecast method that can generate predictions, which are:

- To choose a forecast method in $(\mu_m)_{m=1}^M$ based on some form of selection;
- To choose K forecast methods in $(\mu_m)_{m=1}^M$ and then combine them (either in linear or in non-linear form)

where $(\mu_m)_{m=1}^M$ denotes a set with M predictive methods.

In general, the combination idea may be summarized in the following two stages:

- To consider Δ^K as an ensemble of all forecasts (within and without the sample) originating from K individual predictive methods chosen by the modeler, where $K \leq M$;

- To obtain the map

$$\hat{y}_{CL} : \begin{cases} \nabla^K \rightarrow R \\ (\hat{y}_k(t))_{k=1}^K \in \nabla^K \rightarrow \hat{y}_{CL}(t) \in R \end{cases}$$

of forecasts combination (generic) which transforms a $(\hat{y}_k(t))_{k=1}^K \in \nabla^K$ forecast vector into a combined forecast $\hat{y}_{CL}(t) = \sum_{k=1}^K \rho_k \times \hat{y}_k(t) + \alpha$ for every $t = 1, \dots, T + h$.

where $h \in N$ is the forecast time frame, $\hat{y}_{CL}(t)$ is the linearly combined forecast; $\hat{y}_k(t)$ is the forecast of the k^{th} base method; ρ_k the adaptive weight linearly associated to the $\hat{y}_k(t)$ forecast, and α the added adaptive constant.

Júnior [2013] used mathematic programming in order to calculate the coefficients of linear combination, optimizing a problem of non-linear programming where the objective function was made up by the sum of forecast errors squared. According to Faria and Mubwandarikwa [2008], it is common to impose restrictions $\sum_{k=1}^K \rho_k = 1$ and $\rho_k \geq 0$ on the values that may be assumed by adaptive weights $(\rho_k)_{k=1}^K$ in order to preserve the possibility of its probabilistic interpretation. According to Júnior [2013] the simple, weighted or harmonic averages may also be used in the forecast combination.

To evaluate combined or individual the methods MAE, MAPE, MSE and RMSE errors Hamilton [1994] represented in Equation (17 to 20) were used.

a. Mean Absolute Error (MAE): The mean of the absolute linear distances or absolute linear deviation, calculated point-to-point between the forecast values \hat{y}_t and the original values y_t . It is a measure of the central tendency that centralizes the errores. It is defined by:

$$MAE = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t| \quad (17)$$

b. Mean Absolute Percentage Error (MAPE): The mean of the absolute proportional linear distances, i.e., percentual mean calculated point-to-point between the forecast values \hat{y}_t and the original values y_t . It is defined by:

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_t}{y_t} \right| \quad (18)$$

c. The mean of the quadratic distances, calculated point-to-point between the forecast values \hat{y}_t and the original values y_t .

$$MSE = \frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2 \quad (19)$$

d. Root Mean Square Error (RMSE): The square root of the mean of the quadratic distances, calculated point-to-point between the forecast values \hat{y}_t and the original values y_t . I.e., the proportional distribution of errors where the lesser point differences have less mean error weight while the more significant point differences have more mean error weight. It is defined by:

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2} \quad (20)$$

2.7.1 Application – Combination of forecasting methods in a time series of the Itaipu flow

In the study of Lazzarotto et al. [2015a], considering that there are a relatively small number of methods of combining individual predictive models that make the processing of the data, the linear combination wavelet SARIMA-ANN with multiple stages is proposed. The proposed method can be described in the following basic steps:

1. measurements are obtained in a time series to be modeled;
2. it is conducted at a p level of decomposition, thus generating the first approximation component and p detail components;
3. it is modeled each $p + 1$ Wavelet components, (WC) from step (2) through a model of Box and Jenkins (to map a plausible linear structure), and an artificial neural network (to map a non-linear structure plausible);
4. combine the predictions for each WC originating from both the individual predictors mentioned by traditional linear forecasts, where the adjustment of adaptive parameters is made employing non-linear programming (so long as they are $p + 1$ Wavelet components, linear combinations are required and, therefore, nonlinear programming problem in this step, and
5. linearly combine the predictions of each hybrid WC is generated in step (4), developing the predicted the desired time series. In step (5), solving over a non-linear program is necessary, totaling $p + 2$ nonlinear programs.

To illustrate the hybrid proposed methodology, it was modeled time series of the average monthly flow of Itaipu, which period is from January 1970 until December 2010 (i.e., 492 observations). Their choice was due to its relevance and difficulty in modeling. For experimental purposes, 80% of the first data were used in the training of individual models (Box & Jenkins and artificial neural networks), as well as to obtain the optimum parameters of the adaptive forecasts combinations; 10% of the data were used in subsequent validation sample; and 10% of remaining in the test sample. Short-term predictions were made a step forward, in a horizon of 49 steps ahead. To assess the performance of predictive methods, residuals were MAE (mean absolute error) and MAPE (mean fundamental percentage error).

The choice of level 2 for the Wavelet decomposition was made to preserve parsimony in the number of individual models to be estimated. The base orthonormal Haar wavelet was used to provide the best forecasts.

For comparison, the modeling of time series was carried out above using the methods of artificial neural networks (ANN) and Box & Jenkins (BJ) with a traditional approach. Moreover, the predictive methods used by BJ and ANN integrated with the two-level wavelet decomposition (respectively denoted by ANN-WAVELET and BJ-WAVELET) Donoho and Johnstone [1994a].

Finally, were used the estimates of linear combinations of approaches BJ and ANN (CL1) and the ANN-WAVELET and BJ-WAVELET (CL2) as the approach of Flores and White [2008]. The

process of adjusting the adaptive weights CL1 and CL2 occurred using a PPM, in which the objective function has the minimum mean square error of the training sample residuals (as usual) and the convex weights (i.e., not normalized and negative) Faria and Mubwandarinkwa [2008].

Table 1 note that CL1 was superior to the individual ANN and BJ predictive methods. The same can be checked using the combination CL2 method when compared with the methods based on ANN-WAVELET and BJ-WAVELET.

Table 1: MAPE values, in samples of training, validation and test.

Methods	MAPE (%)		
	Training	Validation	Test
ANN	16.35	15.19	21.78
BJ	16.60	19.46	21.25
CL1	14.95	14.82	19.13
ANN-WAVELET	2.11	2.47	2.80
BJ-WAVELET	2.74	3.13	3.63
CL2	2.06	2.38	2.64
Proposed Method	1.27	1.27	1.36

In Figure 2, there is a comparison between the proposed method and the CL2 approach (which was the best in Table 1 between the benchmark methods). Note that, in every 49 months, in the test sample, the present process showed stability over the time course of the APE values of CL2 (for example, in 6 months, the APE values of the CL2 violate the range of 6% while this is at no time in the proposed method). Moreover, in most instants, the APE proposed combination shows values lower than those of the CL2.

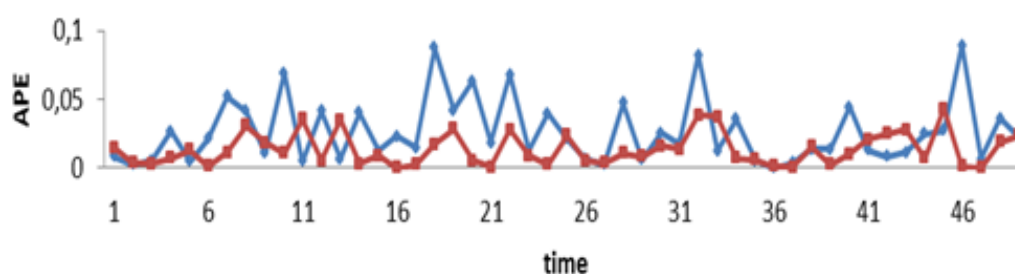


Figure 2: Temporal out-of-sample evolution of Absolute Percentage Error (APE) values of the CL2 model (blue line) and the proposed method (red line).

Source: Lazzarotto et al. [2016]

Analyzing the results, it follows that, in Table 1, the method proposed by the statistics MAPE and MAE were superior to all others, the training samples, validation, and testing – which shows more power learning (the sample workout) and generalization (in the validation and test samples). In Figure 2, in turn, obtained from the second-best method of Table 1, (namely, CL2), greater evolution temporal stability of the APE values (absolute percentage error).

2.7.2 Application – Combination of prediction methods in a time series of dam crest displacement in Itaipu

In the work of Lazzarotto et al. [2015b], a methodology is proposed in which a performance evaluation is made through the prediction of the time series of the readings of an instrument of

the dam of the hydroelectric plant of Itaipu using a hybrid combination of the ARIMA-GARCH predictors and artificial neural networks, using the *Wavelet* decomposition.

In this work, we carried out a comparative study between different approaches to predict the time series of data from the reading of a direct pendulum, which measures the displacement of the dam crest in the direction of the river flow of the dam of the Itaipu hydroelectric plant, located on the Parana River, between Brazil and Paraguay. The objective was to compare the performance of the predictions performed in each approach and to show that the use of the *Wavelet* decomposition and the combination of predictive methods such as ARIMA-GARCH and artificial neural networks can lead to relevant improvements in the performance of the prediction of the behavior of dams.

The autoregressive integrated moving average models (ARIMA), proposed by Box et al. [2008], are linear predictors established in the literature for the analysis and forecasting of time series. In contrast, the generalized autoregressive conditional heteroscedasticity models (GARCHs) are used for the mapping of volatility structures (Gujarati and Porter [2010]).

Artificial neural networks (ANNs) have attracted increasing attention to time series forecasting in the last decades and have achieved successful applications in nonlinear forecasting problems. ANNs can capture non-linear self-dependence structures (Adhikari and Agrawal [2013]).

Hybrid methodologies that combine ARIMA and artificial neural networks for predicting classical time series in the literature were proposed by Zhang [2003]. Their results showed an improvement in prediction performance compared to the two models used separately.

The *Wavelets* functions are used to decompose a time series in terms of an approximation wavelet component and residual detail series. The approximation component is the main one and can be considered a 'smoothed' version of the original series. On the other hand, the detail components represent stochastic fluctuations of low variability and high frequency around the approximation component. The original series is reconstructed by summing all its features generated from itself Kriechbaumer et al. [2014].

This work was limited to a direct pendulum located in the E section of the plant's dam, which is the only instrument responsible for measuring the variable displacement of the dam crest in the direction of the flow and normal to it about its base.

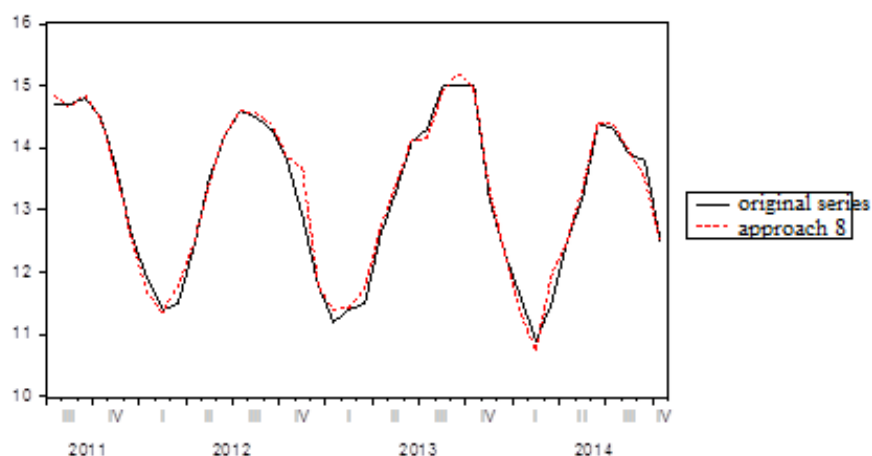


Figure 3: Performance of the best approach in the validation period.
Source: Lazzarotto et al. [2015b]

Eleven comparison approaches were constructed in which the use or not of wavelet decomposition,

ARIMA-GARCH modeling, with neural networks, and hybrid combinations of these techniques were varied. The best performance for the entire validation period was presented by approach 8, which consisted of the wavelet decomposition of the original series followed by a hybrid model, in parallel, of the sum of the optimal ARIMA-GARCH combinations with ANN in each component.

Figure 3 showed the chart in the validation period (July/2011 to October/2014) of the original series and the best approach (approach 8, as described above) according to all performance measures to produce the minor forecast error.

The results showed that the proposed hybrid combination achieved performance in terms of accuracy much better when compared to the individual use of traditional predictive methods and their combinations.

2.7.3 Application ArimaX-Garch Wavelet Neural model

This section presents the main numerical results of applying the traditional Arima-Garch model (benchmark) and the ArimaX-Garch Wavelet Neural model – AXGWN – method to a time series of daily displacement measurements of a block of the dam of the Itaipu hydroelectric plant.

A time series of physical displacement of a dam block that supplies the Itaipu hydroelectric plant is modeled due to its statistical properties and relevance. In Figure 4, its graph shows the temporal behavior of the 2,554 observations, in daily intervals, which were collected from the ADAS automated system.

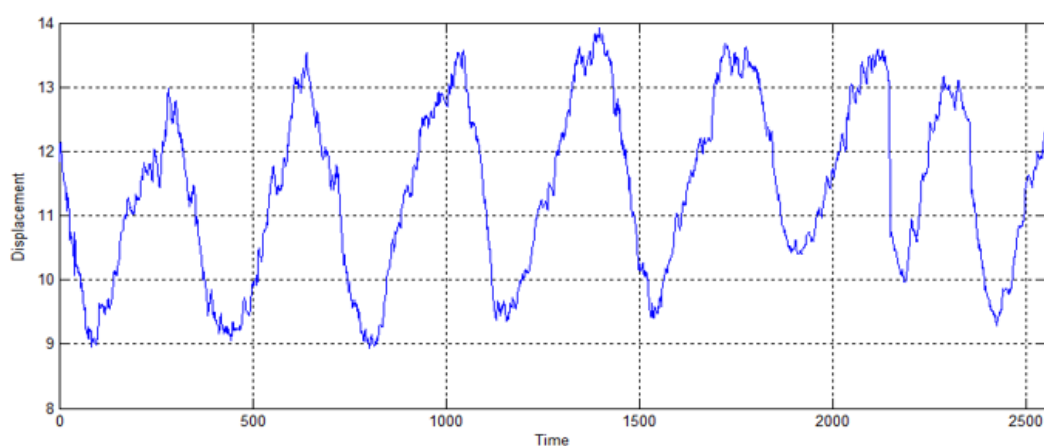


Figure 4: Daily time series of physical displacement of a dam block of Itaipu.

Source:

This section presents the main numerical results of applying the traditional Arima-Garch model (benchmark) and the AXGWN method (proposed) to a time series of daily measurements of the displacement of a block of the dam of the Itaipu hydroelectric plant.

These temporal data comprise the period from October 28, 2005, to October 24, 2012, totaling 2554 points. Of these, the first 2506 observations were used to obtain the optimal parameters of the two predictive methods involved (in other words, such data compose the training sample), while the 48 subsequent observations from the test sample, considering a prediction of 48 steps ahead.

This section presents the main numerical results of applying the traditional Arima-Garch model (benchmark) and the AXGWN method (proposed) to a time series of daily measurements of the displacement of a block of the dam of the Itaipu hydroelectric plant.

For residual diagnoses, in the modeling stage, the following criteria were used: Ljung-Box test (to verify the existence of significant autocorrelation up to lag 36); Durbin-Watson statistics (to analyze empirical evidence of lag one autocorrelation); graphical analyzes of FAC and FACP (to visualize, respectively, the presence of significant superficial and partial linear self-dependence, up to lag 36); ARCH test (for verification of evidence of substantial conditional heteroscedasticity, up to lag 36) (Hamilton [1994]). All Arima-Garch models and their residual tests were performed in software EViews 8 (www.eviews.com/).

The Arima-Garch model with greater adherence to the time series of displacement was the Arima-Garch (3, 2, 4) (2, 1) with logarithmic transformation followed by double differentiation. All estimates of maximum likelihood in both the conditional mean equation and the conditional variance parameters were statistically significant at the 1% level and satisfied the stationary and invertibility conditions. In all residual tests, we accepted the hypothesis of no autocorrelation in the ordinary. We squared standardized residuals at 1% of significance – meaning that the Arima-Garch model captured a plausible auto-dependence structure. In particular, the Durbin -Watson statistic, calculated from ordinary standardized residues, obtained the value of 2.092367, suggesting the lack of lag one autocorrelation (close to the reference value 2). It also stresses that among all identified plausible models, Arima-Garch were adopted, as selection criteria, the following grip statistics, in the training sample: Absolute Percentage Error (APE), Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), e explanation coefficient (R^2).

For the implementation of AXGWN method was used MATLAB software version 2013th (www.mathworks.com/) and EViews software – version 8 (www.eviews.com/). First, it was performed in MATLAB, a level 2 wavelet decomposition of the time series workout sample, i.e., y_t ($t = 1, \dots, 2056$) of the displacement in question, to generate : a Level 2 CW approach, which is denoted by $y_{A_2,t}$ ($t = 1, \dots, 2056$); two levels of detail CWs 2 and 3, which are denoted respectively by $y_{D_2,t}$ ($t = 1, \dots, 2056$) and $y_{D_3,t}$ ($t = 1, \dots, 2056$). Its graphics may be viewed Figure 5. To do this, they used the original scale functions $\phi(\cdot)$ and wavelet $\omega(\cdot)$ of Daubechies family with zero momentum equal to 40 (y_t ($t = 1, \dots, 2506$)). The wavelet family selection criterion was based on the minimization of $\sum_{t=1}^{2056} \varepsilon_t^2$, where: $y_t = y_{A_2,t} + y_{D_2,t} + y_{D_3,t} + \varepsilon^t$, for every $t = 1, \dots, 2056$. Families Daubechies equal to zero moment were tested at 1 to 45 (i.e., 45 different families).

In step 2 of AXGWN method, each of the three components wavelet $y_{A_2,t}$, $y_{D_2,t}$ and $y_{D_3,t}$ ($t = 1, \dots, 2506$) was modeled using an ANN-MLP (described in section 2.3), in order to produce the following sequences forecasts, 48 steps outside of the sample (i.e., $h = 48$): $\hat{y}_{A_2,t}$, $\hat{y}_{D_2,t}$ and $\hat{y}_{D_3,t}$ ($t = 2507, \dots, 2554$). Table 3 shows the optimal settings of MLP's ANN from each of the three CWs.

In step 3, the three components of the wavelet step 1 are supplemented by 48 estimates generated in step 2, generating the three supplemented wavelet components (denoted by $\tilde{y}_{A_2, t,C}$, $\tilde{y}_{D_2, t,C}$ e $\tilde{y}_{D_3, t,C}$ or, equivalently for x_1 , x_2 and x_3). These are interpreted as if they were, in fact, conventional exogenous variables. In step 4, in turn, such wavelet artificial exogenous variables are used in an Arima-Garch model.

In step 4, all three artificial exogenous variables wavelets obtained in step 2 were used, but in a lagged way. Moreover, the prediction of $\hat{\sigma}_t^2$ provides an artificial neural network effect to construct the forecast of the conditional mean \hat{y}_t .

In step 5, errors training forecast ArimaX-Garch model with Wavelet Variables Random (WVR) were modeled by an ANN-MLP to identify a nonlinear structure plausible in residual error – as suggested [8]. For ANN-MLP, we used the software MATLAB. Topological configurations of optimal

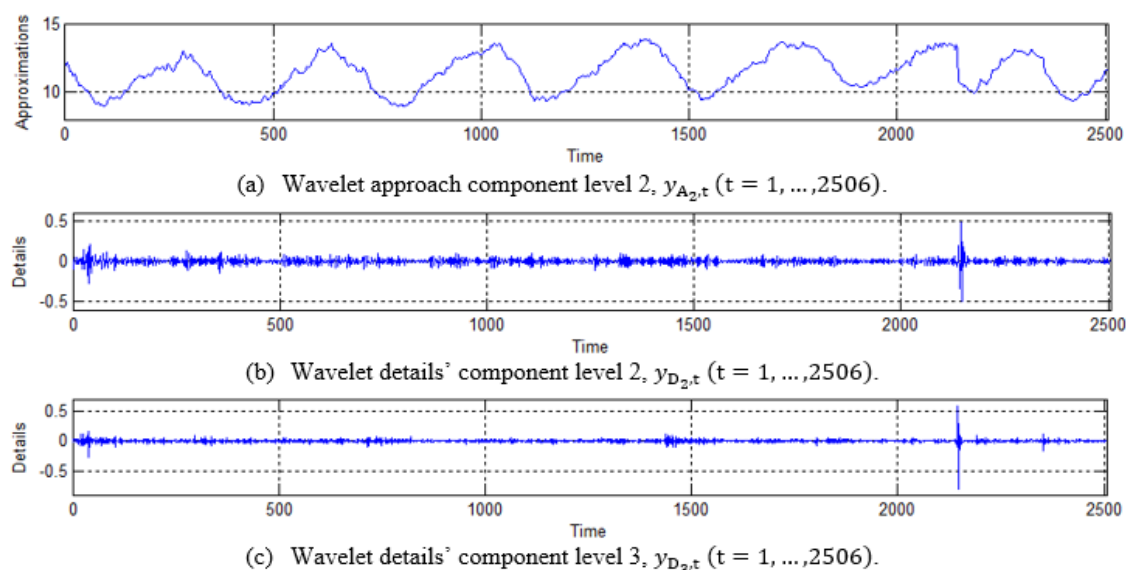


Figure 5: Daily time series of physical displacement of a dam block of Itaipu.

Source:

Table 2: the ANN – MLP's Settings for the wavelet components.

<i>Neural architect. parameters</i>	<i>Approach component level 2</i>	<i>Details' component level 2</i>	<i>Details' component level 2</i>
Transformation	premmx	premmx	premmx
Input window size	5	2	2
Number of hidden layers	1	1	1
Number of neurons in the hidden layer	3	3	5
Activation functioning of the hidden layer	Hyperbolic tangent	Hyperbolic tangent	Hyperbolic tangent
Output layer	Linear activation functioning	Linear activation functioning	Linear activation functioning
Training algorithm	Levenberg-Marquardt [Millar, 2011]	Levenberg-Marquardt [Millar, 2011]	Levenberg-Marquardt [Millar, 2011]

ANN-MLP were: premmx standardization; window size equal to 1; 1 hidden layer consisting of 8 neurons with hyperbolic tangent activation function; output layer with linear activation function; and training algorithm of Levenberg-Marquardt (Haykin [2001]; Millar [2011]). Table 2 Provides the values obtained by adherence statistics MAPE and MAE in the training and test samples.

Table 3: Performance forecasts of Arima – Garch and AXGWN

METHOD	MAPE		MAE	
	<i>In sample</i>	<i>Out of sample</i>	<i>In sample</i>	<i>Out of sample</i>
Arima-Garch	0.719%	0.506%	0.0813	0.0602
AXGWN	0.406%	0.295%	0.0457	0.0350

Please note that, in Table 3, the method AXGWN got better predictive performance relative to the Arima-Garch model, both within the sample (training samples) and out of the sample (test sample).

More precisely, the AXGWN method produced a relative predictive gain over the traditional Arima-Garch model, as the statistical MAPE, equal to 43.53 % in the training sample, equivalent to 41.70 % in the selection of the test; whereas, concerning statistical MAE, it was 43.79 % in the training sample and 41.87 % in the test sample. The estimates should be the most accurate possible so that the operational and financial decision-making, which is excellent, could be run at the right time. It must be stressed that implementing active actions to repair the dam or any unnecessary investments done at inopportune times can lead, among other things, to high financial losses.

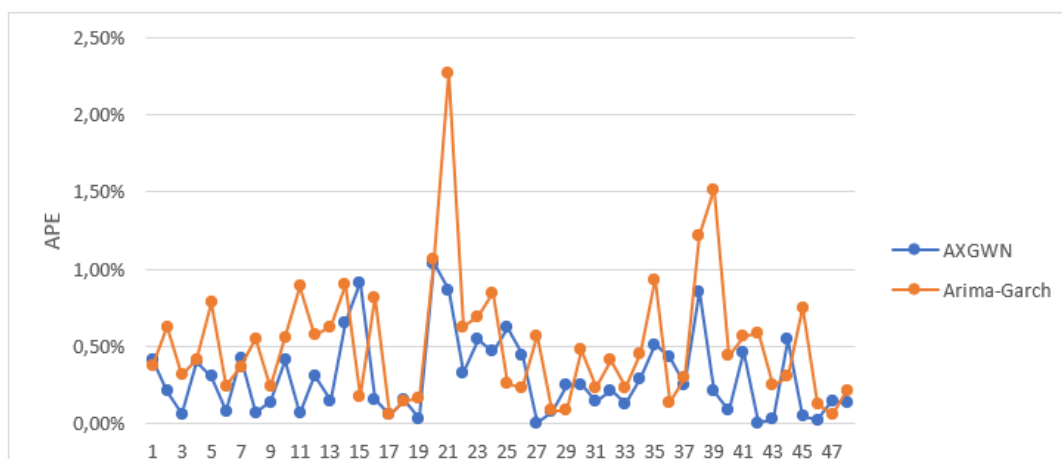


Figure 6: Comparison of temporal evolution of the EPA, in the test sample, the Arima-Garch and AXGWN.

Source:

Figure 6 shows the graph of Absolute Percentage Error (EPAs) calculated for 48 forecasts out of the sample for both models. Please note that the model AXGWN shows that the EPAs were lower than the Arima-Garch model at most 48 points, besides presenting more stability (i.e., smaller “peaks”) in fluctuating errors. For example, at the 1% level, only one value of EPA is violated by AXGWN, while three violations occur for the Arima-Garch model. Considering the level of 0.5%, there are 21 violations for the Arima-Garch model and 9 for AXGWN. Finally, the maximum APE of the Arima-Garch model was approximately 2.25%. In contrast, the maximum EPA AXGWN was slightly more than 1% (which means a gain of predictive accuracy in relative terms, around 50%).

In terms of the ratio R^2 (which measures the variability explained by the models in the training sample), the model Arima-Garch obtained $R^2 = 0.361378$ and AXGWN, of $R^2 = 0.997249$.

2.7.4 Application - Method Support Vector Regression Wavelet of Multiple Kernels (SVRWMN)

In this study, time series predictions were made from measurements of the instrument called direct pendulum denoted by (CO-I-2/X), Located in the key-block (I10) situated on the stretch I of the Itaipu hydroelectric dam, conform Figure 1, and that contributes to the monitoring of the behavior of the block through the hybrid predictive method called SARIMA Support Vector Regression Wavelet of Multiple Kernels (SSVRWMN), to produce predictions that are aggregators of distinct stochastic information captured by different ways. The confidence interval is estimated using the Bootstrap technique with the point predictions. The methodology adopted and the results obtained are described in this section and were published in the work Rodrigues [2015].

The pendulum measures the displacements of the dam; that is, it measures displacements in the

downstream direction, perpendicular to the flow of the river Osako [2002]. The direct pendulum is an instrument set at the top of the dam that makes contact with the concrete rock; that is, it is installed on the structure and is responsible for monitoring displacements of the top of the dam relative to the point considered as fixed in the foundation. Two measurements are obtained from the direct pendulum: one in the upstream-downstream direction (CO-I-2/X) and another in the right-left bank direction (CO-I-2/Y).

In this application, the CO-I-2/X will be considered, which shows the horizontal displacements of the top of the dam in the upstream-downstream direction. The location of the CO-I-2 pendulum in the key-block (I10) can be seen in Figure 7.

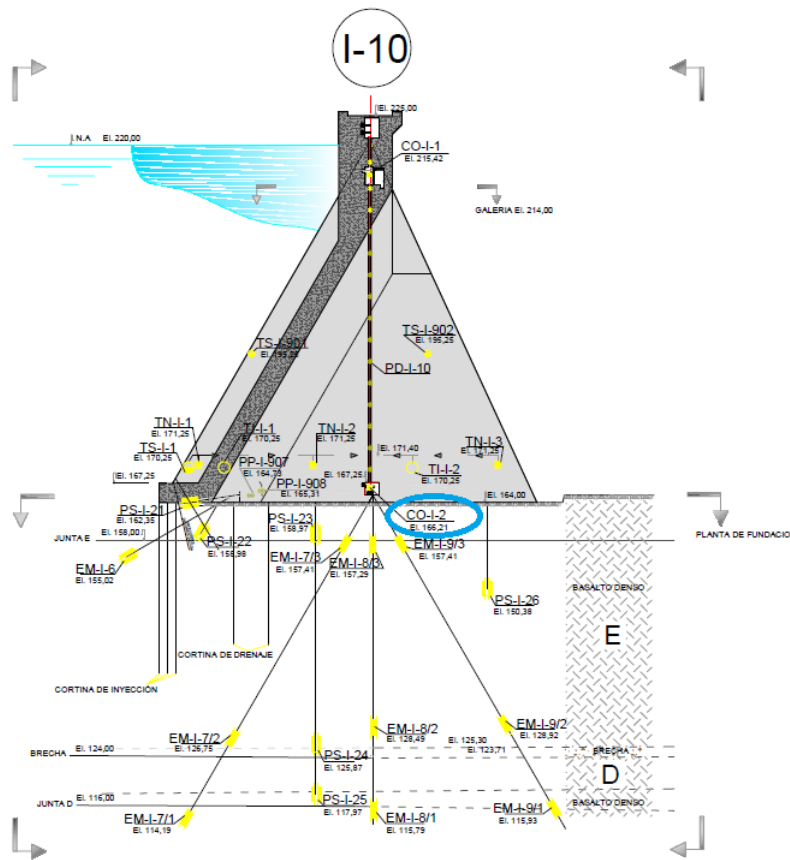


Figure 7: Locating of the CO-I-2 pendulum in the key-block (I10).
Source: Rodrigues [2015]

For the time series from the measurement of the direct pendulum instrument CO-I-2/X, the values are given in millimeters and are shown in Figure 8.

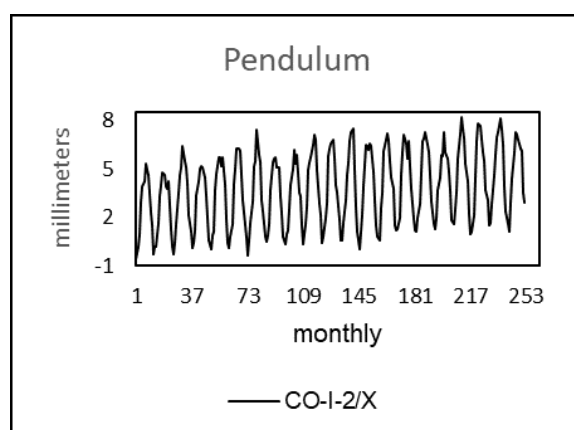


Figure 8: Original series of the pendulum.
Source: Rodrigues [2015]

SARIMA Support Vector Regression Wavelet method of Multiple Kernels considers a time series $(y(t))_{t=1}^T$ to be projected through the average of the forecasts obtained by a linear model (SARIMA) and a non-linear one (Support Vector Regression Wavelet of Multiple Kernels).

The SSVRWMN is given in six stages:

In stage 1 the modeling through SARIMA is proposed, represented in the original time series by $(y(t))_{t=1}^T$. The proposed hybrid model requires the forecast of a linear component $(L(t))_{t=1}^{T+h}$ and for this purpose are considered the forecasts obtained by the SARIMA model which aim at capturing the structures of linear self-dependence, taking into consideration the seasonal effect. In case there is no seasonality, the ARIMA model will be considered.

The estimates of the parameters of the SARIMA model are determined by Maximo Likelihood using Marquardt's algorithm Marquardt [1963], implemented in the software EViews8. The definition of the orders of the models, initially, occurs with the analysis of the charts of the self-correlation functions FAC, partial self-correlation FACP, residues, and tests with several options of orders p , d , and q .

In stage 2 the Wavelet orthogonal decomposition of level $(y(t))_{t=1}^T$ in the original time series is performed, generating $r + 1$ CWs, Donoho and Johnstone [1994b]. In this study, the level of composition $r = 2$ was set, and therefore for each Wavelet base three-time series were obtained: the time series that represents the approximation component indicated by $(\tilde{y}_{A_2}(t))_{t=1}^T$ and two times series that represent the detail components represented by $(\tilde{y}_{D_2}(t))_{t=1}^T$ and $(\tilde{y}_{D_3}(t))_{t=1}^T$. The following bases were utilized: daubechies of null moments from 1 to 45 (db1, db2, . . . , db45), coiflets of null moments from 1 to 5 (coif1, coif2, . . . , coif5), symlets of null moments from 1 to 30 (sym1, sym2, . . . , sym30) and bi-orthogonal of null moments $N_r=1$ and $N_d=1,3$ and 5, $N_r=2$ and $N_d=2,4,6$ and 8, $N_r=3$ and $N_d=1,3,5,7$ and 9, $N_r=4$ and $N_d=4$, $N_r=5$ and $N_d=5$ and $N_r=6$ and $N_d=8$, Daubechies [1992]. Thus, for each time series $(y(t))_{t=1}^T$ to be modeled 255 sub-series were obtained.

The Wavelet components of approximation and detail were obtained in this paper from the software Matlab 8.0, which offers several options of Wavelet base.

In stage 3, the modeling of Wavelet Components is carried out through SVR Multiple Kernels. The methods named SVR Multiple Kernels take into consideration a time series in which the modelings are carried out through the SVR of the RBF kernel, SVR Polynomial kernel, and SVR Sigmoidal

kernel separately. The forecasts obtained through the combination of projections coming from the SVR modelings of SVR kernel RBF, SVR Polynomial kernel, and SVR Sigmoidal kernel is called Support Vector regression of Multiple Kernels (SVRMN). In stage 2, three Wavelet components were obtained for each Wavelet base, modeled through the SVRMN method.

It being so, the decomposed time series in the previous stage and represented by $(y(t))_{t=1}^T \cong (\tilde{y}_{A_2}(t))_{t=1}^T + (\tilde{y}_{D_2}(t))_{t=1}^T + (\tilde{y}_{D_3}(t))_{t=1}^T$ brings forth the following sequence of forecasts within and without the sample, conform Equation (21):

$$\begin{aligned} & (\hat{y}_{RBF_A_2}(t))_{t=T'}^{T+h}, (\hat{y}_{RBF_D_2}(t))_{t=T'}^{T+h}, (\hat{y}_{RBF_D_3}(t))_{t=T'}^{T+h}, \\ & (\hat{y}_{Poly_A_2}(t))_{t=T'}^{T+h}, (\hat{y}_{Poly_D_2}(t))_{t=T'}^{T+h}, (\hat{y}_{Poly_D_3}(t))_{t=T'}^{T+h}, \\ & (\hat{y}_{Sigm_A_2}(t))_{t=T'}^{T+h}, (\hat{y}_{Sigm_D_2}(t))_{t=T'}^{T+h} \text{ and } (\hat{y}_{Sigm_D_3}(t))_{t=T'}^{T+h} \end{aligned} \quad (21)$$

The forecast of the series is given by Equation (22):

$$\begin{aligned} (\hat{y}_{Base\ Wavelet}(t))_{t=T'}^{T+h} = & \frac{1}{3} [\hat{y}_{RBF_A_2}(t) + \hat{y}_{RBF_D_2}(t) + \hat{y}_{RBF_D_3}(t)]_{t=T'}^{T+h} + \\ & + \frac{1}{3} [\hat{y}_{Poly_A_2}(t) + \hat{y}_{Poly_D_2}(t) + \hat{y}_{Poly_D_3}(t)]_{t=T'}^{T+h} + \\ & + \frac{1}{3} [\hat{y}_{Sigm_A_2}(t) + \hat{y}_{Sigm_D_2}(t) + \hat{y}_{Sigm_D_3}(t)]_{t=T'}^{T+h} \end{aligned} \quad (22)$$

where h is the limit of forecast; T' are the degrees of liberty lost by the SVR model; and $T' < T$ (T' not necessarily the same in each forecast).

The following forecasts were generated for each Wavelet base:

$$\begin{aligned} & (\hat{y}_{db_i}(t))_{t=T'}^{T+h}, i = 1, 2, \dots, 45 \\ & (\hat{y}_{sym_i}(t))_{t=T'}^{T+h}, i = 1, 2, \dots, 30 \\ & (\hat{y}_{coif_i}(t))_{t=T'}^{T+h}, i = 1, 2, \dots, 5 \\ & (\hat{y}_{bior_i}(t))_{t=T'}^{T+h}, i = 1.1, 1.3, 1.5, 2.2, 2.4, 2.6, 2.8, 3.1, 3.2, 3.5, 3.7, 3.9, 4.4, 5.5, 6.8 \end{aligned}$$

For modeling by SVR, a library for SVR for Matlab called LIBSVM proposed by Chang and Lin [2011] was used, as well as necessary routines.

In stage 4, the modelings that presented the minor average quadratic error (MSE) are selected based on the validation data, for each Wavelet family, in the previous stage.

Therefore the following forecasts are selected: $(\hat{y}_{db_{better-validation}}(t))_{t=T''}^T$, $(\hat{y}_{sym_{better-validation}}(t))_{t=T''}^T$, $(\hat{y}_{coif_{better-validation}}(t))_{t=T''}^T$, and $(\hat{y}_{bior_{better-validation}}(t))_{t=T''}^T$, with $T'' < T$, so that the linear combination is carried out next in order to obtain the forecast of the non-linear component $(N(t))_{t=1}^T$ of the proposed hybrid method, described in Equation (23):

$$\begin{aligned} (\hat{N}(t))_{t=T''}^T = & \rho_1 \times (\hat{y}_{db_{better-validation}}(t))_{t=T''}^T + \rho_2 \times (\hat{y}_{sym_{better-validation}}(t))_{t=T''}^T + \\ & + \rho_3 \times (\hat{y}_{coif_{better-validation}}(t))_{t=T''}^T + \rho_4 \times (\hat{y}_{bior_{better-validation}}(t))_{t=T''}^T \end{aligned} \quad (23)$$

where $\hat{N}(t)$ is the combined forecast of the $N(t)$ non-linear component; and ρ_i , $i = 1, \dots, 4$ are parameters to be determined with the optimization of the problem of non-linear programming given

in Equation (24) whose function objective consists of the minimization of the average quadratic error (MSE).

$$\begin{aligned} \text{Minimize } MSE &= \frac{1}{T-T''} \sum_{t=T''}^T (y(t) - \hat{N}(t))^2 \\ \text{subject to } \rho_1, \rho_2, \rho_3, \rho_4 & \quad \text{unrestricted} \end{aligned} \quad (24)$$

For the optimization of Problem (24) of non-linear programming, the solver package of the Excel 2013 software was used. The forecasts obtained in this stage are part of the Support Vector Regression Wavelet of Multiple Kernels, or simply SVRWMN.

In stage 5 linear forecast $(\hat{L}(t))_{t=T''}^{T+h}$ and non-linear forecast $(\hat{N}(t))_{t=T''}^{T+h}$ are taken into consideration for obtaining the one-off forecasts for each t instant. Equation (25) presents the generation of forecasts originating from the proposed method, thus generating forecasts within and without the sample, which will be denoted by $(\hat{y}_{SSVRWMN}(t))_{t=T''}^{T+h}$, where each $\hat{y}(t)$ is, in fact, a set of information from different sources.

$$(\hat{y}_{SSVRWMN}(t))_{t=T''}^{T+h} = \frac{1}{2} (\hat{L}(t) + \hat{N}(t))_{t=T''}^{T+h} \quad (25)$$

Therefore, the $(\hat{y}_{SSVRWMN}(t))_{t=T''}^{T+h}$ list of forecasts, within and without the sample (multi step forecasts) consists of the results of the proposed method, mentioned as SARIMA Multiple Kernel Support Vector Regression Wavelet.

And finally, in stage 6, it is expressed that the forecasts within the sample $(\hat{y}_{SSVRWMN}(t))_{t=T''}^T$ present simultaneously information coming from modeler that captured linear and non-linear self-dependence structures and may be considered as a filtered version of the $y(t)$ time series, both through a linear and a non-linear filter.

In this way, the forecast errors within the $(e(t))_{t=T''}^T = (y(t) - \hat{y}_{SSVRWMN}(t))_{t=T''}^T$ sample consist, in fact, of the realization of a stochastic process of the blank noise kind (confirmed by the tests of residual self-correlation, BDS statistical independence and Dickey-Fuller stationarity) Hamilton [1994]. By utilizing the residual Bootstrap technique suggested by Tibshirani [1996] the standard

error of the t^{th} value predicted can be estimated using $\left[\frac{1}{B-1} \sum_{b=1}^B (\hat{y}^b(t) - \bar{\hat{y}}^b(t))^2 \right]^{1/2}$, where

$\bar{\hat{y}}^b(t) = \frac{1}{B} \sum_{b=1}^B \hat{y}^b(t)$ and therefore the trust interval is determined in accordance with Equation (26):

$$\bar{\hat{y}}^b(t) - t_{1-\frac{\alpha}{2}, gl} \hat{\sigma}_g(x) \leq f(\mathbf{x}, \mathbf{w}) \leq \bar{\hat{y}}^b(t) + t_{1-\frac{\alpha}{2}, gl} \hat{\sigma}_g(x) \quad (26)$$

where $\hat{\sigma}_g(x) = \left[\frac{1}{B-1} \sum_{b=1}^B (\hat{y}^b(t) - \bar{\hat{y}}^b(t))^2 \right]^{1/2}$ is the estimated standard error, $t_{1-\frac{\alpha}{2}, gl}$ comes from the t distribution of Student with α significance level and gl are the degrees of liberty. If the degrees of liberty are defined as the difference between the size of the step (number of forecast steps ahead) and the number of models utilized, that is, $gl = h - 2$, $h > 3$ where $h = 12$, it represents the steps ahead to be considered in the forecast. We consider the number of models equal to 2, one linear and the other non-linear.

Initially, the information processing for the time series was carried out to be structured according to the forecasting methodology. The period to be considered was 1994-2014, with monthly values.

Of the 252 values that the time series possesses, the first 228 data were used to adjust the base models (SARIMA and SVR, SSVRWMN); the next 12 values were reserved for validation, and the last 12 values, referring to the year 2014, for testing. The methodology used is a multi-step ahead forecasting, that is, h steps ahead; in this application, $h = 12$; therefore, the limit of forecasting coincides with the total of steps. To evaluate the methods, MAPE (mean absolute percent error), MSE (mean square error), and MAE (mean absolute error) errors Hamilton [1994] represented in Equation (17 to 20) were used.

We shall denote the time series of the direct pendulum by $(y_{co}(t))_{t=1}^{252}$

In stage 1, for the direct pendulum time series, the modeling was carried out through a SARIMA model, generating predictions within the training sample and predictions in the test sample on a horizon of 12 steps ahead; these predictions are considered by the hybrid method as the linear component prediction.

For the SARIMA modeling, the Eviews software was used. The visual analysis of the time series was made, and the appropriate model was identified. The definition of the orders of models was done by analyzing the charts of the self-correlation functions FAC, partial self-correlation FACP, residues, and tests with several options of orders p, d, q . The statistics found through the analysis of the time series of the direct pendulum are represented in Table 4.

Table 4: SARIMA modelling for time series y_{co} .

Variable	Coefficient	Std. Error	t-Statistic	Prob
AR(12)	1.001462	0.006894	145.2634	0.0000
MA(1)	-0.632463	0.049493	-12.77889	0.0000
MA(3)	-0.416386	0.054726	-7.608606	0.0000
MA(9)	0.085333	0.036069	2.365816	0.0189
SMA(12)	-0.843251	0.038201	-22.07423	0.0000
R-squared	0.832369	Mean dependent var		0.009692
Adjusted R-squared	0.829348	S.D. dependent var		1.260694
S.E. of regression	0.520793	Akaike info criterion		1.554850
Sum squared resid	60.21192	Schwarz criterion		1.630290
Log likelihood	-171.4755	Hannan-Quinn criter.		1.585291
Durbin-Watson stat	1.90166			

Source: Rodrigues [2015]

In stage 2 the orthogonal decomposition Wavelet of $r = 2$ level was made for the time series $y_{co}(t)$, considering the daubechies bases (db 1,2,...,45), (coif 1,2,...,5), symlets (sym 1,2,...,30) and bi-orthogonal (1.1, 1.3, ..., 6.8), a total of 95 decompositions, generating an approximation component that is the representation of the original series in low frequency, and two detail components which are high frequency series (more noisy series) in each decomposition carried out, thus obtaining 285 sub-series for $y_{co}(t)$ time serie.

Stage 3 forecasts through Multiple Kernel Support Vector Regression were carried out for each of the 285 sub-series obtained in stage 2. The three main kernels were considered to compose the method: RBF, polynomial, and sigmoidal Smola and Schölkopf [2004].

For example, for the series $(y_{db1_co}(t))_{t=1}^{240}$ the forecasts,

$$\begin{aligned} (\hat{y}_{db1_co_RBF}(t))_{t=12}^{240+12} &= (\hat{y}_{db1_co_RBF A_2}(t))_{t=12}^{240+12} + (\hat{y}_{db1_co_RBF D_2}(t))_{t=12}^{240+12} + \\ &+ (\hat{y}_{db1_co_RBF D_3}(t))_{t=12}^{240+12} \\ (\hat{y}_{db1_co_Poly}(t))_{t=12}^{240+12} &= (\hat{y}_{db1_co_Poly A_2}(t))_{t=12}^{240+12} + (\hat{y}_{db1_co_Poly D_2}(t))_{t=12}^{240+12} + \\ &+ (\hat{y}_{db1_co_Poly D_3}(t))_{t=12}^{240+12} \\ (\hat{y}_{db1_co_Sigm}(t))_{t=12}^{240+12} &= (\hat{y}_{db1_co_Sigm A_2}(t))_{t=12}^{240+12} + (\hat{y}_{db1_co_Sigm D_2}(t))_{t=12}^{240+12} + \\ &+ (\hat{y}_{db1_co_Sigm D_3}(t))_{t=12}^{240+12} \end{aligned}$$

were generated.

$(y_{db1_co}(t))_{t=1}^{240}$ is the forecast of the series, the average forecast, or $(\hat{y}_{db1_co}(t))_{t=12}^{240+12}$. Where $(\hat{y}_{db1_co}(t))_{t=12}^{240+12} = \frac{1}{3} (\hat{y}_{db1_co_RBF}(t))_{t=12}^{240+12} + \frac{1}{3} (\hat{y}_{db1_co_Poly}(t))_{t=12}^{240+12} + \frac{1}{3} (\hat{y}_{db1_co_Sigm}(t))_{t=12}^{240+12}$. This process was repeated for all 95 Wavelet decompositions of each time series.

To obtain the optimum parameters of the SVR model, the data from the time series reserved for adjustment were subdivided into training and validation. The model that presented the lowest Mean squared error (MSE) in the validation set was chosen, being the cardinality of the validation set was equal to 12.

Table 5 shows the two best results obtained for each Wavelet family and the MSE statistics of the validation sample for the time series of the y_{co} pendulum.

Table 5: Two best results for each Wavelet family for series y_{co} .

Families	Base	Validation-Mse
Daubechies	db6	0.02545
	db40	0.01639
Symlets	sym19	0.03968
	sym23	0.02565
Coiflets	coif3	0.17120
	coif4	0.25507
Biorthogonal	bior3.1	0.03982
	bior6.8	0.07165

Source: Rodrigues [2015]

From Table 5, it can be seen that the selected models presented the lowest MSE adherence statistics, and they are for the CO-I-2/X series, the bases db40, sym23, coif3, and bior3.1.

In stage 4, the forecasts of the best modelings were combined when the MSE statistics were compared in the validation set regarding each Wavelet family.

For the time series of pendulum y_{co} the combination is represented by the equation (27):

$$(\hat{y}_{co}(t))_{t=12}^{240} = \left[\begin{array}{l} \rho_1 \times (\hat{y}_{db40_co}(t)) + \rho_2 \times (\hat{y}_{sym23_co}(t)) + \\ + \rho_3 \times (\hat{y}_{coif3_co}(t)) + \rho_4 \times (\hat{y}_{bior3.1_co}(t)) \end{array} \right]_{t=12}^{240} \quad (27)$$

where the adaptive constants assume the following values: $\rho_1 = 0.458526$, $\rho_2 = 0.085159$, $\rho_3 = 0.251612$ and $\rho_4 = 0.221247$, after the minimization of the average quadratic error (MSE).

In this stage, the forecasts of the multiple kernels of the model SVR Wavelet are obtained, that is, of the non-linear component of the modeled time series and which is represented by $(\hat{N}_{co}(t))_{t=12}^{240}$.

In stage 5, the predictions obtained by the linear model (SARIMA) and by the non-linear model (SVRWMN) are taken. The average is calculated to get the point predictions for each t , forming the hybrid method SSVRWMN that presents in its structure of self-dependence linear and non-linear information being considered as a filtered version of the time series $y(t)$, both by a linear filter (stage 1) and by a non-linear filter (stage 3 and 4).

Therefore, the list of forecasts, within and without the sample (multi-step predictions), consists of the results of the proposed method referred to as Multiple Kernel SARIMA Support Vector Regression Wavelet (SSVRWMN).

Then, the list of forecasts for the time series of the direct pendulum is represented by the equation (28):

$$(\hat{y}_{co}(t))_{t=12}^{240+12} = \frac{1}{2} (\hat{L}_{co}(t) + \hat{N}_{co}(t))_{t=12}^{240+12} \quad (28)$$

At this stage, it is possible to compare the forecasts of the proposed method with those obtained through the SARIMA, SVR method, and the combination SARIMA-SVR. The compared adherence statistics are MSE, MAPE, and MAE.

In Table 6, the comparison for the time series of the pendulum is presented.

Table 6: Variables to be considered on the evaluation of interaction techniques.

Methods	MSE	MAPE	MAE
SARIMA	0.3615	12.7%	0.4581
SVR	0.8398	19.8%	0.8419
SARIMA-SVR	0.7146	18.6%	0.7801
SSVRWMN	0.2157	11.8%	0.4149

Source: Rodrigues [2015]

From Table 6 it can be seen that the proposed method SSVRWMN obtained in the forecasts of the time series of instrument CO-I-2/X a reduction in the three adherence statistics MSE, MAPE, and MAE about all other predictive methods listed. Compared with the second-best result (namely, the SARIMA method), the relative reduction was approximately 40.3% in the measure MSE, 9.4% in the action MAE and 7.1% in the measure MAPE.

Such results demonstrate the predictive efficiency of the proposed method against renowned approaches in the time series literature, such as SARIMA, SVR, and the hybrid SARIMA-SVR, in the present case.

In stage 6, after obtaining the residues series, Bootstrap's residual process was used to obtain the trust interval (IC) at the 95% level.

Figure 9 shows the graph of forecasts with the respective trust intervals of the series X, obtained through the hybrid method SSVRWMN – bootstrap.

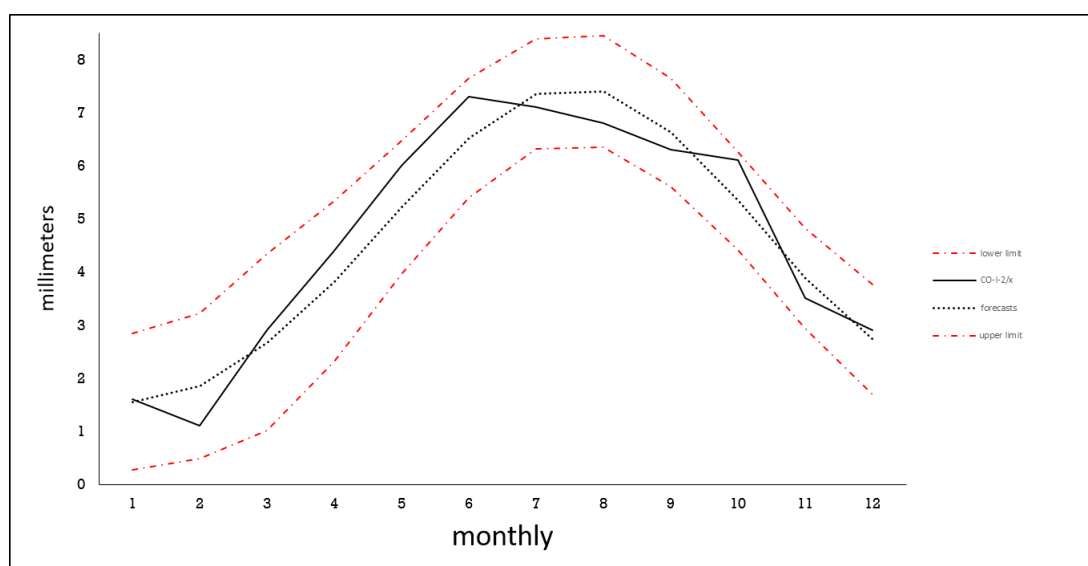


Figure 9: IC graph and forecasts by the SSVRWMN – bootstrap method of the CO-I-2/X.

Source: Rodrigues [2015]

In summary, in generating forecasts, SSVRWMN developed substantially better estimates than those obtained by the predictive methods SARIMA and SVR individually and the composite SARIMA-SVR in all adhesion statistics considered. Therefore, the predictions from the SSVRWMN method add information, thus contributing to decision-making with regard future behavior of the pendulum. Such results encourage the application of this method for time series from other areas. Even if the theoretical base of the hybrid method SSVRWMN – bootstrap is mathematically complex, the use of software and computer routines mentioned in the text permits its utilization.

2.7.5 Application of the SARIMAX-NEURAL Method in Prediction Times Series of the Deformation Concrete Dam

The SARIMAX-NEURAL hybrid method was presented by Ribeiro et al. [2019] as the combination of SARIMAX methods (from the Box & Jenkins methodology) and Recurrent Neural Networks (RNN) Deep Learning with Long Short-Term Memory (LSTM) architecture. The study of concrete deformations from time series measured over 40 years using strain-meter rosettes installed in a buttress concrete block of the Itaipu Dam.

Strain-meter rosettes have recorded the data – RD – with five strain-meters each, installed during construction in five strategic points of a Buttress Block in Itaipu Dam, according to the river flow direction (Fig.10). The rosettes are made up of two trihedral orthogonal. The arms 1, 2, 3 and 4 of the rosette are on the same plane, forming angles of 45° between each component (1 – 2, 2 – 3, and 3 – 4), co-planar to the flow direction, i.e., from upstream to downstream, where the arm 1 is positioned in the downstream order. The arm 5 is perpendicular (90°) to the others (and to the river flow).

The particular modeling methods used were RNN (Recurrent Neural Networks) with hidden layer LSTM (Long Short-Term Memory), SARIMA (Auto-Regressive Integrated Models of Seasonal Moving Average), SARIMAX (SARIMA with Exogenous Variables), and a new hybrid method proposed, called SARIMAX-NEURAL, defined as a linear combination of the SARIMAX and RNN methods.

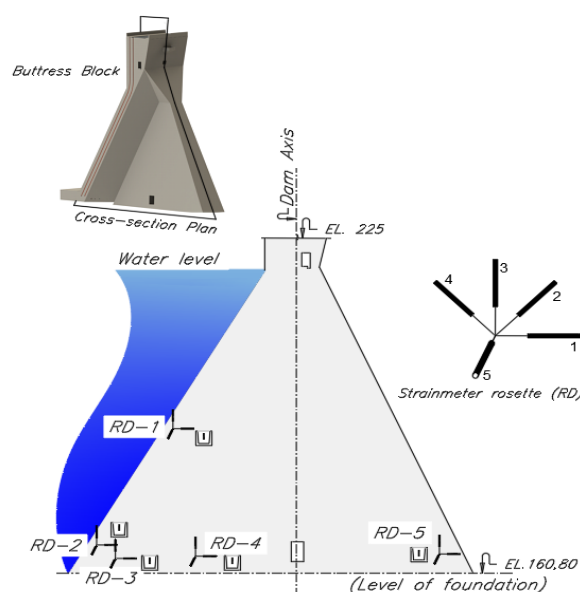


Figure 10: Strain-meter rosettes in the Buttress Block.
Source: Ribeiro et al. [2019]

In the pre-modeling phase of the SARIMA and SARIMAX methods, the lagged correlation between the time series of long-term concrete strains (series of dependent variables) and ambient temperature (series of exogenous variables) was identified. All factors related to the thermal conductivity of the concrete and the geometry of the element were involved in this correlation; that is, the deformations of the concrete were influenced by the ambient temperature sometime later due to the dimensions of the block (64.2 m high; 64.0 m in length) and for each rosette position. The greater the distance between the exposed surface of the block and the part of the rosette, the longer the time required for the deformations to be influenced by the variation in ambient temperature.

The SARIMAX-NEURAL hybrid method consists of two phases. In the first, SARIMAX modeling is performed to determine the linear component. In the second, RNN DL-LSTM modeling is performed in the SARIMAX model residue series to predict the nonlinear element.

Thus, the hybrid method's prediction is obtained as a linear combination of the predictions of the individual models:

$$\widehat{Z}_f(t+h) = \widehat{CL}_f(t) + \widehat{NL}_f(t) \quad (29)$$

Where $\widehat{CL}_f(t)$ is the linear component and $\widehat{NL}_f(t)$ is the non-linear component corresponding to the prediction (out-of-sample) of the time series $\{Z_t\}_{t=1}^N$. Residues of the SARIMAX model within the sample are represented by:

$$\widehat{e}_a(t) = Z_{ta} - \widehat{Z}_{ta} \quad (30)$$

Where Z_{ta} is the real value and \widehat{Z}_{ta} is the predicted value.

The parameters of the Deep Learning neural network models with hidden layer Long Short-Term Memory (RNN DL-LSTM) are the same applied in the individual model.

Regardless of the fit accuracy to in-sample data, some models could not forecast the atypical curve behavior of future data (out-of-sample), according to Fig. 11, for a seasonal period s of 52 weeks per year. This may be related to the functionality of strain meters or the interference of external factors, such as upstream water temperature, deformations due to phenomena other than thermal,

and irregular thermal diffusivity, among others. In the case of RD-2/5 (Fig. 11), for instance, the RNN DL-LSTM prediction model has higher accuracy as compared to other models. Still, it cannot describe the behavior of arm 5 of rosette 2.

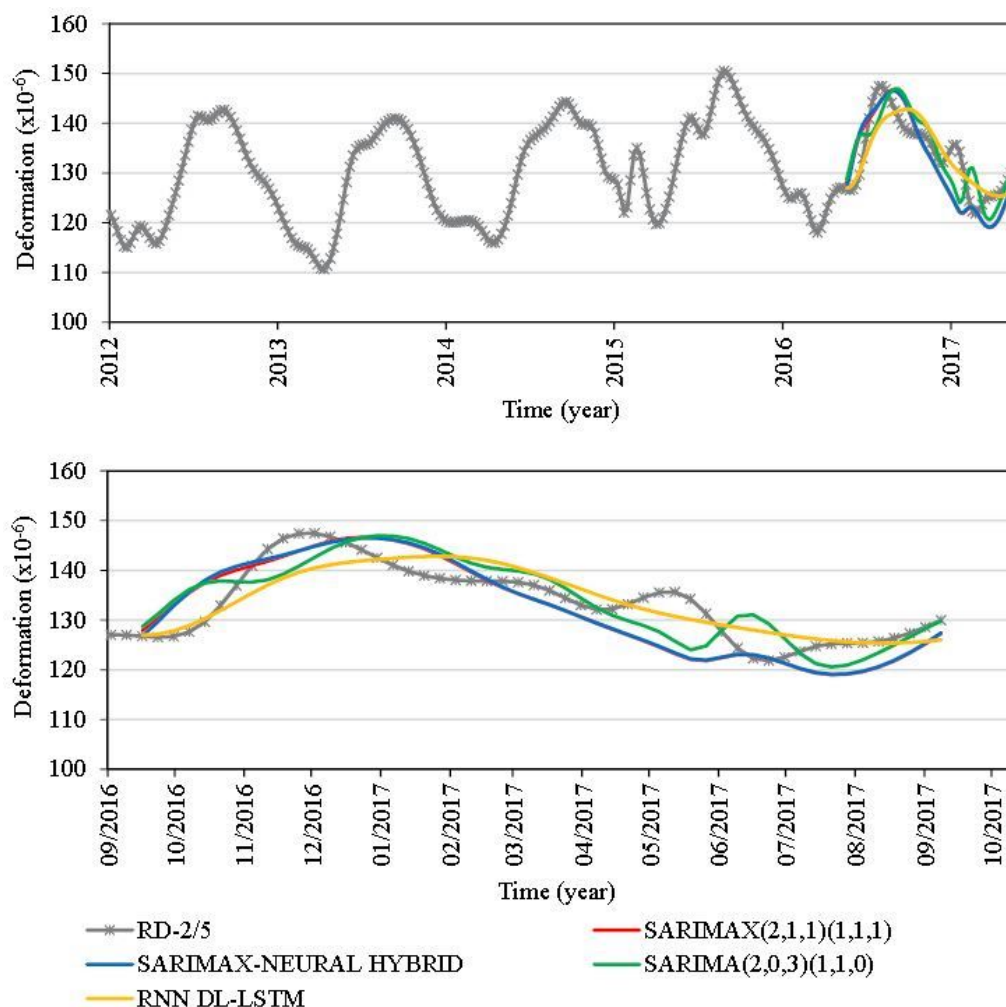


Figure 11: Ability of time series prediction models for RD-2/5

Source: Ribeiro et al. [2019]

Table 7 shows the percentage comparison between the averages of the adherence statistics (MAE, MAPE, RMSE) of the hybrid method about the individual methods for the 25-time series model of the five rosettes in the Fig. 10. Since the hybrid method considers the influence of the exogenous variable (ambient temperature) on the modeling, a predictive gain was obtained from the RNN DL-LSTM and SARIMA neural network models, in which the modeling was performed only on the concrete deformation data.

About SARIMAX, the predictive improvement of the hybrid method was due to the modeling of residues of the SARIMAX model via the RNN DL-LSTM network. The percentage gain was slightly better and indicated that concrete deformation and ambient temperature data have a predominantly linear dependence structure.

The Figure 12 summarizes the mean values of the adherence statistics for each method for the 25-time series modeled. These values correspond to the average of the errors used to calculate

Table 7: Comparison of the mean predictive precision as a percentage of SARIMAX-NEURAL method over individual methods

Individual methods	SARIMAX-NEURAL Method		
	MAE	MAPE	RMSE
RNN DL-LSTM	53.50%	50.31%	50.53%
SARIMA	28.87%	20.71%	27.12%
SARIMAX	1.09%	0.96%	0.51%

Source: Ribeiro et al. [2019]

the percentage predictive improvement of the SARIMAX-NEURAL process about the particular methods of Table 7.

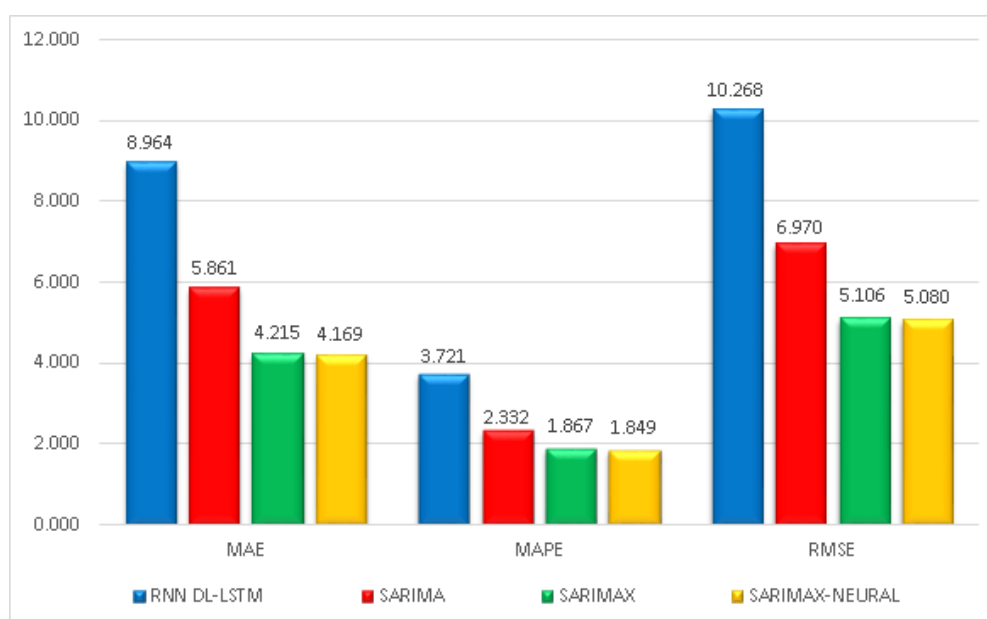


Figure 12: Mean of MAE, MAPE and RMSE adherence statistics of the 25 time series. modeled

Source: Ribeiro et al. [2019]

The results show that temperature-sensitive models as an exogenous variable (SARIMAX and SARIMAX-NEURAL) show better forecasts with actual data than the other two approaches. Therefore, in the hybrid approach (SARIMAX-NEURAL), the RNN DL-LSTM model has been chosen for non-linear mapping effects in the residuals and obtaining more accurate out-of-sample forecasts concerning the SARIMAX model. Thus, the proposed approach yields a slightly better assessment (1%) of concrete long-term deformation behavior forecasting based on monitoring data.

3 Multivariate Analysis

In general, statistical methods of multivariate analysis are applied to analyze multiple measures on each individual or object investigated. These methods have their foundations in matrix algebra.

The multivariate data (qualitative or quantitative) are arranged in a matrix whose columns are the vectors of random variables, and each row corresponds to an object. By random variable is meant

any characteristic observed in an object. Thus, each row of the matrix presents all the features verified in an object of the sample.

A collection of n observations on p different random variables, taken from the same item, composes a multivariate sample. The matrix representation of this sample can be done through a matrix X of order $n \times p$ (31), in such a way that the lines correspond to the observations, the columns to the variables, and the elements x_{ij} to the sample values.

$$X = \begin{pmatrix} x_{11} & \dots & x_{1p} \\ x_{21} & \dots & x_{2p} \\ \vdots & \vdots & \vdots \\ x_{n1} & \dots & x_{np} \end{pmatrix} \quad (31)$$

Statistical analyzes involving multivariate data generally use the degree of the linear relationship between random variables, that is, the covariance between them. The covariance (σ) between two random variables is defined by (32):

$$Cov(X_i, X_j) = \sigma_{ij} = E[(X_i - \mu_i)(X_j - \mu_j)] \quad (32)$$

When the study involves more than two random variables, the covariances between all the variables are calculated (33), and the respective values are stored in the covariance matrix (Σ).

$$Cov(X) = \Sigma = \begin{pmatrix} \sigma_{11} & \dots & \sigma_{1p} \\ \sigma_{21} & \dots & \sigma_{2p} \\ \vdots & \vdots & \vdots \\ \sigma_{n1} & \dots & \sigma_{np} \end{pmatrix} \quad (33)$$

Another measure of great importance in multivariate analysis is the correlation between random variables. The use of correlation is more indicated, especially when the variables have significant differences in measurement scales. Correlation (ρ) is also a measure of the linear relationship between two quantitative variables, with values belonging to the range $[-1, 1]$, determined as follows (34):

$$Cor(X_i, X_j) = \rho_{ij} = \frac{\sigma_{ij}}{\sigma_i \sigma_j} \quad (34)$$

It is also possible to represent the values of the correlations of a set of random variables in matrix form (35):

$$Cor(X) = \mathbf{R} = \begin{pmatrix} 1 & \dots & \rho_{1p} \\ \rho_{21} & \dots & \rho_{2p} \\ \vdots & \vdots & \vdots \\ \rho_{n1} & \dots & 1 \end{pmatrix} \quad (35)$$

3.1 Factorial Analysis

When instrument measurements are correlated with each other and subjected to factor analysis, it is possible to form a new set of unobservable independent variables, called factors. The application of this technique allows the description of extensive data matrices in an exploratory way and the identification of the primary relationships of a large set of instruments with each other or the relationships between measurements and tools.

Let $X_{p \times 1}$ be a vector whose elements are the instruments installed in a dam, with a p -varied normal distribution, with a covariance matrix $Cov(X) = \Sigma_{p \times p}$ and vector of averages $\mu_{p \times 1}$. Generally, these instruments have marked discrepancies between the variances, caused by the different measurement units. In this case, performing a standardization takes all the data to the same measurement scale, through the transformation $Z = (V^{1/2})^{-1}(X - \mu)$, where $V_{p \times p}^{1/2}$ is the diagonal matrix of the standard deviations of the observable random variables. Thus, $Z_{p \times 1}$ will also have a normal distribution p -varied, vector of means $E(Z) = 0_{p \times 1}$ and covariance matrix $Cov(Z) = (V^{-1/2})^{-1}\Sigma(V^{-1/2})^{-1} = \rho_{p \times p}$. In this case, $\rho_{p \times p}$ is the correlation matrix of the original variables (observable).

An assumption for the application of factor analysis is that the measurements of the instruments (random variables) are correlated with each other. Bartlett's sphericity test makes it possible to verify the significance of correlations when the null hypothesis that $\rho_{p \times p}$ is equal to the identity matrix $I_{p \times p}$ is rejected. The statistic T of the Bartlett test [36], calculated based on the sample size n , the number of instruments p and the eigenvalues λ of the sample correlation matrix, has an approximate chi-square distribution with $\frac{1}{2}p(p-1)$ degrees of freedom, for n large enough.

$$T = -\left[n - \frac{1}{6}(2p + 11)\right] \sum_{i=1}^p \ln \lambda_i \quad (36)$$

In the orthogonal factorial model [37], the standardized random variables are linearly dependent on m unobservable random variables ($m < p$), called common factors ($F_{m \times 1}$), with factor loadings L estimated by the maximum likelihood method [38], and from other sources of variation, called specific factors ($\epsilon_{m \times 1}$). The portion of the variance of each instrument that is explained by the common factors is estimated by the commonalities h [39]. Thus, both common and specific factors have a multivariate normal distribution, with $E(F) = 0_{m \times 1}$ and $Cov(F) = I_{m \times m}$, while $E(\epsilon) = 0_{p \times 1}$ and $Cov(\epsilon) = \Psi_{p \times p}$ (Johnson and Wichern [2019]).

$$Z_{p \times 1} = L_{p \times m} F_{m \times 1} + \epsilon_{p \times 1} \quad (37)$$

$$LV(\mathbf{0}, \rho) = \frac{e^{-\frac{1}{2} \sum_{j=1}^n (z_j)' (LL' + \Psi)^{-1} (z_j)}}{(2\pi)^{np/2} |LL' + \Psi|^{n/2}} \quad (38)$$

$$h_i^2 = l_{i1}^2 + l_{i2}^2 + \dots + l_{im}^2, \quad \text{for } i = 1, 2, \dots, p \quad (39)$$

The adequacy of the fit of the orthogonal factor model can be evaluated using the Kaiser-Meyer-Olkin (KMO) criterion and Bartlett's sphericity test [36]. The KMO coefficient varies between 0 and 1. It is calculated [40] based on the sample correlations ρ_{ij} and partial correlations Q_{ij} between the pairs of observable variables X_i and X_j . The closer to unity, the greater the adequacy of the adjustment of the factor model to represent the multivariate set.

$$KMO = \frac{\sum_{i \neq j} \rho_{ij}^2}{\sum_{i \neq j} \rho_{ij}^2 + \sum_{i \neq j} Q_{ij}^2} \quad (40)$$

The maximization of the maximum likelihood function [38] is done by numerical procedures, starting from a random sample of size n from the vector Z , for a fixed value of m , which can be obtained through Horn's Parallel Analysis Method (Çokluk and Koçak [2016]).

If the factor loadings are high by more than one factory for the same instrument, the factors will not be orthogonal, as initially assumed. This makes it difficult to interpret the results. In these situations, Johnson and Wichern [2019] suggests performing the varimax rotation, an orthogonal transformation, to obtain a factor structure in which each factor has a group of variables highly correlated with its negligible correlation with the other elements.

The varimax rotation consists of finding a given matrix T , which maximizes V [41], such that $LL' = LTT'L' = \hat{L}\hat{L}'$.

$$V = \frac{1}{p} \sum_{j=1}^m \left[\sum_{i=1}^p \hat{l}_{ij}^4 - \frac{\left(\sum_{i=1}^p \hat{l}_{ij}^2 \right)^2}{p} \right] \quad (41)$$

For each of the m factors, numerical values are calculated for each sample element, called factor scores (f_j , with $j = 1, \dots, n$), which can be estimated by the method of minimums weighted squares (42), as follows (Johnson and Wichern [2019]):

$$f_j = (L' \Psi^{-1} L)^{-1} L' \Psi^{-1} z_j \quad (42)$$

3.1.1 Application – Representation of the behavior of a buttress block by a statistical index

In the study carried out by Oro et al. [2021], the data from 61 instrumentation sensors used in the monitoring of the Itaipu Dam, including piezometers, extensometer bases, plumbines, and multiple-rod extensometers, were analyzed using the factor analysis to obtain a statistical index, called the Monitoring Index of Joint Responses of the Blocks of Dams (IMCRB), which represents the global behavior of the block responses and can be used in real-time monitoring of the dam and assist in the making process of decision.

The correlations between the sensors were significant ($T = 53949.51 > T_{\alpha=5\%} = 1930.63$) and the sample was considered adequate ($KMO = 0.92$). Horn's Parallel Analysis indicated the extraction of 5 factors. The resulting factor model was able to explain approximately 85.21% of the variability observed in the dataset of the 61 sensors.

The factors were interpreted according to the sensors most correlated with them and their respective factor loadings estimated by the maximum likelihood method. In the first factor, most of the multiple-rod extensometers were considered. In the second, the sensors of the extensometer bases that measure the gap between blocks were considered. Most piezometers had higher correlations

with the third factor. The fourth factor included plumbline sensors that measure radial displacements. Only two piezometers were felt in the last element. Thus, it can be stated that for the analyzed dataset, most of the variability came from foundation creep, measured by extensometers with high factor loadings in F_1 (39.56%), and from the openings/closings of the contraction joints between blocks, measured by the extensometer bases strongly correlated with F_2 (27.68%). The time-series of these two factors (Figure 13) showed seasonal movements resulting from thermal variation with maximum annual values in the winter and minimum in the summer. In addition, F_1 and F_2 showed downward and upward trends, respectively, similarly to the sensors most correlated with each factor.

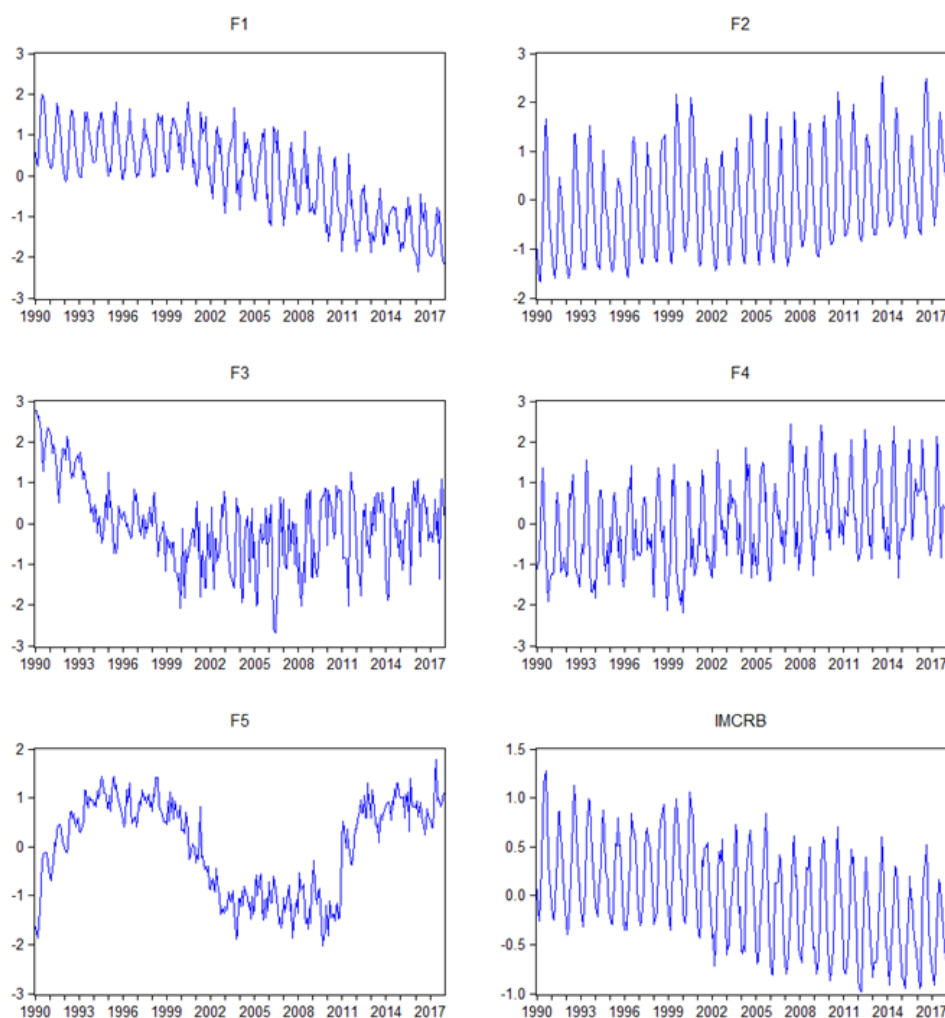


Figure 13: The factors and the statistical index IMCRB.
Source: Oro et al. [2021]

The factor, being a latent variable, cannot be measured directly. Thus, the factor values, called factor scores, were estimated based on the factor loadings and sensor data. Considering the factors as quantitative variables, which jointly represent the variability of the dam's responses to the various phenomena acting on it, these were used as parameters for generating a statistical index IMCRB. The model's coefficients (43) were estimated based on the proportion of variance explained by each factor.

$$IMCRB = 0.3956F_1 + 0.2768F_2 + 0.0858F_3 + 0.0569F_4 + 0.0369F_5 \quad (43)$$

The IMCRB time-series showed combined characteristics of the factors, with similar seasonality to factors F_1 and F_2 and a slight downward trend until 2012. Observing the instruments that had a more excellent representation according to the factorial model, it was possible to list the phenomena that most influenced the behavior of the dam blocks represented by this index, namely: the variation in the level of the reservoir, environmental thermal influences and eventual strains in the rock-mass.

3.2 Canonical Correlation Analysis

The Canonical Correlation Analysis is an interdependence analysis technique that allows researchers to identify and quantify the associations between two groups of variables (X and Y). The basic idea is to find the linear combination of variables X and the linear combination of variables Y that produces the highest correlation between the two groups (Johnson and Wichern [2019]). The first group (X) comprises p decision variables, also called explanatory variables (independents). In comparison, the second (Y) is formed by q response variables (dependents on descriptive).

Given that U and V are linear combinations of the vectors in X and Y , the canonical problem is obtaining the vectors of coefficients a and b that maximize the correlation between U and V . The vectors a and b , in this case, are solutions to a system of equations. The linear combinations U and V , in this case, are called canonical variables.

Thus, each pair of canonical variables have unit variance and maximum correlation and is not correlated with others pairs of canonical variables. The number of pairs of canonical variables that can be obtained equals the lowest value of p and q . In general, we try to get a few pairs of canonical variables that explain much of the interdependence between the two sets of observable variables.

The application of canonical correlation analysis can be performed according to the following sequence of steps:

1. Composition of groups X and Y , with variable p and q , respectively;
2. Estimation of covariance matrices for each group Σ_X and Σ_Y , and cross-covariance matrices Σ_{XY} ;
3. Obtaining the canonical variable pairs U and V , through the linear combinations $U = \underline{a}'\underline{X}$ and $V = \underline{b}'\underline{Y}$, respectively (44), so as to maximize the correlation between U and V (45);

$$\begin{aligned} (\Sigma_{XY}\Sigma_Y^{-1}\Sigma_{YX} - \lambda_k\Sigma_X)\underline{a} &= 0 \\ (\Sigma_{YX}\Sigma_X^{-1}\Sigma_{XY} - \lambda_k\Sigma_Y)\underline{b} &= 0 \end{aligned} \quad (44)$$

$$Corr(U_k, V_k) = \sqrt{\lambda_k} = \frac{\underline{a}'\Sigma_{XY}\underline{b}}{\sqrt{\underline{a}'\Sigma_X\underline{a}}\sqrt{\underline{b}'\Sigma_Y\underline{b}}} \quad (45)$$

Where λ_k is the k -th eigenvalue of the matrix $\Sigma_X^{-1}\Sigma_{XY}\Sigma_Y^{-1}\Sigma_{YX}$ or, equivalently, from the matrix $\Sigma_Y^{-1}\Sigma_{YX}\Sigma_X^{-1}\Sigma_{XY}$;

Similarly, it is possible to perform canonical correlation analysis based on sample correlation matrices.

4. Estimation of the correlations of the canonical variables with the original variables;

5. Calculation of the proportion of the total variation that is explained by each canonical variable;
6. Interpretation of correlations and canonical pairs.

3.2.1 Application – Study of the relationship between two environmental variables with the displacements of the structures of a concrete dam

Two sets of data were subjected to canonical correlation analysis in the study carried out by Oro et al. [2016]. The first was composed of 38 extensometer bases, plumbines, and multiple-rod extensometers to measure different types of displacements of two critical blocks of the Itaipu dam. The second set was formed by four indicators of environmental conditions (temperature on the surface of the concrete downstream) upstream, the temperature of the air, and the water level of the dam reservoir).

Initially, the data matrix composed of the monthly averages (rows) of the values recorded between January 1990 and December 2013 for each of the 42 sensors (columns) considered in the study was organized. The values were standardized as these are measurements on different scales of measures and magnitudes.

All the eigenvalues were considered significant (Table 8), but the authors chose to detail the results only for the first pair of canonical variables, as it was the one with the highest correlation.

Table 8: Canonical correlation of displacement sensors vs environmental conditions.

<i>Eigenvalues</i>	<i>Canonical Correlation</i>	<i>Wilks Lambda</i>	<i>Chi-square</i>	<i>Degrees of Freedom</i>	<i>P-Value</i>
$\lambda_1 = 0.973$	0.986	0.003	1537.08	152	0
$\lambda_2 = 0.717$	0.847	0.114	577.122	111	0
$\lambda_3 = 0.437$	0.661	0.402	242.062	72	0
$\lambda_4 = 0.286$	0.535	0.714	89.492	35	0

Source: Oro et al. [2016]

The canonical variable U_1 (46) explained 38.7% of the variance observed in the displacements, while the proportion of the variance is explained by V_1 (47) for the environmental conditions group was 65.6%. Thus, the groups of *Displacements* and *Environmental Conditions* sensors were well represented by the first pair of canonical variables since the correlation between these groups was 0.986. In contrast, the other teams had lower values.

$$\begin{aligned}
 U_1 = & -0.04Z_1 - 0.09Z_2 - 0.24Z_3 + 0.02Z_4 - 0.04Z_5 - 0.05Z_6 - 0.24Z_7 + \\
 & +0.05Z_8 + 0.12Z_9 - 0.08Z_{10} - 0.04Z_{11} - 0.03Z_{12} + 0.26Z_{13} - 0.07Z_{14} - \\
 & -0.28Z_{15} + 0.11Z_{16} - 0.16Z_{17} - 0.07Z_{18} + 0.08Z_{21} - 0.07Z_{22} + 0.04Z_{23} - \\
 & -0.36Z_{24} + 0.06Z_{25} + 0.10Z_{26} + 0.02Z_{27} - 0.01Z_{28} - 0.04Z_{29} - 0.01Z_{30} - \\
 & -0.02Z_{31} - 0.02Z_{32} + 0.17Z_{33} - 0.28Z_{34} - 0.11Z_{35} + 0.04Z_{36} - 0.12Z_{37} - \\
 & -0.52Z_{38} + 0.06Z_{39} + 0.43Z_{40}
 \end{aligned} \tag{46}$$

$$V_1 = 0.20Z_{41} - 0.02Z_{42} + 0.12Z_{19} + 0.72Z_{20} \tag{47}$$

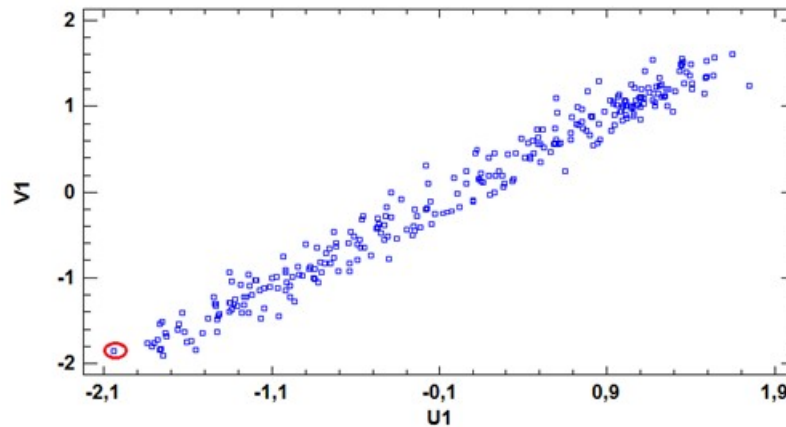


Figure 14: Scatterplot of the canonical variables U_1 and V_1 .
Source: Oro et al. [2016]

Estimates of correlations between sensors that measure displacements and environmental indicators were also obtained. The Table 9 shows the most intense relationships ($\rho > |0.8|$), with a predominance of multiple rod strain gauges (Z_{23} , Z_{24} , Z_{35} , Z_{38} , Z_{40}) and three pendulum sensors (Z_2 , Z_3 and Z_5) strongly related to temperatures (Z_{19} , Z_{20} and Z_{41}).

Table 9: Strongly correlated variables.

<i>Sensors</i>	<i>Correlations</i>
$Z_{35} - Z_{20}$	-0.913
$Z_{24} - Z_{20}$	-0.910
$Z_{24} - Z_{41}$	-0.889
$Z_{23} - Z_{20}$	-0.888
$Z_{23} - Z_{41}$	-0.872
$Z_5 - Z_{41}$	-0.855
$Z_{40} - Z_{19}$	-0.855
$Z_{40} - Z_{41}$	-0.843
$Z_3 - Z_{41}$	-0.829
$Z_2 - Z_{20}$	-0.820
$Z_{35} - Z_{41}$	-0.815
$Z_{23} - Z_{19}$	-0.814
$Z_{24} - Z_{19}$	-0.812
$Z_{38} - Z_{19}$	-0.811

Source: Oro et al. [2016]

In the region where the Itaipu dam is located, the difference between average monthly temperatures can reach 20°C. In addition, the largest displacements, even negligible, occur during periods of low temperature. This justifies the solid negative correlations of the sensors with the confirmed temperature variables. However, the water level of the dam reservoir showed a slight positive correlation individually with only a few sensors. This is probably due to the history of slight variation in the analyzed period, which led to practically constant forces acting on the dam. Another reason may be the need to interact the reservoir water level with the temperature to influence the

displacements.

3.3 Cluster Analysis and Discriminant Analysis

Cluster Analysis seeks to find a small number of homogeneous groups within a heterogeneous data set whose variation within the group is substantially smaller than the total variability of the data set.

First, the researcher must select the methods he will use to calculate the distances between the groups and compare observations on the variables used to characterize the elements to be grouped. The similarity is a measure of proximity between two groups. One way of calculating this measure is the Mahalanobis distance. A widely used comparison method is the agglomerative hierarchical method, which initially considers each observation as a separate group. At each step of the process, groups are joined according to similarities, forming new clusters until only one set remains with the total number of observations. The graphic representation of these steps is known as a dendrogram, characterized by a hierarchical tree structure, generated sequentially in forming the groups.

Discriminant Analysis is a technique that enables, starting from an independent variable, to study the profile, performing classification and differentiation of two or more group elements. The number of groups should be known in advance. The discrimination is based on a mathematical rule, which minimizes the likelihood of incorrect classification errors.

3.3.1 Application – Clustering of monitoring data of a concrete dam by similarity in the environmental conditions

The study carried out by Oro et al. [2016] involved 288 observations collected by 42 sensors used to monitor a block of the Itaipu dam; first, the five factors identified by the application of Factor Analysis were used as input variables for the cluster analysis.

The hierarchical clustering of the 288 observations was performed using the mean linkage method and the Mahalanobis distance. Three homogeneous groups of dates were formed (Figure 15), aggregating 190 elements in the first (G1), 83 in the second (G2) and 15 in the third (G3).

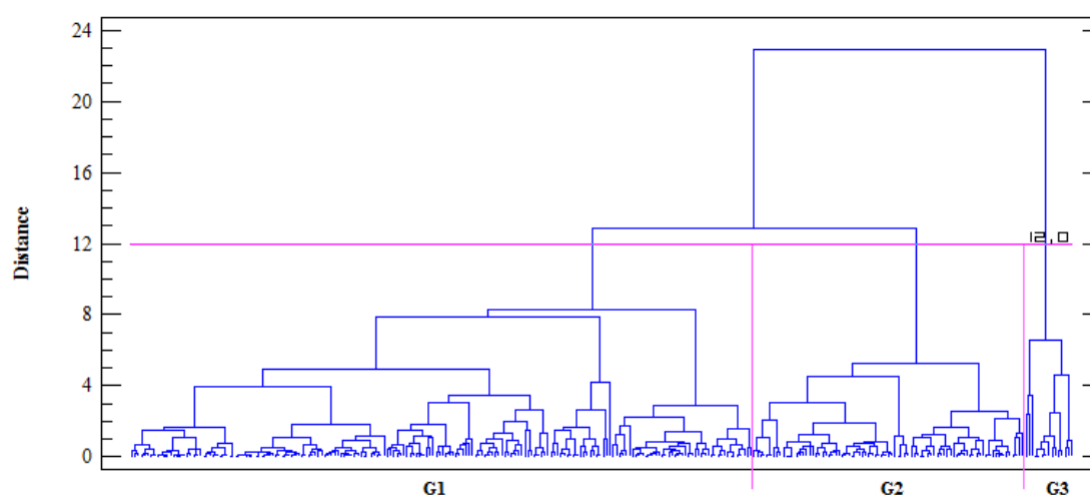


Figure 15: Clustering by the mean linkage method, using the Mahalanobis distance.
Source: Oro et al. [2021]

The first group consisted essentially of observations obtained from November to May, a period characterized by the highest air temperatures. The second was primarily composed of months with

lower air temperatures. The third group gathered the dates on which the lowest water levels in the reservoir were recorded.

The classification of the 288 observations was submitted to discriminant analysis, using as discriminant variables the sensors that presented the highest commonalities in the factor analysis (Z_7 , Z_9 , Z_{33} , Z_{34} , Z_{35} and Z_{42}). Due to the significant difference in the size of the groups, the proportionality of the number of observations per group was considered. The high percentage of correct classification, 94.1%, confirmed the discriminating power of these sensors.

Two functions were considered statistically significant, at the confidence level of 99%, to distinguish the observations belonging to each group (Figure 16). The first function discriminates the "High Temperature" and "Low Temperature" groups, G1 and G2, while the second function discriminates the "Low Water Level of Reservoir" (LWLR) group from the others, G3.

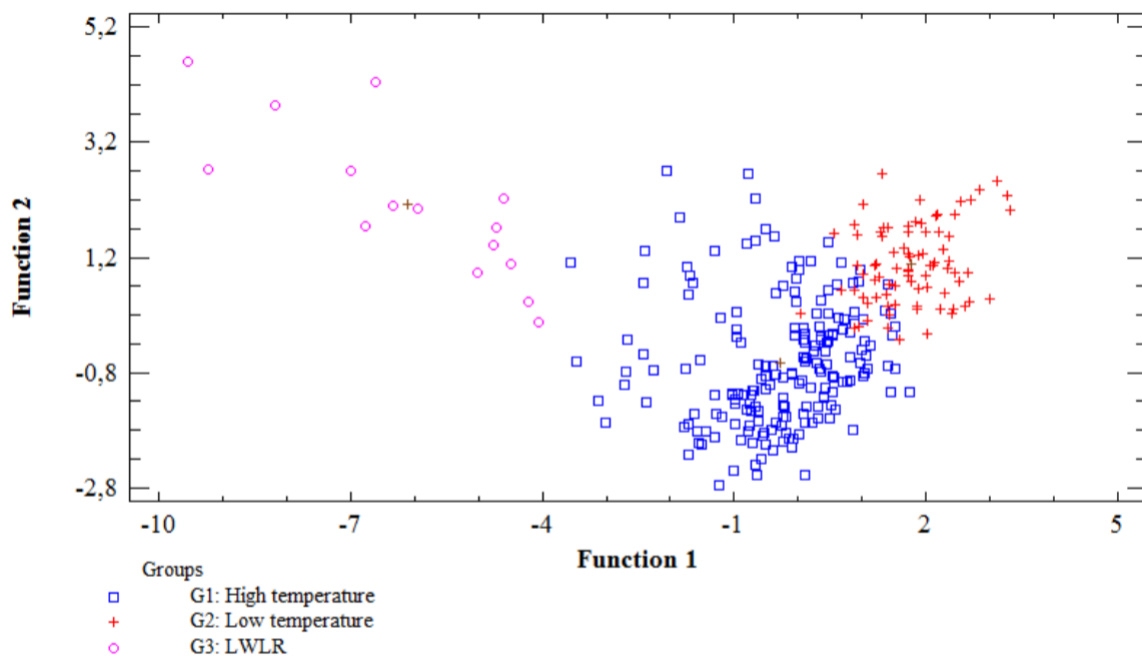


Figure 16: Dispersion of the elements according to discriminant functions between groups.
Source: Oro et al. [2021]

The classification functions of the observations in the groups (48) resulted from the linear combinations of the sensors. To classify new observations, it is sufficient to calculate the scores of each new element in each group and then allocate it to the one with the highest score.

$$\begin{aligned}
 G_1 &= -0.66 + 0.66Z_{34} - 1.33Z_{33} - 0.08Z_{35} + 0.32Z_9 + 0.13Z_{42} - 0.88Z_7 \\
 G_2 &= -3.40 - 0.45Z_{34} + 2.87Z_{33} + 0.07Z_{35} - 1.284Z_9 + 1.656Z_{42} + 2.674Z_7 \\
 G_3 &= -23.81 + 5.94Z_{34} + 1.01Z_{33} + 0.54Z_{35} + 3.01Z_9 - 10.87Z_{42} - 3.64Z_7 \quad (48)
 \end{aligned}$$

3.4 Principal Components Analysis

When the number of variables to be analyzed increases, the parameter that evaluates the average number of samples required to detect changes in the process deteriorates; in this way, if suspected

that the process variability is not equally distributed among all variables, it is helpful to use other methods (Montgomery [2013]).

Techniques reduction data are based on creating sets of latent variables that capture the data's significant variation hidden. The change that the sets of latent variables extract from the process variables is of fundamental importance for evaluating product quality, process security, and, more generally, if the process is in statistical control (Kruger and Xie [2012]).

3.4.1 Application – Principal components multivariate control chart

In the study of Lazzarotto et al. [2016], considering that:

- a high number of instruments that assess various quality characteristics of interest that have an inherent variability monitors hydroelectric plants;
- the readings of these instruments generate time series of data on many occasions correlate;
- each project of a dam plant has characteristics that make it unique and faced with the need to establish statistical control limits for the instrumentation data;

this article approaches multivariate statistical analysis and proposes a model that uses principal components control charts and statistical T^2 and Q to explain the variability and establish a method of monitoring to control future observations.

Because of the the economic, social, and environmental importance of plants, the structural behavior of hydroelectric dams requires that dams are monitored using tools for assessing the safety and ensuring the operation of power generation. Usually, a large dam has hundreds or even thousands of instruments that measure various quality characteristics of the dam. Evaluating the results of measurement instrumentation of a large number of instruments individually may be impractical or even lead to high false alarms and discredit the monitoring system.

In the dam instrumentation, control values are those closer to the physical reality, considering the thermal environmental influences that act on structures and more realistic rheological models of concrete and foundation. Control values are used to monitor the performance of the systems using readings taken and are warning signs for abnormal situations of structural behavior. Values are determined earlier in the design phase, and the reservoir filling is not, in many cases, more applicable during the operation phase Itaipu [1999]. It is necessary to establish operational control values for the instrumentation of hydroelectric plants.

This article aims to propose a multivariate statistical model for monitoring instruments for monitoring dams via control charts and principal components analysis and seeks to separate the effect of environmental variables on the reading of instruments from other sources of variability by use of statistics and and establish control values for monitoring future observations. The method is evaluated in a case study applied accurate data from monitoring a hydroelectric dam plant.

The field of multivariate analysis consists of statistical techniques that consider two or more random variables related to a single entity to produce an overall result that considers the relationship between the variables (Jackson [1991]). Multivariate process control is a methodology based on control charts used to monitor the stability of a multivariate process. Stability is achieved when one or more parameters of interest remain stable on samples (Mason and Young [2002]).

The application of univariate control charts can lead to erroneous and misleading interpretations, and those multivariate methods are a good alternative (Montgomery [2013]; Johnson and Wichern

[2019]); if the variables are correlated increases the probability of emission of false alarms and not receiving an alert when the multivariate process is out of control (Ryan [2011]).

A method that can extract features in the data can be helpful in dam safety studies. As the instrumentation readings are a result of the combination of several factors, methods of multivariate data analysis can provide the following advantages: 1) more profitable by reducing the number of individual analyses, 2) more remarkable ability to explain and separate the variability due to one because of random variability attributable since the random variabilities are, by definition, uncorrelated from one instrument to another and 3) to identify patterns of behavior (Nedushan [2002]).

The PCA is a multivariate data analytical technique in which several related variables are transformed into a set of uncorrelated variables that are linear combinations of the original variables, which is expected to explain the variability between variables with a smaller number of variables (JACKSON, 1991). Its industrial application has contributed to the multivariate statistical process control since only a few multivariate control charts can serve as an index of process performance. PCA improves the early detection of failures in the univariate graphs (Hourti [2005]).

The Q statistic, also known as squared prediction error, represents the changes that the principal component model does not explain. It measures the difference between the sample mean and its projection on the primary components model (Mujica et al. [2011]).

This work was developed with accurate data from the instrumentation section E of the Itaipu hydroelectric plant. Faced with various instruments present, the piezometers were chosen because of the amount and the importance of their role in measuring uplift pressures in the dam. The E section has seven piezometers.

For phase, I, the adjustment test (retrospective) of the model, were selected 300 readings, and for Phase II validation (perspective) of the model, we selected 19 readings. Variables in very different scales and domain variations should be standardized (Johnson and Wichern [2019]). This procedure is because the original variables can have scales and domains quite distinct, giving a false interpretation of its absolute magnitude variability, which is avoided by standardizing (Montgomery [2013]).

The main steps of the method consist of:

- Gathering and data standardization
- Testing the multivariate normality
- Choosing the Type I error probability and the estimate of the covariance matrix
- Construct the graphic T^2 for the full set of variables
- Extraction and selection of the number of principal components and normality test
- Construction of the graph T^2 and the ellipse control of the principal components in phases I and II
- Calculation of residual and Q statistics
- Construction of Q -chart
- Interpretation of results.

For each vector, $X' = (X_1, X_2, \dots, X_7)$ containing the standard readings of piezometers was calculated the quantity

$$T^2 = (X - \bar{X})' S^{-1} (X - \bar{X})$$

where \bar{X} is the sample mean vector, S^{-1} the covariance matrix, and X' represents the transpose matrix of X . This value was compared with the value corresponding to the 95% percentile of the distribution χ^2 with seven degrees of freedom which is the reference distribution for this case, to conclude that the data set can be accepted following a multivariate normal distribution.

Once we were able to accept the hypothesis that the data come from a multivariate normal distribution with a 95% confidence level, for comparison, univariate control Shewhart plots of the sample mean \bar{x} were constructed for phase I generating an unacceptable number of observations outside the control limits for respective univariate Shewhart plots from instruments P_1 to P_7 for phase I data. Note the impossibility of univariate treatment with the high number of observations outside the control limits.

For this study, the usual covariance matrix of equation (49) was selected because it is less sensitive to minor deviations from the mean.

$$S = \frac{1}{m-1} \sum_{i=1}^m (X_i - \bar{X})(X_i - \bar{X})' \quad (49)$$

Extracting the principal components of the data set of phase I, a simulation was performed by selecting $k = 2, \dots, 6$ main features. Data from scores of k main components were retained, and multivariate normality was tested by the test χ_k^2 at a confidence level of 95% that obtained the acceptance of the hypothesis normality,

The control chart of statistical T_4^2 including the data of the two phases is shown in Figure 17. Figure 18 displays the control of the ellipse 95% confidence data T_4^2 for the first two principal components; the point in red represents the single observation out of control, which is outside the ellipse control; we take the components 2 and 3. These plots were constructed for $\alpha = 0.0027$ and the usual covariance matrix at a 95% confidence level. Finally, the Figure 19 shows the behavior of the statistic Q in the period of analysis, for $k = 4$.

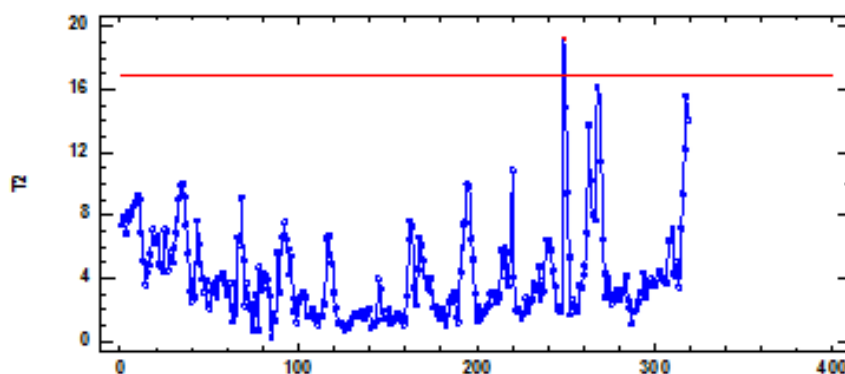


Figure 17: T^2 chart for 4 principal components retained in phases I and II.

Source: Lazzarotto et al. [2016]

This article established a method for building control charts for dam monitoring instruments. Considering the large number of devices that monitored a large dam and the correlation between

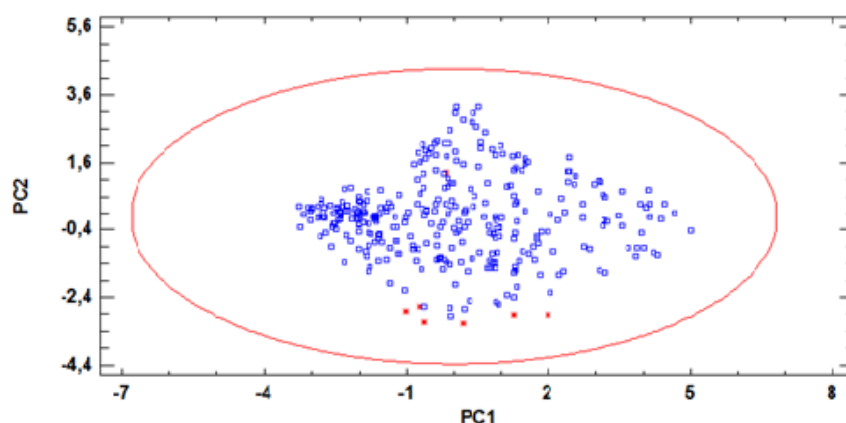


Figure 18: Ellipse control scores the first 2 components of 4 PC's retained.
Source: Lazzarotto et al. [2016]

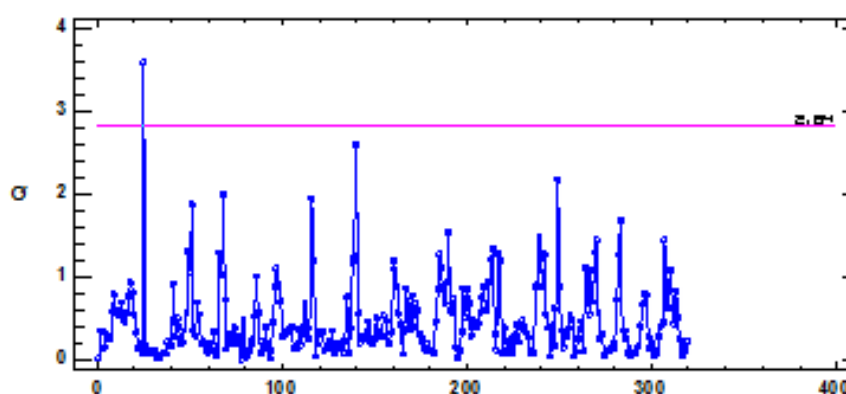


Figure 19: Values of the statistic Q for $k = 4$ principal components in the phases I and II.
Source: Lazzarotto et al. [2016]

them, the individual monitoring of each instrument is not feasible, either because of the high number of individual control charts that would be generated or because of a large number of false alarms could discredit the system.

The results showed that the principal components model combined with the Q statistic was able to fit the phase better I data when at least four main components are taken because, in this case, the observations indicated as out of control for the T^2 graph appear as out of control either for the T^2 graph of the k principal components or for the graph of Q and was able to show that, in the applied case study, the multivariate monitoring system of the piezometers, located in section E of the Itaipu hydroelectric plant, was under statistical control at the 95% confidence level.

The adopted model enabled an interpretation of the principal components as a consequent variability of environmental factors (inherent) to the model. It is essential to understand that the variability of the main components originated from these factors and also can be controlled by the aleatory variability by Q statistic.

4 Finite Element Method (FEM)

According to Silva [2003], and Maken et al. [2013], dams are subject to changes in their internal temperature due to changes in external temperature and heat of hydration, which is a source of heat from chemical reactions of the concrete during the construction phase of the dam, but which dissipates after a specific time. External thermal loads to a surface come from solar radiation and air, foundation, and reservoir temperature variations. Such loads can cause significant deformations and volumetric changes in the structure, which often exceeds the tensile strength of concrete, leading to the development of cracks.

The study of temperature behavior in a concrete dam is linked to the heat transfer processes to which it is subject (conduction, convection, and radiation). And the mathematical models that describe these phenomena can present the following characteristics: they are non-linear; algebraic, partial, and integral differentials; they constitute a coupled system; the properties of the materials involved are generally a function of temperature, and the solution region usually has a complex geometry. (Ozisk [1993], Shukla [2005] and Incropera et al. [2011])

Such characteristics make analytical methods very limited and exact solutions unfeasible, so numerical methods are the most used means of resolution to obtain answers to complex problems of genuine interest.

The finite element method consists of a numerical process to obtain an approximate solution for a mathematical model that describes a physical phenomenon typically. This technique is based on decomposing the solution domain into simple regions, the so-called elements, for which approximate solutions are systematically constructed from variational methods or weighted residuals. The resolution of the problem is obtained from the reconstruction of the domain with the individual keys of each element, guaranteeing the continuity in the borders between the components (Zienkiewicz and Taylor [2005]; Chen and Liu [2018]).

The steps to achieve an approximate solution to a physical model using FEM (according to Zienkiewicz and Taylor [2005]; Chen and Liu [2018] and Lewis et al. [2004]) are:

1. Discretization of continuous: dividing the solution region into subdomains or non-overlapping element (This step can be represented by Figure 20);
2. Selection of basis functions or interpolation: such functions determine the variation of the unknown variable in the study;
3. Formulation of the equations of nodes: take into account the individual properties of each element;
4. Assembling the global matrix for the simultaneous solution of the elements equations;
5. Application of the initial and boundary conditions and loadings;
6. Resolution of the global system for determining the nodal results, such as displacements and temperatures on different nodes.

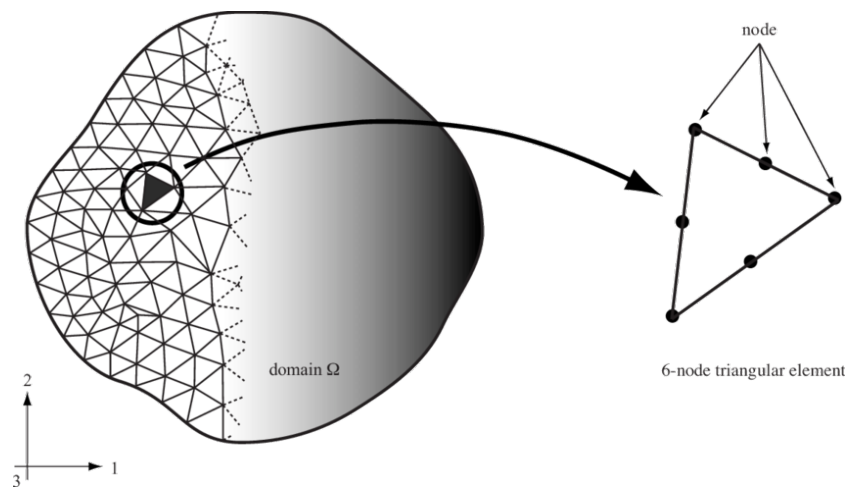


Figure 20: Discretization of continuous.
Source: Bobet et al. [2009]

In the interest of understanding the basic principle of this method, it will be approached in a problem of heat transfer by conduction in three-dimensional geometry.

4.1 Heat Conduction Equation

Conduction heat transfer aims to determine the temperature field of a medium resulting from conditions imposed on its borders. The classical equation for heat diffusion is determined by the first law of thermodynamics and by the Fourier law. It is an empirical relationship involving the heat flow and the temperature gradient through a proportionality constant called the thermal conductivity of the properties of materials.

According to Ozisik [1993], the heat conduction equation for an anisotropic material is given by (50):

$$\frac{\partial}{\partial x} \left(k_x(T) \frac{\partial T}{\partial x} \right) + \frac{\partial}{\partial y} \left(k_y(T) \frac{\partial T}{\partial y} \right) + \frac{\partial}{\partial z} \left(k_z(T) \frac{\partial T}{\partial z} \right) + G = \rho c \frac{\partial T}{\partial t}, \quad (50)$$

where $k_x(T)$, $k_y(T)$ e $k_z(T)$ are the thermal conductivities (W/mK) dependent on the temperature in the directions x , y e z , respectively; ρ is the specific mass (kg/m^3); c is the specific heat ($J/kg K$) of the material e G is the rate of internal heat generation (internal energy) per unit volume (W/m^3).

Briefly, the Equation (50) states that at any point in the medium, the rate of energy transfer by conduction for a unit volume plus the volumetric rate of thermal energy generation must be equal to the rate of change of thermal energy stored inside of the volume.

For an isotropic material in which the thermal coefficients k , ρ and c are all constant, and there is no internal heat generation and Equation (50) can be rewritten as Equation (51):

$$\nabla^2 T = \frac{1}{\alpha} \frac{\partial T}{\partial t}, \quad (51)$$

where ∇^2 is the Laplacian operator, and $\alpha = \frac{k}{\rho c}$, is called thermal diffusivity m^2/s (see e.g. Lewis et al. [2004]).

Yet, according to Lewis et al. [2004], the solution of the Equation [50] is possible if the initial and boundary conditions defined are adequate to the physical problem analyzed. The initial condition consists of a temperature function T to which the structure is subjected, at a certain instant t_0 . And the boundary conditions, which can be of two types: Dirichlet or Neumann, describe, respectively, the temperature ($T|_{\Gamma_1} = T_1$) and the heat flux to which the surface of the structure is subjected ($\frac{\partial T}{\partial n}|_{\Gamma_2} = q$), as can be seen in the Figure 21.

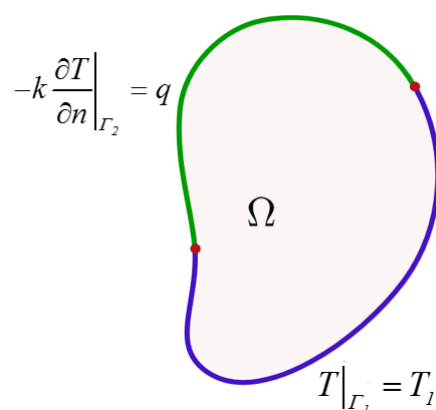


Figure 21: Boundary conditions.

Source: Adapted from Alexandrov [2007]

4.2 Galerkin Method

An expression that describes how temperatures are interpolated in a finite element is given by the Equation (52):

$$T(x, y, z, t) = \sum_{i=1}^n N_i(x, y, z) T_i(t) \quad (52)$$

where N_i are the basis functions, n is the number of nodes in an element and $T_i(t)$ is a nodal temperature time-dependent.

The most widely used basic functions are polynomial type, in particular linear and quadratic, according to the size of the mesh elements. Examples of such functions and applications can be viewed in Lewis et al. [2004] and Oñate [2013].

Multiplying the Equation (50) by a base function and integrating over the domain of analysis, the representation of Galerkin is:

$$\int_{\Omega} N_i \left[\frac{\partial}{\partial x} \left(k_x(T) \frac{\partial T}{\partial x} \right) + \frac{\partial}{\partial y} \left(k_y(T) \frac{\partial T}{\partial y} \right) + \frac{\partial}{\partial z} \left(k_z(T) \frac{\partial T}{\partial z} \right) + G - \rho c \frac{\partial T}{\partial t} \right] d\Omega = 0.$$

Using integration by parts and the initial and boundary conditions mentioned above (21), the Finite Element Method (Galerkin modality) can be written from the matrix Equation (53):

$$[C_{ij}] \left\{ \frac{\partial T_j}{\partial t} \right\} + [K_{ij}] \{T_j\} = \{f_i\}, \quad (53)$$

where:

$$[C_{ij}] = \int_{\Omega} \rho c N_i N_j d\Omega;$$

$$[K_{ij}] = \int_{\Omega} \left[k_x(T) \frac{\partial N_i}{\partial x} \frac{\partial N_j}{\partial x} + k_y(T) \frac{\partial N_i}{\partial y} \frac{\partial N_j}{\partial y} + k_z(T) \frac{\partial N_i}{\partial z} \frac{\partial N_j}{\partial z} \right] d\Omega$$

$$\{f_i\} = \int_{\Omega} N_i G d\Omega - \int_{\Gamma_2} q N_i d\Gamma_2$$

4.3 Application of FEM and ARIMA model for temperature forecast in concrete dam

The main objective of this section is to describe an application of the FEM and the time series forecasting method, called ARIMA, to describe predicted temperature fields. The methodology adopted and the results obtained are described in this section and were published in work: Hickmann et al. [2016].

In chapter 2 some forecasting methods for univariate time series and in this chapter (in the sections above), the finite element method applied to a heat conduction equation in a three-dimensional domain were presented. It is possible to combine these tools to calculate predicted temperature fields.

Since a concrete structure, such as the dams of a hydroelectric power plant, is exposed to the effects of temperature, the future thermal behavior of the dam can help conduct the study of other variables to which the dam is subject, such as deformation and displacement, which are strongly influenced by the dam body temperature. Thus, knowing the future structural response of the dam may be possible, and this will collaborate in the monitoring and predictive actions of the dam.

The sequence describes a methodology that can be used to determine predicted temperature fields.

4.3.1 Methodology

The flowchart (Figure 22) presents a practical scheme that can be adopted to obtain a forecast of the temperature fields, using the ARIMA method (Section 2.1) and the FEM (sections above):

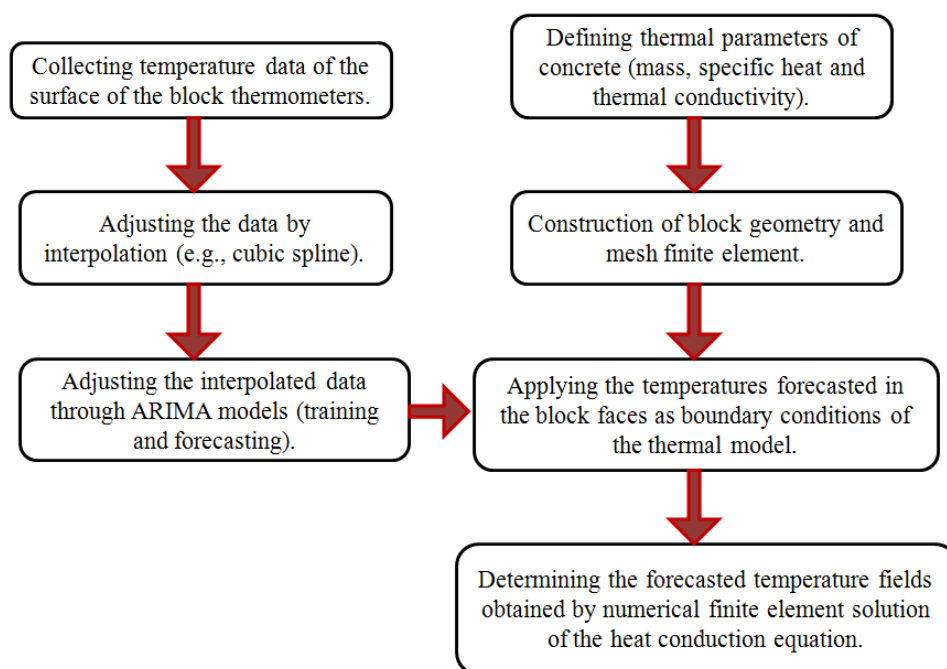


Figure 22: Flowchart of the proposed method.

Source: Hickmann et al. [2016]

The flowchart above displays two initial parts that are independent. One aspect involves the temperature prediction process and the other deals with the concrete block's heat conduction model.

Initially, data is collected, obtained from reading thermometers in the field (surface and internal to the structure), and stored in a database. The time series of surface thermometers provide the boundary conditions for the thermal model, and to obtain the predictions of these series, the data must be periodic. To determine the forecasts, the data were adjusted to monthly periodicity. After the prophecy was made, they were changed to become daily so that the FEM, applied computationally through the commercial software Ansys, is practicable. A simple math technique for completing data is cubic spline interpolation.

The next step consists of forecasting the temperature from the data series of surface thermometers, now with daily periodicity. The thermometer series can be predicted using several techniques, as mentioned in the previous sections of chapter 2. But in this case, only the ARIMA model was used. These predicted data will be used as boundary conditions (in this case, Dirichlet) to compose the concrete block's thermal model together with the heat conduction equation.

The FEM will solve the thermal model (heat equation and initial and boundary conditions). Therefore, it is necessary that: the thermal constants (thermal conductivity, specific mass, and specific heat) are defined for the type of concrete that the structure has; the geometry that represents the block must be the most faithful to the actual geometric model, and the finite element mesh for the geometry is adequate.

Since the boundary conditions are predicted temperature data, the temperature fields, obtained by numerical resolution by FEM of the thermal model, will be expected. This complete process generates indicated temperature fields for the concrete block under analysis. An internal point, where an internal thermometer is located, was monitored, and the MAPE was calculated to assess the methodology's applicability.

4.3.2 Case Study and Results

The methodology presented in the above section was used in a case study of a dam block at the Itaipu Hydroelectric Power Plant. A concrete block (called E-6) was selected for analysis, and temperature data were taken from four surface thermometers (TS-D-903, TS-D-904, TS-E-1, and TS-D-5). Figure 23 shows the block's two-dimensional geometry and the thermometers' location.

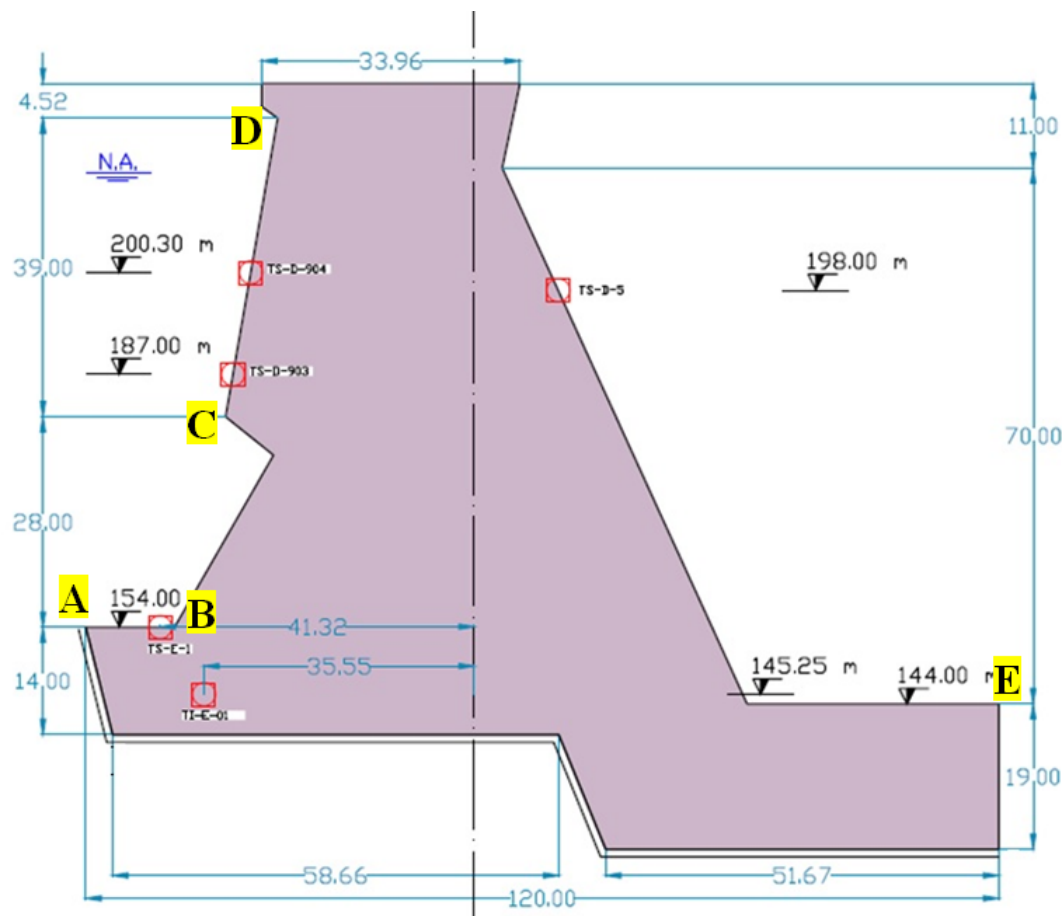


Figure 23: Geometry buttress block, location of thermometers and delimitation of boundary conditions.

Source: Hickmann et al. [2016]

As seen in Figure 23, from point A to point B (clockwise), the adopted temperature function was the thermometer TS E-1, from point B to C of TS-D-903 thermometer, point C to D TS-D-904 thermometer from the point D to E TS-D-5 thermometer and from point D to point E; the face was considered that it had zero heat flow, $q = 0$.

The thermometer data was split into a training in-sample (from 11/01/2010 to 11/11/2013) and a forecasting out-of-sample (09/12/2013 to 27/11/2014). Figure 24 graphically shows the data adjusted by the spline versus the data obtained from the forecast by the ARIMA model. The MAPE-in is an error within the sample, and de MAPE-out is the error out-of-sample:

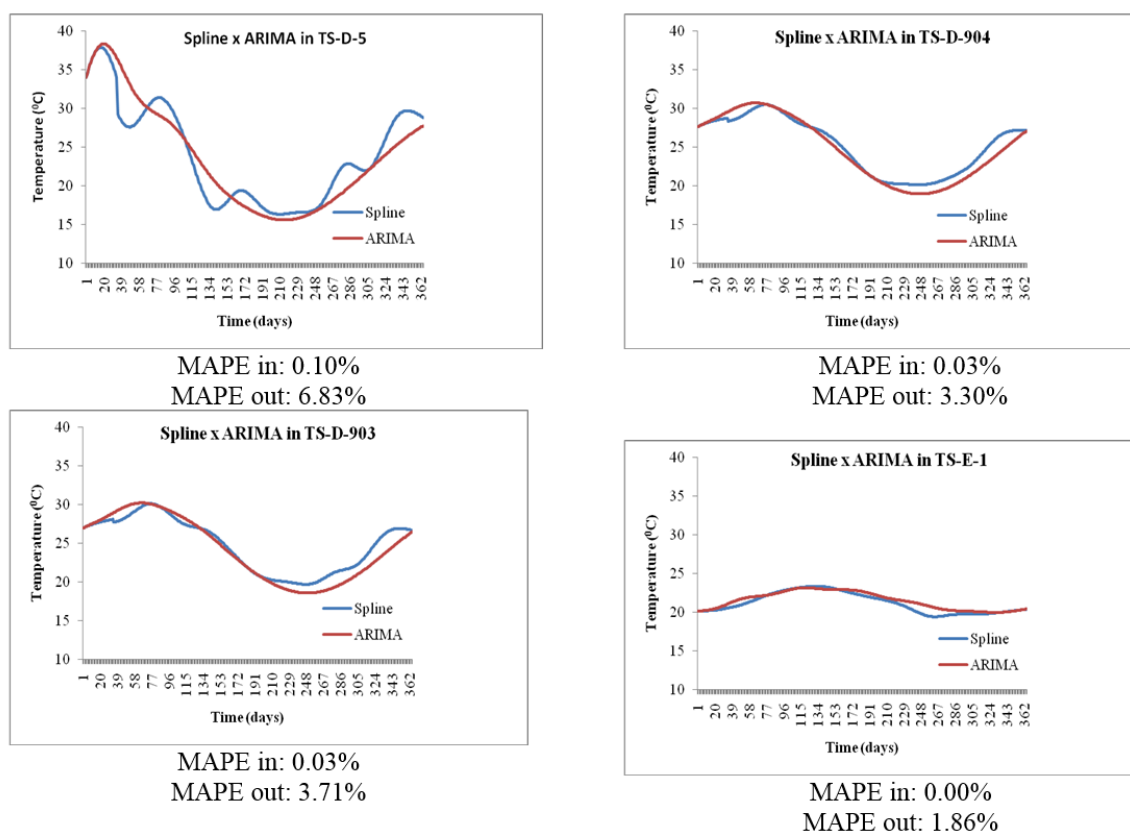


Figure 24: Cubic splines vs ARIMA for the out-of-sample period for each of the surface thermometers (TS-D-5, D-TS-903, TS-904 and TS-D-E-1). Source: Hickmann et al. [2016]

The numerical-computational solution of the thermal model was obtained from some simplifications adopted: no source of heat, the material homogeneity, isotropic and linear elastic behavior, considered the state of plane stress and simplifications in geometric section (chamfers, drainage, supports, and other beams). And the thermal constants, obtained in the block design, were: a thermal conductivity of $1.2W(mK)^{-1}$, a specific heat of $895.38J(kg K)^{-1}$ and a specific mass of $2550kgm^{-3}$.

The simulations were performed in ANSYS – Workbench V16.1 environment, commercial software with the license acquired by CEASB-FPTI, widely used in static and dynamic structural simulations. The mesh had 1386 elements (quadrangular and some triangular). A stationary simulation with the temperature data approximate temperature on 25/11/2013 was performed with the aim being to block an equilibrium temperature range, and this was used as the initial temperature of the block for the transient analysis, following temperature data set of 4 thermometers obtained through the ARIMA model in the period from 26/11/2013 to 31/10/2014 served as boundary conditions to the thermal model. The Figure 25 provides trimonthly temperature fields during this period, considering the last day of the month.

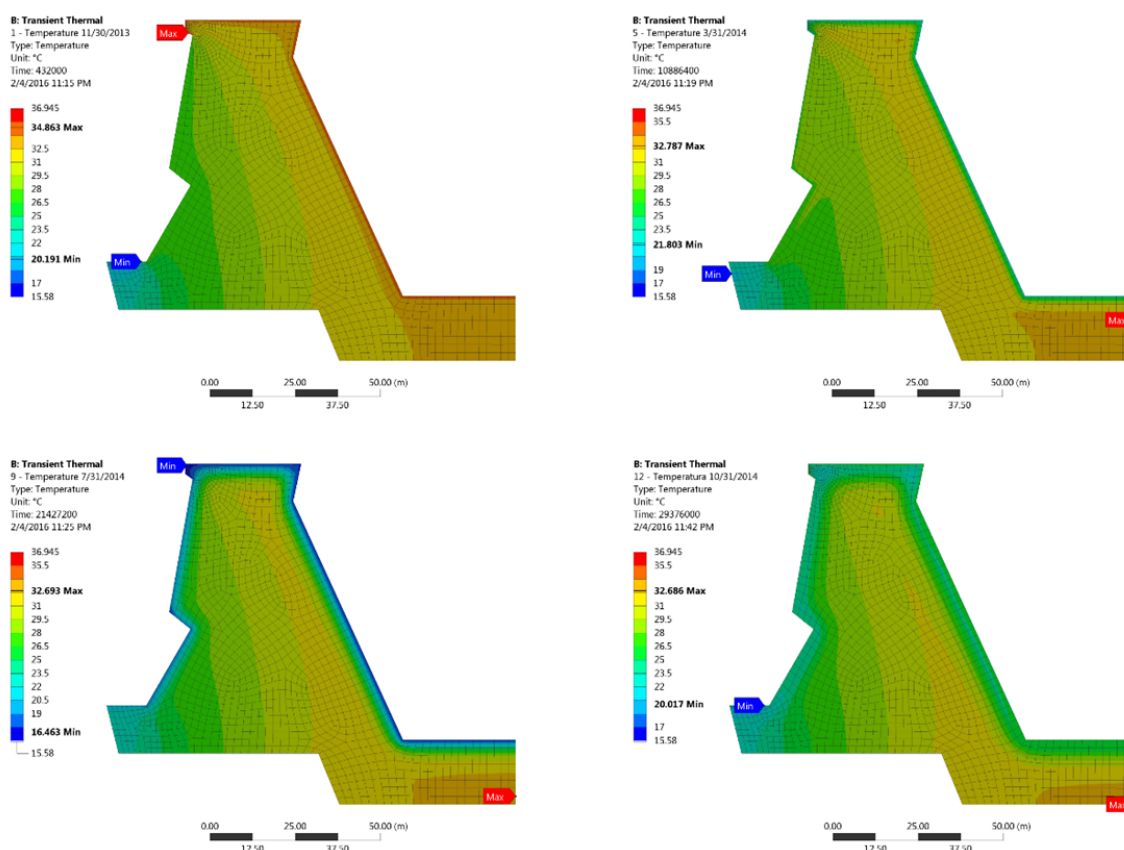


Figure 25: Trimonthly transient temperature fields.
Source: Hickmann et al. [2016]

The color scale at the left-hand side of each of the four diagrams in Figure 25 shows the various temperatures (in Celsius) in increasing order from minimum values at the bottom (dark blue) to maximum values at the top (red). The highest temperatures are in the downstream phase (contact with the environment), while warmer temperatures are located upstream (communication with the reservoir).

The simulation results showed that heat storage in the dam of the soul throughout the year, since the color change in the graph, is closest to the surface. As expected on the type of material, the energy dissipation is gradually slow. Moreover, it has monitored a control point within the structure where the internal thermometer TI-E-01 is located (see Figure 23). The data measured by this thermometer for the out-of-sample forecasting year were compared with those obtained by numerical simulation (the FEM), and the calculated MAPE was 15.37%. With this, the thermal model of the block is within the engineering safety limits, according to the superintendence of the Itaipu Power Plant Engineering.

5 Final Remarks

This chapter had as its primary objective to present some numerical and probabilistic methods with applications aimed at the dam of the Itaipu Hydroelectric Power Plant. The range of applications is certainly not restricted to the above and continues to be the subject of study by the researchers who wrote this chapter.

The instrumentation data are updated daily, either automatically or through manual readings, and therefore all models applied in the mentioned works can be updated and new results obtained. In addition, the methods discussed in this chapter can be used in other dams or concrete structures, respecting their particularities and with the necessary adaptations or simplifications.

The specific knowledge of engineers combined with the result of mathematical methods provides excellent foundations for the safety of any structure, a topic of high relevance, especially in dams, which have a large size and whose economic, environmental, and social damages are enormous in the event of a catastrophe.

Acknowledgements

The authors are grateful for the support of the Itaipu Technological Park – Center for Advanced Studies in Dam Safety – Itaipu Binacional Dam Engineering Area during their original doctoral research carried out in the Postgraduate Program of Numerical Methods in Engineering at the Federal University of Paraná, carried out in partnership between the Itaipu Technological Park and higher education institutions from 2012 to 2016 and the research that continued from 2016 to the present day.

References

- R. Adhikari and R. K. Agrawal. *An introductory study on time series modelling and forecasting*. Lambert Academic Publishing, 2013.
- O. Alexandrov. Mixed boundary conditions. https://commons.wikimedia.org/wiki/File:Mixed_boundary_conditions.svg, 2007. Accessed: 2022-04-10.
- J. M. Bates and C. W. Granger. The combination of forecasts. *Journal of the Operational Research Society*, 20(4):451–468, 1969.
- M. Beltrami, G. V. Loch, and A. Silva. Comparison of support vector regression techniques and neural networks in option pricing. *XLII SBPO*, pages 572–583, 2011. In Portuguese: Comparação das técnicas de support vector regression e redes neurais na precificação de opções.
- Y. Bengio, A. Courville, and P. Vincent. Unsupervised feature learning and deep learning: A review and new perspectives technical report. *Department of computer science and operations research*, 2012.
- A. Bobet, A. Fakhimi, S. Johnson, J. Morris, F. Tonon, and M. R. Yeung. Numerical models in discontinuous media: review of advances for rock mechanics applications. *Journal of geotechnical and geoenvironmental engineering*, 135(11):1547–1561, 2009.
- T. Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3):307–327, 1986.
- B. E. Boser, I. M. Guyon, and V. N. Vapnik. A training algorithm for optimal margin classifiers. In *Proceedings of the fifth annual workshop on Computational learning theory*, pages 144–152, 1992.
- G. E. Box and G. C. Tiao. Intervention analysis with applications to economic and environmental problems. *Journal of the American Statistical Association*, 70(349):70–79, 1975.

- G. E. Box, G. M. Jenkins, and G. C. Reinsel. *Time-series analysis - forecasting and control*. John Wiley and Sons, Inc., 4th edition, 2008.
- C.-C. Chang and C.-J. Lin. LIBSVM: a library for support vector machines. *ACM transactions on intelligent systems and technology (TIST)*, 2(3):1–27, 2011.
- X. Chen and Y. Liu. *Finite element modeling and simulation with ANSYS Workbench*. CRC press, 2018.
- R. T. Clemen. Combining forecasts: A review and annotated bibliography. *International journal of forecasting*, 5(4):559–583, 1989.
- Ö. Çokluk and D. Koçak. Using Horn's parallel analysis method in exploratory factor analysis for determining the number of factors. *Educational Sciences: Theory and Practice*, 16(2):537–552, 2016.
- C. Cortes and V. Vapnik. Support-vector networks. *Machine learning*, 20(3):273–297, 1995.
- G. Cybenko. Approximation by superpositions of a sigmoidal function. *Mathematics of control, signals and systems*, 2(4):303–314, 1989.
- I. Daubechies. *Ten lectures on wavelets*. SIAM, 1992.
- A. C. Davison and D. V. Hinkley. *Bootstrap methods and their application*. Cambridge University Press, 1997.
- D. L. Donoho and I. M. Johnstone. Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 81(3):425–455, 1994a.
- D. L. Donoho and J. M. Johnstone. Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 81(3):425–455, 1994b.
- B. Efron. Bootstrap methods: another look at the jackknife. In *Breakthroughs in statistics*, pages 569–593. Springer, 1992.
- B. Efron and R. J. Tibshirani. *An introduction to the bootstrap*. CRC Press, 1994.
- ENCC. Itaipu hydroelectric power plant dam. Technical report, Department of Civil Engineering and Architecture. Foz do Iguaçu: Itaipu Binacional, 2019.
- R. F. Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the econometric society*, pages 987–1007, 1982.
- A. Faria and E. Mubwandarikwa. The geometric combination of bayesian forecasting models. *Journal of Forecasting*, 27(6):519–535, 2008.
- A. Faria and E. Mubwandarikwa. Multimodality on the geometric combination of bayesian forecasting models. *International Journal of Statistics and Management System*, 3(1):1–25, 2008.
- B. E. Flores and E. M. White. A framework for the combination of forecasts. *Journal of the Academic of Marketing Science*, 16(3-4):95–103, 2008.
- D. N. Gujarati and D. C. Porter. *Econometrics*. McGraw-Hill Interamericana, 2010. In Portuguese: *Econometria*.
- J. D. Hamilton. *Time series analysis*. Princeton University Press, 1994.
- E. Haven, X. Liu, and L. Shen. De-noising option prices with the wavelet method. *European Journal of Operational Research*, 222(1):104–112, 2012.

- S. Haykin. *Redes neurais: princípios e prática*. Bookman Editora, 2001.
- S. Haykin. *Neural networks and learning machines, 3/E*. Pearson Education India, 2009.
- T. Heskes. Selecting weighting factors in logarithmic opinion pools. *Advances in neural information processing systems*, 10, 1997.
- T. Hickmann, L. A. Teixeira Jr, A. Faria, S. B. Rodrigues, J. M. Corrêa, E. L. Garcia, and L. A. S. Aracayo. Temperature field forecast in concrete dam with the use of arima models and the finite element method. In *XXIII SIMPEP*, Vitória - ES, Brazil, 2016.
- T. Hourti. Application of latent variable methods to process control and multivariate statistical process control in industry. *International Journal of adaptive control and signal processing*, 19: 216–246, 2005.
- ICOLD-CIGB. *Dams and The World's Water - An educational book that explains how dams help to manage the world's water*. International Commission on Large Dams/Commission Internationale des Grand Barrages, 2008.
- F. P. Incropera, D. P. Dewitt, et al. *Fundamentals of Heat and Mass Transfer*. John Wiley & Sons, 2011.
- Itaipu. *Studies for the reassessment of control values for civil instrumentation of structures and foundation*. Itaipu Binacional, 1999. In Portuguese: Estudos para a reavaliação dos valores de controle para a instrumentação civil das estruturas e fundações.
- Itaipu. Itaipu binacional. <https://www.itaipu.gov.br/institucional/localizacao>, May 2022a.
- Itaipu. Itaipu binacional. <https://www.itaipu.gov.br/sala-de-imprensa/perguntas-frequentes>, May 2022b.
- Itaipu. Itaipu binacional. <https://www.itaipu.gov.br/sala-de-imprensa/positionpapers/itaipu-referencia-em-seguranca-de-barragem>, May 2022c.
- J. E. Jackson. *A user's guide to principal components*. John Wiley & Sons, Inc., 1991.
- R. Johnson and D. Wichern. *Applied Multivariate Statistical Analysis*. Pearson Modern Classics for Advanced Statistics Series. Pearson, 2019.
- L. A. T. Júnior. *SSA-Wavelet combination of predictive methods with MINIMAX numerical adjustment, in the generation of forecasts and scenarios*. PhD thesis, PUC-Rio, 2013. In Portuguese: Combinação SSA-Wavelet de métodos preditivos com ajuste numérico MINIMAX, na geração de previsões e de cenários.
- A. Khosravi, S. Nahavandi, D. Srinivasan, and R. Khosravi. Constructing optimal prediction intervals by using neural networks and bootstrap method. *IEEE transactions on neural networks and learning systems*, 26(8):1810–1815, 2014.
- T. Kriechbaumer, A. Angus, D. Parsons, and M. Rivas Casado. An improved wavelet-arima approach for forecasting metal prices. *Resources Policy*, 39:32–41, 2014.
- U. Kruger and L. Xie. *Statistical monitoring of complex multivariate process with applications in industrial process control*. John Wiley & Sons Inc., 2012.
- C. Kubrusly and N. Levan. Abstract wavelets generated by hilbert space shift operators. *Advances in Mathematical Sciences and Applications*, 16(2):643, 2006.
- C. S. Kubrusly. *The elements of operator theory*, 2011.

- E. Lazzarotto, L. M. Gramani, A. Chaves Neto, and L. A. Teixeira Junior. A wavelet sarima-ann linear combination with multiple stages in time series prediction. *International Journal of Engineering and Applied Sciences*, 7(4):8–16, 2015a.
- E. Lazzarotto, L. M. Gramani, A. Chaves Neto, L. A. Teixeira Junior, and E. M. C. Franco. Improved time series prediction of dam monitoring instruments via combination of methods. *Cadernos do IME - Serie Estatística*, 38(Junho de 2015):1–20, 2015b. In Portuguese: Melhoria na previsão da série temporal de instrumentos de monitoramento de barragem via combinação de métodos.
- E. Lazzarotto, L. M. Gramani, A. Chaves Neto, and L. A. Teixeira Junior. Principal components in multivariate control charts applied to data instrumentation of dams. *Independent Journal of Management & Production*, 7(1):17–37, 2016.
- N. Levan and C. S. Kubrusly. A wavelet “time-shift-detail” decomposition. *Mathematics and Computers in Simulation*, 63(2):73–78, 2003.
- R. W. Lewis, P. Nithiarasu, and K. N. Seetharamu. *Fundamentals of the Finite Element Method for Heat and Fluid Flow*. John Wiley & Sons, 2004.
- H. Lütkepohl. *New introduction to multiple time series analysis*. Springer Science & Business Media, 2005.
- D. D. Maken, P. Léger, and S. N. Roth. Seasonal thermal cracking of concrete dams in northern regions. *Journal of Performance of Constructed Facilities*, 28(4), 2013.
- S. Mallat. A wavelet tour of signal processing: The sparse way, 805 p, 2008.
- A. C. B. Mancuso and L. Werner. Review of combining forecasts approaches. *Independent Journal of Management & Production*, 4(1):248–277, 2013.
- D. W. Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *Journal of the society for Industrial and Applied Mathematics*, 11(2):431–441, 1963.
- R. L. Mason and J. C. Young. *Multivariate statistical process control with industrial applications*. ASA SIAM - American statistical Association - Society for industrial and applied mathematics, 2002.
- R. B. Millar. *Maximum likelihood estimation and inference: with examples in R, SAS and ADMB*. John Wiley & Sons, 2011.
- D. C. Montgomery. *Introduction to Statistical Quality Control*. LTC, 4th edition, 2013. In Portuguese: Introdução ao Controle Estatístico de Qualidade.
- L. E. Mujica, J. Rodellar, A. Fernandez, and A. Guemes. Q-statistic and t^2 statistic pca-based measures for damage assessment in structures. *Structural Health Monitoring*, 0(5):539–553, 2011.
- B. A. Nedushan. Multivariate statistical analysis of monitoring data for concrete dams. *PhD Thesis, McGill University - Department of CivilEngineering and Apllied Mechanics*, 2002.
- O. Nerrand, P. Roussel-Ragot, L. Personnaz, G. Dreyfus, and S. Marcos. Neural networks and nonlinear adaptive filtering: Unifying concepts and new algorithms. *Neural computation*, 5(2): 165–199, 1993.
- E. Oñate. *Structural analysis with the finite element method. Linear statics: volume 2: beams, plates and shells*. Springer Science & Business Media, 2013.

- S. Oro, T. Mafioleti, A. Chaves Neto, S. Garcia, and C. Neumann Júnior. Study of the influence of temperature and water level of the reservoir about the displacement of a concrete dam. *International Journal of Applied Mechanics and Engineering*, 21(1), 2016.
- S. Oro, S. Garcia, D. Coelho, F. Reinaldo, G. Campos, M. S. Oliveira, L. Hellmann, and F. Peres. Statistical modeling and prediction of the behavior of a buttress block at the Itaipu Dam. *International Journal of Development Research*, 11(12), 2021.
- C. I. Osako. *Maintenance of drains in dam foundations – The case of the Itaipu Hydroelectric Power Plant*. PhD thesis, Dissertação de Mestrado do Programa de Pós-Graduação em Construção Civil, Universidade Federal do Paraná, 2002. In Portuguese: A manutenção dos drenos nas fundações de barragens – O caso da Usina Hidrelétrica de Itaipu.
- M. N. Ozisik. *Heat Conduction*. John Wiley & Sons, 1993.
- L. Pan and D. N. Politis. Bootstrap prediction intervals for linear, nonlinear and nonparametric autoregressions. *Journal of Statistical Planning and Inference*, 177:1–27, 2016.
- A. Pankratz. *Forecasting with dynamic regression models*. John Wiley & Sons, 2012.
- D. J. Reid. Combining three estimates of gross domestic product. *Economica*, 35(140):431–444, 1968.
- L. Ribeiro, V. Wilhelm, Étore Faria, J. Correa, and A. C. Santos. A comparative analysis of long-term concrete deformation models of a buttress dam. *Engineering Structures*, 193:301–307, 2019.
- S. B. Rodrigues. *Interactive hybrid sarima method support multi-core wavelet vector regression in dam instrument time series prediction*. PhD thesis, Universidade Federal do Paraná, 2015. In Portuguese: Método híbrido interativo sarima support vector regression wavelet de múltiplos núcleos na previsão de séries temporais de instrumentos de barragens.
- T. P. Ryan. *Statistical Methods for Quality Improvement*. John Wiley & Sons Inc., 3rd edition, 2011.
- F. Salazar, R. Moran, M. Toledo, and E. Onate. Data-based models for the prediction of dam behaviour: A review and some methodological considerations. *Arch Computat Methods Eng*, 24: 1–21, 2017.
- K. N. Shukla. *Mathematical Principles of Heat Transfer*. Begell House, 2005.
- M. T. Q. S. Silva. Transient thermal and stress analysis in layered structures. Master's thesis, Universidade Federal do Paraná, 2003. In Portuguese: Análise Térmica Transiente e de Tensões em Estruturas Executadas em Camadas.
- J. F. A. Silveira. *Instrumentation and behavior of concrete dam foundations*. Oficina de Textos, 2003. In Portuguese: Instrumentação e comportamento de fundações de barragens de concreto.
- A. J. Smola and B. Schölkopf. A tutorial on support vector regression. *Statistics and computing*, 14 (3):199–222, 2004.
- L. A. Teixeira Júnior, R. M. d. Souza, M. L. d. Menezes, K. M. Cassiano, J. F. M. Pessanha, and R. C. Souza. Artificial neural network and wavelet decomposition in the forecast of global horizontal solar radiation. *Pesquisa Operacional*, 35:73–90, 2015.
- R. Tibshirani. A comparison of some error estimates for neural network models. *Neural computation*, 8(1):152–163, 1996.

- V. Vapnik. *The nature of statistical learning theory*. Springer Verlag, 1995.
- V. Vapnik. *The nature of statistical learning theory*. Springer Science & Business Media, 1999.
- K. F. Wallis. Combining forecasts—forty years later. *Applied Financial Economics*, 21(1-2):33–41, 2011.
- P.-S. Yu, S.-T. Chen, and I.-F. Chang. Support vector regression for real-time flood stage forecasting. *Journal of hydrology*, 328(3-4):704–716, 2006.
- G. P. Zhang. Time series forecasting using a hybrid arima and neural network model. *Neurocomputing*, 50:159–175, 2003.
- O. C. Zienkiewicz and R. L. Taylor. *The finite element method for solid and structural mechanics*. Elsevier, 2005.

Chapter 11

Application of Statistical Processing Techniques to the Impedance-based SHM for the Oil & Gas Industry

Chapter details

Chapter DOI:

<https://doi.org/10.4322/978-65-86503-88-3.c11>

Chapter suggested citation / reference style:

Silva, José W., et al. (2022). “Application of Statistical Processing Techniques to the Impedance-based SHM for the Oil & Gas Industry”. In Jorge, Ariosto B., et al. (Eds.) *Uncertainty Modeling: Fundamental Concepts and Models*, Vol. III, UnB, Brasilia, DF, Brazil, pp. 337–383. Book series in Discrete Models, Inverse Methods, & Uncertainty Modeling in Structural Integrity.

P.S.: DOI may be included at the end of citation, for completeness.

Book details

Book: Uncertainty Modeling: Fundamental Concepts and Models

Edited by: Jorge, Ariosto B., Anflor, Carla T. M., Gomes, Guilherme F., & Carneiro, Sergio H. S.

Volume III of Book Series in:

Discrete Models, Inverse Methods, & Uncertainty Modeling in Structural Integrity

Published by: UnB City: Brasilia, DF, Brazil Year: 2022

DOI: <https://doi.org/10.4322/978-65-86503-88-3>

Application of Statistical Processing Techniques to the Impedance-based SHM for the Oil & Gas Industry

José Waldemar da Silva^{1*}, Quintiliano S. S. Nomelini¹,
Diogo de Souza Rabelo², José dos Reis V. Moura Jr³,
Roberto M. Finzi Neto⁴, Carlos A. Gallo⁴ and Julio E. Ramos⁵

¹Faculty of Mathematics, Federal University of Uberlândia, Brazil. E-mail: zewaldemar@ufu.br; quintiliano.nomelini@ufu.br

²Post-Graduate Program – Industrial Engineering, Fed. University of Goias, School of Sciences and Technology, Aparecida de Goiania, Brazil. e-mail: diogo.rabelo@ufg.br

³Post-Graduate Program – Modeling and Optimization, Fed. University of Catalao, Brazil. e-mail: zereis@ufcat.edu.br

⁴Post-Graduate Program – Mechanical Engineering, Federal University of Uberlandia, Brazil. e-mail: finzi@ufu.br; gallo@ufu.br

⁵Petrobras, Petroleo Brasileiro S.A., RD Center (CENPES), Brazil. e-mail: julio.ramos@petrobras.com.br

*Corresponding author

Abstract

This chapter presents methodologies for preprocessing electromechanical impedance-based structural health monitoring data sets with the objective of error reduction when inferring changes in the structure. Atypical signatures may not represent the actual state of the structure and should be excluded from the inference step. The impedance signature trend in the frequency domain is another aspect that must be considered because divergences between the samples regarding this feature can also lead to mistakes. Polynomial regression models for matching the effect of temperature on impedance signatures are also discussed in this chapter, and methodologies to infer the state of the structure from the CCD metric.

Keywords: Impedance-based structural health monitoring; Statistical preprocessing tools; Regression models

1 Data preprocessing in SHM

Structural health monitoring can be performed from the electromechanical impedance signature, and a set of signatures acquired under the same conditions are expected to be similar. In other words, repeatability of signatures is expected.

The monitored structure may be located in environments where it is impossible to control variables whose effect may interfere with the electromechanical impedance signature, reducing the degree of repeatability among measurements. The repeatability requirement may not be met due to atypical signatures. Abrupt actions resulting from a falling object, somebody walking, or a bird landing is an uncontrolled interference that can affect the electromechanical impedance signature.

The verification of outliers in the damage metrics, calculated within the same group of signature repetitions, can help detect atypical signatures, which should be removed from the database for further analysis. The procedures used for the necessary data cleansing will be called pre-processing (Section 1).

Further analyses, in this chapter, in particular, are defined as the processing (Section 2) of the baseline data for the determination of polynomial models for temperature compensation and the post-processing (Section 3) as the use of data to investigate the integrity of the structure.

After collecting a set of signatures representing the same state of the structure, the median signature can be taken as a reference signature. In this way, it is possible to calculate some damage metrics, such as the CCD (Correlation Coefficient Deviation), which in this situation will depict the degree of variability or similarity between the group's signatures. The median signature is chosen as a reference due to the generally asymmetrical character of the electromechanical impedance distribution at each frequency.

The correlation coefficient is defined as in Eq. 1

$$CC_{Z_1 Z_2} = \frac{1}{n} \sum_{i=1}^n \left\{ \frac{[Z_{1,i} - \bar{Z}_1][Z_{2,i} - \bar{Z}_2]}{S_{Z_1} S_{Z_2}} \right\} \quad (1)$$

where $Z_{1,i}$ and $Z_{2,i}$ are the real parts of the electromechanical impedance of the two signals involved in calculating the metric, at frequency i , with $i = 1, 2, \dots, n$ and n is the number of frequency points used to data acquisition. To calculate the correlation coefficient within the same group of signature repetitions, you can define Z_1 as any signature and Z_2 as the median signature. The averages of each sample set are represented respectively by \bar{Z}_1 and \bar{Z}_2 and, S_{Z_1} and S_{Z_2} are the standard deviations of them. The metric is then determined by Eq. 2.

$$CCD = 1 - CC_{Z_1, Z_2}. \quad (2)$$

Example 1. Consider the 16 impedance signatures as presented in Table 7 (adapted data), measured from a real experiment developed at the Structural Mechanics Lab of the Mechanical Engineering School of the Federal University of Uberlândia. Calculating the CCD values associated with each measurement using the R software:

Solution

The developed code is shown in the following image (Figure 1), whose output is the CCD values rounded to 4 decimal places.

```

> med=vector()
> for(i in 1:nrow(data)){
+   med[i]=median(data[i,])} # median signature calculation
> ccd=vector()
> for(i in 1:ncol(data))
+   {
+     ccd[i]= 1-sum( (data[,i] - mean(data[,i]))*(med - mean(med))/
+                 (sd(data[,i]) * sd(med)) ) /nrow(data)
+   } # Eqs. 1 and 2
> round(ccd,digits=4)
[1] 0.0202 0.0202 0.0201 0.0202 0.0201 0.0201 0.0201 0.0201
[9] 0.0201 0.0202 0.0201 0.0223 0.0209 0.0209 0.0208 0.0207
>

```

Figure 1: Calculating CCD in R language.

It is assumed that outliers values of CCD will be obtained from atypical or discordant signatures about the others in the group. Therefore, the corresponding signatures must be retaken from the database for further analysis. There are some methods of identifying outliers, among them, the Tukey method [Tukey, 1977], better known as boxplot, and the method de Chauvenet [Reddy, 2011]. These two methods will be presented in subsections 1.1 and 1.2 to follow.

1.1 Tukey Method - Boxplot

The box plot or BoxPlot is defined by a rectangle in which, if presented vertically, the lower and upper sides coincide with the first (q_1) and the third (q_3) quartiles, respectively. The inner dash is the median value (q_2). The distance between q_1 and q_3 is the interquartile range, $IIQ = q_3 - q_1$. This interval contains 50% of the central observations, and its amplitude, as well as the full amplitude of the Boxplot, contains information about the dispersion of the data.

A value located between $q_1 - 1.5 \times IIQ$ and $q_1 - 3 \times IIQ$ or between $q_3 + 1.5 \times IIQ$ and $q_3 + 3 \times IIQ$ is called outlier. A point beyond $q_3 + 3 \times IIQ$ or short of $q_1 - 3 \times IIQ$ is called an extreme value. In this text we will not make this distinction and extreme values will also be called outlier (Figure 4).

Example 2. Consider the CCD values presented in the example solution 1 and check the existence of outliers values using Tukey's method.

Solution

Getting $q_1 = \frac{x_{(4)} + x_{(5)}}{2} = 0.0201$ and $q_3 = \frac{x_{(12)} + x_{(13)}}{2} = 0.02075$, values greater than $0.02075 + 1.5 \times (0.02075 - 0.0201) = 0.0217$ or less than $0.0201 - 1.5 \times (0.02075 - 0.0201) = 0.0191$ are classified as outliers. The index in parentheses, $x_{(i)}$, indicates the value of x in the i -th position after all values are sorted in ascending order. In this example, only the value 0.0223 is classified as outliers.

The descriptive analysis presented in the Example 2 can be performed in the Software R [R Core Team, 2021]. However, a different methodology for determining the quartiles. There are several methods for estimating quartiles [Langford, 2006], and they may differ in the way of determining their position and their calculation itself.

The same data explored in the Example 2 were used to illustrate the obtaining of outliers from the standard methodology implemented in the R software (Figure 2).

```
> boxplot(ccd)$out  
[1] 0.0223  
>
```

Figure 2: Calculating outlier in R language.

The outlier point, as well as the behavior of the data distribution, are shown in Figure 3.

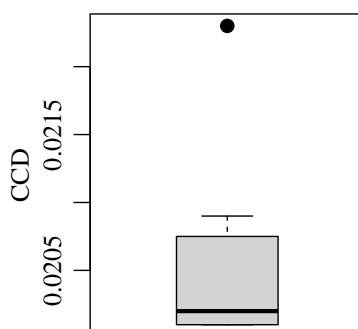


Figure 3: Boxplot of the outlier as well as the CCD distribution.

Based on this criterion, the twelfth signature ($CCD=0.0223$) may have been acquired under the effect of some atypical event and therefore does not represent the actual status of the structure and thus must be removed from the database for future analysis.

Another example of detecting *outliers* at the two extremes of the data distribution, from a set of data arising from the observation of any variable Y , is shown in Figure 4.

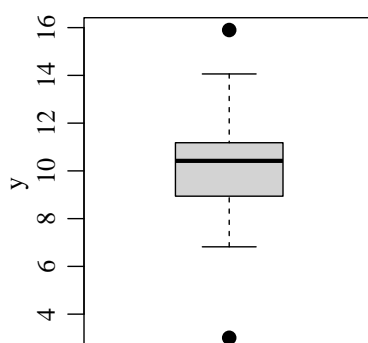


Figure 4: Boxplot of the outliers at the extreme right (top) and left (bottom) of the data distribution.

The values of outliers shown in Figure 4, as well as the positions of these values in the data list, can be obtained with the R commands shown in the following image (Figure 5)


```

> y=c(10.23, 7.39, 10.84, 8.82, 6.82, 15.90, 10.86, 8.94,
+     10.42, 10.04, 11.10,7.47, 11.87, 14.06, 13.43, 10.92,
+     11.18, 12.48, 10.11, 10.42, 3.01, 9.29)
> boxplot(y)$out
[1] 15.90 3.01
> match(boxplot(y)$out,y)
[1] 6 21

```

Figure 5: Identification of outlier values and their respective positions in the data vector.

The position of an outlier damage metric in the list of values makes it easier to identify the atypical impedance signature associated with that value.

1.2 Chauvenet's method

This method is based on the number of standard deviations z (Eq. 3) that a value, (y) is far from the mean (\bar{y}).

$$z = \frac{y - \bar{y}}{s} \implies (y - \bar{y}) = zs \quad (3)$$

If z is greater than the maximum allowable deviation (d_{max}) or less than the minimum permissible deviation (d_{min}), then the corresponding value y is considered an outlier.

The thresholds d_{max} and d_{min} are determined from the normal distribution. Therefore, the hypothesis of normality of the variable (Y), from which measurements or observations are made, cannot be rejected. This is a necessary assumption for applying the Chauvenet criterion [Taylor, 2012].

According to this criterion, a value is considered non-typical if the corresponding number of deviations from the mean z belongs to a range of values, symmetrically delimited by d_{min} and d_{max} around the mean, with probability $1 - 1/2n$ where n is the number of elements in the sample.

Since z , presented in Eq. 3 is a standardized score, the critical values of the Chauvenet test (d_{max} and d_{min}) are determined from the resolution of the integral as presented in Eq. 4.

$$P(-d_{max} < Z < d_{max}) = \int_{-d_{max}}^{d_{max}} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{z^2}{2}\right\} dz = 1 - \frac{1}{2n} \quad (4)$$

Figure 6 shows the area corresponding to the probability presented in Eq. 4, for any n .

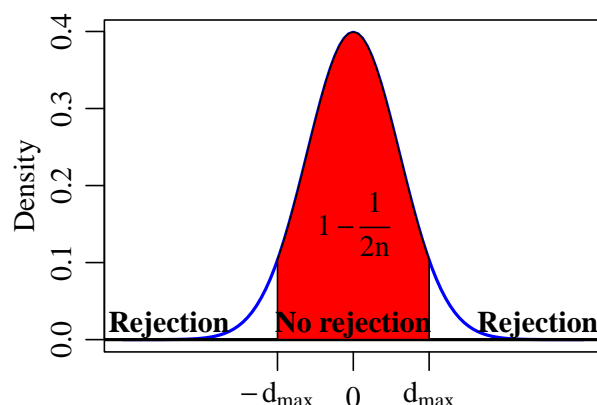


Figure 6: Illustration of non-rejection and rejection data regions using the Chauvenet criterion.

To eliminate the need to determine two thresholds, d_{max} and d_{min} , just use the absolute value of the difference in Eq. 3, that is $z = \frac{|x - \bar{x}|}{s}$ and x will be outliers if $z > d_{max}$. This strategy is justified by the symmetry property of the normal distribution.

Failure to comply with the assumption of normality may result in not-so-promising results. Therefore, in this case, other techniques for detecting outliers should be used, such as Tukey's (Section 1.1).

In practical terms, the critical value d_{max} can be obtained from some software that returns quantiles of the normal distribution for the probability $1 - \frac{1}{2n}$ (Eq. 4).

Example 3. Consider the dataset presented in Figure 5 and check for the presence of outlier values using the Chauvenet method. If they exist, identify their positions in the dataset.

Solution

```
> chauvenet=function(x) {
+   n=length(x)
+   prob=1-1/(2*n)
+   q1=qnorm((1-prob)/2)
+   q2=abs(q1)
+   z=(x-mean(x))/sd(x)
+   out=y[z<q1 | z>q2]
+   c(out)
+ }
> data=c(10.23, 7.39, 10.84, 8.82, 6.82, 15.90, 10.86, 8.94,
+        10.42, 10.04, 11.10, 7.47, 11.87, 14.06, 13.43, 10.92,
+        11.18, 12.48, 10.11, 10.42, 3.01, 9.29)
> out=chauvenet(data);out
[1] 3.01
> match(out,y) # outliers position
[1] 21
```

Figure 7: R source code to identify outliers and respective positions.

Using the Chauvenet function presented in Figure 7 can be obtained the value 3.01. It is classified as an outlier by the Chauvenet method. Match function indicates that this value occupies the 21st position in the dataset.

It is recommended that the Chauvenet criterion be applied only once (Reddy [2011]; Taylor [2012]).

1.3 Trend Analysis

The trending effect on signatures can lead to erroneous inferences about the state of the structure or temperature compensation polynomials if this effect is not the same in all signatures. Therefore, verifying its presence and removal is necessary for further processing or inference.

It is suggested that if trend correction is necessary for some impedance signatures, the same procedure is performed in the others to match them.

The significance of the exponential model (Eq. 5), or its correspondent in linearized form (Eq. 6), for some signature, indicates that the trend characteristic is present and must be adjusted.

The significance of the regression model can be investigated through the analysis of variance. Also, it can be done by verification of the null hypothesis for the parameter associated with the independent regressor variable, in the case of simple linear regression.

The determination coefficient R^2 is another measure that helps in the decision regarding the suitability of the model.

The methodology and basic concepts necessary for regression analysis are presented in the 1.4 section.

1.4 Regression Model for Trend Adjustment

Assuming that the trend effect is exponential, the model presented in Eq. 5 can be used. In this model, impedance is considered a function of frequency:

$$y_i = \beta_0 e^{\beta_1 x_i} \times \epsilon_i; \quad i = 1, 2, \dots, n, \quad (5)$$

where y_i and x_i are respectively the impedance and frequency values obtained in the i -th observation to which the error ϵ_i is also associated; n is the frequency sample at which impedances were collected.

Assuming that the model error described in Eq. 5 is multiplicative, then this model can be approached in its linearized form (Eq. 6). Linearization is advantageous since, in this case, the system of ordinary equations associated with the model has an analytical solution. It, therefore, does not depend on the convergence of algorithms to determine the estimates of the coefficients as in the nonlinear case.

$$\ln(y_i) = \ln(\beta_0) + \beta_1 x_i + \ln(\epsilon_i) \quad (6)$$

If the proposed model is indeed effective then the trend-free signature will be obtained by $y_i - \exp(\widehat{\ln(y_i)})$.

To present some theoretical aspects of the simple linear regression model, consider two variables y and x where y is a variable that linearly depends on x . The statistical model of a simple linear regression is shown in Eq. 7.

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i \tag{7}$$

where y_i : represents the i -th observed value; x_i : the independent variable, $i = 1, 2, \dots, n$; ϵ_i : is the unobservable error associated with i -th observation; β_0 and β_1 : are the intercept and the slope, parameters of the model.

In a linear regression model, the following assumptions must be met:

-
- The x values are fixed, that is, known before the observation of y and therefore x is not a random variable;
- The mean error is zero, ie $E[\epsilon_i] = 0, \forall i = 1, 2, \dots, n$;
- For a given value of x , the error variance is always constant, that is, $V[\epsilon_i] = E[\epsilon_i^2] - (E[\epsilon_i])^2 = E[\epsilon_i^2] = \sigma^2, \forall i = 1, 2, \dots, n$, The error is then said to be homoscedastic;
- The error of one observation is uncorrelated with the error of another observation (the errors are independent), ie $E[\epsilon_i \epsilon_j] = 0$, for $i \neq j$;
- The error has a Normal distribution with zero mean and constant variance (σ^2), that is, $\epsilon_i \sim N(0, \sigma^2)$. Concluding, the errors are independent and identically distributed, that is, $\epsilon \sim iidN(0, \sigma^2)$.

This last item assembles the necessary assumptions so that the inference on the model parameters can be made from the parametric point of view, the ordinary least-squares method. On the other hand, if these assumptions are not met, the inference must be made by alternative means such as the bootstrap method Efron and Tibshirani [1993].

A dependent variable can be modeled as a function of a single regressor variable (Eq. 7), by several regressors (Eq. 8)

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + \epsilon_i, \tag{8}$$

where x_1, x_2, \dots, x_k represent the k explanatory variables considered. Yet another possibility is the modeling of y by a single regressor variable but in polynomial form (Eq. 9) and in this case, k represents the degree of the polynomial

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \dots + \beta_k x_i^k + \epsilon_i. \tag{9}$$

The model presented in Eq. 7 is a particular case of both the model presented in Eq. 8 and Eq. 9. It should be noted that other variations are possible as multiple linear regression models, in addition to the linear terms, quadratic, cubic terms, etc., are considered for some or some regressors in addition to interactions between them. The real need for these terms and even for the variables in the modeling can be verified from the techniques of selection of variables.

The models presented in Eqs. 7, 9 and 8 can be written in matrix form as in Eq. 11 since for each i , there is a line as shown in Eq. 10.

$$\underbrace{\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}}_{\mathbf{Y}} = \underbrace{\begin{pmatrix} 1 & x_{11} & x_{12} & \dots & x_{1k} \\ 1 & x_{21} & x_{22} & \dots & x_{2k} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nk} \end{pmatrix}}_{\mathbf{X}} \underbrace{\begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{pmatrix}}_{\boldsymbol{\beta}} + \underbrace{\begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix}}_{\boldsymbol{\epsilon}} \tag{10}$$

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \quad (11)$$

When $k = 1$ in Eq. 10 it means a simple linear regression model (Eq. 7) and when the columns of the matrix \mathbf{X} correspond to the quadratic terms x^2 , x^3 , etc, from this same variable x , represents the polynomial model.

1.4.1 Least-squares estimators

The least-squares estimator for $\boldsymbol{\beta}$ is determined by minimizing the sum of squared errors, presented in Eq. 12 which can be interpreted as a function of $\boldsymbol{\beta}$, which will be named $Q(\boldsymbol{\beta})$.

$$\begin{aligned} Q(\boldsymbol{\beta}) &= \boldsymbol{\epsilon}'\boldsymbol{\epsilon} = (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}) \\ &= \mathbf{Y}'\mathbf{Y} - \mathbf{Y}'\mathbf{X}\boldsymbol{\beta} - \boldsymbol{\beta}'\mathbf{X}'\mathbf{Y} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta}. \end{aligned} \quad (12)$$

Note that $\mathbf{Y}'\mathbf{X}\boldsymbol{\beta}$ and $\boldsymbol{\beta}'\mathbf{X}'\mathbf{Y}$ are scalar and therefore equal. Then Eq 12 can be presented as in Then Eq 13,

$$Q(\boldsymbol{\beta}) = \mathbf{Y}'\mathbf{Y} - 2\boldsymbol{\beta}'\mathbf{X}'\mathbf{Y} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta}. \quad (13)$$

Differentiating $Q(\boldsymbol{\beta})$ with respect to $\boldsymbol{\beta}$ is obtained:

$$\frac{\partial Q(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = -2\mathbf{X}'\mathbf{Y} + 2\mathbf{X}'\mathbf{X}\boldsymbol{\beta} \quad (14)$$

and equating to Eq. 14 set to zero, specific values are admitted for the vector $\boldsymbol{\beta}$, allocated in a vector \mathbf{b} , which satisfies Eq. 15

$$-2\mathbf{X}'\mathbf{Y} + 2\mathbf{X}'\mathbf{X}\mathbf{b} = 0, \quad (15)$$

and therefore the least-squares estimator of $\boldsymbol{\beta}$ is as shown in Eq. 16

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}. \quad (16)$$

Determining the regression equation or the estimation of regression coefficients does not reflect any information about the significance of the model, that is, how important the estimated model is to explain the response variable. The regression analysis of variance allows this study and is presented in the 1.4.2 section below.

1.4.2 Analysis of variance for regression

The analysis of variance for regression consists of verifying how much variation of the dependent variable is explained by the model adopted. The total variance quantified through the total sum of the square, SS_{Total} , can be written as the sum of the variance explained by the regression model, SS_{Reg} , with the unexplained or residual variance, SS_{Res} .

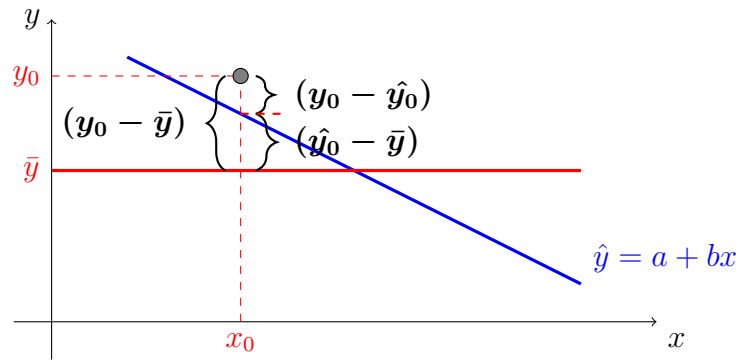


Figure 8: Decomposition of deviations from the mean in a simple linear regression.

Figure 8 illustrates the decomposition of the variation in y considering a simple linear regression. Note, from this Figure, that $(y_0 - \bar{y}) = (y_0 - \hat{y}_0) + (\hat{y}_0 - \bar{y})$ and, generalizing, for any pair (x_i, y_i) obtains,

$$(y_i - \bar{y}) = (y_i - \hat{y}_i) + (\hat{y}_i - \bar{y}) \tag{17}$$

It is possible to show that

$$SSTotal = SSReg + SSEerror$$

or, that

$$\underbrace{\sum_{i=1}^n (y_i - \bar{y})^2}_{SSTotal} = \underbrace{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}_{SSReg} + \underbrace{\sum_{i=1}^n (y_i - \hat{y}_i)^2}_{SSEerror} \tag{18}$$

To show the validity of the equality in Eq. 18, apply the sum to the square of the deviations, term to the right of the equality in Eq. 17, after subtracting and conveniently adding \hat{Y}_i , that is,

$$\begin{aligned} SSTotal &= \sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (y_i - \hat{y}_i + \hat{y}_i - \bar{y})^2 = \sum_{i=1}^n ((y_i - \hat{y}_i) + (\hat{y}_i - \bar{y}))^2 \\ &= \sum_{i=1}^n (y_i - \hat{y}_i)^2 + 2 \underbrace{\sum_{i=1}^n (y_i - \hat{y}_i)(\hat{y}_i - \bar{y})}_{=0} + \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 \end{aligned} \tag{19}$$

and verify that

$$\sum_{i=1}^n (y_i - \hat{y}_i)(\hat{y}_i - \bar{y}) = 0. \tag{20}$$

The validity of Eq. 20 can be shown by noting that

$$\sum_{i=1}^n (y_i - \hat{y}_i)(\hat{y}_i - \bar{y}) = \underbrace{\sum_{i=1}^n (y_i - \hat{y}_i)\hat{y}_i}_I + \bar{y} \underbrace{\sum_{i=1}^n (y_i - \hat{y}_i)}_{II} = 0. \quad (21)$$

Writing I in Eq. 21 in a matrix form obtains

$$\begin{aligned} I &= \sum_{i=1}^n (y_i - \hat{y}_i)\hat{y}_i = (\mathbf{Y} - \mathbf{X}\mathbf{b})' \mathbf{X}\mathbf{b} = \mathbf{Y}' \mathbf{X}\mathbf{b} - \mathbf{b}' \mathbf{X}' \mathbf{X}\mathbf{b} = \mathbf{Y}' \mathbf{X}\mathbf{b} - [(\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}' \mathbf{Y}]' \mathbf{X}' \mathbf{X}\mathbf{b} \\ &= \mathbf{Y}' \mathbf{X}\mathbf{b} - \mathbf{Y}' \mathbf{X} (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}' \mathbf{X}\mathbf{b} = \mathbf{Y}' \mathbf{X}\mathbf{b} - \mathbf{Y}' \mathbf{X}\mathbf{b} = 0 \end{aligned}$$

and further, knowing that the least-squares method consists of minimizing, that is, determining the values β_s that minimize the sum of squares in Eq. 22,

$$\sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (Y_i - [\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki}])^2 \quad (22)$$

applying this method will lead to a system with $k + 1$ equations and $k + 1$ unknowns in which the first equation is obtained by differentiating the given function in Eq. 22 and setting the result to zero. So the derivative concerning β_0

$$\begin{aligned} &\frac{\partial \sum_{i=1}^n (y_i - [\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki}])^2}{\partial \beta_0} \\ &= -2 \sum_{i=1}^n (y_i - [\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki}]) \end{aligned} \quad (23)$$

and equating to zero, specific values are allowed for $\beta_0, \beta_1, \beta_2, \dots, \beta_k$, which $b_0, b_1, b_2, \dots, b_k$ satisfy Eq. 24, below.

$$\sum_{i=1}^n (y_i - [b_0 + b_1 x_{1i} + b_2 x_{2i} + \dots + b_k x_{ki}]) = 0. \quad (24)$$

then,

$$II = \sum_{i=1}^n (y_i - \hat{y}_i) = 0 \quad (25)$$

because,

$$II = \sum_{i=1}^n (y_i - \hat{y}_i) = \sum_{i=1}^n (y_i - [b_0 + b_1 x_{1i} + b_2 x_{2i} + \dots + b_k x_{ki}]) = 0. \quad (26)$$

From the results presented in Eqs. 22 and 26 it follows that the sum in Eq. 20 is in fact zero and therefore the total change in Y , quantified by $SSTotal$ is decomposed into the model-explained change, $SSReg$, and the residual change, $SSError$.

Under the assumption that the errors are independent and identically distributed according to the normal distribution with zero mean and constant variance, σ^2 , the F test of the analysis of variance for the linear regression model (Table 1) allows testing the hypothesis that the variance explained by the model, with p coefficients, is significant. This hypothesis is equivalent to the idea that the slope parameter of the line, or slope of the line, β_1 is zero, in the case of simple linear regression, or that the coefficients of the explanatory variables are all null as shown in Eq. 27. The alternative hypothesis is that at least one of these coefficients differs from zero.

Table 1: Analysis of variance for the complete model

Source	DF	SS	MS	F-Value
Regression	$p - 1$	$SSReg$	$MSReg$	$MSReg/MSError$
Error	$n - p$	$SSError$	$MSError$	
Total	$n - 1$	$SSTotal$		

The hypotheses verified from the analysis of variance can be described as follows,

$$\begin{cases} H_0 : \beta_0 = \beta_1 = \dots = 0 \\ H_a : \beta_j \neq 0 \text{ para pelo menos um } j, \text{ com } j = 0, 1, 2, \dots, k \end{cases} \quad (27)$$

Thus, H_0 is rejected, at the α level of significance, if $F_0 \geq F_{\alpha, p-1, n-p}$ where $F_{\alpha, p-1, n-p}$ is a quantile of the F distribution of *Snedecor* with $p-1$ and $n-p$ degrees of freedom such that $P(F > F_{\alpha, p-1, n-p}) = \alpha$. The values $MSReg$ and $MSError$ are obtained by doing $SSReg/(p-1)$ and $SSError/(n-p)$, respectively.

1.4.3 Goodness of fit

As mentioned earlier, parametric inference on a regression model requires the fulfillment of some assumptions such as normality, independence, and homoscedasticity of errors. The lack of knowledge of the population model makes it necessary to consider the evidence for rejection or non-rejection of these assumptions. This evidence is obtained from the sample data and summarized in the estimated model. The decision on whether or not to reject the assumptions is made based on appropriate hypothesis tests. When any of these assumptions are not met, alternative methodologies must be used, such as data transformations in cases of non-normality and homogeneity or dependence modeling in cases of non-independence, or even non-parametric methodologies such as the bootstrap technique.

The assumptions necessary for the validity of the parametric inference from models estimated via the ordinary least-squares method are:

- **Normality of errors:** errors must be normally distributed. Normality is required so that hypothesis tests based on the normal, t, chi-square, and F distribution are unbiased. There are some normality test methods such as Kolmogorov-Smirnov, Lilliefors, Shapiro-Wilk Royston [1995];
- **Independence of errors or residuals:** the errors must be independent, that is, the probability of occurrence of any error must not depend on other errors. There are some tests to

verify the independence hypothesis, among which the Durbin-Watson Durbin [1969] test, which is widely used;

- **Homogeneity of errors:** The lack of homogeneity can interfere with the hypothesis tests (F test and t-test) and the quality of the estimates of the model coefficients. In the presence of heterogeneity, least squares estimators may not have a minimum variance compared to other estimators, and therefore they are no longer efficient. To verify the hypothesis of homoscedasticity, the Breusch-Pagan [Breusch and Pagan, 1979] or Levene [Fox and Weisberg, 2018] tests can be used.

The Shapiro-Wilk, Durbin-Watson and Breusch-Pagan tests are implemented in R software through the functions *shapiro.test*, *dwtest* and *bptest*. These last two are in the *lmtest* [Zeileis and Hothorn, 2002] package.

The Shapiro-Wilk test is limited to samples with sizes between 3 and 5000, but there are several other tests in the literature with the same purpose [Khatun et al., 2021]. Among them is the Anderson-Darling test [Thode, 2002b] that can be applied from the function *ad.test*, implemented in the package *nortest* [Gross and Ligges, 2015], of the software R.

Graphical methods involving the residuals can also be used to verify the adequacy of the model. Some graphs and their usefulness are presented below.

- Residuals versus adjusted/predicted values:
 - It may indicate poor specification of the model;
 - Evaluate whether the errors have constant variance;
 - Identify *outliers*.

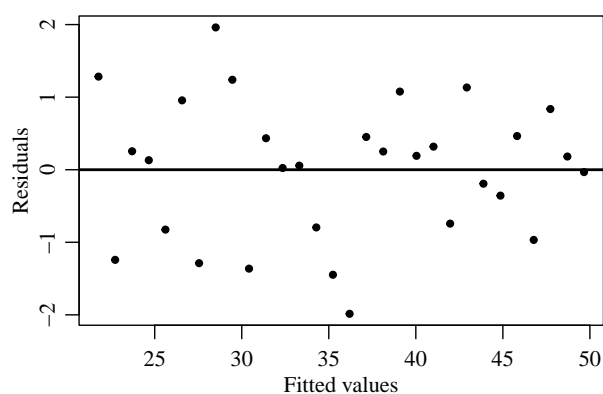


Figure 9: Residuals versus adjusted values.

- Q-Q plot:
 - Check if the errors are approximately normally distributed;
 - Identify *outliers*.

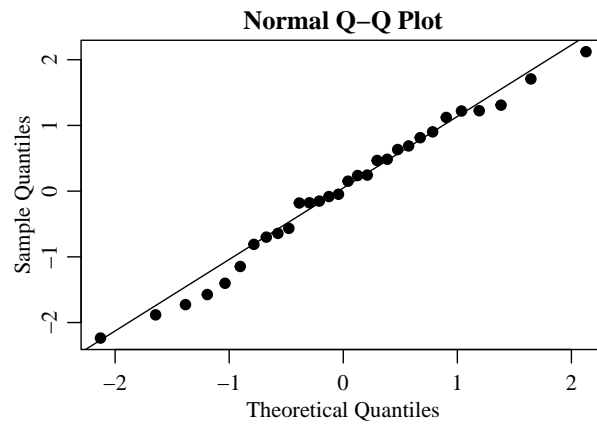


Figure 10: Q-Q plot for residuals.

- Residuals versus collection order:
 - Check the induced correlation by the collection order;

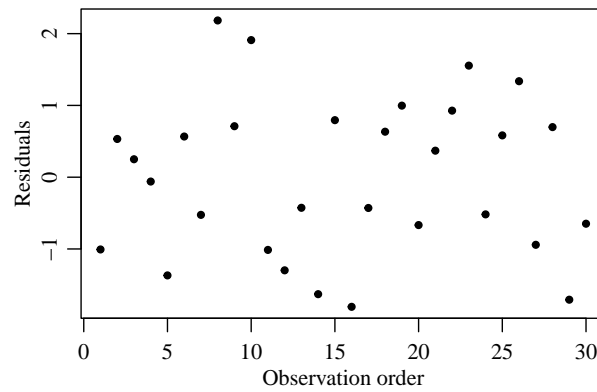


Figure 11: Residuals versus observation collection order.

- Residuals versus omitted variable in the model:
 - Any systematic pattern indicates the need to incorporate the variable into the model.

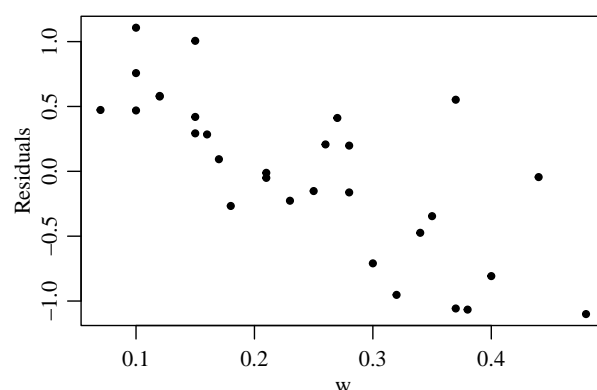


Figure 12: Residuals versus covariate w omitted in the model.

- Residuals versus variable included in the model:
 - Systematic patterns may indicate another relationship between the considered regressor variable and the response variable.
 - Assessment of the homogeneity.

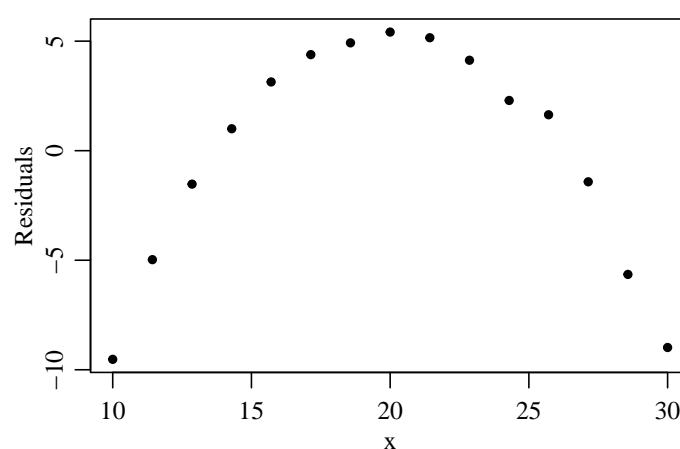


Figure 13: Residuals versus variable x included in the model.

In the Figures 9, 10 and 11 there is a desirable behavior of the residuals when the model is adequate but, on the other hand, Figures 12 and 13 illustrate the inadequacy of the model.

The coefficient of determination R^2 and the adjusted coefficient of determination \bar{R}^2 are measures of model adequacy and express how much of the variation of the response variable is explained by the model.

$$R^2 = \frac{SS_{Reg}}{SST_{Total}} \quad (28)$$

$$\bar{R}^2 = 1 - \left(\frac{n-1}{n-p} \right) (1 - R^2) \quad (29)$$

As the number of parameters, p , increases, with n fixed, \bar{R}^2 decreases when compared to R^2 . This penalty is important because the variation explained by a model with a greater number of parameters will always be greater, which is desirable. On the other hand, the model will be more complex, which is not desirable. Therefore \bar{R}^2 is a better measure of suitability than R^2 when the complexity caused by the greater number of parameters is considered.

Example 4. Let the real part of an electromechanical impedance signature (Ω) be obtained from monitoring a steel plate in an experiment developed at the Gloria campus of the Federal University of Uberlandia. The impedance values were obtained at 3000 frequency points ranging from 40000 to 70000 Hz. Given a large amount of data, only this signature's initial and final values are shown in Table 2.

Table 2: Initial and end values of the impedance signature.

Frequency	Impedance
40000.00	91.10
40010.00	90.86
40020.01	88.53
40030.01	86.53
\vdots	\vdots
69969.99	41.58
69979.99	41.39
69990.00	40.76
70000.00	40.71

a) Estimate the parameters of the linearized exponential model (Eq. 6).

Solution

The referred model can be adjusted from the function \ln of R with the logarithmic transformation in the base e . The log function does this transformation, and the fitted model is:

$$\widehat{\ln(y)} = 5,548 - 2,676 \times 10^{-5} \text{Freq.}$$

The following image (Figure 14) illustrates the R code and its output to obtain this model.

```

> fit=lm(log(y)~Freq)
> summary(fit)

Call:
lm(formula = log(y) ~ Freq)

Residuals:
    Min       1Q   Median       3Q      Max
-0.19681 -0.04099 -0.00412  0.03489  0.33350

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  5.548e+00  6.904e-03   803.6  <2e-16 ***
Freq        -2.676e-05  1.240e-07  -215.8  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.05884 on 2998 degrees of freedom
Multiple R-squared:  0.9395,    Adjusted R-squared:  0.9395
F-statistic: 4.656e+04 on 1 and 2998 DF,  p-value: < 2.2e-16

```

Figure 14: Output of the function *lm* for the linearized exponential model.

- b) Assume that the assumptions of normality, homogeneity, and independence of the residuals are met and verify the significance of the model using the *F* test of the analysis of variance.

Solution

According to the results in item a's figure, the *p*-value associated with the *F* test was less than 2.2×10^{-16} . From this, it is concluded that the model is highly significant in explaining the variation in $\ln(y)$. The value for the *F* test presented in Table 1 (Table of analysis of variance) is defined by the probability $P(F > F_0)$ obtained from an *F* distribution with 1 and 2998 degrees of freedom.

- c) Give the value of the coefficient of determination and interpret it.

Solution

The coefficient of determination was 0.9395, which indicates that the fitted model explains 93.95% of the variation in $\ln(y)$.

- d) Graphically present the signature before and after the trend correction.

Solution

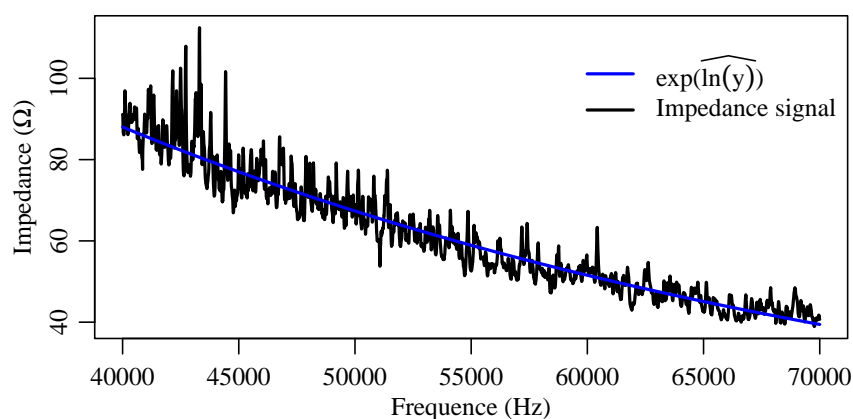


Figure 15: Impedance signature before trend correction.

Since the fitted model satisfactorily models the trend, the impedance signal without trend is obtained by making $y - \exp(\ln(\hat{y}))$ and is shown in Figure 16.

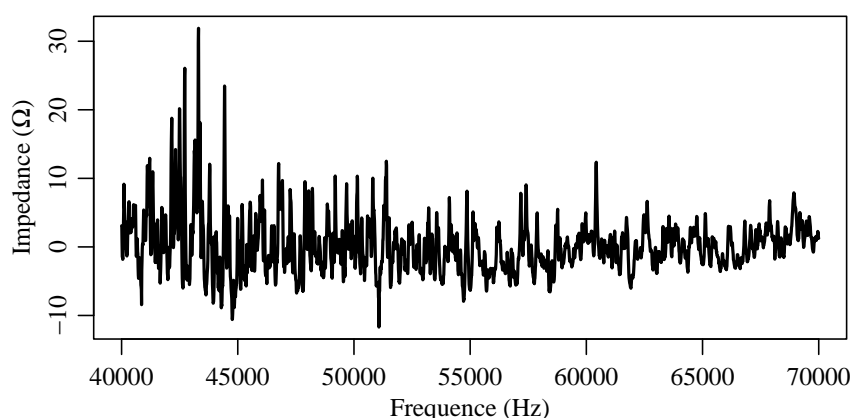


Figure 16: Impedance signature after trend correction.

1.5 Linear Regression With Bootstrapping

The bootstrap resampling method is an alternative to classical parametric inference methodologies. This method calculates the empirical distribution of the statistic of interest, mean, variance, model coefficients, etc., over hundreds or thousands of resamples.

There is no need to worry about assuming a distribution for the statistic in question and, therefore, with the inherent assumptions. The estimates in each resampling allow building the empirical distribution of the distribution of interest.

The central idea of this method is to reproduce several scenarios from the original sample by performing other samples with the same amount of elements but with replacement.

The statistics obtained in each resample are enough to summarize all the information from the mean, variance, and confidence interval, among other possible measures.

Therefore, it is enough to assume that the original sample is reliable and to have enough computational resources to calculate a value of the statistic of interest (θ^*) in each new sample.

Each sample obtained from the original is called a *bootstrap* sample of size n . With the same size as the original sample, from θ^* obtained in each of the hundreds or thousands of samples, it models the distribution of *bootstrap*.

The *bootstrap* method is classified into parametric and non-parametric. In the first, it is necessary to postulate a distribution to generate the samples later; in the second, the samples are generated from the empirical distribution. This chapter will use the non-parametric bootstrap since the necessary assumptions for parametric inference are not always met.

In regression, bootstrapping can be performed by resampling the pairs (dependent and independent variables) and the model residuals. In this chapter, pairwise resampling will be used.

Example 5. Assume that the electromechanical impedance (Ω) is a function of the temperature at a given frequency point (Hz). Consider the following baseline data set on these two variables and estimate the coefficients of the simple linear regression model using the bootstrap technique. Baseline data sets represent the reference state of the specimen and, in this case, are obtained at different temperatures.

Temperature ($^{\circ}C$)	Impedance (Ω)
29,80	123.53
27.25	122.98
25.70	122.74
24.90	122.41
23.90	122.50
27.25	123.87
30.10	124.46
31.10	124.85
28.20	123.89
26.50	123.86
25.20	122.79
24.20	122.33
27.75	123.60
30.00	124.44
26.40	123.34

Solution

The following image illustrates the execution of the R source code to obtain the estimates of the intercept (β_0), the slope (β_1) and the coefficient of determination (R^2), from the simple linear regression model, $y_i = \beta_0 + \beta_1 x_i + e_i$

```

> B <- 5000
> data<-cbind(y,x) # y <- impedance values and x<-temperature
> Tboot <- matrix(NA, nrow = B,ncol = 3) # 2 coef + R2
> for(i in 1:B){
+ u <- sample(1:nrow(data),nrow(data), replace = T)
+ data2 <- data[u,]
+ y1 <- data2[,1]
+ x1 <- data2[,2]
+
+ fit <- lm(y1 ~ x1)
+ Tboot[i,]<-c(summary(fit)$coefficients[, 1],summary(fit)$r.squared)
+ }
> b0 <- mean(Tboot[,1]);b0
[1] 114.8429
> b1 <- mean(Tboot[,2]);b1
[1] 0.3159266
> r2 <- mean(Tboot[,3]);r2
[1] 0.8049488
> b0.IC <- quantile(Tboot[,1], prob = c(0.025, 0.975));b0.IC
      2.5%      97.5%
113.2060 117.0977
> b1.IC <- quantile(Tboot[,2], prob = c(0.025, 0.975));b1.IC
      2.5%      97.5%
0.2310095 0.3760934
> r2.IC <- quantile(Tboot[,3], prob = c(0.025, 0.975));r2.IC
      2.5%      97.5%
0.5242340 0.9654569

```

Figure 17: R source code executed to get bootstrap estimates.

The estimated model is $\hat{y}_i = 114.8316 + 0.3163x_i$ with $R^2 = 80.77\%$ and from the confidence intervals, it is verified that the null hypothesis of the parameters β_0 and β_1 , is rejected since the intervals $b0.IC$ and $b1.IC$ do not include the value zero. The value of R^2 indicates that 80.77% of the change in impedance is explained by the change in temperature.

Figure 18 shows the behavior of impedance as a function of temperature at a given frequency point and the main results obtained from bootstrap estimates.

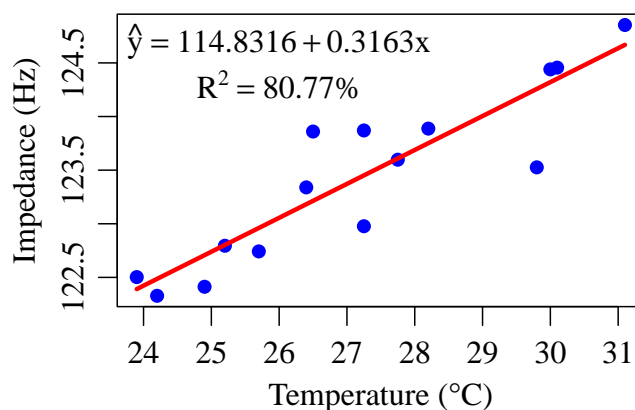


Figure 18: Impedance behavior as a function of temperature.

The bootstrap estimation of the model presented in Eq. 6, for trend correction in the signature, can be obtained similarly to the one shown in the example 5.

Example 6. *Illustrate a case where the assumptions necessary for parametric inference about the example model 4 are not met.*

Solution

```

> fit=lm(log(y)~Freq)
> #summary(fit)
> shapiro.test(residuals(fit))

      Shapiro-Wilk normality test

data:  residuals(fit)
W = 0.97956, p-value < 2.2e-16

> library(lmtest)
> dwtest(modelo)

      Durbin-Watson test

data:  modelo
DW = 0.076243, p-value < 2.2e-16
alternative hypothesis: true autocorrelation is greater than 0

> bptest(modelo)

      studentized Breusch-Pagan test

data:  modelo
BP = 27.069, df = 1, p-value = 1.963e-07

```

Note from the image above that the p-values for the normality (Shapiro-Wilk), independence (Durbin-Watson), and homogeneity of variances (Breusch-Pagan) tests were all less than 0.01. Therefore, the respective tests were significant at 1%, indicating that these assumptions are not met. This finding justifies the use of the bootstrap method for non-parametric inference as applied in the 5 example.

In this chapter, pre-processing of the data set arising from the impedance-based structural health monitoring (SHM) is understood as the investigation and disposal of signatures that are not representative of the actual state of the structure. Matching signatures for trends is also considered a preprocessing technique.

The use of signatures to infer the structure's integrity is classified as post-processing. Therefore, the intermediate step (processing) is defined by using techniques to equate the signatures measured in the baseline acquisition step and the investigation status regarding the effect of temperature. This variable can cause changes in the structure and, if not considered, can lead to erroneous inferences. In the following section, two statistical methodologies can be used for temperature compensation or modeling baseline signatures at a temperature of interest.

2 Data processing

In a sufficiently large number of frequencies, modeling impedance as a function of temperature allows for building reference signatures at temperatures not necessarily observed in the baseline acquisition step. This methodology is called temperature compensation. To apply this methodology, the baseline signatures must be collected in a sufficiently large range of temperatures. In other words, the sampling of temperatures is representative in quantity and amplitude.

The following subsection presents the methodology of polynomial regression, a generalization of simple linear regression, which can be shown in a matrix form as specified in Eq. 11.

2.1 Polynomial regression

The behavior of a dependent variable can be modeled as a function of a single regressor variable by a polynomial model of any order p (Eq. 9). Note that if $p = 1$, we have the simple linear regression model as shown in Eq. 7.

The model shown in Eq. 9 can be written in matrix form as in Eq. 11 and therefore, the vector, least squares estimator for β , is as shown in Eq. 16.

In polynomial regression, the model matrix, \mathbf{X} , is composed of columns that are powers of the same variable. Building this matrix implies increasing the correlation or dependence between the columns.

There is multicollinearity when a regressor variable is dependent on another or others. This means that, at least in part, an independent (regressive) variable can be explained by the others. Multicollinearity is, therefore, not a matter of presence or absence but a matter of intensity. This intensity can be quantified by the coefficient of determination, R_k^2 , obtained from the adjusted model for k -th independent variable as a function of the others. The closer to 1 (100%) the R_k^2 is, the more serious the multicollinearity problem.

The variance inflation factor over the coefficient estimators, VIF_k , is used to measure multicollinearity [Belsley et al., 2005].

$$VIF_k = \frac{1}{1 - R_k^2} \quad (30)$$

Multicollinearity can cause problems such as a high value of the coefficient of determination, non-significance of the parameters (they do not differ significantly from zero), and unexpected (negative/positive) signs for the parameter estimates. Perfect multicollinearity prevents even obtaining least squares estimates since the columns of the matrix \mathbf{X} , in Eq. 11, will be linearly dependent. Therefore, the classic inverse of the matrix $\mathbf{X}'\mathbf{X}$ will not exist. This inverse is necessary for the estimation of least squares.

According to Rawlings et al. [1998], the VIF must be below 10 for the multicollinearity problem to be considered unimportant.

A corrective measure is to use data centered, that is, to replace x with $x^* = x - \bar{x}$. Using orthogonal polynomials is a more efficient alternative. Therefore, it should be preferred when this methodology is applicable.

The study of the model's adequacy in meeting the assumptions of normality, independence, and homogeneity of residual variance, as well as the presentation in the 1.4 section, should also be done for polynomial regression. The significance of the model to explain the response variable and the proportion of explained variation are items that must also be investigated in the polynomial regression. A desirable model is one in which the assumptions are met with relatively high and significant R^2 .

Example 7. Consider the data set on the real part of the electromechanical impedance (Ω) obtained at different temperatures ($^{\circ}C$) at a given frequency (Hz). These data come from an experiment developed at the Structural Mechanics Laboratory (LMEst) of the Faculty of Mechanical Engineering (FEMEC) of the Federal University of Uberlandia (UFU).

Temperature ($^{\circ}C$)	Impedance (Ω)
29.80	631.59
27.25	748.09
25.70	775.73
24.90	775.34
23.90	761.51
27.25	747.76
30.10	605.71
31.10	558.08
28.20	700.31
26.50	767.44
25.20	785.90
24.20	780.12
27.75	719.67
30.00	609.17
26.40	760.08

Fitting the degree 2 polynomial model and checking the multicollinearity problem. In the affirmative case, adopt a strategy to mitigate it. Also, verify that the assumptions of normality, independence, and residual homogeneity are met. Perform all the requested analyses with the help of the R software.

Solution

The following R source code was used to get the questions' required results. In this code, y represents the impedance values and x the respective temperatures.

```

> fit = lm(y~x+I(x^2))
> library(faraway)
> vif(fit)
      x      I(x^2)
736.5912 736.5912
> z=x-mean(x)
> fit2 = lm(y~z+I(z^2))
> vif(fit2)
      z      I(z^2)
1.052605 1.052605
> r=residuals(fit2)
> shapiro.test(r)

      Shapiro-Wilk normality test

data:  r
W = 0.95687, p-value = 0.6382

> require(lmtest)
> dwtest(fit2)

      Durbin-Watson test

data:  fit2
DW = 2.4274, p-value = 0.7653
alternative hypothesis: true autocorrelation is greater than 0

> bptest(fit2)

      studentized Breusch-Pagan test

data:  fit2
BP = 5.5668, df = 2, p-value = 0.06183

```

Figure 19: R source code and output solution for polynomial regression analysis.

The VIF values indicate a severe multicollinearity problem and are the same for the two variables considered. The second is a power of the first. The VIF values' equality is obtained from the regressions of x as a function of x^2 and x^2 as a function of x , which are simple linear regressions. In this case, the R^2 is the square of the Pearson correlation that does not distinguish between the correlations of x and x^2 and between x^2 and x .

On average, the centralization of the explanatory variable x corrected the multicollinearity problem ($VIF < 10$), and the model adjusted with the new variable, z met the assumptions of normality, independence, and homogeneity according to the Shapiro-Wilk tests, Durbin-Watson and Breusch-Pagan, respectively. The p -values in these tests were greater than 0.05.

The image below shows that the model significantly explains the variation in impedance, according to the F test of the analysis of variance (p -value < 0.01). The percentage of explained variation is high ($R^2 = 98.98\%$), and all coefficients are significantly different from zero according to the t -test. (p -value < 0.01).

```

> summary(fit2)

Call:
lm(formula = y ~ z + I(z^2))

Residuals:
    Min       1Q   Median       3Q      Max
-11.8966  -6.7773   0.8046   5.8502  13.1243

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  743.0317     3.1509   235.81 < 2e-16 ***
z           -28.4358     1.0004   -28.42 2.23e-12 ***
I(z^2)       -5.8123     0.4817   -12.06 4.55e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 8.279 on 12 degrees of freedom
Multiple R-squared:  0.9898,    Adjusted R-squared:  0.9881
F-statistic: 582.5 on 2 and 12 DF,  p-value: 1.123e-12

```

Figure 20: R source code and output solution for polynomial regression analysis.

The scatter plot with the fitted model are shown in Figure 21.

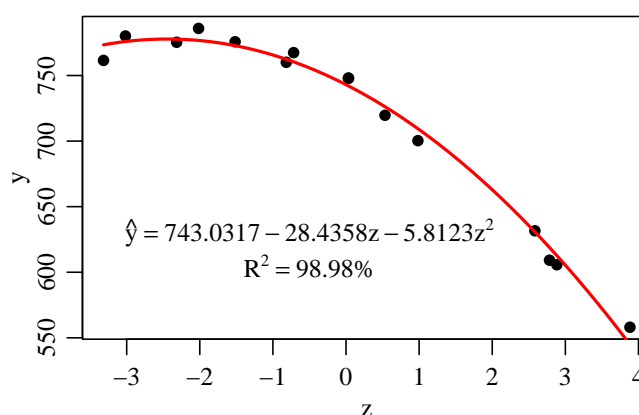


Figure 21: Polynomial regression for the impedance (y) as a function of the temperature centered on the mean ($z = x - \bar{x}$).

The typical characteristic dependence of data set in SHM [Schubert Kabban et al., 2015] motivates the use of statistical methodologies that can consider this autocorrelation in data analysis, for example, in temperature compensation via regression models. Regression models as ARIMA errors is an existing proposal in the literature [Hyndman and Athanasopoulos, 2018] for such an analysis. A summary of this methodology, including the R source codes necessary to obtain results, is presented in the following section.

2.2 Regression with ARIMA errors

The dependence between the impedance values motivates the use of methodologies that can accommodate this characteristic reflected in the residuals of the regression model. The successive nature of the acquisitions carried out by the collection system may justify this dependence.

Regression models with ARIMA errors (Auto-Regressive Integrated Moving Averages Model) are flexible in terms of the residual autocorrelation structure in the sense that it is possible to adopt one among several possibilities (order) for the ARIMA model. The function *auto.arima* of the package *forecast* [Hyndman et al., 2020] of the R software [R Core Team, 2021] is a very useful computational resource that can help in the adjustment and in choosing the most adequate model for each case. This resource will be used in this section to illustrate how to obtain the results in the exemplification analyses.

A polynomial regression with ARIMA errors can be written according to the equation presented in Eq. 31,

$$y_t = \beta_0 + \beta_1 x_t + \beta_2 x_t^2 + \cdots + \beta_k x_t^k + n_t \quad (31)$$

em que $\phi(B)(1 - B)^d n_t = \theta(B)e_t$.

In Eq. 31, n_t are the errors or a variable resulting from an AR (Auto-Regressive), MA (Moving Averages), ARMA (Auto-Regressive Moving Averages), and ARIMA (Auto-Regressive Integrated Moving Averages) processes, among others; e_t is called soft noise, that is, a random variable with zero mean and variance σ_e^2 and B is called a delay operator. The index t in Eq. 31 is purposely used to denote the temporal dependence of data acquisition.

The process n_t is said to be moving averages of order q , or $MA(q)$, if,

$$n_t = e_t + \delta_1 e_{t-1} + \delta_2 e_{t-2} + \cdots + \delta_q e_{t-q}, \quad (32)$$

Using the delay operator, Eq. 32 can be rewritten as

$$n_t = (1 + \delta_1 B + \delta_2 B^2 + \cdots + \delta_q B^q) e_t = \theta(B) e_t, \quad (33)$$

On the other hand, if

$$n_t = \alpha_1 n_{t-1} + \alpha_2 e_{t-2} + \cdots + \alpha_p e_{t-p} + e_t \quad (34)$$

then n_t is called an autoregressive process of order p , or $AR(p)$. Using the delay operator, Eq. 34 is:

$$(1 - \alpha_1 B - \alpha_2 B^2 - \cdots - \alpha_p B^p) n_t = e_t \text{ ou } \phi(B) n_t = e_t \quad (35)$$

and therefore the $AR(p)$ process can be rewritten as:

$$n_t = \phi(B)^{-1} e_t = \psi(B) e_t \quad (36)$$

Another possibility for time-dependent data modeling is the autoregressive moving average models, $ARMA(p, q)$. The expression for a $ARMA(p, q)$ model is:

$$n_t = \alpha_1 n_{t-1} + \alpha_2 n_{t-2} + \cdots + \alpha_p n_{t-p} + e_t + \delta_1 e_{t-1} + \delta_2 e_{t-2} + \cdots + \delta_q e_{t-q} \quad (37)$$

In general, representations through models as described in Eq. 37 requires a smaller number of parameters when compared to representations through models as in Eq. 32 or in Eq. 34. Using the delay operator, the model in Eq. 37 is as in Eq. 38:

$$(1 - \alpha_1 B - \alpha_2 B^2 - \dots - \alpha_p B^p)n_t = (1 + \delta_1 B + \delta_2 B^2 + \dots + \delta_q B^q)e_t \quad (38)$$

or

$$\phi(B)n_t = \theta(B)e_t \quad (39)$$

The models described in Eq. 32, 34 and 37 are suitable for stationary time series. The trend is a characteristic that, if present, characterizes a lack of stationarity and can be removed by taking one or more differences between consecutive values of the series.

An ARMA model, where n_t is replaced by its d -th difference $\Delta^d n_t$, can accommodate some kind of non-stationarity and the differentiated series will be denoted by the expression as shown in Eq. 40.

$$W_t = \Delta^d n_t = (1 - B)^d n_t \quad (40)$$

and the integrated autoregressive process of moving averages, $ARIMA(p, d, q)$ is written according to equation Eq. 41 or Eq. 42,

$$W_t = \alpha_1 W_{t-1} + \alpha_2 W_{t-2} + \dots + \alpha_p W_{t-p} + e_t + \delta_1 e_{t-1} + \delta_2 e_{t-2} + \dots + \delta_q e_{t-q} \quad (41)$$

or, equivalently according to Eq. 42:

$$\phi(B)(1 - B)^d n_t = \theta(B)e_t. \quad (42)$$

After establishing the model, either as defined in Eq. 31, 32, 34, 37 or in Eq. 41 the next step is to estimate its parameters. Considering a polynomial regression model with ARIMA errors, the number of parameters will be $k+1+p+q+1+1$ where $k+1$ (Eq. 31) is the number of coefficients of the polynomial model, $p+q+1$ the number of parameters associated with the ARIMA model (Eq. 41) and 1 more parameter to be estimated is the variance associated with white noise.

Making $\beta = (\beta_0, \beta_1, \dots, \beta_k)$, $\alpha = (\alpha_0, \alpha_1, \dots, \alpha_p)$, $\delta = (\delta_0, \delta_1, \dots, \delta_q)$ and considering $\xi = (\beta, \alpha, \delta, \sigma_e^2)$, the likelihood function for ξ will be $L(\xi|D)$, where D represents the sample data. The maximum likelihood estimators (MLE) of ξ will be the values that maximize L or $l = \log(L)$.

After the adjustment, the study of the model's suitability must be carried out. One possibility is to insert more parameters, obtaining a new model. The significance of the parameter(s) inserted allows verifying whether the more complex model improves the fit of the simpler model.

Among the criteria for selecting models, and in particular the regression models with ARIMA errors, the criteria based on the maximum likelihood function are the most used, most often the Akaike Information Criterion (AIC) [Akaike, 1974], the Akaike Information Corrected Criterion (AICc) [Bozdogan, 1987] and the Bayesian Schwarz Criterion (BIC) [Schwarz, 1978].

The adequacy of the models adjusted via the maximum likelihood method can be quantified using pseudo coefficients of determination and, in the case of time series models, also using the precision of the forecasts.

The adequacy of the model must also be done through the analysis of residuals. If the model is adequate, it is expected that the residuals or white noise are normally distributed, randomly around zero, with constant variance and uncorrelated.

In time series analysis, he traditionally uses the Lilliefors [Thode, 2002a] test to verify the hypothesis of normality, the Box-Ljung [Ljung and Box, 1978] or Breusch-Godfrey [Breusch, 1978, Godfrey, 1978] tests to confirm the independence hypothesis and the Breusch-Pagan test [Breusch and Pagan, 1979] to verify the homogeneity of variance.

Example 8. Consider the data in the table below from an experiment carried out on the Gloria campus from the Federal University of Uberlandia (UFU).

Table 3: Real part of electromechanical impedance signature (Ω) at 36 temperatures ($^{\circ}C$).

Temperature	Impedance	Temperature	Impedance
18.43	84.36	17.14	88.60
17.95	85.66	18.43	86.09
17.95	86.28	18.27	87.26
19.08	83.71	17.63	88.80
17.14	90.26	17.95	87.73
16.66	91.14	19.24	84.06
16.66	91.28	18.92	84.59
17.47	90.11	17.95	87.32
17.31	90.42	18.27	87.34
16.82	90.72	19.72	84.85
15.86	95.33	19.56	85.10
17.14	94.28	18.92	86.46
16.02	95.76	19.24	85.98
15.53	97.28	20.04	84.95
16.02	96.16	19.56	85.58
17.31	90.72	18.59	87.42
17.14	89.94	18.92	87.01
16.50	88.20	18.76	87.04

a) Fit the simple linear regression model for the impedance as a function of temperature and show that the assumptions of normality and homogeneity are met. Still, the same is not valid for the assumption of independence.

Solution

The model was fitted with y representing the impedance values and x the respective temperatures. In the following image, it can be seen that the p -values for the Lilliefors test (`lillie.test`) and the Breusch-Pagan test (`bptest`) were greater than 0.05, and therefore the hypotheses of normality and homogeneity are not rejected at 5% significance. On the other hand, in the Box-Ljung test (`checkresiduals(fit2,test="LB")`), a p -value lower than 0.05 was obtained, and, therefore, the independence assumption is not met.

The two measures R^2 and mean squared error (mean squared error - mse) were calculated and are helpful for decision making regarding the model's adequacy and for comparing it with another proposed model. The values of these measures were respectively 79.21% and 2.6889.


```

> if(!require(nortest)) install.packages('nortest')
> if(!require(forecast)) install.packages('forecast')
> if(!require(Metrics)) install.packages('Metrics')
> fit1=lm(y~x)
> lillie.test(residuals(fit1))

      Lilliefors (Kolmogorov-Smirnov) normality test

data:  residuals(fit1)
D = 0.0745, p-value = 0.8813

> checkresiduals(fit1,test="LB")

      Ljung-Box test

data:  Residuals
Q* = 33.277, df = 5, p-value = 3.315e-06

Model df: 2.   Total lags used: 7

> bptest(fit1)$p.value
      BP
0.1281203
> cor(y,fitted(fit1))^2
[1] 0.7920986
> mse(y,fitted(fit1))
[1] 2.688923

```

Figure 22: R source code and output solution for fitting the simple linear regression model.

b) Fit the simple linear regression model with an ARIMA structure for the residuals. This can be done with the function `auto.arima` of the package `forecast` of R language. Show that the assumptions are met in this case; therefore, the parametric inference from this model is valid.

Solution

With the function `auto.arima`, the autoregressive structure of order one was selected for the residuals, $ARIMA(1,0,0)$ and the assumptions of normality (`lillie.test`), of independence (`checkresiduals(fit2,test="LB")`) and homogeneity (`bptest`) were met because the p -values of the respective tests were greater than 0.05. The value of the pseudo R^2 was 91.60%, and the month was equal to 1.0969.

Then, from the evaluation of the assumptions, the pseudo R^2 , and the mse, it is concluded that the regression model with $AR(1)$ structure for the residuals is more suitable than the simple linear regression model.

```

> fit2=auto.arima(y,xreg=cbind(x))
> summary(fit2)
Series: y
Regression with ARIMA(1,0,0) errors

Coefficients:
      ar1  intercept      x
    0.7993  130.0195 -2.3301
s.e.  0.1022    4.3660  0.2369

sigma^2 = 1.197: log likelihood = -53.26
AIC=114.51  AICc=115.8  BIC=120.84

Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set 0.09035373 1.047317 0.7300343 0.0870417 0.8156447 0.4754428 0.1086486
> lillie.test(residuals(fit2))

      Lilliefors (Kolmogorov-Smirnov) normality test

data: residuals(fit2)
D = 0.1389, p-value = 0.07666

> checkresiduals(fit2,test="LB")

      Ljung-Box test

data: Residuals from Regression with ARIMA(1,0,0) errors
Q* = 3.7959, df = 4, p-value = 0.4343

Model df: 3. Total lags used: 7

> bptest(residuals(fit2) ~ fitted(fit2))$p.value
      BP
0.06984022
> cor(y,fitted(fit2))^2
[1] 0.915992
> mse(y,fitted(fit2))
[1] 1.096872

```

Figure 23: R source code and output solution for fitting the simple linear regression model with ARIMA structure to residuals.

Polynomial models with ARIMA structure for residuals (example 8) appear as an excellent alternative to traditional polynomial models (example 7) when they do not satisfy the necessary assumptions for inference under the parametric optics. These models can predict electromechanical impedance at temperatures not necessarily observed in the baseline data collection step but in the structure investigation steps. The recurrent use of these models at different frequencies (H_z) allows for modeling or predicting a baseline signature.

In this chapter, the processing is understood as the matching of a signature in the investigation of the structure with its respective reference signature regarding the effect of the temperature variable. The post-processing stage of statistical analysis, the inference about the structure's status, is called post-processing. Some possible statistical techniques at this stage are presented in the 3 section below.

3 Data post-processing

3.1 Comparison tests on damage metrics

3.1.1 Hypothesis test on more than two means

The F test, in the analysis of variance, is used to verify whether the null hypothesis, defined as the equality of groups or populations, on average, should be rejected or not. The set of possibilities complementary to the null hypothesis represents the alternative hypothesis. The expression “at least one population differs from the others on average” clearly illustrates the alternative hypothesis. In this case, it is also assumed that the groups are independent.

One-way analysis of variance consists of breaking down the total variance of the data or the response variable into variance between and within factor levels. There is a statistical difference between them, on average. Otherwise, the evidence will be weak, and the different hypotheses will not be supported.

The test statistic used in the analysis of variance is given by:

$$F = \frac{SS_{between}}{\frac{J-1}{SS_{within}}} \quad (43)$$

where $SS_{Between}$ is the sum of squares between levels, SS_{Within} is the sum of squares within levels, J is the number of levels or groups, and I is the number of values or repetitions in each group. Assuming that the assumptions of homogeneity of variance and normality of the residuals are met, the variable defined in (43) follows a distribution F of *Snedecor* with $J - 1$ and $J(I - 1)$ degrees of freedom. When the number of repetitions is not the same in all groups, the amount $J(I - 1)$ must be calculated considering this fact by making $\sum_{j=1}^J (I_j - 1)$ in where I_j is the number of repetitions at each factor level.

The hypothesis test (test F) of the analysis of variance is a right-sided test because the null hypothesis will only be rejected if the observed value of F is located at the extreme right of the F distribution of *Snedecor*. The decision rule or critical value is determined from the F distribution with $J - 1$ and $J(I - 1)$ degrees of freedom and a given significance; if the observed value of F is greater than the critical value, it is said that it is strong enough evidence to reject the hypothesis of equality between the means, that is, the null hypothesis, and it is concluded that at least one mean differs from the others for the established significance.

The analysis of variance is linked to a model that can be defined as follows:

$$y_{ij} = \mu + f_j + e_{ij} \quad (44)$$

y_{ij} is the i -th observation of the response variable at the j -th level of the factor f , μ is constant (generally the average of all observations), and e_{ij} is the error associated with the i -th repetition of the j -th level. For the inferences to be valid, it is necessary that the assumption of normality, independence, and homogeneity of the residuals (constant variance) be valid, in notation $e_{ij} \sim N(0, \sigma^2)$. These assumptions can be verified using the Shapiro-Wilk tests [ROYSTON, 1982] and the Bartlett test [BARTLETT, 1937]. The validity of Bartlett's test is dependent on the normality hypothesis. When this hypothesis is not met, the Levene test [LEVENE, 1960] is an alternative that only requires the existence of any continuous distribution for the variable in question. For the independence of the residuals, we have the Durbin-Watson test [DURBIN and WATSON, 1950].

The hypotheses involved in the analysis of variance do not allow us to verify which levels or groups differ from each other. They only allow us to reject or not the hypothesis of equality of all

means. Therefore, in situations where the decision is for at least one mean different from the others (rejecting the null hypothesis), there is a need for an additional test for two-by-two comparisons or to contrast groups of means. The Student-Newman-Keuls (SNK) [NEWMAN, 1939] and Tukey tests [TUKEY, 1953], among others, are indicated for two-by-two comparisons. The Scheffé Test [SCHEFFÉ, 1953] is flexible as it allows comparing any contrast between means. The two groups (contrasts) involved can have different amounts of means and each mean can come from a different number of repetitions or observations.

Example 9. Consider the following data about the CCD metric under baseline conditions (B), and damage 1, 2, and 3 (D1, D2, and D3). These data come from electromechanical impedance signals obtained from a steel plate from the petrochemical industry.

Table 4: CCD by condition and repetition.

Condition	Replications									
	1	2	3	4	5	6	7	8	9	10
B	0.0234	0.0305	0.0393	0.0474	0.0385	0.0208	0.0243	0.0408	0.0342	0.0206
D1	0.1739	0.1577	0.1840	0.1894	0.1586	0.1999	0.1856	0.1579	0.1623	0.1654
D2	0.2767	0.2659	0.2554	0.2564	0.2718	0.2761	0.2957	0.2546	0.2746	0.2875
D3	0.3500	0.3749	0.3570	0.3693	0.3841	0.3810	0.3624	0.3786	0.3517	0.3637

- a) Check the model assumptions, normality of residuals, homogeneity of variances, and independence of residuals at 5% of significance, using the Shapiro-Wilk, Bartlett, and Durbin-Watson tests.

Solution

```
> shapiro.test(residuals(lm(dados1~danos),data=dados2))
      shapiro-wilk normality test
data:  residuals(lm(dados1 ~ danos), data = dados2)
w = 0.94708, p-value = 0.06021

> bartlett.test(dados1~danos,data = dados2)
      bartlett test of homogeneity of variances
data:  dados1 by danos
Bartlett's K-squared = 2.0527, df = 3, p-value = 0.5616

> durbinwatsonTest(lm(dados1~danos),data=dados2)
lag Autocorrelation D-W Statistic p-value
1      -0.1033981      2.192324      0.91
Alternative hypothesis: rho != 0
```

Figure 24: R source code for verification of assumptions.

From the results presented in Figure 24 there is no significant evidence to reject the hypotheses of normality ($p\text{-value} = 0.06021$), of homogeneity of variances ($p\text{-value} = 0.5616$) and of independence of residuals ($p\text{-value} = 0.91$). Therefore, the assumptions about the residuals of the model are satisfied.

- b) Do the analysis of variance for the model, conclude on the F test and apply Tukey's test for comparison of means, if applicable.

Solution

```

> dic(danos,dados1, quali = TRUE,mcomp = "tukey", sigT = 0.05, sigF = 0.05)
-----
quadro da analise de variancia
-----
              GL      SQ      QM      FC      Pr>Fc
Tratamento  3 0.61534 0.205112 1237.4 2.3526e-36
Residuo      36 0.00597 0.000166
Total       39 0.62130
-----
CV = 6.1 %

-----
Teste de normalidade dos residuos
Valor-p: 0.06020767
De acordo com o teste de Shapiro-wilk a 5% de significancia, os residuos podem ser considerados normais.
-----

Teste de homogeneidade de variancia
valor-p: 0.5615511
De acordo com o teste de bartlett a 5% de significancia, as variancias podem ser consideradas homogeneas.
-----

Teste de Tukey
-----
Grupos Tratamentos Medias
a      dano3      0.36727
b      dano2      0.27147
c      dano1      0.17347
d      danob      0.03198
-----

```

Figure 25: R source code for analysis of variance and multiple comparison.

The result of the F test presented in Figure 25 allows us to conclude that there is significant evidence to reject the hypothesis of equality of means (p -value < 0.05), and at least one differs from the others. Using the Tukey test, all experimental conditions vary, two by two, regarding the average CCD metric. Thus, the methodology effectively-identified structure changes since all damages differed from baseline.

3.1.2 Wilcoxon-Mann-Whitney Test

This test is an alternative to the t test when the normality assumption is not met in at least one of the populations involved. Its application can be made even when there is no information about the distributions, but the variables involved have a measurement scale that is at least ordinal. The two samples must be independent. The interest in applying this test is to verify if a population tends to have higher values than the other or if they have the same median.

The Wilcoxon-Mann-Whitney test is based on the ranks of the values obtained by combining the two samples. The values are ordered from smallest to largest, regardless of which population each value comes from. Let R_1 be the sum of ranks in sample 1 and R_2 be the sum of ranks in sample 2, then define the test statistic as follows:

$$W_1 = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1; W_2 = n_1 n_2 + \frac{n_2(n_2 + 1)}{2} - R_2 \quad (45)$$

where n_1 and n_2 are respectively the sizes of samples 1 and 2. Make W equal to the smallest value between W_1 and W_2 as the value of the test statistic. For values of $n < 20$, the Mann-Whitney U table is used to find the critical values necessary for the decision to reject the null hypothesis. The normal distribution can be approximated if n is greater than 20.

Let R be the sum of the ranks in sample 1:

$$z = \frac{R - \mu_R}{\sigma_R} \quad (46)$$

where $\mu_R = \frac{n_1(n_1+n_2+1)}{2}$, $\sigma_R = \sqrt{\frac{n_1n_2(n_1+n_2+1)}{12}}$, n_1 and n_2 are respectively the sizes of samples 1 and 2. The variable defined in Eq. 49 has an approximately standard normal distribution and therefore, the critical values needed for decision making are determined from this distribution.

Example 10. *Considering the following data set on the CCD metric (baseline and damage), calculated from electromechanical impedance signatures obtained from steel sheet from the petrochemical industry.*

Table 5: CCD metric for damage and repetition.

Condition	Replications									
	1	2	3	4	5	6	7	8	9	10
Baseline	0.0054	0.0469	0.0447	0.0145	0.0162	0.0197	0.0195	0.0048	0.0261	0.0066
Damage	0.1255	0.1051	0.2240	0.4486	0.1718	0.1134	0.1199	0.1064	0.4540	0.3349

a) *Check normality assumptions for populations at 5% significance.*

Solution

```
> shapiro.test(ccdbase)
      shapiro-wilk normality test
data:  ccdbase
W = 0.867, p-value = 0.09221

> shapiro.test(ccddano)
      shapiro-wilk normality test
data:  ccddano
W = 0.78514, p-value = 0.009561
```

Figure 26: R source code and its output solution for the normality test.

According to the results (Figure 26), there is no significant evidence to reject the hypothesis of normality (p -value = 0.09221) for the baseline population. Still, the same does not occur for the damaged population (value $-p = 0.0009561$); Still, a non-parametric alternative, such as the Mann-Whitney test, should be used to compare the two conditions.

b) *Using the Mann-Whitney test, comparing baseline and damage CCD metric populations at 5% significance.*

Solution

$$R_1 = 55; W = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1 = 100 + \frac{10(11)}{2} - 55 = 100 \quad (47)$$

$$R_2 = 155; W = n_1 n_2 + \frac{n_2(n_2 + 1)}{2} - R_2 = 100 + \frac{10(11)}{2} - 155 = 0 \quad (48)$$

Thus, the smallest value for W is zero ($W = 0$). The value of the Mann-Whitney table with $\alpha = 0.05$ and $n_1 = n_2 = 10$ is 23, so as $W < 23$ the null hypothesis is rejected.

Another alternative to solve this question is through the function `wilcox.test`, in the R language, as shown in Figure 27.

```
> wilcox.test(ccdbase, ccddano, correct=FALSE, alternative="two.sided")
      wilcoxon rank sum test

data:  ccdbase and ccddano
w = 0, p-value = 1.083e-05
alternative hypothesis: true location shift is not equal to 0

> median(ccdbase)
[1] 0.01785
> median(ccddano)
[1] 0.14865
> |
```

Figure 27: R source code and output solution for comparison of baseline and damaged groups.

According to the results presented in Figure 27, there is significant evidence to reject the hypothesis that the populations are centered in the same position (p -value < 0.05), that is, the two populations differ, with the median damaged CCD (0.14865) higher than baseline (0.01785).

3.1.3 Kruskal-Wallis Test

The Kruskal-Wallis test is an alternative to conventional analysis of variance when the assumptions of normality and homogeneity of variances are not met.

The procedure for implementing the test consists of sorting the data set from all samples or (K) levels as a single sample and assigning ranks to all values. For each sample or level, determine the sum of ranks R_k with $k = 1, 2, 3, \dots, K$. The following test statistic follows a *chi-square* distribution with $K - 1$ degrees of freedom.

$$H = \frac{12}{N(N + 1)} \sum_{k=1}^K \left(\frac{R_k^2}{n_k} \right) - 3(N + 1) \quad (49)$$

In Eq. 49, N is the total number of observations and n_k is the size or number of observations at the factor level k . If the ranks are similarly distributed among the sample groups, then H will be a relatively small number. The Kruskal-Wallis test is right-sided; that is, the hypothesis of equality between the groups will be rejected if the test statistic, H , is a value located at the extreme right of the *chi-square distribution*. Based on the significance of the test and the degrees of freedom, the critical value for decision-making is determined.

If the necessary assumptions are met, the use of analysis of variance is preferable since, in general, parametric tests are more powerful than non-parametric tests.

Example 11. Let the following values from the baseline CCD metric (*B*) and damage 1, 2, and 3 (*D1*, *D2*, and *D3*) conditions, be calculated from electromechanical impedance signatures for steel sheet from the petrochemical industry.

Table 6: CCD for damage and repetitions

Condition	Replications									
	1	2	3	4	5	6	7	8	9	10
B	0.0383	0.0364	0.0480	0.0385	0.0232	0.0387	0.0434	0.0476	0.0213	0.0496
D1	0.1804	0.1797	0.1792	0.1698	0.1663	0.1926	0.1836	0.1719	0.1995	0.1660
D2	0.2540	0.2671	0.2841	0.2601	0.2508	0.2765	0.2786	0.2539	0.2924	0.2727
D3	0.3534	0.1143	0.4307	0.4080	0.2019	0.2710	0.1899	0.3106	0.3743	0.4862

- a) Show that the assumptions of normality of residuals, homogeneity of variances, and independence of residuals, at 5% significance, are not met. Use the Shapiro-Wilk, Bartlett, and Durbin-Watson tests implemented in R language.

Solution

```
> shapiro.test(residuals(lm(dados1~danos),data=dados2))
      Shapiro-Wilk normality test
data:  residuals(lm(dados1 ~ danos), data = dados2)
W = 0.79106, p-value = 4.434e-06

> bartlett.test(dados1~danos,data = dados2)
      Bartlett test of homogeneity of variances
data:  dados1 by danos
Bartlett's K-squared = 73.809, df = 3, p-value = 6.522e-16

> durbinwatsonTest(lm(dados1~danos),data=dados2)
lag Autocorrelation D-W Statistic p-value
1 -0.08284929 1.939 0.462
Alternative hypothesis: rho != 0
```

Figure 28: Normality, homogeneity and independence tests in R language.

From the results presented in Figure 28, it appears that there is significant evidence to reject both the hypothesis of normality and homogeneity of residues (p -value < 0.05). The independence hypothesis is not rejected (p -value = 0.462). Being of interest, the groups must be compared using non-parametric tests.

- b) Comparing the groups using the Kruskal-Wallis test, at 5% significance.

Solution


```

> comparison<-kruskal(dados1,danos,console=TRUE,alpha=0.05,group=TRUE)

Study: dados1 ~ danos
Kruskal-wallis test's
Ties or no Ties

Critical value: 30.56927
Degrees of freedom: 3
Pvalue Chisq : 1.047413e-06

danos, means of the ranks

      dados1  r
dano1  16.7  10
dano2  29.0  10
dano3  30.8  10
danob   5.5  10

Post Hoc Analysis

t-Student: 2.028094
Alpha    : 0.05
Minimum Significant Difference: 5.131162

Treatments with the same letter are not significantly different.

      dados1 groups
dano3  30.8      a
dano2  29.0      a
dano1  16.7      b
danob   5.5      c

```

Figure 29: R source code and solution for the Kruskal-Wallis test application.

From the results presented in Figure 29, it can be concluded that there is significant evidence to reject the hypothesis that all groups are equal (p -value < 0.05). The Kruskal-Wallis test is a global test, and an additional examination is needed for specific comparisons such as two-by-two comparisons. For this, the modified t test can be used. This test verified that all experimental conditions present significantly different CCD metrics (Figure 30) and ascend from baseline condition to damage 3.

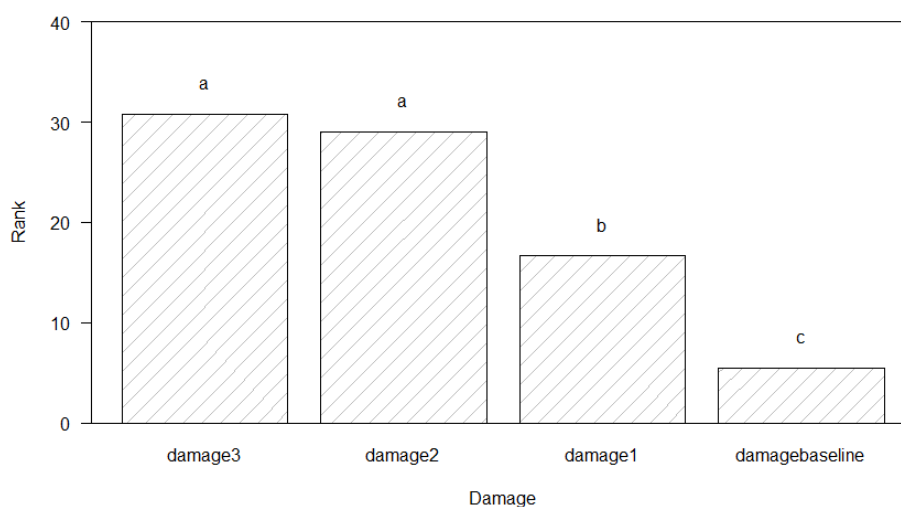


Figure 30: Rankings and Kruskal-Wallis test results for the four experimental conditions.

3.2 Threshold definition

The decision on a significant change in the structure, based on some damage metric such as the CCD metric, for example, is not a deterministic issue, given the inherent variations in data generation. In this way, there will be variation between the CCD values calculated using impedance signatures collected in the baseline measurements and between the CCD values calculated at times of damage investigation even under the same experimental conditions.

These uncertainties must be considered in the decision-making regarding the change in the structure of the baseline measurement step.

The procedure presented in this section consists of establishing a threshold or threshold from the confidence intervals for the mean and standard deviation. The interval estimates for these two parameters include the error associated with the forecast of each one of them.

The threshold is defined by adding to the upper limit of the confidence interval for the mean three times the upper limit of the confidence interval for the standard deviation. A CCD value greater than the threshold will indicate a change in the structure of the pristine condition. On the other hand, a CCD value less than the threshold will suggest that the structure has not changed, at least expressively, when compared to the reference status.

3.2.1 Parametric Threshold

The parametric intervals with $1 - \alpha$ confidence for the mean and the variance of the metric are obtained from Eq. (50) and (51). It should be noted that these intervals are obtained from the baseline data. The denomination parametric intervals is due to the fact that normal distribution is assumed for the sample mean and *chi-square* for the variable $\chi^2 = \frac{(n-1)S^2}{\sigma^2}$, from which the confidence interval for σ^2 is established.

$$\left[\bar{X} - t_{(v, \frac{\alpha}{2})} \frac{S}{\sqrt{n}}; \bar{X} + t_{(v, \frac{\alpha}{2})} \frac{S}{\sqrt{n}} \right] \quad (50)$$

$$\left[\frac{vS^2}{\chi_{(v, \frac{\alpha}{2})}^2}; \frac{vS^2}{\chi_{(v, 1-\frac{\alpha}{2})}^2} \right] \quad (51)$$

where, \bar{X} and S^2 are respectively the mean and variance calculated from the metric values obtained in the pristine condition, n is the sample size, $v = n - 1$ are the degrees of freedom, $t_{(v, \alpha)}$ is a quantile of the *t* distribution of *Student* with v degrees of freedom and probability $\frac{\alpha}{2}$ accumulated to its right, $\chi_{(v, \frac{\alpha}{2})}^2$ and $\chi_{(v, 1-\frac{\alpha}{2})}^2$ are quantiles of the distribution of *Chi-Square* with probabilities $\frac{\alpha}{2}$ and $1 - \frac{\alpha}{2}$ rolled to the right, respectively. The confidence interval for the standard deviation is determined by calculating the square root of the confidence interval limits for the variance Eq. (51), as shown in the equation Eq. (52).

$$\left[\sqrt{\frac{vS^2}{\chi_{(v, \frac{\alpha}{2})}^2}}; \sqrt{\frac{vS^2}{\chi_{(v, 1-\frac{\alpha}{2})}^2}} \right] \quad (52)$$

The threshold or threshold for the decision rule is then determined through Eq. (53)

$$Threshold = \mu_{\max} + 3\sigma_{\max} \quad (53)$$

where μ_{\max} and σ_{\max} are respectively the upper limits for the confidence intervals for the mean and standard deviation built from the baseline data set.

The inference from the interval estimates calculated through Eq. (50) and (52) is valid if the normality hypothesis about the variable in question, the CCD metric, for example, is not rejected. This hypothesis can be verified using the Shapiro-Wilk [Royston, 1995] test, using the experimental data.

Example 12. Let the following CCD damage metric values be calculated from electromechanical impedance signatures for a pristine condition of the steel plate.

0,0223	0,0453	0,0700	0,0632	0,0183
0,0523	0,0550	0,0290	0,0375	0,0057
0,0259	0,0194	0,0359	0,0611	0,0372
0,0098	0,0426	0,0206	0,0311	0,0317

a) Checking the hypothesis of normality, at 5% of significance, using the Shapiro-Wilk test.

Solution

```
> y=c(0.0223, 0.0453, 0.0700, 0.0632, 0.0183, 0.0523, 0.0550,
+      0.0290, 0.0375, 0.0057, 0.0259, 0.0194, 0.0359, 0.0611,
+      0.0372, 0.0098, 0.0426, 0.0206, 0.0311, 0.0317)
> shapiro.test(y)

Shapiro-Wilk normality test

data:  y
W = 0.97289, p-value = 0.8144
```

The p-value associated with the Shapiro-Wilk test is greater than 0.05; Then, the normality hypothesis is not rejected at 5% of significance.

b) Building the interval for the mean with 95% confidence.

Solution

```
> t.interval = function(data, conf.level){
+   n = length(data)
+   df = n - 1
+   t = qt((1 - conf.level)/2, df, lower.tail = FALSE)
+   xbar = mean(data)
+   s = sd(data)
+   c(xbar - t * s/sqrt(n), xbar + t * s/sqrt(n))
+ }
> t.interval(y,0.95)
[1] 0.02731178 0.04407822
```

Figure 31: Confidence interval obtained via parametric inference for the mean of the CCD metric.

With 95% confidence, the range from 0.02731178 to 0.04407822 contains the mean of the CCD variable (Figure 31).

c) Obtaining the range for the standard deviation with 95% confidence.

Solution

```
> var.interval = function(data, conf.level){
+ n = length(data)
+ df = n - 1
+ chi_lower = qchisq((1 - conf.level)/2, df)
+ chi_upper = qchisq((1 - conf.level)/2, df, lower.tail = FALSE)
+ v = var(data)
+ c((df*v)/chi_upper, (df*v)/chi_lower)
+ }
> var.interval(y,0.95) # Conf. interval for the variance
[1] 0.0001855628 0.0006844617
> sqrt(var.interval(y,0.95)) # Conf. interval for the sd
[1] 0.01362214 0.02616222
>
```

Figure 32: Confidence interval obtained via parametric inference for the standard deviation of the CCD metric.

With 95% confidence, the range from 0.01362214 to 0.02616222 contains the standard deviation of the CCD variable (Figure 32).

d) Determine the parametric threshold from the previous results.

Solution

```
> Thershold = 0.04407822 + 3*0.02616222
> Thershold
[1] 0.1225649
>
```

Figure 33: Threshold obtained via parametric inference.

Therefore, CCD values lower than 0.1226 (Figure 33) do not suggest significant changes in the integrity of the structure at investigation moments after the baseline condition.

The purpose of implementing the functions in the R language for calculating the confidence intervals for the mean and variance is that they can be used in other situations where the interest is to obtain these interval estimates.

The statistical methodology for the construction of confidence intervals in Eq. (50) and (51) assumes that the hypothesis of normality of the variable in question is not rejected. This hypothesis was tested based on the available sample data set. If there is no normality, the threshold can be obtained from the non-parametric approach via *bootstrap* resampling estimation.

3.2.2 Non parametric threshold

As mentioned, this method calculates the empirical distribution of the statistic of interest, mean, variance, etc., over hundreds or thousands of resamples. Then, no need to worry about assuming a distribution for the statistic in question and with the inherent assumptions.

With the values of the statistic of interest obtained in each resampling, it is enough to summarize all the information from the mean, variance, and confidence interval, among other possible measures.

The following example calculates the threshold (Eq. 53) in which the intervals for the mean and the standard deviation are defined using the *bootstrap* method.

Example 13. *The CCD values presented below were calculated from 16 impedance signatures measured from an experiment developed at the Structural Mechanics Laboratory of the Faculty of Mechanical Engineering of the Federal University of Uberlandia. The data set was adapted considering the signatures only in the frequencies indicated in Table 7. The median signature of the group was considered as baseline status. Therefore, the CCD metric was calculated using the pairs formed by each signature and the median signature. The variation in this baseline metric is due only to random or uncontrolled factors and not structural changes. The CCD values are as shown below.*

0,0202	0,0202	0,0201	0,0202	0,0201	0,0201	0,0201	0,0201
0,0201	0,0202	0,0201	0,0223	0,0209	0,0209	0,0208	0,0207

a) Checking the hypothesis of normality of the variable via the Shapiro-Wilk test.

Solution

```
> ccd=c(0.0202 ,0.0202, 0.0201, 0.0202, 0.0201, 0.0201, 0.0201, 0.0201,
+       0.0201, 0.0202, 0.0201, 0.0223, 0.0209, 0.0209, 0.0208, 0.0207)
> shapiro.test(ccd)

Shapiro-Wilk normality test

data:  ccd
W = 0.64392, p-value = 4.336e-05
```

Figure 34: Normality test for the CCD variable.

The sample data provide enough evidence to reject the normality hypothesis about the CCD variable at 0.5% of significance (Figure 34).

b) Calculating bootstrap confidence intervals for the mean of the CCD metric.

Solution

```
> boot.ic.mean <- function(data, R,conf.level) {
+   resamples <- lapply(1:R, function(i) sample(data, replace=T))
+   r.mean <- sapply(resamples, mean)
+   ci = quantile(r.mean,probs=c((1-conf.level)/2, (1 + conf.level)/2))
+   c(ci)
+ }
> boot.ic.mean(data=ccd,R=10000,conf.level=0.95)
      2.5%      97.5%
0.02020625 0.02076250
> |
```

Figure 35: Bootstrap confidence interval for the mean.

With 95% confidence, the range from 0.0202 to 0.0208 contains the mean of the CCD variable (Figure 35).

- c) Calculating bootstrap confidence intervals for the standard deviation of the CCD.

Solution

```
> boot.ic.sd <- function(data, R, conf.level) {
+   resamples <- lapply(1:R, function(i) sample(data, replace=T))
+   r.sd <- sapply(resamples, sd)
+   ci = quantile(r.sd, probs=c((1-conf.level)/2, (1 + conf.level)/2))
+   c(ci)
+ }
> boot.ic.sd(data=ccd, R=100000, conf.level=0.95)
      2.5%      97.5%
0.0002272114 0.0008650385
>
```

Figure 36: Bootstrap confidence interval for the standard deviation.

With 95% confidence, the range from 0.0002 to 0.0008 contains the population standard deviation of the CCD metric (Figure 36).

- d) Determining the threshold from the results obtained in the previous items via the non-parametric method of bootstrap.

Solution

```
> Thershold = 0.02075625 + 3*0.0008650385
> Thershold
[1] 0.02335137
>
```

Figure 37: Threshold obtained via bootstrap method.

In later health conditions than the baseline, obtaining a CCD value lower than 0.0233 (Figure 37) indicates that the structural integrity hypothesis is not rejected.

4 Concluding Remarks

The quality of the data set to infer the structure's condition is of principal importance under penalty of erroneous decisions. Thus, atypical signatures should be disregarded in the baseline and posterior signature measuring step. Signatures associated with the CCD metric's outliers should also be considered outliers and discarded.

The trend in the impedance curve along the frequency domain, i.e., a characteristic not necessarily atypical, can affect the results if its effect is different in the different impedance signatures. In this chapter, this characteristic was modeled using the linearized exponential model.

The effect of temperature on impedance was another application explored in this chapter, using regression polynomials. The residual dependence attributed, at least in large part, to the consecutive form of data collection was also considered in the modeling. The status of the structure's

integrity was compared by comparing the CCD metrics obtained for pristine and investigated cases. Thresholds were also determined to verify statistically significant changes in the structure's behavior.

In conclusion, parametric inference techniques cannot always be applied; therefore, both this kind of inference and alternative non-parametric methods were presented. Parametric inference should be preferred as it is more powerful when the conditions for its application are met.

Acknowledgements

The authors are thankful to Petrobras for the financial support via the research project (R&D).

References

- H. Akaike. A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6):716–723, 1974.
- M. S. BARTLETT. Propriedades de suficiência e testes estatísticos. *Anais da Royal Society de Londres. Série A - Ciências Matemáticas e Físicas*, 160(901):268–282, 1937.
- D. A. Belsley, E. Kuh, and R. E. Welsch. *Regression diagnostics: Identifying influential data and sources of collinearity*. John Wiley & Sons, 2005.
- H. Bozdogan. Model selection and akaike's information criterion (aic): The general theory and its analytical extensions. *Psychometrika*, 52(3):345–370, 1987.
- T. S. Breusch. Testing for autocorrelation in dynamic linear models. *Australian economic papers*, 17(31):334–355, 1978.
- T. S. Breusch and A. R. Pagan. A simple test for heteroscedasticity and random coefficient variation. *Econometrica: Journal of the Econometric Society*, pages 1287–1294, 1979.
- J. Durbin. Tests for serial correlation in regression analysis based on the periodogram of least-squares residuals. *Biometrika*, 56(1):1–15, 1969.
- J. DURBIN and G. S. WATSON. Testing for serial correlation in least squares regression. *I. Biometrika, London*, 37:409–428, 1950.
- B. Efron and R. Tibshirani. *An introduction to the bootstrap/bradley* efron and robert j, 1993.
- J. Fox and S. Weisberg. *An R companion to applied regression*. Sage publications, 2018.
- L. G. Godfrey. Testing against general autoregressive and moving average error models when the regressors include lagged dependent variables. *Econometrica: Journal of the Econometric Society*, pages 1293–1301, 1978.
- J. Gross and U. Ligges. *nortest: Tests for Normality*, 2015. URL <https://CRAN.R-project.org/package=nortest>. R package version 1.0-4.
- R. Hyndman, G. Athanasopoulos, C. Bergmeir, G. Caceres, L. Chhay, M. O'Hara-Wild, F. Petropoulos, S. Razbash, E. Wang, and F. Yasmeen. *forecast: Forecasting functions for time series and linear models*, 2020. URL <http://pkg.robjhyndman.com/forecast>. R package version 8.12.

- R. J. Hyndman and G. Athanasopoulos. *Forecasting: principles and practice*. OTexts, 2018.
- N. Khatun et al. Applications of normality test in statistical analysis. *Open Journal of Statistics*, 11(01):113, 2021.
- E. Langford. Quartiles in elementary statistics. *Journal of Statistics Education*, 14(3), 2006.
- H. LEVENE. Robust tests for equality of variances. In: *I. Olkin, et al., Eds., Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*, Stanford University Press, Palo Alto, 1960.
- G. M. Ljung and G. E. Box. On a measure of lack of fit in time series models. *Biometrika*, 65(2): 297–303, 1978.
- M. H. A. NEWMAN. Elementos da topologia de conjuntos de pontos planos. *Cambridge*, 1939.
- R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2021. URL <https://www.R-project.org/>.
- J. O. Rawlings, S. G. Pantula, and D. A. Dickey. *Applied regression analysis: a research tool*. Springer, 1998.
- R. A. Reddy. *Applied data analysis and modeling for energy engineers and scientists*. Springer, Nova Iorque; Dordrecht; Heidelberg; Londres., 2011.
- J. Royston. Shapiro-wilk normality test and p-value. *Applied statistics*, 44(4):547–551, 1995.
- J. P. ROYSTON. An extension of shapiro and wilk's w test for normality to large samples. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 31(2):115–124, 1982.
- H. A. SCHEFFÉ. A method for judging all contrasts in the analysis of variance. *Biometrika*, 40 (1-2):87–110, 1953.
- C. M. Schubert Kabban, B. M. Greenwell, M. P. DeSimio, and M. M. Derriso. The probability of detection for structural health monitoring systems: Repeated measures data. *Structural Health Monitoring*, 14(3):252–264, 2015.
- G. Schwarz. Estimating the dimension of a model. *The annals of statistics*, pages 461–464, 1978.
- J. R. Taylor. *Introdução à análise de erros*. Bookman, Porto Alegre., 2012.
- H. Thode. Testing for normality marcel dekker. *Inc. New York*, pages 99–123, 2002a.
- H. Thode. Testing for normality marcel dekker. *Inc. New York*, pages 99–123, 2002b.
- J. W. TUKEY. The problem of multiple comparisons. *Unpublished manuscript, Princeton University*, 1953.
- J. W. Tukey. *Exploratory data analysis*, volume 2. Reading, Mass., 1977.
- A. Zeileis and T. Hothorn. Diagnostic checking in regression relationships. *R News*, 2(3):7–10, 2002. URL <https://CRAN.R-project.org/doc/Rnews/>.

A Baseline data set from an experiment carried out on steel sheets at the Structural Mechanics Laboratory (LMEst) of the Faculty of Mechanical Engineering (FEMEC) of the Federal University of Uberlandia (UFU).

Table 7: Real part of electromechanical impedance signatures (Ω) by frequency (Hz), in 16 samples and median.

Freq.	Impedance Signatures																Mediana
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
40000.00	25.75	25.93	25.93	25.40	25.77	25.77	25.94	25.94	25.94	25.75	25.76	25.92	25.75	25.21	25.57	25.75	25.77
40010.00	26.26	26.26	26.07	25.90	26.08	26.09	26.26	26.09	26.26	26.26	26.26	26.23	26.06	25.71	26.06	26.08	26.09
40020.01	26.20	26.20	26.21	26.20	26.21	26.38	26.20	26.37	26.20	26.02	26.20	26.02	26.01	26.01	25.83	25.82	26.20
40030.01	26.99	26.98	27.00	26.99	26.99	26.99	27.00	26.98	27.00	26.98	26.99	26.60	26.61	26.60	26.60	26.78	26.98
40040.01	29.17	29.01	29.00	29.01	29.00	29.01	29.01	29.02	29.02	29.01	29.01	28.62	28.95	28.80	28.79	28.78	29.01
40050.02	32.66	32.66	32.66	32.66	32.67	32.66	32.67	32.67	32.67	32.69	32.69	32.64	32.81	32.81	32.63	32.80	32.67
40060.02	36.62	36.46	36.63	36.46	36.44	36.63	36.46	36.63	36.46	36.44	36.62	36.26	36.62	36.63	36.61	36.59	36.60
40070.02	37.53	37.53	37.54	37.54	37.54	37.54	37.54	37.54	37.54	37.54	37.53	36.99	37.35	37.53	37.53	37.53	37.53
40080.03	37.66	37.49	37.66	37.67	37.67	37.85	37.67	37.66	37.66	37.67	37.67	37.25	37.47	37.46	37.62	37.48	37.66
40090.03	38.48	38.45	38.45	38.28	38.27	38.45	38.46	38.46	38.28	38.45	38.27	38.20	38.44	38.25	38.25	38.25	38.36
40100.03	37.44	37.63	37.65	37.66	37.64	37.63	37.66	37.63	37.63	37.63	37.64	37.80	37.81	37.79	37.81	37.78	37.64
40110.04	35.24	35.25	35.24	35.24	35.24	35.25	35.42	35.44	35.25	35.25	35.25	35.61	35.42	35.60	35.42	35.61	35.25
40120.04	32.91	32.72	32.91	32.92	32.72	32.91	32.91	32.91	32.91	33.10	32.92	32.91	32.92	32.92	32.92	32.94	32.91
40130.04	30.35	30.36	30.18	30.54	30.36	30.36	30.54	30.55	30.35	30.55	30.55	30.18	30.18	30.19	30.19	30.01	30.36
40140.05	27.56	27.36	27.56	27.56	27.36	27.37	27.56	27.56	27.56	27.55	27.56	27.19	26.98	26.99	27.01	26.98	27.46
40150.05	25.21	25.21	25.21	25.21	25.21	25.21	25.21	25.21	25.21	25.21	25.21	24.84	24.82	24.84	24.83	24.83	25.21
40160.05	24.53	24.53	24.53	24.35	24.34	24.34	24.34	24.34	24.34	24.35	24.53	24.35	24.14	23.96	24.14	24.32	24.34
40170.06	24.40	24.40	24.60	24.40	24.40	24.58	24.42	24.58	24.58	24.40	24.58	24.03	24.21	24.20	24.21	24.20	24.40
40180.06	25.38	25.55	25.38	25.56	25.56	25.37	25.38	25.38	25.39	25.56	25.39	25.18	25.36	25.36	25.36	25.18	25.38
40190.06	27.45	27.63	27.45	27.45	27.45	27.46	27.46	27.46	27.45	27.44	27.46	27.08	27.26	27.43	27.43	27.26	27.45
40200.07	30.36	30.35	30.36	30.36	30.35	30.36	30.54	30.36	30.35	30.36	30.36	29.99	30.17	30.17	30.35	30.18	30.36

40210.07	32.25	32.28	32.25	32.08	32.26	32.26	32.26	32.26	32.09	32.26	32.08	31.87	31.72	31.71	31.72	31.90	32.17
40220.07	30.33	30.16	30.16	30.15	30.33	30.16	30.16	30.16	30.14	30.16	30.16	29.94	29.75	29.76	29.75	29.75	30.16
40230.08	27.03	27.04	27.05	27.04	27.04	27.05	27.05	27.05	27.05	27.23	27.05	26.66	27.00	27.00	27.00	27.00	27.04
40240.08	24.75	24.75	24.93	24.75	24.75	24.95	24.94	24.95	24.75	24.77	24.95	24.54	24.91	25.10	25.09	24.91	24.91
40250.08	24.08	23.91	24.08	23.91	23.92	24.09	23.90	24.09	23.72	24.10	23.91	23.86	23.91	24.26	24.09	23.90	23.91
40260.09	24.00	24.00	24.00	24.00	24.00	24.18	24.01	24.00	24.01	24.00	24.01	24.17	23.98	23.98	23.98	23.99	24.00
40270.09	24.30	24.48	24.31	24.31	24.32	24.32	24.32	24.31	24.32	24.32	24.32	24.15	24.12	24.13	24.12	24.13	24.31
40280.09	24.29	24.30	24.31	24.48	24.48	24.48	24.31	24.31	24.30	24.48	24.31	23.78	24.11	24.10	24.12	24.12	24.30
40290.10	23.91	23.91	23.91	23.91	23.91	23.91	23.91	23.91	23.91	24.09	23.92	22.99	23.54	23.53	23.53	23.71	23.91
40300.10	23.47	23.64	23.46	23.64	23.47	23.47	23.66	23.49	23.49	23.48	23.48	22.92	23.10	23.27	23.10	23.27	23.47
40310.10	23.91	23.92	23.74	23.92	23.91	23.74	23.92	23.91	23.91	23.92	23.74	23.00	23.54	23.54	23.54	23.53	23.82
40320.11	24.57	24.57	24.57	24.59	24.56	24.59	24.57	24.59	24.56	24.59	24.58	23.65	24.54	24.54	24.55	24.37	24.57
40330.11	25.25	25.23	25.24	25.23	25.23	25.23	25.24	25.24	25.23	25.24	25.23	24.64	25.38	25.20	25.20	25.22	25.23
40340.11	26.39	26.38	26.39	26.39	26.39	26.39	26.40	26.40	26.22	26.40	26.22	26.32	26.53	26.36	26.35	26.35	26.39
40350.12	28.80	28.78	28.79	28.79	28.80	28.80	28.79	28.79	28.79	28.80	28.62	29.23	29.10	28.93	28.93	28.76	28.80
40360.12	32.38	32.39	32.39	32.22	32.37	32.40	32.40	32.22	32.37	32.38	32.40	33.08	32.55	32.55	32.36	32.57	32.39
40370.12	35.15	35.15	34.98	35.14	34.98	35.14	35.14	34.97	35.15	34.98	34.97	35.39	34.99	34.97	34.96	34.97	34.98
40380.13	35.32	35.32	35.33	35.31	35.31	35.31	35.32	35.30	35.32	35.12	35.31	34.61	34.95	35.13	35.12	34.94	35.31
40390.13	34.08	33.90	33.90	33.90	33.90	33.89	33.90	33.89	33.90	33.89	33.89	33.31	33.69	33.70	33.69	33.70	33.89
40400.13	32.02	32.02	32.03	32.02	32.02	32.02	32.03	32.04	32.03	32.02	32.02	31.46	31.46	31.46	31.46	31.47	32.02
40410.14	29.12	29.12	29.12	28.94	29.12	29.30	29.11	29.11	29.12	29.12	29.11	28.73	28.73	28.72	28.54	28.54	29.11
40420.14	27.03	27.03	27.03	26.85	26.85	26.84	27.02	26.85	26.85	27.04	27.03	26.64	26.64	26.64	26.64	26.64	26.85
40430.14	25.85	25.86	25.86	25.86	25.86	25.86	25.86	25.86	25.86	25.85	25.86	25.66	25.48	25.66	25.47	25.66	25.86
40440.15	25.43	25.42	25.61	25.42	25.42	25.43	25.43	25.61	25.44	25.44	25.62	25.41	25.41	25.22	25.40	25.41	25.43
40450.15	25.39	25.22	25.21	25.22	25.21	25.40	25.40	25.40	25.40	25.39	25.40	25.21	25.03	25.03	25.02	25.01	25.22
40460.15	24.66	24.85	24.67	24.85	24.84	24.66	24.85	24.85	24.67	24.85	24.67	24.47	24.48	24.48	24.47	24.48	24.67
40470.16	23.91	23.91	23.92	23.91	23.91	23.91	24.09	23.92	23.91	23.91	24.10	23.71	23.70	23.54	23.72	23.72	23.91
40480.16	23.29	23.29	23.29	23.31	23.30	23.31	23.31	23.30	23.30	23.31	23.30	23.10	23.10	22.92	22.92	23.10	23.30
40490.16	23.55	23.55	23.55	23.72	23.73	23.73	23.55	23.55	23.55	23.55	23.55	23.35	23.52	23.35	23.35	23.34	23.55

Chapter 12

Neuro-Fuzzy Data Normalization Applied for Impedance-based SHM to the Oil & Gas Industry

Chapter details

Chapter DOI:

<https://doi.org/10.4322/978-65-86503-88-3.c12>

Chapter suggested citation / reference style:

Freitas, Fernando A., et al. (2022). “Neuro-Fuzzy Data Normalization Applied for Impedance-based SHM to the Oil & Gas Industry”. In Jorge, Ariosto B., et al. (Eds.) *Uncertainty Modeling: Fundamental Concepts and Models*, Vol. III, UnB, Brasilia, DF, Brazil, pp. 384–426. Book series in Discrete Models, Inverse Methods, & Uncertainty Modeling in Structural Integrity.

P.S.: DOI may be included at the end of citation, for completeness.

Book details

Book: Uncertainty Modeling: Fundamental Concepts and Models

Edited by: Jorge, Ariosto B., Anflor, Carla T. M., Gomes, Guilherme F., & Carneiro, Sergio H. S.

Volume III of Book Series in:

Discrete Models, Inverse Methods, & Uncertainty Modeling in Structural Integrity

Published by: UnB City: Brasilia, DF, Brazil Year: 2022

DOI: <https://doi.org/10.4322/978-65-86503-88-3>

Neuro-Fuzzy Data Normalization Applied for Impedance-based SHM to the Oil & Gas Industry

Fernando Augusto Freitas¹, Rosana Sueli da Motta Jafelice^{2*},
José Waldemar da Silva², Quintiliano S. S. Nomelini²,
Diogo de Souza Rabelo³, José dos Reis V. Moura Jr⁴,
Roberto M. Finzi Neto⁵, Carlos A. Gallo⁵ and Julio E. Ramos⁶

¹Math Graduate Program, Federal University of Uberlândia, Brazil. E-mail: fernandoaugusto.mat@gmail.com

²Faculty of Mathematics, Federal University of Uberlândia, Brazil. E-mail: rmotta@ufu.br; zewaldemar@ufu.br; quintiliano.nomelini@ufu.br

³Post-Graduate Program – Industrial Engineering, Fed. University of Goias, School of Sciences and Technology, Aparecida de Goiania, Brazil. e-mail: diogo.rabelo@ufg.br

⁴Post-Graduate Program – Modeling and Optimization, Fed. University of Catalao, Brazil. e-mail: zereis@ufcat.edu.br

⁵Post-Graduate Program – Mechanical Engineering, Federal University of Uberlandia, Brazil. e-mail: finzi@ufu.br; gallo@ufu.br

⁶Petrobras, Petroleo Brasileiro S.A., RD Center (CENPES), Brazil. e-mail: julio.ramos@petrobras.com.br

*Corresponding author

Abstract

The electromechanical impedance-based approach can perform Structural Health Monitoring (SHM). The impedance signature is used to map any structural change and for follow-up purposes. Two major challenges for SHM are normalizing the collected impedance data and determining strategies to assess the level of damage to engineering structures and equipment over time. The objectives of this chapter are to present a data normalization technique and to illustrate a case study by modeling the damage level of an aluminum beam using Fuzzy Rule-Based Systems (FRBSs) that are generated using the Adaptive Neuro-Fuzzy Inference System (ANFIS). The training is carried out for the first aim with the input variables temperature and frequency, and the output data are impedance baseline signatures. The temperature effect can generate changes in the impedance signature, leading to incorrect structural diagnoses. Because of that, it is necessary to compensate for the impact of this variable for later prediction of impedance signatures without damage at temperatures that were not necessarily observed in the data collection. Results were obtained in the validation, in which a part of the data was used for training the FRBSs, and another

part was intended for confirmation. Both of the baseline data indicate good accuracy of the predicted signatures since the highest Correlation Coefficient Deviation (CCD) damage index was 0,003800021. To evaluate the damage level, FRBSs were constructed, with the input variables being two damage indices established by the electromechanical impedance signatures. The FRBSs average hit percentages are 95%. This result can indicate possible inputs for FRBSs to identify the damage levels when these FRBSs output values are unknown. Finally, the methodology proposed is used for the damage detection process in an experiment to detect corrosion-related damage in metallic structures. .

Keywords: Electromechanical impedance-based SHM; Fuzzy sets; ANFIS; Data normalization; Damage detection

1 Introduction

The Electromechanical Impedance-based Structural Health Monitoring technique (EMI-SHM) allows detecting structural changes using PZT (Lead Zirconate Titanium) ceramic transducers, bonded or embedded in the evaluated host structure. This technique reduces direct and indirect maintenance costs through an online and near real-time monitoring approach. It is also possible to improve by detecting design mistakes in their developing stage. This is possible by allowing a change of potential changes through a shift in excitation that can be changed. These changes are changed to qualitative parameters and can be tested through pattern models, which are either tested pure or untested curvature indices or tested as basal curvature with a verified curvature in the current state [Rabelo et al., 2017] EMI-SHM can be applied to complex structures, such as [Maruo et al., 2015] aircraft panels, steel fiber reinforced concrete structures [Silva et al., 2020] and other types of applications.

This technique has several advantages that make it promising, such as it can be applied to complex structures; transducers can be installed in hard-to-reach places; it has a high sensitivity to incipient failures; measurement data is easy to interpret; it can be implemented online; and is adapted for continuous monitoring, which can reduce the number of maintenance stops [Park et al., 1999].

However, an important limitation of this technique is the choice of the ideal frequency band used for the monitoring process. This parameter is usually determined experimentally in the signature range with high peak density, indicating good electromechanical coupling. It is worth noting that the width of the frequency band can influence the ability to detect structural faults, so a very narrow band can limit the variety of detectable damage types. On the other hand, an extensive frequency range may contain regions of low electromechanical coupling, and small changes may not be found in a specific part of the frequency range, due to a low local electromechanical coupling [Rabelo et al., 2017]. Therefore, it is necessary to use several frequency points and assume that, at least in some of them, the impedance will be modified in the presence of damage. Alternatively, the frequency range can also be determined using artificial intelligence techniques [Moura and Steffen, 2004].

Thus, EMI-SHM allows the early detection of structural failures, which can increase work environment safety and reduce maintenance costs. The presence or absence of structural modifications, in general, is verified from damage indices or metrics computed from the comparison between reference signatures, known as baseline, and investigated signatures.

Among the various existing damage indices or metrics, RMSD1 (root mean square deviation), CCD (correlation coefficient deviation), and ASD (mean square difference) [Palomino, 2008] can

be mentioned. The RMSD1 index can be calculated through the equation (1), the ASD through the equation (2), while the CCD is computed by (3), which also involves the difference between one and the correlation coefficient (CC), given by the equation (4), as follows:

$$RMSD1 = \sqrt{\sum_{j=1}^n \frac{(Z_{B,j} - Z_{I,j})^2}{Z_{I,j}^2}}, \quad (1)$$

$$ASD = \sum_{j=1}^n [(Z_{B,j} - Z_{I,j}) - (\bar{Z}_B - \bar{Z}_I)]^2, \quad (2)$$

$$CCD = 1 - CC, \quad (3)$$

$$CC = \frac{1}{n} \sum_{j=1}^n \left\{ \frac{[Z_{B,j} - \bar{Z}_B] [Z_{I,j} - \bar{Z}_I]}{S_{Z_B} S_{Z_I}} \right\}, \quad (4)$$

where $Z_{B,j}$ is the real part of the baseline signature at frequency j and $Z_{I,j}$ is the real part of the investigated signature at frequency j , both considering n frequency points, \bar{Z}_B and \bar{Z}_I are the averages of the baseline and investigated signatures, respectively, in the selected frequency range, S_{Z_B} is the standard deviation of the baseline signature and S_{Z_I} is the standard deviation of the investigated signature, considering all frequency points of each one.

Suppose the structure does not show significant changes at the time of investigation regarding the state of the baseline signature. In that case, it is expected that the damage index determined from the two signatures will be relatively small.

Fluctuations or changes in the impedance signature due to factors other than the damage itself can inflate the damage index, leading to a false diagnosis. This is especially true if temperature changes are not adequately compensated for, which can lead to false positives as to the presence of damage. Therefore, it is necessary to use methodologies that compensate for temperature effects so that the impedance changes detected in the analysis are mainly due to structural changes and not environmental impacts.

The term data normalization in SHM systems refers to a data processing procedure, which may include variations caused in the impedance signatures due to environmental and operational effects. There are many studies on the normalization of electromechanical impedance signatures, such as Krishnamurthy et al. [1996]; Bhalla et al. [2003]; Zhou et al. [2009]; Sepehry et al. [2011] and Bastani et al. [2012], whose results can be applied under specific conditions. There are also other studies such as Sun et al. [1995], Park et al. [1999] and Koo et al. [2009] on the correction of the variation of the impedance signature caused by the temperature effect, maximizing the correlation between the baseline signature and the signature measured for investigation of the structure. In these approaches, a necessary condition for model validation is that a baseline signature at a specific temperature level must be stored to be used for compensation purposes. Due to weather or climate constraints, this can be an important limitation if impedance signature(s) were not stored at a given temperature during the baseline acquisition procedure.

A technique to remove the temperature effect, based on the fit of regression polynomials, is proposed in Giancesini et al. [2021]. However, applying this technique is conducted in two steps: the first to adjust the horizontal displacement of the signature and the second to adjust the vertical.

However, these techniques have limitations, such as the size of the frequency band and, in the case of adjustment via regression analysis, there are assumptions to be taken into account, such as homogeneity, normality, and independence on the model residuals, which is not always possible to guarantee, and therefore the normalization of the data would be compromised.

Another critical difficulty encountered in traditional SHM methodologies is quantifying damage from these metrics obtained at their respective temperatures since there is no information on

whether the behavior of the damage increment and metric values are or are not directly linked.

Thus, the main objective of this chapter is to present an alternative for the normalization of electromechanical impedance data and a way to quantify the possible structural changes. To this end, we propose using Fuzzy Rule-Based Systems (FRBSs), a methodology that enables the manipulation of uncertain information or the mathematical representation of imprecision. This chapter is based on Freitas [2021].

The mathematical tool used in modeling the problems under study is the theory of Fuzzy Sets, which was proposed by Zadeh [1965]. This theory has proved to be an efficient mathematical tool for quantifying and characterizing uncertainties in various contexts. To exemplify, in the work of Ferreira [2011], the mathematical modeling of luminescence is carried out from experiments with a beam of light focused on a point on the surface of a glassy sample; in Nardez [2015], it is required the determination and positioning of points with high accuracy, using the Global Navigation Satellite System (GNSS), in which it is necessary to know the phase center displacements of the antennas GNSS; in the work of Alfaro [2019], the dynamics of delayed Human Immunodeficiency Virus (HIV) under antiretroviral treatment in seropositive individuals with two fuzzy parameters is simulated. In the first two works, ANFIS was used to determine SBRFs applied in modeling these phenomena.

In the building and elaboration of FRBSs, there is a need for the help of a specialist in the area to find its best configuration and development of its rules. Thus, ANFIS [Fresno et al., 2015] has proved to be an exciting tool as an aid for this FRBS configuration step when the data of the input and output variables are known.

Thus, this chapter aims to determine the FRBSs through ANFIS that make possible the normalization of electromechanical impedance data (Part I) and damage identification (Part II).

- Part I, which deals with normalization, took place in two steps, given by:

Step 1: (**Validation of FRBSs**) In this phase, the baseline impedance signatures are divided into two distinct groups. One group is used for training the model and the other for validation. The validation of the FRBS is performed through some damage metric; the CCD, for example, is calculated from a signature that was left out of training and the corresponding baseline signature built from the FRBSs and, therefore, both at the same temperature. It is noteworthy that the objective is to create baseline signatures and, therefore, it is expected that the damage metrics, under the hypothesis of model validity, will be low because they are comparisons between signatures that represent the structure in the same state, that is, baseline status;

Step 2: (**Building baseline signatures**) The second step consists of determining new FRBSs, taking into account all signatures collected in the baseline stage. These FRBSs build baseline signatures at any temperature, even if not measured but within the collection range. Thus, the problem of the lack of equivalence of the temperature effect between the baseline signatures and the signatures in the investigation step is solved through the FRBSs obtained by ANFIS.

- Part II deals with the identification of the damage level and consists of initially computing the values of the metrics RMSD1 (equation (1)) and CCD (equation (3)) and, from that, using these metrics as FRBS inputs, perform training through ANFIS to generate the FRBS that can identify the level of damage in the structure.

To meet the objectives proposed above, section 2 presents the main concepts of fuzzy set theory, which underlie the uncertainty found in FRBSs, and an example that illustrates this methodology, used in the context of electromechanical impedance at detecting levels of damage. Section

3 presents the primary tool for obtaining the FRBSs built in this work, ANFIS. Section 4 offers a proposal, in the form of an application, for data normalization. In section 5, another application is presented to investigate the structure's level of damage. Finally, in section 6, the conclusions of the chapter are presented.

2 Fuzzy Sets Theory

The fuzzy sets theory was initially proposed by Zadeh [1965], aimed at modeling uncertain situations. Thus, unlike classical set theory, which determines in a very well-defined way the insertion of an element in the set or not, in fuzzy logic, the elements are inserted in the groups with their respective degrees of membership, according to Pourjavad and Shahin [2018]. Thus, a fuzzy set is used to present vague concepts in the characterization of models mathematically.

The following section provides a formal definition of this concept and some examples, according to Jafelice et al. [2012], Barros and Bassanezi [2015].

2.1 Fuzzy Sets

Definition 2.1 A set fuzzy A of the universe U is characterized by a membership function: $\mu_A : U \rightarrow [0, 1]$. This is called the membership function of the set fuzzy A . The value $\mu_A(x)$ of the function is interpreted as the degree to which the element x belongs to the set fuzzy A , that is, $\mu_A(x) = 0$ indicates a no membership of x to A , and $\mu_A(x) = 1$ indicates total membership of the set [Zadeh, 1965]. A classical set of ordered pairs can also represent a set fuzzy: $A = \{(x, \mu_A(x)) | x \in U\}$.

In this context, the intention is to establish, for each element of a set, a value between 0 and 1, which is considered the membership degree and serves to express how much the element belongs or not to a subset of this set. The membership function can be triangular, trapezoidal, Gaussian, or even have other geometric shapes [Barros and Bassanezi, 2015].

Definition 2.2 A singleton set is a fuzzy set with a membership function that is unity at a single point in the considered universe and zero elsewhere.

When it comes to fuzzy sets, it is possible to think about operations that can be performed in this context. Thus, the union, intersection, and complement operations of classical sets can be applied to fuzzy sets and will be covered in the following subsection.

2.2 Operations with Fuzzy Sets

Definition 2.3 Given A and B two fuzzy sets with μ_A and μ_B their respective membership functions, one can define the union operations (5) (Figure 1b), intersection (6) (Figure 1c) and complementary (7) (Figure 1d) of the sets fuzzy A and B , for all $x \in U$, given by:

$$\mu_{(A \cup B)}(x) = \max\{\mu_A(x), \mu_B(x)\}, \quad (5)$$

$$\mu_{(A \cap B)}(x) = \min\{\mu_A(x), \mu_B(x)\}, \quad (6)$$

$$\mu_{A'}(x) = 1 - \mu_A(x), \quad (7)$$

where 1 represents the constant function.

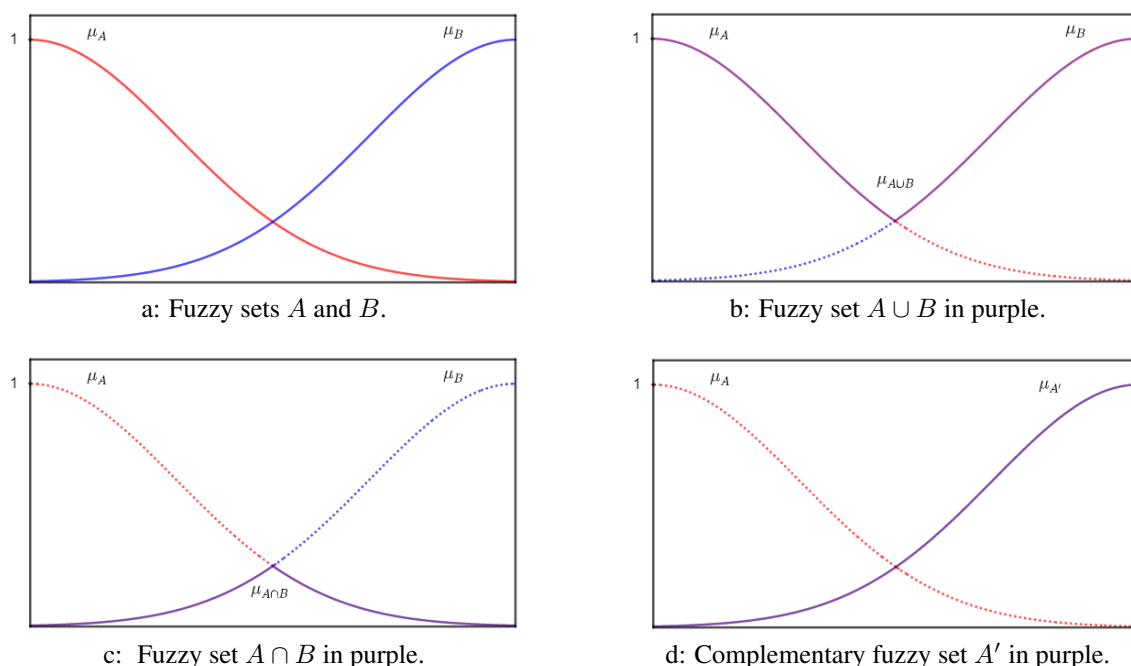


Figure 1: Representation of operations on fuzzy sets through membership functions.
 Source: Freitas et al. [2021c].

In the next subsection, the triangular norms are defined.

2.3 Triangular Norms

Definition 2.4 A triangular conorm (*s*-norm) is a binary operation $s : [0, 1] \times [0, 1] \rightarrow [0, 1]$, denoted by $s(x, y) = xsy$, satisfying the following conditions:

- *Commutativity*: $xsy = ysx$;
- *Associativity*: $xs(ysz) = (xsy)sz$;
- *Monotonicity*: If $x \leq y$ and $w \leq z$ then $xsw \leq ysz$;
- *Boundary conditions*: $xs0 = x$ and $xs1 = 1$.

Example 2.1

- a) *Standard Union*: $s : [0, 1] \times [0, 1] \rightarrow [0, 1]$ with $xsy = \max(x, y)$, (Figure 2a);
- b) *Algebraic Sum*: $s : [0, 1] \times [0, 1] \rightarrow [0, 1]$ with $xsy = x + y - xy$, (Figure 2b);
- c) *Limited Sum*: $s : [0, 1] \times [0, 1] \rightarrow [0, 1]$ with $xsy = \min(1, x + y)$, (Figure 2c);
- d) *Drastic Union*: $s : [0, 1] \times [0, 1] \rightarrow [0, 1]$ com

$$xsy = \begin{cases} x & \text{if } y = 0; \\ y & \text{if } x = 0; \\ 1 & \text{otherwise;} \end{cases} \quad (\text{Figure 2d}).$$

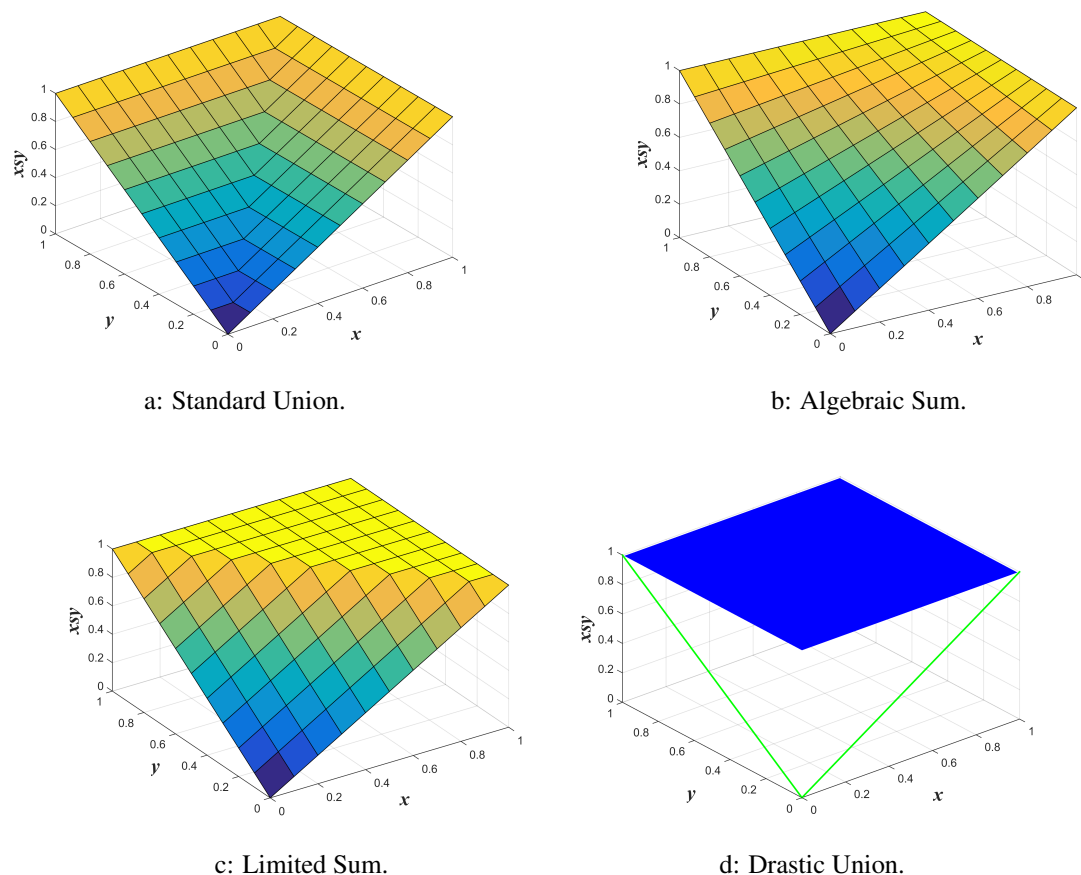


Figure 2: Representation of the surfaces generated by the examples of s-norms. Source: Freitas et al. [2021c].

Definition 2.5 A triangular norm (*t*-norm) is a binary operation $t : [0, 1] \times [0, 1] \rightarrow [0, 1]$, denoted by $t(x, y) = xty$, satisfying the following conditions:

- *Commutativity*: $xty = ytx$;
- *Associativity*: $xt(ytz) = (xty)tz$;
- *Monotonicity*: If $x \leq y$ and $w \leq z$ then $xtw \leq ytz$;
- *Boundary conditions*: $0tx = 0$ and $1tx = x$.

Example 2.2

- a) *Standard Intersection*: $t : [0, 1] \times [0, 1] \rightarrow [0, 1]$ with $xty = \min(x, y)$, (Figure ??);
- b) *Algebraic Product*: $t : [0, 1] \times [0, 1] \rightarrow [0, 1]$ with $xty = xy$, (Figure 3b);
- c) *Limited Difference*: $t : [0, 1] \times [0, 1] \rightarrow [0, 1]$ with $xty = \max(0, x + y - 1)$, (Figure 3c);
- d) *Drastic Intersection*: $t : [0, 1] \times [0, 1] \rightarrow [0, 1]$ with

$$xty = \begin{cases} x & \text{if } y = 1; \\ y & \text{if } x = 1; \\ 0 & \text{otherwise;} \end{cases} \quad (\text{Figure 3d}).$$

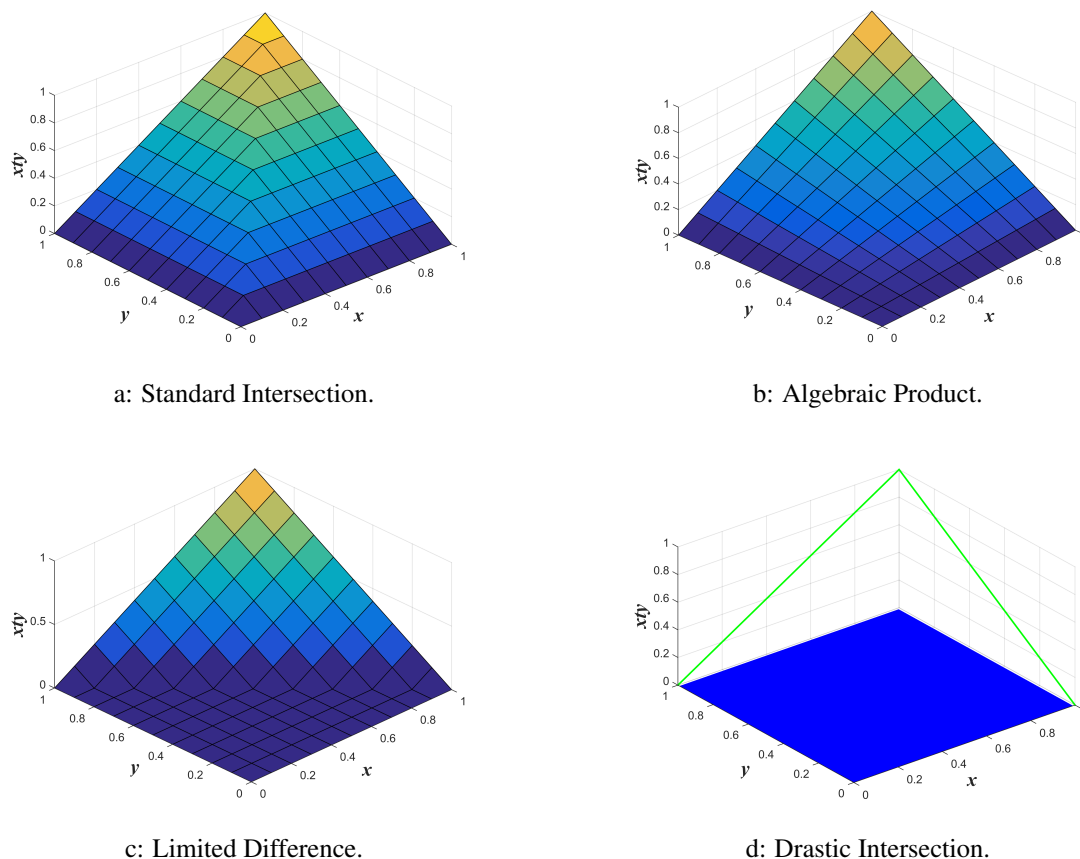


Figure 3: Representation of the surfaces generated by the examples of t-norms.
 Source: Freitas et al. [2021c].

Next, an essential technique of fuzzy set theory for this work, the fuzzy rule-based systems, is discussed.

2.4 Fuzzy Rule-Based System

The Fuzzy Rule-Based System (FRBS) is a process of mapping an input set to an output set. To introduce an FRBS, it is necessary to define linguistic variables.

2.4.1 Linguistic Variables

According to Jafelice et al. [2012], linguistic variables are used to quantify data or information using expressions of human communication.

Definition 2.6 *A linguistic variable is a variable whose value is expressed qualitatively by a linguistic term, providing a concept to the variable and quantitatively through the membership function. Thus, the linguistic variable is composed of a symbolic variable and a numerical value.*

Example 2.3 *Considering the linguistic variable structural change in aluminum plate, ranging from 0 to 1, its linguistic terms are assumed as small, medium, and large. These values are represented by fuzzy sets as shown in Figure 4.*

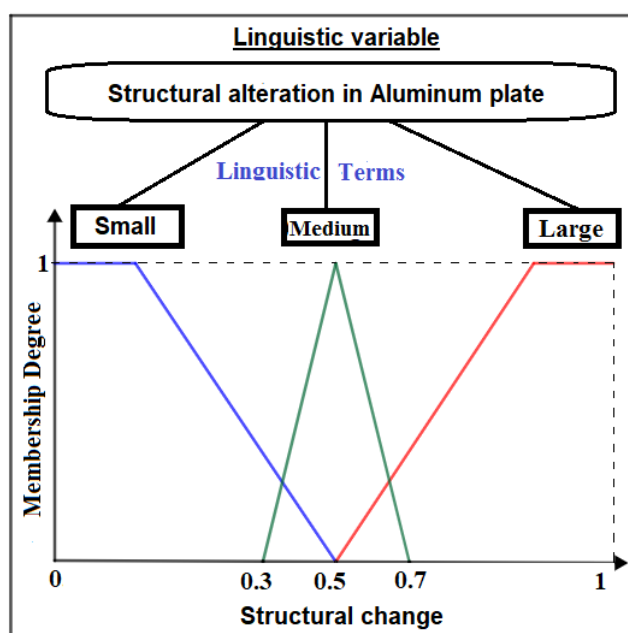


Figure 4: Linguistic variable structural change in an aluminum plate. Source: Adapted from Freitas et al. [2021c].

With the definition of linguistic variables, one can go further to the concept and structuring of FRBS, which will be done in the following subsection.

2.4.2 FRBS

An FRBS is composed of four components: the input processor (fuzzification process), a rule base, a fuzzy inference method, and an output processor (defuzzification process), which are connected as shown in Figure 5.

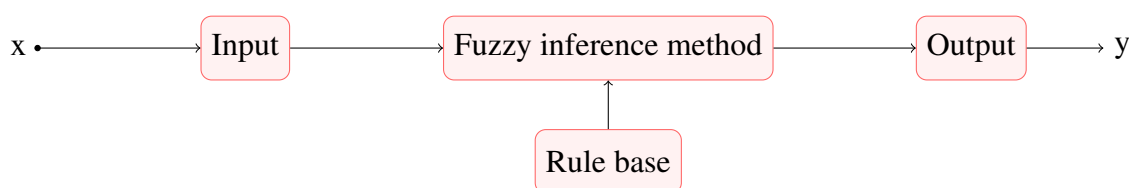


Figure 5: FRBS architecture.

Each component is defined as:

- Input Processor (Fuzzification) The system inputs are translated into fuzzy sets in their respective domains in this component.
- Rule Base This component, together with the inference method, can be considered the core of fuzzy rule-based systems. This is composed of a collection of fuzzy propositions described in the linguistic form:

$$\begin{aligned} &\text{If } x_1 \text{ is } A_1 \text{ and } x_2 \text{ is } A_2 \text{ and } \dots \text{ and } x_n \text{ is } A_n \\ &\text{then } y_1 \text{ is } B_1 \text{ and } y_2 \text{ is } B_2 \text{ and } \dots \text{ and } y_m \text{ is } B_m, \end{aligned}$$

where A_1, \dots, A_n are the fuzzy input sets and B_1, \dots, B_n are the fuzzy output sets.

- Fuzzy Inference Method

It is in this component that each fuzzy proposition is mathematically translated using approximate reasoning techniques. Mathematical operators will be selected to define the rule base's fuzzy relationship. In this way, the fuzzy inference machine is of fundamental importance for the success of the fuzzy system, as it provides the output from each input fuzzy and the relationship defined by the rule base.

- Mamdani Method

A rule is characterized by If (antecedent) then (consequent), provided by the Cartesian product fuzzy of the fuzzy sets that compose its antecedent and consequent. Mamdani's method groups the logical OR operator, arranged by the maximum operation in the s-norm, and the AND operator, configured by the minimum operation in the t-norm, in each rule, as shown in Figure 6. Given the following rules:

Rule 1: If (x_1 is A_1 and x_2 is B_1) then (y is C_1);

Rule 2: If (x_1 is A_2 and x_2 is B_2) then (y is C_2).

In Figure 6 the output y is represented through the Mamdani inference method for the inputs x_1 and x_2 , using the max-min composition.

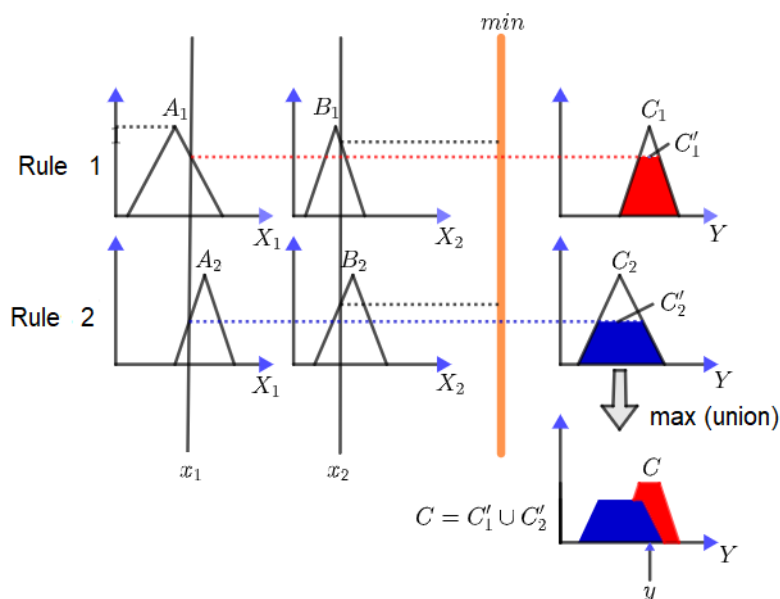


Figure 6: Mamdani inference method with max-min composition. Source: Adapted from Freitas et al. [2021c].

- Takagi-Sugeno method

The consequence characterizes this method in that each rule is a function depending on the input variables. Therefore, given the inputs, x_1 and x_2 , the output for each rule is given as a linear combination of these inputs. Thus, given the following rules:

Rule 1: If (x_1 is A_1 and x_2 is B_1) then ($y = f_1(x_1, x_2)$);

Rule 2: If (x_1 is A_2 and x_2 is B_2) then ($y = f_2(x_1, x_2)$).

A representation of the output y , obtained from this inference method (Takagi-Sugeno), for the inputs x_1 and x_2 , with $k_{ij} \in \mathbb{R}$, $i = 1, 2$ and $j = 0, 1, 2$, is shown in Figure 7

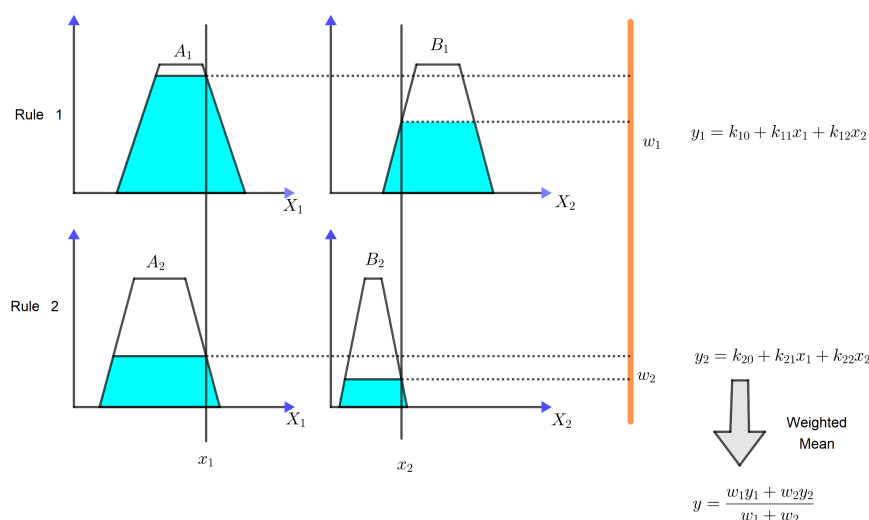


Figure 7: Takagi-Sugeno inference method. Source: Adapted from Freitas et al. [2021c].

The generalization of these inference methods can be found in Barros and Bassanezi [2015].

- Output Processor (Defuzzification)
defuzzification is a process that allows representing a set fuzzy by a crisp number. In fuzzy systems, in general, the output is a fuzzy set. Thus, we must choose a method to defuzzify the output and obtain a real number representing it. The most common method of defuzzification is as follows:

- Center of gravity
This defuzzification method is similar to the weighted average for data distribution, with the difference that the weights are the values $\mu_C(y_i)$ that indicate the degree of compatibility of the value y_i with the concept modeled by the fuzzy C set. For a discrete domain, we have:

$$G(C) = \frac{\sum_{i=0}^n y_i \mu_C(y_i)}{\sum_{i=0}^n \mu_C(y_i)}, \tag{8}$$

where y_i belongs to the domain of C .

The following section will present an example of a system based on fuzzy rules using Mamdani’s inference method.

2.5 FRBS Example: Grinding of Aluminum Beams

The experiment used to obtain the sampled data was carried out on aluminum beams. The structural modification was caused by the grinding process, which is the mechanical wear of the structure. The data set were obtained using the electromechanical impedance-based SHM technique, so

that the experiment was carried out in a climatic chamber, as shown in Figure 8, with temperature and humidity control, at the Structural Mechanics Laboratory (LMEst) of Faculty of Mechanical Engineering of the Federal University of Uberlandia. This experiment was chosen due to the increasing behavior, from the third cycle, of the metrics referring to the electromechanical impedance data set obtained, a condition that occurs, in general, in the actual SHM.



Figure 8: Climatic chamber used in the experiment. Source: Freitas et al. [2021c].

2.5.1 The Experimental Setup

In this experiment, four aluminum beam structures were used. The aluminum beams are 500mm long, 38mm wide, and 3.2mm thick. In each of them, a 1 mm thick and 20 mm diameter PZT patch was bonded 100 mm from the edge of the structure.

The failure mechanisms, or simulated damage, used in the experiment were surface grinding on one of the faces of the 30mm wide beams at 70mm from the opposite end and the same face on which the PZT patch was bonded. Nine levels of damage were considered in total, two baseline and seven levels of thickness removal by machining. Specimen 4 was used as a control mechanism, being removed from the chamber along with the others, but no failure mechanism was introduced in it. The difference between the mass values for specimen 4 (Table 1) comes from a failure in the measurement mechanism since the integrity of the beam was not altered in this case.

Figure 9 shows the four specimens tested in the bi-supported boundary condition. The beams were supported under the thickness, not over the width.



Figure 9: Specimens used in the test. Source: Freitas et al. [2021c].

It is important to emphasize that the stands for positioning the structures inside the chamber were printed in ABS plastic, with rounded feet, to reduce the interference of the chamber floor

in the structures and make it possible to embed connectors in the specimens themselves. The structures were removed from inside the chamber for machining and returned to the chamber for continuity of measurements with the least possible interference from the cables since they were placed in the same position inside the chamber, as shown in Figure 10.



Figure 10: Positioning the specimens inside the chamber. Source: Freitas et al. [2021c].

The temperature range increased with an increment of 3°C , in the range of 13 to 40°C ; therefore, the measurements were obtained at ten different temperatures or in ten cycles.

For each configuration: specimen, temperature, and damage condition, samples of size 30 were obtained. Thus, in a total of 10 temperature cycles, nine damage levels, four specimens, and 30 repetitions, 10800 impedance signatures were gathered.

Since inserting damage was performed manually from a grinder, it was decided to adopt two variables as the response of each specimen after inserting the damage. The first answer is the mass, considering the mass loss from the previous state. For this answer, a precision balance of two decimal places of the gram was used, and eight measurements were taken to obtain the average. The second answer is the thickness in the region delimited for machining, considering the loss of thickness in the previous state. A gauge with a resolution of 0.01mm was used for the second answer. The average of 10 random measurements of the thickness of the area delimited by the machining was recorded. An image of one of the experimental conditions or machining levels is shown in Figure 11, which shows the distance between the edge point and the machined region.

The results of the mass measurements for each specimen are presented in Table 1 and for the thicknesses in each specimen in Table 2, which consist of eight conditions since the baseline con-

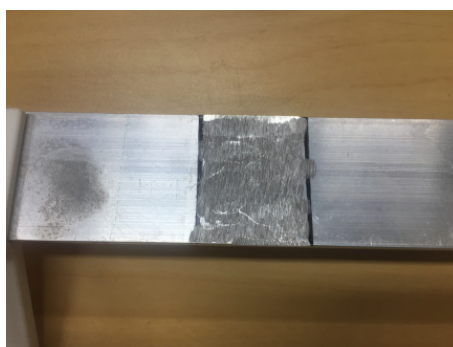


Figure 11: Damage 2 on specimen 1. Source: Freitas et al. [2021c].

dition was placed in the same parameter because they have the same data. In this experiment, we have the data of each damage attributed to the structure through the measurements of mass and thickness; however, the SHM technique is used to verify the structural change from the electromechanical impedance data set without worrying about the values presented in the tables. Thus, these values show that this experiment's loss of mass and thickness was slight.

Table 1: Mass, in grams, of specimens at each damage level.

Damage	Specimen			
	1	2	3	4
No damage ^a	191 .36	192 .42	192 .86	195 .89
1	191 .20	192 .27	192 .69	-
2	191 .16	192 .02	192 .58	-
3	190 .95	191 .82	192 .39	-
4	190 .41	191 .48	192 .02	-
5	189 .96	190 .54	191 .33	-
6	189 .16	189 .42	189 .78	-
7	187 .44	188 .57	188 .37	195 .83

^aBaseline condition

Source: Adapted from Freitas et al. [2021c].

Table 2: Thickness, in millimeters, of the specimens at each damage level.

Damage	Specimen			
	1	2	3	4
No damage ^a	3.17	3.17	3.19	3.18
1	3.09	3.11	3.09	-
2	3.01	2.96	2.99	-
3	2.97	2.93	2.95	-
4	2.85	2.85	2.89	-
5	2.72	2.49	2.63	-
6	2.44	2.23	2.09	-
7	1.93	1.92	1.83	3.18

^aBaseline condition

Source: Adapted from Freitas et al. [2021c].

As can be seen in all structures, by Table 2, the damages inserted in the three specimens never exceeded 50% of the thickness, and the last two machining levels applied were more severe to observe this more remarkable growth (more minor and greater progressions) in thickness loss. The electromechanical impedance technique seeks to analyze structural change; however, when the change is minimal, as found for the initial damages of the experiment, the methodology through the metrics may not show such a structural change. When this change was not included, the structural modification was considered irrelevant. The metric had a value very close to the structure metrics, comparing signatures in the baseline state.

Data from PZT1 patch were not used in the analysis due to problems in acquiring signatures.

2.5.2 Criteria for evaluation and design of FRBSs

A significant difficulty in obtaining FRBSs for damage classification is the enormous variability of signatures according to the material used and experiment settings. An experiment set up in close conditions can vary greatly in its signatures. This interference is smaller in metrics determined from these signatures, which justifies the use of CCD and ASD (Figure 12).

Then, the FRBS was built considering CCD and ASD metrics as input variables.

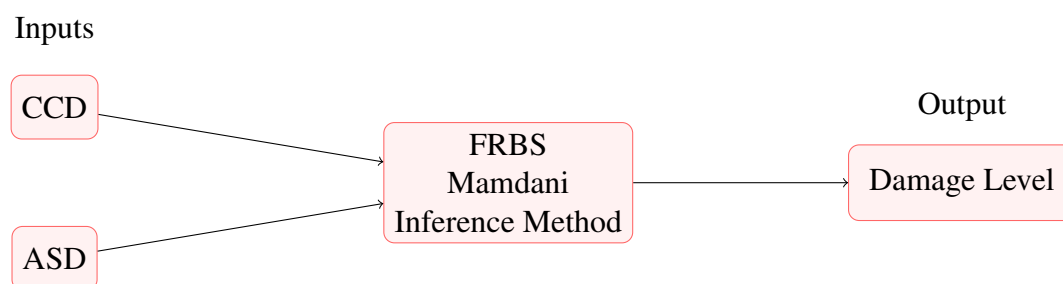


Figure 12: FRBS architecture.

To assess the quality of the FRBS, the combination of CCD and ASD metrics was used to determine the damage that each combination assumes. Three linguistic terms were designed for the damage level output variable: irrelevant, low, and medium. There was no characterization of high damage because the damage to the structure did not exceed 50% change from the original state when looking at the Table 2 of the thickness of the specimens.

2.5.3 Results of the Model Evaluation Step

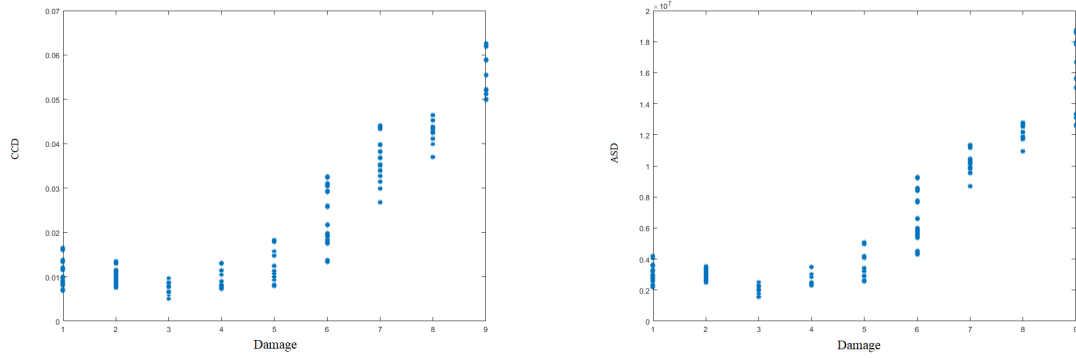
Considering the experiment data for PZT2 and PZT3 patches, the CCD and ASD metrics were calculated from the baseline signatures and those investigated. For the output variable, damage level, the irrelevant linguistic term corresponds to the signature baseline and damages 1, 2, and 3; the small to damage 4 and 5; and the medium to 6 and 7, where the damages from 1 to 7 are those in Table 1.

In Figure 13, the numbers 1 and 2 on the abscissa axis represent baseline experimental condition measurements, and the numbers 3 to 9 are the actual damage, listed in this text, from 1 to 7.

For the construction of the FRBS, CCD and ASD metrics were considered for the inputs, as mentioned, due to the similar behavior of these two metrics, along with the damage, as shown in Figure 13.

From the analysis of the behavior of the CCD and ASD metrics, for PZT2 patch, trapezoidal membership functions were established for each input and output, which are given by:

- Inputs, as shown in Figure 14:
 - CCD values: the domain $[0, 0.1]$, representing the ranges $[0.0, 0.025]$, $[0.02, 0.045]$ and $[0.04, 0.1]$, with the linguistic terms: Low, Medium and High, respectively.
 - ASD values: the domain $[1.2 \times 10^6, 3.5 \times 10^7]$, representing the ranges $[1.2 \times 10^6, 7, 089 \times 10^6]$, $[6.09 \times 10^6, 1.2 \times 10^7]$ and $[1.1 \times 10^7, 3.5 \times 10^7]$, with the linguistic terms: Low, Medium and High, respectively.

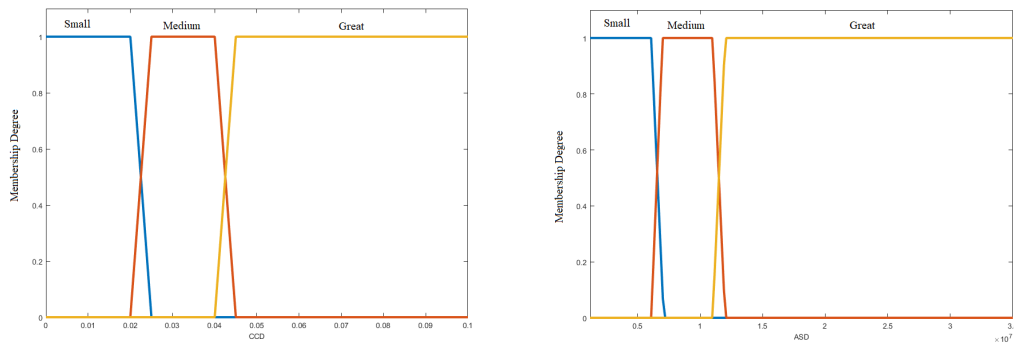


a: CCD metric behavior according to damage.

b: ASD metric behavior according to damage.

Figure 13: Behavior of CCD and ASD metrics with respect to damage. Source: Adapted from Freitas et al. [2021c].

- Output Values: the domain $[0, 2]$, represented by functions singleton in 0, 1 and 2, with the linguistic terms: Irrelevant, Small and Medium, respectively, as shown in Figure 15.



a: Membership function for the CCD variable.

b: Membership function for the ASD variable.

Figure 14: Membership functions for the FRBS inputs. Source: Adapted from Freitas et al. [2021c].

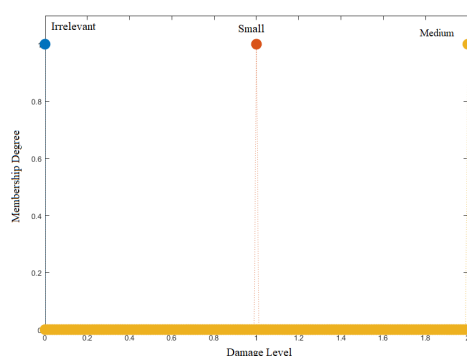


Figure 15: Membership functions for the FRBS output. Source: Adapted from Freitas et al. [2021c].

The following fuzzy rules were established:

- | | |
|---|--|
| Rule 1 : If (CCD is Low) and (ASD is Low) | then the Damage Level is Irrelevant ; |
| Rule 2 : If (CCD is Low) and (ASD is Medium) | then Damage Level is Irrelevant ; |
| Rule 3 : If (CCD is Low) and (ASD is High) | then Damage Level is Small ; |
| Rule 4 : If (CCD is Medium) and (ASD is Low) | then Damage Level is Irrelevant ; |
| Rule 5 : If (CCD is Medium) and (ASD is Medium) | then Damage Level is Small ; |
| Rule 6 : If (CCD is Medium) and (ASD is High) | then Damage Level is Medium ; |
| Rule 7 : If (CCD is High) and (ASD is Low) | then Damage Level is Small ; |
| Rule 8 : If (CCD is High) and (ASD is Medium) | then Damage Level is Small ; |
| Rule 9 : If (CCD is High) and (ASD is High) | then Damage Level is Medium . |

Data validation was performed by comparing the level of damage generated by the FRBS and that expected through the experiment by the absolute value of the difference between the level of damage obtained by the FRBS with that of the investigation. If this value is less than 0.5, it is compatible with the expected linguistic term. Otherwise, it is considered incompatible.

For the PZT2 patch data, the accuracy of 92.21% was found, showing that the methodology could identify the level of damage for a complex problem in most of the data. This same SBRF was used to verify the damage level for the PZT3 patch, and it was found that this sensor had a hitting percentage of 81.93%. Thus, the global hit level, regardless of the two PZT patches, was 87.03%, which is satisfactory, given the difficulty in performing such an analysis.

In the section 3, presented below, the theoretical foundation of ANFIS and relevant examples of numerical methods used for its computational implementation are discussed.

3 Fuzzy Rule-Based System on Adaptive Network

According to Fonseca [2012], the Adaptive Neuro-Fuzzy Inference System (ANFIS) is a hybrid technique of artificial intelligence that infers knowledge using fuzzy set theory principles and adds to this structure the possibility of learning inherent to artificial neural networks. Thus, this hybrid system solves one of the biggest problems of using fuzzy set theory: tuning the input and output functions of an FRBS. This section presents an explanation of artificial neural networks and ANFIS.

3.1 Artificial neural networks

Artificial Neural Networks (ANNs) can be seen as a computational method that, through experience, acquires knowledge based on the structure of biological neurons in the human brain, as shown in Figure 16.

The human brain is a parallel processing system that, even if the processing of an individual neuron is slow, is capable of processing trillions of operations simultaneously [Bezerra, 2016].

According to Valenca [2008], the biological neuron (Figure 16) is composed of a cell body, from which branches, known as dendrites, come out. The cell body and the telodendron, known as the terminal, are connected by the axon. The nerve impulse transmission between two neurons occurs from the axon of one neuron to the dendrite of the next, always in that order. The region where impulse transmission occurs is known as a synapse.

An important concept to understand the functioning of the biological neuron is the intensity of the stimulus, called the excitatory threshold, from which the neuron fires, or not, the nerve impulse. Therefore, if the stimulus is too small, its intensity will be below the excitatory threshold, and no nerve impulse will occur. On the other hand, after this threshold, the neuron's action potential will always be the same, whatever the intensity of the stimulus.

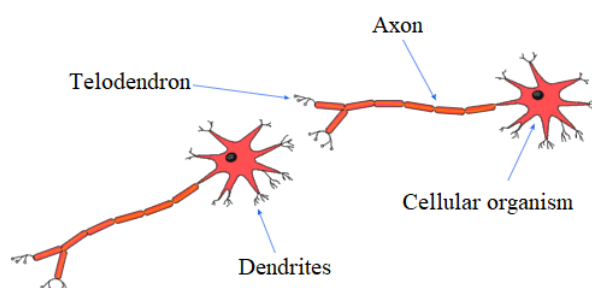


Figure 16: Representation of a biological neuron. Source: [Valenca, 2008].

The first to propose this biological neuron in an artificial model were McCulloch and Pitts [1943], and the model of this neuron is presented in Figure 17. Valenca [2008] states that this model looks for the biological neuron from a representation model.

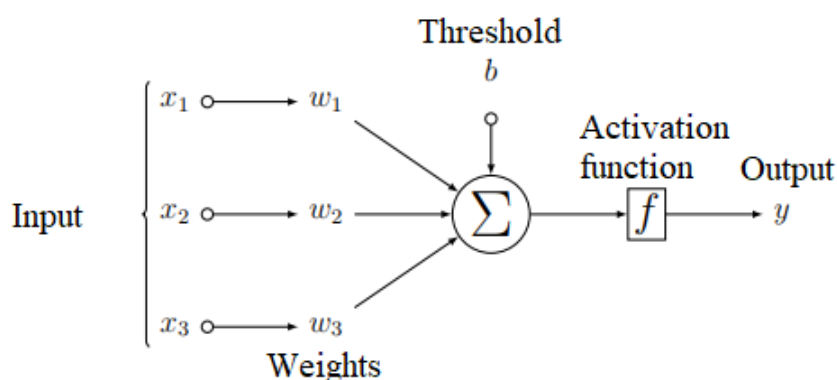


Figure 17: Threshold Logic Unit by McCulloch and Pitts. Source: [Bezerra, 2016].

Thus, consider x_1, x_2, x_3 , to be the input variables of the output neuron y . The following propagation rule gives the Σ entry:

$$\sum_{i=1}^3 w_i x_i - b,$$

where w_i are the synaptic weights and b is the excitatory threshold. From this, it is possible to calculate the $f(\sum)$, known as the activation function, thus giving the desired output $y = f(\sum)$.

Neural networks are commonly structured in layers, where each unit can be related to units in the previous layer. There are the input layers, in which the patterns are indicated to the network, the intermediate or hidden layers, where most of the network’s learning processing is performed, and the output layer, which presents the final result [Antognetti and Milutinovic, 1991].

There are two steps in developing an ANN: learning and use carried out sequentially and separately. In the learning stage, the data entered are visualized in two sets. The first group adjusts the network weights, while the second evaluates the training quality. Then, the utilization step is the mode in which the network corresponds to the input stimulus, not making further adjustments to its parameters [Braga et al., 2000].

In the case of ANFIS, in its training, it is necessary to use the method of least squares and gradient descent .

ANFIS, developed by Jang [1993], is a neuro-fuzzy system used to adjust membership functions and rules of an SBRF to find the most suitable answer for this system [Sodre et al., 2016]. The theory will be presented in the following subsection (subsection 3.2).

3.2 ANFIS

ANFIS is an artificial neural network of five layers interconnected through unit weights. Each layer is responsible for an operation that will produce an output equivalent to the one found at a given stage of a Takagi-Sugeno-type fuzzy system. Figure 18 is shown the equivalent architecture of ANFIS. To facilitate the understanding of the neuro-fuzzy network, we present a schema containing only two inputs, x_1 and x_2 , and an output, y .

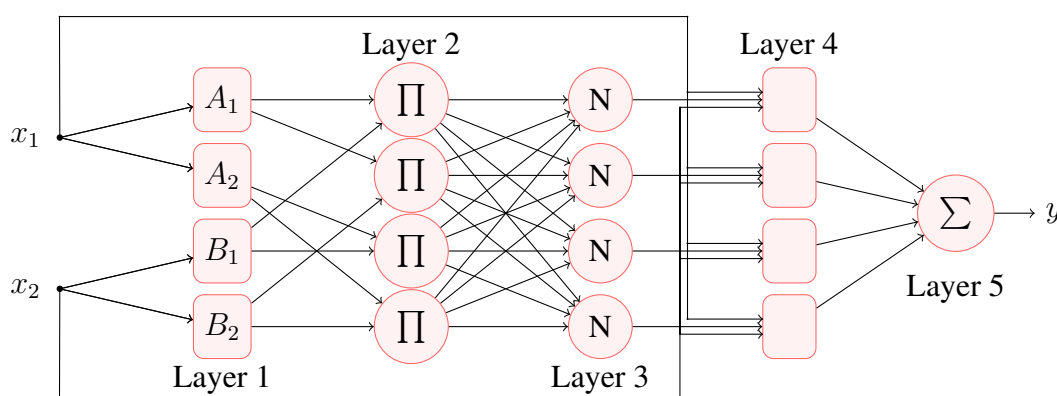


Figure 18: ANFIS architecture.

ANFIS implements four Takagi-Sugeno-type fuzzy rules:

- Rule 1: If (x_1 is A_1) and (x_2 is B_1) then $y = f_1 = k_{10} + k_{11}x_1 + k_{12}x_2$;
- Rule 2: If (x_1 is A_2) and (x_2 is B_2) then $y = f_2 = k_{20} + k_{21}x_1 + k_{22}x_2$;
- Rule 3: If (x_1 is A_2) and (x_2 is B_1) then $y = f_3 = k_{30} + k_{31}x_1 + k_{32}x_2$;
- Rule 4: If (x_1 is A_1) and (x_2 is B_2) then $y = f_4 = k_{40} + k_{41}x_1 + k_{42}x_2$;

where A_1 and A_2 are fuzzy sets of the universe X_1 ; B_1 and B_2 are fuzzy sets of the universe X_2 ; and k_{i0} , k_{i1} and k_{i2} are the parameter sets of rule i . Figure 7 shows the mechanism to obtain the inference method from the rules (9).

According to [Jang, 1993], after reading the data, its values are computed by membership functions that will identify the compatibility of each entry with its respective linguistic terms associated with this node. The function that is calculated in this layer is given by:

$$\begin{aligned} y_i^{(1)} &= \mu_{A_i}(x_1) \text{ for } i = 1, 2 \text{ ou} \\ y_i^{(1)} &= \mu_{B_{i-2}}(x_2) \text{ for } i = 3, 4, \end{aligned}$$

where x_1 or x_2 is an input of node i , A_i or B_{i-2} is the fuzzy set related to this node and $y_i^{(1)}$ is the output on the first layer.

For better understanding, we will assume that the membership functions are of the Gaussian type, according to the following equation:

$$\mu(x) = e^{-\frac{1}{2} \left(\frac{x-a}{b} \right)^2} \quad (9)$$

Note that each neuron in this layer has membership functions as its activation function in which the parameters of each of these functions are adaptive. It means that the parameters a and b are adjustable.

In the second layer, neurons have non-tunable activation functions \prod . In an ANFIS, the conjunction of rule antecedents is evaluated by the product of the operator (using some t -norm), where the output is given by:

$$\begin{aligned} y_1^{(2)} &= w_1 = \mu_{A_1}(x_1) \cdot \mu_{B_1}(x_2), \\ y_2^{(2)} &= w_2 = \mu_{A_1}(x_1) \cdot \mu_{B_2}(x_2), \\ y_3^{(2)} &= w_3 = \mu_{A_2}(x_1) \cdot \mu_{B_1}(x_2), \\ y_4^{(2)} &= w_4 = \mu_{A_2}(x_1) \cdot \mu_{B_2}(x_2), \end{aligned}$$

where $y_i^{(2)}$ is the output of this neuron obtained by the t -norm of the algebraic product.

In the third layer, there is another layer with fixed nodes, called N , non-adjustable, which normalizes the degrees of pertinence of each rule. Its function is given by:

$$y_i^{(3)} = \bar{w}_i = \frac{w_i}{w_1 + w_2 + w_3 + w_4}, i = 1, 2, 3, 4.$$

In the fourth layer, we have adaptive nodes again, whose function is composed of a set of parameters (k_{i0} , k_{i1} and k_{i2}) called consequent parameters. These parameters appear linearly in the output of each rule, according to the following function:

$$y_i^{(4)} = \bar{w}_i f_i = \bar{w}_i (k_{i0} x_1 + k_{i1} x_2 + k_{i2}),$$

where \bar{w}_i is the activation level calculated in the previous layer.

Finally, in the last layer, there is, again, a layer with a single fixed node called \sum , in which the correct output of the neuro-fuzzy system is calculated, encompassing all inputs given by the function:

$$y_i^{(5)} = \sum_{i=1}^4 \bar{w}_i f_i = \frac{\sum_{i=1}^4 w_i f_i}{\sum_{i=1}^4 w_i}$$

ANFIS uses a hybrid learning algorithm that combines the least squares estimator and the [Jang, 1993] gradient descent method. First, initial activation functions are assigned to each associated neuron. The centers of the functions of the neurons connected to the input x_i are defined so that the domain of x_i is divided equally, and the widths and slopes are defined to allow sufficient overlap of the respective functions.

In the ANFIS training algorithm, each epoch comprises one step forward and one step back. In the forward pass, a training set of input patterns (an input vector) is presented to ANFIS. The neuron outputs are calculated layer by layer, and the consequent parameters of the rule are identified by the least-squares estimator. The antecedent parameters are adjusted through the error backpropagation algorithm. The following shows how this training takes place.

In the Takagi-Sugeno inference method, as shown in Figure 7, an output, f , is a linear function. Thus, given the values of the membership parameters and joint training of input-output patterns $p = 1, \dots, P$, we can form linear equations P in terms of the consequent parameters as:

$$\begin{cases} Y_d(1) = \bar{w}_1(1)f_1(1) + \bar{w}_2(1)f_2(1) + \dots + \bar{w}_n(1)f_n(1) \\ Y_d(2) = \bar{w}_1(2)f_1(2) + \bar{w}_2(2)f_2(2) + \dots + \bar{w}_n(2)f_n(2) \\ \vdots \\ Y_d(p) = \bar{w}_1(p)f_1(p) + \bar{w}_2(p)f_2(p) + \dots + \bar{w}_n(p)f_n(p) \\ \vdots \\ Y_d(P) = \bar{w}_1(P)f_1(P) + \bar{w}_2(P)f_2(P) + \dots + \bar{w}_n(P)f_n(P) \end{cases} \tag{10}$$

or

$$\begin{cases} Y_d(1) = \bar{w}_1(1)[k_{10} + k_{11}x_1(1) + k_{12}x_2(1) + \dots + k_{1m}x_m(1)] \\ \quad + \bar{w}_2(1)[k_{20} + k_{21}x_1(1) + k_{22}x_2(1) + \dots + k_{2m}x_m(1)] + \dots \\ \quad + \bar{w}_n(1)[k_{n0} + k_{n1}x_1(1) + k_{n2}x_2(1) + \dots + k_{nm}x_m(1)] \\ Y_d(2) = \bar{w}_1(2)[k_{10} + k_{11}x_1(2) + k_{12}x_2(2) + \dots + k_{1m}x_m(2)] \\ \quad + \bar{w}_2(2)[k_{20} + k_{21}x_1(2) + k_{22}x_2(2) + \dots + k_{2m}x_m(2)] + \dots \\ \quad + \bar{w}_n(2)[k_{n0} + k_{n1}x_1(2) + k_{n2}x_2(2) + \dots + k_{nm}x_m(2)] \\ \vdots \\ Y_d(p) = \bar{w}_1(p)[k_{10} + k_{11}x_1(p) + k_{12}x_2(p) + \dots + k_{1m}x_m(p)] \\ \quad + \bar{w}_2(p)[k_{20} + k_{21}x_1(p) + k_{22}x_2(p) + \dots + k_{2m}x_m(p)] + \dots \\ \quad + \bar{w}_n(p)[k_{n0} + k_{n1}x_1(p) + k_{n2}x_2(p) + \dots + k_{nm}x_m(p)] \\ \vdots \\ Y_d(P) = \bar{w}_1(P)[k_{10} + k_{11}x_1(P) + k_{12}x_2(P) + \dots + k_{1m}x_m(P)] \\ \quad + \bar{w}_2(P)[k_{20} + k_{21}x_1(P) + k_{22}x_2(P) + \dots + k_{2m}x_m(P)] + \dots \\ \quad + \bar{w}_n(P)[k_{n0} + k_{n1}x_1(P) + k_{n2}x_2(P) + \dots + k_{nm}x_m(P)] \end{cases} \tag{11}$$

where m is the number of input variables, n is the number of neurons in the third layer, and $Y_d(p)$ is the desired overall ANFIS output for inputs $x_1(p), \dots, x_n(p)$.

The equation (11) can be written as follows:

$$y_d = Ak, \tag{12}$$

where y_d is a desired $P \times 1$ output vector,

$$y_d = \begin{bmatrix} Y_d(1) \\ Y_d(2) \\ \vdots \\ Y_d(p) \\ \vdots \\ Y_d(P) \end{bmatrix}_{P \times 1},$$

A is an matrix $P \times n(1 + m)$,

$$A = \begin{bmatrix} \bar{w}_1(1) & \bar{w}_1(1)x_1 & \cdots & \bar{w}_1(1)x_m(1) & \cdots & \bar{w}_n(1)x_1(1) & \cdots & \bar{w}_n(1)x_m(1) \\ \bar{w}_1(2) & \bar{w}_1(2)x_1 & \cdots & \bar{w}_1(2)x_m(2) & \cdots & \bar{w}_n(2)x_1(2) & \cdots & \bar{w}_n(2)x_m(2) \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ \bar{w}_1(p) & \bar{w}_1(p)x_1 & \cdots & \bar{w}_1(p)x_m(p) & \cdots & \bar{w}_n(p)x_1(p) & \cdots & \bar{w}_n(p)x_m(p) \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ \bar{w}_1(P) & \bar{w}_1(P)x_1 & \cdots & \bar{w}_1(P)x_m(P) & \cdots & \bar{w}_n(P)x_1(P) & \cdots & \bar{w}_n(P)x_m(P) \end{bmatrix}_{P \times n(1+m)}$$

and k is a vector $n(1 + m) \times 1$ of unknown consequent parameters

$$k = [k_{10} \ k_{11} \ k_{12} \ \cdots \ k_{1m} \ k_{20} \ k_{21} \ k_{22} \ \cdots \ k_{2m} \ \cdots \ k_{n0} \ k_{n1} \ k_{n2} \ \cdots \ k_{nm}]^T.$$

Typically, the number of input-output patterns P used in training is greater than the number of consequent parameters $n(1 + m)$. This means that we are dealing with an overdetermined problem, so the exact solution to the equation (12) may not even exist. Thus, one must find a least-squares solution k^* that minimizes the quadratic error $\|Ak - y_d\|$, determined by:

$$k^* = (A^T A)^{-1} A^T y_d, \tag{13}$$

where, A^T is the transpose of A , and $(A^T A)^{-1} A^T$ is the pseudo-inverse of A if $(A^T A)$ is not singular. Once the consequent parameters of the rule are established, we can calculate a real network output vector, y , and determine the error vector e ,

$$e = y_d - y,$$

so the backward propagation algorithm step is applied. Error signals are propagated back, and antecedent parameters are updated according to the chain rule.

It is considered a correction applied to the parameters a and b of the Gaussian activation function, equation (9), used in the neuron A_i . The chain rule is expressed according to the equations (15) and (16), for the parameters a and b , respectively. The squared error value E for the output neuron of ANFIS is given by:

$$E = \frac{1}{2} e^2 = \frac{1}{2} (y_d - y)^2. \tag{14}$$

According to Sandmann [2006], the partial derivatives are given by:

$$\frac{\partial E}{\partial a} = \frac{\partial E}{\partial e} \cdot \frac{\partial e}{\partial y} \cdot \frac{\partial y}{\partial(\bar{w}_i f_i)} \cdot \frac{\partial(\bar{w}_i f_i)}{\partial \bar{w}_i} \cdot \frac{\partial \bar{w}_i}{\partial w_i} \cdot \frac{\partial w_i}{\partial \mu_{A_i}} \cdot \frac{\partial \mu_{A_i}}{\partial a}. \tag{15}$$

$$\frac{\partial E}{\partial b} = \frac{\partial E}{\partial e} \cdot \frac{\partial e}{\partial y} \cdot \frac{\partial y}{\partial(\bar{w}_i f_i)} \cdot \frac{\partial(\bar{w}_i f_i)}{\partial \bar{w}_i} \cdot \frac{\partial \bar{w}_i}{\partial w_i} \cdot \frac{\partial w_i}{\partial \mu_{A_i}} \cdot \frac{\partial \mu_{A_i}}{\partial b}. \tag{16}$$

note that

$$\frac{\partial E}{\partial e} = \partial \left(\frac{1}{2}(e^2) \right) = e; \tag{17}$$

$$\frac{\partial e}{\partial y} = \frac{\partial(y_d - y)}{\partial y} = -1; \tag{18}$$

$$\frac{\partial y}{\partial(\bar{w}_i f_i)} = \frac{\partial \left(\sum_{i=1}^n \bar{w}_i f_i \right)}{\partial(\bar{w}_i f_i)} = 1; \tag{19}$$

$$\frac{\partial(\bar{w}_i f_i)}{\partial \bar{w}_i} = f_i; \tag{20}$$

$$\frac{\partial \bar{w}_i}{\partial w_i} = \frac{\partial \left(\frac{w_i}{\sum_{j=1}^n w_j} \right)}{\partial w_i} = \frac{\left(\sum_{j=1}^n w_j \right) - w_i}{\left(\sum_{j=1}^n w_j \right)^2} = \frac{\bar{w}_i(1 - \bar{w}_i)}{w_i}; \tag{21}$$

$$\frac{\partial w_i}{\partial \mu_{A_i}} = \frac{w_i}{\mu_{A_i}}; \tag{22}$$

$$\frac{\partial \mu_{A_i}}{\partial a} = \left(e^{-\frac{1}{2} \left(\frac{x-a}{b} \right)^2} \right) a = e^{-\frac{1}{2} \left(\frac{x-a}{b} \right)^2} \left(\frac{(x-a)}{b^2} \right); \tag{23}$$

$$\frac{\partial \mu_{A_i}}{\partial b} = \frac{\partial \left(e^{-\frac{1}{2} \left(\frac{x-a}{b} \right)^2} \right)}{\partial b} = e^{-\frac{1}{2} \left(\frac{x-a}{b} \right)^2} \left(\frac{(x-a)^2}{b^3} \right). \tag{24}$$

$$\tag{25}$$

The expressions of a and b , considering the learning rate α , will be corrected using the equations (15) and (16), given by:

$$a = a + \alpha(y_d - y) f_i \bar{w}_i (1 - \bar{w}_i) \cdot \frac{1}{\mu_{A_i}} \cdot e^{-\frac{1}{2} \left(\frac{x-a}{b} \right)^2} \left(\frac{(x-a)}{b^2} \right), \tag{26}$$

$$b = b + \alpha(y_d - y) f_i \bar{w}_i (1 - \bar{w}_i) \cdot \frac{1}{\mu_{A_i}} \cdot e^{-\frac{1}{2} \left(\frac{x-a}{b} \right)^2} \left(\frac{(x-a)^2}{b^3} \right). \tag{27}$$

Thus, according to ANFIS suggested by Jang [1993], it is verified that both the antecedent parameters and the consequent parameters are optimized. In the forward pass, the consequent parameters are adjusted while the antecedent parameters remain fixed. In the backward pass, the antecedent parameters are adjusted while the consequent parameters are kept fixed. According to Negnevitsky [2005], in some cases, when the input and output datasets are relatively small, membership functions can be described by a human expert. In such situations, these membership functions are kept fixed throughout the training process, and only the consequent parameters are adjusted.

In the sections 4 and 5 are shown ANFIS applications using experiments related to SHM.

4 ANFIS: Application in Temperature Compensation

This section presents an example that aims to perform data normalization, that is, to correct conditions that affect the signatures collected through the electromechanical impedance-based SHM technique, especially temperature.

4.1 The Experiment

The data used in this study come from an experiment carried out to investigate the ability of the electromechanical impedance technique to detect corrosion-associated failures in steel beams. The corrosion process is gradual and permanent and promotes increasing wear. In a long-term experiment, the beams were subjected to corrosion in an environment outside the laboratory through exposure to acid mist, while the impedance was measured cyclically. Acid mist induces corrosion acceleration. Results obtained from the analysis of this experiment were published by Freitas et al. [2021a].

A PZT patch was bonded to each beam to collect impedance signatures. Each impedance signature was measured with 8,000 frequency samples. The impedance analyzer used was developed by Finzi Neto et al. [2010], which is a portable and low-cost device.

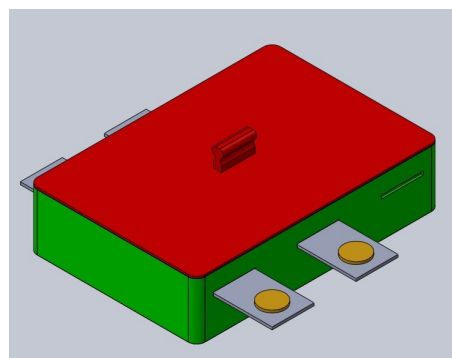
Three specimens (steel beams 1020) with dimensions of $300\text{mm} \times 50\text{mm} \times 3.06\text{mm}$ and a mass of 354.15g were instrumented with PZT patches, with a diameter of 30mm and a thickness of 2mm each. The PZT patches were bonded with epoxy resin-based adhesive and subsequently protected with a sealant. A temperature sensor was coupled to each of the three specimens to compensate the measurements for the effect of temperature variation on the signals obtained. Figure 19a illustrates specimens instrumented with PZT patches.

Then, the samples were lightly sanded to remove the oxide layers formed and to set the initial pristine condition (baseline).

An acid mist was used, which reacted with the specimens, as a corrosive agent. To this end, a structure located outside the laboratory building was adapted to contain the specimens in an airtight environment, in a place with restricted access. In addition, this environment had a cover to protect against gusts of strong winds and rain. For better control of the corrosive mist action environment, a control box was developed, designed in a 3D design software as shown in Figure 19b.



a: Specimens instrumented with PZT patches.



b: Control box used in the acid mist experiment.

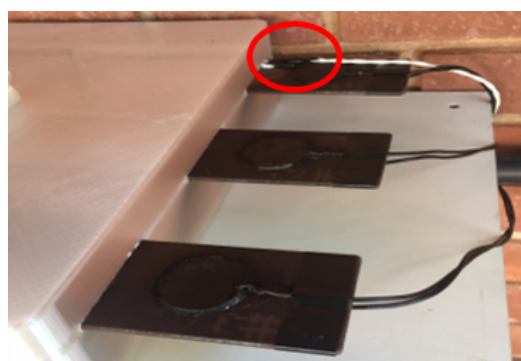
Figure 19: Instrumented specimens and control box. Source: Freitas et al. [2021a].

Figure 19b represents the box that delimits an intermediate and considerable region of the specimens that remain subject to corrosive mist. Inside the box, small reservoirs were developed

to add a small amount of hydrochloric acid, providing a corrosive environment. The control box was built in 3D printing. The final assembly of the experiment can be seen in Figures 20a and 20b.



a: Beams inside the control box.



b: Cable details and temperature sensor.

Figure 20: Control box. Source: Adapted from Freitas et al. [2021a].

The location of the temperature sensor, with the wire in white, can be seen by the red circle in Figure 20b

As stated earlier, temperature changes play an essential role in the damage detection process of the electromechanical impedance-based SHM method. The impedance signature of a structure can be different, even under the same structural conditions but at different temperatures. These differences are seen as horizontal and vertical displacements along with the signatures. Most of this variation is attributed to the temperature variation, as this variable interferes with the impedance value. If there is an important relationship between impedance and temperature, then it is possible to make predictions for impedance at temperatures not recorded in the experiment.

4.2 FRBSs Modeling

The FRBSs were modeled considering temperature and frequency as input variables. It was observed, through empirical tests, that the best setup consists of only two frequency values (points) per FRBSs, totaling 4,000 SBRFs. The output variable was impedance with a value at each frequency and temperature combination.

The data acquisition hardware stores the temperature values for each impedance reading per PZT patch. Therefore, each PZT patch has a set of temperature values corresponding to each impedance signal measured. In the validation process, part of the data set obtained in the baseline measurements was used to train and get the FRBSs. The other part for calculating damage indices, in which the signatures collected and modeled from the data set were used. FRBSs, at the respective temperatures. The temperatures in each of the two groups are shown in Table 3. Figure 21 shows the schematic of the mathematical modeling of the two phases that make up the experiment, in which the red rectangle represents the data obtained in the investigation; the blue rectangles represent the stage of training and construction of the FRBSs; and, in green, the validation step.

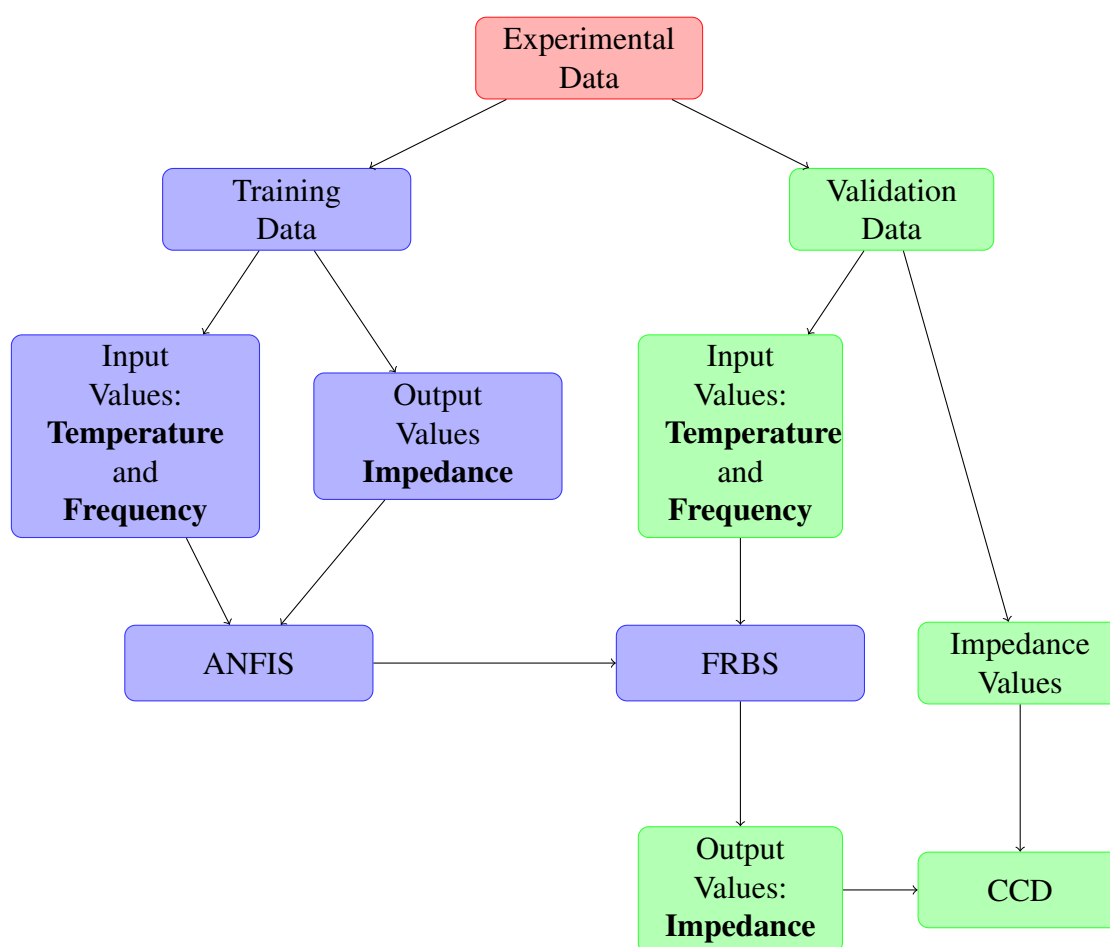


Figure 21: Scheme of the mathematical modeling used for the experiment.

To assess the quality of the models, the CCD metric was calculated from the pairs of baseline signatures measured and modeled from the SBRFs, at the same temperature. Low values of the CCD metric indicate the adequacy of the models or that the constructed signatures configure baseline signatures.

In the damage detection step, all temperatures observed in the baseline collection were used to train the FRBSs. For the second input variable, frequency, as in the validation step, two consecutive values were used in each FRBS. At this step, the output variable's impedance values were also observed in all frequency and temperature combinations.

The automation of building FRBSs for all frequency point pairs is relevant since the number of points used is generally significant. An iterative process was implemented in the software R [R Core Team, 2020] to obtain the 4000 FRBSs built from ANFIS through the package ANFIS [Fresno et al., 2015]. The R software allows the export of the FRBS built in each iteration to, for example, predict the impedance signatures at temperatures that can be observed in future investigations of the state of the structure.

4.3 Results: Model Evaluation

From the experimental data composed of 450 baseline signatures, each one with 8,000 frequency points, collected at temperatures between 23.7 to 31.3 °C, 46 different temperatures were observed. The analysis was implemented from the median signature at each temperature. A median signature is defined by the set of median impedance values at each frequency point. For the three

PZT patches, the ANFIS training routine was programmed 4000 times, that is, to generate 4000 FRBSs. The training for the determination of the FRBSs was carried out considering two input variables and one output, as follows:

- First Input: For the temperature variable (t), the first and last values were considered, in ascending order, which corresponds respectively to $23.7^{\circ}C$ and $31.3^{\circ}C$, in addition to the temperatures corresponding to the exact positions of the 46 values, also arranged in ascending order. In this way, the extreme values were considered in training, ensuring that there was no validation data outside the training range;
- Second Input: For the frequency variable (f), two consecutive values were always used among the 8,000 values;
- Output: For the impedance variable (i), the values collected in the combination of the two input variables were considered, totaling 48 training values.

Temperatures not used in training were assigned for further validation. Table 3 shows the temperature values used in the training and validation of the PZT patches of the experiment.

Table 3: Temperatures ($^{\circ}C$) recorded in the accelerated corrosion experiment.

PZT Patch	Set	Temperatures
1	Training	23.7 23.9 24.2 24.7 25.0 25.4 25.7 26.0 26.3 26.7 27.0 27.3 27.7 28.0 28.3 28.6 29.0 29.3 29.6 30.0 30.3 30.6 30.9 31.3
	Validation	24.0 24.4 24.9 25.2 25.5 25.8 26.2 26.5 26.8 27.2 27.5 27.8 28.1 28.5 28.8 29.1 29.5 29.8 30.1 30.5 30.8 31.1
2	Training	23.7 23.9 24.2 24.7 25.0 25.4 25.7 26.0 26.3 26.7 27.0 27.3 27.7 28.0 28.3 28.6 29.0 29.3 29.6 30.0 30.3 30.6 30.9 31.3
	Validation	24.0 24.4 24.9 25.2 25.5 25.8 26.2 26.5 26.8 27.2 27.5 27.8 28.1 28.5 28.8 29.1 29.5 29.8 30.1 30.5 30.8 31.1
3	Training	23.7 23.9 24.2 24.7 25.0 25.4 25.7 26.0 26.3 26.7 27.0 27.3 27.7 28.0 28.3 28.6 29.0 29.3 29.6 30.0 30.3 30.6 30.9 31.3
	Validation	24.0 24.4 24.9 25.2 25.5 25.8 26.2 26.5 26.8 27.2 27.5 27.8 28.1 28.5 28.8 29.1 29.5 29.8 30.1 30.5 30.8 31.1

Source: Adapted from Freitas et al. [2021a].

Thus, 4000 FRBSs were generated, all of which have two Gaussian membership functions for each input. The membership functions for temperature are the same for all FRBSs, as shown in Figure 22a. For frequency, the membership function centers on the fixed points of the experiment, as represented, for example, in Figure 22b. The four generated fuzzy rules, for PZT1 patch and frequencies 99990 and 100000, are described below:

Rule 1: If (t is High) and (f is High) then $i_1 = 0,3116833t + 0,00025376f + 2,5 * 10^{-9}$

Rule 2: If (t is Low) and (f is High) then $i_2 = 0,3134568t + 0,00025488f + 2,5 * 10^{-9}$

Rule 3: If (t is High) and (f is Low) then $i_3 = -0,02506984t + 0,00035051f + 3,5 * 10^{-9}$

Rule 4: If (t is Low) and (f is Low) then $i_4 = -0,01136947t + 0,00034594f + 3,4 * 10^{-9}$.

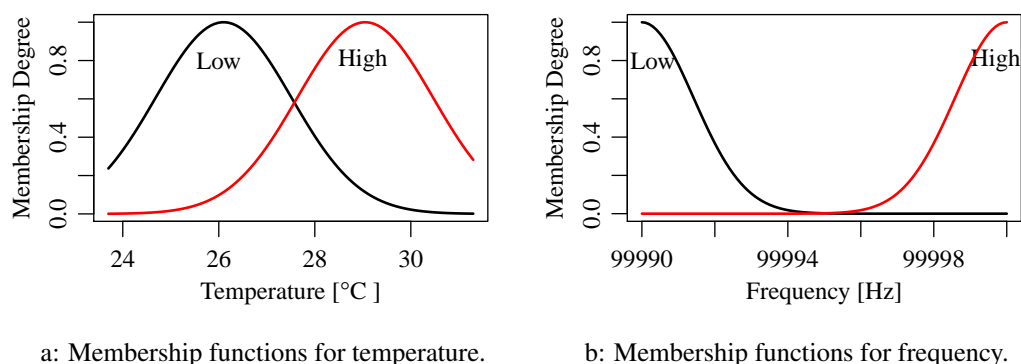


Figure 22: Membership functions of FRBS inputs. Source: Adapted from Freitas et al. [2021a].

Figure 23 shows the surface generated from training step.

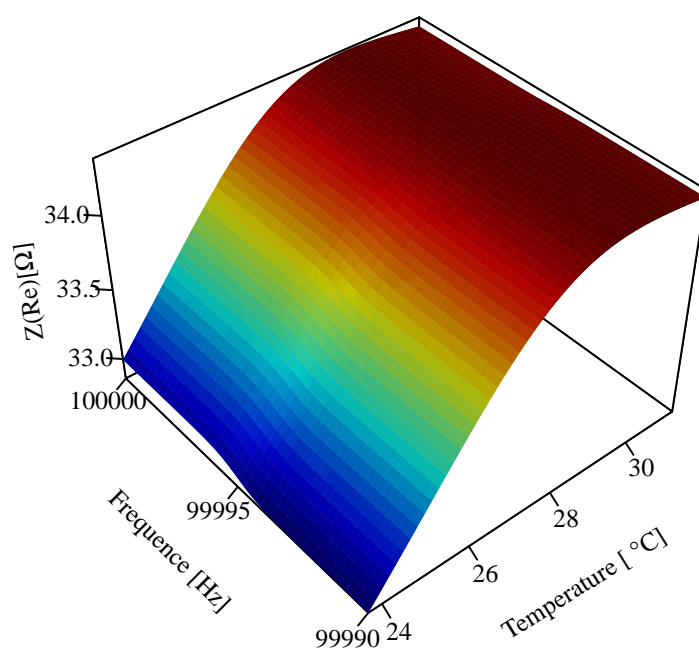


Figure 23: Surface built from an FRBS of the accelerated corrosion experiment, model evaluation step. Source: Freitas et al. [2021a].

The graphical validation was performed from the generated FRBSs, which consists of using an impedance signature of a given validation temperature from Table 3 with the signature predicted by the FRBSs at this same temperature. Three examples are shown in Figures 24, 25 and 26 for temperature values of 30.8 °C, 24.0 °C and 27.2 °C in PZT patches 1, 2 and 3, respectively. In addition to the graphical validation of the model, the validation is performed using the CCD values, according to the equation (3). Figure 27 presents the values obtained for the CCD damage metric for PZT patches 1, 2, and 3.

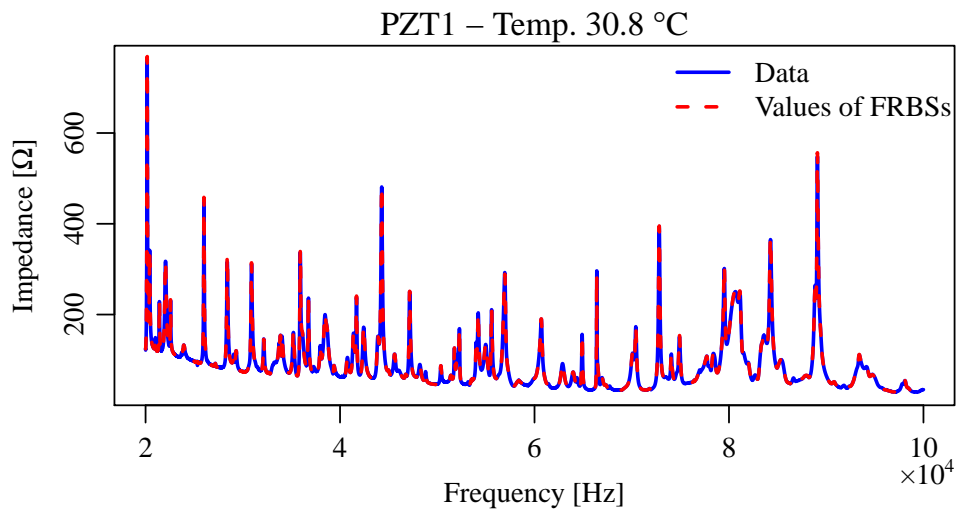


Figure 24: Impedance signatures measured and modeled from FRBSs, at temperature 30.8 °C in PZT1 patch

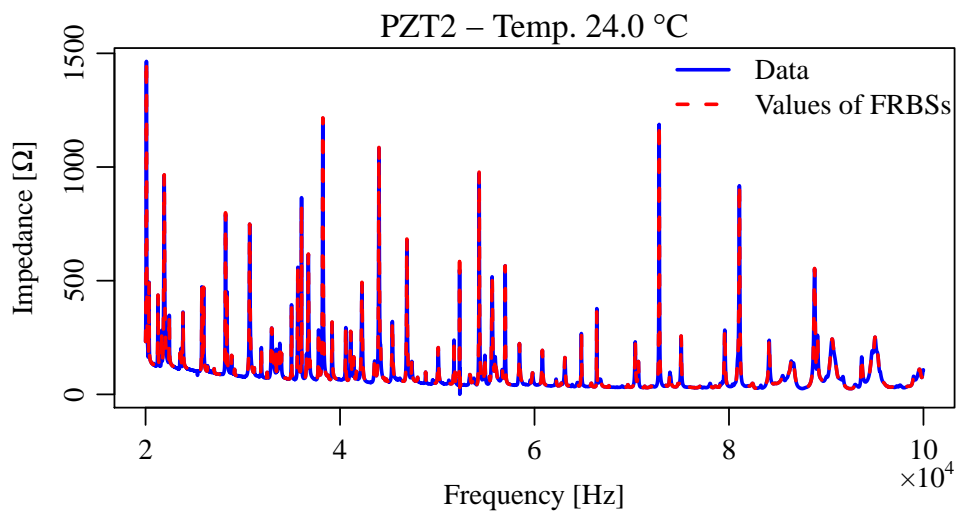


Figure 25: Impedance signatures measured and modeled from FRBSs, at temperature 24 °C in PZT2 patch.

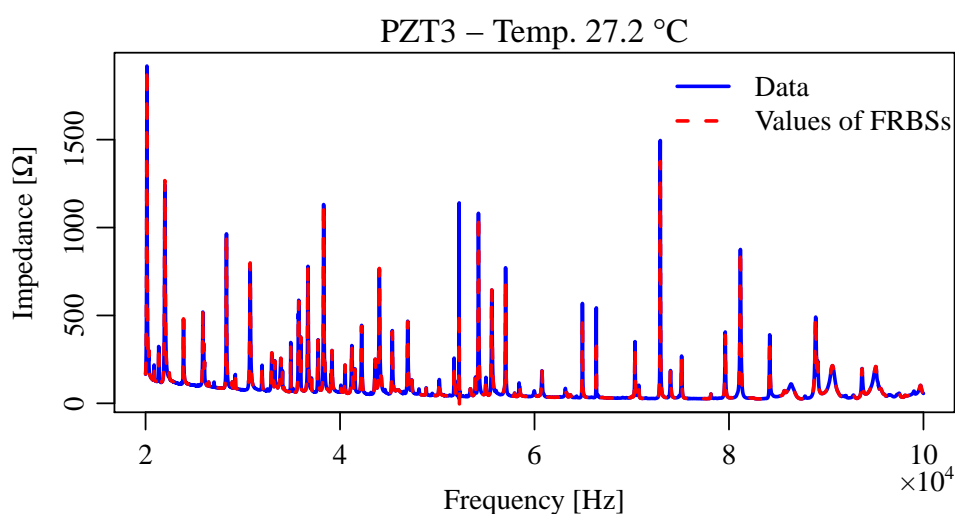


Figure 26: Impedance signatures measured and modeled from FRBSs, at temperature 27.2 °C in PZT3 patch.

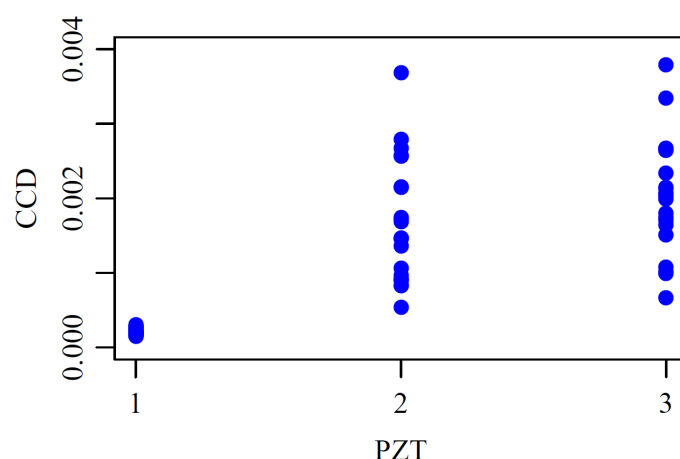


Figure 27: Baselines CCD metrics for the three PZT patches.

The highest CCD value was 0.003800021, and the difference between the impedance signatures collected and those predicted by the FRBSs is not significant (Figures 24, 25 and 26) and therefore built signatures configure baseline signatures.

4.4 Results: Damage Detection

Considering the same data from the first step and using the same settings for the SBRFs, two inputs (temperature and frequency) and one output (impedance), 4000 routines were performed for ANFIS training. All temperatures presented in Table 3 (training and validation), in this situation, were considered for training. Thus, 4000 FRBSs were generated, where the membership function for each entry has the same characteristics as those presented previously.

The four fuzzy rules generated for PZT1 patch and frequencies 99990 and 100000 are given by:

- Rule 1: If (t is High) and (f is High) then $i_1 = 0,4298513t + 0,0002239f + 2,2 * 10^{-9}$
 Rule 2: If (t is Low) and (f is High) then $i_2 = 0,4356641t + 0,0002240f + 2,2 * 10^{-9}$
 Rule 3: If (t is High) and (f is Low) then $i_3 = 0,0542666t + 0,0003258f + 3,2 * 10^{-9}$
 Rule 4: If (t is Low) and (f is Low) then $i_4 = 0,0706055t + 0,0003206f + 3,2 * 10^{-9}$.

Figure 28 shows the surface generated from training.

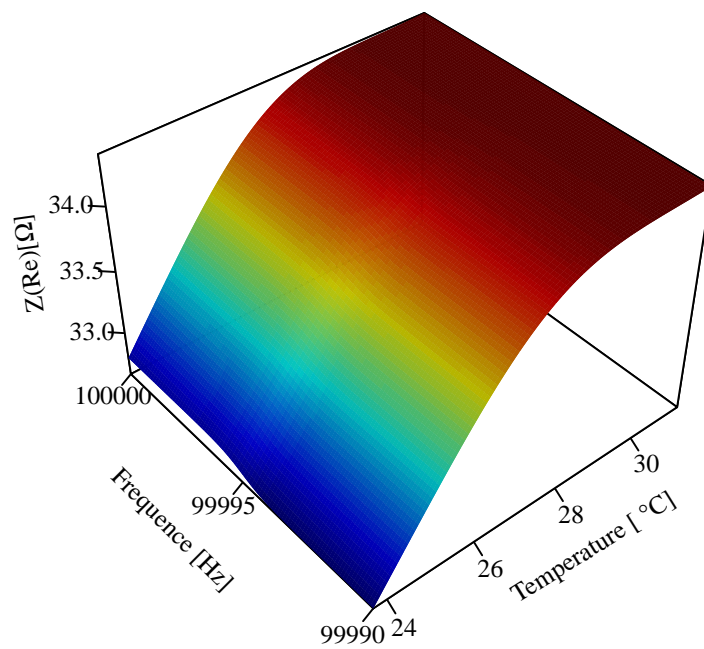


Figure 28: Surface generated from an SBRF from the accelerated corrosion experiment, damage detection step.

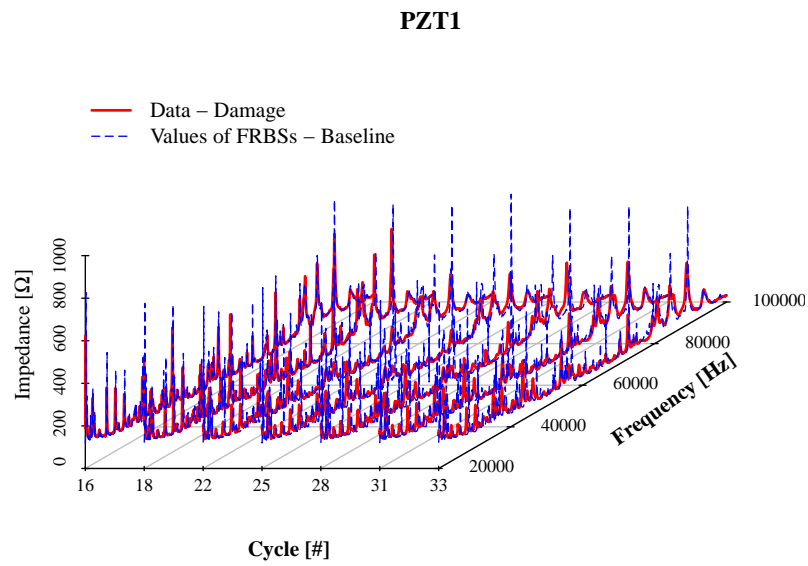


Figure 29: Impedance signatures, in some cycles, in the presence of damage and the corresponding baseline signature modeled from the FRBS, in PZT1 patch. Source: Adapted from Freitas et al. [2021a]

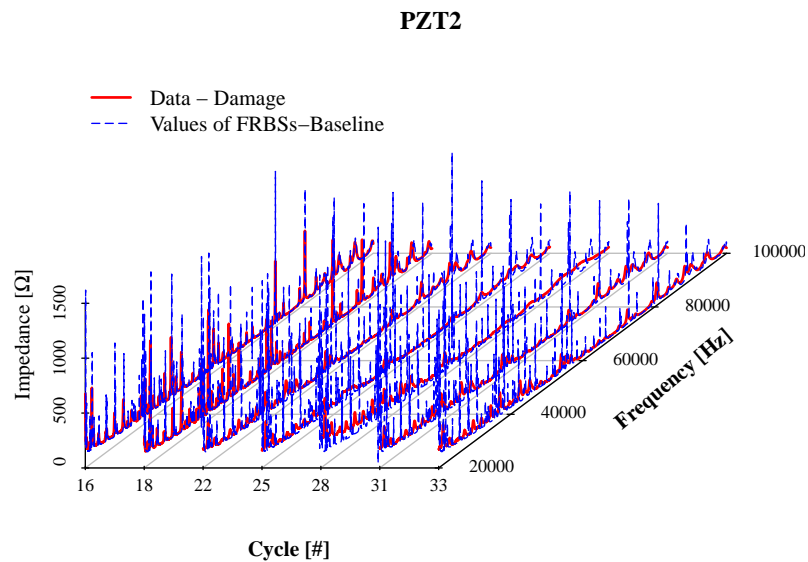


Figure 30: Impedance signatures, in some cycles, in the presence of damage and the corresponding baseline signature modeled from the FRBS, in PZT2 patch. Source: Adapted from Freitas et al. [2021a]

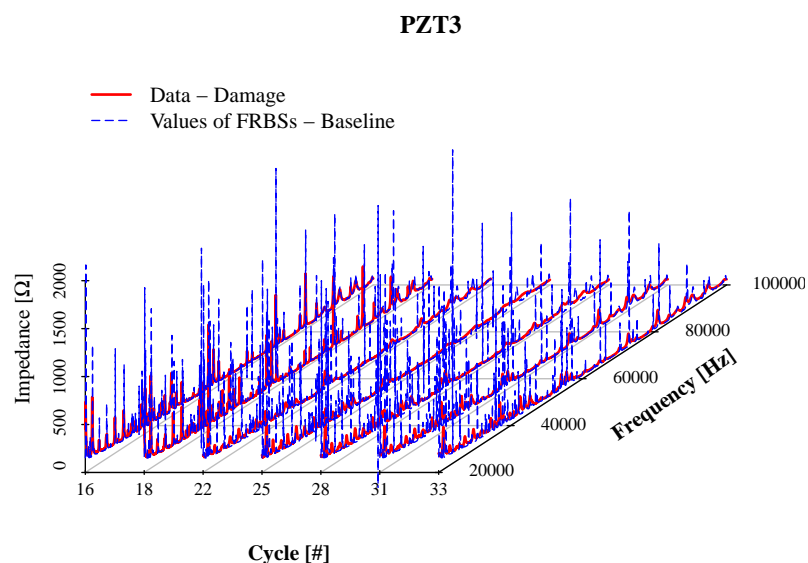


Figure 31: Impedance signatures, in some cycles, in the presence of damage and the corresponding baseline signature modeled from the SBRFs, in PZT3 patch. Source: Adapted from Freitas et al. [2021a]

For the validation part, a data set containing 33 distinct cycles were used, with 30 repetitions each, in which the first 14 cycles correspond to the original condition of the steel beams. In contrast, the acid mist was present in cycles 15 to 33, causing structural changes.

Figures 29, 30 and 31 show, in some corrosion cycles, a comparison of a signature with the damage corresponding to that cycle and a baseline signature modeled by the FRBSs, showing the difference between the measured signature and that provided by the FRBSs. This fact reveals that the methodology was able to distinguish baseline signatures from signatures with the presence of damage.

In Figures 32, 33 and 34 it is possible to observe the values of the CCD metric found according to the corrosion cycle. The cycles represent different moments of collection of impedance signatures due to the time elapsed in the execution of the experiment, called corrosion cycles. The CCD was calculated considering the median of the real signature collected in a given cycle and the signature of the baseline predicted by the SBRFs.

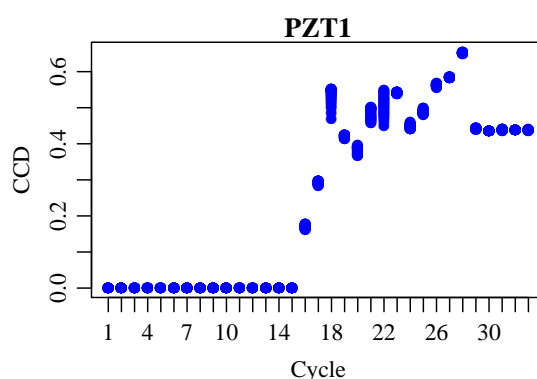


Figure 32: CCD values in the different conditions of the experiment in PZT1 patch.
 Source: Adapted from Freitas et al. [2021a]

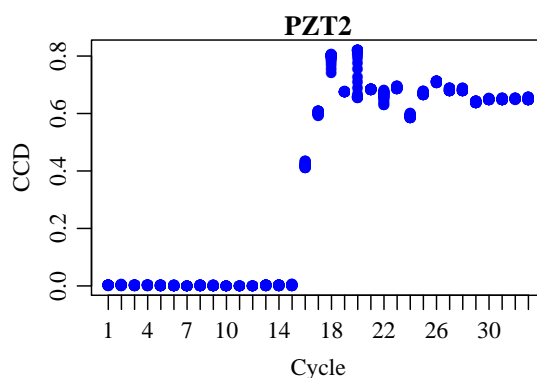


Figure 33: CCD values in the different conditions of the experiment in PZT2 patch.
 Source: Adapted from Freitas et al. [2021a]

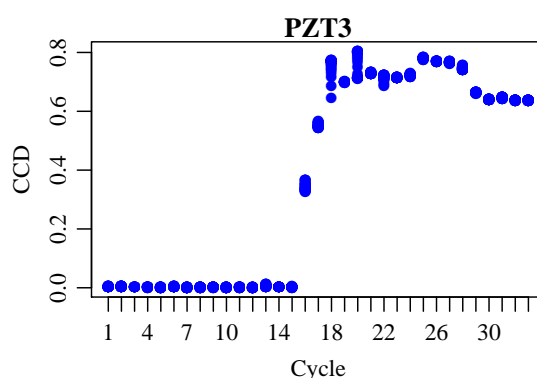


Figure 34: CCD values in the different conditions of the experiment in PZT3 patch.
 Source: Adapted from Freitas et al. [2021a]

The evolution of CCD values over the cycles indicates that the increase in acid caused a significant change in the values of this metric, mainly from the third cycle onwards. In cycles 15 and 16, this change was smaller, which is understandable, given that the acid action time, responsible

for corrosion, was still small when compared to cycles identified by higher numbers. Another influence that must be considered is the accumulated oxides of the acid over time, which can affect the CCD value.

As mentioned, for the first phase, the model validation was perfect since the highest computed CCD was not significant and therefore did not indicate a structural change, which was expected given that the training and validation data were all baselines. The highest CCD baseline value obtained is 0.003800021, a shallow difference, meaning that the simulated signatures would be considered a sample of the actual baselines.

From the results presented in Figures 32, 33 and 34, it is possible to distinguish situations with damage, by the application of acid, from those without damage characterized by the absence of acid from the constructed FRBSs.

Therefore, the two steps of the experiment showed excellent results, providing an effective technique for data normalization.

5 ANFIS: Structural Change from the Increment of Mass

In this experiment is identified the level of damage in aluminum beams.

The experimental data used for the application of the methodology were obtained from eight specimens in an experiment in a climatic chamber (Figure 8), with temperature and humidity control installed in the Structural Mechanics Laboratory (LMEst) from the School of Mechanical Engineering of the Federal University of Uberlandia. This work was presented at the X Scientific Initiation Exhibition of FAMAT Freitas et al. [2021b].

5.1 Beams with added mass

A PZT patch was bonded to each beam to measure impedance signatures. Thus, the setup of the experiment began.

The specimens were aluminum beams 500 mm long, 38 mm wide, and 3.2 mm thick, and in each of these structures a 1 mm thick and 20 mm diameter PZT patch was bonded to 100 mm from the end. The processes of promoting structural changes adopted were the addition of masses glued 380 mm from the center of the PZT patch at the opposite end of the beam. These specimens can be seen in Figure 35.

The experiment was carried out to use the addition of mass to change the structure behavior. Figure 35 shows the eight beams or specimens tested in the bi-supported boundary condition. A polystyrene foam (styrofoam) was used to make the base of the beams to reduce possible vibration transfer from the chamber and external environment to the specimens during measurements.

The experiment was carried out at temperatures -10 , 0 , 10 , and 20°C . For cases of damage by adding masses, five levels of failure were considered, the first being without additional mass (baseline signature) and four progressive levels. For a slight controlled variation in the experiment, the masses were added to the specimens according to Table 4.



Figure 35: Instrumented and identified beams.

Table 4: Mass values by damage level and beam.

	<i>Baseline</i>	Damage 1 (g)	Damage 2 (g)	Damage 3 (g)	Damage 4 (g)
Beam 1	No damage	0.6	1.1	1.6	2.2
Beam 2	No damage	0.6	1.1	1.6	2.2
Beam 3	No damage	0.6	1.1	1.6	2.8
Beam 4	No damage	0.6	1.1	1.6	2.7
Beam 5	No damage	0.6	1.1	1.6	2.2
Beam 6	No damage	0.6	1.1	1.6	2.2
Beam 7	No damage	0.6	1.1	1.6	2.2
Beam 8	No damage	0.6	1.1	1.6	2.2

As represented in Table 4, damage level 4 was not the same mass value in all beams. In Figure 36, it is possible to identify the different configurations, in three beams, for damage level 4.

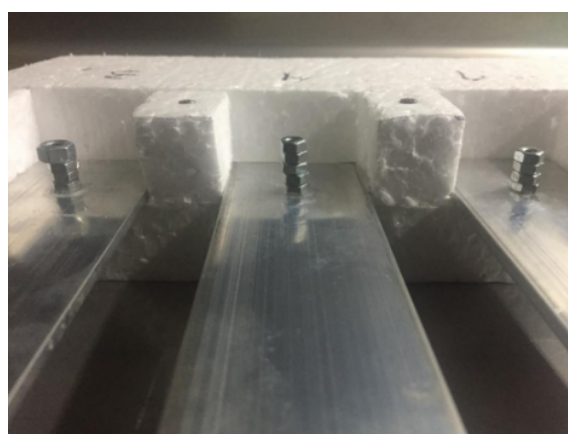


Figure 36: Different mass values: damage level 4, on beams 3, 4 and 1.

For each experimental condition, determined by the temperature value and damage level, 20 impedance signatures were collected in each of the eight beams, totaling 3200 signatures.

Previous exploratory analyses found that the data obtained in the PZT2 patch did not represent typical impedance signatures due to acquisition problems; therefore, these data were not used in the study.

5.2 FRBSs Evaluation and Modeling Criteria

One of the difficulties encountered in this methodology is quantifying the damage from the metrics since their behavior does not follow the change made in the structure.

The SBRFs modeled for this experiment consist of two inputs, CCD and $RMSD1$, calculated from the investigated signature and respective baseline.

In the validation step, the FRBS output value was compared using the pairs CCD and $RMSD1$, continuously computed with identical impedance signatures. Figure 37 presents a schematic of the mathematical modeling for the experiment, in which the red rectangle represents the data obtained in the investigation; the blue rectangles represent the stage of training and construction of the FRBSs, and in green the validation step.

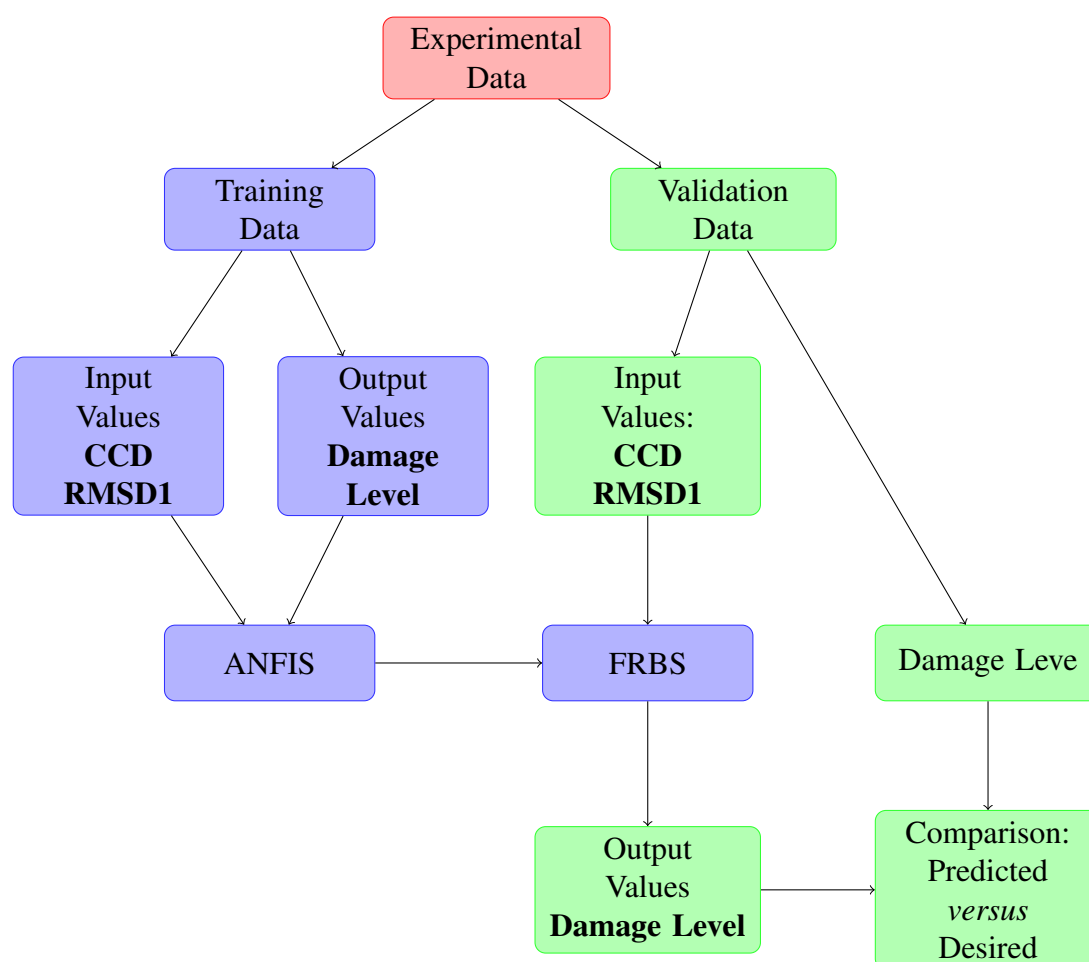


Figure 37: Scheme of the mathematical modeling used for the experiment

5.3 Results: Model evaluation

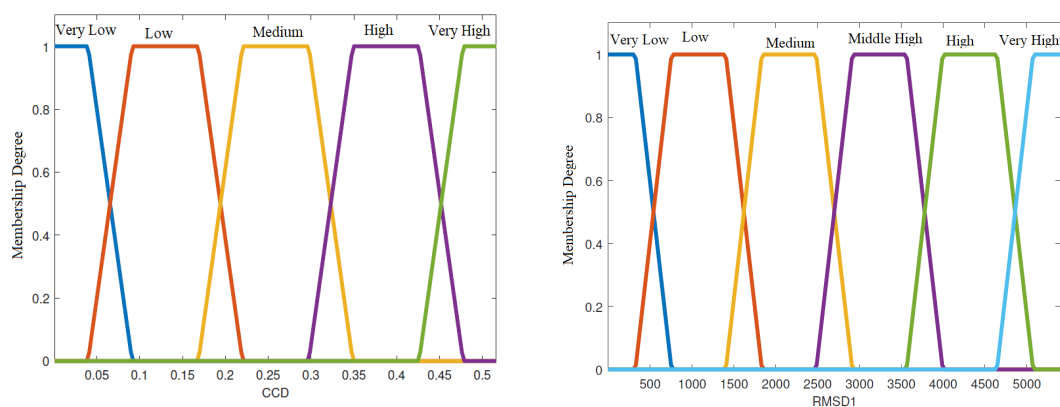
Considering the data set from the experiment for each PZT patch, the CCD (equation (3)) and $RMSD1$ (equation (1)) metrics are computed for each signature, which are the variables of SBRF entries. The output variable is the damage level obtained from Table 4, observed in the experiment by increasing the mass in the structure.

The training step in ANFIS (MATLAB) is performed by taking the even or odd indices of the combinations of CCD and $RMSD1$ metrics, taken according to the order of signature collections

or, in some training, in ascending order of RMSD1 metric. The choice of the data set used in training is the one that best approached the expected results. Thus, data set not used in training were considered for validation. Therefore, the number of FRBSs built is 28, as four temperatures and seven PZT patches were considered in the study.

For the training of each model, different amounts of membership functions are configured. Figure 38 illustrates the membership functions generated by training in the PZT1 patch for a temperature of 20°C . The input variables are:

- CCD values: the domain $[0, 0.52]$, representing the ranges $[0, 0.1]$, $[0.04, 0.22]$, $[0.17, 0.35]$, $[0.3, 0.48]$ and $[0.43, 0.52]$, with the linguistic terms: Very Low, Low, Medium, High and Very High, respectively.
- RMSD1 values: the domain $[1.25, 5402]$, representing the ranges $[1, 25, 757.3]$, $[325.3, 1837]$, $[1405, 2918]$, $[2486, 3998]$, $[3566.5078]$, $[4646.5402]$, with the linguistic terms: Very Low, Low, Medium, Medium High, High and Very High, respectively.



a: Membership functions to input CCD metric. b: Membership functions to input RMSD1 metric.

Figure 38: Membership functions to inputs.

Figure 39 shows the surface generated from training. The system generated 30 possible fuzzy rules and output function parameters.

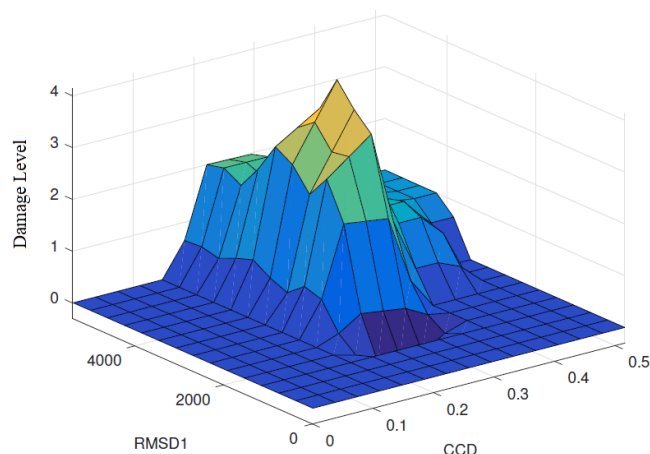


Figure 39: Surface generated from a FRBS. Source: Freitas et al. [2021a]

FRBSs are validated by approximating the output values of FRBSs by the nearest integer to the damage level 0, 1, 2, 3, or 4 and comparing it with the damage observed in the experiment. Furthermore, Table 5 presents the percentages of correct answers for the FRBSs, considering 50 values of CCD and RMSD1 metrics for each temperature and PZT patch.

Table 5: FRBSs hit percentages according to each temperature.

PZT Patch	Temperature [$^{\circ}C$]			
	-10	0	10	20
1	88%	96%	88%	100%
3	100%	90%	100%	98%
4	94%	100%	100%	100%
5	88%	94%	100%	100%
6	80%	96%	100%	100%
7	100%	92%	100%	94%
8	88%	88%	98%	88%

The results presented in Table 5 show that a high proportion of correct answers is obtained from the proposed model. The lowest hit percentage is 80% in the PZT6 patch with $-10^{\circ}C$, while in 12 PZT patches and temperature combinations, there was a 100% hit. The results obtained through the FRBSs showed that the technique could be a promising tool for damage detection since it achieved an average percentage of hits of 95%, comparing the predicted and the expected.

Finally, the next section addresses the conclusion of this chapter.

6 Concluding Remarks

ANFIS combines artificial intelligence and fuzzy set theory, providing efficient models for different problems. The combination of these two tools, through ANFIS, can improve the ability to represent and infer models.

In the application example, the methodology used for temperature compensation was efficient because the values of the CCD metric between the modeled signatures and the measured baseline signatures were low. Therefore, this approach through ANFIS is a suitable implementation of the

baseline state representation. The proposed method is effective for data normalization, and it is possible to build reference signatures in the temperature range collected in the baseline gathering process. In other applications, wider temperature ranges can be considered, and these applications can be extended to different structures such as bridges, buildings, and tanks.

From the second application proposed in this chapter, it can be concluded that the percentage of correctness regarding the state of the structure was high, 95%. In this case, ANFIS was also used to model FRBSs in which the input variables were the CCD and the RMSD1 metrics.

The mathematical approach proposed in this chapter for the study of electromechanical impedance-based SHM provided promising results. This methodology allows the tuning of the functions of the input and output variables. It is a mathematical tool with great potential as it combines the learning characteristics of neural networks with the interpretability of Fuzzy Rule-Based Systems.

Acknowledgements

The authors are thankful to Petrobras for the financial support via the research project (RD).

References

- K. E. I. Alfaro. Aplicações da extensão de zadeh na dinâmica do hiv com retardo fuzzy sob tratamento antirretroviral. Master's thesis, Universidade Federal de Uberlândia-UFU, 2019.
- P. Antognetti and V. Milutinovic. Neural networks: Concepts, applications, and implementations. *New Jersey: Prentice Hall*, page 310, 1991.
- L. C. Barros and R. Bassanezi. *Tópicos de lógica fuzzy e biomatemática*. UNICAMP/IMECC, Campinas, SP, 2015.
- A. Bastani, H. Amindavar, M. Shamshirsaz, and N. Sepehry. Identification of temperature variation and vibration disturbance in impedance-based structural health monitoring using piezoelectric sensor array method. *Structural Health Monitoring*, 11(3):305–314, 2012.
- S. G. T. A. Bezerra. Reservoir computing com hierarquia para previsão de vazões médias diárias. Master's thesis, Universidade de Pernambuco, 2016.
- S. Bhalla, A. S. K. Naidu, and C. K. Soh. Influence of structure-actuator interactions and temperature on piezoelectric mechatronic signatures for nde. In *Smart Materials, Structures, and Systems*, volume 5062, pages 263–269, Bangalore, India, 2003. International Society for Optics and Photonics.
- A. P. Braga, A. C. P. L. F. Carvalho, and T. B. Ludemir. Redes neurais artificiais: Teoria e aplicações. *Rio de Janeiro: Livros Técnicos e científicos Editora S. A*, page 262, 2000.
- D. P. L. Ferreira. Sistema p-fuzzy aplicado às equações diferenciais parciais. Master's thesis, Universidade Federal de Uberlândia - UFU, 2011.
- R. M. Finzi Neto, V. Steffen, D. A. Rade, C. A. Gallo, and L. V. Palomino. A low-cost electromechanical impedance-based SHM architecture for multiplexed piezoceramic actuators. *Structural Health Monitoring*, 10:391–402, 2010. doi: 10.1177/1475921710379518.
- C. A. G. Fonseca. Estrutura ANFIS modificada para identificação e controle de plantas com ampla faixa de operação e não linearidade acentuada. Master's thesis, Universidade Federal do Rio Grande do Norte - UFRGN, 2012.

- F. A. Freitas. Aplicação da teoria dos conjuntos fuzzy no estudo da impedância eletromecânica. Master's thesis, FAMAT, Universidade Federal de Uberlândia, 2021.
- F. A. Freitas, R. M. Jafelice, J. W. d. Silva, D. d. S. Rabelo, Q. S. S. Nomelini, J. d. R. V. d. Moura, C. A. Gallo, M. J. d. Cunha, and J. E. Ramos. A new data normalization approach applied to the electromechanical impedance method using adaptive neuro-fuzzy inference system. *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, 43(11):1–13, 2021a.
- F. A. Freitas, R. S. M. Jafelice, J. W. da Silva, C. A. Gallo, and J. E. Ramos. Detecção de danos em vigas de alumínio via modelagem fuzzy, 2021b. URL http://www.eventos.ufu.br/sites/eventos.ufu.br/files/documentos/caderno_de_resumos_2021_0.pdf. X Mostra de Iniciação Científica da FAMAT, Acesso: 28/06/2021.
- F. A. Freitas et al. Aplicação da teoria dos conjuntos fuzzy no estudo da impedância eletromecânica. 2021c.
- C. Fresno, A. S. Llera, and E. A. Fernandez. *ANFIS: Adaptive Neuro Fuzzy Inference System in R*, 2015. URL <https://CRAN.R-project.org/package=anfis>. R package version 0.99.1.
- B. M. Giancesini, N. E. Cortez, R. A. Antunes, and J. Vieira Filho. Method for removing temperature effect in impedance-based structural health monitoring systems using polynomial regression. *Structural Health Monitoring*, 20(1):202–218, 2021.
- R. S. M. Jafelice, L. C. de Barros, and R. C. Bassanezi. Teoria dos conjuntos fuzzy com aplicações (e-book). São Carlos, SP: SBMAC, 17:119, 2012.
- J.-S. R. Jang. Adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man, and Cybernetics*, 3(23), 1993.
- K.-Y. Koo, S. Park, J.-J. Lee, and C.-B. Yun. Automated impedance-based structural health monitoring incorporating effective frequency shift for compensating temperature effects. *Journal of intelligent material systems and structures*, 20(4):367–377, 2009.
- K. Krishnamurthy, F. Lalande, and C. A. Rogers. Effects of temperature on the electrical impedance of piezoelectric sensors. In *Smart Structures and Materials 1996: Smart Structures and Integrated Systems*, volume 2717, pages 302–310. International Society for Optics and Photonics, 1996.
- I. I. C. Maruo, G. de Faria Giachero, V. S. Júnior, and R. M. F. Neto. Electromechanical impedance - based structural health monitoring instrumentation system applied to aircraft structures and employing a multiplexed sensor array. *Journal of Aerospace Technology and Management*, pages 294–306, 2015. doi: 10.5028/jatm.v7i3.447.
- W. S. McCulloch and W. Pitts. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5:115–133, 1943.
- J. R. V. Moura and V. Steffen. Impedance-based health monitoring: Frequency band evaluation. *Proceedings of the 22nd International Modal Analysis Conference (IMAC XXII)*, 2004.
- N. N. Nardez. Obtenção do parâmetro pco de antenas gnss por meio de sistemas baseado em regras fuzzy. Master's thesis, UFP, 2015.

- M. Negnevitsky. *Artificial Intelligence: A Guide to Intelligent Systems*. Pearson Education Limited, Addison Wesley, 2 edition, 2005.
- L. Palomino. Análise das métricas de dano associadas à técnica da impedância eletomecânica para o monitoramento de integridade estrutural. Master's thesis, Universidade Federal de Uberlândia - UFU, 2008.
- G. Park, K. Kabeya, H. H. Cudney, and D. J. Inman. Impedance-based structural health monitoring for temperature varying applications. *JSME International Journal Series A Solid Mechanics and Material Engineering*, 42(2):249–258, 1999.
- E. Pourjavad and A. Shahin. The application of Mamdani fuzzy inference system in evaluating green supply chain management performance. *International Journal of Fuzzy Systems*, 20(3): 901–912, 2018.
- R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2020. URL <https://www.R-project.org/>.
- D. S. Rabelo, K. M. Tsuruta, D. D. de Oliveira, A. A. Cavalini, R. M. F. Neto, and V. Steffen. Fault detection of a rotating shaft by using the electromechanical impedance method and a temperature compensation approach. *Journal of Nondestructive Evaluation*, 2017.
- H. R. Sandmann. Predição não-linear de séries temporais usando sistemas de arquitetura neuro-fuzzy. Master's thesis, Escola Politécnica, Universidade de São Paulo, São Paulo, 2006.
- N. Sepehry, M. Shamshirsaz, and F. Abdollahi. Temperature variation effect compensation in impedance-based structural health monitoring using neural networks. *Journal of Intelligent Material Systems and Structures*, 22(17):1975–1982, 2011.
- R. N. F. Silva, K. M. Tsuruta, D. S. Rabelo, R. M. F. Neto, A. A. C. Jr., and V. S. Jr. Impedance-based structural health monitoring applied to steel fiber-reinforced concrete structures. *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, page 42:383, 2020. doi: <https://doi.org/10.1007/s40430-020-02458-4>.
- E. D. A. Sodre, W. S. Mota, and B. D. S. Alencar. Sistema híbrido neuro-fuzzy para análise de segurança em tempo real de sistemas de potência. *Congresso Brasileiro de Redes Neurais*, 8, 2016.
- F. P. Sun, Z. Chaudhry, C. Liang, and C. Rogers. Truss structure integrity identification using pzt sensor-actuator. *Journal of Intelligent material systems and structures*, 6(1):134–139, 1995.
- M. Valenca. *Fundamentos das Redes Neurais: exemplos em Java*. Recife, Editora Livro Rápido – Elógica, 2 edition, 2008.
- L. Zadeh. Fuzzy sets. *Information and control*, 8(3):338–353, 1965.
- D. Zhou, J. K. Kim, D. S. Ha, J. D. Quesenberry, and D. J. Inman. A system approach for temperature dependency of impedance-based structural health monitoring. In *Smart Sensor Phenomena, Technology, Networks, and Systems 2009*, San Diego, USA, 2009. International Society for Optics and Photonics.

Chapter 13

On Kriging Techniques & Impedance-based SHM as Applied to Damage Detection in 2D Structures

Chapter details

Chapter DOI:

<https://doi.org/10.4322/978-65-86503-88-3.c13>

Chapter suggested citation / reference style:

Pereira, Paulo E. C., et al. (2022). “On Kriging Techniques & Impedance-based SHM as Applied to Damage Detection in 2D Structures”. In Jorge, Ariosto B., et al. (Eds.) *Uncertainty Modeling: Fundamental Concepts and Models*, Vol. III, UnB, Brasilia, DF, Brazil, pp. 427–458. Book series in Discrete Models, Inverse Methods, & Uncertainty Modeling in Structural Integrity.

P.S.: DOI may be included at the end of citation, for completeness.

Book details

Book: Uncertainty Modeling: Fundamental Concepts and Models

Edited by: Jorge, Ariosto B., Anflor, Carla T. M., Gomes, Guilherme F., & Carneiro, Sergio H. S.

Volume III of Book Series in:

Discrete Models, Inverse Methods, & Uncertainty Modeling in Structural Integrity

Published by: UnB City: Brasilia, DF, Brazil Year: 2022

DOI: <https://doi.org/10.4322/978-65-86503-88-3>

On Kriging Techniques & Impedance-based SHM as Applied to Damage Detection in 2D Structures

Paulo Elias C. Pereira^{1*}, Daniel R. Gonçalves², Stanley W. F. Rezende^{1*}
Jose dos Reis V. Moura Jr² and Roberto M. Finzi Neto¹

¹Post-Graduate Program - Mechanical Engineering, Federal University of Uberlandia, Brazil. e-mail: paulo_elias_carneiro@ufcat.edu.br; stanley_washington@ufu.br; finzi@ufu.br

²Post-Graduate Program - Modeling and Optimization, Federal University of Catalão, Brazil. e-mail: danielresendeg@gmail.com; zereis@ufcat.edu.br

*Corresponding author

Abstract

In Impedance-based Structural Health Monitoring Systems, it is essential to identify the location of possible damage so that corrective measures can be taken promptly. Several approaches have been developed to predict the location of damage, including the one based on indicator kriging. This chapter illustrates the computational packages in Python for carrying out estimates by kriging methods. A case study of the location of structural damage in the center of an aluminum plate using indicator kriging is presented. They are detailed, sequentially, the necessary steps, and discussing each one. The results obtained with the case study showed that it was possible to map the place with the highest probability of occurrence of damage, which proved to be consistent with the actual position of the same, which highlights the ability of the proposed approach to predict the location of structural damage. In addition, perspectives regarding future research involving the use of the technique are discussed.

Keywords: Impedance-based structural health monitoring; Kriging methods; Indicator kriging; Damage location

1 Introduction

The electromechanical impedance technique has been vastly applied to the monitoring of structures due to its low cost - concerning other approaches - and its characteristic of being a non-destructive evaluation (NDE) method (Giurgiutiu [2014]). Further, the method can access the structure's conditions at relatively high frequencies (usually between 30 kHz and 400 kHz), which reduces the environment's interferences, allowing the structure's monitoring at small scales, such

as cracks, delamination, and disbonds, not detectable by other methods (Bhalla and Soh [2012]; Zagrai and Giurgiutiu [2009]).

Post-processed information obtained by ISHM refers to damage metric data (e.g., RMSD and CCD indexes), which can indicate the presence of damage in the structure by comparing the electrical impedance signatures at any moment in time and at the initial (pristine) condition of the structure (Giurgiutiu [2014]; Giurgiutiu and Rogers [1998]). As each damage metric data is related to a specific actuator, it's necessary to interpolate all data collected, each assigned to a point in the space (actuator position), to identify the damage location.

The works of Kravolec et al. [2018] and Cherrier et al. [2013] evidenced the existence of a correlation between the electrical impedance signatures and the distance from the damage, which suggests the presence of spatial correlation between the damage metric data arranged in the form of a network. Results derived from the application of kriging methods (Gonçalves et al. [2021]; Gonçalves et al. [2020]) for the interpolation of RSMD values in an aluminum plate revealed a spatial correlation between these, allowing the identification of the damage location, especially when using the indicator kriging method. Despite the positive results achieved, more research on this approach is necessary, which can allow several new possibilities.

This chapter presents the theoretical concepts related to indicator kriging for applications in structural health monitoring. First, concepts of ISHM and kriging are exposed, followed by computational routines on Python packages for estimations based on kriging methods. Finally, a case study is presented regarding the location of damage from indicator kriging, where the steps involved are sequentially detailed, accompanied by source codes in Python language.

2 Fundamentals of ISHM and Kriging

2.1 Fundamentals of ISHM

There are several techniques for monitoring the occurrence and propagation of structural damage to increase the useful life, improve the performance of structures/equipment and reduce costs (Park et al. [2003]). Among the SHM techniques, there is the Electromechanical Impedance-based Structural Health Monitoring (EMI-SHM), which is based on the electromechanical coupling that results from the bonding of piezoelectric transducers to the structure or equipment that will be monitored, keeping the impedance function defined, which depends on the mechanical characteristics of the transducer and structure to be monitored (Lin and Giurgiutiu [2006]; Park et al. [2003]).

In the EMI-SHM technique, PZT transducers are usually excited by a sinusoidal waveform with an amplitude of approximately 1V RMS (effective voltage) and a frequency range of 10 to 250 kHz (Raju [1997]), or even as high as 1,000 kHz, depending on the structure and type of application (Giurgiutiu et al. [1999]). The lower frequency band covers a larger detection area, while the higher frequency band can determine the location of damage (Sun et al. [1995]). Among the advantages of the high-frequency response is that the wavelength of the signal applied to the structure is short enough to detect even small initial cracks. Such cracks can expand and cause serious failures depending on the structure (Park et al. [2003]).

Electromechanical models were formulated to describe the relationship between the structure and the piezoelectric transducer. Liang et al. [1994] define the electrical impedance of the transducer $Z_E(\omega)$ associated with the mechanical impedance of the structure $Z_s(\omega)$ by Eq. 1.

$$Z_E(\omega) = \frac{1}{j\omega\tau} \left(\epsilon_{33}^T - \frac{Z_s(\omega)}{Z_s(\omega) + Z_p(\omega)} d_{3X}^2 Y_{XX}^E \right)^{-1} \quad (1)$$

where $Z_p(\omega)$ corresponds to the mechanical impedance of the transducer; ω is the angular

frequency; τ , a geometric constant; ϵ_{33}^T , the dielectric constant for constant mechanical stress (T); Y_{XX}^E , Young's modulus for a constant electric field (E); d_{3X}^2 , a dielectric constant; and j , an imaginary unit.

In the one degree of freedom model presented by Liang et al. [1994], the influence of the adhesive layer (glue) between the transducer and the structure is not considered. Thus, the PZT patch is excited at a high frequency and returns the impedance curve of the structure studied. Changes in the mechanical impedance of the structure due to a failure lead to a modification of the signal (electrical impedance) emitted by the PZT patches coupled or embedded in the structure, whose signal is measured within a frequency range. Changes in structural conditions are then identified by comparing the impedance signatures before and after the fault's occurrence (Maruo et al. [2015]; Bhalla and Soh [2012]; Baptista and Vieira Filho [2010]; Zagrai and Giurgiutiu [2009]; Peairs et al. [2007b]; Peairs et al. [2007a]).

The measurement of the difference between the electrical impedance signals before and after the occurrence of changes in the structure's status is done through some indexes, among which we have the Root Mean Square Deviation-RMSD (Eq. 2) and the Correlation Coefficient Deviation-CCD (Eq. 3).

The RMSD (Eq. 2) measures the similarity of impedance attributed to the same frequency and is supported by the Euclidean standard (Giurgiutiu [2014]; Giurgiutiu and Rogers [1998]). This indicator is based on the comparison between the impedance values before the damage, that is, in the pristine condition ($Z_E^0(k)$), whose signal is also called the baseline, and those obtained at some point in the future ($Z_E(k)$), all measured at several k points in the frequency domain, whose range varies from ω_I (initial frequency) to ω_F (final frequency).

$$RMSD = \sqrt{\frac{\sum_{k=\omega_I}^{\omega_F} [Z_E(k) - Z_E^0(k)]^2}{\sum_{k=\omega_I}^{\omega_F} [Z_E^0(k)]^2}} \quad (2)$$

Unlike the RMSD, which measures the average distance/difference between the impedance curves, the CCD (Eq. 3) analyzes the impedance curves globally, based on the correlation coefficient (Marqui et al. [2008]).

$$CCD = 1 - \frac{1}{n} \sum_{i=1}^n \left(\frac{(\text{Re}(Z_{\text{base},i}) - \text{Re}(Z_{\text{base}}^-)) (\text{Re}(Z_{FIS,i}) - \text{Re}(Z_{FIS}^-))}{S_{FIS} S_{\text{base}}} \right) \quad (3)$$

In Eq. 3, n is the number of frequency-domain values, $\text{Re}(Z_{(base,i)})$ is the real part of the impedance signal in the pristine condition, $\text{Re}(Z_{base})$ is the average of the real part of the impedance signal in the pristine status, $\text{Re}(Z_{(FIS,i)})$ is the real part of the PZT patch impedance signal after damage, $\text{Re}(Z_{FIS})$ is the average of the real part of the PZT patch impedance signal after damage, S_{FIS} the standard deviation of the real part of the PZT impedance signal that represents the damage and S_{base} is the standard deviation of the real part of the PZT impedance signal that represents the signal in the intact condition.

An impedance analyzer such as the HP4194A is the simplest way to determine the electromechanical impedance of smart structures (structures containing PZT transducers). However, the equipment cost is approximately US\$ 40,000 and weighs about 30 kg. There are other options with limited functions that cost less than \$2,000. However, they are still bulky (weighing several pounds). As a consequence, researchers have been looking for alternative ways to perform this task (Maruo et al. [2015]).

2.2 Semivariogram

Estimating unsampled points and regions using kriging methods are based on the spatial correlation of the analyzed variable. The spatial continuity means that sample (or input) data close to each other have similar values attributed to them, whose difference increases as the distance between them becomes higher (Revuelta [2018]; Chilès and Delfiner [2012]). In this sense, closer sample (or input data) points have a more significant correlation with each other than those more distant - spatially - among each other.

The quantification of the spatial continuity is done using the semivariogram function ($\gamma(h)$), defined in the Eq. 4, which calculates the mean quadratic difference between the sample (or input) point pairs $z(x_i)$, located at a point x_i of the - stationary - domain being analyzed, and $z(x_i + h)$, located at the $x_i + h$, distant h from the first point, where h refers to the distance vector between the data being compared (Journel and Huijbregts [1978]). The term $N(h)$ represents the number of the sample (or input) data pairs in the direction under analysis.

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} (z(x_i) - z(x_i + h))^2 \quad (4)$$

The calculation of the semivariogram for several classes of distances in a specific spatial direction provides an inventory of the continuity in this one in the form of a graphic (experimental semivariogram) comparing the semivariance values ($\gamma(h)$) and their respective classes of distances (Figure 1). The growth of the function represents the progressive loss of spatial continuity between the compared sample (or input) pairs as the distance between them increases (Abzalov [2016]; Chilès and Delfiner [2012]; Sinclair and Blackwell [2002]). The total loss of continuity between the input points is evidenced at the point from which the growth of the function stops, and the next ones stay oscillating, normally around the variance *a priori*, which recurrently corresponds to the variance of the variable being analyzed within the domain.

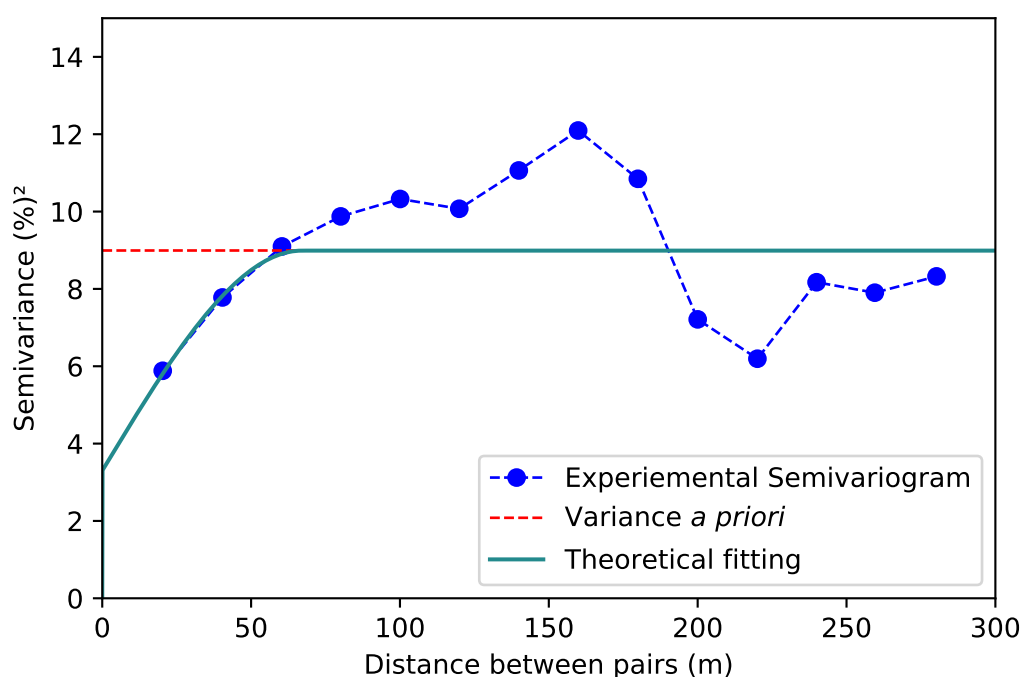


Figure 1: Experimental semivariogram with its respective theoretical fitting

Mathematical functions must adjust experimental semivariograms for their use in the estimation processes (kriging equations). The justifications for this encompass several aspects, whose arguments and discussions can be found in Journel and Huijbregts [1978], Rossi and Deutsch [2014] and Revuelta [2018]. A semivariogram model can be constituted of one or more mathematical functions, depending on the shape of the experimental semivariogram (Kitanidis [1997]). Some mathematical functions for semivariogram fitting include the following (Figure 2): Spherical, exponential, and Gaussian, which are characterized to have a sill (semivariance value) limiting the increase of the function.

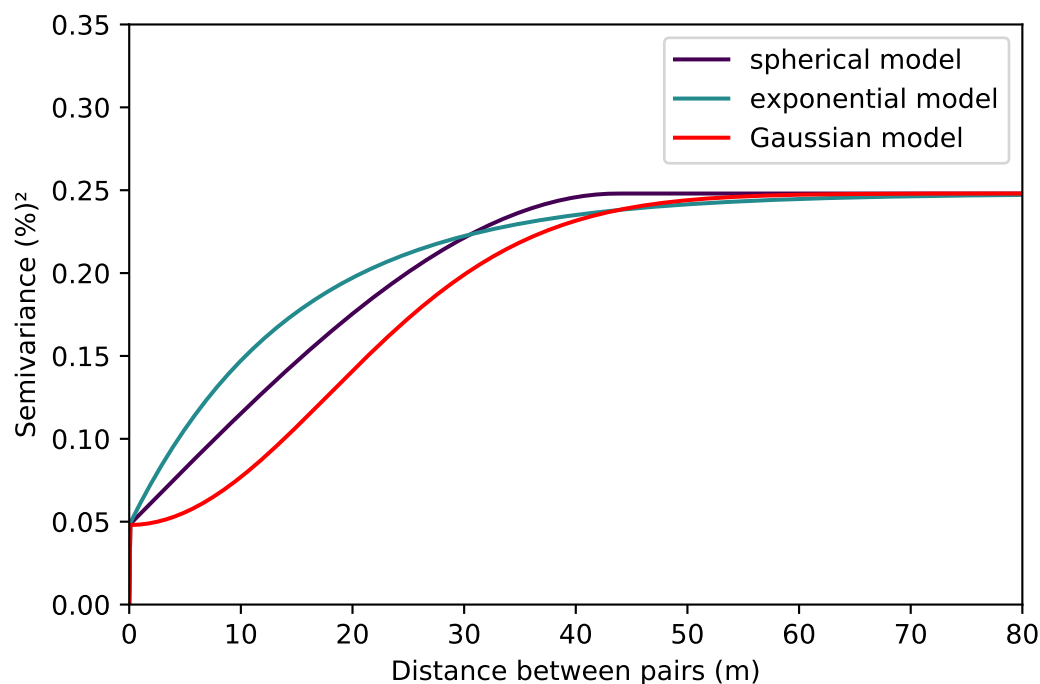


Figure 2: Graphical representation of the spherical, exponential, and Gaussian semivariogram models. All of them are represented using the same parameters.

Three parameters define the semivariogram models (Figure 3): Nugget Effect, Sill, and Range. The first one is a chaotic or unstructured variance. It represents a variability that occurs at a scale lower than the spacing between the sample (or input) data and, therefore, is not mapped by the semivariogram and is also due to sampling errors. The Sill corresponds to the variance of the variable in the direction of the vector h , and the Range is the distance from which there's no more spatial correlation between the pairs of values in the direction being mapped (Journel and Huijbregts [1978]; Hustrulid et al. [2013]; Rossi and Deutsch [2014]).

Calculating the semivariograms in several directions in the analyzed - and stationary - domain quantifies the spatial continuity behavior of the variable under analysis. In this context, directional semivariograms will frequently show different spatial continuities in the function of the spatial direction (vector h), which will indicate the presence of a direction endowed with a greater homogeneity between the values along this one and another one with a lower homogeneity (higher variability). Cases in which the semivariograms' shape are a function of the spatial direction corresponds to the occurrence of anisotropy, among them, the most common is the geometric anisotropy, in which the range varies in function of the spatial direction (vector h), and the sill remains constant (Figure 4) (Armstrong [1998]; Isaaks and Srivastava [1989]; Journel and Huijbregts [1978]).

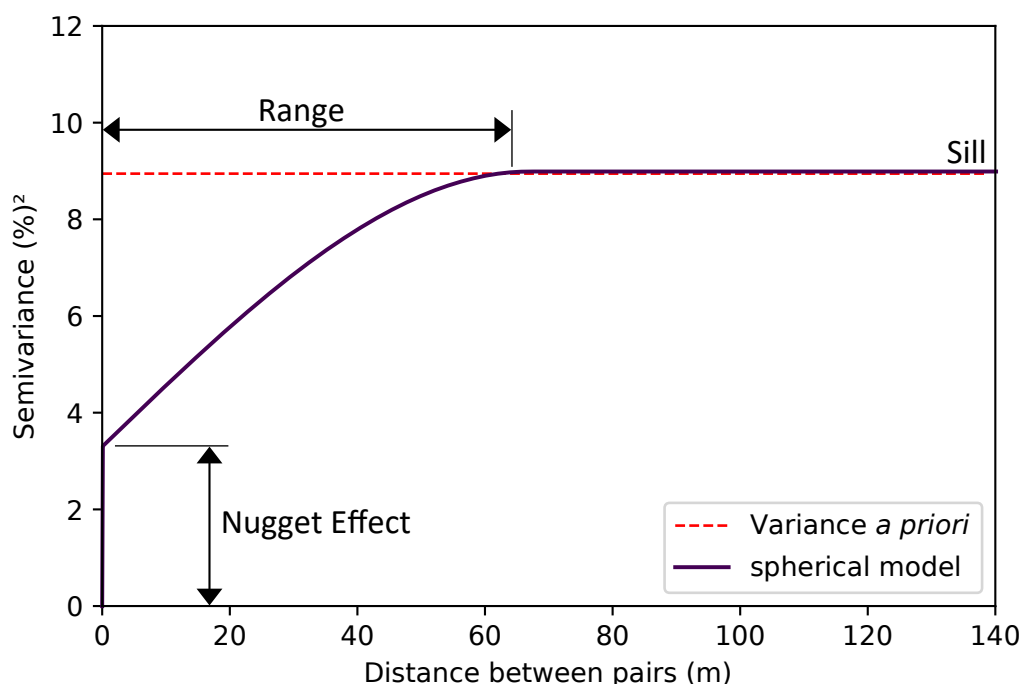


Figure 3: Example of a spherical semivariogram model with its parameters.

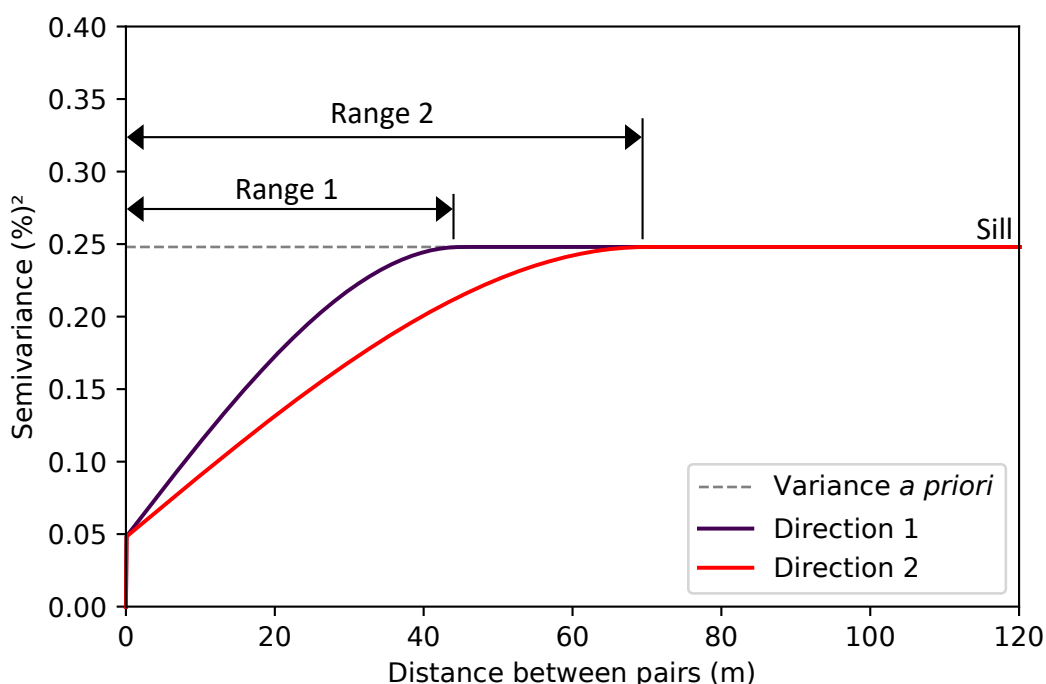


Figure 4: Semivariograms showing a phenomenon with geometric anisotropy.

The possibility to map the spatial continuity of a variable (e.g., damage index) using semivariograms can help identify possible preferential directions of structural damage propagation, which will be shown in Section 4, where there is an occurrence of geometric anisotropy in indicator variables obtained from RMSD-based damage index values.

The information obtained from semivariograms allows the complete characterization of the variable's spatial continuity and is used for estimating its values at the unsampled locations by kriging methods.

2.2.1 Kriging Methods

The term kriging covers a wide variety of estimation methods, characterized to be exact interpolators, whose mathematical formulations of the main types can be obtained in the works of Rossi and Deutsch [2014], Chilès and Delfiner [2012], Sinclair and Blackwell [2002], Kitanidis [1997] and Journel and Huijbregts [1978].

Among the kriging methods, there is the ordinary kriging one, which is a local estimation method, where the estimation of a point at an unsampled location ($Z_{KO}^*(x_0)$) is obtained by a linear combination (Eq. 5) of n sample (or input) values $Z(x_i)$ contained in the search neighborhood, where the weights λ_i reflects the spatial correlation between the unsampled point (or block) and the actual (or input) values within the search neighborhood, and between the sample (or input) values among each other contained in the search neighborhood (Yamamoto and Landim [2013]).

$$Z_{KO}^* = \sum_{i=1}^n \lambda_i Z(x_i) \quad (5)$$

The weights of the linear combination shown in the Eq. 5 are obtained through the solution of the ordinary kriging's system of equations (Eq. 6), where it's intended to get the combination of weights in such a way that the kriging variance is minimized, considering the aspects of spatial continuity of the variable under estimation, being these, reproduced in the semivariogram values between sample (or input) values in the x_i and x_j locations ($\gamma(x_i - x_j)$), and the semivariogram values between the point to be estimated at the x_0 location and the input data at the x_i locations ($\gamma(x_i - x_0)$). Also, a restriction, which refers to the need for the sum of the weights to be unitary, is imposed on the system to obtain unbiased estimates, called an unbiased constraint. This system can be expressed in the matrix form (Eq. 7), from which the optimal weights are calculated (Yamamoto and Landim [2013]).

$$\begin{cases} \sum_{j=1}^n \lambda_j \gamma(x_i - x_j) - \mu = \gamma(x_i - x_0) & \text{for } i = 1, \dots, n \\ \sum_{j=1}^n \lambda_j = 1 \end{cases} \quad (6)$$

$$\begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \\ -\mu \end{bmatrix} = \begin{bmatrix} \gamma(x_1 - x_1) & \gamma(x_1 - x_2) & \dots & \gamma(x_1 - x_n) & 1 \\ \gamma(x_2 - x_1) & \gamma(x_2 - x_2) & \dots & \gamma(x_2 - x_n) & 1 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ \gamma(x_n - x_1) & \gamma(x_n - x_2) & \dots & \gamma(x_n - x_n) & 1 \\ 1 & 1 & \dots & 1 & 0 \end{bmatrix}^{-1} \times \begin{bmatrix} \gamma(x_0 - x_1) \\ \gamma(x_0 - x_2) \\ \vdots \\ \gamma(x_0 - x_n) \\ 1 \end{bmatrix} \quad (7)$$

In addition to the variable's estimated value at the unsampled location x_0 , the kriging methods are capable of providing the estimation variance, which, in the case of the ordinary kriging method, is given by the Eq. 8, where μ is the Lagrange parameter, and n refers to the number of samples (input data) used for the estimation at an unsampled location.

$$\sigma_{KO}^2 = \sum_{i=1}^n \lambda_i \gamma(x_i - x_0) + \mu \quad (8)$$

The ordinary kriging estimator can be used as the base for estimation of indicator variables, characterized to be binary, the situation in which there is the so-called indicator kriging, being a

non-parametric method, in such a way that the statistical distribution of the variable doesn't restrict its application under estimation (Rossi and Deutsch [2014]; Revuelta [2018]). The indicator kriging performs the estimation of binary variables, whose values are determined according to a specific criteria, as shown in the Eq. 9, where, from an original distribution, are indicated K thresholds z_k , and for each of them, are created indicator variables, coding the values according to the value's position of the original variable ($Z(x)$) relative to the specified threshold z_k (Eq. 9).

$$I(x, z_k) = \begin{cases} 1, & \text{if } Z(x) \leq z_k \\ 0, & \text{otherwise} \end{cases} \quad (9)$$

In the case of using the ordinary kriging as the interpolator for the estimation of indicator variables at the unsampled locations, the estimated values $I(x_0)$ at each x_0 location are calculated as a linear combination (Eq. 10) of the n sample (or input) data $I(x_i)$ located at the x_i points and contained in the search neighborhood, generally with an elliptical shape, whose weights λ_i reflects the spatial correlation between the estimated point and the sample values within the search neighborhood, and between the sample (or input) values among each other (Yamamoto and Landim [2013]).

$$I(x_0) = \sum_{i=1}^n \lambda_i I(x_i) \quad (10)$$

The estimated indicator variable values by ordinary kriging are determined in such a way that the variance of the estimation error is minimized, considering the unbiased constraint (average of the estimation error is zero), resulting in optimal weights-based estimations, which minimizes the estimation error variance (Isaaks and Srivastava [1989]).

3 Python packages applied to Kriging

The geostatistical methods and tools (e.g., kriging) have been implemented in several programming languages, including Python. Among the packages used to carry out kriging-based estimates, there is the Geostatspy (Pyrzcz et al. [2021]), whose package is characterized to be reimplementations of the functionalities of the Geostatistics Software Library (GSLIB) in the Python language. The GSLIB is a set of pretty robust codes for building spatial modeling workflows. Therefore, this package offers the opportunity to create projects of 2-D spatial modeling in Python without the necessity of trust in the GSLIB's compiled Fortran code, together with the use of datasets that moves between the GSLIB's Geo-EAS format to Pandas' dataframes and grids to NumPy's n-dimensional arrays.

The operation of the Geostatspy package requires some dependencies, which includes the following libraries: (1) NumPy (Harris et al. [2020]), for n-dimensional arrays; (2) Pandas (McKinney [2010]), for DataFrames; (3) numpy.linalg, for linear algebra; (4) numba, for numerical acceleration; (5) scipy, for quickly nearest neighbor search, and; (6) Matplotlib (Hunter [2007]), for plotting.

The Geostatspy package contains *geostatspy.geostats* and *geostatspy.GSLIB*. The first includes GSLIB functions rewritten in Python, including all semivariograms, distribution transformations, and estimation and simulation-based methods. The second consists of the reimplementations of the GSLIB visualizations and low-tech wrappers of numerical methods, which require access to the GSLIB executables.

The first part, *geostatspy.geostats*, contains numerical methods, for example, the *ik2d* function (GSLIB's *ik3d* program implemented for indicator kriging estimation of 2-D data), *correct_trend*

(fix the order relationships of an indicator-based trend model), and *backtr* (GSLIB's *backtr* program for back-transform a normal-shape distribution to an original one). These and other functions are presented and described in the Table 1.

Table 1: Some functions of the geostatspy.geostats part of the Geostatspy package

Function	Application
correct_trend	fix the order relationships of an indicator-based trend model
nscore	Transform an original distribution to a normal-shape one
backtr	Back-transform a normal-shape distribution to the original one
declus	Execution of sample's declustering (cell-based declustering)
gam	Calculation of semivariograms basing on regularly spaced 2-D data
gamv	Calculation of semivariograms basing on irregularly spaced 2-D data
varmapv	Semivariogram map of irregularly spaced 2-D data
vmodel	Semivariogram fitting
kb2d	Simple and/or ordinary kriging for 2-D data
ik2d	Indicator kriging for 2-D data

The second part, *geostatspy.GSLIB*, contains utility functions that support moves between data tables: DataFrames, n-dimensional arrays, Geo-EAS, and grid data. It also contains data transformation functions, spatial continuity, spatial model re-sampling, and visualization using Matplotlib. Some functions are showed and described in the Table A1, Appendix A.

4 Damage location prediction using Indicator Kriging – A Case Study

Herein it's considered an aluminum plate of 100 cm x 100 cm dimensions containing 100 RMSD-based damage metric values arranged in a regular mesh of 9.09 cm x 9.09 cm, with a known damage located at the plate's center. For the development of the case study it was used the python packages geostatspy (Pyrzcz et al. [2021]), os-sys (Labots [2019]), NumPy (Harris et al. [2020]), Pandas (McKinney [2010]), pygeostat (Deutsch et al. [2020]), and Matplotlib (Hunter [2007]), which were imported according the Code 1.

Code 1: Source code for import the Python packages used in this case study

```

1 import geostatspy.GSLIB as GSLIB
2 import geostatspy.geostats as geostats
3 import os
4 import numpy as np
5 import pandas as pd
6 import pygeostat as gs
7 import matplotlib.pyplot as plt
8 from matplotlib import gridspec
9 %matplotlib inline

```

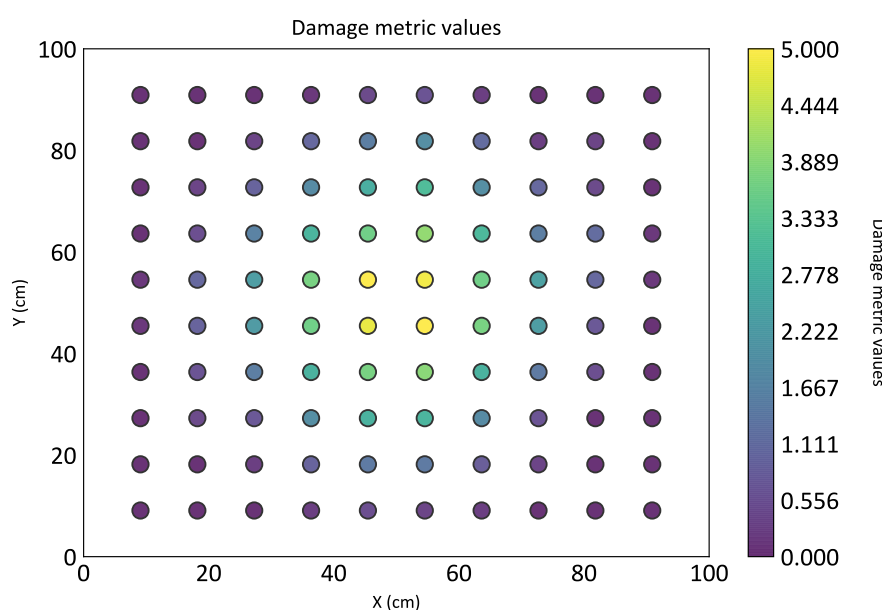
In the sequence, it was selected the working directory using the *chdir* function of the *os-sys* package. The data, in the CSV format, was then imported using the Pandas' *read_csv* function. Subsequently, it was necessary to plot the RMSD values, which was done using the Code 2, resulting in the Figure 5, where one can observe the occurrence of high RMSD values near the damage, with those decreasing gradually from center to the plate's border.

Code 2: Source code for import the dataset and plot the RMSD values in X-Y plane

```

1 os.chdir('C:\\YOUR_DATABASE_PATH')
2 df=pd.read_csv('dataset.csv', header=0, sep=';')
3 xmin = 0.0; xmax = 100.0
4 ymin = 0.0; ymax = 100.0
5 xsiz = 4; ysiz = 4
6 nx = 25; ny = 25
7 xmn = 2; ymn = 2
8 cmap = plt.cm.viridis
9 GSLIB.locmap_st(df, 'X', 'Y', 'PDANO', xmin, xmax, ymin, ymax, 0, 5, 'Damage
   metric values', 'X (cm)', 'Y (cm)', 'Damage metric values', cmap)
10 plt.savefig('map_points.PDF', dpi=800, bbox_inches='tight')
11 plt.show()

```

**Figure 5: Localization map of the PZT patches in the plate**

In the Code 2, the directory was defined in line 1, and the dataset, imported using the *read_csv* function, shown in line 2, where were added the file's name, the number of the header line, and the column's delimiter. Between lines 3 and 7 was defined the followed variables: minimum coordinate in the X direction (*xmin*); maximum coordinate in the X direction (*xmax*), minimum coordinate in Y (*ymin*), maximum coordinate in Y (*ymax*), cell size in the X direction (*xsiz*), cell size in the Y direction (*ysiz*), number of cells in X (*nx*), number of cells in Y (*ny*), origin of the grid in X (*xmn*), and origin of the grid in Y (*ymn*). In the sequence (line 8), the color palette was exposed. Then, using the geostatspy's function named *locmap_st* (line 9), the points (PZTs location) were plotted according to the parameters defined previously.

The statistical analysis of the RSMD values was done by means of boxplot (Figure 6), histogram (Figure 6), and descriptive statistical indicators (Table 2), which were obtained using the Code 3 (boxplot and histogram) and the Code 4 (descriptive statistics). According to the results (Figure 6 and Table 2), one can observe a highly asymmetric, positive skewed distribution, with only 25% of the data higher than 1.9, with presence of outliers, all of them higher than 4.0.

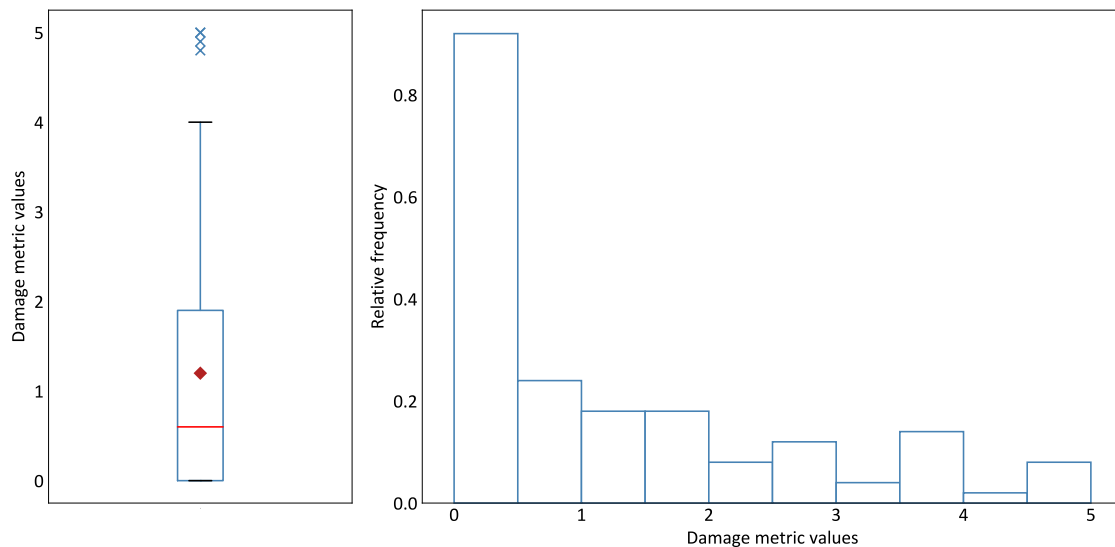


Figure 6: Boxplot and histogram of the RMSD values

Table 2: Descriptive statistics of the damage metric values.

Count	Min	Q1	Median	Mean	Q3	Max	STDEV	Variance	Skewness
100	0.000	0.000	0.600	1.199	1.900	5.000	1.414	1.998	1.151

Code 3: Source code to get the boxplot and histogram of the RMSD values

```

1 fig=plt.figure(figsize=(10,5))
2 gs=gridspec.GridSpec(1,2, width_ratios=[1.3,3])
3
4 ax1=plt.subplot(gs[0])
5 meanpointprops=dict(marker='D', markeredgecolor='black',
6   markerfacecolor='firebrick')
7 flier_props=dict(marker='x', markerfacecolor='magenta', markeredgecolor=
8   ='steelblue')
9 box_props=dict(color='steelblue')
10 whisker_props=dict(color='steelblue')
11 median_props=dict(color='red')
12 ax1.boxplot(df['PDANO'], showmeans=True, meanprops=meanpointprops,
13   flierprops=flier_props, boxprops=box_props, medianprops=
14   median_props, whiskerprops=whisker_props)
15 ax1.tick_params(axis='x', labelsz=0)
16 ax1.tick_params(axis='y', labelsz=12)
17 ax1.set_ylabel('Damage metric values', fontsize=12)
18
19 ax2=plt.subplot(gs[1])
20 ax2.hist(df['PDANO'], density=True, facecolor='white', edgecolor='
21   steelblue', linewidth=1)
22 ax2.set_xlabel('Damage metric values', fontsize=12)
23 ax2.set_ylabel('Relative frequency', fontsize=12)
24 ax2.tick_params(axis='x', labelsz=12)
25 ax2.tick_params(axis='y', labelsz=12)
26
27 plt.subplots_adjust(wspace=0.2, hspace=0.1)

```

```

23 fig.tight_layout()
24 plt.savefig('statistic_analysis.pdf', dpi=800, bbox_inches='tight')
25 plt.show()

```

Code 4: Source code for obtaining the statistical indicators of the results showed in Table 2

```

1 print(df['PDANO'].describe())
2 print(df['PDANO'].skew(skipna=True))
3 print(df['PDANO'].var())

```

In the Code 3, initially was defined the figure's size (line 1), and then the specifications of the grid, containing the boxplot and the histogram, using the *GridSpec* function (line 2). In the sequence (line 4), it was determined the position of the boxplot in the subplot grid (first position), followed by the boxplot properties relative to the mean (line 5), the outliers (line 6), the boxes (line 7), the whiskers (line 8) and median (line 9), which were used to make the boxplot (line 10), whose general formatting features were defined in the lines 11, 12 and 13. In line 15, it was defined the position of the histogram in the subplot grid (second position) and then the histogram (line 16), whose formatting features were stated between lines 17 and 20. Finally, the horizontal and vertical spacing between the plots was established in line 22, followed by the definition of the tight layout (line 23), the *savefig* function to save the figure (line 24), and a command line to show the plots (line 25).

The Code 4 shows the source code for obtaining the descriptive statistics of the RMSD values. In line 1 was used the *describe* function, which returned the following indicators: number of values; mean; standard deviation; minimum and maximum value, and; the first, second, and third quartiles. The skewness was gotten using the *skew* function (line 2) and the sample variance, the *var* function (line 3).

Based on the statistical information (Figure 6), where one can observe the occurrence of outliers, all of them higher than 4.0, it was done a localization map with only the values higher than 4.0 (Figure 7), obtained using the Code 5.

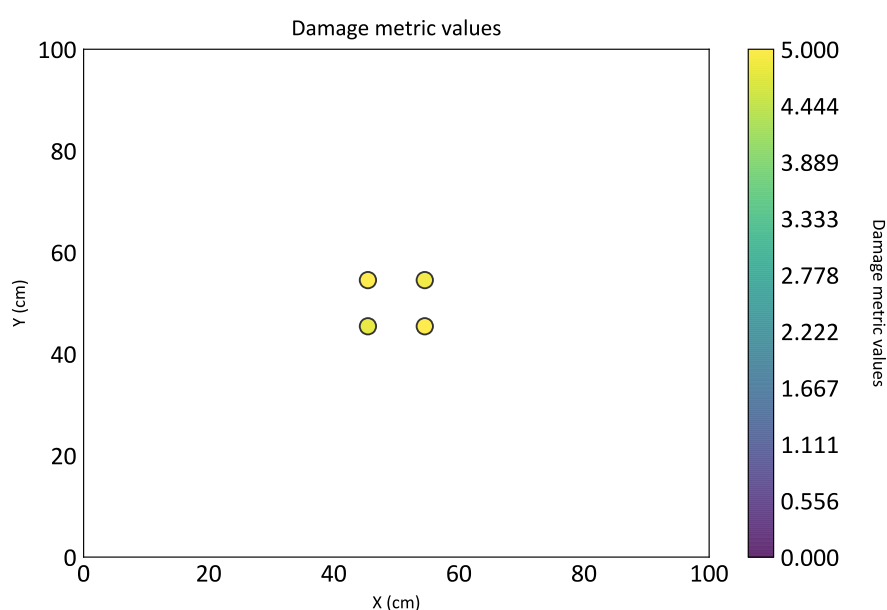


Figure 7: Localization map of the RMSD values higher than 4.0

Table 3: Thresholds and their respective indicator variables.

Threshold	Indicator Variable
1	IND0
2	IND1
3	IND2
4	IND3

Code 5: Source code for make the localization map of the RMSD values higher than 4.0

```

1 df_thr4=df[df['PDANO']>4]
2 df_thr4
3 xmin = 0.0; xmax = 100.0
4 ymin = 0.0; ymax = 100.0
5 xsiz = 4; ysiz = 4
6 nx = 25; ny = 25
7 xmn = 2; ymn = 2
8 cmap = plt.cm.viridis
9 GSLIB.locmap_st(df_thr4,'X','Y','PDANO',xmin,xmax,ymin,ymax,0,5,'Damage
  metric values','X (cm)','Y (cm)','Damage metric values',cmap)
10 plt.savefig('localization_map_thr_4.pdf', dpi=800, bbox_inches='tight')
11 plt.show()

```

In the Code 5, initially (line 1) it was defined a new dataset, containing only the RMSD values higher than 4.0. In the sequence, between lines 3 and 8, it was defined the plot's parameters: minimum (*xmin*) and maximum (*xmax*) coordinates in the X direction (line 3); minimum (*ymin*) and maximum (*ymax*) coordinates in the Y direction (line 4); cell sizes in the X (*xsiz*) and Y (*ysiz*) directions (line 5); number of cells in the X (*nx*) and Y (*ny*) directions (line 6); origin of the grid in the X (*xmn*) and Y (*ymn*) directions (line 7), and; color palette to be used (line 8). Finally, it was used the *locmap_st* function of the Geostatspy package in the line 9 to plot the data, whose resulting figure was saved using the *savefig* function (line 10).

Figure 7 shows an occurrence of values higher than 4.0 near the plate's center, where the known damage is located. Based on this, four thresholds were selected, 1, 2, 3, and 4, to map the occurrence of these values along the plate. The latter threshold (4.0) is particularly interesting, once it's related to the damage position. From the established thresholds, were created four indicator variables, each of them assigned to a specific threshold (Table 3), whose binary values were defined according to Eq. 9, implemented using the Code 6, and indicating the probability below a specific threshold.

Code 6: Source code for defining the indicator variables related to the established thresholds

```

1 thrs=[1,2,3,4]
2 for i in range(1,5):
3     cont=i
4     df["IND"+str(cont-1)]=np.zeros(len(df['PDANO']))
5     df["IND"+str(cont-1)]=np.where(df['PDANO']<=thrs[cont-1],1,0)
6 df

```

The Code 6 started with defining the threshold values (line 1). Then, a loop (line 2) was used for creating the indicator variables, all of them assuming initially zero values (line 4) and with the same length of the variable containing the RMSD values. In the sequence, it was used the *where* function of the *NumPy* package to apply the criterion defined in the Eq. 9, resulting in the

probabilities below the respective thresholds.

To calculate the semivariograms, it was necessary, initially, to obtain the variance of each indicator variable (Table 4); once these values are, in most cases, the variance *a priori* in the semivariogram plot. This step was done with the Code 7, based on the standard deviation calculation and then the variance.

Table 4: Sample variance of each indicator variable.

Indicator Variable	Sample Variance
IND0	0.2356
IND1	0.1824
IND2	0.1131
IND3	0.0384

Code 7: Source code used for calculating the variance of each indicator variable

```

1 var = np.zeros(4)
2 for i in range(4):
3     cont = i
4     x = "IND"+str(cont)
5     var[i] = np.std(df[x])
6     var[i] = (var[i])**2
7     print(var[i])

```

According to Code 7, firstly it was created a variable named *var* (line 1) with a length equal to 4 to receive the variance value of each indicator variable. A *for* loop was then applied in line 2 to calculate all variances. Inside the loop, a variable named *cont* was defined (line 3), being equal to *i*, and in the sequence, it was defined a variable named *x* (line 4), which was equal to the respective indicator variable. Afterward, each value of the *var* variable was defined as the standard deviation of the respective indicator variable (line 5). Finally, these values were squared, resulting in the respective variances, shown in Table 4.

The results obtained for the indicator variables' sample variance showed that the highest variance is related to the indicator variable attributed to the threshold equal to 1, and the lowest, with the threshold equal to 4. This was expected once 62% of the samples were equal or lower than 1.0 (38% higher than 1.0), and 96% of the samples were equal or lower than 4 (only 4% higher than 4), indicating that the indicator variable related to the threshold equal to 1 would have a variability much higher than the other ones. This also means a spatial variability much higher for the indicator variable attributed to the threshold equal to 1 about the others. This characteristic would be manifested in the semivariogram model used to fit the experimental semivariograms in the next step.

The calculating of the experimental semivariograms was done using the Geostatpy's *gamv* function (see Section 3), which calculates the semivariogram for irregularly spaced data (general case). Firstly, an example of the calculation of the semivariogram for the IND0 variable, related to the threshold equal to 1.0, in the North-South direction (Azimuth equal to 0°), along with its variographic fitting. The calculation and fitting of this semivariogram were done using the Code 8, resulting in the semivariogram shown in Figure 8.

Code 8: Source code for calculating the experimental semivariogram for the INDO variable in the North-South direction with its fitting by a theoretical model

```

1 azi=0
2 lag_dist=9.09
3 lagtol=2
4 nlag=8
5 bandh=10
6 atol=10
7 isill=var[0]
8 tmin=-1.0e21
9 tmax=1.0e21
10 lag0, por_gamma0, por_npair0 = geostats.gamv(df, "X", "Y", "INDO", tmin,
      tmax, lag_dist, lagtol, nlag, azi, atol, bandh, isill)
11 plt.plot(lag0[2:-1], por_gamma0[2:-1], color='darkblue', marker='o',
      linestyle='dashed', linewidth=1, markersize=5, label = 'Experimental
      Semivariogram')
12 plt.xlim(0, 80)
13 plt.ylim(0, 0.4)
14 plt.plot([0, 100], [var[0], var[0]], color='#238A8DFF', linewidth=1,
      linestyle='dashed', label='Variance a priori')
15
16 nugg=0.02
17 nst=1
18 it11=1
19 cc11=var[0]-nugg
20 azi1=0
21 hmaj11=52
22 hmin11=52
23 nlag1=160
24 xlag=0.5
25 vario_INDO=GSLIB.make_variogram(nugg, nst, it11, cc11, azi1, hmaj11, hmin11)
26 index_INDO_azi0, h_INDO_azi0, gam_INDO_azi0, cov_INDO_azi0,
      ro_INDO_azi0 = geostats.vmodel(nlag1, xlag, azi1, vario_INDO)
27 plt.plot(h_INDO_azi0, gam_INDO_azi0, color='darkblue', label='
      Theoretical Fitting (spherical model)')
28 plt.xlabel('Distance (cm)', fontsize=10)
29 plt.ylabel('Semivariance', fontsize=10)
30
31 plt.legend(fontsize=12, loc='upper left')
32 plt.savefig('INDO_semivariogram_az0.pdf', dpi=800, bbox_inches='tight')
33 plt.show()

```

The Code 8 starts with the definition of the parameters for calculation of the experimental semivariogram, which were: azimuth direction (line 1); lag separation distance (line 2), usually equal to the lowest spacing between sample pairs (in a regular grid, similar to the sample spacing); lag tolerance (line 3), which should be lower than 50% of the lag separation value; number of lags in the direction (line 4), being equal to the number of lags to cover at least the half of the field being analyzed; horizontal bandwidth (line 5), which can be the sample spacing in the perpendicular direction; angular tolerance (line 6), being lower than 90°, once angular tolerance equal to 90° results in the unidirectional semivariogram; sill of the semivariogram (line 7), usually equal to the variance *a priori*; the minimum value to be considered (line 8), in such a way that values lower than the minimum are ignored, and; the maximum value to be considered (line 9), in such a way that values higher than the maximum are ignored in the semivariogram calculation.

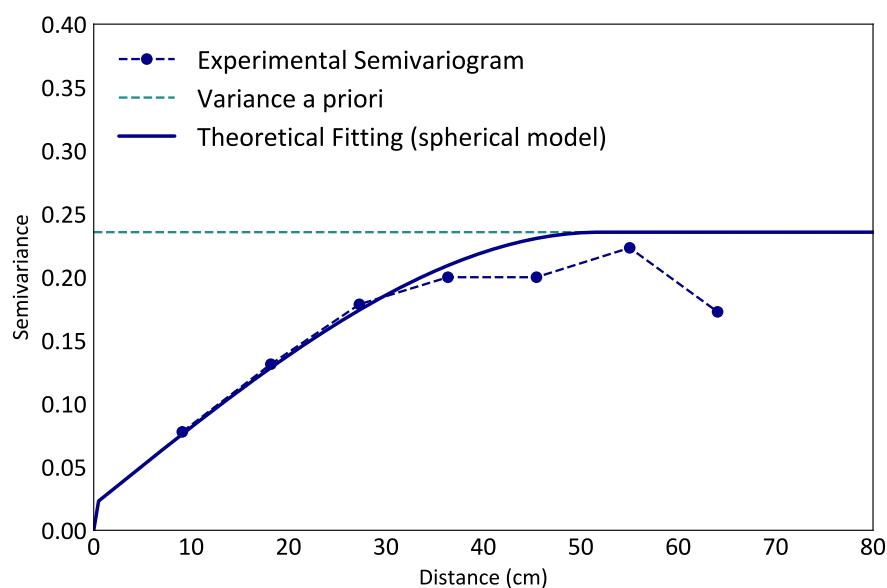


Figure 8: Experimental and theoretical fitting of the semivariogram of the IND0 variable in the North-South direction (Azimuth equal to 0°)

Based on the parameters defined between lines 1 and 9 of the Code 8, the experimental semivariogram was calculated using Geostat's *gamv* function in line 10, resulting in three variables: classes of distances (*lag0* variable), semivariance values (*por_gamma0* variable), and the number of pairs (*por_npair0* variable) used of obtaining each point of the experimental semivariogram. In the sequence, the semivariance and their respective classes of distances were plotted (line 11) using the *plot* function, where only the semivariance values were higher than zero using a filter. Subsequently, the limits of the plot (lines 12 and 13) were plotted, using line 14, the variance *a priori* (normally the semivariogram's sill) together with the experimental semivariogram.

The experimental semivariogram needs to be fitted by theoretical models to firstly identify the preferential directions of continuity, which would be the directions of highest and lowest range values, and then; provide spatial continuity information (based on such preferential directions) for the kriging equations' solving (see Section 2.2.1), resulting in the kriging estimates.

The parameters of the semivariogram's fitting were defined between the lines 16 and 24 of the Code 8, which were: (1) Nugget Effect (variable *nugg* in the line 16), which approximately refers to the interception in the Y axis; (2) position of the nested structure (*nst* in the line 17), in this case, it's the first nested structure; (3) code for the model type (*it11* in the line 18), which can be 1, 2, 3, 4 or 5 for spherical, exponential, Gaussian, Power law, or Cosine hole effect models, respectively; (4) contribution of the structure for the semivariogram's sill (*cc11* in the line 19), which was, in this case, the sill minus the nugget effect; (5) azimuth direction of the semivariogram's major axis (*azi1* in the line 20), in this case, 0° ; (6) range of the structure in the major direction (*hmaj11* in line 21); (7) range of the structure in the minor direction (*hmin11* in the line 22); (8) number of lags (*nlag1* in the line 23), and; (9) lag separation, defined in the variable *xlag*, line 24.

Based on the aforementioned parameters, it was obtained the semivariogram model by means of the *make_variogram* function, described in line 25, whose model was attributed to a variable named *vario_IND0*. To determine the full information for plotting the semivariogram's model, the variable *vario_IND0* was used in the *vmodel* function (line 26), resulting, among others, in the classes of distances (variable *h_IND0_azi0* in line 27) and the theoretical semivariogram values (variable *gam_IND0_azi0* in line 27), which were used for the semivariogram model plotting (line

27). The resulting figure was, then, saved in the PDF format using Matplotlib's *savefig* function (line 32).

Based on the steps for semivariogram calculation, it was obtained experimental semivariograms for several directions for all indicator variables in the form of a report with the plots. This was done using Code 9, in such used loops, automating the semivariogram calculation, and filtering the results obtained in such a way to plot only those that showed a spatial continuity.

Code 9: Source code for obtain the experimental semivariogram of all indicator variables and plotting only those ones which demonstrates a spatial continuity

```

1 lag_dist = np.zeros(4)
2 azi = [0,30,60,90,120]
3 color = ['red', 'green', 'pink', 'darkblue', 'black']
4 lag = [np.zeros(1)]*20
5 por_gamma = [np.zeros(1)]*20
6 por_npair = [np.zeros(1)]*20
7 tmin = -1.0e21; tmax = 1.0e21;
8 lag_tol = 1; nlag = 8; bandh = 10; atol = 10;
9
10 k = 0
11 for i in range(4):
12     cont = i
13     ind = "IND"+str(cont)
14     plt.figure(figsize=(7, 4))
15     plt.xlabel(r'Lag Distance  $\mathbf{h}$ , (cm)')
16     plt.ylabel(r' $\gamma$   $\mathbf{h}$ ')
17     plt.title('Variogram '+ind+' 10x10')
18     plt.plot([0,2000],[var[i],var[i]],color = '#238A8DFF',linewidth=1,
19             linestyle='dashed')
20     plt.xlim(0,80)
21     plt.ylim(0,0.42)
22     for j in range(0,5):
23         lag_dist[i] = 9.09; azi[j]; isill = var[i]
24         lag0, por_gamma0, por_npair0 = geostats.gamv(df, "X", "Y", ind,
25             tmin,tmax,lag_dist[i],lag_tol,nlag,azi[j],atol,bandh,isill)
26         lag[k] = lag0; por_gamma[k] = por_gamma0; por_npair[k] =
27             por_npair0;
28         filtro = lag[k] != 0; filtro2 = por_gamma[k] != 0;
29         aux1 = lag[k][filtro]; aux2 = por_gamma[k][filtro2]
30         if len(aux1)==len(aux2):
31             lag[k] = aux1;por_gamma[k] = aux2
32         if len(lag[k])>2:
33             plt.plot(lag[k][0:len(lag[k])], por_gamma[k][0:len(
34                 por_gamma[k])], color=color[j], marker='o', linestyle='dashed',
35                 linewidth=1,markersize=5,label = 'Porosity')
36             print('Preferential Direction:',azi[j],';cor:',color[j],';
37                 lag['+str(k)+'], por_gamma['+str(k)+''])
38         k +=1
39     plt.show()

```

In Code 9, firstly, it was defined the arrays to receive the lag separation values (*lag_dist* in line 1), the azimuth directions (*azi* in line 2), the color for each direction (*color* in line 3), the classes of distances (*lag* in line 4), the semivariance values (*por_gamma* in line 5), and the number of pairs for each class of distance (*por_npair* in line 6). In the sequence, it was defined the trimming limits (*tmin* and *tmax* in line 7), the lag tolerance (*lag_tol* in line 8), the number of lags (*nlag* in line 8),

the horizontal bandwidth (*bandh* in line 8), and the angular tolerance (*atol* in line 8), all of them used as parameters for the semivariogram calculation.

The information obtained by the experimental semivariograms was used for the semivariogram fitting. Theoretical models fitted those associated with the preferential directions of continuity (highest and lowest range values). The fitting was executed using the *make_variogram* function and then the *vmodel* function for each direction and indicator variable. Taking as an example, the source code used to plot the experimental semivariograms and their respective fittings for the threshold equal to 1 (IND0) is shown in the Code 10, and the resulting plot, in the Figure 9, where one can observe the existence of two semivariogram models (continuous lines), one for each experimental semivariogram.

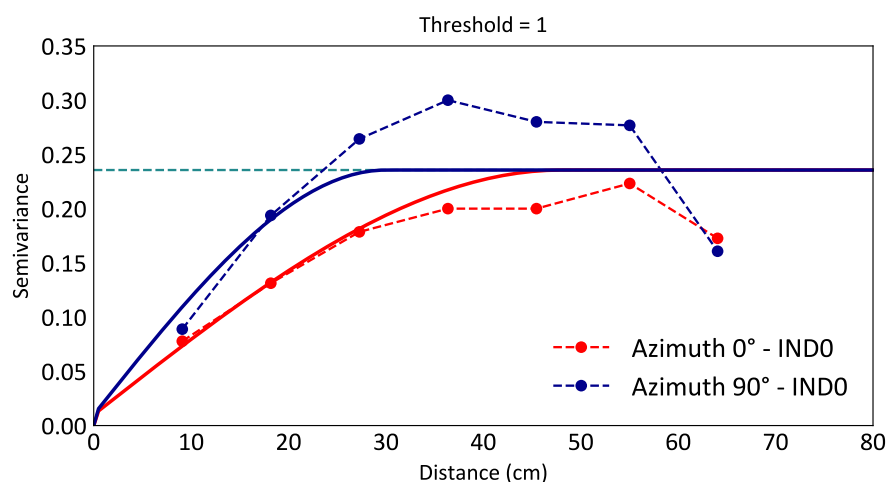


Figure 9: Plot of the experimental semivariograms and its respective fittings in the preferential directions of continuity for the indicator variable related to the threshold equal to 1

The Code 10 starts with determining the figures and the subplots' sizes (lines 1 and 2). In line 4, the position of this specific plot was defined in a grid, in this case, composed of only one plot, using the *subplot2grid* function of the Matplotlib package. Between lines 5 and 10, it was defined the overall parameters of the plot: Position of the variance a priori of the indicator variable under analysis, limits in the X-axis, limitations in the Y-axis, label for the X-axis, label for the Y-axis, and plot's title, respectively. The experimental semivariograms were plotted using the function *plot* (lines 11 and 12) with the respective legends (line 13). In the sequence, the semivariogram models were determined and then plotted, whose procedures and parameters are described between the lines 15 and 20 for the azimuth 0° and between the lines 22 and 27 for the azimuth 90°. Lastly, the figure was saved in PDF format using the *savefig* function (line 29).

Applying similar procedures to that described in the Code 10 for the other semivariogram plots, it was obtained the results shown in the Figure 10, whose fitting parameters used are shown in Table 5.

The analysis of the range values in the experimental semivariograms for each indicator variable (Figure 10 and Table 5) showed the occurrence of geometric anisotropy (differences in the range values, with equal sill), for the thresholds equal to 1 and 3. The preferential directions for these thresholds were the 0° (North-South) and 90° (East-West) azimuths, with North-South being the major axis. The occurrence of anisotropy wasn't identified in the other indicator variables, suggesting an isotropic phenomenon for these.

Code 10: Source code for plotting the experimental semivariogram and its respective fitting for the threshold equal do 1

```

1 plt.figure(figsize=(6,3))
2 fig, axs = plt.subplots(1, 1, figsize=(6,3), constrained_layout=True)
3
4 plt.subplot2grid((1,1), (0,0))
5 plt.plot([0,200],[var[0],var[0]], linestyle='dashed', linewidth=1,
6         color='#238A8DFF')
7 plt.xlim(0,80)
8 plt.ylim(0,0.35)
9 plt.xlabel('Distance (cm)', fontsize=10)
10 plt.ylabel('Semivariance', fontsize=10)
11 plt.title('Threshold = 1', fontsize=10)
12 plt.plot(lag[0][0:len(lag[0])], por_gamma[0][0:len(por_gamma[0])],
13         color='red', marker='o', linestyle='dashed', linewidth=1, markersize
14         =5, label = 'Azimuth 0° - INDO')
15 plt.plot(lag[3][0:len(lag[3])], por_gamma[3][0:len(por_gamma[3])],
16         color='darkblue', marker='o', linestyle='dashed', linewidth=1,
17         markersize=5, label = 'Azimuth 90° - INDO')
18 plt.legend(fontsize=12)
19
20 nugg[0] = 0.01; nst = 1; it11[0] = 1; cc11[0] = var[0]-nugg[0];
21 azith[[0],[0]] = 0; hmaxmin11[[0],[0]]=48; hmaxmin11[[0],[0]] = 48
22 vario_INDO_az0 = GSLIB.make_variogram(nugg[0],nst,it11[0],cc11[0],azith
23 [[0],[0]],hmaxmin11[[0],[0]],hmaxmin11[[0],[0]])
24 nlag = 160; xlag = 0.5;
25 index_INDO_az0, h_INDO_az0, gam_INDO_az0, cov_INDO_az0, ro_INDO_az0 =
26 geostats.vmodel(nlag, xlag,azith[[0],[0]], vario_INDO_az0)
27 plt.plot(h_INDO_az0,gam_INDO_az0,color='red')
28
29 nugg[0] = 0.01; nst = 1; it11[0] = 1; cc11[0] = var[0]-nugg[0];
30 azith[[0],[1]] = 90; hmaxmin11[[0],[1]] = 30; hmaxmin11[[0],[1]] = 30
31 vario_INDO_az90 = GSLIB.make_variogram(nugg[0],nst,it11[0],cc11[0],
32 azith[[0],[1]],hmaxmin11[[0],[1]],hmaxmin11[[0],[1]])
33 nlag = 160; xlag = 0.5;
34 index_INDO_az90, h_INDO_az90, gam_INDO_az90, cov_INDO_az90,
35 ro_INDO_az90 = geostats.vmodel(nlag, xlag,azith[[0],[1]],
36 vario_INDO_az90)
37 plt.plot(h_INDO_az90,gam_INDO_az90,color='darkblue')
38
39 plt.savefig('Exp_adjusted_semivariogram_thr_1.pdf', dpi=800,
40         bbox_inches='tight')
41 plt.show()

```

The existence of preferential directions of continuity can indicate a directional propagation of the failure since the RMSD values vary more gradually in a direction in relation to the other ones. Therefore, this information, only derived from semivariograms, can be helpful in mapping this aspect.

The results of the semivariograms' fitting (Table 5) show the presence of two types of models: spherical (threshold equal to 1) and Gaussian (thresholds equal to 2, 3, and 4). Since the Gaussian model indicates a more continuous (homogeneous) relation to the spherical one, the results show a highly constant phenomenon for the thresholds equal to 2, 3, and 4, also occasioned by the absence

of the nugget effect in these. Because 62% of the RMSD values are lower than 1, the indicator kriging results for the threshold equal to 1 may be characterized to have a spatial variability much higher than in the other ones, reflected by the semivariogram's shape (fitted by spherical model), and their nugget effect, which represents 4,06% of the sill. Furthermore, the fitting of the indicator semivariogram attributed to the threshold equal to 4 (Figure 10) was based on only one point of the experimental semivariogram. However, this is an expected result once semivariograms based on indicator variables related to values situated on the distribution's upper tail may result in deteriorated semivariograms, with few points attributed to the structured region (distances lower than the range).

The input dataset and the information derived from the semivariogram fitting were used to obtain the estimates by the indicator kriging method. In this step, initially, it was necessary to define the general parameters for estimating, described in the Code 11.

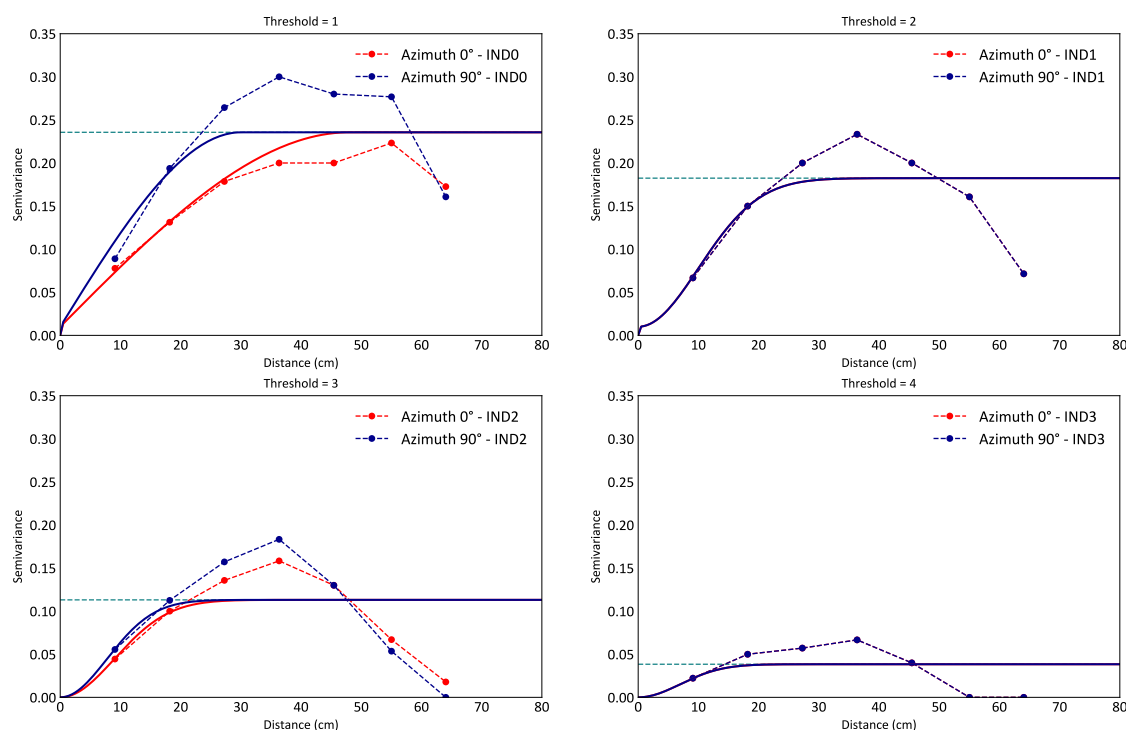


Figure 10: Experimental semivariograms in several directions for all indicator variables

Table 5: Parameters of the indicator semivariograms' fitting.

Threshold	Indicator Variable	Orientation	Nugget Effect	Structure 1		
				Model	Contribution	Range (cm)
1	IND0	Azimuth 0°	0.01	spherical	0.236	48.0
		Azimuth 90°	0.01	spherical	0.236	30.0
2	IND1	Azimuth 0°	0.01	Gaussian	0.182	24.5
		Azimuth 90°	0.01	Gaussian	0.182	24.5
3	IND2	Azimuth 0°	0.00	Gaussian	0.114	22.0
		Azimuth 90°	0.00	Gaussian	0.114	19.0
4	IND3	Azimuth 0°	0.00	Gaussian	0.039	17.0
		Azimuth 90°	0.00	Gaussian	0.039	17.00

Code 11: Source code for determining general parameters to obtain the kriging estimates

```

1 xmin = 0.0; xmax = 100.0
2 ymin = 0.0; ymax = 100.0
3 xsiz = 4; ysiz = 4
4 nx = 25; ny = 25
5 xmn = 2; ymn = 2
6 nxdis = 3; nydis = 3
7 ndmin = 1; ndmax = 4
8 radius = lag_dist[0]
9 ktype = 1
10 ivtype = 1
11 tmin = -999; tmax = 999;
12 vmin = 0.0; vmax = 1.0;
13 ncut = 4
14 pdano = np.array(df['PDANO'])
15 cont1 = 0; cont2 = 0; cont3 = 0; cont4 = 0
16 th1 = 1; th2 = 2; th3 = 3; th4 = 4
17 for i in pdano:
18     if i<=th1:
19         cont1 += 1
20     if i<=th2:
21         cont2 += 1
22     if i<=th3:
23         cont3 += 1
24     if i<=th4:
25         cont4 += 1
26
27 gcdf1 = (cont1/len(pdano))
28 gcdf2 = cont2/len(pdano)
29 gcdf3 = cont3/len(pdano)
30 gcdf4 = cont4/len(pdano)
31 print(gcdf1, gcdf2, gcdf3, gcdf4)

```

In the Code 11 it was necessary to define, initially, the block model's parameters: Minimum and maximum coordinates in the X-axis (line 1); Minimum and maximum coordinates in the Y-axis (line 2); cell's dimensions in the X and Y directions (line 3); the number of cells in the X and Y axis (line 4), and; origin coordinates of the model (line 5). In the sequence, it was specified the general parameters for the kriging process: The number of points to be estimated inside each cell in the X and Y directions (line 6), in such a way that the estimated value in each cell is the average of the estimated points in each one; the minimum and a maximum number of input data inside the search ellipse to be used (line 7); maximum search distance (line 8); kriging estimator to be used, which can be 0 for simple kriging or 1 for ordinary kriging (line 9); type of the variable being estimated, which can be 0 for categorical variable or 1 for a continuous one (line 10); trimming limits (line 11); minimum and maximum value for legend displaying (line 12), and; the number of thresholds (line 13).

Lastly, still in the Code 11, it was calculated the global cumulative distribution function of the variable *PDANO* (lines 14 to 31) related to the damage metric values for the threshold values defined previously (1, 2, 3, and 4), which was executed using a *for* loop to count the number of data lower than the respective threshold value, and then, dividing this amount by the total number of data in the *PDANO* variable, resulting in the following global cumulative distribution function: 0.62, 0.76, 0.87 and 0.96, related to the thresholds equal to 1, 2, 3 and 4, respectively.

The cell's sizes were defined in such a way due to the following: (1) the cell's dimensions

are lower than all ranges in the semivariogram models, and; (2) the cell's sizes are multiple of the plate's dimensions in each direction, allowing an adequate adjustment of the model about the actual dimensions of the plate. Also, the chosen cell's dimensions define the model's resolution (40 cm).

Regarding the kriging estimator used to obtain the estimates (parameter *ktype* in Code 11), it was used the ordinary kriging one, due to its capability to reproduce more properly the local variations of the variable being estimated (local estimation method).

In the sequence, it was defined lists for storing some information for the *ik2d* function of the Geostatpy package, shown in the Code 12.

Code 12: Source code for defining the list objects for the *ik2d* process

```

1 thresh = [th1,th2,th3,th4]
2 gcdf = [gcdf1, gcdf2, gcdf3, gcdf4]
3 varios = []
4 varios.append(GSLIB.make_variogram(nug=nugg[0],nst=1,it1=it11[0],cc1=
    cc11[0],azi1=azith[[0],[0]],hmaj1=hmaxmin11[[0],[0]],hmin1=
    hmaxmin11[[0],[1]]))
5 varios.append(GSLIB.make_variogram(nug=nugg[1],nst=1,it1=it11[1],cc1=
    cc11[1],azi1=azith[[1],[0]],hmaj1=hmaxmin11[[1],[0]],hmin1=
    hmaxmin11[[1],[1]]))
6 varios.append(GSLIB.make_variogram(nug=nugg[2],nst=1,it1=it11[2],cc1=
    cc11[2],azi1=azith[[2],[0]],hmaj1=hmaxmin11[[2],[0]],hmin1=
    hmaxmin11[[2],[1]]))
7 varios.append(GSLIB.make_variogram(nug=nugg[3],nst=1,it1=it11[3],cc1=
    cc11[3],azi1=azith[[3],[0]],hmaj1=hmaxmin11[[3],[0]],hmin1=
    hmaxmin11[[3],[1]]))
8 no_trend = np.zeros((1,1,1,1))
9 ikmap = ik2d(df,'X','Y','PDANO',ivtype,0,4,thresh,gcdf,no_trend,tmin,
    tmax,nx,xmn,xsiz,ny,ymn,ysiz,ndmin,ndmax,radius,ktype,vario=varios)

```

The Code 12 starts with the definition of the following list objects: Thresholds to be used (line 1); a global cumulative distribution function for each threshold (line 2), and; a list for receiving the semivariogram fitting information, which was specified in the sequence, between the lines 4 and 7, and appended in a single list named *varios*. In the sequence, an array of ones was created to indicate the absence of trends (line 8). Then, it was defined a variable named *ikmap* for storage of the results of the *ik2d* process, whose parameters are described in line 9 of the Code 12, where each of these is linked to a specific variable defined in the previous steps.

The indicator kriging results are in the form of *probability below thresholds* due to the procedure performed to create the indicator variables (Eq. 9). To represent the results more appropriately, the results were converted to *probability above thresholds* by applying the Eq. 11, subtracting each probability ($I(x_0)$) by one (maximum probability), resulting, then, in the probabilities of each estimated cell (x_0 location) to be higher than the respective threshold.

$$I_{Above}(x_0) = 1 - I(x_0) \quad (11)$$

The Eq. 11 was applied to the indicator kriging results and each input indicator variable, aiming to represent them in the form of *probability above thresholds*. This was done by adapting the original *locpix_st* function of the Geostatpy package, whose part of the source code is shown in the Code 13.

Code 13: Part of the *locpix_st* function's source code where the adaptations were made

```

65     array = 1-array
66     array[len(array)-1, len(array)-1] = 0
67     array[len(array)-1, 0] = 0
68     array[0, len(array)-1] = 0
69     array[0, 0] = 0
70     cs = plt.contourf(
71         xx,
72         yy,
73         array,
74         cmap=cmap,
75         vmin=vmin,
76         vmax=vmax,
77         levels=np.linspace(vmin, vmax, 100),
78     )
79
80     plt.scatter(
81         df[xcol],
82         df[ycol],
83         s=None,
84         c=1-df[vcol],
85         marker=None,
86         cmap=cmap,
87         vmin=vmin,
88         vmax=vmax,
89         alpha=0.8,
90         linewidths=0.8,
91         verts=None,
92         edgecolors="black",
93     )

```

The conversion of the *probabilities below thresholds* to the *probabilities above thresholds* in the indicator kriging results was done using the Eq. 11, applied in line 65 of the Code 13. In the sequence, the map was centralized in the plot using the procedures shown between lines 66 and 69 since the map was slightly displaced in the original function. The last modification was related to the input indicator variable, subtracted from 1 (maximum probability), done using the procedure shown in line 84. Therefore, all data plotted using the *locpix_st* function was in the form of *probability above thresholds*. The indicator kriging results and the respective input data were plotted using the function above, resulting in the plots shown in Figure 11, and the source code used to generate these are displayed in the Code 14.

For plotting the estimated values and the input data, it was used the *locpix_st* function of the Geostatpy package, as shown in the Code 14, starts with the addition of the information (estimated and input data) related to the threshold equal to 1 (lines 1 to 2). Similarly, the other plots were added by the *subplot* function. In each *locpix_st* function, it's necessary to add the succeeding information, ordered in the following sequence: an array containing the indicator kriging results to be plotted; minimum coordinate in the X-axis; maximum coordinate in the X-axis; minimum coordinate in the Y direction; maximum coordinate in the Y direction; cell's size in the X-axis; minimum value for legend; maximum value for legend; the name of the dataset (in the format of a dataframe) containing the input values; the name of the column in the dataset storing the coordinates in the X-axis; the name of the column in the dataset storing the coordinates in the Y-axis; the name of the variable in the input dataset to be plotted, in this case, the indicator variables; plot's title; label in the X-axis; label in the Y-axis; legend's title, and; color palette to be used.

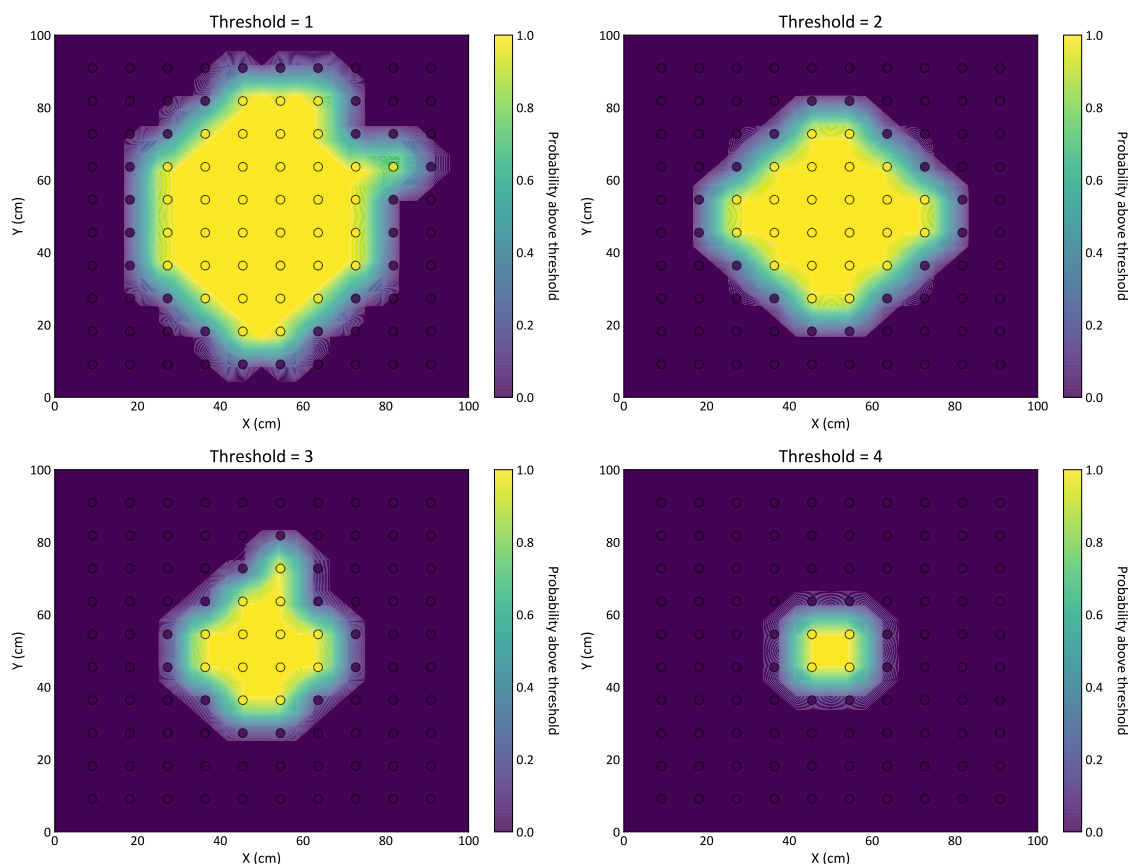


Figure 11: Probability maps of the damage metric values to be higher than each defined threshold, jointly with the respective input data

Code 14: Source code to plot the estimated and the input values of probability above thresholds

```

1 plt.subplot(221)
2 locpix_st(ikmap[:, :, 0], xmin, xmax, ymin, ymax, xsiz, 0.0, 1.0, df, 'X', 'Y', '
  IND0', 'Threshold = 1', 'X (cm)', 'Y (cm)', 'Probability above
  threshold', cmap)
3 plt.subplot(222)
4 locpix_st(ikmap[:, :, 1], xmin, xmax, ymin, ymax, xsiz, 0.0, 1.0, df, 'X', 'Y', '
  IND1', 'Threshold = 2', 'X (cm)', 'Y (cm)', 'Probability above
  threshold', cmap)
5 plt.subplot(223)
6 locpix_st(ikmap[:, :, 2], xmin, xmax, ymin, ymax, xsiz, 0.0, 1.0, df, 'X', 'Y', '
  IND2', 'Threshold = 3', 'X (cm)', 'Y (cm)', 'Probability above
  threshold', cmap)
7 plt.subplot(224)
8 locpix_st(ikmap[:, :, 3], xmin, xmax, ymin, ymax, xsiz, 0.0, 1.0, df, 'X', 'Y', '
  IND3', 'Threshold = 4', 'X (cm)', 'Y (cm)', 'Probability above
  threshold', cmap)
9 plt.subplots_adjust(left=0.0, bottom=0.0, right=2.0, top=2.2, wspace
  =0.1, hspace=0.2)
10 plt.savefig('IK2D_map.pdf', dpi=800, bbox_inches='tight')
11 plt.show()

```

The resulting plots were adjusted using the *subplots_adjust* function (line 9) and then saved in the PDF format using the *savefig* function (line 10), resulting in the plots shown in Figure 11.

Figure 11 shows the probability maps of the RMSD values being higher than their respective thresholds. The results for the threshold equal to 1 exhibit a wide area of high probability, indicating an eventual area affected by the damage. Also, the elongated shape of the area where these high values occur can mean the existence of locations more impacted by the damage than others, providing information for supporting the decision-making process in condition-based maintenance programs.

Since the damage is located at the plate's center, as the thresholds increases, the highest probabilities of the RMSD values to be greater than each threshold are sequentially restricted to the plate's center, converging the existence of these to this location, where the RMSD values were higher than 4.0. Therefore, by mapping the RMSD values to be higher than four, it was possible to predict the structural damage's location.

In the sequence, the average of the probabilities below thresholds in the indicator kriging results and in the input data (indicator variables) was calculated to compare them, analyzing the adherence of the estimates about the respective input indicator variables. The averages were calculated using the Code 15.

The Code 15 starts with the creation of the variables: The mean value of each indicator variable (*mean_inds* in line 1), the mean value of each indicator kriging-based estimated probability (*mean_iks* in line 2), the percentage deviation of the estimated probabilities about the input ones (*desv_means_percent* in line 3). In the sequence, a *for* loop (lines 4 to 7) was used to calculate the means and the percentage deviations, each related to a specific threshold value. Lastly, the resulted percentage deviations were formatted to a maximum decimal place of 2, done by the *format* function in a *for* loop (lines 10 and 11).

The average of the probabilities below thresholds obtained by the indicator kriging approach was compared to its respective input indicator variable, resulting in the plot shown in Figure 12A, whose percentage differences are plotted in Figure 12B. The source code used for preparing these is shown in the Code 16.

Code 15: Source code for calculating of the estimate's and input indicator variables' averages

```

1 mean_inds=np.zeros(len(thrs))
2 mean_iks=np.zeros(len(thrs))
3 desv_means_percent=np.zeros(len(thrs))
4 for i in range(0,len(thrs)):
5     mean_inds[i]=np.mean(df['IND'+str(i)])
6     mean_iks[i]=np.mean(ikmap[:, :, i])
7     desv_means_percent[i]=(mean_iks[i]-mean_inds[i])/mean_inds[i]*100
8
9 desv_means_percent_format=np.zeros(len(desv_means_percent))
10 for i in range(0,len(desv_means_percent)):
11     desv_means_percent_format[i]="{: .2f}".format(desv_means_percent[i],2)

```

Code 16: Source code for plotting the average of the probabilities below thresholds in the estimated and input data, and the percentage differences between these ones

```

1 fig = plt.figure(figsize=(14,5))
2 plt.subplot(121)
3 plt.plot([1,2,3,4],mean_inds,label='Input data',marker='o', color='#440154ff')
4 plt.plot([1,2,3,4],mean_iks, label='Estimated data', marker='o', color='#22A884FF')
5 plt.legend(loc='lower right', fontsize=12)
6 plt.xlabel('Threshold value', fontsize=12)
7 plt.ylabel('Mean', fontsize=12)
8 plt.ylim(0,1)
9 plt.xticks([1,2,3,4])
10 plt.text(1,0.92,'A', fontsize=16, fontweight='bold')
11
12 plt.subplot(122)
13 plt.bar([1,2,3,4],desv_means_percent,color='#2A788EFF')
14 plt.xticks([1,2,3,4])
15 plt.xlabel('Threshold value', size=12)
16 plt.ylabel('Deviation (%)', size=12)
17 plt.ylim(0,12)
18 plt.text(0.6,11,'B', fontsize=16, fontweight='bold')
19 plt.text(0.9,10,''+str(desv_means_percent_format[0])+'', fontsize=12)
20 plt.text(1.9,5,''+str(desv_means_percent_format[1])+'', fontsize=12)
21 plt.text(2.9,2.1,''+str(desv_means_percent_format[2])+'', fontsize=12)
22 plt.text(3.9,0.25,''+str(desv_means_percent_format[3])+'', fontsize=12)
23
24 plt.tight_layout()
25 plt.savefig('difference_results_input.pdf', dpi=800, bbox_inches='tight')
26 plt.show

```

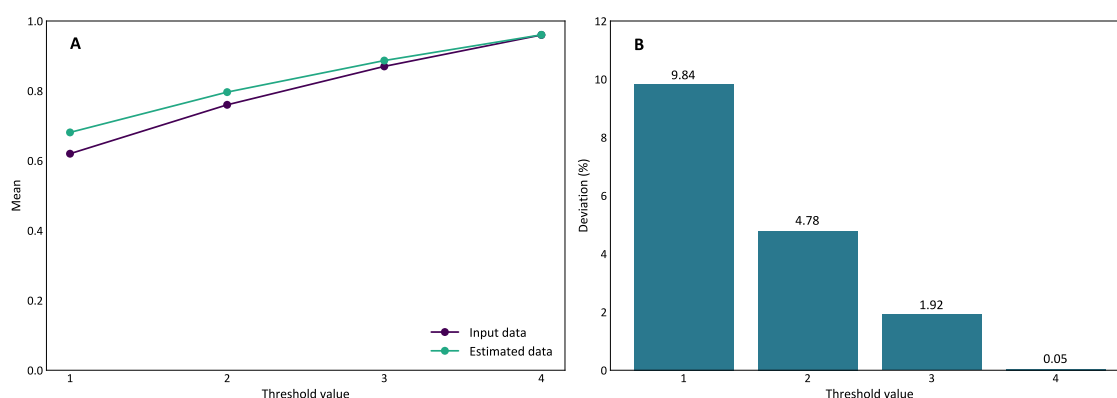


Figure 12: A) Average of the indicator variables in the input data and in the IK-based model (probabilities below thresholds). B) Percentage differences between the averages of the probabilities in the kriging-based model and the input data

In the Code 16, initially, it was defined the plot's size (line 1). In the sequence, using the *subplot* function (line 2), the first plot was positioned in the specified grid (1 row and 2 columns). Then, it was plotted the average of the probabilities below thresholds in the input data against the

threshold values (line 3) and the average of the probabilities in the estimated results against each threshold value (line 4); this first plot was formatted using the functions between the lines 5 and 10. Also, using the *subplot* function, the second plot was positioned in the grid (line 12) and was created using the *bar* function (line 13) and subsequently formatted (lines 14 to 22). Lastly, it was applied the *tight layout* option (line 24) and the final plot was saved using the *savefig* function (line 25).

Comparing the average of the probabilities in the kriging-based model with those of the input data (Figure 12A and Figure 12B), one can ensure that the mean values were approximately reproduced in the model, whose highest deviation was 9.84%, attributed to the threshold equal to 1. Such deviation can be a consequence of the higher spatial variability intrinsic to the indicator variable associated with this threshold than the other ones, observed by the presence of the spherical model in the semivariogram fitting of this variable (IND0). In contrast, in the others, the Gaussian one was used, which indicates a phenomenon with a much higher continuity (Table 10).

Also, it's noted in Figure 12B the occurrence of the smallest deviation (0.05%) in the threshold equal to 4. Despite the low reliability of the experimental semivariograms for this threshold, as stated previously, the occurrence of a smaller deviation for this one, in relation to the others, can be a consequence of the high spatial continuity in this scenario (fitting by Gaussian models), with probabilities higher than zero strictly limited to the plate's center, where the damage is located. Thus, despite instabilities in the semivariograms of indicator variables related to very high thresholds, the results were not inconsistent, allowing the prediction of the structural damage's occurrence.

5 Concluding Remarks

This chapter briefly discussed concepts related to ISHM and kriging methods used to locate structural damages. Using a case study built step by step, the elementary phases were detailed using Python language packages to find structural damage in the center of an aluminum plate. The results showed that this new approach made it possible to estimate the location of the damage, resulting in maps of easy visual identification.

Additional studies are needed to better understand the technique's applicability in structural health monitoring systems, especially regarding the dimensions of the piezoelectric element mesh (spacing and number of elements in the mesh) for the identification of the damage occurrence site, through the proposed approach. In this aspect, other geostatistical tools, such as conditional simulation, could be explored for application in structural integrity monitoring systems regarding sensor mesh determination.

References

- M. Abzalov. *Applied Mining Geology*. Springer Cham, 1 edition, 2016.
- M. Armstrong. *Basic Linear Geostatistics*. Springer Berlin, Heidelberg, 1 edition, 1998.
- F. G. Baptista and J. Vieira Filho. Optimal frequency range selection for pzt transducers in impedance-based shm systems. *IEEE Sensors Journal*, 10(8):1297–1303, 2010.
- S. Bhalla and C.-K. Soh. Electro-mechanical impedance technique. In C.-K. Soh, Y. Yang, and S. Bhalla, editors, *Smart Materials in Structural Health Monitoring, Control and Biomechanics*, pages 17–51. Springer-Verlag, 2012.

- O. Cherrier, P. Selva, V. Pommier-Budinger, F. Lachaud, and J. Morlier. Damage localization map using electromechanical impedance spectrums and inverse distance weighting interpolation: Experimental validation on thin composite structures. *Struct. Health Monit.*, 12(4):311–324, July 2013.
- J.-P. Chilès and P. Delfiner. *Geostatistics: Modeling Spatial Uncertainty*. John Wiley & Sons, 2 edition, 2012.
- J. Deutsch, M. Deutsch, R. Martin, W. Black, T. Acorn, R. Barnett, and M. Hadavand. pygeostat python package, 2020.
- V. Giurgiutiu. *Structural Health Monitoring with Piezoelectric Wafer Active Sensors*. Academic Press, 2 edition, 2014.
- V. Giurgiutiu and C. A. Rogers. Recent advancements in the electromechanical (E/M) impedance method for structural health monitoring and NDE. In M. E. Regelbrugge, editor, *Smart Structures and Materials 1998: Smart Structures and Integrated Systems*. SPIE, July 1998.
- V. Giurgiutiu, A. P. Reynolds, and C. A. Rogers. Experimental investigation of e/m impedance health monitoring for spot-welded structural joints. *Journal of Intelligent Material Systems and Structures*, 10(10):802–812, 1999.
- D. R. Gonçalves, J. d. R. V. d. Moura Junior, and P. E. C. Pereira. MONITORAMENTO DE INTEGRIDADE ESTRUTURAL BASEADO EM IMPEDÂNCIA ELETROMECÂNICA UTILIZANDO O MÉTODO DE KRIGAGEM ORDINÁRIA. *HOLOS*, 36(2):1–16, Apr. 2020.
- D. R. Gonçalves, J. d. R. V. d. Moura Junior, P. E. C. Pereira, M. V. A. Mendes, and H. S. Diniz-Pinto. Indicator kriging for damage position prediction by the use of electromechanical impedance-based structural health monitoring. *Comptes Rendus Mécanique*, 349(2):225–240, Apr. 2021.
- C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, R. Kern, M. Picus, S. Hoyer, M. H. van Kerkwijk, M. Brett, A. Haldane, J. Fernández del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke, and T. E. Oliphant. Array programming with NumPy. *Nature*, 585:357–362, 2020. doi: 10.1038/s41586-020-2649-2.
- J. D. Hunter. Matplotlib: A 2d graphics environment. *Computing in Science & Engineering*, 9(3): 90–95, 2007. doi: 10.1109/MCSE.2007.55.
- W. Hustrulid, M. Kuchta, and R. Martin. *Open Pit Mine Planning and Design*. CRC Press, 3 edition, 2013.
- E. H. Isaaks and R. M. Srivastava. *An Introduction to Applied Geostatistics*. Oxford University Press, 1 edition, 1989.
- A. G. Journel and C. J. Huijbregts. *Mining Geostatistics*. Academic Press, 1 edition, 1978.
- P. Kitanidis. *Introduction to Geostatistics: Applications in Hydrogeology*. Cambridge University Press, 1 edition, 1997.
- C. Kravolec, M. Schagerl, and M. Mayr. Localization of damages by model-based evaluation of electro-mechanical impedance measurements. NDT.net, July 2018.

- M. Labots. os-sys python package, 2019.
- C. Liang, F. P. Sun, and C. A. Rogers. Coupled electro-mechanical analysis of adaptive material systems â determination of the actuator power consumption and system energy transfer. *Journal of Intelligent Material Systems and Structures*, 5(1):12–20, 1994.
- B. Lin and V. Giurgiutiu. Modeling and testing of pzt and pvdf piezoelectric wafer active sensors. *Smart Materials and Structures*, 15(4):1085–1093, 2006.
- C. R. Marqui, D. D. Bueno, F. G. Baptista, R. B. Santos, J. Vieira Filho, and V. Lopes Júnior. External disturbance effect in damage detection using electrical impedance. In *Conference and exposition on structural dynamics*, volume 3, pages 678–687. Society for Experimental Mechanics, 2008.
- I. I. C. Maruo, G. d. F. Giachero, V. Steffen Junior, and R. M. Finzi Neto. Electromechanical impedance - based structural health monitoring instrumentation system applied to aircraft structures and employing a multiplexed sensor array. *Journal of Aerospace Technology and Management*, 7(3):294–306, 2015.
- W. McKinney. Data Structures for Statistical Computing in Python. In Stéfan van der Walt and Jarrod Millman, editors, *Proceedings of the 9th Python in Science Conference*, pages 56 – 61, 2010. doi: 10.25080/Majora-92bf1922-00a.
- G. Park, H. Sohn, C. R. Farrar, and D. J. Inman. Overview of piezoelectric impedance-based health monitoring and path forward. *The Shock and Vibration Digest*, 35(6):451–463, 2003.
- D. M. Peairs, D. J. Inman, and G. Park. Circuit analysis of impedance-based health monitoring of beams using spectral elements. *Structural Health Monitoring*, 6(1):81–94, 2007a.
- D. M. Peairs, P. A. Tarazaga, and D. J. Inman. Frequency range selection for impedance-based structural health monitoring. *Journal of Vibration and Acoustics*, 129(6):701–709, 2007b.
- M. J. Pyrcz, H. Jo, A. Kuppenko, W. Liu, A. E. Gigliotti, T. Salomaki, and J. Santos. Geostatspy python package, 2021.
- V. Raju. Implementing impedance-based health monitoring. Master’s thesis, Virginia Polytechnic Institute and State University, 1997.
- M. B. Revuelta. *Mineral Resources: From Exploration to Sustainability Assessment*. Springer Cham, 1 edition, 2018.
- M. E. Rossi and C. V. Deutsch. *Mineral Resource Estimation*. Springer Dordrecht, 1 edition, 2014.
- A. J. Sinclair and G. H. Blackwell. *Applied Mineral Inventory Estimation*. Cambridge University Press, 1 edition, 2002.
- F. P. Sun, Z. A. Chaudhry, C. A. Rogers, M. Majmundar, and C. Liang. Automated real-time structure health monitoring via signature pattern recognition. In I. Chopra, editor, *Smart Structures and Materials 1995: Smart Structures and Integrated Systems*, volume 2443, pages 236–247. SPIE, 1995.
- J. K. Yamamoto and P. M. B. Landim. *Geoestatística: Conceitos e Aplicações*. Oficina de Textos, 1 edition, 2013.

A. N. Zagrai and V. Giurgiutiu. Electromechanical impedance modeling. In C. Boller, F.-K. Chang, and Y. Fujino, editors, *Encyclopedia of Structural Health Monitoring*, pages 71–89. John Wiley & Sons, 2009.

A Module geostatspy.GSLIB

Table A1: Some functions of the geostatspy.GSLIB

Funtion	Application
ndarray2GSLIB	Utility to convert a NumPy's 1-dimensional or 2-dimensional array to a GSLIB's Geo-EAS file, which can be used in GSLIB methods
GSLIB2ndarray	Utility to convert a GSLIB's Geo-EAS file to a NumPy's 1-dimensional or 2-dimensional array for using in Python methods
Dataframe2GSLIB	Utility to convert a Pandas' Dataframe to a GSLIB's Geo-EAS file for using in GSLIB methods
GSLIB2Dataframe	Utility to convert a GSLIB's Geo-EAS file to a Pandas' DataFrame one for use in Python methods
DataFrame2ndarray	Utility to convert a Pandas' DataFrame to a NumPy's 2-D array
affine	Affine distribution transformation to correct feature mean and standard deviation
nscore	Transform an original distribution to a normal-shape one
declus	Cell-based declustering
make_variogram	Make a dictionary of semivariogram parameters for use in estimation and/or simulation methods
gamv	Calculation of semivariograms in an irregularly spaced 2-D spatial data
varmap	Map of semivariograms for regularly spaced 2-D data
varmapv	Map of semivariograms for irregularly spaced 2-D data
vmodel	Semivariogram fitting
kb2d	Simple and/or ordinary kriging for 2-D data
sgsim_uncond	Non-conditional sequential Gaussian simulation for 2-D data
sgsim	Sequential Gaussian simulation for 2-D or 3-D data
cosgsim_uncond	Non-conditional sequential Gaussian Co-simulation for 2-D data
sample	Utility to create samples from a 2-D model with provided X and Y, and append to a DataFrame
gkern	make a Gaussian kernel for convolution, moving window averaging
regular_sample	Utility to extract samples regularly spaced from a 2-D spatial model
random_sample	Utility to extract samples randomly spaced from a 2-D spatial model
pixelplt	Plotting of 2-D estimated grids. Reimplementation in Python of the GSLIB's pixelplt with Matplotlib methods
pixelplt_st	Plotting of 2-D estimated grids. Reimplementation in Python of the GSLIB's pixelplt with Matplotlib methods, with support for sub plots.
pixelplt_log_st	Plotting of 2-D estimated grids. Reimplementation in Python of the GSLIB's pixelplt with Matplotlib methods, with support for sub plots and log color bar
locpix	Plotting of 2-D estimated grids and sample points. Reimplementation in Python of a GSLIB's MOD with Matplotlib methods
locpix_st	Plotting of 2-D estimated grids and sample points. Reimplementation in Python of a GSLIB's MOD with Matplotlib methods, with support for sub plots
locpix_log_st	Plotting of 2-D estimated grids and sample points. Reimplementation in Python of a GSLIB's MOD with Matplotlib methods, with support for sub plots and log color bar
locmap	Plotting of 2-D sample points. Reimplementation in Python of a GSLIB's MOD with Matplotlib methods
locmap_st	Plotting of 2-D sample points. Reimplementation in Python of a GSLIB's MOD with Matplotlib methods, with support for sub plots
hist	Utility for histogram plotting. Reimplementation in Python of the GSLIB's hist with Matplotlib methods
hist_st	Utility for histogram plotting. Reimplementation in Python of the GSLIB's hist with Matplotlib methods, with support for sub plots

Chapter 14

Time Series Models and Expansion Optimal Linear Estimation for Time- Dependent Reliability Analysis

Chapter details

Chapter DOI:

<https://doi.org/10.4322/978-65-86503-88-3.c14>

Chapter suggested citation / reference style:

Torii, André J., et al. (2022). “Time Series Models and Expansion Optimal Linear Estimation for Time-Dependent Reliability Analysis”. In Jorge, Ariosto B., et al. (Eds.) *Uncertainty Modeling: Fundamental Concepts and Models*, Vol. III, UnB, Brasilia, DF, Brazil, pp. 459–478. Book series in Discrete Models, Inverse Methods, & Uncertainty Modeling in Structural Integrity.

P.S.: DOI may be included at the end of citation, for completeness.

Book details

Book: Uncertainty Modeling: Fundamental Concepts and Models

Edited by: Jorge, Ariosto B., Anflor, Carla T. M., Gomes, Guilherme F., & Carneiro, Sergio H. S.

Volume III of Book Series in:

Discrete Models, Inverse Methods, & Uncertainty Modeling in Structural Integrity

Published by: UnB City: Brasilia, DF, Brazil Year: 2022

DOI: <https://doi.org/10.4322/978-65-86503-88-3>

Time Series Models and Expansion Optimal Linear Estimation for Time-Dependent Reliability Analysis

André J. Torii^{1*}, Eduardo M. Medeiros²,
Henrique Kroetz³, Rafael H. Lopez⁴ and André T. Beck⁵

¹ Latin American Institute for Technology, Infrastructure and Territory (ILATIT), Federal University for Latin American Integration (UNILA), Avenida Tancredo Neves, 6731, Foz do Iguaçu/PR, Brazil, 85867-970. E-mail: andre.torii@unila.edu.br

² Center for Sciences and Agri-food Technology (CCTA), Federal University of Campina Grande (UFCG), Rua Jairo Vieira Feitosa, 1770, Pombal/PB, Brazil, 58840-000. E-mail: eduardo.morais@ufcg.edu.br

³ Center for Marine Studies (CEM), Federal University of Paraná (UFPR), Pontal do Paraná/PR, Brazil, 83255-000. E-mail: henrique.kroetz@ufpr.br

⁴ Center for Optimization and Reliability in Engineering (CORE), Federal University of Santa Catarina (UFSC), Rua João Pio Duarte Silva, s/n, Florianópolis/SC, Brazil, 88040-900. E-mail: rafael.holdorf@ufsc.br

⁵ Department of Structural Engineering, University of São Paulo (USP), Av. Trabalhador São-carlense, 400, São Carlos/SP, Brazil, 13566-590. E-mail: atbeck@sc.usp.br

*Corresponding author

Abstract

This chapter presents conceptual and computational comparisons between EOLE (Expansion Optimal Linear Estimation) and AR (Auto-Regressive) models to represent stochastic processes in the context of time-dependent reliability analysis. Even though expansion techniques, such as EOLE, are appropriate for problems where the properties of the stochastic process are explicitly known, such information is rarely available in practical situations. On the other hand, time series models, such as AR, are widely employed to represent stochastic processes from real time monitoring or available historical data. For this reason, here we compare EOLE and AR in the context of time-dependent reliability analysis. We first demonstrate how AR models can be calibrated to represent a given stochastic process with known properties. It is then demonstrated that similar results for time-dependent reliability can be obtained using the two approaches. This is an important contribution from both conceptual and practical reasons, since it demonstrates that existing AR models (and likely other types of time series models) can be directly employed for time-dependent reliability analysis, without the need to first obtain an equivalent EOLE model.

Keywords: Time Dependent Reliability Analysis; Expansion Optimal Linear Estimation; Time Series Models;

1 Introduction

Several problems from sciences and engineering cannot be appropriately studied in the deterministic context, due to uncertainties concerning its parameters. In these cases, some kind of probabilistic (i.e. stochastic) approach is necessary. In this kind of approach, some problem parameters are modeled as random variables and/or stochastic process, in order to take into account randomness and uncertainty.

A problem of great interest in engineering is that of Reliability Analysis. In this context, one wishes to evaluate the probability of occurrence of undesired events, generally classified as failures. Reliability Analysis is widely employed for design considering uncertainties in several fields of engineering, such as: aerospace, nuclear, structural, mechanical, among others.

Several engineering systems of interest are also subject to uncertain time-dependent effects, such as ageing, deterioration, corrosion and time-evolving demands. In these situations, it is necessary to model the time-dependent uncertainties as Stochastic Processes. This approach leads to what is generally known as Time-dependent Reliability Analysis, a topic of great interest in the scientific community.

Most works concerning time-dependent Reliability Analysis presented in the past employed the same technique to represent the involved stochastic processes: the Expansion Optimal Linear Estimation (EOLE). EOLE was originally presented by Li and Kiureghian [1993] and is very popular because of its generality and computational efficiency. For this reason, most available methods for time-dependent Reliability Analysis are based on the concepts of EOLE. However, it should be noted that EOLE is not, by far, the most popular technique employed to model stochastic process in practice. In fact, the most common models employed in practice are those generally known as Time Series Models, such as: Auto-Regressive (AR), Moving Average (MA), Auto-Regressive with Moving Average (ARMA) and Auto-Regressive Integrated Moving Average (ARIMA) [Cryer and Chan, 2008, Box and Jenkins, 1978]. The ARMA model, in particular, is widely employed in several fields of sciences and engineering to model real time-dependent data. In the engineering community things are not much different. Time Series Models such as AR and ARMA are much more popular in practical applications concerning modeling of time-dependent data.

The popularity of Time Series Models¹ is likely a consequence of the simple mathematical structure of such models. As will be discussed later on this work, the AR model has a much simpler mathematical structure than EOLE. For this reason, Time Series Models are generally in the reach of a larger number of scientists and engineers. However, it must be emphasized that another apparent reason for the popularity of Time Series Models is that this subject has been under careful investigation by researchers for almost a century now. Thus, there are sound mathematical foundations and efficient computational techniques concerning these methods.

Even though Time Series Models were already employed in the context of time-dependent Reliability Analysis, specific comparison to EOLE was only carried out in the works by Kroetz et al. [2020] and de Medeiros [2022]. For this reason, in this chapter we compare the application of Time Series Models (AR in particular) and EOLE in the context of time-dependent Reliability Analysis. Here we wish to answer two main questions: i) how can one apply the AR model for problems that have been traditionally modeled using EOLE; ii) are the results obtained with AR models and EOLE similar? In this work we only investigate stationary Stochastic Process, that can be modeled using the AR model. Also, we do not try to discuss the possible advantages of employing Time Series Models over EOLE, since we understand that this depends much more on the application being investigated.

¹Here we employ the term “Time Series Models” to collectively represent the MA, AR, ARMA, ARIMA models and its modifications.

The rest of this work is organized as follows. In the next section we present an overview of the literature concerning the subject of this work. We then present the main concepts concerning Reliability Analysis and time-dependent Reliability Analysis. Application of Monte Carlo Simulation for evaluation of the probability of failure is detailed in Section 5. The EOLE and the AR techniques for modeling a stochastic process are presented in Sections 6 and 7, respectively. We then present two numerical examples to demonstrate the results obtained. The concluding remarks of this work are finally summarized in Section 9.

2 Literature Overview

In the context of time-dependent Reliability Analysis, stochastic processes are frequently modeled using expansion techniques such as Karhunen-Loève (KL) and Expansion Optimal linear estimation method (EOLE). The KL expansion (see Loève [1977]) is widely employed in several fields of science and engineering. However, the EOLE, originally presented by Li and Kiureghian [1993], presents some numerical advantages in comparison to KL, mainly for the representation of stochastic process with arbitrary distribution. For this reason, EOLE is more popular in the context of time-dependent Reliability Analysis. Zhang and Ellingwood [1994] also compared expansion techniques in the context of Reliability Analysis.

An important method for time-dependent Reliability Analysis is the PHI2, originally proposed by Andrieud-Renaud et al. [2004] and latter improved by Sudret [2008]. The PHI2 method basically estimates an upper bound for the probability of failure by integration of the failure rate. The failure rate, on the other hand, is evaluated using concepts from FORM (First Order Reliability Method) [Madsen et al., 1986, Ditlevsen and Madsen, 1996a]. Some additional methods on the subject were later proposed by Li and Mourelatos [2009], Singh et al. [2010], Hu and Du [2013a], Hu and Du [2013b], Hu and Du [2014], Jiang et al. [2014], Jiang et al. [2017], Jiang et al. [2018], Hawchar et al. [2015], Hawchar et al. [2017], Moarefzadeh and Sudret [2018], Nie and Ellingwood [2005], Gong and Frangopol [2019], Zhou et al. [2017], Yu et al. [2020], Wu et al. [2020], Hu et al. [2021], Zhang et al. [2021]. However, it must be emphasized that all these methods are based on a discretization of the stochastic processes using expansion techniques such as EOLE or KL.

The works by Singh et al. [2011a], Singh et al. [2011b], Walls and Bendell [1987], Mignolet and Spanos [1989], Ho and Xie [1998], Billinton and Wangdee [2007] Sohn et al. [2000], Fugate et al. [2001], Sohn and Farrar [2001], Sohn and Farrar [2001], Owen et al. [2001] Omenzetter and Brownjohn [2006], Carden and Brownjohn [2008], Zheng and Mita [2008], Drignei [2011], Bao et al. [2013], Yao and Pakzad [2013], Goyal and Pabla [2016], Hu and Mahadevan [2017], Datteo et al. [2018], Zhang et al. [2018], Wang et al. [2021], Samaras et al. [1985] and Hu and Mahadevan [2017] employed Time Series models (i.e. AR, ARMA, among others) in the context of structural engineering and Time-dependent Reliability Analysis. These works demonstrate the popularity of Time Series models in practical engineering applications. However, no extensive comparison between Time Series models and EOLE in the context of Reliability Analysis was carried until recently. The work by Kroetz et al. [2020] was the first to extensively compare EOLE and ARMA in the context of time-dependent Reliability analysis. However, in that work the employed ARMA models were calibrated using a sample obtained with EOLE and standard techniques for Time Series calibration, described by Cryer and Chan [2008], Box and Jenkins [1978]. Although the comparisons were enough to demonstrated that ARMA was able to obtain similar results to EOLE, the calibration approach employed included sampling error. The work by Kroetz et al. [2020] was latter extended in the thesis by de Medeiros [2022], were a more accurate approach was conceived to calibrate ARMA models directly from the properties of the stochastic processes. The approach employed by de Medeiros [2022] to calibrate the ARMA is based on

an optimization scheme that searches for the time series coefficients that minimize the distance between the resulting ARMA autocovariance and the autocovariance of the stochastic process. The approach also includes constraints on stationarity and invertibility of the resulting ARMA time series.

This chapter is based on the works by Kroetz et al. [2020] and de Medeiros [2022]. However, here we focus on AR models, that are generally simpler to employ and can be calibrated by solving a system of linear equations.

3 Reliability Analysis

Consider a limit state function g that depends on a vector of random variables \mathbf{X} , such that $g < 0$ indicates failure of the system under analysis. We then define the probability of failure as [Madsen et al., 1986, Ditlevsen and Madsen, 1996b, Melchers and Beck, 2018, Beck, 2019]

$$P_f = P[g(\mathbf{X}) < 0], \quad (1)$$

where $P[\cdot]$ represents the probability of occurrence of a given event.

The most common type of limit state function that appears in Structural Safety is of the kind

$$g = R - S, \quad (2)$$

where R and S represent the resistance and the load effects, respectively. In this case, failure is assumed to occur when $S > R$. Several problems in engineering can be stated in this form, considering that R represents the capacity of the system under analysis and S represents the demand that the system must supply.

Assuming the random vector \mathbf{X} has density given by $f_X(\mathbf{x})$ with support $\Omega \subseteq \mathbb{R}^n$, the probability of failure can also be written as

$$P_f = \int_{\Omega} I(g(\mathbf{x})) f_X(\mathbf{x}) d\mathbf{x}, \quad (3)$$

where

$$I(\zeta) = \begin{cases} 1, & \zeta < 0, \\ 0, & \zeta \geq 0. \end{cases} \quad (4)$$

is the Indicator function. Note that the Indicator function simply returns 1 when its argument is negative and 0 otherwise. $I(g(\mathbf{x}))$ then indicates if g is negative at \mathbf{x} , i.e. if failure occurs for \mathbf{x} .

From the definition of expected value (see Ross [2010]) we then conclude that the probability of failure can also be written as

$$P_f = \int_{\Omega} I(g(\mathbf{x})) f_X(\mathbf{x}) d\mathbf{x} = E[I(g(\mathbf{X}))], \quad (5)$$

where $E[\cdot]$ represents the expected value of its argument. In other words, the probability of failure can also be defined as the expected value of $I(g(\mathbf{X}))$.

4 Time-dependent Reliability Analysis

The reliability analysis problem from Eq. (5) is important for a wide variety of engineering applications. However, when time-dependent effects are considered, the problem is usually restated in order to explicitly account for them. This type of problem is known as Time-Dependent Reliability Analysis. Some examples of time-dependent problems include reliability analysis considering

structures subject to ageing effects, such as corrosion; structures subject to loads with seasonal variation, such as storms; among others.

4.1 Stochastic Processes

In the context of Structural Safety, uncertain time-dependent effects are generally modeled as Stochastic Processes. A Stochastic Process can be understood as a time-dependent random variable $X(t)$, with time-dependent density function $f_X(x, t)^2$. In Figure 1 we present an example of one realization of a given stochastic process $X(t)$. A given realization of $X(t)$ is also known as a “path” and is represented as $x(t)$. Note that a path is a time-dependent function. If one tries to observe a stochastic process several times, different paths would be observed. A group of several paths of a given stochastic process is called an “ensemble” of $X(t)$.

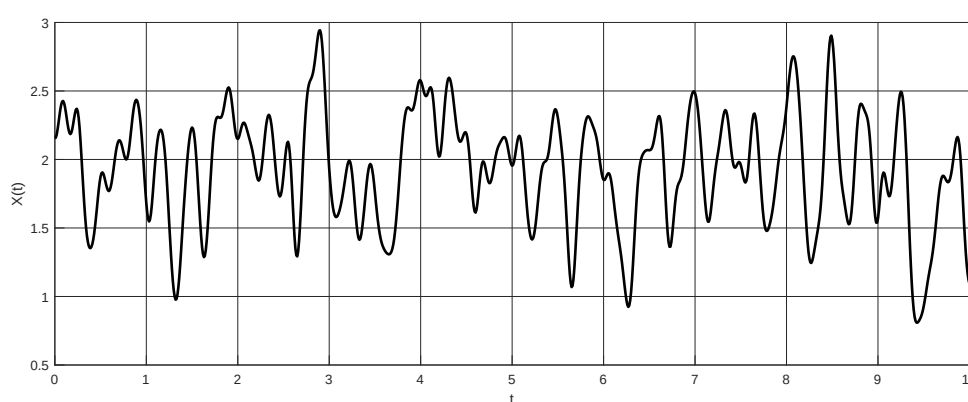


Figure 1: A path of a given stochastic process $X(t)$

Strictly speaking, in order to define a stochastic process $X(t)$ we need to define the time-dependent density $f(x, t)$ for each time instant t [Shriryayev, 1995, Loève, 1977]. This can be done explicitly or implicitly. However, explicit definition of $f(x, t)$ is uncommon, for practical reasons.

A common approach for defining a stochastic process is to define it implicitly (i.e. indirectly), by defining the following properties: a parametric distribution for $X(t)$ (e.g. Normal, Uniform, Gamma, among others); a time-dependent expected value

$$\mu(t) = E[X(t)]; \quad (6)$$

a time-dependent standard deviation

$$\sigma(t) = \sqrt{V[X(t)]} = \sqrt{E[(X(t) - \mu(t))^2]}, \quad (7)$$

where $V[\cdot]$ represents the variance, and the auto-correlation

$$\rho[X(t_a), X(t_b)] = \frac{C[X(t_a), X(t_b)]}{\sigma(t_a)\sigma(t_b)} = \frac{E[(X(t_a) - \mu(t_a))(X(t_b) - \mu(t_b))]}{\sigma(t_a)\sigma(t_b)}, \quad (8)$$

where $C[X(t_a), X(t_b)]$ represents the covariance between $X(t_a)$ and $X(t_b)$. Note that $\rho[X(t_a), X(t_b)]$ measures the correlation between $X(t)$ at time-instants t_a and t_b .

²Higher order density functions are required to describe the correlation structure of the process.

If the density $f(x, t)$ of the stochastic process $X(t)$ does not depend on t the process is called “strictly stationary”. In this case the process has constant expected value

$$\mu = E[X(t)] \quad (9)$$

and constant standard deviation

$$\sigma = \sqrt{V[X(t)]}. \quad (10)$$

Besides, the auto-correlation of a strictly stationary process depends only on the distance between the time instants and thus

$$\rho[X(0), X(\tau)] = \rho[X(t), X(t + \tau)]. \quad (11)$$

For this reason, the auto-correlation of a strictly stationary process is frequently written as

$$\rho(\tau) = \rho[X(0), X(\tau)] = \frac{E[(X(0) - \mu)(X(\tau) - \mu)]}{\sigma^2}, \quad (12)$$

that results

$$\rho(\tau) = \frac{E[X(0)X(\tau)] - \mu^2}{\sigma^2}. \quad (13)$$

A stochastic process that satisfy Eqs. (9), (10) and (13) is called “weakly stationary” or, more often, stationary. Note that a process $X(t)$ that satisfies Eqs. (9), (10) and (13) does not necessarily has time-independent density $f(x, t)$. In other words, a stationary process is not necessarily “strictly stationary”. However, any “strictly stationary” process is also “weakly stationary”. This means that “strict stationarity” is a stronger requirement.

It is also important to highlight that for some specific stochastic process, satisfaction of Eqs. (9), (10) and (13) imply “strict stationarity”. This is the case of the important family of Normal/Gaussian stochastic process, where $X(t)$ has Normal/Gaussian distribution [Shriryaev, 1995, Loève, 1977]. Thus, stationary Normal/Gaussian stochastic process are also “strict stationarity”. However, note that this is not the general rule.

4.2 Time-dependent Probability of Failure

Consider a time-dependent limit state function $g(\mathbf{Y}, X(t), t)$, that depend on a vector of random variables \mathbf{Y} , a stochastic process $X(t)$ and time t . In this case the total (or cumulative) probability of failure in the time interval $[t_1, t_2]$ can be defined as

$$P_f(t_1, t_2) = P[\exists \xi \in [t_1, t_2], \mathbf{y} \in \Omega | g(\mathbf{y}, X(\xi), \xi) < 0], \quad (14)$$

that is the probability of existing some time instant ξ in the time interval and some \mathbf{y} in the support of the random variables for which failure occurs.

On the other hand, the instantaneous probability of failure is defined as

$$p_f(t) = P[g(\mathbf{Y}, X(t), t) < 0]. \quad (15)$$

Note that evaluation of the instantaneous probability of failure is basically a time-independent reliability analysis problem, since the time instant t is fixed.

In order to simplify notation a little, from now on we will represent the above quantities omitting the explicit dependence on the random vector \mathbf{Y} and time t , as is commonly done in literature. In this case we write simply

$$g(X(t)) = g(\mathbf{Y}, X(t), t) \quad (16)$$

and consequently

$$P_f(t_1, t_2) = P[\exists \xi \in [t_1, t_2] | g(X(\xi)) < 0], \quad (17)$$

$$p_f(t) = P[g(X(t)) < 0], \quad (18)$$

where dependence on \mathbf{Y} and t is implicit, i.e. it is implied it should be considered, but it is not written out in the expression.

4.3 Approaches for Time-dependent Reliability Analysis

There exist two main approaches for estimating the cumulative probability of failure $P_f(t_1, t_2)$. The first approach is to generate an ensemble for the stochastic process and then employ a sample-based estimate, such as Monte Carlo Simulation (MCS), that directly evaluates $P_f(t_1, t_2)$ from Eq. (17). In this case, the probability of failure is estimated from the number of cases where g becomes negative at least once in the time interval divided by the total number of cases considered. We will cover this approach in more details later this chapter.

The second approach is based on the fact that (see Ditlevsen and Madsen [1996b] for a proof and more extensive discussion on this subject)

$$P_f(t_1, t_2) \leq p_f(t_1) + \int_{t_1}^{t_2} v^+(t) dt, \quad (19)$$

where $v^+(t)$ is the so-called “up-crossing rate”. Formally, the up-crossing rate is defined as

$$v^+(t) = \lim_{\Delta t \rightarrow 0} \frac{P[g(X(t + \Delta t)) < 0 | g(X(t)) \geq 0]}{\Delta t}. \quad (20)$$

Note that the up-crossing rate represents the time-instant probability of getting $g(X(t + \Delta t)) < 0$ if we have $g(X(t)) \geq 0$. In other words, the up-crossing rate can be understood as the probability of departing for the safety and getting into failure domain.

Note that integration of the up-crossing rate produces an upper bound for the the cumulative probability of failure. However, it turns out that this upper bound becomes sharp (i.e. accurate) for small $v^+(t)$ [Ditlevsen and Madsen, 1996b]. In other words, for most practical situations we can assume the approximation

$$P_f(t_1, t_2) \approx p_f(t_1) + \int_{t_1}^{t_2} v^+(t) dt, \quad (21)$$

when $v^+(t)$ is small. For this reason, several approaches estimate the cumulative probability of failure using the above approximation.

Some methods based on the integration of the up-crossing rate are PHI2 [Andrieud-Renaud et al., 2004], PHI2+ [Sudret, 2008], JUR/FORM [Hu and Du, 2013a], TRPD [Jiang et al., 2014], iTRPD [Jiang et al., 2018] and NEWREL [Gong and Frangopol, 2019], among others. The difference between these methods concerns mainly how the up-crossing rate $v^+(t)$ is evaluated. In fact, very efficient approaches can be conceived by employment of FORM (First Order Reliability Method) [Madsen et al., 1986, Ditlevsen and Madsen, 1996b, Melchers and Beck, 2018] concepts and similar reasoning.

The time-dependent probability of failure can also be estimated using sampling-based schemes. This approach will be covered in detail in the next section.

5 Monte Carlo Simulation

5.1 Time-independent Reliability Analysis

The probability of failure from Eq. (5) can be estimated by MCS as [Madsen et al., 1986, Melchers and Beck, 2018]:

$$P_f \approx \tilde{P}_f = \frac{1}{N} \sum_{i=1}^N I(g(x_i)) = \frac{n_f}{N}, \quad (22)$$

where N is the sample size and n_f is the number of failures observed in the sample. In other words, the probability of failure can be estimated as the ratio between the number of failures observed and the sample size.

The variance and the coefficient of variation of the estimate \tilde{p}_f are given by [Madsen et al., 1986, Melchers and Beck, 2018]

$$Var[\tilde{P}_f] = \frac{\tilde{P}_f(1 - \tilde{P}_f)}{N} \quad (23)$$

$$CV[\tilde{P}_f] = \frac{\sqrt{Var[\tilde{P}_f]}}{E[\tilde{P}_f]} = \frac{1}{\sqrt{N}} \sqrt{\frac{1 - \tilde{P}_f}{\tilde{P}_f}}. \quad (24)$$

The coefficient of variation $CV[\tilde{P}_f]$ is a measure of the relative error of the estimate. Note that this relative error depends on the probability of failure and the sample size N . In particular, small probabilities of failure will lead to large errors unless large samples are employed.

5.2 Time-dependent Reliability Analysis

The time-dependent probability of failure from Eq. (17) can be estimated with MCS as

$$P_f(t_1, t_2) \approx \tilde{P}_f(t_1, t_2) = \frac{1}{N} \sum_{i=1}^N I(g^*(x_i(t))) \quad (25)$$

with

$$g^*(x_i(t)) = \min_{t \in [t_1, t_2]} g(x_i(t)). \quad (26)$$

Note that in this case the indicator function considers the minimum value of the limit state function g in the time interval $[t_1, t_2]$, for a given realization $x_i(t)$ of the stochastic process. This is equivalent to checking if failure occurs for the worst situation observed in the time interval $[t_1, t_2]$.

However, in order to estimate the probability of failure we first need to obtain an ensemble for the stochastic process, composed by several possible paths $x_1(t), x_2(t), \dots, x_N(t)$. Two approaches that can be employed for this purpose will be described in the next sections.

6 Expansion of the Optimal Linear Estimation (EOLE)

In the context of time-dependent reliability analysis, the most popular approach for obtaining an ensemble for the stochastic process is based on the EOLE (Expansion of the Optimal Linear Expansion), proposed by Li and Kiureghian [1993]. The EOLE is an expansion technique, i.e.,

the stochastic process $X(t)$ is represented as a series expansion. Assuming the time interval of analysis is meshed into N_t time instants $\{t_1, t_2, \dots, t_{N_t}\}$, the EOLE representation of $X(t)$ is written as

$$\widehat{X}(t) = \mu(t) + \sum_{i=1}^N \frac{\xi_i}{\sqrt{\lambda_i}} \phi_i^T \Sigma(t), \quad (27)$$

where $\mu(t)$ is the expected value of $X(t)$, ξ_i are independent standard Normal random variables, ϕ_i and λ_i are the eigenvectors and eigenvalues of the autocovariance matrix Σ , with components

$$\Sigma_{ij} = Cov[X(t_i), X(t_j)] \quad (28)$$

and $\Sigma(t)$ is a vector with components

$$\Sigma(t)_i = Cov[X(t), X(t_i)]. \quad (29)$$

Note that Σ is a matrix of covariances among time instants $\{t_1, t_2, \dots, t_{N_t}\}$, while $\Sigma(t)$ is a vector of covariances between time instant t and time instants $\{t_1, t_2, \dots, t_{N_t}\}$. Also note that we can write

$$\Sigma_{ij} = \sigma(t_i)\sigma(t_j)\rho[X(t_i), X(t_j)] \quad (30)$$

and

$$\Sigma(t)_i = \sigma(t)\sigma(t_i)\rho[X(t), X(t_i)], \quad (31)$$

where $\sigma(t)$ is the standard deviation of $X(t)$.

As occurs in expansion schemes, the number of terms of the series can be truncated. Thus, in general we take

$$EOLE : \widehat{X}(t) = \mu(t) + \sum_{i=1}^r \frac{\xi_i}{\sqrt{\lambda_i}} \phi_i^T \Sigma(t), \quad (32)$$

where $r \leq N$ is the number of terms kept in the expansion.

From Eq. (32) we observe that the stochastic process is represented as a series of eigenvectors/eigenvalues and independent standard normal random variables ξ_i . Thus, by sampling values for the random variables ξ_i it is possible to generate several paths for the stochastic process, i.e. to generate an ensemble for $X(t)$. This is an important characteristic of EOLE (and other expansion techniques such as Karhunen-Loève), the stochastic process is written as a series of random variables. This allows sampling of the stochastic process by sampling the random variables with standard schemes.

6.1 Numerical Adjustments

In practice, the autocovariance matrix Σ can become ill conditioned, particularly when N_t increases. When this happens, the eigenvalues λ_i evaluated numerically can become negative, blocking the EOLE algorithm from running (note that EOLE equations assume that all eigenvalues are non-negative, otherwise we get imaginary values). In practice this means that the EOLE equations only hold when Σ is positive semi-definite, i.e. when all eigenvalues satisfy $\lambda_i \geq 0$. It turns out that for numerical reasons Σ may become indefinite in practice, i.e. there may exist some negative eigenvalues arising from numerical errors.

In order to avoid this issue, here we consider a relaxed autocovariance matrix

$$\hat{\Sigma} = \Sigma + \varepsilon \mathbf{I}, \tag{33}$$

where $\varepsilon > 0$ is a small relaxation parameter and \mathbf{I} is the identity matrix. By summing non-negative ε to the diagonal of Σ we ensure that it will be positive definite. This is the procedure employed in this chapter.

7 Auto-Regressive (AR) Model

A stochastic process $X(t)$ can also be represented using Time Series Models. The most common models are: Auto-Regressive (AR), Moving Average (MA), Auto-Regressive with Moving Average (ARMA) and Auto-Regressive Integrated Moving Average (ARIMA). Note that the AR, the MA and the ARMA models are appropriate to represent stationary stochastic processes, while ARIMA can be used to model non-stationary processes. See Cryer and Chan [2008], Box and Jenkins [1978] for a detailed discussion on Time Series.

Consider a stationary stochastic process $X(t)$ with expected value μ . We first define

$$Y(t) = X(t) - \mu, \tag{34}$$

i.e., $Y(t)$ is the deviation from the expected value. Assuming the time interval of analysis is meshed into N_t time instants $\{t_1, t_2, \dots, t_{N_t}\}$ uniformly spaced with

$$\Delta t = t_i - t_{i-1}. \tag{35}$$

The AR model can be written as

$$\hat{Y}_i = \varphi_1 \hat{Y}_{i-1} + \varphi_2 \hat{Y}_{i-2} + \dots + \varphi_p \hat{Y}_{i-p} + \epsilon_i, \tag{36}$$

where $\hat{Y}_i = Y(t_i)$, $\{\varphi_1, \varphi_2, \dots, \varphi_p\}$ are the parameters of the model, p is the order of the model and ϵ_i is a gaussian white noise with null expected value and standard deviation σ_ϵ . Note that the AR model is a linear combination of the past p values of Y plus a white noise. This is the reason why the model is called Auto-regressive: it explicitly depends on it's own past values.

Consider the autocorrelation between time instants $t, t + n\Delta$ (i.e. the lag n autocorrelation)

$$\rho_n = \rho(n\Delta t), \tag{37}$$

where ρ is as defined in Eq. (13). It can be demonstrated that ρ_n satisfy [Cryer and Chan, 2008]

$$\begin{cases} \rho_1 = \phi_1 + \phi_2 \rho_1 + \phi_3 \rho_2 + \dots + \phi_p \rho_{p-1} \\ \rho_2 = \phi_1 \rho_1 + \phi_2 + \phi_3 \rho_1 + \dots + \phi_p \rho_{p-2} \\ \vdots \\ \rho_p = \phi_1 \rho_{p-1} + \phi_2 \rho_{p-2} + \phi_3 \rho_{p-3} + \dots + \phi_p \end{cases}, \tag{38}$$

that are known as Yule-Walker equations.

The Yule-Walker equations are generally employed to predict the discrete autocorrelation of the AR model \hat{Y} as a function of the parameters $\phi_1, \phi_2, \dots, \phi_p$. Here we employ the Yule-Walker equations to obtain these parameters assuming the autocorrelation function is known. For this purpose, note that we can write the Yule-Walker equations as

$$\mathbf{R}\phi = \mathbf{r}, \tag{39}$$

with

$$\mathbf{r} = \{\rho_1, \rho_2, \dots, \rho_p\}^T \tag{40}$$

$$\boldsymbol{\phi} = \{\phi_1, \phi_2, \dots, \phi_p\}^T \tag{41}$$

$$\mathbf{R} = \begin{bmatrix} 1 & \rho_1 & \rho_2 & \rho_3 & \dots & \rho_{p-1} \\ \rho_1 & 1 & \rho_1 & \rho_2 & \dots & \rho_{p-2} \\ \rho_2 & \rho_1 & 1 & \rho_1 & \dots & \rho_{p-2} \\ \vdots & \vdots & \vdots & 1 & \ddots & \vdots \\ \rho_{p-1} & \rho_{p-2} & \rho_{p-3} & \rho_{p-4} & \dots & 1 \end{bmatrix}. \tag{42}$$

Besides, the components of \mathbf{R} can be written as

$$R_{ij} = \rho_{j-i}, \tag{43}$$

since $\rho_0 = 1$ and $\rho_n = \rho_{-n}$.

Both \mathbf{R} and \mathbf{r} can be evaluated if the autocorrelation function of the stochastic process is known. In this case, by solving the system of linear equations from Eq. (39) we are able to obtain the parameters $\boldsymbol{\phi}$ of the AR(p) model directly from the autocorrelation of the stochastic process. Finally, the standard deviation of the white noise can be evaluated from [Cryer and Chan, 2008]

$$\sigma_e^2 = \sigma \left(1 - \sum_{i=1}^p \phi_i \rho_i \right), \tag{44}$$

where σ is the standard deviation of the stationary stochastic process.

7.1 Stationarity of Auto-Regressive (AR) Model

The AR model only results in a stationary time series if the roots of the characteristic equation

$$1 - \phi_1 x - \phi_2 x^2 - \dots - \phi_p x^p = 0 \tag{45}$$

are all larger than unity in magnitude (note that some roots may be imaginary) [Cryer and Chan, 2008]. Dividing by $-\phi_p$ we can write the above equation as

$$-\frac{1}{\phi_p} + \frac{\phi_1}{\phi_p} x + \frac{\phi_2}{\phi_p} x^2 + \dots + x^p = 0. \tag{46}$$

The roots of the above polynomial equation can be obtained from the eigenvalues of the companion matrix [Tismenetsky, 1992]

$$\hat{\mathbf{C}}_L = \begin{bmatrix} 0 & 0 & \dots & 0 & 1/\phi_p \\ 1 & 0 & \dots & 0 & -\phi_1/\phi_p \\ 0 & 1 & \dots & 0 & -\phi_2/\phi_p \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & -\phi_{p-1}/\phi_p \end{bmatrix}. \tag{47}$$

Thus, the AR model will be stationary if

$$|\lambda_i| > 1, i = 1, 2, \dots, p, \tag{48}$$

where λ_i are the eigenvalues of

$$\hat{\mathbf{C}}_L \mathbf{v} = \lambda \mathbf{v}. \quad (49)$$

Tismenetsky [1992] demonstrates that the condition from Eq. (48) is satisfied if \mathbf{R} is positive definite. It is known that the correlation matrix is always positive semi-definite, otherwise negative variance should occur. However, \mathbf{R} is a Toeplitz matrix that can become singular when its order is increased. Thus, in order to ensure that the Yule-Walker equations always consider a positive definite matrix some numerical adjustment can be employed.

In order to ensure stationarity of the AR models, in this work we obtain the coefficients of the AR models from the relaxed system of linear equations

$$(\mathbf{R} + \varepsilon \mathbf{I}) \boldsymbol{\phi} = \mathbf{r}, \quad (50)$$

where $\varepsilon > 0$ is a small relaxation parameter and \mathbf{I} is the identity matrix. By summing non-negative ε to the diagonal of \mathbf{R} we ensure that it will be positive definite. Thus, the resulting AR model will be stationary.

It is interesting to note that both the AR models and the EOLE expansion require some kind of relaxation in order to be numerically robust in practice. Besides, in both cases we observe that the relaxation employed must ensure positive definiteness of the autocovariance/autocorrelation matrix. Here we do not pursue this interesting idea further.

8 Numerical Examples

We now present two numerical examples in order to compare the results obtained using EOLE and the AR model to represent the stochastic processes in the context of time-dependent reliability analysis. In the examples, the autocorrelation function is assumed to be of the form

$$\rho(\tau) = \exp \left[- \left(\frac{\Delta t}{\alpha} \right)^2 \right], \quad (51)$$

where α is known as the correlation length. The autocorrelation function from Eq. (51) was chosen because it is widely employed in the literature in the context of time-dependent reliability analysis.

The relative difference between the results obtained with the AR model and EOLE was evaluated as

$$e(\%) = \left| \frac{\bar{P}_f - P_{REF}}{P_{REF}} \right| 100\%, \quad (52)$$

where P_{REF} is a reference value for the probability of failure and \bar{P}_f an estimate.

In both the EOLE and the AR approaches we employed a relaxation parameter equal to $\varepsilon = 10^{-12}$, that was observed to make the algorithms numerically stable without affecting accuracy of the results. It is highlighted that if no relaxation parameter is employed, both EOLE and AR become unstable if the order of the models increases too much.

8.1 Example 1: Constant Barrier

In the first example we consider a stationary Gaussian process $X(t)$ with expected value $\mu = 2$ and standard deviation $\sigma = 0.4$. The autocorrelation is given by Eq. (51) with correlation length $\alpha = 1.0$. The limit state function is given by

$$g = 3 - X(t), \quad (53)$$

i.e. failure is assumed to occur if $X(t) > 3$. The cumulative probability of failure was evaluated for the time interval $[0, 10]$, using a discretization $\Delta t = 0.1$ (i.e. 101 uniformly spaced time instants).

The reference solution $P_f = 0.099799$ was obtained using EOLE with all terms in the expansion (i.e. $r = 101$ in Eq. (32)) and a sample size $N = 10^6$. The results obtained using AR and EOLE models of different orders and a sample size $N = 10^5$ are presented in Tables 1 and 2.

Table 1: Results obtained with AR models (Example 1)

AR	P_f	$e(\%)$
$p = 2$	0.08604	0.13786711289692
$p = 4$	0.09261	0.07203478992775
$p = 8$	0.09809	0.01712442008437
$p = 16$	0.09819	0.01612240603613
$p = 32$	0.10027	0.00471948616719
$p = 64$	0.10025	0.00451908335754
$p = 100$	0.09942	0.00379763324281

Table 2: Results obtained with EOLE models (Example 1)

EOLE	P_f	$e(\%)$
$r = 2$	0.00252	0.97474924598442
$r = 4$	0.01856	0.81402619264722
$r = 8$	0.06916	0.30700708423932
$r = 16$	0.09842	0.01381777372518
$r = 32$	0.10040	0.00602210442990
$r = 64$	0.10062	0.00822653533602
$r = 100$	0.10023	0.00431868054790

The relative differences to the reference solution $e(\%)$ versus the order of the models employed is plotted in Figure 2 in log-log scale. We observe that EOLE gives very poor results with less than 8 terms in the expansion. The AR, on the other hand, is able to obtain consistent results even with models of order $p = 2$ and 4. As the order of the model is increased, both the EOLE and the AR converge to similar values.

8.2 Example 2: Beam Subject to Corrosion

In this example, adapted from Andrieud-Renaud et al. [2004], we consider a steel beam subject to corrosion. The limit state function is taken as

$$g = \frac{(b_0 - 2\kappa t)(h_0 - 2\kappa t)^2 \sigma_y}{4} - \left(\frac{F(t)L}{4} + \frac{\rho_{st} b_0 h_0 L^2}{8} \right), \quad (54)$$

where b_0 and h_0 are the initial cross section dimensions, L is the beam length, κ is the corrosion rate t is the time measured in years, σ_y is the yielding stress, ρ_{st} is the density of the material. It is assumed that the corrosion depth evolves in time as

$$d_c(t) = \kappa t. \quad (55)$$

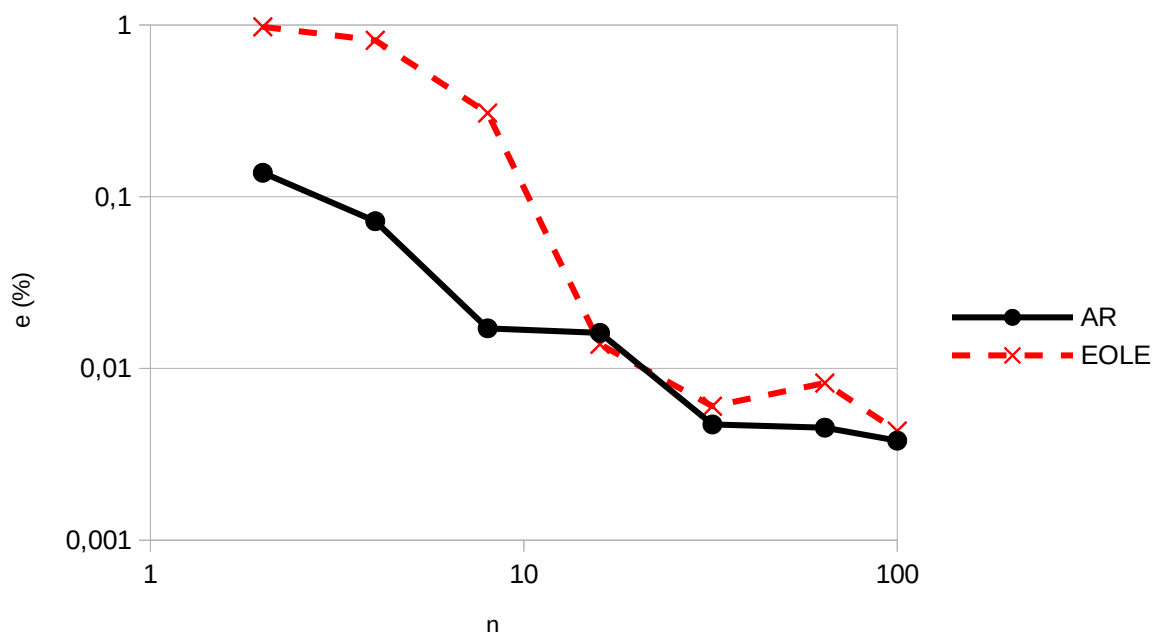


Figure 2: Convergence of AR and EOLE models (Example 1)

The random variables are as defined in Table 3. The correlation length of the Stochastic Process (i.e. applied force) is $\alpha = 1/12$ years.

Table 3: Parameters of Example 2

Parameter	Distribution	Expected value	CV*
Applied force F (kN)	Gaussian Stochastic Process	5000	20%
Yielding stress σ_y (MPa)	Lognormal	240	10%
Initial cross section width b_0 (m)	Lognormal	0.20	5%
Initial cross section height h_0 (m)	Lognormal	0.04	10%
Corrosion rate κ (mm/ano)	Deterministic	0.03	-
Beam length L (m)	Deterministic	5.00	-
Steel density ρ_{st} (kN/m ³)	Deterministic	78.50	-

* CV: coefficient of variation

The time interval of analysis was taken as $[0, 20]$ years. This interval was divided into 200 uniformly spaced sub-intervals with $\Delta t = 0.1$ years. The reference solution $P_f = 0.022138$ was obtained using EOLE with all terms in the expansion (i.e. $r = 201$ in Eq. (32)) and a sample size $N = 10^6$. The results obtained using AR and EOLE models of different orders and a sample size $N = 10^5$ are presented in Tables 4 and 5.

The relative differences to the reference solution $e(\%)$ versus the order of the models employed is plotted in Figure 3 in log-log scale. We observe that EOLE gives very poor results unless almost all terms are included in the expansion (i.e. $r = 200$). The AR, on the other hand, is able to obtain consistent results even with models of order $p = 2$ and 4. For this reason, it is not possible to identify advantages of employing high order AR models in this example.

Table 4: Results obtained with AR models (Example 2)

AR	P_f	$e(\%)$
$p = 2$	0.02178	0.01617128918601
$p = 4$	0.02172	0.01888156111663
$p = 8$	0.02197	0.00758876140572
$p = 16$	0.02242	0.01273827807390
$p = 32$	0.02119	0.04282229650374
$p = 64$	0.02115	0.04462914445749
$p = 128$	0.02123	0.04101544855000
$p = 200$	0.02185	0.01300930526696

Table 5: Results obtained with EOLE models (Example 2)

EOLE	P_f	$e(\%)$
$p = 2$	2.60E-04	0.98825548830066
$p = 4$	2.30E-04	0.98961062426596
$p = 8$	4.00E-04	0.98193152046255
$p = 16$	7.50E-04	0.96612160086728
$p = 32$	0.00177	0.92004697804679
$p = 64$	0.00534	0.75878579817508
$p = 128$	0.01385	0.37437889601590
$p = 200$	0.02188	0.01165416930165

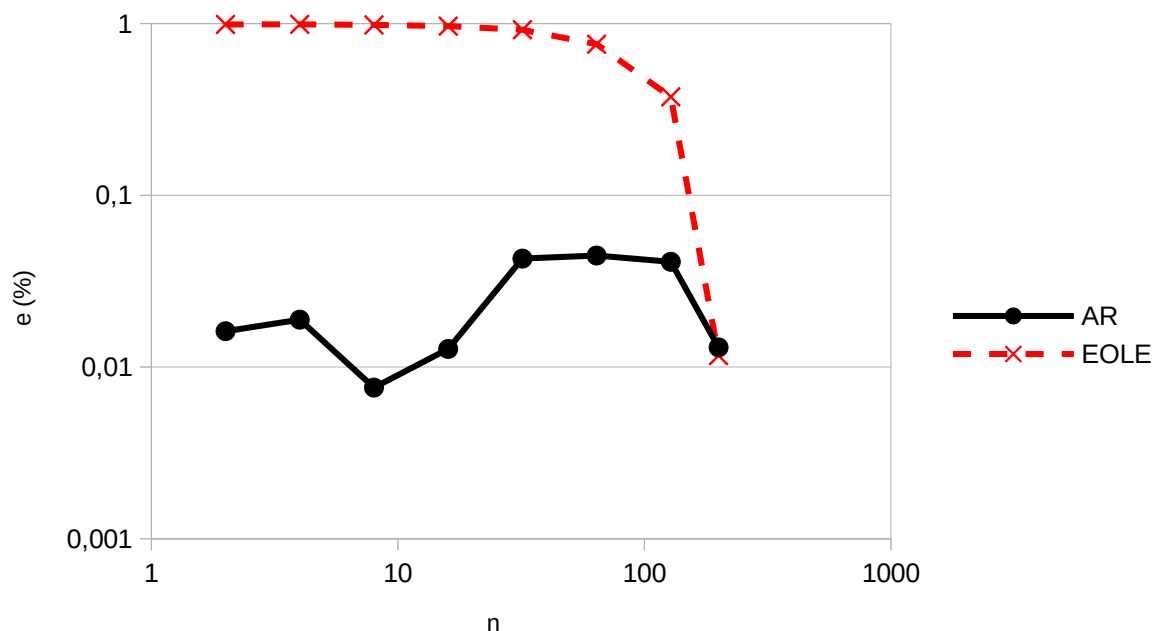


Figure 3: Convergence of AR and EOLE models (Example 2)

9 Concluding Remarks

We presented a conceptual and computational comparison between EOLE (Expansion Optimal Linear Estimation) and AR (Auto-Regressive) models to represent stochastic processes in the

context of time-dependent reliability analysis. Although EOLE is by far the most widespread technique to represent stochastic process in this context, it can be applied only if the properties of the stochastic process are explicitly known. However, such information is rarely available in practice. Time Series Models such as AR, on the other hand, are efficient to represent existing data obtained from real time monitoring or available historical databases. For this reason, in this chapter we investigated if it is possible to obtain similar results using EOLE and AR models in the context of time-dependent reliability analysis.

The results obtained here demonstrated that EOLE and AR models are able to provide similar results for time-dependent reliability analysis. This is an important conclusion, since AR (and other Time Series Models) are widely employed in practice. However, some questions should be further investigated in future works. First, it is necessary to investigate if the conclusions obtained here also hold when non-stationary stochastic processes are considered. Besides, it is also interesting to study which Time Series Models (AR, ARMA, ARIMA, among others) are more fit for reliability analysis problems.

Acknowledgements

The authors would like to acknowledge the financial support of CNPq and FAPESP.

References

- C. Andrieud-Renaud, B. Sudret, and M. Lemaire. The phi2 method: a way to compute time-variant reliability. *Reliability Engineering and System Safety*, 84:75–86, 2004.
- C. Bao, H. Hao, and Z. Li. Vibration-based structural health monitoring of offshore pipelines: numerical and experimental study. *Structural Control and Health Monitoring*, 20(5):769–788, 2013. doi: 10.1002/stc.1494.
- A. T. Beck. *Confiabilidade e Segurança das Estruturas*. Elsevier, São Paulo, 2019.
- R. Billinton and W. Wangdee. Reliability-based transmission reinforcement planning associated with large-scale wind farms. *IEEE Transactions on power systems*, 22(1):34–41, 2007.
- G. E. P. Box and G. M. Jenkins. *Time series analysis : forecasting and control*, volume Revised Edition. Holden-Day, 1978.
- E. P. Carden and J. M. Brownjohn. Arma modelled time-series classification for structural health monitoring of civil infrastructure. *Mechanical Systems and Signal Processing*, 22(2):295–314, 2008. doi: 10.1016/j.ymsp.2007.07.003.
- J. D. Cryer and K. S. Chan. *Time series analysis : with applications in R*, volume Second Edition. Springer, 2008.
- A. Datteo, G. Busca, G. Quattromani, and A. Cigada. On the use of ar models for shm: A global sensitivity and uncertainty analysis framework. *Reliability Engineering and System Safety*, 170: 99–115, 2018. doi: 10.1016/j.ress.2017.10.017.
- E. M. de Medeiros. *Análise de Confiabilidade dependente do tempo usando modelos de Séries Temporais*. PhD thesis, Universidade Federal da Paraíba, João Pessoa-PB, 2022.
- O. Ditlevsen and H. O. Madsen. *Structural Reliability Methods*. John Wiley & Sons, Chichester, 1996a.

- O. Ditlevsen and H. O. Madsen. *Structural Reliability Methods*. John Wiley & Sons, Chichester, 1996b.
- D. Drignei. A general statistical model for computer experiments with time series output. *Reliability Engineering and System Safety*, 96(4):460–467, 2011. doi: 10.1016/j.ress.2010.11.006.
- M. L. Fugate, H. Sohn, and C. R. Farrar. Vibration-based damage detection using statistical process control. *Mechanical Systems and Signal Processing*, 15(4):707–721, 2001. doi: 10.1006/mssp.2000.1323.
- C. Gong and D. M. Frangopol. An efficient time-dependent reliability method. *Structural Safety*, 81(64):10–18, 2019.
- D. Goyal and B. S. Pabla. The vibration monitoring methods and signal processing techniques for structural health monitoring: A review. *Archives of Computational Methods in Engineering*, 23(2):585–594, dec 2016. doi: 10.1007/s11831-015-9145-0.
- L. Hawchar, C. P. E. Soueidy, and F. Shoefs. Time-variant reliability analysis using polynomial chaos expansion. In *12 International conference on applications of statistics and probability in civil engineering. ICASP12*, 2015.
- L. Hawchar, C. P. E. Soueidy, and F. Shoefs. Principal component analysis and polynomial chaos expansion for time-variant reliability problems. *Reliability Engineering and System Safety*, 167:406–416, 2017.
- S. Ho and M. Xie. The use of arima models for reliability forecasting and analysis. *Computers and Industrial Engineering*, 35(1-2):213–216, 1998.
- Y. Hu, Z. Lu, X. Jiang, N. Wei, and C. Zhou. Time-dependent structural system reliability analysis model and its efficiency solution. *Reliability Engineering and System Safety*, 216:108029, 2021. doi: 10.1016/j.ress.2021.108029.
- Z. Hu and X. Du. Time-dependent reliability analysis with joint upcrossing rates. *Struct Multidisc Optim*, 48:893–907, 2013a.
- Z. Hu and X. Du. A sampling approach to extreme value distribution for time-dependent reliability analysis. *Journal of Mechanical design*, 135:071003, 2013b.
- Z. Hu and X. Du. First order reliability method for time-variant problems using series expansions. *Structure Multidisciplinary Optimization*, 51:1–21, 2014.
- Z. Hu and S. Mahadevan. Time-dependent reliability analysis using a vine-arma load model. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering*, 3(1):011007, 2017. doi: 10.1115/1.4034805.
- C. Jiang, X. P. Huang, X. Han, and D. Q. Zhang. A time-variant reliability analysis method based on stochastic process discretization. *Journal of Mechanical Design*, 136:091009, 2014.
- C. Jiang, X. P. Huang, X. P. Wei, and N. Y. Liu. A time-variant reliability analysis method for structural systems based on stochastic process discretization. *International Journal of Mechanics and Materials in Design*, 13:173–193, 2017.
- C. Jiang, X. P. Wei, B. Wu, and Z. L. Huang. An improved trpd method for time-variant reliability analysis. *Structural and Multidisciplinary Optimization*, 58:1935–1946, 2018.

- H. M. Kroetz, E. M. Medeiros, and A. J. Torii. On the applicability of time-series models for structural reliability analysis. In *Proceedings of XLI CILAMCE*, page 7, 2020.
- C. C. Li and A. Kiureghian. Optimal discretization of random fields. *Journal of Engineering Mechanics*, 1136(119):6, 1993.
- J. Li and Z. P. Mourelatos. Time-dependent reliability estimation for dynamic problems using a niching genetic algorithm. *J. Mech. Des.*, 131(7):071009, 2009.
- M. Loève. *Probability Theory II*. Springer-Verlag, New York, 4th edition, 1977.
- H. O. Madsen, S. Krenk, and N. C. Lind. *Methods of structural safety*. Prentice Hall, Englewood Cliffs, 1986.
- R. E. Melchers and A. T. Beck. *Structural Reliability Analysis and Prediction*, volume 1. John Wiley and Sons, 2018.
- M. Mignolet and P. Spanos. Arma monte carlo simulation in probabilistic structural analysis. *The Shock and Vibration Digest*, 21(11):3–14, 1989.
- M. R. Moarefzadeh and B. Sudret. Implementation of directional simulation to estimate outcrossing rates in time-variant reliability analysis of structures. *Qual Reliab Engng Int*, 34:1–10, 2018.
- J. Nie and B. R. Ellingwood. Finite element-based structural reliability assessment using efficient directional simulation. *Journal of Engineering Mechanics*, 131(3):259–267, 2005.
- P. Omenzetter and J. M. W. Brownjohn. Application of time series analysis for bridge monitoring. *Smart Materials and Structures*, 15(1):129–138, jan 2006. doi: 10.1088/0964-1726/15/1/041.
- J. Owen, B. Eccles, B. Choo, and M. Woodings. The application of auto-regressive time series modelling for the time-frequency analysis of civil engineering structures. *Engineering Structures*, 23(5):521–536, 2001. doi: 10.1016/S0141-0296(00)00059-6.
- S. M. Ross. *A First Course in Probability*. Prentice Hall, Upper Saddle River, 8th edition, 2010.
- E. Samaras, M. Shinzuka, and A. Tsurui. Arma representation of random processes. *Journal of Engineering Mechanics*, 111(3):449–461, 1985. doi: 10.1061/(ASCE)0733-9399(1985)111:3(449).
- A. N. Shriryaev. *Probability*. Springer-Verlag, New York, 2nd edition, 1995.
- A. Singh, Z. P. Mourelatos, and J. Li. Design for lifecycle cost using time-dependent reliability. *ASME Journal of Mechanical Design*, 132(9):091008, 2010.
- A. Singh, Z. P. Mourelatos, and E. Nikolaidis. An importance sampling approach for time-dependent reliability. In *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, 2011a.
- A. Singh, Z. P. Mourelatos, and E. Nikolaidis. Time-dependent reliability of random dynamic systems using time-series modeling and importance sampling. *SAE international Journal of Materials and Manufactur*, 4(1):929–946, 2011b.
- H. Sohn and C. R. Farrar. Damage diagnosis using time series analysis of vibration signals. *Smart Materials and Structures*, 10(3):446–451, jun 2001. doi: 10.1088/0964-1726/10/3/304.

- H. Sohn, M. L. Fugate, and C. R. Farrar. Damage diagnosis using statistical process control. In *Proc., 7th Int. Conf. on Recent Advances in Structural Dynamics, Southampton, UK*. Citeseer, 2000.
- B. Sudret. Analytical derivation of the outcrossing rate in time-variant reliability problems. *Structure and Infrastructure Engineering: Maintenance, management, Life-Cycle Design and Performance*, 4(5):353–362, 2008.
- M. Tismenetsky. Some properties of solutions of yule-walker type equations. *Linear Algebra and its Applications*, 173:1–17, 1992. doi: 10.1016/0024-3795(92)90419-B.
- L. Walls and A. Bendell. Time series methods in reliability. *Reliability Engineering*, 18(1):239–265, 1987.
- C. Wang, M. Beer, and B. M. Ayyub. Time-dependent reliability of aging structures: Overview of assessment methods. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, 7(4):03121003, 2021. doi: 10.1061/AJRUA6.0001176.
- H. Wu, Z. Hu, and X. Du. Time-Dependent System Reliability Analysis With Second-Order Reliability Method. *Journal of Mechanical Design*, 143(3), 11 2020. doi: 10.1115/1.4048732.
- R. Yao and S. N. Pakzad. Damage and noise sensitivity evaluation of autoregressive features extracted from structure vibration. *Smart Materials and Structures*, 23(2):025007, dec 2013. doi: 10.1088/0964-1726/23/2/025007.
- S. Yu, Y. Zhang, Y. Li, and Z. Wang. Time-variant reliability analysis via approximation of the first-crossing pdf. *Structural and Multidisciplinary optimization*, 62:2653–2667, 2020.
- J. Zhang and B. Ellingwood. Orthogonal series expansions of random fields in reliability analysis. *Journal of Engineering Mechanics*, 120(12):2660–2677, 1994. doi: 10.1061/(ASCE)0733-9399(1994)120:12(2660).
- X. Zhang, Y. Ju, and F. Wang. Statistical Analysis of Wind-Induced Dynamic Response of Power Towers and Four-Circuit Transmission Tower-Line System. *Shock and Vibration*, 2018, 3 2018. doi: 10.1155/2018/5064930.
- Y. Zhang, C. Gong, and C. Li. Efficient time-variant reliability analysis through approximating the most probable point trajectory. *Structural and Multidisciplinary Optimization*, 63(1):289–309, 2021.
- H. Zheng and A. Mita. Damage indicator defined as the distance between arma models for structural health monitoring. *Structural Control and Health Monitoring*, 15(7):992–1005, 2008. doi: 10.1002/stc.235.
- W. Zhou, C. Gong, and H. P. Hong. New perspective on application of first-order reliability method for estimating system reliability. *Journal of Engineering Mechanics*, 173(9):04017074, 2017.

Chapter 15

Bayesian Optimal Experimental Design

Chapter details

Chapter DOI:

<https://doi.org/10.4322/978-65-86503-88-3.c15>

Chapter suggested citation / reference style:

Carlton, Andre G., et al. (2022). “Bayesian Optimal Experimental Design”. In Jorge, Ariosto B., et al. (Eds.) *Uncertainty Modeling: Fundamental Concepts and Models*, Vol. III, UnB, Brasilia, DF, Brazil, pp. 479–501. Book series in Discrete Models, Inverse Methods, & Uncertainty Modeling in Structural Integrity.

P.S.: DOI may be included at the end of citation, for completeness.

Book details

Book: Uncertainty Modeling: Fundamental Concepts and Models

Edited by: Jorge, Ariosto B., Anflor, Carla T. M., Gomes, Guilherme F., & Carneiro, Sergio H. S.

Volume III of Book Series in:

Discrete Models, Inverse Methods, & Uncertainty Modeling in Structural Integrity

Published by: UnB City: Brasilia, DF, Brazil Year: 2022

DOI: <https://doi.org/10.4322/978-65-86503-88-3>

Bayesian Optimal Experimental Design

Andre G. Carlon¹, Rafael H. Lopez^{2*}, Leandro F. F. Miguel²,
André J. Torii³

¹Computer, Electrical and Mathematical Sciences & Engineering Division, King Abdullah University of Science and Technology, Saudi Arabia. E-mail: agcarlon@gmail.com

²Center for Optimization and Reliability in Engineering (CORE), Department of Civil Engineering, Federal University of Santa Catarina, Federal University of Santa Catarina, Brazil. E-mail: rafael.holdorf@ufsc.br, leandro.miguel@ufsc.br

³ Latin American Institute for Technology, Infrastructure and Territory (ILATIT), Federal University for Latin American Integration (UNILA), Avenida Tancredo Neves, 6731, Foz do Iguaçu/PR, Brazil, 85867-970. E-mail: andre.torii@unila.edu.br

*Corresponding author

Abstract

Experiments are of fundamental importance both in academic sciences and the industry. In many cases, performing experiments is an expensive task requiring resources and time. Therefore, it is essential to design experiments as best as possible to obtain the most information. However, in many cases, the optimal design of an experiment is not obvious, being even counter-intuitive in some situations. The goal of optimal experimental design is to formalize the statistical information obtained by experiments as an objective function to be maximized using optimization methods. Here, we focus on the Bayesian setting, using a discrepancy between the information before and after an experiment as its quality. Specifically, we use the Shannon's Expected Information Gain (SEIG), the mutual information between the experimental observations and the parameters of interest, to measure the quality of experiments.

Keywords: Optimal Experimental Design; Expected Information Gain; Uncertainty Quantification; Bayesian Inference; Monte Carlo Sampling

1 Introduction

Many fields of science heavily rely on information obtained by experiments. Statistics about quantities of interest are inferred from data resulting from experiment observations. Hence, it is important that experiments provide informative data. For example, in structural engineering, it is useful to have statistical information about the material properties, such as the Young modulus, Poisson modulus, yield stress, so that the engineer can take informative decisions. However, to obtain statistically relevant data, experiments must have large enough samples. For example, if a three-point flexural test is performed on concrete beams to evaluate the fracture toughness of concrete, a significant number of beams must be built beforehand and let to cure for 28 days. Even

if a hundred beams are built, cured, and tested, the standard error in the fracture toughness estimate is still the standard deviation of one of the samples divided by ten. To reduce the standard error of the mean by one digit requires increasing the sample size by a hundred times. A three-point bending experiment for fracture toughness determination and its design parameters are presented in Figure 1. Properly defining design parameters can reduce dispersion of the quantity of interest

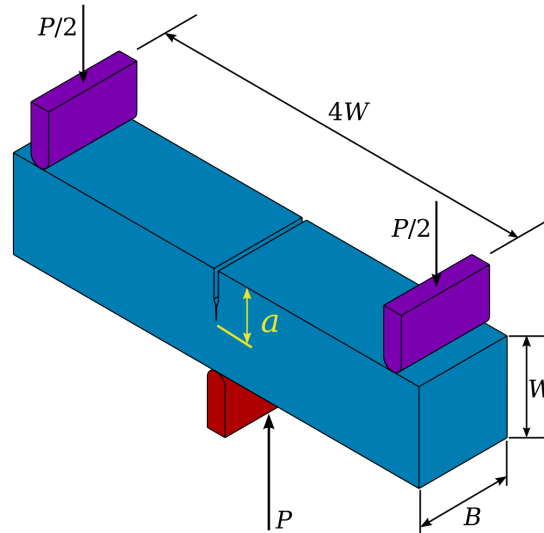


Figure 1: Example of fracture toughness testing. Source: Wikimedia commons Commons [2011]

estimation or reduce the cost required to achieve the same precision.

Another example of practical interest is the one of verifying the quality of composite laminates. Some composite materials can have improved mechanical properties in some directions by alternating plies of orthotropic materials with some specific angles between them. One way of testing these laminates is by electrical impedance tomography (EIT). However, the information obtained depends heavily on the currents imposed on the electrodes, as can be seen in the results obtained by Beck et al. [2018]. The set-up of an EIT experiment is presented in Figure 2, where the electrodes are illustrated in black, the first ply in blue and the second ply in red. We optimize the currents in an EIT experiment in Section ??.

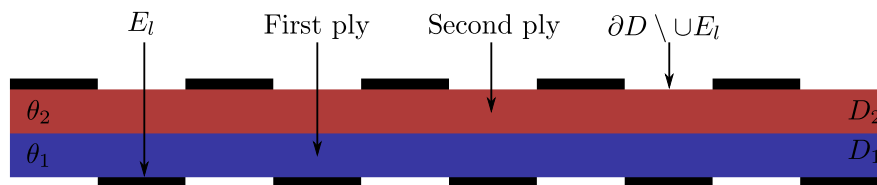


Figure 2: Electrical impedance tomography example. Source: adapted from Beck et al. [2018]

Given the difficulties and the costs involved in the experimental process, it might be interesting to optimally tune experiments parameters to maximize their efficiency, which is the main concern of optimal experimental design (OED). Thus, in OED, we seek the optimum design parameters that maximize some measure of efficiency of an experiment.

The classical optimality criteria for experiments are known as the alphabetic optimality criteria. However, they require the model to be linear with respect to the random parameters. The al-

phabetic optimality criteria use the Fisher information matrix, a measure of information inversely proportional to the covariance matrix, to qualify an experiment. The better known of these criteria, D-optimality, consists in minimizing the determinant of the Fisher information matrix, as presented in Chaloner and Verdinelli [1995].

In this chapter, we aim at presenting an OED framework that can optimize experiments with nonlinear black-box forward models. Thus, we opt to measure the performance of an experiment by its Shannon's expected information gain (EIG) [Shannon, 1948]. The EIG is related to the relative entropy of information and, in the OED context, is a measure on how much information an experiment provides [Huan and Marzouk, 2013]. Lindley [1956] was the first to use EIG as an utility function for OED. However, for a general experiment with nonlinear model, estimating EIG is cumbersome.

Ryan [2003] develop an EIG estimator based on Monte Carlo integration (MCI), however, his estimator requires the evaluation of two nested integrals. We refer to Ryan's estimator as the double-loop Monte Carlo (DLMC). The DLMC estimator has the disadvantages of being expensive to evaluate and numerically unstable. The evaluation of the two nested MCI requires a large number of experiment simulations. Indeed, if the two outer and inner MCI have sample size of, respectively, N and M , the DLMC estimator requires $N(M + 1)$ forward model evaluations to approximate the EIG. Moreover, for some cases, M needs to be large to avoid numerical underflow [Beck et al., 2018]. Long et al. [2013] propose a method to estimate the EIG using a Laplace approximation, furnishing the Monte Carlo with Laplace approximation (MCLA) estimator. The MCLA estimator does not require the evaluation of one of the two nested integrals, thus, being less expensive. The disadvantage of MCLA is that the Laplace approximation introduces a bias that might not be acceptable depending on the situation. Beck et al. [2018] propose an importance sampling for DLMC that uses Laplace approximation to draw more informative samples, reducing the cost of DLMC without adding the bias of MCLA. The DLMC with the importance sampling is called double-loop Monte Carlo with importance sampling (DLMCIS).

The main goal of this chapter is to present the DLMC, MCLA and DLMCIS estimators for the SEIG in a didactic manner. To achieve this goal, we first present the basis of Bayesian inference and how to measure information using the SEIG. Then, the procedure for the numerical evaluation of the DLMC, MCLA and DLMCIS estimators is presented, together with their the pseudocodes.

2 Optimal Experimental Design (OED)

The goal of OED is to find the best experimental setup [Chaloner and Verdinelli, 1995]. From the perspective of optimization, it is important to define a criterion to measure the efficiency of experiments: the objective function to be maximized. Here, we use the Shannon's SEIG [Shannon, 1948] of an experiment to evaluate its performance. Thus, on the next section, we introduce SEIG and other concepts related to it.

2.1 Experiment model

We model the evaluation of N_e repetitive experiments as

$$\mathbf{y}_i(\boldsymbol{\xi}, \boldsymbol{\theta}_t, \boldsymbol{\epsilon}_i) = \mathbf{g}(\boldsymbol{\xi}, \boldsymbol{\theta}_t) + \boldsymbol{\epsilon}_i, \quad i = 1, \dots, N_e, \quad (1)$$

where $\mathbf{y}_i \in \mathbb{R}^q$ is the vector of experiment observations, $\boldsymbol{\theta}_t \in \mathbb{R}^d$ is the vector of parameters of interest to be recovered, $\boldsymbol{\xi}$ is the vector with experiment parameters to be optimized, \mathbf{g} is the experiment model, and $\boldsymbol{\epsilon}$ is the additive noise from measurements. For N_e experiments performed with the same setup $\boldsymbol{\xi}$, the set of observed data is gathered in a set $\mathbf{Y} = \{\mathbf{y}_i\}_{i=1}^{N_e}$. Since we cannot observe $\boldsymbol{\theta}_t$ directly, we use the random variable $\boldsymbol{\theta}: \Theta \rightarrow \mathbb{R}^m$ with prior distribution $\pi(\boldsymbol{\theta})$ instead

of θ_t . Thus, through observations \mathbf{Y} , we calculate statistics about θ . Here, our goal is to find the optimal experimental design $\xi \in \Xi \subset \mathbb{R}^n$ that provides more information about θ , where Ξ is the space of experiment designs. Moreover, we consider $\epsilon \sim \mathcal{N}(\mathbf{0}, \Sigma_\epsilon)$ to be an additive noise, i.e. independent of g , ξ and θ , for some positive-definite and symmetric covariance matrix Σ_ϵ . This model for experiments is the same used in [Long et al., 2013, Beck et al., 2018].

For example, consider the case of a single three-point test where the stiffness modulus (E) of some material is estimated, as illustrated in Figure 3. In this case, the parameter of interest (θ) is

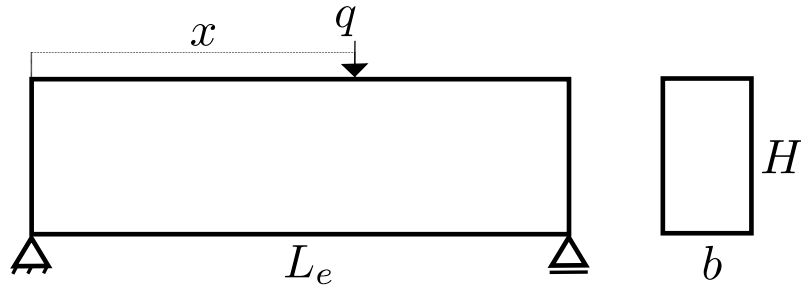


Figure 3: Example of the setup for a three-point flexural test.

E , the measurement y is the observed deflection in mm, and ϵ is the error in the observation of y . The design parameter of the experiment to be optimized is $\xi = x$, the position where the point load q is applied, and where the deflection is measured. The model relating E and x to the observation y is $g(x, E) = \frac{qx}{6L_eEI_e}(2L_e - x)$, where I_e is the moment of inertia of the cross-section of the beam. This simple problem is used to illustrate the concepts introduced in the present chapter; the proof that the optimum is at $x^* = L_e/2$ is trivial.

2.2 Bayesian Inference

To evaluate the quality of experiments, we use a Bayesian framework of analysis. Thus, in this section, we introduce essential concepts of Bayesian inference. The main idea behind Bayesian inference is to use new data to update some previously known probability distribution of a parameter. This update is done using Bayes’ theorem:

$$\pi(\theta|\mathbf{Y}, \xi) = \frac{p(\mathbf{Y}|\theta, \xi)\pi(\theta)}{p(\mathbf{Y}|\xi)}, \tag{2}$$

where $\pi(\theta|\mathbf{Y}, \xi)$ is the posterior distribution, the updated probability density function (pdf) of some random variable θ given observations \mathbf{Y} ; $\pi(\theta)$ is the prior distribution, the pdf of θ before the experiment; $p(\mathbf{Y}|\theta, \xi)$ is the likelihood, the probability of observing \mathbf{Y} given the previously known prior $\pi(\theta)$; and $p(\mathbf{Y}|\xi)$ is the evidence, the probability of \mathbf{Y} being observed.

On the context of experiments, Bayesian inference is used to estimate a posterior pdf of a parameter of interest given a prior pdf and some observations \mathbf{Y} provided by the experiment. For example, consider the previously mentioned three-point flexural test; in Figure 4 we present the pdf for the stiffness both before and after the flexural experiment, i.e., the prior and posterior pdfs of (2). It can be seen that the dispersion is reduced after the experiment, meaning that the experiment provided useful information about the parameters of interest.

In this example, we modeled the prior knowledge about the parameter of interest as being Gaussian-distributed; however, other distributions can be used as well. Using the Bayes’ theorem in (2), the posterior can be obtained by multiplying the prior by the division between the likelihood and the evidence. For the experiment model with additive noise in (1), the likelihood of observing

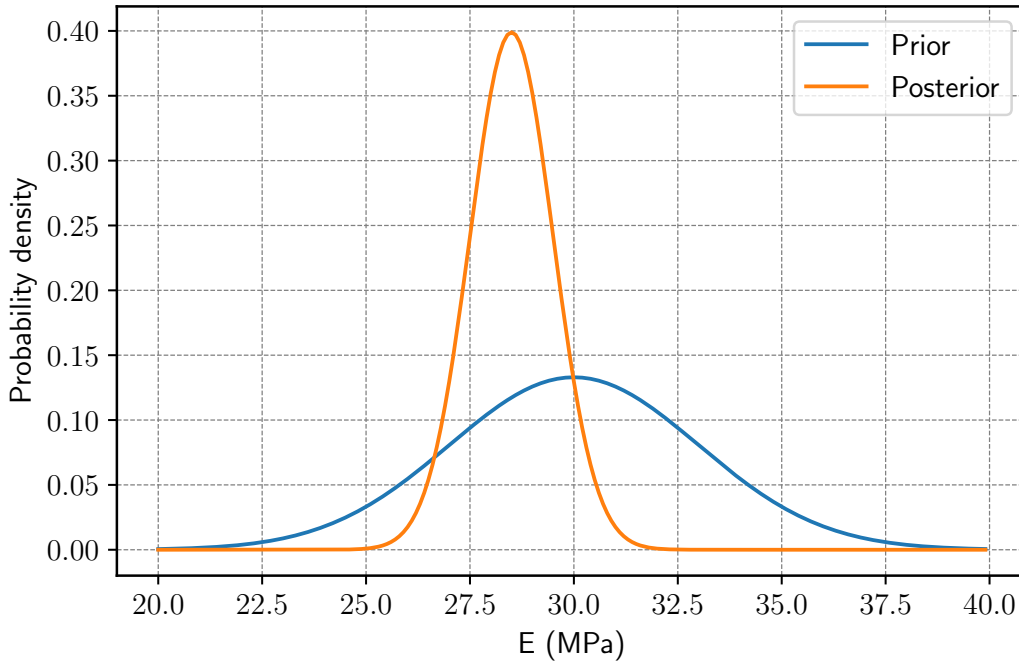


Figure 4: Prior and posterior pdfs for the three-point flexural experiment.

\mathbf{Y} given $\boldsymbol{\theta}$ and $\boldsymbol{\xi}$ is a multivariate Gaussian with mean $\mathbf{g}(\boldsymbol{\xi}, \boldsymbol{\theta})$ and covariance matrix $\boldsymbol{\Sigma}_\epsilon$,

$$p(\mathbf{Y}|\boldsymbol{\theta}, \boldsymbol{\xi}) = \det(2\pi\boldsymbol{\Sigma}_\epsilon)^{-\frac{N_e}{2}} \exp\left(-\frac{1}{2} \sum_{i=1}^{N_e} \|\mathbf{y}_i - \mathbf{g}(\boldsymbol{\xi}, \boldsymbol{\theta})\|_{\boldsymbol{\Sigma}_\epsilon^{-1}}^2\right), \quad (3)$$

where $\boldsymbol{\theta}$ is not necessarily the same used to calculate \mathbf{Y} . Using (3), the likelihood in (2) can be evaluated, yet, the evidence $p(\mathbf{Y}|\boldsymbol{\xi})$ is not known. To estimate the evidence, we follow the same procedure as Ryan [2003]. We introduce a new variable $\boldsymbol{\theta}^*$ that is independent and identically distributed with respect to $\boldsymbol{\theta}$ and marginalize its likelihood with respect to $\boldsymbol{\theta}^*$ as

$$p(\mathbf{Y}|\boldsymbol{\xi}) = \int_{\Theta} p(\mathbf{Y}|\boldsymbol{\theta}^*, \boldsymbol{\xi})\pi(\boldsymbol{\theta}^*)d\boldsymbol{\theta}^*. \quad (4)$$

The likelihood in (4) can be calculated from (3) as

$$p(\mathbf{Y}|\boldsymbol{\theta}^*, \boldsymbol{\xi}) = \det(2\pi\boldsymbol{\Sigma}_\epsilon)^{-\frac{N_e}{2}} \exp\left(-\frac{1}{2} \sum_{i=1}^{N_e} \|\mathbf{y}_i - \mathbf{g}(\boldsymbol{\xi}, \boldsymbol{\theta}^*)\|_{\boldsymbol{\Sigma}_\epsilon^{-1}}^2\right). \quad (5)$$

Thus, substituting in the Bayes' equation in (2) the likelihood and the evidence respectively presented in (3) and (4), one can obtain the posterior distribution of the experiment, i.e., the probability distribution for the parameter of interest after the experiment.

To estimate how much information an experiment provides, we use the Kullback–Leibler divergence (D_{KL}) between the prior and posterior pdfs.

2.3 Kullback–Leibler Divergence

According to Cover and Thomas [2012], entropy is a measure of uncertainty of a random variable. Cover and Thomas define the differential entropy of a random variable $\boldsymbol{\theta}$ with pdf f_θ as

$$-\int_{\Theta} f_\theta(\boldsymbol{\theta}) \log f_\theta(\boldsymbol{\theta})d\boldsymbol{\theta}. \quad (6)$$

The larger the entropy of $\pi(\theta)$ is, the larger its uncertainty with respect to θ is. The D_{KL} is the entropy of a probability measure with respect to another [Kullback and Leibler, 1951], and, for two probability measures f_θ and g_θ on θ with the same support Θ , the D_{KL} is defined as

$$\begin{aligned} D_{KL}(f_\theta(\theta)||g_\theta(\theta)) &= - \int_{\Theta} f_\theta(\theta) \log g_\theta(\theta) d\theta + \int_{\Theta} f_\theta(\theta) \log f_\theta(\theta) d\theta \\ &= - \int_{\Theta} f_\theta(\theta) \log \left(\frac{g_\theta(\theta)}{f_\theta(\theta)} \right) d\theta \\ &= \int_{\Theta} \log \left(\frac{f_\theta(\theta)}{g_\theta(\theta)} \right) f_\theta(\theta) d\theta. \end{aligned} \tag{7}$$

The larger the D_{KL} of $f_\theta(\theta)$ with respect to $g_\theta(\theta)$ is, the more different is $f_\theta(\theta)$ with respect to $g_\theta(\theta)$; the D_{KL} between two distributions is a measure of the discrepancy between them. As a way of measuring the efficiency of an experiment, we use the D_{KL} of the posterior with respect to the prior:

$$D_{KL}(\pi(\theta|Y, \xi)||\pi(\theta)) = \int_{\Theta} \log \left(\frac{\pi(\theta|Y, \xi)}{\pi(\theta)} \right) \pi(\theta|Y, \xi) d\theta. \tag{8}$$

For the sake of simplicity, the Kullback–Leibler divergence between the prior and posterior pdfs is referred simply as Kullback–Leibler divergence and denoted as D_{KL} . In Figure 5, at the left, we present the prior and posterior pdf of the three-point flexural test. The regions where the probability density of the posterior distribution is greater than the one of the prior are shaded in green, whereas the regions where the posterior pdf is less than the prior pdf are shaded in red. In Figure 5, at the right, we present the integrand in (8) over the domain of θ for the three-point flexural test. The regions where the integrand is positive are shaded in green, and the regions where the integrand is negative are shaded in red. The more informative an experiment is, the

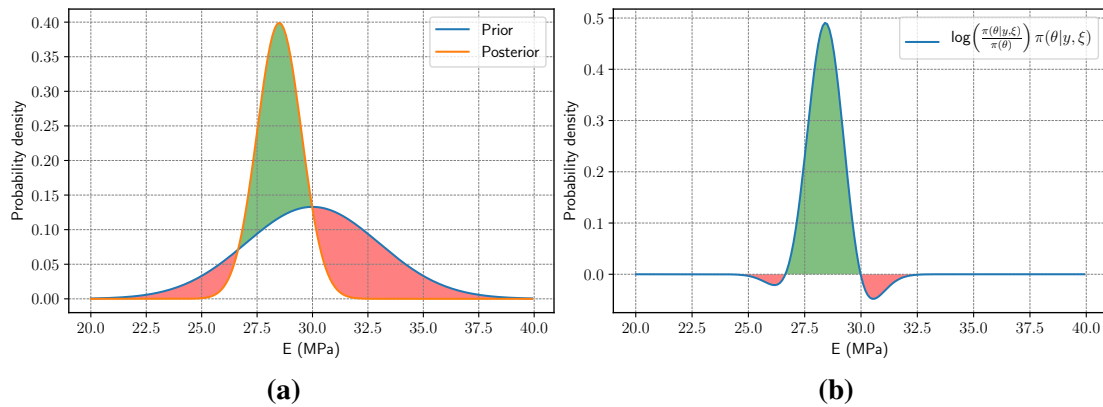


Figure 5: For the three-point flexural test: prior and posterior distributions (a), and the integrand of the D_{KL} (b).

larger is the integral of the function in Figure 5 (b). Thus, we aim to find the ξ^* that provides the more informative observations Y^* ; the Y^* that maximize difference between the green area and the red area in Figure 5 (b).

In Figure 6, we present a comparison between the optimized and non-optimized cases for the three-point flexural test. The non-optimized case is evaluated with $x = L_e/4$, whereas, in the optimized case, the load and the measurement are in the middle of the beam, i.e., $x = x^* = L_e/2$. In the left plot of Figure 6, it can be observed that the posterior pdf for the optimized configuration is more concentrated than the posterior pdf before before optimization. In the right plot of Figure 6, the integrand of the D_{KL} is presented for both the optimized and non-optimized configurations. It

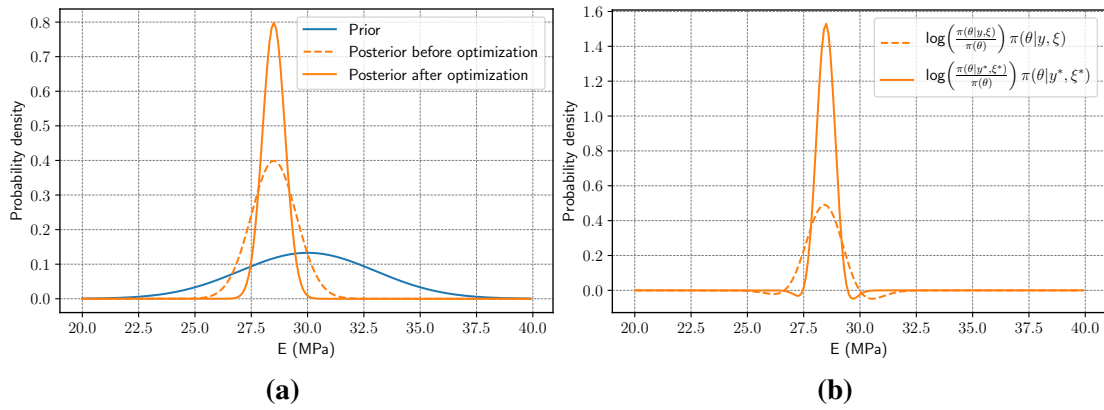


Figure 6: The pdfs for prior, posterior, and optimized posterior (a) and the D_{KL} integrand for the non-optimized and for the optimized cases (b).

can be observed that the D_{KL} for the optimized case is greater than the D_{KL} for the non-optimized configuration, as can be inferred from the areas under the integrands.

When modeling a real experiment, one might need to consider the noise inherent to experiment observations. To estimate the information gain considering the noise in observations, we use the SEIG [Shannon, 1948].

2.4 Shannon’s Expected Information Gain

The D_{KL} does not consider the noise in observations due to measurement uncertainties. To estimate the information gain on this context, we need to marginalize the D_{KL} with respect to \mathbf{Y} , thus, obtaining the Shannon’s expected information gain [Shannon, 1948] as

$$I = \int_{\mathcal{Y}} \int_{\Theta} \log \left(\frac{\pi(\boldsymbol{\theta}|\mathbf{Y}, \boldsymbol{\xi})}{\pi(\boldsymbol{\theta})} \right) \pi(\boldsymbol{\theta}|\mathbf{Y}, \boldsymbol{\xi}) d\boldsymbol{\theta} p(\mathbf{Y}|\boldsymbol{\xi}) d\mathbf{Y}. \tag{9}$$

Lindley [1956] is the first to use SEIG as an utility function for optimal experimental design. An example of the integrand in (9) with respect to both the parameter of interest and the observations for the three-point flexural test is presented in Figure 7. Estimating I requires the solution of a double integral over the sample space of the parameters of interest (Θ), and the observations (\mathcal{Y}).

Using the Bayes’ equation presented in (2), we can rewrite (9) as

$$I = \int_{\Theta} \int_{\mathcal{Y}} \log \left(\frac{p(\mathbf{Y}|\boldsymbol{\theta}, \boldsymbol{\xi})}{p(\mathbf{Y}|\boldsymbol{\xi})} \right) p(\mathbf{Y}|\boldsymbol{\theta}, \boldsymbol{\xi}) d\mathbf{Y} \pi(\boldsymbol{\theta}) d\boldsymbol{\theta}. \tag{10}$$

Presenting SEIG as in (10) has the advantage that, for our experiment model, the likelihood can be calculated directly using (3). Moreover, using Eq.4 to estimate the evidence results in

$$I = \int_{\Theta} \int_{\mathcal{Y}} \log \left(\frac{p(\mathbf{Y}|\boldsymbol{\theta}, \boldsymbol{\xi})}{\int_{\Theta} p(\mathbf{Y}|\boldsymbol{\theta}^*, \boldsymbol{\xi}) \pi(\boldsymbol{\theta}^*) d\boldsymbol{\theta}^*} \right) p(\mathbf{Y}|\boldsymbol{\theta}, \boldsymbol{\xi}) d\mathbf{Y} \pi(\boldsymbol{\theta}) d\boldsymbol{\theta}. \tag{11}$$

Here, we use (11) to calculate the SEIG.

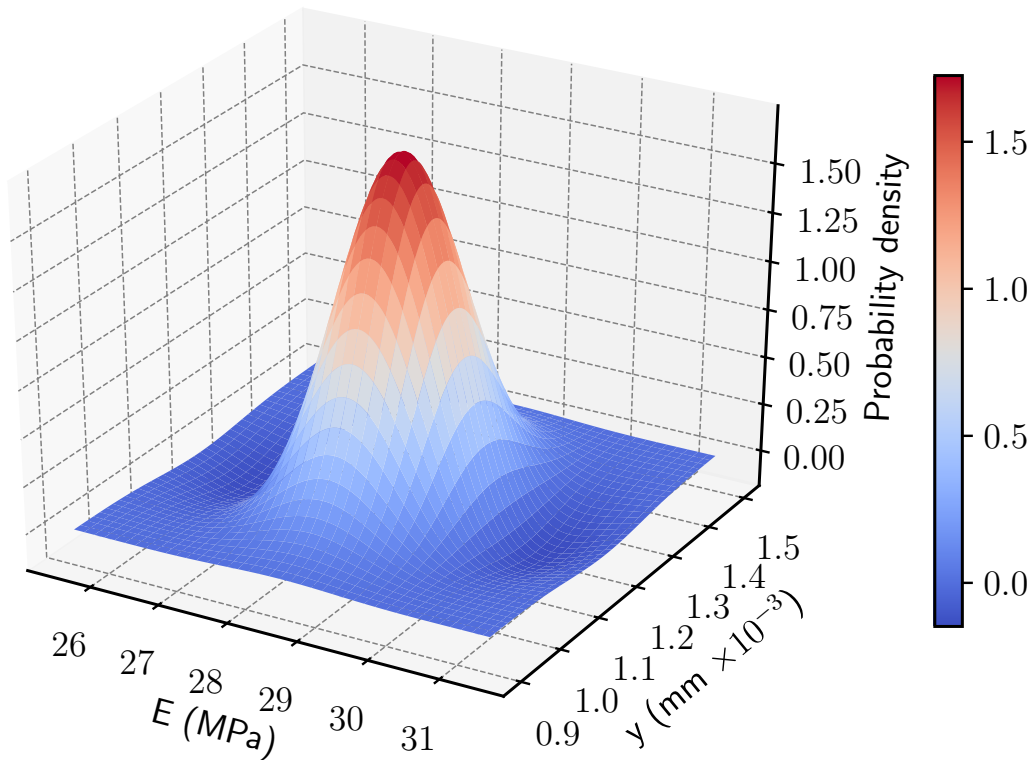


Figure 7: The SEIG integrand.

3 Shannon’s Expected Information Gain Estimators

Evaluating I in (11) requires the solution of the double integral over θ and Y , moreover, it requires the solution of the integral used to marginalize the evidence. In most cases, these integrals do not have closed form, thus, numerical methods are needed to approximate them. When the number of random parameters is small, quadrature methods can be used to approximate the integrals in (11), however, as pointed out by Robert and Casella [2013], these methods suffer from the curse of dimensionality. Hamada et al. [2001] note that using deterministic integration methods to approximate (11) becomes unfeasible if the dimensionality of θ exceeds three.

3.1 Error analysis

We focus on the case where the integrals in (11) do not have closed form solution and Monte Carlo integration (MCI) is needed to approximate them. In this section, we introduce Monte Carlo-based estimators for SEIG that can deal with high dimensional parameter spaces. We present the error analysis and complexity of the SEIG estimators with respect to the evaluation of g by FEM with a mesh discretization parameter h . As $h \rightarrow 0$, the discretization bias asymptotically converges as

$$\mathbb{E} [\|g(\xi, \theta) - g_h(\xi, \theta)\|_2] = \mathcal{O}(h^\eta), \tag{12}$$

where g_h is the FEM approximation of g using h , and $\eta > 0$ is the rate of convergence of the discretization error. For the computational effort analysis, we assume that the cost of evaluating g_h is $\mathcal{O}(h^{-\varrho})$, for the constant $\varrho > 0$. Both η and ϱ are constants that depend only on the numerical approach used in approximating g by g_h .

In the present case, the error of estimators can be decomposed in two terms: their bias and variances. The variance of the estimator is a measure of the dispersion from different and independent

estimations of the same quantity, e.g., for an estimator \mathcal{I} , its variance is $\mathbb{V}[\mathcal{I}] \stackrel{\text{def}}{=} \mathbb{E} [(\mathcal{I} - \mathbb{E}[\mathcal{I}])^2]$. The bias of an estimator is the difference between its expected value and the true value that is being estimated, e.g., for an estimator \mathcal{I} of a quantity I , the bias of \mathcal{I} is $|I - \mathbb{E}[\mathcal{I}]|$. Figure 8 illustrates the bias and variance of an hypothetical estimator of I , where the curve in blue is the probability density that a value of \mathcal{I} is estimated. Moreover, the distance between the expected value of \mathcal{I} and I , i.e., the bias, is presented, as is the square root of the variance, σ , the standard error of the estimator. The bias of the estimator includes the bias from numerical approximation of the model.

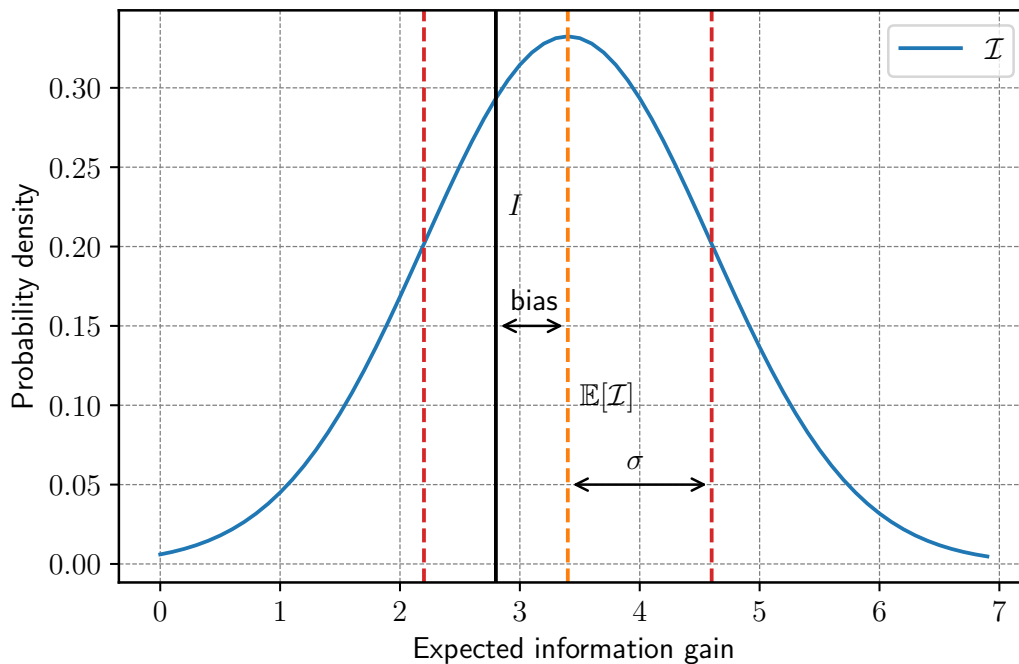


Figure 8: Bias and variance of an estimator

For further information about the bias and variance of each SEIG estimator, the reader is referred to Beck et al. [2018].

3.2 Monte Carlo integration

The MCI is a method for approximating integrals by sampling the integrand and averaging the samples. For example, let $\theta \in \Theta \subset \mathbb{R}^n$, and $f : \theta \rightarrow \mathbb{R}$ with θ with probability $\pi(\theta)$. In this case, the expectation of $f(\theta)$ can be approximated as

$$\mathbb{E}[f(\theta)] = \int_{\Theta} f(\theta)\pi(\theta)d\theta \approx \frac{1}{N} \sum_{i=1}^N f(\theta_i), \tag{13}$$

where θ is sampled from $\pi(\theta)$. From the strong law of large numbers, the Monte Carlo estimator converges to the real value of the integral as $N \rightarrow \infty$ [Rubinstein and Kroese, 2016]. Moreover, from the central limit theorem, as $N \rightarrow \infty$, the error in approximating an integral by MCI converges to zero with rate $1/\sqrt{N}$ [Rubinstein and Kroese, 2016].

3.3 Double-loop Monte Carlo estimator

The double integral in (10) can be approximated using MCI by sampling θ from the prior pdf and Y from the likelihood (the measures that the integrands are being integrated over in (10)). Then,

for N samples, the Monte Carlo estimator for OED is defined as

$$\mathcal{I}_{MC}(\boldsymbol{\xi}) \stackrel{\text{def}}{=} \frac{1}{N} \sum_{n=1}^N \log \left(\frac{p(\mathbf{Y}_n | \boldsymbol{\theta}_n, \boldsymbol{\xi})}{p(\mathbf{Y}_n | \boldsymbol{\xi})} \right). \tag{14}$$

Also, we use (4) and approximate the evidence integral by a MCI as

$$\int_{\Theta} p(\mathbf{Y} | \boldsymbol{\theta}^*, \boldsymbol{\xi}) \pi(\boldsymbol{\theta}^*) d\boldsymbol{\theta}^* \approx \frac{1}{M} \sum_{m=1}^M p(\mathbf{Y} | \boldsymbol{\theta}_m^*, \boldsymbol{\xi}). \tag{15}$$

Thus, the DLMC estimator for SEIG is defined as

$$\mathcal{I}_{DLMC}(\boldsymbol{\xi}) \stackrel{\text{def}}{=} \frac{1}{N} \sum_{n=1}^N \log \left(\frac{p(\mathbf{Y}_n | \boldsymbol{\theta}_n, \boldsymbol{\xi})}{\frac{1}{M} \sum_{m=1}^M p(\mathbf{Y}_n | \boldsymbol{\theta}_m^*, \boldsymbol{\xi})} \right). \tag{16}$$

The first to use DLMC in OED is Ryan [2003].

Monte Carlo estimators are generally not biased, however, the DLMC estimator has a bias resulting from the logarithm of the inner loop. However, DLMC is a consistent estimator because the bias goes to zero as the number of inner loop samples M goes to infinity. The DLMC estimator has bias and variance respectively given by

$$|I - \mathbb{E}[\mathcal{I}_{DLMC}]| \leq C_{DL,1} h^\eta + \frac{C_{DL,2}}{M} + o(h^\eta) + \mathcal{O}\left(\frac{1}{M^2}\right), \quad \text{and} \tag{17}$$

$$\mathbb{V}[\mathcal{I}_{DLMC}] = \frac{C_{DL,3}}{N} + \frac{C_{DL,4}}{NM} + \mathcal{O}\left(\frac{1}{NM^2}\right) \tag{18}$$

for the constants $C_{DL,1}$, $C_{DL,2}$, $C_{DL,3}$, and $C_{DL,4}$ (cf. [Beck et al., 2018]).

To evaluate the DLMC estimator in (16), one needs the likelihood of observing \mathbf{Y}_n given $\boldsymbol{\theta}_n$ and the likelihoods of observing \mathbf{Y}_n given each $\boldsymbol{\theta}_m^*$. Using (3) to estimate the likelihood of observing \mathbf{Y}_n given $\boldsymbol{\theta}_n$ furnishes

$$p(\mathbf{Y}_n | \boldsymbol{\theta}_n, \boldsymbol{\xi}) = \det(2\pi \boldsymbol{\Sigma}_\epsilon)^{-\frac{N_\epsilon}{2}} \exp \left(-\frac{1}{2} \sum_{i=1}^{N_\epsilon} \left\| \mathbf{y}_i^{(n)}(\boldsymbol{\xi}) - \mathbf{g}(\boldsymbol{\xi}, \boldsymbol{\theta}_n) \right\|_{\boldsymbol{\Sigma}_\epsilon^{-1}}^2 \right), \tag{19}$$

$$= \det(2\pi \boldsymbol{\Sigma}_\epsilon)^{-\frac{N_\epsilon}{2}} \exp \left(-\frac{1}{2} \sum_{i=1}^{N_\epsilon} \left\| \mathbf{g}(\boldsymbol{\xi}, \boldsymbol{\theta}_n) + \boldsymbol{\epsilon}_i - \mathbf{g}(\boldsymbol{\xi}, \boldsymbol{\theta}_n) \right\|_{\boldsymbol{\Sigma}_\epsilon^{-1}}^2 \right), \tag{20}$$

$$= \det(2\pi \boldsymbol{\Sigma}_\epsilon)^{-\frac{N_\epsilon}{2}} \exp \left(-\frac{1}{2} \sum_{i=1}^{N_\epsilon} \left\| \boldsymbol{\epsilon}_i \right\|_{\boldsymbol{\Sigma}_\epsilon^{-1}}^2 \right). \tag{21}$$

Since \mathbf{Y} and \mathbf{g} are evaluated using the same $\boldsymbol{\theta}$, the model evaluation \mathbf{g} in (20) cancels out. Thus, evaluating the likelihood of observing \mathbf{Y}_n given $\boldsymbol{\theta}_n$ does not require any model evaluation. However, for the evidence evaluation, we must evaluate the model \mathbf{g} . For example, consider the likelihood in (3), where \mathbf{Y}_n is evaluated from $\boldsymbol{\theta}_n$ and ϵ (sampled in the outer loop), and $\boldsymbol{\theta}_m^*$ is sampled in the inner loop (independently from $\boldsymbol{\theta}_n$),

$$p(\mathbf{Y}_n | \boldsymbol{\theta}_m^*, \boldsymbol{\xi}) = \det(2\pi \boldsymbol{\Sigma}_\epsilon)^{-\frac{N_\epsilon}{2}} \exp \left(-\frac{1}{2} \sum_{i=1}^{N_\epsilon} \left\| \mathbf{y}_i^{(n)}(\boldsymbol{\xi}) - \mathbf{g}(\boldsymbol{\xi}, \boldsymbol{\theta}_m^*) \right\|_{\boldsymbol{\Sigma}_\epsilon^{-1}}^2 \right). \tag{22}$$

Each observation of $p(\mathbf{Y}_n | \boldsymbol{\theta}_m^*, \boldsymbol{\xi})$ requires an evaluation of $\mathbf{g}(\boldsymbol{\xi}, \boldsymbol{\theta}_m^*)$.

Algorithm 1 Pseudocode for the DLMC estimator for SEIG.

```

1: function DLMC( $\xi, N, M$ )
2:   for  $n = 1, 2, \dots, N$  do ▷ Outer loop
3:     Sample  $\theta_n$  from  $\pi(\theta)$ 
4:     Evaluate  $g(\xi, \theta_n)$ 
5:     for  $i = 1, 2, \dots, N_e$  do
6:       Sample  $\epsilon_i$  from  $\mathcal{N}(0, \Sigma_\epsilon)$ 
7:        $y_i \leftarrow g(\xi, \theta_n) + \epsilon_i$ 
8:     end for
9:      $\mathbf{Y}_n \leftarrow \{y_i\}_{i=1}^{N_e}$ 
10:     $p(\mathbf{Y}_n | \theta_n, \xi) \leftarrow \det(2\pi \Sigma_\epsilon)^{-\frac{N_e}{2}} \exp\left(-\frac{1}{2} \sum_{i=1}^{N_e} \|\epsilon_i\|_{\Sigma_\epsilon^{-1}}^2\right)$ 
11:    for  $m = 1, 2, \dots, M$  do ▷ Inner loop
12:      Sample  $\theta_m^*$  from  $\pi(\theta)$ 
13:      Evaluate  $g(\xi, \theta_m^*)$ 
14:       $p(\mathbf{Y}_n | \theta_m^*, \xi) \leftarrow \det(2\pi \Sigma_\epsilon)^{-\frac{N_e}{2}} \exp\left(-\frac{1}{2} \sum_{i=1}^{N_e} \left\|y_i^{(n)}(\xi) - g(\xi, \theta_m^*)\right\|_{\Sigma_\epsilon^{-1}}^2\right)$ 
15:    end for
16:     $p(\mathbf{Y}_n | \xi) \leftarrow \frac{1}{M} \sum_{m=1}^M p(\mathbf{Y}_n | \theta_m^*, \xi)$ 
17:  end for
18:   $\mathcal{I}_{DLMC}(\xi) \leftarrow \frac{1}{N} \sum_{n=1}^N \log\left(\frac{p(\mathbf{Y}_n | \theta_n, \xi)}{p(\mathbf{Y}_n | \xi)}\right)$ 
19:  return  $\mathcal{I}_{DLMC}(\xi)$ 
20: end function

```

In Algorithm 1, we present the pseudocode for DLMC, where the inputs are the experiment setup, ξ , the sample-size for the outer loop, N , and the sample-size for the inner loop, M . Problem parameters, e.g., Σ_ϵ , $\pi(\theta)$, g , N_e , are considered to be known. The DLMC returns the estimation of I , \mathcal{I}_{DLMC} .

From Algorithm 1, it can be seen that the cost of evaluating the DLMC estimator is $N(M+1)$ forward model evaluations. Considering that each model evaluation requires the solution of a PDE using FEM with a mesh of size h , the computational cost of evaluating the DLMC estimator is of order $N(M+1)h^{-\rho}$.

The DLMC estimator can suffer from numerical instabilities, namely, *numerical underflow* [Beck et al., 2018]. The marginalization of the likelihood of observing \mathbf{Y} (evaluated using θ^* sampled in the outer loop) given θ sampled in the inner loop can result in numerical underflow. If the likelihood is zero for all the M inner samples, the evidence also becomes zero. If $g(\xi, \theta_m^*)$ is too distant to each $y^{(n)}$, or if Σ_ϵ has small eigenvalues, the likelihood can get smaller than floating-point precision. Thus, to avoid numerical underflow, M needs to be large enough as to guarantee that at least one of the M likelihoods is not numerically evaluated to zero. If the evidence is evaluated to zero, then (14) cannot be evaluated. Figure 9 illustrates the numerical underflow for the flexural test for five inner loop samples; for each loop the model is evaluated and the likelihood of y being observed is drawn in red. It can be observed that, for all g evaluated at each inner loop, the likelihood of observing the $y^{(n)}$ evaluated at the outer loop is near to zero. In the next sections we present two estimators of SEIG that are able to overcome the high cost and numerical instabilities of DLMC.

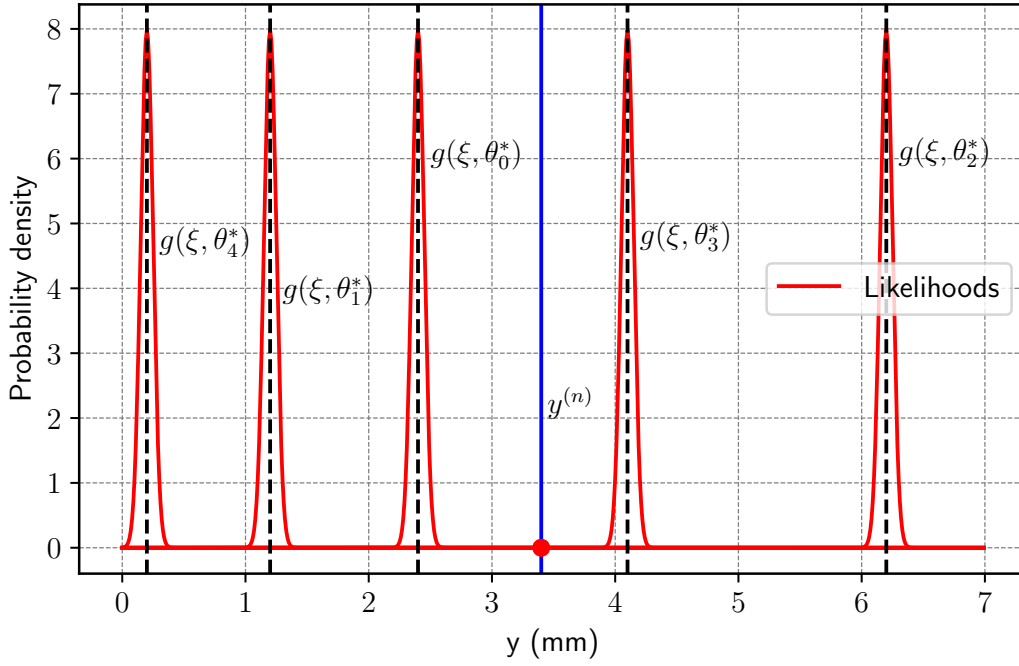


Figure 9: Numerical underflow illustrated for the three-point flexural test.

3.4 Monte Carlo estimator with Laplace approximation

The cost of solving the two-loop Monte Carlo required for DLMC can be large. Since MCI is a method for approximating integrals, one might think of alternative methods to approximate one of the integrals. Long et al. [2013] propose the use of the Laplace method to approximate the the logarithm of the posterior distribution by a second-order Taylor expansion, thus, avoiding the evaluation of the evidence. We follow the same approach as Long et al. [2013] to devise the Monte Carlo with Laplace method (MCLA) estimator. One advantage of the MCLA is that the approximated posterior pdf is a Gaussian function

$$\pi(\theta|Y, \xi) \approx \det(2\pi\Sigma(\xi, \hat{\theta}))^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\|\theta - \hat{\theta}(\xi, Y)\|_{\Sigma^{-1}(\xi, \hat{\theta})}^2\right), \quad (23)$$

where $\hat{\theta}$ is the maximum a posteriori (MAP) of θ , and $\Sigma(\xi, \hat{\theta})$ is the covariance matrix of the posterior at the MAP. The MAP is the θ that maximizes the posterior pdf, i.e., the θ more likely to be θ_t after the experiment data is considered. Thus, $\hat{\theta}$ is the θ that solves

$$\hat{\theta}(\xi, Y) \stackrel{\text{def}}{=} \arg \min_{\theta \in \Theta} \left[\frac{1}{2} \sum_{i=1}^{N_e} \|\mathbf{y}_i - \mathbf{g}(\xi, \theta)\|_{\Sigma_\epsilon^{-1}}^2 - \log(\pi(\theta)) \right]. \quad (24)$$

Long et al. [2013] show that

$$\hat{\theta} = \theta_t + \mathcal{O}_{\mathbb{P}}\left(\frac{1}{\sqrt{N_e}}\right). \quad (25)$$

The covariance matrix of the posterior $\Sigma(\xi, \hat{\theta})$ is the Hessian matrix of the negative logarithm of the posterior pdf evaluated at ξ and $\hat{\theta}$,

$$\Sigma^{-1}(\xi, \hat{\theta}) = N_e \nabla_{\theta} (\mathbf{g}(\xi, \hat{\theta}))^T \Sigma_{\epsilon}^{-1} \nabla_{\theta} \mathbf{g}(\xi, \hat{\theta}) - \nabla_{\theta} \nabla_{\theta} \log(\pi(\hat{\theta})) + \mathcal{O}_{\mathbb{P}}\left(\sqrt{N_e}\right). \quad (26)$$

A detailed deduction of $\hat{\theta}$ and $\Sigma(\xi, \hat{\theta})$ is presented in Beck et al. [2018].

According to Long et al. [2013], using the approximation $\hat{\theta} \approx \theta_t$ yields the Laplace-approximated SEIG as

$$I(\xi) = \int_{\Theta} \left[-\frac{1}{2} \log(\det(2\pi\Sigma(\xi, \theta_t))) - \frac{\dim(\theta)}{2} - \log(\pi(\theta_t)) \right] \pi(\theta_t) d\theta_t + \mathcal{O}\left(\frac{1}{N_e}\right). \quad (27)$$

Compared to (9), the approximated SEIG in (27) has only one integral, not requiring the integration over \mathbf{Y} nor the integral to calculate the evidence. This SEIG with Laplace approximation is consistent with Chaloner and Verdinelli [1995], according to whom, maximizing SEIG is equivalent to minimizing the determinant of the posterior covariance matrix.

Estimating I in (27) by MCI results in the MCLA:

$$\mathcal{I}_{\text{MCLA}}(\xi) \stackrel{\text{def}}{=} \frac{1}{N} \sum_{n=1}^N \left[-\frac{1}{2} \log(\det(2\pi\Sigma(\xi, \theta_n))) - \frac{\dim(\theta)}{2} - \log(\pi(\theta_n)) \right]. \quad (28)$$

According to Beck et al. [2018], the bias and variance of the MCLA estimator are, respectively,

$$|I - \mathbb{E}[\mathcal{I}_{\text{MCLA}}]| \leq C_{LA,1} h^\eta + \frac{C_{LA,2}}{N_e} + o(h^\eta), \quad \text{and} \quad (29)$$

$$\mathbb{V}[\mathcal{I}_{\text{MCLA}}] = \frac{C_{LA,3}}{N} \quad (30)$$

for the constants $C_{LA,1}$, $C_{LA,2}$, and $C_{LA,3}$. Although DLMC is a consistent estimator, MCLA is not, because the bias from the Laplace approximation does not vanish as the number of samples increases, being dependent on the number of repetitive experiments N_e .

Algorithm 2 presents the pseudocode for the MCLA estimator. Since MCLA does not have an inner loop, it only needs the outer loop sample-size, N . The evaluation of (26) requires a Jacobian of the model g with respect to the parameters θ ; the main cost of MCLA evaluation.

Algorithm 2 Pseudocode for the MCLA estimator for SEIG.

```

1: function MCLA( $\xi, N$ )
2:   for  $n = 1, 2, \dots, N$  do ▷ Outer loop
3:     Sample  $\theta_n$  from  $\pi(\theta)$ 
4:     Evaluate  $\nabla_{\theta} g(\xi, \theta_n)$ 
5:     Use  $\nabla_{\theta} g(\xi, \theta_n)$  to evaluate  $\Sigma(\xi, \theta_n)$  using (26)
6:   end for
7:    $\mathcal{I}_{\text{MCLA}}(\xi) \leftarrow \frac{1}{N} \sum_{n=1}^N \left[ -\frac{1}{2} \log(\det(2\pi\Sigma(\xi, \theta_n))) - \frac{\dim(\theta)}{2} - \log(\pi(\theta_n)) \right]$ 
8:   return  $\mathcal{I}_{\text{MCLA}}(\xi)$ 
9: end function

```

If forward-Euler is used for the estimation of the Jacobian in (26), the evaluation of the MCLA estimator cost has order $N(\dim(\theta) + 1)h^{-\ell}$. Thus, in comparison to DLMC, MCLA requires lower computational effort if $\dim(\theta)$ is less than M . For most cases, the M required to achieve a certain precision with DLMC is large, therefore, for these cases, MCLA is computationally more efficient than DLMC.

3.5 Double-loop Monte Carlo estimator with Laplace-based importance sampling

Based on the method proposed by Beck et al. [2018], we use an importance sampling in the MCI of the evidence. Instead of sampling θ^* from $\pi(\theta)$, we sample θ^* from $\tilde{\pi}(\theta)$, where

$\tilde{\pi}(\boldsymbol{\theta}) \sim \mathcal{N}(\hat{\boldsymbol{\theta}}, \boldsymbol{\Sigma}(\hat{\boldsymbol{\theta}}))$, and $\hat{\boldsymbol{\theta}}$ and $\boldsymbol{\Sigma}$ are given in (24) and (26). Therefore, we are using a Gaussian approximation of the posterior at its MAP to draw more informative samples for the inner loop. The advantage of using the importance sampling is that it avoids the approximation error from the Laplace approximation and the numerical underflow problem from DLMC with the drawback of requiring, for each outer loop, to solve the sub-problem of finding the MAP $\hat{\boldsymbol{\theta}}$. The DLMCIS estimator is defined as

$$\mathcal{I}_{DLMCIS}(\boldsymbol{\xi}) \stackrel{\text{def}}{=} \frac{1}{N} \sum_{n=1}^N \log \left(\frac{p(\mathbf{Y}_n | \boldsymbol{\theta}_n, \boldsymbol{\xi})}{\frac{1}{M} \sum_{m=1}^M \mathcal{L}(\mathbf{Y}_n; \boldsymbol{\xi}; \boldsymbol{\theta}_m^*)} \right), \quad (31)$$

where

$$\mathcal{L}(\mathbf{Y}; \boldsymbol{\xi}; \boldsymbol{\theta}) = \frac{p(\mathbf{Y} | \boldsymbol{\theta}, \boldsymbol{\xi}) \pi(\boldsymbol{\theta})}{\tilde{\pi}(\boldsymbol{\theta})}. \quad (32)$$

The bias and variance of the DLMCIS estimator are deduced on the original paper by Beck et al. [2018] and are proven to be the same as of DLMC for a given tolerance, thus the DLMCIS estimator is also consistent. However, to achieve the desired tolerance, the inner loop sample-size M is significantly smaller for DLMCIS than for DLMC. Like in MCLA, the evaluation of (26) requires a Jacobian of the model \mathbf{g} with respect to the parameters $\boldsymbol{\theta}$, moreover, finding $\hat{\boldsymbol{\theta}}$ requires solving the optimization problem in (24). We use a steepest descent search to find $\hat{\boldsymbol{\theta}}$. For that we use the gradient of the function to be minimized in (24),

$$\hat{\nabla}_{\boldsymbol{\theta}} \stackrel{\text{def}}{=} -(\nabla_{\boldsymbol{\theta}} \mathbf{g}(\boldsymbol{\xi}, \hat{\boldsymbol{\theta}}))^T \boldsymbol{\Sigma}_{\epsilon}(\mathbf{Y} - \mathbf{g}(\boldsymbol{\xi}, \hat{\boldsymbol{\theta}})) - \nabla_{\boldsymbol{\theta}} \log \pi(\hat{\boldsymbol{\theta}}). \quad (33)$$

Algorithm 3 Pseudocode for finding MAP using steepest descent.

```

1: function FINDMAP( $\boldsymbol{\xi}, \boldsymbol{\theta}, \mathbf{Y}, \alpha_{\boldsymbol{\theta}}, \text{TOL}$ )
2:    $\hat{\boldsymbol{\theta}} \leftarrow \boldsymbol{\theta}$ 
3:   for  $j = 1, 2, 3..$  do
4:      $\hat{\nabla}_{\boldsymbol{\theta}} \leftarrow -(\nabla_{\boldsymbol{\theta}} \mathbf{g}(\boldsymbol{\xi}, \hat{\boldsymbol{\theta}}))^T \boldsymbol{\Sigma}_{\epsilon}(\mathbf{Y} - \mathbf{g}(\boldsymbol{\xi}, \hat{\boldsymbol{\theta}})) - \nabla_{\boldsymbol{\theta}} \log \pi(\hat{\boldsymbol{\theta}})$ 
5:      $\hat{\boldsymbol{\theta}} \leftarrow \hat{\boldsymbol{\theta}} - \alpha_{\boldsymbol{\theta}} \hat{\nabla}_{\boldsymbol{\theta}}$ 
6:     if  $\|\hat{\nabla}_{\boldsymbol{\theta}}\|_2 < \text{TOL}$  then
7:       Break
8:     end if
9:   end for
10:  return  $\hat{\boldsymbol{\theta}}$ 
11: end function

```

If forward-Euler is used for the estimation of the Jacobians in (26) and Algorithm 3, the evaluation of the DLMCIS estimator costs $N((C_{MAP} + 1)(\dim(\boldsymbol{\theta}) + 1) + M)h^{-e}$, where C_{MAP} is the number of iterations to estimate $\hat{\boldsymbol{\theta}}$ in Eq. in Algorithm 3.

Besides the advantage of reducing the number of samples of the inner loop, M , the DLMCIS estimator is more robust to numerical underflow than the DLMC estimator. The change of measure in the sampling of $\boldsymbol{\theta}^*$ guarantees that $\mathbf{g}(\boldsymbol{\xi}, \boldsymbol{\theta}^*)$ is close to \mathbf{Y} , evaluated using the $\boldsymbol{\theta}$ sampled in the outer loop. Thus, it is not likely that, for all inner loops, the likelihood of observing \mathbf{Y} given $\boldsymbol{\theta}^*$ is numerically evaluated to zero. Figure 10 illustrates how the importance sampling mitigates the numerical underflow problem for the three-point flexural point; the likelihoods of observing $y^{(n)}$ for each g are larger than zero.

A pseudocode for the DLMCIS is presented in Algorithm 4, where the line where importance sampling happens is shaded in gray.

Algorithm 4 Pseudocode for the DLMCIS estimator for SEIG.

```

1: function DLMCIS( $\xi, N, M$ )
2:   for  $n = 1, 2, \dots, N$  do ▷ Outer loop
3:     Sample  $\theta_n$  from  $\pi(\theta)$ 
4:     for  $i = 1, 2, \dots, N_e$  do
5:       Sample  $\epsilon_i$  from  $\mathcal{N}(0, \Sigma_\epsilon)$ 
6:        $y_i \leftarrow g(\xi, \theta_n) + \epsilon_i$ 
7:     end for
8:      $\mathbf{Y}_n \leftarrow \{y_i\}_{i=1}^{N_e}$ 
9:      $p(\mathbf{Y}_n | \theta_n, \xi) \leftarrow \det(2\pi \Sigma_\epsilon)^{-\frac{N_e}{2}} \exp\left(-\frac{1}{2} \sum_{i=1}^{N_e} \|\epsilon_i\|_{\Sigma_\epsilon^{-1}}^2\right)$ 
10:    Find  $\hat{\theta}_n(\xi, \mathbf{Y}_n)$  using Algorithm 3
11:    Evaluate  $\nabla_{\theta} g(\xi, \hat{\theta}_n)$ 
12:    Use  $\nabla_{\theta} g(\xi, \hat{\theta}_n)$  to evaluate  $\Sigma(\xi, \hat{\theta}_n)$  using (26)
13:    for  $m = 1, 2, \dots, M$  do ▷ Inner loop
14:      Sample  $\theta_m^*$  from  $\tilde{\pi}(\theta) \sim \mathcal{N}(\hat{\theta}_n, \Sigma(\xi, \hat{\theta}_n))$  ▷ Importance sampling
15:      Evaluate  $g(\xi, \theta_m^*)$ 
16:       $p(\mathbf{Y}_n | \theta_m^*, \xi) \leftarrow \det(2\pi \Sigma_\epsilon)^{-\frac{N_e}{2}} \exp\left(-\frac{1}{2} \sum_{i=1}^{N_e} \left\| y_i^{(n)}(\xi) - g(\xi, \theta_m^*) \right\|_{\Sigma_\epsilon^{-1}}^2\right)$ 
17:       $\mathcal{L}(\mathbf{Y}_n; \xi; \theta_m^*) \leftarrow p(\mathbf{Y}_n | \theta_m^*, \xi) \pi(\theta_m^*) / \tilde{\pi}(\theta_m^*)$ 
18:    end for
19:     $p(\mathbf{Y}_n | \xi) \leftarrow \frac{1}{M} \sum_{m=1}^M \mathcal{L}(\mathbf{Y}_n; \xi; \theta_m^*)$ 
20:  end for
21:   $\mathcal{I}_{DLMCIS}(\xi) \leftarrow \frac{1}{N} \sum_{n=1}^N \log\left(\frac{p(\mathbf{Y}_n | \theta_n, \xi)}{p(\mathbf{Y}_n | \xi)}\right)$ 
22:  return  $\mathcal{I}_{DLMCIS}(\xi)$ 
23: end function

```

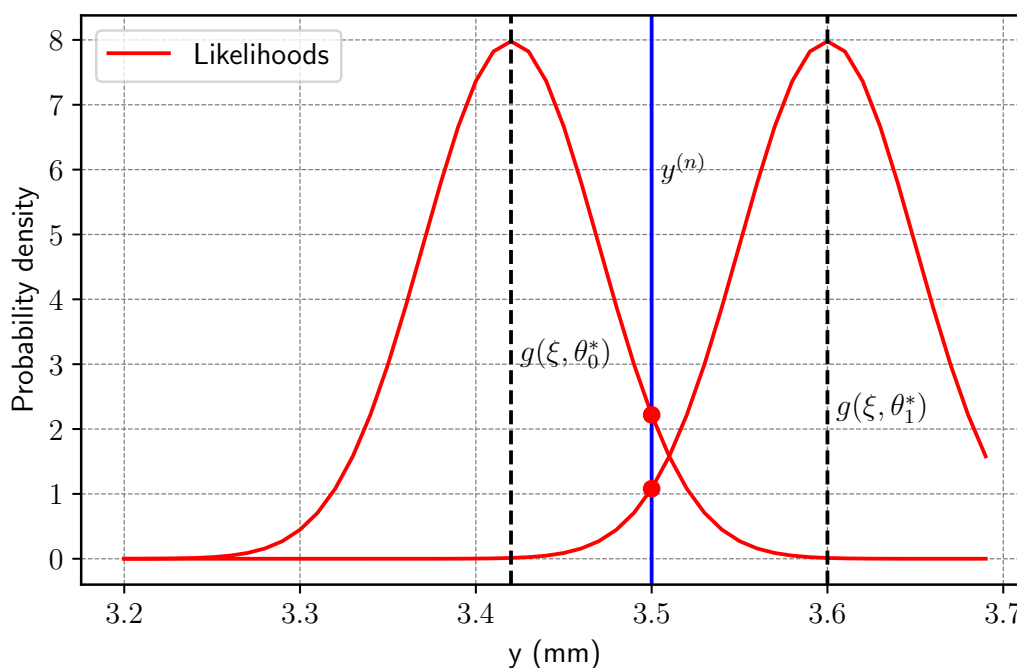


Figure 10: Avoiding numerical underflow using importance sampling illustrated for the three-point flexural test.

4 Numerical example

This numerical example addresses the problem of positioning a strain gauge on a beam modeled following Timoshenko theory for estimating the beam’s mechanical properties. Its purpose is to show that the proposed OED framework can reproduce the engineer’s intuition. In other words, the optimized design found is consistent with what an engineer would intuitively expect.

4.1 Strain gauge positioning on Timoshenko beam

The OED problem here consists in finding the optimal placement of a strain-gauge on a beam to estimate the material’s Young and shear deformation moduli. It is opted to employ the MCLA estimator of the EIG, given its lower cost in comparison to DLMC and DLMCIS.

It is assumed that the strain-gauges provide (noisy) strains observations in the vertical and longitudinal axes for a given point of the domain of the beam. The beam mechanical properties to be estimated are the the Young modulus E and the shear modulus G , given measurements obtained from the strain gauge using the Timoshenko beam model. The beam’s length, height and width are, respectively, $L_e = 10$ m, $H = 2$ m and $b = 0.1$ m. Although the beam’s geometry is not consistent with engineering practice, it is devised to furnish an interesting OED problem: we want an optimization problem where the optimum is not in a vertex of the beam. A uniform load q_o of 1.00 kN/mm is imposed on the beam’s vertical axis and distributed along its main axis. The geometry of the beam, the load, and the position of the strain gauge are illustrated in Figure 11.

We aim to locate a strain gauge on the beam that maximizes the information on E and G . We employ the Timoshenko’s theory Timoshenko [1921] as the forward model of the problem, a mechanical model that captures the strains resulting from both normal and shear stresses. Such a

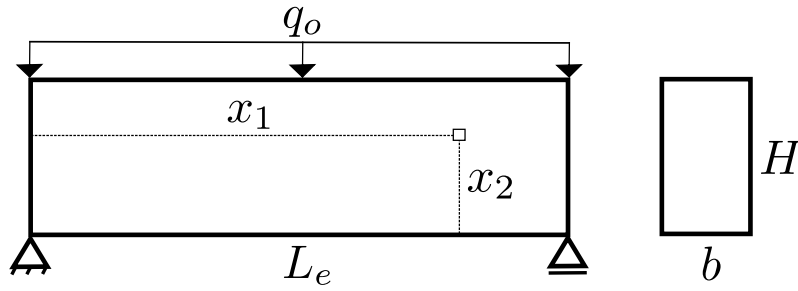


Figure 11: (Example 3) Geometry of the Timoshenko beam.

model is defined as

$$\begin{cases} K_s G A_r \varepsilon_{12} = \frac{q_o L_e}{2} - q_o x_1, \\ E I_n \varepsilon_{11} = \frac{q_o x_1 (L_e - x_1)}{2} x_2, \end{cases} \quad (34)$$

where ε_{11} is the normal strain, ε_{12} is the shear strain, x_1 and x_2 are the positions of the strain gauge on the horizontal and vertical axes respectively, q_o is the uniform load, L_e is the length of the beam, I_n is the inertia moment of the cross section, K_s is the Timoshenko constant ($K_s = 5/6$ in all test cases), and A_r is the cross-section area.

4.2 Bayesian formulation

The optimal position for the strain gauge that provides the maximum information about E and G is denoted by $\xi^* = (x_1^*, x_2^*)$. The longitudinal strain on the main axis of the beam, denoted by ε_{11} , together with the transverse strain ε_{12} , compose the output of the forward model. Therefore, based on (34), we find that

$$\begin{aligned} g(\xi, \theta) &= (\varepsilon_{11}(\xi, \theta), \varepsilon_{12}(\xi, \theta)) \\ &= \left(\frac{\xi_2 (q_o L_e \xi_1 - q_o \xi_1^2)}{2 \theta_1 I_n}, \frac{\frac{L_e}{2} q_o - q_o \xi_1}{K_s \theta_2 A_r} \right), \end{aligned} \quad (35)$$

where (x_1, x_2) and (E, G) are replaced by (ξ_1, ξ_2) and (θ_1, θ_2) , respectively. The additive error of the measurement is Gaussian $\epsilon \sim \mathcal{N}(0, \Sigma_\epsilon)$, where the noise covariance matrix is $\Sigma_\epsilon = \text{diag} \{ \sigma_{\epsilon_1}^2, \sigma_{\epsilon_2}^2 \}$.

4.3 Test cases

We investigate the EIG of this problem in four test cases, in which we attempt to locate the optimal strain-gauge placement on a beam. We test all the different cases, changing the variance of the prior pdf of θ , the dispersion of the measurement noise, and the number of experiments. The prior pdf of θ is Gaussian with the distribution $\pi(\theta) \sim \mathcal{N}((\mu_{pr}^E, \mu_{pr}^G)^T, \text{diag} \{ (\sigma_{pr}^G)^2, (\sigma_{pr}^E)^2 \})$, where $\mu_{pr}^E = 30.00$ GPa and $\mu_{pr}^G = 11.54$ GPa.

Table 1 presents the parameters used in each of the four cases.

We devised the parameters of the four cases with the intention of having four different OED problems. In the first two cases, both E and G have standard deviation of, respectively, 30% and 20% of their means. Also, case 1 has larger observation noise than case 2, and more experiments, 3. The third and fourth cases are exactly like case 2, except that case 3 has significantly less dispersion in G , a coefficient of variation of 4%, and case 4 has the same coefficient of variation for E .

Table 1: Parameters for the Timoshenko beam problem (Example 3).

Parameter	N_e	σ_{pr}^E (GPa)	σ_{pr}^G (GPa)	$\sigma_{\epsilon_1} (\times 10^{-4})$	$\sigma_{\epsilon_2} (\times 10^{-4})$
Case 1	3	9.00	3.46	6.25	1.30
Case 2	1	6.00	2.31	3.75	0.78
Case 3	1	6.00	0.46	3.75	0.78
Case 4	1	1.20	2.31	3.75	0.78

We evaluate the EIG of each case using a grid of uniformly distributed points, which provides contour plots of the expected information gain across the beam's domain. These plots are illustrated in Figure 12.

In cases 1 and 2, the optima are similarly located near the bottom of the beam, between the middle and the end. In case 3, the optimum is located in the bottom-middle of the beam; in case 4, the optimum is located on the supports. These placements are expected, as the Young modulus depends on the bending moment (that is maximum at the middle of the beam ($x_1 = L_e/2$)), and the shear modulus depends on the shear stress (that is maximum at the beam supports ($x_1 = 0$ and $x_1 = L_e$)). In case 3, the prior information about G is more accurate; consequently, the optimum is at the middle of the beam where more information about E can be collected. Conversely, in case 4, the optimum is at the supports, where data is more informative about G .

In Table 2, we present the non-optimized points randomly chosen on the design domain (Non-Opt.), the optimized setups (Opt.), the respective expected information gains in relation to the prior, and the standard deviations of the posterior pdfs of the parameters E and G for the four cases. The posteriors are evaluated at $\hat{\theta} = (\mu_{pr}^E, \mu_{pr}^G)$ for the four cases are presented in Figure 13.

We observe a reduced variance in the optimized experiment, compared to the original, reflecting the importance of an informative experiment. In cases 3 and 4, no information is acquired about G and E , respectively, since the variances in the axes are not reduced, compared to the prior.

Table 2: Results from the Timoshenko beam problem (Example 3).

		x_1^* (mm)	x_2^* (mm)	\mathcal{I}_{MCLA}	σ_{post}^E (GPa)	σ_{post}^G (GPa)
Case 1	Non-Opt.	5500.00	-100	0.14	8.00	2.40
	Opt.	8022.59	-1000.00	2.43	2.48	0.54
Case 2	Non-Opt.	5500.00	-100	0.23	2.38	1.38
	Opt.	7962.77	-1000.00	3.35	1.60	0.74
Case 3	Non-Opt.	5500.00	-100	0.06	5.70	0.46
	Opt.	5004.47	-1000.00	1.28	1.72	0.46
Case 4	Non-Opt.	5500.00	-100	0.22	1.20	1.93
	Opt.	10000.00	-1000.00	1.94	1.20	0.33

Because we use the biased and inconsistent MCLA estimator of the gradient, as a sanity check, we evaluate the gradient at the optima we found (the first two cases) using the full gradient of the DLMCIS estimator with $N = 10^3$ and $M = 10^2$. In both cases, the gradient norm is below 10^{-3} , meaning that the bias of the Laplace approximation is considerably small at the optima. We conclude that the biased solutions found are not significantly distant to the real optima.

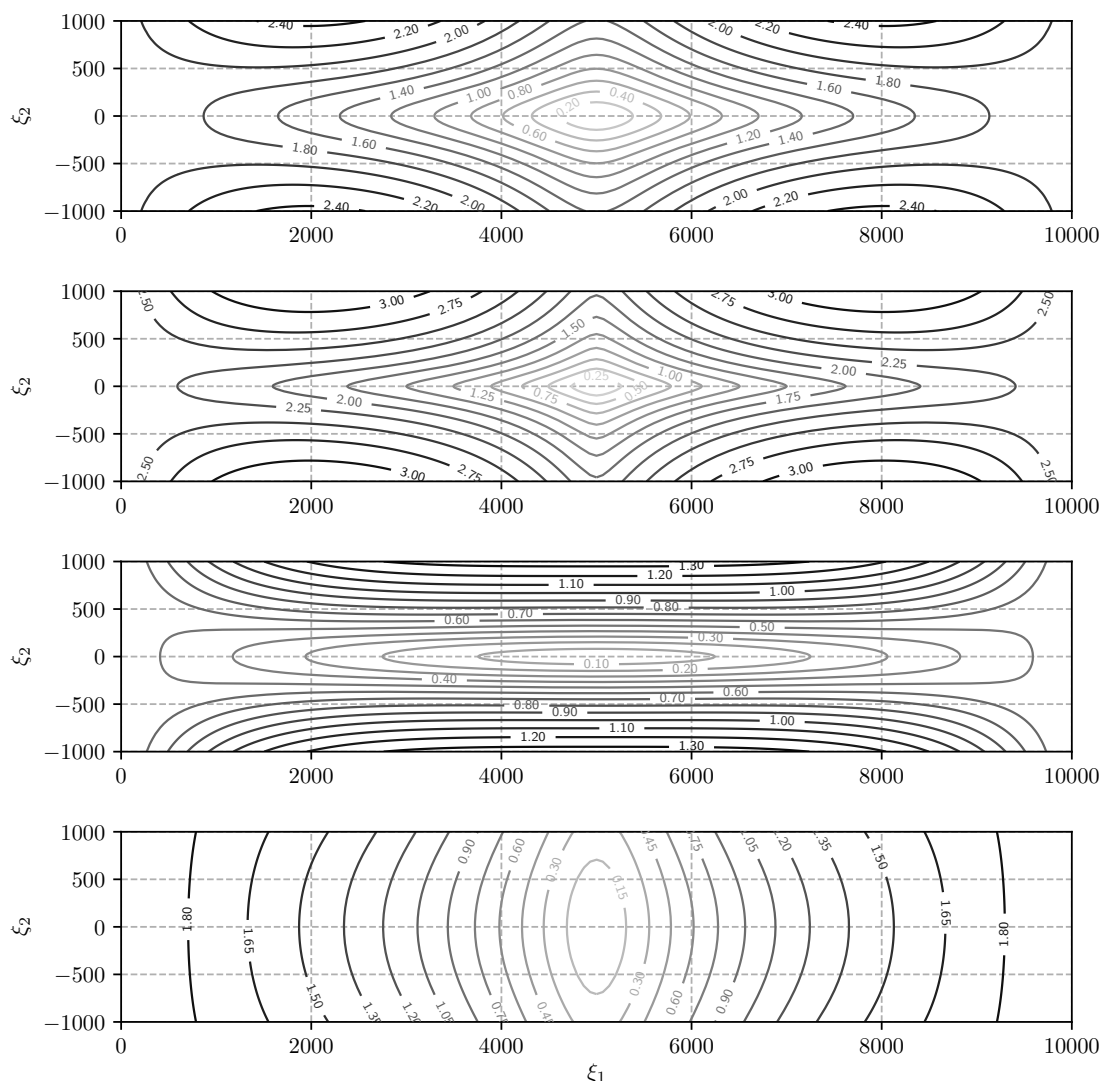


Figure 12: From top to bottom, cases 1 to 4 (summarized in Table 2). Expected information gain contours computed with MCLA.

The main goal of this example is to visualize how the optimum results are consistent with the input data given.

5 Chapter summary

In this chapter, we introduced key concepts related to Bayesian inference and OED. In Section 2.1, we presented the model of experiments with additive noise. Also, in Section 2.1, we defined a measure of efficiency of experiments: D_{KL} between the prior and posterior pdfs. Moreover, we presented, from an experiment with additive noise, the SEIG, the quantity we want to maximize.

In Section 3, we presented SEIG estimators based in MCI. Applying MCI in SEIG furnishes the DLMC estimator, introduced by Ryan [2003]. The DLMC estimator is consistent and can estimate SEIG for any general case, however, depending on the problem it can be expensive and unstable [Beck et al., 2018]. As an alternative to DLMC, we presented the MCLA estimator pro-

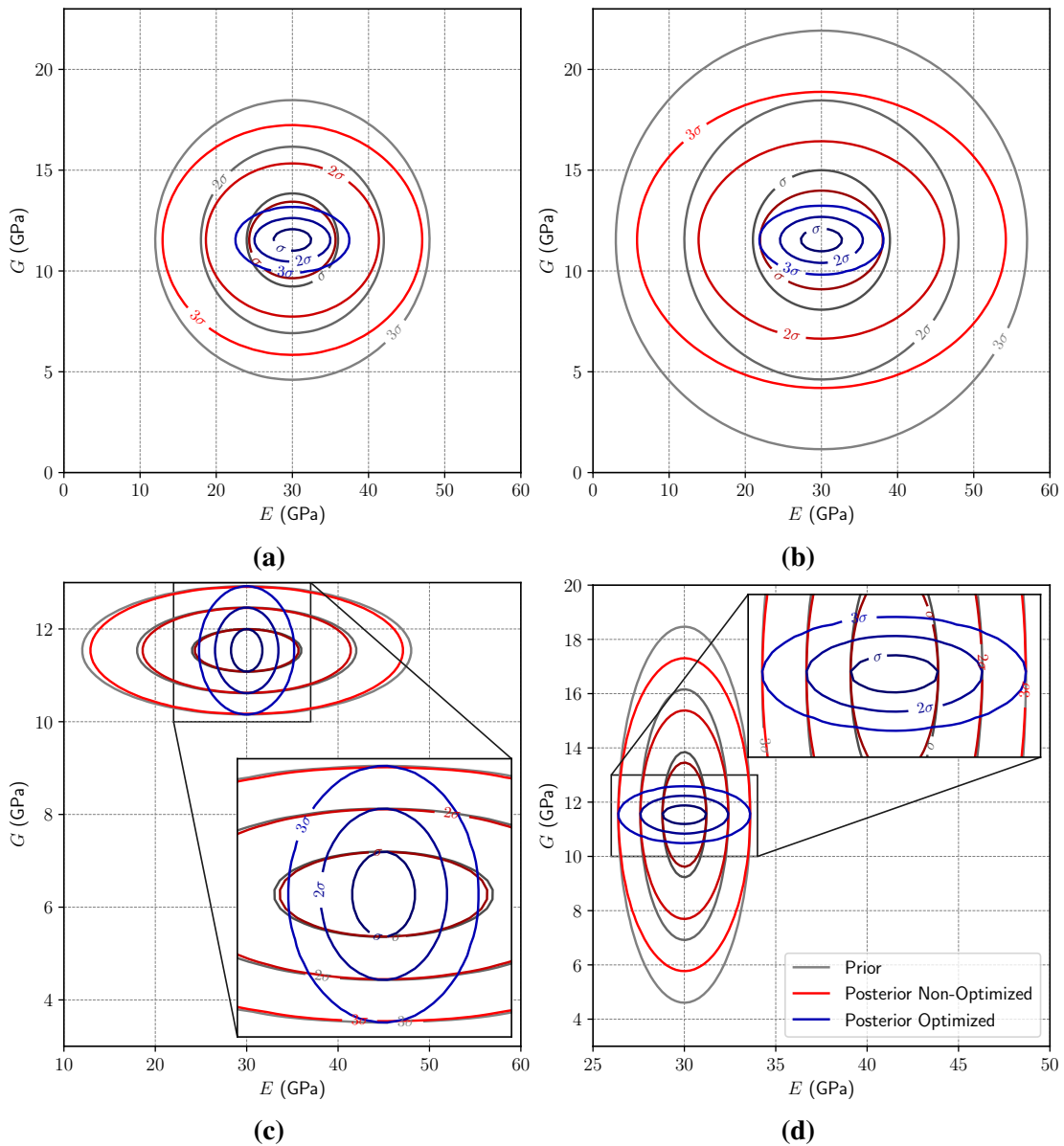


Figure 13: Prior, posterior, and optimized posterior pdfs for the Young modulus E and the shear modulus G for cases 1 (a), 2 (b), 3 (c), and 4 (d).

posed by Long et al. [2013] in Section 3. This estimator uses the Laplace approximation of the posterior, thus, avoiding the inner-loop in DLMC. The resulting algorithm, MCLA, is an inexpensive estimator for SEIG, with the drawback of introducing bias due to the Laplace approximation of the posterior distribution. This bias resulting from the Laplace approximation does not vanish as sample-sizes go to infinity, thus, MCLA is an inconsistent estimator. This fact is then investigated in the numerical example section. For the cases where the bias of MCLA is not acceptable, we introduced an importance sampling based on the Laplace approximation, proposed by Beck et al. [2018], resulting in the DLMCIS estimator.

Acknowledgements

Acknowledgements are optional, at the discretion of the chapter co-authors.

References

- J. Beck, B. M. Dia, L. F. R. Espath, Q. Long, and R. Tempone. Fast Bayesian experimental design: Laplace-based importance sampling for the expected information gain. *Computer Methods in Applied Mechanics and Engineering*, 334:523–553, 2018.
- K. Chaloner and I. Verdinelli. Bayesian experimental design: A review. *Statistical Science*, pages 273–304, 1995.
- W. Commons. Single-edge notch-bending specimen (also called three-point bending specimen) for fracture toughness testing., 2011. URL <https://commons.wikimedia.org/wiki/File:SingleEdgeNotchBending.svg>.
- T. M. Cover and J. A. Thomas. *Elements of information theory*. John Wiley & Sons, 2012.
- M. Hamada, H. Martz, C. Reese, and A. Wilson. Finding near-optimal bayesian experimental designs via genetic algorithms. *The American Statistician*, 55(3):175–181, 2001.
- X. Huan and Y. M. Marzouk. Simulation-based optimal bayesian experimental design for nonlinear systems. *Journal of Computational Physics*, 232(1):288–317, 2013.
- S. Kullback and R. A. Leibler. On information and sufficiency. *The annals of mathematical statistics*, 22(1):79–86, 1951.
- D. V. Lindley. On a measure of the information provided by an experiment. *The Annals of Mathematical Statistics*, pages 986–1005, 1956.
- Q. Long, M. Scavino, R. Tempone, and S. Wang. Fast estimation of expected information gains for bayesian experimental designs based on laplace approximations. *Computer Methods in Applied Mechanics and Engineering*, 259:24–39, 2013.
- C. Robert and G. Casella. *Monte Carlo statistical methods*. Springer Science & Business Media, 2013.
- R. Y. Rubinstein and D. P. Kroese. *Simulation and the Monte Carlo method*, volume 10. John Wiley & Sons, 2016.
- K. J. Ryan. Estimating expected information gains for experimental designs with application to the random fatigue-limit model. *Journal of Computational and Graphical Statistics*, 12(3): 585–603, 2003.

- C. E. Shannon. A mathematical theory of communication, part i, part ii. *Bell Syst. Tech. J.*, 27: 623–656, 1948.
- S. P. Timoshenko. Lxvi. on the correction for shear of the differential equation for transverse vibrations of prismatic bars. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 41(245):744–746, 1921.

Chapter 16

Airplane Optimal Robust Design Combined with an Entropy-Statistics Unsupervised Learning Classification Algorithm

Chapter details

Chapter DOI:

<https://doi.org/10.4322/978-65-86503-88-3.c16>

Chapter suggested citation / reference style:

Mattos, Bento S., et al. (2022). “Airplane Optimal Robust Design Combined with an Entropy-Statistics Unsupervised Learning Classification Algorithm”. In Jorge, Ariosto B., et al. (Eds.) *Uncertainty Modeling: Fundamental Concepts and Models*, Vol. III, UnB, Brasilia, DF, Brazil, pp. 502–551. Book series in Discrete Models, Inverse Methods, & Uncertainty Modeling in Structural Integrity.

P.S.: DOI may be included at the end of citation, for completeness.

Book details

Book: Uncertainty Modeling: Fundamental Concepts and Models

Edited by: Jorge, Ariosto B., Anflor, Carla T. M., Gomes, Guilherme F., & Carneiro, Sergio H. S.

Volume III of Book Series in:

Discrete Models, Inverse Methods, & Uncertainty Modeling in Structural Integrity

Published by: UnB City: Brasilia, DF, Brazil Year: 2022

DOI: <https://doi.org/10.4322/978-65-86503-88-3>

Airplane Optimal Robust Design Combined with an Entropy-Statistics Unsupervised Learning Classification Algorithm

Bento S. de Mattos^{1*} José A. T. G. Fregnani¹ Ariosto B. Jorge²

¹Instituto Tecnológico de Aeronáutica. E-mail: bmattos@ita.br, eng.fregnani@hotmail.com

²Post-Graduate Program - Integrity of Engineering Materials, University of Brasilia, Brazil. E-mail: ariosto.b.jorge@gmail.com

*Corresponding author

Abstract

The present work describes an airliner design platform that has been developed over several years and its application to design airliners based on an efficiency index developed by the authors of the present work. Among other characteristics, the design tool can handle airplane geometries with detail and count on a unique and accurate neural network system to predict aerodynamic coefficients. A new feature for product classification was incorporated into the framework. It is based on entropy statistics and was calibrated for jet transport airplanes and the results of classification are analyzed in the present work. The design of a Middle of Market airplane was then carried out with the classification provided by entropy statistics set as a constraint. Middle of the Market is a term that was established to address the commercial aircraft market segment that encompasses long-haul sectors from 3500 nm to 5000 nm, and airplanes with single-class passenger capacity between 220 and 250. Until now, this segment was not extensively explored, and some airlines have been showing interest to purchase airplanes better suited to it = this category. Multi-objective optimizations were carried out to determine the best aircraft design that fits into this concept's range and capacity specifications. The unsupervised learning algorithm based on entropy statistics was applied for airplane classification into four labels, classic, niche, failure, and breakthrough designs. A higher fidelity engine model optimization task was also carried out to search for optimal classic designs. In addition, a robust optimization considering a fuel price variation on Direct Operating Cost was carried out. All results raised an important discussion about the benefits and drawbacks of a trijet configuration for transport airplanes with a capacity of over 200 passengers.

Keywords: Commercial aviation, conceptual aircraft design, artificial neural network, multi-disciplinary design, and optimization, entropy statistics

Symbols and abbreviations

AAT	Airplane Analysis Tool
ANN	Artificial neural network
AR	Aspect ratio
ATC	Air Traffic Controller
BEW	Basic Empty Weight
BPR	Engine by-pass ratio
CAD	Computer-Aided Design
CAE	Computer Aided Engineering
CFD	Computational fluid dynamics
DOC	Direct operating cost [US\$/nm]
D_E	Engine external diameter
EI	Efficiency index for airliners
EI_{BJ}	Efficiency index for business jets
ETOPS	Extended-range Twin-engine Operations Performance Standards
FPR	Fan pressure ratio
HT	Horizontal tail
L/D	Lift-to-drag ratio
L_f	Fuselage length [m]
MDO	Multi-disciplinary design and optimization
MMA	Middle of the market aircraft
MMO	Maximum operating Mach number
MTOW	Maximum takeoff weight [kg]
n_e	Number of engines
NPAX	Passenger capacity in 2-class configuration (10% in business class)
OEW	Operating empty weight [kg]
OPR	Engine overall pressure ratio
PAX	Passenger or passengers
T	Overall takeoff thrust of an airplane [kN]
TBO	The time between overhaul for engines [h]
TFL	Takeoff field length
T_s	The thrust of an engine @ takeoff [kN]
TSFC	Thrust specific fuel consumption [kg/h/kg]
TW	Thrust-to-weight ratio
VT	Vertical tail
wAR	Wing aspect ratio
wS	Wing area [m ²]
wTR	Wing taper ratio
wSw14	Wing quarter-chord sweepback angle

1. Introduction

Aviation experienced a huge sprung when jet airplanes were introduced into commercial service in the late 1950s. Since then, worldwide passenger transportation has steadily increased at an exponential pace [1] thanks to more efficient airliners and to the Deregulation Act of 1978, which provided a boost for low-cost/low-fare airlines. Indeed, aviation is steadily evolving and is currently characterized by acute competition between two giant aircraft manufacturers in the medium and high-capacity market segment. Both companies introduced airliners to fulfill passenger and cargo demands in the 150-220 passenger capacity segment, airplanes that compose the fleet of most airlines around the world, especially those labeled as low-cost/low-fare airlines. In the high-capacity segment, high-performance airplanes enabled the establishment of new long-range routes [2].

As propulsion, material, and avionics technologies evolved, as well as computational tools (such as CFD, MDO, and electronic mockups) manufacturers could develop more fuel-efficient, long-range aircraft with a larger payload. Today, airplanes like Boeing 787 and Airbus A350 are 70% more fuel-efficient, in terms of kilograms of fuel burned per passenger, than the first-generation jet transports [3].

Figure 1 shows a range capacity graph where a gap between the short- and middle-range exists for single-aisle airliners. More specifically, there is an unfulfilled region for passenger capacity between 230 and 310 seats and a range above 3700 nm. In fact, until now there are no models available on the market that could accommodate 250 passengers into a single-class cabin layout offering a range beyond 4500 nm for shot/medium-haul operations. The MMA gap could be filled with airplanes ranging from 3700 to 5000 nm, capable of transporting 220-270 passengers. Originally this market was serviced by the Boeing 757 Series, which was designed to replace the venerable Boeing 707 for cross-country routes in the United States. The production of this airplane was discontinued, and its design was outclassed by newer, more fuel-efficient aircraft.

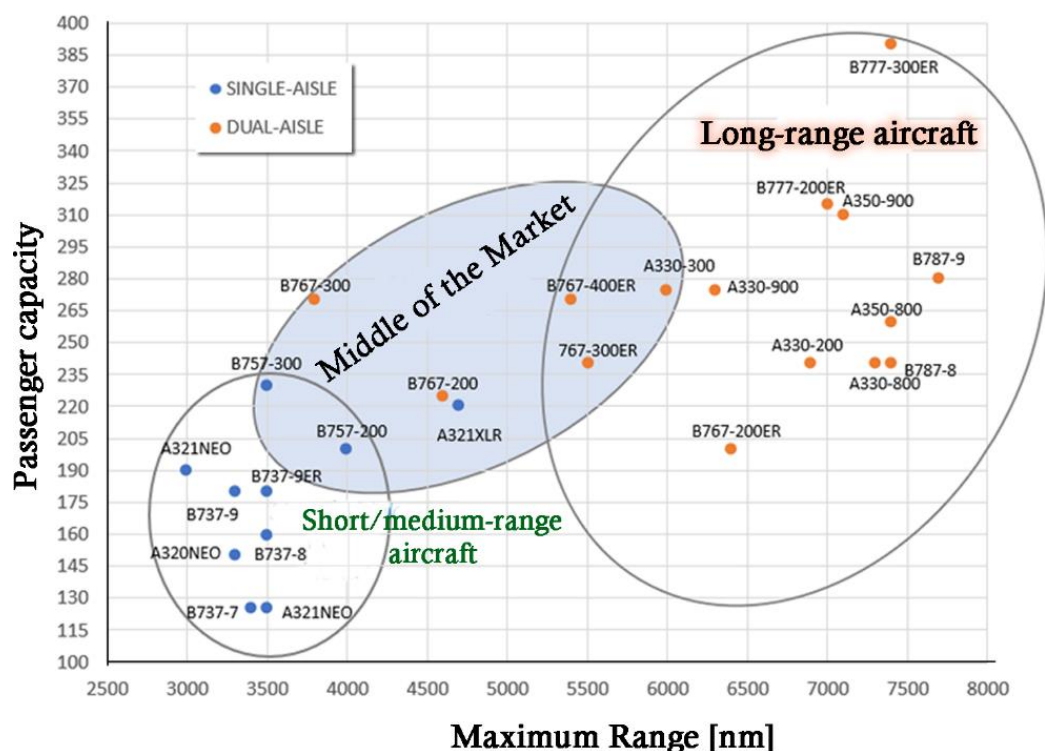


Figure 1: The middle of the market aircraft is waiting for a new airliner [4]

It is important to specify the kind of city pairs that are being considered for MMA operations. A non-stop concept for long-range direct routes was introduced by the new high-efficient wide-body designs like Boeing 787 and Airbus A350. This was due to the emergence of some markets that could be serviced by those ultra-long-range airliners. With the expansion of low-cost carriers into more distant destinations, using improved engine and airframe technologies present in current narrow-body airplanes, it could be possible to reach such markets with lower seating capacity and with narrow-body operating costs. Routes such as Honolulu-Seattle, Los Angeles-New York, New York-Manchester, London-Berlin, Dubai-Mumbai, and Beijing-Hong Kong could be served this way. These routes as well as some others would be feasible with such new, more efficient airliners, as shown in Figure 2 [5].



Figure 2: Potential routes with new generation narrow bodies [5]

Multi-disciplinary optimization tools are part of the everyday activities of aircraft manufacturers, whatever size they are. Thanks to the increase of computational power and research on algorithms and optimization methods, aircraft conceptual design is nowadays carried out by aircraft manufacturers using Multidisciplinary Design Optimization frameworks and sophisticated CAD/CAE software. Currently, many design variables (and types) may be employed, and advanced optimization algorithms that enable complex models of the aeronautical disciplines may be considered. Empirical formulae like the Class-I and -II methodologies proposed by Torenbeek [6] and Loftin [7] for drag and weight estimation are still used because they provide faster results [8]. Because of the increased complexity of such kind of methodology, significant efforts of computational analysis are required to obtain optimum configurations that satisfy many objectives and design constraints.

The extensive investigation of efficient airliners in the 150-250 seating capacity class is the objective of the present work by using an advanced MDO platform, encompassing aircraft which fit into the MMA market segment and their economic feasibility. The optimization tasks properly incorporated an entropy statistics tool for aircraft classification [9] [10]. This tool is based on an unsupervised machine learning algorithm that classifies products into four categories: scaled trajectories (classic); niche; breakthrough, and failure. This criterium was envisaged by Frtenken in his Ph.D. thesis [11]. Aircraft classification was then used as a constraint in the simulations. The multi-objective *gamultiobj* algorithm from MATLAB® was employed in all simulations that were carried out in the present work [12]. Two objectives were considered here: direct operating cost per nautical mile and an efficiency index for the configuration that was elaborated by the authors and will be discussed later in the following sections.

2. Methodology

This Section describes the MDO framework used in the present computations for the tasks to design airliners in the 150-240 seating capacity range. An airplane calculator module is an essential part of this computational tool; therefore, it will be addressed in some level detail. Initially, a description of the entropy statistics module, an important part of the present study, is described.

2.1 Entropy statistics

Historic development and description

German physicist Rudolf Clausius created the entropy concept in 1850 as a thermodynamic property that predicts the irreversibility or impossibility of certain spontaneous processes [13]. In statistical mechanics, entropy is formulated as a statistical property using probability theory. Statistical entropy or entropy statistics (ES) was introduced in 1870 by Ludwig Boltzmann, i.e., the statistical interpretation of the second law of thermodynamics. ES provides a linkage between the macroscopic observation of nature and the microscopic view based on the rigorous treatment of a large ensemble of microstates that constitute thermodynamic systems [13].

Entropy is a physic quantity that measures the degree of system diversity. An important parameter, relative entropy, was introduced by Kullback in 1951 [14]. According to Cover [15], relative entropy is a measure of the distance between two distributions. In statistics, it arises as an expected logarithm of the likelihood ratio. The relative entropy or Kullback-Leiber distance between two probability mass functions $p(x)$ and $q(x)$ is defined by:

$$I(p|q) = \sum_{x \in X} p(x) \log \left(\frac{p(x)}{q(x)} \right) \quad (1)$$

Indeed, the relative entropy $I(p|q)$, also known as Kullback–Leibler divergence is a special case of a broader class of divergences called f-divergences, considered a class of Bregman divergences [16]. It is the only such divergence over probabilities that belongs to both classes. Although it is often intuited as a way of measuring the distance between probability distributions, the Kullback–Leibler divergence is not a true metric for this purpose. It does not obey the triangle inequality, and therefore $I(p|q)$ does not equal $I(q|p)$. Nonetheless, it is often useful to consider the relative entropy index as a “distance” between distributions [15]. In its infinitesimal form, specifically its Hessian form, the relative entropy gives a metric tensor known as the Fisher information metric. The relative entropy is always non-negative, and it becomes zero if and only if $p = q$. For applications in the evaluation of product evolution, in Eq. 3, p is a posteriori distribution and q is a priori one. Based on this, two important parameters can be derived, which are important for product categorization. These parameters are called convergence and diffusion. Convergence is obtained when we consider an individual p and look at a certain period in the past and consider the characteristics of another individual q in the same timeframe. For the computation of the diffusion coefficient of a design that appeared in a determined time, designs that appeared after it in a determined timeframe must be selected [9]. Here, we consider the service entry of the aircraft as a time variable. The diffusion of a reference product can then be measured by its distance calculated by Eq. 3 to all the members of the technological population that were selected before. The final I-values are then obtained by dividing the sum of I-values for each product pair by the number of comparisons in the timeframe of observation [14].

The combination of diffusion and convergence indexes is useful for product classification. According to Frenken [9], industrial products can be classified as innovations, scaled trajectories, niches, and failures.

Innovative products can be recognized if they present new, redesigned, or substantially improved characteristics that are well distinguished from their predecessors. Innovative products influence their successors and introduce dominant designs. They are categorized by a low value of diffusion and a high value of convergence indexes. Scaled trajectories or classical designs follow patterns defined by previous concepts and influence their successors. Scaled trajectories are characterized by a low diffusion value as well as a low convergence value.

Products classified as “failures” may even be a market success, as will be shown by the application of the ES tool developed by Ref. [17]. This just means that they did not inherit major characteristics of their predecessors in a stipulated timeframe, and they did not transmit their ones to successors.

Products that fit into market niches are those complying with standards defined by previous designs and do not exert a profound influence on designs that come after them. They may be characterized by the combination of high diffusion and low convergence indexes. Failures are concepts that differ from their predecessors and do not influence their successors, characterized by a high value of its diffusion coefficient and a high value of the convergence index. The kind of design classification of industrial products described before was introduced by Frenken [9]. Figure 3 provides a compilation of this classification. Frenken originally designated designs in the Northeast quadrant as failures. Niche products are in the Southeast quadrant, and the scaled trajectories are in the Southwest one. An ellipse with a center at the confluence of the quadrants defines a region where the configurations are undefined, denominated here fuzzy designs (Figure 3). In the present work, scaled trajectories are referred to as classic airplanes.

Breakthrough products usually feature advanced technology. Sometimes, the creativity employed to define these products rather than technology provides them with unique characteristics. Breakthrough products are intended to fulfill consumer needs to an extended or greater degree than existing products, while most of the time featuring more advanced technology and innovation. These products enjoy the advantage of being the first movers on the market and can establish an early market share lead or a higher market share by outperforming existing products. Although all breakthrough products are first movers, they may not succeed in the marketplace in the long run.

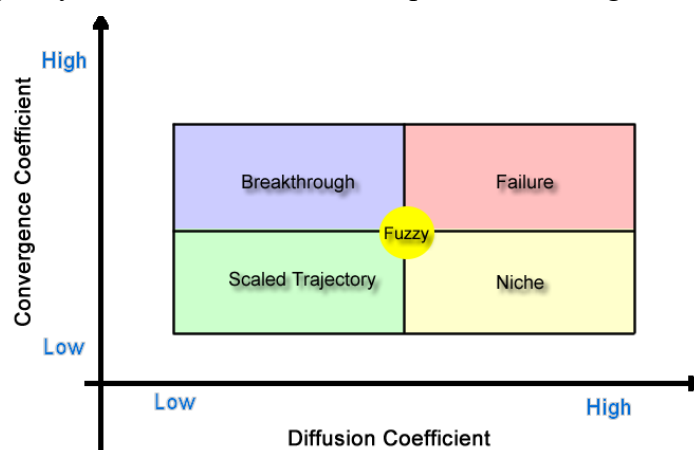


Figure 3: Product classification according to diffusion and convergence indexes

Application to jet airliners market

An Excel worksheet containing data for 123 jet airliners was compiled to feed the code with all data it needs. The first column of the worksheet was filled with airplane names; the second one contains the year of service entry of the airliners. The remaining columns contain the parameters employed for airplane description which are shown in Table 1. Note that the number of airplanes of each type that were delivered is not a variable.

Table 1: Variable stored in the airplane databank for entropy statistics analysis

1) MTOW [kg]	16) Horizontal tail sweepback angle
2) Maximum Zero-Fuel Weight [kg]	17) Takeoff thrust-to-MTOW ratio
3) Operating empty weight [kg]	18) Maximum operating Mach number
4) Fuel capacity [kg]	19) Maximum range with typical passenger payload [nm]
5) Wing reference area [m ²]	20) Service ceiling [ft]
6) Wing aspect ratio	21) Engine configuration (1= two underwing; 2= two at rear fuselage; 3=three at rear; 4= four underwing; 5= four at rear fuselage, 6= two overwing;7=two underwing + one at rear;8=two buried in wing;9=four buried in wing)
7) Wing taper ratio	22) Typical two-class passenger capacity at 32" pitch in the economy class
8) Wing quarter-chord sweepback angle [degrees]	23) Seating abreast in the economy class
9) Passenger cabin external width [m]	24) Wing position (1= low; 2=high)
10) Fuselage length [m]	25) Tail configuration (1=conventional; 2= "T" tail; 3=Cruciform tail)
11) Vertical tail aspect ratio	26) Number of decks
12) Vertical tail taper ratio	27) Number of corridors in passenger main cabin
13) Vertical tail quarter-chord sweepback angle	
14) Horizontal tail aspect ratio	
15) Horizontal tail taper ratio	

Mattos et al. implemented the entropy statistics methodology for product classification [17]. Table 2 contains the classification results for some airliners. Market data for each airplane is also shown in Table 2 to help to evaluate to some extent the degree of success of the airplanes. Most airplanes were properly classified. With the addition of some airliners and the incorporation of more accurate data regarding others, a new classification diagram was generated (Figure 4).

Table 2 – Airliner classification by entropy statistics [17] [18]

Airplane	Year of introduction*	Classification	Units delivered or built*
Boeing 747-400ER	2002	Failure	6
Boeing 767-300ER	1988	Scaled trajectory	583
Boeing 777-200LR	2006	Failure	59
Boeing 777-300	1998	Niche	60
Boeing 777-300ER	2004	Failure	839
Airbus A340-600	2002	Breakthrough	97
Boeing 737-800	1998	Scaled trajectory	4991
Boeing 737-900	2001	Scaled trajectory	52
Boeing 787-8	2012	Scaled trajectory	378
Caravelle VI	1955	Scaled trajectory	109
CRJ-100ER	1992	Failure	47
CRJ-200LR	1996	Breakthrough	298
ERJ145EP/ER/EU/MP	1997	Failure	146
ERJ145LR/LU	1998	Breakthrough	423
ERJ145XR	2002	Niche	110
ERJ135ER	1999	Failure	31
ERJ135LR/LU	2000	Failure/Niche?	94
ERJ140LR	1998	Niche	74
VFW-614	1975	Failure	19
* Data from Wikipedia and Aifleets.net			

The twinjet for 45 passengers EMB-145 found development in the late 1980s. The first concept outlined engines mounted over the wing, which was inherited from the twin-turboprop Brasília. Due to performance issues and a financial collapse of the Brazilian company, the development of the regional jet was lengthened, and the type entered service only in April 1997. By this time the airplane featured rear-mounted engines and the capacity was increased to 50 passengers. Later the type was redesignated ERJ145.

The entropy statistics provided the following results for the three ERJ-145 versions:

- ERJ145ER was positioned in the breakthrough region but remarkably close to that belonging to the failure one. Indeed, this seems to be correct due to the shorter range of this version. The ER was the first ERJ145 version and entered service in 1997. Embraer soon perceived this shortcoming and quickly developed and offered the LR version with a greater range, already in 1998.
- ERJ145LR. This version was labeled as a breakthrough design which also seems to be correct. Its direct competitor, the CRJ100 twinjet, entered service considerably earlier, in 1992. The CRJ 100 presents a range of 1,000 nm with a payload of 50 passengers; while the CRJ 100ER had its range increased to 1,650 nm with the same payload [19]. Why the Canadian CRJ 100 was not also categorized as a breakthrough? This can be possibly credited to its shorter range and heavier MTOW (23,133 kg). Probably in this category, the ERJ145LR seems to present the right characteristics to be classified as a breakthrough. Bombardier introduced the CRJ-200LR in 1996 [20], which features a 1,800-nm range (50 PAX) and better airfield performance when compared to the -100ER version. Alongside the ERJ145LR, the CRJ-200LR airliner appears in the classification box as a breakthrough concept.
- ERJ 145XR. This airplane seems to have been designed to fulfill a market niche. This version of the ERJ 145 presenting improved performance was purchased by ExpressJet (former Continental Express) only, a subsidiary of Continental Airlines. The Newark airport was the main operational base for this type. Thanks to the better performance when compared to the LR version, the ERJ 145XR enabled ExpressJet to service more distant cities from Newark.

For the high-capacity market, the classification that resulted for the B777-300ER and the A340-600 provides good stuff for discussion. Boeing B777-300.

- The long-haul A340-600 airliner can transport 361 passengers for a 7,540 nm stage length. This performance is comparable with that of B777-300ER when accommodating 339 passengers in a typical two-class layout (Figure 5). The A340-600 entered service in 2002 and the B777-300ER two years later [21]. Although with a poor sales record, the A340-600 was categorized as a breakthrough and its direct competitor as a failure. According to Table 2, the B777-300ER presents an outstanding sales record. It is worth mentioning that a breakthrough design is not a synonym for a successful product, therefore it could be a failure. On the other hand, regarding the B777-300ER airplane, based on the fundamentals of ES its classification makes sense because, in the timeframe of 5 years (and even longer), there is no other similar airplane to the B777-300ER. Thus, it received the “failure” label. The question that could be placed here is why the A340-600 was not as successful as the Boeing airplane. This probably resides in the number of engines; airlines prefer twin-engine airplanes for easy maintenance and maintainability.
- Launched in June 1995, the Boeing 777-300 major assembly started in March 1997, and it was rolled out on September 8 and made its first flight in October. The B777-300 was designed to be stretched accommodating 451 passengers in two classes, or up to 550 in an all-economy like the 747SR. Contrary to the extended range version, its sales record is not successful, indicating that airlines preferred the range instead of the larger capacity. It was classified as a niche by the ES application.

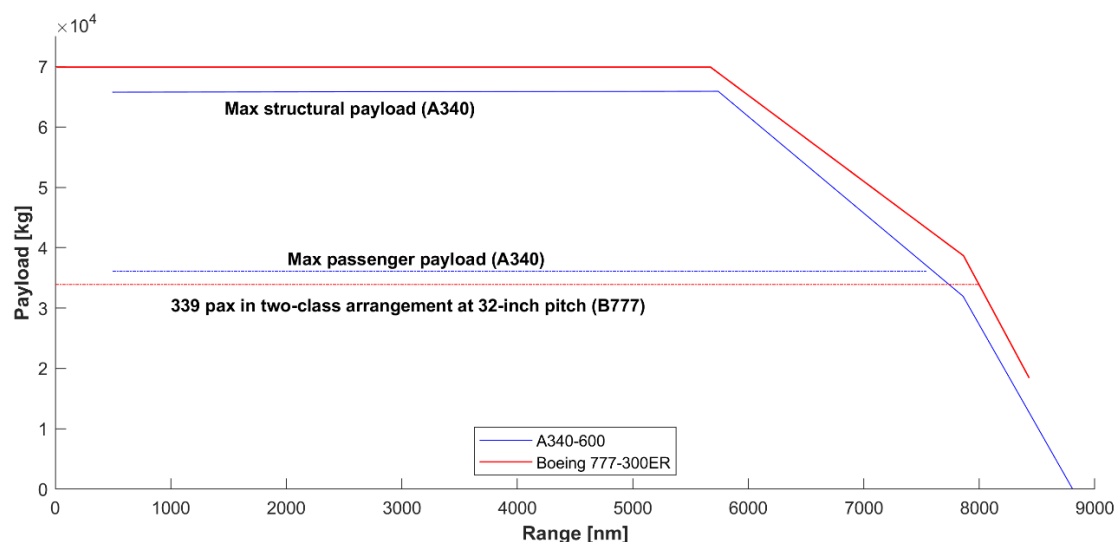


Figure 5: Payload-range diagrams for A340-600 and B777-300ER [22] [23]

Classification tree

The Classification Tree Method is a method for test design, as it is used in different areas of product development. It was elaborated by Grimm and Grochtmann in 1993 [24]. Classification Trees in terms of methodology cannot be confused with decision trees.

Classification tree methods greatly enhance the analysis capability the way of considering survey data and other information into it. There are three methods for producing classification trees: AID (automatic interaction detection), CHAID (chi-squared automatic interaction detection), and CHAID/CART (CHAID and classification and regression tree).

The classification tree method consists of two major steps:

- Identification of test-relevant aspects (so-called classifications) and their corresponding values (called classes) as well as
- Combination of different classes from all classifications into test cases.

The identification of test-relevant aspects usually follows a functional specification of the system under evaluation. These aspects form the input and output data space of the test object. The second step of test design follows combinatorial test design principles.

The chi-square automatic interaction detection (CHAID) is currently the most popular classification tree method. CHAID is much broader in scope than AID and can also be applied when the dependent variable is categorical. The algorithm that is used in the CHAID model splits records into groups with the same probability of the outcome, based on values of independent variables. Branching may be binary, ternary, or more.

When the relationship between a set of predictor variables and a response variable is linear, methods like multiple linear regression can produce accurate predictive models. However, for a system highly non-linear and complex, methods to handle non-linear behavior is a must. One such example of a non-linear method is classification and regression trees, often abbreviated as CART. As the name implies, CART models use a set of predictor variables to build decision trees that predict the value of a response variable.

Regarding node splitting rules, the MATLAB® command *fitctree* uses these processes to determine how to split node t [25]. For standard CART and for all predictors $x_i = 1, \dots, p$:

- *fitctree* computes the weighted impurity of node t , i_t .
- *fitctree* estimates the probability that an observation is in node t using

$$P(T) = \sum_{j \in T} w_j \tag{2}$$

with w_j being the weight of observation j , and T is the set of all observation indices in node t . If you do not specify prior or weights, then $w_j = 1/n$, where n is the sample size. There are additional important rules that are available in Ref. [25].

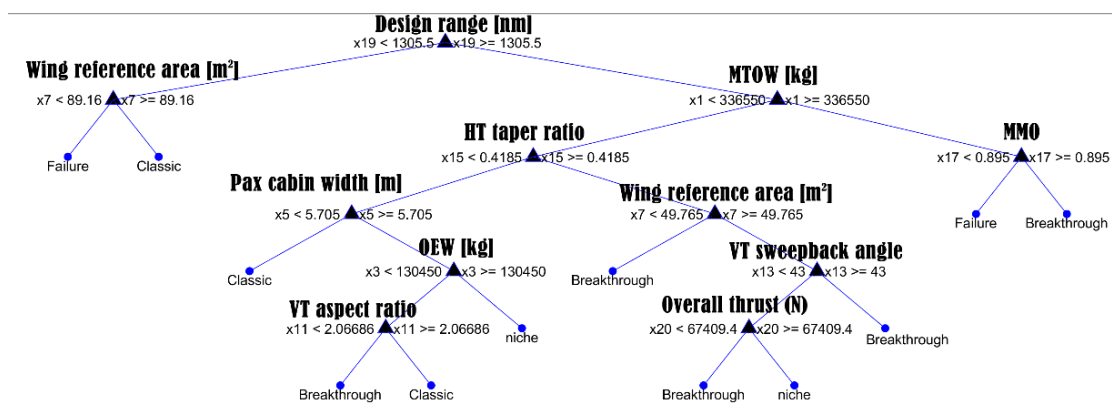


Figure 6: Decision tree classification based on ES

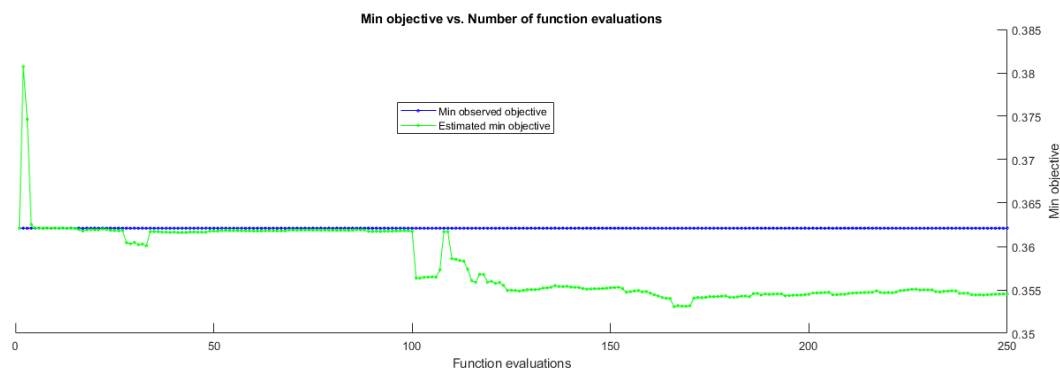


Figure 7: Convergence log of the Bayesian optimization task

The *fitctree* MATLAB® command was used to construct the tree seen in Figure 6. The convergence process with a Bayesian optimizer can be seen in Fig. Y. *fitctree* generates a classification tree model that can be an input for a predictor algorithm that delivers a label. Data of some airplanes already classified by ES were input into the predictor algorithm. Their labels of the classification tree were the same as those provided by ES.

It was expected that the configuration variables (see Table 1) would be part of the tree in Figure 6, mainly related to engine configuration. On the other hand, the appearance of the VT sweepback angle could be an indication that the classification tree of Figure 6 needs great improvements and is not ready to be employed in simulations, as a surrogate of the ES method. This utilization would be very interesting, considering that saving in computational cost will be considerable and the surrogate would deliver a classification independent of the time of aircraft service entry.

Dendrogram

Dendrograms are a good tool for group products, and they fit into the hierarchical clustering category [26]. MATLAB® has some algorithms that are easily able to compose a dendrogram. Considering the information shown in Table 1 and normalizing it, the Euclidian distance is calculated between all airliners from the databank, and then the Dendrogram is built (Fig. 4). This figure reveals the 10 closest configurations to the 767-200 high-gross sub-version (MTOW of 115,650 kg [27]). The cluster analysis using the Dendrogram approach correctly identified the larger 757-300 version as the closest configuration to the Boeing 757-200 airplane. Airbus A321-200 and a variant A321LR were placed in a group that is remarkably close to Boeing's airplane. Thus, this tool was then certified to be used for finding out individuals in the Pareto front closest to any reference airplane of choice, such as the Boeing 757-200.

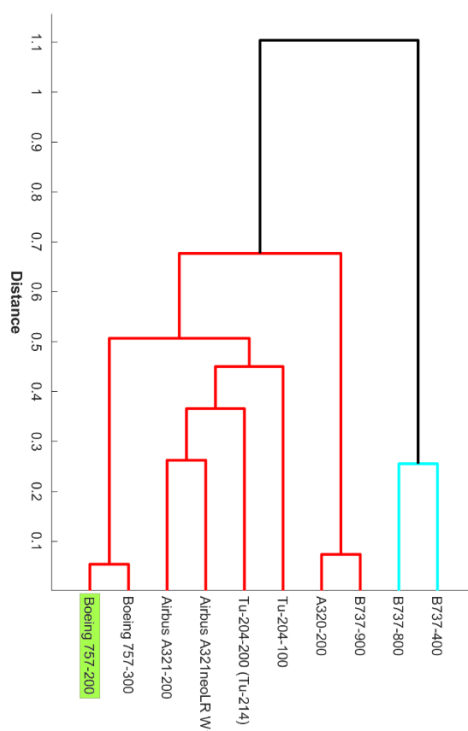


Figure 8: Dendrogram plot of the closest airplanes to Boeing 757-200

2.2 Optimization approaches

Thanks to the increase in computational power and advanced algorithms, and optimization methods, aircraft conceptual design is nowadays carried out by aircraft manufacturers using multi-disciplinary design and optimization frameworks and sophisticated CAD/CAE equipment. A great deal of computational analysis is required to obtain optimum configurations to satisfy many objectives and design constraints. Disciplines such as aerodynamics, propulsion, flight mechanics, structures, and aeroelasticity, among others, are frequently considered in the optimization framework to obtain more realistic geometries of flight vehicles and, recently, the addition of flight profiles and mission analysis [28]. Nowadays, many design variables (and types) may be employed, and advanced optimization algorithms that enable complex models of the aeronautical disciplines may be considered. Empirical formulae like the Class-I and -II methodologies proposed by Torenbeek [6] and Loftin [7] for drag and weight estimation are still used because they provide faster results [8]. Because of the increased complexity of such kind of methodology, significant efforts of computational analysis are required to obtain optimum configurations that satisfy many objectives and design constraints.

CFD analysis of complex aircraft configurations is still time-consuming, thus hampering in part its application in aircraft conceptual and preliminary design. Aerodynamic analysis requires considerable computational resources and labor work for the construction of complex 3D computational meshes, necessary to produce credible results. The case is even worse for multi-disciplinary optimizations, which deal with high-dimension problems that may reach thousands of design variables and their relationships. However, surrogate techniques have been increasingly adopted to model the physics with enough precision and at the same time reduce the computational time of complex and heavy calculations in MDO frameworks [28] [29].

The use of gradient-based optimization algorithms associated with adjoint methods is an alternative to cut considerably the processing time of optimization problems since their computational cost does not strongly scale up with the number of design variables compared to evolutionary algorithms [30]. “There are two approaches for the adjoint method: continuous adjoint, by deriving the exact adjoint equation and then linearizing it; or the discrete adjoint, by directly deriving it from the discretized flow equations” [30]. “Both approaches compute the adjoint variables by an additional linear system of equations. Afterward, together with the mesh sensitivities, the final sensitivities, or gradients, are evaluated and transferred to an optimizer” [30].

On other hand, there are some limitations of the adjoint methodology:

- Huge and long tasks must be carried out in preparation for several third-party codes to be used for automatic differentiation algorithms (Their source codes sometimes are not available).
- Usually, mono-objective optimization for aircraft conceptual design is possible. For objectives encompassing several disciplines, the approach employed by adjoint optimization frameworks is to proceed with a weighted sum of objectives.
- It is challenging to consider optimization problems with discrete variables.
- It is characterized by low flexibility in changing constraints and objectives.
- Adjoint methods can handle tens of thousands of variables to define a wing geometry. However, airfoil geometry can be defined by ten or a few more variables by using parametrization like PARSEC [31]. **Fig. 12** illustrates the excellent modeling capability performed for the NACA21010 airfoil by a PARSEC parametrization with just 12 parameters.

Genetic algorithms, an evolutionary approach based on their metaheuristic of reproduction of the fittest, have widespread and effective use in the optimization of aircraft configuration [32] [33]. The procedure of reproduction of the fittest is repeated until some established criterium to close the process is reached. Evolutionary algorithms solve both constrained and unconstrained optimization problems its great advantage lies in the absence of the necessity of gradient calculation of the objective function with regards to the design variables. Opposed to this, is that this necessity is present in classic gradient optimization methods such as gradient descent and quasi-newton methods.

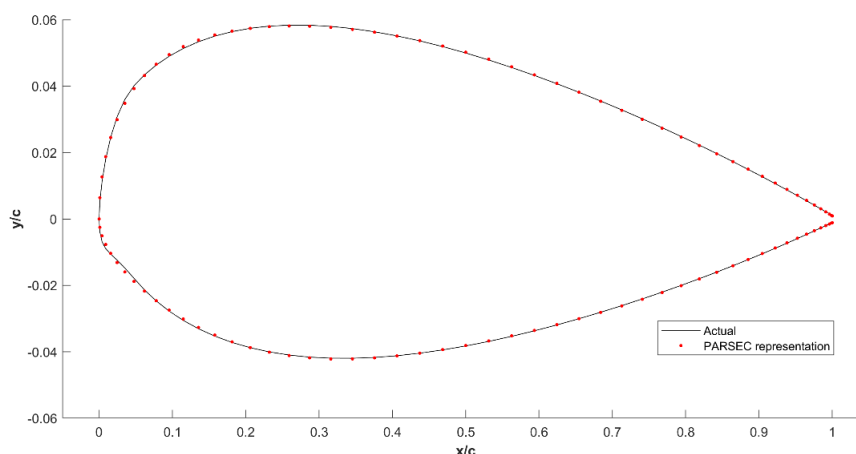


Figure 9: PARSEC representation of NACA21010 airfoil

Bayesian optimization is another approach that offers advantages over the adjoint and heuristics methods. It employs a sequential design strategy for global optimization of black-box functions that does not assume any functional forms. Bayesian optimization is usually employed to optimize expensive-to-evaluate functions [34] [35]. Since the objective function is unknown, the Bayesian strategy is to treat it as a random function and place a prior over it. The prior captures confidence about the behavior of the function. After gathering the function evaluations, which are treated as data, the prior is updated to form the posterior distribution over the objective function. The posterior distribution, in turn, is used to construct an acquisition function (often also referred to as infill sampling criteria) that determines the next query point.

Priem et al. [34] employed Bayesian optimization for the optimal design of a business jet (BJ) designated Bombardier Research Aircraft. Conceptual and preliminary design were addressed with the Bayesian optimization approach. For the conceptual design, 11 design variables were selected, with 6 of them related to rear spar chordwise location and airfoil maximum relative thickness along definition stations along the wingspan.

The constrained Bayesian optimization procedure is summarized below:

Problem setup: objective and constraints functions, starting DoE and setting the maximum number of iterations
Loop i=1 to max iterations
Build the surrogate models using Gaussian Process
Find x^{i+1} a solution to the enrichment maximization sub-problem
Evaluate objectives and constraints
Update de DoE
End loop
Final solution: the optimal point from DoE

2.3 Robust Optimization

Robust optimization (RO) has undoubtedly become an effective method of choice for optimization under uncertainty in process systems engineering with practical applications ranging from production scheduling to airfoil and wing design. RO does not presume that probability distributions are known, but instead, data uncertainty ordinarily resides in a user-specified set of uncertain parameter values, called the uncertainty set [36]. In RO the modeler aims to find decisions that are optimal for the worst-case realization of the uncertainties within a given set. Typically, the original uncertain optimization problem is converted into an equivalent deterministic form, called the robust counterpart, using strong duality arguments and then solved with standard optimization algorithms.

However, robust optimization methodologies have an important drawback. By the very nature of the methodology, robust optimal solutions are highly conservative, aiming to hedge against all possible worst-case realizations of the uncertainty. In some cases, this can lead to solutions that are too conservative. For example, if a chemical plant or biorefinery is designed too conservatively, then the project may be too expensive or needlessly complex. Researchers have looked for methods to control the robustness or conservativeness of robust optimization techniques, and some promising methods, such as adaptive or adjustable robust optimization have recently surfaced.

The main techniques of robustness used to solve optimization problems with uncertainty are summarized as follows [37]:

- *Strict robustness or static robustness.* Static RO intends to find solutions that are resistant to all perturbations of the data in a so-called uncertainty set. First, consider that $x \in S$ is a solution to the optimization problem with uncertainty $P(\xi)$, with $\xi \in U$. The solution is strict if x is feasible for all possible scenarios of U , that is if $F(x, \xi) \leq 0$ for all $\xi \in U$. This technique is very intuitive to solve the optimization problem robustly.
- *Cardinality constrained robustness.* One approach to relaxing the strict robustness is to restrict the space of uncertainty. This way, it allows for accounting for a certain degree of uncertainty with a reasonable computational effort, providing a trade-off between computational time and robustness. In addition, it can be tuned to consider the specific degree of risk the decision maker accepts [38].
- *Light robustness.* A completely different way of relaxing the concept of strict robustness corresponds to instead of reducing the space of uncertainty, we can relax the constraints in favor of the quality of the solution.
- *Adjustable robustness.* Another way to relax the space of uncertainty of strict robustness is to split the space into groups of variables. In this technique, some of the decision variables are adjusted after some portion of the uncertain data reveals itself [36].
- *Recoverable robustness.* Recoverable robustness uses the concept of a recovery algorithm, and, like adjustable robustness, it obtains the solution in two stages. Give a family of algorithms A . A solution x is recovery robust concerning A if it exists for every scenario $\xi \in U$, an algorithm A applied to the solution x , and a scenario ξ allows you to build a solution $A(x, \xi) \in F(\xi)$ [37].

2.4 Airplane analysis tool

Introduction

The airplane analysis tool (AAT) is the masterpiece of the present design framework. It provides an embracing description and characterization of the airplane. Many levels of fidelity for discipline representation are available. AAT also calculates the noise and emissions footprint. Detailed mission performance and range evaluations are derived from flight profiles set by the user and obeying air traffic constraints, all based on the aerodynamics, mass, and engine characteristics of the airplane. It is possible to examine any design mission or off-design missions corresponding to specific takeoff weights or required block distances.

The airplane mission setup is highly detailed with a flight profile that makes accurate calculations of major airplane masses possible like MTOW, OEW, and MZFW. Air traffic constraints can be incorporated into the calculation, if required [39]. Several aerodynamic methods are available in the framework with different levels of fidelity, including an artificial neural network surrogate model, which is the object of the present work. An in-house generic turbofan engine model is also part of the MDO framework [40]. It is possible to analyze the performance, emissions, and operational costs for missions inside a complex airline network with different takeoff weights or block distances [41]. The fuel storage capacity is precisely calculated considering the actual airplane geometry and structural layout.

All mission segments are analyzed, and optimal and transitional climb altitudes are calculated considering buffet margins, available engine thrust, rate of climb, and accurate airplane lift-to-drag ratio. Alternate airport, hold and descent patterns, reserve and maneuvering fuels, and weight allowances are additional parameters. Typically, after the second segment climb, the in-route climb uses airspeed schedules, namely, a calibrated airspeed (CAS) segment followed by a constant previously stipulated Mach number until the initial cruise altitude. The cruise phase can be flown at a constant altitude or following a step-increasing altitude pattern. There are two cruise speed profiles, one with a constant Mach number, the other being a long-range pattern.

Fuel storage

Fuel is stored in the wingbox and there is a routine that computes the available volume, and therefore it requires information on wing spar location. Depending on user choice, fuel can also be stored in a fuel tank located inside the fuselage close to the position where the wing-fuselage junction is.

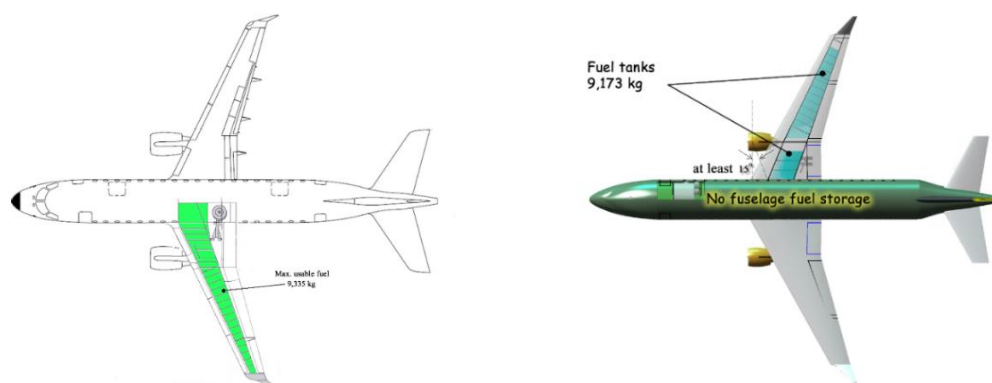


Figure 10: Fuel storage in actual aircraft (on the left) and for the model of the present work (on the right)

For purposes of illustration, the fuel storage approach of Embraer E-jets is depicted in **Fig. 5**. Besides the storage inside the wing, another region where the fuel can optionally be stored is a place inside the fuselage. It is worth mentioning that the boundary of the wing outboard tank is close to the wingtip [42]. According to the manufacturer, the E170 has a usable fuel capacity of 9,428 kg. The airplane analysis tool offers two options already outlined before: the first one considers that the fuel is stored only in the wings and the tanks extend until 90% of the wingspan (**Fig. 5**). In this case, for the E170, the resulting fuel capacity is 9,178 kg. As the airfoil geometry for this twinjet is not available, this figure is just an approximate value; the other possibility in the design framework is to consider a fuel tank inside the fuselage. Its capacity is then dependent on the cabin's central fuselage dimensions and the location of the cabin floor.

Maximum lift coefficient

High-lift characteristics have a huge impact on the sizing, economics, and safety of an airplane. Airplanes featuring lower approach speed are safer and complex high-lift systems increase manufacturing costs and reduce flight dispatchability. It is important to achieve the highest possible maximum lift coefficient at takeoff and landing to avoid the necessity of higher wing areas.

Applied analysis of the Navier-Stokes equations is very time-consuming due to the complex nature of the flow in combination with complex geometries. In addition, considering regions of highly separated flow convergence of the solution poses an immense difficulty for the code. Several analytical methods that both account for stall effects and enable the prediction of the maximum lift coefficient were discussed and a trade-off between the methods was conducted in Ref [43]. From this trade-off, it was found that both the pressure difference rule and the so-called Critical Section Method are good options in terms of accuracy and versatility.

The pressure difference rule, on the other hand, determines wing stall by implementing an empirical chordwise pressure difference criterion. Thus, wing stall is established when the wing section's chordwise pressure difference, at an angle of attack, equals the critical pressure difference which is empirically derived and depends on both the Mach- and Reynolds number.

The critical section method utilizes a two-dimensional airfoil analysis code for the maximum lift coefficient calculation at many prescribed wing sections. By using three-dimensional wing-fuselage combinations C_{L0} and C_{La} coefficients for the wing sections are obtained. After a sweeping procedure, a stall is reached when the lift coefficient distribution along the semispan is locally equal to the wing section's maximum lift coefficient.

Singh [43] validated both the pressure difference rule and critical section method against wind-tunnel experiments, both for wing-only samples and wing-fuselage configurations. The approximation of the clean maximum lift coefficient showed to be in good agreement, with about a 3% error margin, which is a particularly good figure.

For the present computations, a MATLAB® code was elaborated to automatically run the panel code XFOIL [44] for the calculation of section maximum lift coefficient. A full potential code for wing-fuselage combinations then calculates the lift coefficient distribution along the wingspan at two conditions with different angles of attack. After the clean wing configuration is calculated, the Datcom method [45] is used to estimate the maximum lift coefficient at takeoff and landing from flap configuration and some wing geometric parameters.

Empty weight calculation

For the calculation of the maximum takeoff weight (MTOW), the estimation of the empty weight is needed, which can be supplied according to the MTOW (class I weight method) itself or it can be calculated by the sum of the individual structural weights and aircraft systems (class II method). Within the Class I methodology, the following relationship between MTOW and *OEW* was developed by the authors:

$$\frac{OEW}{MTOW} = \left[a + b \left(\widehat{MTOW} \widehat{AR} TW \widehat{S}_w MMO \widehat{L}_f F_w \widehat{\Psi} \widehat{SC} \right)^{c0} + c \left(\widehat{MTOW} \right)^{c1} \left(\widehat{AR} \right)^{c2} \left(TW \right)^{c3} \left(\widehat{S}_w \right)^{c4} \left(MMO \right)^{c5} \left(\widehat{L}_f \right)^{c6} \left(F_w \right)^{c7} \left(\Psi \right)^{c8} \left(\widehat{SC} \right)^{c9} \right] \tag{3}$$

Eq. 3 was tailored for a mix of units of the international and English systems. The *MMO*, *TW*, and *F_w* symbols refer to the maximum Mach of operation, the thrust-to-weight ratio, and the diameter of the cross-section of the passenger cabin, respectively.

Normalizations that were considered in **Eq. 3**, are given in Table 3.

Table 3: Variable normalization for Eq. 3

$\widehat{MTOW} = \frac{MTOW}{50000}$	<i>MTOW</i> = Maximum Takeoff Weight [kg]
$\widehat{AR} = \frac{AR}{8}$	<i>AR</i> = Wing aspect ratio
$\widehat{WL} = \frac{WL}{100}$	<i>WL</i> = Wing loading [kg/m ²]
$\widehat{L}_f = \frac{L_f}{30}$	<i>L_f</i> = Fuselage length [m]
$\widehat{\Psi} = \frac{\Psi}{20}$	<i>Ψ</i> = Quarter-chord sweepback angle [degree]
$\widehat{SC} = \frac{SC}{40000}$	<i>SC</i> = Service Ceiling [ft]

Table 4: Values of the coefficients present in Eq. 3

a	b	c	C1	C2	C3	C4	C5	C6	C7	C8	C9
-0.2359	-0.3065	0.8900	-0.1731	0.0871	0.0857	0.1050	-0.0931	0.1220	0.2232	0.0857	0.0418

To obtain the exponents and multipliers of the terms of **Eq. 3**, the authors used a database of line aircraft containing information from 123 models, with entry into service from the 1950s to others with commercial operation expected in 2020. **Fig. 6** shows the relationship between the MTOW and the reference area of the wings for the aircraft in the database. An optimization algorithm was elaborated for the minimization of the mean quadratic error between the *OEW/MTOW* estimated by **Eq. 3** and the actual values of the airplanes in the database. The exponents and coefficients of **Eq. 3** are the design variables. The mono-objective genetic algorithm of MATLAB® 2019a was used in optimization simulations and after 5500 generations the simulation was stopped. Table 4 shows the coefficients and exponents that the optimization run provided. Table 5 contains some estimation errors for the weight fraction.

Table 5: Estimation error of weight fraction for some jet airliners

Airplane	Actual OEW/MTOW	Calculated OEW/MTOW	Percent error
Boeing 737-700	0.54157	0.53548	1.12
Airbus A320-200	0.54494	0.53829	1.22
B747-400	0.49077	0.4976	1.39
Boeing 757-300	0.53874	0.49639	7.86
Boeing 767-300	0.55681	0.55400	0.505
CRJ200ER	0.59806	0.59552	0.425
EMBRAER E170LR	0.55645	0.56798	2.07
Fokker 100	0.57074	0.55416	2.90
Tupolev Tu-104A	0.54737	0.56043	2.39
VFW 614	0.61048	0.60329	1.18

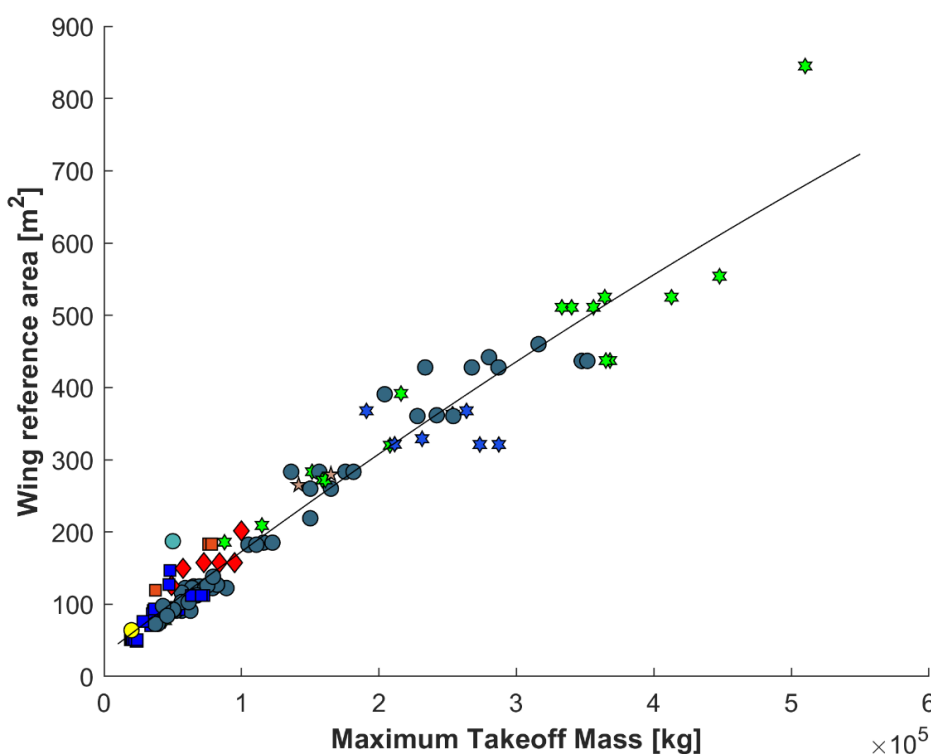


Figure 11: MTOW plotted against wing reference area considering airplanes utilized for the elaboration of Eq. 2. The different symbols in the graph reflect the different engine configurations

Engine weight calculation

There are some simple methods as well as some more sophisticated methods for the estimation of turbofan engine weight [46]. But for WATE++, all of them are not accurate enough to be used in the optimization framework for airplane design. Thus, a methodology developed by the authors was adopted for turbofan engine weight estimation. This approach utilizes a formulation based on some engine geometric and thermodynamic parameters (Eq. 4).

$$W_E = \frac{T1 * (BPR/\overline{BPR})^a (OPR/\overline{OPR})^b (T_s/\overline{T})^c (D_E/\overline{D_E})^d (L_E/\overline{L_E})^e (\dot{m}/\overline{m})^f + T2 * (BPR/\overline{BPR})(OPR/\overline{OPR})(T_s/\overline{T})(D_E/\overline{D_E})(L_E/\overline{L_E})(\dot{m}/\overline{m}) + T3}{(4)}$$

The coefficients and exponents of **Eq. 3** were obtained by optimization using a genetic algorithm [47]. The objective was the minimization of the square mean error of known engine weights. A database, comprised of 20 engines, was elaborated which embraced a large variety of turbofan engines [48]. Table 6 shows the coefficients and A, B, and C parameters obtained with the optimization process. Table 7 shows the average parameters used for normalization in **Eq. 4**.

Table 6 - Obtained exponents and coefficients for Eq. 4

Coefficient/exponent	Value
T1	2587.2461
T2	50.1920
T3	154.6179
a	-0.1965
b	-0.0718
c	1.0435
d	0.2493
e	-0.3444
f	-0.1455

Table 7: Parameters used for normalization in Eq. 4

Parameter	Value
\overline{BPR}	4.6911
\overline{OPR}	25.4000
\overline{D}_E [m]	1.7906
\overline{L}_E [m]	3.3276
\overline{T} [kN]	148.1217
\overline{m} [kg/s]	464.7333

Table 8 contains weight estimation errors for some known turbofan engines.

Table 8: Weight error estimation for some known engines

Engine	Percent error
AE3007A1P	10.94
CF6-50C	2.55
JT8D-219	0.74
GE CF-34-10A	6.48
R&R RB211-535C	0.97
Trent 800-875	3.12
Williams FJ-44	4.29
Pratt & Whitney PW2040	0.44
GE-90/77B	0.31
R&R Tay 620	2.65

Workflow of the airplane analysis tool

The airplane calculator workflow is shown in Figure 12. From airplane geometry and topology an iterative process for MTOW calculation is performed. The formulation for the OEW/MTOW weight fraction of Eq. 3 and the methodology outlined in Eq. 4 for the estimation of the engine weight is employed in this process. To evaluate the accuracy of the present design workflow three airliners of different categories were analyzed and Tables VIII-XI contain their characteristics and those estimated by the airplane calculator tool.

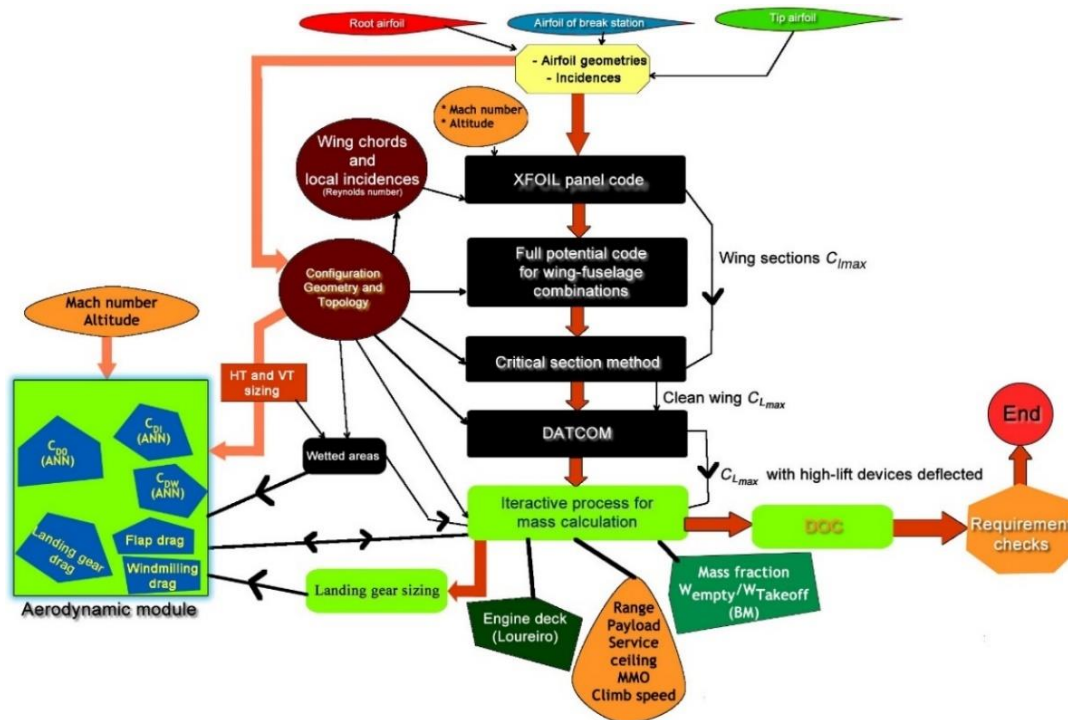


Figure 12: Workflow of the airplane analysis tool

Calculating the A320 airliner

The A320 twinjet presents a six-abreast cross-section and is powered by either CFM56 or IAE V2500 turbofans. The A320 family, which includes the A318 and A321, pioneered the use of digital fly-by-wire and side-stick flight controls in airliners. Variants offer maximum take-off weights from 68 to 93.5 t, to cover a 5,740–6,940 km (3,100–3,750 nm) range [49]. Some of the A320 main characteristics are presented in Table 9 and the MTOW and OEW calculated by the present methodology are given in Table 10.

Table 9: Characteristics of A320-200

Design Range	L_f	wAR	wS	wSw14	BPR	Engine Dry Weight	D_E
2,350 ^α nm	37.57 m	10.33	122.4 m ²	25°	4.8:1	2,400 kg	1.75 m
MTOW	OEW	Usable fuel	T	Aisle Width Y-class	MMO		
73,500 ^θ kg	42,600 ^ψ kg	21,005 ^β kg	220 kN	0.48 m	0.82		
α @ FL370, a payload of 15,000 kg, ISA+0 °C, takeoff with MTOW, IAE V2500-A5							
β Includes one fuselage tank							
θ Airbus A320-200 WV015 [50]							
ψ Ref. [51]							

Table 10: Estimated values for the A320-200 airliner by the present design framework

MTOW	OEW	Fuel Capacity	Range	Categorized as
73,440 kg	41,371 ^α kg	19,795 ^α kg	2,350 nm ^α	Classic
^α @ FL370, a payload of 15,000 kg, ISA+0 °C, takeoff with MTOW, IAE V2527-A5, 200 nm to an alternate airport				

The B757 has a six-abreast cross-section and is powered by either RB211 or PW2030 and PW2040 Series turbofans. The B757-200 is the basic version, which entered service in 1983.

In August of 1978, Eastern Airlines and British Airways announced orders for the B757 and chose the RB211-535 to equip the airplane [52]. Designated RB211-535C, the 3-spool engine entered service in January of 1983 when the first B757-200 was delivered to Eastern Airlines. The engine has a nominal thrust of 37,400 lbf (166.36 kN) and an OPR of 21.2:1 [53]. In 1979, Pratt & Whitney launched its PW2000 engine, claiming 8% better fuel efficiency than the -535C for the PW2037 version [52]. The English engine manufacturer reacted and using the -524 core as a basis, the company developed the 40,100 lbf (178 kN) thrust RB211-535E4, which entered service in October of 1984. There are differences in appearance between the two versions like a mixed exhaust nozzle and a bigger fan cone for the RB211-535E4. BPR was slightly reduced to 4.40:1 from 4.46 for the 535C [53]. There is another version of the Rolls&Royce engine designated RB-211E4B, which has a takeoff thrust of 43,500 lb [54]. The 535E4 engine was also the first to use the wide chord fan which increased efficiency, reduced noise, and gave increased protection against damage from foreign object ingestion [52]. As a result, a relatively small number of -535Cs was installed on production aircraft in May of 1988. American Airlines ordered 50 B757s powered by the -535E4 emphasizing the engine's low noise as an important factor in their choice. The stretched version B757-300 entered service with Condor Flugdienst in 1999. With a length of 54.5 m, the type is the longest single-aisle twinjet ever built.

According to Ref. [27], The B757-200 has several sub-versions presenting different MTOW, OEW, and MZFW. Two configurations fitted with RB211-535E4B engines were selected as references for the present work, and their payload-range diagrams are shown in Figure 13 and some characteristics are given in Table 11 [54]. Two ranges signaled by the dashed lines in the payload-range diagrams are related to a mission with a payload of 192 passengers.

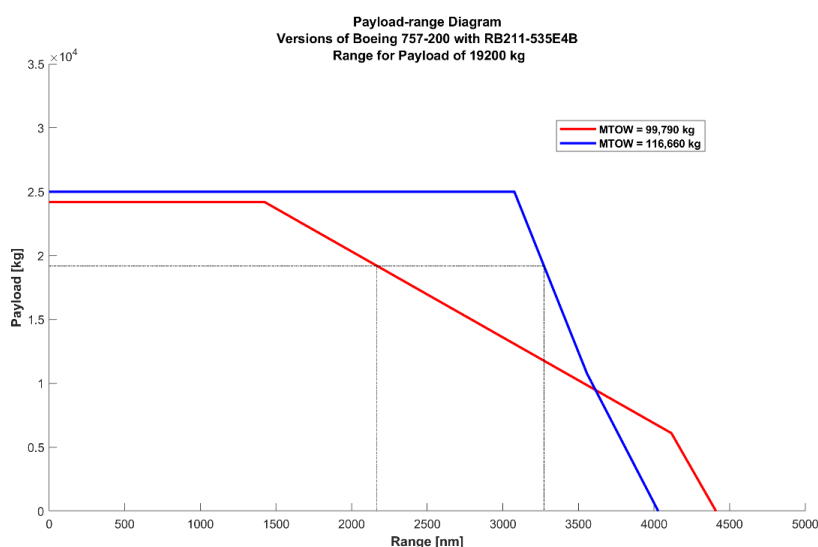
**Figure 13: Payload-range diagram for some sub-versions of B757-200 [54]**

Table 11: Characteristics of two B757-200 sub-versions fitted with RB211-535E4B engines [22] [54] [53] [52]

Sub-version	MTOW	OEW	MZFW	Usable fuel	Takeoff Thrust	TFL
1	99,790 kg	58,570 kg	83,460 kg	42,680 kg	37,400 lbf	1,660 m
2	115,650 kg	58,570 kg	84,360 kg	43,490 kg	43,500 lbf	2,070 m
L_f	wAR	wS	wSw14	MMO	Aisle width of Y-class	
46.97 m	7.82	185.25 m ²	25°	0.86	0.508 m	

Some of the B757-200 main characteristics calculated by the present methodology are given in Table 12. Again, the agreement between actual and estimated weights is particularly good. The year 1984 was set for the year of service entry. As expected, the airplane was categorized as a classic design (scaled trajectory). However, just for testing, if the year 1981 were adopted for that parameter, the airplane would be labeled by the entropy statistics classification program as a breakthrough, as also is expected. The reason behind this is simple: this way the configuration being analyzed passes its unique characteristics to the actual airplanes that entered service in 1983 and 1984, which are part of the database.

Table 12: Characteristics of B757-200 model fitted with RB211-535E4

MTOW	OEW	Fuel Capacity	Range	Categorized as
115,100 kg	58,495 ^α kg	41,138 ^β kg	3,272 nm ^α	Classic
α @ FL370 and FL390, Mach of 0.80, payload of 19,200 kg, ISA+0 °C, takeoff with MTOW, RB211-535E4, 200 nm to an alternate, 30 min loiter @ 1,500 ft				
β wing capacity of 30,918 kg / ε Jet A1 price of US\$ 2.387 US\$/gal				

2.5 Multi-disciplinary and multi-objective optimization

Introduction

Thanks to the increase of computational power and research on algorithms and optimization methods, aircraft conceptual design is nowadays carried out by aircraft manufacturers using MDO frameworks and sophisticated CAD/CAE equipment. A great deal of computational analysis is required to obtain optimum configurations to satisfy many objectives and design constraints. Disciplines such as aerodynamics, propulsion, flight mechanics, structures, and aeroelasticity, among others, are frequently considered in the optimization framework to obtain more realistic geometries of flight vehicles and, recently, the addition of flight profile and mission analysis has been added [28]. Nowadays, many design variables (and types) may be employed, and advanced optimization algorithms that enable complex models of the aeronautical disciplines may be considered. Empirical formulae like the Class-I and -II methodologies proposed by Torenbeek [6] and Loftin [7] for drag and weight estimation are still used because they provide faster results [8].

Despite the rapid computer development and computational techniques, CFD analysis of complex aircraft configurations is still time-consuming, thus hampering in part its application in aircraft conceptual and preliminary design. Aerodynamic analysis requires considerable computational resources and labor work for the construction of complex 3D computational meshes, necessary to produce credible results. The case is even worse for multi-disciplinary optimizations, which deal with high-dimension problems that may reach thousands of design variables and their relationships. However, surrogate techniques have been increasingly adopted to model the physics with enough precision while at the same time reducing the computational time of complex and heavy calculations in MDO frameworks [28] [29].

The use of gradient-based optimization algorithms associated with adjoint methods is an alternative to considerably cut the processing time of optimization problems since their computational cost does not strongly scale up with the number of design variables compared to evolutionary algorithms [30]. There are two approaches for the adjoint method: continuous adjoint, by deriving the exact adjoint equation and then linearizing it; or the discrete adjoint, by directly deriving it from the discretized flow equations [30]. Both approaches compute the adjoint variables by an additional linear system of equations. Afterward, together with the mesh sensitivities, the final sensitivities, or gradients, are evaluated and transferred to an optimizer. On the other hand, there are some disadvantages to the adjoint methodology:

- Huge and long tasks must be carried out in preparation for several third-party codes to be used for automatic differentiation algorithms (Their source codes sometimes are not available).
- Usually, only mono-objective optimization for aircraft conceptual design is possible. For objectives encompassing several disciplines, the approach employed by adjoint optimization frameworks is to proceed with a weighted sum of objectives.
- It is challenging to consider optimization problems with discrete variables.
- It is characterized by low flexibility in changing constraints and objectives.
- Adjoint methods can handle tens of thousands of variables to define a wing geometry. However, airfoil geometry can be defined by ten or a few more variables by using parametrization like PARSEC [55].

Genetic algorithms, an evolutionary approach based on their metaheuristic of reproduction of the fittest, have widespread and effective use in the optimization of aircraft configuration [32] [33]. The procedure of reproduction of the fittest is repeated until some established criterium to close the process is reached. Evolutionary algorithms solve both constrained and unconstrained optimization problems whose great advantage lies in the absence of the necessity of gradient calculation of the objective function concerning the design variables. Opposed to this, is that this necessity is present in classic gradient optimization methods such as gradient descent and quasi-newton methods.

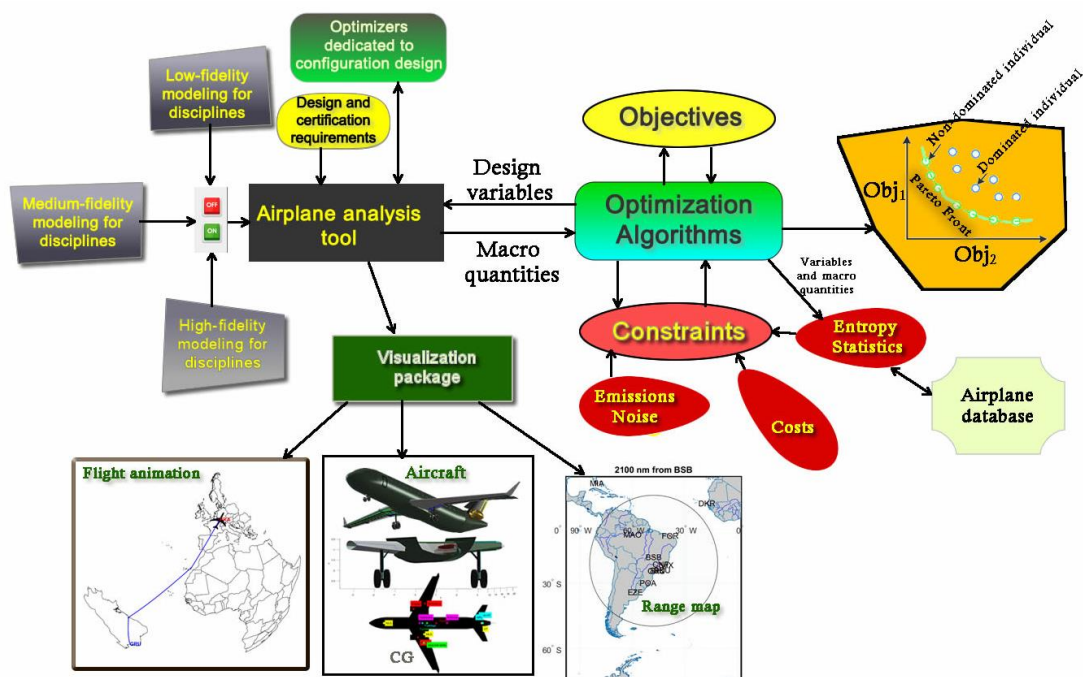


Figure 14: The MDO platform workflow of the present methodology

The workflow of the airplane design framework of the present work is shown in Figure 14. There are many possibilities in terms of fidelity of the discipline modeling that is used for airplane sizing. For example, lift-drag characteristics can be calculated by Class-I formulations as well as CFD codes, or even by surrogate ANN models [29]. In the present work, multi-objective genetic algorithms were chosen to perform the optimization runs.

Important design requirements must be considered because when satisfied, they will generate a feasible airplane configuration. Certification requirements such as climb gradient in 2nd segment, cruise and missed approach thrust requirements, rate of climb at service ceiling, and balanced takeoff field length are some of them.

The airplane classification provided by the entropy statistics approach was incorporated into the MDO platform. For the optimization runs carried out in the present work a category among the four shown in Figure 3 was set as a constraint. The service entry is a time variable used by the entropy statistics algorithm to perform the calculation of the convergence and diffusion indexes required for product classification. The year 2017 was selected as a reference year for all airplane configurations handled by the genetic algorithm to perform the classification procedure. The algorithm searches the airplanes in a database consisting of 123 jet airliners and selects configurations with service entry before and after the reference year within a prescribed timeframe (5 years in the present work).

Objectives

Objective 1: efficiency index

The authors elaborated a parameter that synthesizes the value of an airliner and that was designated here as the efficiency index:

$$EI_{Liner} = \frac{\widehat{Rnm} \times MMO \times \widehat{NPAX}}{(TW \times \widehat{OEW})} \quad (5)$$

The symbols in **Eq. 4** are:

$$\widehat{NPAX} = \frac{NPAX}{150} \quad \widehat{Rnm} = \frac{Rnm}{2500} \quad \widehat{OEW} = \frac{OEW}{40000}$$

A low-efficiency index may be an indication of a low range, low cruise speed, or a high OEW for the mission it was conceived for. A low speed may appear not to be important, but it can be translated into a longer block time that impacts DOC. Besides, slower airplanes are obligated by air traffic control to fly at lower altitudes to avoid spoiling faster airplanes, causing in general higher fuel consumption. For business jet, which presents a different mission profile than that for airliners, the following efficiency index is proposed:

$$EI_{BJ} = \frac{\widehat{Rnm} \times MMO \times \widehat{CV}}{(TW \times \widehat{OEW})} \quad (6)$$

In **Eq. 5**, \widehat{CV} is the passenger cabin volume, given in m³, divided by a normalization parameter. This figure excludes the baggage compartment. The value for the normalization of this parameter adopted here is 30 m³. This parameter replaces the passenger capacity used in the composition of the EI_{Liner} .

The efficiency index for three ultra-long-range airplanes was calculated to verify if there is some degree of dependency of the index with the number the engines. The airplanes are the Boeing 747-400, the McDonnell Douglas DC-10-30, and the Boeing 777-200. Table 13 shows the results of the EI calculations and testifies that a three-engine configuration here presents a higher value than a two-engine airplane. An efficiency index comparison for some business jets was performed (Table 14).

Table 13: Efficiency index calculated for long-range airplanes with different motorizations

Airplane	n_e	MTOW [kg]	Single engine thrust [lb]	Two-class seating	Range ^β [nm]	EI
B747-400 [56]	4	362,874	57,000	358 ^α	6,100	1.436
McDonnell Douglas DC-10-30 [57]	3	251,774	51,000	270	5,200	1.212
Boeing 757-200 ^γ [58]	2	115,650	43,500	192	3,272	0.862
B777-200 [22]	2	229,500	71,200	375	3,750	1.050
Boeing 767-300 [59]	2	159,700	57,000	261	3,900	0.996

α The-class configuration: 24 first, 74 business, 302 in economy β Typical seating, MTOW, γ High-gross option

Table 14: Efficiency index calculated for long-range airplanes with different motorizations

Airplane	n_e	MTOW [kg]	Single engine thrust [lb]	Range [nm]/pax	EI _{BJ}
Falcon 900 LX [60]	3	22,225	4,750	4000/8	6.99
G550 [61]	2	41,314	15,385	6750/8	6.10
Falcon 6X [62]	2	35,135	13,460	4900/8	4.26
Legacy 650E [63]	2	24,100	9,020	3800/8	4.85

Most airlines in current operation are fitted with turbofan engines due to some airline requirements. A third central engine like the one from McDonnell Douglas DC-10-30 is difficult to remove for overhaul. Besides, crews during pre-flight inspections must inspect the engines to assure that there are no problems like instance oil or fuel leaks and no visible debris. For an engine located in the vertical tail root, this presents a formidable task. However, another drawback of the trijet transport airplanes is that the wings must be located further aft on the fuselage due to the weight of the central engine that moves the center of gravity afterward. This, compared to twin jets and quad jets with all wing-mounted engines, will require vertical and horizontal tails of larger areas. On the other hand, a factor for sizing the vertical tail is the failure of an engine at takeoff. Considering that the thrust asymmetry of a trijet is lower than that of a twinjet, this will contribute to a smaller vertical tail.

Engines are priced according to the thrust they deliver established by contract clauses. If a trijet configuration could demand a lower overall thrust for a given set of requirements, this would contribute to lowering the manufacturing cost of such configurations. Besides, there is the important benefit of reduced emissions. However, the integration cost for assembling the central engine at the vertical tail root could nullify those benefits.

Unlike transport airplanes fitted with two engines, trijets are not required to land immediately at the nearest suitable airport if one engine fails. This is advantageous if the aircraft is not near one of the operator's maintenance bases, as the pilots may then proceed

with the flight and land at an airport where there are more resources for maintenance. Additionally, trijets on the ground with one engine inoperative can be allowed to perform two-engine ferry flights. Before the introduction of ETOPS, only trijets and quad jets were able to perform long-range international flights over areas without any diversion airports. However, this advantage has largely disappeared since two-engine aircraft were certified for such operations.

Dassault claims that trijet design enhances comfort and performance and allows for a large cabin within a compact fuselage. The tri-engine airplanes present lower critical speeds for safer landings and superior short-field agility. Besides, the French manufacturer adds that trijets are more maneuverable on the tarmac, require less hangar space – and provide greater peace of mind and more direct routing on long routes over water.

Objective 2: direct operating cost

Although it is widely discussed which cost elements do belong to *Direct* Operating Costs and which do not, it is generally accepted that DOC includes those cost elements which depend on the equipment (aircraft) and crew that are necessary to perform a determined flight. *Indirect* Operating Costs (IOC), in contrast, are dependent on the way an airline is administrated [64]. The ATA 67 DOC method [65], a good reference, considers aircraft-dependent and hence part of DOC:

- cockpit crew costs,
- fuel costs,
- maintenance costs,
- depreciation,
- insurance (against hull loss).

Cockpit crew cost is heavily dependent on MTOW, and it is calculated in terms of flight hours. Training costs for the crew or maintenance personnel traditionally are considered as a percentage of the crew's fixed cost. For the engine, it is necessary to provide to the DOC calculation routine the number of engines, the engine weight, and the time between overhauls that are considered.

Figure 15 displays the result of a direct operating cost estimation performed for the hypothetical 76-seat ITA76ADV airliner with a design range of 2100 nm. The calculation considered two jet-fuel prices, US\$ 1.70 and US\$ 2.389, a 40.6% increase. The contribution to DOC increased from 25% to 31%, surpassing the crew slice related to that cost. Thus, as expected, the DOC is indeed sensitive to fuel price.

The values here calculated for the DOC have estimated values and no comparison to actual values practiced by airlines is possible, because they handle this information as strategical and confidential. Nevertheless, it offers a relatively good basis for comparison among the airplanes that are an object of investigation. It is worth mentioning, that changing the fuel price will, as shown in Figure 15, significantly alter the way the airplanes are evaluated, and eventually it could be part of a Pareto front. Ref. [66] provides graph information about the annual variation in jet fuel price.

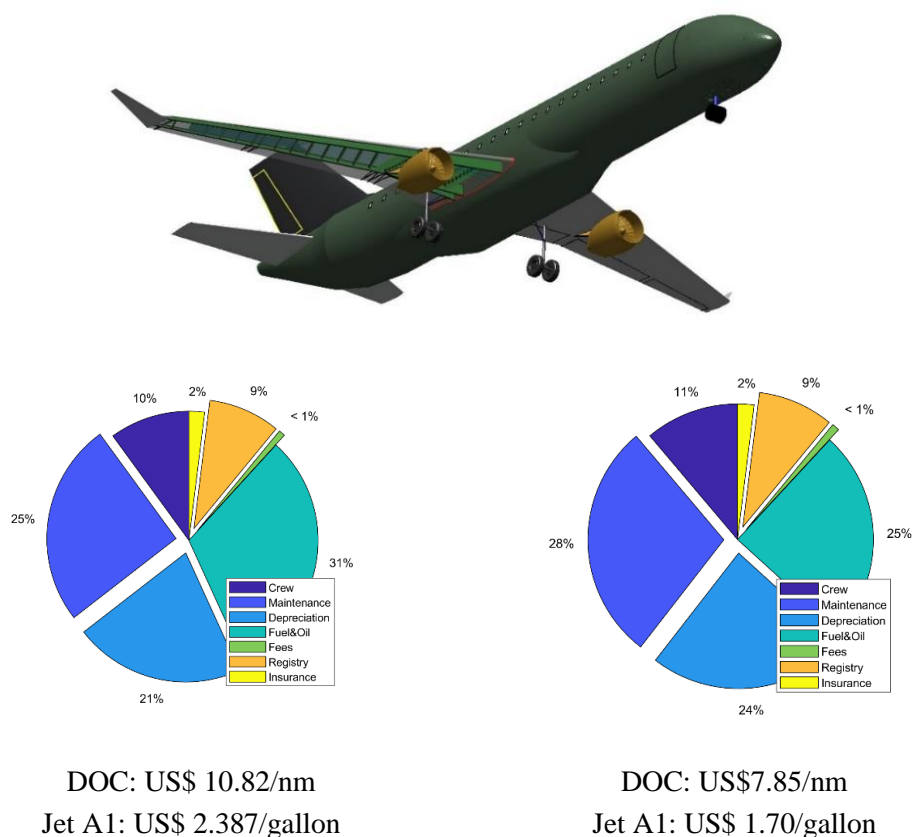


Figure 15: DOC breakdown for the ITA76ADV considering two Jet A1 fuel prices

Optimization tasks

Optimization algorithms are employed for the following tasks:

- To find the weights for airfoil geometries, so that they can be described by an airfoil database. The purpose of this task lies in providing wing geometry input for aerodynamic coefficient calculation by a neural network.
- To obtain the intersection surface between a high wing configuration and the fuselage.
- For optimization of the wing structural layout.
- For sizing of the vertical tail.

For wing airfoil optimization when required by the user.

For all multidisciplinary optimization runs of the present work two objectives were set and will be detailed later. The optimizations can be categorized into two profiles:

Type I. Here, the maximum takeoff thrust is one of 13 design variables. Engine diameter and overall pressure ratio are fixed parameters provided to the airplane calculator program. The variation of thrust and altitude were modeled according to Howe's approach [67]. Four runs were performed. Each one of them had a different classification constraint for the airplane configuration: classic, breakthrough, failure, and niche.

Type II. This is an improvement for airplane modeling when compared to the previous simulations. An engine deck was introduced as part of the aircraft analysis tool package. There are five definition variables plus some additional parameters for engine characterization. Thus, the engine is designed jointly with the airframe. A total of 17 design variables are considered in this simulation.

The boundaries of the variables for the Type I optimization problems are shown in Table 15. The second optimization type utilizes another approach for engine modeling, and it is an improvement for airplane representation when compared to the simple engine model used in the previous type. It is important to emphasize the fact that, all engines in Type I simulations made use of a hypothetical turbofan whose reference TSFC at cruise is 0.55 lb/h/lb. Now, a generic turbofan engine deck was introduced as part of the aircraft analysis tool package to improve aircraft modeling. Its theoretical approach and validation are given in Ref. [40]. A map for compressor efficiency calculation and other improvements and features were added such as engine noise and emission estimation [10]. The present modeling does not contemplate new engine concepts like geared turbofans. There are five variables for the engine and additional parameters for engine characterization, such as the Mach number and altitude of the design point. Thus, the engine is designed jointly with the airframe. A total of 17 design variables were considered in this simulation. The variables and their boundaries of the present optimization problem are depicted in Table 16

Table 15: Boundaries of variables of the optimization problem considering overall maximum thrust as a design variable

Variable	Lower boundary	Upper boundary
Wing reference area [m ²]	120	300
Wing aspect ratio	8.0	10.9
Wing taper ratio	0.240	0.450
Wing quarter-chord sweepback angle	22°	32°
Fuselage length [m]	36	60
Number of seating rows	6	7
Range with typical passenger payload (ISA, no winds)	2100	3700
Height/width of central fuselage [m]	0.90	1.10
Number of engines ^α	2	3
MMO	0.78	0.83
Overall engine thrust @ takeoff [kN]	220	500
Service ceiling [ft]	35,000	42,000
Number of corridors in the passenger cabin	1	2

α the trijet configurations have two underwing engines and a third one located on the vertical tail root

Table 16: Boundaries of variables for Type II optimization problem

Variable	Lower boundary	Upper boundary
Wing reference area [m ²]	120	300
Wing aspect ratio	8.	10.9
Wing taper ratio	0.240	0.450
Wing quarter-chord sweepback angle	22°	32°
Two-class passenger accommodation	150	280
Number of seating rows	6	7
Range with typical passenger payload (ISA, no winds)	2300	4000
Height/width of central fuselage [m]	0.90	1.10
Number of engines	2	2
Service ceiling [ft]	35,000	42,000
Number of corridors in the passenger cabin	1	2
Engine by-pass ratio	4.4	6.2
Engine diameter [m]	1.50	2.80
OPR	26:1	33:1
FPR	1.20	1.83
Turbine inlet temperature [K]	1250	1455

Some aircraft parameters that were set for all simulations simulation are:

- Year of service entry: 2017.
- Engine TBO of 3000 h.
- Container to be accommodated in cargo compartment: LD3-45W.
- The relative maximum thickness of root airfoil: 15.5%.
- The relative maximum thickness of root airfoil: 11.0%.
- Passenger cabin height of 2 m.
- Y-class seat width of 0.46 m.
- Y-class seat pitch of 0.8128 m.
- Business-class seat pitch of 0.9144 m.
- Aisle width of 0.48 m.
- Vertical tail volume coefficient of 0.09.
- Horizontal tail volume coefficient of 1.
- Fuselage tank in front of the wing leading-edge extension.
- Takeoff field length lower than 2400 m (MTOW, ISA, sea level).
- Must comply with FAR requirements for 2nd segment climb.
- Must comply with FAR121 requirements for balked landing.
- Must comply with FAR119 requirements.
- Enough thrust to cruise at the prescribed altitude.
- Jet A1 fuel price US\$ 2.387 US\$/gal.
- Location of the aft wing spar at 72% of chord.
- Takeoff flap deflection of 25°.
- Landing flap deflection of 40°.
- Long-range cruise profile.
- No emissions or noise constraints.

3. Entropy-statistics classification as a constraint in an optimization problem

Type I and II optimization runs were carried out here. For Type II, the classification constraint classic was the only one used just to compare its results with the related Type I simulation.

The Type I simulation additional parameters are:

- Three types of product classification: classic, niche, and failure.
- Engine by-pass ratio of 6.
- Engine external diameter of 1.715 m.
- Engine overall pressure ratio of 32.2.
- TSFC at long-range cruise of 0.55 lb/h/lb.

Although attempted, no feasible configuration emerged from the Type I simulation with the airplane classification set as a breakthrough. As already explained in the preceding paragraphs, this result is due to the year of service entry chosen for the present simulations, namely 2017, the airplanes present in the database, and the limitations of the aircraft modeling employed here.

No engine deck was employed in the simulations where the overall thrust was provided. The engine parameters here were used to estimate the engine length, and therefore the engine weight, which is necessary for the DOC calculation, and the wetted area, which is important for drag calculation.

3.1 Multi-objective optimization with overall thrust as a design variable

Classic designs

A multi-objective optimization design run was carried out where the constraint for airplane classification by entropy statistics was set as classic. After 50 generations with the MATLAB® multi-objective genetic algorithm, the number of feasible individuals is 1980 – which is the number that does not violate any constraint (Figure 16). Some configurations with two aisles and presenting six or seven seating rows emerged in the computation and were categorized as classic but none with seven rows or two corridors are part of the Pareto front. In addition, all feasible airplanes that emerged from the simulation are fitted with two engines. The configurations with two aisles are the heaviest and are expected to present a higher DOC. The main advantage of this configuration is to propitiate more comfort to the passenger. However, the present simulation also allows for more fuel capacity of the tank located in the fuselage.

This is mainly due to the engine characteristics that were chosen for the optimization tasks. Those characteristics are closer to modern and fuel-efficient engines in the category of the A320 and B737 airplanes. The engine that was then selected, presents a fuel consumption at long-range cruise conditions approximately 8% lower than the one of IAE V2500-A5. The departure of the A320 basic design was minimal among the airplanes with the same range and capacity as of A320-200.

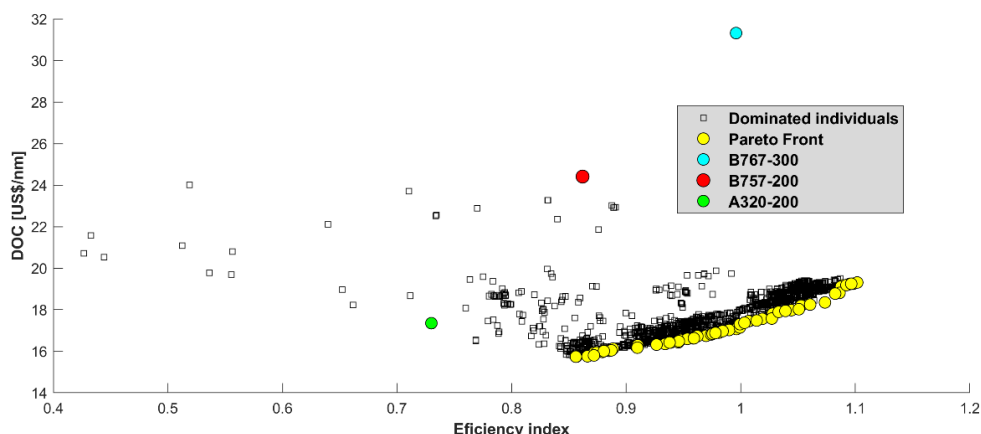


Figure 16: Pareto front for classic designs. Calculated values by the present design framework for B757 and A320 are plotted for comparison purposes, only

As was performed for A320-200, the efficiency index and DOC values for B757-200 are displayed in Figure 16 for comparison purposes, only. Table 17 contains some featured individuals from the Pareto front, characterized for EI varying from 0.856 to 1.102. For the same efficiency index, the B757-200 airliner has a DOC US\$ 5.1/nm higher than the one from ID 1845C, featuring the highest DOC from the configuration shown in Table 17.

Figure 17 shows some fuselage cross-sections from the individuals presenting the lowest and the highest DOCs and now a significant difference between them becomes evident. Figure 18 shows the top view for two individuals from the Pareto front featuring the lowest and highest DOCs. The geometric characteristics of ID 1031C are remarkably like those from A320-200 but with the exception that they present a design range of 3,005 nm which is considerably higher.

Table 17: Some individuals belonging to the Pareto front of the classic aircraft optimization

ID	MTOW [kg]	OEW [kg]	Design range [nm]	Total thrust [kN]	NPAX	Aisle	Seating rows	DOC	EI
1031C	74,000	42,092	3,005	225	140	1	6	15.74	0.856
1963C	73,975	41,666	3,006	225	140	1	6	15.80	0.872
1138C	82,222	45,843	3,124	242	162	1	6	16.97	0.984
1704C	85,997	47,983	3,177	245	172	1	6	17.56	1.019
1868C	90,380	50,153	3,274	263	180	1	6	18.19	1.053
1738C	95,759	52,637	3,392	279	192	1	6	19.17	1.089
1845C	96,576	52,867	3,415	281	192	1	6	19.31	1.102

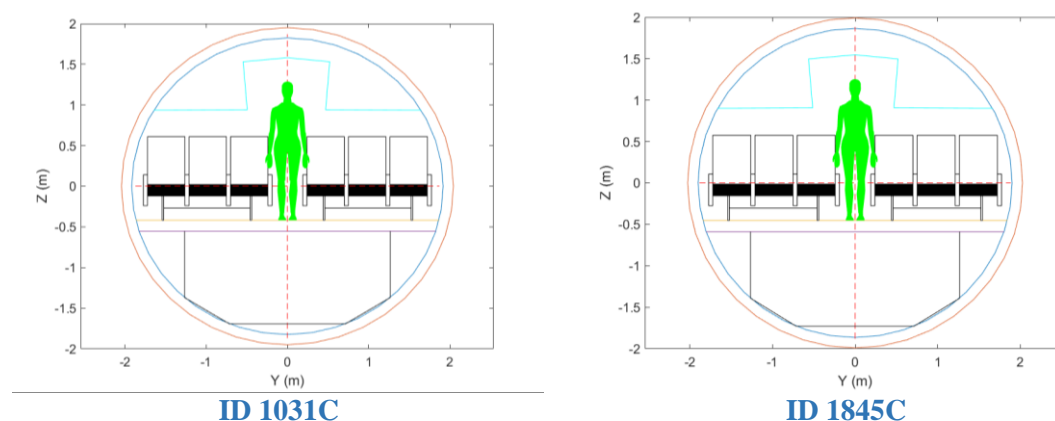


Figure 17: Cross-sections of some individuals in the Pareto front from the optimization run for classic airplanes

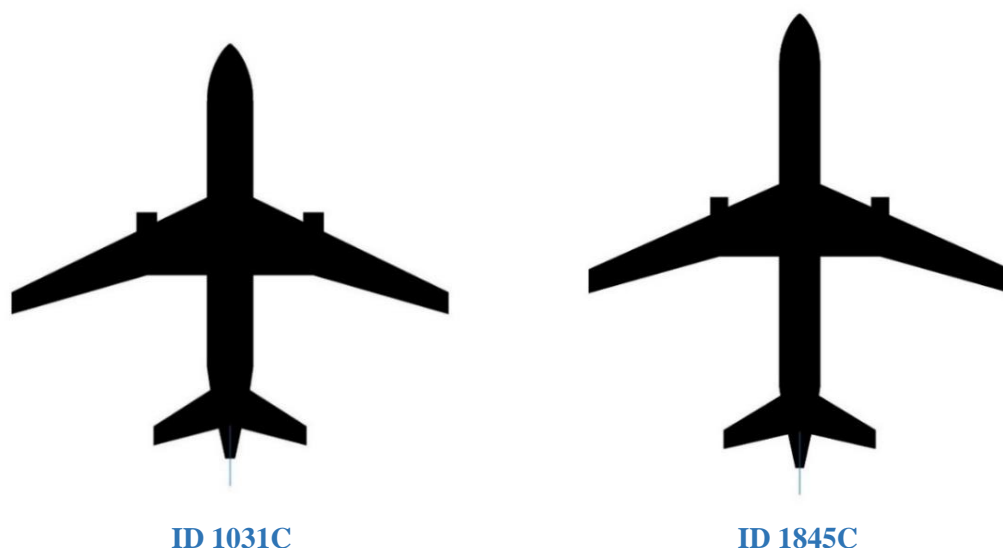


Figure 18: Top view of some individuals in the Pareto front of classic design simulation (not on the same scale)

Figure 19 shows the DOC divided by passenger accommodation for all individuals from the simulation. There is an almost linear relationship between the efficiency and the DOC per passenger. All individuals in the Pareto front present the DOC per passenger lower than that registered by the B757-200.

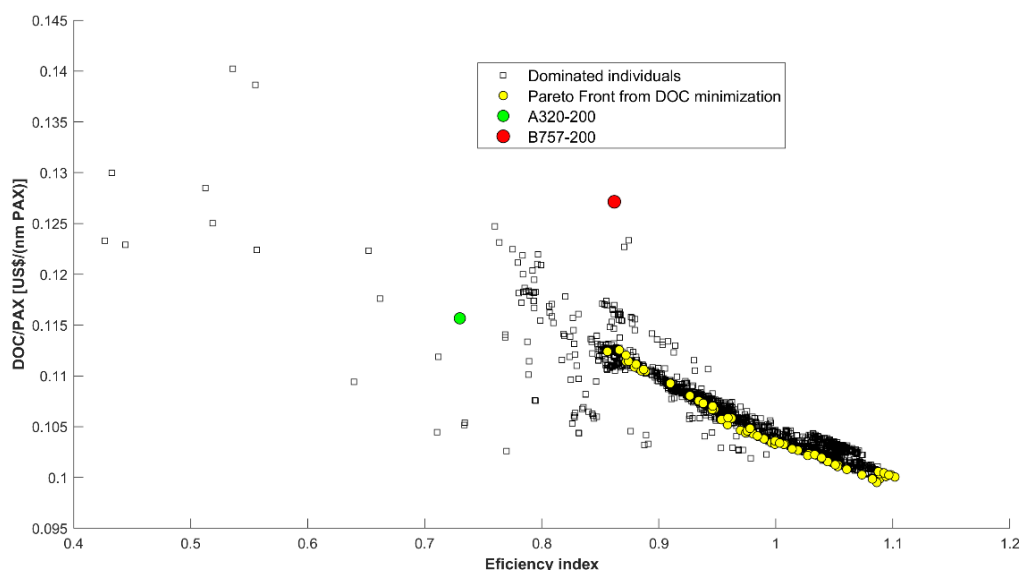


Figure 19: DOC per passenger two-class capacity plotted against efficiency index for all feasible airplanes from the optimization run

To verify which individuals from the Pareto front are closer to B757-200, a Dendrogram approach was elaborated. This time the variables used to build up the distance matrix are related to configuration geometric characteristics and performance parameters, but not to weights and engines. The reason for this is because the point of interest is to find out an airplane with higher efficiency index and lower DOC that is related to a configuration with lower MTOW and OEW.

The variables that were used to compose the dendrogram are the following:

- Wing area
- Wing aspect ratio
- Wing quarter-chord sweepback angle
- L_f
- Design range
- Fuselage height-to-width ratio
- MMO
- TW
- Wing taper ratio
- Service ceiling
- Two-class passenger accommodation
- Number of aisles
- Engine configuration
- Seating rows in Y-class

Figure 20 shows the resulting Dendrogram for the four closest individuals to the B757-200 configuration.

Table 18 contains a comparison among some characteristics of the closest individuals to B757-200. The configuration of ID 1785C (Figure 21) was found to be the most resemblance to the B757-200. It surpasses the veteran airliner in all aspects: lower DOC, 20 t lower MTOW, and considerably lower OEW for the same passenger capacity and almost the same range. MMO of 0.82 is lower than that of B757-200.

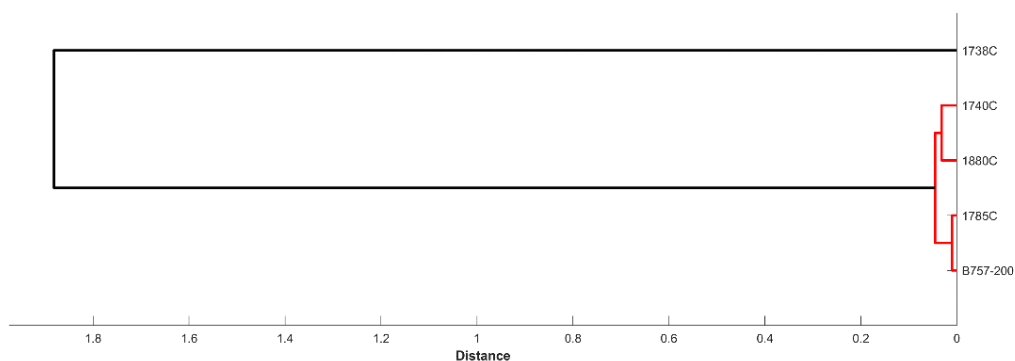


Figure 20: Dendrogram of the closest configurations to B757-200

Table 18: Closest individuals to B757-200

ID	wS [m ²]	wAR	MTOW [kg]	OEW [kg]	Design range [nm]	Thrust [kN]	NPAX	DOC [US\$/nm]	EI
1740C	167.4	10.5 ₃	96,067	52,754	3,397	279	192	19.22	1.094
1880C	168.6	10.3 ₂	95,743	52,486	3,399	279	192	19.19	1.092
1785C	167.4	10.3 ₅	95,253	52,276	3,387	277	190	19.11	1.088
B757-200	185.25	7.8	115,650	58,570	3,476	351	192	24.41	1.014

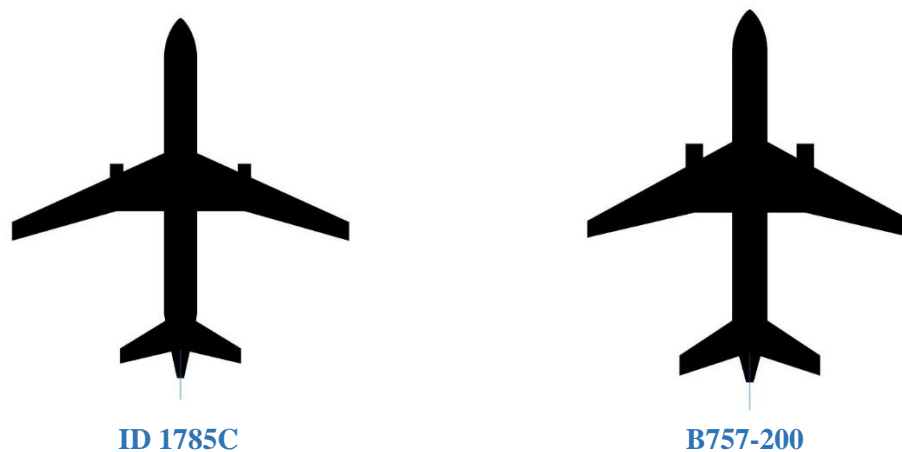


Figure 21: Top view of B757-200 and its closest potential replacement that emerged from optimization run for classic airliners (not on the same scale).

Niche designs

The optimization run with a constraint stipulating that the airplanes must obey the classification as the niche was carried out and was stopped after 25 generations. Each generation was populated with 55 individuals. Just 693 individuals were revealed to be feasible, and 40 of them are part of the resulting Pareto front. Some characteristics of airplanes found in the Pareto front are:

- All are fitted with 3 engines.
- All feature 2 aisles and a T-tail.
- Two-class passenger capacity in the 177-236 range.
- DOC is equal to or higher than 17.88 US\$/nm.

Figure 22 shows the objective values of the feasible individuals from the optimization run for niche airplanes. The niche airplanes belong to the Pareto front individuals from the optimization with the classification constraint *niche*. Table 19 contains characteristics of some individuals from the Pareto front sorted by ascending DOC. Considerably higher EI was found than those that emerged from the classic airplane simulation.

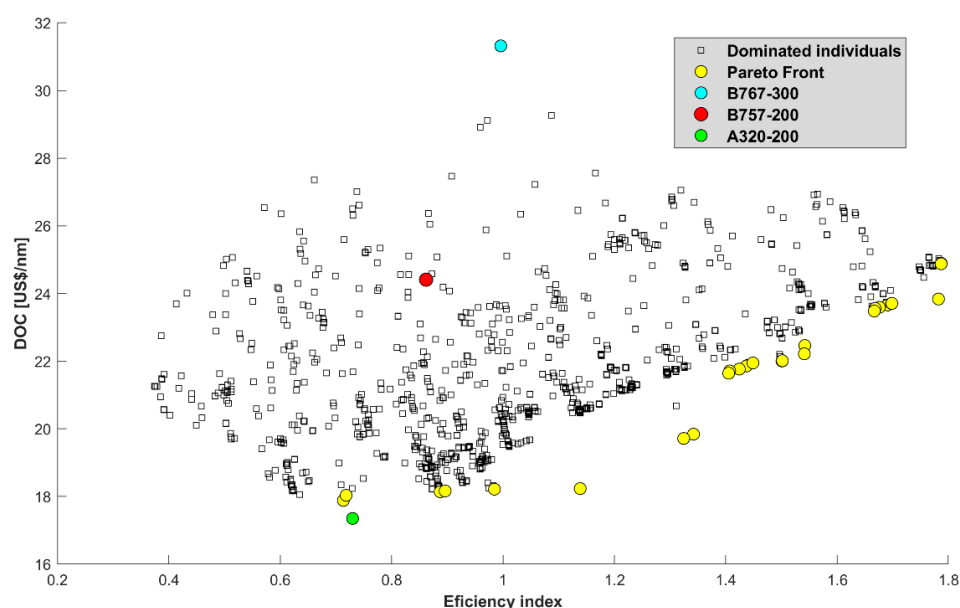
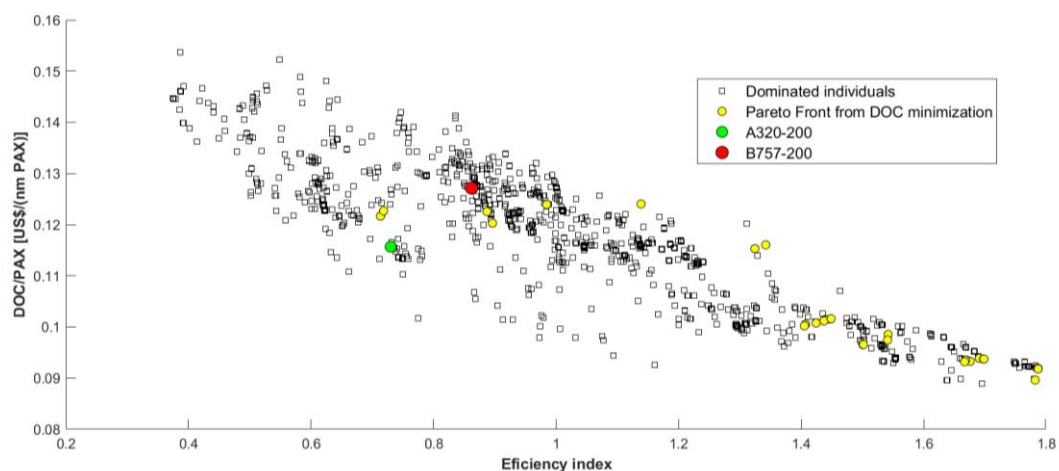


Figure 22: Optimization considering niche airplanes

Table 19: Individuals belonging to Pareto front of optimization for niche aircraft

ID	MTOW [kg]	OEW [kg]	Range [nm]	T [kN]	NPAX	Rows	DOC [US\$/nm]	EI
1393N	81,676	48,715	2,708	238	150	6	17.88	0.713
1025N	89,217	49,269	3,632	286	148	6	18.14	0.887
1356N	90,155	49,269	3,632	286	150	6	18.16	0.896
1321N	116,630	63,679	3897	286	216	6	21.65	1.405
986N	116,450	62,165	3,897	286	216	6	21.86	1.437
1132N	129,020	63,045	3,897	286	228	6	22.46	1.542
1074N	129,540	68,579	3,897	286	252	7	23.71	1.698
1332N	138,400	74,557	3,897	286	273	7	24.89	1.787
Boeing 757-200 [27]	115,650	58,570	3,272	387	192	6		
Boeing 767-200 [59]	136,078	80,286	2,750	454	216	7		

Figure 23 shows the DOC divided by passenger accommodation for all individuals from the simulation. Contrary to the previous simulation, there is not a direct relationship between the efficiency index and the DOC per passenger. All individuals in the Pareto front present the DOC per passenger lower than the one registered by the B757-200. However, the individuals from the Pareto front with EI higher than 1.4 present better figures and should be chosen for further analysis by a decision-making team.

**Figure 23: DOC per passenger two-class capacity plotted against efficiency index for all feasible airplanes from the optimization run with niche constraint**

In terms of capacity, the ID 1321N configuration is the one closer to B757-200 with the lowest DOC. The range with 216 passengers is 3,897 nm and for the reference capacity of 192 passengers, it will fly farther than 4,000 nm. Its maximum engine thrust is also considerably lower than that of B757-200 fitted with RB2111-535E4 engines (Table 11), but the wing reference area is higher, 228.41 m². Cross-sections of three individuals from the Pareto front are compared in Figure 24 and Figure 25 displays their top views. It is important to stress that airplane classification by entropy statistics depends on the year of the aircraft's entry into service. Considering the airplanes present in the database, there are no three-engine aircraft in the timeframe of 5 years around 2017. For this reason, airplanes with this configuration will be considered anything other than classic. No breakthrough design emerged from the optimization simulation with this constraint. Thus, there are only three simulations with optimal configurations: classic, niche, and failure.

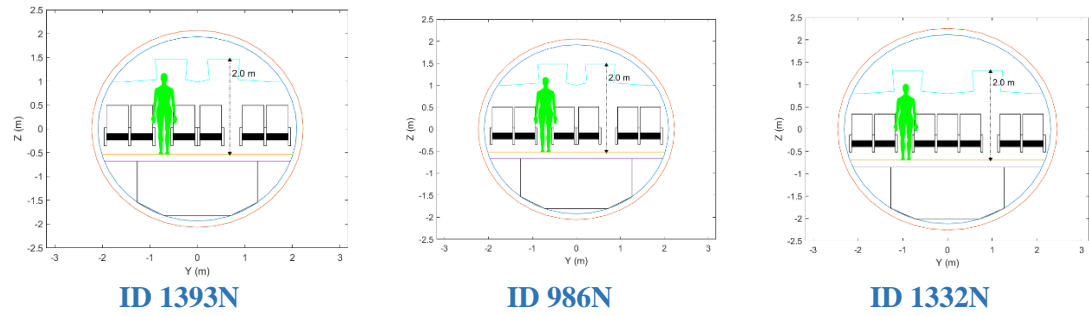


Figure 24 – Cross-sections of some individuals in the Pareto front from the optimization run for niche airplanes



Figure 25: Top view of some individuals in the Pareto front of niche design simulation (not on the same scale)

Failure aircraft

All failure designs that emerged in the simulation in the Pareto front (Figure 26) share the following characteristics:

- Six rows of seats
- Single aisle.
- Three engines and a T-tail.

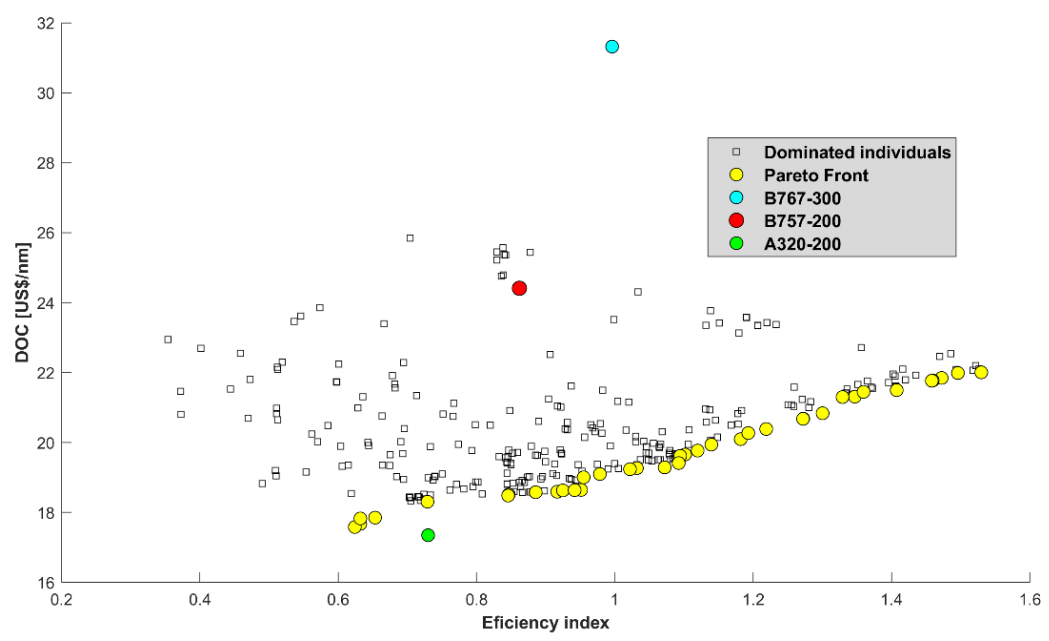


Figure 26: Optimization considering failure airplanes

For the same EI, Airbus A320-200 presents a DOC lower than its counterpart in the Pareto front. Table 20 contains some characteristics of selected individuals from the Pareto front. There are airplanes in the 148-234 range. The airplane with the same capacity as the B757-200 reference aircraft, carrying the ID 241F, presents higher efficiency and lower DOC in comparison. The top view of some optimal configurations is shown in Figure 27.

Table 20 – Individuals belonging to Pareto front of optimization for failure aircraft

ID	n_c	MTOW [kg]	OEW [kg]	Range [nm]	T [kN]	NPAX	Aisle	Rows	DOC [US\$/nm]	EI
320F	3	78,552	45,879	2,817	286	144	1	6	17.59	0.624
290F	3	90,944	50,070	3,473	289	168	1	6	18.64	0.42
163F	3	91,703	49,928	3,713	262	162	1	6	18.65	0.951
200F	3	97,035	52,718	3,714	293	180	1	6	19.29	1.072
241F	3	100,000	53,201	3,714	294	192	1	6	19.94	1.139
362F	3	116,900	63,593	3,713	262	234	1	6	22.01	1.530

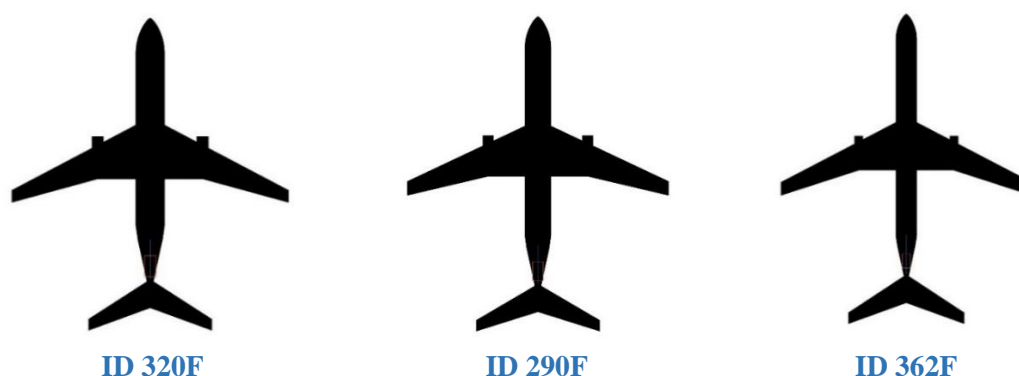


Figure 27: Top view of some individuals in the Pareto front of failure design simulation (not on the same scale)

3.2 Optimization of airframe engine design concomitantly

The Pareto front is shown in Figure 28, with the EI and DOC of reference airplanes for comparison purposes. Table 21 contains the characteristics of some relevant individuals in the Pareto front. Further selections of individuals of Pareto front that present passenger capacity like the one of Boeing 757-200 is given in Table 22. All of them are fitted with engines featuring by-pass ratios of 6:1, presenting ranges slightly longer than the Boeing-757-200, with considerably lower MTOW, as well as lower engine maximum thrust. Figure 29 displays some top views of three configurations from the Pareto front.

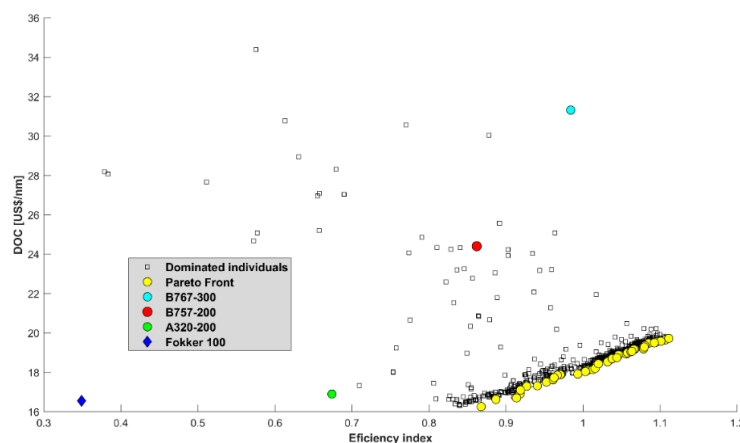


Figure 28: Optimization run considering a turbofan engine generic deck

Table 21: Individuals belonging to Pareto front of optimization (Engine deck)

ID	n_e	MTOW [kg]	OEW [kg]	Range [nm]	T [kN]	NPAX	Aisle	Rows	DOC [US\$/nm]	EI
470CE	2	75,872	43,119	3,027	215.74	142	1	6	16.26	0.868
727CE	2	79,055	44,514	3,090	225.93	150	1	6	16.70	0.914
662CE	2	86,236	47,504	3,171	247.40	168	1	6	17.92	0.993
755CE	2	90,113	49,332	3,282	261.41	174	1	6	18.44	1.020
744CE	2	93,752	50,838	3,360	271.33	180	1	6	19.05	1.063
725CE	2	96,040	51,836	3,403	278.95	186	1	6	19.37	1.080
435CE	2	98,125	52,729	3,421	285.67	192	1	6	19.66	1.107
731CE	2	98,600	52,934	3,431	287.22	192	1	6	19.73	1.112

Table 22: Optimal designs with passenger capacity and range similar to B757-200

ID	wS [m ²]	wAR	wTR	MTOW [kg]	OEW [kg]	Design range [nm]	Thrust [kN]	NPAX	DOC [US\$/nm]	EI
731CE	176.4	9.94	0.295	98,600	52,934	3,431	287.2	192	19.73	1.112
726CE	170.85	10.08	0.309	95,657	51,799	3,351	275.9	186	19.29	1.080
303CE	175.5	9.97	0.296	98,124	52,729	3,420	285.6	192	19.67	1.107
B757-200	185.25	7.8	0.243	115,650	58,570	3,476	387	192	24.41	1.014
ID	wSw14	OPR	BPR	Engine Diameter [m]						
731CE	22.26	32.18:1	6.02:1	1.93						
726CE	23.14	32.42:1	5.96:1	1.89						
303CE	22.29	32.21:1	6.00:1	1.93						
B757-200	25.0	25:1	4.40:1	2.20						

**Figure 29: Top view of some individuals in the Pareto front of classic design simulation (not on the same scale)**

3.3 Featured choices

The main difference between classic and failure configurations concerns their number of engines, both present configurations with 6 seating rows. The niche and failure designs share the same number of engines, which is three. The characteristic that distinguishes the two groups is the number of corridors: two for the niche airplanes; one for those considered a failure. The Airbus A350 is present in the airplane database in the timeframe considered and this airplane has two corridors in its cabin layout. This characteristic was transmitted to some aircraft, and they could then be classified as niche. No airplane in the database, in the timeframe around 2017 that was set for the optimization runs, has three engines combined with a single corridor and this was enough to characterize the failure airplanes.

Figure 30 on the left contains all Pareto fronts of the four optimization runs. The Pareto front of the failure aircraft does not intercept any of the classic ones, intercepting the niche aircraft at approximately an EI of 1.42. The two optimizations for the classic airplanes are remarkably close to one another. The result observed here is that the niche

configuration 1332N has the best efficiency index among the individuals of the three Pareto fronts. Due to this and its reasonable DOC for its category, it deserves a deeper analysis. This airplane is fitted with three engines. It was classified as a niche because there were no tri-engine airplanes before and after it, considering the timeframe of 5 years utilized by the classifier algorithm. The DOC takes into consideration the number of engines and their thrust. The 2nd segment climb requirement is very stringent for two-engine airplanes. Thus, depending on aerodynamics and weight, a three-engine airplane may present a lower thrust-to-weight ratio. The tradeoff between TW and the number of engines seems to be more favorable to configurations with three engines in some cases.

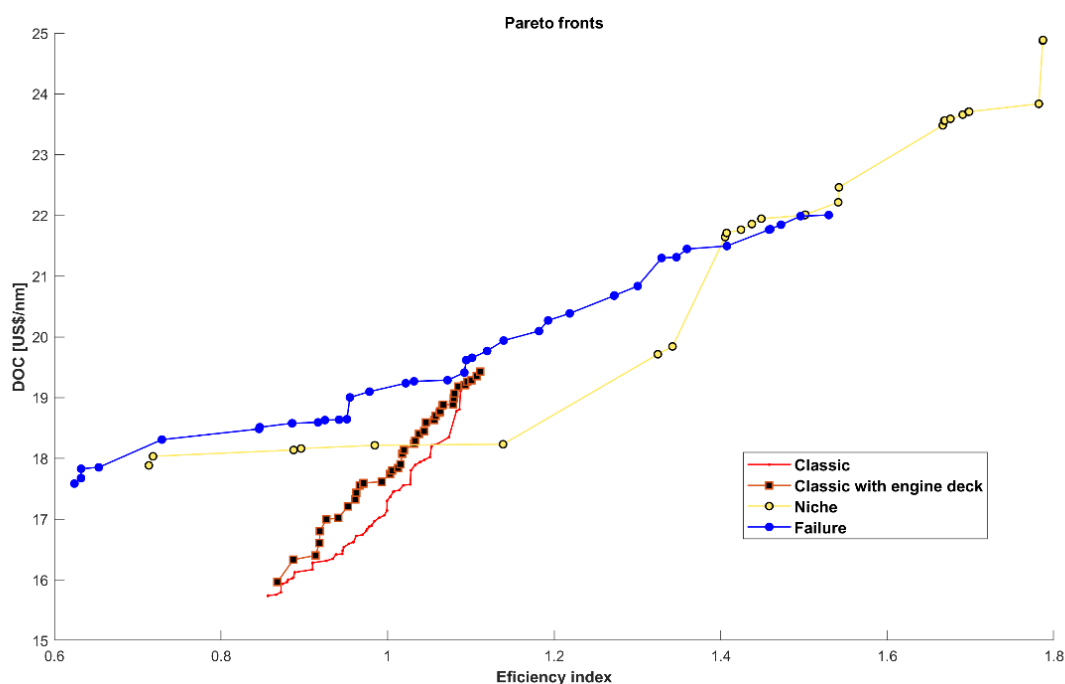


Figure 30: Plot combining all Pareto fronts from the present simulations

According to the simulation results, there are two choices for the possible fulfillment of MMA:

- Classic design: the ID 303CE is an extraordinarily strong candidate due to its considerably lower MTOW, OEW, and lower maximum takeoff thrust when compared to the B757-200 reference airplane. Its configuration is like the one of ID1845C.
- Niche design: 1332N is also a strong candidate due to its impressive efficiency index enabling a great range and passenger capacity. It will fulfill another part of the market for MMA. Niche candidates with EI above 1.4 present a better Doc per passenger (Fig. 25).

4. Robust optimization of a chosen individual of a previous optimization

This optimization task considers the fuel price ranging from 0.887 to 4.387 US\$/nm and its impact on DOC for a mission of 3420 nm with 192 pax. The objectives are just the average DOC and its standard deviation. The intent is to evaluate whether an airplane configuration resilient to fuel price variation can be derived. An improved ANN system based on full potential formulation with viscous correction was employed in this simulation [68]. This time, the root, and the break station airfoils are part of the design effort. Table 23 presents the design variables and their boundaries for the present design task.

Table 23: Design variables and their boundaries

Variable	Lower boundary	Upper boundary
Wing reference area [m ²]	165	195
Wing aspect ratio	9	9.90
Wing taper ratio	0.240	0.380
Wing quarter-chord sweepback angle	22°	28°
Engine diameter [m]	1.62	2.20
Engine by-pass ratio	5.0	6.2
OPR	25:1	33:1
FPR	1.25	1.80
Turbine inlet temperature [K]	1250	1455
Root airfoil weights (14 in total)	0	1
Break station airfoil weights (14 in total)	0	1

By a slight reduction of the wing aspect ratio of the Design ID 303CE (Table 22), keeping the remained parameters but the root airfoil section of the wing unchanged, a baseline airplane was created for this last optimization task (Table 24). Mattos et al. [69] applied a flutter speed constraint to an MDO framework for airliner conceptual design. As optimization results, a noticeable reduction of wing aspect ratio was observed when compared with an optimization where no flutter speed constraint was applied.

Table 24: Baseline airplane for the robust optimization

MTOW [kg]	99277
OEW [kg]	53282
Fuel capacity [kg]	31121
wS [m²]	175.45
wAR	9.90
wSw14	22.29°
BPR	6.00
OPR	32.21
Avg DOC [US\$/nm]	20.960
Range with 192 pax at FL390 [nm]	3,420
DOC Standard deviation [US\$/nm]	2.9986

The optimization task was stopped after 35 generations and slightly over 1300 feasible configurations emerged from the simulation. Figure 31 shows the relationship between MTOW and DOC for all feasible individuals. A linear correlation between the two variables can easily be recognized. Figure 32 shows the objective diagram containing optimal airplanes, feasible individuals, the baseline airplane, and the figures obtained in the present work for the B757-200 airliner. There are two individuals in the Pareto front with similar characteristics. This means a high degree of dependence on the two objectives chosen for this optimization task. Table 25 contains major information regarding one of the optimal airplanes. Its MTOW is 1,760 kg lower than that of the baseline airplane. This represents slightly over 18 t less than the Boeing 757-200 MTOW. All this was achieved thanks to the airfoil design and an engine featuring a higher by-pass ratio, with no “advanced” technology being incorporated into the configuration.

Regarding the baseline aircraft, the average DOC was decreased by 0.473 US\$/nm. For a fleet comprised of 10 airplanes performing two flights a day operating 300 days/year, this will represent a cost saving of US\$ 9.7 million. A direct comparison between the optimal and reference airfoil for the tip station is shown in Figure 33. Figure 34 shows the probability density function that resulted from all feasible airplanes analyzed in the robust optimization simulation.

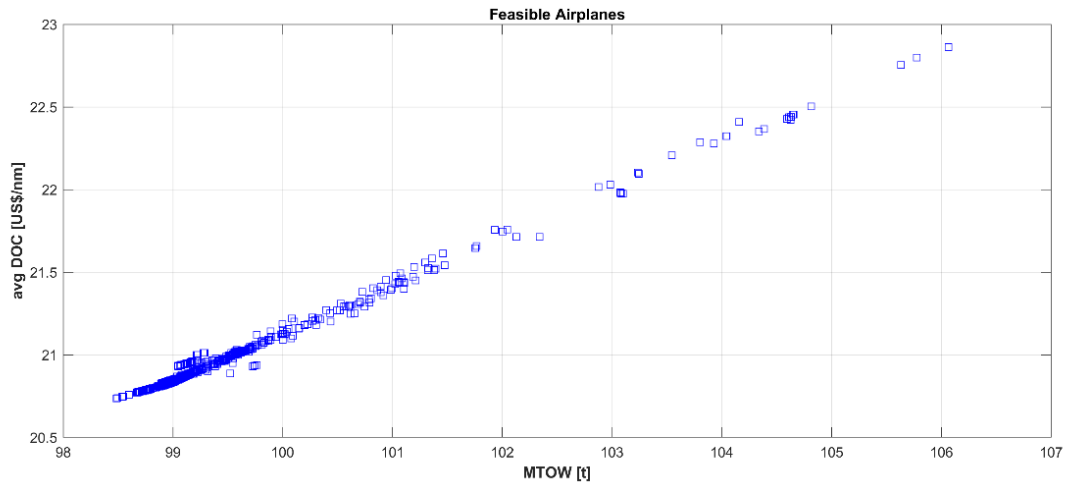


Figure 31: Linear dependence between DOC and MTOW

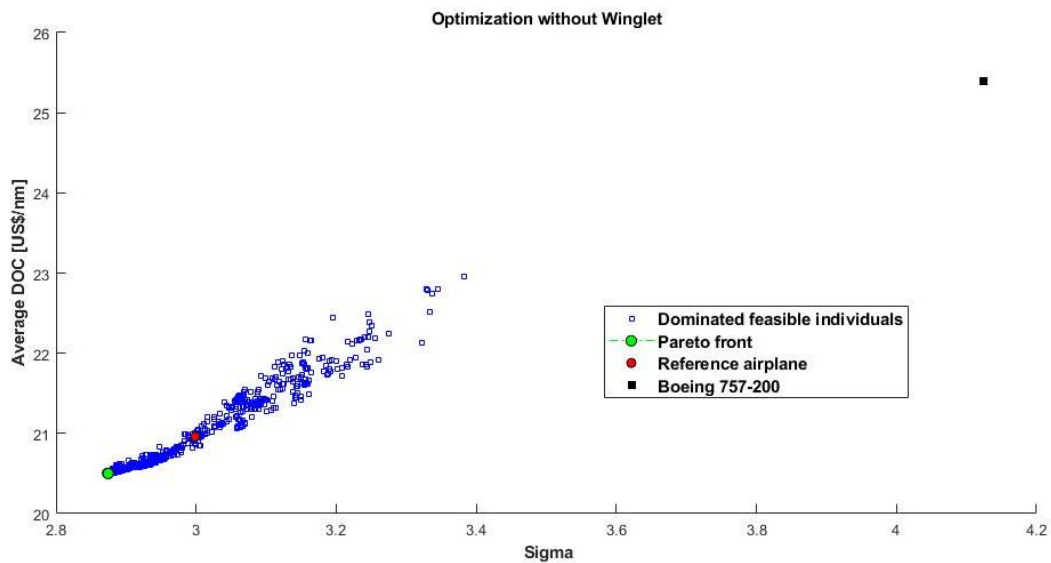


Figure 32: Objectives and feasible individuals

Table 25: Optimal airplane characteristics

MTOW [kg]	97517
OEW [kg]	52637
Fuel capacity [kg]	32118
wS [m²]	175.51
wAR	9.886
wSw14	22.04°
Engine diameter [m]	1.925
BPR	6.19
OPR	32.60
Avg DOC [US\$/nm]	20.487
DOC Standard deviation [US\$/nm]	2.8724

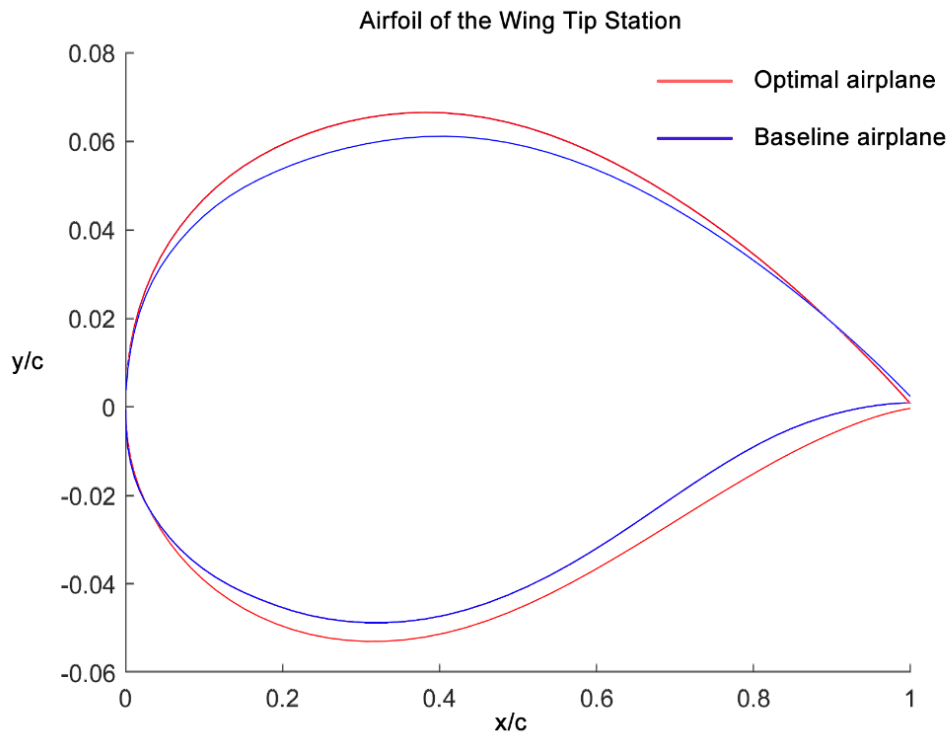


Figure 33: Comparison between optimal and reference airfoil of the wing tip station

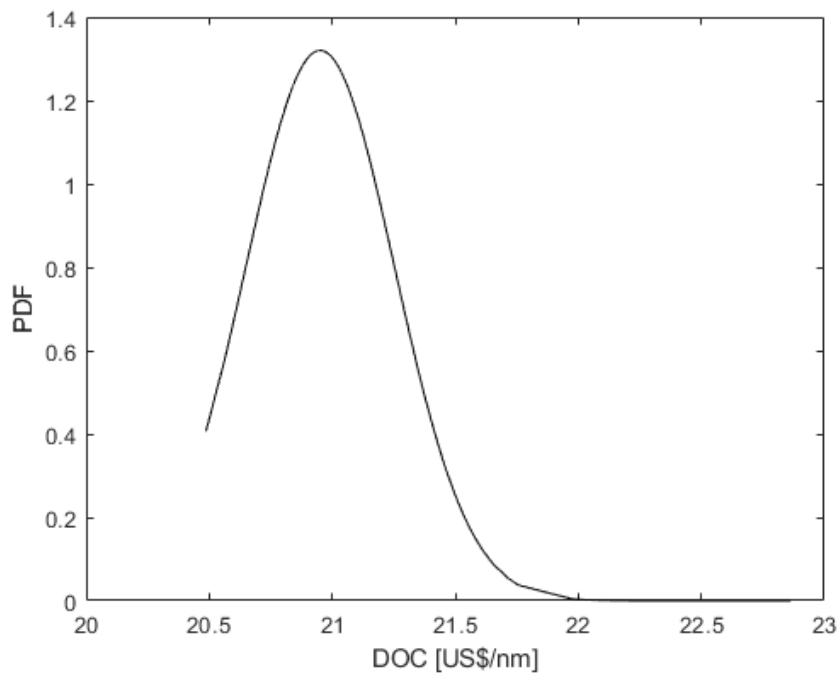


Figure 34: Probability density function of the average DOC

5. Conclusions

An unsupervised learning classification computational tool based on entropy statistics was implemented. It was applied to jet airliners that saw service since the 1950s and proved to be reasonable in most cases. Although correct according to the fundamentals of diffusion and convergence coefficients, the resulting classification of the Boeing 777-300ER airliner indicates that the classification provided by that tool must be confronted with other methodologies of different nature to provide an ultimate label for a product.

The entropy statistics variables that were employed here for a product definition are a considerable improvement from those utilized by Frenken in his inspirational work. The variables of the present work are related to propulsion, airplane size and capacity, airplane performance, and some configuration topology. Certainly, additional variables are needed to map the technological evolution of airplanes, variables linked to navigation, structures, and hydraulics systems, among many others. Another important observation refers to the importance to represent all airplanes that entered service in the database. The omission of a single airplane or its versions may lead to incorrect labeling of other aircraft, noticeably those that saw service short before and after it.

Optimization tasks were successfully carried out to find out high-efficiency airliners presenting lower DOC than the ones related to existing airplanes, some of them being best sellers but whose basic designs are dated back decades ago. The task was focused on the middle of the market segment. An unsupervised learning algorithm based on entropy statistics was utilized for product classification, comprising four categories, classic, niche, failure, and breakthrough designs.

A classic and a niche design were selected for an MMA, with 192 and 273 passengers accommodated into two classes, respectively.

The classic optimal airplane selected is closer to Boeing 757-200 fitted with RB211-535E4B engines and the passenger cabin layout with a single aisle and 6 seating rows. Its engine by-pass ratio of 6:1 is considerably higher than the one for the B757. For the same mission of the out-of-production MMA airliner, the classic design has an MTOW of approximately 17,500 kg lower than B757-200. The present work does not consider advanced technology such as geared turbofan engines, laminar flow, advanced wing airfoils, and others. Thus, the potential for further improvement is huge.

The niche aircraft features 7 seating rows and three engines. It is closer to a B767-300 configuration in terms of capacity and performance. An analysis of the advantages of a three-engine transport airplane was carried out in the present work.

Future work will consider tail-surface sizing by utilizing stability and controllability criteria, turbofan engines with water injection for lower fuel consumption and NOx emission, and fuselage sizing incorporation ditching regulations. Considering that the DOC will be demanding even more computational effort, Bayesian optimization becomes a more efficient way to handle the optimization problems that were carried out in the present work with genetic algorithms.

References

- [1] World Bank, "Air Transport, passenger carried," World Bank, October 2019. [Online]. Available: <https://data.worldbank.org/indicator/IS.AIR.PSGR?end=2018&start=2004&year=2016>. [Accessed October 2019].
- [2] G. Rich, "Boeing's 787 Has Realigned Air Travel; These 5 Routes Show How," Investor's Business Daily, 2018. [Online]. Available: <https://www.investors.com/news/boeing-787-new-routes-air-travel-realignment-airbus/>. [Accessed April 2021].
- [3] International Air Transport Association (IATA), "Aviation Technology Roadmap," IATA, Montreal, Canada, 2013.
- [4] Aviation Week, "Boeing's New Midsize Airplane: Low Development Cost, Price Are Key," March 2018. [Online]. Available: <https://aviationweek.com/air-transport/boeings-new-midsize-airplane-low-development-cost-price-are-key>. [Accessed 09 03 2021].
- [5] Great Circle Mapper, "Great Circle Distances Calculator," [Online]. Available: <http://www.gcmap.com/>. [Accessed 09 03 2021].
- [6] E. Torenbeek, *Advanced Aircraft Design*, Wiley & Sons, 2013.
- [7] L. K. Loftin, *Subsonic aircraft: Evolution and the matching of size to performance* (NASA reference publication), National Technical Information Service, 1980, p. 441.
- [8] Y. Sun and H. Smith, "Supersonic Business Jet Conceptual Design in a Multidisciplinary Design Analysis Optimization Environment," in *2018 AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, Kissimmee, Florida, 2018.
- [9] K. Frenken and L. Leydesdorff, "Scaling Trajectories in Civil Aviation (1913-1997)," *Research Policy*, vol. 29, pp. 331-348, 1999.
- [10] B. S. Mattos, J. A. Fregnani and P. C. Magalhães, *Conceptual Design of Green Transport Airplanes*, Sharjah, UAE: Bentham Books, 2018.
- [11] K. Frenken, "Understanding product innovation using Complex Systems Theory," University of Amsterdam, Amsterdam, 2001.
- [12] Mathworks, Inc., "gamultiobj," Mathworks, Inc, 2018. [Online]. Available: <https://www.mathworks.com/help/gads/gamultiobj.html>. [Accessed August 2019].
- [13] Wikipedia, The Free Encyclopedia, "Entropy (statistical thermodynamics)," Wikipedia, The Free Encyclopedia, September 2022. [Online]. Available: [https://en.wikipedia.org/wiki/Entropy_\(statistical_thermodynamics\)](https://en.wikipedia.org/wiki/Entropy_(statistical_thermodynamics)). [Accessed September 2022].
- [14] S. Kullback and R. A. Leibler, "On Information and Sufficiency," *Annals of Mathematical Statistics*, vol. 22, pp. 79-86, 1951.
- [15] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, New York: Wiley&Sons, 1991.
- [16] Wikipedia, The Free Wikipedia, "Bregman divergence," Wikipedia, The Free Wikipedia, 2019. [Online]. Available: https://en.wikipedia.org/wiki/Bregman_divergence. [Accessed April 2019].

- [17] B. S. Mattos, J. A. Fregnani and J. T. Tomita, "Effects of the 1978 Airline deregulation Act on Aircraft Industry Measured by Entropy Statistics," *Transportes*, pp. 1-19, 2022.
- [18] Airfleets.Net, "Aircraft Production List," Airfleets.Net, 2019. [Online]. Available: <https://www.airfleets.net/home/>. [Accessed 21 08 2019].
- [19] Bombardier Aerospace, "Canadair CRJ 100/200 Airport Planning Manual," Bombardier Aerospace, Montreal, 2016.
- [20] Airfleets.net, "Production List: Canadair Regional Jet," Airfleets.net, 2017.
- [21] Wikipedia, The Free Encyclopedia, "Boeing 777," Wikipedia, The Free Encyclopedia, 2022. [Online]. Available: https://en.wikipedia.org/wiki/Boeing_777. [Accessed September 2022].
- [22] Boeing Commercial, "777-200/300 Airplane Characteristics for Airport Planning," Boeing Co., Seattle, 1998.
- [23] Airbus SE, "Airbus A340-500/-600 Airplane Characteristics," January 2010. [Online]. Available: https://www.airbus.com/sites/g/files/jlcbta136/files/2021-11/Airbus-Commercial-Aircraft-AC-A340-500_600.pdf. [Accessed August 2022].
- [24] M. Grotchmann and K. Grimm, "Classification trees for partition testing," *Journal of Software: Testing, Verification and Reliability*, pp. 63-82, June 1993.
- [25] Mathworks, Inc., "fitctree - Fit binary decision tree for multiclass classification," Mathworks, Inc., September 2022. [Online]. Available: <https://www.mathworks.com/help/stats/fitctree.html>. [Accessed 2022].
- [26] J. Chen, A. M. MacEachren and D. J. Pequet, "Constructing Overview + Detail Dendrogram-Matrix Views," *IEEE Transactions on Visualization and Computer Graphics*, pp. 889-896, 2009.
- [27] Boeing Commercial Airplanes, "Boeing 757 Airplane Characteristics for Airport Planning," Boeing, Seattle, 2002.
- [28] J. A. Fregnani, B. Mattos and J. A. Hernandez, "Multidisciplinary and Multi-Objective Optimization Considering Aircraft Program Cost and Airline Network," *Journal of Air Transportation*, pp. 21-41, January 2021.
- [29] N. Secco and B. Mattos, "Artificial neural networks to predict aerodynamic coefficients of transport airplanes," *Aircraft Engineering and Aerospace Technology*, pp. 211-230, March 2017.
- [30] M. Widhalm, A. Ronzheimer and M. Hepperle, "Comparison between Gradient-free and Adjoint Based Aerodynamic Optimization of a Flying Wing Transport Aircraft in the Preliminary Design," in *25th AIAA Applied Aerodynamics Conference*, Miami, 2012.
- [31] Z. Saad, "Genetic Algorithm Optimization on an Airfoil using PARSEC Method to Describe the Shape of the Airfoil," Oct 2018. [Online]. Available: https://www.mathworks.com/matlabcentral/fileexchange/69095-optimization-of-an-airfoil-shape-using-genetic-algorithm-ga?s_tid=srchtitle. [Accessed June 2021].
- [32] J. Fregnani, B. S. Mattos and J. A. Hernandez, "An Innovative Approach for integrated Airline Network and Aircraft Family Optimization," in *AIAA Aviation Forum 2019*, Dallas, DOI: <https://doi.org/10.2514/6.2019-2865>, 2019.

- [33] I. Chakraborty, T. Nam, J. Gross, D. Mavris, J. Schetz and R. Kapania, "Comparative Assessment of Strut-Braced and Truss-Braced Wing Configurations Using Multidisciplinary Design Optimization," *Journal of Aircraft*, pp. 2009-2020, 2015.
- [34] R. Priem, H. Gagnon, I. Chittick, S. Dufresne, Y. Diouane and N. Bartoli, "An Efficient Application of Bayesian Optimization to an Industrial MDO framework for Aircraft Design," in *AIAA AVIATION 2020 FORUM*, Virtual Event, 2020.
- [35] R. Garnett, *Bayesian Optimization*, Cambridge: Cambridge University Press, 2022.
- [36] İ. Yanıkoğlu, B. Gorissen and D. den Hertog, "A survey of adjustable robust optimization," *European Journal of Operational Research*, pp. 799-813, September 2019.
- [37] J. Garcia and A. Peña, "Robust Optimization: Concepts and Applications," in *Nature-inspired Methods for Stochastic, Robust and Dynamic Optimization*, United Kingdom, InterchOpen, 2018.
- [38] D. Bertsimas and M. Sim, "The Price of Robustness," *Operations Research*, pp. 35-53, 2004.
- [39] A. Rios, J. Fregnani, B. Mattos and M. Condé, "Identification of the Actual Mission Profiles and Their Impact on the Integrated Aircraft and Airline Network Optimization," 2021.
- [40] V. Loureiro and B. S. Mattos, "Powerplant as Design Variable for Multi-disciplinary Design Optimization of Transport Airplane," in *27th AIA Applied Aerodynamics Conference*, San Antonio, 2009.
- [41] J. Fregnani, B. Mattos and J. Hernandez, "Multidisciplinary and Multi-Objective Optimization Considering Aircraft Program Cost and Airline Network," *Journal of Air Transportation*, pp. 27-41, January 2021.
- [42] R. Langton, C. Clark, M. Hewitt and L. Richards, *Aircraft Fuel Systems*, John Wiley & Sons, 2009.
- [43] B. Singh, "A Medium-Fidelity Method for Rapid Maximum Lift Estimation," Delft University, Delft, 2017.
- [44] M. Drela, "XFOIL - Subsonic Airfoil Development System," October 2018. [Online]. Available: <http://web.mit.edu/drela/Public/web/xfoil/>. [Accessed October 2018].
- [45] S. Gudmundsson, "Maximum Lift Coefficient," Elsevier, 2018. [Online]. Available: <https://www.sciencedirect.com/topics/engineering/maximum-lift-coefficient>. [Accessed September 1029].
- [46] R. C. Quintero, "TECHNO-ECONOMIC AND ENVIRONMENTAL Techno-Economic and Environment Risk Assessment of Innovative Propulsion Systems for Short-Range Civil Aircraft," Cranfield, 2009.
- [47] MathWorks, "Genetic Algorithm," MathWorks, December 2017. [Online]. Available: <https://www.mathworks.com/discovery/genetic-algorithm.html>. [Accessed December 2017].
- [48] Jane's, "Jane's Aero Engines," IHS Markit, December 2017. [Online]. Available: <https://janes.ihs.com/>.
- [49] Wikipedia, The Free Encyclopedia, "Airbus A320 family," Wikipedia, The Free Encyclopedia, 2021. [Online]. Available: https://en.wikipedia.org/wiki/Airbus_A320_family. [Accessed March 2021].

- [50] Airbus SE, "Airbus A320 Aircraft Characteristics and Maintenance Planning," Airbus, Blagnac, 2017.
- [51] Modern Airlines, "Airbus A320 Specs – What is behind one of the most popular short-haul airliners?," Modern Airlines, March 2021. [Online]. Available: <https://modernairliners.com/airbus-a320-introduction/airbus-a320-specs/>. [Accessed 2021].
- [52] Wikipedia, The Free Encyclopedia, "Rolls-Royce RB211," Wikipedia, The Free Encyclopedia, 2021. [Online]. Available: https://en.wikipedia.org/wiki/Rolls-Royce_RB211. [Accessed May 2021].
- [53] European Union Aviation Safety Agency, "Type-Certificate Data Sheet for Rolls&Royce for RB211-535 Series Engine," EASA, Cologne, 2020.
- [54] Boeing Commercial Airplanes, "Boeing Resources," 2021. [Online]. Available: https://www.boeing.com/resources/boeingdotcom/company/about_bca/startup/pdf/historical/757_passenger.pdf. [Accessed May 2021].
- [55] Z. Saad, "Genetic Algorithm Optimization on an Airfoil using PARSEC Method to Describe the Shape of the Airfoils," Oct 2018. [Online]. Available: https://www.mathworks.com/matlabcentral/fileexchange/69095-optimization-of-an-airfoil-shape-using-genetic-algorithm-ga?s_tid=srchtitle. [Accessed June 2021].
- [56] Boeing Commercial, "747-400 Airplane Characteristics for Airport Planning," Boeing Co., Seattle, 2002.
- [57] Wikipedia, The Free Encyclopedia, "McDonnell Douglas DC-10," Wikipedia, The Free Encyclopedia, March 2021. [Online]. Available: https://en.wikipedia.org/wiki/McDonnell_Douglas_DC-10. [Accessed March 2021].
- [58] Boeing Commercial Airplanes, "757-200/300 Airplane Characteristics for Airport Planning," Boeing Co., Seattle, 2002.
- [59] Boeing Commercial Airplanes, "767 Airplane Characteristics for Airport Planning," Boeing Co., Seattle, 2005.
- [60] "Falcon 900LX," Dassault Aviation, 2021. [Online]. Available: <https://www.dassaultfalcon.com/en/Aircraft/Models/900LX/Pages/overview.aspx>. [Accessed March 2021].
- [61] Gulfstream Aerospace, "Gulfstream G550," Gulfstream Aerospace, 2021. [Online]. Available: <https://www.gulfstream.com/en/aircraft/gulfstream-g550/>. [Accessed March 2021].
- [62] Dassault Aviation, "Falcon 6X," Dassault Aviation, 2021. [Online]. Available: <https://www.dassaultfalcon.com/en/Aircraft/Models/6X/Pages/overview.aspx>. [Accessed March 2021].
- [63] Embraer Executive Jets, "Embraer Legacy 650," 2021. [Online]. Available: <https://www.basjets.com/de/Flugzeugkauf/Neuflugzeuge/Long-Range-Jets/Embraer-Legacy-650>. [Accessed March 2021].
- [64] G. W. van Bondergraven, "Commercial Aircraft DOC Methods," in *AIAA/AHS/ASEE Aircraft Design Systems and Operations Conference*, Dayton, 1990.

- [65] Air Transport Association of America, "Standard Method of Estimating Comparative Operating Direct Costs of Turbine Powered Transport Airplanes," Air Transport Association of America, Washington, D. C., 1967.
- [66] "Jet Fuel Price Monitor," IATA, May 2021. [Online]. Available: <https://www.iata.org/en/publications/economics/fuel-monitor/>. [Accessed May 2021].
- [67] D. Howe, Aircraft Conceptual Design Synthesis, London and Bury St Edmunds: Professional Engineering Publishing Limited, 2000.
- [68] M. Gumes e L. D. Andrade Pereira, "Utilização de Redes Neurais para Estimação de Coeficientes Aerodinâmicos e Derivadas de Estabilidade," Undergraduate thesis, Instituto tecnológico de Aeronáutica,, São José dos Campos, 2019.
- [69] B. S. Mattos, F. A. Bortolete, J. A. T. Fregnani, A. B. Jorge, W. M. Alves and R. V. Cruz, "Application of Machine Learning and Multi-Disciplinary/Multi-Objective Optimization Techniques for Conceptual Aircraft Design," in *Model-Based and Signal-Based Inverse, Vol. 1*, Brasília, D.F., University of Brasilia, 2022, pp. 143-236.

Chapter 17

Random Fatigue Under Uniaxial and Multiaxial Loading Conditions

Chapter details

Chapter DOI:

<https://doi.org/10.4322/978-65-86503-88-3.c17>

Chapter suggested citation / reference style:

Ferreira, Jorge L. A., et al. (2022). “Random Fatigue Under Uniaxial and Multiaxial Loading Conditions”. In Jorge, Ariosto B., et al. (Eds.) *Uncertainty Modeling: Fundamental Concepts and Models*, Vol. III, UnB, Brasilia, DF, Brazil, pp. 552–612. Book series in Discrete Models, Inverse Methods, & Uncertainty Modeling in Structural Integrity.

P.S.: DOI may be included at the end of citation, for completeness.

Book details

Book: Uncertainty Modeling: Fundamental Concepts and Models

Edited by: Jorge, Ariosto B., Anflor, Carla T. M., Gomes, Guilherme F., & Carneiro, Sergio H. S.

Volume III of Book Series in:

Discrete Models, Inverse Methods, & Uncertainty Modeling in Structural Integrity

Published by: UnB City: Brasilia, DF, Brazil Year: 2022

DOI: <https://doi.org/10.4322/978-65-86503-88-3>

Random Fatigue Under Uniaxial and Multiaxial Loading Conditions

Jorge Luiz de A. Ferreira^{1*}, Felipe M. de Lima² and Emanuely U. Cardoso³

¹ Post-Graduate Program - Integrity of Engineering Materials, University of Brasilia, Brazil. Email:jorge@unb.br

² Post-Graduate Program - Mechanical Sciences (Fatigue, Fracture and Materials Group, GFFM/PCMEC), University of Brasilia, Brazil. E-mail:felipe.maganhalima@gmail.com

³ Post-Graduate Program - Integrity of Engineering Materials, University of Brasilia, Brazil, E-mail:emanuely.ugulino@gmail.com

Abstract

Fatigue life prediction under multiaxial loading has become a fundamental part of product development for industries in areas such as aeronautics, offshore, automotive, and power generation, among others. Among the reasons can be mentioned the need to increase the safety and performance levels of structural components, together with the popularization of the use of computational tools capable of allowing the modeling of complex component geometries under any type of load. Although the computational packages used in these analyzes have several fatigue models, research is still needed to improve their life prediction capacity and computational efficiency, especially in the presence of random multiaxial loads. This is mainly because fatigue analysis under multiaxial random loading is still a task that is erroneously considered to be extremely complex. In this sense, this monograph is dedicated to presenting a comprehensive review of the concepts needed to understand the multiaxial random fatigue criteria available in the literature. Such a review is mainly devoted to stress-based criteria for the evaluation of fatigue life in high-cycle regime using the Maximum Variance Method. Time and frequency domain approaches are examined. To level the knowledge necessary to use the methodologies, a detailed review of the literature is presented on i) stochastic processes and characterization of loading histories (sections 1 and 2), ii) Material Fatigue and Uniaxial Random Fatigue Analysis (sections 3 to 5), iii) Fatigue multiaxial and Maximum Variance Methods using the Critical Plane approach (sections 6 to 8) iv) Fatigue multiaxial and Projection-by-Projection criteria (PbP) (section 9)

Keywords: life prediction; multiaxial loading; multiaxial random fatigue criteria; stochastic processes; characterization of loading histories; maximum variance method; Projections-by-Projection critical planes.

1 Introduction

Many mechanical and civil structures are subject to loads that vary essentially randomly throughout their lifetime. These loadings are, in general, caused by the most diverse sources (such as:

sea waves, winds, earthquakes, vehicle traffic, etc.). Fig. 1 are exemplified some of the possible loading histories observed in real situations.

These loading histories can be induce a mode of structural failure known as fatigue. Roughly speaking, fatigue can be defined as a process of localized damage and the cumulative one consisting of the beginning of the crack, the propagation, and the final fracture of a component. Typically, the final fracture occurs without notice, which makes the prevention of fatigue fractures a vital aspect of the design of machines, vehicles and structures that are subject to repeated loads or vibrations.

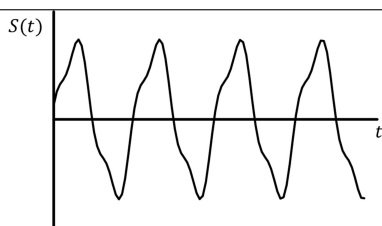
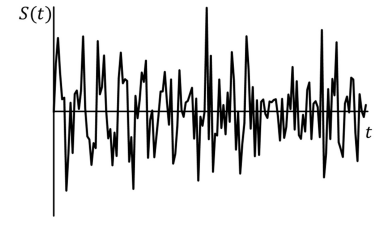
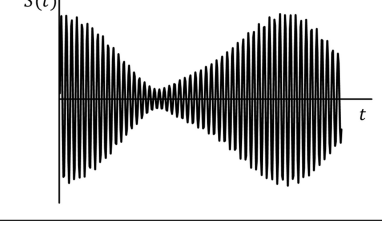
Shape of loading	Type of structural component
	Simple mechanism
	Rigid bridge decks tides winds
	Overhead Conductor

Figure 1: Some typical loading stories acting on structural components.

Fatigue life prediction models of components subjected to random loading are divided into time domain methods and the frequency domain methods. Life prediction methods in the time domain are generally based on identification and cycle counting algorithms associated with a damage accumulation rule. The frequency domain methods, on the other hand, use the spectral density of the loading to estimate the amplitude distribution of the cycles present in the history. In particular, in this work will be discussed and presented the approaches of modeling of fa say that consider that the phenomenon of fatigue is controlled only by stress, that is, characterized by the presence of plastic deformations only at the level of grain (minimum level).

2 Loadings histories and their classifications

In general, loading histories can be classified as **deterministic and non deterministic or random loadings**. Deterministic loading histories are those that can be described by an explicit mathematical relationship. A classic example is the oscillation of a mass-spring system composed of an ideal spring, of elastic constant, attached to a mass body (see Fig. 2).

If the body is moved from its equilibrium position by a distance X and released at time $t = 0$, from basic laws of mechanics or through repeated observations, it can be established that the following relationship between the body's position and the time t .

$$x(t) = X \cos \left(\sqrt{\frac{k}{m}} t \right) \quad (1)$$

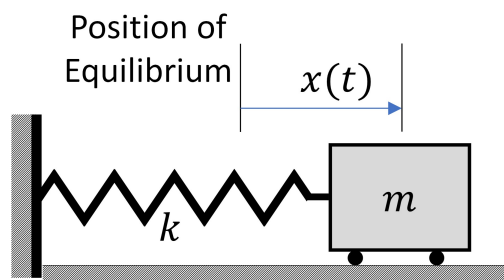


Figure 2: Single spring mass system.

Thus, once the initial offset is X defined, the Eq. (1) sets the exact location of the body at any time in the future. Therefore, the information representing the mass movement is deterministic.

In other physical systems, there is no functional relationship that allows explaining the behavior of the data that characterizes the behavior of the system (There is no way to predict the exact value of the parameter in a future instant). In these situations, the system is called nondeterministic. In such cases the data are random in nature and should be described in terms of probability of occurrence and/or statistical means, rather than explicit equations. Examples of non-deterministic systems are i) the height of waves in a sea, ii) the dynamic pressure applied on a structure at a specific time instant, iii) the displacement of an automotive suspension system subject to the iterations between the tire and the tread.

2.1 Deterministic histories

Deterministic histories can be classified as periodic or non-periodic. Periodic histories can be further classified as:

- Sinusoidal;
- Complex periodic;
- Quasi-periodic;
- Transients.

2.1.1 Sinusoidal Histories

The histories sinusoidal are histories that can be described by a mathematical function that describes a smooth repetitive oscillation, defined to mathematically by a variable function in the time of the form:

$$x(t) = X \sin (2\pi f t + \theta) \quad (2)$$

where X , f and θ represents, respectively, the amplitude of history, frequency (in Hertz), and the initial phase angle in relation to the origin of time (in radians).

Note that the sinusoidal function belongs to a class of functions called periodic. Periodic functions are those in which the values of the function ($x(t) = y$) are repeated for certain values of the variable t , that is, for each given period, T , we will obtain repeated values for the function, such that: $y = x(t) = x(t + T) = x(t + 2T) = x(t + nT)$ (where T is the repetition period of y and n is an integer).

2.1.2 Complex Periodic Histories

There are per-iodic loading histories that need to be defined mathematically by a variable function in time more complex than that presented in Eq. (2). With few exceptions, these histories (called complex periodic histories) can be expanded into a Fourier series according to the following formula:

$$x(t) = \frac{a_0}{2} + \sum_{(n=1)}^{\infty} (a_n \cos(\omega_n t) + b_n \sin(\omega_n t)) \tag{3}$$

where a_0 , a_n and b_n are called Fourier coefficients, obtained by the expressions:

$$a_0 = \frac{1}{T} \int_0^T x(t) dt \tag{4}$$

$$a_n = \frac{2}{T} \int_0^T x(t) \cos(\omega_n t) dt \tag{5}$$

$$b_n = \frac{2}{T} \int_0^T x(t) \sin(\omega_n t) dt \tag{6}$$

where $n = 1, 2, 3, 4, \dots$

where $\omega_n = 2\pi f_n$ for $n = 1$, $f_n = f$ is a fundamental frequency.

Considering that $x(t)$ is real, the Eq. (3) can be written as follows:

$$x(t) = X_0 + \sum_{(n=1)}^{\infty} (X_n \cos(\omega_n t + \theta_n)) \tag{7}$$

where:

$$X_0 = \frac{a_0}{2} \tag{8}$$

$$X_n = \sqrt{a_n^2 + b_n^2} \tag{9}$$

$$\theta_n = \tan^{-1} \left(\frac{-b_n}{a_n} \right) \tag{10}$$

The identification of harmonic components present in complex periodic histories is very convenient because it allows the graphic representation of these histories through a graph, called frequency spectrum, which correlates the frequencies, and the respective amplitudes present in the loading history.

Example 2.1 Determine the coefficients of the Periodic Signal Fourier Series, $x(t)$, and present the loading history and frequency spectrum of the signal

$$x(t) = x_1 \sin(2\pi f_1 t) + x_2 \sin(2\pi f_2 t) + x_3 \sin(2\pi f_3 t)$$

where $[x_1, x_2, x_3] = [10, 5, 2]e[f_1, f_2, f_3] = [\sqrt{3}, 18, 25]$.

Solution:

Assuming $f = 1, T = 1s$ and solving the equations Eq. (5) and Eq. (6) for n ranging from 1 to 30, we found the values represented in Figure 2. 2(a) pp the coefficients a_n and b_n . The graphs describing the loading history and frequency spectrum of the signal are shown in figures Fig. 3(b) and Fig. 3(c), respectively.

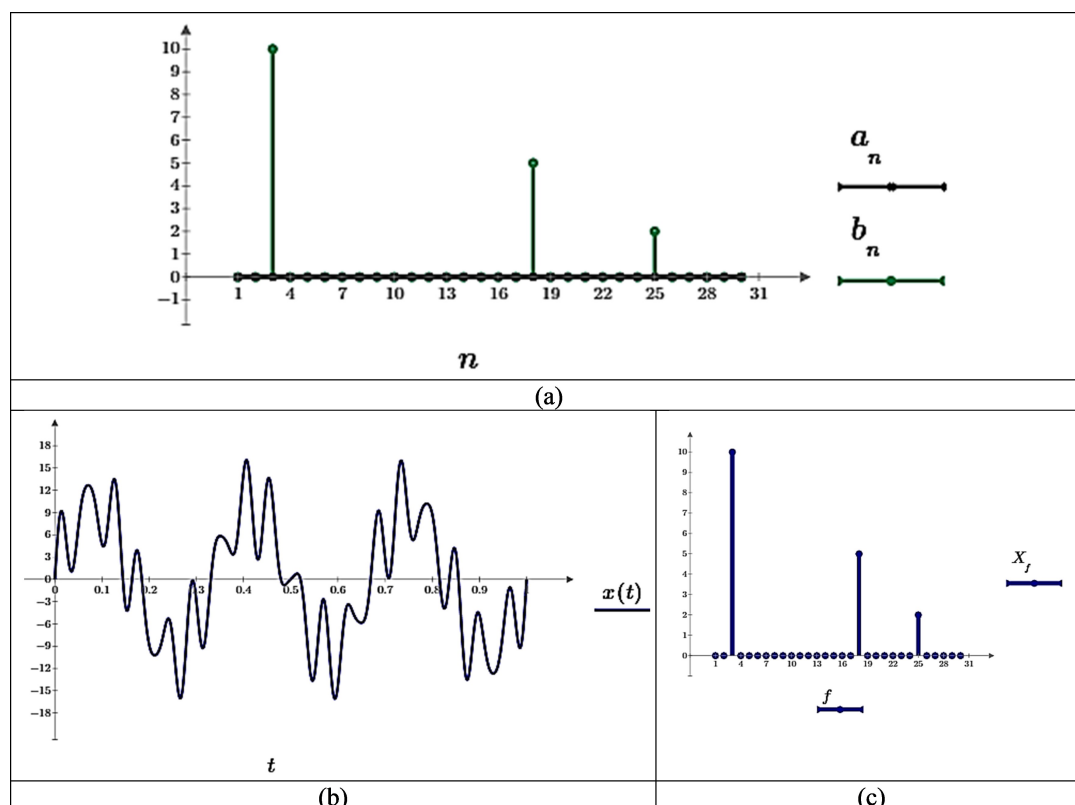


Figure 3: (a) Coefficients and the Fourier series a_n and b_n , (b) representation of history $x(t)$, and (c) representation of the frequency spectrum of history

2.1.3 Quasi-periodic Histories

The sum of two or more sine waves is periodic only when the ratio between all possible frequency pairs are rational numbers. This indicates that there is a fundamental period that will satisfy the periodicity requirement, i.e $x(t) = (t + nT)$. For example, considering the frequencies used in **Example 2.1**, $([f_1, f_2, f_3] = [\sqrt{3}, 18, 25])$, it is observed that the ratios f_2/f_1 and f_3/f_1 are rational numbers (indicating that the story is periodic!!). However, in some situations the loading history can be composed of frequency components that belong to the set of irrational numbers (i.e., numbers that cannot be represented as a fraction). In these specific situations, the fundamental period is infinitely long and, as a consequence, the resulting time history will not meet the basic requirement of periodicity. When we analyze quasi-periodic histories using the Fourier series, the differences are observed: a a) the frequencies of the components are not related by rational numbers (i.e., the frequencies observed in the frequency spectrum are approximations of the real

values), and b) spurious frequency components may appear around frequencies present in the signal, as illustrated in Fig. 4.

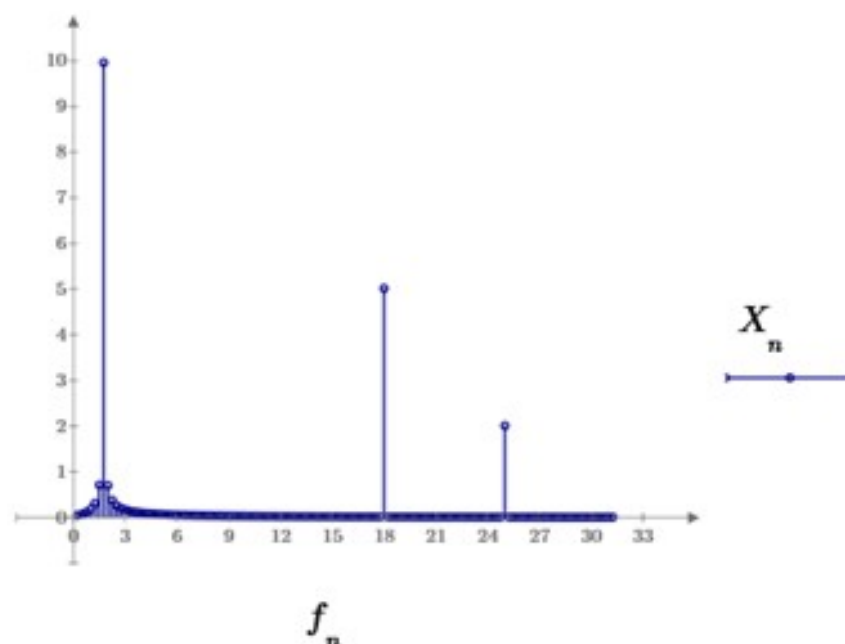


Figure 4: Representation of the frequency spectrum of history of Example 1.

2.1.4 Transient Histories

Determinist histories to transients usually reflect unique events that do not have apparent periodicity, that is, they have no characteristic period. The vibratory response of a mass-spring-damper system resulting from an initial condition $v(t = 0) = v_0 \neq 0$ ou $x(t, = 0) = X_0 \neq 0$, can be given by the expression:

$$x(t) = X_0 e^{-\alpha t} \cos(2\pi f t + \theta) \quad (11)$$

2.2 Stochastic Histories

Stochastic (or random) stories represent physical phenomena that vary to some degree, unpredictably, as time passes, and cannot be described by explicit mathematical relationships because each observation of the phenomenon will be unique. Thus, one can understand a random history as a random variable sequence defined in a probability space. Thus, one can understand a random history as a random variable sequence defined in a probability space. Thus, a single sample representing the observed record over a finite time interval of a random phenomenon is defined as a sample function (represented by each graph of Fig. 5). The collection of all possible sample functions that the random phenomenon may have produced is called a random process or stochastic process (represented by all graphs of Fig. 5). Thus, we can admit that a record observed in a finite time interval of a random phenomenon can be considered as a physical realization of a random process.

Similar to what was observed for deterministic histories, random histories also show differences in terms of their behavior in terms of time. In this sense, the random histories can be classified as stationary or nonstationary.

Conceptually, a time series is said to be stationary when its statistical characteristics (mean, variance, autocorrelation, probability distribution, statistical moments, etc.) are constant over time.

That is, it is a series that develops randomly in time, around a constant mean, reflecting that the probability laws that act in the process do not change over time. Note that for the characterization of a history as stationary is extremely restrictive. For this reason, the definition of this concept takes two forms, namely: a) stationary in the strict sense and b) stationary in the wide sense.

2.2.1 Stationarity in the strict sense (or strongly stationary)

The stochastic process $x(t)$ is said stationary in the strict sense if its statistical properties are invariant to any translation of the origin of the times. This means that processes $X(t)$ and $X(t+\tau)$ have the same statistics whatever the time shift $\Delta\tau$. Formally this means that the joint distribution function obeys equality when it k tends to infinity:

$$F_{X(t_1),X(t_2),\dots,X(t_k)}(X_1, X_2, \dots, X_k) = F_{X(t_1+\tau),X(t_2+\tau),\dots,X(t_k)+\tau}(X_1, X_2, \dots, X_k) \quad (12)$$

for any temporal translation, any and all instant choice $t_1 < t_2 < \dots < t_k$.

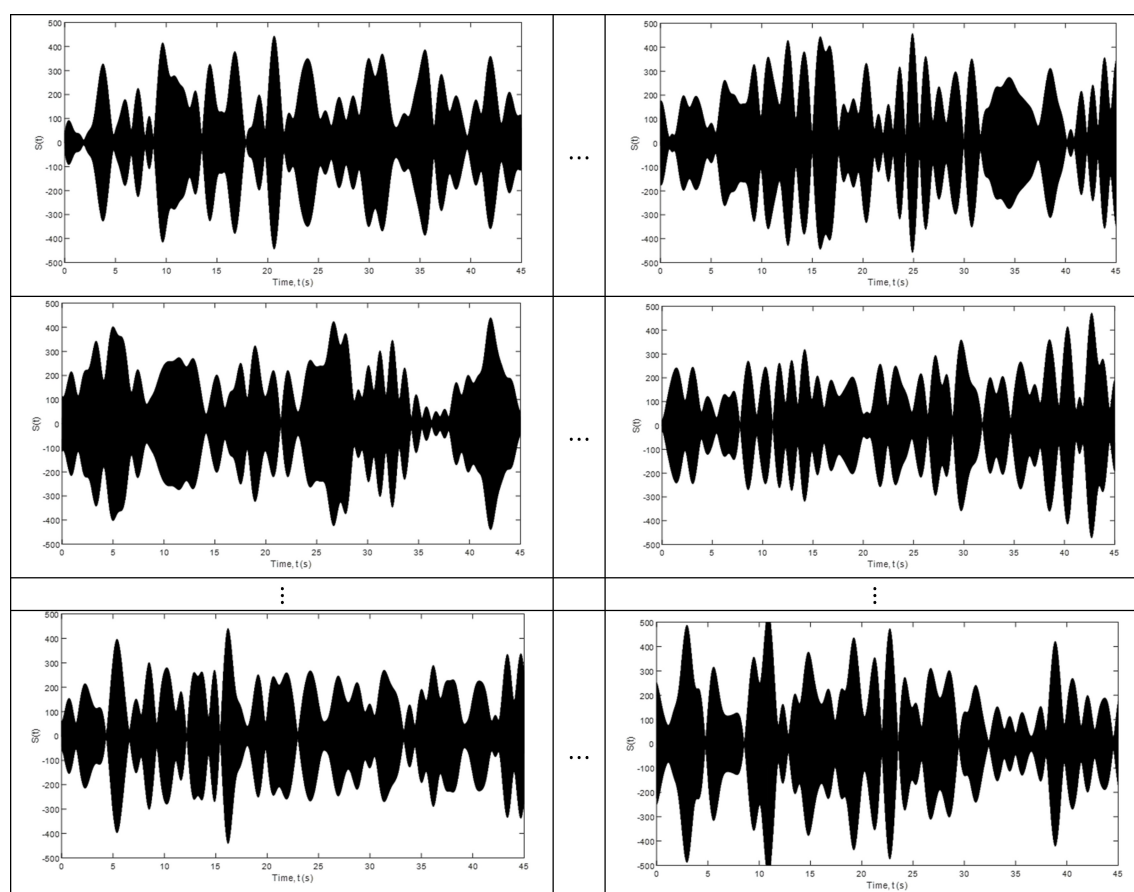


Figure 5: Representation of sample functions of a random process (collection of all possible sample functions).

Figure 2.5 illustrates a set with n sample functions that form a random process. In this figure can be observed the set formed by the realization observed in the moments $t([X_t^1, X_t^2, \dots, X_t^k])$ and $t + \tau([X_{t+\tau}^1, X_{t+\tau}^2, \dots, X_{t+\tau}^k])$.

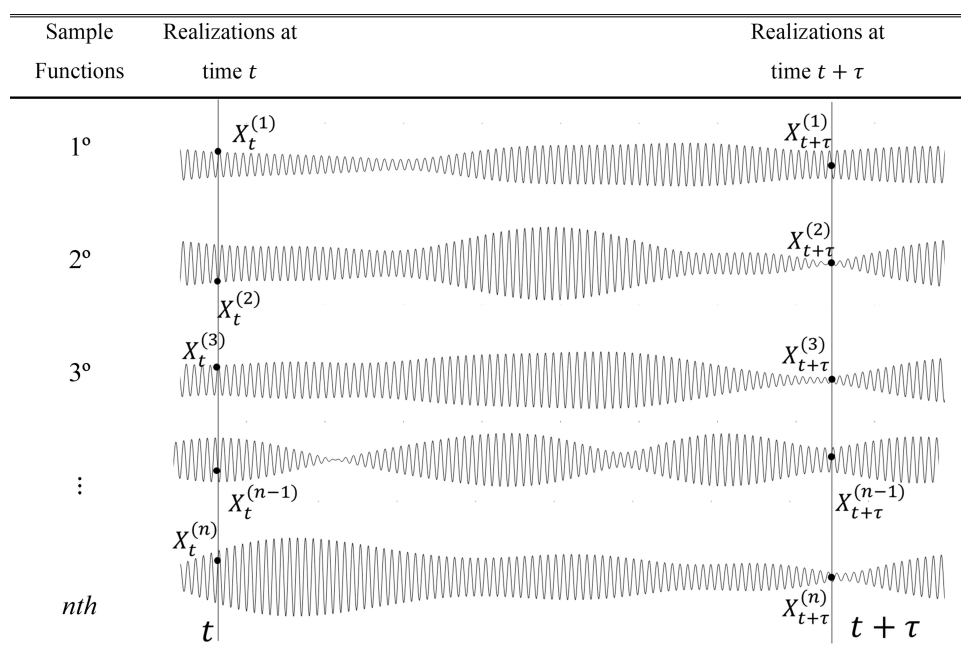


Figure 6: Schematic representation of the acquisition process of the points used to estimate the function of joint distribution of the history observed in two distinct moments of time.

From these two sets of points, you can infer about the behavior of these set of points. For example, the mean value (first-order statistical moment) of the random process at some time t can be calculated by estimating the instantaneous value of each sample function of the ensemble at that particular instant (i.e., the realizations at t), summing the values and dividing by the number of sample functions. Similarly, a correlation (set moment) between the values of the random process in the moments t and $t + \tau$, called the autocorrelation function, $R_{xx}(t_1, t_{1+\tau})$ can be estimated considering the equation Eq. (13):

$$R_{xx}(t, t + \tau) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N X_t^n X_{t+\tau}^n \tag{13}$$

where τ represents the time interval separating the time series realizations.

When the history is discretized, τ it is calculated as $r\Delta t$ where Δt represents the time interval used to sample the history realizations, and r is called *lag* and represents the number of time periods separating the time series data (The maximum number of lags, as suggested by Box and Jenkins, is approximately $N/4$ for a history with less than 240 realizations or the largest integer greater than $\sqrt{N} + 45$ for a series of more than 240 observations, where N is the number of realizations). Note that when τ it equals zero to Eq. (13) takes the form of the variance estimator.

2.2.2 Stationarity in the wide sense (or weakly stationary)

A less rigorous way of defining the concept of stationarity is known as weak stationarity, stationarity in the wide sense, or covariance stationarity. The random histories that fall as weakly stationary require only that the first moment (i.e., the mean) be constant over time and that the function of autocorrelation, $R_{xx}(t_1, t_1 + \tau)$, (estimated according to the Eq. (13) depends only on τ .

2.2.3 Stationarity of n-order

In some practical situations (for example in fatigue analysis), ensuring that a history is stationary in a wide sense is not enough to perform an adequate modeling of the problem. In these cases, it is interesting to check which moments. In these cases, it is interesting to verify which statistical moments do not change over time, that is, to evaluate the value n of which the Equation satisfies Eq. (12)). In summary, we can conclude that for any random process $X(t)$ where the mean, $\mu(t_1)$, and the autocorrelation function, $R_{xx}(t_1, t_1 + \tau)$, vary with time t_1 , the random process $X(t)$ is said to be nonstationary. For the special case where $\mu(t_1)$ and $R_{xx}(t_1, t_1 + \tau)$ do not vary as time t_1 varies, the random process $X(t)$ is said to be weakly stationary or stationary in the wide sense. For weakly stationary random processes, the mean value is a constant, and the autocorrelation function depends only on the lag, τ . That is, $\mu(t_1) = \mu_x$ and $R_{xx}(t_1, t_1 + \tau) = R_{xx}(\tau)$.

Example 2.2 Determine the auto-correlation functions of the sample functions presented in Fig. 7

Development:

For the history, the outputs of the sample functions were organized in the form of a data matrix DADOS(IxJ) with 1379 (3/0.002175) rows and 17 columns.

Considering that the data are discretized, the estimation of the autocorrelation function is performed using the routine presented in Table 1.

Table 1 – Script MATLAB: function used for calculating the autocorrelation function

```
function [R,t_lag] = AUTO_CORR_M(DATA,LAG,delta_t)
[I J] = size(DADOS);
for layer=1:LAYER+1
    team = team - 1;
    t_lag(layer) = (layer)*delta_t;
    SAMPLES = I-lag;
    DUMMY = 0;
    for i = 1: I-layer
        D1 = DATA(i,:);
        D2 = DATA (i+laga,:);
        D1_D2 = D1*D2';
        DUMMY = DUMMY + (D1_D2);
    end
    R_1(team+1.1) = DUMMY/(J*(I-team));
end
MAX = max(R_1);
R_1 = R_1 /MAX;
end
```

For each history, 17 sample functions of the recorded for 3s are available at a 0.002175s sampling rate.

The routine has as output parameters the vectors t_{lag} (which contains the time intervals that separate the achievements from the time series) and R (which contains the estimate of the history self-correlation function), and as input parameter: a) the data matrix (DATA), b) maximum number of lags (LAG), and the sampling rate (delta_t). As a result of data processing, the auto-correlation functions presented in Fig. 8 are estimated.

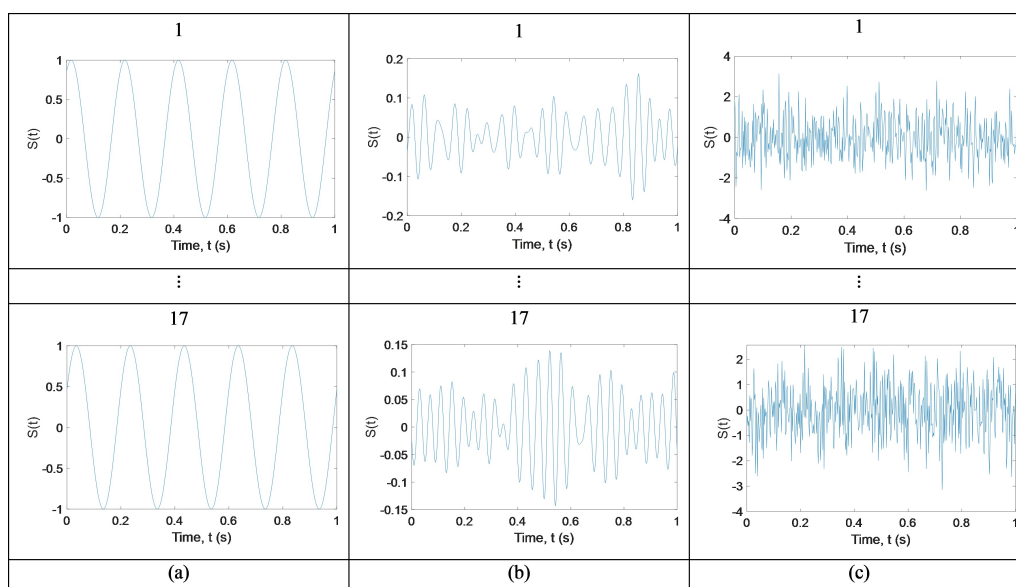


Figure 7: Histories (a) sinusoidal, (b) Narrowband and (c) Noise.

Looking at the charts of Fig. 8 It is verified that the function of autocorrelation of sine history is also a sinusoidal function, which indicates that the correlation between achievements varies cyclically as a function of lag. The function of autocorrelation of the narrow band history (to be defined in following topics) presents a large peak Lag 0 ($R_{xx}(\tau = 0) = Var[S(t)]$) in the followed by a decreasing wave that alternates between positive and negative values, indicating that the correlation between achievements is very strong for when they are very close and weakened as they move away. On the other hand, the function of noise autocorrelation also has a high value in the first or second lag, followed by correlations that are not significant, indicating a strong statistical independence between the achievements of the process.

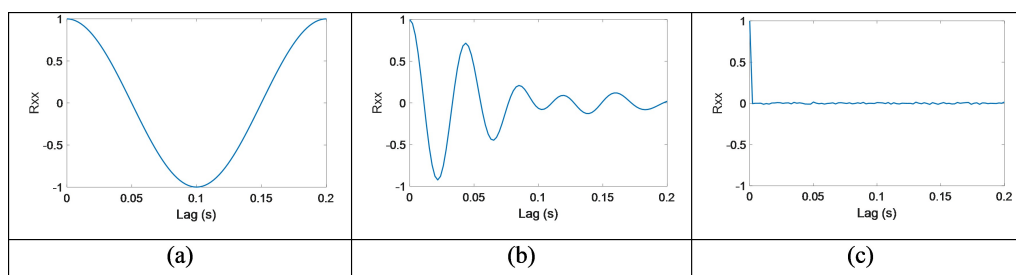


Figure 8: Function of Autocorrelation of histories (a) sine. (b) Complex Periodic, (c) Noisy

Ergodicity

Ergodicity is an important property of a stochastic process, related to stationarity. A process is defined as ergodic in relation to a statistical parameter, if the parameter characterized over time considering a randomly selected sample function is equivalent to the same parameter characterized in the whole set of sample functions in a given period of time. For example, a process is said to be ergodic on the mean if the mean of a randomly selected sample function is equal to the mean of the process over the entire set at any time t or, in other words, if the time mean is the same as the mean of the sample function. Certain statistical conditions must be satisfied to determine whether a process has the property of ergodicity. For example, considering measurements taken over a period T of various sample functions of the process $X(t)$ and assuming that N represents the total number of sample functions, M total number of realizations taken on each sample function over a period T , and $x_m^{(n)}$ is the value of the m th realization of the N -th sample function, we can define:

A - Process Ergodic on Average (or ergodic for the First Moment)

If the mean of the sample function (Ensemble average) approaching the mean over time (time average), we characterize this random process as “process ergodic in the mean”.

Where the mean of the sample functions at time t_i , $\hat{\mu}_{X(t_i)}$, is estimated by averaging the n points of the sample functions, that is:

$$\hat{\mu}_{X(t_m)} = \frac{1}{N} \sum_{n=1}^N x_i^n \quad (14)$$

and the mean over time is called the *time average* and is estimated by Eq. (15):

$$\hat{\mu}_m^n = \frac{1}{M} \sum_{m=1}^M x_m^n \quad (15)$$

B - Ergodic Process in relation to Autocorrelation Function (or Ergodic for the Second Moment)

A process is ergodic in relation to the autocorrelation function if the autocorrelation function of the sample function (*Ensemble autocorrelation*), calculated using Eq. (13) approach the *temporal autocorrelation function* $R_X^j(\tau)$, estimated using Eq. (16).

$$\hat{\mu}_m^n = \frac{1}{M} \sum_{m=1}^M x_m^n \quad (16)$$

In more general situations, one is not interested in all the ensemble statistics of a random process. Typically, it is only necessary to guarantee the mean and autocorrelation function (which represents the definition of ergodicity in a more limited sense). In fatigue analysis, to the best of our knowledge, we are also not interested in all ensemble statistics, but in order to be able to use a single sample function to predict fatigue life, we need to ensure that there is ergodicity with respect to the 4th statistical moment.

Example 2.3 Evaluate whether the random process representing a random phase sine wave, represented by the function presented in Eq. (17), is ergodic in relation to the autocorrelation function.

$$s(t) = A \sin(2\pi ft + \phi) \quad (17)$$

where A is the amplitude of the wave ($A = 10$), f is the frequency, t is the time, and ϕ is the random phase with uniform distribution between 0 and 2π .

Solution:

To answer this question, we will generate a set of sample functions using the routine presented in Table 2. The Fig. 9 shows a set of sample functions generated from the routine described in Table 2.

Table 2 – Script MATLAB: routine for the generation of sample functions representative of a sine wave with random phase

```

J=200; % J defines the number of sample functions

f = 10; %f is the frequency

T=5/f; %T is the observation time (in this case 5 wavelengths will be analyzed)

N=10000; %N is the number of points used to describe each sample function

dt = T/N; %dt is the sampling rate

a = 2*pi*f; %a is the argument of the periodic function

phi = 2*pi*rand(J,1); %phi represents a vector containing the J random phases

%time vector calculation

for n =1:N

    t(n,1) = (n-1)*dt;

end

% calculation of nxj matrix containing sample functions

for j=1:J

    for n =1:N

        s(n,j) = 10*sin(a*t(n,1)+phi(j,1));

    end

end

```

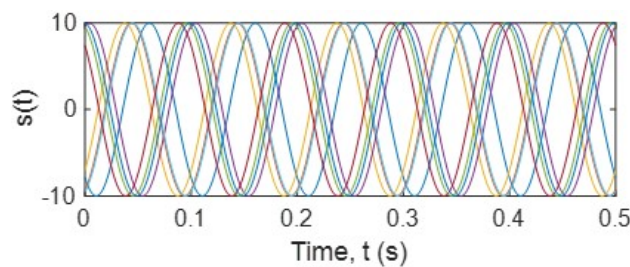


Figure 9: Sine wave sampling functions with random phase.

Once the matrix containing the sample functions is generated, the AUTO_CORR_M function is used to estimate both the autocorrelation function of the sample function (vector R) and the

estimates for the temporal autocorrelation function (vector R_t). As a result, the graphs presented in Fig. 10 showed the evolution of the autocorrelation function with the lag time were obtained.

Comparing the estimates of the autocorrelation function of the sample function and the temporal autocorrelation function, it is verified that both have the same behavior. This indicates that random phase sine waves present the characteristic of and ergodicity in relation to the autocorrelation function.

Table 3 – Script MATLAB: AUTO_CORR_M

```
function [R_t, R_t_lag] = AUTO_CORR_M(DATA,LAG,delta_t)

[I J] = size(DATA); %returns the number of rows, I, and the number of columns, J, of the Data array

%Estimation of the matrix containing the temporal autocorrelation functions of each %sample function

for j = 1:J

    for layer=1:LAYER

        t_lag(team) = (team-1)*delta_t;

        team = team -1;

        SAMPLES = I-lag;

        DUMMY = 0;

        for i = 1: I-layer

            D1 = DATA(i,j);

            D2 = DATA (i+laj,j);

            D1_D2 = D1 *D2;

            DUMMY =DUMMY +(D1_D2);

        end

        R(lag+1,j) = DUMMY/SAMPLES;

    end

    MAX = max(R(:,j));

    R(:,j) = R(:,j) /MAX;

%Estimate of vector containing the estimate of the autocorrelation function of the sample function

for layer=1:LAYER

    team = team -1;

    SAMPLES = I-lag;

    DUMMY = 0;

    for i = 1: I-layer

        D1 = DATA(i,:);

        D2 = DATA (i+laj,:);

        D1_D2 = D1 *D2;

        DUMMY =DUMMY +(D1_D2);

    end

    R_1(team+1.1) = DUMMY/(J*(I-team));

end

MAX = max(R_1);

R_1 = R_1 /MAX;

end
```

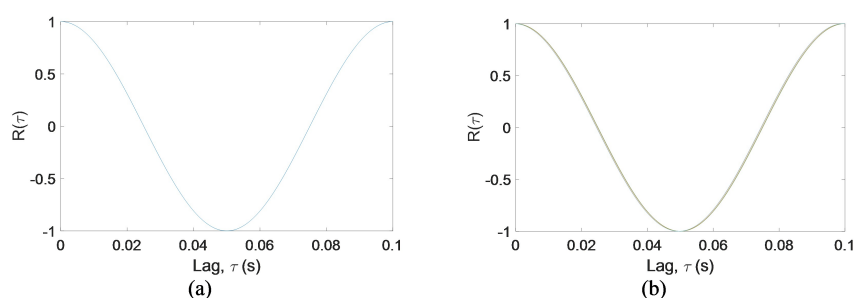


Figure 10: Graphs showing the estimates of (a) autocorrelation function of sample functions and (b) temporal autocorrelation functions.

2.2.4 Non-Stationary Loadings

When random processes cannot be framed as stationary, we say that they are nonstationary. Formally a process is assumed to be nonstationary when the mean and/or variance vary with time. Nonstationary behavior is characterized by the exclusive or combined presence of the following components:

- **Trend:** The *trend* of a random process indicates its "long-term" behavior, that is, whether it grows, decreases or remains stable, and how fast these changes are. For example, increased vibratory response of a system due to increased unbalance, or loss of effectiveness of a damper, etc.
- **Cycles:** *Cycles* are characterized by the oscillations of ascent and fall in the series, smoothly and repeatedly, along the trend component. For example, tidal cycles, winds, operational cycles of equipment.
- **Random walking:** The *random walk* is a statistical phenomenon in which a variable does not follow a discernible trend and apparently moves randomly. The Brownian movement is a classic example of random walking.

In aerospace, automotive and offshore structures, for example, a strategy widely used in the characterization of the loadings (typically non-stationary), consists of modeling this process as a sequence of independent modes of operation that occur during a certain period of time. That is, the loads acting on these structures can be described as a sequence of stationary processes. For example, when analyzing the forces acting on an aircraft, it is common to divide the flight stages into a sequence of loading histories related to the flight phases/conditions/aircraft type (e.g. taxi, takeoff, gusts, cruise, maneuver, landing). Fig. 11 shows a sketch illustrating a typical loading history of aircraft transport aircraft.

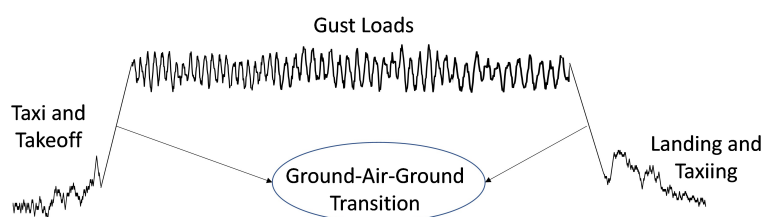


Figure 11: Schematic representation of a typical history of carrying aircraft loading.

3 Fatigue of Materials: Introduction to Stress-Based Approach

As discussed in the previous chapter, during their useful lives it is very likely that structural components are subject to cyclic loads. These loads can have a constant amplitude or amplitude that varies over time. Regardless of the type of load applied, there are three main approaches to analyze and project against fatigue failures. The traditional stress-based approach was developed essentially to its current form in 1955. In this approach, the analysis is based on nominal stresses working at the critical points of the component. The nominal stress that can be applied under cyclic load is determined considering the average stresses and adjusting the effects of stress increases (due to the presence of stress concentrators). Another approach is the deformation-based approach, which involves a more detailed analysis of localized flow that can occur at critical points during cyclic loading. Finally, there is the approach of fracture mechanics, which specifically treats the increasing cracks by the methods of fracture mechanics.

Whereas the service lives of real structures are relatively long (*High number of cycles to failure*), that the stress levels are lower than the material flow stress (*little plastic deformation due to cyclic loading*), and due to the ease of implementation, the stress-based approach is the most traditional and most frequently used in applications involving fatigue under random conditions. This approach assumes that the amplitude of the stress acting at the critical point is responsible for the process of initiation and propagation of fatigue cracks and that fatigue life, expressed in terms of the number of load cycles, is a direct function of the intensity of the stress amplitude.

The stress-based method has its origins in the investigations carried out by August Wöhler in 1858 and resulted in the widely known Wöhler curve, or S-N curve (i.e. stress, S , versus number of cycles for failure, N).

The simplest way to characterize the fatigue behavior of a material is to subject standardized specimens to efforts with sinusoidal characteristics with prescribed amplitudes (with zero mean) and count the number of cycles or stress inversions necessary for complete rupture of the specimen. From the experimental results, one S-N diagram (called diagram or Wöhler diagram) is constructed, which displays very conveniently the basic fatigue data in the elastic stress range (fatigue under high cycle conditions) on a graph that correlates the amplitude of the actuating stress (S_a) versus number of cycles until failure (N).

The S-N diagram covers a life range ranging from 1 cycle to a very high number of cycles. Due to the different physical phenomena that may be related to the fatigue process, it is observed that the relationship between stress and life presents different behaviors for different life ranges. For this reason, as illustrated in Fig. 12, the S-N diagram is usually divided into three regions: (i) low-cycle fatigue (LCF) for life lower than 10³ cycles, (ii) high cycle fatigue (HCF) for life up to 10⁷ cycles, and (iii) very high cycle fatigue (VHCF) for cycles greater than 10⁷ (the VHCF region is also called the Giga cycle fatigue region (GCF) for life upper to 10⁹ cycles). In general, the region of high-cycle fatigue is located between the two knees observed and the region known as the finite life region begins at the end of the 1st knee and ends at the beginning of the 2nd and then, as shown in Fig. 12. Typically, fatigue analyses using the stress-based approach adopt the relationship between stress and life in the finite life region based on describing fatigue behavior. The analytical representation that describes the behavior of material fatigue in this region is commonly expressed in the form:

$$S = AN^{-b} \quad (18)$$

$$NS_a^{\frac{1}{b}} = k = \left(\frac{1}{A}\right)^{\frac{1}{b}} \quad (19)$$

where b , k and A are material parameters estimated from test data obtained using identical samples. The basic methodology for characterizing the statistical behavior of these parameters can be found

in ASTM E739-10 [2015] - *Standard Practice for Statistical Analysis of Linear or Linearized Stress-Life (S-N) and Strain-Life (ε-N) Fatigue Data.*

For ferrous materials in general, it is often adopted a threshold stress amplitude (endurance limit, S_e), below which it is assumed that the material does not fail by fatigue (has infinite life). For engineering purposes, this infinite life is generally regarded as cycles. 10^6 Assuming the existence of endurance limit, the relationship S-N takes the following form:

$$N = \begin{cases} N = AS^{-b}, & S_a > S_e \\ \infty, & S_a \leq S_e \end{cases} \quad (20)$$

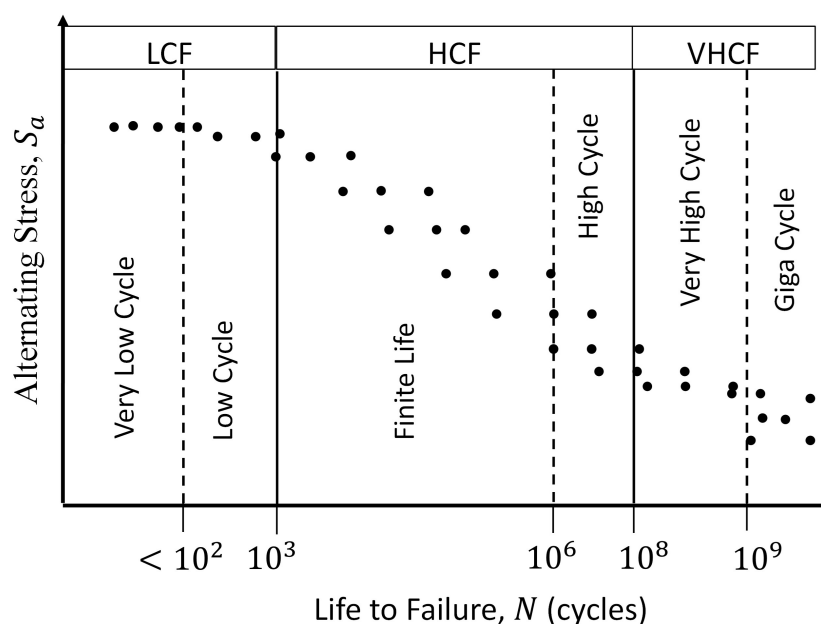


Figure 12: Schematic representation of the S-N diagram.

3.1 Mean Stress Effect

Stress amplitude strongly influences the fatigue behavior of a structural component. However, the presence of an average component in the loading history is also important, as a secondary way) in the characterization of fatigue behavior.

In general, the characterization of the mean stress effect on fatigue behavior, the S-N diagrams should be constructed considering that the experimental points are performed under constant mean stress conditions. Another procedure often used in the construction of the S-N diagram that represents the effect of the mean stress is to use test results performed under stress ratio conditions, constant $R = \sigma_{min}/\sigma_{max}$.

A very usual way of representing the effect of the mean stress on the fatigue behavior is through the Haigh diagram (or constant life diagram). This diagram allows the identification of alternating stress pairs, S_a , and mean, S_m , responsible for fatigue failure after the application of N load cycles. The ordinate of the Haigh diagram represents the stress amplitude, S_a , normalized by the fatigue limit, S_e , and the abscissa represents the mean stress normalized by the ultimate tensile strength, S_u (see Fig. 13). In theory, the curve that best fits the experimental data represents the combinations between S_a and S_m responsible for fatigue failure after N cycles. Since the tests required to generate a Haigh diagram is expensive and do not allow one to obtain a curve that

represents the typical behavior of the material (since the fitted curve will always depend on the quality of the experimental data), several empirical relations have been developed to generate the curve that defines the failure condition for a given life (typically this procedure is used to define the curve that represents the design conditions for infinite life). These methods use various curves to connect the endurance limit on the alternating stress axis to either the yield strength, S_y , ultimate strength, S_u , or true fracture stress, σ_f , on the mean stress axis.

The curve that connects these two points in each of the coordinate axes hypothetically represents the combinations S_a between S_m and responsible for fatigue failure after the application N of loading cycles. Eq. (21) presents a generalized expression to represent empirical relationships widely used to evaluate the effect of mean stress on fatigue behavior:

$$\left(\frac{S_a}{\sigma_f(N)}\right)^n + \left(f\frac{S_m}{S_u}\right)^m = 1 \tag{21}$$

where f , n and m are constants that assume different values according to the empirical relationship considered and S_u represents the tensile strength limit of the material. Without further details, the most adopted criteria can be derived considering the constants presented in Table 3.1.

Table 1: Constants that characterize the empirical relationships used to evaluate the effect of mean stress on fatigue behavior.

Relation	Parameter		
	m	n	f
Geber	2	1	1
Goodman	1	1	1
Morrow	1	1	$S_u/\tilde{\sigma}_f B$
Sailor	2	2	1
Dietmann	1	2	1

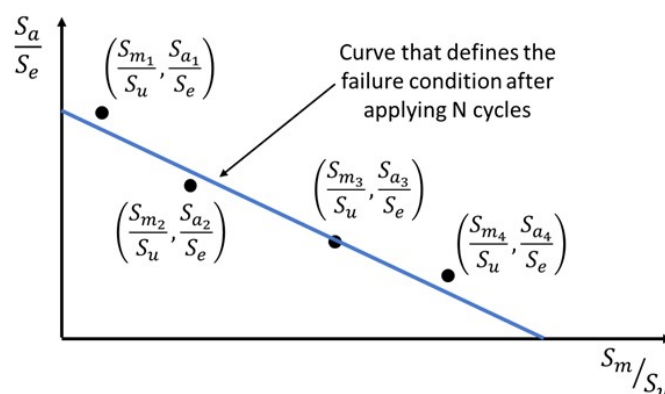


Figure 13: Schematic representation of the Constant-life diagram (Haigh diagram).

3.2 Accumulated Damage Concept and Miner Rule

The fatigue failure process in ductile metals is intrinsically related, at a first stage, to the nucleation of microcracks resulting from the movement of disagreements during cyclic loading. This cyclic movement leads to the appearance of thin steps on the surface (intrusions and extrusions) that act as stress amplifiers where micro cracks start. In a second step of the fatigue process, microcracks coalesce and begin to propagate through the material. With the continued cyclic loading,

the growth of the cracks progresses until the still intact section of the material can no longer withstand the imposed load and experiences a rapid fracture. All this process of degradation structural can only be monitored and measured in highly controlled laboratory environments. And in practical applications, the estimation of component lives subject to random loadings is made based on data from fatigue studies performed for constant amplitude loads, and the law of linear damage accumulation proposed by Miner.

In the formulation of the law of accumulation of Miner damage, it is considered that the damage, induced by a cycle, of the stress amplitude S_a is estimated by the following expression:

$$d = \frac{n}{N_f} \quad (22)$$

where N_f it represents the number of cycles required for a periodic load with a constant amplitude equal S_a to induce component failure in the structure.

Considering the history of stresses is composed K of stress amplitude classes, each being applied times n_k , the Miner rule is written as:

$$d = \sum_{k=1}^K d_k = \sum_{k=1}^K \frac{n_k}{N_{f_k}} \leq D_c \quad (23)$$

where D it represents the accumulated damage D_c and represents critical damage and must be determined experimentally.

In the absence of further information, it should be assumed that the failure will occur when the critical damage reaches a nominal value of 1,0. However, and all these studies show that the median values of damage observed in specimens at certain loading conditions vary from 0,15 to 1,6.

Example 3.1 A mechanical component was manufactured with steel that has a tensile strength limit, S_u , equal to 755 MPa and a fatigue behavior described by the following Wöhler function $S_a = 1440N^{-0.096}$. Under operating conditions, the stresses at the critical point of this component are applied in the form of a sequence of loading blocks with constant amplitudes, as shown in Fig. 14. Considering this information, time out the number of load blocks that can be applied until the component fails.

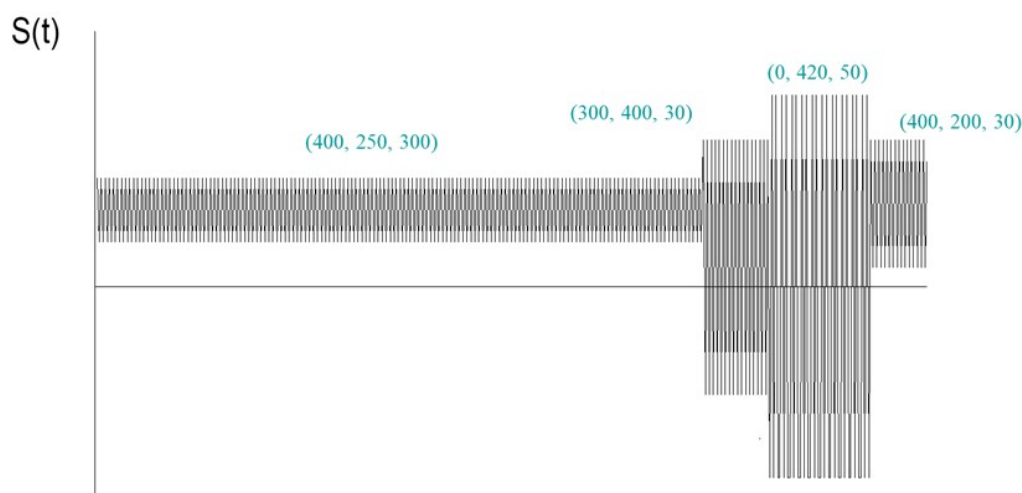


Figure 14: Sequences of block loading.

Solution:

Characteristics of the sinusoidal loading sequence S_m, S_a, n :

- (1) (400, 250, 300);
- (2) (300, 400, 30);
- (3) (0, 420, 50);
- (4) (400, 200, 30).

The estimate of the damage accumulated by this loading block will be performed considering Eq. (23). For this particular case, the sum of the damage increments will take the following form:

$$D_{block} = \frac{300}{N_f(250, 400)} + \frac{30}{N_f(400, 300)} + \frac{50}{N_f(420, 0)} + \frac{30}{N_f(200, 400)} \quad (24)$$

where $N_f(S_a, S_m)$ represents the number of cycles that could be applied if the pair of alternating stresses, S_a , and average, S_m were applied continuously until the failure of the mechanical component.

Considering, for example, the Goodman ratio the value of N_f will be calculated through the following equation:

$$N_f(S_a, S_m) = \left[\frac{S_a}{A \left(1 - \frac{S_m}{S_u}\right)} \right]^{\frac{1}{b}} = \left[\frac{S_a}{1440 \left(1 - \frac{S_m}{755}\right)} \right]^{-\frac{1}{0.0961}} \quad (25)$$

where S_u is the limit of tensile strength, and A and b represent the constant and exponent of the curve of Wöhler.

Substituting the mean and alternating stresses into Eq. (25), one arrives at the results presented in Table 2.

Table 2: Estimation of the damage produced by each constant amplitude load present in the loading block shown in Fig. 14

Event	S_m	S_a	n	$N_f(S_m, S_a)$	d
1	400	250	300	3,18E+04	9,43E-03
2	300	400	30	3,16E+03	9,48E-03
3	0	420	50	3,70E+05	1,35E-04
4	400	200	30	3,24E+05	9,25E-05

As a consequence of the application of a loading block, it is verified that the damage accumulated with each application of this load sequence is equal to 1.91E-3. Assuming that the number of blocks that can be applied is equal to:

$$Num.blocks = \frac{D_c}{D_{blocks}} = \frac{1}{1.91 \cdot 10^{-3}} = 52 \quad (26)$$

4 Random Fatigue Analysis in the time domain

As commented in Section 1 to Section 3, over their useful life, many structural components are subjected to cyclic loads. These loads can be applied under constant amplitude conditions or under amplitude conditions that vary over time. We also show that in situations where structural components are subjected to constant amplitude loads (with or without average component) or in conditions of application of loading blocks with varying amplitudes, the determination of fatigue life is a relatively simple task. However, if the loading amplitude varies at random over time, it is more difficult to characterize a loading cycle and its corresponding amplitude. This difficulty forced fatigue researchers to develop cycle identification and counting methodologies to reduce a complicated loading history in a series of simple and discrete constant amplitude loading events. These methods were initially developed for the study of fatigue damage generated in aeronautical structures and, over the last decades, several counting methods have been proposed, leading to different results and, therefore, to erroneous predictions of fatigue life. However, there was a consensus that the best approach is the procedure called Rainflow, introduced by Endo in the late 1960s, being the first article in English by Matsuichi and Endo [1968]. The procedures for using the Rainflow cycle counting method were standardized together with other counting methods in ASTM E1049-85 [2017] Standard Practices for Cycle Counting in Fatigue Analysis.

For the implementation of the rainflow method, it is necessary to initially transform the history of stresses into a sequence of peaks¹ and valleys², which represent the points at which the first derivative of the stress history changes sign (note that peaks and valleys are specific definitions for reversals), as illustrated in Fig. 15.

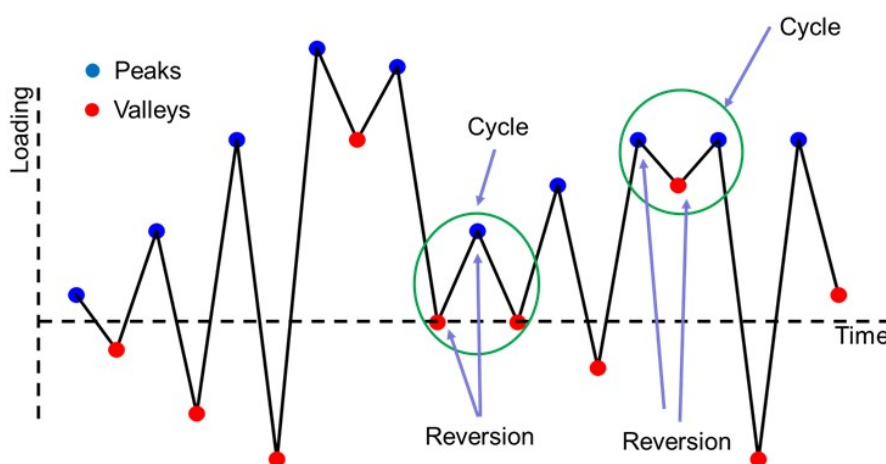


Figure 15: Identification of the basic parameters of a history used by the Rainflow method for the identification and counting of cycles.

From empirical observations of the stress-strain curve estimated from tests under variable amplitude conditions, it is verified that when the material is subjected to a history of stresses or deformations, small load reversals occur appear in the $\sigma - \epsilon$ diagram, but without affecting the large reversals and the response of the material. This behavior is represented in Fig. 16, in which it is observed that the loading sequence 2-3-4 generates the 2-3-2' cycle that does not interfere with

¹The point at which the first derivative of the load-time history changes from a positive to a negative sign; the point of maximum load in constant amplitude loading.

²The point at which the first derivative of the load-time history changes from a negative to a positive sign (also known as trough); the point of minimum load in constant amplitude loading.

the evolution of the 4-5 section. By analogy, the rainflow technique treats the same way the closed stress-strain curves (hysteresis cycles), in which the small reversals of load in a fatigue cycle with greater amplitude will also not interfere with the accumulated damage by fatigue.

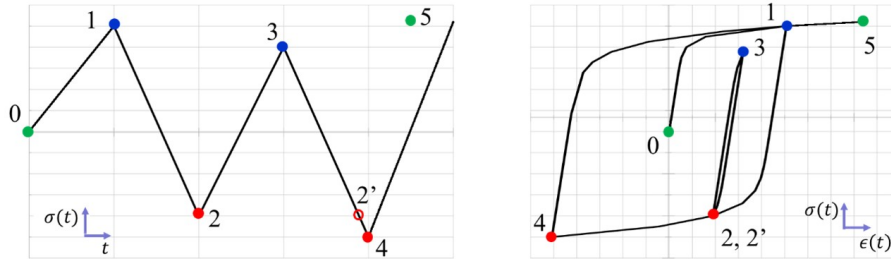


Figure 16: Example of the behavior of the stress-strain curve for a given stress history.

In essence, the Rainflow method identifies reversals and not loading cycles, but as can be seen in Fig. 15 and 16, two segments of equal amplitudes that have a common reversal represent a loading cycle.

A series of rules are used to identify load cycles in the Rainflow method. However, the core of the method of identifying the reversals present in history consists of comparing the bands formed by three consecutive reversals (S_i, S_j, S_k) , see Fig. 17. Thus, the three consecutive reversions (S_i, S_j, S_k) are used to define the stress ranges: $R_l = |S_i - S_j|$ (left track) and $R_r = |S_j - S_k|$ (right track). If the left band R_l , is less than or equal to the right band, the R_r material shall be considered to have been subjected to 1 or 1/2 loading cycles with alternating stress, S_a , and mean stress, S_m , calculated by the following expressions:

$$S_a = \frac{R_l}{2} \tag{27}$$

$$S_m = \frac{S_i + S_j}{2} \tag{28}$$

If the condition $R_l \leq R_r$ is not met, it is assumed that the loading (S_i, S_j, S_k) run does not induce loading cycles in the material.

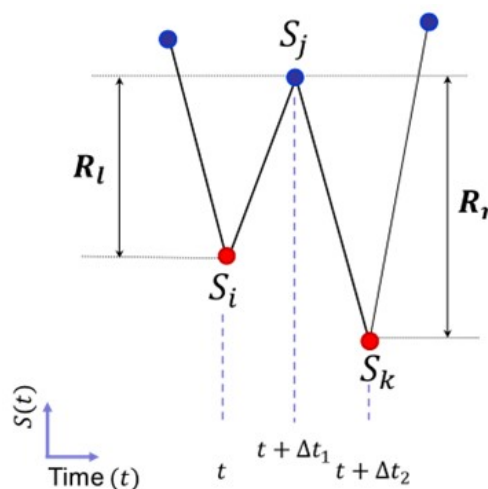


Figure 17: Primal criteria for the identification of a reversal according to the rainflow method.

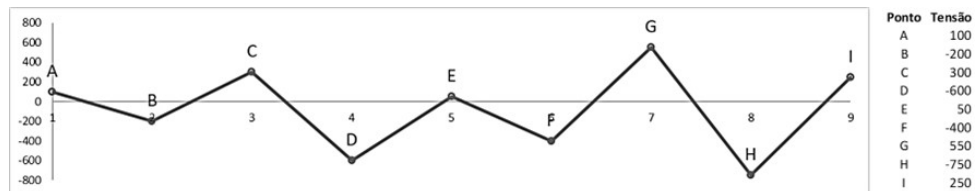
From the comparison between R_l and R_r , a few more rules are required to implement the Rainflow method.

1. Turn the loading history into a sequence of peaks and valleys, quantify the number of peaks and valleys present in the sequence, L , and set the counter l (at the beginning of the counting process $l = 1$).
2. Define an $l \times 3$ matrix in which each row will be stored the information about the alternating stress, S_a , the mean stress, S_m , and the number of reversals, n , identified during the cycle counting process.
3. The count is performed initially from the first position of the peak sequence or valley present in history.
4. If $L > 2$, form the first triple $(S_i, S_j, S_k) = (S_l, S_{(l+1)}, S_{(l+2)})$, $l = 1$.
5. Calculate the tracks R_l and R_r using the reversals (S_i, S_j, S_k) and compare tracks R_l and R_r .
 - 5.1. If $R_l \leq R_r$:
 - 5.1.1. If the reversal S_i is in the first position of the sequence of peaks and valleys ($l = 1$):
 - 5.1.1.1. Calculating S_a and S_m using Eq. (27) and Eq. (28) and transport S_a , S_m and $n = 1$ to the rainflow matrix.
 - 5.1.1.2. Delete point $S_{(l=1)}$ from the sequence of peaks and valleys, rearrange it, and make $L = L - 1$.
 - 5.1.1.3. If $L > 2$:
 - 5.1.1.3.1. Form the term $(S_i, S_j, S_k) = (S_1, S_2, S_3)$ and go to Step 5.
 - 5.1.1.4. If $L = 2$:
 - 5.1.1.4.1. Calculating S_a and S_m using Eq. (27) and Eq. (28) and transport S_a , S_m and $n = 1$ reversion to the rainflow matrix.
 - 5.1.1.4.2. End counting process.
 - 5.1.2. If the reversal S_i is not in the first position of the sequence of peaks and valleys ($l \neq 1$):
 - 5.1.2.1. Calculate S_a and S_m using Eq. (27) e Eq. (28) and transport the values of S_a , S_m , $n = 2$ and reversals to the rainflow matrix
 - 5.1.2.2. Delete points S_i and S_j from the sequence, rearrange the sequence, make $L = L - 2$, $l = l - 1$.
 - 5.1.2.3. Form the triple $(S_i, S_j, S_k) = (S_l, S_{(l+1)}, S_{(l+2)})$ and go to Step 5.
 - 5.2. If $R_l > R_r$
 - 5.2.1. If $l \leq L - 3$.
 - 5.2.1.1. $l = l + 1$
 - 5.2.1.2. Form the triple $(S_i, S_j, S_k) = (S_1, S_2, S_3)$ and go to Step 4
 - 5.2.2. If $l > L - 3$.
 - 5.2.2.1. Calculate for each remaining range (S_i, S_j) S_a and S_m and transport the values to S_a , S_m and $n = 1$ to the rainflow matrix
 - 5.2.2.2. Finish the counting process.

Example 4.1 Identify the carga cycles present in the stretch of the loading section shown in Fig. 18

Solution:

The Fig. 19 presents the sequence of steps required to implement rainflow count of the history illustrated in Fig. 18, that presents the result of the cycle count by applying the rainflow algorithm to loading.



Events		Stress		S_a	S_m	n
Start	End	Start	End			
A	B	-2	1	1.5	-0.5	1
B	C	1	-3	2	-1	1
E	F	-1	3	2	1	2
C	D	-3	5	4	1	1
D	G	5	-4	4.5	0.5	1
G	H	-4	4	4	0	1

Figure 18: Example of a service load-time history and the results of the cycle counting

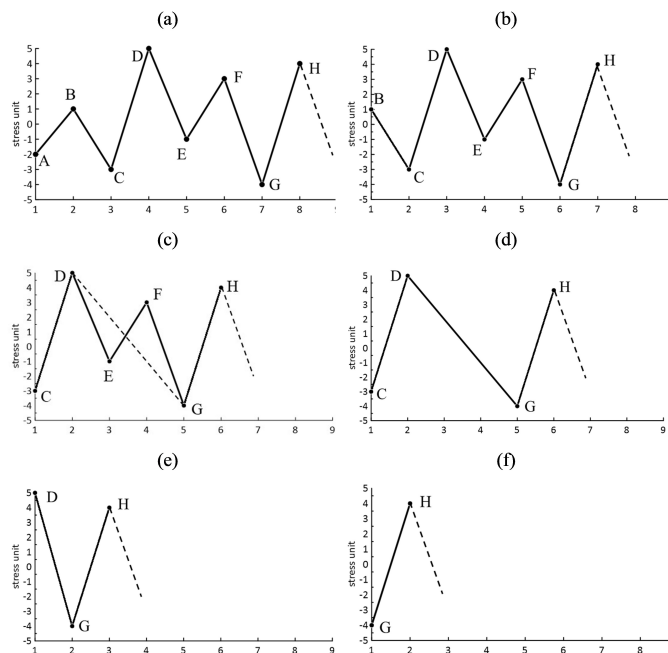


Figure 19: Example 1 of the rainflow counting method using the rules presented in this chapter.

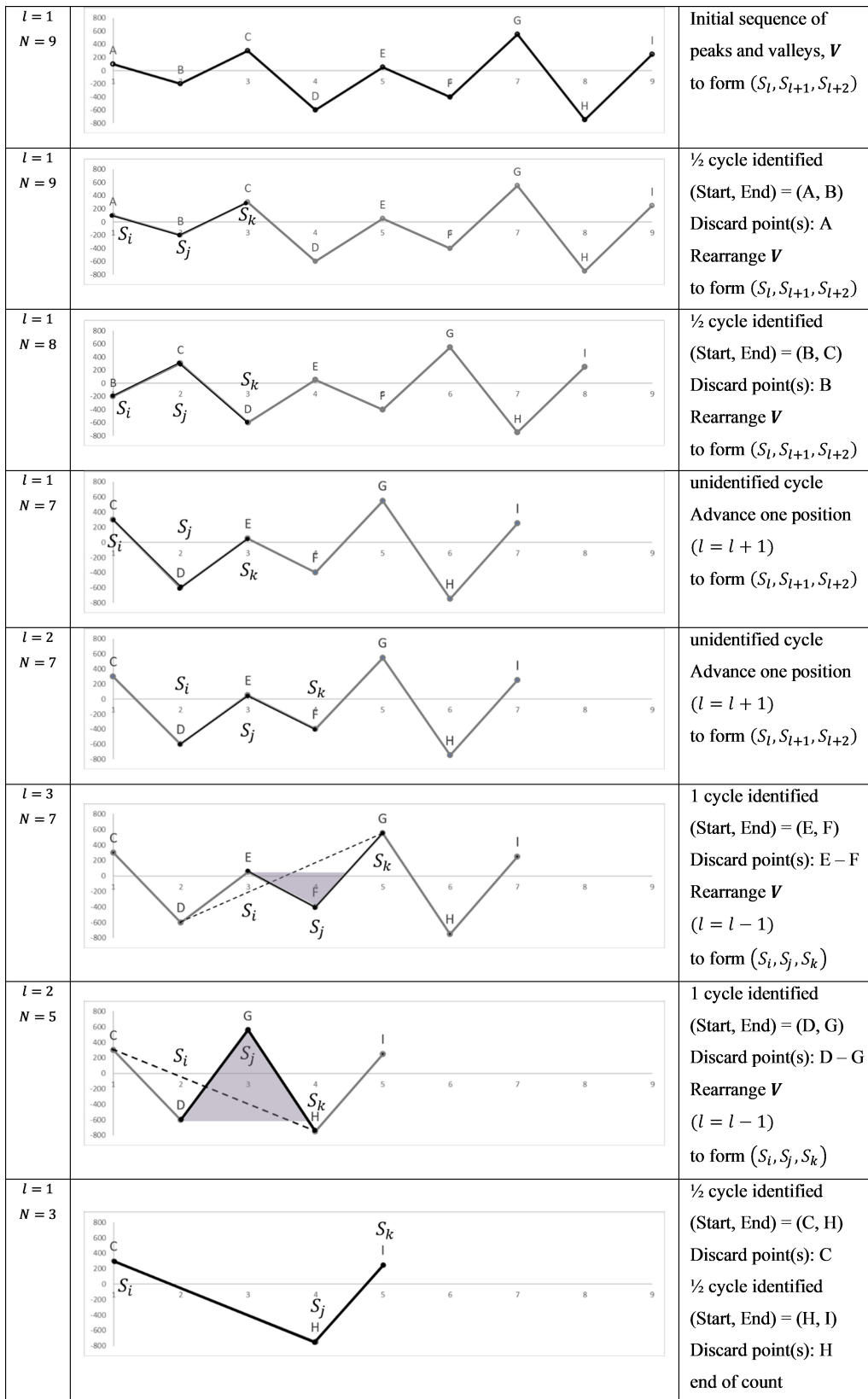


Figure 20: Example 2 of the rainflow counting method using the rules presented in this chapter.

5 Random Fatigue Analysis in the Frequency Domain

5.1 Spectral Properties

Fatigue analysis is often performed in the time domain, where all input and output loading, stress response or strain, are time-based signals. In some situations, however, the stress response and input loading are preferably expressed as frequency-based signals, usually in the form of a power spectral density (PSD) graph. In this case, a system function is required to relate a loading input PSD to the response output PSD Lee [2005].

For linear systems, fatigue calculations performed entirely in the frequency domain can shorten the calculation time considerably, which corresponds to the so-called frequency domain approach (Dirlik and Benasciutti [2021])

The PSD represents the energy of the time signal at different frequencies and is another way to denote the charging signal in the time domain. The Fast Fourier Transform (FFT) of a signal in time can be used to obtain the loading PSD, while the Inverse Fourier Transform (IFT) can be used to transform the frequency-based signal for time-based loading. The transformation of the loading history between the time domain and the frequency domain is subject to certain requirements, according to which the signal must be stationary, random, and Gaussian, i.e., not varying its statistical properties over time and presenting normal distribution (Lee [2005]). An ergodic process is a stationary process whose expected value (envelope) statistics are identical to time statistics (Dirlik [1985]).

In practice, a one-sided spectral density, $G(f)$, is often used, limited only to the positive frequency range of the two-sided spectral density, $S(f)$, as negative frequencies have no physical meaning (Bishop [1988]). A random signal can be classified based on its frequency content, i.e., the shape of its PSD. The signal is called narrowband if the PSD peaks around a single frequency, usually the resonant frequency of a vibrating system. In all other cases where the PSD covers a wider range of frequencies, the random signal is termed broadband (Dirlik and Benasciutti [2021]).

Fig. 21 compares sample histories representative of two distinct types of random processes, showing the differences between the time histories based on their corresponding PSD. Power spectra with rectangular blocks were considered which are just idealizations of the smoother spectra typically found in practical applications.

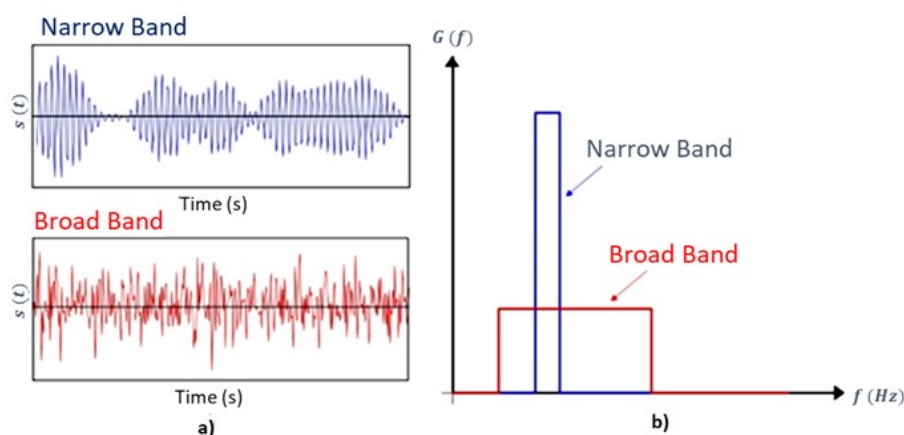


Figure 21: a) Time history samples b) Narrowband and broadband PSD of the history samples.

The narrowband or broadband process is defined by means of several spectral parameters,

among them we define the spectral moments of n-order, λ_n , as:

$$\lambda_n = \int_{-\infty}^{+\infty} |f|^n S(f) df = \int_{-\infty}^{+\infty} f^n G(f) df \quad , \quad n = 0, 1, 2, \dots \quad (29)$$

The zero-order moment, λ_0 , can be used as an alternative method of calculating the root mean square (RMS) value, which is a good indication of the intensity of a loading history.

$$RMS = \sigma_S = \sqrt{\lambda_0} \quad (30)$$

and also represents an important feature in the time domain, which is the uniaxial loading history variance $s(t)$:

$$\lambda_0 = \int_{-\infty}^{+\infty} G(f) df = Var(s(t)) \quad (31)$$

From the spectral moments, some spectral characteristics can be calculated, such as the expected rate of zero upcrossings v_0 per unit of time.

$$v_0 = \sqrt{\frac{\lambda_2}{\lambda_0}} \quad (32)$$

and the expected level upcrossing rate v_p per unit of time.

$$v_p = \sqrt{\frac{\lambda_4}{\lambda_2}} \quad (33)$$

The irregularity factor will essentially measure how many peaks a signal has for each time the signal crosses at zero stress level (Lee [2005]). The ratio between the zero-crossing rate and the peak-crossing rate is known as the irregularity factor α .

When the irregularity factor tends to one, it means that for each peak, there will be approximately one zero level crossing and the process will have a narrow bandwidth, that is, v_0 is equal to v_p . However, when the bandwidth is wide, it is said that the irregularity factor will tend to zero, thus there will be infinite peaks for each zero-level crossing, characterized by a more irregular process (Lee [2005]).

The irregularity factor, which varies between 0 and 1, can be used to characterize the shape and geometric distribution of the frequency of a given spectrum as a function of spectral moments (Benasciutti and Tovo [2005]). In general, it can be calculated as:

$$\alpha_n = \frac{\lambda_n}{\sqrt{\lambda_0 \lambda_{2n}}} \quad (34)$$

The process is considered narrowband if its spectral density has only a narrow range of frequencies. In contrast, a broadband process is one whose spectral density covers a wide band of frequencies, this bandwidth is called the width parameter (Lee [2005]). The spectral width parameter, γ , is introduced as:

$$\gamma = \sqrt{1 - \alpha^2} \quad (35)$$

Thus, narrowband processes must have an irregularity factor equal to 1 and a width parameter equal to 0, while wideband processes must have an irregularity factor approaching 0 and a spectral width parameter approaching 1 (Lee [2005]).

5.2 Fatigue Damage Based on Power Spectral

For the fatigue life of gaussian and stationary processes, the Miner rule will be used to estimate the fatigue damage per unit time from the amplitude distribution of the stress cycles present in the loading history, $f_{(S_a)}(S_a)$. This is achieved by following the following procedures:

Assume a history of stresses whose amplitudes vary over time, $S(t)$. Through the rainflow counting technique, for example, this variable-amplitude load is transformed into a sequence of constant-amplitude blocks, S_{a_i} . So, using the Palmgren-Miner, the damage D can be estimated through the following equation:

$$D = \sum_{i=1}^K \frac{n_i}{N_{f_i}} \tag{36}$$

where n_i is the total number of cycles in the i th block of constant-stress amplitude $\sigma_{a,i}$, N_{f_i} is the number of cycles to failure under $\sigma_{a,i}$, and I is the total number of blocks. Recall that the relationship between constant stress amplitude $\sigma_{a,i}$ (or constant stress range $\sigma_{r,i}$) and the fatigue life N_{f_i} has the following expression:

$$\sigma_{a,i} = AN_{f_i}^{-b} \tag{37}$$

$$N_{f_i} \sigma_{a,i}^m = k \tag{38}$$

where A is the fatigue strength coefficient and b is the fatigue strength exponent. If $m = 1/b$, $k = A^m$.

The amplitude distribution of the stress amplitude present in the loading history, $f_{\sigma_a}(\sigma_a)$ can be established performing the rainflow cycle counting technique (see Fig. 22).

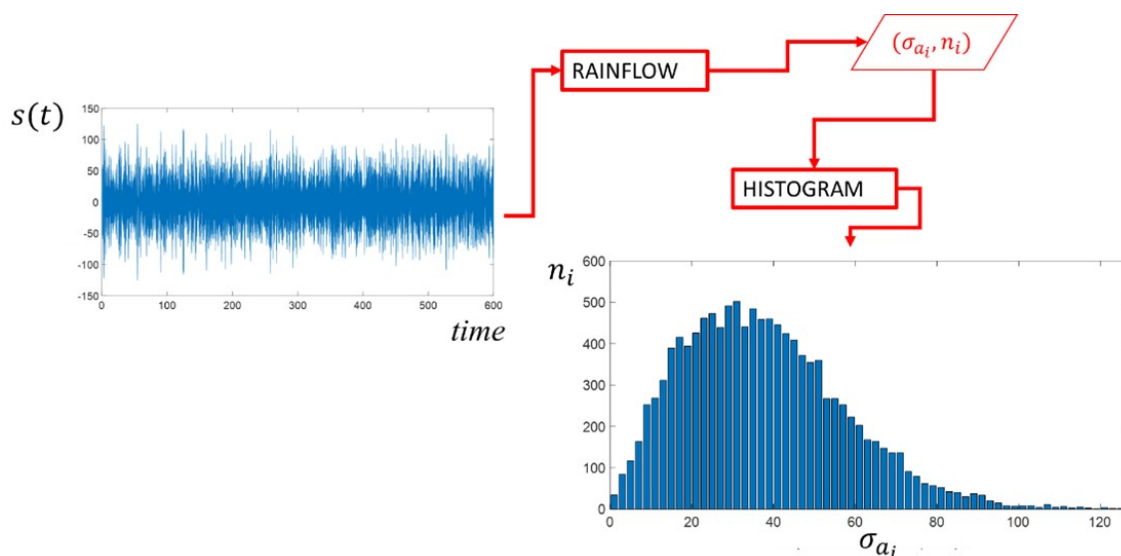


Figure 22: Basic strategy for obtaining the histogram representative of the distribution of stress amplitudes present in a random loading history.

Assuming that the total number of stress cycles $\sigma_{a,i}$ present in the process $S(t)$ is calculated as $N_{total} = \sum_{i=1}^I n_i$, the probability f_i that a stress amplitude, $\sigma_{a,i}$, occur will be approximated by the relation:

$$f_i = \frac{n_i}{N_{total}} \tag{39}$$

So, substituting Eq. (39) in Eq. (36), an approximation for the total fatigue damage can be written as:

$$D = \sum_{i=1}^I \frac{f_i N_{total}}{N_{f_i}} \quad (40)$$

By using the linear S–N model (Eq. (37) and Eq. (38)), the expression for fatigue damage is:

$$D = \frac{N_{total}}{k} \sum_{i=1}^I \sigma_{a_i}^m f_i \quad (41)$$

Recalling the concept of n-order statistical moment, we can express the sum of Eq. (41) as:

$$E(\sigma_{a_i}^m) = \sum_{i=1}^I \sigma_{a_i}^m f_i \quad (42)$$

As a result, the Eq. (41) will take the following form:

$$D = \frac{N_{total}}{k} E(\sigma_{a_i}^m) \quad (43)$$

Assume that the PSD of stress amplitude σ_a can be treated as a continuous random variable, the expected value of σ_a is:

$$E(\sigma_{a_i}^m) = \int_0^{\infty} \sigma_a^m f_{\sigma_a}(\sigma_a) d\sigma_a \quad (44)$$

That is, the total accumulated damage can be calculated using the following general equation:

$$D = \frac{N_{total}}{k} \int_0^{\infty} \sigma_a^m f_{\sigma_a}(\sigma_a) d\sigma_a \quad (45)$$

5.2.1 Fatigue damage under narrowband Gaussian loading

If the loading history is stationary narrowband Gaussian, the probability density function of the stress amplitude, $f_{\sigma_a}(\sigma_a)$, follows a Rayleigh distribution. Consequently, the total accumulated damage will be calculated using the following equation:

$$D = \frac{N_{total}}{K} E(\sigma_{a_i}^m) = \frac{N_{total}}{K} (\sqrt{2\lambda_0})^m \Gamma\left(\frac{m}{2} + 1\right) \quad (46)$$

where, $\Gamma(\cdot)$ is the gamma function, and λ_0 is the variance of $S(t)$. For narrowband histories, N_{total} can be obtained as the product of the zero-crossing rate, v_0 , and T is the total lifetime.

Finally, the damage caused by the history of narrowband loading and the expected life of the component can be obtained through the equations:

$$D_{NB} = \frac{v_0 T}{K} (\sqrt{2\lambda_0})^m \Gamma\left(\frac{m}{2} + 1\right) \quad (47)$$

$$T_{NB} = \frac{K D_{NB}}{v_0 (\sqrt{2\lambda_0})^m \Gamma\left(\frac{m}{2} + 1\right)} \quad (48)$$

5.2.2 Fatigue damage under broadband Gaussian loading

When the spectral characteristics of the stress history do not have characteristics of a narrow-band Gaussian process, the probability density function of the stress amplitudes, $f_{\sigma_a}(\sigma_a)$, does not follow a Rayleigh distribution.

In such situations, the function $f_{\sigma_a}(\sigma_a)$ tends to be dependent on the irregularity factor, α , as illustrated in Figure 5.3. This Figure presents 3 Gaussian histories with their respective probability density functions, $f(S(t))$ and $f_{\sigma_a}(\sigma_a)$, and spectral density function, $G(f)$. Looking at this figure, it can be noted that only the functions $f_{\sigma_a}(\sigma_a)$ and $G(f)$ vary with the irregularity factor, α . This implies that for a broadband process the relation of the peak distribution and cycle amplitudes is much more complex. For this reason, several empirical solutions have been proposed. Below we present some of the many models available in the literature.

Wirsching–Light method

Wirsching and Light [1980] used an additional spectral width parameter α_2 (see Eq. (34)) to correct the narrow-band approximation with the empirical factor ζ :

$$D_{BB} = \gamma D_{NB} \tag{49}$$

where D_{BB} is the damage accumulated due to the application of a broadband process (our wide-band process), D_{NB} is the damage accumulated is the accumulated damage calculated through Eq. (47) and ζ is the correction factor, calculated using the following equation:

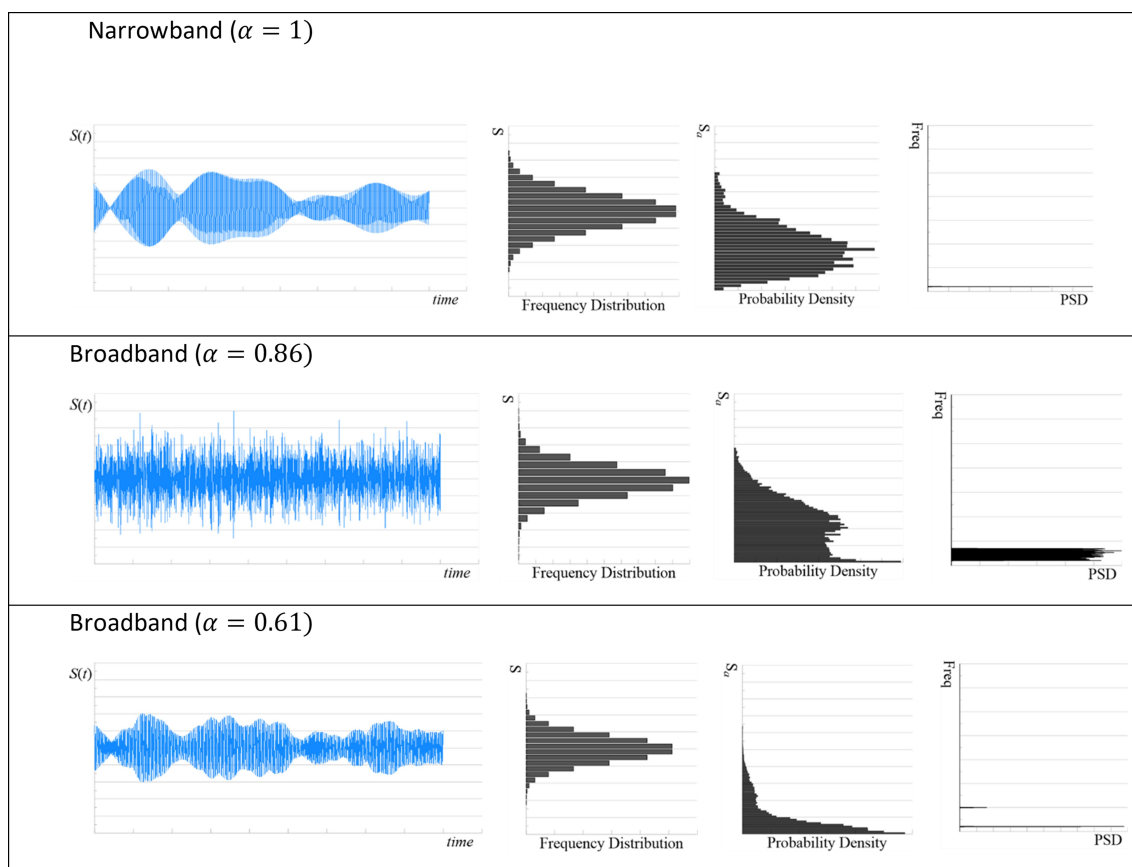


Figure 23: Characteristics of the Gaussian loading as a function of the irregularity factor.

$$\zeta = a + [1 - a](1 - \gamma)^b \tag{50}$$

$$a = 0,926 - 0,033k \tag{51}$$

$$b = (-1) \cdot (2.323) + 1.587k \tag{52}$$

where γ is the spectral width parameter, $\gamma = \sqrt{1 - \alpha^2}$, and a (Eq. (51)) and b (Eq. (52)) are the best fitting parameters dependent on the slope of the S–N curve, k (see Eq. (19)).

Tovo–Benasciutti method

Benasciutti and Tovo [2005] first proposed an approach where the fatigue life is calculated as a linear combination of the upper and lower fatigue-damage intensity limits. The final expression for the estimated fatigue-damage intensity is:

$$D_{TB} = \eta_{NB} D_{NB} \tag{53}$$

where $\eta_{TB} = w + (1 - w)\alpha_2^{(k-1)}$ considered as a narrowband damage correction factor. The weighting factor w is conditioned to the power spectrum through the parameters of the irregularity factor as follows:

$$w = \frac{(\alpha_1 - \alpha_2) [1, 112(1 + \alpha_1\alpha_2 - (\alpha_1 + \alpha_2))e^{2,11\alpha_2} + (\alpha_1 - \alpha_2)]}{(\alpha_2 - 1)^2} \tag{54}$$

The analysis of Eq. (53) indicates that, indicates that, according to the model of Tovo and Benasciutti, the damage estimate is linked to four spectral moments through parameters α_1 and α_2 , showing a clear dependence of the spectral properties.

Dirlik method

Dirlik [1985] based on extensive Monte Carlo simulations of load histories, it developed a closed form empirical expression from the approximation of the probability density function of the "rainflow amplitude" as the sum of an Exponential distribution and two Rayleigh distributions. That is, the Dirlik damage model for a time period studied T is shown by the equation:

$$E[D] = D_{DK} = \frac{v_p T}{C} \int_0^\infty S_a^k f_{S_a}(S_a) dS_a \tag{55}$$

where the probability density function, $f_{S_a}(S_a)$, takes the following form:

$$f_{DK}(S_a) = \frac{1}{\sigma_s} \left(\frac{D_1}{Q} e^{\left(\frac{-Z}{Q}\right)} + Z \frac{D_2}{R^2} e^{\left(\frac{-Z^2}{2R^2}\right)} + Z D_3 e^{\left(\frac{-Z^2}{2}\right)} \right) \tag{56}$$

where the function parameters are defined as:

$$Z = \frac{S_a}{\sqrt{M_0}} \tag{57}$$

$$X_m = \frac{M_1}{M_0} \sqrt{\frac{M_2}{M_4}} \tag{58}$$

$$D_1 = \frac{2(X_m - \gamma^2)}{1 + \gamma^2} \tag{59}$$

$$R = \frac{\gamma - X_m - D_1^2}{1 - \gamma - D_1 + D_1^2} \tag{60}$$

where Z is the normalized stress amplitude (S_a) and X_m, D_1, D_2, D_3, R and Q are best-fitting parameters. Substituting these parameters in Eq. (53), the expected value of the cumulative damage over a time T is estimated by Eq. (61) - The steps necessary to obtain this formula are presented in Appendix B.

$$D_{DK} = \frac{N_{DK}}{C} \sigma_S^k \left[D_1 Q^k \Gamma(1+k) + 2^{\left(\frac{k}{2}\right)} \Gamma\left(1 + \frac{k}{2}\right) (D_2 |R|^k + D_3) \right] \tag{61}$$

in which N_{DK} is the total number of cycles in $S(t)$ over a period of time T ($N_{DK} = v_P T$).

Zhao and Baker method

With the objective to describe the stress range distribution, Zhao and Baker [1992] have proposed that the peak distribution function can be approximated by a linear combination between the distribution functions of Weibull and Rayleigh, with parameters directly estimated from the power spectral density function, PSD, of the process.

$$f_{ZB}(Z) = \omega \beta \lambda Z^{\beta-1} e^{-\lambda Z^\beta} + (1 - \omega) Z^{\frac{Z^2}{2}} \tag{62}$$

where Z is the normalized amplitude defined above, ω is a weighting factor ($0 \leq \omega \leq 1$), and b, λ are the Weibull parameters ($b > 0, \lambda > 0$). As commented, previous parameters depended on the spectral density of process $S(t)$ and are approximated on the basis of both simulation results and theoretical arguments. Specifically, β, λ , and ω , are defined as:

$$\beta = \begin{cases} 1.1, & \text{if } \gamma \leq 0.9 \\ 1.1 + 9(\gamma - 0.9), & \text{if } \gamma > 0.9 \end{cases} \tag{63}$$

$$\lambda = 8 - 7\gamma \tag{64}$$

$$\omega = \frac{1 - \gamma}{1 - \Gamma\left(1 + \frac{1}{\beta}\right) \lambda^{\frac{-1}{\beta}} \sqrt{\frac{2}{\pi}}} \tag{65}$$

Thereby, considering this hypothesis, a Gaussian process, $S(t)$, with average zero, the expected value of the cumulative damage over a time T is estimated by Eq. (66).

$$E[D] = D_{ZB} = \frac{N_{ZB}}{A} \sigma_s^m \left[\omega \lambda^{\frac{-m}{\beta}} \Gamma\left(1 + \frac{m}{\beta}\right) + (1 - \omega) 2^{\frac{m}{2}} \Gamma\left(1 + \frac{m}{2}\right) \right] \tag{66}$$

in which N_{ZB} is the total number of cycles in $S(t)$ over a period of time T ($N_{ZB} = v_P T$).

Example 5.1 A hot-rolled component made of 1045 steel Q&T (BHN=410) is subjected to random loading process. The fatigue strength coefficient, S_f , and the fatigue strength exponent, b , are 3906 MPa and -0.167, respectively. The stress response at a critical location is calculated in terms of the PSD in Fig. 24. The PSD has a frequency band varying between 2 and 6 Hz with $G(f) = 12,500$ and another one varying between 10 and 20 Hz with $G(f) = 5,000$. Determine the fatigue life of this component using fatigue prediction models under wide-band Gaussian loading.

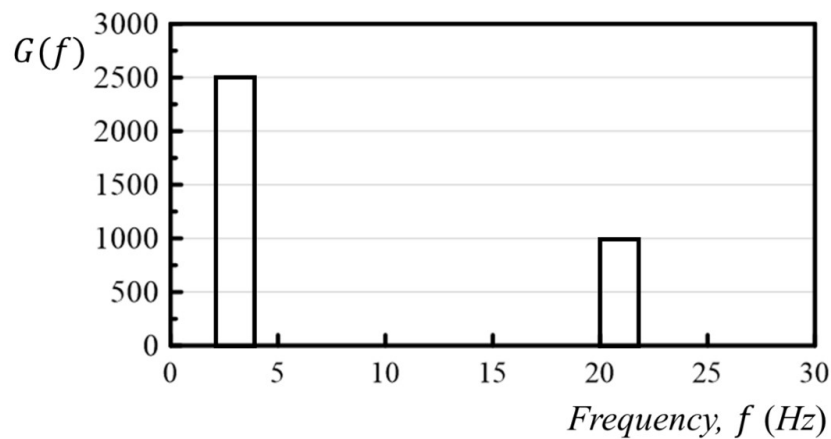


Figure 24: PSD of the stress response of a component.

Development:

We will first calculate the component fatigue damage in the frequency domain and then we will make the life estimation considering the time domain method (Rainflow) that we will use as a reference standard to compare the spectral methods. For each model we will calculate the cumulative damage per second, $\Delta_{(**)}$, and then calculate the total lifetime of the component, $T_{(**)}$.

Definition of S-N curve parameters used in spectral models:

The parameters of the S–N curve of the material presented in the test follow Eq 5.10a:

$$\sigma_{a_i} = AN_{f_i}^{-b} \quad (67)$$

Whereas the material S–N curve for vibration fatigue usually follows the following expression:

$$N_{f_i} \sigma_{a_i}^m = k \quad (68)$$

where A is the fatigue strength coefficient and b is the fatigue strength exponent. And:

$$m = \frac{1}{b} \quad (69)$$

$$k = A^m \quad (70)$$

Thus, from these considerations we will find the following values for the constants C and k :

$$k = 5,988 \quad (71)$$

$$C = 3,1216 \cdot 10^{21} [\text{cycle} \cdot \text{MPa}^k] \quad (72)$$

Determination of spectral moments of n-order, M_n :

The parameters used in the methods based on the frequency domain are obtained from the spectral moments extracted from the PSD of the load history acting on the component. Such moments are easily calculated using the following expressions:

$$M_0 := \int_{L1_i}^{L1_f} f^0 \cdot G_1(f) df + \int_{L2_i}^{L2_f} f^0 \cdot G_2(f) df = 7 \cdot 10^3$$

$$M_1 := \int_{L1_i}^{L1_f} f^1 \cdot G_1(f) df + \int_{L2_i}^{L2_f} f^1 \cdot G_2(f) df = 5.7 \cdot 10^4$$

$$M_2 := \int_{L1_i}^{L1_f} f^2 \cdot G_1(f) df + \int_{L2_i}^{L2_f} f^2 \cdot G_2(f) df = 9.293 \cdot 10^5$$

$$M_3 := \int_{L1_i}^{L1_f} f^3 \cdot G_1(f) df + \int_{L2_i}^{L2_f} f^3 \cdot G_2(f) df = 1.871 \cdot 10^7$$

$$M_4 := \int_{L1_i}^{L1_f} f^4 \cdot G_1(f) df + \int_{L2_i}^{L2_f} f^4 \cdot G_2(f) df = 3.912 \cdot 10^8$$

where $L1_i = 2Hz$, $L1_f = 4Hz$, $L2_i = 20Hz$, $L2_f = 22Hz$, $G_1(f) = 2,500$ and $G_2(f) = 1,000MPa^2/Hz$.

Determination of spectral Parameters:

“root mean square”	$RMS := \sqrt{M_0} = 83.666$
“zero crossing rate”	$\nu_0 := \sqrt{\frac{M_2}{M_0}} = 11.522$
“Peak occurrence rate”	$\nu_p := \sqrt{\frac{M_4}{M_2}} = 20.518$
“Irregularity Factor”	$\alpha_2 := \frac{\nu_0}{\nu_p} = 0.562$
“Spectral Width Parameter”	$\gamma := \sqrt{1 - \alpha_2^2} = 0.827$

Applying the Methods:

1. Narrowband Method

$$\Delta_{NB} := \frac{\nu_0}{C} \cdot (2 \cdot M_0)^{\frac{k}{2}} \cdot \Gamma\left(\frac{k}{2} + 1\right) = 5.529 \cdot 10^{-8}$$

$$\Delta_{NB} = 5.529 \cdot 10^{-8} \quad T_{NB} := \frac{1}{\Delta_{NB} \cdot 24 \cdot 3600} = 209$$

2. Wirsching and Light Method

$$a := 0.929 - 0.033 \cdot k = 0.731$$

$$b := -2.323 + 1.587 \cdot k$$

$$\zeta := a + (1 - a) \cdot (1 - \gamma^2) = 0.816$$

$$\Delta_{WL} := \zeta \cdot \Delta_{NB} = 4.512 \cdot 10^{-8}$$

$$\Delta_{WL} = 4.512 \cdot 10^{-8} \quad T_{WL} := \frac{1}{\Delta_{WL} \cdot 24 \cdot 3600} = 257$$

3. Tovo–Benasciutti Method

$$\alpha_1 := \frac{M_1}{\sqrt{M_0 \cdot M_2}} = 0.707 \quad \alpha_2 := \frac{M_2}{\sqrt{M_0 \cdot M_4}} = 0.562$$

$$w := \frac{(\alpha_1 - \alpha_2) \cdot ((1.112 \cdot (1 + \alpha_1 \cdot \alpha_2 - (\alpha_1 + \alpha_2))) \cdot e^{2.11 \cdot \alpha_2} + (\alpha_1 - \alpha_2))}{(\alpha_2 - 1)^2} = 0.463$$

$$\eta_{TB} := w + (1 - w) \cdot \alpha_2^{k-1} = 0.493$$

$$\Delta_{TB} := \eta_{TB} \cdot \Delta_{NB} = 2.725 \cdot 10^{-8}$$

$$\Delta_{TB} = 2.725 \cdot 10^{-8} \quad T_{TB} := \frac{1}{\Delta_{TB} \cdot 24 \cdot 3600} = 425$$

4. Zhao and Baker Method

$$\beta(\gamma) := \begin{cases} \text{if } \gamma \leq 0.9 \\ R - 1.1 \\ \text{else} \\ R - 1.1 + 9 \cdot (\gamma - 0.9) \\ R \end{cases} = 1.1 \quad \lambda = 8 - 7 \cdot \gamma$$

$$\omega := \frac{1 - \gamma}{1 - \Gamma\left(1 + \frac{1}{\beta(\gamma)}\right) \cdot \lambda^{\frac{1}{\beta(\gamma)}} \cdot \sqrt{\frac{2}{\pi}}} = 0.558$$

$$\Delta_{ZB} := \frac{\nu_r \cdot M_0^{\frac{k}{2}}}{C} \cdot \left(\omega \cdot \lambda^{\frac{k}{\beta(\gamma)}} \cdot \Gamma\left(1 + \frac{k}{\beta(\gamma)}\right) + (1 - \omega) \cdot 2^{\frac{k}{2}} \cdot \Gamma\left(1 + \frac{k}{2}\right) \right) = 4.361 \cdot 10^{-8}$$

$$\Delta_{ZB} = 4.361 \cdot 10^{-8} \quad T_{ZB} := \frac{1}{\Delta_{ZB} \cdot 24 \cdot 3600} = 265$$

5. Dirlik Method

$$X_m := \frac{M_1}{M_0} \cdot \sqrt{\frac{M_2}{M_4}} = 0.397 \quad D_1 := \frac{2 \cdot (X_m - \gamma^2)}{1 + \gamma^2} = 0.124$$

$$R := \frac{\gamma - X_m - D_1^2}{1 - \gamma - D_1 + D_1^2} = 0.453 \quad D_2 := \frac{1 - \gamma - D_1 + D_1^2}{1 - R} = 0.603$$

$$D_3 := 1 - D_1 - D_2 = 0.273 \quad Q := \frac{1.25 \cdot (\gamma - D_3 - R \cdot D_2)}{D_1} = 0.155$$

$$\Delta_{Dk} := \frac{\nu_r \cdot M_0^{\frac{k}{2}}}{C} \cdot \left(D_1 \cdot Q^k \cdot \Gamma(1 + k) + 2^{\frac{k}{2}} \cdot \Gamma\left(1 + \frac{k}{2}\right) \cdot (D_2 \cdot |R|^k + D_3) \right) = 2.742 \cdot 10^{-8}$$

$$\Delta_{Dk} = 2.742 \cdot 10^{-8} \quad T_{Dk} := \frac{1}{\Delta_{Dk} \cdot 24 \cdot 3600} = 422$$

Applying the time domain method:

A basic routine for predicting fatigue life using the time domain method is presented in Appendix 1. This routine uses together the rainflow counting technique and Miner’s rule. White noise filtering was used to generate the load history. As a result of using this methodology, the following results were found: $\Delta_{TD} = 2.8649e^{-08}$ and $T_{TD} = 404$ dias.

Results Comparison:

Table 3: Table of Results Comparison.

Frequency Domain Method	$\Delta_{(**)}$	$T_{(**)}$	Error in relation to the time domain method (%)
NarrowBand	5.53E-08	209	-48,3
Wirsching and Light	5.51E-08	257	-36,4
Tovo–Benasciutti	2.73E-08	425	5,2
Zhao and Baker	4.36E-08	265	-34,4
Dirlik	2.74E-08	422	4,5

Comparing the results of the frequency-based life prediction models with the result obtained using the time-based life prediction model (gold standard), it is verified that the methods of Dirlik and of Tovo and Benasciutti are the methods that present results very close to those obtained by the methodology considered gold standard.

6 Introduction: Multiaxial fatigue e Maximum Variance Method (MVM)

The life prediction for mechanical components subject to the phenomenon of fatigue has evolved with new fatigue models to ensure reliability and safety in mechanical system designs. According to Papadopoulos [2001], it is of greater interest to industries to build projects that ensure high cycle fatigue strength and, consequently, a finite fatigue life around thousands or millions of loading cycles.

Multiaxial fatigue is a general term that can be used to describe a structural failure process associated with mechanical component loads and geometry that are responsible for generating complex states of stress and strain in preferred planes, which induce the initiation and propagation of cracks. Therefore, evaluating mechanical components against fatigue is a complex problem that must be handled in an appropriate manner during a mechanical design in order to avoid catastrophic failures.

There are several examples of engineering scenarios that induce the multiaxial fatigue process, such as mechanical structures subject to cyclic loads (aircraft in flight, wind regime in wind turbines, wave action in offshores, liquid flow and vibrations in industrial pipes, etc. ..) and the presence of a notch in the component geometry that can change the stress state to multiaxial.

Over the years, in order to understand the multiaxial fatigue failure mechanism, three main types of approach have been developed: Critical Plane Method, Dissipated Energy Method and Mesoscopic Deformation Models. In this chapter, the multiaxial fatigue models will be discussed, which are based on the critical plane approach and consider that the cracks, originated due to material fatigue, originate in certain material planes where the combination of normal and shear stresses or strains are extremely severe. .

Among the varieties of experimentally validated solutions for mechanical problems involving multiaxial fatigue, the Maximum Variance Method (MVM) hypothesizes that the critical plane coincides with the material plane that experiences the maximum variance of a given driving force, such as the maximum normal stress and the maximum amplitude of resolved shear stress in a given direction. This simplification allows to work with one-dimensional quantities if the problem is treated in terms of the shear stress in the direction of maximum variance $\tau_{MV}(t)$ and the stress normal $\sigma_n(t)$. Based on modern studies of variable amplitude cyclic loading, the next topics will present MVM-based multiaxial fatigue criteria:

- **MVM based on equivalent normal stress and power density function of equivalent normal stress:** consists in the analysis of load history in the time domain and spectral analysis in the frequency domain in order to compare the critical planes and obtain the equivalent

normal stress and the power spectral density function equivalent, as shown in Table 4. In addition, it also has the application of the cycle counting method such as Rainflow for the time domain and, using the spectral parameters, the spectral method with correction, as Wirsching and Light [1980] for example, for the frequency domain. The driving force that induces the failure process for the time and frequency domains are respectively the equivalent normal stress and the equivalent power spectral density function, as presented in the references Łagoda et al. [2005], Niesłony and Macha [2007] and Niesłony et al. [2020].

Table 4: Approaches used by MVM based on equivalent normal stress and power density function of equivalent normal stress.

Task	Variable
Material Plane	ϕ and θ
Maximum Variance Method (MVM)	σ_{eq} and G_{eq}
Damage Accumulation	Rainflow and Palmgren-Miner Rule

- MVM based on shear stress resolved in the direction of maximum variance:** consists of a reformulation of the Maximum Variance Method (MMV) to make it suitable in multiaxial fatigue evaluations by numerical post-processing of loading histories as complex as those that damage mechanical components real, as shown in the references Susmel and Lazzarin [2002], Susmel [2010], Susmel and Taylor [2012] and Luo et al. [2020]. The methodology presents an addition of a third orientation angle to a generic vector \mathbf{q} , following the Taylor nomenclature, and the driving force that induces fatigue failure is the shear stress, as shown in Table 5. The cycle counting method is maintained, again using the Rainflow method and Miner's Rule. The results will be compared with a time domain model and life prediction against failure proposed by Lagoda.

Table 5: Approaches used by MVM based on shear stress resolved in the direction of maximum variance.

Task	Variable
Material Plane	ϕ , θ and α
Maximum Variance Method (MVM)	τ_q
Damage Accumulation	Rainflow, Modified Wöhler Curves Method and Palmgren-Miner Rule

The advantage of MVM is that, once the variance and covariance are known, the computational time required to determine the direction that has the maximum variance of the shear stress doesn't depend strictly on the full size of the load history, but a sufficient amount of load history is required to have a stability of statistical parameters such as variance and covariance.

7 Multiaxial Fatigue - Maximum Variance Method Approach based on equivalent normal stress and power density function of equivalent normal stress

7.1 Time Domain

The stresses acting on a given plane can be deduced as long as the necessary stresses are known so that the balance of external forces belonging to the infinitesimal tetrahedron of the material

body is maintained. Assuming a material point O on the inclined cutting plane as the origin of the $X'-Y'-Z'$ coordinate system as shown in Figure 25. This point could be interpreted as the critical point in relation to fatigue strength depending on the magnitude, direction and direction of the stress vector in which it is oriented at that point. The cutting plane Δ contains perpendicularly a unit normal vector \mathbf{n} or (n_x, n_y, n_z) , being possible to determine the normal stress $\sigma_{x'}$ and the shear components $\tau_{x'y'}$ and $\tau_{x'z'}$. This plane can be referenced by its spherical coordinates θ and ϕ between 0 a 2π , however, due to the property of symmetry, the variation θ, ϕ is adopted in $[0, \pi]$ (Socie and Marquis [2000]).

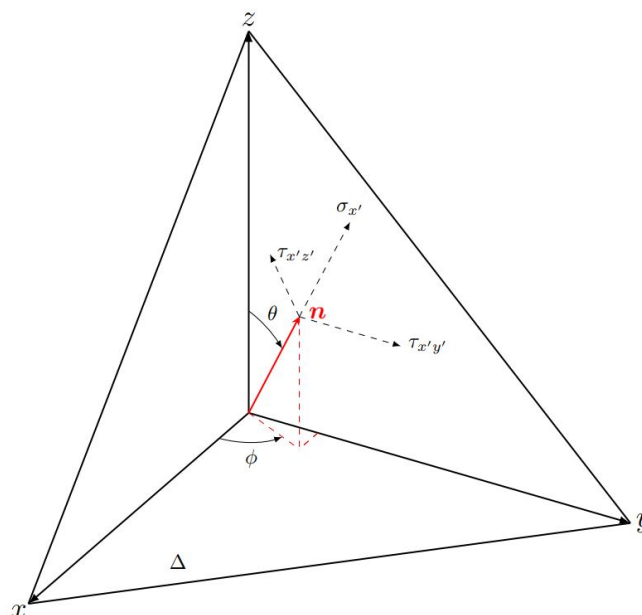


Figure 25: Stresses acting in a plane with a three-dimensional coordinate system.

Suppose a multiaxial state of stress, given by $\sigma(t)$:

$$\sigma(t) = \begin{bmatrix} \sigma_{xx}(t) & \tau_{xy}(t) & \tau_{xz}(t) \\ \tau_{yx}(t) & \sigma_{yy}(t) & \tau_{yz}(t) \\ \tau_{zx}(t) & \tau_{zy}(t) & \sigma_{zz}(t) \end{bmatrix} \tag{73}$$

To represent the normal stress $\sigma_{x'}$, the director cosines in function of θ and ϕ that are given according to the relations below:

$$a_{11} = \cos(\phi) \sin(\theta) \tag{74}$$

$$a_{12} = \sin(\phi) \sin(\theta) \tag{75}$$

$$a_{13} = \cos(\theta) \tag{76}$$

$$a_{21} = -\sin(\phi) \tag{77}$$

$$a_{22} = \cos(\phi) \tag{78}$$

$$a_{23} = 0 \tag{79}$$

$$a_{31} = -\cos(\phi) \cos(\theta) \tag{80}$$

$$a_{32} = -\sin(\phi) \cos(\theta) \tag{81}$$

$$a_{33} = \sin(\theta) \tag{82}$$

While, the normal vector can be given by Eq. (83):

$$\mathbf{n} = a_{11}\mathbf{i} + a_{12}\mathbf{j} + a_{13}\mathbf{k} \tag{83}$$

Therefore, the transformation equation for normal stress in any plane is given by:

$$\sigma_{x'} = \sigma_x a_{11}^2 + \sigma_y a_{12}^2 + \sigma_z a_{13}^2 + 2(\tau_{xy} a_{11} a_{12} + \tau_{xz} a_{11} a_{13} + \tau_{yz} a_{13} a_{12}) \tag{84}$$

where a_{11} , a_{12} and a_{13} are the cosines of the angles between the axes x and x' , x and y' and x and z' , respectively.

The Eq. (84) can be rewritten as a linear combination of the stress tensor multiplied by director cosines in the direction normal to the plane, given by Eqs. (74) to (76). Rearranging the terms:

$$\sigma_{x'} = \boldsymbol{\sigma} \cdot \mathbf{a} \tag{85}$$

where $\boldsymbol{\sigma}$ is the stochastic process stress vector and \mathbf{a} is the vector of the direction cosines:

$$\mathbf{a} = [a_{11}^2, a_{12}^2, a_{13}^2, 2a_{11}a_{12}, 2a_{11}a_{13}, 2a_{13}a_{12}] \tag{86}$$

$$\boldsymbol{\sigma}(t) = [\sigma_{xx}(t), \sigma_{yy}(t), \sigma_{zz}(t), \tau_{xy}(t), \tau_{xz}(t), \tau_{yz}(t)] \tag{87}$$

Applying the maximum normal stress criterion in the critical plane, it is stated that fatigue failure is caused by the maximum normal stress acting on the critical plane and that the critical plane is perpendicular to the mean direction of the maximum principal stress σ_1 , according to the bibliographic reference Macha [1989]. In this way, the critical plane is defined by the direction of the normal vector \mathbf{n} , which coincides with the direction of the maximum principal stress σ_1 , and, with the application of the MMV, the identification of the critical plane will be the material plane in which the variance of the equivalent stress reaches its maximum value.

Based on the normal stress transformation equation in any plane, Eq. (85) can be rewritten as a linear combination of the components of the vector \mathbf{a} , represented by a_k in the `crefeq:top7vector_a` and modified with the directions of the normal vector \mathbf{n} :

$$\sigma_{eq}(t) = \sum_{k=1}^6 a_k \sigma_k(t) \quad (k = 1, \dots, 6) \tag{88}$$

With the application of the maximum normal stress criterion, the Maximum Variance Method (MVM) allows determining the plane in which the equivalent stress variance reaches its maximum value, and this plane will be considered as a critical plane and will be subject to fatigue failure:

$$var[\sigma_{eq}(t)] = \max \left(\sum_i \sum_j a_i a_j Cov[\sigma_i(t), \sigma_j(t)] \right) = \max \left(\sum_i \sum_j a_i a_j [C] \right) \tag{89}$$

where:

$$[C] = \begin{pmatrix} C_{xx,xx} & C_{xx,yy} & C_{xx,zz} & C_{xx,xy} & C_{xx,xz} & C_{xx,yz} \\ C_{yy,xx} & C_{yy,yy} & C_{yy,zz} & C_{yy,xy} & C_{yy,xz} & C_{yy,yz} \\ C_{zz,xx} & C_{zz,yy} & C_{zz,zz} & C_{zz,xy} & C_{zz,xz} & C_{zz,yz} \\ C_{xy,xx} & C_{xy,yy} & C_{xy,zz} & C_{xy,xy} & C_{xy,xz} & C_{xy,yz} \\ C_{xz,xx} & C_{xz,yy} & C_{xz,zz} & C_{xz,xy} & C_{xz,xz} & C_{xz,yz} \\ C_{yz,xx} & C_{yz,yy} & C_{yz,zz} & C_{yz,xy} & C_{yz,xz} & C_{yz,yz} \end{pmatrix} \tag{90}$$

where the covariance of two equal components of the tension tensor vector, in other words, with the same subindex ($m = n$), is equal to the variance of the component itself. Therefore, the

covariance matrix can be redefined as follows:

$$[C] = \begin{pmatrix} V_{xx} & C_{xx,yy} & C_{x,zz} & C_{xx,xy} & C_{xx,xz} & C_{xx,yz} \\ C_{yy,xx} & V_{yy} & C_{yy,zz} & C_{yy,xy} & C_{yy,xz} & C_{yy,yz} \\ C_{zz,xx} & C_{zz,yy} & V_{zz} & C_{zz,xy} & C_{zz,xz} & C_{zz,yz} \\ C_{xy,xx} & C_{xy,yy} & C_{xy,zz} & V_{xy} & C_{xy,xz} & C_{xy,yz} \\ C_{xz,xx} & C_{xz,yy} & C_{xz,zz} & C_{xz,xy} & V_{xz} & C_{xz,yz} \\ C_{yz,xx} & C_{yz,yy} & C_{yz,zz} & C_{yz,xy} & C_{yz,xz} & V_{yz} \end{pmatrix} \quad (91)$$

Critical angles are determined from the critical plane, where the variance of $\sigma_{eq}(t)$ has reached its maximum value as seen in Eq. (92).

$$\theta^c, \phi^c = \max[\text{var}_{\sigma_{eq}}(\theta, \phi)] \quad (92)$$

7.2 Frequency Domain

In the frequency domain, from spectral methods, there is application of spectral power density functions (PSD) to evaluate the characteristics of frequencies in the history of stress. In this way, the multiaxial stress states are described by the power spectral density function matrix $\mathbf{G}(f)$:

$$\mathbf{G}(f) = \begin{bmatrix} G_{11}(f) & G_{12}(f) & G_{13}(f) & G_{14}(f) & G_{15}(f) & G_{16}(f) \\ G_{21}(f) & G_{22}(f) & G_{23}(f) & G_{24}(f) & G_{25}(f) & G_{26}(f) \\ G_{31}(f) & G_{32}(f) & G_{33}(f) & G_{34}(f) & G_{35}(f) & G_{36}(f) \\ G_{41}(f) & G_{42}(f) & G_{43}(f) & G_{44}(f) & G_{45}(f) & G_{46}(f) \\ G_{51}(f) & G_{52}(f) & G_{53}(f) & G_{54}(f) & G_{55}(f) & G_{56}(f) \\ G_{61}(f) & G_{62}(f) & G_{63}(f) & G_{64}(f) & G_{65}(f) & G_{66}(f) \end{bmatrix} \quad (93)$$

As shown in the bibliographic reference Nieslony and Macha [2007], from the power spectral density matrix $\mathbf{G}(f)$, $G_{eq}(f)$ can be obtained by Eq. (94):

$$G_{eq}(f) = \mathbf{a}\mathbf{G}(f)\mathbf{a}^T \quad (94)$$

Solving the inner product, Eq. (94) will be represented as follows:

$$G_{eq} = a_1^2 G_{11} + a_1 a_4 G_{41} + a_1 a_4 G_{14} + a_4^2 G_{44} = a_1^2 G_{11} + a_4^2 G_{44} + (a_1 a_4)(G_{41} + G_{14}) \quad (95)$$

$$G_{eq} = a_1^2 G_{11} + a_4^2 G_{44} + 2a_1 a_4 \Re(G_{41}) \quad (96)$$

The application of the MMV in the frequency domain occurs in a similar way to the time domain, with identification of the maximum value of variance of power density function of equivalent normal stress, proposed in Eq. (97), in which the vector of the coefficients coincides with the \mathbf{a} vector of the linear combination:

$$\text{var}[G_{eq}(f)] = \sum_i \sum_j a_i a_j \text{var}(G_{kl}) \quad (97)$$

$$\text{var}(G_{kl}) = \Re \left[\int_0^\infty G_{kl}(f) df \right] \quad (k,l=1,\dots,6) \quad (98)$$

The power spectral density matrix is given by real parts of zero-order spectral moments of power spectral density functions. By MMV method, identifying the critical angles where the variance has its maximum value:

$$\theta^c, \phi^c = \max[\text{var}_{G_{eq}}(\theta, \phi)] \quad (99)$$

It is worth mentioning that the variance obtained is also the spectral moment of the equivalent power spectral density function as shown by Eq. (100):

$$\text{var}(G_{eq}) = m_{o,G_{eq}} = \Re \left[\int_0^{\infty} G_{eq}(f) df \right] \quad (100)$$

Therefore, the number of oscillations, also known as crossover rate, and the number of local peaks are given respectively by Eq. (101) and Eq. (102). While the number of oscillations and local peaks of the equivalent spectral density function is given by Eq. (103) and Eq. (104).

$$N_o^+ = \sqrt{\frac{m_2}{m_0}} \quad (101)$$

$$M^+ = \sqrt{\frac{m_4}{m_2}} \quad (102)$$

$$N_o^+ = \sqrt{\frac{m_{2,G_{eq}}}{m_{0,G_{eq}}}} \quad (103)$$

$$M^+ = \sqrt{\frac{m_{4,G_{eq}}}{m_{2,G_{eq}}}} \quad (104)$$

8 Multiaxial Fatigue - Maximum Variance Method Approach based on shear stress resolved in the direction of maximum variance

To formalize the theory, we consider a body in a frame of reference O_{xyz} subjected to a complex system of forces and moments that result in a three-dimensional state of stress that varies over time. The generic material plane, as shown in Fig. 26, can be oriented by the angles θ and ϕ through a normal vector \mathbf{n} , while the second frame of reference $O_{e_a e_b n}$, a local coordinate system belonging to a generic material plane where the normal vector passes through the point of origin, has its orientation defined by other unit vectors e_a and e_b whose cross product of these vectors results in the normal vector, as shown in Fig. 26. Using the critical plane method, it is possible to identify the orientation of a generic plane in the form of the previous coordinate system, xyz .

According to Figure Fig. 26, Fig. 27 and Fig. 28, the unit vectors that define the orientation of the axes \mathbf{n} , e_a e e_b ³ plane can be expressed as follows:

$$\mathbf{n} = \begin{bmatrix} n_x \\ n_y \\ n_z \end{bmatrix} = \begin{bmatrix} \sin \theta \cos \phi \\ \sin \theta \sin \phi \\ \cos \theta \end{bmatrix} \quad (105)$$

$$\mathbf{e}_a = \begin{bmatrix} e_{a_x} \\ e_{a_y} \\ e_{a_z} \end{bmatrix} = \begin{bmatrix} \sin \phi \\ -\cos \phi \\ 0 \end{bmatrix} \quad (106)$$

$$\mathbf{e}_b = \begin{bmatrix} e_{b_x} \\ e_{b_y} \\ e_{b_z} \end{bmatrix} = \begin{bmatrix} \cos \theta \cos \phi \\ \cos \theta \sin \phi \\ -\sin \theta \end{bmatrix} \quad (107)$$

³ e_b : to project this unit vector it is necessary to project in the plane $z b_\theta$ for later decompose it into the xy .

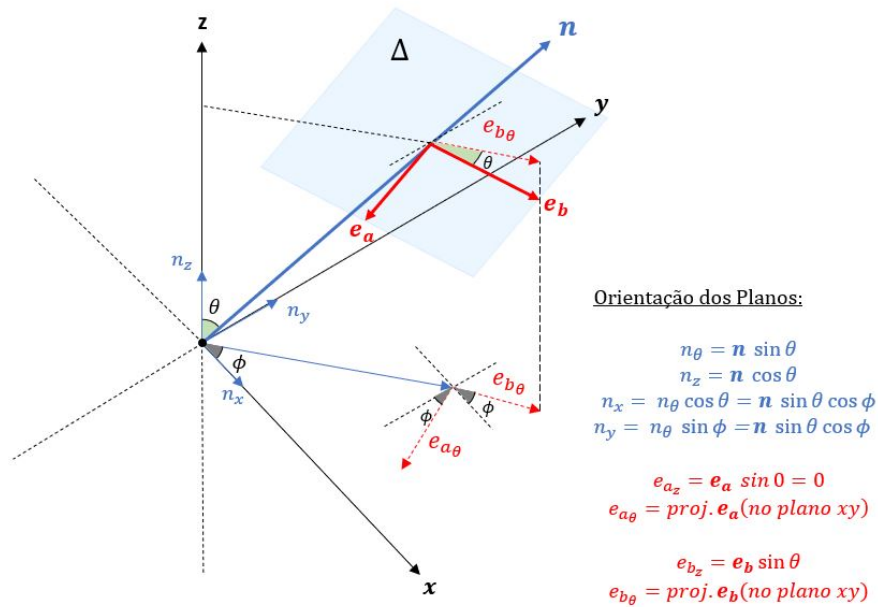


Figure 26: Stresses acting in a plane with a three-dimensional coordinate system (Courtesy of Lima [2021])

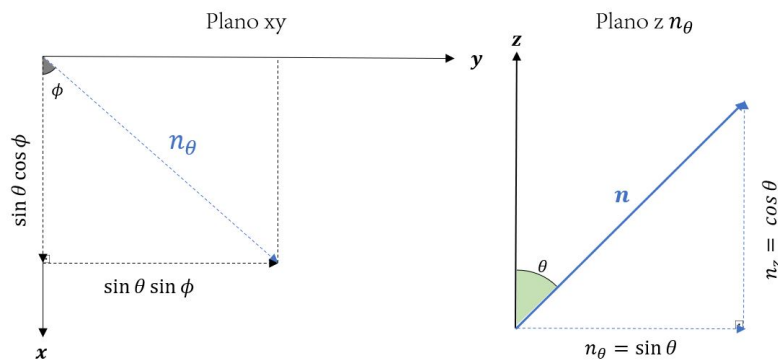


Figure 27: Projection of unit vector n (Courtesy of Lima [2021]).

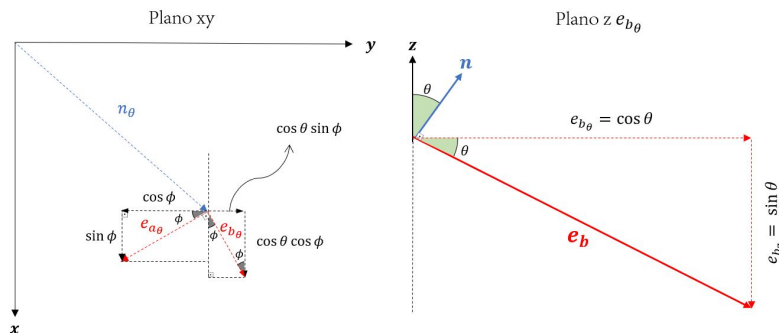


Figure 28: Projection of unit vector e_a and e_b (Courtesy of Lima [2021]).

Based on Fig. 28, Susmel assumed a generic direction m for the unit vector \mathbf{q} in a given material plane situated in the new coordinate system and oriented by the angle α between unit vector \mathbf{a} and unit vector \mathbf{q} . As Fig. 29, the components of \mathbf{q} vector can be written in the following vector form:

$$\mathbf{q} = q_{e_a} + q_{e_b} \quad (108)$$

$$\mathbf{q} = \mathbf{e}_a \cos \alpha + \mathbf{e}_b \sin \alpha \quad (109)$$

$$\mathbf{q} = \begin{bmatrix} \cos \alpha \sin \phi + \sin \alpha \cos \theta \cos \phi \\ -\cos \alpha \cos \phi + \sin \alpha \cos \theta \sin \phi \\ -\sin \alpha \sin \theta \end{bmatrix} \quad (110)$$

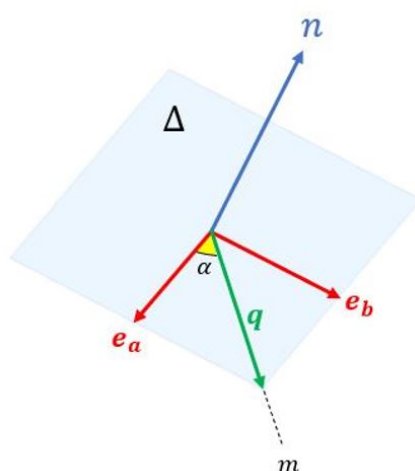


Figure 29: Projection of vector \mathbf{q} in vector \mathbf{e}_a and \mathbf{e}_b (Courtesy of Lima [2021]).

$$\sigma_n(t) = \begin{bmatrix} n_x & n_y & n_z \end{bmatrix} \begin{bmatrix} \sigma_{xx}(t) & \tau_{xy}(t) & \tau_{xz}(t) \\ \tau_{yx}(t) & \sigma_{yy}(t) & \tau_{yz}(t) \\ \tau_{zx}(t) & \tau_{zy}(t) & \sigma_{zz}(t) \end{bmatrix} \begin{bmatrix} n_x \\ n_y \\ n_z \end{bmatrix} \quad (111)$$

$$\tau_q(t) = \begin{bmatrix} q_x & q_y & q_z \end{bmatrix} \begin{bmatrix} \sigma_{xx}(t) & \tau_{xy}(t) & \tau_{xz}(t) \\ \tau_{yx}(t) & \sigma_{yy}(t) & \tau_{yz}(t) \\ \tau_{zx}(t) & \tau_{zy}(t) & \sigma_{zz}(t) \end{bmatrix} \begin{bmatrix} n_x \\ n_y \\ n_z \end{bmatrix} \quad (112)$$

Developing Eq. (112), the instantaneous value of the resolved shear stress along the m direction:

$$\tau_q(t) = \mathbf{q}^T \bar{\boldsymbol{\sigma}}(t) \mathbf{n} \quad (113)$$

$$\tau_q(t) = n_x \cdot (q_z \cdot \tau_{xz} + q_y \cdot \tau_{xy} + q_x \cdot \sigma_{xx}) + n_y \cdot (q_z \cdot \tau_{yz} + q_x \cdot \tau_{xy} + q_y \cdot \sigma_{yy}) + n_z \cdot (q_y \cdot \tau_{yz} + q_x \cdot \tau_{xz} + q_z \cdot \sigma_{zz}) \quad (114)$$

Rearranging Eq. (114), isolating the stress terms:

$$\tau_q(t) = \sigma_x \cdot (n_x \cdot q_x) + \sigma_{yy} \cdot (n_y \cdot q_y) + \sigma_{zz} \cdot (n_z \cdot q_z) + \tau_{xy} \cdot (n_x \cdot q_y + n_y \cdot q_x) + \tau_{xz} \cdot (n_x \cdot q_z + n_z \cdot q_x) + \tau_{yz} \cdot (n_y \cdot q_z + n_z \cdot q_y) \quad (115)$$

Rewriting tensor in Voigt's Notation, the representation of $\boldsymbol{\sigma}(t)$ would be a six-dimensional vector process and, therefore, the shear stress can be expressed as:

$$\tau_q(t) = \mathbf{d} \cdot \mathbf{s}(t) \quad (116)$$

Where the tensor in Voigt notation is given as follows:

$$\mathbf{s}(t) = \boldsymbol{\sigma}^T(t) = [\sigma_{xx}(t) \ \sigma_{yy}(t) \ \sigma_{zz}(t) \ \tau_{xy}(t) \ \sigma_{xz}(t) \ \sigma_{yz}(t)] \quad (117)$$

Where \mathbf{d} is known as the vector of director cosines, being expressed according to Eq. (118):

$$\mathbf{d} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \\ d_5 \\ d_6 \end{bmatrix} = \begin{bmatrix} n_x \cdot q_x \\ n_y \cdot q_y \\ n_z \cdot q_z \\ n_x \cdot q_y + n_y \cdot q_x \\ n_x \cdot q_z + n_z \cdot q_x \\ n_y \cdot q_z + n_z \cdot q_y \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \left[\sin(\theta) \sin(2\phi) \cos(\alpha) + \sin(\alpha) \sin(2\theta) \cos(\phi)^2 \right] \\ \frac{1}{2} \left[-\sin(\theta) \sin(2\phi) \cos(\alpha) + \sin(\alpha) \sin(2\theta) \sin(\phi)^2 \right] \\ -\frac{1}{2} \sin(\alpha) \sin(2\theta) \\ \frac{1}{2} \sin(\alpha) \sin(2\phi) \sin(2\theta) - \cos(\alpha) \cos(2\phi) \sin(\theta) \\ \sin(\alpha) \cos(\phi) \cos(2\theta) + \cos(\alpha) \sin(\phi) \cos(\theta) \\ \sin(\alpha) \sin(\phi) \cos(2\theta) - \cos(\alpha) \cos(\phi) \cos(\theta) \end{bmatrix} \quad (118)$$

The model formulation is based on the assumption that fatigue damage maximizes in a plane that experiences the maximum amplitude of shear stress. From shear stress of a generic direction m , τ_q , the MVM identifies the critical plane from the variance of the resolved shear stress history in the direction in which it reaches its maximum value. With this, the variance of shear stress, $\tau_q(t)$, resolved along the direction m can then be calculated directly as:

$$\text{var}[\tau_q(t)] = \mathbf{d}^T [C] \mathbf{d} \quad (119)$$

$$\text{var}[\tau_q(t)] = [d_1 \ d_2 \ d_3 \ d_4 \ d_5 \ d_6] \begin{bmatrix} C_{xx,xx} & C_{xx,yy} & C_{xx,zz} & C_{xx,xy} & C_{xx,xz} & C_{xx,yz} \\ C_{yy,xx} & C_{yy,yy} & C_{yy,zz} & C_{yy,xy} & C_{yy,xz} & C_{yy,yz} \\ C_{zz,xx} & C_{zz,yy} & C_{zz,zz} & C_{zz,xy} & C_{zz,xz} & C_{zz,yz} \\ C_{xy,xx} & C_{xy,yy} & C_{xy,zz} & C_{xy,xy} & C_{xy,xz} & C_{xy,yz} \\ C_{xz,xx} & C_{xz,yy} & C_{xz,zz} & C_{xz,xy} & C_{xz,xz} & C_{xz,yz} \\ C_{yz,xx} & C_{yz,yy} & C_{yz,zz} & C_{yz,xy} & C_{yz,xz} & C_{yz,yz} \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \\ d_5 \\ d_6 \end{bmatrix} \quad (120)$$

Or by Taylor series expansion as:

$$\text{var}[\tau_q(t)] = \text{var} \left[\sum_k d_k \sigma_k \right] = \sum_i \sum_j d_i d_j \text{Cov}[\sigma_i, \sigma_j] \quad (121)$$

Another way to simplify of Eq. (119) would be to consider the product of the covariance matrix and the vector \mathbf{d} as a column matrix $[M]$.

$$[M] = \begin{bmatrix} M_1 \\ M_2 \\ M_3 \\ M_4 \\ M_5 \\ M_6 \end{bmatrix} = \begin{bmatrix} C_{xx,xx} & C_{xx,yy} & C_{xx,zz} & C_{xx,xy} & C_{xx,xz} & C_{xx,yz} \\ C_{yy,xx} & C_{yy,yy} & C_{yy,zz} & C_{yy,xy} & C_{yy,xz} & C_{yy,yz} \\ C_{zz,xx} & C_{zz,yy} & C_{zz,zz} & C_{zz,xy} & C_{zz,xz} & C_{zz,yz} \\ C_{xy,xx} & C_{xy,yy} & C_{xy,zz} & C_{xy,xy} & C_{xy,xz} & C_{xy,yz} \\ C_{xz,xx} & C_{xz,yy} & C_{xz,zz} & C_{xz,xy} & C_{xz,xz} & C_{xz,yz} \\ C_{yz,xx} & C_{yz,yy} & C_{yz,zz} & C_{yz,xy} & C_{yz,xz} & C_{yz,yz} \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \\ d_5 \\ d_6 \end{bmatrix} \quad (122)$$

And in the form of Taylor series expansion, the variance can be expressed as:

$$\text{var}[\tau_q(t)] = \sum_{i=1}^6 d_i M_i \quad (123)$$

In this way, the variance of shear stress in the direction of vector \mathbf{q} of Eq. (119) can be represented as follows:

$$\text{var}[\tau_q(t)] = [d_1 \ d_2 \ d_3 \ d_4 \ d_5 \ d_6] \begin{bmatrix} M_1 \\ M_2 \\ M_3 \\ M_4 \\ M_5 \\ M_6 \end{bmatrix} \quad (124)$$

Critical angles are determined from the critical plane, where the variance of $\tau_q(t)$ has reached its maximum value as seen in Eq. (125).

$$\theta^c, \phi^c, \alpha^c = \max[\text{var}_{\tau_q}(\theta, \phi, \alpha)] \quad (125)$$

9 Multiaxial Fatigue - Maximum Variance Method Approach as a function of the invariants of deviatoric stress tensor: Projection-by-Projection (PbP)

Among the stress invariant-based methods to estimate fatigue life, the approach called “Projection by Projection – Projection-by-Projection (PbP)” is suitable for multiaxial fatigue analysis. The proposed criterion is particularly useful for the evaluation of the durability of complex structures under random loading. The advantages of the PbP method can be obtained in the time domain, as well as for applications in the frequency domain (Benasciutti et al. [2013])

In 2008, the first works published with the PbP method were developed by Cristofori, Susmel and Tovo (Cristofori et al. [2008]), who described the time domain methodology to estimate fatigue damage in the presence of complex multiaxial loads. In 2011, Cristofori, Benasciutti and Tovo (Cristofori et al. [2011]) continued the development of the method and published the formulation for the frequency domain, which can be applied to stationary multiaxial random loads. In 2014, a numerical example for analyzing a steel beam excited by a random input acceleration was proposed by Cristofori and Benasciutti (Cristofori and Benasciutti [2014]), who performed a finite element analysis to calculate the PSD stress matrix at the nodes of the structure and applied the spectral model of the PbP method to estimate the fatigue damage. A practical guide for implementing the method in the frequency domain was presented by Benasciutti, Zanellati and Cristofori in 2019 (Benasciutti et al. [2019]), which detailed all the steps of the method. These are the most relevant works developed so far for the PbP method, which include numerical examples and/or procedures for implementing the method.

9.1 Projection-by-Projection (PbP)

This method makes use of the amplitude of the second invariant of the deviator tensor $\sqrt{(J_{2a})}$ and the hydrostatic pressure $\sigma_H(t)$ to evaluate the damage due to a generic fatigue loading. The fatigue life of a real component is estimated considering the loading history described in 5-dimensional Euclidean space, in which the damage related to a generic loading path can be determined from the contributions calculated by the projection of the loading trajectory itself to the along the axes of a chosen frame of reference based on Euclidean space (Cristofori et al. [2008]).

From a geometric point of view, the loading trajectory can be fully represented in terms of its projections along the directions of maximum variance defined by the axes of the chosen frame of reference and adopted based on the 5-dimensional deviator space. In this way, analysis of a complex loading trajectory can be performed from its projections, such that the damage relative to each projection can be calculated first, and then the overall damage from the loading history can

be obtained as a function of the different contributions given by the projections (Cristofori et al. [2008]).

The procedure adopted by the PbP method to estimate fatigue life consists of defining the reference of the projection system, evaluating the damage along each projection and, finally, evaluating the total damage.

In this sense, considering that $\sigma(t)$ represents the tensor of the time-varying stress components

$$\sigma(t) = \begin{bmatrix} \sigma_{xx}(t) & \tau_{xy}(t) & \tau_{xz}(t) \\ \tau_{xy}(t) & \sigma_{yy}(t) & \tau_{yz}(t) \\ \tau_{xz}(t) & \tau_{yz}(t) & \sigma_{zz}(t) \end{bmatrix} \rightarrow \sigma(t) = \begin{bmatrix} \sigma_{xx}(t) \\ \sigma_{yy}(t) \\ \sigma_{zz}(t) \\ \tau_{xy}(t) \\ \sigma_{xz}(t) \\ \sigma_{yz}(t) \end{bmatrix} \quad (126)$$

The six non-redundant stress components of the $\sigma(t)$ can be reordered into a stress vector which, in the frequency domain, can be characterized by the PSD matrix, $\mathbf{G}(f)$

$$\mathbf{G}(f) = \begin{bmatrix} G_{11}(f) & G_{12}(f) & G_{13}(f) & G_{14}(f) & G_{15}(f) & G_{16}(f) \\ G_{21}(f) & G_{22}(f) & G_{23}(f) & G_{24}(f) & G_{25}(f) & G_{26}(f) \\ G_{31}(f) & G_{32}(f) & G_{33}(f) & G_{34}(f) & G_{35}(f) & G_{36}(f) \\ G_{41}(f) & G_{42}(f) & G_{43}(f) & G_{44}(f) & G_{45}(f) & G_{46}(f) \\ G_{51}(f) & G_{52}(f) & G_{53}(f) & G_{54}(f) & G_{55}(f) & G_{56}(f) \\ G_{61}(f) & G_{62}(f) & G_{63}(f) & G_{64}(f) & G_{65}(f) & G_{66}(f) \end{bmatrix} \quad (127)$$

where $\mathbf{G}(f)$ is the matrix of autospectral and cross-spectral density functions of stress vector $\sigma(t)$.

The set of zero-order spectral moments of the PSD matrix ($\mathbf{G}(f)$) is represented in the covariance matrix:

$$\text{Var}(G_{kl}) = \Re \left[\int_0^\infty G_{kl}(f) df \right] \quad (k,l=1,\dots,6) \quad (128)$$

$$\mathbf{C} = \begin{bmatrix} V_x & C_{xx,yy} & C_{xx,zz} & C_{xx,xy} & C_{xx,xz} & C_{xx,yz} \\ C_{yy,xx} & V_y & C_{yy,zz} & C_{yy,xy} & C_{yy,xz} & C_{yy,yz} \\ C_{zz,xx} & C_{zz,yy} & V_z & C_{zz,xy} & C_{zz,xz} & C_{zz,yz} \\ C_{xy,xx} & C_{xy,yy} & C_{xy,zz} & V_{xy} & C_{xy,xz} & C_{xy,yz} \\ C_{xz,xx} & C_{xz,yy} & C_{xz,zz} & C_{xz,xy} & V_{xz} & C_{xz,yz} \\ C_{yz,xx} & C_{yz,yy} & C_{yz,zz} & C_{yz,xy} & C_{yz,xz} & V_{yz} \end{bmatrix} \quad (129)$$

where the diagonal elements are the variance terms of the stress $x_i(t)$, $C_{ii} = \text{Var}(x_i(t))$, and the off-diagonal elements are the covariance terms between the stress $x_i(t)$ and $x_j(t)$, $C_{ij} = \text{Cov}(x_i(t), x_j(t))$. The coefficient of correlation r_{ij} , $r_{ij} = C_{ij} / \sqrt{(C_{ii}C_{jj})}$, represents, for multiaxial random loads, the degree of proportionality between $x_i(t)$ and $x_j(t)$. When $r_{ij} = 1$, the processes are perfectly correlated, that is, proportional stresses, when $r_{ij} = 0$, they are uncorrelated processes, that is, non-proportional stresses (Cristofori et al. [2011], Benasciutti et al. [2019])

Fatigue damage from a multiaxial random stress is addressed by the PbP method in the deflector space. Thus, the vector of stresses in the deviator space, $\mathbf{s}(t)$, is related to the stress tensor $\sigma(t)$ by the following transformation rules:

$$\mathbf{s}(t) = \begin{bmatrix} s_1(t) \\ s_2(t) \\ s_3(t) \\ s_4(t) \\ s_5(t) \end{bmatrix} \quad (130)$$

where:

$$s_1(t) = \frac{1}{2\sqrt{3}}(2\sigma_{xx}(t) - \sigma_{yy}(t) - \sigma_{zz}(t)) \quad (131)$$

$$s_2(t) = \frac{1}{2}(\sigma_{yy}(t) - \sigma_{zz}(t)) \quad (132)$$

$$s_3(t) = \tau_{xy}(t) \quad (133)$$

$$s_4(t) = \tau_{zx}(t) \quad (134)$$

$$s_5(t) = \tau_{yz}(t) \quad (135)$$

The stochastic vector $\mathbf{s}(t)$ that describes the stress components in the deviator space can be characterized by the covariance matrix \mathbf{C}' and in the frequency domain by the PSD 5x5 matrix $\mathbf{G}'(f)$ (Cristofori and Benasciutti [2014]). The axes of the projections are defined in the direction of maximum variance of the deviator space, so from the eigenvectors \mathbf{U} of the covariance matrix \mathbf{C}' the projections of the stress histories in time are described as:

$$\mathbf{\Omega}(t) = \mathbf{U}^T \mathbf{s}(t) \quad \text{where} \quad \mathbf{\Omega}(t) = (\Omega_{(p,1)}(t), \Omega_{(p,2)}(t), \Omega_{(p,3)}(t), \Omega_{(p,4)}(t), \Omega_{(p,5)}(t)) \quad (136)$$

$$\mathbf{C}'_p(t) = \mathbf{U}^T \mathbf{C}'(t) \mathbf{U} \quad (137)$$

where $\mathbf{\Omega}(t)$ represents the projections in the direction of maximum variance and $\mathbf{C}'_p(t)$ is a diagonal matrix of variance of the vector $\mathbf{\Omega}(t)$:

$$C'_{p,ii} = Var[\Omega_{(p,i)}] \quad i = 1, 2, \dots, 5. \quad (138)$$

In the frequency domain, the diagonal PSD matrix $\mathbf{G}'_p(f)$ determines the power spectra in the projections:

$$\mathbf{G}'_p(f) = \mathbf{U}^T \mathbf{G}'(f) \mathbf{U} \quad (139)$$

Therefore, the projections of random stresses $\Omega_{(p,i)}(t)$ of the vector $\mathbf{\Omega}(t)$ are totally uncorrelated, where each component represents an independent random stress, and its fatigue damage can be estimated separately either in the time domain and frequency (Cristofori and Benasciutti [2014]).

9.2 PbP Method for a Plane Stress State

In the plane or biaxial stress state, the stress tensor becomes three non-redundant stress components:

$$\boldsymbol{\sigma}(t) = \begin{bmatrix} \sigma_{xx}(t) & \tau_{xy}(t) \\ \tau_{xy}(t) & \sigma_{yy}(t) \end{bmatrix} = \begin{bmatrix} \sigma_{xx}(t) \\ \sigma_{yy}(t) \\ \tau_{xy}(t) \end{bmatrix} \quad (140)$$

where $\boldsymbol{\sigma}(t)$ can be represented as stress vector, where each stress is considered a stationary random variable Benasciutti et al. [2019]. In the frequency domain, the PSD matrix is now represented as:

$$\mathbf{G}(f) = \begin{bmatrix} G_{11}(f) & G_{12}(f) & G_{13}(f) \\ G_{21}(f) & G_{22}(f) & G_{23}(f) \\ G_{31}(f) & G_{32}(f) & G_{33}(f) \end{bmatrix} = \begin{bmatrix} G_{xx}(f) & G_{xx,yy}(f) & G_{xx,xy}(f) \\ G_{yy,xx}(f) & G_{yy}(f) & G_{yy,xy}(f) \\ G_{xy,xx}(f) & G_{xy,yy}(f) & G_{xy}(f) \end{bmatrix} \quad (141)$$

Consequently, the covariance matrix will take the form:

$$\mathbf{C} = \begin{bmatrix} C_{xx} & C_{xx,yy} & C_{xx,zz} \\ C_{yy,xx} & C_{yy} & C_{yy,zz} \\ C_{xy,xx} & C_{xy,yy} & C_{xy} \end{bmatrix} = \begin{bmatrix} V_{xx} & C_{xx,yy} & C_{xx,zz} \\ C_{yy,xx} & V_{yy} & C_{yy,zz} \\ C_{xy,xx} & C_{xy,yy} & V_{xy} \end{bmatrix} \quad (142)$$

The stress vector in the deviator space becomes three dimensional, $\mathbf{s}(t) = [s_1(t), s_2(t), s_3(t)]$, which can be determined as:

$$\mathbf{s}(t) = \mathbf{A} \cdot \mathbf{x}(t) \quad (143)$$

where:

$$\mathbf{A} = \begin{bmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{2\sqrt{3}} & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (144)$$

The covariance matrix and the PSD of $\mathbf{s}(t)$ can be obtained from the matrix \mathbf{A} through the following equations:

$$\mathbf{C}' = \mathbf{A} \cdot \mathbf{C} \cdot \mathbf{A}^T = \begin{bmatrix} C'_{11} & C'_{12} & C'_{13} \\ & C'_{22} & C'_{23} \\ sim & & C'_{33} \end{bmatrix} \quad (145)$$

$$\mathbf{G}'(f) = \mathbf{A} \cdot \mathbf{G}(f) \cdot \mathbf{A}^T \quad (146)$$

Thus, from the eigenvectors \mathbf{U} of the covariance matrix \mathbf{C}' , the loadings in the projections with the direction of maximum variance are obtained:

$$\mathbf{\Omega}(t) = \mathbf{U}^T \mathbf{s}(t) \quad \text{where} \quad \mathbf{\Omega}(t) = [\Omega_{(p,1)}(t), \Omega_{(p,2)}(t), \Omega_{(p,3)}(t)] \quad (147)$$

The covariance matrix of the vector $\mathbf{\Omega}(t)$ is determined as:

$$\mathbf{C}'_p = \mathbf{U}^T \cdot \mathbf{C}' \cdot \mathbf{U} = \begin{bmatrix} C'_{p,11} & 0 & 0 \\ 0 & C'_{p,22} & 0 \\ 0 & 0 & C'_{p,33} \end{bmatrix} \quad (148)$$

and in the frequency domain the PSD matrix is defined:

$$\mathbf{G}'_p(f) = \mathbf{U}^T \mathbf{G}'(t) \mathbf{U} = \begin{bmatrix} G'_{p,11}(f) & 0 & 0 \\ 0 & G'_{p,22}(f) & 0 \\ 0 & 0 & G'_{p,33}(f) \end{bmatrix} \quad (149)$$

The principle of the PbP method can be illustrated through Fig. 30, in which a random loading history of bending and torsion $[\sigma_{xx}(t), \tau_{xy}(t)]$ is represented in the deviator space with two non-zero components $[s_1(t), s_3(t)]$ that generate a random loading path Ψ . To define the reference of the projection system $[s_1, 0, s_3, 0]$ the direction of maximum variance of the loading Ψ is considered, thus obtaining the projections of stress $[\Omega_{(p,1)}(t), \Omega_{(p,3)}(t)]$.

For fatigue life estimation under multiaxial conditions, the Modified Wöhler Curve Method (MWCM) proposed by Susmel and Lazzarin [2002]. The MWCM takes as its starting point the assumption that crack initiation is Mode II dominated, and it holds true independently of both stress concentration feature and degree of multiaxiality of the stress field damaging the fatigue process zone (Susmel and Taylor [2003], Susmel [2004]). In more detail, the MWCM assumes that fatigue damage depends on the shear stress amplitude, τ_a , and on the maximum normal stress, $\sigma_{(n,max)}$, relative to the critical plane, which is defined as the plane experiencing the maximum shear stress amplitude. From a practical point of view, the MWCM postulates that fatigue damage depends on two stress parameters: τ_a and ρ . The latter parameter is defined as the ratio between the maximum normal stress, $\sigma_{(n,max)}$, and the shear stress amplitude, τ_a , relative to the critical plane (Araújo et al. [2007]).

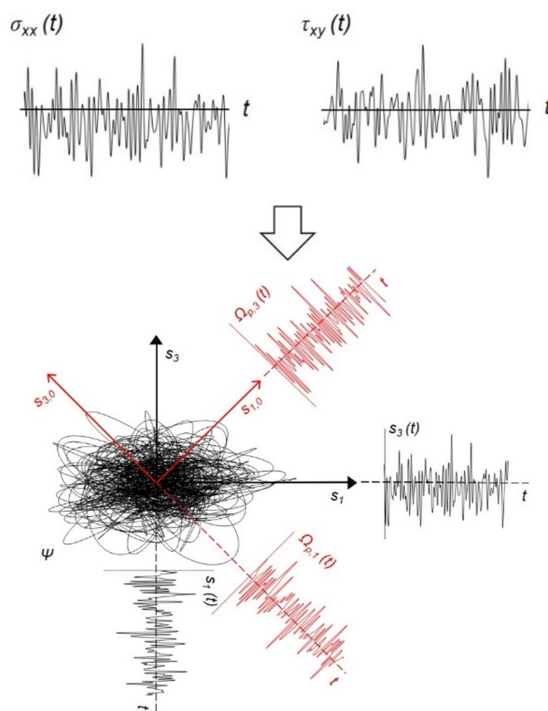


Figure 30: Example of a random loading path Ψ in the deflector space, generated from flexural-torsional stresses, to determine the stresses in the reference frame of the projection system

The modified Wöhler diagram used to define the parameters of the S-N curve in multiaxial loading situations considers two fatigue curves that are generally available for materials: a normal stress versus life curve obtained under fully reverse loading conditions, characterized by $\rho = 1$ (just normal stress and see Fig. 31), and a fatigue curve obtained under fully reverse torsional conditions, characterized by $\rho = 0$ (just shear stress and see Fig. 31).

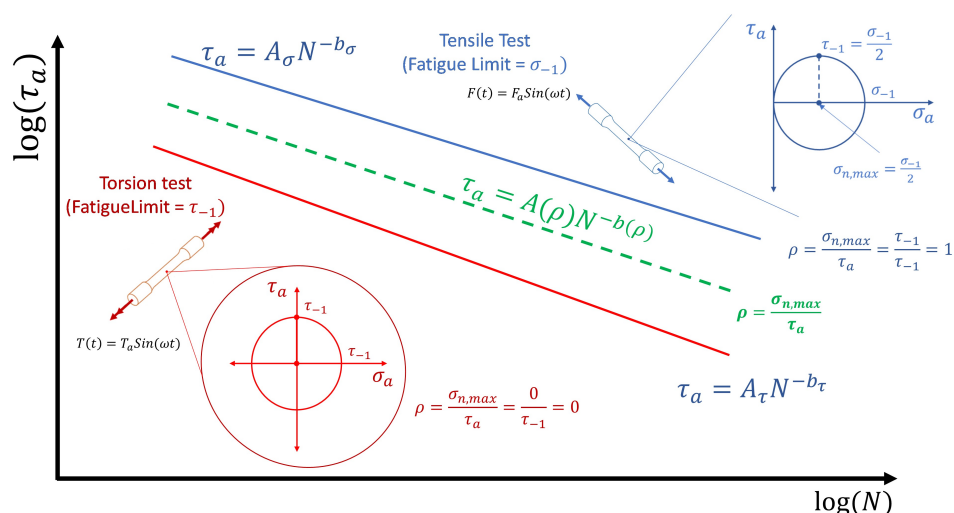


Figure 31: Modified Wöhler diagram for defining the reference S-N curve for a generic multiaxial loading.

Hypothetically, it is assumed that the reference S-N curve in the modified Wöhler diagram for any given value of ρ is defined by linear interpolation of the fatigue parameters for axial and torsional loading (Benasciutti et al. [2019]). From the definition of the multiaxiality degree in the critical plane (ρ_{crit}), the parameters that characterize the curve τ_a versus N_f respectively, $A_{\rho_{crit}}$ (coefficient of the τ_a -N curve) and $b_{\rho_{crit}}$ (curve exponent), can be estimated by means of Eq. (150) and Eq. (151).

$$A_{\rho_{crit}} = [A_{\sigma} - A_{\tau}]\rho_{crit} + A_{\tau} \quad (150)$$

$$b_{\rho_{crit}} = [b_{\sigma} - b_{\tau}]\rho_{crit} + b_{\tau} \quad (151)$$

9.3 Life Prediction Strategies Using the PbP Method

Life prediction by the PbP method can be performed in both time and frequency domains, the main difference between the two domains being the determination of the S-N reference curve in the modified Wöhler diagram. Once the ρ_{ref} parameter is defined, according to each domain, the parameters of the S-N curve are defined by Equation (9.17) and then the damage estimate is performed, which in time considers the rainflow technique associated with the Palmgren-Miner rule and follows the development presented in Section 4, for the frequency domain, it analyzes the power spectra as shown in Section 5.

9.3.1 Application of the Life Estimation in the Time Domain

The rainflow counting method can be applied to each projection of the loading history in the U_i directions, where the amplitude of the cycles defined in terms of the deviator stress $J_{a_{ij}}$, and the values of the maximum hydrostatic pressure $\sigma_{H_{ij}}$, are recorded for each cycle. counted (identified by the subscript j) along each considered projection (identified by the subscript $i = 1, 2, \dots, 5$) (Cristofori et al. [2008]). In order to estimate fatigue damage, the reference curve in the modified Wöhler diagram can be located from the stress ratio ρ_{ref} defined as:

$$\rho_{ref} = \sqrt{3} \frac{\sigma_{(H_{ref})}}{(\sum_i (J_{a_{iref}})^2)^{1/2}} \quad (152)$$

where $J_{a_{iref}}$ is the average value of the amplitude for the cycles counted in relation to the considered projection and $\sigma_{H_{ref}}$ the average value of the maximum hydrostatic stresses of the cycles, defined by the following relations:

$$\sigma_{H_{ref}} = \text{mean}_{ij}[\sigma_{(H_{ij})}] \quad (153)$$

$$J_{a_{iref}} = \text{mean}_{ij}[J_{(a_{ij})}] \quad (154)$$

where $\sigma_H = \text{trace}(\boldsymbol{\sigma})/3$ is the hydrostatic stress.

The fatigue damage, D_{ij} , due to each identified cycle can then be estimated using the following equation:

$$D_{ij} = \left[\frac{J_{a_{ij}}}{J_{a_{ref}}} \right]^{k_{ref}} \frac{1}{N_A} \quad (155)$$

and the total damage relative to the i -th projection D_i , considering the linear cumulative rule proposed by Palmgren-Miner, results in the estimation of the damage $D(\Omega_{p,i})$ related to the projected loading $\Omega_{p,i}$:

$$D(\Omega_{p,i}) = \sum_j D_{ij} \quad (156)$$

Considering that the damage caused by the histories of multiaxial loads acting in the directions of the deflector tensor is decoupled, Cristofori et al. [2008] proposed the following quadratic rule to calculate the accumulated damage at the analyzed point:

$$D(\Omega) = \left[\sum_i (D(\Omega_{p,i}))^{\frac{2}{k_{ref}}} \right] = \left[\sum_i \left(\sum_j D_{ij} \right)^{\frac{2}{k_{ref}}} \right]^{\frac{k_{ref}}{2}} \quad (157)$$

9.3.2 Application of the Life Estimation in the Frequency Domain

To take full advantage of a frequency domain approach, the definition of the stress ratio ρ_{ref} must be properly reformulated for stationary multiaxial random loading. This is accomplished by considering the following equation:

$$\rho_{ref} = \sqrt{3} \frac{\frac{\sigma_{H,m}}{\sqrt{2}} + \sqrt{Var(\sigma_H(t))}}{\sqrt{\sum_i Var(\Omega_{p,i}(t))}} = \sqrt{3} \frac{\sigma_{H,m} + \sqrt{2\lambda_{0,H}}}{\sqrt{2 \sum_i \lambda_{0,i}}} \quad (158)$$

where $\sigma_H(t)$ is the hydrostatic stress, a uniaxial stationary random process, with constant mean $\sigma_{H,m}$ and variance $Var(\sigma_H(t)) = \lambda_{0,H}$ (which coincides with the zero-order moment $\lambda_{(0,H)}$ of the power spectral density $G_H(f)$ of $\sigma_H(t)$). Likewise, every projection $\Omega_{p,i}(t)$ is a stationary random process, with variance $Var(\Omega_{p,i}(t)) = \lambda_{0,i}$ (where $\lambda_{0,i}$ is the zero-order moment of the power spectral density $G'_{p,i}(f)$ of the loading history in each of the "i" projections, which is an element of the matrix $G'_p(f)$) (Cristofori et al. [2011]). If the multiaxial loading has stress components with an average value of zero, then $\sigma_{H,m} = 0$. In the case of sinusoidal loading of constant amplitude, Eq. (158) reduces to the expression of Eq. (152).

In the frequency domain an estimate of the expected damage in a time interval can be provided by spectral methods. If the random projection $\Omega_{p,i}(t)$ is a narrow band process the damage is defined as:

$$D_{NB}(\Omega_{p,i}) = \frac{v_{0,i}T}{C_{ref}} (\sqrt{2\lambda_{0,i}})^{k_{ref}} \Gamma \left(1 + \frac{k_{ref}}{2} \right) \quad (159)$$

and the total damage of the multiaxial loading histories is defined from the quadratic accumulation rule:

$$D_{NB}^{PbP} = \frac{T}{C_{ref}} \Gamma \left(1 + \frac{k_{ref}}{2} \right) \left[\sum_i \left(2\lambda_{0,i} (v_{0,i})^{\frac{2}{k_{ref}}} \right) \right]^{\frac{k_{ref}}{2}} \quad (160)$$

If the random projection $\Omega_{p,i}(t)$ is a broadband process, the damage can be estimated by other spectral methods, such as the Tovo-Benasciutti method (cited in Section 5.2.2).

$$D_{TB}(\Omega_{p,i}) = \eta_{TB,i} \frac{v_{0,i}T}{C_{ref}} (\sqrt{2\lambda_{0,i}})^{k_{ref}} \Gamma \left(1 + \frac{k_{ref}}{2} \right) \quad (161)$$

so that the total damage is obtained from the consideration of all projections

$$D_{TB}^{PbP} = \frac{T}{C_{ref}} \Gamma \left(1 + \frac{k_{ref}}{2} \right) \left[\sum_i \left(2\lambda_{0,i} (\eta_{TB,i} v_{0,i})^{\frac{2}{k_{ref}}} \right) \right]^{\frac{k_{ref}}{2}} \quad (162)$$

Moreover, the theory of the PbP method can be easily extended to other spectral methods, such as Dirlik's method ((cited in section 5.2.2).). For this method the damage can be obtained as:

$$D_{DK}(\Omega_{p,i}) = \frac{v_{p,i}T \sigma_{S,i}^{k_{ref}}}{C_{ref}} \left[D_{1,i} Q_i^{k_{ref}} \Gamma(k_{ref} + 1) + (\sqrt{2})^{k_{ref}} \Gamma \left(\frac{k_{ref}}{2} + 1 \right) (D_{2,i} |R_i|^{k_{ref}} + D_{3,i}) \right] \quad (163)$$

where the total damage of the loading story consists of:

$$D_{DK}^{PbP} = \frac{T}{C_{ref}} \left[\sum_i \left(v_{p,i}^{\frac{2}{k_{ref}}} \sigma_{S,i}^2 \left[D_{1,i} Q_i^{k_{ref}} \Gamma(k_{ref}+1) + (\sqrt{2})^{k_{ref}} \Gamma\left(\frac{k_{ref}}{2}+1\right) (D_{2,i} |R_i|^{k_{ref}} + D_{3,i}) \right] \right) \right]^{\frac{k_{ref}}{2}} \quad (164)$$

10 Final Remarks

This chapter sought to present a review of the state of art and the concepts necessary to understand the criteria of multiaxial random fatigue using the Maximum Variance Method (MVM) in medium and high cycle fatigue regimes. In this sense, approaches based on time domain and frequency domain were examined. We tried, as far as possible, to structure the material presented in a amicable way and with easy-to-understand language so that it can contribute, especially for researchers who wish to start studies on this topic.

Despite the theoretical review on the phenomenology of random fatigue under uniaxial and multiaxial loading conditions and the deductions presented in the methodologies involved in this chapter, we must emphasize that all the topics discussed are in continuous development and improvement - An example is the contribution recently published in the article Ferreira et al. [2022] which presents an innovative and highly efficient strategy for the search for the critical plane based on the maximum variance of resolved shear stress method. Comparative tests showed that the proposed algorithm is, on average, 150 times faster than the search strategy based on the Gradient Ascent Method and at least 230 times faster than the Direct Search Method of critical planes.

References

- J. Araújo, L. Susmel, D. Taylor, J. Ferro, and E. Mamiya. On the use of the theory of critical distances and the modified wöhler curve method to estimate fretting fatigue strength of cylindrical contacts. *International Journal of Fatigue*, 29(1):95–107, 2007. ISSN 0142-1123. doi: <https://doi.org/10.1016/j.ijfatigue.2006.02.041>.
- ASTM-E1049-85. Standard practices for cycle counting in fatigue analysis. Technical report, ASTM International, West Conshohocken, PA, 2017.
- D. Benasciutti and R. Tovo. Spectral methods for lifetime prediction under wide-band stationary random processes. *International Journal of Fatigue*, 27:867–877, 08 2005. doi: 10.1016/j.ijfatigue.2004.10.007.
- D. Benasciutti, A. Cristofori, and R. Tovo. Analogies between spectral methods and multiaxial criteria in fatigue damage evaluation. *Probabilistic Engineering Mechanics*, 31:39–45, 2013. ISSN 0266-8920. doi: <https://doi.org/10.1016/j.probengmech.2012.12.002>.
- D. Benasciutti, D. Zanellati, and A. Cristofori. The "projection-by-projection" (pbp) criterion for multiaxial random fatigue loadings: Guidelines to practical implementation. *Frattura ed Integrità Strutturale*, 47:348–366, 01 2019. doi: 10.3221/IGF-ESIS.47.26.
- N. Bishop. The use of frequency domain parameters to predict structural fatigue. *Tese, Department of Engineering, University of Warwick, England*, 01 1988.
- A. Cristofori and D. Benasciutti. “projection-by-projection” approach: A spectral method for multiaxial random fatigue. *SAE Technical Paper 2014-01-0924*, 1, 04 2014. doi: 10.4271/2014-01-0924.

- A. Cristofori, L. Susmel, and R. Tovo. A stress invariant based criterion to estimate fatigue damage under multiaxial loading. *International Journal of Fatigue*, 30(9):1646–1658, 2008. ISSN 0142-1123. doi: 10.1016/j.ijfatigue.2007.11.006.
- A. Cristofori, D. Benasciutti, and R. Tovo. A stress invariant based spectral method to estimate fatigue life under multiaxial random loading. *International Journal of Fatigue*, 33:887–899, 07 2011. doi: 10.1016/j.ijfatigue.2011.01.013.
- T. Dirlik. Application of computers in fatigue analysis. *Ph.D. thesis. The University of Warwick*, 1985.
- T. Dirlik and D. Benasciutti. Dirlik and tovo-benasciutti spectral methods in vibration fatigue: A review with a historical perspective. *Multidisciplinary Digital Publishing Institute 11*, 11:1333, 08 2021. doi: 10.3390/met11091333.
- E739-10. Standard practice for statistical analysis of linear or linearized stress-life (s-n) and strain-life (e-n) fatigue data. Technical report, ASTM International, West Conshohocken, PA, 2015.
- J. Ferreira, J. Dias, E. Cardoso, J. Araújo, and C. da Silva. A contribution to the identification of the critical plane using the maximum variance method. *International Journal of Fatigue*, page 107228, 2022. ISSN 0142-1123. doi: <https://doi.org/10.1016/j.ijfatigue.2022.107228>.
- Y.-L. Lee. Fatigue analysis in the frequency domain. *Fatigue Testing and Analysis: Theory and Practice*, Yung-Li Lee, J. Pan, R. B. Hathaway e M. E. Barkey (Ed.). Elsevier Inc., pages 369–394, 12 2005. doi: 10.1016/B978-075067719-6/50011-7.
- F. M. d. Lima. Fadiga multiaxial: Análise de espectro para previsão de vida em componentes sujeitos a carregamentos de amplitude variável. In *Projeto de Graduação (Engenharia Mecânica), Departamento de Engenharia Mecânica, Universidade de Brasília, Brasília, DF*, page 127 p., 2021.
- P. Luo, W. Yao, and L. Susmel. An improved critical-plane method and cycle counting method to assess damage under variable amplitude multiaxial fatigue loading. *Fatigue and Fracture of Engineering Materials and Structures*, 43:2024–2039, 05 2020. doi: 10.1111/ffe.13281.
- E. Macha. Generalization of fatigue criteria for multiaxial sinusoidal loadings in the range of random loadings. *Biaxial and Multiaxial Fatigue*, EGF 3:425–36, 1989.
- M. Matsuichi and T. Endo. Fatigue of metals subjected to varying stress. *Paper presented to Japan Soc. Mech. Engrs, Jukvoka, Japan*, 1968.
- A. Nieslony and E. Macha. *Spectral Method in Multiaxial Random Fatigue*. Lecture Notes in Applied and Computational Mechanics. Springer Berlin Heidelberg, 2007.
- A. Nieslony, M. Böhm, and R. Owsiniński. Formulation of multiaxial fatigue failure criteria for spectral method. *International Journal of Fatigue*, 135:105519, 06 2020. doi: 10.1016/j.ijfatigue.2020.105519.
- I. Papadopoulos. Long life fatigue under multiaxial loading. *International Journal of Fatigue - INT J FATIGUE*, 23:839–849, 2001.
- D. Socie and G. Marquis. *Multiaxial fatigue*. Premiere Series Bks. Society of Automotive Engineers, 2000.

- L. Susmel. A unifying approach to estimate the high-cycle fatigue strength of notched components subjected to both uniaxial and multiaxial cyclic loadings. *Fatigue & Fracture of Engineering Materials & Structures*, 27:391 – 411, 04 2004. doi: 10.1111/j.1460-2695.2004.00759.x.
- L. Susmel. A simple and efficient numerical algorithm to determine the orientation of the critical plane in multiaxial fatigue problems. *International Journal of Fatigue*, 32:1875–1883, 11 2010.
- L. Susmel and P. Lazzarin. A bi-parametric wöhler curve for high cycle multiaxial fatigue assessment. *Fatigue & Fracture of Engineering Materials & Structures*, 25:63 – 78, 01 2002. doi: 10.1046/j.1460-2695.2002.00462.x.
- L. Susmel and D. Taylor. Two methods for predicting the multiaxial fatigue limits of sharp notches. *Fatigue & Fracture of Engineering Materials & Structures*, 26:821 – 833, 09 2003. doi: 10.1046/j.1460-2695.2003.00683.x.
- L. Susmel and D. Taylor. A critical distance/plane method to estimate finite life of notched components under variable amplitude uniaxial/multiaxial fatigue loading. *International Journal of Fatigue*, 38:7–24, 2012.
- P. H. Wirsching and M. C. Light. Fatigue under wide band random stresses. *Journal of the Structural Division*, 106:1593–1607, 1980.
- W. Zhao and M. J. Baker. On the probability density function of rainflow stress range for stationary gaussian processes. *International Journal of Fatigue*, 14:121–135, 1992.
- T. Łagoda, E. MACHA, and A. Niesłony. Fatigue life calculation by means of the cycle counting and spectral methods under multiaxial random loading. *Fatigue, Fracture of Engineering Materials and Structures*, 28:409 – 420, 2005.

APPENDIX A

DEDUCTION OF THE DIRLIK DAMAGE EQUATION

The expected total fatigue damage by Dirlik's method, D_{DK}

$$D_{DK} = \frac{v_p T}{K} \int_0^{\infty} S_a^m P_{rr}(S_r) dS_r. \quad (A1)$$

where $P_{rr}(S_r)$ is the probability density function of rainflow intervals in terms of the auxiliary variable Z defined as

$$P_{rr}(S_r) = \frac{D_1}{2\sigma_S Q} e^{-\frac{Z}{Q}} + \frac{D_2 Z}{2\sigma_S R^2} e^{-\frac{Z^2}{2R^2}} + \frac{D_3 Z}{2\sigma_S} e^{-\frac{Z^2}{2}}, \quad (A2)$$

where $Z = \frac{S_r}{2\sigma_S} = \frac{S_a}{\sigma_S}$.

Therefore, we have the integral to be solved

$$I = \int_0^{\infty} S_a^m P_{rr}(S_r) dS_r = \int_0^{\infty} \left(\frac{S_r}{2}\right)^m \left[\frac{D_1}{2\sigma_S Q} e^{-\frac{S_r}{2\sigma_S Q}} + \frac{D_2 S_r}{4\sigma_S^2 R^2} e^{-\frac{S_r^2}{8\sigma_S^2 R^2}} + \frac{D_3 S_r}{4\sigma_S^2} e^{-\frac{S_r^2}{8\sigma_S^2}} \right] dS_r. \quad (A3)$$

Separating the integral into parts

$$I = \int_0^{\infty} \left(\frac{S_r}{2}\right)^m \frac{D_1}{2\sigma_S Q} e^{-\frac{S_r}{2\sigma_S Q}} dS_r + \int_0^{\infty} \left(\frac{S_r}{2}\right)^m \frac{D_2 S_r}{4\sigma_S^2 R^2} e^{-\frac{S_r^2}{8\sigma_S^2 R^2}} dS_r + \int_0^{\infty} \left(\frac{S_r}{2}\right)^m \frac{D_3 S_r}{4\sigma_S^2} e^{-\frac{S_r^2}{8\sigma_S^2}} dS_r, \quad (A4)$$

so that

$$I_1 = \int_0^{\infty} \left(\frac{S_r}{2}\right)^m \frac{D_1}{2\sigma_S Q} e^{-\frac{S_r}{2\sigma_S Q}} dS_r; \quad I_2 = \int_0^{\infty} \left(\frac{S_r}{2}\right)^m \frac{D_2 S_r}{4\sigma_S^2 R^2} e^{-\frac{S_r^2}{8\sigma_S^2 R^2}} dS_r; \quad (A5)$$

$$I_3 = \int_0^{\infty} \left(\frac{S_r}{2}\right)^m \frac{D_3 S_r}{4\sigma_S^2} e^{-\frac{S_r^2}{8\sigma_S^2}} dS_r,$$

that is,

$$I = I_1 + I_2 + I_3. \quad (A6)$$

Solving the integral I_1

$$I_1 = \frac{D_1}{2\sigma_S Q} \int_0^\infty \left(\frac{S_r}{2}\right)^m e^{-\frac{S_r}{2\sigma_S Q}} dS_r. \quad (\text{A7})$$

Performing a change of variable

$$m = a - 1; \quad \frac{S_r}{2\sigma_S Q} = x \quad e^{-\frac{S_r}{2\sigma_S Q}} dS_r = 2\sigma_S Q dx. \quad (\text{A8})$$

The integral I_1 can be rewritten as

$$I_1 = \frac{D_1}{2\sigma_S Q} \int_0^\infty (\sigma_S Q x)^{a-1} e^{-x} 2\sigma_S Q dx = D_1 (\sigma_S Q)^{a-1} \int_0^\infty x^{a-1} e^{-x} dx. \quad (\text{A9})$$

Given that the gamma function is defined as

$$\Gamma(a) = \int_0^\infty x^{a-1} e^{-x} dx, \quad (\text{A10})$$

the integral I_1 can be expressed based on the gamma function

$$I_1 = D_1 (\sigma_S Q)^{a-1} \Gamma(a). \quad (\text{A11})$$

Returning from the variable change

$$I_1 = D_1 (\sigma_S Q)^m \Gamma(m + 1). \quad (\text{A12})$$

The integral I_2 has the following form

$$I_2 = \frac{D_2}{2\sigma_S^2 R^2} \int_0^\infty \left(\frac{S_r}{2}\right)^{m+1} e^{-\frac{S_r^2}{8\sigma_S^2 R^2}} dS_r. \quad (\text{A13})$$

Again, performing a change of variable

$$m + 1 = a - 1; \quad \frac{S_r^2}{8\sigma_S^2 R^2} = x \quad e^{-\frac{S_r^2}{8\sigma_S^2 R^2}} dS_r = \frac{\sqrt{2} \sigma_S R}{\sqrt{x}} dx. \quad (\text{A14})$$

which allows you to write it as

$$I_2 = \frac{D_2}{2\sigma_S^2 R^2} \int_0^\infty (\sqrt{2}\sigma_S R \sqrt{x})^{a-1} e^{-x} \frac{\sqrt{2} \sigma_S R}{\sqrt{x}} dx = \frac{D_2}{2\sigma_S R} \sqrt{2} (\sqrt{2}\sigma_S R)^{a-1} \int_0^\infty (\sqrt{x})^{a-1} e^{-x} \frac{1}{\sqrt{x}} dx. \quad (\text{A15})$$

Simplifying the expression

$$I_2 = D_2 (2)^{\frac{a}{2}-1} (\sigma_S R)^{a-2} \int_0^{\infty} (\sqrt{x})^{a-2} e^{-x} dx . \quad (\text{A16})$$

Performing another variable change

$$a = 2b , \quad (\text{A17})$$

The I_2 can be write as

$$I_2 = D_2 (2)^{b-1} (\sigma_S R)^{2b-2} \int_0^{\infty} (\sqrt{x})^{2b-2} e^{-x} dx = D_2 (2)^{b-1} (\sigma_S R)^{2b-2} \int_0^{\infty} x^{b-1} e^{-x} dx . \quad (\text{A18})$$

Using the gamma function shown in Equation (A10)

$$I_2 = D_2 (2)^{b-1} (\sigma_S R)^{2b-2} \Gamma(b) . \quad (\text{A19})$$

Returning from the variable change

$$I_2 = D_2 (2)^{\frac{a}{2}-1} (\sigma_S R)^{a-2} \Gamma\left(\frac{a}{2}\right) = D_2 (2)^{\frac{m}{2}} (\sigma_S R)^m \Gamma\left(\frac{m}{2} + 1\right) . \quad (\text{A20})$$

Solving the integral I_3

$$I_3 = \frac{D_3}{2\sigma_S^2} \int_0^{\infty} \left(\frac{S_r}{2}\right)^{m+1} e^{-\frac{S_r^2}{8\sigma_S^2}} dS_r . \quad (\text{A21})$$

Performing a change of variable

$$m + 1 = a - 1; \quad \frac{S_r^2}{8\sigma_S^2} = x \quad e^{-\frac{S_r^2}{8\sigma_S^2}} dS_r = \frac{\sqrt{2} \sigma_S}{\sqrt{x}} dx . \quad (\text{A22})$$

the integral I_3 can be rewritten as

$$I_3 = \frac{D_3}{2\sigma_S^2} \int_0^{\infty} (\sqrt{2}\sigma_S \sqrt{x})^{a-1} e^{-x} \frac{\sqrt{2} \sigma_S}{\sqrt{x}} dx = \frac{D_3}{2\sigma_S} \sqrt{2} (\sqrt{2}\sigma_S)^{a-1} \int_0^{\infty} (\sqrt{x})^{a-1} e^{-x} \frac{1}{\sqrt{x}} dx . \quad (\text{A23})$$

Simplifying the expression

$$I_3 = D_3 (2)^{\frac{a}{2}-1} (\sigma_S)^{a-2} \int_0^{\infty} (\sqrt{x})^{a-2} e^{-x} dx . \quad (\text{A24})$$

Performing another variable change

$$a = 2b , \quad (\text{A25})$$

the integral I_3 can be rewritten as

$$I_3 = D_3 (2)^{b-1} (\sigma_S)^{2b-2} \int_0^\infty (\sqrt{x})^{2b-2} e^{-x} dx = D_3 (2)^{b-1} (\sigma_S)^{2b-2} \int_0^\infty x^{b-1} e^{-x} dx. \quad (\text{A26})$$

Using the gamma function shown in Equation (A10)

$$I_3 = D_3 (2)^{b-1} (\sigma_S)^{2b-2} \Gamma(b). \quad (\text{A27})$$

Returning from variable change

$$I_3 = D_3 (2)^{\frac{a}{2}-1} (\sigma_S)^{a-2} \Gamma\left(\frac{a}{2}\right) = D_3 (2)^{\frac{m}{2}} (\sigma_S)^m \Gamma\left(\frac{m}{2} + 1\right). \quad (\text{A28})$$

Adding all solved integrals together

$$I = D_1 (\sigma_S Q)^m \Gamma(m+1) + D_2 (2)^{\frac{m}{2}} (\sigma_S R)^m \Gamma\left(\frac{m}{2} + 1\right) + D_3 (2)^{\frac{m}{2}} (\sigma_S)^m \Gamma\left(\frac{m}{2} + 1\right). \quad (\text{A29})$$

Simplifying the equation

$$I = \sigma_S^m \left[D_1 Q^m \Gamma(m+1) + (\sqrt{2})^m \Gamma\left(\frac{m}{2} + 1\right) (D_2 |R|^m + D_3) \right]. \quad (\text{A30})$$

Substituting the integral into Equation (A1), the Dirlik damage at a time interval T can be determined as

$$D_{DK} = \frac{v_p T}{K} \sigma_S^m \left[D_1 Q^m \Gamma(m+1) + (\sqrt{2})^m \Gamma\left(\frac{m}{2} + 1\right) (D_2 |R|^m + D_3) \right]. \quad (\text{A31})$$

APPENDIX 2

MATLAB ROUTINE FOR LIFE ESTIMATION BASED ON THE TIME DOMAIN METHOD

```

clc
clear
close all

% T = SIGNAL ACQUISITION WINDOW TIME
% RMS_sigmoid = RMS VALUE OF NORMAL STRESS COMPONENT
% RMS_tau = RMS VALUE OF SHEAR STRESS COMPONENT
% ro = CORRELATION COEFFICIENT BETWEEN SIGMA AND TAU STRESS HISTORIES
% BANDA = SPECTRAL BANDWIDTH OF SIGNALS SIGNAL CENTRAL FREQUENCY
% PT_PERIODO = NUMBER OF POINTS PER PERIOD
% Fmin = LOWER LIMIT OF SIGNAL FREQUENCY BAND
% Fmax = UPPER LIMIT OF SIGNAL FREQUENCY BAND
% TAMANHO = SPECTRUM SIZE
% PERCENTUAL = PERCENTAGE OF SPECTRUM OVERLAY FOR ESTIMATE WELCH PERIODOGRAM

%MATERIAL
k = 5.988;
C = 3.216E21;

% SAMPLING TIME SETTING
PT_PERIODO = 2;

Fmin1 = 2;
Fmax1 = 4;
G1 = 2500
RMS_1 = sqrt(G1*(Fmax1-Fmin1));

Fmin2 = 20;
Fmax2 = 22;
G2 = 1000;
RMS_2 = sqrt(G2*(Fmax2-Fmin2));

%DEFINITION OF SPECTRUM CHARACTERISTICS
TAMANHO = 2^19;
PERCENTUAL = 99;

F_INF = min(Fmin1, Fmin2);
F_SUP = max(Fmax1, Fmax2);
TAXA = F_SUP*2*PT_PERIODO;
FC = (F_SUP+F_INF)/2;
BANDA = (F_SUP-F_INF);
TAXA = F_SUP*2*PT_PERIODO;

SOBRE = round(TAMANHO*PERCENTUAL/100);

T = TAMANHO/TAXA;
TEMPO = T;

[t,Saa] = FILTRA_SINAL_2(T,RMS_1,RMS_2,Fmin1,Fmax1,Fmin2,Fmax2,TAXA,TAMANHO);

media = mean(Saa)
desv = std(Saa)
kurt = kurtosis(Saa)

pd_g = makedist('Normal','mu',media,'sigma',desv);

[c,hist,edges,rmm,idx] = rainflow(Saa/2);

```

```

[LINHAS, COLUNAS] = size(hist);

for i = 1:LINHAS
    NSa(i) = sum(hist(i,:));
    Sa(i) = (edges(i)+edges(i+1))/2;
    d(i,1) = (NSa(i)*Sa(i)^k)/C;
end
nsa = NSa;
NSa = NSa/sum(NSa);

figure(1)
plot(Sa,NSa);

figure(2)
plot(t,Saa);

figure(3)
h = histogram(Saa,20)
U = transpose(h.BinEdges);
[LINHAS, COLUNAS] = size(U);

for i = 1:LINHAS-1
    St(i) = (U(i)+U(i+1))/2;
    y(i) = pdf(pd_g,St(i));
end

DELTA_St = St(2)-St(1);
NSt = h.Values/(sum(h.Values)*DELTA_St);

plot(St,y,St,NSt)

V(:,1) = St;
V(:,2) = y;
V(:,3) = NSt;

G(:,1) = Sa;
G(:,2) = NSa/(G(2,1)-G(1,1));

figure(4)
bar(G(:,1),G(:,2))

[pxx,w]= pwelch(Saa,2048,32,2048,'onesided',1/t(2));

figure(5);
bar(w,pxx);

P(:,1) = w;
P(:,2) = pxx;

Saaa = Saa(10000:30000);
Dac = sum(d)
D = Dac/TEMPO
Tf = 1/(24*3600*D)

Function [t,Saa] = FILTRA_SINAL_2(T,RMS_1,RMS_2,Fmin1,Fmax1,Fmin2,Fmax2,TAXA,TAMANHO);

function [t,Z] = FILTRA_SINAL(T,RMS_SINAL1,RMS_SINAL2,Fmin1,Fmax1,Fmin2,Fmax2,TAXA,TAMANHO)

% T = SIGNAL ACQUISITION WINDOW TIME
% RMS_SINAL = RMS VALUE OF THE SIGNAL TO BE GENERATED
% Fmin = LOWER LIMIT OF SIGNAL FREQUENCY BAND
% Fmax = UPPER LIMIT OF SIGNAL FREQUENCY BAND
%TAXA = SIGNAL ACQUISITION RATE
% TAMANHO = SPECTRUM SIZE
% PERCENTUAL = PERCENTAGE OF SPECTRUM OVERLAY FOR WELCH PERIODOGRAM ESTIMATE

```

```
t = linspace(0,T,TAMANHO);

R = rand(1);
rng(R);
X1 = randn(TAMANHO,1);
Y1 = bandpass(X1,[Fmin1 Fmax1],TAXA,'Steepness',0.9999,'StopbandAttenuation',10);

R = rand(1);
rng(R);
X2 = randn(TAMANHO,1);
Y2 = bandpass(X2,[Fmin2 Fmax2],TAXA,'Steepness',0.9999,'StopbandAttenuation',10);
Z = RMS_SINAL2*Y2/std(Y2) + RMS_SINAL1*Y1/std(Y1);
```

Chapter 18

Random Field Discretization in Stochastic Finite Elements: Element Size Effects and Error Indicators in Reliability Structural Analysis

Chapter details

Chapter DOI:

<https://doi.org/10.4322/978-65-86503-88-3.c18>

Chapter suggested citation / reference style:

Jorge, Ariosto B., et al. (2022). “Random Field Discretization in Stochastic Finite Elements: Element Size Effects and Error Indicators in Reliability Structural Analysis”. In Jorge, Ariosto B., et al. (Eds.) *Uncertainty Modeling: Fundamental Concepts and Models*, Vol. III, UnB, Brasilia, DF, Brazil, pp. 613–639. Book series in Discrete Models, Inverse Methods, & Uncertainty Modeling in Structural Integrity.

P.S.: DOI may be included at the end of citation, for completeness.

Book details

Book: Uncertainty Modeling: Fundamental Concepts and Models

Edited by: Jorge, Ariosto B., Anflor, Carla T. M., Gomes, Guilherme F., & Carneiro, Sergio H. S.

Volume III of Book Series in:

Discrete Models, Inverse Methods, & Uncertainty Modeling in Structural Integrity

Published by: UnB City: Brasilia, DF, Brazil Year: 2022

DOI: <https://doi.org/10.4322/978-65-86503-88-3>

Random Field Discretization in Stochastic Finite Elements: Element Size Effects and Error Indicators in Reliability Structural Analysis

Ariosto B. Jorge^{1a*}, Guilherme F. Gomes²,
Sergio H. Carneiro^{1b}, and Carla T. M. Anflor^{1c}

¹Post-Graduate Program - Integrity of Engineering Materials, University of Brasilia (UnB), Brazil. E-mail: (a) ariosto.b.jorge@gmail.com; (b) shscarneiro@gmail.com; (c) anflor@unb.br

²Mechanical Engineering Institute, Federal University of Itajubá (UNIFEI), Brazil. E-mail: guilhermefergom@unifei.edu.br

*Corresponding author

Abstract

Stochastic finite element methods are used for structural reliability analysis. The element properties and parameters in element equations may involve randomness. One extra dimension is added to the deterministic finite element method to account for the variability in the parameters. Discretization of the structure can be performed and the distributed parameters can be modeled as random fields. Different methods of discretization of the random fields are reviewed in this work. Depending on the discretization method, and on the particular problem being studied, the element size will influence the accuracy and numerical stability of the results. A correlation model is assumed for the random variables that represent the random field in the discretized problem. The correlation models adopted in this work are the triangular and the exponential ones. These models depend on specific parameters: the scale of fluctuation (θ/L) and the correlation length (λ). The type of correlation model and the value of these parameters also influence the reliability index results. Depending on the correlation model adopted, a simple error indicator for the reliability index (β) can be used, by taking, for a specific mesh, the difference between the solution for two different discretization models, one that is known to overestimate the covariance matrix, and another that underestimates this matrix. For the numerical example, the reliability index results showed that both error indicators pointed out to the same range of values, for the parameter in the discretization model, in which bigger errors are expected, or in which the reliability index practically does not change. Values for the correlation model parameter that minimize the differences in the reliability index, for the various discretization models and meshes, are convenient, because, for these values, the model would be more robust and less dependent on external factors, such as the element size of the random field.

Keywords: SFEM - Stochastic Finite Element Methods, random fields, structural reliability, error estimators.

1 Stochastic Finite Elements: Introduction

Stochastic Finite Elements are structural models represented by finite elements, the properties of which involve randomness [Schueller, 1997]. These properties are parameters or variables in the element equations. Examples are the stiffness matrix in static analysis, shape imperfection in buckling analysis, damping and mass properties in structural dynamics. One extra dimension is added to the deterministic finite element method (FEM) to account for the variability in the parameters.

Two categories of Stochastic Finite Elements procedures can be distinguished [Schueller, 1997]. The first category is the representation of Stochastic Finite Elements and their global assembly as random structural matrices. Several representations or methods of discretization have been developed to describe spatial random fluctuations within the element. Some of these representations are the midpoint method [Hisada and Nakagiri, 1985], the nodal point method [Hisada and Nakagiri, 1981], the interpolation method [Liu et al., 1986], the local average method [Vanmarcke and Grigoriu, 1983] and the weighted integral method [Takada, 1990, Deodatis, 1991]. As a tendency [Schueller, 1997], the mid-point methods leads to an overestimation of the variance of the response, the local average method to an underestimation and the weighted integral method leads to the most accurate results. The random field may also be discretized by expansion into a series using basis random variables [Lawrence, 1986], Karhunen-Loeve expansion [Ghanem and Spanos, 1991, Spanos and Ghanem, 1988], or Neumann expansion [Spanos and Ghanem, 1988, Shinozuka and Deodatis, 1988]. The discretization of the random field can be different from the geometric discretization of the deterministic FEM. Depending on the discretization method, and on the particular problem being studied, the element size will influence the accuracy and might lead to singular matrices and numerical instability of the results.

The second category is the evaluation of the stochastic response of the Finite Elements model due to its randomness. The most widely used procedure [Schueller, 1997] for evaluating the stochastic response is the perturbation approach [Hisada and Nakagiri, 1985, Liu et al., 1986, Haldar, 1997]. It is well adapted to the finite element formulation and capable to evaluate first and second moment properties of the response in an efficient manner, for small deviation from the center value. For an accurate reliability assessment, importance sampling and the response surface method might be employed, where the perturbation procedure might be utilized to estimate the response in the neighborhood of some specified center value (design point). Perturbation algorithm is used in this way for the establishment of the sampling distribution in an adaptive procedure.

2 Random Field Discretization Aspects: Introduction

The choice of the random field discretization and of the solution refinement approaches depend on various aspects. The first aspect is what is the problem to be solved and what stochastic finite element formulation will be used. Section 4 describes some existing formulations. The choice will have to consider if the problem is linear or nonlinear, if it is static or time-dependent, if it has or not a closed-form solution, what are the random fields to be considered, if some hypothesis can be done, like assuming Gaussian variables or assuming known covariances or assuming known moments of the distributions. The choice of the formulation may depend also on what information is wanted as a response of the system, and on the available CPU-time to solve the problem.

The second aspect is what kind of discretization will be done for the random fields and to what extent mesh refinement can be done. Section 3 describes some representations of random fields and some existing discretization methods. Depending on the stochastic finite element formulation being used, discretization may be intrinsic to the method, like when using numerical integration. Depending on the assumed or known types of random variables, certain types of discretization may

fit better, or give more accurate results, or demand less CPU-time. Depending on the discretization method adopted, some parameters may be influential in the problem, like the correlation length, or give rise to numerical instabilities, leading to upper and lower bounds for the mesh element size.

A third aspect that may be considered is the influence of the random fields in the response of the system. A sensitivity analysis may be performed [Mahadevan and Haldar, 1991, Mahadevan, 1997] to evaluate this influence. Section 5 describes the sensitivity analysis approach. Depending on the problem, a sensitivity index may be defined so that the relative influence of the various random fields in the response of the system can be assessed. If this index is very small compared to the other random fields, the corresponding random field could be approximated by a single random variable, without much loss in accuracy. Further, if this sensitivity index is very low – less than a cutoff value, say 0.05 – the stochasticity of the random field could be simply ignored, and the field treated as a deterministic variable. This sensitivity analysis may be dependent on the first two aspects, the SFEM formulation and the random field discretization method adopted. If the assumptions behind a discretization method for one random field are not reasonably satisfied, results may be inaccurate, and that may include also the sensitivity index. This means that a decision to disregard a random field could be done based on inaccurate information, and maybe with a different discretization method this decision would not be made.

3 Representation of Random Fields

3.1 Stochastic discretization

This review of the different discretization procedures follows Matthies [Matthies et al., 1997]. In deterministic finite elements, continuous functions are represented by a system of equations written in terms of a finite set of nodal point parameters. In stochastic finite elements, random fields or random processes are represented in terms of a set of discrete random variables. Their statistical properties are expected to depend on the chosen finite element mesh. The discretization of the random fields may be chosen independently of the one for the deterministic part. The selection criteria for the mesh of the random fields and of the geometry are different.

3.2 Element size and mesh refinement in random field discretization

The discretization of the random fields depends on three factors [Liu, 1990]:

The first factor is accuracy. In view of accuracy, the element size is controlled by a parameter, the correlation length, that describes the rate of fluctuation of the random field. The correlation length of a random field is a measure of the distance over which the correlation coefficient approaches a small value. Hence, smaller random field elements should be used if the correlation length is large. If this correlation length is infinity, the random field is, in fact, a single random variable. If the correlation length is small, the element size has to be short enough so that a good approximation of the random field by the random variable in within the element is obtained. According to Liu, [1990], an element size of one half of the correlation length suffices for a satisfactory representation of the random field. Another parameter used by Hisada and Nakagiri, [1981] is λ_u , the largest wave number that cannot be disregarded in the power spectral density $S_{HH}(\lambda)$. They concluded that the nodal interval should be taken to be less than $0.25\lambda_u$. Any of these parameters provide an upper bound for the distance between two adjacent random variables.

The second factor is the numerical stability of the transformation to the standard normal space. An excessively fine mesh yields highly correlated random variables. Therefore, the correlation matrix is nearly singular and, in cases where a transformation to the standard normal space is required, this may cause numerical difficulties. Hence, this factor provides a lower bound on the element size. In the local average method [Vanmarcke and Grigoriu, 1983], it was observed that

the reduction of the point variance under local averaging was inversely proportional to T , the averaging interval, for large values of T . This proportionality constant, the scale of fluctuation θ , is the approximate length over which strong correlation persists in the random field. Element sizes can be chosen larger than the scale of fluctuation to avoid numerical difficulties.

The third factor is the efficiency of the computational method. A smaller number of basic variables would require less computation time, especially when partial derivatives of the functions are to be computed. Thus, the elements should be as large as possible from this viewpoint. Matrix condensation [Mahadevan, 1997] can be done to reduce the computational effort in deterministic analysis as well as in the computation of response sensitivity.

These controlling factors are different from the factor in the selection of the finite element mesh, which is basically governed, in the static problem, by the gradient in the displacement field.

Der Kiureghian and Ke, [1988] suggested separate meshes for the finite element and for the random fields. In general, it is appropriate to use a finite element mesh that satisfies the requirements based on the displacement gradient and the correlation lengths of all random fields, and then for each random field choose a mesh that is coincident with or coarser than the finite element mesh such that the corresponding probability transformation remains stable. In their work, the influence of the size of the random field elements in the accuracy and stability of the results was not explored.

Extra difficulties in investigating the effect of element size on the stability and accuracy of the results may arise due to numerical errors dependent on the solution algorithm or in problems where no exact solution is available. This is the case in reliability analysis, which does not incorporate random fields. Hence, there is no exact solution that can be used to check the accuracy of the analysis for a certain random field mesh, and only comparisons between results of a coarse mesh and a refined mesh can be done.

An analysis of the error resulting from discretizing random fields in the SFEM was presented by Spanos and Zeldin, [1997]. This error analysis is performed for the shift invariant subspace discretization method which encompasses the midpoint method, the interpolation method and the local averaging method, giving a simple criterion of estimating the accuracy of these random field discretization schemes. Their work investigates the influence of the random field discretization on the response variability of the stochastic mechanics problem, deriving an a-posteriori bound for the response variability.

The kind of discretization method affects the quality and the properties of the representation. For different random fields in a structural problem, optimal mesh sizes might be different. These meshes might be different from the mesh of the deterministic problem. And for each random field this optimal mesh size may also vary depending on the type of discretization method being used. Upper and lower bounds for the element size for a certain random field may exist and be different from one discretization method to another.

3.3 Point discretization methods

3.3.1 General

Point discretization methods represent the random field $H(\mathbf{x})$ by its values at one or more points. The random field evaluated at a certain point is a random variable. At point i the discretized value is given by

$$H_i = H(\mathbf{x}_i) \quad (1)$$

The mean, the variance and the marginal distribution of the random variable $H(\mathbf{x}_i)$ correspond to those of the random field at the point \mathbf{x}_i . Covariance (matrix) between two points is computed in

terms of a given autocovariance function $C_{HH}(\mathbf{x}, \mathbf{y})$:

$$C_{H_i H_j} = C_{HH}(\mathbf{x}_i, \mathbf{x}_j) \quad (2)$$

All point discretization methods have advantages and disadvantages.

The advantages are:

- i) the covariance matrix can be easily computed;
- ii) the covariance matrix is positive definite;
- iii) same distribution function in both the discretized and in the continuous case. Therefore, there is no restriction of this method to Gaussian random fields only.

The disadvantages are:

- i) The mesh size has to be small compared to the correlation length, so that random properties could be considered as constant within the element;
- ii) The shape and size of all the elements should be the same;
- iii) A fine mesh requires much CPU-time. A coarse mesh will not fulfill the accuracy requirements. These methods will only fulfill accuracy requirements with coarse meshes for medium to long correlation distances.

3.3.2 The midpoint method

The value at the center point of the element is used to represent the random field. The coordinates of the centroid are given in terms of the nodal coordinates $\mathbf{x}_j^{(i)}$ of the i^{th} element:

$$\mathbf{x}_i^* = \frac{1}{N} \sum_{j=1}^N \mathbf{x}_j^{(i)} \quad (3)$$

where N denotes the number of nodes of the element. The midpoint method tends to over-represent the variability of the random field within each element.

3.3.3 The nodal point method

In this method, the coordinates of the points used to represent the random field are given by

$$\mathbf{x}_i^* = \mathbf{x}_i \quad (4)$$

where \mathbf{x}_i denotes the coordinate vector of the i^{th} global node.

3.3.4 The integration point method

If numerical integration is applied to the problem, then the random field is implicitly discretized by means of the integration points. This discretization is not done in the first step, but when integration is needed. An example is the evaluation of covariances and correlations when using local averages. The coordinates of the points depend on the rules of the integration scheme adopted. A common scheme is the Gauss quadrature. Also, this method has been used directly by [Brenner, \[1993\]](#).

3.4 The local averaging method

The local averaging method [[Vanmarcke and Grigoriu., 1983](#)] provides discretized values of the random field given by

$$H_i = \frac{1}{vol(\Omega_i)} \int_{\Omega_i} H(\mathbf{x}) d\mathbf{x} \quad (5)$$

where Ω_i is the domain over which the integration has to be performed. The domain represents an averaging time, in the case of a stochastic process, or the area of the i^{th} element in the case of stochastic finite elements. If only functions of space are involved, the method is called the spatial averaging method. The spatial averaging method tends to underrepresent the variability within each element, and yields accurate results even for rather coarse meshes [Der Kiureghian and Ke., 1988]. This method is practically limited to cases where the field is Gaussian. For other fields, the distribution function of the random variable H_i is difficult or even impossible to obtain.

3.5 The interpolation method

Liu et al., [1986] suggested an approximation of the random field $H(\mathbf{x})$ by q nodal values using shape functions $N_i(\mathbf{x})$ by

$$H(\mathbf{x}) = \sum_{i=1}^q N_i(\mathbf{x}) H_i \quad (6)$$

where H_i are the nodal values of $H(\mathbf{x})$ at suitable nodes \mathbf{x}_i . The shape functions need not to be the same from the finite element shape functions. Nodal points need not to coincide with the ones of finite elements. The expected value and the variance of the random field written with this discretization are, respectively:

$$E[H(\mathbf{x})] = \sum_{i=1}^q N_i(\mathbf{x}) E[H_i] \quad (7)$$

$$Var[H(\mathbf{x})] = \sum_{i,j=1}^q N_i(\mathbf{x}) N_j(\mathbf{x}) C_{HH}(\mathbf{x}_i, \mathbf{x}_j) \quad (8)$$

where $C_{HH}(\mathbf{x}, \mathbf{y})$ is the auto-covariance function of the random field. Depending on the shape functions, good approximations are to be expected even for coarse meshes and elements of quite different size. Choosing the nodes to be the centroids of the random field elements and the shape functions to be unity inside each element and zero elsewhere, reduces the interpolation method to the midpoint method.

3.6 The Kriging technique

The Kriging technique [Der Kiureghian et al., 1991] (optimum linear estimation method) can be seen as the linear regression of the original random field onto some linear functionals of the field, usually simple [Ditlevesen., 1995], which make use of a discretizing mesh different from the one from the deterministic part. The covariance function of the random field is then employed for an interpolation. If the stochastic and deterministic meshes become identical, the method reduces to the weighted integral method Deodatis, [1991]. This method is valid for Gaussian random fields only, because this is one of the assumptions in linear regression.

3.7 Series expansion method

3.7.1 General

In the above methods the defining parameters of the discretization were linear functionals of the random field. These methods are convenient if the covariance function of the random field is known and if the random field itself is Gaussian. If either the random field is not Gaussian (thus, the random field is modeled as a nonlinear function of a Gaussian field) or the deterministic quantity of interest is a nonlinear function of other random variables (hence, the covariance function is

not easy to compute), these methods are not directly applicable anymore. Nonlinear methods such as the homogeneous chaos expansion method [Ghanem and Spanos, 1991] and the perturbation method [Hisada and Nakagiri, 1985, Liu et al., 1986] may be used in this case.

3.7.2 The basis random variable method

Lawrence, [1986, 1987] represents the random field $H(\mathbf{x})$, of which the first and second moment functions are known, by a series

$$H(\mathbf{x}) = \sum_{j=1}^{\infty} h_{0j} \phi_j(\mathbf{x}) + \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} h_{ij} e_i \phi_j(\mathbf{x}) \tag{9}$$

where the $\phi_j(\mathbf{x})$ are a set of linearly independent and preferably orthogonal deterministic functions satisfying all the kinematic boundary conditions. The statistically independent basis random variables e_i have the properties

$$E[e_i] = 0 \text{ for } i = 1, 2, \dots, \tag{10}$$

$$E[e_i e_j] = \delta_{ij} \tag{11}$$

where δ_{ij} is the Kronecker delta. The series is truncated by using only $i = N$ basis random variables (necessary to model the properties of the random field), and $j = M$ linearly independent deterministic functions. Approximating $H(\mathbf{x})$ over the interval $[a, b]$, the coefficients h_{ij} can be determined by a least-square fit of the first and second moments of the truncated random field to the first and second moments of $H(\mathbf{x})$.

3.7.3 The kernel expansion method

The kernel expansion method, also known as Karhunen-Loeve expansion [Ghanem and Spanos, 1991, Spanos and Ghanem, 1988] is a type of spectral representation of the random field $H(\mathbf{x}, \theta)$ by

$$H(\mathbf{x}, \theta) = \sum_{i=1}^{\infty} \sqrt{\lambda_i} \xi_i(\theta) \phi_i(\mathbf{x}) \tag{12}$$

where $\xi_i(\theta)$ denotes a denumerable set of random variables to be determined, λ_i is some constant and $\phi_i(\mathbf{x})$ are orthonormal deterministic functions which can be related to the covariance kernel of $H(\mathbf{x}, \theta)$. The symbol θ denotes an element of the space of random events. After truncating the series at N terms, the random field is expressed by

$$H(\mathbf{x}, \theta) = \bar{H}(\mathbf{x}) + \sum_{i=1}^N \sqrt{\lambda_i} \xi_i(\theta) \phi_i(\mathbf{x}) \tag{13}$$

where $\bar{H}(\mathbf{x})$ is the expected value of the random field and λ_i and $\phi_i(\mathbf{x})$ are the eigenvalues and eigenfunctions of the covariance kernel, respectively, obtained as the solutions to the eigenvalue problem:

$$\int_{\Omega} C(\mathbf{x}_1, \mathbf{x}_2) \phi_i(\mathbf{x}_2) d\mathbf{x}_2 = \lambda_i \phi_i(\mathbf{x}_1) \tag{14}$$

where $C(\mathbf{x}_1, \mathbf{x}_2)$ is the auto-covariance function of the stochastic field. The random variables $\xi_i(\theta)$ have properties similar to those of the basis random variables:

$$E[\xi_i(\theta)] = 0 \tag{15}$$

$$E[\xi_i(\theta) \xi_j(\theta)] = \delta_{ij} \tag{16}$$

This expansion minimizes the mean-square error introduced by the truncation. For a decreasing correlation length the number of terms has to grow for a certain required accuracy [Ghanem and Spanos, \[1991\]](#), [Spanos and Ghanem, \[1988\]](#). If the covariance function of the field to be expanded is not known another kind of spectral representation can be done, the homogeneous chaos expansion [Ghanem and Red-Horse, \[1999\]](#), [Ghanem, \[1999\]](#).

3.7.4 The homogeneous chaos expansion method

In this method the solution field can be expressed as

$$u = h[\xi_i(\theta), \mathbf{x}], \tag{17}$$

where u denotes the solution field and $h[.]$ is a nonlinear functional of its arguments. First, the random fields are replaced by their corresponding Karhunen-Loeve expansions. Next, a nonlinear expansion of $h[.]$ is done using the set of all orthogonal polynomials Γ_p not exceeding order p , called the polynomial chaos of order p , in terms of a set of orthogonal, uncorrelated, Gaussian random variables $\{\xi_i(\theta)\}_{i=1}^{\infty}$. Assuming symmetry of the polynomials, which is always possible [[Ghanem and Spanos, 1991](#)], the random variables $\xi(\theta)$ of the random field, whose statistical properties are not known, can be expressed as

$$\begin{aligned} \chi(\theta) = & a_0\Gamma_0 + \sum_{i_1=1}^{\infty} a_{i_1}\Gamma_1(\xi_{i_1}(\theta)) + \sum_{i_1=1}^{\infty} \sum_{i_2=1}^{i_1} a_{i_1 i_2}\Gamma_2(\xi_{i_1}(\theta), \xi_{i_2}(\theta)) \\ & + \sum_{i_1=1}^{\infty} \sum_{i_2=1}^{i_1} \sum_{i_3=1}^{i_2} a_{i_1 i_2 i_3}\Gamma_3(\xi_{i_1}(\theta), \xi_{i_2}(\theta), \xi_{i_3}(\theta)) \\ & + \sum_{i_1=1}^{\infty} \sum_{i_2=1}^{i_1} \sum_{i_3=1}^{i_2} \sum_{i_4=1}^{i_3} a_{i_1 i_2 i_3 i_4}\Gamma_4(\xi_{i_1}(\theta), \xi_{i_2}(\theta), \xi_{i_3}(\theta), \xi_{i_4}(\theta)) + \dots \end{aligned} \tag{18}$$

where Γ_p are successive polynomial chaoses of their arguments and a_{i_1, \dots, i_p} are the deterministic constants. The random variable series, truncating after the J^{th} polynomial reads

$$\chi(\theta) = \sum_{i_1=1}^J \kappa_i \Psi_i[\{\xi_r\}] \tag{19}$$

where κ_i and $\Psi_i[\{\xi_r\}]$ are identical to a_{i_1, \dots, i_p} and $\Gamma_p(\xi_{i_1}, \dots, \xi_{i_p})$, respectively. A computational implementation of the method uses a finite set of random variables, the n -dimensional polynomial chaos of order p , which is a subset out of the polynomial chaos of order p , function of only n of the random variables ξ_i . The convergence properties of the representation depend on the choice of the ξ_i out of the infinite set as well as on the number of random variables. This expansion is restricted to multidimensional functions of Gaussian variables. For non-Gaussian fields, an approximation via the polynomial chaos could be used.

3.7.5 The perturbation method

A Taylor series representation [Ghanem and Spanos., 1991] of a random field $H(\mathbf{x})$ about the mean value of the m random variables $\{h_i(\mathbf{x})\}_{i=1}^m$ yields

$$\begin{aligned}
 H(\mathbf{x}) &= g(h_1(\mathbf{x}), \dots, h_m(\mathbf{x})) \\
 &= g(\bar{h}_1(\mathbf{x}), \dots, \bar{h}_m(\mathbf{x})) + \sum_{i=1}^m \frac{\partial g(\bar{h}_1(\mathbf{x}), \dots, \bar{h}_m(\mathbf{x}))}{\partial h_i(\mathbf{x})} [h_i(\mathbf{x}) - \bar{h}_i(\mathbf{x})] \\
 &+ \sum_{i=1}^m \sum_{j=1}^m \frac{\partial^2 g(\bar{h}_1(\mathbf{x}), \dots, \bar{h}_m(\mathbf{x}))}{\partial h_i(\mathbf{x}) \partial h_j(\mathbf{x})} [h_i(\mathbf{x}) - \bar{h}_i(\mathbf{x})][h_j(\mathbf{x}) - \bar{h}_j(\mathbf{x})] + \dots \\
 &= \bar{g} + \sum_{i=1}^m \bar{g}_i \epsilon_i + \sum_{i=1}^m \sum_{j=1}^m \bar{g}_{ij} \epsilon_i \epsilon_j + \dots
 \end{aligned} \tag{20}$$

Evaluating all random fields and functions involved in the equation leads to a multidimensional polynomial in ϵ . Equating polynomials of the same order on both sides yields a set of equations to be solved. This method is well suited for small fluctuations (in comparison with the mean value), in which a first-order perturbation analysis shows good results. Second-order perturbation analysis improves first-order results only a little and may require a huge amount of CPU-time. The perturbation method requires many terms for representing random fields with a large coefficient of variation. But the use of higher order terms of the expansion may induce numerical rounding error even for small coefficients of variation, affecting the accuracy of the representation [Ghanem and Spanos., 1991].

4 Stochastic Finite Element Formulations

This review of the different Stochastic finite element formulations follows Matthies et al., [1997]. The general matrix form of the equation of motion is

$$\mathbf{M}\ddot{\mathbf{v}}(t) + \mathbf{C}\dot{\mathbf{v}}(t) + \mathbf{K}\mathbf{v}(t) = \mathbf{f}(t) \tag{21}$$

with proper initial conditions. The matrices \mathbf{K} , \mathbf{C} and \mathbf{M} and the force vector \mathbf{f} may involve randomness. In the case of static problems, this equation reduces to

$$\mathbf{K}\mathbf{v} = \mathbf{f} \tag{22}$$

The stiffness matrix can be split into two parts, the non-fluctuating part \mathbf{K}_0 and the fluctuating part \mathbf{K}_* . To solve the static problem, several methods were proposed, as follows.

4.1 Stochastic global stiffness matrix expansions

4.1.1 The weighted integral method

The weighted integral method [Takada., 1990, Deodatis., 1991] does not require a separate discretization of the random field. There is no independent mesh for the stochastic part, but everything is determined by the chosen mesh for the deterministic part. The stochastic field can be written as

$$\alpha(\mathbf{x}) = \alpha_0(\mathbf{x}) \cdot [1 + f_*(\mathbf{x})] \tag{23}$$

where $f_*(\mathbf{x})$ is a zero mean homogeneous isotropic Gaussian field with auto-covariance function C_{ff} and $\alpha_0(\mathbf{x})$ is the mean of the random field at \mathbf{x} . The fluctuating part of the element stiffened

matrix can be written as

$$\mathbf{K}_*^{(e)}(i, j) = \mathbf{K}_{*,ij}^{(e)} = a_{ij} \int_{\Omega_e} x^r y^s z^t \alpha_0(x, y, z) dx dy dz \tag{24}$$

where the weighted integral is a random variable, $X_i^{(e)}$, and the associate weight $a_{ij} = \Delta \mathbf{K}^{(e)}$ is a deterministic constant matrix. Polynomial shape functions are used. The number of weighted integrals N_w depends on the type of element used, which determines the shape function to use. The random field is projected onto the space spanned by the shape functions. The fluctuating part for element stiffness matrix is

$$\mathbf{K}_*^{(e)} = \sum_{i=1}^{N_w} X_i^{(e)} \Delta \mathbf{K}_i^{(e)} \tag{25}$$

and the random global stiffness matrix is

$$\mathbf{K} = \sum_{e=1}^{M_e} \mathbf{K}^{(e)} = \sum_{e=1}^{M_e} (\mathbf{K}_0^{(e)} + \sum_{i=1}^{N_w} X_i^{(e)} \Delta \mathbf{K}_i^{(e)}) \tag{26}$$

The stochastic part of the stiffness matrix is described in terms of $N_w \cdot M_e$ weighted integrals. The response statistics can be determined by a first-order Taylor series expansion of the response vector \mathbf{v} about E , where the mean values of the random variables are $\bar{X}_i^{(e)}$

$$\mathbf{v} \approx \mathbf{v}_0 + \sum_{e=1}^{M_e} \sum_{i=1}^{N_w} \left. \frac{\partial \mathbf{v}}{\partial X_i^{(e)}} \right|_E (X_i^{(e)} - \bar{X}_i^{(e)}) \tag{27}$$

Where $\left. \frac{\partial \mathbf{v}}{\partial X_i^{(e)}} \right|_E$ is obtained by partially differentiating Eq. 22

$$\left. \frac{\partial \mathbf{v}}{\partial X_i^{(e)}} \right|_E = -\mathbf{K}_0^{-1} \left. \frac{\partial \mathbf{K}}{\partial X_i^{(e)}} \right|_E \mathbf{v}_0 \tag{28}$$

From Eq. 27 the statistics $E[\mathbf{v}]$ and $Cov[\mathbf{v}, \mathbf{v}]$ can be obtained [Matthies et al., 1997]. The major drawback of this method is the first order Taylor series expansion which partially ruins the improvement in quality of representation of the method. This method yields good results in the domain of short correlation distances, where the midpoint discretization method fails. For correlation distance about equal or larger than the geometric size of the problem, both methods approach each other. If numerical integration is applied, instead of analytical integration, then there is an implicit discretization done by the integration points used in the quadrature.

4.1.2 The Neumann series expansion

In conjunction with stochastic finite elements, the expansion of the inverse of the stochastic global stiffness matrix in a Neumann series can be done [Ghanem and Spanos., 1991, Spanos and Ghanem., 1988, Shinozuka and Deodatis., 1988]. The stochastic static problem Eq.22 can be written

$$\mathbf{v} = (\mathbf{K}_0 + \mathbf{K}_*)^{-1} \mathbf{f} = (\mathbf{I} + \mathbf{K}_0^{-1} \mathbf{K}_*)^{-1} \mathbf{K}_0^{-1} \mathbf{f} = (\mathbf{I} + \mathbf{J})^{-1} \mathbf{K}_0^{-1} \mathbf{f} \tag{29}$$

The term in parentheses can be expanded in a kind of geometric series, called the Neumann series

$$\mathbf{K}^{-1} = (\mathbf{K}_0 + \mathbf{K}_*)^{-1} = (\mathbf{I} - \mathbf{J} + \mathbf{J}^2 - \mathbf{J}^3 + \dots) \mathbf{K}_0^{-1} = \sum_{k=0}^{\infty} (-\mathbf{K}_0^{-1} \mathbf{K}_*)^k \mathbf{K}_0^{-1} \tag{30}$$

This expansion has the advantage that only the non-fluctuating part of the stiffness matrix has to be inverted and this only once. The response vector is now represented by the series

$$\mathbf{v} = (\mathbf{I} - \mathbf{J} + \mathbf{J}^2 - \mathbf{J}^3 + \dots) \mathbf{v}_0 = \mathbf{v}_0 - \mathbf{v}_1 + \mathbf{v}_2 - \mathbf{v}_3 + \dots \quad (31)$$

This series solution is equivalent to the recursive formula

$$\mathbf{K}_0 \mathbf{v}_i = \mathbf{K}_* \mathbf{v}_{i-1} \quad i = 1, 2, \dots \quad (32)$$

Once \mathbf{v}_0 is obtained from $\mathbf{v}_0 = \mathbf{K}_0^{-1} \mathbf{f}$, every subsequent \mathbf{v}_i is determined. [Shinozuka and Deodatis, \[1988\]](#) derived analytical formulas for the response variability using first-order Neumann expansion, $\mathbf{v} = (\mathbf{I} - \mathbf{J}) \mathbf{v}_0$. The improved Neumann expansion [\[Ghanem and Spanos., 1991\]](#) is obtained when the Karhunen-Loeve expansion is used to represent the stochastic field describing the randomness of the stiffness matrix and then Neumann series expansion of the response vector is written. The number of terms included in both Karhunen-Loeve and Neumann expansions will dictate the cost in CPU-time and the accuracy of the model.

4.1.3 The Taylor series expansion

The perturbation method [\[Hisada and Nakagiri., 1985, Liu et al., 1986\]](#) can be applied on the stochastic stiffness matrix, where the random field $\alpha(x, y, z) = H(\mathbf{x})$ has to be discretized by a set of N zero mean random variables $\{\alpha_i\}_{i=1}^N$, as follows:

$$\mathbf{K} = \mathbf{K}_0 + \sum_{i=1}^N \mathbf{K}_i^I \alpha_i + \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \mathbf{K}_{ij}^{II} \alpha_i \alpha_j + \dots \quad (33)$$

where \mathbf{K}_0 is the stiffness matrix evaluated at the mean value of the vector of random variables $\alpha = \mathbf{0}$, $\mathbf{K}_i^I = \left. \frac{\partial \mathbf{K}}{\partial \alpha_i} \right|_{\alpha=\mathbf{0}}$, $\mathbf{K}_{ij}^{II} = \left. \frac{\partial^2 \mathbf{K}}{\partial \alpha_i \partial \alpha_j} \right|_{\alpha=\mathbf{0}}$ and so on. This method can be applied to static and dynamic problems, linear and nonlinear. The response vector \mathbf{v} and the external force vector \mathbf{f} can be expanded in a similar way as the stiffness matrix [\[Matthies et al., 1997\]](#). In comparison with the Neumann series expansion, the Taylor series expansion has the severe disadvantage that partial derivatives with respect to the random variables are required. Computational effort increase dramatically when including higher-order terms. The results are in general quite satisfying for small coefficients of variation of the underlying stochastic fields, using just first- and second-order perturbation methods. For a wider range of fluctuations, a successive perturbation method [\[Hisada and Nakagiri., 1985\]](#) can be applied.

4.2 Monte-Carlo simulation

Monte Carlo simulation (MCS) can be used to generate samples of the global stiffness matrix and, hence, by solving the linear equation system, samples of the response can be obtained. As the generation of random variables according to a specific distribution is an easy task in the uncorrelated space, usually the covariance matrix of all random variables involved in the problem has to be calculated [\[Matthies et al., 1997\]](#). Integrals have to be performed to compute the covariance matrix between two element stiffness matrix entries, with extensive need of computational time even for small problems. The number of random variables has to be decreased by one of the methods in Section 3, which involve a point discretization, or by the weighted integral method in Section 4.1.1. In this case, as the integrations would only be performed numerically, this method will also involve a point discretization. The total number of random variables involved in a stochastic problem is at most equal to the total number of different integration points used. The distance between two integration points has to be less than the upper bound discussed in Section 3.1.

4.3 Implicit time integration with Hermite-Gauss quadrature

This method was developed by Liu et al., [1986] and applied to a stochastic finite element system described by Eq. 21, where \mathbf{M} and $\mathbf{f}(t)$ are deterministic, $\mathbf{C} = \mathbf{0}$, and \mathbf{K} is a stochastic stiffness matrix involving a set of random variables $\{\alpha_i\}_{i=1}^N$. The system is solved using a finite difference recursive algorithm. Numerical integrations have to be performed at every time step. Accuracy and the need of CPU-time are determined by the number of integration points used and also by the observation time.

4.4 SFEM using the homogeneous chaos expansion

Homogeneous chaos expansion can be applied to stochastic finite elements [Ghanem and Spanos., 1991]. In Eq. 22, the Karhunen-Loeve expansion is used to represent the random field involved in the global stiffness matrix. The Karhunen-Loeve expansion is then applied to the solution vector \mathbf{v} . As the covariance C_{vv} of the response is not known, the homogeneous chaos expansion is applied to the random variables χ_j describing the response \mathbf{v} . Multiplying the obtained equation by all the polynomial chaoses and requiring the mean-square representation error to be minimal leads to an equation that, after imposing the kinematic boundary conditions will supply a set of algebraic equations to be solved for the coefficients of the polynomial chaoses. Properties of these chaoses, nominally, the zero mean and the orthogonality, are then used to evaluate the mean and the covariance of the response.

4.5 Basis random variable-based SFEM

Considering a static finite element problem, the involved random functions can be replaced by a series using basis random variables as described in Section 3.7.2. The first- and second- moments of these functions are assumed to be known. The vector of random displacement functions \mathbf{v} is expanded as

$$\mathbf{v} = \mathbf{N}^T \mathbf{d}_0 + \sum \mathbf{N}^T \mathbf{d}_i e_i \tag{34}$$

where \mathbf{N}^t is the matrix of deterministic shape functions defined within the boundaries of the respective element, \mathbf{d}_0 is the mean nodal displacement vector and the \mathbf{d}_i comprise the second-moment properties of the nodal displacement vector. In the case of linear elasticity, the material stiffness is a linear operator represented by the constitutive matrix \mathbf{D} in $\mathbf{D}\epsilon = \sigma$. The basis random variable expression of the matrix \mathbf{D} and of the nodal force \mathbf{f} give

$$\mathbf{D} = \mathbf{D}_0 + \sum \sum \mathbf{D}_{ij} e_i \psi_j \tag{35}$$

$$\mathbf{f} = \mathbf{f}_0 + \sum \mathbf{f}_i e_i \tag{36}$$

where the matrix \mathbf{D}_0 is the mean constitutive matrix, \mathbf{D}_{ij} is a deterministic coefficient matrix, where the coefficients are defined by the least-squares fit described in Section 3.7.2 and ψ_j is a matrix of deterministic shape functions defined over the entire structure. The strain-displacement relation is $\epsilon = \mathbf{B}\mathbf{v}$. The random strain energy density is $\frac{1}{2}\epsilon^T \sigma = \frac{1}{2}\epsilon^T \mathbf{D}\epsilon$. The random strain energy of the element is the integral of the random strain energy density over the volume of the element. Integrating over the volume of an element, taking the expected value and differentiating with respect to the \mathbf{d}_i yields the element stiffness matrix. For the integration, the function matrix ψ_j can be evaluated at the center of the element (midpoint method) or at the Gauss points (integration point method). The assembled finite element equation system is similar to the deterministic system, with an extra set of dimensions, included due to the random character of the problem. The global stiffness matrix is sparse and can be solved by an iterative method.

5 Sensitivity Analysis

The SFEM formulation and the discretization method may require huge computational effort. In order to ensure efficiency of the computations, without loss of accuracy, some random fields related to some of the distributed parameters of the problem may be considered, as an approximation, as random variables. A sensitivity analysis may be performed [Mahadevan and Haldar, 1991, Mahadevan, 1997] to evaluate the influence of the random fields in the response of the system. In the sensitivity analysis approach, depending on the problem, a sensitivity index may be defined so that the relative influence of the various random fields in the response of the system can be assessed. If this index is very small compared to the same index from other random fields, the corresponding random field could be approximated by a single random variable, without much loss in accuracy of the response. Further, if this sensitivity index is very low – less than a cutoff value, say 0.05 – the stochasticity of the random field could be simply ignored, and the field treated as a deterministic variable. This sensitivity analysis may be dependent on the SFEM formulation and on the random field discretization method adopted. If the assumptions behind a discretization method for one random field are not reasonably satisfied, results may be inaccurate, and that may include also the sensitivity index. This means that a decision to disregard a random field could be done based in inaccurate information, and maybe with a different discretization method this decision would not be made.

Sensitivity analysis is not the only guideline to reduce the size of the problem without much loss in accuracy. The representation of a distributed parameter by a single random variable implies a perfectly correlated random field, i. e., a random field with infinite correlation length. Therefore, if a distributed parameter has a very large correlation length, it may be modeled as a random variable, despite its having a high sensitivity index.

6 Error Indicators For the Random Field Discretization in Reliability Structural Analysis

The previous sections review the existing stochastic finite element formulations, the discretization methods proposed up to now and some aspects already discussed in the literature concerning element sizes and mesh refinement.

Discretization has to be addressed even for the SFEM methods that does not require it explicitly if numerical integration is to be performed.

Upper and lower bounds of the error exist only for certain discretization methods. They were obtained in the context of some specific problems, like static formulation (structures) or reliability analysis, and they may be problem-specific.

Error estimates that could lead to an adaptive remeshing of the random field exist only for the shift invariant subspace discretization method, which uses an spectral approach.

Depending on the discretization method, and on the particular problem being studied, the element size will influence the accuracy and numerical stability of the results. A correlation model is assumed for the random variables that represent the random field in the discretized problem. The correlation models adopted in this work are the triangular and the exponential ones. These models depend on specific parameters: the scale of fluctuation (θ/L) and the correlation length (λ). The type of correlation model and the value of these parameters also influence the reliability index results.

Two different error indicators for the reliability index (β) are proposed in this work. These error indicators can be useful to evaluate the influence of the random field discretization model and mesh for stochastic finite element methods (SFEM) as applied to structural reliability analysis.

1. Depending on the correlation model adopted, a simple error indicator for the reliability index (β) can be used, by taking, for a specific mesh, the difference between the solution for two different discretization models, one that is known to overestimate the covariance matrix, and another that underestimates this matrix. In this work, the methods used to over- and under-estimate the covariance matrix are the midpoint method and the local average method, respectively.
2. Another error indicator can be easily established by taking differences between two different random field meshes, for the same deterministic finite element mesh and the same discretization model.

It has to be pointed out that it is necessary to first change some characteristic of the discretization of the problem and then compare the solutions for the reliability index for both cases. In the above error indicators, either the random field discretization model is changed, or the mesh is refined. If Monte Carlo Simulation (*MCS*) for correlated variables were performed for a certain discretization model / mesh, it will give the same results as in *FORM method 1*, because to perform the correlated *MCS*, it is necessary to get the same set of uncorrelated reduced normal variables Y as in *FORM*. This means that *MCS* will be simulating the very same limit state equation then *FORM*, and the results for the reliability index will be the same. This means that the comparison between *FORM* and *MCS* does not give any other error indicator that could be used.

7 Numerical Example

A numerical example is performed to compare the two error indicator approaches. This example is presented in Mahadevan and Haldar, [1991]. A cantilever beam made of structural steel, with Young's modulus E and moment of inertia I , has a uniform load W acting on it vertically downward. For this example, the beam is discretized in one, two and four elements. The discretization in four elements is shown in Fig. 1.

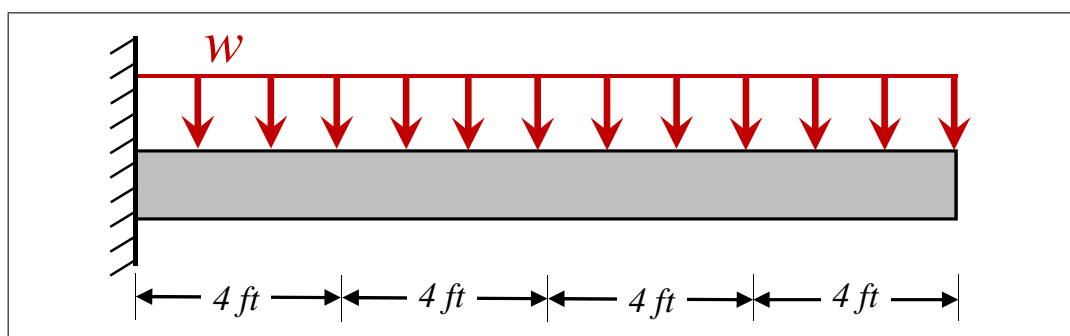


Figure 1: Numerical example: cantilever beam. Figure shows discretization with 4 elements: W1, W2, W3 & W4 - adapted from Jorge, [2004]

7.1 SFEM random field discretization for the beam

7.1.1 Beam considered as one element only

Assuming all parameters are deterministic, the solution for the vertical deflection at the free end of the beam is $y_{max} = \frac{WL^4}{8ET}$. The finite element solution for one element gives also the exact solu-

Table 1: Reliability analysis of cantilever beam: W, E and I as random variables - adapted from Jorge, [2004]

Quantity	Units	Sensitivity Index	Initial Checking point	Final checking point
W	k/in	-1.0000	0.040	0.048
E	ksi	0.0000	29000.000	28396.361
I	in ⁴	0.0002	301.000	296.678

$\beta=1.053$

Number of iterations = 3

tion. This means that no improvement in the deterministic solution is obtained while subsequently discretizing the beam. When considering the stochasticity of these parameters, the difference in the reliability index results for the various random field meshes or for the different discretization models is not affected by the deterministic FEM mesh. Any changes in the solution can be related to the discretization of the random fields only.

The limit state considered is that the vertical deflection at the free end of the beam is less or equal than length/200, with length in inches. Considering W , E and I as random variables, the limit state equation is

$$g = R - S = EI(y_{adm} - y_{max}) = EI \frac{L}{200} - \frac{WL^4}{8} \tag{37}$$

Assuming the statistical description of the basic variables as

$$W \sim N\left(0.04 \frac{K}{in}; 0.008 \frac{K}{in}\right)$$

$$E \sim N(29000 \text{ksi}; 1740 \text{ksi})$$

$$I \sim N(301 \text{in}^3; 15.05 \text{in}^3)$$

A reliability analysis *FORM Method 1* is performed as in Haldar and Mahadevan, [2000]. The results are shown in Table 1.

The sensitivity indexes in Table 1 indicated that the random variables E and I could be considered as deterministic in Eq. 37. Reliability analysis results are, in this case:

Checking point: $W = 0.049 \frac{K}{in}$
 $\beta = 1.166$
 Iterations: 2

7.1.2 Beam discretized in two elements

The parameters W , E and I are now random fields. Let's assume E and I as deterministic parameters. Let's assume that the random field W has normal distribution at every point with same mean, but different variance as a function of the longitudinal coordinate x , for $0 \leq x \leq L$. Let's consider a linear variation of the coefficient of variation (δ) with respect to x : $\delta = 0.048 + \frac{x}{L} 0.024$ so that its average is $\delta = 0.06$, which is the value that was used when considering W as a random variable, previously.

The random field W is then discretized into two elements, W_1 and W_2 , while parameters E and I are considered as deterministic. A closed form for the vertical deflection at the free end of the beam y_{max} can be obtained in terms of W_1 and W_2 , so that limit state equation can be written as

$$g = R - S = EI(y_{adm} - y_{max}) = EI \frac{L}{200} - \frac{L^4}{6144} (112W_1 + 656W_2) \tag{38}$$

where the random variables W_1 and W_2 are correlated. A reliability analysis with correlated variables is performed by rewriting the limit state equation with respect to a set of uncorrelated reduced normal variables Y and then using *FORM Method 1* [Haldar and Mahadevan, 2000].

At this point, a model for the correlation of the random variables W_1 and W_2 has to be assumed. The model expresses the decay of the variance function with the distance between two points. Two models are used in this work: the triangular and the exponential [Mahadevan and Haldar, 1991].

In the triangular model, the approximation for the decay of the variance function is given by

$$\begin{aligned} \gamma_x(T) &= 1 - \frac{T}{3\theta_x}, T < \theta_x \\ \gamma_x(T) &= \frac{\theta_x}{T} \left(1 - \frac{\theta_x}{3T}\right), T > \theta_x \end{aligned} \quad (39)$$

where $\gamma_x(T)$ is the variance function of the random field X , θ_x is the scale of fluctuation, and T is the distance. Using this approximation, the correlation coefficient between two discretized random variables in the i^{th} and j^{th} elements may be expressed as

$$\rho(x_i, x_j) = \frac{1}{2} \left\{ (k-1)^2 \gamma_x \left[\frac{(k-1)L}{N} \right] - 2k^2 \gamma_x \left[\frac{kL}{N} \right] + (k+1)^2 \gamma_x \left[\frac{(k+1)L}{N} \right] \right\} \quad (40)$$

in which $k = |i - j|$.

In the exponential model, the correlation coefficient between any two random variables E_i and E_j obtained by the discretization of E is computed using

$$\rho(E_i, E_j) = \exp \left[- \left(\frac{\Delta x}{\lambda_x L} \right)^2 \right] \quad (41)$$

where Δx is the difference between the x coordinates of the midpoints of the elements i and j on the random field discretization and $\lambda_x L$ is the correlation length. λ_x is a dimensionless constant that characterizes the length over which the autocorrelation function decays to a pre specified value.

For each of these correlation functions, two random field discretization models are performed: the midpoint method and the local average method. In the midpoint method, the correlation function is obtained directly from each of the correlation models Equations 40 and 41. In the local average method, the correlation function is obtained as [Spanos and Zeldin, 1997]

$$\rho(x_i, x_j)|_{ave} = \frac{1}{|\Omega_i||\Omega_j|} \int_{\Omega_i} \int_{\Omega_j} \rho(x_i, x_j) dx_i dx_j \quad (42)$$

7.1.3 Beam discretized in four elements

Random field W is discretized into four elements, W_1, W_2, W_3 and W_4 , while parameters E and I are considered as deterministic. A closed form for the vertical deflection at the free end of the beam y_{max} can be obtained in terms of W_1, W_2, W_3 and W_4 , so that limit state equation can be written as

$$g = R - S = EI(y_{adm} - y_{max}) = EI \frac{L}{200} - \frac{L^4}{6144} (15W_1 + 97W_2 + 239W_3 + 417W_4) \quad (43)$$

where the random variables W_1 and W_2 are correlated. Similarly to the previous case, a reliability analysis with correlated variables is performed by rewriting the limit state equation with respect

o a set of uncorrelated reduced normal variables Y and then using *FORM Method 1* [Haldar and Mahadevan, 2000]. Triangular and exponential correlation models are performed, using the midpoint method for the random field discretization.

7.2 Reliability index and error indicator results

Different values are assumed for the parameters θ_x (scale of fluctuation) and $\lambda_x L$ (correlation length) of the triangular and exponential models, respectively. Figure ?? shows the influence of the mesh refinement when the random field is discretized using the midpoint method. The results for the reliability index are plotted for the mesh with two elements and for the mesh with four elements.

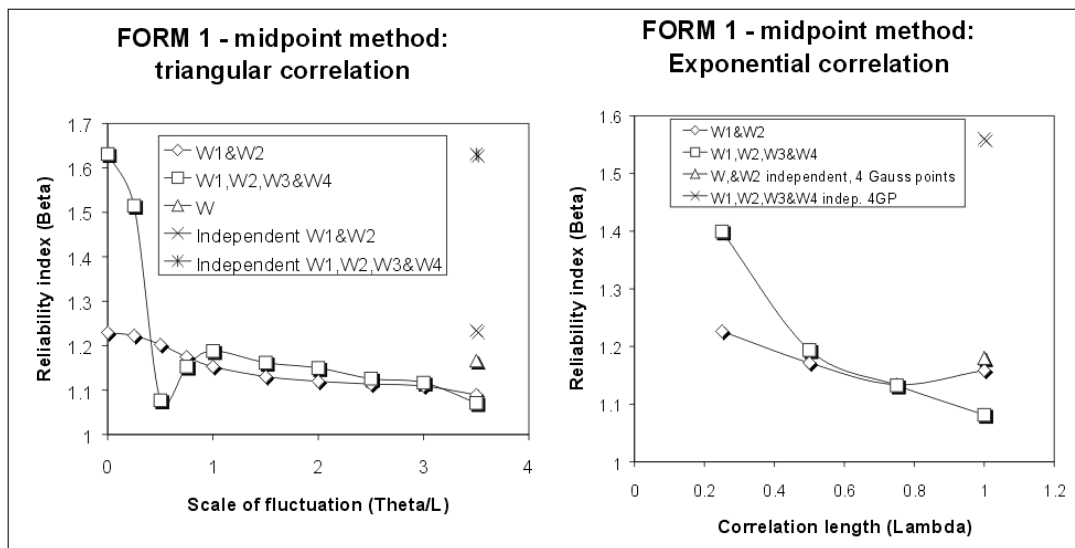


Figure 2: Influence of random field mesh refinement on reliability index for triangular and exponential correlation models (adapted from Jorge, [2004]).

Figure 3 shows the influence of the random field discretization method for a particular mesh. The results for the reliability index are plotted for the mesh with two elements for the midpoint method and for the local average method.

Figure 4 shows the error estimates obtained both from the change in the random field discretization method (for a particular mesh) and from the mesh refinement (for a particular random field discretization method). The differences in the reliability index are obtained starting from the mesh with two elements, with the midpoint method, and either changing the method to local average (first error indicator) or keeping the midpoint method and refining the mesh (second error indicator).

For the numerical example performed, the reliability index results for the error in Fig. 4 showed that both error indicators pointed out to the same range of values, for the parameter in the discretization model, in which bigger errors are expected, or in which the reliability index practically does not change. Values for the correlation model parameter that minimize the differences in the reliability index, for the various discretization models and meshes, are convenient, because, for these values, the model would be more robust and less dependent on external factors, such as the element size of the random field.

These error indicators are qualitatively comparable. although they do not shoe the same magnitude for the error, they show big errors in the same locations. To locate the range of the pa-

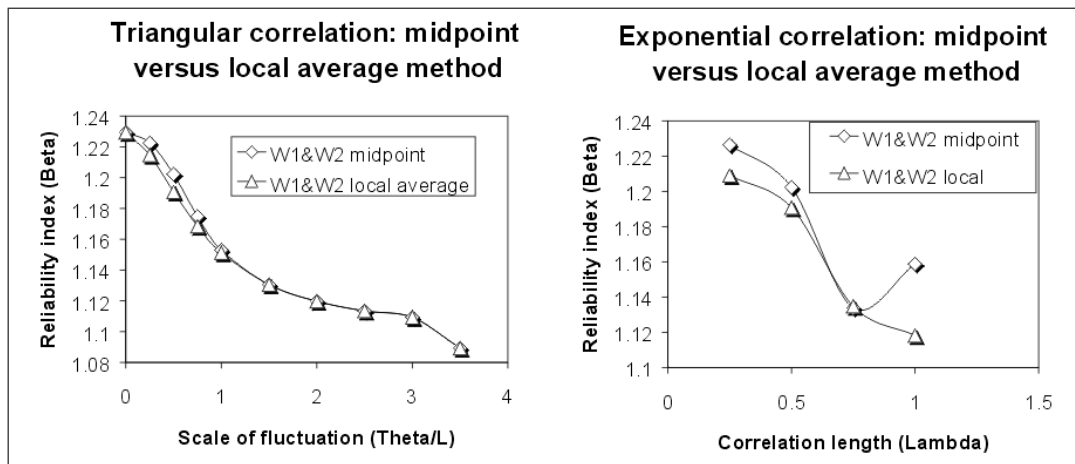


Figure 3: Influence of random field discretization method on reliability index for triangular and exponential correlation models (adapted from Jorge, [2004]).

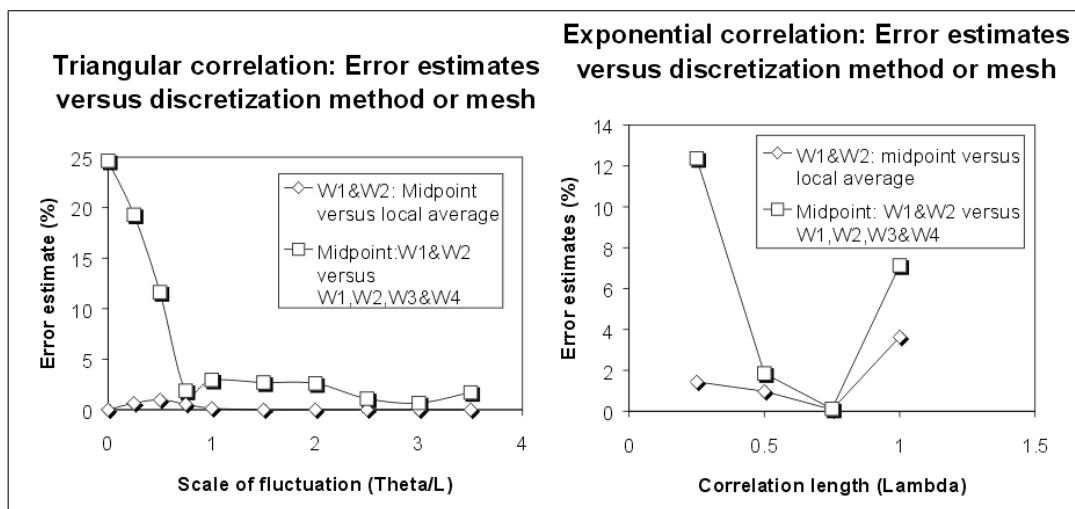


Figure 4: Error estimates (%) obtained by changing the discretization model or by random field mesh refinement (adapted from Jorge, [2004]).

parameter values in which the error is small, both methods are equivalent. In this case, changing the discretization method for a coarse mesh can be an option to its refinement. The change from midpoint method to local average method gives, with less computational effort, the range of values of the parameters of the correlation model in which this model is more robust with respect to the random field mesh refinement. There is no need to make random field mesh refinements *a-priori* to find the appropriate values for the corresponding parameters.

8 Concluding Remarks

This work reviews the existing stochastic finite element formulations, the discretization methods proposed up to now and some aspects already discussed in the literature concerning element sizes and mesh refinement.

Discretization has to be addressed even for the SFEM methods that does not require it explicitly if numerical integration is to be performed.

Upper and lower bounds of the error exist only for certain discretization methods. They were obtained in the context of some specific problems, like static formulation (structures) or reliability analysis, and they may be problem-specific.

Error estimators for the reliability index (β) were proposed in which comparisons were done between two methods of random field discretization (for the same random field mesh) and between two different random field meshes (for the same random field discretization method). It is necessary to first change some characteristic of the discretization of the problem and then compare the solutions for the reliability index for both cases. The comparison between *FORM* and Monte Carlo Simulation *MCS* for correlated variables does not give any other error indicator that could be used.

The reliability index results for the error showed that both error indicators pointed out to the same range of values, for the parameter in the discretization model, in which bigger errors are expected, or in which the reliability index practically does not change. These error indicators pointed out the values for the correlation model parameter that minimize the differences in the reliability index, for the various discretization models and meshes. These are the values of the parameters to be used for the discretization model to be more robust and less dependent on the chosen model or on the element size of the random field.

Also, as the error indicators were comparable, changing the discretization method for a coarse mesh can be an option to its refinement. The change from midpoint method to local average method gives, with less computational effort, the range of values of the parameters of the correlation model in which this model is more robust with respect to the random field mesh refinement. There is no need to make random field mesh refinements *a priori* to find the appropriate values for the corresponding parameters.

A safer design of a structure could include a study on the correlation parameters based on the outlined error indicator approaches, and then the evaluation of the lowest value of the reliability index for the appropriate range of the correlation parameters, which will give the biggest value of the probability of failure of the structure. The approximation of the random field by a system of random variables gives, using this approach, a conservative value for the reliability index.

9 Perspectives for Future Work in Random Field Discretization

Regarding random field discretization, there are some open fields of study, such as:

- A general study of what SFEM formulations are best adapted to a random field discretization;
- A study of what method of random field discretization is well suited to each of the SFEM formulations;
- A study on error bounds, error estimation and adaptivity for the various methods of random field discretization;
- A study on computational efficiency of random field discretization methods and remeshing approaches, including how different random field discretizations may interfere, modify or interact with the sensitivity index.

9.1 Perspectives for Application in Aerospace Structures

There are several potential applications of the random field approach discussed herein. A very good example is the field of aerospace structures, which usually involves demanding design requirements due to the challenging environmental and operational conditions. One interesting problem that has recently gained increased attention from researchers is the development of reliable damage identification procedures for structural health monitoring for aerospace structures. This is an application that involves stochasticity and uncertainties and could certainly benefit from the topics covered in the present chapter.

As an example, consider the problem of detecting possible damaged areas in a metallic honeycomb sandwich plate, such as the one illustrated in Figure 5 (a). One interesting approach is the use of variations in the modal strain energy due to the presence of damage in order to detect and locate regions with possible damage, such as delaminations, manufacturing imperfections due to flawed curing process, etc. [Domingues, 2019a,b]. The developed method relies on a numerical description of the stiffness distribution over the panel and on the modal parameters of the pristine and potentially damaged panels. A finite element model (FEM) is built in order to generate a stiffness matrix required for the calculations. Domingues [2019b] developed detailed 3D FEM models using shell elements, as depicted in Figure 5 (b), and also simplified, 2D models using orthotropic plate elements, as shown in Figure 5 (c). In both cases, numerical models were adjusted using the results obtained from experimental modal analysis. Figure 5 (d) is a representation of the 9 candidate regions for damage investigation. It is worth noting that, for the small 30cm by 30cm specimen used in the investigation, even the simplified, 2D model involved over 50,000 degrees of freedom (DOF), whereas the experimental model was based on 49 DOFs from the measurement points. Therefore calculation of the damage index based on modal strain energy demands a dynamic condensation procedure.

Figure 6 presents the results of a simulated damage identification for the described reduced specimen. Damage was simulated as a 1% reduction of the elastic module in each region. The damage index was calculated for all 9 regions for each case of damage. It can be noticed that the largest calculated index value is clearly well correlated for every damage scenario, even for such a small level of damage severity, successfully detecting and locating the damaged region. However, the implementation was based on a completely deterministic approach. It is expected that the presence of uncertainties from a more realistic application will impact the method's ability to produce reliable results. An interesting possible extension of the method is to represent the strain energy over the panel as a random field, where each region will have its strain energy described by a random variable. The variables may be considered, in a first approximation, as uncorrelated, rendering a diagonal covariance matrix. Different damage scenarios and sources of uncertainties from both the numerical and the experimental data may result in more complex cases where the covariance matrix could produce off-diagonal terms.

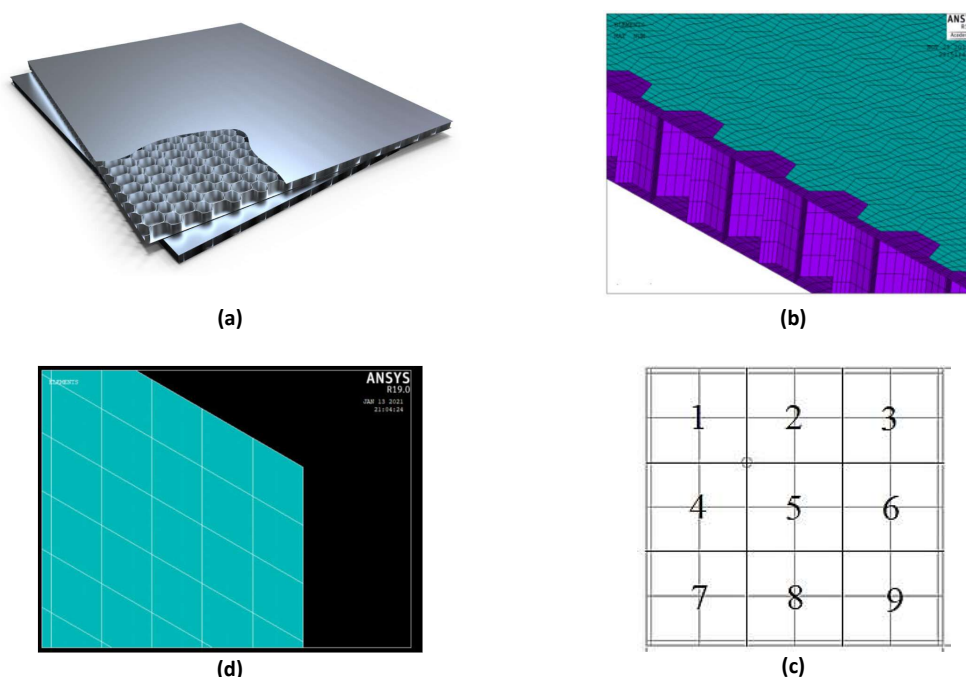


Figure 5: Models for damage detection in aerospace structures (a) Al-Al honeycomb panel for aerospace application; (b) detailed 3D FEM model (c) equivalent orthotropic 2D FEM model (d) Sub-regions for damage interrogation

9.2 Perspectives for Application in Aeronautical Composite Structures

Regarding the subject of random field, there is still a range of application possibilities, mainly in relation to composite materials applied (or not) in the aeronautical industry. Such materials have a high strength/weight ratio, however, they have greater complexity in relation to their manufacture and failure modes.

Figure 7 shows, by way of example, some suggestions and application perspectives related to the theme. Figure 7a and 7b show optimization studies of isogrid-type structures that are widely applied in the aerospace sector. As it is known that boundary conditions can present substantial variability, adopting adequate uncertainty models is imperative. Equally important, Figure 7c shows a structural (multi-objective) optimization study of a composite aircraft wing. Once again, uncertainties must be considered in the mood to obtain an effective and reliable final component. Figure Xe also shows a case of structural optimization already in the field of additive manufacturing of unconventional (auxetic) structures with negative Poisson ratio. Similarly, load variations are also present and must be duly considered.

Figure 7d also provides an example of application in the field of damage identification. This field is quite susceptible to noise (environmental, sensors, etc.). A field of study within Structural Health Monitoring concerned sensor optimization. Figures 7f and 7h show examples of sensor optimization application in aeronautical structures in order to obtain the highest quality of information (modal or static) possible. Since data are noisy in *in situ* works, once again it is necessary to use uncertainty studies for correct decision making.

Optimization under uncertainty is of paramount importance in the structural field. Figure 7g

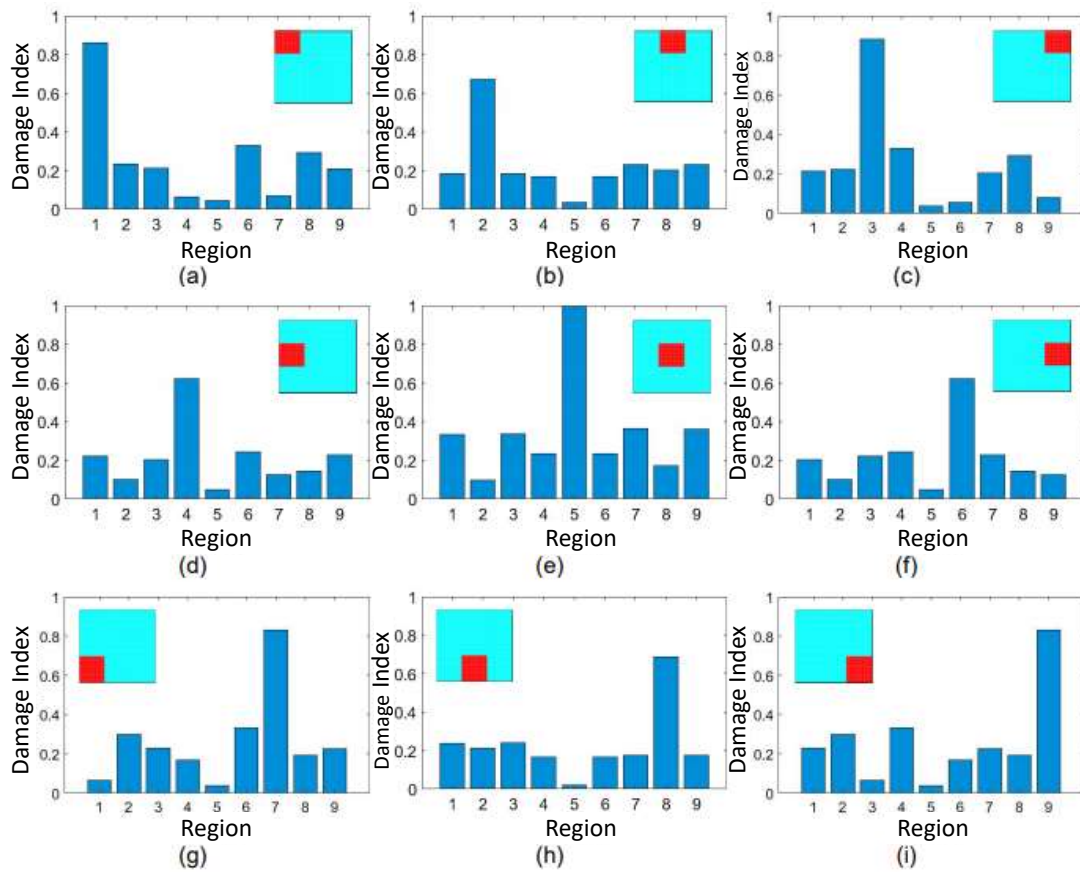
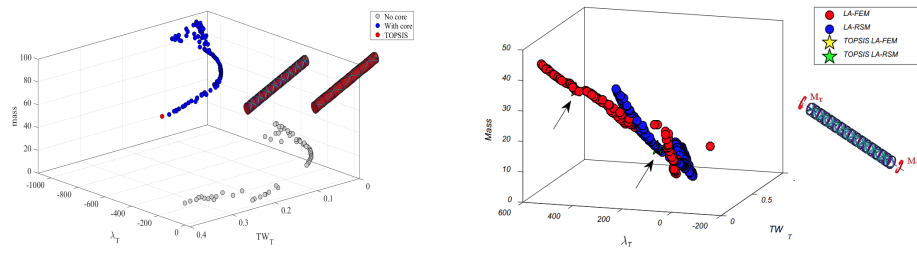


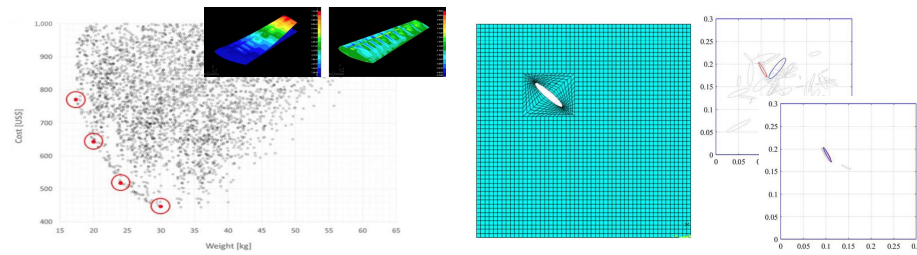
Figure 6: Modal strain energy damage index for honeycomb sandwich panel. Damage in regions (a) 1,(b) 2, (c) 3,(d) 4,(e) 5 ,(f) 6,(g) 7,(h) 8,(i) 9

shows an application perspective in optimizing the manufacture of composites with interrupt layers or simply drop-off ply.

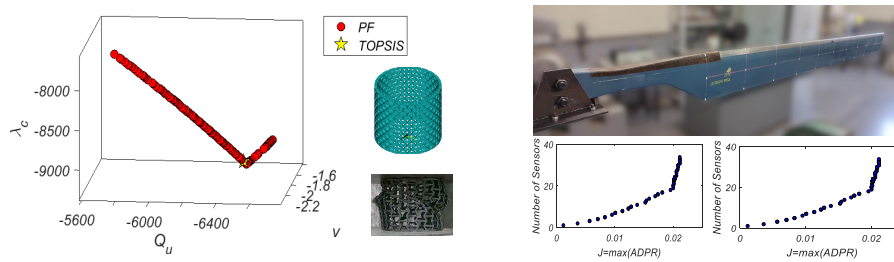
Still on the subject of damage and vibration data, Figures 7i and 7j show two different perspectives of application in methods that make use of signal processing techniques. Signal processing techniques (e.g.: wavelet, short-term Fourier transform, etc.) are vital in order to extract the greatest quantity and/or quality of information hidden in a given signal. The effectiveness of a given technique directly depends on the clarity of a signal. Therefore, properly treating a vibration signal is very important in order to obtain a robust and effective method of detecting, classifying and/or identifying damage.



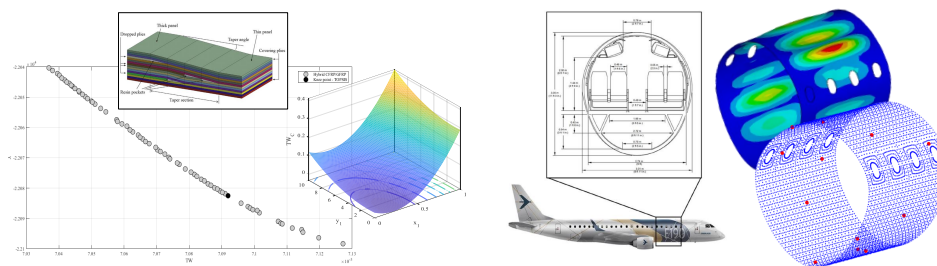
(a) Isogrid optimization (Francisco, [2021]) (b) Isogrid optimization (Pereira, [2022])



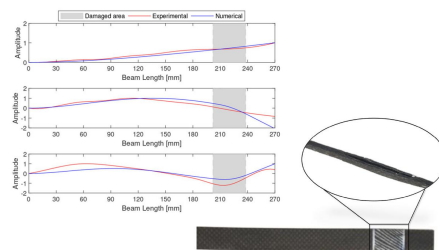
(c) Aircraft wing optimization (Ricco, [2021]) (d) Crack and delamination identification (de Assis, [2021])



(e) Auxetic structure optimization (f) Sensor placement optimization (Pereira, [2022])



(g) Ply drop-off optimization (Diniz, [2022]) (h) Sensor Placement Optimization (Gomes, [2020])



(i) Wavelet Based Damage Identification (Oliver, [2022])

Figure 7: Some examples in aeronautical composite structures.

Acknowledgements

The first author would like to acknowledge the support from the Structural Reliability Research Group at Vanderbilt University, where he was a Post-doctoral Research Associate in 2002 [Structural Reliability Research Group - Faculty and Staff](#), [n.d.].

References

- Ribeiro Junior, Ronny Francis and dos Santos Areias, Isac Antônio and Campos, Mateus Mendes and Teixeira, Carlos Eduardo and da Silva, Luiz Eduardo Borges and Gomes, Guilherme Ferreira. Fault Detection and Diagnosis in Electric Motors Using Convolution Neural Network and Short-Time Fourier Transform, *Journal of Vibration Engineering & Technologies*, pp. 1-12, 2022.
- Oliver, Guilherme Antonio and Pereira, João Luiz Junho and Francisco, Matheus Brendon and Gomes, Guilherme Ferreira. The influence of delamination parameters on the wavelet based damage index in CFRP structures, *Mechanics of Advanced Materials and Structures*, pp. 1-11, 2022.
- Gomes, Guilherme Ferreira and Pereira, João Vitor Purcino. Sensor placement optimization and damage identification in a fuselage structure using inverse modal problem and firefly algorithm, *Evolutionary Intelligence*, 13, 4, pp. 571-591, 2020.
- de Assis, Felipe Mouallem and Gomes, Guilherme Ferreira. crack identification in laminated composites based on modal responses using metaheuristics, artificial neural networks and response surface method: a comparative study, *Archive of Applied Mechanics*, 91, 10, pp. 4389-4408, 2021.
- Diniz, Camila Aparecida and Pereira, João Luiz Junho and da Cunha, Sebastião Simões and Gomes, Guilherme Ferreira. Drop-off Location Optimization in Hybrid CFRP/GFRP Composite Tubes Using Design of Experiments and SunFlower Optimization Algorithm, *Applied Composite Materials*, pp. 1-30, 2022.
- Pereira, João Luiz Junho and Francisco, Matheus Brendon and Ribeiro, Ronny Francis and Cunha, Sebastião Simões and Gomes, Guilherme Ferreira. Deep multiobjective design optimization of CFRP isogrid tubes using lichtenberg algorithm, *Soft Computing*, pp. 1-15, 2022.
- Pereira, João Luiz Junho and Francisco, Matheus Brendon and de Oliveira, Lucas Antônio and Chaves, João Artur Souza and Cunha Jr, Sebastião Simões and Gomes, Guilherme Ferreira. Multi-objective sensor placement optimization of helicopter rotor blade based on Feature Selection, *Mechanical Systems and Signal Processing*, 180, 2022.
- Francisco, Matheus Brendon and Pereira, João Luiz Junho and Oliver, Guilherme Antonio and da Silva, Fernando Helton Sanches and da Cunha Jr, Sebastião Simões and Gomes, Guilherme Ferreira. Multiobjective design optimization of CFRP isogrid tubes using sunflower optimization based on metamodel, *Computers & Structures*, 249, 2021.
- Ricco, Juan Tamassia and Coimbra, Rogerio Frauendorf Faria and Gomes, Guilherme Ferreira. DMultiobjective optimization of the LASER aircraft wing's composite structural design, *Aircraft Engineering and Aerospace Technology*, 2021.

- A.B. Jorge: Discretization of random fields in reliability structural analysis: error indicator methodology, Proc. VI Simpósio Mineiro de Mecânica Computacional, Itajubá, MG - Brazil, 2004
- Vanderbilt University: Structural Reliability Research Group - Faculty and Staff, url = http://www.structural-reliability.vanderbilt.edu/faculty_staff.htm, urldate = 2022-06-29
- G. I. Schuëller: Structural reliability – recent advances, Freudenthal Lecture, 7th Int. Conf. on Structural Safety and Reliability (ICOSSAR '97), Kyoto, Japan, 1997.
- T. Hisada and S. Nakagiri: Role of the stochastic finite element method in structural safety and reliability, Proc. 4th Int. Conf. on Structural Safety and Reliability (ICOSSAR '85), Kobe, Japan, Vol.1, pp. 385-394, 1985.
- T. Hisada and S. Nakagiri: Stochastic finite element method developed for structural safety and reliability, Proc. 3rd Int. Conf. on Structural Safety and Reliability (ICOSSAR '81), Trondheim, Norway, pp. 395-408, 1981.
- W. K. Liu, T. Belytschko and A. Mani: Random field finite elements, Int. J. Numer. Meth. Eng. , 23, pp. 1831-1845, 1986.
- E. H. Vanmarcke and M. Grigoriu: Stochastic finite element analysis of simple beams, J. Eng. Mech. Div. , ASCE, Vol. 109, 5, pp. 1203-1214, 1983.
- T. Takada: Weighted integral method in stochastic finite element analysis, Prob. Eng. Mech. , Vol. 5, 3, pp. 146-156, 1990.
- G. Deodatis: Weighted integral method. I: Stochastic stiffness matrix, J. Eng. Mech. Div. , ASCE, Vol. 117, 8, pp. 1851-1864, 1991.
- M. Lawrence: A basis random variable approach to stochastic structural analysis, PhD Thesis, University of Illinois, Urbana, Illinois, 1986
- R. Ghanem and P. Spanos: Stochastic finite elements: a spectral approach, Springer Verlag, Berlin, 1991.
- P. D. Spanos and R. Ghanem: Stochastic finite element expansion for random media, Research Report No. NCEER-88-0005, March 1988.
- M. Shinozuka and G. Deodatis: Response variability of stochastic finite element systems, J. Eng. Mech., Vol. 114, 3, pp. 499-519, 1988.
- A. Haldar: Reliability evaluation of structures using nonlinear SFEM, in: Uncertainty modeling in finite element, fatigue and stability of systems, Series on Stability, Vibration and control of Systems, Series B, Vol. 9, pp. 23-50, World Scientific, 1997.
- S. Mahadevan and A. Haldar: Practical random field discretization in stochastic finite element analysis, Structural Safety , 9, pp. 283-304, 1991.
- S. Mahadevan: Probabilistic finite element analysis of large structural systems, in: Uncertainty modeling in finite element, fatigue and stability of systems, Series on Stability, Vibration and control of Systems, Series B, Vol. 9, pp. 1-21, World Scientific, 1997.

- H. G. Matthies, C. E. Brenner, C. G. Bucher and C. G. Soares: Uncertainties in probabilistic numerical analysis of structures and solids - Stochastic finite elements, *Structural Safety*, Vol. 19, 3, pp. 283-336, 1997.
- P. L. Liu: Size effect of random field elements on finite element reliability methods, Proc. 3rd IFIP WG 7.5 Conference, in *Reliability and Optimization of Structural Systems '90*, pp. 223-239, 1990.
- A. Der Kiureghian and J. -B. Ke: The stochastic finite element method in structural reliability, *Probabilistic Engineering Mechanics*, 3(2), pp. 83-91, 1988.
- P. Spanos and B. Zeldin: Discretization aspects of stochastic finite elements, Proc. 7th Int. Conf. on Structural Safety and Reliability (ICOSSAR '97), 9, pp. 835-874, Kyoto, Japan, 1997.
- C. E. Brenner: A stochastic method for nonlinear dynamic problems, *ZAMM*, 73(7/8), T900-T902, 1993.
- A. Der Kiureghian, C. -C. Li and Y. Zhang: Recent developments in stochastic finite elements, In *Lecture Notes in Engineering*, IFIP 76, Proc. 4th IFIP WG 7.5 Conf., Munich, Springer, pp. 19-38, 1991.
- O. Ditlevesen: Discretization of random fields in beam theory, 7th Int. Conf. Applications of Statistics and Probability in Civil Engineering, 1995.
- M. A. Lawrence: Basis random variables in finite element analysis, *International Journal for Numerical Methods in Engineering*, 24, pp. 1849-1863, 1987.
- R. Ghanem and J. Red-Horse: Propagation of probabilistic uncertainty in complex physical systems using a stochastic finite element approach, *Physica D*, 133, pp. 137-144, 1999.
- R. Ghanem: Ingredients for a general purpose stochastic finite elements implementation, *Comput. Methods Appl. Mech. Engrg.*, 168, pp. 19-34, 1999.
- W. K. Liu, T. Belytschko and A. Mani: Probabilistic finite elements for nonlinear structural dynamics, *Comput. Methods Appl. Mech. Engrg.*, 56, pp. 61-81, 1986.
- A. Haldar and S. Mahadevan: *Probability, Reliability and Statistical Methods in Engineering Design*, Wiley, 2000
- Domingues, A. C., and Anflor, C. T. M. and Carneiro, S. H. S.: Damage Identification in Light Panels with Orthotropic Properties Using Modal Strain Energy, *Proceedings of the Congreso Iberoamericano de Ingeniería Mecánica*, 2019.
- Domingues, A. C., and Anflor, C. T. M. and Carneiro, S. H. S.: Damage identification in light aerospace structures using modal strain energy method. (In Portuguese: Identificação de dano em estruturas aeroespaciais leves utilizando o método da energia de deformação modal, Master's Thesis, Integrity of Engineering Materials Program, University of Brasilia, 2019

Chapter 19

On Error Estimators and Bayesian Approaches in Computational Model Validation

Chapter details

Chapter DOI:

<https://doi.org/10.4322/978-65-86503-88-3.c19>

Chapter suggested citation / reference style:

Jorge, Ariosto B., et al. (2022). “On Error Estimators and Bayesian Approaches in Computational Model Validation”. In Jorge, Ariosto B., et al. (Eds.) *Uncertainty Modeling: Fundamental Concepts and Models*, Vol. III, UnB, Brasilia, DF, Brazil, pp. 640–743. Book series in Discrete Models, Inverse Methods, & Uncertainty Modeling in Structural Integrity.

P.S.: DOI may be included at the end of citation, for completeness.

Book details

Book: Uncertainty Modeling: Fundamental Concepts and Models

Edited by: Jorge, Ariosto B., Anflor, Carla T. M., Gomes, Guilherme F., & Carneiro, Sergio H. S.

Volume III of Book Series in:

Discrete Models, Inverse Methods, & Uncertainty Modeling in Structural Integrity

Published by: UnB City: Brasilia, DF, Brazil Year: 2022

DOI: <https://doi.org/10.4322/978-65-86503-88-3>

On Error Estimators and Bayesian Approaches in Computational Model Validation

Ariosto B. Jorge^{1a*}, Sergio H. S. Carneiro^{1b}, Jhon N. V. Goulart^{1c},
Marcus V. G. Morais^{1d}, Guilherme F. Gomes², Carla T. M. Anflor^{1e},
Jalusa M. S. Ferrari^{1f}, and José L. C. R. Vila^{1g}

¹Post-Graduate Program - Integrity of Engineering Materials, University of Brasilia, Brazil. E-mail: a) ariosto.b.jorge@gmail.com; b) shscarneiro@unb.br; c) jvaz@unb.br; d) mvmorais@unb.br; e) anflor@unb.br; f) jalusaferrari@gmail.com; g) jl.vila@hotmail.com

²Mechanical Engineering Institute, Federal University of Itajubá, Brazil. E-mail: guilhermefergom@unifei.edu.br

*Corresponding author

Abstract

This chapter presents an overview of computational models and formulations, a description of some error estimators in Boundary Element Methods (BEM) and in Finite Element Methods (FEM), and some perspectives of the use of these error estimators in Bayesian approaches for Model Validation.

Keywords: Error Estimators, Finite Element Methods, Boundary Element Methods, Bayesian Approaches, Model Validation.

1 Introduction

Modeling and simulation is an important field in engineering and has been the object of relevant research and publications. To illustrate, a non-exhaustive list and overview of related research work in modeling, simulation, applications, etc. in this area may include, among others, (Chong et al., 2004) (Oberkampf & Trucano, 2002) (Oberkampf et al., 2004) (Oberkampf & Trucano, 2006) (Oberkampf & Trucano, 2008) (Oberkampf, 2009)(R. G. Hills & Trucano, 1999) (R. G. U. Y. Hills & Trucano, 2002) (Easterling, 1998) (Sokolowski & Banks, 2010) (Kerns et al., 1992) (Oberkampf & Barone, 2006) (Roy & Oberkampf, 2010) (Oberkampf & Roy, 2010) (Swiler et al., 2017) (Inman et al., 2005).

This chapter presents an overview of computational models and formulations, a description of some error estimators in Boundary Element Methods (BEM) and in Finite Element Methods (FEM), and some perspectives of the use of these error estimators in Bayesian approaches for Model Validation.

Section 2 presents an overview of modeling of physical systems.

Section 3 discusses some error estimators in computational / numerical methods, with focus in error estimators for Finite Element Methods (FEM) and for Boundary Element Methods (BEM).

Section 4 discusses some aspects on model validation under uncertainty, with focus on a consistent error estimation strategy for model validation and on combining error estimators and Bayesian hypothesis testing in a model validation procedure.

Section 5 discusses some perspectives for the use of Validation Techniques in FEM Condensation Models, and also some perspectives for aeronautical structures applications, such as modeling of structural damage detection/localization problems and modeling of damage detection/identification problems in structural fasteners / joints.

Section 6 presents some concluding remarks.

2 Modeling of Physical Systems: Overview

This overview of the different mathematical models of physical systems and numerical formulations used in engineering problems follows (Basu, 2002) (Basu et al., 2003), (Basu, 2007), (Basu & Jorge, 2009). Current state of affairs and future of modeling and simulation in engineering design is discussed. Various modeling schemes and all the associated aspects like multiscale modeling, risk and reliability, multi-disciplinary modeling and optimization, and the role of digital computers are also discussed.

2.1 Introduction

Engineers and scientists routinely use mathematical models for physical systems to simulate the real behavior. Such a model embodies the mathematical interpretation of a phenomenon based on well-defined physical laws of nature and is based on some conceptual description of the physical problem. Such models are normally ordinary or partial differential equations with spatial (in one-, two-, or three-dimensions) and/or temporal dependence. For instance, the mathematical model of a bar under distributed axial traction is based on Hooke's law and the concept of equilibrium of forces acting on the body. For instance in the case of the bar shown in Fig. 1, the one-dimensional mathematical model can be shown to be

$$-\frac{d}{dx}\left(AE\frac{du}{dx}\right) + Ac\frac{\partial u}{\partial t} + A\rho\frac{\partial^2 u}{\partial t^2} = b \quad (1)$$

Here: A = cross-sectional area of the bar, which may be function of x

E = Young's modulus of bar material

(Hooke's Law: $\sigma = E\varepsilon$, where σ = stress, and ε = strain)

ρ = mass density per unit volume

c = viscous damping coefficient per unit volume

u = displacement of a point in the bar in the positive x -direction

b = time and space dependent distributed axial force in the bar

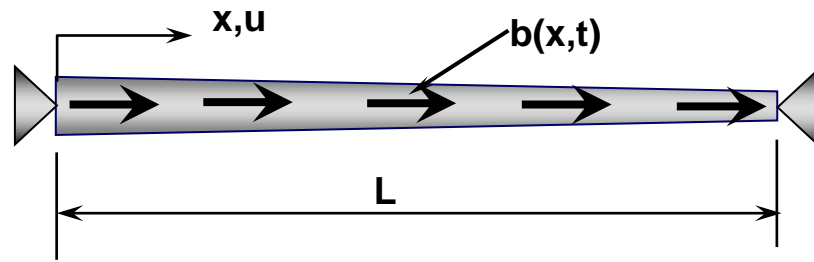


Figure 1: axial bar problem - adapted from (Basu & Jorge, 2009)

By applying the principle of conservation of energy and Fourier law, namely, $Q_i = -k_i \frac{\partial T}{\partial x_i}$, the mathematical model for a two-dimensional heat transfer problem, shown in Fig. 2, is:

$$\frac{\partial}{\partial x_i} \left(k_i \frac{\partial T}{\partial x_i} \right) - \rho c \frac{\partial T}{\partial t} = f \quad \text{in } \Omega, i=1,2 \tag{2}$$

With the boundary conditions;

$$T = T_g \text{ on } \Gamma_g$$

$$k_i \frac{\partial T}{\partial x_i} n_i = q_n \text{ on } \Gamma_n$$

$$k_i \frac{\partial T}{\partial x_i} n_i = \beta(T_\infty - T) \text{ on } \Gamma_\beta$$

$$k_i \frac{\partial T}{\partial x_i} n_i = \alpha q_r - \kappa \epsilon T^4 \text{ on } \Gamma_r$$

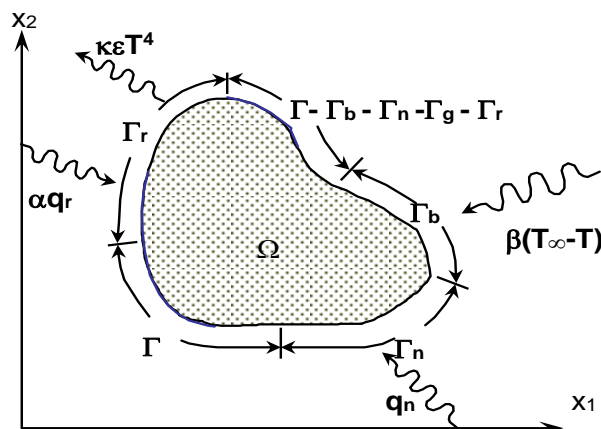


Figure 2: Two-dimensional heat transfer - adapted from (Basu & Jorge, 2009)

In Eq. (2):

- T_g = the specified temperature on Γ_g
- q_n = specified flow of heat per unit area on Γ_n
- β = coefficient of convective heat transfer at Γ_β
- T_∞ = ambient temperature of surrounding medium
- α = coefficient of surface absorpivity at Γ_r
- κ = Steffan - Boltzman constant
- ε = coefficient of surface emissivity
- q_r = radiant heat flow per unit area
- k_x, k_y = coefficients of conductivity in x - and y - directions
- f = heat input (or generation) per unit volume

It is only in the simplest (or, trivial) cases a closed form analytical solution of such mathematical models is possible. In most real life situations, a theoretical solution of such models becomes overly complex, well-nigh impossible. Complexities may be introduced by geometry, material characteristics, external influences, and other physical attributes. This calls for the use of discrete numerical methods, which essentially discretize the problem domain into small enough sub-regions on which the mathematical model is exercised using local attributes in space and time.

Additional complexities affecting the model and the solution are the presence of multiphysics effects – namely, the interaction of electromagnetic, thermal and mechanical effects. At the present time, there is an increased emphasis to depart from traditionally deterministic treatment of problems and use stochastic approach instead, accounting for the random nature of the problem parameters. The above possibilities are summarized in the block chart shown in Fig. 3.

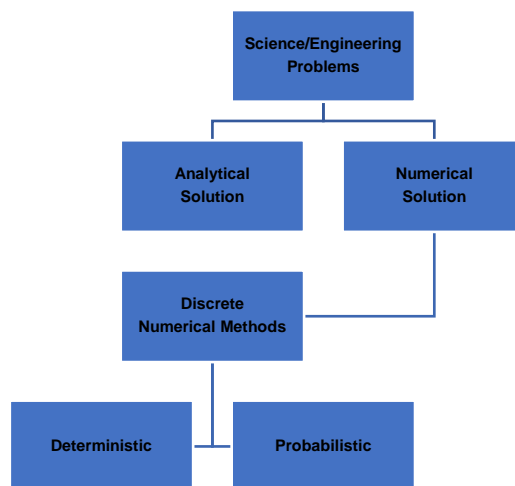


Figure 3: Solution of science and engineering problems - adapted from (Basu & Jorge, 2009)

To successfully compete in this global economy, it is necessary to create safe and reliable products which are economic too, by using optimization methods in the design process.

Safety evaluation requires consideration of performance under service conditions as well as at limit states, when the behavior becomes highly nonlinear.

2.1.1 Continuum vs. Multi-scale Modeling

Traditionally, a mathematical model for modeling and simulation treats the problem domain as a continuum (or, uses a macro model) both under service conditions and at limit states. As the limit state is approached incrementally, the degradation in the continuum is reflected approximately in the problem parameters in the average sense by the smearing process. In reality, with the progress of the degradation process the response traverses through meso-, micro-, and nano-scale resolutions, as seen in Fig. 4.

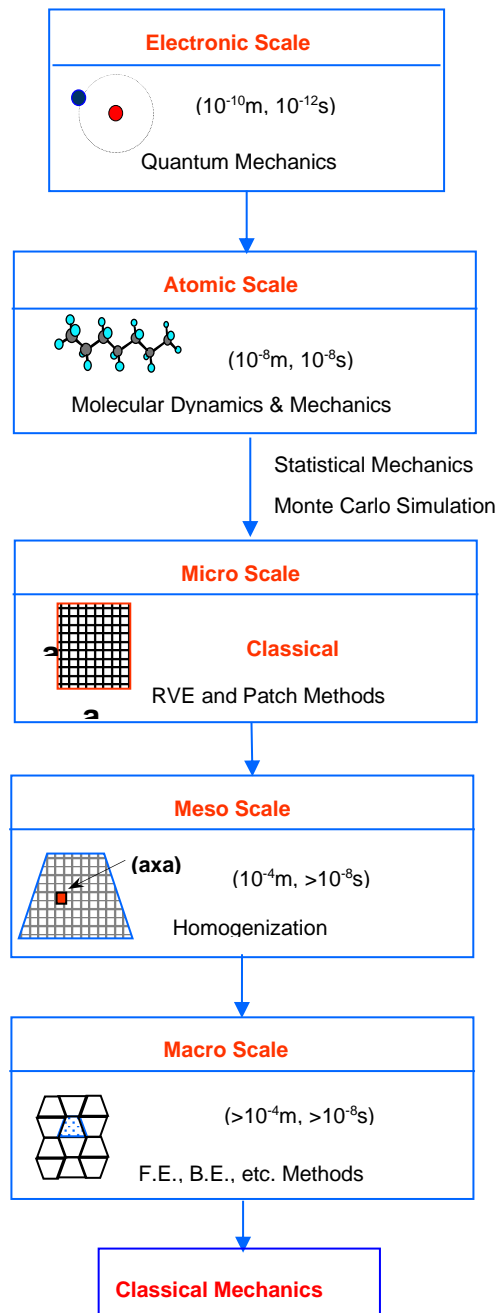


Figure 4: Multi-scale Modeling - adapted from (Basu & Jorge, 2009)

In other words, the material behavior is actually controlled by processes on wide range of length and time scales. At least in the critical (say, damaged) regions, this calls for multiscale modeling of the problem to capture the effect of highly localized behavior on problem parameters for use in the macro scale. Micro-scale modeling requires accounting for the local behavior of the material in the presence of damage and then coming up with homogenized problem parameters for use in the continuum model. For instance, to accurately simulate the deterioration process in complex materials, multiscale modeling is very important. Also, in the case of problem of contaminant transport in groundwater, it is necessary to use multiscale modeling to capture the effect of highly inhomogeneous nature of the substrata.

The need for using very high resolution (or, nano-scale) will be determined by the nature of the problem. Typical examples of such need are in the case of weathering resistance of cement mortar reinforced with carbon nanotubes, corrosion resistance of nano-crystalline coating on alloys, or material behavior near the interface of two bonded dissimilar materials. In nano-scale modeling, continuum physics is no more applicable and one needs to resort to molecular dynamics calculations considering the motion of atoms along with the application of Newtonian Mechanics to atomic systems. The results of molecular dynamics simulation can be used to find material properties at the continuum level by the application of statistical mechanics via Monte Carlo simulations. In most cases, however, damage characterization at the micro-scale leads to the desired level of accuracy in the predictions at the macro- or continuum scale.

Apart from multiple spatial scales, many problems in nature involve multiple temporal scales. For instance, if we examine the temperature variations experienced by a bridge structure, we can notice hourly variations, daily variations, seasonal variations, and annual variations. Another example is that of chemical reactions that may take seconds or hours. On the other hand, the vibrations of chemical bonds may occur at a time scale of femtoseconds (10⁻¹⁵s). In the case of most of such problems, the problems are considered at the scale of interest, ignoring the influence of other scales. It is evident from Fig. 4 that the smaller the physical scale smaller is the associated time scale.

2.1.2 Computational Tools

Far reaching changes in modeling and simulation techniques have been fueled by spectacular advances in digital computational technology, which has radically changed the way we do engineering. In developing new and more powerful techniques for modeling and simulation, the analysts and designers have always played the catch-up game with advances in digital technology. Meaningful molecular dynamics simulations would not have been possible without the power of today's massively parallel machines. Moore's law shows that computer speed (FLOPS) and memory increase approximately by a factor of 2 per 1.5 years. Consequently, the boundary between the power of personal computer and workstation and between network of workstations and powerful supercomputer has become blurred. As of June 2009, the No. 1 position as the fastest computer in the world with 129,600 processors belonged to IBM's Roadrunner System located at DOE's Los Alamos National Laboratory, California. It has reached the speed of 1.059 petaflops/s (quadrillions of floating point calculations per second) and continues to be the only system ever to break the petaflop/s. Linpack barrier. Many of these high powered systems consist of PC clusters. As PC is getting more powerful so are the new clusters. New PCs with Intel's Core i7 or AMD's Phenom II processor with speed

exceeding 3GHz are proving to be significantly more powerful than the high-powered workstations of just a few years before.

2.2 Discrete Numerical Methods

Discrete numerical methods are being almost exclusively used in the numerical solution of mathematical models of the continua expressed as initial and/or boundary value problems of differential calculus. The associated partial differential equations are formulated in space, either in the Lagrangian reference frame, or Eulerian reference frame.

In the first case, the selected reference points change with the continuum and hence more suited for solids; whereas, in the latter case, the selected reference points in the continua remain stationary and are more suited for fluids. The finite difference and Ritz type methods of the pre-computer era have largely been replaced in the computer era by finite element, boundary element, mesh-free, finite volume, and other evolving methods.

The first two of these methods were available to the general users in the seventies of the last century; whereas, the mesh-free methods became known during the early nineties. Finite volume method is specially used in fluid dynamics simulation or forms the basis of, what is termed as, hydro code. Wavelet and a host of other less known methods have appeared in the scene more recently.

As a matter of fact, all the new developments, in one way or other, involved modifying the first three mentioned methods. In all these methods the problem is reduced into one of finding numerical values at a discrete number of points covering the problem domain.

An essential difference between the boundary element method and other methods is that in this method the discrete points need to be selected at the boundary of the domain only; whereas, in the remaining methods the points are selected in the volume of the domain. For instance, in the case of three-dimensional problems, the boundary is a surface and hence the sub-regions selected on this bounding surface are two-dimensional. This fact introduces the possibility of higher computational efficiency in the boundary element method, especially, when the volume of the problem domain becomes significantly large compared to the area of the enclosing surface. In this method, an additional requirement is to append the fundamental solution of the problem, which is essentially the homogeneous solution of the differential equation due to a source term.

In the case of mesh-free method, the approximating function is not confined within a sub-region. For each integration point, it is defined over a predefined region of influence for that point. A more recent method, the wavelet method can be viewed as a method in which the approximating function is defined using multi-resolution technique based on the use of wavelets, similar to those used in signal processing.

In general, discrete numerical methods involve global discretization in space as well as time, coupled with the spatial approximation of the unknown function in terms of its values at discrete points within a sub-region, or in the neighborhood of a point of interest, and piecewise temporal approximation of the unknown function in terms of values at discrete instants of time in a temporal window.

As the spatial variation of the approximating function is represented by polynomial functions, difficulties associated with all the methods in modeling singular behavior is avoided in the boundary element method by introducing it through the fundamental

solution. In all other cases, the use of special functions near the singular point becomes necessary.

The ability to predict solution quality through a posteriori analysis and adaptive mesh refinement has assumed great importance in recent times and significant efforts are being put to advance in this area.

2.2.1 Finite Element Method (FEM)

After discretization of the problem domain into finite elements (Fig. 5) and appropriately defining the piecewise polynomial approximation over each finite element, a set of linear or nonlinear simultaneous equations is arrived at by applying the weighted residual method, if differential equations are used or, by applying the stationary condition to the functional for the problem, where applicable.

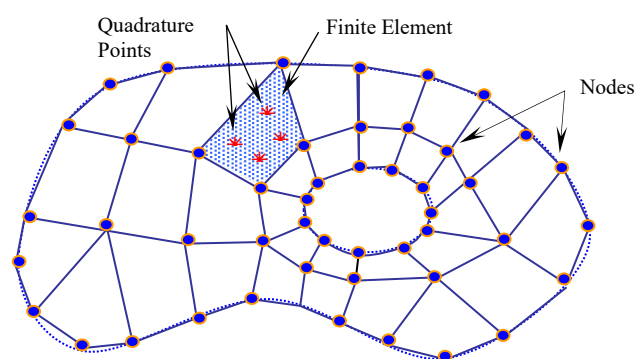


Figure 5: Multi-scale Modeling - adapted from (Basu & Jorge, 2009)

Sometimes, other approaches, like least squares method, are found to be more appropriate. The quality of the resulting approximate solution is controlled by the closeness of the finite element model to that represented by the mathematical model.

Solution can be improved by selective spatial refinement in the case of elliptic problems and local refinement in the case of hyperbolic problems, as the present may not affect the future.

In elements bordering a singular point, the polynomial basis can be appended by singular terms as shown below (Basu & Shah, 2000). In recent times, some investigators are terming it as x-FEM (Chessa & Belytschko, 2004). This is also conceptually similar to the, so-called, generalized finite element (Ivo Babuška et al., 2003) in which the choice of type of shape function is controlled by the nature of the problem to be solved.

Quality of the solution obtained can be assessed in a posteriori sense on the basis of the extent of violation of natural boundary conditions, and continuity requirements of functional derivatives at inter-element boundaries.

A number of schemes have been proposed to estimate the solution error in different norms, as discussed in subsequent sections.

Discrepancies between raw and improved estimates are accounted for through model refinement, with strategies such as h -, p -, hp -, etc, refinements.

2.2.2 Boundary Element Method (BEM)

In the case of boundary-element method (BEM), the governing partial differential equations are transformed to a set of integral equations which are numerically solved for the unknown boundary variables by discretizing the boundary of the element, as shown in Fig. 6.

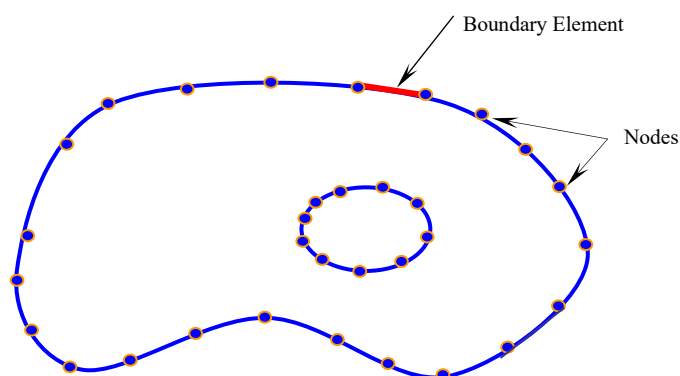


Figure 6: Boundary element method - adapted from (Basu & Jorge, 2009)

Apart from fewer unknown variables to be handled during the global solution, BEM handles singular problems, like those in linear fracture mechanics calculations, quite efficiently giving accurate point wise results close to the singular points; whereas, FEM gives damped global solution valid in the sense of energy norm.

Also, BEM is adept in handling problems of half-plane, or far-field problems. FEM is, however, better suited to handle complex domains and the presence of material nonlinearities.

In BEM, the system of equations are arrived at through the discretization of an integral identity, like Green's second identity in potential theory, or Somigliana's stress identity in elasticity theory, which is an exact representation of the original boundary-value problem. The system of equations in BEM arises from the discretization of integral identities, which are an exact representation of the original boundary-value problem.

A number of different BEM formulations have been proposed, for potential, elastostatics, fracture mechanics, acoustics, etc. problems, in 2D and in 3D. One can cite, for example, self-regular BEM formulation for potential problems (Jorge, Ribeiro, Cruse, et al., 2001), near-crack contour behaviour and extraction of log-singular stress terms of the self-regular traction boundary integral equation (Jorge, Cruse, & Fisher, 2003), variational self-regular traction-BEM formulation (Jorge, Cruse, Fisher, et al., 2003), variational self-regular flux-BEM formulation (Porto et al., 2005), non-singular BEM algorithms for potential problems (Ribeiro et al., 2009).

2.2.3 Mesh-Free Methods

In real-life large scale computations, generation, adaptation, and upgrading of finite element mesh is the costliest part of computational effort, especially, in situations involving crack propagation, penetration, large deformations causing significant mesh distortions, complex geometries, advanced materials, material degradation leading to discontinuities, etc.

In such situations, a group of methods called mesh-free methods may be the answer. In such methods, instead of associating the nodal points, to finite or boundary elements, the problem domain is covered by a network of nodal points, or, reference points of unknowns of the problem, with no fixed connectivity between them, as shown in Fig. 7.

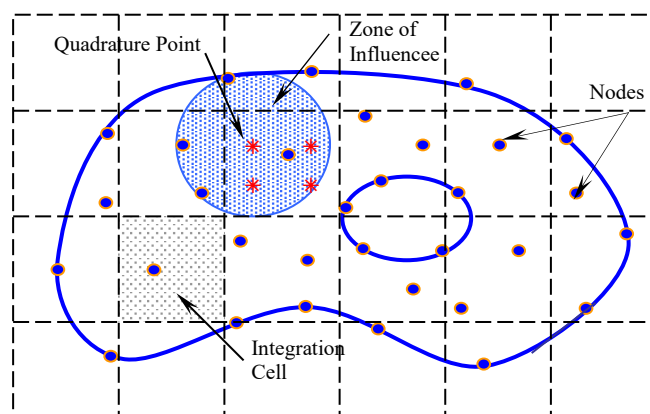


Figure 7: Element-Free Galerkin Method - adapted from (Basu & Jorge, 2009)

(Duarte, 1995) gives a comprehensive review of this method. The method can be applied either as a volume integral scheme, or as a boundary integral scheme. In both cases, the weak form of the governing differential equation can be satisfied at selected sampling or quadrature points based on weighted nodal values within a predefined sub-domain of influence assigned to the sampling point or, alternatively, the collocation method can be used.

The main idea of point collocation approach is to satisfy the governing differential equation at each of the nodal points covering the problem domain. At boundary points, as needed, the Dirichlet or Neumann boundary conditions are satisfied instead.

Depending upon the way the functional approximation is defined for the point of interest, a number of methods can be identified. According to the way this functional approximation is defined one may follow one of the three schemes – (a) Shepard's scheme (Shepard, 1968) which has a consistency of first order only, (b) Moving Least Squares Scheme (Lancaster & Salkauskas, 1981) which generalizes the Shepard's scheme, and (c) Partition of Unity p-version scheme (I. Babuška & Melenk, 1997), which generalizes Shepard's approach to higher consistency orders. A few of the methods based on these are:

- (a) Diffuse Element Method (DEM) (Nayroles et al., 1992);
- (b) Element-Free Galerkin Method (EFGM) (Belytschko et al., 1994);
- (c) Natural Element Method (NEM) (Sukumar et al., 1998);
- (d) Material Point Method (MPM) (Sulsky et al., 1995);
- (e) hp-Clouds Method (HPCM) (Armando Duarte & Tinsley Oden, 1996);
- (f) Finite Cloud Method (FCM) (Aluru & Li, 2001);
- (g) Boundary Node Method (BNM) (Xie Mukherjee & Mukherjee, 1997);
- (h) Boundary Cloud Method (BCM) (G. Li & Aluru, 2002), etc.

In the moving least squares method (MLS), the best interpolation in terms of nodal values is defined in a (moving) weighted least-squares sense. Diffuse element method uses the MLS approximation in the context of Galerkin method (Nayroles et al., 1992).

The EFG method is an extension of DEM (Belytschko et al., 1994). The primary differences between the two are that the EFG method (Fig. 7) includes certain extra terms in the derivatives of the approximating function that are not present in DEM. In addition, the EFG method uses Lagrange multipliers to enforce the essential boundary conditions.

In the Natural Element Method, natural neighbor interpolants are used to construct the trial and test functions. Such a multivariate data interpolation scheme is primarily used in data interpolation and modeling of geophysical phenomena (Sukumar et al., 1998).

In the Material Point Method (MPM) a solid continua is defined in terms of a collection of material points (Sulsky et al., 1995). Each material point is contained in a sub-domain assigned to it and carries information like mass, volume, velocity, strain, etc. of its sub-domain. Unlike the regular mesh-free method, MPM has a computational mesh of sub-domains associated with it. In this method, the material points do not interact with each other directly, rather the material point information is interpolated on to the mesh as the equation of motion is allowed to march in the time-domain. At any instant of time, the information on the new position of the mesh can be translated back onto the material points.

The hp-clouds method (Armando Duarte & Tinsley Oden, 1996) uses partition of unity scheme to construct radial basis functions of varying size of supports as circular patches (Fig. 8) or spherical volumes attached to each point and with polynomial reproducing properties of arbitrary order. The resulting basis functions are used in a Galerkin or collocation sense setting up a linear system of equations. The union of patches or volumes forms an open covering of the problem domain.

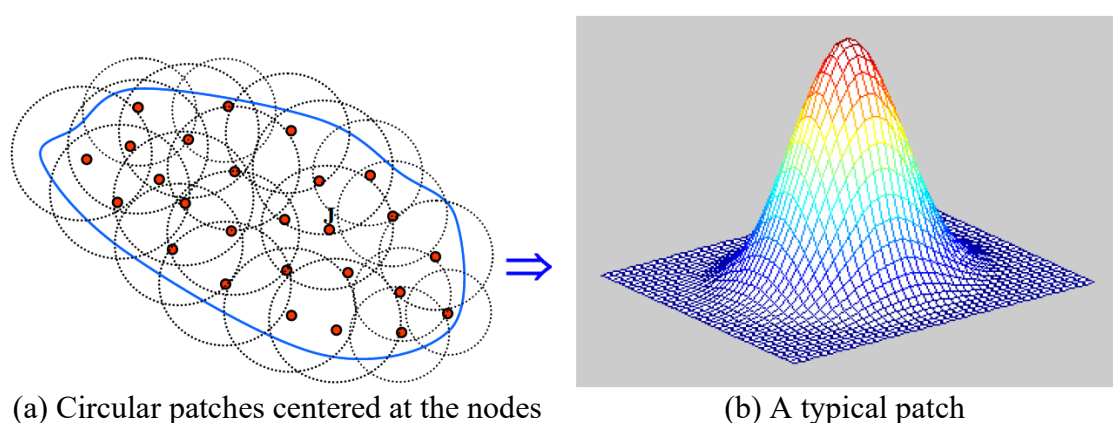


Figure 8: hp-clouds method - adapted from (Basu, 2007)

The finite cloud method (Aluru & Li, 2001) is a point collocation method based on reproducing kernel approximation with inherent multiresolution capability for error estimation and adaptivity. The basic idea of reproducing kernel approximation theory is to define the approximate function as the convolution of the original function with a kernel function, where the kernel function is sometimes called as the weighting function or the smoothing function.

The boundary node method (Xie Mukherjee & Mukherjee, 1997) is a boundary-only mesh-free method which combines MLSM interpolation scheme with the standard Boundary Integral Equations for a boundary value problem. Curvilinear boundary coordinates were originally proposed and used in this method for both two- and three-dimensional problems.

The boundary cloud method (BCM) was introduced by (G. Li & Aluru, 2002) as an improvement of BNM. Further improvements to BCM have also been proposed leading to the Extended Boundary Node Method (EBNM).

Variations of the mesh-free / meshless methods are in constant evolution, and can be used as tools for modeling and simulation of challenging problems of the continua, noting that they are suitable in multiscale modeling as well as in handling multiphysics problems.

2.2.4 Finite Volume Method

This method is an extension of the classical finite difference method. Here also the problem space is covered by a mesh of discrete number of points where the unknown function is supposed to be determined.

As a small volume is assumed to surround each such point in the mesh, it is termed as the finite volume method. The method is characterized by the conversion of divergence terms in the governing partial differential equation to surface integrals. The method is highly suited for unstructured meshes and is very popular for computational fluid dynamics computations.

2.2.5 Wavelet Method

An important aspect of discrete numerical solution of differential equations of the continua is the characterization of the variation of the unknown function over a sub-domain of the problem space. Like Fourier series expansion but without its restrictions, a wavelet is based on a series of scaled linear combination of a set of approximating functions (Burrus et al., 1998) (Dahmen, 1997) (Cohen et al., 1992) (Cohen & Masson, 1999).

Wavelets allow accurate representation of functions with rapid oscillations, or, possibly, with discontinuities, in localized regions. Although, initially wavelets found primary application in signal and image processing, its unique properties makes it a worthwhile candidate for the numerical solution of the partial differential equations of the continua having different kinds of irregularities (Basu & Shah, 1998) (Basu et al., 2003).

Wavelets based discrete numerical method for the solution of such equations may offer some unique advantages not available in other established discrete numerical methods, like those described above. In these methods, the presence of irregularities (or discontinuities) requires proper attention in spatial discretization, and, sometimes, requires the use of special approximating functions. But it is expected that the use of approximating functions based on wavelets will eliminate or reduce such restrictions.

In the finite element model, the element boundaries have to be located at the internal points of discontinuity. On the other hand, in the mesh-free method, the sub-domain boundaries cannot cross the same points. In the case of boundary element method applied to two-dimensional problems, similar consideration as in the case of finite element method is necessary. In view of positive experiences with wavelets in modeling localized

behavior related to signal and image processing, it is expected that traditional modeling restrictions in the case of problems of the continua can be relaxed. This is possible because this approach is so designed that it adapts itself to the local nature of the problem by retaining the finer wavelet components in the vicinity of an irregular point (say, near a singular point) whereas, the coarser components are retained in the smooth regions. The unknown function can be approximated by a series expansion and Galerkin method can be applied to the weighted residual of the equation. The equation expression for the unknown variable includes scaling functions, wavelet functions, coefficients, scaling parameters, translation. Of the many available scaling functions, two most commonly used ones in signal and image processing are the Haar functions, and the Hat functions. Hat wavelets are continuous and form orthonormal basis and can be obtained by integrating the Haar wavelets with proper scaling. Typical Hat scaling and wavelet functions are shown in Fig. 9.

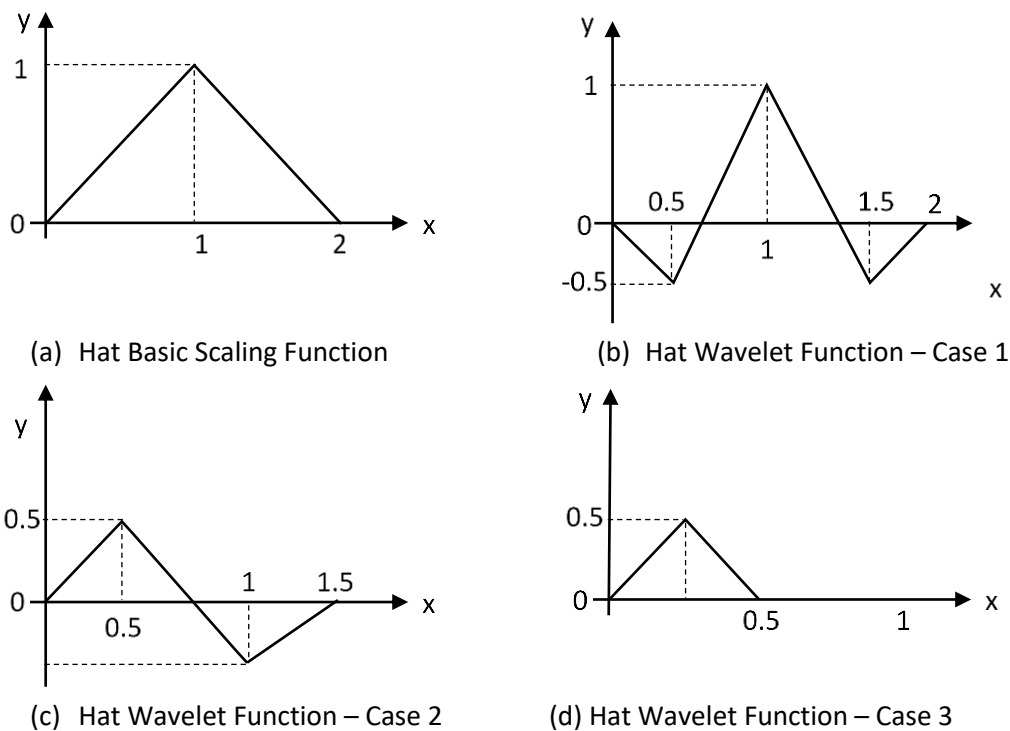


Figure 9: Typical Hat scaling and wavelet functions - adapted from (Basu et al., 2003)

A recent class of wavelets, known as multiwavelets (Aldroubi & Unser, 1993) (Aldroubi & Papadakis, 1998) are able to combine many of the properties of orthogonal wavelets, such as the famous Daubechies wavelets (Daubechies, 1988) (Daubechies, 1992) (Cohen et al., 1992) with the advantages of piecewise polynomial wavelets (such as the Hat wavelet). In particular, in (Donovan et al., 1996), families of orthogonal, smooth (continuous, and continuously differentiable), piecewise polynomial multi-wavelets are constructed.

2.2.6 Cellular Automata Inspired Modeling of Systems

In modeling and simulation of complex fluid systems as found in contaminant transport process in ground water, in slurry wall construction, in the use of fluid slurry during drilling for oil, industrial processes involving thermal variations, particular interactions, and colloidal mixtures, it is difficult to capture the mesoscopic features embedded in macroscopic phenomena of such fluids by using computational fluid dynamics (CFD).

In the CFD based continuum model, the macroscopic characteristics of fluids is based on Navier-Stokes type equations satisfying conservation of mass, energy, and momentum as well as continuity of flow and are discretized in space and time by finite element and finite difference methods.

In the other extreme, one may get most accurate simulation by using molecular dynamics requiring numerical solution of a set of ordinary differential equations in time for each molecule. While marching along time, the locations and velocities of each molecule are updated. The property values for a system of particles are arrived at by averaging in space and time the molecular positions and velocities and also by evaluating the moments and correlations.

However, one practical problem is that using even some of the most massively parallel computer only a few thousand molecules can be modeled at a time to get the results in reasonable time. So, in order to understand interactions between microscopic structure and macroscopic flows, it is necessary to use mesoscopic model, which is significantly less compute intensive than full blown molecular level modeling. At mesoscale an assemblage of fluid molecules termed as particle is used as discrete elements of the model.

Complex fluids exhibiting significant irregularity at molecular level appear homogeneous at the continuum level but exhibit an ordered structure at mesoscopic scale. In recent times it has been shown that such characterization can be effectively handled by using Lattice-Boltzman method (LBM) and Dissipative Particle method (DPM). The primary difference between LBM and DPM is that the latter includes thermal fluctuations but the former does not.

Lattice-Boltzman method (LBM) is essentially an extension of the concepts of cellular automata (Wolf-Gladrow, 2000). In 1940, Neuman and Ulam put forward the concepts of cellular automata. In this method the problem domain is represented by an array of cells. Each cell can be in one of the few possible states. At an instant of time, the state of each cell is determined by the prevailing state of the cells in its immediate neighborhood as well as its previous state. The process is repeated across all the cells in an array repeatedly in successive instants of time. In updating the state of the cells in the array, the same rule is applied to all. In other words, the model is homogeneous with respect to the rules. Updating of all the cells in the array leads to a new generation.

Cellular automata are best used to model situations where local interactions are predominant. This modeling technique has found applications in many areas of physical science, biology, engineering, mathematics, and social science. A popular example is lattice gas modeling (LGM). In this model (Frisch et al., 1986) (Chopard & Droz, 1998), the fluid 'particles' (a conglomerate of large number of molecules but is significantly smaller than the length scale of the simulation) are restricted to move on the links of the array of cells and the motion evolves in discrete time-steps.

LBM resulted from modifications made to LGM by removing the shortcomings. In LBM, instead of considering single particle occupation variables, particle distribution functions are used, linearization of the collision operator and the use of a single time relaxation approximation. As a result of these improvements, it was possible to eliminate statistical noise, and achieve enhanced computational efficiency. LBM, similar to LGA, however, is restricted to uniform lattice structures (Fig. 10), which severely limits its potential application to many practical problems, namely, flow through a porous media, where representations of complex pore geometry require a very fine uniform lattice, leading to computational inefficiency.

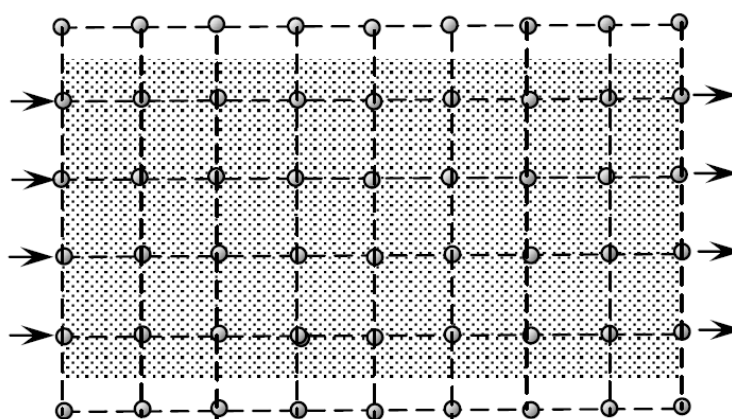


Figure 10: Uniform lattice structure for LBM - adapted from (Basu, 2007)

Least-squares finite element (LSFE) method (Y. Li et al., 2005), on the other hand, was recently shown to be a robust and efficient way to solve non-self-adjoint equations where convection operators are of first order (Jiang, 2000), always leading to symmetric, positive definite linear systems of equations without eliminating the need to use upwinding, staggered grids and operator splitting techniques (Ding & Tsang, 2001). Compared with Taylor-Galerkin-based FE methods, LSFE method possesses improved stability. Furthermore, for more complex systems, Taylor-Galerkin-based FE method may promote oscillations at discontinuities (Jiang, 2000) or at solid-liquid interfaces. Those oscillations may be suppressed by artificially adding dissipation terms like those in 'upwind' and 'artificial viscosity' schemes, which, however, are dependent on the specific parameters of the problem. The model includes a statistical parameter (called as distribution function, from which the macroscopic properties of the fluid can be extracted), the particle velocity distribution (the nine possible velocity directions are shown in Fig. 11), the body force per unit mass, and a collision function. The collision operator is commonly approximated by the Bhatnagar-Gross-Krook (BGK) model (Bhatnagar et al., 1954).

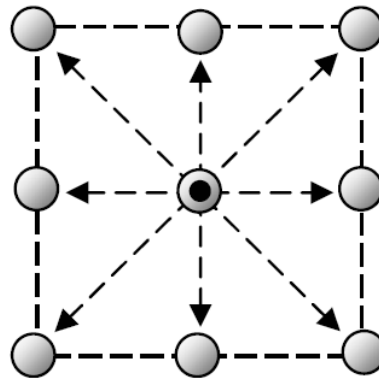


Figure 11: Nine possible directional velocities - adapted from (Basu, 2007)

An application for flow in a porous media is presented in (Y. Li et al., 2005), in which the porous media is considered as a statistical distribution of non-overlapping circular disks representing soil particles distributed in a rectangular two-dimensional uniform continuum representing the pore space through which a fluid flows. Simulation was conducted for randomly generated particle diameters, for both traditional LBM with uniform mesh and LSFE-LBM with unstructured mesh. Fig. 12 shows an irregular triangular mesh used for LSFE-LBM (Y. Li et al., 2005).

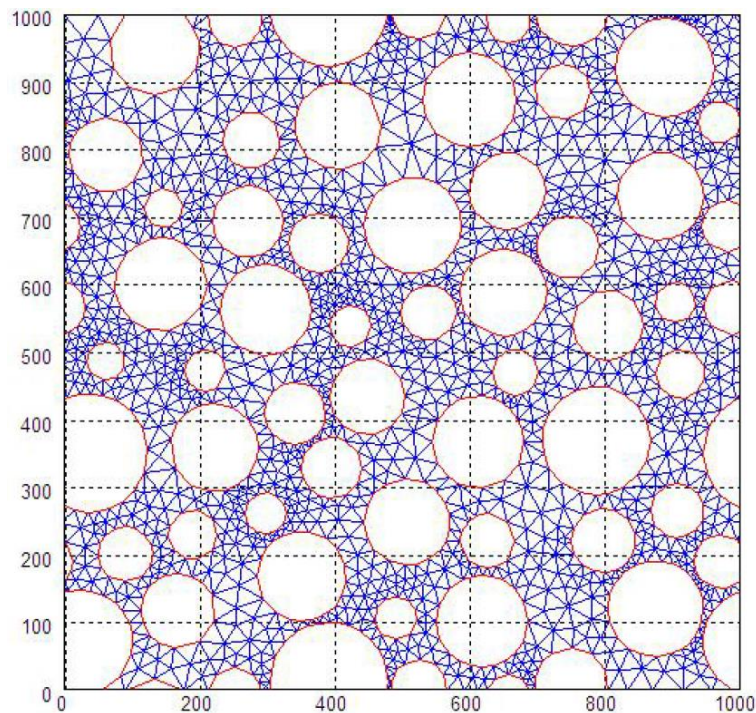


Figure 12: LSFE-LBM mesh used in permeability problem - adapted from (Basu, 2007)

2.2.7 Molecular Dynamics-Based Modeling

Modern materials, especially the complex composites, have hierarchical property in the sense that such materials have constituents which span across different length scales. The failure of such materials is controlled by the behavior at different length scales as well as the process of property transfer between different length-scales. Composite structures are

hierarchical in nature and are made up of substructures distributed across several length scales. The present day technology allows detection of single molecules.

The molecular scale (with nano scale as its subset) is the smallest scale at which modeling can be undertaken realistically. Nano-scale modeling has received further boost due to the recent advances in the synthesis of high-performance materials (like high performance concrete) using nanoscale material constituents. This development will eventually remove the knowledge gap on the interplay of molecular characteristics and continuum characteristics of engineering materials.

At nano scale, modeling can be undertaken by applying either the principles of molecular mechanics (Tomlinson, 1929) (Kontorova & Frenkel, 1938) or molecular dynamics. In molecular mechanics the molecule are treated as rigid balls connected by springs and depends strongly on concepts of bonding and involves energy minimization. Molecular dynamics involves deriving a set of equations of motion and solving these equations by discretization in time.

The equations of molecular (or atomic) motion are based on Newton's laws of motion, and the use of Lennard-Jones 6-12 Potential (Carey, 1999) (Liu et al., 2006). In other words, it shows how the molecules (or, atoms) in a system move with time in the nanosecond timescale (say, 10-15s).

If the coupled behavior of electrons and nucleus are considered, one needs to use complex quantum laws, such as Schrodinger equation. But as per Born-Oppenheimer approximation, the effect of electrons can be reasonably neglected. Thus, the movement of atoms involves the motion of nucleus only, and classical Newton's laws can be applied. As in many practical applications, it is necessary to consider the interaction of nanoscale behavior with the macro-scale behavior, multiscale modeling becomes necessary.

In Molecular Dynamics modeling, it is first necessary to determine the initial configuration of molecules (or atoms) through NMR spectroscopy, or X-ray crystallography. Initial molecular (or atomic) velocities are then assigned assuming random Gaussian distribution with zero mean and determine force $F(x_i)$ on individual molecule (or atom) (Fig. 13).

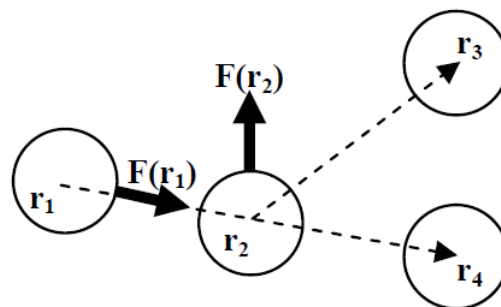


Figure 13: molecular velocities and forces - adapted from (Basu, 2007)

Newton's second law of motion is then applied to each molecule (or atom). The system is now advanced by a small time step during which the molecular forces are assumed to remain constant. The time steps may sometimes be of the order of femtosecond (1 fs = 10⁻¹⁵sec). The molecular velocities and forces can now be calculated. In this manner the process can be repeated, as the simulation marches along time.

A typical plot of the Lennard-Jones potential V_{LJ} is shown in Fig. 14. Here the Lennard-Jones parameters σ and ϵ represent the collision diameter (the distance at which $V_{LJ} = 0$) and the distortion energy, respectively. These parameters are different for different interacting molecules. As for example, for water molecules, $\sigma = 0.316555$ nm and $\epsilon = 0.6501696$ kJ/ml. The attraction results from induced dipole-dipole moment interaction between the particles.

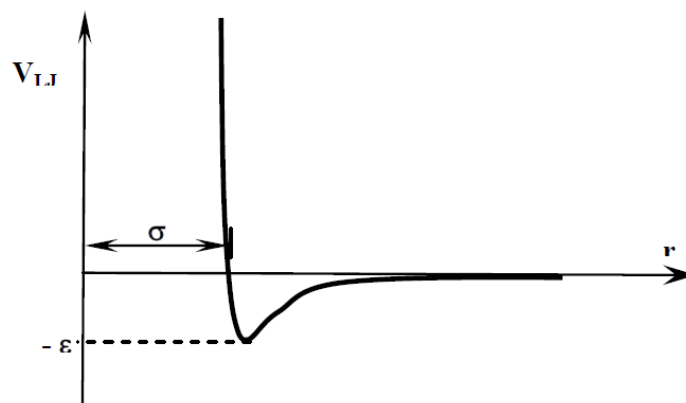


Figure 14: Lennard-Jones potential V_{LJ} - adapted from (Basu, 2007)

The steps involved in MD simulation are

1. Initialize values (Choose initial position and velocities of atoms)
2. Update molecular forces (Calculate forces from positions and potentials)
3. Set up equations of motion and solve for the selected time step, Δt (Calculate positions and velocities after time step Δt)
4. For $t=t+\Delta t$, compute physical quantities of interest (Calculate new forces)
5. Repeat step 2 onwards till maximum simulation time is reached (As the system evolves compute the average quantities of interest, like stresses, strains, pressure, volume, temperature, energy, etc.)
6. Compute the final results.

In problems of mechanics, the results of MD simulation can be used to compute stresses averaged over the N molecules using virial theorem (Tsai, 1979) (Zimmerman et al., 2004) and therefrom the mechanical properties.

2.2.8 Multiscale Modeling

A close examination of modern engineering materials will reveal different characteristics at different length scales. The objective of multiscale analysis is to qualify the load transfer between different length scales.

As for example, we can consider the case of laminated composites used in aircraft structures. In the macro-scale the property of the laminate as a whole is of interest. The next level is individual lamina, which in turn consists of high-strength fibers embedded in a lower strength matrix material. As long as the lamina is perfect, that is the fibers are straight, parallel, uniformly spaced, continuous (or unbroken), perfectly bonded with the matrix material, and are of uniform diameter and strength, the lamina properties are not difficult to define. The problem arises when one or more of these conditions are violated requiring further refinement of scale to account for the flaws which often cause geometrical singularities in the system.

Similar situation exists in the case of high performance concrete used in modern bridge structures because of the constituents used in such concrete. In the case of subsurface contaminant transport phenomenon, one needs to account for the effect of grossly inhomogeneous character and complex material structure of the substrata on the large scale flow characteristics through multiscale modeling. So, it is necessary to develop methods to define load transfer processes between length scales in deterministic sense and in statistical sense for systems with uncertainties. This will allow the definition of accurate macro-scale properties reflecting the behavior in finer scale, say, at the microstructure scale.

First, this will involve up scaling of the microscale characteristics to the macroscale through homogenization process. Very often engineering materials have a periodic microstructure which can be taken advantage of in the multiscaling process. The homogenization process enables the computation of average stresses and strains in the limit as the size of the periodic cell goes to zero. Next it will involve down scaling of the average response at the macroscale to the microscale to predict behavior at the micro scale so that damage initiation or propagation can be monitored.

Apart from multiple spatial scales, multiple temporal scales may also need to be considered. For instance, in the case of fatigue and fracture problems, multiple temporal scales coexist multiple spatial scales due to the wide difference between the period of a single load cycle acting on a structural component, which may be in the order of seconds, and the overall design life, which may span years. The multiple spatial scales exist in such problems, due to the presence of micro- size cracks or inclusions with respect to the large size of a structural component.

As the behavior of many modern engineering materials is controlled by processes subject to wide range of length and time scales, the objective of multiscale modeling is to capture relevant physics from all the scales by coupling the computations at different scales. This can be undertaken by adopting one of the following two schemes.

- a) Hierarchic (or sequential) scheme: Undertake separate simulations at each selected scale passing on the results to the neighboring scale – this can happen from coarse to fine scale, or vice versa (Fig. 15). In this approach the computational simulations are undertaken at one scale and quantities that can be used to define parameters of the model applicable on a longer scale are extracted.
- b) Undertake simultaneous coupled simulations with different scale resolutions in different regions of the problem domain (or discrete patches covering it), as needed. The patches are overlapped at the boundaries with each other applying the boundary values of one patch as the boundary condition of the other patch and vice-versa. This approach is considered to be more efficient.

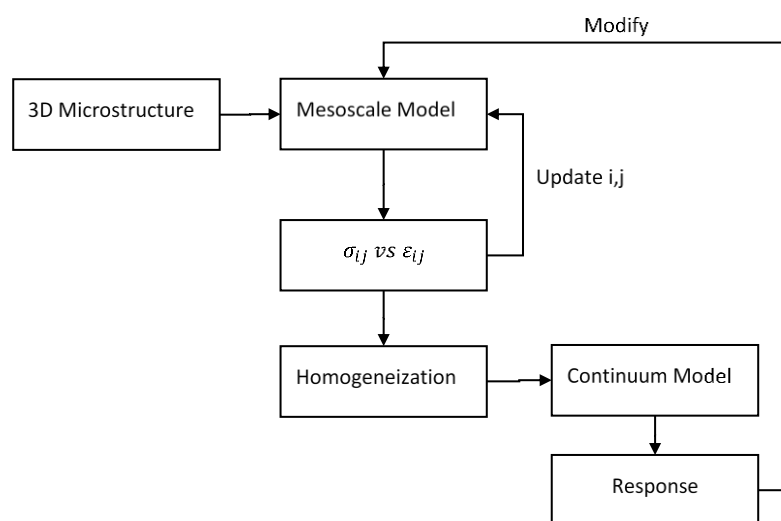


Figure 15: Hierarchic Scheme for Mechanics Problems - adapted from (Basu, 2007)

In analyzing heterogeneous medium, the mathematical homogenization theory (Bensoussan et al., 1979) (Sanchez-Palencia & Zaoui, 1985) has normally been used. The hierarchic scheme uses the assumption of microstructural periodicity and uniformity of macroscopic field within a unit cell domain, and the homogenization process decomposes the heterogeneous medium problem into the cell problem in the micro scale, and the continuum problem in the macro scale (Aboudi, 1991) (Aboudi, 1996). As shown in Fig. 18, the computational steps at an instant of time consist of

- Identification the representative volume element (RVE) and discretization into a regular grid of cells.
- Calculation of the cell problem
- Determination of the homogenized material properties
- Calculation of the continuum problem
- Evaluation of the results in the micro-scale
- Repetition of the above steps till convergence

In the case of composites, the constitutive properties of the matrix material, the fibers, and fiber-matrix interface needs to be considered. For the interface, (Aboudi, 1991) used a bilinear model but more recently a model based on cohesion due to Needleman (Needleman, 1987) (Needleman, 1990b) (Needleman, 1990a) has become more popular. Accuracy of the above simulation scheme is largely dependent upon the validity of the assumptions of periodicity and uniformity. For instance, near steep stress gradient regions, the assumption of local uniform stress field in a unit cell is no more valid requiring the use of higher norms of local stress field. In general, these can be effectively addressed in the context of risk and reliability analysis.

2.2.9 Multiphysics Modeling

In engineering design it is necessary to allow for simultaneous influence of multiple physical and chemical phenomena, sometimes, requiring consideration of mutual interactions between them. The interactions may need to be considered at governing equations level, or at the constitutive laws level.

An example of the latter is diffusion process controlled by temperature. The solutions from an individual discipline may act as input for another discipline. This approach may be acceptable in the case of weakly coupled systems, but in the case tightly coupled systems this may not be acceptable. Typical examples of multidisciplinary coupling in modeling and simulation in science/engineering are given below (Basu, 2002).

- Structural-thermal: Thermal stress analysis of Highway Bridge structure allowing for coupling between stress analysis and heat transfer.
- Fluid-structure: Aeroelastic analysis of airplane structure allowing for interaction between airflow and structural response; or, action of waves on maritime structures or ships.
- Fluid-structure-thermal: Welded joint analysis beginning with weld metal deposition to cooling down to ambient temperature.
- Magneto-thermal: Induction heating analysis of, say, annealing process of electrically conducting material allowing for coupling between electromagnetic field and heat transfer.
- Acoustic-structure: Acoustic response of aeronautical structures and automotive structures.
- Magneto-structure: Structural response caused by transient magnetic field such as in the case of a speaker.
- Electro-structure: Piezoelectric sensor modeling allowing coupling between deformation and voltage distribution.
- Advection-diffusion-dispersion: Modeling of groundwater flow associated with pollution transport and remediation.
- MD-DD-CM: Multiphysics/multiscale modeling and simulation of radiation damage to nuclear reactor pressure vessel steel by using coupled molecular dynamics, damage dynamics, and continuum models.
- QM-MD-CM: Coupling of quantum mechanics, molecular dynamics and continuum modeling and simulation of carbon nano-tube devices.

The groundwater pollution transport problem mentioned above is a multiphysics, multiscale, and multiphase inhomogeneous porous media problem. The processes involved are advection, dispersion, adsorption, and decay requiring hierarchic modeling from pore-throat scale to the pore scale, to the core scale and beyond culminating at the field scale.

The problem involves three phases – fluid, solid, and vapor. During the remediation process, other physics components may be involved wrought by thermal enhancement, intrinsic bioremediation, electro-kinetics, etc. Number of, so called; grand challenge problems are also highly complex multiphysics/multiscale problems. For instance, the modeling of earth's crust primarily involves coupling between rock mechanics (elastic-plastic rheology), fluid flow (Darcy's law), thermal transport (Fourier's law), and chemical processes.

Another example is the weld cooling problem at a structural joint. This in fact is a coupled fluid-thermal-structural interaction problem. The fluid flow phenomenon during the weld deposition process is usually ignored, and the problem is traditionally treated as a two-stage process involving phase change in weld metal from liquid to solid. In the first stage, the heat transfer problem is solved to determine the temperature distribution. In the

second stage, thermal stress analysis is undertaken using the temperature distribution disclosed in the first as input.

If the governing equations are tightly coupled, the equations for all the disciplines are considered simultaneously, often, require allowing for nonlinearities. This approach may also be used in the case of loosely coupled system, if a large amount of data needs to be exchanged between different disciplines.

Other reasons for doing so in the case of loosely coupled and moderately coupled systems are:

- a) difficulty in properly synchronizing the discipline modules;
- b) proper mapping of transferred results from one module to another;
- c) difficulties in monitoring of error and convergence characteristics of the results.

A big advantage favoring loose coupling is the existence of codes for individual disciplines. A major disadvantage of tightly coupled systems is that the code is very difficult to implement, especially, in the case of highly complex systems. Whereas, in the case of a moderately coupled system, the dominant disciplines is treated as primary, and the effect of other disciplines are added on after introducing some simplifications.

For instance, in the case of fluid-structure interaction simulation of a submarine, computational fluid dynamics (CFD) is the dominant feature and the simplified structural aspects of the problem will be added to CFD code. In the simplest model, the submarine can be treated as a rigid body and interaction with deformations is not accounted for.

2.2.10 Integrated Systems

The global economy of this millennium requires the development of superior products cheaply and quickly. Also, for instance, in the case of civil infrastructure, apart from safety and functionality, the emphasis is on life-cycle cost reflecting factors like projected life, risks, first cost, maintenance cost, replacement cost at the end of useful life, etc. In a nutshell, the driving force is the development of an optimal solution allowing for all the factors involved in the multi-disciplinary environment.

Another interesting example is that of a commercial aircraft design, which requires accurate prediction of its flight characteristics controlled by the interaction of disciplines like aerodynamics, structural dynamics, solid mechanics, acoustics, propulsion, thermodynamics, controls, and manufacturing processes, and the involvement of stakeholders like owners, pilots, flight attendants, maintenance people, passengers, environmental agencies, etc.

In designing Model 777, and its, decedent Model 787, commercial aircraft, Boeing fully exploited the interdisciplinary approach when 238 cross-functional design-build teams produced a paperless design that allowed engineers to design every component, electronically simulate the assembly process to eliminate errors in fitness and assembly process.

In the process, alternative designs could be evaluated as single integrated units before actual assemble and test flights. This eased the transition from the design stage to production for actual commercial introduction. Boeing has used similar concept design-build concept to the International Space Station with a mass of about one-half million

kilogram representing one of the largest and, perhaps, the most complex engineering project in history involving interdisciplinary effort of international scope.

It is obvious from above that in the global market place, in undertaking a new project, time-consuming trial-error approach require to be minimized. It is important that all aspects of realization of a new project (or product) – design, manufacture or construction, cost, reliability, and support process, be considered at the earliest stage.

To avoid unforeseen problems late in the design process, it is important to undertake design evaluations throughout the design process accounting for all possible interactions. It may be emphasized that the design process in many modern projects is prone to complex multidisciplinary interactions of a number of physical disciplines over several length and time scales. This no doubt is in sharp contrast to the conventional practice of using one discipline at a time.

As for example, in the case of a highly complex structure like high-speed airplane, the practice has been to simulate for each discipline, as stated earlier, individually and the effects of mutual interaction sorted out mostly through numerous expensive prototype tests over a considerable length of time. In opposition to this approach, modeling and simulation allowing for interaction amongst these disciplines will lead to significant savings in cost and development time can be achieved (Basu, 2002). For example, in the coupled dynamic stability analysis of an airplane, the interaction between aerodynamics, elastodynamics, and control theory are used (Crisfield, 2001). In addition, if piezoelectric controls are used, electromechanical coupling will also need to be considered.

2.2.11 Risk and Reliability

Engineering design is based on mathematical models, which are driven by all kinds of information based on experiments, observations, intuition, experience, assumptions, etc. which cannot by any stretch of imagination be considered as highly accurate. The response of the system being designed is predicted by exercising (or making simulation runs) on the mathematical model will have a degree of uncertainty due to inaccuracies in the information supplied.

Additional uncertainties are introduced during the execution of a project due to poor workmanship due to deficiencies in materials and methods used, errors, etc. Under these circumstances, it is not possible to predict the performance of a designed and built system in absolute terms or deterministically.

In order to account for the uncertainty of the parameters controlling the response of a system, it is necessary to undertake probabilistic simulation assuming that the sensitive parameters of the system are random in nature. In order for this to be possible, it is necessary to have enough information about the random nature (or distribution characteristics) of a parameter of interest. In the absence of adequate experimental data for a random variable, one of the available standard distributions based on qualitative judgment is used.

In reliability-based design, the objective is to proportion a structure so that the probability of failure, P_f , does not exceed certain threshold value. In structural engineering terms, the common formal procedures of reliability analysis are first order reliability methods (FORM), second order reliability methods (SORM), and Monte Carlo method (Ang & Tang, 1984) (Haldar & Mahadevan, 1999) (Shinozuka, 1983).

Of these, the last method is very compute intensive and essentially is a brute force method. In this method, once the probability function is defined, random sampling is followed by Monte Carlo simulation. After large enough number of such simulations, statistical analysis of the outputs of interest is undertaken to obtain quantities like expected value and standard deviation. This can be followed by the computation of reliability index β given by

$$\beta = \frac{\mu_R - \mu_Q}{\sqrt{\sigma_R^2 + \sigma_Q^2}} \quad (3)$$

Here, μ_R and μ_Q are expected values of resistance (R) and load effect (Q), and σ_R^2 and σ_Q^2 are variance of resistance and load effect. Graphically, reliability index signifies the shortest distance from the origin to the limit state function, $g(Z_R, Z_Q)$, where Z_R and Z_Q are the reduced variables defined as below. Larger the reliability index smaller is the probability of failure. For instance, if $\beta=2.33$, the probability of failure is 1 in 100.

$$Z_R = \frac{R - \mu_R}{\sigma_R} \quad \text{and} \quad Z_Q = \frac{Q - \mu_Q}{\sigma_Q} \quad (4)$$

Risk and reliability methods have lately matured lately. Component level reliability evaluations are being extended to systems level. Significant efforts are being made towards reliability-based optimization of systems, including multidisciplinary problems. Uncertainties introduced by the choice of mathematical model as well as the discrete numerical technique used in predicting the response are also being looked at.

In civil engineering design, purely deterministic design approach based on the use of safety factors have been replaced by load and resistance factor design (LRFD) which uses different uncertainty factors (load factors) with different external effects representing the demands on the system, and another set of uncertainty factors (resistance factors) applied to capacity measures (limit states) of the system.

These factors indirectly allow for the randomness of various design parameters without the need for undertaking formal risk and reliability (or probabilistic) analysis (Ellingwood & Galambos, 1982) (Thoft-Christensen & Baker, 1982). In railroad and automotive industries there is some interest to use probabilistic methods mainly for investigative purposes. In the aerospace industry also there is an ongoing effort to look at the possibility of adopting some of the risk and reliability methods in aerospace vehicle design, primarily, prompted by the US Department of Defense.

A number of reliability analysis, methods, formulations and applications have been proposed, for different types of problems. One can cite, for example, a study on formulations of the discretization of random fields in reliability structural analysis (Jorge, 2004), and a study of an application of reliability analysis for the design of a helicopter composite armor, with correlated design variables (Santos et al., 2016).

2.2.12 Optimization

Designing a system is an iterative process in which the design variables are adjusted to reach the levels which will lead to the best possible outcome. In other words, the final objective is to improve an existing design or replace an unacceptable design by one that meets the given requirements, that is, choosing the best possible design that satisfies all the special constraints.

This is called optimum design, which can be arrived at intuitively (that is, indirectly) or logically (that is, directly). The intuitive approach is based on past engineering experience and involves random selection and evaluation. On the other hand, the formal design optimization follows a logical approach to arrive at the optimal design.

The formalities include defining the behavioral constraints, the side constraints, and the objective function which is minimized or maximized. The objective may be weight minimization, cost minimization, quality maximization, or maximization of value, which is combination of cost and value. Other objectives can also be used, like minimizing vibrations in a vehicle (Wang et al., 2003) (Wang et al., 2004). The final objective is to achieve competitive advantage in today's global economy.

Mathematical programming algorithms include linear programming, nonlinear programming, geometric programming, dynamic programming, etc. (Rao, 2019) (Arora, 2012). Lately, genetic algorithms and neural networks have been used quite efficiently for certain types of optimization problems.

Genetic algorithms are evolutionary algorithms inspired by evolutionary biology and use techniques like mutation, selection, inheritance, and crossover in the abstract representation of candidate optimal solutions that evolve towards better solutions. Genetic algorithms are implemented as a computer simulation in which a population of abstract representations (called chromosomes or the genotype or the genome) of candidate solutions (called individuals, creatures, or phenotypes) to an optimization problem evolves toward better solutions.

A neural network consists of a network of simple processing elements with complex global characteristics, determined by the connections between the processing elements and element parameters. The genetic algorithm technique has been found to be extremely effective for small to medium sized constrained optimization problems.

However, in the case of more computational resource intensive complex problems, a combination of neural networks and genetic algorithm tend to improve the situation. In linear programming both objective function and the constraints are linear functions of design variables.

In civil engineering design, the objective function and/or the constraints are mostly nonlinear functions of design variables. In most cases, linear programming cannot be used, except when the nonlinear programming problem can be transformed into an equivalent linear programming problem.

A large number of software have been developed based on these algorithms and also encoded in many commercial modeling and simulation software. Also, software like GENESIS by VR&D was especially developed with the primary objective of design optimization. In addition, software for optimal design search, such as DOT by VR&D (Ghosh & Vanderplaats, 1997), which can be integrated with regular modeling and simulation software, are also available. In view of the fact that much of the design variables are random in nature, there is a surge towards reliability-based optimum design. The objective is to arrive at the optimum values of the design variables with a certain specified level of reliability. Software undertaking such optimization should also be able to undertake design sensitivity analysis to identify the critical design variables.

Another twist to optimization process appears when the objective functions and constraints represent a number of disciplines which may be coupled in some manner. In

this case, instead of a single objective function, multiple objective functions may need to be satisfied. Getting an optimal solution for such multidisciplinary problem is a more challenging proposition, especially when randomness of design variables is also considered. Another area of current interest allied to optimization process is the inverse problem in which for a given system response it is necessary to arrive at the design variables which will enable it.

There is significant increase in research effort research in the field of structural optimization in recent decades. The increasing interest in this field has been strongly boosted by the advent of reliable discrete numerical modeling tools, advances in design sensitivity analysis, and methods of mathematical programming, along with the exponentially increasing speed and capacity of digital computers.

2.3 Some Remarks in Modeling and Simulation Tools

This probably is the most challenging time for those using modeling and simulation tools as part of the engineering design process to effectively respond to the strong technological competition, which demands reduced design time and cost of products with greater functionality and higher quality, coupled with present day emphasis on energy savings, material recycling, and sensitivity to environment.

With ever-increasing power of digital computers coupled with continuous improvement of mathematical models and the significant advancement of numerical simulation techniques, the future for more accurate prediction of system behavior leading to optimal design of efficient systems meeting the global challenge seems to be bright. This section may be considered simply as an attempt to scratch the surface of this vast and expanding field.

3 On Error Estimators in Computational / Numerical Methods

The various discretized integral representations of Boundary Value Problems (BVP), including, for example, Finite Element Methods (FEM) and Boundary Element Methods (BEM) will produce results with some error with respect to the exact solution.

Some possible sources of error are inherent to almost all formulations, when these formulations are applied to a non-trivial problem. Several sources of error may be present in the numerical model of a problem. Some sources of error may play an important role in specific formulations, or may be inherent to a certain formulation. What source of error is the most important in a particular problem may depend on the representation being adopted. The same applies to the sensitivity of the response error due to a specific source.

The first source of error may be in the modeling of the problem. The importance of an error source in a certain representation depends on the problem being modeled. For example, corners may have different influences in the problem, depending if it is a potential problem or an elastostatics one, because corners will lead to discontinuous derivatives, and these discontinuous derivatives satisfy different relationships in each problem.

The potential problem is modeled by a scalar formulation derived from Laplace's equation. Normal and tangential derivatives will be discontinuous at these corners. The gradient of the potential is a continuous function in the interior domain. This gradient can

be written into components in two or three orthogonal directions, depending if it is a 2-D or 3-D problem. When taking the limit to the boundary, if the boundary is smooth, this gradient can be written in tangential and normal components, with the tangential and normal directions considered at the point in the boundary to where the limit was taken. The smoothness of the boundary guarantees that these directions are unique at some point in the boundary, so the decomposition of the unique gradient of the potential is also done in unique components. Now, if there is a corner, although the gradient could be considered as continuous in the limit process from an interior point to the boundary, its decomposition would be in different tangential and normal directions for the different elements sharing that point of the boundary. This circumstance leads to a different decomposition of the same gradient vector. Therefore, a relationship exists between tangent and normal components of the gradient vector in the two elements sharing the corner, by means of a simple rotation in the local coordinate system.

In elastostatics, normal and tangential derivatives of the displacement are the tractions, and they will also be discontinuous at the corners. A different relationship exists between these derivatives in the two boundary elements sharing the corner. Hooke's law can be written for a small domain element having the two adjacent boundary elements as parts of its own boundary. As the stress in the interior of this domain element is continuous, a relationship appears between the tractions in the two boundary elements sharing the corner, to satisfy equilibrium in the domain element.

The second source is the round-off error. The round-off error may be different for different integral representations of the same problem. The parameter that controls this error is the condition number of the matrices obtained for a certain formulation. The condition number is the ratio between the greatest and the lowest eigenvalue of the obtained matrix of the system of equations. If the magnitude of this number is written as a power of 10, say, 10^m , then m is the number of digits that shall be discarded in the numerical solution, because these digits are not significant, and contain only the error. A comparison between the condition number of the matrices as obtained from two different BEM formulations may show, for instance, the formulation where more round-off error is introduced.

The third source is errors in the evaluation of parameters that may be necessary in some formulations, like the tangential derivatives of the potential or the displacement. Tangential derivatives are obtained, in general, as a post-processed result, by taking local derivatives of the displacements in the tangential direction. For a 2-D element, the derivative in the tangential direction is the derivative with respect to the intrinsic coordinate ξ . Then the gradient is obtained combining the normal and the tangential derivatives. For isoparametric elements, the tangential derivative is, then, one degree less than the normal derivative. For example, if the element is quadratic, the displacement and the normal derivative are written in terms of quadratic shape functions, while the tangential derivative, which is the derivative of the displacement in the ξ direction, has linear shape functions. The gradient is, then, obtained in a somewhat "unbalanced" manner. The error introduced by the approximation in the tangential derivative is likely to be more significant in the regions of the boundary in which the tangential derivative is changing very fast, as its representation with shape functions that are one degree less is poorer than the representation of the normal derivative.

The fourth source is the discretization error. Usually this discretization error can be global or local. When a local error estimator is used, then an adaptive procedure to mesh refinement can be adopted.

Other sources are errors due to approximations of the boundary shape and errors due to simplifications or approximations on the intrinsic necessary hypothesis of a specific representation. For example, not all representations allow for discontinuities or non-smoothness in the boundary. So, if corners are present, an error may be introduced in some representations.

Besides the above-mentioned deterministic errors, the inherent randomness of the problem may be itself a source of additional stochastic error terms, which might appear due to uncertainties in the various parameters and variables of the problem, such as the material properties, the geometry, the boundary conditions, the loading, etc. These uncertainties may be treated through modeling these parameters and/or variables as:

- random variables (mean and standard deviation are constant over time and over position);
- random fields (mean and standard deviation are constant over time, but variable over position)
- stochastic processes (mean and standard deviation varying over time, but constant over position)
- a mixture of the above cases.

3.1 Error Estimators in Finite Element Methods (FEM)

3.1.1 Introduction

This section presents a summary of the discussion as presented in (Huebner et al., 2008). If the exact 2-D solution for a problem $T(x, y)$ is known, then the error associated to the finite element solution for this problem $T_{FE}(x, y)$ can be written as:

$$E(x, y) = T(x, y) - T_{FE}(x, y) \quad (5)$$

One can note that the error varies over the solution domain.

Global error estimation

Global error estimation procedures:

- need to know *a-priori* the exact solution;
- use some norm of the error over the domain. Example:

$$L_2 \text{norm} : \|E\|_{L_2} = \left[\int_{\Omega} E^2 dA \right]^{\frac{1}{2}} \quad (6)$$

- evaluate this integral over each element and sum it over all elements.

For elliptic boundary value problems:

$$\|E\|_{L_2} = Ch^p \quad (7)$$

where:

- C is a constant;

- h is a parameter that depends on mesh size. It is related to some characteristic element dimension;
- p is an integer. It depends on the order of the interpolating functions.

If the equation in vector form is

$$[L]\{u\} = \{b\} \tag{8}$$

Where $[L]$ is an adjoint linear operator, then:

$$\|E\|_{L_2} = \left[\int_{\Omega} \{E\}^T [L]\{E\} dA \right]^{\frac{1}{2}} \tag{9}$$

Possible refinements to make $\|E\| \rightarrow 0$ are:

- h - refinement. It is performed decreasing the element dimensions;
- p - refinement. It is performed increasing the order of the interpolating functions;
- r - refinement. Remeshing is done by repositioning the nodes, for same h and p .

Local error estimation

If the exact solution is not known a local error estimation procedure can be adopted. One example of this procedure is the local error estimation based on the interpolation error.

One-dimensional problem

The following development is done, for simplicity, for a one-dimensional problem.

Expansion in Taylor series about a point:

$$E(x) = E(\bar{x}) + \frac{dE}{dx} \Big|_{\bar{x}} (x - \bar{x}) + \frac{1}{2} \frac{d^2E}{dx^2} \Big|_{\bar{x}} (x - \bar{x})^2 + \dots \tag{10}$$

Where the higher order terms for x are neglected in a vicinity of \bar{x} . Take $\bar{x} = x_M$, the point in one element where error is a local maximum. Then $\frac{dE}{dx}(x_M) = 0$. Figure 16 illustrates the case.

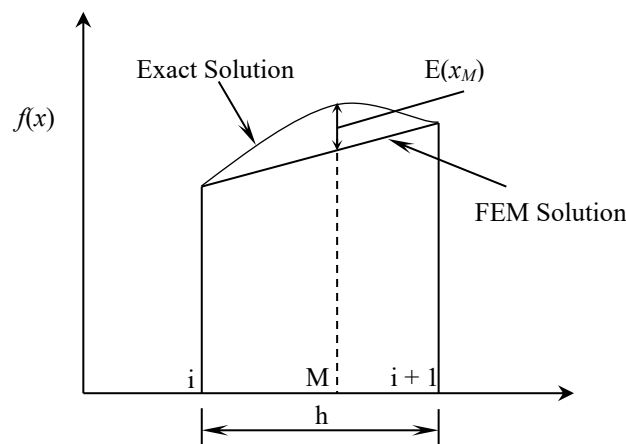


Figure 16: FEM solution error - adapted from (Huebner et al., 2008)

Assume that the FE solution is exact at the nodes. This is valid for $\frac{d^2T}{dx^2} = Q$ (and, in particular, for Laplace's equation, which is this equation when $Q = 0$). But this is only approximate for other equations.

For $x = x_i$:

$$0 = E(x_i) = E(x_M) + \frac{1}{2} \frac{d^2E}{dx^2} \Big|_{x_M} (x_i - x_M)^2 \Rightarrow E(x_M) = -\frac{1}{2} \frac{d^2E}{dx^2} \Big|_{x_M} (x_i - x_M)^2 \tag{11}$$

For $x = x_{i+1}$:

$$0 = E(x_{i+1}) = E(x_M) + \frac{1}{2} \frac{d^2E}{dx^2} \Big|_{x_M} (x_{i+1} - x_M)^2 \Rightarrow E(x_M) = -\frac{1}{2} \frac{d^2E}{dx^2} \Big|_{x_M} (x_{i+1} - x_M)^2 \tag{12}$$

Where either $|x_i - x_M| \leq \frac{h}{2}$ or $|x_{i+1} - x_M| \leq \frac{h}{2}$. This leads to

$$|E|_{(e)} \leq \frac{1}{8} \left| \frac{d^2E}{dx^2} \right|_{x_M} h^2 \tag{13}$$

which is an estimate for the interpolation error at element (e). As $\left| \frac{d^2E}{dx^2} \right|_{x_M}$ is unknown, it is approximated by

$$\left| \frac{d^2E}{dx^2} \right|_{x_M} \cong \left| \frac{d^2T_{FE}}{dx^2} \right|_e \tag{14}$$

This is not anymore an estimate of the error, but just an error indicator. If, in particular, T satisfies Laplace's equation, then

$$\frac{d^2E}{dx^2} = \frac{d^2T}{dx^2} - \frac{d^2T_{FE}}{dx^2} = -\frac{d^2T_{FE}}{dx^2} \tag{15}$$

Because $\frac{d^2T}{dx^2} = 0$. In this case

$$\left| \frac{d^2E}{dx^2} \right|_{x_M} = \left| \frac{d^2T_{FE}}{dx^2} \right|_{x_M} \tag{16}$$

But even in this case, Equation 14 is still just an approximation as $\left| \frac{d^2T_{FE}}{dx^2} \right|_{x_M}$ is in this equation considered to be the same (constant) for the entire element (e).

Two-dimensional problem

Similarly, for a two-dimensional problem:

$$|E|_{(e)} \leq \frac{1}{8} \left(\left| \frac{\partial^2 T_{FE}}{\partial x^2} \right|_{x_M} + 2 \left| \frac{\partial^2 T_{FE}}{\partial x \partial y} \right|_{x_M} + \left| \frac{\partial^2 T_{FE}}{\partial y^2} \right|_{x_M} \right) h^2 \tag{17}$$

From the above equation, it can be seen that 2nd derivatives of FE solution have to be computed, for error estimation. A procedure to smooth the gradient to compute these derivatives with good accuracy is called gradient recovery.

Gradient recovery

If FE solution is linear, then nodal 2nd derivatives are evaluated from a weighted average of element derivatives surrounding the node. The weighting factors are the element areas.

From $\{T^{(e)}\}$ it can be written:

$$\left(\frac{\partial T}{\partial x}\right)^e = \left[\frac{\partial N}{\partial x}\right] \{T\}^{(e)} \quad (18)$$

The derivatives are now written in an average sense

$$\left(\frac{\partial T}{\partial x}\right)_{AVE}^e = [N] \left\{\frac{\partial T}{\partial x}\right\}_{AVE}^{(e)} \quad (19)$$

to find $\left\{\frac{\partial T}{\partial x}\right\}^e$. It can now be written:

$$\int_{\Omega^{(e)}} \{N\} [N] dA \left\{\frac{\partial T}{\partial x}\right\} = \int_{\Omega^{(e)}} \{N\} dA \frac{\partial T^{(e)}}{\partial x} \quad (20)$$

This is a system of equations from which the vector of nodal derivatives $\left\{\frac{\partial T}{\partial x}\right\}^{(e)}$ can be obtained. The same approach can be used for the 2nd derivatives. From $\left\{\frac{\partial^2 T}{\partial x^2}\right\}^{(e)}$:

$$\left(\frac{\partial^2 T}{\partial x^2}\right)^e = \left[\frac{\partial N}{\partial x}\right] \left\{\frac{\partial T}{\partial x}\right\}^{(e)} \quad (21)$$

2nd derivatives can now be evaluated in this same average sense:

$$\int_{\Omega^{(e)}} \{N\} [N] dA \left\{\frac{\partial^2 T}{\partial x^2}\right\} = \int_{\Omega^{(e)}} \{N\} dA \frac{\partial^2 T^{(e)}}{\partial x^2} \quad (22)$$

This is a system of equations from which the vector of nodal 2nd derivatives $\left\{\frac{\partial^2 T}{\partial x^2}\right\}^{(e)}$ can be obtained. Figures 17 and 18 illustrate the smoothing process to obtain the refined gradient from the FE solution.

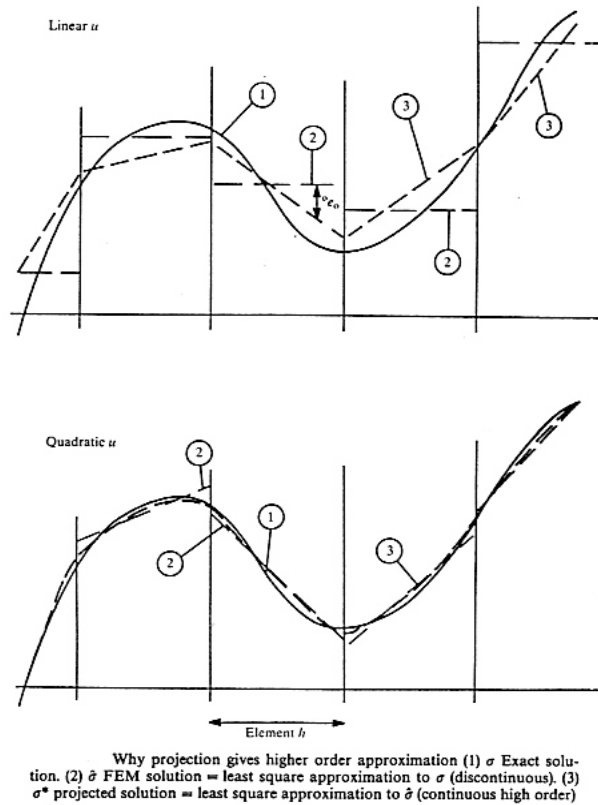
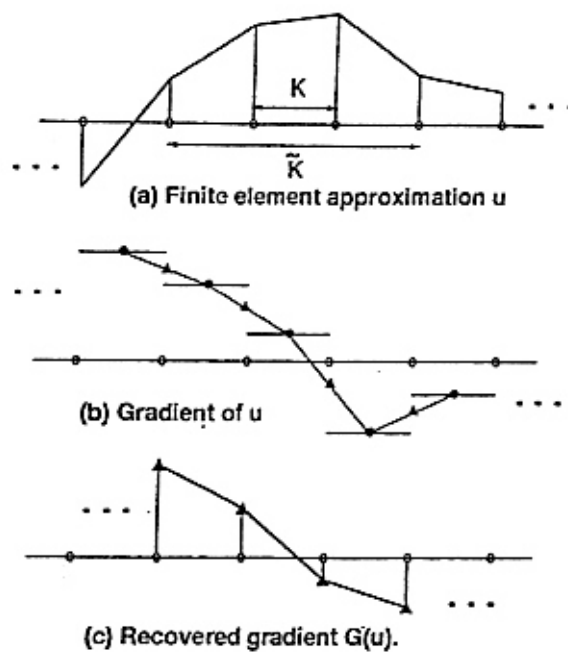


Figure 17: Higher order approximation for the gradient - adapted from (Olgierd C. Zienkiewicz & Taylor, 1994)



Construction of recovery operator $G_X = G$ from piecewise linear approximation in one dimension.

Figure 18: Construction of recovered gradient - adapted from (Ainsworth & Oden, 1997)

3.1.2 Error Estimates and Adaptive Finite Element Refinement

This section follows the discussion as presented in (Huebner et al., 2008) and in (Olgiard C. Zienkiewicz & Taylor, 1994). Error estimators are used in adaptive remeshing. The goal is to have constant error throughout the domain, or to have the error uniformly distributed throughout the mesh.

- 1-D: find h such that

$$h^2 \left| \frac{d^2 T}{dx^2} \right| = Constant \tag{23}$$

- 2-D: find eigenvalues λ_1, λ_2 and corresponding eigenvectors $\mathbf{x}_1, \mathbf{x}_2$ of the matrix

$$\begin{pmatrix} \frac{\partial^2 T}{\partial x^2} & \frac{\partial^2 T}{\partial x \partial y} \\ \frac{\partial^2 T}{\partial x \partial y} & \frac{\partial^2 T}{\partial y^2} \end{pmatrix} \tag{24}$$

such that

$$h_1^2 |\lambda_1| = h_2^2 |\lambda_2| = Constant \tag{25}$$

Where h_i is the element size in the direction of the eigenvector \mathbf{x}_i corresponding to the eigenvalue λ_i . With this, an h refinement process is obtained. Figure 19 illustrates the process.

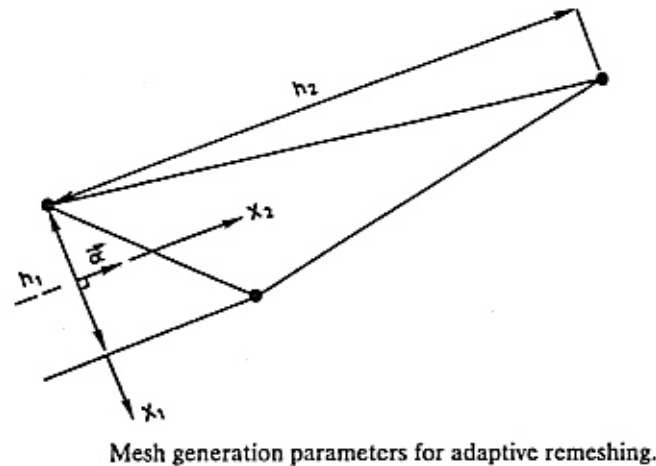


Figure 19: Resizing the mesh - adapted from (Huebner et al., 2008)

Errors and adaptivity

The general procedure is

- Find the error, i.e., the *a-posteriori* estimate. For an elasticity problem, the errors could be:
 - in displacement: $e_u = u_{Exact} - u_{FE}$
 - in stress: $e_\sigma = \sigma_{Exact} - \sigma_{FE}$

(b) Use this information on the error to refine the approximation. Refinement methods could be *h*-, *p*-, *hp*- or *r*-. Adaptive refinement process is used for optimal results. Figure 20 illustrates the refinement process.

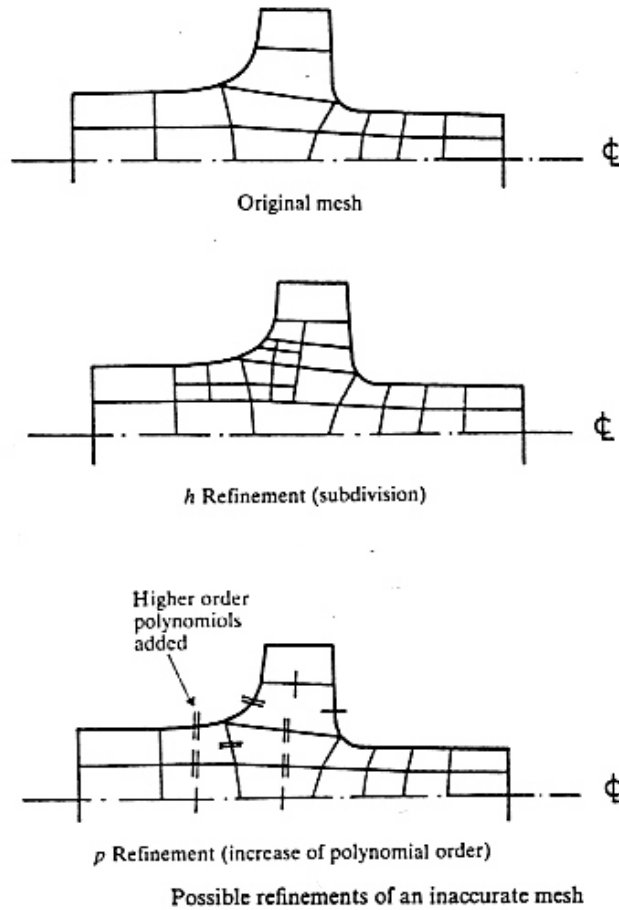


Figure 20: h- and p- refinements -adapted from (Olgierd C. Zienkiewicz & Taylor, 1994)

Error norms and convergence rates

Let $Lu + P = 0$ and $\sigma = D\varepsilon$ be the equations for elasticity problem. Different norms can be defined:

- Energy norm:

$$\|e\|^2 = \int_{\Omega} e^T L e d\Omega = \int_{\Omega} (u_{Exact} - u_{FE})^T L (u_{Exact} - u_{FE}) d\Omega \int_{\Omega} (\sigma_{Exact} - \sigma_{FE})^T D^{-1} (\sigma_{Exact} - \sigma_{FE}) d\Omega \tag{26}$$

- L_2 norm for displacement error:

$$\|e_u\|_{L_2}^2 = \int_{\Omega} (u_{Exact} - u_{FE})^T (u_{Exact} - u_{FE}) d\Omega \tag{27}$$

- L_2 norm for stress error:

$$\|e_{\sigma}\|_{L_2}^2 = \int_{\Omega} (\sigma_{Exact} - \sigma_{FE})^T T (\sigma_{Exact} - \sigma_{FE}) d\Omega \tag{28}$$

Total error in the entire domain is the sum of the errors in all elements.

$$\|e\|^2 = \sum_{i=1}^m \|e\|_i^2 \tag{29}$$

Convergence rate is the slope of the curve error (in non-dimensional form) and the number of degrees of freedom of the FEM mesh used. The bigger this slope, the faster convergence is attained with the same refinement. Figures (21) and (22) illustrate the differences in convergence rates when there is a singularity. This example shows that there is a need for highly refined meshes in a vicinity close to a singularity.

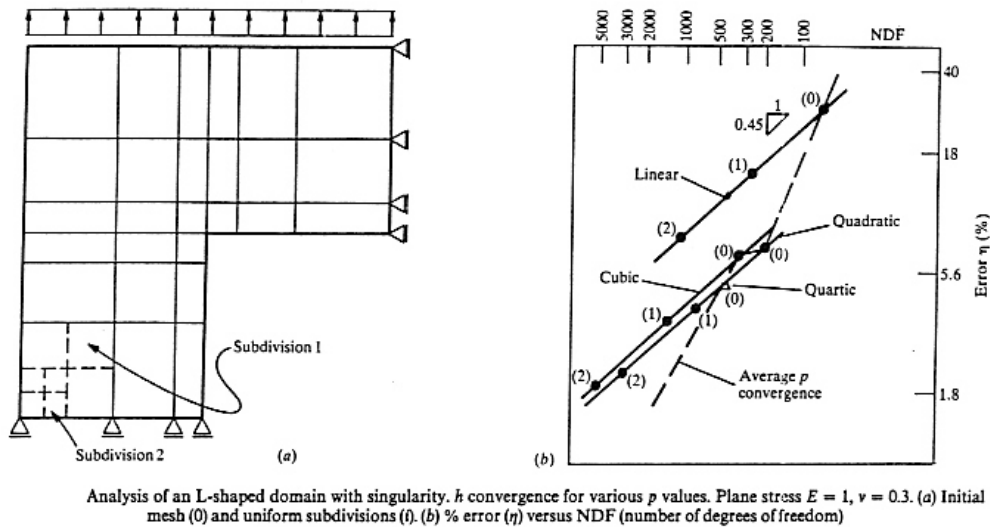


Figure 21: L-shaped domain with singularity (Olgierd C. Zienkiewicz & Taylor, 1994)

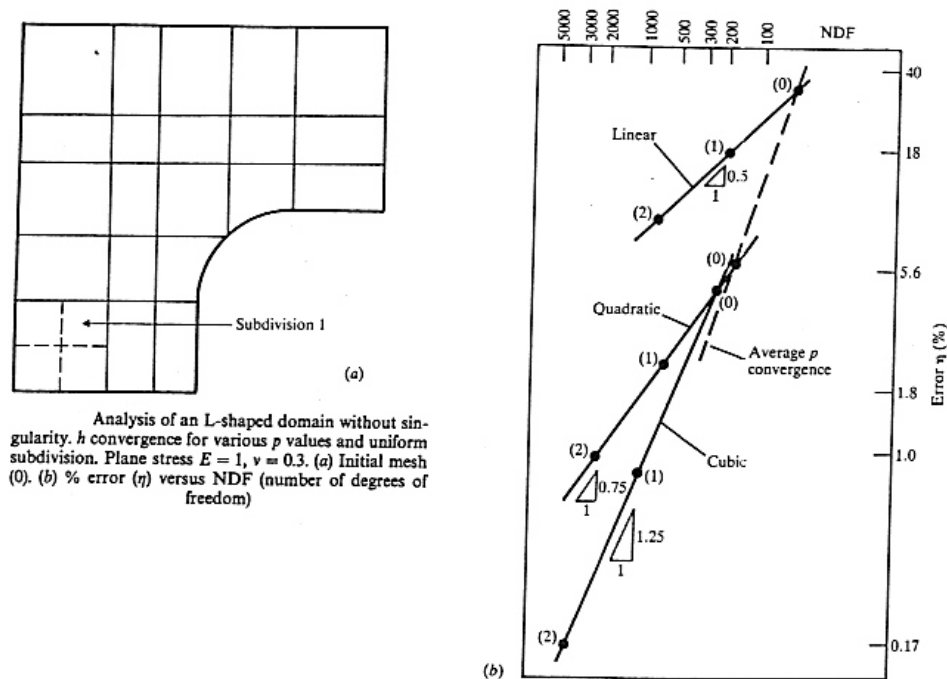


Figure 22: L-shape domain without singularity (Olgierd C. Zienkiewicz & Taylor, 1994)

3.1.3 *a-posteriori* Error Estimation in Finite Element Analysis

This section follows the discussion as presented in (Ainsworth & Oden, 1997).

3.1.3.1 Status and scope

- earliest estimates
 - *a-priori* : for example, errors $\sim Ch^p$
 - predictor corrector algorithms, as for ODE's
 - error η_k on energy norm on element: explicit algorithm; basis for adaptivity
 - complementary energy (global)
 - element complementary plus equilibrated boundary data
 - element residual (plus equilibrated boundary data)
- 1980's
 - *ad-hoc* estimators, for example, based on interpolation. Used in CFD; not accurate for boundary layer problems
 - recovery based methods (smoothed gradients) (superconvergent patch recovery method)
 - extrapolation (global / hierarchical)
- 1990's
 - application to general problems
- current
 - robustness of estimators
 - limits on their performance

3.1.3.2 Reference elements

Standard elements:

$$\mathbf{K} = \{(x, y) : 0 \leq x \leq 1; 0 \leq y \leq 1 - x\} \text{ (triangle)}$$

$$\mathbf{K} = \{(x, y) : -1 \leq x \leq 1; -1 \leq y \leq 1\} \text{ (quadrilateral)}$$

Polynomial spaces of degree $p \in \mathbb{N} : P_k$

Finite element space:

$$x = \left\{ v \in C(\Omega) : v|_k = \mathbf{v} \circ F_k^1 \text{ for some } \mathbf{v} \in P_k \text{ for all } k \in P \right\} \tag{30}$$

3.1.3.3 Model problem

Elliptic BVP: $-\Delta u + cu = f(x)$ in Ω , with BC: $\frac{\partial u}{\partial n} = q$ on Γ_N ; $u = 0$ on Γ_D .

Variational form: $B(u, v) = L(v)$, $\forall v \in V$,

Where $V = \{v \in H^1(\Omega) : v = 0 \text{ on } \Gamma_D\}$

$$B(u, v) = \int_{\Omega} (\nabla u \cdot \nabla v + cuv) dx \quad l(v) = \int_{\Omega} f v dx + \int_{\Gamma_N} g v ds \tag{31}$$

FE approximation: $u_x \in X$ such that $B(u_x, v_x) = L(v_x)$

Error: $e = u - u_x$; $e \in V$ Residual equation:

$$B(e, v) = B(u, v) - B(u_x, v) = L(v) - B(u_x, v) \quad \forall v \in V \tag{32}$$

Orthogonality condition:

$$B(e, v_x) = 0, \forall v_x \in X \tag{33}$$

3.1.3.4 Estimators based on gradient recovery

In this approach, u_x is post-processed to obtain more accurate representations of the gradient $G(u_x)$.

An estimate of the error is $G(u_x) - \nabla u_x$.

This estimator makes no use of information from original problem. In Section 3.1.1 this procedure was explained with more detail.

3.1.3.5 Explicit a posteriori estimators

Here, a residual equation similar to Eq. 28 is obtained using integration by parts.

$$\int_k (\nabla e \cdot \nabla v + cev) dx = \int_k r v dx + \int_{\partial k} v (n_k \cdot \nabla u - n_k \cdot \nabla u_x) ds \tag{34}$$

Where: k is the element;

∂k is the boundary;

n_k is the unit exterior normal;

and the residual is: $r = f + \Delta u_x - cu_x$. Solution is bounded:

$$\|e\|_k \leq C_1 \|r\|_{L_2(k)} + C_2 \|n_k \cdot \nabla e\|_{L_2(\partial k)} \tag{35}$$

Where C_1 and C_2 are non-negative constants. If flux on ∂k is replaced by a suitable approximation, a bound on the error is found. This leads to the implicit error residual method.

Error estimate in energy norm

Upper bound on discretization error for the above bilinear model problem

$$\|e\|^2 \leq C \sum_{k \in P} \left\{ h_k^2 \|r\|_{L_2(k)}^2 + \frac{1}{2} h_k \|R\|_{L_2(\partial k)}^2 \right\} = C \sum_{k \in P} \eta_k^2 \tag{36}$$

Where

- $r = f_\Lambda u_x - Cu_x$ in k is the interior residual
- $R = g - \frac{\partial u_x}{\partial n_k}$ on ∂k on Γ_N is the boundary residual
- η_k^2 is the local error indicator in energy norm

Error estimate in L_2 norm

The upper bound is:

$$\|e\|_{L_2(\Omega)}^2 \leq C \sum_{k \in P} \left\{ h_k^4 \|r\|_{L_2(k)}^2 + \frac{1}{2} h_k^3 \|R\|_{L_2(\partial k)}^2 \right\} = C \sum_{k \in P} \eta_{L_2(k)}^2 \tag{37}$$

Where $\eta_{L_2(k)}^2$ is the local error indicator in L_2 norm

Equivalence of estimators

There should be C such that

$$\sum_{k \in P} \eta_k^2 \leq C \|e\|^2 \tag{38}$$

Two sided bound for the error:

$$\frac{1}{C} \|e\|^2 \leq \sum_{k \in P} \eta_k^2 \leq C \|e\|^2 + \sum_{k \in P} h_k^2 \|f - \pi_P f\|_{L_2(k)}^2 + \sum_{\gamma \in \Gamma_N} h_k \|g - \pi_P g\|_{L_2(\gamma)}^2 \tag{39}$$

The effect of numerical quadrature

Actual error is different from η_k^2 . Also, r_k and R_k are finite dimensional approximations on actual data.

3.1.3.6 Implicit a posteriori estimators

The implicit error residual method is obtained by replacing the flux on ∂k in Eq. 34 by a suitable approximation. In this case, a bound on the error is found.

Error residual problem must be solved over k to determine error estimator $\|\phi_k\|$.

To get rid of C_1 and C_2 a local BVP has to be solved for the error:

$$\int_k (\nabla \phi_k \cdot \nabla v + c \phi_k v) dx = \int_k r v dx + \int_{\partial k} v (g_k - n_k \cdot \nabla u_x) ds \tag{40}$$

Where:

- g_k = approximation of flux on ∂k
- $\|\phi_k\|^2 = \int_k (\nabla \phi_k \cdot \nabla \phi_k + c \phi_k^2) dx$
- an infinite-dimensional space is approximated by a finite dimensional sub-space
- boundary flux $n_k \cdot \nabla u$ must be approximated in some effective way
- if $c = 0$, solution may not exist unless:

$$\int_k r dx + \int_{\partial k} (g_k - n_k \cdot \nabla u_x) ds = 0 \tag{41}$$

Subdomain residual method

- homogeneous essential boundary data over subdomains of Ω
- requires the exact solution of the local problem over the subdomain
- the method is seldom used
- integrating by parts leads to $\|e\| < C \sum_{k \in P} \eta_k^2$, which is an explicit estimator in energy norm.

Element residual method

- uses a larger space than the original FE subspace x
- replaces the original problem by a number of independent local problems posed over each k
- BC for elements in the interior of Ω will interfere on existence of solution. For instance, homogeneous essential BC does not guarantee existence of solution. In fact, in this case only for very particular problems the solution will exist.

3.1.3.7 The equilibrated residual method

This approach evaluates implicit error estimates in which boundary fluxes are specially constructed so that problem is well posed.

Let η_k be the local error estimate. Let η be the global error estimate. Then

$$\eta = \left\{ \sum_{k \in P} \eta_k^2 \right\}^{\frac{1}{2}} \tag{42}$$

General property for a successful error estimator.

$$C_1 \|e\| \leq \eta \leq C_2 \|e\|. \tag{43}$$

Where $\|e\|$ is the global error in energy norm.

Effectivity indices:

$$\text{global : } \frac{\eta}{\|e\|} \qquad \text{local : } \frac{\eta_k}{\|e\|} \tag{44}$$

This section reviewed issues of error estimation for the finite element method. Different types of error measures were discussed. Error estimators based on error norms or residuals are helpful in adaptive remeshing. To obtain an error estimator suited for further stochastic treatment, the procedures for error estimation need to be adapted or modified in order for the error estimate to be allowed to assume either positive or negative values. Error estimators are needed that represent directly the errors in the response quantity of interest, and not errors obtained from energy norms related to this quantity. Section 4.1 describes several types of error estimators adapted to give directly the desired information of error in stress for the elasticity problem.

3.2 Error Estimators in Boundary Element Methods (BEM)

This section is based on new error estimation and adaptivity approaches for BEM, presented in (Jorge, Ribeiro, et al., 2003), for 2D potential-BEM, and in (Jorge et al., 2005), for 2-D elastostatics BEM. These approaches were originally presented in the PhD Dissertation (Jorge, 2002) at Vanderbilt University.

3.2.1 Introduction

Boundary integral equation (BIE) formulations are exact representations of a boundary value problem (BVP), but errors may appear in the discretization process leading to the boundary element method (BEM). To evaluate these local errors, two error estimator approaches are presented. An adaptive mesh refinement procedure using these local error estimators is also presented.

To solve the BIE numerically, the boundary is divided into elements. The collocation boundary element method (BEM) is a discretization of the BIE so that the solution is only valid at a finite number of boundary collocation points. For the other boundary points, an approximation is written using interpolating functions. Several sources of error exist in BEM problems. First, the discretized boundary may not represent the original one anymore. Second, the polynomial interpolation is only an approximation of the variable inside the element. Third, the continuity of the boundary elements may not allow the

smoothness requirements for the boundary in the original BIE to be satisfied exactly. Fourth, numerical integration using Gaussian quadrature is only an approximation of the integral being evaluated. Local and global contributions for the error due to an element can be estimated so that refinements of the mesh of the discretized problem can be performed in order to minimize these errors. A global error estimator can be used to determine the need for refinement, while a local error estimator can be used to determine which portions of the boundary require refinement, in an adaptive process.

This section briefly gives some background on existing error estimation approaches and algorithms for the collocation Boundary Element Method (BEM). This brief review includes error estimators for BEM in two different representations: symmetric and collocation. Two different error estimator approaches for the collocation BEM are considered in detail, for 2-D potential problems. The first approach is a local error estimator based on a gradient recovery procedure, wherein the smoothed or recovered functions are rates of change of the boundary variables in the local tangential direction.

The gradient recovery approach as used in the Finite Element Method (Huebner et al., 2008) (Olgierd C. Zienkiewicz & Taylor, 1994) (Ainsworth & Oden, 1997) considers the evaluation of the gradient in global coordinates, allowing for continuous derivatives of the field variable in the fixed global directions. On the other hand, in the BEM the variables of interest are written in local coordinates, so that the flux is the boundary variable associated to the outward local normal derivative of the potential. The potential is continuous throughout the boundary, while the flux and the tangential derivative of the potential may be discontinuous, such as in the case where the boundary is not smooth. The recovery procedure of the tangential derivatives in BEM shall allow these discontinuities to appear in the formulation. The adopted approach was to include double nodes at corners for the recovery approach.

The second error estimator approach is associated to an external problem formulation and gives both a local and a global measure of the error, depending on an appropriate choice of the external point where the equation is applied. The potential should vanish at points located in the exterior of the domain due to the fact that the swept angle that appears in the boundary integral formulation is zero. Nonzero values indicate that the formulation is in error. By collocating the external point close to the boundary, the influence of the nearby elements in the error estimator will be dominant, so that the information on the error can be considered to be local.

These two error estimator approaches represent a post-processing step, after the BEM solution is already known. No information on the exact solution is required for both error estimators. The two error estimators are computed for a 2-D potential problem containing a singularity, so that a comparison is made between both error estimators for comparable boundary meshes. The consistency with mesh refinement is also compared for both error estimators.

3.2.2 Current Error Estimators for Collocation and Symmetric BEM

Several types of a posteriori error estimators for BEM have been proposed in the literature. A review can be found in (S. Liapis, 1994). These error estimators include various versions of element residual methods, flux projection estimators, extrapolation error estimators and estimators measuring the sensitivity of the numerical solution to a shift in the collocation point. In (Kelly et al., 1987) a global error estimator is proposed.

In (Liang et al., 1999) local error estimators are proposed based on dual BIE formulations. In (Paulino et al., 1996) and (Menon et al., 1999) a *pointwise* error estimator is proposed, based on the difference between two formulations, the standard formulation and the hypersingular formulation, obtained by differentiation of the first one. This approach is applied to Symmetric-Galerkin BEM formulation in (Glaucio H. Paulino & Gray, 1999). An error estimator for the direct collocation method in the elastostatics problem is proposed in (Denda & Dong, 1999). In this case, a comparison with an equivalent representation of the same problem, obtained for the exterior domain, is done. Upper and lower bounds of the error are found using a residual approach in (Jou & Liu, 1999). Nodal sensitivities are used in (G. H. Paulino et al., 1997). Other nodal sensitivities to variations in node location are found in (Shi et al., 1995) and (Guiggiani, 1996). Other error estimators are proposed in (Varfolomeyev et al., 1998) (Zhan & Yokoyama, 1997) (Karafiat, 1997) (Stergios Liapis, 1996) (Charafi & Wrobel, 1997) (Chao & Lee, 1999) (Zhao, 1998) (Ammons & Vable, 1998).

Some current approaches for error estimators or error indicators already available in other BEM formulations can be adapted or transposed to the collocation BEM formulation. Usually, the error estimator approaches deal with the discretization error, looking for an optimal adaptive procedure for mesh refinement, if a local error measure is available. The approaches for error estimators and adaptivity for the collocation BEM can be grouped as follows.

Gradient recovery

Using a local average approach, an updated or more accurate gradient could be obtained. This average could be an element average or a domain average. If a domain average is used, then the gradient could be evaluated in some interior region in the vicinity of the boundary, and then a limit to the boundary could be done. The difference between this refined gradient and the actual gradient as found for a certain mesh would be a measure of the error. The components of the gradient are the normal and the tangential derivatives. The tangential derivative is obtained in a post-processed manner, by taking the derivative of the shape functions with respect to the intrinsic variable ξ , while the normal derivative is written in terms of the shape functions themselves. For the usual polynomial shape functions, the tangential derivative is written in a representation that has one degree less than the normal derivative. So, another possibility, which would be simpler, would be the recovery of only the tangential derivative, so that the gradient components (normal and tangential) would now be written with shape functions of same order. A measure of the error could be the difference between the refined and the actual tangential derivatives. This simpler approach would be better adapted to the parts of the boundary where the tangential derivative varies very fast, and its representation, with shape functions of one degree less, is inaccurate. On the other hand, for parts of the boundary where the normal derivative is changing very fast, as could be the case, for example, if the radius of curvature of that part of the boundary is small, then the contribution of the normal derivative to the error might be significant and could not be neglected *a-priori*.

Average in the norm

The numerical solution is the exact solution plus some error. The error satisfies the equation modeling the problem. If numerical solution is not known, the numerical solution for a refined mesh could be adopted as the "exact" solution for the purpose of finding this difference. Errors in displacement and traction (in the elastostatics problem)

or in potential and flux (in the potential problem) can be evaluated for all nodal points. Local (in the element) and global (in the entire boundary) averages of this error could be written, in the L_2 norm sense or in some energy norm, and a local relative error in this norm could be obtained. This relative error could be used to guide adaptive mesh refinement. The error in one variable (say, traction or flux) could be dominant with respect to the error on the other variable (say, displacement or potential). Another procedure could be to first find the error in the tangential derivatives from the error in displacement (or in potential), and then evaluate the error in some norm sense, the norm being some average of the tangential and normal derivatives. The same as before, local (element) and global measures of the error shall be obtained, so that a relative error is assessed in the element. This alternative procedure could be adopted in conjunction with the tangent derivative recovery, to have both tangential and normal derivatives written with shape functions of the same order, before the error norm is evaluated.

Sensitivity analysis

A sensitivity analysis could be performed in two ways, leading to two different sensitivity measures. The first measure is related to regularization sensitivity. The singular or hypersingular kernels in certain BIEs can be regularized before a numerical procedure to solve the BIE is implemented. Simple solutions may be imposed in this regularization procedure to obtain self-regular formulations for the BIEs. The sensitivity measure would be on the response to variations in the simple solution being imposed in the self-regular formulation of the integral representation of the problem. If these simple solutions are considered as a truncated Taylor series expansion about some point, then small variations about this point could be considered when higher order terms were included in this series expansion. For example, for the self-regular potential-BIE formulation, with local coordinates ζ_i , the general format of the variation about the simple constant potential solution would be to include the Taylor expansion first order term, which means to substitute $\phi(x)$ by $\left[\phi(x) + \frac{\partial \phi}{\partial \zeta_i} \Big|_x \zeta_i \right]$.

Similarly, in the self-regular gradient-BIE formulation, the variation about the simple linear potential solution would be to include the Taylor expansion second order term, which means to substitute $\left[\phi(x) + \frac{\partial \phi}{\partial \zeta_i} \Big|_x \zeta_i \right]$ by $\left[\phi(x) + \frac{\partial \phi}{\partial \zeta_i} \Big|_x \zeta_i + \frac{1}{2} \frac{\partial^2 \phi}{\partial \zeta_i^2} \Big|_x \zeta_i^2 \right]$.

The variations in the self-regular displacement and traction formulations in the elastostatics problem would have similar approaches.

The second measure is related to nodal sensitivity. The sensitivity of the response to variations of location of nodes could be evaluated. Small variations of the position of the central node could be done. To evaluate this sensitivity, two approaches are possible: allowing the element size to change (Lagrangian approach) or imposing the change in position of the node to occur in the same deformed shape of the element, i.e. without changing its length (Eulerian approach).

Difference between two equivalent formulations

The exact solution for the BVP as written in one integral representation shall satisfy any other possible different integral representation of the same BVP. When a numerical solution is found for some discretization of one formulation, this solution can be plugged into the discretized form (with same discretization) of the other possible formulation. As

the numerical solution is not exact, the identities in the discretized form of this second formulation will not be satisfied. The difference will be a measure of the error in the discretization of both formulations. For example, for a closed domain, the exterior problem representation is equal to zero in both the potential and in the elastostatics problems. So, any difference from zero could be related to the discretization error. As the fundamental solutions are functions of the distance between source and field points, one representation in which the exterior point is very close to the boundary might give a different error measure than another external formulation in which the exterior point is considered far away from the boundary. If different exterior points are considered, different external formulations may be obtained. The error measure might be sensitive to the external point position, for some particular problem. Another possibility is to obtain an equivalent formulation by taking derivatives of an existing formulation. Derivatives taken in different directions are all different equivalent formulations of the same problem. All scalar equations that are components of the vector equation obtained in some representation, are different equivalent formulations and can be used for this error measure purpose. This is the case in the gradient formulation, which was transformed into the flux formulation by taking the component of this equation in one direction. So, the component of this equation in the other direction (say, tangential) is an equivalent formulation that can be used. The same applies to the various components of the vector equation in the elasticity problem.

Various equivalent formulations could be used, for some specified boundary discretization. For example, for the potential problem, a comparison could be done between the potential-BIE and the gradient-BIE formulations. As the gradient-BIE formulation is a vector equation, one component of this equation, say, the flux equation, could be retained. The decomposition of the gradient equation in a different direction will lead to a different, but equivalent, equation, because the derivative of the potential equation was taken in a different direction. Other possible comparisons could be between the potential formulation and potential formulation for an exterior problem, or between the gradient formulation and the gradient formulation for an exterior problem. Again, as the gradient formulation is a vector equation, all scalar components of this equation are valid equivalent formulations.

Similar comparisons for the elastostatics problem could also be done, for example, between the displacement-BIE and the traction-BIE formulations. As the traction formulation is a vector equation (the projection of the tensor of stresses in one direction) any component of this equation can be retained. The traction equation in a different direction will lead to a different, but equivalent, equation, because the derivative of the displacement equation was taken in a different direction. Other possible comparisons could be between the displacement formulation and the displacement formulation for an exterior problem, or between the traction formulation and the traction formulation for an exterior problem. As before, as the traction formulation is a vector equation, all scalar components of this equation are valid equivalent formulations.

Adaptivity

When a local measure of the discretization error, in one element, is available, it could be compared to an average global measure of this error. If the relative error is bigger than 1 for a certain element, a refinement shall be applied to this element. Adaptive refinement of the mesh could be done by adopting either h -refinement or p -refinement procedures.

In the h -refinement procedure, the element where the local error is big is divided into two small elements (usually it is divided into two equal elements). In the p -refinement procedure, the degree of the shape functions used to represent the variables in this element is increased. For this p -refinement procedure, the use of hierarchical shape functions would be advantageous, because computations already done for the coarse mesh will be kept and don't need to be done again in the refined mesh.

3.2.3 Proposed Error Estimators for Potential-BEM

Two error estimators for BEM are proposed, one based on the gradient recovery approach and the other based on the error on the external formulation. These error estimators are post-processing approaches that give qualitative indicators of the local errors on the boundary solution for any BEM formulation and for a certain boundary discretization into a particular mesh. These error estimators are suitable to be used for the collocation BEM, where other error estimators such as the ones obtained using energy norms may not be available.

3.2.3.1 The gradient recovery approach

The following discussion will be limited to potential theory in 2D, but the ideas can be adapted to the 3D problem and to the elasticity theory. The idea of recovery of derivatives is borrowed from error estimators for finite element methods, where to compute some error indicators it was necessary first to obtain expressions for the derivatives of the finite element solution. Even small errors in the finite element solution could lead to large errors or inaccuracy in the derivatives of the solution, evaluated in a post-processing manner by taking derivatives of the shape functions. Smoothed or refined approximations for the derivatives were developed so that the error indicators depending on these derivatives would be more credible. A procedure to smooth the gradient to compute these derivatives with good accuracy is called gradient recovery (Huebner et al., 2008) (Olgiard C. Zienkiewicz & Taylor, 1994) (Ainsworth & Oden, 1997). The derivatives are evaluated from a weighted average of element derivatives surrounding the node. The weighting factors are the element areas. The FEM procedure is illustrated for a one-dimensional problem, for simplicity. For each element (e), the first derivatives are first expressed in terms of the derivatives of the shape functions N and of the nodal values of the potential solution $\{\phi\}^{(e)}$ as:

$$\left(\frac{\partial\phi}{\partial x}\right)^e = \left[\frac{\partial N}{\partial x}\right]\{\phi\}^{(e)} \quad (45)$$

The derivatives are now written in an average sense

$$\int_{\Omega^{(e)}} \{N\}[N]dA \left\{\frac{\partial\phi}{\partial x}\right\} = \int_{\Omega^{(e)}} \{N\}dA \frac{\partial\phi^{(e)}}{\partial x} \quad (46)$$

where the goal is to find the vector of smoothed or recovered nodal derivatives $\{\partial\phi/\partial x\}^e$. For that, an equivalence equation can now be written in a region $\Omega^{(e)}$ surrounding an element (e) as:

$$\int_{\Omega^{(e)}} \{N\} [N] dA \left\{ \frac{\partial \phi}{\partial x} \right\} = \int_{\Omega^{(e)}} \{N\} dA \frac{\partial \phi^{(e)}}{\partial x} \tag{47}$$

The element equations as in Eq. (47) are then assembled for the entire domain Ω . A system of equations is obtained from which the vector of smoothed nodal derivatives $\{\partial \phi / \partial x\}^{(e)}$ can be obtained. An important feature in this procedure is that the x -coordinate belongs to a global coordinate system, so that not only the potential field at the nodes $\{\phi\}^{(e)}$ is continuous in Ω , but also all the elements of the smoothed gradient components $\{\partial \phi / \partial x\}^{(e)}$ vector in a general problem are continuous in Ω too, so long as linear or higher order elements are used.

This post-processing approach to find error estimators from gradient recovery can be adapted to a BEM problem, after the solution is known for all nodes on the boundary. Consider first an element (e) of the boundary region in which the gradient is being recovered. Let ϕ be the potential, $\partial \phi / \partial s$ be the tangent derivative of ϕ and $q = \partial \phi / \partial n$ be the flux. From the BEM solution, the nodal values $\phi_i^{(e)}$ and $q_i^{(e)}$ are known. Both the flux q and the tangential derivative $\partial \phi / \partial s$ are allowed to be discontinuous at boundary points where the normal unit vector is not unique. The tangent and normal unit vectors are uniquely defined only locally in the interior of a boundary element. Usual BEM solutions allow for these discontinuities to occur, and the flux solution on the boundary is given both as a function of the node number i and of the element number (e) so that on a certain element i the flux on the element before the node $q_i^{(e)}$ might be different from the flux for the element after the node $q_i^{(e+1)}$. The same discontinuities will apply for both the original and the recovered or smoothed values of the tangential derivative of the potential $\partial \phi / \partial s$ at the same boundary nodes where the flux is discontinuous.

To be able to write a system of equations similar to the one obtained with the FEM procedure as outlined above, continuity would be required for the derivatives of the potential. But the derivatives of the potential would be continuous only in a global coordinates system, and not in local coordinates systems along the boundary curve, if the boundary has corners. The approach adopted in this work was to introduce double nodes at all the discontinuity points, such as the corner nodes on the boundary. Elements with zero length connect these double nodes, so that no local or global contribution for the error from these fictitious elements is obtained. It must be pointed out that the gradient recovery approach is a post-processing procedure that is completely independent from the BEM method that was used to obtain the boundary solution. The original BEM method may or may not use double nodes, and this is irrelevant for the post-processing procedure, because the double nodes being introduced here came from the fact that discontinuities exist in the values of q and $\partial \phi / \partial s$ on these corner nodes locations, and not due to the original BEM procedure to evaluate the boundary solution, that remains unchanged.

The solution at any point inside the element is $\phi^{(e)}(\xi) = \sum_{i=1}^m N_i(\xi) \phi_i^{(e)}$ and $q^{(e)}(\xi) = \sum_{i=1}^m N_i(\xi) q_i^{(e)}$, where ξ is the intrinsic coordinate and m is the number of element nodes corresponding to the polynomial interpolation being used, such as $m = 2$ for linear elements, $m = 3$ for quadratic elements, etc. Taking derivatives with respect to the intrinsic coordinate gives

$$\frac{d\phi^{(e)}(\xi)}{d\xi} = \sum_{i=1}^m \frac{dN_i(\xi)}{d\xi} \phi_i^{(e)} \quad ; \quad \frac{dq^{(e)}(\xi)}{d\xi} = \sum_{i=1}^m \frac{dN_i(\xi)}{d\xi} q_i^{(e)} \quad (48)$$

These derivatives can now be written with polynomial shape functions one degree higher as

$$\begin{aligned} \frac{d\phi^{(e)(*)}(\xi)}{d\xi} &= \sum_{i=1}^m N_i(\xi) \frac{d\phi_i^{(e)(*)}}{d\xi} \quad ; \quad \frac{dq^{(e)(*)}(\xi)}{d\xi} \\ &= \sum_{i=1}^m N_i(\xi) \frac{dq_i^{(e)(*)}}{d\xi} \end{aligned} \quad (49)$$

where $d\phi_i^{(e)(*)}/d\xi$ and $dq_i^{(e)(*)}/d\xi$ are the sets of nodal values of the corresponding derivatives, equivalent in a local averaging sense, to the original ones, $d\phi_i^{(e)}/d\xi$ and $dq_i^{(e)}/d\xi$. The element equivalence is enforced by setting that both formulations for the derivatives give same element integrals, defined by

$$\begin{aligned} I_\phi^{(e)} &= \int_S^{(e)} \{N\} \frac{d\phi_i^{(e)}(\xi)}{d\xi} J(\xi, s) dS = \int_S^{(e)} \{N\} \frac{d\phi_i^{(e)(*)}(\xi)}{d\xi} dS \\ I_q^{(e)} &= \int_S^{(e)} \{N\} \frac{dq_i^{(e)}(\xi)}{d\xi} J(\xi, s) dS = \int_S^{(e)} \{N\} \frac{dq_i^{(e)(*)}(\xi)}{d\xi} dS \end{aligned} \quad (50)$$

where $\frac{d\phi_i^{(e)}(\xi)}{d\xi} J(\xi, s) = \frac{\partial \phi_i^{(e)}(s)}{\partial s} J(\xi, s)$, $J(\xi, s)$ being the Jacobian of the transformation between the real element (with coordinate s) and the standard element (with the intrinsic coordinate ξ). The derivatives of $\phi^{(e)(*)}$, $q^{(e)}$ and $q^{(e)(*)}$ follow similar transformation of coordinates. The integration is performed on S which is the element length in 2-D or the element area in 3-D. If the element for integration is now written as $dS = J(\xi, s)d\xi$, all terms in the integrands are now only function of the intrinsic variable ξ , and all element integrations are now performed for $\xi \in [-1,1]$, for the usual Lagrangian interpolating functions.

The element equations as in Eq. (50) are then assembled for the entire boundary S . A system of equations is obtained from which the vectors of smoothed nodal derivatives $\{\partial \phi_i / \partial s\}^{(e)}$ and $\{\partial q_i / \partial s\}^{(e)}$ can be obtained. It must be pointed out that the system matrix for the coefficients of the recovered or smoothed quantities is not singular, and admits an inverse, after all elements are assembled, including the zero-length elements that link the double nodes.

Several approaches could be envisaged to use the assembly of the element integrals as obtained in Eq. (50). By assembling the element integrals of the type $I_\phi^{(e)}$ only, an expression for the local error related to the tangent derivative of ϕ could be obtained. Similarly, by assembling the integrals of the type $I_q^{(e)}$ only, an expression for the local error related to the tangential derivative of the flux q could be obtained, noting that the flux is the normal derivative of the potential ϕ . A more elaborate expression for the error related to both the tangential and the normal derivatives of the potential could be obtained by combining both integrals. In this case, normalizing or scaling the integrals first, before

adding or combining them, might be necessary, so that they will be of the same order of magnitude.

In any of these approaches, the error is already normalized with respect to the size of the element, the normalization factor being the Jacobian $J(\xi, s)$ that appeared naturally in the formulation when the chain rule was applied for the derivatives of the transformation equations from local variables to intrinsic element variables. So the obtained the quantity of interest is a measure of the relative error, already normalized with respect to the element size, and not a measure of the error that could be dependent on the size of the element.

The approach follows several steps. The first step is to solve the system of equations corresponding to the integral (or to the combination of integrals) chosen, obtaining the equivalent set of nodal values for the derivatives. For any point with generic coordinate S in the element, a difference will exist between the derivative as written in the original system and the derivative using the recovered equivalent nodal values. By construction, these differences will add to zero if the entire boundary is considered, because adding for all positions S is the same as integrating in the entire boundary, obtaining again the equivalence equation that was enforced. A quadratic expression for a residual is proposed such that the differences between the derivatives point by point are first squared and then added to obtain an expression for the local error. The proposed residuals are, then, functions of the position S within the element:

$$r_\phi^{(e)}(s) = \left[\frac{\partial \phi^{(e)(*)}}{\partial s}(s) - \frac{\partial \phi^{(e)}}{\partial s}(s) \right]^2 ; \quad r_q^{(e)}(s) = \left[\frac{\partial \phi^{(e)(*)}}{\partial s}(s) - \frac{\partial q^{(e)}}{\partial s}(s) \right]^2 \quad (51)$$

Local error expressions in the original elements can now be defined as:

$$E_\phi^{(e)} = \frac{1}{S^{(e)}} \int_{S^{(e)}} r_\phi^{(e)}(s) dS^{(e)} \quad ; \quad E_q^{(e)} = \frac{1}{S^{(e)}} \int_{S^{(e)}} r_q^{(e)}(s) dS^{(e)} \quad (52)$$

where $S^{(e)}$ is the total size, or length, of the element, appearing as the scaling factor equivalent to the Jacobian that would appear if the integration were performed with respect to the intrinsic variable ξ .

The residuals in an element (e) at the position s , obtained from the two sources of error, $r_\phi^{(e)}(s)$ and $r_q^{(e)}(s)$ can be combined to give a unique information about the local error in the element, coming from both the tangential and the normal derivatives of the potential.

As the order of magnitude of the quantities involved may vary, these residuals are proposed to be normalized with respect to the average values of the quantities in the element, before they are added to obtain a unique residual, now normalized, for the element.

The average values can be obtained directly from the equivalence equation, Eq. (50) as

$$\begin{aligned} \left. \frac{\partial \phi^{(e)}}{\partial s} \right|_{AVE} &= \frac{1}{S^{(e)}} \int_{S^{(e)}} \frac{\partial \phi_i^{(e)}}{\partial s}(s) dS^e = \frac{1}{S^{(e)}} \int_{S^{(e)}} \frac{\partial \phi_i^{(e)(*)}}{\partial s}(s) dS^e \\ \left. \frac{\partial q^{(e)}}{\partial s} \right|_{AVE} &= \frac{1}{S^{(e)}} \int_{S^{(e)}} \frac{\partial q_i^{(e)}}{\partial s}(s) dS^e = \frac{1}{S^{(e)}} \int_{S^{(e)}} \frac{\partial q_i^{(e)(*)}}{\partial s}(s) dS^e \end{aligned} \quad (53)$$

The normalized residuals are then defined as:

$$r_{\phi}^{(e)}(s)|_{norm} = \left[\frac{\frac{\partial \phi^{(e)(*)}}{\partial s}(s) - \frac{\partial \phi^{(e)}}{\partial s}(s)}{\frac{\partial \phi^{(e)}}{\partial s}|_{AVE}} \right]^2 ; \quad r_q^{(e)}(s)|_{norm} = \left[\frac{\frac{\partial \phi^{(e)(*)}}{\partial s}(s) - \frac{\partial q^{(e)}}{\partial s}(s)}{\frac{\partial q^{(e)}}{\partial s}|_{AVE}} \right]^2 \quad (54)$$

so that a single residual for the element (e) at the position s and the resulting local error element can now be defined as

$$r^{(e)}(s) = r_{\phi}^{(e)}(s)|_{norm} + r_q^{(e)}(s)|_{norm} \quad ; \quad E^{(e)} = \frac{1}{S^{(e)}} \int_{S^{(e)}} r^{(e)}(s) dS^{(e)} \quad (55)$$

It has to be pointed out that the residuals were normalized with respect to the average values of the corresponding variables, and also with respect to the size of the elements, so that these residuals are now related only to the local error due to the corresponding source, either the tangential derivative or the normal derivative of the potential.

It may occur that one source of error might be dominant, so that the other residual could be negligible in a certain problem. Numerical experiments have to be performed to establish relationships between these residuals, or conditions in which one of them, if any, could be discarded without loss in the accuracy of the error estimator.

In the current work, only the error obtained from the recovery of the tangential derivative of the potential, denoted as $Error_{\phi}$, is evaluated. Thus, in this gradient recovery approach, the smoothed or recovered function is not the gradient itself, but its component in the locally tangent direction, because this component is not obtained from the integral equation directly, but from the tangential derivative of the potential.

The reason for not considering the error from the recovery of the flux in this first study is that the tangential derivative of the potential is being approximated, using isoparametric elements, by a polynomial that is one degree less than the polynomial that approximates the normal derivative of the potential, which was obtained directly in the solution of the boundary integral equation. The smoothed tangential derivative and the original normal derivative are, then, approximated using polynomials of the same degree.

The recovery procedure could be refined in a future work, so that the rate of change in tangential direction of the normal derivative could be smoothed too. With this refinement in the gradient recovery procedure, a comparison could be performed between the normal and tangential contributions for the element error. This refinement is proposed but not implemented in this work, as the tangential derivative of the potential, being written with a polynomial with one degree less than the flux, appears to be the component of the gradient that gives the most significant contribution for the total element error. (Jorge, Ribeiro, Cruse, et al., 2001) shows that accurate results for the potential and the flux were obtained for coarse graded meshes where the tangential derivatives of the potential were better approximated.

A global error estimator can now be defined simply by $E_{Global} = \sum_{e=1}^N E^{(e)}$, where N is the number of boundary elements, so that an average error for the entire boundary and the normalized local element errors can be defined respectively as:

$$E_{AVE} = \frac{1}{N} E_{Global} = \frac{1}{N} \sum_{e=1}^N E^{(e)} \quad ; \quad E^{(e)}|_{norm} = \frac{E^{(e)}}{E_{AVE}} \quad (56)$$

so that this normalized local error indicator can now be used in an adaptive procedure. If $E^{(e)}|_{norm} > 1$ then the local error in the element is bigger than the average, and a mesh refinement is needed, either changing the size of the elements in this boundary region to smaller ones (h - refinement) or increasing the order of the polynomial interpolant in the element (p - refinement). If $E^{(e)}|_{norm} \cong 1$, the local error has an average value and no refinement is needed. If $E^{(e)}|_{norm} < 1$, no refinement is needed, either. In this case, in fact, the local error being small shows that the element size is already small enough and it could be made bigger without much loss in the solution accuracy.

3.2.3.2 The external domain approach

Usually, the error estimator approaches deal with the discretization error, looking for an optimal adaptive procedure for mesh refinement, if a local error measure is available. Existing ideas, approaches and error estimators for collocation BEM are the starting point for the error estimators to be proposed. From these ideas, the difference between two equivalent formulations seems to be one of the easiest to be adopted, due to its simplicity. The exact solution for the BVP as written in one integral representation shall satisfy any other possible different integral representation of the same BVP. When a numerical solution is found for some discretization of one formulation, this solution can be plugged into the discretized form (with same discretization) of the other possible formulation. As the numerical solution is not exact, the identities in the discretized form of this second formulation will not be satisfied. The difference will be a measure of the error in the discretization of both formulations.

For example, for a closed domain, the exterior problem representation is equal to zero in both the potential and in the elastostatics problems. So, any difference from zero could be related to the discretization error. As the fundamental solutions are functions of the distance between source and field points, one representation in which the exterior point is very close to the boundary might give a different error measure than another external formulation in which the exterior point is considered far away from the boundary. If different exterior points are considered, different external formulations may be obtained. The error measure might be sensitive to the external point position, for some particular problem. Two equivalent formulations can be used, for some discretization, for the potential problem: evaluate the solution in the boundary with a potential formulation and then check the expected vanishing of the results of the potential formulation for an exterior problem.

So, an error estimator is implemented for the potential problem, based on the fact that the potential vanishes at external points. In the potential formulation for an exterior problem, a numerical solution at the boundary is first obtained for the discretized mesh using the collocation BEM or any other BEM formulation. This boundary solution is then substituted on the exterior form of Green's identity given by

$$0 = - \int_S \phi(s) \vec{\nabla} \ln \left(\frac{1}{r(s,y)} \right) \cdot \vec{n}(s) dS + \int_S \vec{\nabla} \phi(s) \cdot \vec{n}(s) \ln \left(\frac{1}{r(s,y)} \right) dS \quad \forall y \in \bar{R} \quad (57)$$

As it was discussed in Section 3.2.2, other choices of integral representations could also have been adopted for the exterior equation. By implementing this methodology to other exterior problem equations, optimal choices for this equation could be looked at.

The exterior point $y \in \bar{R}$ is proposed to be located at a distance to the boundary (the distance from the exterior point to the closest point on the boundary) that is $1/4$ of the size of the element closest to this exterior point. This distance is measured in the direction normal to the element. This distance was chosen to avoid the quasi-singularity problem that may appear if the external point is too close to the boundary and also to avoid round-off error that may appear in the computations of very small fundamental solutions when the exterior point is too far away from the boundary. It has to be noted that other choices for this position could have been adopted, and numerical experiments could have been performed to determine optimum minimal distances for different cases.

3.2.3.3 Implementation of the proposed Error Estimators

A numerical example is performed to demonstrate the implementation of the two proposed error estimators (Jorge, Ribeiro, & Fisher, 2001). This example was proposed by Motz (Jaswon et al., 1978). The problem consists of a rectangular domain, as shown in Fig. 23 (a). The flux is singular at point O in the element where the potential is prescribed.

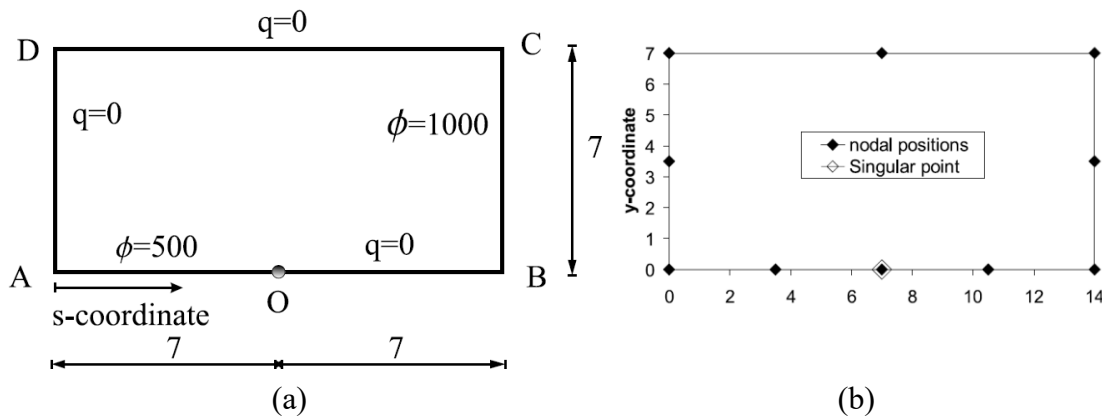


Figure 23: The Motz problem: (a) a singularity is created at point O by the boundary conditions; (b) the coarse mesh: 10 elements - adapted from (Jorge, Ribeiro, et al., 2003)

All BEM results were obtained with the code BETIS from (París & Cañas, 1997), which uses linear elements. Two meshes are implemented to compare the error estimators. The first mesh is coarse and has only 2 elements in each linear segment in which a specific boundary condition is prescribed. Then, a finer mesh is studied in which there are eight elements on the side where the flux is known to have a discontinuity. Figure 23 (b) illustrates the coarse mesh that was implemented.

The implementation of the tangent derivative recovery for linear elements is detailed as follows. A standard linear element has two end nodes with corresponding shape functions $N_1(\xi) = (1 - \xi)/2$ and $N_2(\xi) = (1 + \xi)/2$ for $\xi \in [-1,1]$. The Jacobian for this 1-D straight element is $J(\xi, s) = \partial\xi / \partial s = 2/L^{(e)}$, where $L^{(e)}$ is the length of the boundary element (e) , so the element of integration is $dS^{(e)} = (L^{(e)}/2)d\xi$. The original and the recovered (denoted by $*$) tangent derivatives of the potential are written for the element (e) so that the equivalence in the element can be enforced as follows.

$$\int_{-1}^1 \begin{Bmatrix} N_1(\xi) \\ N_2(\xi) \end{Bmatrix} [N_1(\xi) \quad N_2(\xi)] \frac{L^{(e)}}{2} d\xi \begin{Bmatrix} \frac{\partial \phi_1^{(e)(*)}}{\partial s} \\ \frac{\partial \phi_2^{(e)(*)}}{\partial s} \end{Bmatrix} \tag{58}$$

$$= \int_{-1}^1 \begin{Bmatrix} N_1(\xi) \\ N_2(\xi) \end{Bmatrix} \frac{L^{(e)}}{2} d\xi \frac{2}{L^{(e)}} [N_1'(\xi) \quad N_2'(\xi)] \begin{Bmatrix} \phi_1^{(e)} \\ \phi_2^{(e)} \end{Bmatrix}$$

Now all element matrices given by Eq. (58) can be assembled for both the coarse and the refined meshes. By solving the assembled system of equations, the vector of nodal values for the smoothed tangential derivatives of the potential $\partial \phi_i^{(*)} / \partial s$ is obtained. The assembled matrices are sparse. The vector of element interpolations is now obtained as

$$\left\{ \frac{\partial \phi^{(i)(*)}}{\partial s}(\xi) \right\} = \left(\frac{1 - \xi}{2} \right) \frac{\partial \phi_i^{(*)}}{\partial s} + \left(\frac{1 + \xi}{2} \right) \frac{\partial \phi_{i+1}^{(*)}}{\partial s} \tag{59}$$

The residuals r_ϕ^i are now evaluated for each element as the square of the differences between the smoothed solution obtained in Eq. (59) and the original non-smoothed solution.

$$r_\phi^i(\xi) = \left[\frac{\partial \phi^{(i)(*)}}{\partial s}(\xi) - \frac{2}{L^i} [N_i'(\xi) \quad N_{i+1}'(\xi)] \begin{Bmatrix} \phi_i \\ \phi_{i+1} \end{Bmatrix} \right]^2 \tag{60}$$

so the element errors R_ϕ^i , the average error E_{AVE} and the normalized element errors $E_\phi^i = R_{\phi_{norm}}^i$ are found as

$$R_\phi^i = \int_{-1}^1 r_\phi^i(\xi) d\xi \quad E_{AVE} = \frac{1}{N} \sum_i R_\phi^i \quad E_\phi^i = \frac{R_\phi^i}{E_{AVE}} \tag{61}$$

where the element errors E_ϕ^i can now be compared with the local element errors E_{EXT}^i , obtained from the external error approach. Figure (24) shows the results for the normalized local error estimators for both the coarse and the fine meshes. The horizontal axis of the four plots corresponds to the s- coordinate of the middle of each element along the boundary.

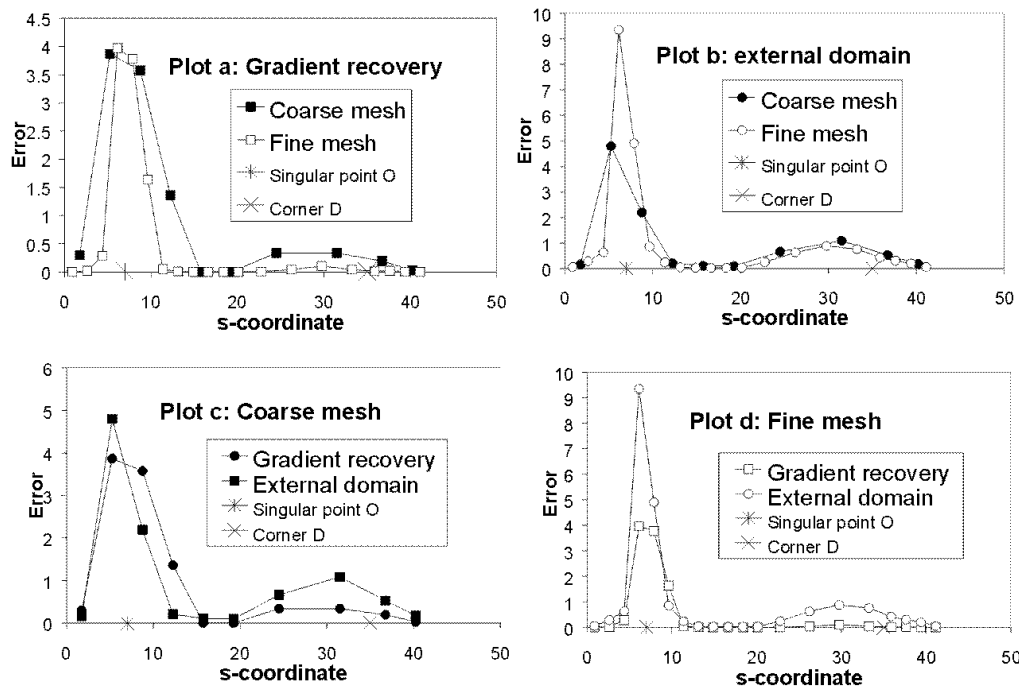


Figure 24: Error estimator results. (a) and (b): Influence of mesh refinement for each estimator. (c) and (d): Comparison between estimators for each mesh. - adapted from (Jorge, Ribeiro, & Fisher, 2001)

Plot (a) shows the results for the error on the tangential derivative of the recovered gradient $Error_{\phi}$ while plot (b) shows the results for the error obtained using the exterior domain formulation $Error_{EXT}$. The boundary region that shows the highest values for the normalized error in both cases is the region that contains the elements adjacent to the singular point O, as expected. The error results were qualitatively consistent with the mesh refinement. Plots (c) and (d) compare the error estimators for the coarse and the refined meshes, respectively. Both error estimators indicate the same region about the singular point O with highest values for the normalized error.

Also, the boundary region immediately before corner node D presents relatively high secondary errors for both estimators. These errors appear to be related to the fact that this region corresponds to the boundary corner the farthest distance away from any boundary region where the potential is prescribed.

It must be pointed out that the element before the singular point has the potential prescribed as constant. So, any contribution for the error could only come from the error in the flux, in this element. Similarly, the element after the singular point has the flux prescribed as constant. So, any contribution for the error could only come from the error in the potential, in this element. This means that both error estimators were capable of recovering the local information of the total error regardless of the cause of the error being the boundary potential or the boundary flux.

Although the adopted gradient recovery procedure does not explicitly recover the errors from the tangential derivative of the flux, high values for the error that could only be related to the local flux were encountered in an element adjacent to the singular point. An

explanation for this behavior is that the gradient recovery procedure is based on an algorithm that includes all the boundary elements. The recovery of the tangential derivative of the potential is not made independently for each element, in a local basis. Instead, a unique value for the smoothed or recovered nodal values of the tangential derivative of the potential is only obtained when all the element matrices are assembled and the solution for the obtained system of equations is calculated. The redistribution of the tangential derivatives of the potential throughout the boundary appears to implicitly account for the local error in the flux.

3.2.4 Proposed Error Estimators for Elastostatics-BEM

New Error Estimators based on Gradient Recovery and External Domain Approaches were extended to 2D Elastostatics Problems in (Jorge et al., 2005). These error estimator approaches were derived for 2D potential problems in (Jorge, Ribeiro, et al., 2003), and the extension of these error estimators for a problem with a different integral formulation highlights their generality of use.

Error estimators based on residuals and error norms, similar to the approaches available in the finite element method (FEM), have been used for symmetric BEM formulations, but the error estimators literature is scarce for the collocation BEM formulation. Reviews for error estimation and adaptivity methods for BEM can be found in (S. Liapis, 1994) and in (Kita & Kamiya, 1994) (Kita & Kamiya, 2001).

The two local error estimator approaches presented in this work can be used for both Galerkin and collocation BEM formulations. The first approach involves a gradient recovery procedure, wherein the smoothed or recovered functions are rates of change of the boundary variables in the local tangential direction. The gradient recovery approach is commonly used in the finite element method (Ainsworth & Oden, 1997) and considers the evaluation of the gradient of the field variable in global coordinates, with continuous derivatives of this field variable in fixed global directions.

On the other hand, in the BEM the variables of interest are written in local coordinates, the displacements are continuous throughout the boundary, while the tractions and the tangential derivatives of the boundary variables may be discontinuous, such as at points where the boundary is not smooth or where the boundary conditions are discontinuous.

The second error estimator approach is associated with an external problem formulation and gives both a local and a global measure of the error depending on the external point where the equation is applied. (Denda & Dong, 1999) presented an error estimator for elasticity based on the strain energy in the exterior region. In the current work, the error estimator is obtained using directly the error in the displacements and in the stresses obtained from the exterior problem. Displacements and stresses should vanish at points located in the exterior of a closed domain, and consequently, nonzero values indicate errors. If the external evaluation point is near the boundary, the influence of the nearby boundary elements in the error estimate will be dominant, and the error information assumes a local characteristic. To eliminate any influence of the choice of the global directions on the values of the stress error estimates, the stresses at the exterior point are first combined into stress invariants, the mean stress and the octahedral stress, before any error information is evaluated. In (Jorge et al., 2005), the two error estimators were computed for two 2D elastostatics problems, one of which contains a singularity. A comparison was made between global and local error estimates, and the error consistency

with mesh refinement is also compared. Some details are shown in the following subsections.

3.2.4.1 The gradient recovery approach

The concept of recovery of derivatives for error estimation originated with finite element methods, in which error indicators were derived from derivatives of the computed finite element solution. A procedure to smooth the gradient to compute these derivatives with good accuracy is called gradient recovery (Ainsworth & Oden, 1997). The derivatives are evaluated from a weighted average of element derivatives surrounding a node. The weighting factors are the element areas.

This post-processing approach to estimate error from gradient recovery was adapted by the authors to 2D potential BEM problems in (Jorge, Ribeiro, et al., 2003). An important feature in this procedure when applied to FEM is that the derivatives of the field variables (displacements) are taken with respect to a global coordinate system. Thus, not only are the field variables continuous at the domain nodes, but also the smoothed gradient vectors in a general elastostatics problem are continuous. On the other hand, when the gradient recovery procedure is applied to BEM, the tangential and normal directions are written in local coordinates, and discontinuities in the tangential derivatives need to be allowed in both the original and the recovered tangential derivatives of the density fields (boundary displacements and tractions).

To write a system of equations similar to the one obtained with the FEM procedure, continuity would be required for the derivatives of the densities. The approach adopted in (Jorge, Ribeiro, et al., 2003) and extended in this work is to introduce double nodes at all the discontinuous points, such as the corner nodes on the boundary. Elements with zero length connect these double nodes, so that no local or global contribution for the error from these fictitious elements is incurred. The concept of double nodes to account for discontinuities in the tractions when obtaining the boundary solution is well known in the BEM literature (París & Cañas, 1997) but is not used in this work.

The double nodes introduced here originate from discontinuities in the values of the tangential derivatives of the densities at corner nodes in the post-processing algorithm to obtain the element errors and not from the original BEM procedure to evaluate the boundary solution.

The basic boundary integral equation for the elastostatics problem (Wendland & Cruse, 1990), gives the displacement $u_j(x)$ as a function of boundary integrals. This displacement-BIE, can be written, for an homogeneous domain Ω as:

$$C_{ji}(p)u_i(p) = \int_{\partial\Omega} \Psi_{ji}(p, Q)t_i(Q)dS(Q) - \int_{\partial\Omega} T_{ji}^\psi(p, Q)u_i(Q)dS(Q) \quad (62)$$

where $Q \in \partial\Omega$. The functions Ψ_{ji} and T_{ji}^ψ are associated with fundamental solutions of the Navier equation, and C_{ji} is associated with the location of the collocation point p , as an interior, exterior, or boundary point. The integrals in Eq. (62) provide a constraining equation for the boundary variables when evaluated for $p \in \partial\Omega$, in which case C_{ji} depends on the boundary smoothness at the point.

Four different densities must be considered when writing a boundary integral equation for 2D elastostatics: the two components of the displacements, $u_1 = u$ and $u_2 = v$, and

the two components of the tractions, $t_1 = SG$ and $t_2 = TAU$. In this work, recovery of the tangential derivatives of each density is made independently, so that separate element error information is obtained for each of the four boundary densities. In a later step, the element residuals are combined into element errors that contain information from the recovery of the tangential derivatives of all four densities.

In what follows, a boundary density is denoted by ϕ ($\phi = u, v, SG$ or TAU) and its tangential derivative by $\partial\phi/\partial s$. The solution at any point inside the element (e) is

$$\phi^{(e)}(\xi) = [N]\{\phi\}^{(e)} = \sum_{i=1}^m N_i(\xi)\phi_i^{(e)}$$

where ξ is the intrinsic coordinate and m is the number of element nodes corresponding to the polynomial interpolation being used. The symbols $[\dots]$ and $\{\dots\}$ denote row and column vectors, respectively. Taking derivatives with respect to the intrinsic coordinate gives

$$\left(\frac{\partial\phi}{\partial s}\right)^{(e)} = \left(\frac{\partial\phi}{\partial\xi}\right)^{(e)} \left(\frac{\partial\xi}{\partial s}\right)^{(e)} = \left[\frac{\partial N}{\partial\xi}\right]\{\phi\}^{(e)} \left(\frac{\partial\xi}{\partial s}\right)^{(e)} \tag{63}$$

where $J(\xi, s) = (\partial\xi/\partial s)^{(e)}$ is the Jacobian of the transformation between the real element (with coordinate s) and the standard element (with the intrinsic coordinate ξ). These derivatives can be written with polynomial shape functions of one degree higher order as

$$\left(\frac{\partial\phi}{\partial s}\right)^{(e)*} = [N] \left\{\frac{\partial\phi}{\partial s}\right\}^{(e)*} \tag{64}$$

where $\{\partial\phi/\partial s\}^{(e)*}$ is the set of nodal values of the corresponding derivatives, equivalent in a local averaging sense, to the original ones, $\{\partial\phi/\partial s\}^{(e)}$. The element equivalence is enforced by requiring that both formulations for the derivatives give same element integrals, defined by

$$I_\phi^{(e)} = \int_{S^{(e)}} \{N\} \left(\frac{\partial\phi}{\partial s}\right)^{(e)*} ds = \int_{S^{(e)}} \{N\} \left(\frac{\partial\phi}{\partial\xi}\right)^{(e)} J(\xi, s) ds \tag{65}$$

where the integration is performed on $S^{(e)}$, which is the element length in 2D or the element area in 3D. If the element under integration is now written as $ds = J(s, \xi)d\xi$, where $J(s, \xi) = \partial s/\partial\xi = 1/J(\xi, s)$, all terms in the integrands are only functions of the intrinsic variable ξ , and all element integrations are performed for $\xi \in [-1,1]$, for the usual Lagrangian interpolating functions.

The element equations in Eq. (65) are then assembled for the entire boundary S , leading to an expanded system of equations that includes the original elements and the new zero-length elements at the double nodes.

A simple procedure is implemented in this work to account automatically for these new zero-length elements when assembling the system of equations. The procedure consists of three steps. In the first step, a test is performed at each node i , for discontinuities in the tractions or in the normal unit vector between the elements that share the node. A corner locator variable is set to a value of one wherever a corner is located or a discontinuity in traction exists.

The second step is to renumber the current nodes and elements to include the zero-length elements at the located “corner” nodes. A new augmented connectivity matrix is thus created relating node/element numbering, including the added zero-length elements. With this procedure, single values for the tractions and for the tangential derivatives of both the displacements and the tractions can be assigned to each node in the augmented system. An assembled system of equations is thus obtained from which the vectors of smoothed nodal derivatives $\{\partial\phi/\partial s\}^{(e)}$ can be obtained in this augmented system. The system matrix for the coefficients of the recovered or smoothed quantities is not singular and admits an inverse after assembling all elements.

The third step is to evaluate the element errors still in the augmented system and then renumber backwards the nodes/elements to their original numbering, by excluding the zero-length elements, so that only the error information of the original elements is retained.

The recovery of all four densities ($\phi = u, v, SG$ or TAU) is implemented in this work so that error estimators are obtained separately from the smoothed tangential derivatives of the displacements and of the tractions, and also a combined error estimator is obtained when the individual contributions from the tangential derivatives of the various densities are added. A scaling parameter is introduced so that the individual contributions are scaled before being added. Additional details for the implementation of the tangential derivative recovery for linear elements are described in (Jorge et al., 2005).

3.2.4.2 The external domain approach

The external domain error estimator approach is based on the difference between two equivalent formulations. The exact solution for one integral representation a problem satisfies any other possible different integral representation of the same problem. When a numerical solution is found for one formulation, this solution can be substituted into the discretized form (with same discretization) of the other possible formulation. Because numerical solutions are not exact, the identities in the discretized form of this second formulation will not be satisfied. The difference is a measure of the discretization error of both formulations.

The external domain approach uses the fact that, for a closed domain, the exterior boundary integral representations for both the displacements and the stresses equal zero in the elastostatics problem. Thus, any difference from zero can be related to the discretization error. Because fundamental solutions are functions of the distance between source and field points, one representation in which the exterior point is very close to the boundary might give a different error measure than the case in which the point is far away from the boundary. If different exterior points are considered, several external formulations may be obtained.

This error measure, estimated from the residual of the integral equation held at the outer source point, is sensitive to the external point position for non-trivial problems. For small but finite distances to the boundary, the estimator gives a local measure of the discretization error, but if the distance to the boundary becomes too small, errors associated to the location of the source point, coming from the numerical evaluation of the near-singular integrals, might become significant.

To implement the procedure, first a numerical solution at the boundary is obtained for the discretized mesh using a collocation BEM formulation. This boundary solution is then

substituted into the discretized integral representations for the interior displacements and stresses, and finally these displacements (u and v) and stresses (σ_x , σ_y and τ_{xy}) are evaluated at an exterior point $y^{(e)}$ in a close vicinity to a boundary element (e). An external error estimator for the total displacement ($Ext: D$) is constructed as

$$E_{Ext:D}^{(e)} = \left[u(y^{(e)})^2 + v(y^{(e)})^2 \right]^{1/2} \tag{66}$$

External error estimators can be defined from the stresses at the exterior point, as well. To eliminate any influence of the choice of global directions on the values of the stress error estimates, the stresses at the exterior point are combined into stress invariants to obtain error estimators. Two external errors estimators are adopted in this work: the octahedral stress ($Ext: T$) and the mean stress ($Ext: M$), given by

$$E_{Ext:T}^{(e)} = \frac{\sqrt{2}}{3} \left[\sigma_x^2 - \sigma_x \sigma_y + \sigma_y^2 + 3\tau_{xy}^2 \right]^{1/2} \tag{67}$$

$$E_{Ext:M}^{(e)} = (\sigma_x + \sigma_y) / 2$$

with σ_x , σ_y and τ_{xy} evaluated at the exterior point $y^{(e)}$. The mean and octahedral stresses were chosen because they are independent stress invariants. Naturally, if the numerical boundary solution is exact, all external error estimators are zero.

The exterior evaluation point $y^{(e)}$ is located at a distance from the boundary (the distance from the exterior point to the closest point on the boundary) that is a fraction α of the size of the nearest boundary element (e). This distance is measured from the element center in the normal direction to the element at this point.

In this work, numerical experiments are done with various values of the fraction α in order to evaluate the most appropriate location for the external point y . The closer the external point to the boundary, the more “local” the error information is.

On the other hand, if the external point is too close to the boundary, numerical errors may appear due to quasi-singularities.

The external domain error estimators can also be combined to obtain a more general error estimator ($Ext: D + T + M$) with error information from both external displacements and stresses. The individual contributions for element error are also scaled before adding, with the scaling factors defined as the average boundary total displacement $(u^2 + v^2)^{1/2}$ and total traction $(SG^2 + TAU^2)^{1/2}$, respectively.

An implementation of the proposed error estimators was presented in (Jorge et al., 2005), for the element errors $E_{Ext}^{(e)}$ for various values of the fraction α . Comparisons were also made between $E_{Ext}^{(e)}$ and $E_{GR}^{(e)}$ to show the consistency of the local information given by these estimates with mesh refinement.

One must point out that the external domain error estimator can be used only for domains in which an exterior region exists. For example, if using a traction-BIE without a multiregion approach for the case of a crack embedded in an infinite elastic body, no exterior region could be defined in which the evaluation point would be at a finite distance from the boundary.

In this case, the external domain error estimator could not be used, while the gradient recovery approach can always be used, for closed or open boundaries, and for finite or infinite domains.

3.2.4.3 A final note on the proposed Error Estimators for elastostatics BEM

Error estimators based on gradient recovery and external solution have shown consistency with mesh refinement and have given similar qualitative results. The error estimators were also proven to be useful to compare the efficiency of different types of mesh refinement (such as the r - and h -refinement schemes) for a particular problem.

The error estimator using the gradient recovery approach presents a more general characteristic as this formulation does not rely on an “optimal” choice of an external parameter, such as in the case of the external domain error estimator. Also, the external domain error estimator can be used only for domains in which an exterior region exists, while the gradient recovery approach can be used on a broader range of elastostatics problems.

4 Model Validation Under Uncertainty: Consistent Error Estimators and Bayesian Approaches

This section is based on research work in errors estimators in the context of modeling and simulation for design under uncertainty (Jorge et al., 2002), performed under the supervision of Dr. Sankaran Mahadevan, head of the multidisciplinary NSF-IGERT program (Multidisciplinary Doctoral Studies in Reliability and Risk Engineering and Management). The research objectives comprised:

- Review of error estimators for finite element methods.
- Coding of two error error estimators for the 2-D displacement-FEM. The first estimator is based on a gradient recovery procedure; the second estimator is based on the lack of inter-element stress continuity for the discretized results.
- Study the correlation / consistency between different error estimators, including both estimators coded (see previous item) and two other estimators previously obtained: error estimator based on Richardson extrapolation (uses two meshes: coarse and fine) and maximum error bound for the error (based on second derivatives).
- Outline a global-to-local approach for validation as follows: globally, the error estimator is investigated for consistency with the measured error or with other available estimators. Then, weak local regions are chosen to construct a reliability model to validate the numerical code.

A new model validation approach has been developed (Zhang & Mahadevan, 2003) (Rebba et al., 2006) (Mahadevan, 2002), in which error estimators from the discretized model were combined into an appropriate stochastic response surface approach, and a Bayesian approach for hypothesis testing was subsequently used, leading to decisions on acceptance/rejection of the model. The methods were demonstrated for application to the probabilistic life prediction of a composite helicopter rotor hub component, shown in Fig. (25).

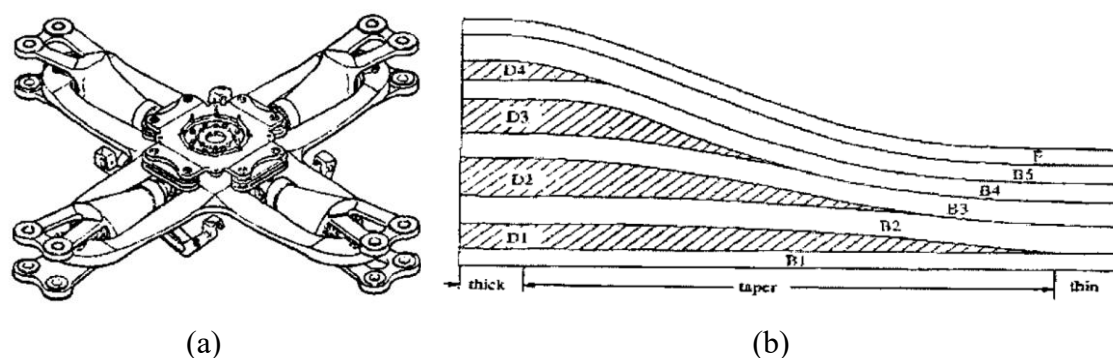


Figure 25: (a) Rotor hub assembly (b) Half of the symmetric section of tapered composite test specimen - adapted from (Mahadevan, 2002)

In (Jorge et al., 2002), the consistency of different error estimators evaluated from different response quantities was investigated. Mean and octahedral stresses at a particular location were used as response quantities to obtain independent error estimates. Consistency between predicted weak locations of the structure was also investigated. The gradient recovery and the stress continuity error estimators were shown to be highly correlated and to indicate similar weak regions for the structure, while error estimators based on Richardson extrapolation and maximum error bound didn't shown similar consistency.

Also in this work, model validation results were investigated for consistent and nonconsistent error estimators, using conventional and stochastic response surface methods (RSM). Stochastic RSM was shown to be a more accurate model (for the response quantity of interest) than conventional RSM. Also, model validation results were found to be consistent (i.e., the probability of rejecting the code was similar) for consistent error estimators using the stochastic RSM.

For related research work in these topics see also (Rebba et al., 2003) (Rebba, 2005) (Rebba & Mahadevan, 2006b) (Rebba et al., 2006) (Rebba & Mahadevan, 2006b) (Rebba & Mahadevan, 2006a).

4.1 Consistent Error Estimation Strategy for Model Validation

4.1.1 Introduction

A 3-step methodology for numerical model validation under uncertainty is presented in this paper, in which a) a consistent error estimator is selected; b) an appropriate response surface method is constructed for the error function to replace the numerical model; and c) a Bayesian approach is used for decisions on model acceptance. This section concentrates on the strategy for the choice and use of appropriate error estimators for model validation purposes.

The details of the proposed methodology are included in subsection 4.1.2. For the purpose of illustrating the consistent error estimator strategy, the discussion in this work uses a displacement-based finite element model for a 2D elastostatics problem. Surveys of the extensive literature existing on error estimation for finite element methods can be found in (Noor & Babuška, 1987) (Olgierd C. Zienkiewicz & Taylor, 1994) (J. Tinsley Oden & Demkowicz, 1989) (Zhu, 1997) (Ainsworth & Oden, 1997). Mostly, the literature deals with discretization error, and the development of methods for estimating modeling error

has appeared very recently (J. Tinsley Oden & Vemaganti, 2000) (J. Tinsley Oden et al., 2003).

Adaptive mesh refinement is a common application of error estimators (O. C. Zienkiewicz & Zhu, 1991) (Verfürth, 1994) (Zhu & Zienkiewicz, 1997). For that purpose, error approaches usually include the evaluation of error norms and / or some type of residuals (J. T. Oden et al., 1990) (Ainsworth & Oden, 1993) (I. Babuška et al., 1995). Usually, residuals, norms and other quantities used for error estimation are not explicit variables of a code. Instead, residuals are usually obtained after numerical manipulations of the code variables, and thus they cannot be measured directly. In this work, error estimators are sought that provide local error information on variables obtained explicitly from the code, so that these variables can be studied as response quantities of interest, for further stochastic treatment.

Four existing discretization error estimator approaches are investigated in this work, for their usefulness for further stochastic treatment of the error. The first error estimator is based on Richardson's extrapolation (Richards, 1997), in which the local error is evaluated from the differences between the response quantities obtained for two different meshes. The second error estimator gives in fact not an estimate, but just an upper bound for the magnitude of the error, and is called the maximum error bound error estimator (Huebner et al., 2008). This maximum error bound is obtained by evaluating second derivatives of the smoothed numerical solution. The third error estimator procedure is based on inter-element continuity requirements for the stress field, not satisfied by the finite element solution (O. C. Zienkiewicz & Zhu, 1987). The fourth error estimator is based on a gradient recovery procedure (O. C. Zienkiewicz & Zhu, 1987) (O. C. Zienkiewicz & Zhu, 1992a) (O. C. Zienkiewicz & Zhu, 1992b). The recovery of displacement derivatives is rewritten as stress recovery, using Hooke's law. The error estimators used in this work are detailed in subsection 4.1.3.

The key feature in the approach is the study of the consistency of some error estimator with respect to other available information. In this work, the available information is represented by other independent error estimator approaches and / or error estimators obtained for other independent response quantities. The error estimators are investigated for consistency in Section 4.1.4.

Model validation can be performed also by comparing the numerical information with some available measured / experimental results. The procedure to investigate error consistency is analogous, with the error estimator results being compared with the experimental results, in this case. The consistency of the error estimators with respect to measurement results (if available) is studied using the same global-to-local methodology, but comparisons between measurement and numerical results can be made only in regions where measurement results are available.

The consistent error estimator approach (first step of the 3-step approach) is detailed in subsection 4.1.4. Depending on the choice of the error estimator and response quantity, the model validation criterion could lead to different decisions of accepting or rejecting the model. The strategy adopted in this work is to choose different error estimators and check for the consistency of the error throughout the domain being modeled. For a particular problem, a measure of the consistency between the available error estimators can be obtained by the correlation between the 2D plots of the local error estimates. Similarly, this correlation study can be performed between estimated error results and the

measured error, if some information from measurements of the physical problem is available. Typical results for the errors are included in subsection 4.1.4, below.

The response surface methods (RSM) and the Bayesian approach used in this work (for the time-independent case without statistical uncertainty in the model) are detailed in subsection 4.2, below. The Bayesian approach is applied to obtain Bayes factors, needed for model-acceptance decisions. The obtained model-acceptance decisions are compared in subsection 4.2, below, for the cases of the conventional and stochastic RSMs and also for the cases of consistent and non-consistent error estimators.

4.1.2 A 3-Step Methodology for Numerical Model Validation under Uncertainty

The proposed methodology for numerical model validation under uncertainty consists of three steps. In the first step, a global-to-local strategy is used to locate regions with high error values and to choose an error estimator that consistently represents the discretization error in these regions of interest.

For a particular problem being modeled, local error estimators could be evaluated for different response quantities. Also, different error estimation approaches could be adopted for the same response quantity. Also, different locations in the domain being modeled could be chosen to generate the prior estimates of model failure probability. Depending on so many choices for the error estimator, the Bayesian model validation could lead to different decisions of accepting or rejecting the model, for different error estimators, response quantities and locations.

The strategy adopted in this work is to check for the consistency of the distributed error throughout the domain being modeled, for different error estimators. The correlation between 2D plots for the local error obtained from different estimators is used in this work as a measure of the consistency of the estimator.

From the correlation between the estimators, an error estimator is chosen based on its global consistency when compared with the other available error sources.

Different error estimators are compared by means of 2D dispersion plots. Also compared are the errors obtained from different response quantities. By using 3D plots, errors are shown to indicate weak regions of the structure, where high errors coexist with high, unacceptable, values of the response quantity. Consistent error estimators are highly-correlated, and thus are expected to identify and locate similar weak regions. A certain number of representative points are selected from these weak regions.

The second step constructs a metamodel for the error at the selected points using response surface techniques, to estimate the statistical distribution of the prediction error for the finite element model. The statistical distribution of the model error at a particular location is used to compute prior estimates of model failure probability, to be used in Bayesian model validation.

Finally, Bayesian hypothesis testing-based model validation procedure can be applied for the selected critical locations, and decisions of accepting or rejecting the model are made based on the Bayes factor computed at the critical locations. The Bayesian approach (third step of the 3-step approach) has been developed earlier in (Zhang & Mahadevan, 2003), for time-independent and time-dependent cases, and for cases with or without statistical uncertainty in the model.

The obtained Bayes factors were used as acceptance/rejection criteria for the model. This Bayesian approach is extended in this work to metamodels corresponding of response surfaces representing the chosen consistent estimator of the numerical error at the selected points.

Basically, first the approach looks for error estimator consistency from a global point of view. Then, weak regions are located where the Bayesian approach is applied, so model validation is considered from a local perspective, only at regions where important errors were found.

The goals are to establish enough confidence on the error estimator first, and then to use the error information to reduce the number of points where model validation will be tested.

4.1.3 Estimators of Discretization Error for Model Validation

Errors in the prediction of a certain response quantity are expected to exist when using approximate numerical models. Different error estimators can be envisaged for a particular problem. The current work focuses on a strategy for the choice and use of appropriate error estimators for model validation purposes.

Some error estimators well suited for adaptive refinement provide estimates that are always positive. This is the case, for example, of error estimators based on residuals. In this work, the envisaged stochastic distribution of the error may be able to assume either positive or negative values, as the numerical model over-predicts or under-predicts the average response quantity in the element.

Some of the error estimators in this work were derived so that they can assume positive or negative values, to allow further stochastic treatment for the error.

Four error estimator approaches are investigated in this work:

- error estimator based on Richardson's extrapolation (Richards, 1997);
- maximum error bound error estimator (Huebner et al., 2008);
- error estimator based on lack of inter-element stress continuity (O. C. Zienkiewicz & Zhu, 1987); and
- gradient recovery error estimator (O. C. Zienkiewicz & Zhu, 1987) (O. C. Zienkiewicz & Zhu, 1992a) (O. C. Zienkiewicz & Zhu, 1992b).

These estimators were selected because they are able to provide error information for variables explicitly obtained from the code, without the need to indirectly obtain other quantities such as the residuals. These code variables can be understood as response quantities, for further stochastic treatment of the error.

For the elastostatics problem being studied, all estimators provide local error information for the stresses. Stress invariants can be written so that the stress information is recast in terms of scalar quantities, independent of the directions. The mean stress and the octahedral stress are used for this purpose. Because in 2D these two stresses are independent quantities, error estimators based on mean stress and on octahedral stress are also independent.

After the various distributions of element errors are obtained, relationships between the different error indicators, or between an error indicator and the measured error, can be obtained. In subsection 4.1.4, 2D plots are adopted, to study the 2D dispersion or

correlation of those error measures. Linear regression can be used to quantify this correlation. Points in the 2D plot are generated from local error information at the various elements. As far as the spatial error distribution is concerned, each element error is associated with a point corresponding to the element centroid.

A square plate with a hole in the center subjected to uniform distributed loading is considered as a numerical example. One quarter of the plate is modeled, as shown in Fig. (26). The plate is discretized using a mesh of 133 nodes, with constant strain triangular elements (CST), i.e., elements with linear distribution for the displacements. A second mesh with 280 nodes is considered for the Richardson extrapolation error estimator.

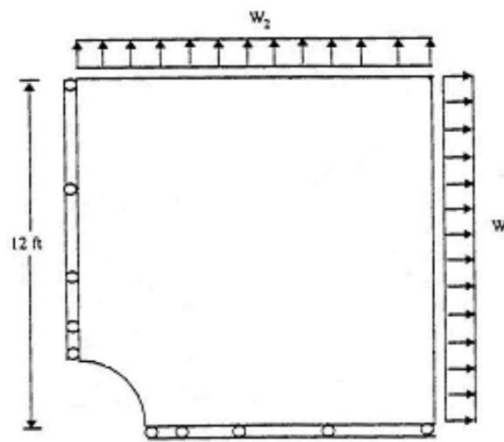


Figure 26: Plate with a central hole under uniform distributed loading - adapted from (Jorge et al., 2002)

4.1.3.1 Richardson extrapolation error estimator

In the Richardson extrapolation for estimating the error in grid refinement, consider an analysis code that predicts the output response f as

$$f = f_{h=0} + g_1h + g_2h^2 + g_3h^3 + \dots \tag{68}$$

where $f_{h=0}$ is the exact solution or the continuum value given for the mesh size $h = 0$. The functions g_1, g_2, g_3 , etc., are independent of the grid size. Assuming a second order solution and computing the response at two different mesh sizes h_1 and h_2 , with h_1 being the size of the finer mesh, neglecting higher order terms in the above expansion and solving for $f_{h=0}$,

$$f_{h=0} \cong f_1 + \frac{f_1 - f_2}{r^2 - 1} \tag{69}$$

where the grid refinement ratio $r = \frac{h_2}{h_1}$ and f_1, f_2 are the solutions at two mesh sizes. By generalizing the Richardson extrapolation for a p^{th} order method,

$$f_{h=0} \cong f_1 + \frac{f_1 - f_2}{r^p - 1} \tag{70}$$

The ordered error estimate, given by $(f_{h=0} - f_1)$, is a good approximation of the discretization error. The order of convergence p can be obtained from the equation:

$$p = \ln\left(\frac{f_3 - f_2}{f_2 - f_1}\right) / \ln(r) \tag{71}$$

where f_3 is the solution with the finest grid size.

4.1.3.2 Maximum error bound

If the exact solution for a problem in two dimensions, $f(x, y)$ is known, then the error associated with the finite element solution $f_{fem}(x, y)$ can be written as (Verfürth, 1994):

$$E(x, y) = f(x, y) - f_{fem}(x, y) \tag{72}$$

Thus the error varies across the domain. If the exact solution is not known, local error estimation based on interpolation error is adopted. To illustrate the derivation of the local maximum error bound, a one-dimensional case is shown subsection 3.1.1, for simplicity (see also Fig. (16), above). For a two-dimensional problem,

$$E(e) \cong \frac{1}{8} \left(\left| \frac{\partial^2 f_{fem}}{\partial x^2} \right|_{x_M} + 2 \left| \frac{\partial^2 f_{fem}}{\partial x \partial y} \right|_{x_M} + \left| \frac{\partial^2 f_{fem}}{\partial y^2} \right|_{x_M} \right) h^2 \tag{73}$$

4.1.3.3 Error based on inter-element continuity for stresses

An error estimation procedure for displacement-based FEM was developed by (O. C. Zienkiewicz & Zhu, 1987) to obtain energy norms of errors derived from the lack of inter-element stress continuity. In this work, instead of evaluating element residuals that are always positive, an average value for the element error is pursued, which can assume either positive or negative values.

The usual C^0 continuity assumption for the displacements in displacement-based finite element formulations results in a stress field that is continuous inside the elements, while discontinuous in the interface between elements. Except for problems with physical discontinuities, such as cracks in fracture mechanics problems, the usual linear elastostatics problems deal with stress fields that are physically continuous. Thus, differences between stress values evaluated at an inter-element node, from the various elements sharing this node, represent local error indicators.

To construct an element-based measure for the stress error at a node, the average stress at this node, as evaluated from all elements sharing the node, is taken as a better approximation for the nodal stress. Thus, the stress error vector at node n of element i is given by

$$\Delta \sigma_n^i = \sigma_n^i - \sigma_n^{ave} \tag{74}$$

where σ_n^i represents a stress component (either σ_x , σ_y or τ_{xy}) at node n of element i , and the corresponding averaged stress is evaluated as

$$\sigma_n^{ave} = \frac{\sum_{i=1}^{N_e^n} \sigma_n^i}{N_e^n} \tag{75}$$

where N_e^n is the number of elements sharing node n . An element error for each stress quantity is constructed as the element average of the nodal contributions. For that, the stress error is interpolated inside element i using appropriate shape functions.

$$\Delta\sigma_i(x, y) = [N_n(x, y)] \{ \Delta\sigma_n^i \} \tag{76}$$

The element error for each stress quantity is obtained from the stress error average.

$$E_\sigma^i = \frac{1}{S^i} \iint_{S^i} \Delta\sigma_i(x, y) dx dy \tag{77}$$

In this work, the stress error is interpolated with the same shape functions used for the interpolation of the displacements. In other words, the stress error is interpolated inside each element with shape functions one degree higher than the shape functions necessary to interpolate the original stress components. In the case that $\Delta\sigma_i(x, y)$ is interpolated with linear shape functions, the element error in Equation (77) reduces to the stress error at the element centroid G

$$E_\sigma^i = \Delta\sigma_i(x_G, y_G) \tag{78}$$

The resulting element error has dimensions of stress, and in terms of spatial error distribution, this error can be associated with a point in the domain corresponding to the centroid of the element. Also, different from (O. C. Zienkiewicz & Zhu, 1987), the error is not obtained from an energy norm expression. The goal here is to obtain estimates that can assume positive or negative values, to allow further stochastic treatment for the error. The stress tensor components are direction-dependent quantities. To obtain scalar quantities to be used as response quantities, the stress tensor invariants can be evaluated. In this work, the mean stress (σ_m) and the octahedral stress (τ_o) are adopted as response quantities, as they are independent invariants obtained from the stress tensor.

$$\sigma_m = \frac{\sigma_x + \sigma_y}{2} \tag{79}$$

$$\tau_o = \frac{\sqrt{3}}{2} [\sigma_x^2 - \sigma_x \sigma_y + \sigma_y^2 + 3\tau_{xy}^2]^{1/2} \tag{80}$$

Thus, from Equation (77) (general case) or from Equation (78) (linear interpolations for the stress errors), the element stress errors can be obtained as $E_{\sigma_m}^i$ and $E_{\tau_o}^i$. Thus, independent error estimators are obtained from the two independent response quantities, mean stress and the octahedral stress.

4.1.3.4 Error based on a gradient recovery approach

The procedures of error estimation using stress or gradient recovery were introduced by (O. C. Zienkiewicz & Zhu, 1987) and are detailed in subsection 3.3.1, above. In this procedure, element residuals are obtained from differences between smoothed and non-

smoothed derivatives of the field variables. In the current work, instead of evaluating element residuals that are always positive, an average value for the element error is pursued, which can assume either positive or negative values.

Nodal smoothed derivatives are evaluated from a weighted average of element derivatives surrounding the node. The weighting factors are the element areas. The derivative of the finite element solution in the element can be written using derivatives of the shape functions. The derivatives are written in an average sense to find the vector of nodal smoothed derivatives. An element equivalence between smoothed and non-smoothed derivatives is written and a system of equations is obtained by assembling the element equations, from which the vector of nodal smoothed derivatives is obtained.

Also, similar to the inter-element continuity error procedure, the response quantities of interest are two independent stress invariants, the mean stress and the octahedral stress. To obtain a better approximation of nodal stresses, the first step is to apply the gradient recovery approach to displacement derivatives. Smoothed nodal values for the displacement derivatives are obtained by imposing the smoothed solution to be equivalent to the original solution in an average sense. The obtained nodal values for the smoothed displacement derivatives are converted into smoothed strains which are subsequently converted into stresses by means of Hooke's law. Then, a better approximation for the mean and octahedral stresses is obtained from Equations (79) and (80). Finally, the stress error vector at node n of element i is given by

$$\Delta\sigma_n^i = \sigma_n^i - \sigma_n^{i*} \quad (81)$$

where σ_n^i represents a stress component (either σ_x , σ_y or τ_{xy}) at node n of element i , and the smoothed stresses are denoted by an asterisk (*). Then, similar to the inter-element continuity error procedure, the element error is obtained by Equation (77), for both the mean and the octahedral stresses. Thus, two independent local error estimators, one for each response quantity, are obtained using the recovery procedure for the displacement derivatives.

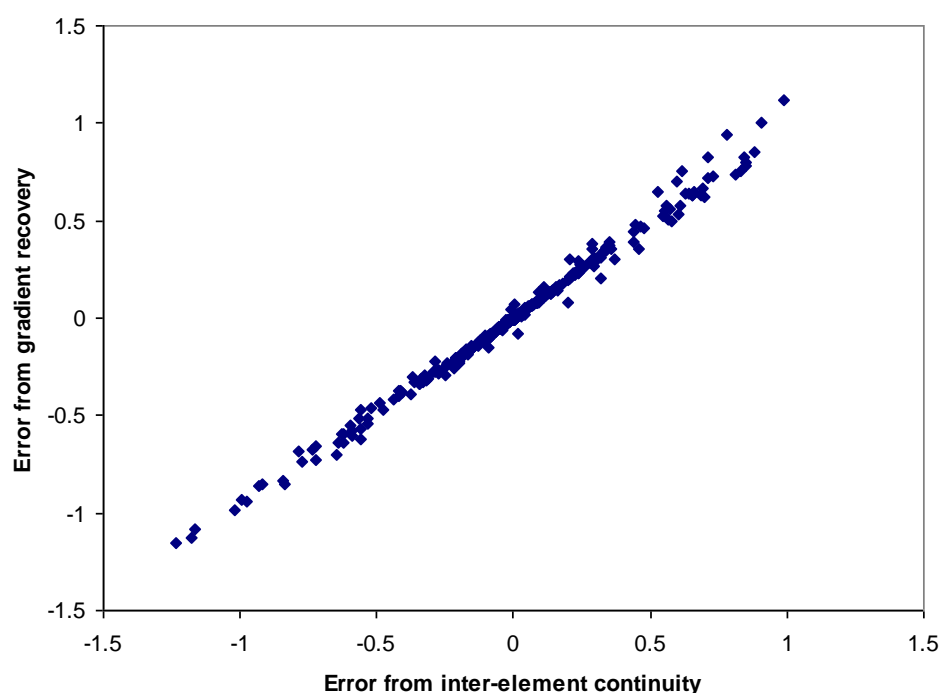
The gradient recovery error estimator was coded for 2D elastostatics problems modeled by displacement-based FEM. The FEM code uses CST constant strain triangular elements with linear distribution of displacements. The recovery of the four types of derivatives $\frac{\partial u}{\partial x}$, $\frac{\partial u}{\partial y}$, $\frac{\partial v}{\partial x}$, and $\frac{\partial v}{\partial y}$ is performed independently. Also, to obtain the smoothest solution for the displacement derivatives, element matrices obtained from the equivalence equation are assembled for the entire domain, and not only for a region surrounding a node.

4.1.4 Consistency of the Error Estimators

The four error estimators presented in this work are investigated for consistency. The correlation between different error estimator results for the response quantity is investigated. Also investigated is the correlation between error results for the same error estimator and different response quantities.

4.1.4.1 Consistency between error estimators for the same response quantity

Different error estimation approaches can lead to different local error estimates for a given response quantity. Consistency through different error estimators is investigated for the two response quantities used in this work: mean stress and octahedral stress in the region of interest. Figure 27 shows the good agreement between the errors in mean stress from two different error estimation approaches: gradient recovery and lack of inter-element continuity in stress. The points present a very good correlation, with $\rho=0.9927$, showing that the gradient recovery error estimator is consistent with the inter-element continuity error estimator for the response quantity considered.



**Figure 27: Errors in Mean Stress from two different error estimation approaches. Each point corresponds to an element centroid. Mesh of 280 nodes
- adapted from (Jorge et al., 2002)**

The four error estimators for the octahedral stress were compared two by two, to investigate the correlation between the pairs of estimators. Figures 28 (a) to (f) show the various plots, including the regression equation and the R -squared parameter obtained from linear regression.

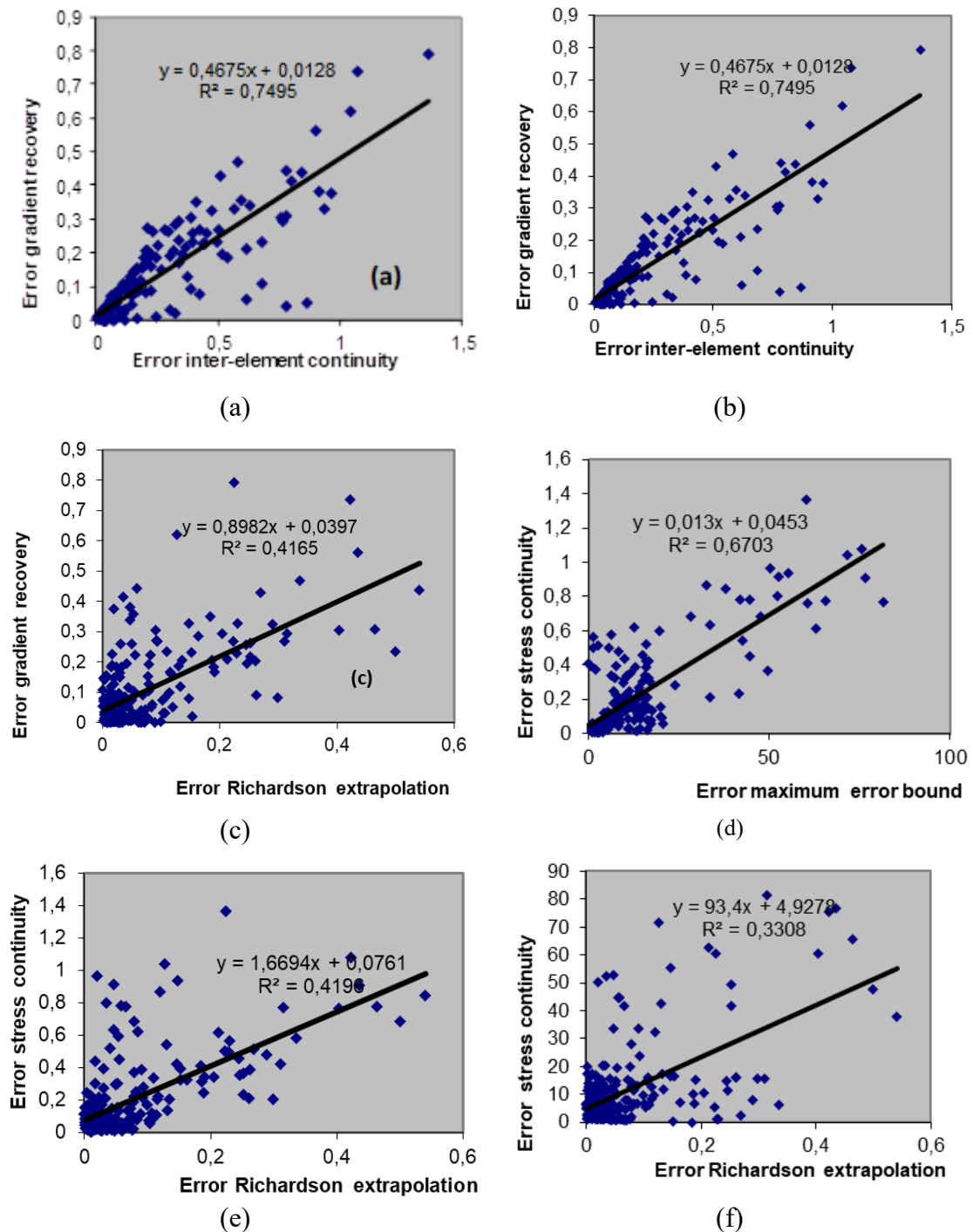


Figure 28: error estimators for the octahedral stress, 133-node mesh (a) Gradient recovery vs. stress continuity; (b) Gradient recovery vs. maximum error bound; (c) Gradient recovery vs. Richardson extrapolation; (d) Stress continuity vs. maximum error bound; (e) Stress continuity vs. Richardson extrapolation; (f) Maximum error bound vs. Richardson extrapolation - adapted from (Jorge et al., 2002)

From the six cases, the diagram for gradient recovery and stress continuity presents the best correlation, with $\rho = 0.8657$. When the comparison is done including either Richardson extrapolation or maximum error bound error estimators, the plots present poorer correlation coefficients.

Table 1 summarizes the results for the correlation.

Table 1: octahedral stress error estimators correlation - adapted from (Jorge et al., 2002)

Estimators being compared	Correlation coefficient ρ
Gradient recovery versus stress continuity	0.8657
Gradient recovery versus maximum error bound	0.7049
Gradient recovery versus Richardson extrapolation	0.6454
Stress continuity versus maximum error bound	0.8187
Stress continuity versus Richardson extrapolation	0.6478
Maximum error bound versus Richardson extrapolation	0.5752

The poor correlation results shown in Table 1 whenever either Richardson extrapolation or maximum error bound is included in the 2-D plot, helps demonstrate the lack of consistency when comparing these error estimators to any other estimator among those evaluated in this work.

4.1.4.2 Consistency between response quantities for the same error estimator

The consistency of different (and independent) error estimators evaluated from different response quantities is investigated. For example, the mean stress and the octahedral stress at a particular point are independent stress invariants and can be used to obtain independent error indicators. Figure 29 shows the plot of the errors from the gradient recovery approach as evaluated for the two response quantities of interest, the mean stress and the octahedral stress. The points present a reasonably good correlation, with $\rho = 0.8512$, showing that the gradient recovery error estimators are consistent for the two independent response quantities considered.

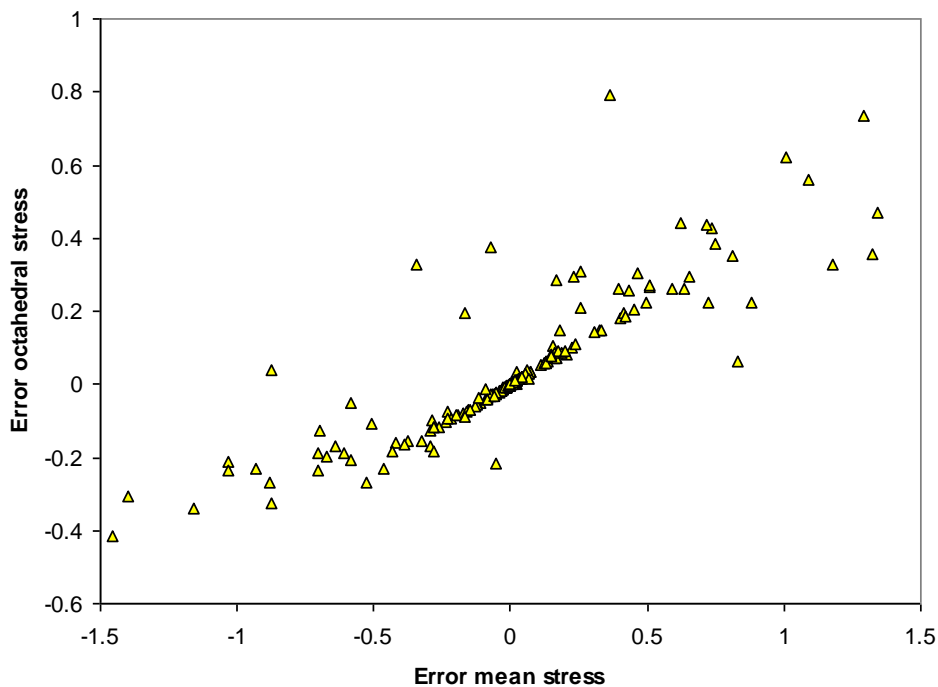


Figure 29: Stress errors from gradient recovery for two independent quantities. Points are associated with element centroids. Mesh with 133 nodes - adapted from (Jorge et al., 2002)

4.1.4.3 Consistency of weak locations for response quantities and error estimators

The consistency between predicted weak locations of the structure is investigated. Weak locations are expected to be consistent for different (but consistent) error indicators. Also, predicted weak locations should be consistent with experimentally measured weak locations. In case weak locations are not consistent, an issue remains to be investigated, namely, to verify if the error indicators were able to predict this inconsistency.

Figure 30 shows the plot of the errors in mean stress (from the lack of inter-element continuity in stresses) versus the mean stress values at element centroids. The points present a very correlation, with $\rho = 0.9641$, showing that these errors are consistent with the variable values. High values of the errors were found at locations with high stress values.

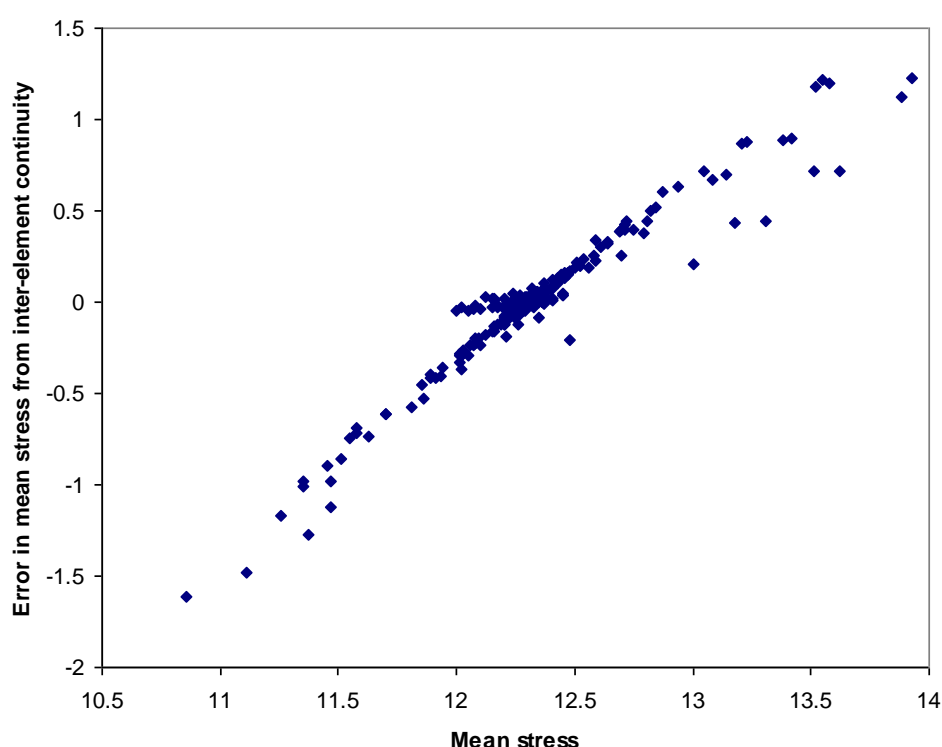


Figure 30: Error in mean stress (from the lack of inter-element continuity in stresses) versus mean stress values at element centroids. Mesh with 133 nodes – adapted from (Jorge et al., 2002)

Figure 31(a) shows the plot of the errors due to the lack of inter-element continuity in the octahedral stresses versus the Cartesian coordinates of the plate. The locations for the points with high errors are consistent with the locations with high values for the octahedral stress itself, as shown in Figure 31(b). The gradient recovery error estimator (not shown) gives similar results.

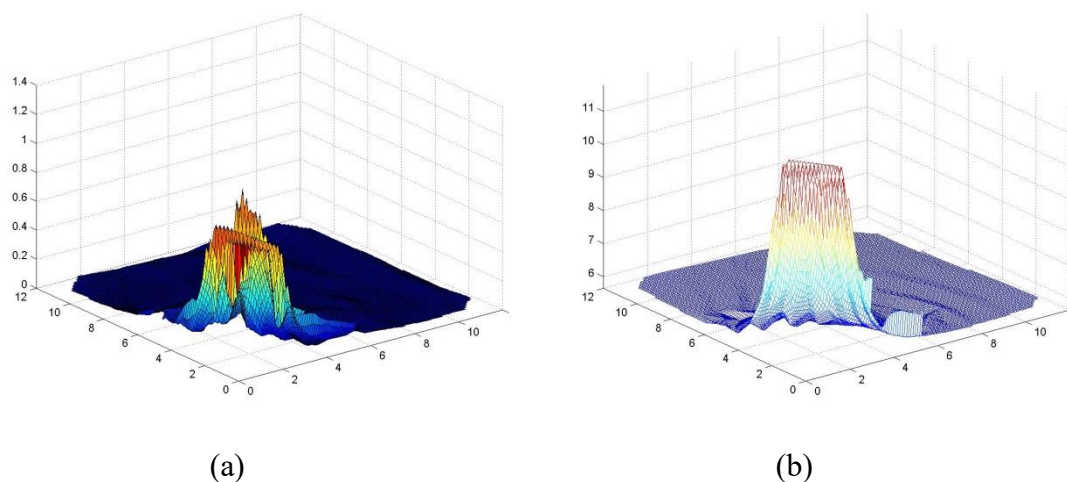


Figure 31: Consistency of weak locations of the plate. (a) Errors in Octahedral Stress from lack of inter-element stress continuity. (b) Octahedral Stress – adapted from (Jorge et al., 2002)

In this problem, the region close to the circular hole corresponds to high values of both a response quantity and the discretization error relative to this quantity.

4.1.4.4 Consistency between error estimators and model validation

An important issue that needs to be addressed is to verify the desired correlation between error consistency and model validation consistency. If error indicators are consistent, the corresponding model validation results are expected to be consistent, too. In other words, the probability of rejecting the code is expected to be similar in the case of two consistent error indicators. Furthermore, the decisions to accept/reject the model are also expected to be similar. This issue is addressed in subsection 4.2, below.

4.1.5 Global-to-Local Approach for Using Consistent Error Estimators in Model Validation

Error estimators quantify the local discretization error. With random inputs to the finite element model, the output becomes a random quantity. Thus the prediction error is a random variable, whose statistics combined with Bayesian hypothesis testing (I. Babuška et al., 1995) lead to a model validation criterion. However, the choice of the error estimator may control the acceptance or rejection of the model to some extent.

One estimator may imply that the model is valid but another estimator may imply the opposite. Also, the error in model predictions may vary over the domain. A consistent error estimator is needed for model validation, so that error predictions for two different but consistent estimators are equivalent.

A global-local approach for using consistent errors for model validation can be envisaged as follows: globally, an error indicator for a particular response quantity is investigated for consistency with other available error estimators or with a measured error, if available. Then, local weak regions are identified, where the error is large and the values of the response quantity are also large. One or more points can be selected in this ‘‘weak’’ region so that a metamodel for the error at this location can be constructed using response surface methods.

Finally, the Bayesian approach can be applied at the chosen points in the weak region, to obtain a minimum value for the Bayes factor, to be used for model validation. The Bayesian approach is applied in subsection 4.2, below, for two metamodels, one constructed using conventional response surface methods (RSM) and another using stochastic RSM. Also, the Bayesian approach is applied in subsection 4.2, below, for metamodels obtained using consistent and non-consistent error estimators.

4.1.6 Some Remarks on Consistent Error Estimators

Error consistency can be qualitatively demonstrated by means of 2D dispersion plots. A statistical model to measure error consistency is useful to quantify this consistency. In this work, the evaluation of the regression parameters from the 2D plots was performed with this purpose.

Typical consistency results are obtained by comparing the correlation between different error estimators and also by comparing errors from different response quantities. In this work, two error estimators didn't show consistent behavior (Richardson extrapolation and maximum error bound) while two other estimators (inter-element continuity and gradient recovery) presented a good correlation, for a give response quantity. Also, two different response quantities were shown to be reasonably correlated (the mean and the octahedral stress)

From this approach, one or more weak locations are selected to apply the Bayesian approach on subsection 4.2, below. Consistent error estimators lead to the same decisions in terms of the choice of the weak regions of the structure.

Consistency of the error estimators with respect to measurement results (if available) can also be studied using the same global-to-local methodology. Comparisons between measurement and numerical results can be made only in regions were measurement results are available.

4.2 Combining Error Estimators and Bayesian Hypothesis Testing in a Model Validation procedure

4.2.1 Introduction

In subsection 4.1, a 3-step methodology for numerical model validation under uncertainty is presented, in which

- a) a consistent error estimator is selected;
- b) an appropriate response surface method is constructed for the error function to replace the numerical model; and
- c) a Bayesian approach is used for decisions on model acceptance.

Also, subsection 4.1 concentrates on the strategy for the choice and use of appropriate error estimators for model validation purposes. The consistent error estimator approach (first step of the 3-step approach) is detailed in subsection 4.1, wherein four error estimator approaches were investigated: error estimator based on Richardson's extrapolation (Richards, 1997), maximum error bound error estimator (Huebner et al., 2008), error estimator based on inter-element stress continuity requirements (O. C. Zienkiewicz & Zhu, 1987), and error estimator based on a gradient recovery procedure (O. C. Zienkiewicz & Zhu, 1987) (O. C. Zienkiewicz & Zhu, 1992a) (O. C. Zienkiewicz & Zhu, 1992b).

These local error estimators were applied to two independent response quantities: the octahedral stress and the mean stress. The error estimator based on inter-element continuity for stress and the gradient recovery error estimator were shown to be consistent for a given response quantity. Also, the two response quantities were shown to be consistent for a given error estimator.

The Bayesian approach for model validation (third step of the 3-step approach) was introduced in (Zhang & Mahadevan, 2003). This approach requires the knowledge, or at least an estimate, of the statistical distribution of the error in the computational model. This information is used to compute a prior estimate of model failure probability, an essential quantity in Bayesian model validation. The error estimators used in deterministic finite element analysis need to be extended to stochastic analysis for systems with uncertainties. To attain this objective, two possible approaches would be Monte Carlo simulation and response surface methods (RSM).

In this subsection, response surface models are employed for computational efficiency in predicting the stochastic distribution of error. The conventional RSM, as well as a stochastic collocation point-based RSM are explored to obtain prior error distribution to be used in the Bayesian approach. The error in predicting the response is obtained in terms of the input random variables and the distribution of this error is used to calculate the Bayes factor, an estimator of model validity.

Finally, in this subsection, only the Bayesian approach for the time-independent case without statistical uncertainty in the model is applied, to obtain the Bayes factors needed for model-acceptance decisions. The obtained model-acceptance decisions are compared for the cases of the conventional and stochastic RSMs and also for the cases of consistent and non-consistent error estimators.

4.2.2 Stochastic Distribution of Error

For a certain problem, the output of the model is given by some function of response quantities. In the 2D plate presented in subsection 4.1, the octahedral and the mean stress are local response quantities obtained from the finite element model. These response quantities are evaluated at certain locations, either the nodes or the element centroids, depending on the error estimator.

Due to physical, model and data uncertainty, the obtained response quantity at a certain location is not a single value but follows a statistical distribution. The type of distribution depends on the statistics of the input random variables, and also on how these uncertainties propagate through the model to become uncertainties in the response quantity.

The uncertainties in the response need to be quantified using probabilistic methods. One possible approach to quantify these uncertainties would be to perform Monte Carlo simulation. For that, statistical parameters are assigned to the input quantities, and the response is obtained by repeating the finite element analysis for a large sample input parameters. For the plate problem, the loads are the random inputs of the model. The scatter in the response can be used to estimate the scatter in error. In order to determine the type of distribution for the response, we need typically, a large number of finite element runs. For complex finite element models, this could be practically impossible and very time intensive, and is not pursued in this work.

For computational efficiency, a sufficiently accurate computational model may be used to replace the finite element model. The variabilities in the input variables are propagated to the output through this computational model. Various methods are available to carry out such uncertainty analysis. Available uncertainty propagation models can be classified into three categories:

- (a) analytical methods;
- (b) sampling based methods; and
- (c) response surface methods.

The selection of the method depends on the nature of model used for predicting the output. In this subsection, response surface-based methods are used to predict the output of the response quantity at a certain location and determine the distribution of the error in this predicted output. The error estimators discussed in subsection 4.1 can be extended to estimating the error in the prediction of the response in stochastic analysis, in which the response is written in terms of input random variables by a response surface technique. Thus, the error calculated is stochastic and its distribution can be easily obtained by simulating the input random variables.

4.2.3 Response Surface Methods

Response surface methods (Pike et al., 1988) consist in selecting different values of input parameters, running the computer or simulation model to get samples of response, and then fitting a closed-form model which acts as a surrogate to the actual input-output relationship. The new adopted surrogate model or metamodel can be used to predict the response and carry out the uncertainty analysis. In this subsection, two stochastic response surface based methods are investigated, specifically, conventional response surface method (CRSM) and stochastic response surface method (SRSM). The details in implementation of the two methods are discussed below.

4.2.3.1 Conventional Response Surface Method

The basic idea in using the response surface is to propagate the uncertainties in the input parameters to the output response. The full-scale complex physical model of interest e.g. finite element model or the exact model can be replaced by a response surface equation, consisting of the input parameters, with data obtained from an appropriate design of experiments. This model can describe the variability in the response. A simple second order response surface model in terms of the input variables x_i is given by

$$y(x) = b_o + \sum_j b_j x_j + \sum_i \sum_j b_{ij} x_i x_j \tag{82}$$

For the implementation of the response surface methodology to the plate problem defined in subsection 4.1, the coefficients of the above equation are estimated from a factorial design, using the statistics for the load variables given in Table 2.

Table 2: Statistics of the input variables - adapted from (Jorge et al., 2002)

Variable	Type	Parameters
W_1	Normal	(12, 2.4) kips/in
W_2	Normal	(12, 2.4) kips/in

Sample points for fitting the conventional RSM are obtained when the response quantity (in this work, the error in the mean stress or in the octahedral stress at the chosen location) is evaluated for W_1 and W_2 assuming the following values: low= $12-3*2.4=4.8$, average= 12 , and high= $12+3*2.4=19.2$.

4.2.3.2 Stochastic Response Surface Method

The metamodel or response surface could also be constructed in a different way. A stochastic response surface method is proposed here with the objective of providing a better response surface, which is accomplished by approximating both the input and output of the system with uncertainties through series expansions of standard random variables. Each of the input random variables may be expressed as a function of standard normal variables ξ_i , because of mathematical tractability. For example, a normal random variable can be expressed in terms of its parameters as $\mu + \sigma\xi$. A uniform random variable bounded between a and b is expressed as $a + \frac{b-a}{2} \left(1 + \operatorname{erf} \left(\frac{\xi}{\sqrt{2}} \right) \right)$. The output can be approximated by a polynomial chaos expansion, given by:

$$y = a_o + \sum_{i_1=1}^n a_{i_1} \Gamma_1(\xi_{i_1}) + \sum_{i_1=1}^n \sum_{i_2=1}^{i_1} a_{i_1 i_2} \Gamma_2(\xi_{i_1}, \xi_{i_2}) + \sum_{i_1=1}^n \sum_{i_2=1}^{i_1} \sum_{i_3=1}^{i_2} a_{i_1 i_2 i_3} \Gamma_3(\xi_{i_1}, \xi_{i_2}, \xi_{i_3}) + \dots \quad (83)$$

where y is the output and $\Gamma_p(\xi_{i_1}, \dots, \xi_{i_p})$ are multi-dimensional Hermite Polynomials of degree p given by

$$\Gamma_p(\xi_{i_1}, \dots, \xi_{i_p}) = (-1)^p e^{\frac{1}{2}\xi^T \xi} \frac{\partial^p}{\partial \xi_{i_1} \dots \partial \xi_{i_p}} e^{-\frac{1}{2}\xi^T \xi} \quad (84)$$

where ξ is a vector of independent standard normal variables $\{\xi_{i_k}\}_{k=1}^p$. The series could be truncated to a finite number of terms. The accuracy of the computational model depends on the order of the expansion. Additional transformations are necessary if the variables are correlated. The unknown coefficients may be estimated by various methods. One approach is to use the Galerkin method to minimize the residual obtained by requiring the residual to be orthogonal to the approximation space (Ghanem & Spanos, 1991). Another approach is to use the collocation method (Isukapalli & Georgopoulos, 1999). Here the model outputs are selected at a set of collocation points to estimate the coefficients. These collocation points can be the roots of the Hermite polynomial of a higher order. In this case, this way of estimating coefficients will capture points from regions of high probability (Tatang et al., 1997). Compared to the conventional response surface method, the stochastic response surface method has several advantages. The first advantage is that the polynomial chaos is a mean square convergent series expansion. This series is optimal in the Fourier sense, as it minimizes the mean square error from truncation after a finite number of the terms in the series (Ghanem & Spanos, 1991).

When the inputs are expressed as functions of standard normal variables and the number of the standard normal variables is defined as "number of the degrees of the freedom" in input uncertainty, the output are deterministic function of the the model inputs, so they have the same number of degrees of the freedom with respect to uncertainty. Polynomial chaos in terms of the standard normal variables provides a complete orthogonal basis for

the L^2 space (space for second order random function). Therefore any random function in the space can be represented by the polynomial chaos basis with corresponding coefficients. The second advantage is that the error from the truncation in the polynomial chaos series can be quantified. Thus, the accuracy of the metamodel itself can be quantified. This metamodel error will be accounted for in the response surface error. The third advantage is that polynomial chaos is a Hermite polynomial of random variables. The roots of the Hermite polynomial provide efficient collocation points to evaluate the model output and to compute the coefficients in the response surface. In general, the stochastic response surface method is founded on a solid mathematical basis and it is expected to provide a more accurate response surface model for predicting the response than the model from conventional RSM.

4.2.4 Bayesian Approach for Model Validation

When evaluating model uncertainty, if several possible models exist to describe a phenomenon, a Bayesian approach can be used to include all the candidate models by assigning a model weight (the probability of each model being correct) and integrating the effects of all models. When there is observation/data available, the model weights may be updated and shifted to the more appropriate model. This approach has been applied to probability distribution type uncertainty and linear regression model uncertainty problems in statistics (Edwards, 1984) (Soares, 1989) (Draper, 1995) (Volinsky et al., 1997), and was recently used to account for mechanical model uncertainty (Zhang & Mahadevan, 2000). The Bayesian approach can be adapted to problems in which there is only one model available, either due to the complexity of the problem, or because of the limited resources for model implementation. The Bayesian hypothesis testing uses assumptions on the prior distribution under the alternative hypothesis (Jeffreys, 1961).

A methodology for the problem of acceptance testing of mechanical reliability computation models, based on the concepts of Bayesian hypothesis testing was proposed in (Zhang & Mahadevan, 2003), for both time-independent and time-dependent problems, and also for problems both with and without statistical uncertainty. An outline of the Bayesian approach (only the case of time-independent problem without statistical uncertainty in the model) is reproduced below from (Zhang & Mahadevan, 2003), for clarity. First, consider two models M_i and M_j , with prior probabilities to be the correct model denoted by $P(M_i)$ and $P(M_j)$. By Bayes' rule, when an event/data Y is observed, the relative posterior probabilities of two hypotheses are obtained as:

$$\frac{P(M_i | \text{observation})}{P(M_j | \text{observation})} = \left[\frac{P(\text{observation} | M_i)}{P(\text{observation} | M_j)} \right] \left[\frac{P(M_i)}{P(M_j)} \right] \quad (85)$$

The data-dependent term in the first set of square brackets on the right hand side of Eq. (85) is called the "Bayes factor" (Jeffreys, 1961). This Bayes factor can be viewed as the "weighted" likelihood ratio of M_i to M_j and hence can be interpreted solely in terms of comparative support of the data for the two models. The data are said to favor M_i relative to M_j if the Bayes factor exceeds one, that is, if the observed data is more likely under hypothesis M_i than it is under hypothesis M_j . To make use of the Bayes' rule and Bayes factor in the case where only one model M is available, an alternative model M_a is constructed to represent the case in which the model M is not correct.

For the case without statistical uncertainty, the failure probability x of an engineering system predicted by model M is a deterministic value denoted by x_0 . It can be considered as a null hypothesis, under which the density of the parameter x is sharply concentrated at x_0 . To make use of Bayes' rule, we need to constitute an alternative hypothesis. Assume that there is no prior information about x . Therefore, the uniform distribution between $[0, 1]$ is taken as $P(x|M_a)$, the prior distribution under the alternative hypothesis. In the Bayesian framework, the distribution of x may be taken to be the weighted summation of the prediction of the two models:

$$P(x) = P(x|M)P(M) + P(x|M_a)(1 - P(M)) \tag{86}$$

where $P(M)$ is the prior probability that we believe model M is correct. Assuming $P(M)=a$, Eq. (86) may also be written as

$$P(x) = \begin{cases} a & (x = x_0) \\ (1-a)dx & (x \neq x_0) \end{cases} \tag{87}$$

where $x = x_0$ if the model M is correct and $x \neq x_0$ if the model M not is not correct. If n experiments are undertaken, and k failures are observed out of n samples, the probability of this result is:

$$P_n(k) = \binom{n}{k} x^k (1-x)^{n-k} \tag{88}$$

And the Bayes factor is obtained as

$$B = \frac{P(\text{data}|M)}{P(\text{data}|M_a)} = (n+1) \binom{n}{k} x^k (1-x)^{n-k} \tag{89}$$

Thus in this reliability prediction problem, Bayes factor B is a function of k and n , and is independent of the prior probability of the model being correct. The Bayes factor B in Eq. (89) attains its maximum at $x = k/n$ (Zhang & Mahadevan, 2003). Using the above-mentioned criterion for the Bayes factor, a decision can be made to reject or accept the model M . In what follows, several Bayes factors are evaluated, using prior error distribution obtained from response surface models. These RSM are metamodels for the error of the plate problem, defined in subsection 4.1. One quarter of the plate is modeled, generating corners between the interior hole and the lines representing the axis of symmetry. The discretization error in stress of the finite element model at a location near these corners is selected to be modeled using RSM.

With the availability of the distribution and the parameters for the prior error E , the probability that the computational model will fail to predict accurately or the error exceeding a desired value can be computed. In this work, this desired value is assumed to be 3% of the average value of the predicted stress. Thus, the prior estimate of the probability of failure of the model using one of the possible error estimators is given by $x = P(-0.03 \sigma < E < 0.03 \sigma)$.

4.2.4.1 Comparison of Bayes factors for different response surface methods

A 2D elastostatics problem was defined in subsection 4.1, consisting of a plate with a hole submitted to external loading. For the plate problem being studied, the region of

interest is one of the corners of the central hole. The octahedral stress at the centroid of the element closest to this corner is the response quantity of interest, and the error in this quantity is evaluated using the gradient recovery (GR) error estimator and the inter-element stress continuity (SC) error estimators. These estimators were shown previously to be consistent error estimators. The local GR and SC error results for the octahedral stress are modeled using conventional and stochastic response surface methods, and the Bayes factors from both cases are compared. The first-order conventional response surface for the error E in the response quantity of interest is given by

$$E = a_o + a_1W_1 + a_2W_2 \tag{90}$$

In this work, the input load variables are written in terms of standard normal variables ξ_1, ξ_2 as: $W_1 = 12 + 2.4\xi_1$ and $W_2 = 12 + 2.4\xi_2$, so that these conventional response surfaces can be written directly in terms of the corresponding standard normal variables for the input variables. The local error in the region of interest is written for the conventional response surface using the GR and SC error estimators as

$$\begin{aligned} E_C^{GR} &= a_{o_c}^{GR} + a_{1_c}^{GR} \xi_1 + a_{2_c}^{GR} \xi_2 \\ E_C^{SC} &= a_{o_c}^{SC} + a_{1_c}^{SC} \xi_1 + a_{2_c}^{SC} \xi_2 \end{aligned} \tag{91}$$

2-factorial experimental design is done, obtaining nine sample points, shown in Table 3.

Table 3: Sample points for fitting the conventional RSM. and are the gradient recovery and the stress continuity error estimates in the octahedral stress in the element centroid closest to the corner of the hole – adapted from (Jorge et al., 2002)

Point	ξ_1	ξ_2	w_1	w_2	E_C^{GR}	E_C^{SC}
1	-3	-3	4.8	4.8	0.1529	0.3658
2	3	3	19.2	19.2	0.6116	1.463
3	-3	3	4.8	19.2	0.8351	2.06
4	3	-3	19.2	4.8	2.054	3.188
5	0	0	12	12	0.3823	0.9145
6	0	-3	12	4.8	1.106	1.743
7	-3	0	4.8	12	-0.6515	0.9829
8	0	3	12	19.2	-0.3715	1.04
9	3	0	19.2	12	1.339	2.22

Polynomial regression is subsequently performed to fit these sample points into the corresponding curves. In this case, linear fitting is adopted, with equations given by:

$$\begin{aligned} E_C^{GR} &= 0.606433333 + 0.203783333 \xi_1 - 0.124316667 \xi_2 + \varepsilon_C^{GR} \\ E_C^{SC} &= 1.553022222 + 0.19235 \xi_1 - 0.040766667 \xi_2 + \varepsilon_C^{SC} \end{aligned} \tag{92}$$

where ε_C^{GR} and ε_C^{SC} are the residual errors of the regression model adopted, which follow normal distribution with zero mean values and standard deviations of 0.441936688333333 and 0.622834690092593, respectively. Also, the R^2 values for these linear-fit equations are 0.537131 and 0.358417, respectively, showing the poor fit of this set of data. The first-order stochastic response surfaces for the errors E_S^{GR} and E_S^{SC} are written, respectively, in terms of the standard normal variables, as

$$\begin{aligned} E_S^{GR} &= a_{o_s}^{GR} + a_{1_s}^{GR} \xi_1 + a_{2_s}^{GR} \xi_2 \\ E_S^{SC} &= a_{o_s}^{SC} + a_{1_s}^{SC} \xi_1 + a_{2_s}^{SC} \xi_2 \end{aligned} \tag{93}$$

The samples for the linear regression are shown in Table 4. These collocation points are obtained as roots of the Hermite polynomial of one degree higher than the polynomial expansion being used. There are 25 possible points to be generated, for each root of the Hermite polynomial. The first nine points are adopted to generate the linear regression model. These points correspond to the nine roots closest to the origin. The remaining points can be used to generate a more accurate linear model for comparison purposes.

Table 4: Sample points for fitting the stochastic RSM. and are the gradient recovery and the stress continuity error estimates in the octahedral stress in the element centroid closest to the corner of the hole – adapted from (Jorge et al., 2002)

Point	ξ_1	ξ_2	w_1	w_2	E_S^{GR}	E_S^{SC}
1	1	1	14.4	14.4	0.4587	1.097
2	1	0	14.4	12	0.7025	1.316
3	1	-1	14.4	9.6	0.9442	1.591
4	0	1	12	14.4	0.1355	0.8039
5	0	0	12	12	0.3823	0.9145
6	0	-1	12	9.6	0.6259	1.139
7	-1	1	9.6	14.4	-0.1942	0.7466
8	-1	0	9.6	12	0.05853	0.6406
9	-1	-1	9.6	9.6	0.3058	0.7316
10	1	$\sqrt{3}$	14.4	16.16	0.2782	0.9952
11	1	$-\sqrt{3}$	14.4	7.84	1.12	1.814
12	0	$\sqrt{3}$	12	16.16	-0.0481	0.8244
13	0	$-\sqrt{3}$	12	7.84	0.8028	1.345
14	-1	$\sqrt{3}$	9.6	16.16	-0.3841	0.9154
15	-1	$-\sqrt{3}$	9.6	7.84	0.4845	0.8942
16	$\sqrt{3}$	1	16.16	14.4	0.6939	1.383
17	$\sqrt{3}$	0	16.16	12	0.9363	1.639
18	$\sqrt{3}$	-1	16.16	9.6	1.177	1.932
19	$\sqrt{3}$	$\sqrt{3}$	16.16	16.16	0.5148	1.232
20	$\sqrt{3}$	$-\sqrt{3}$	16.16	7.84	1.353	2.161
21	$-\sqrt{3}$	1	7.84	14.4	-0.4466	0.889
22	$-\sqrt{3}$	0	7.84	12	-0.1842	0.6295
23	$-\sqrt{3}$	-1	7.84	9.6	0.06859	0.5233
24	$-\sqrt{3}$	$\sqrt{3}$	7.84	16.16	-0.6477	1.119
25	$-\sqrt{3}$	$-\sqrt{3}$	7.84	7.84	0.2497	0.5975

Polynomial regression is subsequently performed to fit these sample points into the corresponding curves.

In this case, linear fitting using the first nine points is adopted, with equations given by:

$$\begin{aligned}
 E_S^{GR} &= 0.379914 + 0.322545\xi_1 - 0.245983333\xi_2 + \varepsilon_S^{GR} \\
 E_S^{SC} &= 0.9978 + 0.3142\xi_1 - 0.135683333\xi_2 + \varepsilon_S^{SC}
 \end{aligned}
 \tag{94}$$

where ε_s^{GR} and ε_s^{SC} are the residual errors of the regression model adopted, which follow normal distribution with zero mean values and standard deviations of 0.0000116068675925925 and 0.0143696230555556, respectively. Also, the R^2 values for these linear-fit equations are 0.999929465 and 0.890726, respectively, showing the good linear fit for this set of data. The prior estimate for the probability of failure (x) of the model is evaluated for both the conventional and the stochastic response surfaces, for the first order models. In all cases, the maximum allowed error is assumed to be 0.72ksi, which corresponds to 3% of the average predicted mean stress in the location of interest, the corner of the hole. The mean value for the local error in the obtained equations using the error estimates should be zero, corresponding to the case with no external loading. Equations (92) and (94) show that the linear fitting adopted for the error estimators presents a bias from zero. This mean value represents the inability of the error estimator in representing the zero-mean property of the real error. The real error is assumed to have zero mean with same variability than the estimated errors. Thus, to evaluate the probability of failure, the mean value of the error is dropped from Equations (92) and (94). After evaluating the probability of failure, a set of $n=10$ experiments is conducted. Failure occurs when the predicted and the observed stresses at the location of interest differ more than 3%. Assuming that failure occurs in a number k of these experiments ($k=0, \dots, 10$), the Bayes factors can be evaluated using Eq. (89). The tests are based on the pass/fail criterion and hence the test data k (number of failures) follows binomial distribution with the parameters n (number of tests) and the probability of failure of the each test given by x . Table 5 shows the results for the probability of failure x and for the corresponding Bayes factors for each value of x .

Table 5: Prior estimates of probability of failure (x) and corresponding Bayes factors (B) for various k values. Probabilities of failure x obtained for the cases of conventional (C) and stochastic (S) RSM, first order model, for the GR and SC estimators of error in the octahedral stress in the region of interest – adapted from (Jorge et al., 2002)

K	B_C^{GR}	B_C^{SC}	B_S^{GR}	B_S^{SC}
	$x = 0.3112$	$x = 0.3741$	$x = 0.0731$	$x = 0.0469$
0	0.264438	0.101494	5.148984	6.804205
1	1.194732	0.606631	4.060748	3.348203
2	2.429011	1.631625	1.44113	0.74141
3	2.926474	2.600589	0.303079	0.097289
4	2.313818	2.720148	0.041829	0.008378
5	1.25446	1.950997	0.003959	0.000495
6	0.472304	0.971758	0.00026	2.03E-05
7	0.121936	0.331897	1.17E-05	5.7E-07
8	0.020659	0.07439	3.47E-07	1.05E-08
9	0.002074	0.009881	6.08E-09	1.15E-10
10	9.37E-05	0.000591	4.79E-11	5.66E-13

As shown in Table 5, the conventional and stochastic RSM lead to different results in terms of probability of failure. The large difference between the probabilities of failure for the conventional and stochastic RSM can be explained by the poor linear fitting in the conventional RSM case. Thus, only the results of probability of failure for the stochastic response surface case are acceptable. In this case, the Bayes factors for both error estimators evaluated (GR and SC) are consistently higher than 1 for $K=0$ and $K=1$, are in close vicinity of 1 for $K=2$, and are below 1 for $K=3$ and beyond.

4.2.4.2 Comparison of Bayes factors for consistent error estimators

In this subsection, the Bayes factors are compared for different consistent error estimators applied to different independent response quantities. The stochastic RSM is used to generate response surfaces for two different error estimators, the gradient recovery (GR) error estimator and the error estimator obtained from the lack of inter-element continuity (SC) for the stress. The errors in two independent response quantities, the mean stress and the octahedral stress in the region of interest, are evaluated. The linear expressions for the error using the GR and SC error estimators for the octahedral stress were already obtained in Equations (94). Table 6 shows the GR and SC error data for the mean stress.

Table 6: Sample points for fitting the stochastic RSM. $E_{S,Mean}^{GR}$ and $E_{S,Mean}^{SC}$ are the gradient recovery and the stress continuity error estimates in the mean stress in the element centroid closest to the corner of the hole – adapted from (Jorge et al., 2002)

Point	ξ_1	ξ_2	w_1	w_2	$E_{S,Mean}^{GR}$	$E_{S,Mean}^{SC}$
1	1	1	14.4	14.4	0.8989	0.8027
2	1	0	14.4	12	1.099	1.162
3	1	-1	14.4	9.6	1.298	1.521
4	0	1	12	14.4	0.5494	0.3096
5	0	0	12	12	0.7491	0.6689
6	0	-1	12	9.6	0.9488	1.028
7	-1	1	9.6	14.4	0.1998	-0.1834
8	-1	0	9.6	12	0.3995	0.1758
9	-1	-1	9.6	9.6	0.5992	0.5351
10	1	$\sqrt{3}$	14.4	16.16	0.7524	0.5392
11	1	$-\sqrt{3}$	14.4	7.84	1.445	1.785
12	0	$\sqrt{3}$	12	16.16	0.4029	0.04615
13	0	$-\sqrt{3}$	12	7.84	1.095	1.292
14	-1	$\sqrt{3}$	9.6	16.16	0.05339	-0.4469
15	-1	$-\sqrt{3}$	9.6	7.84	0.7457	0.7986
16	$\sqrt{3}$	1	16.16	14.4	1.155	1.164
17	$\sqrt{3}$	0	16.16	12	1.355	1.524
18	$\sqrt{3}$	-1	16.16	9.6	1.555	1.883
19	$\sqrt{3}$	$\sqrt{3}$	16.16	16.16	1.009	0.9008
20	$\sqrt{3}$	$-\sqrt{3}$	16.16	7.84	1.701	2.146
21	$-\sqrt{3}$	1	7.84	14.4	-0.05646	-0.545
22	$-\sqrt{3}$	0	7.84	12	0.1432	-0.1857
23	$-\sqrt{3}$	-1	7.84	9.6	0.3429	0.1736
24	$-\sqrt{3}$	$\sqrt{3}$	7.84	16.16	-0.2029	-0.8085
25	$-\sqrt{3}$	$-\sqrt{3}$	7.84	7.84	0.4894	0.437

Polynomial regression is performed to fit the sample points for the gradient recovery (GR) and stress continuity (SC) error estimators into the corresponding curves. In this case, linear fittings are adopted, with equations given by:

$$\begin{aligned}
 E_{S,Mean}^{GR} &= 0.749078 + 0.349567 \xi_1 - 0.19965 \xi_2 + \varepsilon_{S,Mean}^{GR} \\
 E_{S,Mean}^{SC} &= 0.668856 + 0.493033 \xi_1 - 0.3592 \xi_2 + \varepsilon_{S,Mean}^{SC}
 \end{aligned}
 \tag{95}$$

where $\varepsilon_{S,Mean}^{GR}$ and $\varepsilon_{S,Mean}^{SC}$ are the residual errors of the regression model adopted, which follow normal distribution with zero mean values and standard deviations of 3.9×10^{-8} and 5.93×10^{-9} , respectively. Also, the R^2 values for these linear-fit equations are both equal to 1, showing the good linear fit for this set of data.

The prior estimate for the probability of failure (x) of the model is evaluated for the first order models for the stochastic response surface. Again, the maximum allowed error is assumed to be 0.72ksi, and a set of $n=10$ experiments is conducted, with failure occurring in k of these experiments ($k=0, \dots, 10$). The Bayes factors are evaluated using Eq. (89). Table 7 shows the results for the probability of failure x and for the corresponding Bayes factors for each value of x .

Table 7: Prior estimates of probability of failure (x) and corresponding Bayes factors (B) for various k values. Probabilities of failure x obtained for the gradient recovery (GR) and stress continuity (SC) error estimators, for first order stochastic RSM, for two response quantities: mean and octahedral stress – adapted from (Jorge et al., 2002)

K	$B_{S,Oct}^{GR}$	$B_{S,Oct}^{SC}$	$B_{S,Mean}^{GR}$	$B_{S,Mean}^{SC}$
	$x = 0.0731$	$x = 0.0469$	$x = 0.0706$	$x = 0.2333$
0	5.148984	6.804205	5.289558	0.772053
1	4.060748	3.348203	4.018106	2.34929
2	1.44113	0.74141	1.373523	3.216906
3	0.303079	0.097289	0.278232	2.610336
4	0.041829	0.008378	0.036987	1.390029
5	0.003959	0.000495	0.003372	0.507568
6	0.00026	2.03E-05	0.000213	0.128707
7	1.17E-05	5.7E-07	9.26E-06	0.02238
8	3.47E-07	1.05E-08	2.64E-07	0.002554
9	6.08E-09	1.15E-10	4.45E-09	0.000173
10	4.79E-11	5.66E-13	3.38E-11	5.25E-06

As shown in Table 7, different (but consistent) error estimators evaluated for different (and independent) response quantities lead to similar results in terms of probability of failure in most of the cases tested. In three of the four cases tested, the Bayes factors for both error estimators evaluated (GR and SC) are consistently higher than 1 for $K=0$ and $K=1$, are in close vicinity of 1 for $K=2$, and are below 1 for $K=3$ and beyond. Only the case for the SC error in the mean stress presented a higher probability of failure, with different conclusions in terms of Bayes factors. For the majority of the cases tested, similar model acceptance results are obtained with consistent error estimators.

4.2.4.3 Comparison of Bayes factors for non-consistent error estimators

Two error estimators, Richardson extrapolation (R) and maximum error bound (MEB) were previously found to be non-consistent. The octahedral stress near the corner node at the plate hole is used as the output response quantity to be modeled, while the input quantities are the loads W1 and W2 given in Table 1, above. Similarly to the previous subsection, results can be later obtained for the R and MEB error data for the octahedral stress, and are not included in this chapter, at this time.

4.2.5 Some Remarks on Error Estimators and Bayesian Hypothesis Testing in a Model Validation

The results obtained in this preliminary study indicate that:

- stochastic RSM is might be preferable than conventional RSM (more cases need to be studied to confirm this statement);
- consistent error estimators appear to lead to same decisions on model acceptance. On the other hand, non-consistent error estimators might lead to different probabilities of failure of the structure and consequently to different model acceptance decisions. Thus, a key part in the model validation procedure is the appropriate selection of a consistent set of:
 - a) weak location;
 - b) response quantity; and
 - c) error estimator.

4.3 Perspectives for Future work in Error Estimators and Bayesian Hypothesis Testing in Model Validation

The perspectives for future work in error estimators and Bayesian hypothesis testing in Model Validation appear quite promising, and several important topics for future research can be envisaged, such as:

- Selection criteria among error estimators;
- Including measurement error in model validation;
- Extrapolation of validation decision parameters to unmeasured regions;
- Extrapolation of input parameters to regions where data is not available;
- Consistency in model validation decision across various response quantities;
- Incorporate random field nature of error in model validation;
- Quantifying confidence in validation;
- Extension of consistent error estimator technique to time-dependent problems.

5 Perspectives: Validation Techniques in FEM Structural Models

This section presents potential applications in FEM structural modeling that may benefit from the model validation techniques addressed throughout the chapter. Among the selected topics, the reader will find brief discussions on FE model condensation problems, models for structural health monitoring applications in aeronautical structures, modeling of structures obtained via additive manufacturing, and composite modeling for aerospace applications.

5.1 Perspectives for FEM Condensation Models

The recent advances in computer technology have made it possible to build extremely detailed FE models for structural applications. It is not uncommon to come across FE models with hundreds of thousands of degrees of freedom. However, the development of a reliable model often involves validation, updating, or adjustment steps, usually through the correlation of numerical and experimental models. Such experimental models are significantly smaller in size when compared to FE models. There are several methodologies developed over the past few decades that address the problem of reducing the number of degrees of freedom in a numerical model for a variety of purposes.

Figure 32 illustrates numerical and experimental models developed for the structural integrity investigations of the aluminum-aluminum honeycomb sandwich panel used in the construction of the Brazilian Geostationary Satellite. Detailed finite element models were built to study different aspects of the panel structural integrity when subjected to the static, dynamic, and environmental operational loads. Figure 32 (b) shows a small portion of a very detailed 3-D model built with shell elements, with a very large number of degrees of freedom, even for smaller specimens used in the experimental models. Figure 32 (d) presents one typical configuration for an experimental modal analysis that produced a good experimental model with only 136 degrees of freedom (Domingues, 2019). Once the test is carefully conducted with all the required expertise, the experimental model is a much more reliable representation of the real behavior of the panel. However, for detailed design simulations, the more refined model is extremely useful. The task is then to validate the numerical model using the experimental one. In order to allow an efficient correlation of the two models and to adjust or update the numerical model, reduction techniques must be used (Wang et al., 2017). The process of condensing a numerical model will certainly introduce errors that may impose a tremendous challenge that demands the applications of the theory and methodologies discussed in the present chapter in order to minimize possible errors induced by the process of model reduction.

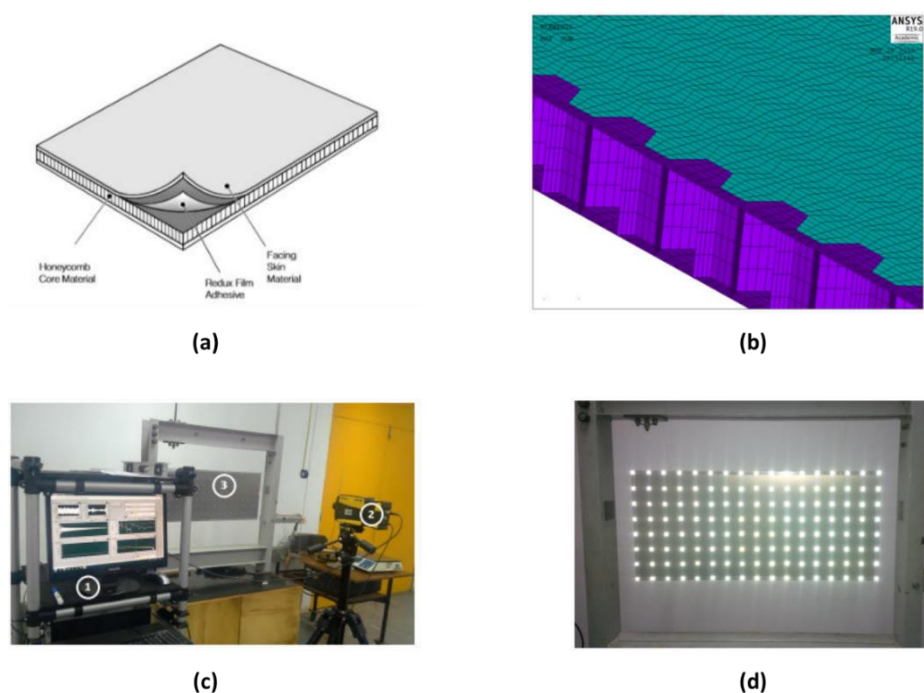


Figure 32: Size differences for numerical and experimental models of a Al-Al honeycomb sandwich panel (a) schematic representations of the panel; (b) detailed finite element model with an order of 10^5 degrees of freedom. (c) experimental apparatus for a modal analysis test; (d) measurement points that generate an experimental dynamic model of 136 degrees of freedom.

5.2 Perspectives for Aeronautical Structures Applications

Structural damage detection methods, sometimes also referred to as Structural Health Monitoring (SHM) methods have received tremendous attention from researchers in the

past few decades. Although some SHM techniques have been implemented in real structures over the years, most practical problems still need further development and offer interesting research challenges and opportunities. The ideal solution of having one general, fully automated, and dependable SHM method which is capable of solving damage problems in a wide variety of structural systems doesn't seem to be feasible in the foreseeable future. Detecting the presence of damage may be possible with a limited amount of measured physical quantities that may describe the structure status during operation, such as strains, deflections, accelerations, etc. However, due to the intrinsic nature of the problem, the tasks of localizing and quantifying the damage are much more complex and usually will require the inclusion of a dependable, validated model of the structure, usually a numerical model. Such a model will be most likely used in the solution of an inverse problem, it will be evaluated iteratively in some sort of "error" minimization procedure, either classical or heuristic. Therefore, the models need to be, at the same time, reliable and computationally light. In that sense, the subject of the present chapter is of utmost importance. The next two subsections present potential applications in current research for damage identification in aeronautical structures.

5.2.1 Perspectives for Modeling of Structural Damage Detection/Localization Problems

The problem of localizing damage in aeronautical structures has gained attention in recent years. Many different strategies have been developed in order to solve the corresponding inverse problem. One interesting approach is the use of wavelet transform to treat measurements of the behavior of the potentially damaged structure (Katunin, 2010, 2021). For example, the deflection caused by the operational loads can be monitored to develop an indicator that can reveal the existence and location of changes in the physical characteristics of the structure due to the existence of damaged areas.

Figure 33 illustrates the preliminary results of an ongoing investigation on a discrete wavelet transform (DWT) based method for localizing damage in a beam-like structure, modeled by finite elements. In the depicted example, the beam is discretized with 100 elements and 101 nodes and fixed at both ends. The presence of damage is simulated as a local reduction of stiffness in one single element. Static deflections of the damaged and undamaged beams are calculated for different damage values and locations, and different loading configurations.

A damage indicator was developed using the normalized differences between the calculated coefficients of the DWT of the deflections for the damaged and undamaged beam over its entire length. Several combinations of wavelets and levels of decomposition were tested in order to establish the best candidates for the implementation of the method. The SYM8 wavelet at one level of decomposition was found to be the more suitable choice for implementing the method. The index is computed along the entire FE mesh of the structure capturing the statistical variations of the index in the near vicinity of each node by computing the median and the boxplot height (range from the first and third quartiles) of the normalized differences of the DWT coefficients.

It can be noticed from Figure 33 the potential of the proposed index for damage detection and localization. It should be noted that damage is successfully detected and localized even for small amounts of damage. In such scenarios, the deflection curves of the damaged and the undamaged beams are virtually indistinguishable. The DWT-based damaged index is able to provide a "statistical amplification" of the effect of the damaged

element over the structural behavior of the entire beam. Further investigation of the method is aimed at extending the idea to more complex geometries and including DTW of dynamic responses in the calculation of the statistical damage index.

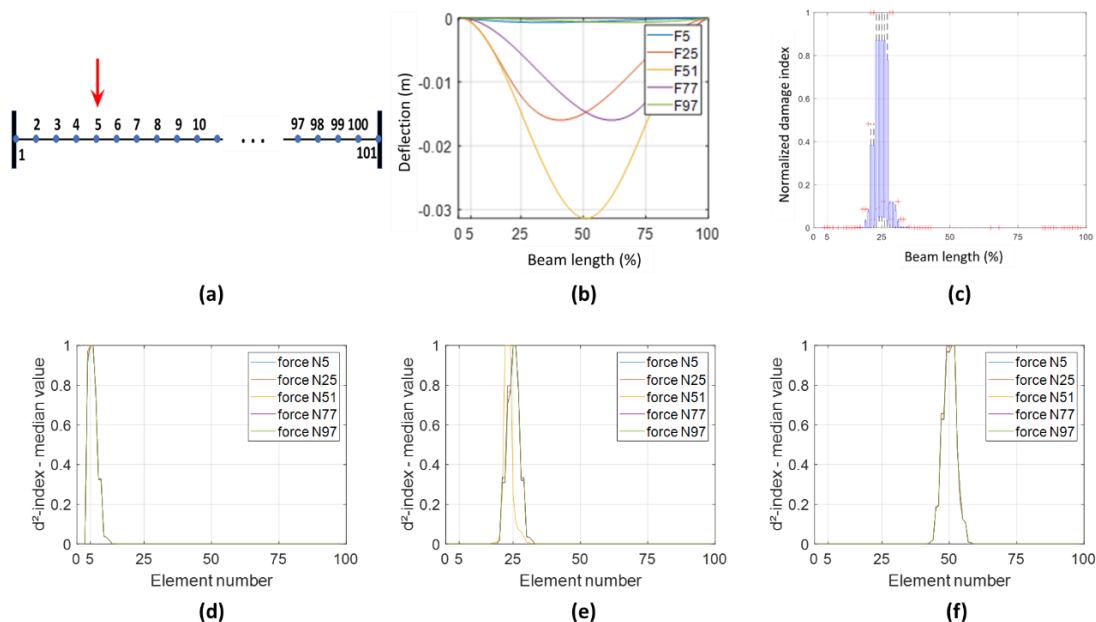


Figure 33: DWT based damage localization in a double clamped beam (a) Mesh with 101 nodes, 100 elements and force at node 5 (b)deflection plots; (c) typical boxplot for 50% damage at 25% beam length; (d), (e), (f) normalized median of damage index for different loading positions and 10% damage at 5%, 25% and 50% of beam length, respectively.

5.2.2 Perspectives for Modeling of Damage Detection/Identification Problems in Structural Aeronautical Fasteners / Joints

Consider the problem of identifying possible failure of fasteners in a structural joint. One possible approach is to monitor the structures in order to evaluate eventual changes in their dynamic behavior (Sinou, 2009). Measured data may be used in the solution of a model-based inverse problem in order to identify potential damage in the joint. Figure 34 depicts the conceptual construction of a simplified, low-fidelity model representing an aircraft wing. The model is intended to be light enough to feed an interactive numerical procedure to solve the inverse problem but must be able to include a mathematical description of both the distributed properties of the wing, in this case using beam finite elements, and the mechanical properties of the joint, modeled as concentrated elastic elements representing the transverse and rotational stiffnesses of the connection. These stiffness parameters can be related to a more detailed model of the joint, which can include geometrical, and material information, as well as torque of a series of bolts, for instance. The model can be used to reproduce the dynamic responses of the wing when submitted to controlled excitations in different positions of its span. The time responses of the structure can be monitored by a series of sensors over the wing. This simplified model is extremely useful for simulations in the development of SHM strategies, allowing for the investigation of types of excitation functions, number and position of excitors and sensors,

selection of damage index candidates based on the ability to evidence damage under uncertainties avoiding false positive and false negative results, etc.

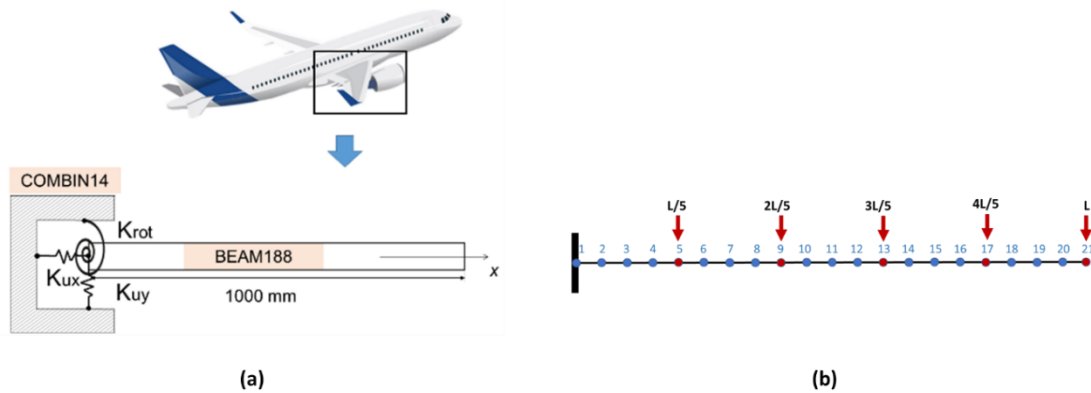


Figure 34: Simplified model of an aircraft wing for SHM procedures. (a) conceptual representation of the joint; (b) finite element mesh with beam elements and representation of excitation pints.

Figure 35 illustrates the study of a damage indication index based on the normalized differences of time responses measured over the wing for damaged and undamaged joints. It can be noticed that for the configurations studied, the damaged index is sensitive to variations in the joint integrity for different combinations of excitation and measurement points, as illustrated by Figures 35 (a) through (c). For all the simulated cases, confidence-bound curves were obtained, correlating the variations in damage index to the loss of torque in a simulated bolted joint. Figure 35 (d) indicates that results show low dispersion of the results for a 95% confidence level, indicating that the proposed strategy is a viable candidate for more detailed studies aiming at a reliable SHM procedure for aeronautical structures with fastened joints.

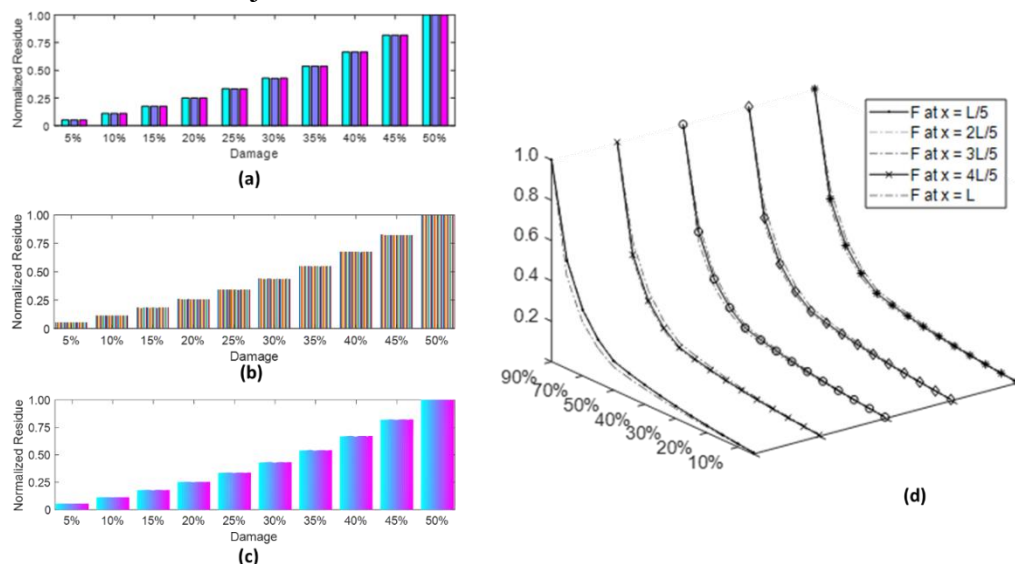


Figure 35: Damage index based on normalized differences of time responses for one excitation point and equally spaced sensors over the wing span (a) 3 sensors; (b) 5 sensors (c) 21 sensors (d) Confidence bound curves for 95% confidence level for all sensor configurations and different excitation points.

5.3 Perspectives in FEM Models for Additive Manufacturing in Aeronautical / Aerospace / Materials Applications

Additive manufacturing (AM), also known as three-dimensional printing, is an innovative manufacturing process based on the deposition of material in a localized manner, originating directly from a CAD/CAE model. This enables the fabrication of complex structures using metallic, ceramic, polymer and hybrid materials. The developed AM processes are the technique of Fused Deposition Modeling (FDM) (Gibson et al. 2010), Selective Laser Sintering (SLS), Selective Laser Sintering (SLS), Selective Laser Melting (SLM) (Gibson et al. 2010, Smith et al. 2014, Smith et al. 2014), and Electron Beam Melting (EBM) (Smith et al. 2017). The current AM technology for metal alloys enables the construction of complex monolithic parts, with high mechanical strength and precision, being possible to be used in operating conditions. In addition, it is important to emphasize that the use of the material is optimized and the production of waste is minimal or often nil, which are positive aspects of the environmental impact of the project. Currently, the negative aspects of the technique are the time and high cost of production; surface quality of the part, and the uncertainties generated by the AM process. Despite being extremely applicable and versatile, AM in metals has brought several challenges in the uncertainties field that must be explored to expand their application in primary structures. Some challenges concerning the MA process can be treated by the methods introduced in this chapter, such as:

- In the influence of manufacturing parameters. The material behavior is sensitive to the set of parameters, such as electron beam power, number of exposures, chamber temperature, layer width among others. Some studies (Lopez et al. 2016 and Hu & Mahadevan 2017) have been addressed the uncertainties associated with mechanical properties due to the aforementioned diversity of manufacturing parameters. Another factor directly affected by manufacturing parameters is the dimensional tolerance of the final part, generally affected by thermal expansion caused by beam power and speed.
- In the computational modeling of the additive manufacturing process. Numerical modeling is attractive in the study of the parameters influence on the final part to be printed. From the analysis of the input parameters for the print, it is possible to evaluate possible final distortions in the geometry, such as the one caused by the beam power responsible for generating thermal expansion and effect of warping in the final geometry (Prabhakar et al. 2015). Figure 36 depicts this type of analysis, where as a next step the behavior of the printed geometry can be numerically evaluated in relation to other engineering problems, such as modal analysis, hysteresis effect, fatigue, among others. The use of numerical simulation makes the experimental AM process more assertive as long as the associated uncertainties are addressed, generating a more efficient and financially viable prototype.

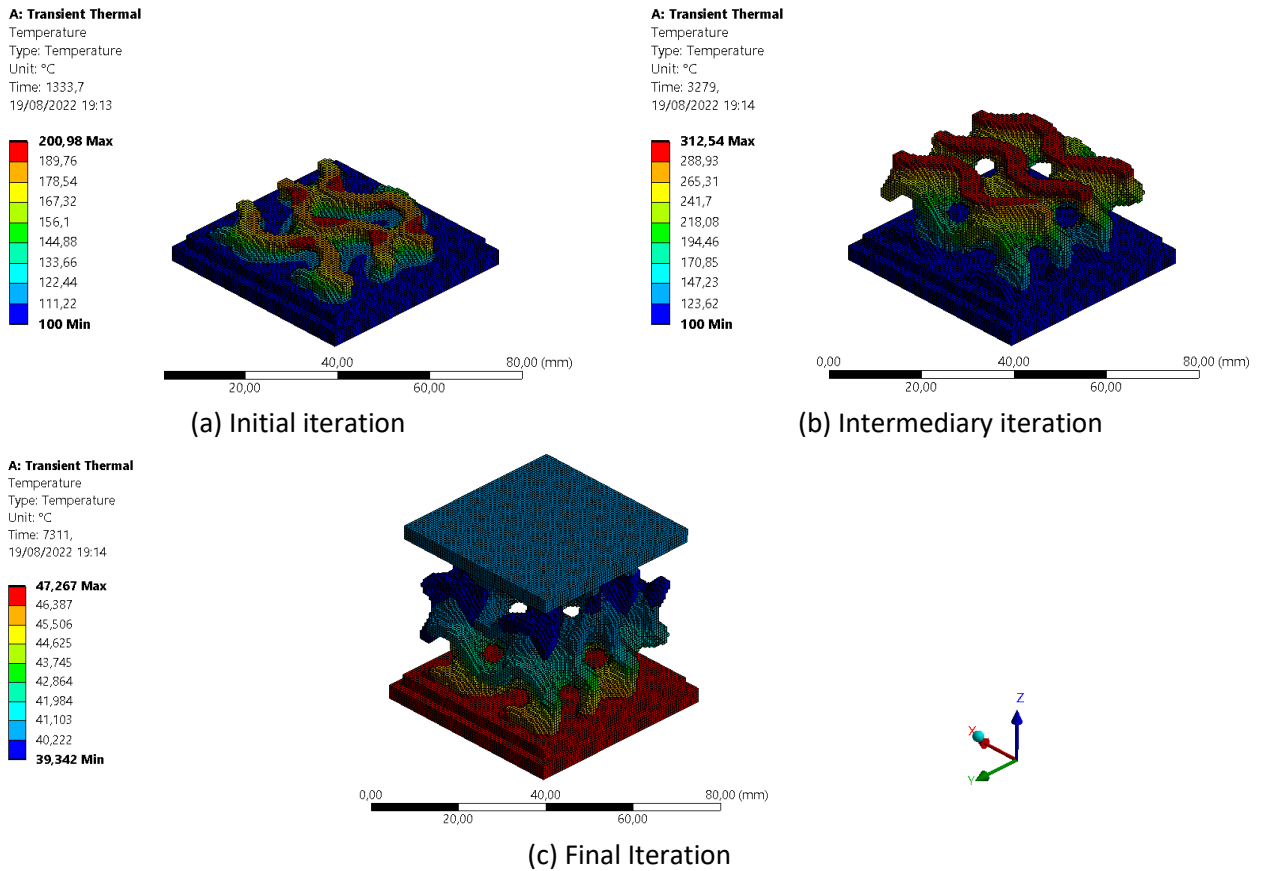


Figure 36: MA modeling by Finite Element Method.

- In the development of complex cellular structures called gyroids. These structures are configuration examples for combinations via MA because of their complexity. This type of structure is indicated for impact absorption (Ramos et al. 2019), being generally manufactured by MA (Figure 37) in aluminum alloys (SLM process). The uncertainties are straight associated in the mechanical performance of the structure, once the geometric parameters such as porosity, width and helix angle affect in the capacity for energy absorption.

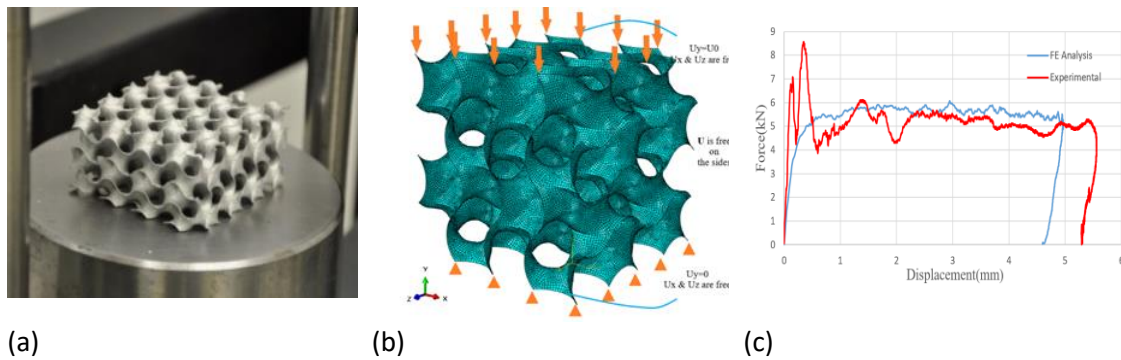


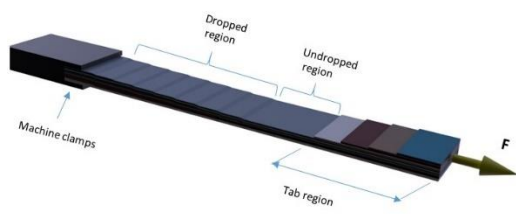
Figure 37: Gyroid-like cell structure: (a) MA fabricated gyroid structure, (b) numerical model of the structure and, (c) numerical-experimental comparison of gyroid behavior at impact. Ramos et al. (2019).

5.4 Perspectives for Composite Modeling in Aeronautical / Aerospace Applications

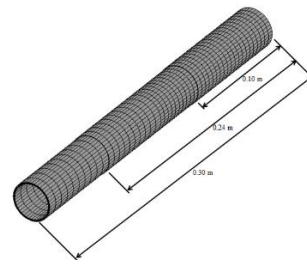
The field of studies in composite materials with or without aeronautical application is vast. It is extremely important to address modeling uncertainties in these materials in order to obtain increasingly robust (numerical) models that are faithful to reality.

Some examples and perspectives of application of the methods discussed in this chapter can be cited, namely:

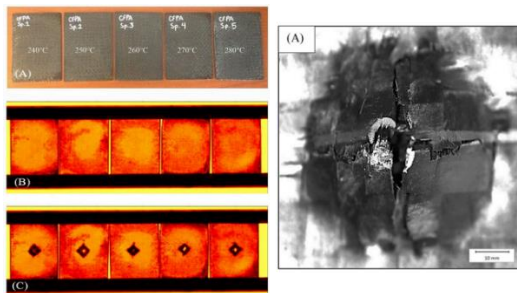
- In the modeling of composite structures in the presence of drop-off ply. These advanced structures are important in the field of composite materials due to their variable thickness. However, interrupting layers can lead to undesirable failure modes (Diniz et al., 2021; Kane et al., 2022; Diniz et al., 2022), as shown in Figure 38(a) and 38(b).
- In the development of materials with high impact absorption capacity. Interest in the use of structural composite materials in components for the automotive and aerospace industry has been increasing considerably in recent years. For this reason, improvements on structural composites crashworthiness are directly related to the human safety in a possible automotive crash and it depends on the energy absorption capability of the composite components (Di Benedetto et al., 2019; Di Benedetto et al., 2021; Di Benedetto et al., 2022), the examples shown in Figure 38(c).
- In the study of aeroelasticity in composite materials. The companies' incessant search for lighter aircraft has led to the use of several new technologies and design solutions in new aircraft models. Two of them have been the use of composites and more flexible structures. However, such structures lead to two main problems: (i) difficulty in identifying damage and defects in composite structures and (ii) very flexible aircraft are more subject to flutter. In view of the new challenges imposed, the present work seeks to assess the influence of structural reduction caused by damage/defects in the flutter speed. (Mendes & Gomes, 2022), as examples shown in Figure 38(d)
- It is also known that the dynamic behaviors of mechanical structures, particularly composite ones, are critical to the effective operation of numerous applications such as automotive, civil and aeronautic industries. With appropriate structural behavior knowledge, failure can be avoided, and undesirable resonance can also be eliminated or even reduced. In this sense, it is extremely important to effectively understand the model and associated uncertainties in the modeling of sandwich structures with a core composed of magnetorheological material (de Souza Eloy *et al.*, 2018; de Souza Eloy *et al.*, 2019, Faria *et al.*, 2020), as examples shown in Figure 38(e).



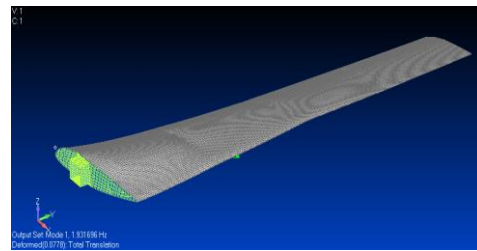
(a) Drop-off modelling
(Kane *et al.*, 2022)



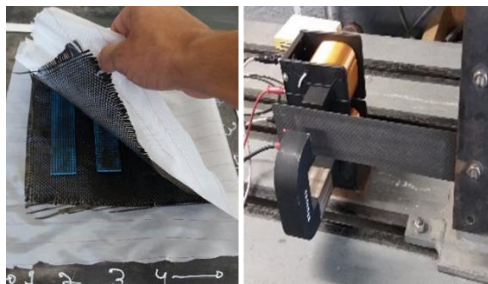
(b) Drop-off modelling in tubes
(Diniz *et al.*, 2022)



(c) Impact absorption
(Di Benedetto *et al.*, 2019)



(d) Flutter in composite wings
(Mendes & Gomes, 2022)



(e) Sandwich composite structure (de Souza Eloy *et al.*, 2018)

Figure 38: Some examples of application perspectives in composite structures.

6 Concluding Remarks

This chapter presented an overview of computational models and formulations, a description of some error estimators in Boundary Element Methods (BEM) and in Finite Element Methods (FEM), and some perspectives of the use of these error estimators in Bayesian approaches for Model Validation.

Acknowledgements

The first author would like to acknowledge the support from the Structural Reliability Research Group at Vanderbilt University, led by Dr. Sankaran Mahadevan, where he was a Post-doctoral Research Associate in 2002 (*Structural Reliability Research Group - Faculty and Staff*, n.d.), and also to express this gratitude to Dr. Prodyot K. Basu, Civil Engineering, for the continuous Department support, advising, and assistance, during the first author's Post-Doctoral research period at Vanderbilt University, in 2009-2010.

References

- Aboudi, J. (1991). Mechanics of composite materials: a unified micromechanical approach. *Studies in Applied Mechanics*, 29.
- Aboudi, J. (1996). Micromechanical analysis of composites by the method of cells - Update. *Applied Mechanics Reviews*, 49(10). <https://doi.org/10.1115/1.3101981>
- Ainsworth, M., & Oden, J. T. (1993). A unified approach to a posteriori error estimation using element residual methods. *Numerische Mathematik*, 65(1). <https://doi.org/10.1007/BF01385738>
- Ainsworth, M., & Oden, J. T. (1997). A posteriori error estimation in finite element analysis. In *Computer Methods in Applied Mechanics and Engineering* (Vol. 142, Issues 1–2). [https://doi.org/10.1016/S0045-7825\(96\)01107-3](https://doi.org/10.1016/S0045-7825(96)01107-3)
- Aldroubi, A., & Papadakis, M. (1998). *Characterization and parameterization of multiwavelet bases*. <https://doi.org/10.1090/conm/216/02967>
- Aldroubi, A., & Unser, M. (1993). Families of multiresolution and wavelet spaces with optimal properties. *Numerical Functional Analysis and Optimization*, 14(5–6). <https://doi.org/10.1080/01630569308816532>
- Aluru, N. R., & Li, G. (2001). Finite cloud method: A true meshless technique based on fixed reproducing kernel approximation. *International Journal for Numerical Methods in Engineering*, 50(10). <https://doi.org/10.1002/nme.124>
- Ammons, B. A., & Vable, M. (1998). An HR-method of mesh refinement for boundary element method. *International Journal for Numerical Methods in Engineering*, 43(6). [https://doi.org/10.1002/\(SICI\)1097-0207\(19981130\)43:6<979::AID-NME451>3.0.CO;2-J](https://doi.org/10.1002/(SICI)1097-0207(19981130)43:6<979::AID-NME451>3.0.CO;2-J)
- Ang, A. H.-S., & Tang, W. H. (1984). *Probability Concepts in Engineering Planning and Design, Vol. 2: Decision, Risk, and Reliability*. John Wiley & Sons.
- Armando Duarte, C., & Tinsley Oden, J. (1996). H-p Clouds - An h-p Meshless Method. *Numerical Methods for Partial Differential Equations*, 12(6). [https://doi.org/10.1002/\(sici\)1098-2426\(199611\)12:6<673::aid-num3>3.0.co;2-p](https://doi.org/10.1002/(sici)1098-2426(199611)12:6<673::aid-num3>3.0.co;2-p)
- Arora, J. S. (2012). Introduction to Optimum Design Chapter 2. In *Introduction to Optimum Design*.
- Babuška, I., & Melenk, J. M. (1997). The partition of unity method. *International Journal for Numerical Methods in Engineering*, 40(4). [https://doi.org/10.1002/\(SICI\)1097-0207\(19970228\)40:4<727::AID-NME86>3.0.CO;2-N](https://doi.org/10.1002/(SICI)1097-0207(19970228)40:4<727::AID-NME86>3.0.CO;2-N)
- Babuška, I., Strouboulis, T., Upadhyay, C. S., & Gangaraj, S. K. (1995). A model study of element residual estimators for linear elliptic problems: The quality of the estimators in the interior of meshes of triangles and quadrilaterals. *Computers and Structures*, 57(6). [https://doi.org/10.1016/0045-7949\(95\)00075-R](https://doi.org/10.1016/0045-7949(95)00075-R)
- Babuška, Ivo, Banerjee, U., & Osborn, J. E. (2003). Survey of meshless and generalized finite element methods: A unified approach. In *Acta Numerica* (Vol. 12). <https://doi.org/10.1017/S0962492902000090>

- Basu, P. K. (2002). Multiphysics/Multiscale Modeling in Engineering Design. *Proceedings of Second Annual International Conference on Advances in Structural Engineering and Mechanics*.
- Basu, P. K. (2007). Modeling and simulation in engineering design in the new millennium. *Proc. Int. Conf. on Civil Engineering in the New Millennium: - Opportunities and Challenges*, 108–141.
- Basu, P. K., & Jorge, A. B. (2009). *Modeling of Physical Systems*.
- Basu, P. K., Jorge, A. B., Badri, S., & Lin, J. (2003). Higher-Order Modeling of Continua by Finite-Element, Boundary-Element, Meshless, and Wavelet Methods. *Computers and Mathematics with Applications*. [https://doi.org/10.1016/S0898-1221\(03\)90078-2](https://doi.org/10.1016/S0898-1221(03)90078-2)
- Basu, P. K., & Shah, S. (1998). Modeling singular problems of solid continua. *American Society of Mechanical Engineers, Petroleum Division (Publication) PD*.
- Basu, P. K., & Shah, S. (2000). Accurate Modeling of Singular Problems of the Continua. *Int. J. Of Applied Science and Computations*, 7, 1–19.
- Belytschko, T., Lu, Y. Y., & Gu, L. (1994). Element-free Galerkin methods. *International Journal for Numerical Methods in Engineering*, 37(2). <https://doi.org/10.1002/nme.1620370205>
- Bensoussan, A., Lions, J.-L., Papanicolaou, G., & Caughey, T. K. (1979). Asymptotic Analysis of Periodic Structures. *Journal of Applied Mechanics*, 46(2). <https://doi.org/10.1115/1.3424588>
- Bhatnagar, P. L., Gross, E. P., & Krook, M. (1954). A model for collision processes in gases. I. Small amplitude processes in charged and neutral one-component systems. *Physical Review*, 94(3). <https://doi.org/10.1103/PhysRev.94.511>
- Burrus, C. S., Gopinath, R. A., & Guo, H. (1998). Introduction to Wavelets and Wavelet Transforms: A Primer. In *Recherche*.
- Carey, V. P. (1999). Statistical Thermodynamics and Microscale Thermophysics. In *Statistical Thermodynamics and Microscale Thermophysics*. <https://doi.org/10.1017/cbo9780511626395>
- Chao, R. M., & Lee, S. Y. (1999). An H-adaptive refinement BEM procedure using modified sample point error analysis in two-dimensional elastic problems. *Advances in Engineering Software*, 30(4). [https://doi.org/10.1016/S0965-9978\(98\)00085-4](https://doi.org/10.1016/S0965-9978(98)00085-4)
- Charafi, A., & Wrobel, L. C. (1997). A new h-adaptive refinement scheme for the boundary element method using local reanalysis. *Applied Mathematics and Computation*, 82(2–3). [https://doi.org/10.1016/S0096-3003\(96\)00028-8](https://doi.org/10.1016/S0096-3003(96)00028-8)
- Chessa, J., & Belytschko, T. (2004). Arbitrary discontinuities in space-time finite elements by level sets and X-FEM. *International Journal for Numerical Methods in Engineering*, 61(15). <https://doi.org/10.1002/nme.1155>
- Chopard, B., & Droz, M. (1998). Cellular Automata Modeling of Physical Systems. In *Cellular Automata Modeling of Physical Systems*. <https://doi.org/10.1017/cbo9780511549755>

- Cohen, A., Daubechies, I., & Feauveau, J. -C. (1992). Biorthogonal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics*, 45(5). <https://doi.org/10.1002/cpa.3160450502>
- Cohen, A., & Masson, R. (1999). Wavelet methods for second-order elliptic problems, preconditioning, and adaptivity. *SIAM Journal of Scientific Computing*, 21(3). <https://doi.org/10.1137/S1064827597330613>
- Crisfield, M. A. (2001). Finite Element Multidisciplinary Analysis K.K. Gupta J.L. Meek American Institute of Aeronautics and Astronautics, 1801 Alexander Bell Drive, Reston, VA 20191, USA. 2000. 352pp. Illustrated. ISBN 1-56347-393-3. \$64.95 (AIAA members), \$89.95 (non-members). *The Aeronautical Journal*, 105(1046). <https://doi.org/10.1017/s0001924000025513>
- Dahmen, W. (1997). Wavelet and multiscale methods for operator equations. *Acta Numerica*, 6(2). <https://doi.org/10.1017/S0962492900002713>
- Daubechies, I. (1988). Orthonormal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics*, 41(7). <https://doi.org/10.1002/cpa.3160410705>
- Daubechies, I. (1992). Ten Lectures on Wavelets. In *Ten Lectures on Wavelets*. <https://doi.org/10.1137/1.9781611970104>
- de Souza Eloy, F., Gomes, G. F., Ancelotti Jr, A. C., da Cunha Jr, S. S., Bombard, A. J. F., & Junqueira, D. M. (2018). Experimental dynamic analysis of composite sandwich beams with magnetorheological honeycomb core. *Engineering Structures*, 176, 231-242.
- de Souza Eloy, F., Gomes, G. F., Ancelotti Jr, A. C., da Cunha Jr, S. S., Bombard, A. J. F., & Junqueira, D. M. (2019). A numerical-experimental dynamic analysis of composite sandwich beam with magnetorheological elastomer honeycomb core. *Composite Structures*, 209, 242-257.
- Denda, M., & Dong, Y. F. (1999). An error estimation measure in the direct BEM formulation by the external Somigliana's identity. *Computer Methods in Applied Mechanics and Engineering*, 173(3-4). [https://doi.org/10.1016/S0045-7825\(98\)00296-5](https://doi.org/10.1016/S0045-7825(98)00296-5)
- Di Benedetto, R. M., Botelho, E. C., Gomes, G. F., Junqueira, D. M., & Junior, A. A. (2019). Impact energy absorption capability of thermoplastic commingled composites. *Composites Part B: Engineering*, 176, 107307.
- Di Benedetto, R. M., Botelho, E. C., Janotti, A., Junior, A. A., & Gomes, G. F. (2021). Development of an artificial neural network for predicting energy absorption capability of thermoplastic commingled composites. *Composite Structures*, 257, 113131.
- Di Benedetto, R. M., Gomes, G. F., Janotti, A., Junior, A. C. A., & Botelho, E. C. (2022). Statistical approach to optimize crashworthiness of thermoplastic commingled composites. *Materials Today Communications*, 103651.

- Diniz, C.A., Méndez, Y., de Almeida, F.A., da Cunha Jr, S.S. and Gomes, G.F. (2021), "Optimum design of composite structures with ply drop-offs using response surface methodology", *Engineering Computations*, Vol. 38 No. 7, pp. 3036-3060. <https://doi.org/10.1108/EC-07-2020-0354>
- Diniz, C.A., Pereira, J.L.J., da Cunha, S.S. et al. Drop-off Location Optimization in Hybrid CFRP/GFRP Composite Tubes Using Design of Experiments and SunFlower Optimization Algorithm. *Appl Compos Mater* (2022). <https://doi.org/10.1007/s10443-022-10046-z>
- Ding, X., & Tsang, T. T. H. (2001). Large eddy simulation of turbulent flows by a least-squares finite element method. *International Journal for Numerical Methods in Fluids*, 37(3). <https://doi.org/10.1002/flid.172>
- Domingues, A. C. – Damage identification in light aerospace structures using modal strain energy method. (In Portuguese: Identificação de dano em estruturas aeroespaciais leves utilizando o método da energia de deformação modal). Master's Thesis, Integrity of Engineering Materials Program, University of Brasilia, 2019.
- Domingues, A. C.; Anflor, C. T. M.; Carneiro, S. H. S. Damage Identification in Light Panels with Orthotropic Properties Using Modal Strain Energy In: Congreso Iberoamericano de Ingeniería Mecánica, 2019, Cartagena. *XIV Congreso Iberoamericano de Ingeniería Mecánica*, 2019. pp276-280. <https://www.uis.edu.co/webUIS/es/academia/facultades/fisicoMecanicas/escuelas/ingenieriaMecanica/cibim/MEMORIAS-CIBIM/A.pdf>
- Domingues, A. C.; Anflor, C. T. M.; Carneiro, S. H. S. Damage identification using modal strain energy in a structure with orthotropic properties In: International Congress of Mechanical Engineering, 2019, Uberlândia. *25th ABCM International Congress of Mechanical Engineering*, 2019. <https://abcm.org.br/proceedings/view/COB2019/0089>
- Domingues, A. C. ; Santos, T. ; Anflor, C. T. M. ; Carneiro, S. H. S. . Damage Identification in a Cantilever Beam using Modal Strain Energy.. In: *18th International Conference on Machine Design and Production, 2018, 2018*, Eskişehir. 18th International Conference on Machine Design and Production, 2018.
- Donovan, G. C., Geronimo, J. S., & Hardin, D. P. (1996). Intertwining multiresolution analyses and the construction of piecewise-polynomial wavelets. *SIAM Journal on Mathematical Analysis*, 27(6). <https://doi.org/10.1137/S0036141094276160>
- Draper, D. (1995). Assessment and Propagation of Model Uncertainty. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1). <https://doi.org/10.1111/j.2517-6161.1995.tb02015.x>
- Duarte, C. A. (1995). *A review of some meshless methods to solve partial differential equations, Technical report 95-06, TICAM.*
- Edwards, G. (1984). A Bayesian procedure for drawing inferences from random data. *Reliability Engineering*, 9(1). [https://doi.org/10.1016/0143-8174\(84\)90002-7](https://doi.org/10.1016/0143-8174(84)90002-7)
- Ellingwood, B., & Galambos, T. V. (1982). Probability-based criteria for structural design. *Structural Safety*, 1(1). [https://doi.org/10.1016/0167-4730\(82\)90012-1](https://doi.org/10.1016/0167-4730(82)90012-1)

- Faria, L. E. R., Gomes, G. F., de Sousa, S. R. G., Bombard, A. J. F., & Anceleti Jr, A. C. (2020). Dynamic experimental behavior of sandwich beams with honeycomb core filled with magnetic rheological gel: a statistical approach. *Smart Materials and Structures*, 29(11), 115044.
- Frisch, U., Hasslacher, B., & Pomeau, Y. (1986). Lattice-gas automata for the Navier-Stokes equation. *Physical Review Letters*, 56(14). <https://doi.org/10.1103/PhysRevLett.56.1505>
- Ghanem, R. G., & Spanos, P. (1991). *Stochastic Finite Elements: A Spectral Approach*. Springer-Verlag. <https://doi.org/10.1007/978-1-4612-3094-6>
- Ghosh, D. K., & Vanderplaats, G. N. (1997). Development of a flexible design optimization capability. *Proceedings of the Conference on Optimization in Industry*.
- Gibson, I., Rosen, D. W., Stricker, B. Additive Manufacturing Technologies. Springer, 2010.
- Guiggiani, M. (1996). Sensitivity analysis for boundary element error estimation and mesh refinement. *International Journal for Numerical Methods in Engineering*, 39(17). [https://doi.org/10.1002/\(SICI\)1097-0207\(19960915\)39:17<2907::AID-NME982>3.0.CO;2-0](https://doi.org/10.1002/(SICI)1097-0207(19960915)39:17<2907::AID-NME982>3.0.CO;2-0)
- Haldar, A., & Mahadevan, S. (1999). *Probability, Reliability, and Statistical Methods in Engineering Design*. Wiley.
- Hu, Z., & Mahadevan, S. (2017). Uncertainty quantification and management in additive manufacturing: current status, needs, and opportunities. *The International Journal of Advanced Manufacturing Technology*, 93(5), 2855-2874.
- Huebner, K. H., Dewhurst, D. L., Smith, D. E., & Byrom, T. G. (2008). *The finite element method for engineers*. John Wiley & Sons.
- Isukapalli, S. S., & Georgopoulos, P. G. (1999). Computational Methods for the Efficient Sensitivity and Uncertainty Analysis of Models for Environmental and Biological Systems Towards a Framework for an Exposure and Dose Modeling and Analysis System. *Occupational Health*.
- Jaswon, M. A., Symm, G. T., & Cruse, T. A. (1978). Integral Equation Methods in Potential Theory and Elastostatics. *Journal of Applied Mechanics*, 45(4). <https://doi.org/10.1115/1.3424468>
- Jeffreys, H. (1961). *Theory of Probability* (3rd ed.). Oxford University Press.
- Jiang, B. (2000). The Least-Squares Finite Element Method-Theory and Applications in Computational Fluid Dynamics and Electromagnetics. *Journal of fluid mechanics*, 411(1).
- Jorge, A. B. (2002). *Self-regular BEM approaches and error estimation for linear elastic fracture mechanics in two dimensions* [Dissertation, ProQuest Dissertations Publishing]. https://catalog.library.vanderbilt.edu/discovery/fulldisplay?docid=cdi_proquest_journals_305487649&context=PC&vid=01VAN_INST:vanui&lang=en&search_scope=MyInst_and_CI&adaptor=PrimoCentral&tab=Everything&query=any,contains,Self-Regular BEM approaches and e

- Jorge, A. B. (2004). Discretization of random fields in reliability structural analysis : error indicator methodology. *VI Simpósio Mineiro de Mecânica Computacional*. https://www.researchgate.net/publication/344286184_Discretization_of_random_fields_in_reliability_structural_analysis_error_indicator_methodology
- Jorge, A. B., Cruse, T. A., & Fisher, T. S. (2003). Near-crack contour behaviour and extraction of log-singular stress terms of the self-regular traction boundary integral equation. *International Journal for Numerical Methods in Engineering*. <https://doi.org/10.1002/nme.818>
- Jorge, A. B., Cruse, T. A., Fisher, T. S., & Ribeiro, G. O. (2003). A new variational self-regular traction-bem formulation for inter-element continuity of displacement derivatives. *Computational Mechanics*. <https://doi.org/10.1007/s00466-003-0506-4>
- Jorge, A. B., Rebba, R., Mahadevan, S., & Huang, S. (2002). Model validation under uncertainty. *Part i: consistency issues in error estimation*.
- Jorge, A. B., Ribeiro, G. O., Cruse, T. A., & Fisher, T. S. (2001). Self-regular boundary integral equation formulations for Laplace's equation in 2-D. *International Journal for Numerical Methods in Engineering*. <https://doi.org/10.1002/nme.138>
- Jorge, A. B., Ribeiro, G. O., & Fisher, T. S. (2003). New approaches for error estimation and adaptivity for 2D potential boundary element methods. *International Journal for Numerical Methods in Engineering*. <https://doi.org/10.1002/nme.559>
- Jorge, A. B., Ribeiro, G. O., & Fisher, T. S. (2005). Application of new error estimators based on gradient recovery and external domain approaches to 2D elastostatics problems. *Engineering Analysis with Boundary Elements*. <https://doi.org/10.1016/j.enganabound.2005.03.004>
- Jorge, A. B., Ribeiro, G. O., & Fisher, T. S. (2001). Error estimation and adaptivity approaches in BEM. *BETEQ - Joint Meeting of International Conference on Boundary Element Techniques and International Association for Boundary Element Methods. Advances in Boundary Element Techniques II*, 117–126.
- Jou, J., & Liu, J. L. (1999). A posteriori boundary element error estimation. *Journal of Computational and Applied Mathematics*, 106(1). [https://doi.org/10.1016/S0377-0427\(99\)00049-7](https://doi.org/10.1016/S0377-0427(99)00049-7)
- Kane, D., Gomes, G., Macanhan, V. and Ancelotti Jr, A. (2022), "The effect of ply drop-off on tensile strength of thermoplastic carbon fiber composite: a numerical and experimental study", *Engineering Computations*, Vol. 39 No. 6, pp. 2284-2305. <https://doi.org/10.1108/EC-07-2021-0383>
- Karafiati, A. (1997). On hp- Error estimation in the BEM for a three-dimensional Helmholtz exterior problem. *Computer Methods in Applied Mechanics and Engineering*, 150(1–4). [https://doi.org/10.1016/S0045-7825\(97\)00091-1](https://doi.org/10.1016/S0045-7825(97)00091-1)
- Katunin, Andrzej (2010). Identification of multiple cracks in composite beams using discrete wavelet transform. *Scientific Problems of Machines Operation and Maintenance*, 45(2), 41-52. <https://www.researchgate.net/profile/Andrzej-Katunin/publication/237085291>

- Katunin, A., dos Santos, J. V. A., & Lopes, H. "Damage identification by wavelet analysis of modal rotation differences". In *Structures* (Vol. 30, pp. 1-10, 2021). <https://www.sciencedirect.com/science/article/abs/pii/S2352012421000102>
- Kelly, D. W., Mills, R. J., Reizes, J. A., & Miller, A. D. (1987). A posteriori estimates of the solution error caused by discretization in the finite element, finite difference and boundary element methods. *International Journal for Numerical Methods in Engineering*, 24(10). <https://doi.org/10.1002/nme.1620241008>
- Kita, E., & Kamiya, N. (1994). Recent studies on adaptive boundary element methods. *Advances in Engineering Software*, 19(1). [https://doi.org/10.1016/0965-9978\(94\)90043-4](https://doi.org/10.1016/0965-9978(94)90043-4)
- Kita, E., & Kamiya, N. (2001). Error estimation and adaptive mesh refinement in boundary element method, an overview. *Engineering Analysis with Boundary Elements*, 25(7). [https://doi.org/10.1016/S0955-7997\(01\)00018-2](https://doi.org/10.1016/S0955-7997(01)00018-2)
- Kontorova, T., & Frenkel, J. (1938). On the theory of plastic deformation and twinning. I. *Zhurnal Eksperimentalnoi I Teoreticheskoi Fiziki*, 8(March).
- Lancaster, P., & Salkauskas, K. (1981). Surfaces generated by moving least squares methods. *Mathematics of Computation*, 37(155). <https://doi.org/10.1090/s0025-5718-1981-0616367-1>
- Li, G., & Aluru, N. R. (2002). Boundary cloud method: A combined scattered point/boundary integral approach for boundary-only analysis. *Computer Methods in Applied Mechanics and Engineering*, 191(21–22). [https://doi.org/10.1016/S0045-7825\(01\)00415-7](https://doi.org/10.1016/S0045-7825(01)00415-7)
- Li, Y., Leboeuf, E. J., & Basu, P. K. (2005). Least-squares finite-element scheme for the lattice Boltzmann method on an unstructured mesh. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 72(4). <https://doi.org/10.1103/PhysRevE.72.046711>
- Liang, M. T., Chen, J. T., & Yang, S. S. (1999). Error estimation for boundary element method. *Engineering Analysis with Boundary Elements*, 23(3). [https://doi.org/10.1016/s0955-7997\(98\)00086-1](https://doi.org/10.1016/s0955-7997(98)00086-1)
- Liapis, S. (1994). A review of error estimation and adaptivity in the boundary element method. *Engineering Analysis with Boundary Elements*, 14(4). [https://doi.org/10.1016/0955-7997\(94\)90061-2](https://doi.org/10.1016/0955-7997(94)90061-2)
- Liapis, Stergios. (1996). An adaptive boundary element method for the solution of potential flow problems. *Engineering Analysis with Boundary Elements*, 18(1). [https://doi.org/10.1016/S0955-7997\(96\)00041-0](https://doi.org/10.1016/S0955-7997(96)00041-0)
- Liu, W. K., Karpov, E. G., & Park, H. S. (2006). Nano Mechanics and Materials: Theory, Multiscale Methods and Applications. In *Nano Mechanics and Materials: Theory, Multiscale Methods and Applications*. <https://doi.org/10.1002/0470034106>
- Lopez, F., Witherell, P., & Lane, B. (2016). Identifying uncertainty in laser powder bed fusion additive manufacturing models. *Journal of Mechanical Design*, 138(11).

- Mahadevan, S. (2002). Modeling and simulation for design under uncertainty. In K. P. Chong, S. Saigal, S. Thynell, & H. S. Morgan (Eds.), *Modeling and Simulation-Based Life Cycle Engineering* (pp. 276–290). Spon Press. <https://doi.org/10.1201/9781482264715-29>
- Mendes, P. V. M., & Gomes, G. F. (2022). Analysis of the influence of damage on flutter speed in CFRP structures. *Composite Structures*, 280, 114931.
- Menon, G., Paulino, G. H., & Mukherjee, S. (1999). Analysis of hypersingular residual error estimates in boundary element methods for potential problems. *Computer Methods in Applied Mechanics and Engineering*, 173(3–4). [https://doi.org/10.1016/S0045-7825\(98\)00297-7](https://doi.org/10.1016/S0045-7825(98)00297-7)
- Nayroles, B., Touzot, G., & Villon, P. (1992). Generalizing the finite element method: Diffuse approximation and diffuse elements. *Computational Mechanics*, 10(5). <https://doi.org/10.1007/BF00364252>
- Needleman, A. (1987). A continuum model for void nucleation by inclusion debonding. *Journal of Applied Mechanics, Transactions ASME*, 54(3). <https://doi.org/10.1115/1.3173064>
- Needleman, A. (1990a). An analysis of decohesion along an imperfect interface. *International Journal of Fracture*, 42(1). <https://doi.org/10.1007/BF00018611>
- Needleman, A. (1990b). An analysis of tensile decohesion along an interface. *Journal of the Mechanics and Physics of Solids*, 38(3). [https://doi.org/10.1016/0022-5096\(90\)90001-K](https://doi.org/10.1016/0022-5096(90)90001-K)
- Noor, A. K., & Babuška, I. (1987). Quality assessment and control of finite element solutions. *Finite Elements in Analysis and Design*, 3(1). [https://doi.org/10.1016/0168-874X\(87\)90030-8](https://doi.org/10.1016/0168-874X(87)90030-8)
- Oden, J. T., Demkowicz, L., Rachowicz, W., & Westermann, T. A. (1990). A posteriori error analysis in finite elements: The element residual method for symmetrizable problems with applications to compressible euler and Navier-stokes equations. *Computer Methods in Applied Mechanics and Engineering*, 82(1–3). [https://doi.org/10.1016/0045-7825\(90\)90164-H](https://doi.org/10.1016/0045-7825(90)90164-H)
- Oden, J. Tinsley, Browne, J. C., Babuška, I., Liechti, K. M., & Demkowicz, L. F. (2003). A computational infrastructure for reliable computer simulations. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2660. https://doi.org/10.1007/3-540-44864-0_40
- Oden, J. Tinsley, & Demkowicz, L. (1989). Advances in adaptive improvements: Survey of adaptive finite element methods in computational mechanics. In *State-of-the-Art Surveys on Computational Mechanics*.
- Oden, J. Tinsley, & Vemaganti, K. S. (2000). Estimation of Local Modeling Error and Goal-Oriented Adaptive Modeling of Heterogeneous Materials: I. Error Estimates and Adaptive Algorithms. *Journal of Computational Physics*, 164(1). <https://doi.org/10.1006/jcph.2000.6585>
- París, F., & Cañas, J. (1997). *Boundary Element Method: Fundamentals and Applications*. Oxford University Press.

- Paulino, G. H., Gray, L. J., & Zarijian, V. (1996). Hypersingular residuals - a new approach for error estimation in the boundary element method. *International Journal for Numerical Methods in Engineering*, 39(12). [https://doi.org/10.1002/\(sici\)1097-0207\(19960630\)39:12<2005::aid-nme940>3.0.co;2-d](https://doi.org/10.1002/(sici)1097-0207(19960630)39:12<2005::aid-nme940>3.0.co;2-d)
- Paulino, G. H., Shi, F., Mukherjee, S., & Ramesh, P. (1997). Nodal sensitivities as error estimates in computational mechanics. *Acta Mechanica*, 121. <https://doi.org/10.1007/BF01262532>
- Paulino, Glaucio H., & Gray, L. J. (1999). Galerkin Residuals for Adaptive Symmetric-Galerkin Boundary Element Methods. *Journal of Engineering Mechanics*, 125(5). [https://doi.org/10.1061/\(asce\)0733-9399\(1999\)125:5\(575\)](https://doi.org/10.1061/(asce)0733-9399(1999)125:5(575))
- Pike, D. J., Box, G. E. P., & Draper, N. R. (1988). Empirical Model-Building and Response Surfaces. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 151(1). <https://doi.org/10.2307/2982196>
- Porto, P. A. C., Jorge, A. B., & Ribeiro, G. O. (2005). Extension of the variational self-regular approach for the flux boundary element method formulation. In *CMES - Computer Modeling in Engineering and Sciences*. <https://doi.org/10.3970/cmcs.2005.010.065>
- Prabhakar, P., Sames, W. J., Dehoff, R., & Babu, S. S. (2015). Computational modeling of residual stress formation during the electron beam melting process for Inconel 718. *Additive Manufacturing*, 7, 83-91.
- Ramos, H., Santiago, R., Alves, M., Theobald, P., Soe, S., & Brasil, R. (2019). Finite Element Modeling of Gyroid Structures Subjected to Impact Loadings. *MECSOL 2019 - Proceedings of the 7th International Symposium on Solid Mechanics*, ABCM, Sao Carlos, SP, Brazil.
- Rao, S. S. (2019). *Engineering Optimization: Theory and Practice* (5th ed.). John Wiley & Sons. <https://www.wiley.com/en-us/Engineering+Optimization%3A+Theory+and+Practice%2C+5th+Edition-p-9781119454793>
- Rebba, R. (2005). Model validation and design under uncertainty. *PhD Thesis*.
- Rebba, R., & Mahadevan, S. (2006a). Model predictive capability assessment under uncertainty. *AIAA Journal*, 44(10). <https://doi.org/10.2514/1.19103>
- Rebba, R., & Mahadevan, S. (2006b). Statistical methods for model validation under uncertainty. *Collection of Technical Papers - AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference*, 7. <https://doi.org/10.2514/6.2006-1997>
- Rebba, R., Mahadevan, S., & Huang, S. (2006). Validation and error estimation of computational models. *Reliability Engineering and System Safety*, 91(10-11). <https://doi.org/10.1016/j.ress.2005.11.035>
- Rebba, R., Mahadevan, S., & Zhang, R. (2003). Validation of uncertainty propagation models. *Collection of Technical Papers - AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference*, 6. <https://doi.org/10.2514/6.2003-1913>

- Ribeiro, G. O., Ribeiro, T. S. A., Jorge, A. B., & Cruse, T. A. (2009). Evaluation of non-Singular BEM algorithms for potential problems. *Journal of the Brazilian Society of Mechanical Sciences and Engineering*. <https://doi.org/10.1590/S1678-58782009000300012>
- Richards, S. A. (1997). Completed Richardson extrapolation in space and time. *Communications in Numerical Methods in Engineering*, 13(7). [https://doi.org/10.1002/\(SICI\)1099-0887\(199707\)13:7<573::AID-CNEM84>3.0.CO;2-6](https://doi.org/10.1002/(SICI)1099-0887(199707)13:7<573::AID-CNEM84>3.0.CO;2-6)
- Sanchez-Palencia, E., & Zaoui, A. (1985). Homogenization Techniques for Composite Media. In *Homogenization Techniques for Composite Media*. Springer. <https://link.springer.com/book/10.1007/3-540-17616-0>
- Santos, V. C., Jorge, A. B., & Cunha Jr, S. S. (2016). A Reliability Analysis Procedure for the Design of a Helicopter Composite Armor with Correlated Design Variables. *International Journal of Engineering Applied Sciences and Technology*, 1(8), 7–18. <http://www.ijeast.com/papers/7-18,Tesma108,IJEAST.pdf>
- Shepard, D. (1968). A two-dimensional interpolation function for irregularly-spaced data. *Proceedings of the 1968 23rd ACM National Conference, ACM 1968*. <https://doi.org/10.1145/800186.810616>
- Shi, F., Ramesh, P., & Mukherjee, S. (1995). Adaptive mesh refinement of the boundary element method for potential problems by using mesh sensitivities as error indicators. *Computational Mechanics*, 16(6). <https://doi.org/10.1007/BF00370560>
- Shinozuka, M. (1983). Basic Analysis of Structural Safety. *Journal of Structural Engineering*, 109(3). [https://doi.org/10.1061/\(asce\)0733-9445\(1983\)109:3\(721\)](https://doi.org/10.1061/(asce)0733-9445(1983)109:3(721))
- Smith, M., Guan, Z., & Cantwell, W. J. (2013). Finite element modelling of the compressive response of lattice structures manufactured using the selective laser melting technique. *International Journal of Mechanical Sciences*, 67, 28-41.
- Sinou, Jean-Jacques. "A review of damage detection and health monitoring of mechanical systems from changes in the measurement of linear and non-linear vibrations." *Mechanical vibrations: measurement, effects, and control*, 643-702, 2009.
- Smith, M., Mines, R.A.W., Cantwell, W.J., Static and impact behavior of a wing leading edge configuration with a micro lattice core. ICEM16, Cambridge, 2014.
- Smith, C. J., Tammis-Williams, S., Hernandez-Nava, E., & Todd, I. (2017). Tailoring the thermal conductivity of the powder bed in Electron Beam Melting (EBM) *Additive Manufacturing. Scientific Reports*, 7(1), 1-8.
- Soares, C. G. (1989). *Bayesian Prediction of Design Wave Heights*. https://doi.org/10.1007/978-3-642-83828-6_22
- Structural Reliability Research Group - Faculty and Staff. (n.d.). Vanderbilt University. Retrieved June 29, 2022, from http://www.structural-reliability.vanderbilt.edu/faculty_staff.htm
- Sukumar, N., Moran, B., & Belytschko, T. (1998). The natural element method in solid mechanics. *International Journal for Numerical Methods in Engineering*, 43(5). [https://doi.org/10.1002/\(SICI\)1097-0207\(19981115\)43:5<839::AID-NME423>3.0.CO;2-R](https://doi.org/10.1002/(SICI)1097-0207(19981115)43:5<839::AID-NME423>3.0.CO;2-R)

- Sulsky, D., Zhou, S. J., & Schreyer, H. L. (1995). Application of a particle-in-cell method to solid mechanics. *Computer Physics Communications*, 87(1–2). [https://doi.org/10.1016/0010-4655\(94\)00170-7](https://doi.org/10.1016/0010-4655(94)00170-7)
- Tatang, M. A., Pan, W., Prinn, R. G., & McRae, G. J. (1997). An efficient method for parametric uncertainty analysis of numerical geophysical models. *Journal of Geophysical Research Atmospheres*, 102(18). <https://doi.org/10.1029/97jd01654>
- Thoft-Christensen, P., & Baker, M. J. (1982). Structural Reliability Theory and Its Applications. In *Structural Reliability Theory and Its Applications*. <https://doi.org/10.1007/978-3-642-68697-9>
- Tomlinson, G. A. (1929). CVI. A molecular theory of friction. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 7(46). <https://doi.org/10.1080/14786440608564819>
- Tsai, D. H. (1979). The virial theorem and stress calculation in molecular dynamics. *The Journal of Chemical Physics*, 70(3). <https://doi.org/10.1063/1.437577>
- Varfolomeyev, I. V., Busch, M., & Petersilge, M. (1998). Characterization of the computational accuracy in surface crack problems. *International Journal for Numerical Methods in Engineering*, 41(4). [https://doi.org/10.1002/\(SICI\)1097-0207\(19980228\)41:4<721::AID-NME307>3.0.CO;2-G](https://doi.org/10.1002/(SICI)1097-0207(19980228)41:4<721::AID-NME307>3.0.CO;2-G)
- Verfürth, R. (1994). A posteriori error estimation and adaptive mesh-refinement techniques. *Journal of Computational and Applied Mathematics*, 50(1–3). [https://doi.org/10.1016/0377-0427\(94\)90290-9](https://doi.org/10.1016/0377-0427(94)90290-9)
- Volinsky, C. T., Madigan, D., Raftery, A. E., & Kronmal, R. A. (1997). Bayesian model averaging in proportional hazard models: Assessing the risk of a stroke. *Journal of the Royal Statistical Society. Series C: Applied Statistics*, 46(4). <https://doi.org/10.1111/1467-9876.00082>
- Wang, L., Basu, P. K., & Leiva, J. P. (2004). Automobile body reinforcement by finite element optimization. *Finite Elements in Analysis and Design*, 40(8). [https://doi.org/10.1016/S0168-874X\(03\)00118-5](https://doi.org/10.1016/S0168-874X(03)00118-5)
- Wang, L., Leiva, J. P., & Basu, P. K. (2003). Design optimization of automobile welds. *International Journal of Vehicle Design*, 31(4). <https://doi.org/10.1504/IJVD.2003.003352>
- Wendland, W., & Cruse, T. A. (1990). Boundary Element Analysis in Computational Fracture Mechanics. *Mathematics of Computation*, 55(192). <https://doi.org/10.2307/2008455>
- Weng, S.; Tian, W.; Zhu, H.; Xia, Y.; Gao, F.; Zhang, Y. 2017. Dynamic condensation approach to calculation of structural responses and response sensitivities. *Mechanical Systems and Signal Processing* Vol. 88 pp 302-317.
- Wolf-Gladrow, D. A. (2000). 5. Lattice Boltzmann Models 5.1 From lattice-gas cellular automata to lattice. In *Lattice-Gas Cellular Automata and Lattice Boltzmann Models, An Introduction*.
- Xie Mukherjee, Y., & Mukherjee, S. (1997). The boundary node method for potential problems. *International Journal for Numerical Methods in Engineering*, 40(5). [https://doi.org/10.1002/\(sici\)1097-0207\(19970315\)40:5<797::aid-nme89>3.0.co](https://doi.org/10.1002/(sici)1097-0207(19970315)40:5<797::aid-nme89>3.0.co)

- Zhan, J., & Yokoyama, M. (1997). An adaptive method for the determination of element degree for acquiring the desired accuracy in 2-D BEM. *Advances in Engineering Software*, 28(4). [https://doi.org/10.1016/s0965-9978\(96\)00054-3](https://doi.org/10.1016/s0965-9978(96)00054-3)
- Zhang, R., & Mahadevan, S. (2000). Model uncertainty and Bayesian updating in reliability-based inspection. *Structural Safety*, 22(2). [https://doi.org/10.1016/S0167-4730\(00\)00005-9](https://doi.org/10.1016/S0167-4730(00)00005-9)
- Zhang, R., & Mahadevan, S. (2003). Bayesian methodology for reliability model acceptance. *Reliability Engineering and System Safety*, 80(1). [https://doi.org/10.1016/S0951-8320\(02\)00269-7](https://doi.org/10.1016/S0951-8320(02)00269-7)
- Zhao, Z. (1998). A simple error indicator for adaptive boundary element method. *Computers and Structures*, 68(5). [https://doi.org/10.1016/S0045-7949\(98\)00081-9](https://doi.org/10.1016/S0045-7949(98)00081-9)
- Zhu, J. Z. (1997). A posteriori error estimation - The relationship between different procedures. *Computer Methods in Applied Mechanics and Engineering*, 150(1-4). [https://doi.org/10.1016/S0045-7825\(97\)00076-5](https://doi.org/10.1016/S0045-7825(97)00076-5)
- Zhu, J. Z., & Zienkiewicz, O. C. (1997). A posteriori error estimation and three-dimensional automatic mesh generation. *Finite Elements in Analysis and Design*, 25(1-2 SPEC. ISS.). [https://doi.org/10.1016/s0168-874x\(96\)00037-6](https://doi.org/10.1016/s0168-874x(96)00037-6)
- Zienkiewicz, O. C., & Zhu, J. Z. (1987). A simple error estimator and adaptive procedure for practical engineering analysis. *International Journal for Numerical Methods in Engineering*, 24(2). <https://doi.org/10.1002/nme.1620240206>
- Zienkiewicz, O. C., & Zhu, J. Z. (1991). Adaptivity and mesh generation. *International Journal for Numerical Methods in Engineering*, 32(4). <https://doi.org/10.1002/nme.1620320409>
- Zienkiewicz, O. C., & Zhu, J. Z. (1992a). The superconvergent patch recovery and a posteriori error estimates. Part 1: The recovery technique. *International Journal for Numerical Methods in Engineering*, 33(7). <https://doi.org/10.1002/nme.1620330702>
- Zienkiewicz, O. C., & Zhu, J. Z. (1992b). The superconvergent patch recovery and a posteriori error estimates. Part 2: Error estimates and adaptivity. *International Journal for Numerical Methods in Engineering*, 33(7). <https://doi.org/10.1002/nme.1620330703>
- Zienkiewicz, Olgierd C., & Taylor, R. L. (1994). The finite element method. Volume 1: Basic formulation and linear problems. In *The finite element method* (Vol. 3).
- Zimmerman, J. A., Webb, E. B., Hoyt, J. J., Jones, R. E., Klein, P. A., & Bammann, D. J. (2004). Calculation of stress in atomistic simulation. *Modelling and Simulation in Materials Science and Engineering*, 12(4). <https://doi.org/10.1088/0965-0393/12/4/S03>

About the Editors



Ariosto Bretanha Jorge (Book Series Leading Editor)

Visiting Professor at Post-Graduate Program - Integrity of Engineering Materials, University of Brasilia, Brazil

Research interests include: helicopter technologies, mechanical vibrations, computational mechanics, numerical methods, optimization, reliability, aircraft structures, fracture mechanics, inverse problems.

More info: lattes.cnpq.br/3558866397613277, orcid.org/0000-0002-8631-1381

✉ ariosto.b.jorge@gmail.com, ariosto.jorge@unb.br



Carla Tatiana Mota Anflor

Professor at Post-Graduate Program - Integrity of Engineering Materials, University of Brasilia, Brazil

Research interests include: optimization, boundary element method, mechanical vibrations and solid mechanics

More info: lattes.cnpq.br/0526742760439036, orcid.org/0000-0003-3941-8335

✉ ctanflor@gmail.com; anflor@unb.br



Guilherme Ferreira Gomes

Professor at Mechanical Engineering Institute, Federal University of Itajubá, Itajubá, Brazil.

Research interests include: structures, vibration and modal testing, structural health monitoring, composite structures, optimization and applied artificial intelligence

More info: guilherme.unifei.edu.br, lattes.cnpq.br/4963257858781799, orcid.org/0000-0003-0811-6334

✉ guilhermefergom@gmail.com; guilhermefergom@unifei.edu.br



Sergio Henrique da Silva Carneiro

Collaborating Professor at Post-Graduate Program - Integrity of Engineering Materials, University of Brasilia, Brazil.

Research interests include: modal testing, structural analysis, finite element method, dynamics, fracture mechanics and damage detection.

More info: lattes.cnpq.br/6280300531787552, orcid.org/0000-0001-6669-2255

✉ shscarneiro@gmail.com, shscarneiro@unb.br

Book Series in Discrete Models, Inverse Methods, & Uncertainty Modeling in Structural Integrity

VOLUME I

MODEL-BASED AND SIGNAL-BASED INVERSE METHODS

VOLUME II

FUNDAMENTAL CONCEPTS AND MODELS FOR THE DIRECT PROBLEM

VOLUME III

UNCERTAINTY MODELING: FUNDAMENTAL CONCEPTS AND MODELS

This book series is an initiative of the Post Graduate Program in Integrity of Engineering Materials from UnB, organized as a collaborative work involving researchers, engineers, scholars, from several institutions, universities, industry, recognized both nationally and internationally. The book chapters discuss several direct methods, inverse methods and uncertainty models available for model-based and signal based inverse problems, including discrete numerical methods for continuum mechanics (Finite Element Method, Boundary Element Method, Mesh-Free Method, Wavelet Method). The different topics covered include aspects related to multiscale modeling, multiphysics modeling, inverse methods (Optimization, Identification, Artificial Intelligence and Data Science), Uncertainty Modeling (Probabilistic Methods, Uncertainty Quantification, Risk & Reliability), Model Validation and Verification. Each book includes an initial chapter with a presentation of the book chapters included in the volume, and their connection and relationship with regard to the whole setting of methods and models.

The Book Series is an initiative supported by:



UNIVERSITY OF BRASILIA - UnB
www.unb.br

With the kind encouragement of:



BRAZILIAN ASSOCIATION OF COMPUTATIONAL METHODS IN ENGINEERING - ABMEC
www.abmec.org.br



BRAZILIAN SOCIETY OF MECHANICAL SCIENCES AND ENGINEERING - ABCM
www.abcm.org.br



LATIN AMERICAN JOURNAL OF SOLIDS AND STRUCTURES - LAJSS
www.lajss.org