


Species traits shape the relationship between local and regional species abundance distributions

MURILO DANTAS DE MIRANDA,^{1,2} LUÍS BORDA-DE-ÁGUA,^{3,4}
HENRIQUE MIGUEL PEREIRA,^{1,2,3,4,†} AND THOMAS MERCKX^{5,6} 

¹*Institute of Biology, Martin Luther University Halle-Wittenberg, Am Kirchtor 1, 06108 Halle (Saale), Germany*

²*German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Deutscher Platz 5e, 04103 Leipzig, Germany*

³*CIBIO/InBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, Universidade do Porto, Campus Agrário de Vairão, Rua Padre Armando Quintas, 4485-601 Vairão, Portugal*

⁴*CEABN/InBIO, Centro de Ecologia Aplicada “Professor Baeta Neves”, Instituto Superior de Agronomia, Universidade de Lisboa, Tapada da Ajuda, 1349-017 Lisboa, Portugal*

⁵*Behavioural Ecology and Conservation Group, Biodiversity Research Centre, Earth and Life Institute, UCLouvain, Croix du Sud 4-5, bte L7.07.04, BE-1348 Louvain-la-Neuve, Belgium*

⁶*Department of Ecology and Genetics, University of Oulu, 3000, FI-90014 Oulu, Finland*

Citation: Dantas de Miranda, M., L. Borda-de-Água, H. M. Pereira, and T. Merckx. 2019. Species traits shape the relationship between local and regional species abundance distributions. *Ecosphere* 10(5):e02750. 10.1002/ecs2.2750

Abstract. The species abundance distribution (SAD) depicts the relative abundance of species within a community, which is a key concept in ecology. Here, we test whether SADs are more likely to either follow a lognormal-like or follow a logseries-like distribution and how that may change with spatial scale. Our results show that the shape of SADs changes from logseries-like at small, plot scales to lognormal-like at large, landscape scales. However, the rate at which the SAD's shape changes also depends on species traits linked to the spatial distribution of individuals. Specifically, we show for oligophagous and small macro-moth species that a logseries distribution is more likely at small scales and a lognormal distribution is more likely at large scales, whereas the logseries distribution fits well at both small and large scales for polyphagous and large species. We also show that SAD moments scale as power laws as a function of spatial scale, and we assess the performance of Tchebichef moments and polynomials to reconstruct SADs at the landscape scale from information at local scales. Overall, the method performed well and reproduced the shapes of the empirical distributions.

Key words: body size; countryside biogeography; host-plant specialization; Lepidoptera; moths; multi-habitat landscapes; raw moments; spatial scale; species abundance distribution; species traits; Tchebichef moments; Tchebichef polynomials.

Received 22 October 2018; revised 2 April 2019; accepted 3 April 2019. Corresponding Editor: Robert R. Parmenter.

Copyright: © 2019 The Authors. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

† **E-mail:** hpereira@idiv.de

INTRODUCTION

The species abundance distribution (SAD) describes the relative abundance of species within a community, which is a central concept in ecology and essential for theories on biodiversity and biogeography (McGill et al. 2007, Matthews and Whittaker 2014, Arellano et al. 2017). Analyses of SADs that enable the identification of patterns in both commonness and rarity of

species can also be useful in applied ecology and biodiversity management (Matthews and Whittaker 2015). As such, both theoretical and empirical studies have examined the influence on SADs of several environmental and biological variables, such as elevational and latitudinal gradients, niche differentiation, dispersal, and ecological disturbance (Whittaker 1975, Hubbell 1979, 2001, Magurran 2004, Arellano et al. 2017).

Species abundance distributions—the distribution of species into abundance classes—can take a variety of shapes. Fisher et al. (1943) first derived the logseries distribution to fit the SAD of an abundance data set of Lepidoptera, thus implicitly stating that it was a monotonically decreasing distribution, with the maximum frequency for the singleton class. This was challenged by Preston (1948, 1962) who pointed out that his data on bird species abundance led to a SAD that had a bell-shaped curve with the maximum no longer occurring for the singleton class, the rarest species. Furthermore, he pointed out that this maximum for intermediate classes was progressively revealed as more data were being collected, that is, the concept of the veil line, in the process revealing the tail of rare species. Therefore, Preston clearly acknowledged that the shape of SADs was a function of the sample size, and he suggested that SADs converge on a log-normal distribution as sample size increases. The veiling line can arise for several reasons. One is that some species may first go undetected and an increase in sampling effort may reveal more species; Chao et al. (2015) have suggested methods to deal with this issue. Another reason is that species are not present everywhere, often showing some degree of aggregation, and hence, an increase in the size of the sampled area leads to an increase in the number of species and, obviously, of individuals. In this paper, we deal solely with the latter issue.

The unveiling process results from increasing sampling. Since collecting data to construct large-scale SADs is a time-consuming and economically expensive endeavor, it remains elusive what the true shape of the SAD is for very large samples, such as entire biogeographical (meta)-communities (but see ter Steege et al. 2006 and Hubbell et al. 2008). It is a major challenge in ecology to describe the regional SAD using samples characterized by a relatively small number of individuals or with data that cover only small areas relatively to the extent of the entire community. A key point that will move the study of regional SADs forward is understanding how SADs in general are affected by spatial scale (McGill et al. 2007). Some studies have focused on downscaling (i.e., how species aggregation affects the shape of local SADs within a region; Dewdney 1998, Green and Plotkin 2007), while

others have focused on upscaling, trying to predict the regional SAD from SADs obtained at smaller scales using various methods, such as maximum-entropy and Bayesian methods (Magurran 2004, Harte et al. 2009, Zillio and He 2010). The method that we apply here, developed by Borda-de-Água et al. (2012), uses the raw moments of SADs (see Eq. 1 below) to describe the scaling properties of the SAD followed by the application of the discrete Tchebichef moment and polynomial method to reconstruct it at large scales.

The description of a SAD based on its moments was first applied to data on tree and shrub species in a 50-ha plot of tropical rainforest (Borda-de-Água et al. 2012). By testing the method within the 50-ha plot and then extrapolating the SAD for areas up to 5 km², the authors demonstrated the advantages of the Tchebichef moment method over other upscaling methods. The basic idea of this method is to describe the SAD shape at different scales using the (raw) moments of the SAD (the more the better): The first raw moment is the mean, the second raw moment is related to the variance, and the third and fourth raw moments are related to the skewness and the kurtosis, respectively. When using this method, we are not concerned with the SAD at one spatial scale but, instead, with how the moments change as a function of the size of the area sampled and, especially, at the patterns exhibited by the moments as a function of the size of the area sampled. Previous work has shown that the moments have a power law behavior as a function of area size (Borda-de-Água et al. 2012, 2017). If the moments exhibit a pattern, then their values can be extrapolated for larger areas and the reconstruction of the SAD for hitherto unsampled area sizes. Different alternatives exist for the reconstruction of a distribution using (some) of its moments. Here, we use a method based on discrete Tchebichef moments and polynomials (Mukundan et al. 2001, Borda-de-Água et al. 2012). In summary, our approach consists of calculating the raw moments of the SADs at different scales, determining which tendencies the moments exhibit when plotted as a function of the sampled area, extrapolating the moments' values for larger areas, and then using the discrete Tchebichef moments and polynomials to reconstruct the SAD.

Actually, the above approach is closely related to the one which is often adopted by ecologists studying the species–area relationship (SAR). These studies typically focus on how the number of species changes as a function of area, a relationship that is well characterized by a power law. While the SAR relates to one single variable—the number of species—and is described by a single curve, for the description of the evolution of the SAD across scales, we need to describe the evolution of a distribution function. Because a distribution is characterized by its moments, such as the mean and variance, we need a curve to describe the evolution of each moment: one to describe the evolution of its mean, another to describe the evolution of its variance, and so on. Thus, we use several raw moments, and in analogy with SAR studies, we plot the raw moments of the SAD as a function of area. Our emphasis is on how the SAD changes as a function of area and not on which distribution gives the best fit of the SAD at all scales. Because moments describe any distribution, our approach provides a continuous description of the progression of any empirical SAD across different area sizes.

Here, we test whether SADs are more likely to follow lognormal-like or logseries-like distributions, and whether this is affected by scale. By logseries-like, we mean distributions that are monotonically decreasing functions, and by lognormal-like, we mean distributions that when plotted on a logarithmic scale exhibit a maximum for intermediate classes and are approximately bell-shaped; hereafter, we refer to these distributions as logseries and lognormal, respectively. The logseries SAD implies that most individuals belong to a few species and that most species are represented by a few individuals, whereas the lognormal SAD implies a higher number of species with intermediate abundance and smaller numbers of rare and abundant species (Magurran 2004). Also, since some species traits (e.g., mobility, niche preference) are linked to the spatial distribution of individuals, we test whether and how such traits affect the scale dependency of SADs. Moreover, we analyze how SAD moments relate to spatial scale, and we assess the performance of the Tchebichef moments and polynomials to reconstruct SADs from the local plot scale to the landscape scale.

We do so using a nested multi-site design in three multi-habitat landscapes. Our focus is on the species-rich group of macro-moths, which allows us to test the effect of two traits linked to spatial distribution: body size and host-plant specialization.

METHODS

Study area and sampling

We gathered species abundance data on macro-moths in the Castro Laboreiro area (approximately 42°2' N, 8°10' W) within the Peneda-Gerês National Park, NW Portugal (Fig. 1). The landscape within our study area consists mainly of scrub (78.4%), forest (10.5%), and meadows (9.8%). Sampling was conducted in three 1.6-km² multi-habitat landscapes that differed in dominant habitat type: a meadow-dominated, scrub-dominated, and forest-dominated landscape. For each landscape, 28 fixed light-trap sites were selected using a semi-nested design with four levels, each corresponding to a different spatial scale (Pronça and Pereira 2013; Fig. 1). This sampling design follows an almost fractal arrangement: There is a trap-level scale of ~20 × 20 m (i.e., based on an attraction-to-light radius of ~10 m; Merckx and Slade 2014; then there are four of these samples forming a 80 × 80 m², i.e., 4 × 20, which is nested in a 320 × 320 m², i.e., 4 × 80, and this arrangement is then embedded in a 1280 × 1280 m², i.e., 4 × 320; Fig. 1). Henceforth, expressions such as “we sampled (extrapolated to) an area of size X” are shorthand for “we sampled (extrapolated to) a set of plots with total area corresponding to the total area of the sampling traps, following a design as shown in Fig. 1, and whose corners form a square of area X.”

During both 2011 and 2012, macro-moths were light-trapped between May and September, using identical equipment (i.e., heath pattern 6W actinic light traps). Each of the 84 sites was sampled three times per year, resulting in a total of six samples per site over both years. Per sampling night seven sites were simultaneously sampled so that they were sampled under identical weather conditions, covering three spatial scales: (1) trap-level scale, (2) 80 × 80 m, and (3) 320 × 320 m (Fig. 1). Moreover, all 84 sites (i.e.,

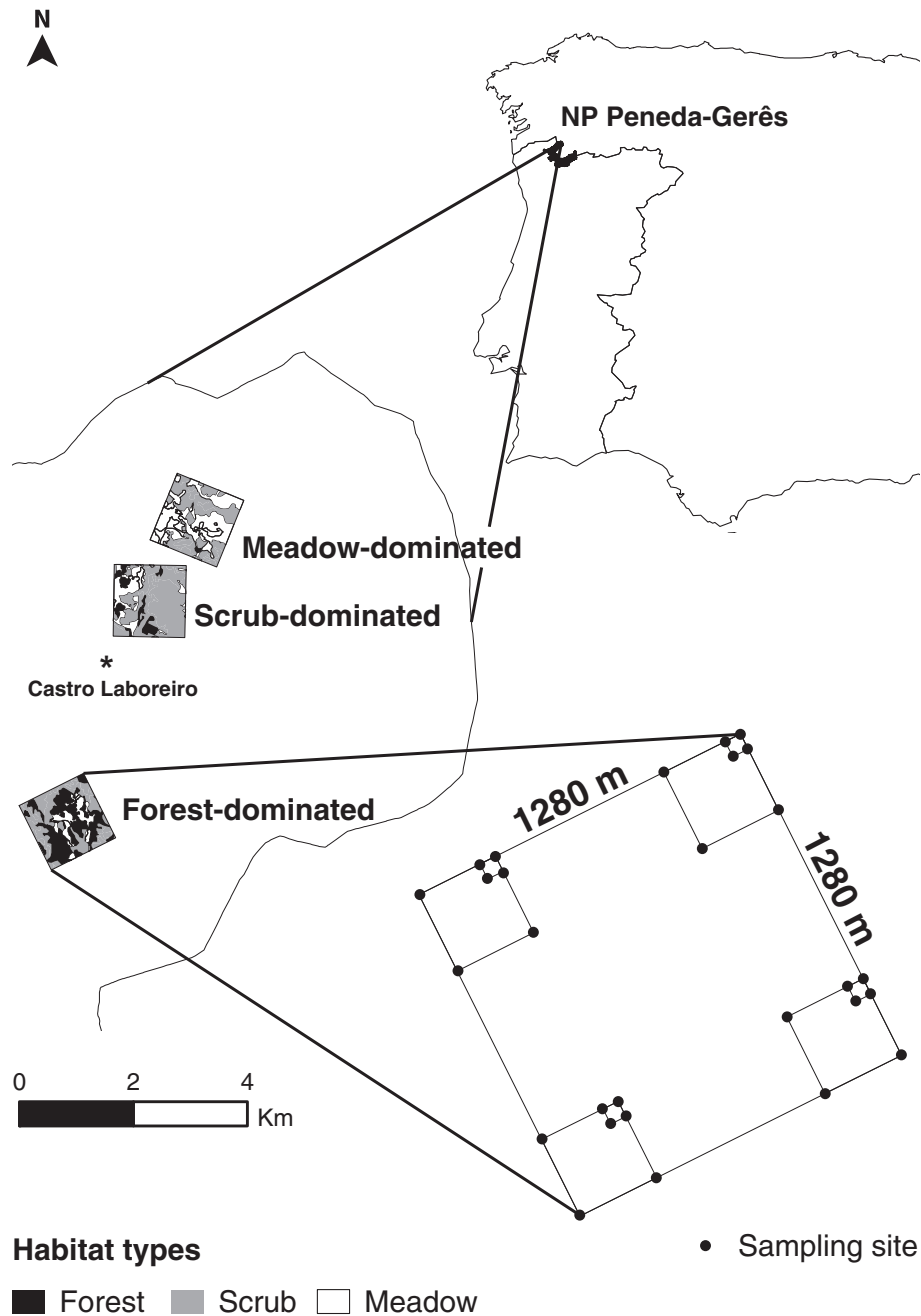


Fig. 1. Location of the study area in Peneda-Gerês National Park, Northern Portugal. Eighty-four fixed light-trap sampling sites ($\sim 400 \text{ m}^2$) were part of a semi-nested ($80 \times 80 \text{ m}$, $320 \times 320 \text{ m}$) sampling design in three multi-habitat study landscapes ($1280 \times 1280 \text{ m}$).

twelve sampling nights) were sampled in as short a period as possible so as to avoid seasonal differences in species composition, while the overnight sampling only took place during

weather conditions favorable for moth flight activity (i.e., minimum night temperature $>10^\circ\text{C}$; maximum wind speed $<20 \text{ km/h}$; no persistent rain; see Merckx et al. 2012a, b).

Species abundance distribution models

For each sampling site, which corresponds to the trap-level scale, macro-moth abundance data were lumped and then fitted to histogram-type data with logseries and truncated lognormal distributions using maximum likelihood. For the next spatial scale (i.e., 80×80 m), we lumped all sampling sites within these areas and fitted logseries and lognormal distributions for each area. This procedure was repeated three more times for the other spatial scales (i.e., 320×320 m, 1280×1280 m, and the sum of the three 1280×1280 m landscapes). Because we were merely interested in assessing the relative quality of fit among the logseries and lognormal distribution across scales, we used corrected Akaike information criteria (AIC_c) in order to select the best-fitting model. Specifically, we calculated the difference between the AIC_c values (ΔAIC_c) corresponding to the lognormal and the logseries distributions, using the threshold of $\Delta AIC_c > |2|$ to establish whether both distributions are significantly different (Burnham and Anderson 2002, Slik et al. 2015). Accordingly, the lognormal SAD model is considered to provide a better fit when $\Delta AIC_c \leq -2$ while $\Delta AIC_c \geq 2$ indicates a better fit for the logseries model. Models characterized by $-2 < \Delta AIC_c < 2$ were classified as intermediate ones, with both the lognormal and logseries having a similar level of support. We then created ordinal logistic regression models (Guénette and Villard 2004, Rutherford et al. 2007) to test the relation between spatial scale ($\ln(\text{area})$) and the probability for these lognormal, intermediate, and logseries SADs. Next, we evaluated the goodness of fit and deviance. The strength of the association (McFadden's R^2) was calculated as $1 - (L_{\text{mod}}/L_{\text{null}})$, where L_{mod} is the log-likelihood value for the fitted model and L_{null} is the log-likelihood for the null model which includes only an intercept.

Ordinal logistic regression models too were used to test how species traits (body size and host-plant specialization) affect the SAD shape probability (three classes: lognormal, intermediate, and logseries). Species-specific average wingspan (mm) was used as a proxy for body size (Hambäck et al. 2007). Mean values were obtained from www.lepidoptera.eu, topped up with values from Manley (2009) for a few species where information was missing. Species

with average wingspan smaller than the overall median (32.5 mm) were classified as small, the others as large. Host-plant specialization was classified into two classes: oligophagous species, whose larvae only feed on plant species from the same family, and polyphagous species, whose larvae are able to feed on several plant families. Data on host-plant specialization were obtained from the same sources as above. An overview of all species classifications can be found in Appendix S1: Table S1. All statistical analyses were run in the statistical software environment R version 3.1.1 with packages mass and sads (R Development Core Team 2014, Prado et al. 2018).

SAD raw moments, Tchebichef moments, and Tchebichef polynomials

In principle, a probability distribution can be reconstructed directly from its moments (Borda-de-Água et al. 2012). However, in practice this is not viable because of the large number of moments required and the numerical instabilities associated with the high order moments. Thus, other methods have to be sought (Mukundan et al. 2001). Following Borda-de-Água et al. (2012), we reconstructed the probability density function using the estimated discrete Tchebichef moments and polynomials. Tchebichef moments are a general non-parametric tool to describe the shape of any distribution. The idea is to fit a histogram, $f(x)$ with a sum of Tchebichef polynomials, $\tilde{t}_n(x)$, weighed by the corresponding Tchebichef moments, T_n , according to the formula

$$f(x) = \sum_{n=0}^{N-1} T_n \tilde{t}_n(x)$$

In principle, we need as many Tchebichef moments as the number of bins of the histogram. Further information about discrete Tchebichef moments and polynomials can be found in Mukundan et al. (2001) and Borda-de-Água et al. (2012).

The Tchebichef moments are related to the raw moments (Appendix S3). The raw moment of order n of the SAD from a given community is calculated as

$$M_n = \frac{1}{S} \sum_{j=1}^S x_j^n \quad (1)$$

where S is the number of species, and x_j is the \log_2 -transformed number of individuals of species j . For each spatial scale, we calculated the average of the moments obtained from each set of lumped sampling sites and then plotted this average as a function of the corresponding spatial scale. As mentioned previously, the number of Tchebichef moments is limited to the number of bins of the histogram. Specifically, if the histogram has N bins, then the maximum number of Tchebichef moments is N and it requires the same number of raw moments (Mukundan et al. 2001, Borda-de-Água et al. 2012). Although in our analysis the number of moments varied between 0 and 11—because there were 12 bins in the histogram of all landscapes—we limited the number of moments to 9, both separately for each landscape and for all landscapes together, as the extrapolations appeared sensitive for the highest two moments. In practice, we observe that the calculation of higher moments has numerical instabilities, because of the very large numbers that result from the large exponents involved in the computation of higher order moments (Mukundan 2004). Thus, the number of moments used has to be chosen judiciously, often by trial and error, and currently, there are no rules to determine the number of moments to be used. However, these numerical instabilities are very noticeable, leading to distributions with large oscillations with negative values. In order to better quantify the best number of moments to use, we determined the number of moments that minimizes the sum of the difference in the histogram of the real data and that obtained with the Tchebichef method.

For each site, data were lumped. We calculated the moments for each sampling site (Eq. 1), and then, for each moment, we calculated the average from all sampling sites. For the next spatial scale (i.e., 80×80 m), we lumped all sampling sites within such areas, calculated the moments for each area, and for each moment we then averaged the values across all these areas. This procedure was repeated one more time for the next spatial scale (i.e., 320×320 m). Next, we fitted a linear regression in order to assess the relationship between the \ln -transformed area and the \ln -transformed moment, $\ln(M_n(A)) = a_n + b_n \ln(A)$, where a_n and b_n are parameters estimated from the regression of that particular moment. Then, we extrapolated the moments for the largest two

scales (i.e., 1280×1280 m or the individual landscapes, and the sum of the three landscapes) using the moments obtained at the trap level, 80×80 m and 320×320 m scales; that is, we used the smallest three scales of the sampling design to extrapolate the SADs (1) to the full data for each landscape and (2) to the entire data set. We used chi-square tests to compare the extrapolated and observed SADs on histograms and did this separately for each landscape as well as for the sum of the three landscapes.

RESULTS

In total, we collected 22,825 individuals belonging to 378 species. Most species belonged to two families: Noctuidae (39.4%) and Geometridae (38.9%). The two most abundant species were the noctuid *Xestia agathina* and the geometrid *Pachycnemia hippocastanaria*, comprising 8.2% and 6.9% of all individuals collected, respectively, 15.3% of the species were singletons, and 38.3% had five individuals or less.

For all macro-moth species, ordinal logistic regression showed that the probabilities of moving from a logseries to intermediate/lognormal distribution (or from logseries/intermediate to lognormal) increased as \ln -area increases (Fig. 2). This means that the probability of being logseries distributed is highest at small areas and that the probability of being lognormal distributed is highest at large areas (\ln -area = 0.395; $t = 4.436$; $P < 0.0001$; Fig. 2; Appendix S2: Fig. S1; Appendix S1: Table S3).

Our results also show the importance of host-plant specialization ($R_{\text{Oligophagy}}^2 = 0.14$; $R_{\text{Polyphagy}}^2 = 0.00$) and body size ($R_{\text{Small}}^2 = 0.27$; $R_{\text{Large}}^2 = 0.01$), with the difference in R^2 values indicating a fit of the model only for oligophagous and small species. This shows that for oligophagous and small species, a logseries distribution is more likely at small areas, and a lognormal distribution is more likely at large areas ($\text{Area}_{\text{Oligophagy}} = 0.412$; $t = 4.463$; $P < 0.0001$; $\text{Area}_{\text{Small}} = 0.717$; $t = 5.786$; $P < 0.0001$). For polyphagous and large body size species groups, the logseries distribution generally provided the best relative fit at both small and large spatial scales (Fig. 2; Appendix S1: Table S3).

Plotting the \ln -transformed moments of order 1–9 as a function of the \ln -transformed area

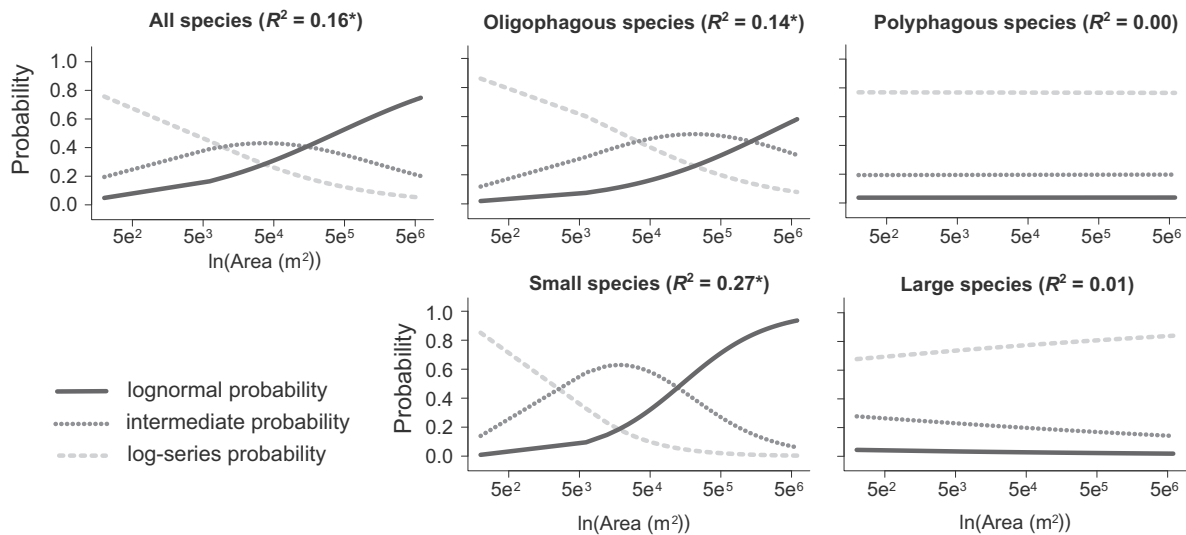


Fig. 2. Probabilities for three types of species abundance distribution (SAD)—lognormal, logseries, and an intermediate combination of both ($-2 < \Delta AIC < 2$)—as a function of area ($\ln m^2$). Panels depict relationships for all macro-moth species as well as for separate species groups, contrasting oligophagous versus polyphagous species, and small versus large species. Solid, dotted, and dashed lines represent lognormal, intermediate, and log-series SADs, respectively, based on ordinal logistic regression model output. R^2 values and significance levels of the probabilities when changing distribution as function of \ln -area ($*P < 0.05$) are given for each panel.

shows that there is an almost linear relationship (Fig. 3; Appendix S1: Table S2), similar to previous findings (Borda-de-Água et al. 2012, 2017). We used this relationship to extrapolate the SAD to larger scales. For the meadow-, forest-, and scrub-dominated landscapes, the SAD was extrapolated to an area of 1280×1280 m (Fig. 1). Next, the SAD was also extrapolated to the sum of all three landscapes combined. Consistently, the predicted extrapolated curves do not statistically differ from the empirical distributions (extrapolated SADs: gray lines in Fig. 4, Appendix S2: Figs. S2, S3; $\chi^2_{\text{Meadow}} = 120$, $df = 110$, $P = 0.24$; $\chi^2_{\text{Scrub}} = 108$, $df = 99$, $P = 0.25$; $\chi^2_{\text{Forest}} = 132$, $df = 121$, $P = 0.23$; $\chi^2_{\text{Three landscapes}} = 108$, $df = 99$, $P = 0.25$). Also, note that the distributions at smaller spatial scales, which contain the information used to forecast, are very different from the extrapolated distributions (Appendix S2: Fig. S1).

DISCUSSION

The main goals of this study were to assess how well Tchebichef moments and polynomials are able to predict the regional SAD using SADs

from smaller spatial scales and whether spatial scale and species traits affect the shape of SADs. We showed that the shape of the SADs changes across spatial scales, but we also showed that this is less noticeable for polyphagous and large species.

A simple qualitative model helps interpret our results for the different groups of species. Imagine two communities with exactly the same number of species and relative abundance, or in other words with the total number of individuals occurring at the same density. In one of the communities, individuals of the same species tend to be strongly aggregated, while in the other community individuals of different species are spatially well mixed. We do not invoke any particular reason for the level of clustering of both communities, but it could be due to dispersal (Borda-de-Água et al. 2017), which is likely the case for small versus large-bodied macro-moth species, or it could be due to niche preferences, which is likely the case for oligophagous moths. For small sample sizes, if one increases the area sampled in the aggregated community, we find a small number of species, but each represented by several individuals. Furthermore,

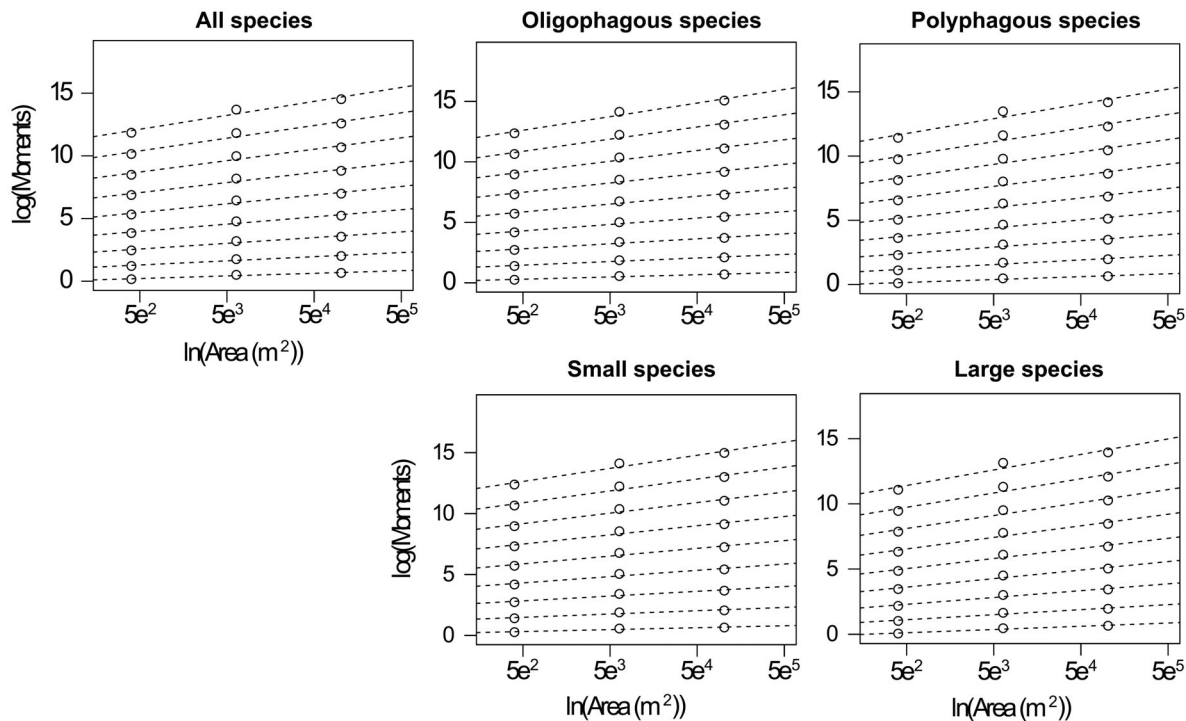


Fig. 3. Plots of \ln -transformed moments (1–9) of the species abundance distribution as a function of spatial scale ($\ln m^2$). Panels depict relationships for all macro-moths as well as for separate species groups, contrasting oligophagous versus polyphagous species, and small versus large species. Dashed lines represent the best fit of linear regression models.

when we increase the sample size one tends to find more individuals of the same species. Thus, as we increase the sampling scale, the SAD quickly develops a maximum for intermediate abundance classes, although we should obviously not exclude the presence of some rare species in the sample. On the other hand, if one increases sample size in the mixed community from small to intermediate sample sizes, we tend to find more species, but all with small abundances because species are well mixed, the SAD classes corresponding to rare species keep increasing, and the maximum for intermediate classes only occurs for larger sample sizes than it does for the case of the aggregated community. Hence, most species tend to remain rare and the SAD classes corresponding to rare species keep increasing. This trend will eventually stop because the rarest species will accumulate more individuals and not many more rare species will be found. Thus, overall, most species will have intermediate abundances, but it requires a larger

sample size for the SAD to finally start developing a maximum for intermediate abundance classes; Borda-de-Água et al. (2007) performed simulations that illustrate these behaviors.

In line with our previous explanations, complexity arises as different species traits interact to determine SADs (Gaston et al. 2000). For instance, although polyphagous moth species typically have wider distribution ranges than oligophagous species (Spitzer et al. 1984, Spitzer and Lepš 1988, Quinn et al. 1997), the latter are also typically smaller than polyphagous species. Also, large moth species tend to occur at lower densities at the local scale than small species (Peters 1986, Nieminen et al. 1999). Moreover, large macro-moth species (e.g., noctuids) are typically relatively mobile (Ockinger et al. 2010, Sekar 2012, Slade et al. 2013). This agrees with our previous simple model and it may explain why the SADs of small species followed the log-series distribution at the local scale, and changed to the lognormal distribution at larger scales,

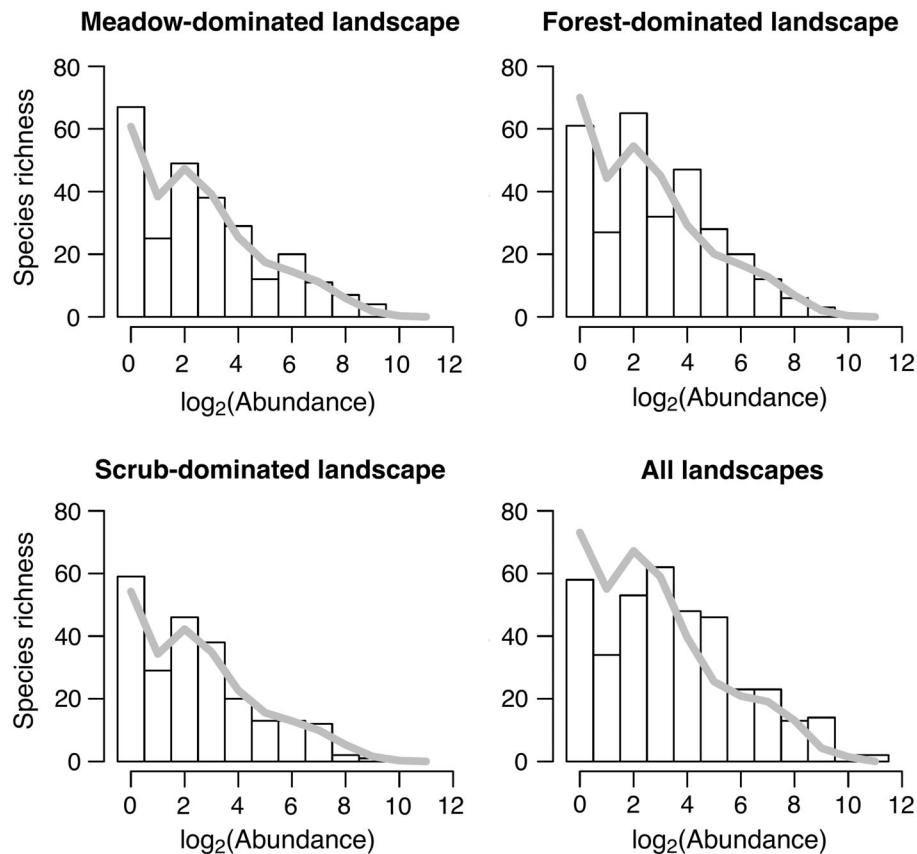


Fig. 4. Species abundance distributions (SAD) for all macro-moth species, both separated according to landscape type (forest-, scrub-, and meadow-dominated), and for all three landscapes combined. Histograms represent the observed data at these spatial scales while gray lines represent the number of species as predicted by the moments ($n = 9$). For the meadow-, forest-, and scrub-dominated landscapes, the SAD (gray line) was extrapolated to an area of 1280×1280 m and to an area that summed these three landscapes (Fig. 1).

while the SADs for the large moth species followed the logseries distribution at all spatial scales. As we mentioned before, the higher dispersal ability of large species tends to homogenize them in the community, making the SAD steeper and placing the mode at the rare species. On the other hand, small species tend to be more aggregated, and hence more common at a given spatial scale, due to their low to intermediate mobility levels (Nieminen et al. 1999, Borda-de-Água et al. 2017). As a consequence, the SADs of this species group show a lognormal shape at larger spatial scales. Likewise, Mouquet and Loreau (2003) observed that species rank-abundance distributions were strongly affected by the level of dispersal between spatial scales. Dornelas and Connolly (2008) also

showed that evenly spaced species are rare in abundance and dominate the rare mode of the SAD, while clustered species dominate the intermediate mode.

Some studies have shown that plant species distributions are important predictors of moth abundance, diversity, and distribution (Kitching et al. 2000, Hilt and Fiedler 2006, Novotny et al. 2006). If host plants are generally characterized by a patchy distribution, we would expect that oligophagous species are more aggregated than polyphagous ones (Lindström et al. 1994). This interpretation would explain why the shape of the SADs for oligophagous species also changed from logseries at a local scale to lognormal at larger scales, while no change was observed for polyphagous species.

The capability of forecasting the SAD for larger areas than those sampled is one of the most interesting features of our method. There are, however, two important assumptions that need to be met (1) the correct identification of the pattern associated with the moments within the sampled spatial scales and (2) this pattern needs to remain valid for the extrapolated areas. Here, and contrary to previous applications (Borda-de-Água et al. 2012, 2017), we only have a small number of area sizes as a base to fit the pattern. However, because the sampling design follows a fractal pattern, it is ideal to assess the patterns on a logarithmic scale and, in particular, to identify possible power laws. Note nevertheless that the method also works when the sampling scheme is not fractal, such as the scheme used in the paper where the method was originally developed (Borda-de-Água et al. 2012) or in Borda-de-Água et al. (2017). Although we recognize that we used only three points based on which we fitted the moments, they nevertheless follow approximately straight lines for several orders of magnitude. More work, both empirical and theoretical, needs to be carried out in order to assess the generality of this pattern across taxa and across space. With regard to the second assumption, the pattern remains valid for the extrapolated areas. However, extrapolating for even larger areas should be done with care given the uncertainty on the extrapolation of the coefficients of the Tchebichef moments and the increasing habitat heterogeneity at larger scales. Indeed, although our method partially reveals the inherent heterogeneity in the landscape, at very large scales, we risk combining very different communities which may translate into sudden changes in the scaling patterns of the moments or of the SAR.

Overall, we believe that our results are in agreement with Preston's veil line concept, and for those groups studied here that did not exhibit it, the reason is likely to be due to paucity of data. However, we do not assume that the distributions are converging to a lognormal distribution. The reason is that empirical (ter Steege et al. 2006), simulation (Borda-de-Água et al. 2007), and theoretical (Hubbell 2001) results have hinted at the possibility of the distributions reverting to a logseries type of curve for very large areas.

CONCLUSIONS

We show that the shape of SADs generally changes across spatial scales and that species traits are able to affect the rate of this change; small species are characterized by SADs that change more rapidly than those of large species, while SADs of polyphagous species do not exhibit any change contrary to those of oligophagous species. Although studying species richness across spatial scales—SARs—is common practice (Rosenzweig 1995), our findings demonstrated the importance of considering relative species abundance via the SAD across spatial scales. Indeed, our approach is similar to that used by ecologists when analyzing species richness across spatial scales, but we added the analysis of species abundance across scales. We predict this analysis to be important both from a theoretical perspective, thanks to the patterns it reveals, and from a practical perspective, thanks to possible applications to conservation and management of species.

ACKNOWLEDGMENTS

We thank Gabriel Arellano and Tom Matthews for their constructive comments. The Brazilian National Counsel of Technological and Scientific Development CNPq provided financial support to Murilo Dantas de Miranda (process 237206/2012-9) while the Portuguese Science Foundation FCT provided financial support to Thomas Merckx (postdoc grant SFRH/BPD/74393/2010). Luís Borda-de-Água was funded by the FCT project MOMENTOS (PTDC/BIA-BIC/5558/2014). Permits were granted to carry out light trapping within the Peneda-Gerês National Park.

LITERATURE CITED

- Arellano, G., M. N. Umaña, M. J. Macía, M. I. Loza, A. Fuentes, V. Cala, and P. M. Jørgensen. 2017. The role of niche overlap, environmental heterogeneity, landscape roughness and productivity in shaping species abundance distributions along the Amazon-Andes gradient. *Global Ecology and Biogeography* 26:191–202.
- Borda-de-Água, L., P. A. Borges, S. P. Hubbell, and H. M. Pereira. 2012. Spatial scaling of species abundance distributions. *Ecography* 35:549–556.
- Borda-de-Água, L., S. P. Hubbell, and F. He. 2007. Scaling biodiversity under neutrality. Pages 347–375 *in* D. Storch, P. A. Marquet and J. H. Brown, editors.

- Scaling biodiversity. Cambridge University Press, Cambridge, UK.
- Borda-de-Água, L., R. J. Whittaker, P. Cardoso, F. Rigal, A. M. Santos, I. R. Amorim, A. Parmakelis, K. A. Triantis, H. M. Pereira, and P. A. V. Borges. 2017. Dispersal ability determines the scaling properties of species abundance distributions: a case study using arthropods from the Azores. *Scientific Reports* 7:3899.
- Burnham, K. P., and D. R. Anderson. 2002. Model selection and multimodel inference: a practical information-theoretic approach, Second edition. Springer, New York, New York, USA.
- Chao, A., T. C. Hsieh, R. L. Chazdon, R. K. Colwell, and N. J. Gotelli. 2015. Unveiling the species-rank abundance distribution by generalizing the Good-Turing sample coverage theory. *Ecology* 96:1189–1201.
- Dewdney, A. K. 1998. A general theory of the sampling process with applications to the “Veil Line”. *Theoretical Population Biology* 54:294–302.
- Dornelas, M., and S. R. Connolly. 2008. Multiple modes in a coral species abundance distribution. *Ecology Letters* 11:1008–1016.
- Fisher, R. A., A. S. Corbet, and C. B. Williams. 1943. The relation between the number of species and the number of individuals in a random sample of an animal population. *Journal of Animal Ecology* 12:42–58.
- Gaston, K. J., T. M. Blackburn, J. J. Greenwood, R. D. Gregory, R. M. Quinn, and J. H. Lawton. 2000. Abundance-occupancy relationships. *Journal of Applied Ecology* 37:39–59.
- Green, J. L., and J. B. Plotkin. 2007. A statistical theory for sampling species abundances. *Ecology Letters* 10:1037–1045.
- Guénette, J.-S., and M.-A. Villard. 2004. Do empirical thresholds truly reflect species tolerance to habitat alteration? *Ecological Bulletins* 1:163–171.
- Hambäck, P. A., K. S. Summerville, I. Steffan-Dewenter, J. Krauss, G. Englund, and T. O. Crist. 2007. Habitat specialization, body size, and family identity explain lepidopteran density-area relationships in a cross-continental comparison. *Proceedings of the National Academy of Sciences USA* 104:8368–8373.
- Harte, J., A. B. Smith, and D. Storch. 2009. Biodiversity scales from plots to biomes with a universal species-area curve. *Ecology Letters* 12:789–797.
- Hilt, N., and K. Fiedler. 2006. Arctiid moth ensembles along a successional gradient in the Ecuadorian montane rain forest zone: How different are sub-families and tribes? *Journal of Biogeography* 33:108–120.
- Hubbell, S. P. 1979. Tree dispersion, abundance, and diversity in a tropical dry forest. *Science* 203:1299–1309.
- Hubbell, S. P. 2001. The unified neutral theory of biodiversity and biogeography. Princeton University Press, Princeton, New Jersey, USA.
- Hubbell, S. P., F. He, R. Condit, L. Borda-de-Água, J. Kellner, and H. ter Steege. 2008. How many tree species are there in the Amazon and how many of them will go extinct? *Proceedings of the National Academy of Sciences USA* 105:11498–11504.
- Kitching, R. L., A. G. Orr, L. Thalib, H. Mitchell, M. S. Hopkins, and A. W. Graham. 2000. Moth assemblages as indicators of environmental quality in remnants of upland Australian rain forest. *Journal of Applied Ecology* 37:284–297.
- Lindström, J., L. Kaila, and P. Niemelä. 1994. Polyphagy and adult body size in geometrid moths. *Oecologia* 98:130–132.
- Magurran, A. E. 2004. Measuring biological diversity. Blackwell Publishing, Malden, Massachusetts, USA.
- Manley, C. 2009. British moths and butterflies: a photographic guide. Bloomsbury Publishing, London, UK.
- Matthews, T. J., and R. J. Whittaker. 2014. Fitting and comparing competing models of the species abundance distribution: assessment and prospect. *Frontiers in Biogeography* 6:67–82.
- Matthews, T. J., and R. J. Whittaker. 2015. On the species abundance distribution in applied ecology and biodiversity management. *Journal of Applied Ecology* 52:443–454.
- McGill, B. J., et al. 2007. Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework. *Ecology Letters* 10:995–1015.
- Merckx, T., R. E. Feber, D. J. Hoare, M. S. Parsons, C. J. Kelly, N. A. Bourn, and D. W. Macdonald. 2012a. Conserving threatened Lepidoptera: towards an effective woodland management policy in landscapes under intense human land-use. *Biological Conservation* 149:32–39.
- Merckx, T., L. Marini, R. E. Feber, and D. W. Macdonald. 2012b. Hedgerow trees and extended-width field margins enhance macro-moth diversity: implications for management. *Journal of Applied Ecology* 49:1396–1404.
- Merckx, T., and E. M. Slade. 2014. Macro-moth families differ in their attraction to light: implications for light-trap monitoring programmes. *Insect Conservation and Diversity* 7:453–461.
- Mouquet, N., and M. Loreau. 2003. Community patterns in source-sink metacommunities. *American Naturalist* 162:544–557.
- Mukundan, R. 2004. Some computational aspects of discrete orthonormal moments. *IEEE Transactions on Image Processing* 13:1055–1059.

- Mukundan, R., S. H. Ong, and P. A. Lee. 2001. Image analysis by Tchebichef moments. *IEEE Transactions on Image Processing* 10:1357–1364.
- Nieminen, M., H. Rita, and P. Uuvana. 1999. Body size and migration rate in moths. *Ecography* 22:697–707.
- Novotny, V., P. Drozd, S. E. Miller, M. Kulfan, M. Janda, Y. Basset, and G. D. Weiblen. 2006. Why are there so many species of herbivorous insects in tropical rainforests? *Science* 313:1115–1118.
- Öckinger, E., et al. 2010. Life-history traits predict species responses to habitat area and isolation: a cross-continental synthesis. *Ecology Letters* 13:969–979.
- Peters, R. H. 1986. *The ecological implications of body size*. Cambridge University Press, Cambridge, UK.
- Prado, P. I., de Dantas Miranda M. and A. Chalom. 2018. 'sads': maximum likelihood models for species abundance distributions. <https://cran.r-project.org/package=sads>
- Preston, F. W. 1948. The commonness, and rarity, of species. *Ecology* 29:254–283.
- Preston, F. W. 1962. The canonical distribution of commonness and rarity: part I. *Ecology* 43:185–215.
- Proença, V., and H. M. Pereira. 2013. Species-area models to assess biodiversity change in multi-habitat landscapes: the importance of species habitat affinity. *Basic and Applied Ecology* 14:102–114.
- Quinn, R., K. Gaston, T. I. M. Blackburn, and B. Everham. 1997. Abundance-range size relationships of macrolepidoptera in Britain: the effects of taxonomy and life history variables. *Ecological Entomology* 22:453–461.
- R Development Core Team. 2014. *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Rosenzweig, M. L. 1995. *Species diversity in space and time*. Cambridge University Press, Cambridge, UK.
- Rutherford, G. N., A. Guisan, and N. E. Zimmermann. 2007. Evaluating sampling strategies and logistic regression methods for modelling complex land cover changes. *Journal of Applied Ecology* 44:414–424.
- Sekar, S. 2012. A meta-analysis of the traits affecting dispersal ability in butterflies: Can wingspan be used as a proxy? *Journal of Animal Ecology* 81:174–184.
- Slade, E. M., T. Merckx, T. Riutta, D. P. Bebber, D. Redhead, P. Riordan, and D. W. Macdonald. 2013. Life-history traits and landscape characteristics predict macro-moth responses to forest fragmentation. *Ecology* 94:1519–1530.
- Slik, J. W. F., et al. 2015. An estimate of the number of tropical tree species. *Proceedings of the National Academy of Sciences USA* 112:7472–7477.
- Spitzer, K., and J. Lepš. 1988. Determinants of temporal variation in moth abundance. *Oikos* 53:31–36.
- Spitzer, K., M. Rejmánek, and T. Soldán. 1984. The fecundity and long-term variability in abundance of noctuid moths (Lepidoptera, Noctuidae). *Oecologia* 62:91–93.
- ter Steege, H., et al. 2006. Continental-scale patterns of canopy tree composition and function across Amazonia. *Nature* 443:444–447.
- Whittaker, R. H. 1975. *Communities and ecosystems*. MacMillan Publishers, New York, New York, USA.
- Zillio, T., and F. He. 2010. Inferring species abundance distribution across spatial scales. *Oikos* 119:71–80.

SUPPORTING INFORMATION

Additional Supporting Information may be found online at: <http://onlinelibrary.wiley.com/doi/10.1002/ecs2.2750/full>