

Long Database Report

The Brazilian Program for Biodiversity Research (PPBio) Information System

Flávia Fonseca Pezzini, Pablo Hendrigo Alves de Melo, Dayane Mayely Silva de Oliveira, Rainer Xavier de Amorim, Fernando Oliveira Gouvêa de Figueiredo, Debora Pignatari Drucker, Flávio Rogério de Oliveira Rodrigues, Gabriela Zuquim, Thaise Emilio, Flávia Regina Capellotto Costa, William Ernest Magnusson, Adeilza Felipe Sampaio, Albertina Pimentel Lima, Ana Raquel de Mesquita Garcia, Angelo Gilberto Manzatto, Anselmo Nogueira, Cândida Pereira da Costa, Carlos Eduardo de Araújo Barbosa, Carolina Bernardes, Carolina Volkmer de Castilho, Catia Nunes da Cunha, Cintia Gomes de Freitas, Claymir de Oliveira Cavalcante, Diego Oliveira Brandão, Domingos de Jesus Rodrigues, Eliana Celestino da Paixão Rodrigues dos Santos, Fabricio Beggiato Baccaro, Françoise Yoko Ishida, Fernanda Antunes Carvalho, Gabriel Massaine Moulatlet, Jean-Louis Bernard Guillaumet, José Luiz Purri Veiga Pinto, Juliana Schietti, Julio Daniel do Vale, Lauren Belger, Luciano Martins Verdade, Marcelo Petratti Pansonato, Marcelo Trindade Nascimento, Márcia Cléia Vilela dos Santos, Mariana Souza da Cunha, Rafael Arruda, Reinaldo Imbrozio Barbosa, Ricardo Laerte Romero, Susamar Pansini & Tania Pena Pimentel

Abstract: The database of the Brazilian Program for Biodiversity Research (PPBio; GIVD ID SA-BR-001) includes data on the environment and biological groups such as plants. It is organized by site, which is usually a grid with 10 to 72 uniformly-distributed plots, and has already surveyed 1,638 relevés across different Brazilian ecosystems. The sampling design is based on the RAPELD system to allow integration of data from diverse taxa and ecosystem processes. RAPELD is a spatially-explicit sampling scheme to monitor biodiversity in long-term ecological research sites and during rapid appraisals of biodiversity that has attracted support from many management agencies, which are using it as their long-term monitoring system. Vegetation surveys include measurements of cover, biomass and number of individuals from woody and herbaceous vascular plants, along with environmental data. We have recently migrated to a metadata catalog and data repository which allows searching for specific groups across all sites. All RAPELD data have been collected since 2001, though the site also allows data from other long-term plots to be archived as associated projects.

Keywords: Amazonia; assessment; data repository; impact; landscape; monitoring; multidisciplinary; remote sensing.

Received: 5. October 2010 – Accepted: 18. July 2011 – Co-ordinating Editor: Gabriela Lopez-Gonzalez.

Introduction

Mindful of its obligations under the Convention on Biological Diversity, the Brazilian Federal Government launched the Program for Biodiversity Research (PPBio) in 2004, after three years of workshops and meetings to plan the program. The program involves many actions, ranging from support for biological collections, training of taxonomists and production of field guides to support field

surveys (<http://ppbio.inpa.gov.br/guias>), to field surveys themselves (Costa & Magnusson 2010) and bioprospecting (Beattie et al. 2010). However, one aspect that permeates all components of the program is that data should be made publicly available as soon as possible, and this is made explicit in the data-policy document (<http://ppbio.inpa.gov.br/politicadados>).

Most field surveys follow the RAPELD methodology (Magnusson et al. 2005, Costa & Magnusson 2010, briefly de-

scribed below), which facilitates integration of vegetation and faunal data with environmental variables. The name RAPELD originated from the combination of two words that represent the working scales of the method: RAP, which stands for rapid assessments and PELD, which is the acronym for Long Term Ecological Research (LTER) in Portuguese. Trail grids link permanent plots at a scale that facilitates analysis of landscape diversity and complementarity of manage-

ment units, such as conservation reserves, forestry coupes, and sustainable-development reserves. Because the system generates data useful to management agencies as well as academic researchers, it is being implemented by many different organizations, such as state conservation agencies, the Federal Forest Service, and agencies responsible for environmental-impact studies. Since the methodology is standardized, field infrastructure installed by an agency interested in monitoring physical-chemical parameters can be used by researchers studying plants or fauna. This has resulted in the system becoming the fastest growing biodiversity monitor-

ing system in Brazil, and possibly of the world (Fig. 1).

Because of the huge areas involved, simple random sampling is not an option for a large megadiverse country such as Brazil, and is probably not economically viable in most regions of the world. Designing a monitoring system is much more complicated than designing a monitoring study (Watson & Novelty 2004). The RAPELD system is designed to allow modeling of species distributions with nonrandom sampling (see section on Design), based on environmental characteristics (see section on Ancillary Information). As the data are publicly available

(see section on Data Management), they can be used by government agencies, NGOs and the private sector for management decisions. Management decisions are usually based on extrapolation of plot data to larger scales using remote sensing, so it is important to maximize the suitability of data for calibrating remote sensing techniques (see section on Remote Sensing). The standardized relevés allow investigation of relationships among different plant life forms and relationships between plant composition or forest structure and the distribution of fauna (see section on faunal studies and vegetation data).

GIVD Database ID: SA-BR-001		Last update: 2012-05-18	
Brazilian Program for Biodiversity Research (PPBio) Information System			
Scope: Amazon Plants			
Status: completed and continuing		Period: 2001-2011	
Database manager(s): Flávia Fonseca Pezzini (flaviapezzini@gmail.com)			
Owner: [NA]			
Web address: http://ppbio.inpa.gov.br			
Availability: after blocking period		Online upload: yes	Online search: yes
Database format(s): CSV file, TXT file		Export format(s): CSV file	
Publication: [NA]			
Plot type(s): normal plots		Plot-size range: 1-10000 m ²	
Non-overlapping plots: 1,843	Estimate of existing plots: [NA]	Completeness: [NA]	
Total plot observations: 1,843	Number of sources: [NA]	Valid taxa: [NA]	
Countries: BR: 100.0%			
Forest: [NA] — Non-forest: [NA]			
Guilds: all vascular plants: 100%			
Environmental data: altitude: 100%; slope inclination: 35%; soil pH: 92%; other soil attributes: 92%			
Performance measure(s): cover: 100%; number of individuals: 100%; biomass: 100%			
Geographic localisation: GPS coordinates (precision 25 m or less): 100%			
Sampling periods: 2000-2009: 69.3%; 2010-2019: 30.7%			
<i>Information as of 2012-07-19; further details and future updates available from http://www.givd.info/ID/SA-BR-001</i>			

Design

Collection of data under the PPBio is undertaken on a standardized RAPELD system of trails and plots. Trails are organized into grids, and the standard RAPELD grid system for intensive studies of local processes is a 25-km² 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. Stakes with tags containing the coordinates on the grid system are placed at 50 m intervals along the trails. Altitude is recorded for each stake by a professional topographer. Plots are installed along trails in one direction (usually East-West), at 1-km intervals, and the standard grid has 30 regularly-spaced plots.

The RAPELD system also has smaller modules for assessments over larger areas using methods that are comparable to those used in the regular grids. These modules are smaller than the standard grid, but have the same structure: trails and plots with the same size and the same distance apart. The most frequently used module consists of only two parallel 5 km trails 1 km apart, with 10 regularly-spaced plots.

Regularly-spaced plots do not have a fixed shape, but have a 250 m center line that follows the elevational contour. The width of the plots varies according to the biological group being sampled (Fig. 2). Since plots do not have a fixed regular shape, the area is not exactly the width multiplied by the length, and must be calculated for each plot to allow calculation of correct density estimates. However, the

difference is small, and generally less than the difference between the plan area and the ground area in conventional square plots. As the plots follow the contour line, variation in altitude within each plot is small, and this allows the use of altitude as a predictor variable. Although altitude per se probably does not directly affect organisms in the tropical lowlands, 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.

The vegetation sampling is conducted on both sides of the center line. Along this line there is a 0.6 to 1 m wide access trail for access by researchers. All surveys are undertaken along the entire long axis of the plot (250 m), but the width varies for each life-form depending on its size. The

objective is to adjust the plot size to have enough representation for that life-form, in a reasonable time and cost. Trees with a diameter at breast height (DBH) greater or equal to 30 cm and lianas with diameter (D) \geq 5 cm at 130 cm from the rooting point are surveyed in a 40 m wide strip, 20 meters to each side of the center line. The strip is 20 m wide for trees with DBH of 10 cm or more, and 1.5 to 4 m for small trees and shrubs with DBH \geq 1 cm. Lianas with $D \geq$ 1 cm at 130 cm from the rooting point are surveyed in 10 m wide plots. Herbs are usually surveyed in 1.5 to 2 m wide plots. Each tree, shrub and liana stem has its diameter measured providing information to calculate density, basal area and biomass. Biomass is calculated using published allometric equations. Herbs are counted considering clumps as individuals, and cover measurements are made by the point method (Buckland 1997). Because plots are narrow, it is relatively easy to locate marked individuals within a plot. It generally takes about four days to mark all woody plants $>$ 1 cm DBH in the subplots.

The taxonomy of the first metadata was documented using the Cronquist system for plant names. Currently, plant names follow the APG II and the researchers and data managers are instructed to revise the names according to the International Plant Name Index (<http://www.ipni.org/>), but we intend to consult other initiatives in near future, such as The Plant List (<http://www.theplantlist.org/>).

Riparian plots are also established where drainages cross the trails, because regularly spaced plots 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. In some biomes, other plots are placed where special vegetation or landform features intersect trails, such as plots in forest patches in the Pantanal wetlands. The objective is to maintain the general conformation of regularly spaced plots, while allowing distribution of plots to intercept strata that may be of special interest to individual researchers, but that are under represented by regular sampling.

Ancillary information

Basic data on abiotic variables available for all plots include plot coordinates, soil type, altitude and slope. Soil samples are taken in each plot at 50 m intervals along the main axis of the plot, totaling 6 sam-

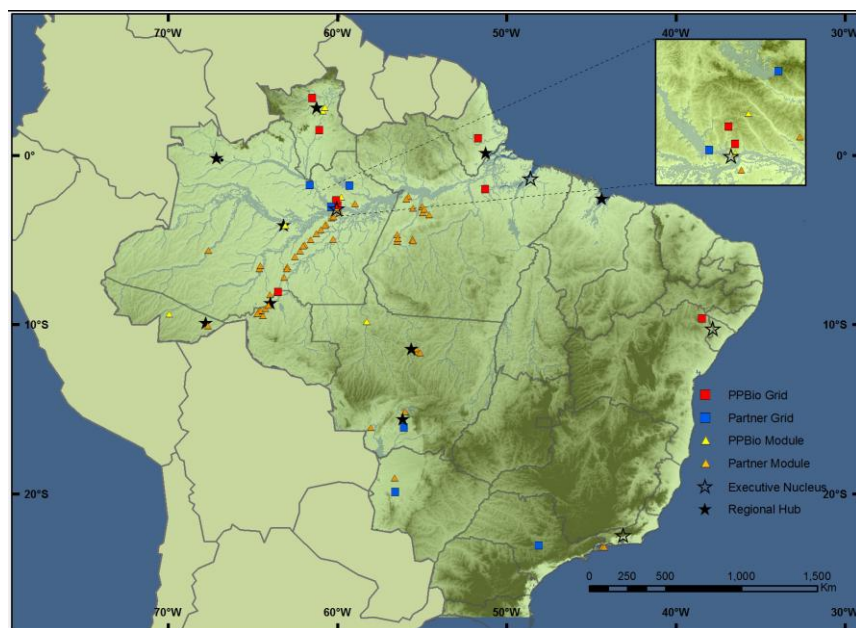


Fig. 1: Distribution of RAPELD 25 km² grids (squares) and modules (triangles), which are smaller units, generally 5 km², in Brazil. Red squares and yellow triangles represent research sites financially installed by PPBio and blue squares and orange triangles by partners. Stars represent regional hubs.

ples, which are analyzed for texture and chemical properties. Slope measurements are taken with a clinometer perpendicularly to the main axis of the plot at the same points that soil samples are collected. Altitude where the plot adjoins the trail is measured with a theodolite, by a professional topographer.

Additional abiotic data available for some of the plots include distance to the nearest water course (measured with a tape laid perpendicularly to the main axis of the plot to the margin of the water course); water table fluctuation (measured with dip wells); understory light (measured with hemispherical photographs or light sensors); litter quantity (measured as dry mass per area or depth), and percentage of burned plot area in fire prone areas.

The environmental measures from each plot can be used to investigate the distribution of plant taxa throughout the landscape or along environmental gradients. Analysis of the mesoscale response of trees, palms, shrubs, herbs and ferns in relation to topography, soil and watersheds has already been undertaken for some sites (Costa et al. 2005, 2009, Kinnup & Magnusson 2005, Costa 2006, Carvalho et al. 2007, Zuquim et al. 2007, 2009 Drucker et al. 2008).

We found that composition of all plant groups studied to date varies along topographic gradients, with many species restricted to sandy lowland areas and others

to clayey uplands. However, variation in composition is not homogeneous along topographic gradients. Gradients within gradients are found in riparian plots, linked to the distance from watercourses (Drucker et al. 2008, Costa et al. 2009). An unexpected finding was a significant difference in composition between neighboring watersheds within a single grid, which differ little in soil properties, which is probably attributable to landscape configuration (Costa et al. 2005, Kinnup & Magnusson 2005).

Data management

An aspect that permeates all components of the program is that data should be made publicly available as soon as possible, and this is made explicit in the data-policy document. The Information Management Committee, the Scientific Committee, and the core of the Information Technology group, plan, maintain and execute the Data Policy of the PPBio. According to the Data Policy, all metadata should be available to the Information Management Committee within 30 days of data collection, and the preliminary data within 12 months of data collection (<http://ppbio.inpa.gov.br/politicadados>). Researchers interested in using the data for publications should contact those

listed as responsible for the data in the metadata to discuss authorship.

Promoting a data-sharing culture has always been an important goal of the PPBio. Standardized RAPELD field surveys generate heterogeneous data sets, which make data documentation an issue of great importance to enable integration with data from other sources. Therefore, researchers have been encouraged to provide metadata describing their data sets based on the EML specification (Ecological Metadata Language, Fegeus et al. 2005). Metadata and data collected since 2005 in more than 1,600 RAPELD relevés are available in the PPBio website (<http://ppbio.inpa.gov.br/>) (Fig. 3). Some

researchers were reluctant to make data available early in the program, but the data-sharing culture is now entrenched in the PPBio community. The program provides an open-access database on Amazonian biodiversity freely available to society. The possibility to analyze data integrated across temporal and regional scales has led researchers to increase the range of answerable questions and encourage more researchers to participate.

A key difference from other programs is that the PPBio has a full-time data manager, exclusively responsible for guaranteeing that time limits of the data policy are respected, checking data quality and uploading metadata and data.

Therefore, the researchers do not upload the data directly into the database, a process that is prone to errors and idiosyncrasies. Interaction between the data manager and researchers occurs mostly through email or, in some cases, through direct conversation, and usually takes three weeks between the first contact, revision, and the metadata and data being made available online. All researchers are also encouraged to store their field data (field sheets, notebooks) scanned in PDF format in the website as a security copy. Provision of a dedicated data-repository manager has avoided the “empty-archive” syndrome (Nelson 2009).

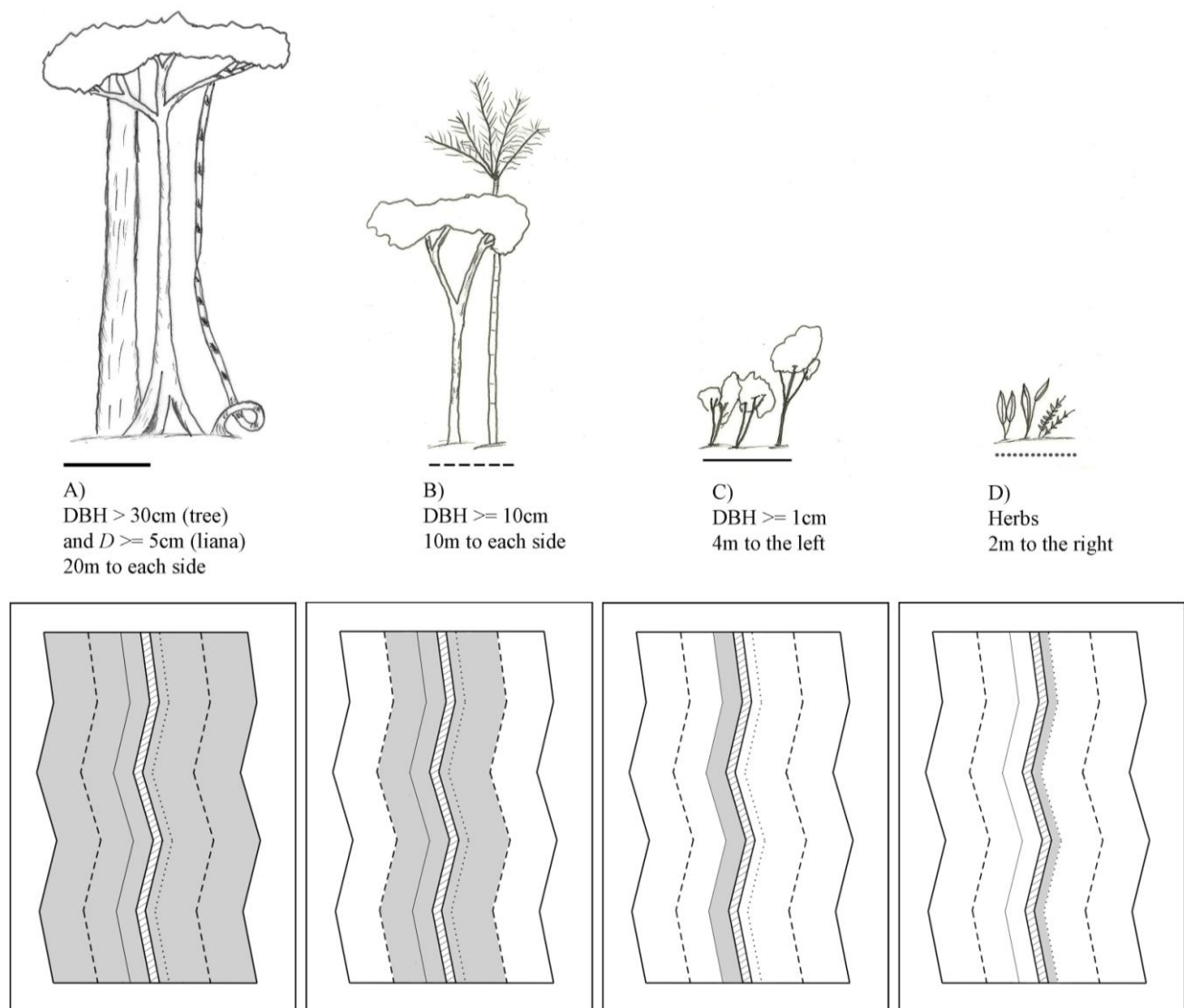


Fig. 2: Schematic representation of a RAPELD plot. The 250 m center line follows the altitudinal contour in straight-line segments of 10 m. Different taxa are surveyed in plots of different width. For example, trees with diameter at breast height (DBH) ≥ 30 cm and lianas with $D \geq 5$ cm at 130 cm from the rooting point are surveyed in a plot that is 40 m wide (discounting the 1 m wide access trail in the middle) (A). Trees (including palms) with DBH ≥ 10 cm is surveyed in a 20 m wide plot (B). Small trees and shrubs with DBH ≥ 1 cm have been surveyed in plots that are 4 m wide. Herbs are surveyed in still narrower plots (D). Also lianas with $D \geq 1$ cm at 130 cm from the rooting point are surveyed in a plot that is 10 m wide.



Plate: Vegetation types featured by the PPBio database GIVD ID SA-BR-001.

A: Flooded forest at BR-319 research site (Photo: F. Penna Espinelli).

B: Stream at Ducke Forest Reserve (Photo: D. Drucker).

C: Vegetation at Viruá National Park (Photo: F. Costa).

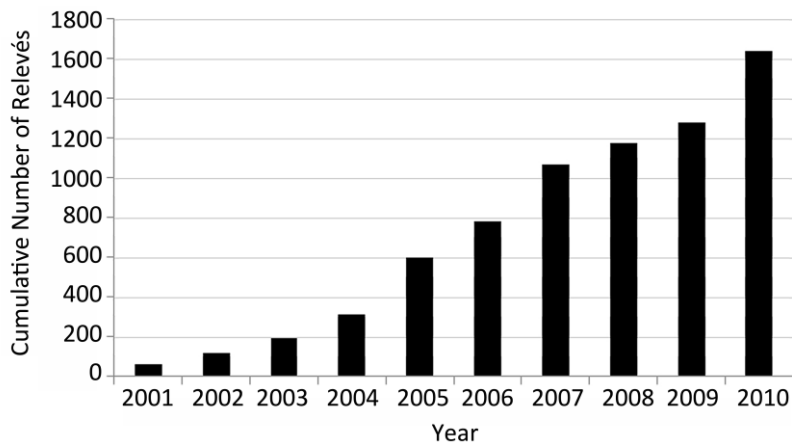


Fig. 3: Cumulative increase in RAPELD relevés in the PPBio data repository with time.

To address complications concerning the heterogeneity of ecological data and data integration, the datasets are stored in a repository with some pre-defined attributes common for every data set, such as research site, plot identifier and date. Other attributes are specific to each data set. Therefore our database is suitable to any type of data. The plot identifiers in the pre-defined attributes are also standardized to avoid misspelling. As relevés are based on the plot identifier, use of idiosyncratic names (e.g. LO_1500 and L-1500) for the same plot will create confusion in the future. Use of generic identifiers, such as “Plot 1”, which could refer to plots in other systems, is potentially even more misleading. In addition, manipulation generally degrades data in a way that the information in the raw data cannot be recovered, and this limits future analyses. Therefore, we strongly recommend that researchers make available raw data that can be easily manipulated to other formats when required.

Data validation consists in a careful examination of outliers, errors and missing data using graphical exploratory data analysis. When an outlier or extreme value is found, the researcher responsible for the data set is alerted. If the value is correct, the data set is published in the website. If the value is an error, the correction is made and only then the data is made available on the website. No outlier or extreme values are discarded. In addition, a major concern is the correct recording of missing data. Often the distinction between a numeric value of zero (e.g. the absence of individuals of a given species in a plot) and absence of data is not clear. Zero values are sometimes repre-

sented by blank cells, which confound zeros with missing data.

Metadata and data, after their validation by the data manager, are stored in a metadata repository on a web server and are available for viewing and downloading directly from the PPBio web site. Metadata were initially stored in manually created pages in HTML format, the data tables attached in PDF and CSV formats, and organized by research site. This system has proved to be useful for the amount of information gathered in the first five years of the program. However, new features were developed and have recently been incorporated into the PPBio information system as a result of the increased value given to data sharing and

the rapid increase in the volume of data, which required new data-querying tools. To facilitate data searches, all the metadata were converted to XML, and the PPBio has installed a METACAT server (<http://knb.ecoinformatics.org/software/metacat/>) to integrate with the Knowledge Network for Biocomplexity (KNB), a network which aims to assist ecological and environmental research. The METACAT server is a repository for metadata and data that makes datasets discoverable through consistently described metadata. This system allows the storage of any kind of data as an attached file in the metadata, whereas storing the data itself in a database system such as MySQL or PostgreSQL requires defined and fixed attributes. Such database systems can be developed for specific purposes from the data repository (e.g. DB for tree biomass measurements). Some metadata are still in Portuguese, but translation of all metadata to English will be available shortly.

Remote sensing

The uniformly-distributed RAPELD plots are 250 m in length for all organisms surveyed, and this allows evaluation of the possibility of using remote sensing to model distributions. RAPELD sites are being used by the remote-sensing community for a variety of calibrations. Schietti et al. (2007) used ground elevation data from several RAPELD Long Term Ecological Research (LTER) sites with distinct vegetation types to calibrate

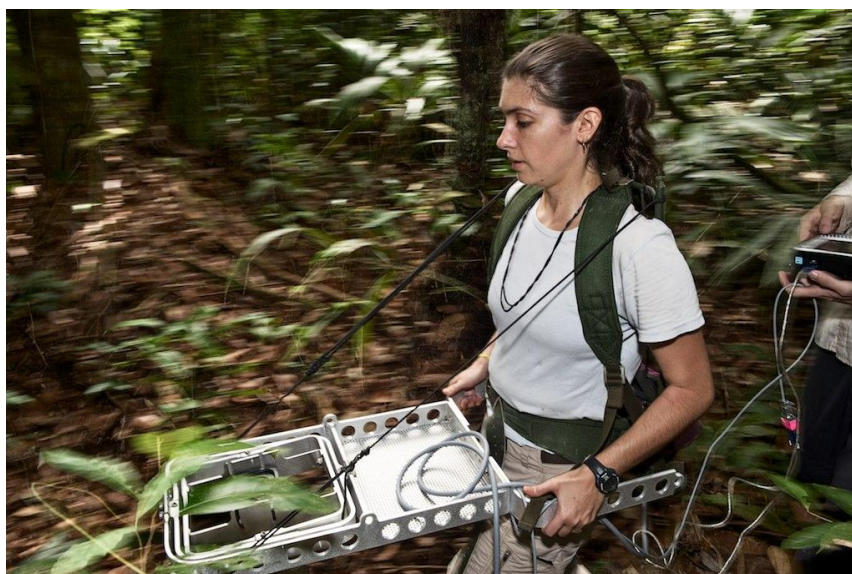


Fig. 4: Ground LIDAR survey used for calibration of relevés based on remote sensing technology (Photo: J. Bryant; <http://www.envirofoto.com>).

Shuttle Radar Topography Mission (SRTM) data and evaluate the use of this radar for association with species distribution patterns. A consortium of Brazilian and North American researchers (<http://www.amazonpire.org/>) are using data of tree live biomass, tree growth and mortality rates and coarse wood debris from RAPELD and other long-term monitoring sites in Amazonia to calibrate air-borne Light Detection and Ranging (LIDAR) data. Airborne LIDAR and RAPELD vegetation data are also being calibrated against ground LIDAR (Fig. 4), and in the future will be used to calibrate satellite-borne LIDAR (Vierling et al. 2011). The calibration among vegetation, ground LIDAR and airborne LIDAR data is being undertaken in the RAPELD plots at Adolpho Ducke Forest Reserve and calibrations between vegetation structure and ground LIDAR data are being conducted in 11 RAPELD sites.

Faunal studies and vegetation data

The PPBio standardized relevés allow not only investigation of the spatial distribution of plant species but also the landscape patterns of diversity and species distribution of animals and their relationships with vegetation structure. As in plant studies, different widths of plots or subsamples are taken along the 250 m center line of each plot to adjust the scale of sampling organisms in accordance with body size, local abundance and movement patterns. Some examples of these studies are the effects of sessile palm density on scorpion abundance (Araujo 2007) and oribatid mites assemblage composition (Moraes 2011); the influence of tree density on the spatial distribution of frog species (Menin et al. 2007) and frog assemblage composition (Condrati 2009); the relationship between tree, shrub and herb dominance and the diversity patterns of coprophagous scarab beetles (Tissiani 2009). Most of these studies are based on Masters and PhD theses and can be freely accessed through the PPBio web site (<http://ppbio.inpa.gov.br/public>).

Conclusion

The PPBio database has provision for making available data collected under other systems. The possibility of integrated biodiversity surveys that are useful to management agencies (RAPELD sys-



Fig. 5: Field course for federal management agencies to show data collection procedures and to teach how data from field surveys can be used for making management decisions.

tem) has revolutionized the relationship between scientists and practicing conservationists. Biologists are now having difficulty keeping pace with the RAPELD field-survey infrastructure being installed by conservation agencies and the private sector (Fig. 5). RAPELD grids and modules have been installed in Brazilian savannas, the Pantanal wetland, agricultural areas in São Paulo State, and plans are underway to install them in the sea. The system has been adopted in Australia and Nepal (<http://ppbio.inpa.gov.br/ppbio/inter/>). Biologists in other countries, such as Bolivia and Peru are seeking funds to install RAPELD infrastructure, and a grid is planned to be installed in Oregon in the USA in 2012 (M. Hero, pers. comm.). The effort to work at scales that are of interest to land managers, such as park administrators and governmental environmental agencies, using techniques that allow integration of data from different taxa and ecosystem processes has not reduced the academic quality of the studies. In fact, most researchers that have worked within the RAPELD system have actually increased their academic output (see publications in <http://ppbio.inpa.gov.br/public>). The RAPELD system has been used as the basis for programs financed by the Federal Education Ministry to increase scientific output of post-graduate students and professors. Perspectives for the future include increased demand for information on biodiversity from profes-

sionals in various areas associated with land-use planning, which will put pressure on politicians for maintenance of herbaria and zoological collections, and taxonomists training. We invite readers to consider the installation of the RAPELD infrastructure in their areas.

Ecological data sharing is necessary to measure the success of research projects and to enable integration in large-scale, long-term multi-disciplinary ecological projects (Reichman et al. 2011). Some funding agencies already require a Data Management Plan for project applications; many important journals have dedicated special issues to address advantages and the main challenges, and most of these require that the datasets of published papers are available in an electronic data repository. However, this culture is still not widely accepted by all groups within the scientific community. Technological issues are not barriers to open access to data. In PPBio, as in many other projects, the cultural barrier is the principal difficulty, and this is the reason that the presence of a data manager has a major impact in the project's success in data sharing.

Since its creation in 2004, the PPBio Amazonia has made available more than 270 metadata and 180 datasets. It has produced 11 field guides, published books on research in the grids installed in Reserva Ducke in the Amazon and the Pirizal in the Brazilian Pantanal (Oliveira et al.

2008, Fernandes e. al. 2010), coordinated more than 70 outreach events and contributed to the production of numerous collaborative research papers, all of which are available in the website. The metadata and data summaries provided by the PPBio are contributing to collaborative strategic actions by several Brazilian institutions to meet the goals set by the Convention on Biological Diversity.

Acknowledgements

The authors would like to thank the financial sponsors MCT, CNPq, FAPEAM, FAPEMAT, Programa Biota/FAPESP (Proc. No. 2006/60954-4), CENTRO DE PESQUISAS DO PANTANAL (CPP); and the institutions that provide political and logistic support: INPA, ICMBio, IBAMA, Serviço Florestal Brasileiro and the associated projects: Projeto Cenários, Pronex, Fronteiras, PELD Banco de Dados, PELD Amazônia, PNPd, ProSET, Projeto Igarapés.

References

- Araujo, J.S. (2007): Métodos de amostragem, influência das variáveis ambientais e guia de identificação das espécies de escorpiões (*Chelicerata*, *Scorpiones*) da Reserva Ducke, Manaus, Amazonas, Brasil. – MSc thesis in Entomology, INPA/UFAM, Manaus.
- Beattie, A.J., Mark, H., Magnusson, W.E., De Nys, R., Smeathers, J., Vincent, J.F.V. (2010): Ecology and bioprospecting. – *Austral Ecology* **36**: 341–356. [CrossRef](#)
- Buckland, J. (1997): Plants. – In: Sutherland, W.J. [Ed.]: *Ecological census techniques*: 186–212. Cambridge: Cambridge University Press.
- Carvalho, F.A., Costa, F.R.C., Salino, A. (2007): Determinantes da estrutura da comunidade de pteridófitas na BR 319, interflúvio Purus Madeira Amazonas, Brasil. – *Revista Brasileira de Biociências* **5**: 1074–1076.
- Condrati, L.H. (2009): Padrões de distribuição e abundância de anuros em áreas ripárias e não ripárias de floresta de Terra Firme na Reserva Biológica do Uatumã – Amazônia Central. – MSc thesis in Ecology, INPA/UFAM, Manaus.
- Costa, F.R.C. (2006): Mesoscale Gradients of Herb Richness and Abundance in Central Amazonia. – *Biotropica* **38**: 711–717. [CrossRef](#)
- Costa, F.R.C., Guillaumet, J.-L., Lima, A.P., Pereira, O.S. (2009): Gradients within gradients: The mesoscale distribution patterns of palms in a central Amazonian forest. – *Journal of Vegetation Science* **20**: 69–78. [CrossRef](#)
- Costa, F.R.C., Magnusson, W.E. (2010): The need for large-scale, integrated studies of biodiversity – The experience of the Program for Biodiversity Research in Brazilian Amazonia. – *Natureza & Conservação* **8**: 1–5.
- Costa, F.R.C., Magnusson, W.E., Luizão, R.C.C. (2005): Mesoscale distribution patterns of Amazonian understory herbs in relation to topography, soil and watersheds. – *Journal of Ecology* **93**: 863–878. [CrossRef](#)
- Drucker, D.P., Costa, F.R.C., Magnusson, W.E. (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.
- Fegraus, E.H., Andelman, S., Jones, M.B., Schildhauer, M. (2005): Maximizing the value of ecological data with structured metadata: An introduction to ecological metadata language (EML) and principles for metadata creation. – *Bulletin of the Ecological Society of America* **86**: 158–168. [CrossRef](#)
- Fernandes, I. M., Signor, C. A., Penha, J. (2010) [Eds.]: *Biodiversidade no Pantanal de Poconé*. – Manaus: Attema Design Editorial.
- Kinupp, V.F., Magnusson, W.E. (2005): Spatial patterns in the understory shrub genus *Psychotria* in Central Amazonia: effects of distance and topography. – *Journal of Tropical Ecology* **21**: 363–374. [CrossRef](#)
- Magnusson, W.E., Lima, A.P., Luizão, R., Luizão, F., Costa, F.R.C., De Castilho, C.V., Kinupp, V.P. (2005): RAPELD: a modification of the Gentry method for biodiversity surveys in long-term ecological research sites. – *Biota Neotropica* **5**: 19–24.
- Menin, M., Lima, A.P., Magnusson, W.E., Waldez, F. (2007): Topographic and edaphic effects on the distribution of terrestrially reproducing anurans in Central Amazonia: mesoscale spatial patterns. – *Journal of Tropical Ecology* **23**: 539–547. [CrossRef](#)
- Moraes, J., Franklin, E., Morais, J.W., Souza, J.L.P. (2011). Species diversity of edaphic mites (*Acar*: *Oribatida*) and effects of topography, soil properties and litter gradients on their qualitative and quantitative composition in 64 km² of forest in Amazonia. – *Experimental & Applied Acarology* **55**: 39–63. [CrossRef](#)
- Moraes, J. (2010): Diversidade de ácaros edáficos (*Acar*: *Oribatida*) e redução do esforço amostral no gradiente ecológico de 10 mil ha de Floresta de Terra-firme da Reserva Ducke, Manaus, AM, Brasil. – MSc thesis in Entomology, INPA/UFAM, Manaus.
- Nelson, B. (2009): Data sharing: empty archives. – *Nature* **461**: 160–163. [CrossRef](#)
- Oliveira, M., Baccaro, F. B., Braga-Neto, R., Magnusson, W. E. (2008) [Eds.]: *Reserva Ducke – A biodiversidade amazônica através de uma grade*. – Manaus: Attema Design Editorial.
- Reichman, O.J., Jones, M.B., Schildhauer, M.P. (2011): Challenges and opportunities of open data in ecology. – *Science* **331**: 703–705. [CrossRef](#)
- Schiatti, J., Drucker, D., Keizer, E., Carneiro-Filho, A., Magnusson, W. (2007): Avaliação do uso de dados SRTM para estudos ecológicos na Amazônia central. – In: *Anais Online do XIII Simpósio Brasileiro de Sensoriamento Remoto (XIII SBSR)*, Florianópolis, Brasil: 6949–6955. São José dos Campos: INPE.
- Tissiani, A.S.O. (2009): Composição da comunidade de Scarabaeidae Coprófagos (Insecta, Coleoptera), em uma área na região norte do Pantanal de Mato Grosso. – MSc thesis in Ecologia e Conservação da Biodiversidade, UFMT, Cuiabá, Mato Grosso.
- Vierling, L. A., Martinuzzi, S., Asner, G.P., Stoker, J., Johnson B. (2011): LiDAR: providing structure. – *Frontiers in Ecology and Environment* **9**: 261–262. [CrossRef](#)
- Watson, I., Novelty, P. (2004): Making the biodiversity monitoring system sustainable: design issues for large-scale monitoring systems. – *Austral Ecology* **29**: 16–30. [CrossRef](#)
- Zuquim, G., Costa, F.R.C., Prado, J. (2007): Fatores que determinam a distribuição de espécies de pteridófitas da Amazônia Central. – *Revista Brasileira de Biociências* **5**: 360–362.
- Zuquim, G., Costa, F.R.C., Prado, J., Braga-Neto, R. (2009): Distribution of pteridophyte communities along environmental gradients in Central Amazonia, Brazil. – *Biodiversity and Conservation* **18**: 151–166. [CrossRef](#)

Flávia Fonseca Pezzini^{1,17,18*}
(flaviapezzini@gmail.com)

Pablo Hendrigo Alves de Melo^{1,17}
(pablopains@gmail.com)

Dayane Mayely Silva de Oliveira^{1,17}
(daya.mso@gmail.com)

Rainer Xavier de Amorim^{1,17}
(rx.a.eng@gmail.com)

Fernando Oliveira Gouvêa de Figueiredo^{1,17}(nandoco06@gmail.com)

Debora Pignatari Drucker^{2,17}
(debora@cnpm.embrapa.br)

Flávio Rogério de Oliveira Rodrigues^{1,17}
(frdor85@gmail.com)

Gabriela Zuquim^{1,17}
(gabizuquim@gmail.com)

Thaise Emilio^{1,17}
(thaise.emilio@gmail.com)

Flávia Regina Capelotto Costa^{1,17}
(flaviacosta001@gmail.com)

William Ernest Magnusson^{1,17}
(bill@inpa.gov.br)

Adeilza Felipe Sampaio³
(adeilzafelipesampaio@yahoo.com.br)

Albertina Pimentel Lima^{1,17}
(lima@inpa.gov.br)

Ana Raquel de Mesquita Garcia⁴
(narraque@gmail.com)

Angelo Gilberto Manzatto^{3,17}
(manzatto@unir.br)

Anselmo Nogueira⁵
(anselmoeco@yahoo.com.br)

Cândida Pereira da Costa⁶
(candidapcosta@gmail.com)

Carlos Eduardo de Araújo Barbosa^{1,17}
(cadubarbosa@gmail.com)

Carolina Bernardes⁷
(carolina1601@yahoo.com)

Carolina Volkmer de Castilho⁸
(cvcastilho@yahoo.com.br)

Catia Nunes da Cunha⁶
(catianc@cpd.ufmt.br)

Cintia Gomes de Freitas^{1,17}
email:(cintuca@gmail.com)

Claymir de Oliveira Cavalcante^{9,17}
(mestreclams@yahoo.com.br)

Diego Oliveira Brandão^{1,17}
(dbrandao13@gmail.com)

Domingos de Jesus Rodrigues^{10,17}
(djmingo23@gmail.com)

Eliana Celestino da Paixão Rodrigues dos Santos⁶
(nanypaixao11@yahoo.com.br)

Fabricio Beggiano Baccaro^{1,17}
(fabricera@gmail.com)

Françoise Yoko Ishida^{11,17}
(fyishida@gmail.com)

Fernanda Antunes Carvalho¹²
(antunesfc@gmail.com)

Gabriel Massaine Moulatlet^{1,17}
(mandaprogabriel@gmail.com)

Jean-Louis Bernard Guillaumet¹³
(jlguillau@club-internet.fr)

José Luiz Purri Veiga Pinto^{1,17}
(zecapurri@gmail.com)

Juliana Schiatti^{1,17}
(jujuschiatti@gmail.com)

Julio Daniel do Vale^{1,17}
(juliodovale@gmail.com)

Lauren Belger¹⁴
(laurenbelger@hotmail.com)

Luciano Martins Verdade¹⁵
(lmverdade@usp.br)

Marcelo Petratti Pansonato^{1,17}
(marcelopansonato@hotmail.com)

Marcelo Trindade Nascimento¹⁶
(mtn@uenf.br)

Márcia Cléia Vilela dos Santos^{10,17}
(marciacleia@gmail.com)

Mariana Souza da Cunha⁹
(marybio@ig.com.br)

Rafael Arruda^{10,17}
(rafael.soares.arruda@gmail.com)

Reinaldo Imbrozio Barbosa⁹
(reinaldo@inpa.gov.br)

Ricardo Laerte Romero^{1,17}
(ecoromero@gmail.com)

Susamar Pansini³
(susa_unir@yahoo.com.br)

Tania Pena Pimentel^{1,17}
(tania@inpa.gov.br)

¹ Coordenação de Pesquisas em Ecologia – CPEC, Instituto Nacional de Pesquisas da Amazônia –INPA, Avenida Efigênio Sales 2239 69011-970, Manaus, Amazonas, BRAZIL

² Embrapa Satellite Monitoring Av. Soldado Passarinho, 303 13070-115, Campinas, São Paulo, BRAZIL

³ Laboratório de Biogeoquímica Ambiental, Núcleo de Ciência e Tecnologia, Universidade Federal de Rondônia Rodovia BR 364 Km 9,5 s/n Rural 78905-000, Porto Velho, Rondônia, BRAZIL

⁴ Ministério da Saúde. Esplanada dos Ministérios– 8º andar – Sala 843 Eixo Monumental, Brasília, Distrito Federal, BRAZIL

⁵ Departamento de Botânica, Instituto de Biociências, Universidade de São Paulo, Rua do Matão, 277, Cidade Universitária 05422-970, São Paulo, BRAZIL

⁶ Laboratório de Ecologia Vegetal, Instituto de Biociências, Universidade Federal de Mato Grosso, Av. Fernando Correia s/n, Coxipó 78060-900, Cuiabá, Mato Grosso, BRAZIL

⁷ Departamento de Engenharia Civil e Ambiental, Faculdade de Tecnologia, Universidade de Brasília, 70910-900, Brasília, Distrito Federal, BRAZIL

⁸ Centro de Pesquisa Agroflorestal de Roraima (EMBRAPA-CPAFRR) BR 174, km 8, Distrito Industrial 69301-970, Boa Vista, Roraima, BRAZIL

⁹ Coordenação de Pesquisas em Ecologia – CPEC, Núcleo de Pesquisas de Roraima – NPRR, Instituto Nacional de Pesquisas da Amazônia – INPA Rua Coronel Pinto 315 – Centro 69301-150, Boa Vista, Roraima, BRAZIL

¹⁰ Núcleo de Estudo da Biodiversidade da Amazônia Mato-Grossense – NEBAM, Instituto de Ciências Naturais Humanas e Sociais – ICNHS, Universidade Federal de Mato Grosso – UFMT Av. Alexandre Ferronato, 1200 78557-267, Sinop, Mato Grosso, BRAZIL

¹¹ Instituto Nacional de Pesquisas da Amazônia – INPA, Programa LBA, Avenida Efigênio Sales 2236, campus II 69060-000, Manaus, Amazonas, BRAZIL

¹² Systematische Botanik und Mykologie, Universität München (LMU), Menzinger Str. 67 80638 München, GERMANY

¹³ Laboratoire de Phanérogamie, Muséum National D'histoire Naturelle 16, rue Buffon 75005 Paris, FRANCE

¹⁴ Fundação Municipal do Meio Ambiente de Itajaí. Rua XV de Novembro, 215, Centro 88301-420, Itajaí, Santa Catarina, BRAZIL

¹⁵ Laboratório de Ecologia Isotópica / CENA / USP Piracicaba, Cx. P. 09 134186-000, Piracicaba, São Paulo, BRAZIL

¹⁶ Laboratório de Ciências Ambientais, Centro de Biociências & Biotecnologia – CBB, Universidade Estadual do Norte Fluminense – UENF Av. Alberto Lamego 2000 28013-620, Campos dos Goyatacazes, Rio de Janeiro, BRAZIL

¹⁷ Instituto Nacional de Ciência e Tecnologia de Estudos Integrados da Biodiversidade Amazônica – CENBAM, Ministério de Ciência e Tecnologia, Conselho Nacional de Desenvolvimento Científico e Tecnológico, INPA Manaus, Amazonas, BRAZIL

*Corresponding author