

1 **Drivers of metacommunity structure diverge for**
2 **common and rare Amazonian tree species**

3 Polyanna da Conceição Bispo^{1,2*}, Heiko Balzter^{1,2}, Yadvinder Malhi³, J.W. Ferry Slik⁴,
4 João Roberto dos Santos⁵, Camilo Daleles Rennó⁶, Fernando D. Espírito-Santo⁷, Luiz E.
5 O. C. Aragão^{5,8}, Arimatéa C. Ximenes⁹, Pitágoras da Conceição Bispo¹⁰

6 ¹Leicester Institute for Space and Earth Observation, Centre for Landscape and Climate
7 Research, Department of Geography, University of Leicester, Leicester, UK

8 ²NERC National Centre for Earth Observation at the University of Leicester, Leicester,
9 UK

10 ³Environmental Change Institute, School of Geography and the Environment, University
11 of Oxford, Oxford, UK

12 ⁴Faculty of Science, Universiti Brunei Darusallam, Gadong, Brunei

13 ⁵Divisão de Sensoriamento Remoto, Instituto Nacional de Pesquisas Espaciais, São José
14 dos Campos, São Paulo, Brazil

15 ⁶Divisão de Processamento de Imagens, Instituto Nacional de Pesquisas Espaciais, São
16 José dos Campos, São Paulo, Brazil

17 ⁷Lancaster Environment Centre), Lancaster University, UK

18 ⁸College of Life and Environmental Sciences, University of Exeter, UK

19 ⁹Université Libre de Bruxelles, le Département de Biologie des Organismes, Brussels,
20 Belgium

21 ¹⁰Department of Biological Sciences, Faculty of Sciences and Languages of Assis, State
22 University of São Paulo (UNESP), Assis, São Paulo, Brazil

23 * polybispo@gmail.com

1 **Abstract**

2 We analysed the flora of 46 forest inventory plots (25m x 100m) in old growth
3 forests from Amazon region in order to identify the role of environmental (topography)
4 and spatial variables (obtained using Principal Coordinates of Neighborhood Matrix,
5 PCNM) for rare and common species. For the analyses, we used Multiple Partial
6 Regression in order to partition the specific effects of the topography and spatial variables
7 on the univariate data (standardised richness, total abundance and total biomass) and
8 partial RDA (Redundancy Analysis) for partitioning these effects on composition
9 (multivariate data) based on incidence data, abundance and biomass. Our results showed
10 that different attributes (richness, abundance, biomass and composition based in
11 incidence, abundance and biomass) used to study this metacommunity responded
12 differently to environmental and spatial processes. Considering standardised richness,
13 total abundance (univariate) and composition based on biomass, the results for common
14 species differ from those obtained for all species. On the other hand, for total biomass
15 (univariate) and for compositions based on incidence and abundance, there was
16 correspondence between the data obtained for the total community and for common
17 species. Except for total abundance, the environmental and spatial factors measured were
18 insufficient to explain the attributes of communities of rare species. These results indicate
19 that predicting the attributes of rare species tree communities based on environmental and
20 spatial factors is a huge challenge. Our data show that niche-related processes are
21 important; however, the spatial component is also important, possibly because of the mass
22 effect. As the spatial component was relevant for several community attributes, our results
23 demonstrate the importance of a metacommunities approach when attempting to
24 understand the main ecological processes underlying the diversity of forest communities
25 in the Amazon region.

26 Keywords: tropical forest, diversity, rarity, topography

27

28

1 Introduction

2 Throughout the history of ecology, researchers have sought to understand the
3 effects of environmental and spatial process on biodiversity. This quest has led to the
4 collection of a large number of datasets and the formulation of competing theories, such
5 as the niche and the neutral theories. Niche theory has been one of the most important
6 theoretical approaches in ecology and assumes that the combination of resources and
7 conditions determines the local characteristics of a community [1]. In this context, it is
8 expected that communities structured by niche-related processes have similar values of
9 local attributes (e.g. species richness, species abundance, biomass and composition) when
10 in similar habitat patches. In contrast, the neutral theory is based on functional
11 equivalence between species and considers dispersion and demographic stochasticity as
12 central phenomena [2,3]. According to the neutral theory, it is expected that similar values
13 for local attributes of a community in a particular patch can be determined by influence
14 from communities of nearby patches, stressing the importance of spatial variables.

15 Megadiverse tropical forests can be important model ecosystems for
16 understanding the relative roles of environmental and spatial variables on communities.
17 The endeavour to understand these roles may help clarifying the importance of different
18 processes in maintaining species diversity of these forests. Concerning the niche theory,
19 variables representative of topography are likely to play an important role in local
20 environmental conditions which may determine species diversity, functional and
21 structural attributes and composition of tree communities [4–7]. One advantage of using
22 topography in vegetation studies is that it can be assessed at large scales by remote
23 sensing. Moreover, it is considered a good surrogate to several important variables for
24 vegetation structure that would be difficult to measure on a larger scale, including nutrient
25 availability, soil moisture and texture, insolation, etc. [6].

26 Despite their known importance, the influence of spatial process has long been
27 ignored in ecological studies, and until the 1990s, the main focus of ecology was the study
28 of niche-related processes [8]. Since the 1990s, the spatial process has been strongly
29 embedded in ecology [9]. Since then new methods and analytical strategies have been
30 proposed and empirical and theoretical data have been collected, generating a new body
31 of knowledge about the structuring factors of communities [10–12]. This knowledge,

1 along with other information (functional, phylogenetic, etc.), has provided a much more
2 detailed understanding of the processes involved in community structure [13–15],
3 supporting the search for better conservation and biodiversity monitoring strategies [16].
4 Thus, understanding the role of spatial factors in community structure at different
5 landscape scales is essential, especially given the current scenario of rapid biodiversity
6 loss due to habitat degradation and fragmentation [17].

7 Metacommunity theory recognises that communities are not isolated entities, but
8 they are connected by movements of individuals of different species [18,19]. Thus, the
9 metacommunity can be defined by groups of communities connected to each other by
10 dispersal of individuals. Metacommunity theory has benefited from discussions regarding
11 niche and neutral theory, and both have helped to understand the role of environmental
12 and spatial variables on the diversity and composition of the biota [18,19]. According to
13 the theory, the group of species occurring in a community is determined both by a
14 combination of local factors (interaction between species, interaction of species with local
15 factors) and by the ability of the species to reach that community (by dispersion) [18,20].
16 In metacommunities structured by the principles of niche theory, it is expected that the
17 environmental component plays a more important role. On the other hand, in
18 metacommunities subject to the principles of neutral theory, it is expected that the spatial
19 component plays a more important role. The metacommunity theory encompasses four
20 main models (species sorting, mass effect, patch dynamics and neutral) which represent
21 points of a continuum formed by different combinations of environmental (niche) and
22 dispersal (spatial) processes [18,19,21,22].

23 A remarkable feature of most communities is the presence of a few common
24 species and many rare species [23,24]. Rare and common species can respond to
25 ecological processes differentially [25], depending on the features of organisms (e.g.
26 competition and dispersion capacities) and spatial temporal dynamics. For example,
27 based on the niche theory, Tokeshi [8] proposed the composite niche model, arguing that
28 more than one process may be acting on the community. According to this model, rare
29 species should fit a random assortment model, while common species should fit any
30 model of niche apportionment. On the other hand, Siqueira et al. [24] studied
31 metacommunities of aquatic macroinvertebrates and showed that common and rare
32 species responded similarly and both were mainly structured by niche processes.

1 The analysis of different attributes of communities while considering groups of
2 common and rare species allows testing some hypotheses about the processes involved in
3 the structuring of metacommunities [24,26]. Empirical studies have shown that most of
4 the metacommunities are structured principally by niche processes [21,24]. Our first
5 hypothesis is that the niche processes are more important for structuring the
6 metacommunity studied [24], at least for the common species, which we expect to be
7 those most affected by competition [27]. In the case of rare species, as they can be more
8 affected by ecological drift [27], we expect spatial factors to be more relevant. Taking
9 into account that habitat generalist and habitat specialist species differ in terms of
10 population dynamics, we also propose an alternative hypothesis [24,28]. While generalist
11 species occupy habitats with a wide environmental variation, specialist species
12 preferentially occupy habitats with specific environmental characteristics, which are
13 generally rare in the landscape [28,29]. In this context, assuming that common species
14 are habitat generalists and rare species are habitat specialists, our alternative hypothesis
15 is that spatial factors are more important for common species, while environmental factors
16 are crucial for rare species [28].

17 To test these hypotheses and to identify the role of environmental (topography)
18 and spatial variables for rare and common species, our aim was to analyse the flora of 46
19 forest inventory plots in the old growth forests of the eastern Amazon region. These
20 analyses were based on vegetation data collected in the field and topographic variables
21 obtained by remote sensing data.

22 **Material and Methods**

23 **Study area**

24 This study was conducted in the Tapajós National Forest (TNF). The TNF is a
25 large protected area of about 545,000 ha, located in Amazon biome, west of Pará State,
26 Brazil (Fig 1). This area has average annual temperature of 25.5°C and average annual
27 rainfall of 1,820 mm. The local topography ranges from flat to strongly undulating terrain.
28 Predominant soil types in the area are dystrophic oxisol (US classification) or dystrophic
29 yellow latosol (Brazil classification)) and red-yellow podzol. Vegetation is mainly
30 ombrophilous dense forest and ombrophilous open forest [30].

1 **Fig 1.** Study area in the Tapajós National Forest (FNT), Pará State, Brazil with detail of
 2 the five geomorphometric variables (elevation, slope, HAND, profile curvature and plan
 3 curvature) of four areas where the 46 plots are distributed.

4 We sampled 46 forest inventory plots (Fig 1) of 25 x 100 m in TNF. Our plots
 5 representing a sampling area of 11.5 ha. We installed the plots in different vegetal
 6 typologies and topographies [5] using the phyto-ecologic classes from RADAM-BRASIL
 7 project [31]. Our plots encompassed different floristic and geomorphological
 8 characteristics [31]. We sampled and identify all individual trees with diameter at breast
 9 height (dbh) ≥ 10 cm. The abundance was represented by number of individuals and
 10 aboveground biomass (hereafter referred just as biomass) was calculated by the
 11 allometric equation [32], considering the measurements of DBH (diameter at breast
 12 height) and TH (total height).

13
$$\text{Biomass} = 0.044 * ((\text{DBH}^2) * \text{TH})^{0.9719}$$

14 **Topographic data derived from SRTM**

15 We used geomorphometric attributes (elevation, slope, profile curvature and plane
 16 curvature) from Brazilian Geomorphometrics Database (TOPODATA) [33]. The
 17 TOPODATA is based on SRTM (Shuttle Radar Topography Mission-version 1, NASA,
 18 2006) and has different neighbourhood operations to calculate geomorphometric
 19 variables [33]. The TOPODATA is free and the layers are easily accessible
 20 (<http://www.dsr.inpe.br/topodata/acesso.php>). We also used the vertical distance to the
 21 nearest drainage or HAND (height above the nearest drainage). Hand was derivate from
 22 SRTM and describes the vertical distance of each point regarding the nearest drainage
 23 channel detected by remote sensing [34]. All data used in this study has 30 m of spatial
 24 resolution (Table 1).

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

Topographic variables	Description
Elevation (<i>h</i>)	Terrain altitude. This is related to the altitude distribution of soil and climate, determining different landscape vegetation patterns.
Slope (<i>G</i>)	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 (<i>kv</i>)	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 (<i>kh</i>)	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 (<i>HAND</i>)	Describes the vertical distance of each point regarding the nearest drainage. The lower the <i>HAND</i> value, the closer the water table is to the surface and, consequently, the more easily the area could be flooded.

1 Topographic variables obtained on the basis of SRTM have been used to explain
2 or predict the properties of vegetation [35, 36]. These studies have helped understanding
3 the effects of topography on the distribution of different types of vegetation [37–41],
4 floristic composition [5,42] and forest structure [7,43], particularly in tropical areas.

5 **Data analysis**

6 We defined common and rare species using the criterion of the inflection point of
7 the curve of species abundance (or species biomass) [24]. We defined the inflection point
8 visually; species left of this point were considered as common and those ones to the right
9 as rare [24]. As matrices with different amounts of information can affect the results, we
10 made comparisons considering the same information content. First of all, the information
11 content of the matrices of common and rare species was calculated considering the
12 binomial variance of the incidence matrix, $\sum p_i(1-p_i)$, where p_i is the proportion of plots
13 occupied by i^{th} species [24,44]. As the matrix of rare species had a higher information
14 content, we removed rare species, following species rank, until this matrix had the same
15 information content than that of common species.

16 After defining the common and rare species with the same information content,
17 we performed data analyses considering univariate and multivariate community
18 attributes. The univariate attributes were standardised richness (residuals of regression
19 between abundance and richness), total abundance (sum of the abundance of all species

1 per plot) and total biomass (total biomass of all species per plot). The multivariate
2 attributes (species x plots) were represented by three different matrices of composition
3 (1. composition based on incidence; 2. composition based on abundance, and 3.
4 composition based on biomass)

5 The richness of species is usually positively correlated with abundance. As
6 abundance can explain part of the variation in richness, without due caution, we can
7 erroneously conclude that similar factors are important in explaining both community
8 attributes. Thus, disentangling the richness from abundance is necessary to understand
9 the real effect of topography on richness (free of abundance). For this reason, we used the
10 residuals of the regression between abundance and richness as standardised measure of
11 richness (standardised richness). In this case, the residuals indicate the part of variation
12 in richness that cannot be explained by abundance, in other words, richness free of
13 abundance.

14 In general, the community matrix based on abundance has many zeros, which is a
15 problem for multivariate analysis based on Euclidian distances, such as Principal
16 Components Analysis (PCA) and Redundancy Analysis (RDA). A strategy to minimise
17 this problem is using the Hellinger transformation [45]. Thus, in the case of composition
18 based on abundance, prior to analyses, we transformed the data matrix using the Hellinger
19 method [46].

20 For the analyses, we used a multiple partial regression in order to partition the
21 specific effects of the topography and spatial variables on the univariate response
22 variables (standardised richness, abundance and total biomass) and partial RDA for
23 partitioning these effects on multivariate response matrices represented by incidence data,
24 abundance and biomass [10,47,48]. The RDA is a direct gradient analysis based on
25 multiple regression that deals with the variation in a multivariate response matrix (in our
26 case, composition based on incidence, abundance and biomass) and one or more matrices
27 of explanatory variables (in our case, topographic and spatial variables) [48].

28 Spatial variables were obtained using the PCNM method (Principal Coordinates
29 of Neighborhood Matrix) [48]. The PCNM is based on Coordinates Principal Analysis
30 obtained from a geographic distance matrix. The eigenvectors (axes) obtained from this
31 analysis are called PCNMs, are uncorrelated and represent different spatial patterns, from

1 coarse (axes with higher eigenvalues) to more refined (axes with smaller eigenvalues)
2 [48,49]. Thus, in this paper, we extracted the spatial variables (PCNMs) from a Euclidean
3 distance matrix between plots, which were represented by eigenvectors with positive
4 eigenvalues and with spatial autocorrelation according to Moran's I index [50]. For the
5 analyses, we selected variables using the forward selection method to evaluate only the
6 environmental and spatial variables that were more related to the studied metacommunity.

7 We assessed the following fractions: environmental (topography) component
8 independent of the space (a), environmental component inseparable of the spatial
9 component (b), spatial component independent of the environment (c) and component not
10 explained (d). As the coefficient of determination (R^2) is influenced by the sample size
11 and number of predictor variables, we used the adjusted R^2 to obtain the importance of
12 each assessed fraction [51]. We performed the analyses in the computing environment R
13 version 2.13 [52], associated with the PCNM package [53] to obtain the spatial variables,
14 Packfor [54] for variable selection and Vegan [55] for multiple regression and RDA.

15 **Results**

16 The results revealed that, when considering the inflection points of the abundance
17 curves, 22 species are considered common and 208 rare (Fig 2a). The 93 rarest species
18 have the same information content as the 22 most common species. When the inflection
19 of the curves of species biomass is considered, 35 species are considered common and
20 195 rare (Fig 2b). In this case, the 94 rarest species have the same information content as
21 the 35 most common species.

22 **Fig 2.** Rank of abundance (a) and rank of biomass (b) of Amazonian tree species of a
23 metacommunity of Tapajós National Forest, Pará State, Brazil. * indicate the rare species
24 with the same information content than common species.

25 **Standardized richness, abundance and total biomass**

26 Total standardised richness (without the abundance effect) and standardised
27 richness of rare species cannot be explained by any of the measured factors (topographic
28 or spatial variables) (Table 2). Nevertheless, standardised richness of common species is
29 significantly explained by topography (Table 2). The data also show that total abundance

1 is explained by spatial variables. Moreover, the abundance of common species is
 2 significantly explained both by topographic and spatial variables, whereas rare species
 3 are explained by topography (Table 2). The data show that in general, the topography
 4 explains the total tree biomass, both for total and common species, but not for rare species
 5 (Table 2).

6 **Composition (incidence, abundance and biomass)**

7 Our data reveal that topographic and spatial variables explained relevant
 8 proportions of the variability, both when the analysis is based on an incidence matrix and
 9 on an abundance matrix (Table 2), and both for total and common species (Table 2).
 10 Topographic and spatial variables do not explain the variation of rare species (Table 2).
 11 When the analysis is based on biomass, its variability is explained significantly by
 12 topographic and spatial variables for total species, only by topography for common
 13 species and by no factor for rare species (Table 2).

14 **Table 2.** Results of Partial Multiple Regression and Partial Redundancy Analysis with
 15 the coefficient of determination (R^2) for common and rare species.

	Topography (%)	Shared (%)	Space (%)	Not explained (%)
<i>Std Richness</i>				
Total	-	-	-	100.00
Common (1-22)	10.4*	-	-	89.6
Rare (137-230)	-	-	-	100.00
<i>Abundance (Total)</i>				
Total	-	-	25.9**	74.1
Common (1-22)	11.4**	-	10.3**	78.3
Rare (137-230)	-	-	26.3**	73.7
<i>Biomass (Total)</i>				
Total	27.4***	15.1	-	57.5
Common (1-35)	14.2**	12.5	-	73.3
Rare (136-230)	-	-	-	100.00
<i>C. Incidence</i>				
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
<i>C. Abundance</i>				
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
<i>C. Biomass</i>				
Total (1-230)	3.1**	3.2	1.6	92.1
Common (1-35)	4.0**	4.2	1.9ns	89.9
Rare (136-230)	-	-	-	100.00

1 Topography refers to the effects of geomorphometric variables without spatial factors;
2 shared refers to the effects of common variation between topographic and spatial factors;
3 and space refers to the spatial effects (PCNMs) without topography. Common and rare
4 species are delimited based on the inflection point species x abundance curve (in the case
5 of abundance), or species x biomass curve (in the case of biomass). Numbers in
6 parentheses refer to the rank position of the species. Std Richness (standardised richness,
7 residuals of regression between abundance and richness); composition based on incidence
8 (C. Incidence); composition based on abundance (C. Abundance) and composition based
9 on biomass (C. Biomass). *p < 0.05; **p < 0.01; ***p < 0.001; ns non-significant.

10 Discussion

11 Our results show that different attributes (richness, abundance, biomass and
12 composition based on incidence, abundance and biomass) used to study this
13 metacommunity respond differently to environmental and spatial processes. This
14 indicates that the studied attributes represent different dimensions of the community and
15 that their analysis may provide insights about the main processes that structure the studied
16 metacommunity. Common and rare species differ in terms of biological traits [56–58] and
17 how they relate to environmental factors [28,29,59]. Our expectation was that the role of
18 environment and spatial variables differed between common and rare species. Our results
19 confirm our expectation, suggesting that common and rare species are subject to different
20 combinations of environmental and spatial variables and/or different ecological
21 processes.

22 The communities are composed of few common species and many rare species.
23 Due to the greater number of individuals, common species interact strongly with the
24 various components of the system. A common question is therefore whether the common
25 species are sufficient to describe the attributes (e.g. richness, abundance and composition)
26 of the whole community [44,59,60]. If this is the case, studies on communities could focus
27 on common species, which are more easily sampled. Most studies go in this direction,
28 since, in general, the results found for all species are equivalent to those found using only
29 common species [59–61]. Our results only partially confirm this expectation and show
30 that this depends on the analysed attribute. For example, for standardised richness, total

1 abundance (univariate) and for composition based on biomass, the results for common
2 species differ from those obtained for all species. However, for total biomass (univariate)
3 and for compositions based on incidence and abundance, there is a correspondence
4 between the data obtained for the total community and for common species. Thus, our
5 data suggest that for these last attributes, it is possible to draw appropriate conclusions
6 for a whole community based on common species.

7 Considering the standardised richness, our results are in disagreement with the
8 results of studies at different scales, which have shown that the richness patterns (total
9 community) can be predicted by richness of common species [43,62]. In our study,
10 topography only significantly explained the variability of common species, and neither
11 topographic nor spatial variables explained the variability of the total community or of
12 rare species. According to Lennon et al. [59], the richness of common species can be more
13 easily explained by simple environmental gradients when compared to the richness of
14 rare species. As the richness of rare species can be associated with rare environments
15 [29,59], it is harder to predict it. Thus, the absence of the effect of the environment on the
16 richness of rare species may be due to the fact that a part of the environmental factors
17 important for these species is not registered, since these must be associated to uncommon
18 niches [29], determined mainly by environmental factors that are difficult to measure. In
19 this context, our results show that on the scale of our study, part of the variation of the
20 standardised richness of common species can be predicted by the environmental gradient
21 (in our case, topography), which does not happen with rare species.

22 Total abundance (univariate) and biomass (univariate) were explained by different
23 processes when the total community, common species and rare species were considered.
24 For total abundance, in all combinations (total community, common and rare species),
25 spatial variables were important in explaining variability. These results show that spatial
26 processes determine a relevant part of total abundance variation, suggesting that the
27 effects of mass are important when abundance is considered. For common species, in
28 addition to spatial processes, the environmental variables were also important. In the case
29 of total biomass of the whole community, only environmental factors were important.
30 Topography influences other extremely important variables, such as soil texture,
31 availability of nutrients and water [4,63,64]. This may explain the results found. Our
32 observations may have practical consequences. For example, we have sought ways to

1 predict and monitor the biomass at larger scales, and this relationship with topography
2 can be important since it influences other extremely important factors for the
3 accumulation and maintenance of biomass. Therefore, topography can be a surrogate of
4 several variables difficult to be measured in building predictive models that facilitate
5 monitoring of the biomass and carbon stocks in tropical forests.

6 Considering the composition data based on incidence, abundance and biomass,
7 one of our initial expectations was that the niche-related factors were the most relevant to
8 explain data variability and that the studied metacommunity would follow the model
9 species sorting (SS) [21]. Our results showed that both topography and spatial variables
10 were relevant to explain variability, both for total and for common species (except for
11 biomass, which was explained only by environmental processes). These results suggest
12 that within the range studied, data fit the model species sorting (SS) + mass effect (ME).
13 The SS and SS + ME have been the most frequently adjusted models to empirical data.
14 For example, Cottenie [21] studied 158 metacommunities and found that 44% of them fit
15 the SS model and 29% the SS + ME model. These patterns (SS or SS + ME) have been
16 confirmed by most studies since Cottenie [21]. It is important to point out that the
17 inclusion of other relevant environmental variables could increase the percentage of
18 explanation of the environment and reduce the importance of spatial variables, whose
19 effects may be a reflection of both mass effects and dispersion difficulties, as well as
20 responses to no measured environmental factors [11]. Therefore, many SS + ME results
21 observed in the literature may indeed be SS, since many important variables may not have
22 been measured, suggesting that niche processes are highly relevant in structuring
23 metacommunities.

24 The data presented here show that the ecological processes underlying
25 composition based on incidence, abundance and biomass differ between common and
26 rare species, agreeing with Tsang & Bonebrake [61], which studied the composition of
27 butterflies. On the other hand, this result disagrees with other data for different organisms
28 (e.g. aquatic macroinvertebrates [24,60] and macrophytes [26]), which show that
29 common and rare species are governed by the same processes, in these cases by processes
30 related to the niche. Specifically for vegetation, Wang et al. [65], studying the effects of
31 topography on the species composition of a subtropical forest, also verified that the
32 environmental factors are important determinants of the variation of the composition of

1 common and rare species. On the other hand, this relation was much weaker for rare
2 species. Our initial hypothesis was that the composition of rare species could be explained
3 by the environment or by spatial factors. Our results refute this hypothesis, since none of
4 the components (environmental or spatial) explained the variation of the composition of
5 rare species. This may be a reflection of stochastic factors and of the non-inclusion of
6 specific variables important for rare species.

7 In this study, we found that much of the variability was neither explained by
8 environment (represented by the topography) nor by space. This is a relatively common
9 result in studies of metacommunities [10,12]. Two main factors can help explain this
10 common result: 1) there are a lot of environmental factors in tropical forests that affect
11 the biota, and often only one portion of them is measured, in our case, notably factors
12 related to topography; 2) tropical forests have a large number of biotic interactions which,
13 despite having the potential to affect the biota structure, are impossible to be measured to
14 capture their complexity. Thus, the proportion of unexplained variability is probably due
15 to stochastic and unmeasured factors. Baldeck et al. [11] showed that in addition to
16 topography, the inclusion of variables such as nutrients can improve a model's
17 explanatory power. Thus, despite the topography being a good substitute for other
18 variables that are difficult to measure, the inclusion of additional relevant variables should
19 decrease the proportion of unexplained variability by the model.

20 Our results revealed that for the common species, the variation of the attributes
21 measured was explained by topography and/or spatial variables. On the other hand, except
22 for total abundance, none of the factors measured explained the measured attributes of
23 the rare species. These results indicate that predicting the attributes of rare species tree
24 communities from environmental and spatial variables is a huge challenge. Our data show
25 that niche-related processes are important; however, the spatial component, possibly
26 because of the mass effect, was also important. As the spatial component was important
27 for several community attributes, our results demonstrate the importance of a
28 metacommunity approach when attempting to understand the main ecological processes
29 underlying the diversity of forest communities in the Amazon region.

30

1 **Authorization for the field work**

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

8 **Funding**

9 PCB was supported by European Union’s Horizon 2020 research and innovation
10 programme under the Marie Skłodowska-Curie grant agreement N° 660020 (‘MF-
11 RADAR’). HB was supported by Royal Society Wolfson Research Merit Award
12 (2011/R3) and by the Natural Environment Research Council’s National Centre for Earth
13 Observation. PITCB, JRS and LEOCA were supported by CNPq National Council for
14 Scientific and Technological Development) productivity fellowships (grants
15 305275/2014-3, 303228/2013-0 and 305054/2016-3, respectively). FES was supported by
16 Natural Environment Research Council (NERC) grants (‘BIO-RED’ NE/N012542/1 and
17 ‘AFIRE’ NE/P004512/1) and Newton Fund (‘The UK Academies/FAPESP Proc. N°:
18 2015/50392-8 Fellowship and Research Mobility’).

19 **Acknowledgments**

20 The authors acknowledge Brazil’s National Institute for Space Research (INPE)
21 and INPA-NAPPA/LBA/Santarém for logistic support during the field work, and
22 ICMBio/MMA for provide the authorization (SISBIO n. 20591-1) to carry out the field
23 work at Tapajós National Forest. The authors also thank Erly Pedroso (in memoriam) for
24 assisting during the field campaign and Dr. Márcio de Morisson Valeriano (INPE) and
25 his team for having developed and made available the TOPODATA (Brazilian
26 Geomorphometrics Database).

27

1 **References**

- 2 1. Hutchinson GE. Population studies - animal ecology and demography - concluding
3 remarks. *Cold Spring Harbor Symposia on Quantitative Biology*. 1957; 22: 415–
4 427.
- 5 2. Hubbell SP. *The Unified Neutral Theory of Biodiversity and Biogeography*.
6 *Monographs in Population Biology* 32. Princeton University Press 2001, 375 pp.;
7 ISBN 0-691-02128-7
- 8 3. Hubbell SP. Neutral theory in community ecology and the hypothesis of functional
9 equivalence. *Functional Ecology*. 2005; 19: 166–172. doi:10.1111/j.0269-
10 8463.2005.00965.x
- 11 4. Castilho C V, Magnusson WE, Araújo RNO, Luizão RCC, Luizão FJ, Lima AP,
12 et al. Variation in aboveground tree live biomass in a central Amazonian Forest:
13 Effects of soil and topography. *Forest Ecology and Management*. 2006; 234: 85–
14 96. doi:10.1016/j.foreco.2006.06.024
- 15 5. Bispo PC, Valeriano MM, Santos JR. Effects of the geomorphometric
16 characteristics of the local terrain on floristic composition in the central Brazilian
17 Amazon. *Austral Ecology*. 2012; 37: 491–499. doi:10.1111/j.1442-
18 9993.2011.02310.x
- 19 6. Liu J, Yunhong T, Slik JWF. Topography related habitat associations of tree
20 species traits, composition and diversity in a Chinese tropical forest. *Forest*
21 *Ecology and Management*. 2014; 330: 75–81. doi:10.1016/j.foreco.2014.06.045
- 22 7. Bispo PC, Santos JR, Valeriano MM, Graça PMLA, Balzter H, França H, et al.
23 Predictive models of primary tropical forest structure from geomorphometric
24 variables based on SRTM in the Tapajós region, Brazilian Amazon. *PLoS One*.
25 2016;11. doi:10.1371/journal.pone.0152009
- 26 8. Tokeshi M. Niche Apportionment or Random Assortment: Species Abundance
27 Patterns Revisited. *Journal of Animal Ecology* 1990; 59: 1129–1146.
28 doi:10.2307/5036

- 1 9. Legendre P. Spatial autocorrelation: trouble or new paradigm? *Ecology*. 1993. pp.
2 1659–1673. doi:10.2307/1939924
- 3 10. Beisner BE, Peres-Neto PR, Lindström ES, Barnett A, Longhi ML. The role of
4 environmental and spatial processes in structuring lake communities from bacteria
5 to fish. *Ecology*. 2006; 87: 2985–2991. doi:10.1890/0012-
6 9658(2006)87[2985:TROEAS]2.0.CO;2
- 7 11. Baldeck CA, Harms KE, Yavitt JB, John R, Turner BL, Navarrete H, et al. Soil
8 resources and topography shape local tree community structure in tropical forests
9 Soil resources and topography shape local tree community structure in tropical
10 forests. *Proceedings of the Royal Society London B Biol Sci*. 2013; 280:
11 20122532. doi:10.1098/rspb.2012.2532
- 12 12. Branco CCZ, Bispo PC, Peres CK, Tonetto AF, Branco LHZ. The roles of
13 environmental conditions and spatial factors in controlling stream macroalgal
14 communities. *Hydrobiologia*. 2014; 732: 123–132. doi:10.1007/s10750-014-1852-
15 2
- 16 13. Leibold MA, Economo EP, Peres-Neto P. Metacommunity phylogenetics:
17 Separating the roles of environmental filters and historical biogeography. *Ecology*
18 *Letters*. 2010; 13: 1290–1299. doi:10.1111/j.1461-0248.2010.01523.x
- 19 14. Spasojevic MJ, Copeland S, Suding KN. Using functional diversity patterns to
20 explore metacommunity dynamics: A framework for understanding local and
21 regional influences on community structure. *Ecography (Cop)*. 2014; 37: 939–949.
22 doi:10.1111/ecog.00711
- 23 15. Saito VS, Soinenen J, Fonseca-Gessner AA, Siqueira T. Dispersal traits drive the
24 phylogenetic distance decay of similarity in Neotropical stream metacommunities.
25 *Journal of Biogeography*. 2015; 42: 2101–2111. doi:10.1111/jbi.12577
- 26 16. Siqueira T, Bini LM, Roque FO, Cottenie K. A metacommunity framework for
27 enhancing the effectiveness of biological monitoring strategies. *PLoS One*. 2012.
28 doi:10.1371/journal.pone.0043626

- 1 17. Barlow J, Lennox GD, Ferreira J, Berenguer E, Lees AC, Nally RM, Thomson JR,
2 Ferraz SFDB.; Louzada J, Oliveira VHF, et al. Anthropogenic disturbance in
3 tropical forests can double biodiversity loss from deforestation. *Nature*. 2016; 535:
4 144–147. doi:10.1038/nature18326.
- 5 18. Leibold MA, Holyoak M, Mouquet N, Amarasekare P, Chase JM, Hoopes MF, et
6 al. The metacommunity concept: A framework for multi-scale community
7 ecology. *Ecology Letters*. 2004. pp. 601–613. doi:10.1111/j.1461-
8 0248.2004.00608.x
- 9 19. Logue JB, Mouquet N, Peter H, Hillebrand H. Empirical approaches to
10 metacommunities: a review and comparison with theory. *Trends in Ecology &
11 Evolution*. 2011; 26: 482–91. doi:10.1016/j.tree.2011.04.009
- 12 20. Santos AS, Saraiva DD, Müller SC, Overbeck GE. Interactive effects of
13 environmental filtering predict beta-diversity patterns in a subtropical forest
14 metacommunity. *Perspectives in Plant Ecology, Evolution and Systematics*.
15 Elsevier GmbH.; 2015; 17: 96–106. doi:10.1016/j.ppees.2015.01.002
- 16 21. Cottenie K. Integrating environmental and spatial processes in ecological
17 community dynamics. *Ecology Letters*. 2005; 8: 1175–1182. doi:10.1111/j.1461-
18 0248.2005.00820.x
- 19 22. Alexander HM, Foster BL, Ballantyne F, Collins CD, Antonovics J, Holt RD.
20 Metapopulations and metacommunities: Combining spatial and temporal
21 perspectives in plant ecology. *Journal of Ecology*. 2012; 100: 88–103.
22 doi:10.1111/j.1365-2745.2011.01917.x
- 23 23. Spitale D. A Comparative Study of Common and Rare Species in Spring Habitats.
24 *Ecoscience*. 2012; 19: 80–88. doi:10.2980/19-1-3435
- 25 24. Siqueira T, Bini LM, Roque FO, Marques Couceiro SR, Trivinho-Strixino S,
26 Cottenie K. Common and rare species respond to similar niche processes in
27 macroinvertebrate metacommunities. *Ecography (Cop)*. 2012; 35: 183–192.
28 doi:10.1111/j.1600-0587.2011.06875.x

- 1 25. László Z, Rákosy L, Tóthmérész B. Landscape and local variables benefit rare
2 species and common ones differently. *Journal of Insect Conservation*. 2014; 18:
3 1203–1213. doi:10.1007/s10841-014-9734-5
- 4 26. Alahuhta J, Johnson LB, Olker J, Heino J. Species sorting determines variation in
5 the community composition of common and rare macrophytes at various spatial
6 extents. *Ecological Complexity*. 2014; 20: 61–68.
7 doi:10.1016/j.ecocom.2014.08.003
- 8 27. Chase JM, Amarasekare P, Cottenie K, Gonzalez A, Holt RD, Holyoak M, et al.
9 Competing theories for competitive metacommunities. In Holyoak M, Leibold
10 MA, Holt RD, eds. *Metacommunities spatial dynamics and ecological*
11 *communities*. 2005. pp. 335–354. Chicago University Press, Chicago.
- 12 28. Pandit SN, Kolasa J, Cottenie K. Contrasts between habitat generalists and
13 specialists: An empirical extension to the basic metacommunity framework.
14 *Ecology*. 2009; 90: 2253–2262. doi:10.1890/08-0851.1
- 15 29. Markham J. Rare species occupy uncommon niches. *Scientific Reports*. 2014; 4:
16 6012. doi:10.1038/srep06012
- 17 30. Veloso HP, Rangel Filho ALR, Lima JCA. *Classificação da Vegetação Brasileira*
18 *Adaptada a um Sistema Universal*. Rio de Janeiro, IBGE, Departamento de
19 Recursos Naturais e Estudos Ambientais. 1991. doi:ISBN 85-240-0384-7
- 20 31. RADAMBRASIL Departamento Nacional de Produção Mineral. No TitleFolha
21 AS.21- Santarém. Geologia, geomorfologia, pedologia, vegetação e uso potencial
22 da terra. Rio de Janeiro: DNPM (Levantamento dos Recursos Naturais v. 10).
23 1976.
- 24 32. Brown S, Gillespie AJR, LUGO A. E. Biomass estimation methods for tropical forest
25 with applications to forest inventory data. *Forest Science*, 1989; 35: (4): 881-902.
- 26 33. Valeriano MM, Rossetti DF. Topodata: Brazilian full coverage refinement of
27 SRTM data. *Applied Geography*. 2012; 32: 300–309.
28 doi:10.1016/j.apgeog.2011.05.004

- 1 34. Rennó CD, Nobre AD, Cuartas LA, Soares JV, Hodnett MG, Tomasella J, et al.
2 HAND, a new terrain descriptor using SRTM-DEM: Mapping terra-firme
3 rainforest environments in Amazonia. *Remote Sensing of Environment*. 2008;
4 112: 3469–3481. doi:10.1016/j.rse.2008.03.018
- 5 35. Kellndorfer J, Walker W, Pierce L, Dobson C, Fites JA, Hunsaker C, et al.
6 Vegetation height estimation from Shuttle Radar Topography Mission and
7 National Elevation Datasets. *Remote Sensing of Environment*. 2004; 93: 339–358.
8 doi:10.1016/j.rse.2004.07.017
- 9 36. Simard M, Rivera-Monroy VH, Mancera-Pineda JE, Castañeda-Moya E, Twilley
10 RR. A systematic method for 3D mapping of mangrove forests based on Shuttle
11 Radar Topography Mission elevation data, ICESat/GLAS waveforms and field
12 data: Application to Ciénaga Grande de Santa Marta, Colombia. *Remote Sensing
13 of Environment*. 2008; 112: 2131–2144. doi:10.1016/j.rse.2007.10.012
- 14 37. Bispo, P C, Valeriano, M M, Kuplich TM. Variáveis geomorfométricas locais e
15 sua relação com a vegetação da região do interflúvio Madeira-Purus (AM-RO).
16 *Acta Amazonica*. 2009; 39: 81–90.
- 17 38. Bispo PC, Valeriano MM, Kuplich TM. Relação da vegetação de caatinga com a
18 condição geomorfométrica local. *Revista Brasileira de Engenharia Agrícola e
19 Ambiental*. 2010; 14: 523–530. doi:10.1590/S1415-43662010000500010
- 20 39. Bispo PC, Valeriano MM, Kuplich TM. Relação entre as variáveis morfométricas
21 extraídas de dados SRTM (Shuttle Radar Topography Mission) e a vegetação do
22 Parque Nacional de Brasília. *Acta Botanica Brasilica*. 2010; 24: 96–103.
23 doi:10.1590/S0102-33062010000100010
- 24 40. Muñoz V, Valeriano MM, Bispo P. Surveying the Topographic Height from
25 SRTM Data for Canopy Mapping in Brazilian Pantanal. *Geografia*. 2013; 38: 139–
26 155. Available:
27 http://www.geopantanal.cnptia.embrapa.br/publicacoes/4geo/p10_139-155.pdf
- 28 41. Guadagnin PMA, Trentin R, Alves FDS. Relação entre as variáveis
29 Geomorfométricas e a Vegetação Florestal na Bacia Hidrográfica do Arroio

- 1 Caverá - Oeste do RS. Revista do Departamento de Geografia – USP. 2015; 29:
2 246. doi:10.11606/rdg.v29i0.102131
- 3 42. Guitet S, Pélissier R, Brunaux O, Jaouen G, Sabatier D. Geomorphological
4 landscape features explain floristic patterns in French Guiana rainforest.
5 Biodiversity and Conservation. 2015; 24: 1215–1237. doi:10.1007/s10531-014-
6 0854-8
- 7 43. Bispo PC, Santos JR, Valeriano MM, Touzi R, Seifert FM. Integration of
8 polarimetric PALSAR attributes and local geomorphometric variables derived
9 from SRTM for forest biomass modeling in central Amazonia. Canadian Journal
10 of Remote Sensing. 2014; 40: 26–42. doi:10.1080/07038992.2014.913477
- 11 44. Lennon JJ, Koleff P, Greenwood JJD, Gaston KJ. Contribution of rarity and
12 commonness to patterns of species richness. Ecology Letters. 2004; 7: 81–87.
13 doi:10.1046/j.1461-0248.2004.00548.x
- 14 45. Legendre P, Gallagher ED. Ecologically meaningful transformations for
15 ordination of species data. Oecologia. 2001; 129: 271–280.
16 doi:10.1007/s004420100716
- 17 46. Borcard D, Gillet F, Legendre, Legendre P. Numerical Ecology with R. Springer.
18 2011. doi:10.1017/CBO9781107415324.004
- 19 47. Borcard D, Legendre P, Drapeau P. Partialling out the spatial component of
20 ecological variation. Ecology. 1992. pp. 1045–1055. doi:10.2307/1940179
- 21 48. Legendre P, L. L. Numerical ecology, 2nd English Edition. Numerical ecology,.
22 1998. doi:10.1017/CBO9781107415324.004
- 23 49. Griffith DA, Peres-Neto PR. Spatial modeling in ecology: The flexibility of
24 eigenfunction spatial analyses. Ecology. 2006; 87: 2603–2613. doi:10.1890/0012-
25 9658(2006)87[2603:SMIETF]2.0.CO; 2
- 26 50. Borcard D, Legendre P. All-scale spatial analysis of ecological data by means of
27 principal coordinates of neighbor matrices. Ecological Modelling. 2002; 153: 51–
28 68. doi:10.1016/S0304-3800(01)00501-4

- 1 51. Peres-Neto PR, Legendre P, Dray S, Borcard D. Variation partitioning of species
2 data matrices: Estimation and comparison of fractions. *Ecology*. 2006;87: 2614–
3 2625. doi:10.1890/0012-9658(2006)87[2614:VPOSDM]2.0.CO;2
- 4 52. R Development Core Team R, Team RDC, R Development Core Team R. R: A
5 Language and Environment for Statistical Computing. R Foundation for Statistical
6 Computing. 2011. p. 409. doi:10.1007/978-3-540-74686-7
- 7 53. Legendre P, Cáceres M, Borcard D. Community surveys through space and time:
8 Testing the space-time interaction in the absence of replication. *Ecology*. 2010; 91:
9 262–272. doi:10.1890/09-0199.1
- 10 54. Dray S, Legendre P, Blanchet G. packfor: Forward selection with permutation [R
11 package version 0.0-8/r100]. Canoco. 2011. p. 46.
- 12 55. Oksanen J, Blanchet FG, Kindt R, Legendre P, O’Hara RB, Simpson GL, et al.
13 vegan: Community Ecology Package. R package version. 2007. p. R package
14 version 1.17-9.
- 15 56. Kunin WE, Shmida A. Plant Reproductive Traits as a Function of Local, Regional,
16 and Global Abundance. *Conservation Biology*. 1997; 11: 183–192.
17 doi:10.1046/j.1523-1739.1997.95469.x
- 18 57. Cornwell WK, Ackerly DD. A link between plant traits and abundance: evidence
19 from coastal California woody plants. *Journal of Ecology*. 2010; 98: 814–821.
20 doi:10.1111/j.1365-2745.2010.01662.x
- 21 58. Espeland EK, Emam TM. The value of structuring rarity: The seven types and links
22 to reproductive ecology. *Biodiversity and Conservation*. 2011; 20: 963–985.
23 doi:10.1007/s10531-011-0007-2.
- 24 59. Lennon JJ, Beale CM, Reid CL, Kent M, Pakeman RJ. Are richness patterns of
25 common and rare species equally well explained by environmental variables?
26 *Ecography (Cop)*. 2011; 34: 529–539. doi:10.1111/j.1600-0587.2010.06669.x
- 27 60. Heino J, Soininen J. Are common species sufficient in describing turnover in
28 aquatic metacommunities along environmental and spatial gradients? *Limnology*

- 1 and Oceanography. 2010; 55: 2397–2402. doi:10.4319/lo.2010.55.6.2397
- 2 61. Tsang TPN, Bonebrake TC. Contrasting roles of environmental and spatial
3 processes for common and rare urban butterfly species compositions. *Landscape*
4 *Ecology*. 2016: 1–11. doi:10.1007/s10980-016-0427-1
- 5 62. Mazaris A, Tzanopoulos J, Kallimanis A, Matsinos Y, Sgardelis S, Pantis J. The
6 contribution of common and rare species to plant species richness patterns: the
7 effect of habitat type and size of sampling unit. *Biodiversity and Conservation*.
8 2008; 17: 3567–3577. doi:10.1007/s10531-008-9480-7
- 9 63. Luizão RCC, Luizão FJ, Paiva RQ, Monteiro TF, Sousa LS, Kruijt B. Variation of
10 carbon and nitrogen cycling processes along a topographic gradient in a central
11 Amazonian forest. *Global Change Biology*. 2004; 10: 592–600.
12 doi:10.1111/j.1529-8817.2003.00757.x
- 13 64. Balvanera P, Quijas S, Perez-Jimenez a. Distribution Patterns of Tropical Dry
14 Forest Trees Along a Mesoscale Water Availability Gradient. *Biotropica*. 2011;
15 43: 414–422. doi:DOI 10.1111/j.1744-7429.2010.00712.x
- 16 65. Wang Q, PUNCHI-MANAGE R, Lu Z, Franklin SB, Wang Z, Li Y, et al. Effects of
17 topography on structuring species assemblages in a subtropical forest. *Journal of*
18 *Plant Ecology*. 2016; 1–10.

19

20 **Supporting information**

21 **S1 Table.** Selected topographic and spatial variables for Partial Multiple Regression and
22 Partial Redundancy according to forward selection.

23

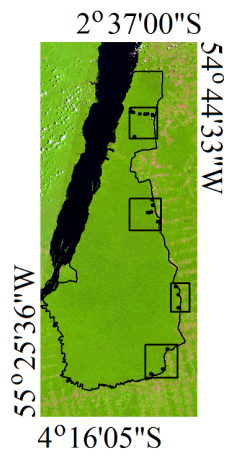
24

S1 Table. Selected topographic and spatial variables for Partial Multiple Regression and Partial Redundancy according to forward selection.

	Topographic variables	Spatial variables
<i>Std Richness</i>		
Total (1-230)	None selected	None selected
Common (1-22)	<i>h, HAND</i>	None selected
Rare (137-230)	None selected	None selected
<i>Abundance(Total)</i>		
Total (1-230)	None selected	PCNM 2, PCNM 3
Common (1-22)	<i>kh</i>	PCNM 4
Rare (137-230)	None selected	PCNM 2, PCNM 3
<i>Biomass (Total)</i>		
Total (1-230)	<i>h, G</i>	PCNM 1, PCNM 2
Common (1-22)	<i>h</i>	PCNM 1
Rare (136-230)	None selected	None selected
<i>C. Incidence</i>		
Total (1-230)	<i>h HAND, kv</i>	PCNM 1, PCNM 2, PCNM 3, PCNM 4, PCNM 5
Common (1-22)	<i>h, G</i>	PCNM 1
Rare (137-230)	<i>G</i>	PCNM 1
<i>C. Abundance</i>		
Total (1-230)	<i>h, HAND, kv</i>	PCNM 1, PCNM 2, PCNM 3, PCNM 4, PCNM 5
Common (1-22)	<i>G, HAND, kv</i>	PCNM 1, PCNM 2, PCNM3, PCNM 5
Rare (137-230)	<i>G, HAND</i>	PCNM 1
<i>C. Biomass</i>		
Total (1-230)	<i>h, G</i>	PCNM 1, PCNM 2
Common (1-35)	<i>h, G</i>	PCNM 1, PCNM2
Rare (136-230)	None selected	None selected

Composition based on incidence (*C. Incidence*); Composition based on abundance (*C. Abundance*) and Composition based on biomass (*C. Biomass*). Variables: Elevation (*h*); Slope (*G*); Profile curvature (*kv*); Plan curvature (*kh*); Height above the nearest drainage (*HAND*); Spatial variables (PCNM) based on Coordinates Principal.

Tapajós National Forest, Pará, Brazil



Legend:
■ Limit of TNF
● Plots

