

# The Need for Large-Scale, Integrated Studies of Biodiversity – the Experience of the Program for Biodiversity Research in Brazilian Amazonia

Flávia Regina Capellotto Costa\* & William Ernest Magnusson

*Coordenação de Pesquisas em Ecologia, Instituto Nacional de Pesquisas da Amazônia, CP 478, CEP 69011-970, Manaus, AM, Brasil*

## Abstract

Brazilian Amazonia covers around 5 million km<sup>2</sup>. If we could sample it regularly, with one sample for each 1 degree cell (10,000 km<sup>2</sup>), we would need 500 sample sites. It is clear that financial costs limit the amplitude of biodiversity studies. As the area to be explored is so large, financial resources limited, and demands varied, a program on Amazonian biodiversity research must be cost-effective. Integrated and large-scale studies, using standardized protocols represent the only way to achieve these goals. In this review, we report on experience gathered at two model sites in Amazonia, Reserva Ducke and Alter do Chão, which were used to design the RAPELD system, which is the principle basis for the Inventory Component of the Program on Biodiversity Research (PPBio) of the Brazilian government (<http://ppbio.inpa.gov.br>). We address the following issues: 1) how scale, size, shape and distribution of sampling units affect the outcome of biodiversity studies, in terms of the estimation of biotic complementarity between sites, estimation of organismal abundance, and modeling of species distributions; 2) how different sampling needs, from different taxonomic groups, can be adjusted in integrated protocols; 3) how costs can be reduced through sub-sampling. Use of the method in other research sites in Amazonia is being conducted successfully, and a large network of standardized plots is being constructed (see <http://ppbio.inpa.gov.br>). We also show that other methodologies currently used in large-scale biodiversity studies can be integrated into the RAPELD design.

**Key words:** Biodiversity, Survey, Monitoring, Conservation, Sampling Design.

## Introduction

Brazilian Amazonia covers around 5 million km<sup>2</sup>. If we could sample it regularly, with one sample for each 1 degree cell (about 10,000 km<sup>2</sup>), we would need 500 sample sites. It is clear that financial costs limit the amplitude of biodiversity studies. As the area to be explored is so large, financial resources limited, and demands varied, a program on Amazonian biodiversity research must be cost-effective. Use of integrated and large-scale studies, with standardized protocols, is the only way to achieve these goals.

Many biodiversity studies have been carried in Brazilian Amazonia. However, these studies can not be used to achieve the goals outlined above, because of two main problems. The first is that each researcher, interested in a particular biological group, has sampled their own sites, distributed in the manner he/she thought best. Sometimes, researchers have

very large data bases on the particular group being studied. Unfortunately, these data bases do not coincide with data bases for other groups and most lack environmental data collected as intensively as the biological data. The absence of complete data bases has induced the development of many modeling strategies to fill the gaps. However, the models still suffer from our absence of knowledge of the basic factors controlling species distributions (Brooks *et al.* 2004), and in practice, managers and planners usually have to use coarse predictors of biodiversity (Lopes *et al.* 2010). Algorithms used to plan land use or to select areas for conservation work better with complete data bases (*e.g.* Table 1). When the data base is incomplete, especially in relation to sampling effort, missing values are either treated as evidence of absence, or the data for sites without a full complement of variables measured are deleted (Margules *et al.* 2002). The second problem is that sampling has rarely been designed to cover large, representative areas of the landscape. As we will show, lack of such coverage impairs generalization and often can lead to erroneous conclusions. Therefore, effective biodiversity sampling programs need to be integrated and to cover relevant scales. Many other

\*Send correspondence to: Flávia Regina Capellotto Costa  
Coordenação de Pesquisas em Ecologia,  
Instituto Nacional de Pesquisas da Amazônia, CP 478,  
CEP 69011-970, Manaus, AM, Brasil  
E-mail: [flaviacosta001@gmail.com](mailto:flaviacosta001@gmail.com)

issues need to be addressed to achieve data bases that are really many-purpose, but if the first two requirements are not met, these details do not matter.

Reserva Ducke, from which many examples will be shown to illustrate the integrated large-scale system now in use, is an example of the situation described above. During 50 years since it was established, hundreds of biodiversity studies were conducted there. Some very intensive studies made possible the production of very useful books, such as the *Flora da Reserva Ducke* (Ribeiro *et al.* 1999) or the guide for Amazonian Arachnida and Myriapoda (Adis 2002). Researchers from the Ecology Department of the Instituto Nacional de Pesquisas da Amazonia (INPA) have collected intensive long-term data on caimans and amphibians and researchers from the Forest Science Department at INPA have a long-term data base on tree phenology. When researchers tried to collate this data to produce a management plan for the reserve, they discovered that the data could not be integrated and it would be easier to start again. This conclusion led to the development of the system we now call RAPELD\* and the posterior implementation of this method in other research sites around the Brazilian Amazonia, as the main survey strategy in the Program for Biodiversity Research (Programa de Pesquisa em Biodiversidade – PPBio), funded by the Brazilian Ministry of Science and Technology (<http://ppbio.inpa.gov.br>).

The RAPELD system was not designed to replace other survey designs for specific studies, but is complementary, allowing comparisons among sites at a range of scales that are generally of interest to land managers. Standardized biological-survey systems involve an enormous number of interrelated issues, and we cannot treat all of them here. Therefore, we will concentrate on those issues that have generated most questions from conservationists and land managers. As the PPBio survey strategy aims to make the sampling effective and efficient for a wide range of stakeholders, the sampling design must be compatible with the following principles: 1) Be standardized; 2) Permit integrated surveys for all taxa; 3) Be large enough for monitoring all elements of the biodiversity and ecosystem processes; 4) Be modular, to permit comparisons with less intensive samples taken over very large areas; 5) Be compatible with existing initiatives; 6) Be implementable with existing manpower; and 7) Make the data available quickly in a usable form to managers and other stakeholders. More detailed explanations of the need for each of these principles can be found in the PPBio site <http://ppbio.inpa.gov.br/Eng/inventarios/>.

\* The method is called RAPELD, because of the two scales it has to accommodate. For comparisons among sites, we need rapid assessments, such as the RAP surveys carried out by Conservation International. To understand ecological processes within sites, we need long-term ecological research (LTER) sites. The Brazilian acronym for LTER is PELD; hence the name for the method, RAPELD.

**Table 1.** Often useless and Generally useful data for land-use or conservation planning. Detailed analyses can only be made with complete data sets.

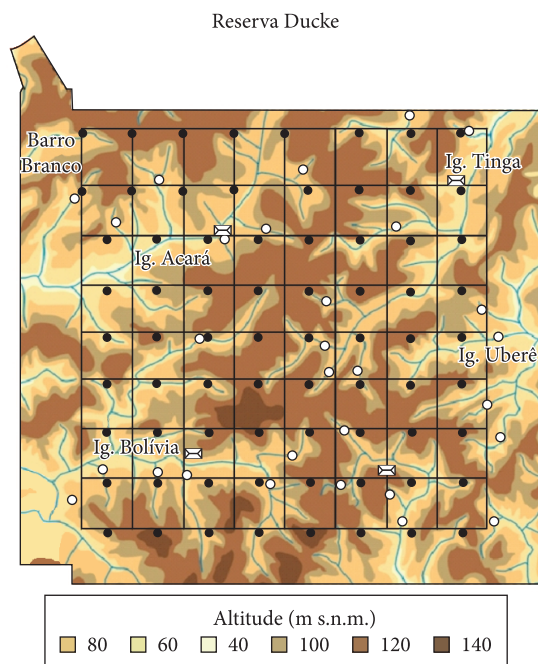
Site	Density of Sp1	Density of Sp2	Soil texture	Vegetation structure
<b>Often useless</b>				
A	45	-	30	1.5
B	68	7	-	-
C	1	-	-	2.3
D	32	-	-	-
E	9	-	21	-
<b>Generally useful</b>				
A	45	1	30	1.5
B	68	7	22	1.4
C	1	10	45	2.3
D	32	15	10	6.7
E	9	25	21	8.9

## Method Outline

The standard grid system used by the PPBio for intensive studies of local processes is a 25 km<sup>2</sup> square, crossed by 6 trails in the east-west and 6 trails in north-south direction, crossing at 1 km intervals. Trails are marked with a geodesic GPS system, so that the measured length along trails is the map length. At each 50 or 100 m there is a pole with a tag informing the coordinates on the grid system. At each of these poles there is a measure of altitude, taken by a professional topographer. The grid system installed at Reserva Ducke is shown in Figure 1. The original grid was made larger (64 km<sup>2</sup>) than in the subsequent sites, to allow testing of sampling effort.

Regularly spaced plots were established along the trails with 1 km between them. These plots do not have a regular shape, but have a 250 m center line that follows the elevational contour. Width of the plot is not fixed, but varies according to the biological group being sampled (Figure 2). Since plots do not have a fixed regular shape, the size is not exactly the width × length, and must be calculated for each plot to allow correct density estimates. As the plot follows the contour line, variation in altitude within the plot is negligible. This allows the use of altitude as a predictor variable. Although altitude per se probably does not directly affect the organisms (variation within most of lowland Amazonia is less than 150 m), it is related to many other characteristics which may directly affect organisms, such as drainage, soil, light and litter deposition, and is easily retrieved from maps or satellite images.

In most areas, soils correlate with altitude and therefore, by minimizing variation in altitude, variation of soil properties inside each plot is also minimized. Variation in soil properties in plots with other formats can be high. For example, a square plot with the same total area as a 250 × 2 m (0.05 ha) plot laid across the contour line at Reserva Ducke,

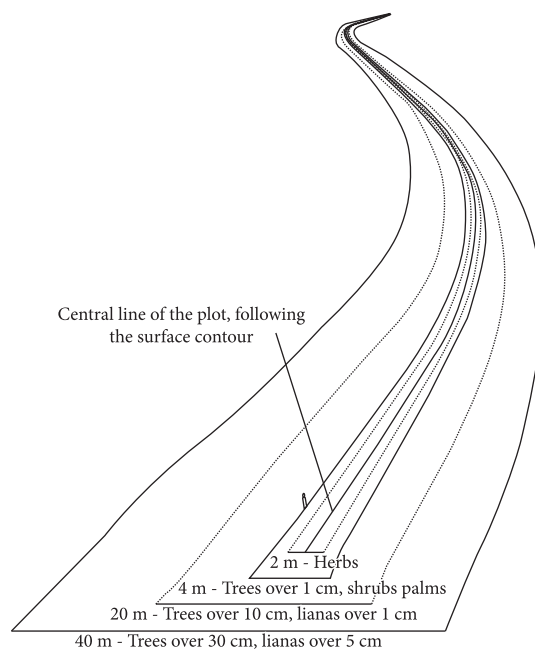


**Figure 1.** The design of the grid system at Reserva Ducke, with regularly spaced plots (black dots) spaced 1 km from each other, aquatic/riparian plots (open circles), and field camps (squares).

could sample a variation of 20 m in altitude. The correlation between altitude and soil clay content at Reserva Ducke ( $r = 0.94$ ) means that a variation of 20 m in altitude could cause a difference of up to 20% in clay content within a single square plot. The strong correlation of altitude and soil properties may not occur in areas subject to strong faulting and distortion of sedimentary layers, but is likely to occur in most parts of the world.

Small-scale soil maps are very rare for Amazonia, which makes the use of soil as a predictor for organismal abundance or distribution difficult. However, topographic variables such as altitude and slope are available in maps or can be estimated from images derived from STRM (Shuttle Radio Topographic Mission) radar. These variables can be used as predictors for larger areas than would be possible if we had to rely only on soil data.

Soil samples are taken for each plot at 50 m intervals along the main axis of the plot, totalling 6 samples. Samples are bulked, cleaned of roots, sieved and analysed for texture – sand, silt and clay fractions, and for chemical components C, N,  $\text{Ca}^{+2}$ ,  $\text{Mg}^{+2}$ ,  $\text{Na}^+$ , P,  $\text{K}^+$ , Cu, Zn, Fe,  $\text{Al}^{+3}$ , ( $\text{Al}^{+3}+\text{H}^+$ ) and  $\text{Mn}^{+2}$ . Soil analyses follow the protocols established by EMBRAPA (Silva 1999). Slope measurements are taken at the same points that soil samples are collected. Slope is measured with a clinometer, perpendicularly to the main axis of the plot. Distance to the nearest water course, when relevant, is measured with a tape laid perpendicularly to the main axis, to the margin of the water course. In the first studies



**Figure 2.** Schematic representation of a terrestrial plot, showing different plot sizes used for sampling plants. In this figure, the central line is shown as completely coincident with the contour line, but actual center lines are constructed from contiguous 10 m long straight line segments.

there were no light measurements in the plots, but they are now being done at some of the sites (Zuquim *et al.* 2009). All of these measures are summarized as the mean of the 6 sub-samples per plot.

Aquatic plots cover 50 m stretches of water-courses. These can not be placed as regularly as the terrestrial plots, because they depend on the distribution of water courses in each grid system. Structural measures of the watercourse, and chemical measures of the water are taken in each plot. For most sites, mean width, depth, water velocity, discharge, canopy openness and substrate type are available as structural measures, and pH,  $\text{O}_2$  and conductivity as chemical measures. Descriptions of the methods used to collect this data are available in Mendonça *et al.* (2005).

In regions crossed by small streams, riparian plots were also established, since regularly spaced plots usually do not fall frequently in this important habitat. Riparian plots are similar to the regularly spaced plots, but instead of following the contour line, they follow the stream margin.

Surveys at Alter do Chão (Pará, Brazil) were undertaken before the PPBio design was elaborated in detail. However, the basic idea was the same and the results of that study are useful to illustrate the reasons for the PPBio design. Alter do Chão is a relictual savanna site in Amazonia, and savanna vegetation occurs interspersed with forest patches. Detailed description of the site can be found in Magnusson *et al.* (2002) and Vasconcelos *et al.* (2006). Forty

terrestrial plots were established in the savanna, and 31 in forest patches, in order to cover the area as uniformly as feasible. Each plot was a group of 4 straight-line transects, each of which was 250 m long.

## How Scale, Size, Shape and Distribution of Sampling Units Affect the Outcome of Biodiversity Studies

Data collected on different geographical scales generally cannot be compared (Urban 2005). Biodiversity measures, such as species richness, community composition, genetic variability, biomass change, and productivity are all strongly scale dependent. This is the reason that it is virtually impossible to use the extensive data banks that have been developed over the past four decades for management applications or the understanding of what determines the distribution of biodiversity.

The size, shape and orientation of plots often determine the questions that can be answered (Caughley & Sinclair 1994; Stern 1998; Magnusson & Mourão 2004). If the objective is to describe the variability within sites, the plot system has to cover the whole site. If the objective is to compare sites, the pattern and spacing of plots within plot systems have to be similar between sites. In this section we will examine and discuss some examples of how scale, size, shape and distribution of sampling units affect the outcome of biodiversity studies, in terms of the estimation of biotic complementarity between sites, estimation of organismal abundance, and modeling of species distributions.

### Estimation of biotic complementarity between sites

Land use planning requires information on the biotic complementarity between sites (Margules and Pressey 2000). However, estimation of complementarity is sensitive to false absences (MacKenzie *et al.* 2003), to the number of samples taken in each planning unit (Albernaz 2001) and to the size and shape of sampling units at each site (Condit *et al.* 1996).

Incomplete coverage makes evaluation of biotic complementarity difficult or impossible (Reddy & Dávalos 2003). Albernaz (2001) applied an iterative algorithm for reserve selection, using data bases of plant and animal groups, to plan the conservation of forest and savanna areas around Alter do Chão. Her results showed that the order of priority for conservation, assigned to each area, is strongly associated with the number of samples taken from each area. Areas selected as higher priorities for conservation were always those areas with higher numbers of samples. This result was independent of the group of organisms examined. The distributions of some taxonomic groups, in some parts of the world, is known with considerable confidence, and these can be used in reserve-selection algorithms. However, naive

application of these algorithms for megadiverse taxa, or even low-diversity taxa in most tropical regions, will result in priorities for conservation that are biased toward areas with more samples. To use reserve-selection algorithms for most taxa, for which false absences can be expected, it is necessary to have spatially standardized sampling units with a known sampling effort.

When we make surveys, we may or may not detect a species that occurs there. When we detect it, we have no doubt that it is there, but when we do not, the situation is not that simple. A species may not be detected in a sample because it really does not occur there. However, for many reasons, the species may occur and not be detected. Sampling may have occurred outside the season when the species is active, the technique used was not appropriate, or even the vagaries of sampling. This situation generates false absences. Distinguishing a true from a false absence is not simple, and only can be done when effort is standardized and distribution of samples is regular. Standardized sampling allows the evaluation of false absences (MacKenzie *et al.* 2002; Field *et al.* 2005; Huettmann 2005), or the analysis of the data using methods that are little affected by false absences (*e.g.* Reyers *et al.* 2002), and is essential for efficient surveying. Standardization is a prerequisite to quantify uncertainty, and estimation of false absences is the first step towards managing landscapes for conservation under uncertainty (Burgman *et al.* 2005).

At Reserva Ducke, Albertina Lima and William Magnusson have monitored frogs for the last 25 years. However, they had access to only about a quarter of the reserve, before the grid system was established. During this period of intensive searches, they never detected *Atelopus spumarius*. When the grid system was opened, this species was readily found in the base camp situated in one of the corners of the reserve. Despite 20 years of natural-history studies, with only this data, it was impossible to tell if the species was present in the intensively searched area but was never detected, or if it was truly absent. However, standardized searches in the 72 plots by a Ph.D student, in less than a year, showed that *Atelopus spumarius* was consistently absent from plots in the western drainage basin but consistently present in plots from the eastern drainage (Menin *et al.* 2008). Because the latter study was standardized and repeatable, it is possible to use the techniques developed by MacKenzie *et al.* (2002), to show that the probability of the species occurring in the western drainage is very small.

Mendonça *et al.* (2005) sampled the fish community of the small streams in Reserva Ducke, in 38 aquatic plots. With data from one year of sampling, there was an indication that 28% of the species were restricted to the western watershed and 22% to the eastern watershed. This suggested a pattern of high turnover of species from one main watershed to the other, implying a high level of complementarity, and initial management plans for the reserve assumed this high divergence between basins. However, analyses taking into account the probability of false absences indicated that

there are only about 3 species restricted to a single drainage basin. Estimates of biotic complementarity are very sensitive to false absences, and our studies showed that large and systematic cover in both terrestrial and aquatic habitats is essential to detect and correct for them.

Even plants may not be always detected, especially when small and when the habitat is complex. At Alter-do-Chão, Fadini (2010) has shown that observers searching for mistletoes failed to detect them 26% of the times, and the plants not detected were seedlings or small juveniles. By using segments of plots as replicates, it is possible to calculate detection probabilities even for sedentary species. Line transect methods can also be used to take into account detectability in estimates of density of individual species.

Environmental variation occurs at different scales, but its association with variation in species distributions tends to increase with scale. At Reserva Ducke, a 1-ha plot could potentially be selectively placed in an area with relatively homogenous soil conditions, but the situation rapidly changes as the size increases. Even with biased positioning, it would not be possible to locate a 10 ha or larger plot on homogeneous soil. Even within vegetation “types”, such as terra firme forest, there is much small-scale (1-10 km) beta diversity in tropical forest vegetation (Clark *et al.* 1999; Phillips *et al.* 2003; Tuomisto and Ruokolainen 1994; Tuomisto *et al.* 2003; Costa *et al.* 2005). Therefore, large grid systems are essential to capture variation in environmental conditions that tend to manifest only at the landscape scale. Small plots (1-100 ha) capture a very small part of the species diversity of the site for many taxonomic groups. The number of tree species registered in the Reserva Ducke’s Flora is around 1176 (Ribeiro *et al.* 1999), but a 1 ha plot in this region will probably have less than 285 species, and nearby plots may have only ~40% shared species (Oliveira & Mori 1999). For herbs, a 0.1 ha plot has 22 species on average and 32 at most, but 108 species have been registered for the whole reserve (Costa 2006; Costa *et al.* 2008).

Although 25 km<sup>2</sup> is much larger than the area generally used for intensive studies of ecosystem processes, it is not enough to capture all landscape variation in management units, such as large parks or watersheds. For example, the Viruá National Park, 500 km north of Manaus, is in the transition of Amazonian forests and savannas. To cover all the vegetation types visible in Landsat images of that park, the grid would need to be > 2,000 km<sup>2</sup>, and the standard 25 km<sup>2</sup> grid captures only a limited amount of that variation (<http://ppbio.inpa.gov.br/Eng/inventarios/nrrr/virua/>). In management units such as the Viruá National Park, 25 km<sup>2</sup> units or smaller modules of the grid must be used as replicates.

The RAPELD system used by the PPBio also has smaller modules for rapid assessment over larger areas using methods that are comparable to those used in the regular

grids (Magnusson *et al.* 2005). These modules are smaller combinations of the same elements present in a grid: trails and plots with the same size and the same distance apart. For questions relating to very large areas, it is more useful to have several smaller modules spread over the area, instead of a single grid, in order to maximize the probability of covering the relevant heterogeneity. This strategy was used to sample the large areas around the federal roads BR 163 and BR 319, in Pará and Amazonas, respectively (<http://ppbio.inpa.gov.br>). Modules are cheaper and this reduces the financial losses if some of them become unavailable due to development pressures.

Even in areas considered homogeneous, there may be variation in community composition that will not be recognized if the area sampled is small. At Reserva Ducke, most of the terrestrial biological groups sampled so far have different complements of species in the eastern and western watersheds (Costa *et al.* 2005; Kinup & Magnusson 2005; Garcia 2005; Drucker *et al.* 2008; Menin 2006). This result was not expected, since there are no barriers between these drainage systems and the soils change very little (Costa *et al.* 2005). There were many additions to the species lists since the grid system was opened and the eastern watershed became accessible. It became clear that much small-scale variation in environments and communities may be present in terra-firme forests and the correct estimation of biotic complementarity will only be possible if the scale of studies is large enough to detect them.

### *Estimation of organismal abundance or growth rates*

Correct estimation of abundance or growth rate of certain organisms or groups of organisms is necessary when these have commercial value or are at extinction risk. However, these can only be estimated correctly when changes in abundance associated with environmental variability is taken into account. To determine how abundance or growth rate vary in relation to the environment, samples must systematically cover the environmental variability in the landscape. At the same time, each sample unit must be homogeneous in terms of the environmental gradients considered most relevant to the distribution of organisms in each site. Samples covering only part of the environmental gradients of a landscape and/or with high internal heterogeneity can lead to under or over estimates of abundance or growth rates.

Castilho *et al.* (2010) studied the growth rate of trees in 72 1-ha terrestrial plots at Reserva Ducke. Overall, biomass increased through the 2 years of study (from 321 to 324 t.ha<sup>-1</sup>). However, as plots covered all the range of soils of the Reserve, the study showed that, during some periods, plots on clay soils increased in biomass, but plots in sandy soils decreased. Previous studies may have over estimated rates of biomass change, since plots for studies of forest dynamics tend to be established on flat high-clay

plateaus. Negotiations in the carbon market and management decisions on exploitation depend heavily on this kind of information, so incorrect estimations have a high cost. A large part of the uncertainty on carbon flux estimates in tropical regions is a consequence of inconsistencies in estimates of Amazon forest biomass (Houghton *et al.* 2001; Saatchi *et al.* 2007), and many relatively small plots but with a large total area per site are needed to correctly estimate the effects of climatic changes (Clark 2007; Fisher *et al.* 2008). RAPELD methodology allows many of those uncertainties to be reduced.

The area covered by sampling also needs to be large because many organisms use very large areas, and studies in small areas have a very low probability of capturing them. Jaguars and harpy eagles are examples of top predators that use large areas. In Reserva Ducke, harpy-eagles were thought to be extinct, but a nest was discovered by a botanist a few months after the opening of the trail system. Simply knowing that the species was present mobilized local organizations to undertake education campaigns to avoid the eagles being killed by ignorant hunters. Some tree species also occur at very low densities, and estimation of their abundance is only possible in large areas. For instance, Wilson Spironello, based on a sample of 59 km of trails at Reserva Ducke estimated the density of rosewood (*Aniba roseodora*) to be 0.54 individuals per hectare, while a previous estimate based on plots was of only 0.03 ind/ha (Costa *et al.* 2008). This was because only 2 individuals of this species were found in 72 plots. Medium to large mammals (*e.g.* primates and ungulates) can only be surveyed over large areas, and most researchers use trails that are 5 km long or larger. A sampling grid needs to be large enough to include 5 km trails, or many taxa can not be included in surveys.

### Modeling species distributions

Modeling species distributions is difficult for many reasons, from false absences (discussed above) to limited geographical coverage of samples. Size and shape of sampling units can also affect the outcome of models (Magnusson & Mourão 2004). Useful sampling units for modeling must balance two conflicting needs: the need to record enough individuals (or another unit of measure of abundance of organisms) of the species or group of species concerned, and the need to accurately measure the environment. Precision of sampling for organisms increases with area, since more individuals can be detected. However, measurements for environmental variables are more representative of the whole sampling unit in smaller areas, since heterogeneity increases with area. Therefore, the ideal size of a plot is as small as it can be and still capture enough numbers of individuals.

Shape of a plot is also crucial for the correct determination of the association between species or assemblages and the environment. Most studies of terrestrial sessile organisms have been done in square or rectangular plots. When plots are small, they are generally laid out along the ground without consideration of slope. The horizontal (map) area

they occupy is thus generally less than their purported size. Larger plots are generally installed by topographers and their nominal area corresponds to map area, while their ground areas are larger. Unfortunately, metadata for most plots is not sufficiently detailed to know whether plot sizes reflect map or ground area. Results from Reserva Ducke, and elsewhere, indicate that this can cause errors of up to 17% in only moderately undulating terrain (Castilho *et al.* 2006; Macedo 2008).

When the plot is small, shape is less important since any form of plot will cover only a small area, and hence little environmental variability. As size increases, shape becomes limiting, because a fixed shape cannot take into account spatial variation in the environment. To model species distributions in relation to the major landscape features, and therefore be able to extrapolate the results to larger areas, samples must control for these features. Therefore, the RAPELD plots do not have a fixed shape, but instead follow the surface contour lines, which are the major landscape features affecting soil, drainage and flooding patterns, and are often associated with light and litter conditions.

Control of the size and shape of plots is important because samples with high internal heterogeneity impair the detection of the factors determining distributions. We can, for example, represent the soil of a plot by the mean of sub-samples taken within it, but if the plot spanned conditions from high clay in the flat plateaus to high sand in bottomlands, the mean will be a value experienced by few if any individuals inside the plot. Therefore, it would not be possible to assign a true value of an environmental predictor to the recorded abundance of a species. This implies in a very poor capacity to model the species distribution.

Using the long thin plots that follow contour lines at Reserva Ducke, Costa *et al.* (2005) showed that altitude and surface slope could explain 16 to 29% of the variance in community composition of three herb assemblages. Slope and altitude could also predict 20% of variation in composition of *Psychotria* (Rubiaceae) shrubs (Kinupp and Magnusson 2005). For palms, soil and topography could explain 60 to 73% of the variance in species composition (Costa *et al.* 2009). For ants, soil texture could predict 3 to 31% (depending on the collection method) of species composition (Oliveira *et al.* 2009). Some biological groups, such as pseudo-scorpions (Aguir *et al.* 2006), were not associated with soil and topography, and the tight control on plot environmental conditions means that these results are not artifacts of sampling.

### Ecosystem processes

Ecosystem processes, such as biomass accumulation, erosion, pollution and sedimentation act over large areas, and can only be evaluated at such scales. Evaluation of ecosystem processes can not be done in small plots, because they can not integrate the processes occurring through watersheds or other functional units. The relationship between arboreal

biomass and ecosystem drivers, such as soil and climate (Castilho *et al.* 2010), described in the preceding sections could not have been obtained using conventional plots.

Using the RAPELD plots spread over a low-impact commercial logging area, Dias *et al.* (2010) were able to detect short term changes in water chemistry and stream structure after logging, but no mean significant changes over longer times. Parallel studies in Reserva Ducke, which was not logged, showed that the chemical changes, but not the stream-structural changes, could have been due to natural seasonal variation (Espírito-Santo *et al.* 2009). The absence of logging effects was possibly because plots were spread in order to capture the landscape structure that includes a large amount of unaffected forest even in logging coupes. Most studies of effects of forestry, including some of ours (Costa and Magnusson 2002), have been conducted on scales too small to detect meaningful effects. The RAPELD grids were designed to measure effects at landscape scales important to land managers. As such they can be used for applied purposes, such as forestry certification or monitoring of marine reserves.

### How Different Sampling Needs, from Different Taxonomic Groups, Can Be Adjusted in Integrated Protocols

Organisms vary in size, and this attribute strongly affects their density. Smaller organisms tend to have higher density because more individuals can be packed in the same area (especially in the case of plants) or because they tend to require smaller areas to survive (especially in the case of animals). Patterns of movement also affect density or frequency of occurrence (MacKenzie 2005). It is not practical or economical to sample all sizes of organisms in the same size of plot, because plots would need to be large to capture enough numbers of the larger organisms, but plots of that size would include numbers of smaller organisms that would be impractical to count, or collect in the field, and too time consuming to sort in the laboratory. This, however, does not make integration of sampling of different groups in the same plot unviable, because width of plot or distribution of sub-samples can be adjusted for each biological group.

Organisms as diverse as soil invertebrates and canopy trees have been sampled in RAPELD plots. Sampling for each biological group covers the whole length of the plot (250 m). However, the width of the plot varies, or sub-samples are taken along it. For plants, a different width is used for each size class or habit. Very small plants, such as herbs, are sampled in a width of 2 m (which gives about 0.05 ha for each plot), large trees, above 30 cm in diameter, are sampled in a width of about 40 m (~1 ha). Soil invertebrates are sampled in soil cores taken at regular intervals along the plot, and the same scheme is used for ants, and flying insects. Frogs and lizards are also sampled in plots of different width, larger for lizards and smaller for frogs. The scheme can be adapted for any organism for which the plot scale is adequate (Figure 2).

### How Costs Can Be Reduced Through Sub-Sampling

Complete surveys are financially unviable for most taxonomic groups (Magnusson *et al.* 2005). However, complete surveys are not necessary to understand community turnover patterns and the relationships between these compositional patterns and the environment. Studies of groups as diverse as soil mites and plants, have shown that sub-sampling is an effective way to collect data at reasonable costs, while still retaining most of the relevant biological information.

Although most animals are invertebrates, most taxonomic groups of invertebrates are not used in faunal surveys (Oliver & Beattie 1996). Instead, vertebrates and plants are used as surrogates for overall biodiversity in biological surveys. However, distributions of invertebrates may be poorly correlated with distributions of vertebrates and plants (Oliver *et al.* 1998) and invertebrates may be better indicators of environmental degradation (van Straalen 1998; Behan-Pelletier 1999). One of the reasons that invertebrates are not used is that it can take a lot of time and money to sort bulk samples of invertebrates (Oliver *et al.* 1998; Lawton *et al.* 1998).

Santos *et al.* (2007) sampled soil mites in plots in savanna sites at Alter-do-Chão. The cost to collect, process, sort, identify and register all the material contained in 10 sub-samples for each of 38 plots was around US\$ 98,000. Reduction in the number of compound sub-samples from 10 to 7 retained more than 80% of the information of dissimilarity among plots. More than 75% of the information was retained when only 12.5% of sub-samples were sorted. Moreover, reduction in the proportion of sub-samples sorted from 50 to 12.5% had little effect on the ability of multivariate inferential methods, based on indirect ordination, to detect the effects of ecological variables on community composition. The cost (~US\$ 5,000) of the reduced effort (7 sub-samples and 12.5% sorting) was 88% lower than the cost of analyzing 100% of the material. Therefore, even biological groups considered difficult can be included in biodiversity assessments if rarefaction studies are conducted to set the minimum sampling needs.

Souza (2010) tested how the number of collection methods and the number of sub-samples could be reduced for ants. He showed that an economy of 53% is possible through a reduction of 40% in sub-sample number, and the use of only two collection methods. This combination would still recover around 90% of the relative differences between plots in ant composition present in the complete effort (3 collection methods and 10 sub-samples per plot).

Zuquim *et al.* (2007) investigated the consequences of a reduction in plot area on the ability to detect the effects of environmental predictors on fern community composition. They sampled 38 RAPELD plots distributed in 3 reserves around 100 km from Manaus. The full size of the plot for ferns was 625 m<sup>2</sup> (2.5 × 250 m). Dissimilarity matrices for

both 20 and 60% reduction in size were highly correlated with the matrix for full size plots. Reductions in size to 20 or 60% of the full size did not affect the ability to detect the strongest environmental effects (soil) on community composition. However, more subtle environmental effects, such as light, were not consistently retained. The 60% reduction in the size of the plot would imply a 24% reduction in sampling time.

Castilho *et al.* (2006) used samples of 4 × 250 m (0.1 ha) for trees 1 to 10 cm of diameter. The number of stems included in each plot was on average 504 (±113), which is half the recommended by Condit *et al.* (1996) for comparisons of community parameters between sites. Measurement, permanent marking and collection cost were US\$ 827 per plot, but were estimated to cost 4 times more if extended to the entire 1 ha plot that was sampled for trees larger than 30 cm diameter. Castilho *et al.* (2010) estimated biomass in 72 plots of up to 1 ha in area twice, for an estimated total cost to measure and identify plants of US\$ 56,000. This is much less than the minimum of US\$ 300,000 for only one survey of a 50 ha plot installed by CTFS (Condit 1998).

The above examples lead us to believe that most if not all organisms can be included in long-term studies for a variety of questions as long as preliminary studies are undertaken to determine financially realistic sampling strategies.

## Integration with Other Methodologies Currently Used in Biodiversity Studies

There have been many initiatives for biodiversity assessment and monitoring of tropical ecosystems, for example the CTFS 50-ha plots, the Conservation International TEAM Initiative, the Large-scale Biosphere-Atmosphere (LBA) project, the Biological Dynamics of Forest Fragments (PDBFF), the Wildlife Conservation Society jaguar project, etc. These experiences all produced important results for particular questions, but by themselves do not provide a comprehensive system for biodiversity studies that can be used by land managers.

The initiatives listed above have been extremely expensive and focused on only a limited range of organisms, or focused on only a limited area. The CTFS plots sample only vascular plants. The TEAM plots sample only those organisms that are so uniformly distributed in the landscape and small enough that they can meaningfully be sampled in a few plots ranging from 1 ha to 1 km<sup>2</sup>. LBA evaluates only ecosystem processes and has no direct link to biodiversity. The PDBFF carries long-term ecological studies of a great variety of organisms, but in relation to only one environmental impact (forest fragmentation) in only one habitat (tropical terra-firme forest). The WCS plots focus on jaguars and their prey. All of these initiatives are presently suffering from the present retraction in conservation support by international donors. It is unlikely that they can be expanded in the near future

to provide the geographical coverage necessary for large areas, such as Amazônia.

Conservation International and The Nature Conservancy undertake rapid assessment (RAP) surveys to evaluate biodiversity, but they are of limited taxonomic coverage, they are not comparable with long-term studies undertaken at other sites, and arguably are not even comparable among sites. The RAPELD methodology, using smaller modules for short-term studies that may or may not become the foci of long-term ecological research (LTER), was designed to allow monitoring of most elements considered in RAP and LTER projects in an integrated manner, at a tiny fraction of the overall cost. The RAPELD methodology was not developed in isolation. It benefited from the experience of other partners working within the PELD site #1 of the Brazilian LTER network. The PELD site #1 includes a CTFS plot, 6 TEAM plots, a LBA eddy-flux tower and model catchment, and the PDBFF plots. The PPBio not only benefited from the experience of its PELD partners, it designed the RAPELD system to allow insertion of these much more expensive initiatives when funding is available. As a whole, the PELD partners cover most of the most current questions in tropical ecology. They provide the benchmarks, and RAPELD tries to adapt the methodology to reduce costs, increase geographic coverage, and obtain spatial standardization at a few scales of interest to land managers.

## Concluding Remarks

RAPELD plots have been installed in Amazonian rainforest, tropical savannas, Pantanal wetlands, and there is reason to believe that they would be equally effective in other areas where organismal abundance is likely to vary over large areas, and be affected by altitude (or depth), such as marine reserves. The RAPELD system is probably the fastest growing biodiversity monitoring system in the world, and there are now RAPELD plots being installed in Australia and Nepal (<http://www.griffith.edu.au/environment-planning-architecture/environmental-futures-centre/projects/program-for-planned-biodiversity-studies-australasia>). RAPELD infrastructure can be used to answer a wide range of questions on a wide range of scales, and was designed to facilitate integration of researchers from different fields and working in different time periods. Many different stakeholders have opted to use RAPELD methodology, because it answered their questions, and because it allows them to compare their results with those obtained in other areas (see <http://ppbio.inpa.gov.br/Port/nregionais/> for an idea of the variety of stakeholders involved). RAPELD infrastructure cannot be used to answer all questions on all scales. However, it is likely to be useful, and reduce costs for the majority of studies. We recommend it to anyone who needs to monitor all elements of biodiversity or environmental degradation at the landscape level, at a reasonable cost.



## Acknowledgements

The RAPELD methodology was initially developed with funding from the CNPq and MCT, but many partners have funded or provided logistic support for specific studies leading to the results presented here. Among those organizations that provided major support were CNPq, MCT, CAPES, FAPEAM, PPG7, PELD/CNPq, IBAMA, ICMBio, Mil Madeireira, PDBFF and CL. Many other organizations and people, including dozens of students, helped to develop the methods and to run the studies, and it is not possible to name all of them here.

## References

- Adis J, 2002. *Amazonian Arachnida and Myriapoda*. Sofia-Moscow: Pensoft.
- Aguiar NO, Gualberto TL & Franklin E, 2006. A medium-spatial scale distribution pattern of Pseudoscorpionida (Arachnida) in a gradient of topography (altitude and inclination), soil factors, and litter in a central Amazonia forest reserve, Brazil. *Brazilian Journal of Biology*, 66:791-802.
- Albernaz ALKM, 2001. *Zoneamento da região de Alter do Chão, Pará: um exercício de planejamento para uma unidade de conservação de uso direto*. [PhD Thesis]. Manaus: Instituto Nacional de Pesquisas da Amazônia.
- Behan-Pelletier VM, 1999. Oribatid mite biodiversity in agroecosystems: role for bioindication. *Agriculture, Ecosystem and Environment*, 74: 411-423.
- Brooks TM, Fonseca GAB & Rodrigues ASL, 2004. Protected areas and species. *Conservation Biology*, 18:616-618.
- Burgman MA *et al.*, 2005. Managing landscapes for conservation under uncertainty. *Ecology*, 86:2007-2017.
- Castilho CV *et al.*, 2006. Variation in aboveground tree live biomass in a central Amazonian forest: effects of soil and topography. *Forest Ecology and Management*, 34:85-96.
- Castilho CV *et al.*, 2010. Short-term temporal changes in tree live biomass in a central Amazonian forest, Brazil. *Biotropica*, 42:95-103.
- Coughley G & Sinclair ARE, 1994. *Wildlife management and ecology*. Cambridge: Blackwell Science.
- Clark DB, Palmer MW & Clark DA, 1999. Edaphic factors and the landscape scale distributions of tropical rain forest trees. *Ecology*, 80:2662-267.
- Clark DA, 2007. Detecting tropical forests' responses to global climatic and atmospheric change: current challenges and a way forward. *Biotropica*, 39:4-19.
- Condit R, 1998. *Tropical forest census plots*. Berlin:Springer.
- Condit RG *et al.*, 1996. Species-area and species-individual relationships for tropical trees: A comparison of three 50-ha plots. *Journal of Ecology*, 84:549-562.
- Costa FRC, 2006. Mesoscale Gradients of Herb Richness and Abundance in Central Amazonia. *Biotropica*, 38:711-717.
- Costa FRC & Magnusson WE, 2002. Selective logging effects on abundance, diversity and composition of tropical understory herbs. *Ecological Applications*, 12:807-819.
- Costa FRC, Magnusson WE & Luizão RC, 2005. Mesoscale distribution patterns of Amazonian understory herbs in relation to topography, soil and watersheds. *Journal of Ecology*, 93:863-878.
- Costa FRC *et al.*, 2008. Flora. In Oliveira *et al.* (eds). *Reserva Ducke: A biodiversidade através de uma grade*. Manaus: Áttema Design Editorial. p. 21-30.
- Costa FRC *et al.*, 2009. Gradients within gradients: the mesoscale distribution patterns of palms in a central Amazonian forest. *Journal of Vegetation Science*, 20:69-78.
- Dias MS, Magnusson WE & Zuanon J, 2010. Effects of reduced-impact logging on fish assemblages in Central Amazonia. *Conservation Biology*, 24:278-286.
- Drucker DP, Costa FRC & Magnusson WE, 2008. How wide is the riparian zone of small streams in tropical forests? A test with terrestrial herbs. *Journal of Tropical Ecology*, 24:65-74.
- Espírito-Santo HMV *et al.*, 2009. Seasonal variation in the composition of fish assemblages in small Amazonian forest streams: evidence for predictable changes. *Freshwater Biology*, 54:536-548.
- Fadini RF, 2010. *Especificidade por hospedeiro, abundância e prevalência de ervas-de-passarinho (Psittacanthus – Loranthaceae) em uma savana amazônica afetada por queimadas*. [PhD Thesis]. Manaus: Instituto Nacional de Pesquisas da Amazônia.
- Field SA, Tyre AJ and Possingham H, 2005. Optimizing allocation of monitoring effort under economic and observational constraints. *Journal of Wildlife Management*, 69:473-482.
- Fisher JI *et al.* 2008. Clustered disturbances lead to bias in large-scale estimates based on forest sample plots. *Ecology Letters*, 11:554-563.
- Garcia ARM, 2005. *Distribuição e abundância de arbustos do gênero Piper (Piperaceae) na Reserva Florestal Adolpho Ducke*. [MSc Thesis]. Manaus: Instituto Nacional de Pesquisas da Amazônia.
- Houghton RA *et al.*, 2001. The spatial distribution of forest biomass in the Brazilian Amazon: a comparison of estimates. *Global Change Biology*, 7:731-746.
- Huettmann F, 2005. Databases and science-based management in the context of wildlife and habitat: toward a certified ISO standard for objective decision-making for the global community by using the internet. *Journal of Wildlife Management*, 69:466-472.
- Kinupp VF & Magnusson WE, 2005. Spatial patterns in the understory shrub genus *Psychotria* in central Amazonia: effects of distance and topography. *Journal of Tropical Ecology*, 21:363-374.
- Lawton JH *et al.*, 1998. Biodiversity inventories, indicator taxa and effects of habitat modification in a Cameroon tropical forest. *Nature*, 391:72-76.
- Lopes TE, Schiatti J, Nelson B, Desmoulière SJM, Espírito Santo HM & Costa FRC, 2010. Assessing the relationship between forest types and canopy tree beta diversity. *Ecography* (Copenhagen), doi: 10.1111/j.1600-0587.2009.06139.x
- Macedo RC, 2008. Modelo digital de elevação e sua importância em inventários florestais. R63. In Oliveira YMM *et al.* (eds). *Anais do Seminário Nacional sobre Dinâmica de Florestas*; Curitiba. Curitiba: EMBRAPA Florestas. CD-ROM.

- MacKenzie DI, 2005. What are the issues with 'presence/absence' data for wildlife managers? *Journal of Wildlife Management*, 69:849-860.
- MacKenzie DI *et al.*, 2002. Estimating site occupancy rates when detection probabilities are less than one. *Ecology*, 83:2248-2255.
- MacKenzie DI *et al.*, 2003. Estimating site occupancy, colonization and local extinction when a species is detected imperfectly. *Ecology*, 84:2200-2207.
- Magnusson WE & Mourão G, 2004. *Statistics Without Math*. Londrina: Editora Planta & Sinauer Assoc.
- Magnusson WE *et al.*, 2005. RAPELD: a modification of the Gentry method of biodiversity surveys in long-term ecological research sites. *Biota Neotropica*, 5.
- Magnusson WE *et al.*, 2002. A comparison of  $\delta^{13}\text{C}$  ratios of surface soils in savannas and forests in Amazônia. *Journal of Biogeography*, 29:857-863.
- Margules RC & Pressey LR, 2000. Systematic conservation planning. *Nature*, 405:243-253.
- Margules CR, Pressey RL & Williams PH, 2002. Representing biodiversity: data and procedures for identifying priority areas for conservation. *Journal of Bioscience*, 27:309-326.
- Mendonça FP, Magnusson WE & Zuanon J, 2005. Relationships between habitat characteristics and fish assemblages in small streams of central Amazonia. *Copeia*, 4:751-754.
- Menin M, Waldez F & Lima AP, 2008. Temporal variation in the abundance and number of species of frogs in 10,000 ha of a forest in Central Amazonia, Brazil. *South American Journal of Herpetology*, 3:68-81.
- Oliveira AA & Mori SA, 1999. A Central Amazonian terra firme forest. I. High tree species richness on poor soils. *Biodiversity and Conservation*, 8:1219-1244.
- Oliveira PY *et al.*, 2009. Ant species distribution along a topographic gradient in a "terra-firme" forest reserve in Central Amazonia. *Pesquisa Agropecuária Brasileira*, 44:852-860.
- Oliver I & Beattie AJ, 1996. Designing a cost-effective invertebrate survey. A test of methods for the rapid assessment of biodiversity. *Ecological Applications*, 6:596-607.
- Oliver IA *et al.*, 1998. Spatial fidelity of plant, vertebrate, and invertebrate assemblages in multiple use forest in eastern Australia. *Conservation Biology*, 12:822-835.
- Phillips OL *et al.*, 2003. Efficient plot-based floristic assessment of tropical forests. *Journal of Tropical Ecology*, 19:629-645.
- Reddy S & Davalos LM, 2003. Geographical sampling bias and its implications for conservation priorities in Africa. *Journal of Biogeography*, 30:1719-1727.
- Reyers B *et al.*, 2002. A multicriteria approach to reserve selection: addressing long-term biodiversity maintenance. *Biodiversity and Conservation*, 11:769-793.
- Ribeiro JELS *et al.* 1999. *Flora da Reserva Ducke: Guia de Identificação das Plantas Vasculares de uma Floresta de Terra Firme na Amazônia Central*. Manaus: INPA/DFID.
- Saatchi SS, Houghton RA, Dos Santos ARC, Soares JV & Yu Y, 2007. Distribution of aboveground live biomass in the Amazon basin. *Global Change Biology*, 13:816-837.
- Santos EMR, Franklin E & Magnusson E, 2007. Cost-efficiency of sub-sampling protocols to evaluate Oribatid-Mite communities in an Amazonian Savanna. *Biotropica*, 40:728-735.
- Silva FC (org.), 1999. *Manual de análises químicas de solos, plantas e fertilizantes*. Brasília: EMBRAPA.
- Souza JLP, 2010. *Avaliação do esforço amostral, captura de padrões ecológicos e utilização de taxa substitutos em formigas (Hymenoptera, Formicidae) de serrapilheira com três métodos de coleta na floresta amazônica, Brasil*. [PhD Thesis]. Manaus: Instituto Nacional de Pesquisas da Amazônia.
- Stern MJ, 1998. Field comparisons of two rapid survey vegetation assessment techniques with permanent plot inventory data in Amazonian Peru. In Dallmeier F and Comiskey JA (eds). *Forest Biodiversity Research, Monitoring and Modelling*. Paris: UNESCO & Parthenon Publishing. p. 269-283.
- Tuomisto H & Ruokolainen K, 1994. Distribution of Pteridophyta and Melastomataceae along an edaphic gradient in an Amazonian rain forest. *Journal of Vegetation Science*, 5:25-34.
- Tuomisto H *et al.* 2003. Floristic patterns along a 43-km long transect in an Amazonian rain forest. *Journal of Ecology*, 91:743-756.
- Urban DL, 2005. Modeling ecological processes across scales. *Ecology*, 86:1996-2006.
- van Straalen NM, 1998. Evaluation of bioindicator systems derived from soil arthropod communities. *Applied Soil Ecology*, 9:429-437.
- Vasconcelos H *et al.*, 2006. Long-term effects of forest fragmentation on Amazonian ant communities. *Journal of Biogeography*, 33:1348-1356.
- Zuquim G, Costa FRC & Prado J, 2007. Redução de esforço amostral vs. retenção de informação em inventários de pteridófitas na Amazônia Central. *Biota Neotropica*, 7.
- Zuquim G, Costa FRC & Prado J, 2009. Distribution of pteridophyte communities along environmental gradients in Central Amazonia, Brazil. *Biodiversity and Conservation*, 18:151-166.

Received: February 2010  
 First Decision: March 2010  
 Accepted: May 2010