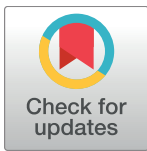# Drivers of metacommunity structure diverge for common and rare Amazonian tree species

**Polyanna da Conceição Bispo[1,2]\***, **Heiko Balzter[1,2]**, **Yadvinder Malhi[3]**, **J. W. Ferry Slik[4]**, **João Roberto dos Santos[5]**, **Camilo Daleles Rennó[6]**, **Fernando D. Espírito-Santo[7]**, **Luiz E. O. C. Aragão[5,8]**, **Arimatéa C. Ximenes[9,10]**, **Pitágoras da Conceição Bispo[11]**

**1** Leicester Institute for Space and Earth Observation, Centre for Landscape and Climate Research, Department of Geography, University of Leicester, Leicester, United Kingdom, **2** NERC, National Centre for Earth Observation at the University of Leicester, Leicester, United Kingdom, **3** Environmental Change Institute, School of Geography and the Environment, University of Oxford, Oxford, United Kingdom, **4** Faculty of Science, Universiti Brunei Darussalam, Gadong, Brunei, **5** Remote Sensing Division, National Institute for Space Research (INPE), São José dos Campos, Brazil, **6** Image Processing Division, National Institute for Space Research (INPE), São José dos Campos, Brazil, **7** Lancaster Environment Centre, Lancaster University, Lancaster, United Kingdom, **8** College of Life and Environmental Sciences, University of Exeter, Exeter, United Kingdom, **9** Laboratory of Systems Ecology and Resource Management, Department of Biology of Organisms, Université Libre de Bruxelles, Brussels, Belgium, **10** Laboratory of Plant Biology and Nature Management, Ecology & Biodiversity, Vrije Universiteit Brussel, Brussels, Belgium, **11** Department of Biological Sciences, Faculty of Sciences and Languages of Assis, State University of São Paulo (UNESP), Assis, Brazil

\* polybispo@gmail.com

## Abstract

We analysed the flora of 46 forest inventory plots (25 m x 100 m) in old growth forests from the Amazonian region to identify the role of environmental (topographic) and spatial variables (obtained using PCNM, Principal Coordinates of Neighbourhood Matrix analysis) for common and rare species. For the analyses, we used multiple partial regression to partition the specific effects of the topographic and spatial variables on the univariate data (standardised richness, total abundance and total biomass) and partial RDA (Redundancy Analysis) to partition these effects on composition (multivariate data) based on incidence, abundance and biomass. The different attributes (richness, abundance, biomass and composition based on incidence, abundance and biomass) used to study this metacommunity responded differently to environmental and spatial processes. Considering standardised richness, total abundance (univariate) and composition based on biomass, the results for common species differed from those obtained for all species. On the other hand, for total biomass (univariate) and for compositions based on incidence and abundance, there was a correspondence between the data obtained for the total community and for common species. Our data also show that in general, environmental and/or spatial components are important to explain the variability in tree communities for total and common species. However, with the exception of the total abundance, the environmental and spatial variables measured were insufficient to explain the attributes of the communities of rare species. These results indicate that predicting the attributes of rare tree species communities based on environmental and spatial variables is a substantial challenge. As the spatial component was relevant for several community attributes, our results demonstrate the importance of

using a metacommunities approach when attempting to understand the main ecological processes underlying the diversity of tropical forest communities.

## Introduction

Throughout the history of studying ecology, researchers have sought to understand the effects of environmental and spatial processes on biodiversity. This quest has led to the collection of a large number of datasets and the formulation of competing theories, such as niche and neutral theories. The niche theory was developed during the 20th century and has been one of the most important theoretical approaches in ecology. This theory assumes that the species in a community are different and that the combination of available resources and environmental conditions determines the local diversity [1]. The coexistence of species within a community, therefore, can be mainly explained by the way in which species with different resource needs and environmental requirements partition existing niches. In this context, it is expected that communities structured by niche-related processes have similar values for local attributes (e.g., species richness, species abundance, biomass and composition) in similar habitat patches. In contrast with the niche theory, the neutral theory is based on the functional equivalence among species and considers dispersion and demographic stochasticity as central phenomena [2,3]. The neutral theory considers that the diversity of a community is a result of the dynamic balance between immigration and extinction [4,5]. According to this theory, it is expected that similar values for local attributes of a community in a particular patch can be determined by the influence of communities of nearby patches, stressing the importance of spatial processes.

In the past few decades, the neutral theory has brought new force to discussions about the processes that determine diversity in communities. In this debate, the proponents of the niche theory have reacted strongly against the assumption that species are equivalent, which is advocated by the neutral theory. For their part, supporters of the neutral theory argue that, in fact, they do not believe that species are equivalent but that this assumption (in addition to the use of stochastic elements) allows the building of simple models with good predictive ability [6]. The supporters of the two theories are usually on opposite sides [5–9]; however, there are researchers who have tried to reconcile them [10]. Although the neutral theory is controversial because it radicalises the assumptions [5], it brings important elements such as the limitation of dispersion and ecological drift, which, along with the niche theory, can help us understand the processes involved in the diversity of species in communities. Perhaps the best way to consider these two theories is as representative of two extremes along a *continuum* [10], among which communities are structured by the relative balance between the mechanisms emphasised by the niche (e.g. environmental filters) and neutral (e.g. dispersion and ecological drift) theories.

Megadiverse tropical forests can be important model ecosystems for understanding the relative roles of environmental and spatial variables on communities. Efforts to understand these roles may help clarify the importance of different processes in maintaining the species diversity of these forests. In terms of the niche theory, variables representative of topography are likely to play an important role in local environmental conditions, which may determine the species diversity, functional and structural attributes and composition of tree communities [11–14]. One advantage of using topography in vegetation studies is that it can be assessed at large scales by remote sensing. Moreover, topography is considered a good surrogate for several important variables of vegetation structure that would be difficult to measure on a larger scale, including nutrient availability, soil moisture and texture, and insolation [13].

The spatial processes, despite their known importance, were long ignored in ecological studies, and until the 1990s, the main focus of ecology was the study of niche-related processes [15]. Since then, spatial processes have been strongly embedded in ecology [16] and new methods and analytical strategies have been proposed, generating a new body of knowledge about the structuring factors of communities [17–19]. This knowledge, along with other information (functional, phylogenetic, etc.), has provided a much more detailed understanding of the processes involved in community structure [20–22], supporting the search for better conservation and biodiversity monitoring strategies [23]. Understanding the role of spatial component in the community structure at different landscape scales is essential, especially given the current scenario of rapid biodiversity loss due to habitat degradation and fragmentation [24].

Metacommunity can be defined by groups of communities that are connected to each other by the movements of individuals of different species [25,26]. This approach recognises that communities are not isolated entities. The metacommunity theory has benefited from discussions regarding niche and neutral theories, and both have helped to understand the effects of environmental and spatial variables on the diversity of species in communities [25,26]. According to the theory, the group of species occurring in a community is determined both by a combination of local factors (interactions among species and interactions of species with local environmental factors) and by the ability of the species to reach that community (by dispersion) [25,27]. In metacommunities structured by the principles of niche theory, it is expected that the environmental component plays a more important role. In contrast, in metacommunities subject to the principles of neutral theory, it is expected that the spatial component plays a more important role. The metacommunity theory encompasses four main models (species sorting, mass effect, patch dynamics and neutral models), which represent points along a *continuum* formed by different combinations of environmental and dispersal processes in different ecological scenarios [25,26,28,29].

A remarkable feature of most communities is the presence of a few common species and many rare species [30,31]. Rare and common species can respond differently to ecological processes [32], depending on the features of the organisms (e.g., competition and dispersion capacities) and the spatial temporal dynamics. For example, based on the niche theory, Tokeshi [15] proposed the composite niche model, arguing that more than one process may be acting on the community. According to this model, common species should fit any model of niche apportionment, while rare species should fit a random assortment model. On the other hand, Siqueira et al. [31] studied metacommunities of aquatic macroinvertebrates and showed that common and rare species responded similarly and that both were mainly structured by niche processes.

The analysis of common and rare species allows the testing of some hypotheses about the processes involved in the structuring of metacommunities [31,33]. Empirical studies have shown that most metacommunities are structured principally by niche processes [28,31]. Our first hypothesis is that niche processes are more important for structuring the metacommunity studied [31], at least for the common species, which we expect to be most affected by competition [34]. In the case of rare species, we expect spatial variables to be more relevant, as these species can be more affected by ecological drift [34]. On the other hand, taking into account that habitat generalist and habitat specialist species differ in terms of population dynamics, we also propose an alternative hypothesis [31,35]. While generalist species occupy habitats with broad environmental variation, specialist species preferentially occupy habitats with specific environmental characteristics, which are generally rare in the landscape [35,36]. In this context, assuming that common species are habitat generalists and rare species are habitat specialists, our alternative hypothesis is that spatial component is more important for common species, while environmental component is crucial for rare species [35].

To test these hypotheses and to identify the role of environmental (topographic) and spatial variables for common and rare species, we analysed the flora of 46 forest inventory plots in the old growth forests of the eastern Amazon region. These analyses were based on vegetation data collected in the field and topographic variables obtained by remote sensing data.

## Material and methods

### Study area

This study was conducted in the Tapajós National Forest (TNF). The TNF is a large protected area of approximately 545,000 ha, located in Amazon biome, western part of Pará State, Brazil (Fig 1). This area has an average annual temperature of 25.5°C and average annual rainfall of 1,820 mm. The local topography ranges from flat to strongly undulating terrain. Predominant soil types in the area are dystrophic oxisol (US classification) or dystrophic yellow latosol (Brazil classification) and red-yellow podzol. Vegetation is mainly ombrophilous dense forest and ombrophilous open forest [37].

We sampled 46 forest inventory plots (Fig 1) of 25 x 100 m in the TNF. Our plots represented a sampling area of 11.5 ha. We installed the plots in different vegetal typologies and topographies [12] using the phyto-ecologic classes from the RADAM-BRASIL project [38]. Our plots encompassed different floristic and geomorphological characteristics [38]. We sampled and identified all individual trees with diameter at breast height (DBH) ≥ 10 cm. The abundance can be represented by the number of individuals and by biomass. To differentiate these two attributes throughout the text, hereafter, we use "abundance" to refer to the number of individuals and "biomass" to refer to above ground biomass. The biomass was calculated by the allometric equation [39], using the measurements of DBH and TH (total height).

$$\text{Biomass} = 0.044 * ((\text{DBH}^2) * \text{TH})^{0.9719}$$

### Topographic data derived from SRTM

We used geomorphometric attributes (elevation, slope, profile curvature and plane curvature) from the Brazilian Geomorphometrics Database (TOPODATA) [40]. TOPODATA is based on the SRTM (Shuttle Radar Topography Mission-version 1, NASA, 2006) and has different neighbourhood operations to calculate geomorphometric variables [40]. TOPODATA is free, and the layers are easily accessible (http://www.dsr.inpe.br/topodata/acesso.php). We also used the vertical distance to the nearest drainage or HAND (height above the nearest drainage). HAND was derived from the SRTM and describes the vertical distance of each point regarding the nearest drainage channel detected by remote sensing [41]. All data used in this study have a spatial resolution of 30 m (Table 1).

Topographic variables obtained on the basis of the SRTM have been used to explain or predict the properties of vegetation [42,43]. These studies have helped in the understanding of the effects of topography on the distribution of different types of vegetation [44–48], floristic composition [12,49] and forest structure [14,50], particularly in tropical areas.

### Data analysis

We defined common and rare species using the criterion of the inflection point of the curve of species abundance (or species biomass) [31]. We defined the inflection point visually; species left of this point were considered as common and species to the right as rare [31]. As matrices with different amounts of information can affect the results, we made comparisons
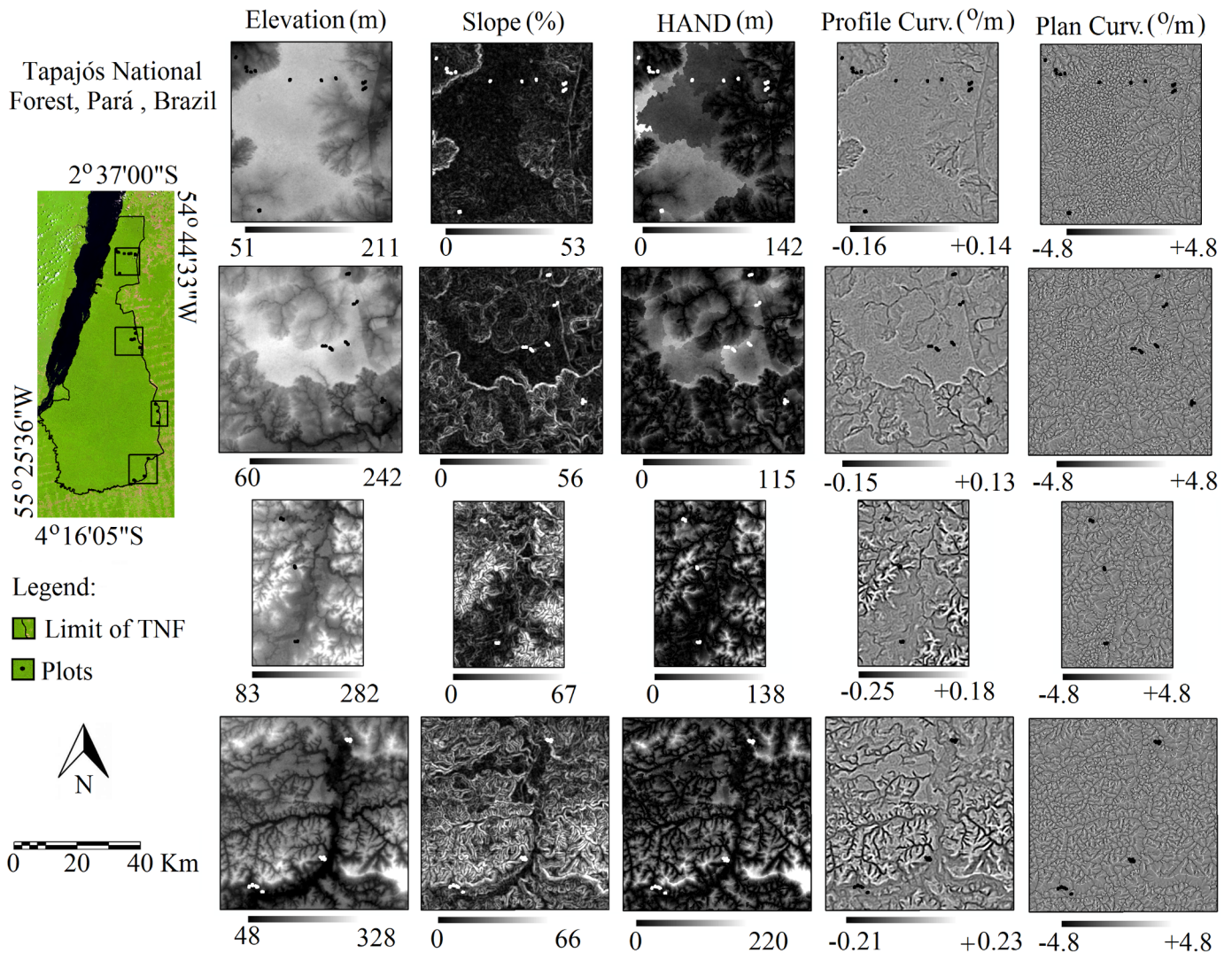
**Fig 1. Study area in the Tapajós National Forest (TNF), Pará State, Brazil, with details of the five geomorphometric variables (elevation, slope, HAND, profile curvature and plan curvature) of the four areas where the 46 plots were distributed.**

considering the same information content. First, the information content of the matrices of common and rare species was calculated based on the binomial variance of the incidence matrix, $\Sigma p_i(1-p_i)$, where $p_i$ is the proportion of plots occupied by $i^{th}$ species [31,51]. As the matrix of rare species had a higher information content, we removed rare species, following species rank, until this matrix had the same information content as that of common species.

After defining common and rare species with the same information content, we performed data analyses considering univariate and multivariate community attributes. The univariate attributes were standardised richness (residuals of regression between abundance and richness), total abundance (sum of the abundance of the species per plot) and total biomass (total biomass of the species per plot). The multivariate attributes (species x plots) were represented by three different matrices of composition (1. composition based on incidence or presence-absence; 2. composition based on abundance; and 3. composition based on biomass).

The richness of species is usually positively correlated with abundance. As abundance can explain part of the variation in richness, without due caution, we can erroneously conclude

**Table 1. Definitions of the topographic variables used in this study.**

| Topographic variables | Description |
| --- | --- |
| Elevation (*h*) | Terrain altitude. This is related to the altitude distribution of soil and climate, determining different landscape vegetation patterns. |
| Slope (*G*) | Inclination angle of the local surface. This has a direct effect on the balance between soil water infiltration and surface runoff and controls the intensity of flows of matter and insolation. This set of factors results in environments with different physical and biological characteristics, allowing the establishment of different types of vegetation. |
| Profile curvature (*kv*) | Concave/convex character of the terrain. This characterizes the land surface, which is directly associated with hydrological and transport properties and may directly influence the distribution and development of vegetation. |
| Plan curvature (*kh*) | Divergent/convergent character of flows of matter on the ground when analysed on a horizontal projection. As with the profile curvature, the plan curvature characterises the land surface, which is directly associated with hydrological and transport properties and may indirectly influence vegetation. |
| Height above the nearest drainage (*HAND*) | Describes the vertical distance of each point regarding the nearest drainage channel. It can reveal the local water table conditions (the lower the HAND value, the closer the water table is of the surface). |

that similar factors are important in explaining both community attributes. Disentangling richness from abundance is necessary to understand the real effect of topography on richness (free of abundance). For this reason, we used the residuals of the regression between abundance and richness as a standardised measure of richness (standardised richness). In this case, the residuals indicate the part of the variation in richness that cannot be explained by abundance, in other words, richness free of abundance.

In general, the community matrix (species x plots) based on abundance has many zeros, which is a problem for multivariate analysis based on Euclidian distances, such as Principal Component Analysis (PCA) and Redundancy Analysis (RDA). A strategy to minimise this problem is using the Hellinger transformation [52]. In the case of composition based on abundance, prior to the analyses, we transformed the data matrix using the Hellinger method [53].

For the analyses, we used a multiple partial regression to partition the specific effects of the topographic and spatial variables on the univariate response variables (standardised richness, abundance and total biomass), and we used partial RDA to partition these effects on the multivariate response matrices represented by incidence, abundance and biomass [17,54,55]. The RDA is a direct gradient analysis based on multiple regression that addresses the variation in a multivariate response matrix (in our case, composition based on incidence, abundance and biomass) and one or more matrices of explanatory variables (in our case, topographic and spatial variables) [55].

Spatial variables were obtained using the PCNM (Principal Coordinates of Neighbourhood Matrix) method [55]. The PCNM is based on Principal Coordinate Analysis obtained from a geographic distance matrix. The eigenvectors (axes) obtained from this analysis are called PCNMs, are uncorrelated and represent different spatial patterns, from coarse (axes with higher eigenvalues) to more refined (axes with smaller eigenvalues) [55,56]. In this paper, we extracted the spatial variables (PCNMs) from a Euclidean distance matrix between plots, which were represented by eigenvectors with positive eigenvalues and with spatial autocorrelation according to Moran's I index [57]. For the analyses, we selected variables using the forward selection method to evaluate only the environmental and spatial variables that were more related to the studied metacommunity.

We assessed the following fractions: the environmental (topography) component independent of the space (a), environmental component inseparable of the spatial component (b), spatial component independent of the environment (c) and component not explained (d). As the coefficient of determination ($R^2$) is influenced by the sample size and number of predictor variables, we used the adjusted $R^2$ to determine the importance of each assessed fraction [58]. We performed the analyses in the computing environment R version 2.13 [59], associated with the PCNM package [60] for obtaining the spatial variables, Packfor [61] for variable selection and Vegan [62] for multiple regression and RDA.

## Results

The results revealed that when considering the inflection points of the abundance curves, 22 species were considered common, and 208 species were considered rare (Fig 2a). The 93 rarest species had the same information content as the 22 most common species. When the inflection of the curves of species biomass was considered, 35 species were considered common, and 195 were considered rare (Fig 2b). In this case, the 94 rarest species had the same information content as the 35 most common species.

### Standardised richness, abundance and total biomass

Total standardised richness (without the abundance effect) and standardised richness of rare species could not be explained by any of the measured variables (topographic or spatial variables) (Table 2). Nevertheless, the standardised richness of common species was explained significantly by topography (Table 2). The total abundance was explained significantly by spatial variables (Table 2). Moreover, the abundance of common species was explained by both topographic and spatial variables, whereas rare species were explained just by the spatial variables (Table 2). The topography explains significantly the total tree biomass for all and for common species (Table 2). The total biomass of rare species could not be explained by any of the measured variables (Table 2).

### Composition (incidence, abundance and biomass)

Our data reveal that topographic and spatial variables explained significant proportions of the variability when the analysis was based on an incidence matrix and on an abundance matrix, both for total and common species (Table 2). In these cases, the fraction explained by spatial variables was larger than the fraction explained by environmental variables (Table 2). Topographic and spatial variables did not explain the variation of rare species. When the analysis was based on biomass, the variability was explained significantly by topographic (larger fraction) and spatial variables (smaller fraction) for total species, but by only topography for common species and by no factor for rare species (Table 2).

## Discussion

Our results show that different attributes (richness, abundance, biomass and composition based on incidence, abundance and biomass) used to study this metacommunity respond differently to environmental and spatial processes. Common and rare species differ in terms of biological traits [63–65] and how they relate to environmental and spatial components [35,36,66]. Thus, our expectation was that the role of environmental and spatial variables differed between common and rare species. This expectation is confirmed by our study, suggesting that common and rare species are subject to different processes.
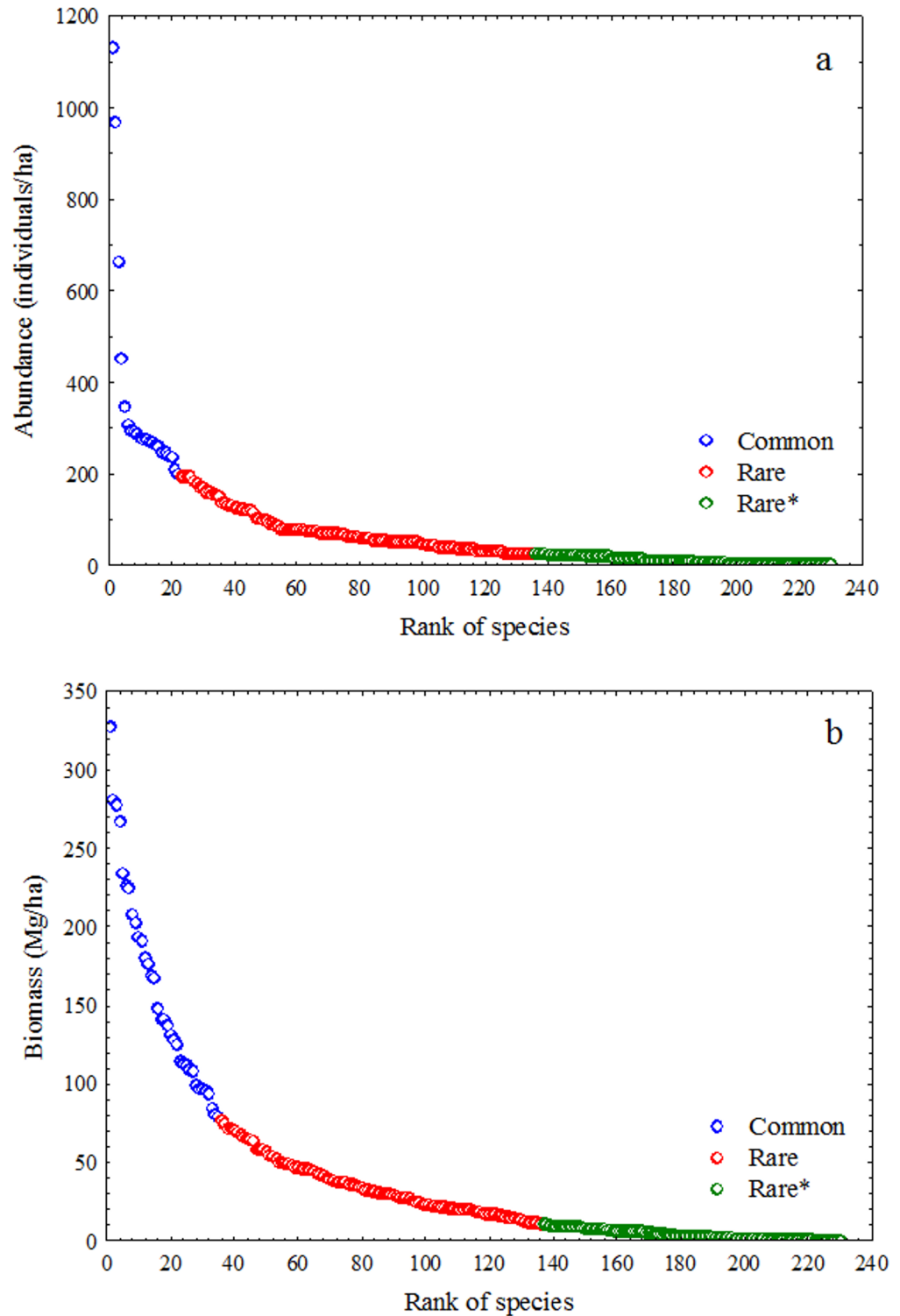
**Fig 2. Rank of abundance (a) and rank of biomass (b) of Amazonian tree species of a metacommunity in the Tapajós National Forest, Pará State, Brazil.** * indicates the rare species with the same information content as common species.

**Table 2. Results of the partial multiple regression and partial RDA with the coefficient of determination ($R^2$) for whole community (total), common and rare species.** Topography refers to the effects of geomorphometric variables without spatial component, shared refers to the effects of common variation between topographic and spatial components, and space refers to the spatial effects without topography. Common and rare species have the same information content and were delimited based on the inflection point of the species x abundance curve (in the case of abundance) or species x biomass curve (in the case of biomass).

| | Topography (%) | Shared (%) | Space (%) | Not explained (%) |
|---|---|---|---|---|
| *Std. Richness* | | | | |
| Total | - | - | - | 100.00 |
| Common (1–22) | 10.4* | - | - | 89.6 |
| Rare (137–230) | - | - | - | 100.00 |
| *Abundance* | | | | |
| Total | - | - | 25.9** | 74.1 |
| Common (1–22) | 11.4** | - | 10.3** | 78.3 |
| Rare (137–230) | - | - | 26.3** | 73.7 |
| *Biomass* | | | | |
| Total | 27.4*** | 15.1 | - | 57.5 |
| Common (1–35) | 14.2** | 12.5 | - | 73.3 |
| Rare (136–230) | - | - | - | 100.00 |
| *C. Incidence* | | | | |
| Total | 2.8*** | 1.8 | 6.6*** | 88.8 |
| Common (1–22) | 2.7** | 2.7 | 5.4*** | 89.2 |
| Rare (137–230) | - | 0.9 | 0.3ns | 98.8 |
| *C. Abundance* | | | | |
| Total | 6.2*** | 0.4 | 10.4*** | 83.0 |
| Common (1–22) | 5.1** | 2.7 | 10.6*** | 81.6 |
| Rare (137–230) | 0.1ns | 0.8 | 0.4ns | 98.7 |
| *C. Biomass* | | | | |
| Total | 3.1** | 3.2 | 1.6* | 92.1 |
| Common (1–35) | 4.0** | 4.2 | 1.9ns | 89.9 |
| Rare (136–230) | - | - | - | 100.00 |

Numbers in parentheses refer to the rank position of the species. Univariate attributes: 1) std. richness (standardised richness, residuals of regression between abundance and richness); 2) abundance (sum of the abundance of the species per plot); and biomass (sum of biomass of the species per plot). Multivariate attributes (species x plots): 1) C. Incidence (composition based on incidence); C. Abundance (composition based on abundance); and C. Biomass (composition based on biomass).

*p < 0.05.

**p < 0.01.

***p < 0.001.

ns non-significant.

https://doi.org/10.1371/journal.pone.0188300.t002

The communities are composed of few common species and many rare species. Due to the greater number of individuals, common species interact strongly with the various components of the system. Therefore, a common question is whether the common species are sufficient to describe the attributes (e.g., richness, abundance and composition) of the whole community [51,66,67]. If this is the case, studies on communities could focus on common species, which are more easily sampled. Several studies have shown that the conclusions found for all species are equivalent to those found using only common species [66–68]. Our study only partially confirms this expectation and it adds complexity to the picture by showing that this depends on the analysed attribute. For example, for standardised richness, total abundance (univariate) and composition based on biomass, the results for common species differ from those obtained

for all species. For total biomass (univariate) and compositions based on incidence and abundance, there is a correspondence between the results obtained for the total community and common species. Our data suggest that for these last attributes, it is possible to draw appropriate conclusions for an entire community based on only common species.

The results to standardised richness are in disagreement with other studies which have shown that the richness patterns (total community) can be predicted by the richness of common species [12,69]. In our study, topography only explained the variability of the richness of common species, and neither topographic nor spatial variables explained the variability of the richness of the total community or of rare species. According to Lennon et al. [66], the richness of common species can be more easily explained by simple environmental gradients when compared to the richness of rare species. As the richness of rare species can be associated with rare environments [36,66], it is harder to predict it. Thus, the absence of the effect of the environment on the richness of rare species may be due to no inclusion of important but difficult to measure environmental variables, which must be associated with uncommon niches [36]. In this context, part of the variation of the standardised richness of common species can be predicted by the environmental gradient (in our case, topography), which does not occur with rare species.

Total abundance (univariate) and biomass (univariate) were explained by different processes when the total community, common species and rare species were considered. For total abundance, in all combinations (total community, common and rare species), spatial variables explained part of the variability. These results show that spatial processes determine a part of the total abundance variation, suggesting that the effects of mass are relevant when abundance is considered. For common species, in addition to spatial processes, the environmental variables were also important. In the case of the total biomass of the entire community and of the common species, only environmental variables were relevant. The topography influences other extremely important variables, such as soil texture and the availability of nutrients and water [11,70,71], and this may explain the results found for biomass. Our observations may have practical consequences. For example, we have sought ways to predict and monitor biomass at larger scales, and the relationship with topography can help since it influences other extremely important variables to the vegetation structure. Therefore, topography can be a surrogate of several variables that are difficult to measure in building predictive models that facilitate the monitoring of biomass and carbon stocks in tropical forests.

The data presented here show that the ecological processes underlying composition differ between common and rare species, in agreement with the results of Tsang and Bonebrake [68], who studied the composition of butterflies. On the other hand, this result disagrees with other data for different organisms (e.g., aquatic macroinvertebrates [31,67] and macrophytes [33]), which show that common and rare species are governed by the same processes, in these cases, by processes related to the niche. Specifically for vegetation, Wang et al. [72], who studied the effects of topography on the species composition of a subtropical forest, also verified that environmental variables are important determinants of the variation of the composition of common and rare species. However, this relation was much weaker for rare species. Our initial hypothesis was that the composition of rare species could be explained by the environmental and/or by spatial variables. Our results refuted this hypothesis since none of the components (environmental or spatial) explained the variation of the composition of rare species. This result may be a reflection of stochastic factors and of the non-inclusion of specific variables important for rare species.

Our initial expectations to the composition (based on incidence, abundance and biomass) were that the niche-related processes were the most relevant to explain data variability and that the studied metacommunity would follow the species sorting (SS) model [28]. The results

showed that both environmental and spatial variables were relevant to explain variability for both total and common species (except for biomass of common species, which was explained by only environmental processes). These results suggest that on the scale studied, the data fit the species sorting (SS) + mass effect (ME) models. The SS and SS + ME models have been the most frequently adjusted models for empirical data. For example, Cottenie [28] studied 158 metacommunities and found that 44% of them fit the SS model and that 29% the SS + ME model. These patterns (SS or SS + ME) have been confirmed by most studies since Cottenie [28]. It is important to note in our study that in the case of composition based on incidence and abundance, the spatial component was more important than environmental component. In this context, several studies have shown that spatial variables explain a relevant part of the variability in both tropical and temperate forests [73,74]. However, the inclusion of other important environmental variables could increase the percentage explained by the environment and reduce the importance of spatial variables, whose effects may be a reflection of mass effects and dispersion difficulties, as well as responses to spatially structured but not measured environmental variables [25].

The result that the compositions based on incidence and abundance were explained more by spatial variables than by environmental variables is in agreement with a study carried out in the Bolivian Amazon Forest [75]. Myers et al. [75], who studied two forests in different latitudes, found that beta diversity in a tropical forest was explained predominantly by spatial variables, while beta diversity in a temperate forest was mostly explained by environmental variables. These authors suggest that in megadiverse systems with many rare species, such as tropical forests, intraspecific aggregation is more related to the limitation of the dispersion; while in temperate forests with fewer species, intraspecific aggregation is more related to environmental filters. Therefore, intraspecific aggregation, which generates beta diversity, is influenced by different processes in rich and poor metacommunities [75]. We also suggest other explanation, as forests in different latitudes may differ in complexity, a possible result is that the same set of environmental variables explain a smaller fraction of beta diversity in tropical forests (more complex system) than in temperate forests (less complex system). Therefore, a greater proportion of important unmeasured environmental variables in more complex forests could also increase the fraction explained by spatial variables in tropical forests. In this context, we suggest that in tropical forests, the spatial component tends to be larger than the environmental component, at least when compositions are based on incidence and abundance. In the case of composition based on biomass, the fraction explained by the environment is greater than the fraction explained by spatial variables.

In this study, we found that much of the variability was not explained by the environmental (represented by the topography) or by spatial variables. This is a relatively common result in studies of metacommunities [17,19]. Two main factors can help explain this result: 1) there are many environmental variables in tropical forests that affect the biota, and often only one portion of them (in our case, topographic variables) is measured; 2) tropical forests have a large number of biotic interactions, which despite having the potential to affect the community, are impossible to measure to capture their complexity. De Caceres at al. [73], who studied tropical, subtropical and temperate forests, found that the unexplained fraction was negatively correlated with latitude. The proportion of unexplained variability is probably due to stochastic and unmeasured variables. Baldeck et al. [18] showed that in addition to topography, the inclusion of variables such as nutrients can improve a model's explanatory power. Although the topography is a good substitute for other variables that are difficult to measure, the inclusion of additional relevant variables should decrease the proportion of unexplained variability in the model.

This study revealed that for total biomass (univariate) and for compositions based on incidence and abundance, there was a correspondence between the results obtained for the total community and for common species. The possibility of monitoring the variability of the tropical forests based on only common species is highly relevant, especially given the current high deforestation rates. The variation of the measured community attributes of common species was explained by topographic and/or spatial variables. None of the components explained the measured community attributes (except for the total abundance) of rare species. These results indicate that predicting the attributes of rare species tree communities from environmental and spatial variables is a considerable challenge. In summary, our data show that in general, depending on the attribute, environmental and/or spatial variables are important to explain the variability in tree metacommunities. However, there are still doubts regarding whether the spatial component and large fractions of unexplained variability in forest metacommunities are due to insufficient data or are a feature of these systems [18,74]. In this study, the spatial component was important for several community attributes demonstrating the importance of a metacommunity approach when attempting to understand the main ecological processes underlying the diversity of tropical forest communities.

## Authorization for the field work

The study was carried out in the Tapajós National Forest (TNF) and dendrometric measurements (diameter at breast height and height) as well as botanical identification of the trees were done, just inside of this area. The authorization to carry out the field work at TNF was provided by the Instituto Chico Mendes de Conservação da Biodiversidade-ICMBio/MMA (SISBIO n. 20591–1). This study did not involve endangered or protected species and no biological samples were taken.

## Supporting information

**S1 Table. Selected topographic and spatial variables for partial multiple regression and partial redundancy according to forward selection.**
(DOCX)

## Acknowledgments

## Author Contributions

**Conceptualization:** Polyanna da Conceição Bispo, Pitágoras da Conceição Bispo.

**Data curation:** Polyanna da Conceição Bispo, Heiko Balzter, Camilo Daleles Rennó, Pitágoras da Conceição Bispo.

**Formal analysis:** Polyanna da Conceição Bispo, Pitágoras da Conceição Bispo.

**Funding acquisition:** Polyanna da Conceição Bispo, Heiko Balzter.

**Investigation:** Polyanna da Conceição Bispo, João Roberto dos Santos, Pitágoras da Conceição Bispo.

**Methodology:** Polyanna da Conceição Bispo, Pitágoras da Conceição Bispo.

**Project administration:** Polyanna da Conceição Bispo, Heiko Balzter.

**Resources:** Polyanna da Conceição Bispo, Heiko Balzter, João Roberto dos Santos, Pitágoras da Conceição Bispo.

**Software:** Polyanna da Conceição Bispo, Heiko Balzter, João Roberto dos Santos, Pitágoras da Conceição Bispo.

**Supervision:** Polyanna da Conceição Bispo.

**Validation:** Polyanna da Conceição Bispo, Pitágoras da Conceição Bispo.

**Visualization:** Polyanna da Conceição Bispo.

**Writing – original draft:** Polyanna da Conceição Bispo, Pitágoras da Conceição Bispo.

**Writing – review & editing:** Polyanna da Conceição Bispo, Heiko Balzter, Yadvinder Malhi, J. W. Ferry Slik, João Roberto dos Santos, Camilo Daleles Rennó, Fernando D. Espírito-Santo, Luiz E. O. C. Aragão, Arimatéa C. Ximenes, Pitágoras da Conceição Bispo.

# References

1. Hutchinson GE. Population Studies—Animal Ecology and Demography—Concluding Remarks. Cold Spring Harbor Symposia on Quantitaive Biology. 1957; 22: 415–427. https://doi.org/10.1101/SQB.1957.022.01.039

2. Hubbell SP. Neutral theory in community ecology and the hypothesis of functional equivalence. Functional Ecology. 2005. 19: 166–172. https://doi.org/10.1111/j.0269-8463.2005.00965.x

3. Hubbell SP. The Unified Neutral Theory of Biodiversity and Biogeography [Internet]. Monographs in Population Biology 32. Princeton University Press. 2001, 375pp; ISBN 0-691-02128-7. https://doi.org/10.1016/S0006-3207(02)00228-8

4. Etienne RS, Alonso D. Neutral community theory: How stochasticity and dispersal-limitation can explain species coexistence. Journal of Statistical Physics. 2007; 128: 485–510. https://doi.org/10.1007/s10955-006-9163-2

5. Halley JM, Iwasa Y. Neutral theory as a predictor of avifaunal extinctions after habitat loss. Proc Natl Acad Sci. 2011; 108: 2316–2321. https://doi.org/10.1073/pnas.1011217108 PMID: 21262797

6. Rosindell J, Hubbell SP, He F, Harmon LJ, Etienne RS. The case for ecological neutral theory. Trends in Ecology and Evolution. 2012; 27: 203–208. https://doi.org/10.1016/j.tree.2012.01.004 PMID: 22341498

7. Clark JS. The missing link: From island extinction to Neutral Theory (a reply to Halley and Iwasa). Trends in Ecology and Evolution. 2012; 27: 364. https://doi.org/10.1016/j.tree.2012.04.009

8. Clark JS. The coherence problem with the Unified Neutral Theory of Biodiversity. Trends in Ecology and Evolution. 2012; 27: 198–202. https://doi.org/10.1016/j.tree.2012.02.001 PMID: 22401902

9. Wennekes PL, Rosindell J, Etienne RS. The Neutral-Niche Debate: A Philosophical Perspective. Acta Biotheor. 2012; 60: 257–271. https://doi.org/10.1007/s10441-012-9144-6 PMID: 22302362

10. Furniss TJ, Larson AJ, Lutz JA. Reconciling niches and neutrality in a subalpine temperate forest. Ecosphere. 2017; 8. https://doi.org/10.1002/ecs2.1847

11. Castilho CV, Magnusson WE, de Araújo RNO, Luizão RCC, Luizão FJ, Lima AP, et al. Variation in aboveground tree live biomass in a central Amazonian Forest: Effects of soil and topography. For Ecol Manage. 2006; 234: 85–96. https://doi.org/10.1016/j.foreco.2006.06.024

12. Bispo PC, Valeriano MM, Santos JR. Effects of the geomorphometric characteristics of the local terrain on floristic composition in the central Brazilian Amazon. Austral Ecol. 2012; 37: 491–499. https://doi.org/10.1111/j.1442-9993.2011.02310.x

13. Liu J, Yunhong T, Slik JWF. Topography related habitat associations of tree species traits, composition and diversity in a Chinese tropical forest. For Ecol Manage. 2014; 330: 75–81. https://doi.org/10.1016/j.foreco.2014.06.045

14. Bispo PC, Santos JR, Valeriano MM, Graça PMLA, Balzter H, França H, et al. Predictive models of primary tropical forest structure from geomorphometric variables based on SRTM in the Tapajós region, Brazilian Amazon. PLOS ONE. 2016; 11. https://doi.org/10.1371/journal.pone.0152009 PMID: 27089013

15. Tokeshi M. Niche Apportionment or Random Assortment: Species Abundance Patterns Revisited. J Anim Ecol. 1990; 59: 1129–1146. https://doi.org/10.2307/5036

16. Legendre P. Spatial autocorrelation: trouble or new paradigm? Ecology. 1993; 74: 1659–1673. https://doi.org/10.2307/1939924

17. Beisner BE, Peres-Neto PR, Lindström ES, Barnett A, Longhi ML. The role of environmental and spatial processes in structuring lake communities from bacteria to fish. Ecology. 2006; 87: 2985–2991. https://doi.org/10.1890/0012-9658(2006)87[2985:TROEAS]2.0.CO;2 PMID: 17249222

18. Baldeck CA, Harms KE, Yavitt JB, John R, Turner BL, Navarrete H, et al. Soil resources and topography shape local tree community structure in tropical forests Soil resources and topography shape local tree community structure in tropical forests. Proc R Soc London B Biol Sci. 2013; 280: 20122532.

19. Branco CCZ, Bispo PC, Peres CK, Tonetto AF, Branco LHZ. The roles of environmental conditions and spatial factors in controlling stream macroalgal communities. Hydrobiologia. 2014; 732: 123–132. https://doi.org/10.1007/s10750-014-1852-2

20. Leibold MA, Economo EP, Peres-Neto P. Metacommunity phylogenetics: Separating the roles of environmental filters and historical biogeography. Ecol Lett. 2010; 13: 1290–1299. https://doi.org/10.1111/j.1461-0248.2010.01523.x PMID: 20735465

21. Spasojevic MJ, Copeland S, Suding KN. Using functional diversity patterns to explore metacommunity dynamics: A framework for understanding local and regional influences on community structure. Ecography (Cop). 2014; 37: 939–949. https://doi.org/10.1111/ecog.00711

22. Saito VS, Soininen J, Fonseca-Gessner AA, Siqueira T. Dispersal traits drive the phylogenetic distance decay of similarity in Neotropical stream metacommunities. J Biogeogr. 2015; 42: 2101–2111. https://doi.org/10.1111/jbi.12577

23. Siqueira T, Bini LM, Roque FO, Cottenie K. A metacommunity framework for enhancing the effectiveness of biological monitoring strategies. PLOS ONE. 2012; 7. https://doi.org/10.1371/journal.pone.0043626 PMID: 22937068

24. Barlow J, Lennox GD, Ferreira J, Berenguer E, Lees AC, Nally RM, et al. Anthropogenic disturbance in tropical forests can double biodiversity loss from deforestation. Nature. 2016; 535: 144–147. https://doi.org/10.1038/nature18326 PMID: 27362236

25. Leibold MA, Holyoak M, Mouquet N, Amarasekare P, Chase JM, Hoopes MF, et al. The metacommunity concept: A framework for multi-scale community ecology. Ecology Letters. 2004; 7: 601–613. https://doi.org/10.1111/j.1461-0248.2004.00608.x

26. Logue JB, Mouquet N, Peter H, Hillebrand H. Empirical approaches to metacommunities: A review and comparison with theory. Trends in Ecology and Evolution. 2011; 26: 482–491. https://doi.org/10.1016/j.tree.2011.04.009 PMID: 21641673

27. Santos AS, Saraiva DD, Müller SC, Overbeck GE. Interactive effects of environmental filtering predict beta-diversity patterns in a subtropical forest metacommunity. Perspect Plant Ecol Evol Syst. Elsevier GmbH. 2015; 17: 96–106. https://doi.org/10.1016/j.ppees.2015.01.002

28. Cottenie K. Integrating environmental and spatial processes in ecological community dynamics. Ecol Lett. 2005; 8: 1175–1182. https://doi.org/10.1111/j.1461-0248.2005.00820.x PMID: 21352441

29. Alexander HM, Foster BL, Ballantyne F, Collins CD, Antonovics J, Holt RD. Metapopulations and metacommunities: Combining spatial and temporal perspectives in plant ecology. J Ecol. 2012; 100: 88–103. https://doi.org/10.1111/j.1365-2745.2011.01917.x

30. Spitale D. A Comparative Study of Common and Rare Species in Spring Habitats. Ecoscience. 2012; 19: 80–88. https://doi.org/10.2980/19-1-3435

31. Siqueira T, Bini LM, Roque FO, Marques Couceiro SR, Trivinho-Strixino S, Cottenie K. Common and rare species respond to similar niche processes in macroinvertebrate metacommunities. Ecography (Cop). 2012; 35: 183–192. https://doi.org/10.1111/j.1600-0587.2011.06875.x

32. László Z, Rákosy L, Tóthmérész B. Landscape and local variables benefit rare species and common ones differently. J Insect Conserv. 2014; 18: 1203–1213. https://doi.org/10.1007/s10841-014-9734-5

33. Alahuhta J, Johnson LB, Olker J, Heino J. Species sorting determines variation in the community composition of common and rare macrophytes at various spatial extents. Ecol Complex. 2014; 20: 61–68. https://doi.org/10.1016/j.ecocom.2014.08.003

34. Chase JM, Amarasekare P, Cottenie K, Gonzalez A, Holt R, Holyoak M, et al. Competing theories for competitive metacommunities. In: Metacommunities spatial dynamics and ecological communities (eds Holyoak M, Leibold M and Holt R). 2005. pp. 335–354, University of Chicago Press, Chicago, IL.

35. Pandit SN, Kolasa J, Cottenie K. Contrasts between habitat generalists and specialists: An empirical extension to the basic metacommunity framework. Ecology. 2009; 90: 2253–2262. https://doi.org/10.1890/08-0851.1 PMID: 19739387

36. Markham J. Rare species occupy uncommon niches. Sci Rep. 2014; 4: 6012. https://doi.org/10.1038/srep06012 PMID: 25110113

37. Veloso HP, Rangel Filho ALR, Lima JCA. Classificação da Vegetação Brasileira Adaptada a um Sistema Universal. Rio de Janeiro, IBGE, Departamento de Recursos Naturais e Estudos Ambientais. 1991. ISBN 85-240-0384-7

38. RADAMBRASIL Departamento Nacional de Produção Mineral. No TitleFolha AS.21- Santarém. Geologia, geomorfologia, pedologia, vegetação e uso potencial da terra. Rio de Janeiro: DNPM (Levantamento dos Recursos Naturais v. 10). 1976.

39. Brown S, Gillespie AJR, Lugo AE. Biomass estimation methods for tropical forests with applications to forest inventory data. Forest science. 1989; 35: 881–902.

40. Valeriano MM, Rossetti D F. Topodata: Brazilian full coverage refinement of SRTM data. Appl Geogr. 2012; 32: 300–309. https://doi.org/10.1016/j.apgeog.2011.05.004

41. Rennó CD, Nobre AD, Cuartas LA, Soares JV, Hodnett MG, Tomasella J, et al. HAND, a new terrain descriptor using SRTM-DEM: Mapping terra-firme rainforest environments in Amazonia. Remote Sens Environ. 2008; 112: 3469–3481. https://doi.org/10.1016/j.rse.2008.03.018

42. Kellndorfer J, Walker W, Pierce L, Dobson C, Fites JA, Hunsaker C, et al. Vegetation height estimation from Shuttle Radar Topography Mission and National Elevation Datasets. Remote Sens Environ. 2004; 93: 339–358. https://doi.org/10.1016/j.rse.2004.07.017

43. Simard M, Rivera-Monroy VH, Mancera-Pineda JE, Castañeda-Moya E, Twilley RR. A systematic method for 3D mapping of mangrove forests based on Shuttle Radar Topography Mission elevation data, ICESat/GLAS waveforms and field data: Application to Ciénaga Grande de Santa Marta, Colombia. Remote Sens Environ. 2008; 112: 2131–2144. https://doi.org/10.1016/j.rse.2007.10.012

44. Bispo P. C.; Valeriano M. M.; Kuplich TM. Variáveis geomorfométricas locais e sua relação com a vegetação da região do interflúvio Madeira-Purus (AM-RO). Acta Amaz. 2009; 39: 81–90.

45. Bispo PDC, Valeriano MDM, Kuplich TM. Relação da vegetação de caatinga com a condição geomorfométrica local. Rev Bras Eng Agrícola e Ambient. 2010; 14: 523–530. https://doi.org/10.1590/S1415-43662010000500010

46. Bispo PDC, Valeriano MDM, Kuplich TM. Relação entre as variáveis morfométricas extraídas de dados SRTM (Shuttle Radar Topography Mission) e a vegetação do Parque Nacional de Brasília. Acta Bot Brasilica. 2010; 24: 96–103. https://doi.org/10.1590/S0102-33062010000100010

47. Muñoz V, Valeriano MM, Bispo P. Surveying the Topographic Height from SRTM Data for Canopy Mapping in Brazilian Pantanal. Geografia. 2013; 38: 139–155.

48. Guadagnin PMA, Trentin R, Alves FDS. Relação entre as variáveis Geomorfométricas e a Vegetação Florestal na Bacia Hidrográfica do Arroio Caverá—Oeste do RS. Geogr Dep Univ Sao Paulo. 2015; 29: 246. https://doi.org/10.11606/rdg.v29i0.102131

49. Guitet S, Pélissier R, Brunaux O, Jaouen G, Sabatier D. Geomorphological landscape features explain floristic patterns in French Guiana rainforest. Biodivers Conserv. 2015; 24: 1215–1237. https://doi.org/10.1007/s10531-014-0854-8

50. Bispo PC, Santos JR, Valeriano MM, Touzi R, Seifert FM. Integration of polarimetric PALSAR attributes and local geomorphometric variables derived from SRTM for forest biomass modeling in central Amazonia. Can J Remote Sens. 2014; 40: 26–42. https://doi.org/10.1080/07038992.2014.913477

51. Lennon JJ, Koleff P, Greenwood JJD, Gaston KJ. Contribution of rarity and commonness to patterns of species richness. Ecol Lett. 2004; 7: 81–87. https://doi.org/10.1046/j.1461-0248.2004.00548.x

52. Legendre P, Gallagher ED. Ecologically meaningful transformations for ordination of species data. Oecologia. 2001; 129: 271–280. https://doi.org/10.1007/s004420100716 PMID: 28547606

53. Borcard D, Gillet F, Legendre P. Numerical Ecology with R. Springer. 2011.

54. Borcard D, Legendre P, Drapeau P. Partialling out the spatial component of ecological variation. Ecology. 1992; 73: 1045–1055. https://doi.org/10.2307/1940179

55. Legendre P, L L.. Numerical ecology, 2nd English Edition. Numerical ecology,. 1998.

56. Griffith DA, Peres-Neto PR. Spatial modeling in ecology: The flexibility of eigenfunction spatial analyses. Ecology. 2006; 87: 2603–2613. https://doi.org/10.1890/0012-9658(2006)87[2603:SMIETF]2.0.CO;2 PMID: 17089668

57. Borcard D, Legendre P. All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. Ecol Modell. 2002; 153: 51–68. https://doi.org/10.1016/S0304-3800(01)00501-4

58. Peres-Neto PR, Legendre P, Dray S, Borcard D. Variation partitioning of species data matrices: Estimation and comparison of fractions. Ecology. 2006; 87: 2614–2625. https://doi.org/10.1890/0012-9658 (2006)87[2614:VPOSDM]2.0.CO;2 PMID: 17089669

59. R Development Core Team R. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. 2011. p. 409.

60. Legendre P, De Cáceres M, Borcard D. Community surveys through space and time: Testing the space-time interaction in the absence of replication. Ecology. 2010; 91: 262–272. https://doi.org/10. 1890/09-0199.1 PMID: 20380215

61. Dray S, Legendre P, Blanchet G. packfor: Forward selection with permutation [R package version 0.0-8/r100]. Canoco. 2011. p. 46.

62. Oksanen J, Blanchet FG, Kindt R, Legendre P, O'Hara RB, Simpson GL, et al. vegan: Community Ecology Package. R package version. 2012. p. R package version 2.0–4.

63. Kunin WE, Shmida A. Plant Reproductive Traits as a Function of Local, Regional, and Global Abundance. Conserv Biol. 1997; 11: 183–192. https://doi.org/10.1046/j.1523-1739.1997.95469.x

64. Cornwell WK, Ackerly DD. A link between plant traits and abundance: evidence from coastal California woody plants. J Ecol. 2010; 98: 814–821. https://doi.org/10.1111/j.1365-2745.2010.01662.x

65. Espeland EK, Emam TM. The value of structuring rarity: The seven types and links to reproductive ecology. Biodivers Conserv. 2011; 20: 963–985. https://doi.org/10.1007/s10531-011-0007-2

66. Lennon JJ, Beale CM, Reid CL, Kent M, Pakeman RJ. Are richness patterns of common and rare species equally well explained by environmental variables? Ecography (Cop). 2011; 34: 529–539. https:// doi.org/10.1111/j.1600-0587.2010.06669.x

67. Heino J, Soininen J. Are common species sufficient in describing turnover in aquatic metacommunities along environmental and spatial gradients? Limnol Oceanogr. 2010; 55: 2397–2402. https://doi.org/10. 4319/lo.2010.55.6.2397

68. Tsang TPN, Bonebrake TC. Contrasting roles of environmental and spatial processes for common and rare urban butterfly species compositions. Landscape Ecology. 2016; 32: 47–57. https://doi.org/10. 1007/s10980-016-0427-1

69. Mazaris AD, Tzanopoulos J, Kallimanis AS, Matsinos YG, Sgardelis SP, Pantis JD. The contribution of common and rare species to plant species richness patterns: The effect of habitat type and size of sampling unit. Biodivers Conserv. 2008; 17: 3567–3577. https://doi.org/10.1007/s10531-008-9480-7

70. Luizao RCC, Luizao FJ, Paiva RQ, Monteiro TF, Sousa LS, Kruijt B, et al. Variation of carbon and nitrogen cycling processes along a topographic gradient in a central Amazonian forest. Glob Chang Biol. 2004; 10: 592–600. https://doi.org/10.1111/j.1529-8817.2003.00757.x

71. Balvanera P, Quijas S, Perez-Jimenez A. Distribution Patterns of Tropical Dry Forest Trees Along a Mesoscale Water Availability Gradient. Biotropica. 2011; 43: 414–422. https://doi.org/10.1111/j.1744-7429.2010.00712.x

72. Wang Q, Punchi-Manage R, Lu Z, Franklin SB, Wang Z, Li Y, et al. Effects of topography on structuring species assemblages in a subtropical forest. J Plant Ecol. 2016; 10: 440–449.

73. De Cáceres M, Legendre P, Valencia R, Cao M, Chang LW, Chuyong G, et al. The variation of tree beta diversity across a global network of forest plots. Glob Ecol Biogeogr. 2012; 21: 1191–1202. https://doi. org/10.1111/j.1466-8238.2012.00770.x

74. Murphy SJ, Audino LD, Whitacre J, Eck JL, Wenzel JW, Queenborough SA, et al. Species associations structured by environment and land-use history promote beta-diversity in a temperate forest. Ecology. 2015; 96: 705–715. PMID: 26236867

75. Myers JA, Chase JM, Jiménez I, Jørgensen PM, Araujo-Murakami A, Paniagua-Zambrana N, et al. Beta-diversity in temperate and tropical forests reflects dissimilar mechanisms of community assembly. Ecol Lett. 2013; 16: 151–157. https://doi.org/10.1111/ele.12021 PMID: 23113954