RESEARCH PAPER

Global Ecology and Biogeography — A Journal of Macroecology — WILEY

Environment- and trait-mediated scaling of tree occupancy in forests worldwide

Haibao Ren[1,2] | Petr Keil[3] | Xiangcheng Mi[1] | Keping Ma[1] | Zhanqing Hao[4] |
Wanhui Ye[5] | Luxiang Lin[6] | Renato Valencia[7] | Christine Dawn Fletcher[8] |
Duncan W. Thomas[9] | Robert W. Howe[10] | James Lutz[11] | Norman A. Bourg[12] |
Sheng-Hsin Su[13] | I-Fang Sun[14] | Li Zhu[1] | Li-Wan Chang[13] | Xihua Wang[15,16] |
Xiaojun Du[1] | David Kenfack[17,18] | George B. Chuyong[19] | Walter Jetz[2]

[1]State Key Laboratory of Vegetation and Environmental Change, Institute of Botany, Chinese Academy of Sciences, Beijing

[2]Department of Ecology and Evolutionary Biology, Yale University, New Haven, Connecticut

[3]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany

[4]Key Laboratory of Forest Ecology and Management, Institute of Applied Ecology, Chinese Academy of Sciences, Shengyang

[5]Key Laboratory of Vegetation Restoration and Management of Degraded Ecosystems, South China Botanical Garden, Chinese Academy of Sciences, Guangzhou

[6]Key Laboratory of Tropical Forest Ecology, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Kunming

[7]Herbario QCA, Laboratorio de Ecología de Plantas, Escuela de Ciencias Biológicas, Pontificia Universidad Católica del Ecuador, Quito, Ecuador

[8]Forest Research Institute Malaysia, Kepong, Kuala Lumpur, Malaysia

[9]School of Biological Sciences, Washington State University, Vancouver, Washington, USA

[10]Department of Natural and Applied Sciences, University of Wisconsin-Green Bay, Green Bay, Wisconsin, USA

[11]Wildland Resources Department, Utah State University, Logan, UT, USA

[12]Conservation Ecology Center, Smithsonian Conservation Biology Institute, National Zoological Park, Washington, DC, USA

[13]Taiwan Forestry Research Institute, Taipei

[14]Department of Natural Resources and Environmental Studies, National Dong Hwa University, Hualien

[15]Tiantong National Forest Ecosystem Observation and Research Station, Ningbo

[16]School of Ecological and Environmental Sciences, East China Normal University, Shanghai

[17]Center for Tropical Forest Science–Forest Global Earth Observatory, Smithsonian Tropical Research Institute, Panama, Republic of Panama

[18]Department of Botany, National Museum of Natural History, Washington, DC, USA

[19]Department of Botany and Plant Physiology, University of Buea, Buea, Cameroon

Correspondence
Haibao Ren, State Key Laboratory of Vegetation and Environmental Change, Institute of Botany, Chinese Academy of Sciences, Beijing 100093
E-mail: renhb@ibcas.ac.cn

Walter Jetz, EEB Department, Yale University, New Haven, CT 06511, USA.
E-mail: walter.jetz@yale.edu

Funding information
Strategic Priority Research Program of Chinese Academy of Sciences, Grant/Award Number: XDB31030000; National Natural

Aim: The relationship between the proportion of sites occupied by a species and the area of a site [occupancy–area relationship (OAR)] offers key information for biodiversity management and has long fascinated ecologists. We quantified the variation in OAR for 3,157 woody species in 17 forest plots worldwide and tested the relative importance of environment and species traits for explaining this variation and evaluated overall model predictive ability.

Location: Global.

Time period: Early 21st century.

Major taxa studied: Woody plants.

Methods: We used mixed-effect regression to examine the observed shape of the OAR (its "slope") against species-specific and plot-wide predictors: coarse-grain occupancy, tree size, plot species richness, energy availability and topographic complexity.

Results: We found large variation in OAR slopes, and the variation was strongest among species within plots. The OAR slopes showed a latitudinal trend and were steeper near the equator. As predicted, coarse-grain occupancy and tree size negatively affected OAR slopes, whereas species richness had a positive effect and explained most of the variance between plots. Although hypothesized directionalities were broadly confirmed, traits and environment had relatively limited overall predictive power.

Main conclusions: These results document the variation of the OAR for 3,157 species at near-global extent. We found a latitudinal gradient in OAR slopes and confirmed key hypothesized predictors. But at this global extent and over the large set of species analysed, the remaining unexplained variation in OAR slopes was substantial. Nevertheless, this large-scale empirical analysis of the OAR offers an initial step towards a more general use of OARs for the fine-scale prediction of species distributions and abundance.

KEYWORDS

abundance, actual evapotranspiration, biodiversity conservation, Nachman model, occupancy–area relationship, scaling, spatial aggregation, topography, tree size

1 | INTRODUCTION

Knowledge about the spatial distributions of organisms is fundamental for both the study of biodiversity and its management, but is usually unavailable at sufficiently fine spatial detail (Hartley & Kunin, 2003; Jetz, McPherson, & Guralnick, 2012). Although ideally the spatial locations of all individuals of a species would be known, in reality the spatial grain (i.e., resolution) of species occurrence information is rather coarse, usually from tens to hundreds of square kilometres (Hurlbert & Jetz, 2007). One way to unify the coarse- and fine-grained occurrence of species is through the concept of "occupancy", usually measured as the proportion of grid cells of a given area in which a species occurs. For a landscape with nested grids of different areas, the occupancy–area relationship (OAR) describes how this proportion varies with the area of the grain at which it is measured (He & Condit, 2007; Kunin, 1998). This relationship is also known as the range–area relationship (Harte, Conlisk, Ostling, Green, & Smith, 2005), area–area curve (IUCN, 2010) or scaling pattern of occupancy (Hui et al., 2009).

The OAR is always monotonically increasing, usually in a nonlinear way (Hartley & Kunin, 2003), and the rate of its increase characterizes the spatial distribution of species across grains. Hereafter, we follow others (He & Condit, 2007) and refer to the rate of increase in occupancy with increasing grain size as an "OAR slope",

acknowledging that this is a measure of the steepness of a curve rather than a slope in the strict sense. When the grain is very small, with grid cells approximately the size of one tree, occupancy is simply the number of individuals divided by the number of all cells (Azaele, Cornell, & Kunin, 2012). Using nested grids of increasing cell area (i.e., increasing grain size), the OAR describes how individuals aggregate to occupy larger areas (Kunin, 1998). If a species is highly aggregated, it occupies few larger cells, thus occupancy increases slowly with increasing grain size. The same number of individuals scattered uniformly across cells would mean that occupancy increases rapidly with grain size (Figure 1), and a smaller number of individuals with the same degree of aggregation would exhibit an even steeper increase (He & Condit, 2007). The OAR slope thus captures information on both species abundance and spatial aggregation (Figure 1). It is also directly linked to the box-counting fractal dimension of spatial distribution (Halley et al., 2004). The OAR can thus be extrapolated to grains at which we have no data on occupancy (Azaele et al., 2012; Kunin, 1998), allowing for downscaling from coarse-grain occupancies to abundances or, conversely, upscaling from abundances to geographical occupancies.

Despite the fundamental role of OARs in connecting biological patterns across scales, accurate estimation of OARs is difficult for most species because of the scarcity of detailed information on spatial distribution (Barwell, Azaele, Kunin, & Isaac, 2014; Hui et al.,
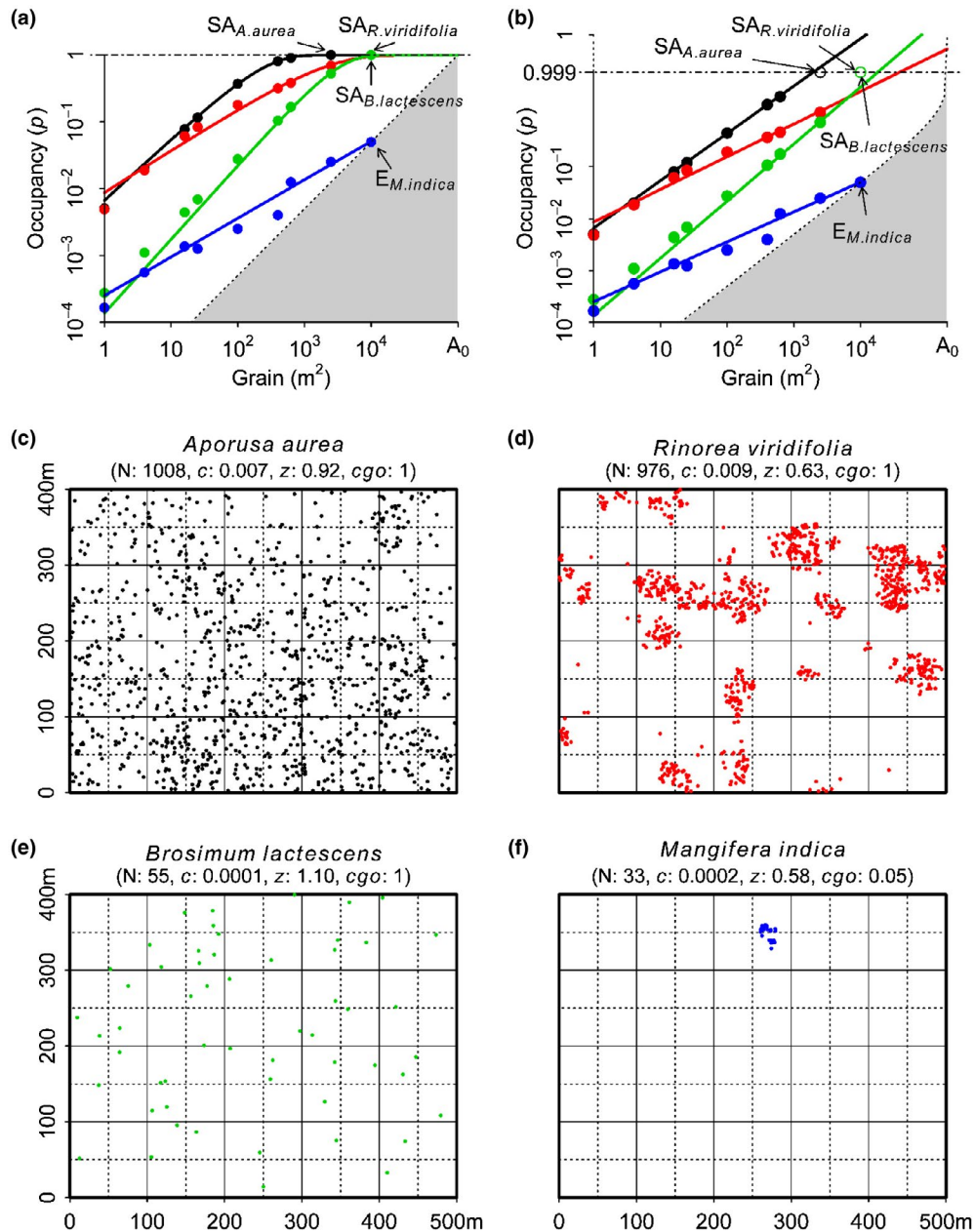
FIGURE 1 Occupancy–area relationships (OAR) for four tree species from the 20 ha (500 m × 400 m) tree plots. Fitted lines follow the Nachman model ($p = 1 - e^{-ca^z}$), where p is the proportional occupancy at a grid cell, a is area, and c and z are the parameters to fit. The value of z represents the OAR slope used for further analysis. (a) The OAR in log–log space, illustrating how saturated occupancies are fitted by the Nachman model. (b) The OAR in Nachman-transformed space {x axis: $\ln(a)$; y axis: $\ln[-\ln(1 - p)]$}; the intercept corresponds to the parameter c, and the slope to the parameter z. The filled circles are occurrences at 1 m grain (N is their total number). (c–f) Variation in OAR slopes can arise either from differences in spatial aggregation or from variation in N. The widespread species *Aporusa aurea*, *Rinorea viridifolia* and *Brosimum lactescens* all reach saturated occupancy ($p = 1$) at grains below the coarsest possible; this is the grain of saturation (SA). *Mangifera indica* is a rare species that reaches the so-called grain of endemism (E), at which all of their individuals are contained in a single grid cell. Occupancy values cannot be lower than A/A_0 (the hypotenuse of the shaded grey triangle in panels a and b), where A is the area of a grid cell at a given grain, and A_0 is the total area of the plot ($A_0 = 2 \times 10^5$ m^2), because species cannot occupy less than one grid cell at each scale. In (b), the y axis is positively infinite (dotted line), because the proportional occupancy is Nachman-transformed; we assume .999 to be the approximation of the first saturated occupancy ($p = 1$) for illustrating the grain of saturation (open circles) in the Nachman transformation; cgo is the proportional occupancy at 100 m grain (i.e., 10^4 m^2)

2009; IUCN, 2010). A generalized characterization of variation in OARs and evaluation of potential drivers of the relationship would have both basic and applied significance, including more accurate estimation of OARs for downscaling or upscaling occupancy of species. However, the potential mechanistic underpinnings and variation in OARs have remained surprisingly understudied, with empirical

demonstrations limited to a relatively small number of regions and species. Briefly, Hartley, Kunin, Lennon, and Pocock (2004) examined how OARs vary across grains in 16 British plant species, and Pocock, Hartely, Telfer, Preston, and Kunin (2006) and Barwell et al. (2014) investigated how particular habitats and species traits are correlated with OAR in rare British plants and dragonflies.

In this study, we set out to explore the variation in OARs and test key potential predictors of among-species and among-site OAR variation and to evaluate the generality of these predictors. Specifically, we sought to quantify the OAR for trees worldwide and document the consistency and predictability of this relationship among species and forests. We hypothesized that environmental (biotic and abiotic) attributes of landscapes, biological traits of species, and emergent species attributes, such as their coarse-grain occupancy (*cgo*), all affect the OAR in a predictable way (Table 1).

1.1 | Environmental drivers of OAR

Environmental conditions can affect population size and its spatial structure (Chesson, 2003; Condit, Engelbrecht, Pino, Pérez, & Turner, 2013). For example, habitats with high energy availability and more resources may support more individuals, increase local population sizes (Evans, Warren, & Gaston, 2005), and thus result in a lower rate of increase in occupancies with increasing grain size (i.e., a decrease in OAR slope). But high energy or resource availability may depress species abundance and spatial aggregation via increasing intraspecific or interspecific competition for resources, or via increasing mortality of species or decreasing their recruitment owing to natural enemies (pathogens and herbivores) (Schemske, Mittelbach, Cornell, Sobel, & Roy, 2009; Wright, 2002). Elevated energy or resource availability would, in that case, steepen the OAR slope.

TABLE 1 Hypothesized rate of increase in occupancy with grain [i.e., how the slope of the occupancy–area relationship (OAR) might be influenced by plot- and species-specific attributes]

Predictors	Rationale for hypotheses	Expected effect	Our results
Plot level			
	I. Total available resources or energy		
Energy availability	High energy supports more individuals, increases population sizes and depresses the OAR slope. Alternatively, it intensifies intraspecific competition, reduces population sizes, disperses individuals randomly and steepens the OAR slope	Steeper or shallower	Steeper
	II. Heterogeneity of abiotic conditions		
Elevation range	Large elevation difference produces patchy habitats and increases spatial aggregation of individuals, reducing the OAR slope	Shallower	Shallower
Variability of slope, convexity and elevation	High variability of topography creates more habitats in which species specialize, causing individuals to be more spatially aggregated, lowering the OAR slope	Shallower	Shallower
Fractal dimension (*F*) of terrain slope, convexity and mean elevation	If individuals track environment, the OAR slope should increase with increasing random distribution of habitats (i.e. fractal dimension, *F*)	Steeper	Steeper, except for *F* of terrain slope
	III. Biotic conditions		
Species richness	Species richness is positively correlated with intraspecific competition, the increase of which decreases conspecific abundance and makes the species more dispersed, which would steepen the OAR slope	Steeper	Steeper
Mean number of individuals per species	A larger number of individuals means higher occupancy at fine grains and thus shallower OAR	Shallower	Shallower
Species level			
Tree size	Large-statured tree species out-compete small ones in use of energy or resources, and thus are more abundant, which depresses the OAR slope. On the contrary, large trees facilitate seed dispersal, making individuals more dispersed, steepening the OAR slope.	Shallower or steeper	Shallower, with levelling off with increasing tree size
Coarse-grain occupancy	Spatially rarer species (i.e., those with small coarse-grain occupancy) are increasingly unlikely to have shallow OAR, owing to a hard boundary on minimum occupancy at the finest grain measured, whereas spatially common species (i.e., those with large coarse-grain occupancy) can have either small or large fine-gain occupancy, thus either steep or shallow OAR	Shallower, with levelling off as coarse-grain occupancy increases	Shallower, with levelling off with increasing coarse-grain occupancy

In addition to the total available energy and resources, their spatial variability and structure (arrangement) can influence the distribution of individuals on the landscape, because most species have specific environmental requirements (Condit et al., 2013; Harms, Condit, Hubbell, & Foster, 2001). Environmental conditions are often spatially autocorrelated or structured (Legendre, 1993). High spatial variability or dispersion of environmental conditions can translate into a greater diversity of habitats, thus increasing the probability of species–habitat association and spatial aggregation of species, thus weakening the increase of occupancies with grains. In turn, spatial structure of environments defines how they are arranged in space; patchy or spatially structured environments can reinforce the spatial aggregation of species, thus depressing the increase of occupancies with grains.

Detailed maps of fine-grain energy and resources are rarely available, but it is possible to use topographic factors, such as elevation, convexity and slope, as a comprehensive proxy for habitats (Cáceres et al., 2012; Legendre et al., 2009). Large elevation difference generates heterogeneity of environmental conditions, increasing spatial aggregation of species (Chesson, 2003; Davies et al., 2007), thus depressing the OAR slope.

The species richness is often higher in areas with high energy availability and environmental heterogeneity, both of which are seen as facilitators of coexistence (Evans et al., 2005; Stein, Gerstner, & Kreft, 2014). But beyond these broad-scale richness–energy availability associations, globally, conspecific negative density dependence is positively correlated with species richness (Johnson, Beaulieu, Bever, & Clay, 2012; LaManna et al., 2017). Given that the density dependence can reduce both population size and its spatial aggregation, species richness may be positively correlated with OAR slope. We expect the potential effect of species richness on the occupancy–area slope to be equivocal and dependent on the prevalence and directionality of niche segregation and density dependence processes. Given that a smaller number of individuals often corresponds to a steeper OAR slope (He & Condit, 2007), OAR slopes are expected to decrease with the mean number of individuals per species.

1.2 | Traits as drivers of OAR

Traits are known to influence species abundances and spatial distributions (Rüger, Wirth, Wright, & Condit, 2012; Westoby & Wright, 2006) and thus by extension might also affect the scale dependence of their occurrence. For example, tree size at maturity reflects both the strategy of species to compete for light, water and nutrients and also their ability to disperse (Díaz et al., 2016; Westoby, 1998). We hypothesize that because large-sized trees tend to disperse farther (Thomson, Moles, Auld, & Kingsford, 2011), they might show a more regular spatial distribution, resulting in a steeper occupancy–area relationship. Alternatively, a large-sized tree might more readily outcompete smaller ones for resources at small scales (Lutz et al., 2018), leading the large-sized tree species to be more abundant and have a shallower increase of occupancy with area.

Coarse-grain occupancy (i.e., the proportion of cells occupied at coarser spatial scale) is information that is often more readily gathered than fine-scale occurrence. For a given analysis grid (e.g., of 100 m × 100 m cells in the case of tree plots; variable cgo in Figure 1), it represents the maximum occupancy attained by a species. It distinguishes species that are spatially rare (those occurring in only a small portion of the landscape, $cgo \ll 1$) from those that are common (e.g., occurring in all cells, $cgo = 1$). Coarse-grain occupancy thus emerges as a joint species property, or emergent attribute, from both species traits and the local environment. Spatially common species ($cgo \approx 1$) can have either steep or shallow OARs, but with increasing spatial rarity of species ($cgo < 1$) shallow relationships become increasingly unlikely, because of a hard boundary on minimum occupancy owing to the finest spatial grain measured (see Figure 1). Given that the occurrence of species is much more readily surveyed and known at coarser spatial grains, the information value of coarse-grain information for predicting fine-grain occupancy has practical significance but has seen only limited evaluation over large spatial scales.

In this study, we quantify OARs for a large dataset of 3,157 tree species based on 17 fully mapped forest plots of 20 ha worldwide. We use this compilation to ask the following questions:

1. How much does the OAR vary among species and among plots?
2. Do OAR slopes vary as expected with key species traits and environmental conditions?
3. How large is the predictive ability of these environment and trait associations, and do they offer scope for the prediction of OAR for new species or locations?

2 | METHODS

We selected 17 forest plots ≥ 20 ha from the Smithsonian Forest Global Earth Observatory (ForestGEO; https://forestgeo.si.edu/) and the Chinese Forest Biodiversity Monitoring Network (CForBio; http://www.cfbiodiv.org/english) (Figure 1). Each of the plots is in relatively undisturbed forest. For each plot, we selected a 20 ha (400 m × 500 m) subplot (starting at the coordinate origin of the plot), except at the Yosemite and Wind River plots, where we used a 21 ha subplot, 300 × 700 m. Topography differs greatly among plots: the elevation difference within individual plots is as low as 17 m and as high as 299 m. Plot censuses follow the centre for tropical forest science (CTFS) protocol (Condit, 1998). All individuals ≥ 1 cm diameter at breast height (d.b.h.) were measured, mapped and identified taxonomically. To unify taxonomy among plots, we used the Taxonomic Name Resolution Service v.3.0 (http://tnrs.iplantcollaborative.org) and the Chinese Virtual Herbarium (www.cvh.ac.cn). We excluded 289 unidentified species: 107 from the Banna and Korup plots (leaving 358 and 316 identified species, respectively), 57 from the Pasoh plot (723 identified), 11 from the Yasuni plot (1,015 identified), four from the Lienhuachih plot (137 identified) and one from each of the

Baotianman, Dongling and Yosemite plots. Most (70%) of unidentified species were rare, with occurrence of < 1/ha. We also excluded species that were represented by only one stem. After pruning, the resulting dataset for this study had 3,157 unique species, with 3,546 different OARs owing to the species that occurred in multiple plots.

2.1 | Occupancy–area relationship

In each plot, we examined the occupancy of each species in grids of square cells having sides of 1, 2, 4, 5, 10, 20, 25, 50 and 100 m. To characterize the OAR, we quantified occupancy as the proportion (p) of occupied cells and related it to the area (α) of the grid cell, or the grain. The slope (z) of OAR quantifies how quickly occupancy aggregates from fine to coarse grain and defines the shape of OAR.

To choose the most appropriate function for the OAR, we performed a preliminary analysis to simulate and assess three commonly fitted models: Nachman, power-law and logistic models (Azaele et al., 2012; see Supporting Information Appendix S1). Analytically, and based on a simulation of complete spatial randomness (CSR; see Supporting Information Appendix S1, Fig. S1.1), we found that the slopes (z) of the power-law and logistic models depended on grain and population density. Only the Nachman model consistently showed a slope of one under CSR. Based on the simulation of extreme spatial aggregation (ESA; see Supporting Information Appendix S1, Fig. S1.1), we found that the slope (z) of all the three models decreased with increasing abundance, and approached one as abundances decreased. Those simulations indicated negative effects of both species abundances and spatial aggregation on OAR slope with strong interaction: OAR slopes increasingly decreased with species abundances as spatial aggregation increased, and vice versa (see Supporting Information Appendix S1, Fig. S1.2). Therefore, only the Nachman model clearly defined aggregation, with the slope $z < 1$ indicating spatial aggregation. We therefore used only the Nachman model in further analyses.

The Nachman model is often fitted using linear regression on transformed scale–area data, {i.e., log[−log(1 − p)] against log (α)}. However, this transformation has been shown to give too much weight to the occupancies at small scales (Packard, Birchard, & Boardman, 2011), and it fails as the occupancy approaches one. Nonlinear regression, in contrast, allows an asymptote of one when $c\alpha^z \gg 1$ (where c is a parameter to fit). We thus adopted nonlinear regression to fit the Nachman model over the grains of 1, 2, 4, 5, 10, 20, 25, 50 and 100 m (Figure 1). Furthermore, we used slopes (z values) emerging from this fitting as the response variable to characterize the variation in slope among species and among plots and to explore their ecological predictors.

2.2 | Predictors of occupancy–area slopes

We considered variables correlating with z, the occupancy–area "slope", at two levels: species and plots. At the plot level, we used three groups of environmental predictors to characterize the differences in abiotic and biotic conditions (Table 1). I. Energy and resource

availability. We used ambient energy, measured as actual evapotranspiration extracted from the MODIS (Moderate Resolution Imaging Spectroradiometer) MOD 16A3 product that provides average annual mean evapotranspiration for the years 2000–2013 at 1 km resolution (Mu, Zhao, & Running, 2011), which captures both heat and water essential for plants. II. Heterogeneity of abiotic conditions. We assessed a total of three variables: (a) elevation range, the maximal difference of elevation that the plot spans, measuring the topographic complexity of each plot; (b) spatial structure (arrangement) of three topographic variables (slope, mean elevation and convexity), quantified by fractal dimension of the variables based on variograms in each plot, with high values meaning more random distribution (Constantine & Hall, 1994); and (c) spatial variability of the three topographic variables, defined as the standard deviation of the variables among grid cells in each plot. All the three variables were calculated for each 20 m × 20 m cell in each plot. The mean elevation of a cell was defined as the mean of the elevation values at its four corners. Slope was the mean angular deviation from horizontal of each of the four triangular planes formed by connecting three of its corners. Convexity was the elevation of the cell of interest minus the mean elevation of the eight surrounding cells. For the edge cells, convexity was the elevation of the centre point minus the mean of the four corners. III. Biotic conditions. Here, we evaluated: (a) species richness, counted as the total number of woody species in each plot; and (b) the mean number of individuals per species, calculated as the sum of the individual number for all species in a plot divided by the total number of species.

As species-level predictors (traits) we included tree size, measured as the maximal diameter at breast height for a species in each plot, and coarse-grain occupancy, measured as the occupancy at a grain of 100 m × 100 m.

2.3 | Statistical models

We used linear mixed-effect models (Pinheiro & Bates, 2000) to evaluate the effect of plot conditions and species attributes as predictors on the slopes (z) of the OARs. Both plot and species were assigned random intercepts, and we assumed that those random intercepts were independent of each other. Furthermore, random slopes by plot were added to the species-specific predictors of interest.

First, we started with a null model for the OAR slope (z), having no predictors, only plot and species random intercepts. Second, to test the effect (including significance and direction) of plot-level variables on OARs, we examined models with all the possible combinations of 10 plot-level predictors and selected those with each predictor having a variance inflation factor (VIF) less than two and $p \le .05$ (i.e., all predictors involved in the models are significant and independent of each other). Third, to test the effect of species-specific predictors on OARs, we built three separate types of models for the two species-level predictors: two with a single species-specific predictor, and the other combining tree diameter and coarse-grain occupancy as predictors. Fourth, we also constructed a model

combining both sets of predictors. For each model involving species-level variables, we examined all the possible combinations of the fixed- and random-slope terms of interest (the terms assigned random slope must also be represented as fixed effects) and selected the one with the smallest corrected Akaike information criterion (AICc; see Supporting Information Appendix S2).

Model residuals using raw values of slope z were negatively skewed (Figure 2); therefore, we instead used the transformation –log(θ – z), where θ is a parameter that we adjusted to get the skewness of residuals as close to zero as possible. Tree diameter was also highly skewed; therefore, we first $\log_{10}$-transformed it. All the variables were then standardized to zero mean and variance of one. We used second-order polynomials to model the potentially nonlinear

effect of tree diameter and *cgo*. We set the VIF of two as the threshold for testing collinearity among predictors (Zuur, Ieno, & Elphick, 2010), and no models exceeded that threshold.

All mixed-effect models and model selection were run using the R package "lmerTest" v.2.0-29 (Kuznetsova, Brockhoff & Christensen, 2015) and "MuMIn" v.1.14.0 (Bartoń, 2015), and we visualized the effects of individual predictors on OAR slope using the R package "visreg" v.2.2-0 (Breheny & Burchett, 2015), and calculated the fractal dimension of topographic variables using R function "RFfractaldim" (Schlather et al., 2017). Effect sizes of all predictors were measured by: (a) marginal coefficients of determination (R^2_m), the proportion of total variance explained by the fixed effects alone; and (b) the so-called partial conditional coefficients of determination R^2_{pc}, the
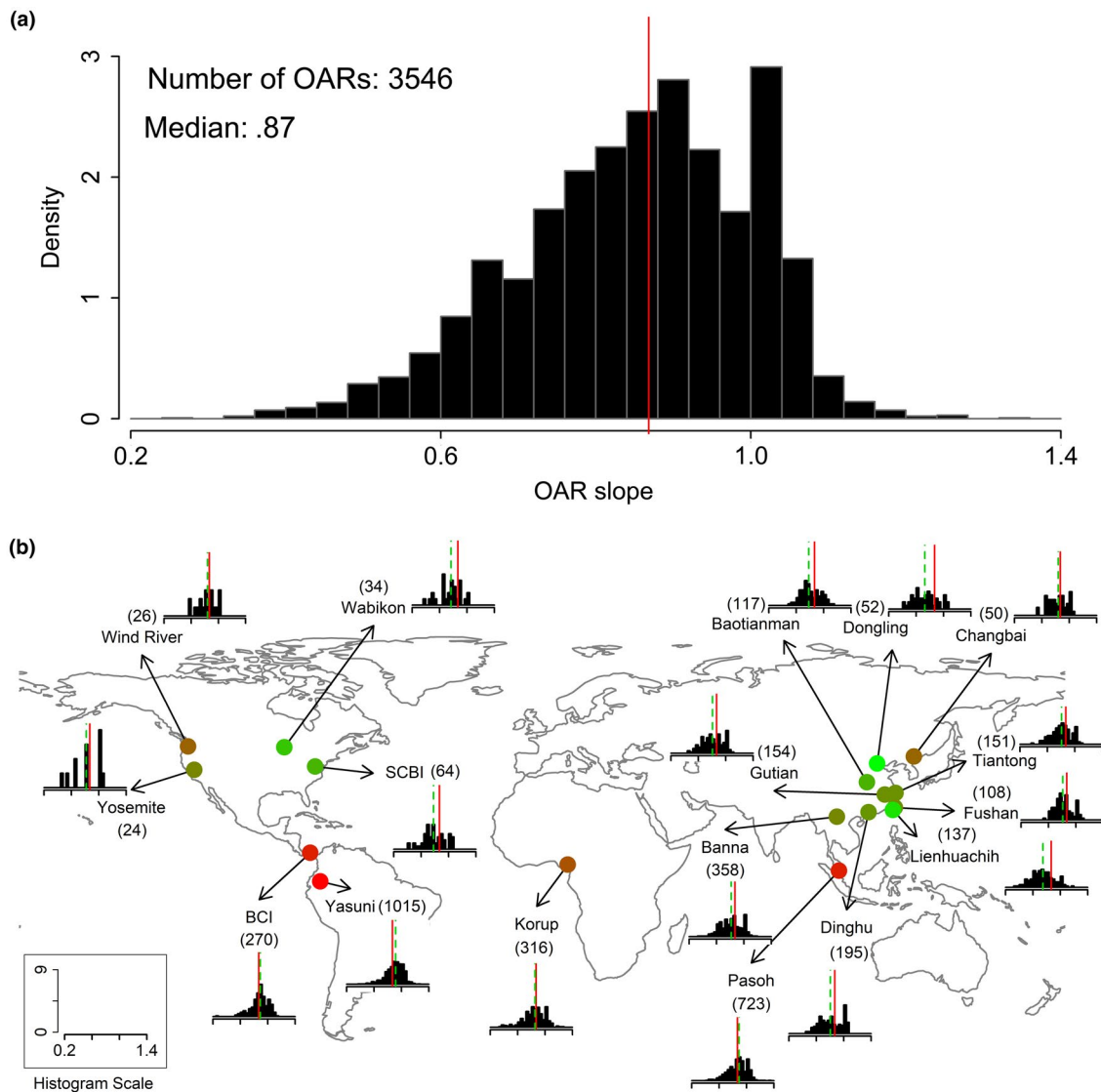

FIGURE 2 Locations of study forest plots worldwide and variations in the slope (z) of the occupancy–area relationship (OAR slope) among species and across plots. (a) Overall frequency distribution of the slopes from 17 forest plots worldwide (3,546 OARs for 3,157 unique species). (b) Geographical locations and frequency distributions of the slopes and number of species (in parenthesis) per plot. Red vertical lines show the overall median of the slope across all plots; dashed green vertical lines are medians of the slopes in individual plots. The medians of individual plots are also represented by the point colours from low (green) to high (red) values. The y axes for frequency distribution are all in density. The bottom left inset shows the axis scales for all OAR slope frequency distributions shown in (b)
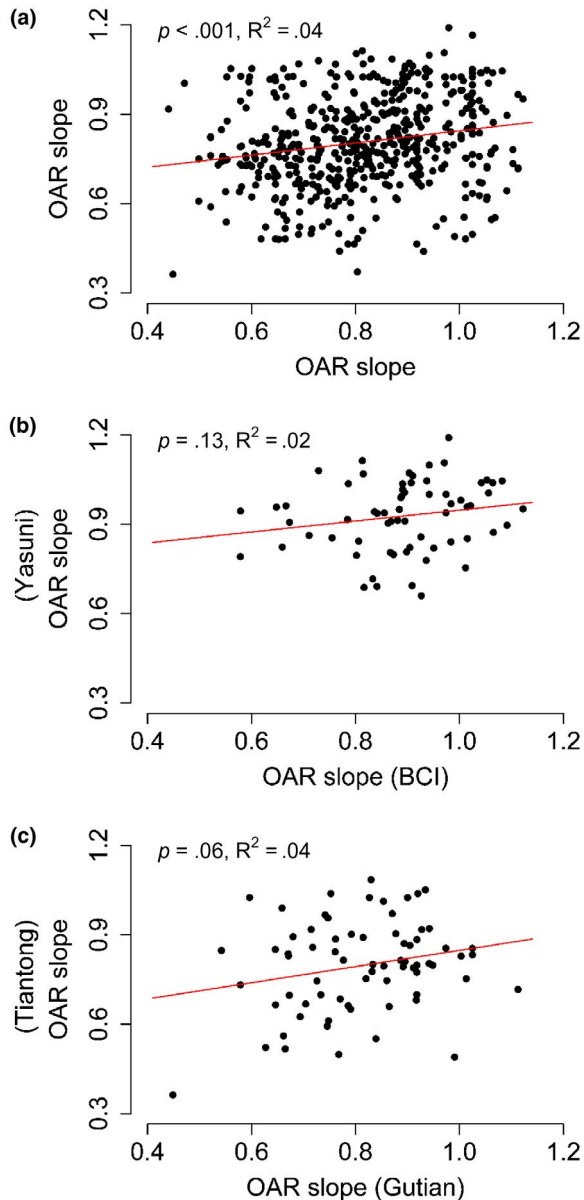
FIGURE 3 The pairwise relationships of the occupancy–area slope (*z*) for all the species shared in multiple forest plots (a), for the species occurring in the Barro Colorado Island (BCI) and Yasuni plots (b) and for the species common in the Gutian and Tiantong plots (c) [Colour figure can be viewed at wileyonlinelibrary.com]

proportion of total variance accounted for by both the fixed and random slope effects of variables of interest. The R^2_{pc} is calculated by subtracting the random intercept effects from the conditional coefficient of determination (the total variance explained by both, all fixed and all random factors; Johnson, 2014; Nakagawa & Schielzeth, 2013). The fixed effect is the mean effect of species- or plot-specific predictors over the entire set of species or plots, whereas the random slope effect represents the species-specific deviation from the mean. For R scripts of our main analyses, see the Supporting Information (Appendix S2).

3 | RESULTS

3.1 | Variation in OARs among species and plots worldwide

The OAR, measured by its slope (*z*; i.e., rate of increase in occupancy with grains), showed strong variation among species, ranging from the shallowest increase in occupancy observed in *Anaxagorea panamensis* (Family: Annonaceae) on the Barro Colorado Island plot (*z* = .248) to *Aquilaria malaccensis* (Family: Thymelaeaceae) at Pasoh with the steepest slope (*z* = 1.331) (Figure 2a). The largest proportion of the variance of OAR slope was from within-group (plot and species) error (i.e., the residuals in the null model), at 77.5%. Next was the variance between the groups by species (i.e., the random species intercepts), at 14.9%. The between-plot variance (i.e., the random plot intercepts) was only 7.6%. The OAR slope varied geographically (Figure 2b), with its median value per plot increasing towards the equator (*r* = −.68, *p* = .003) from .73 at latitude 40° N (the Dongling plot) to .92 at latitude .69° S (the Yasuni plot) (Supporting Information Table S1). Plot-level skewness of the OAR slopes changed from negative to positive from the equator to higher latitudes (range: −.89 to .26, *r* = .50, *p* = .04). There were 302 (out of 3,157) species occurring in multiple plots. The OAR slope had a very weak correlation between plots for the same species [Figure 3a; i.e., the slope varied greatly among plots within the same species, even for the species shared in the plots close to each other (Figure 3b,c)].

3.2 | Species traits and plot-level environment as predictors of the OAR

We found significant associations of the observed OAR slopes with species- and plot-level predictors (Table 2), but a large amount (88%; see model 19 in Table 2) of the slope variance remained unexplained, highlighting the importance of local or stochastic processes. The OAR slope declined significantly and levelled off slightly with increasing coarse-grain occupancy, accounting for 10.7% of the variance in OAR slope, of which 10.5% was within plots (model 17 in Table 2; Figure 4a). Adult tree size also had a nonlinear effect on the OAR slope, but weaker than coarse-grain occupancy: it explained 8.9% of the variance of the slope, and 6.1% when examined within plots (model 16 in Table 2). Smaller species tended to have steeper slopes, but this diameter–slope association leveled off towards large diameters (Table 2; Figure 4b). These two traits together explained 13.8% of total variance in the OAR slope, and 11% when examined within plots (model 18 in Table 2).

Plot-level attributes significantly affected the OAR slope and explained a small portion of total variance in OAR slope but a majority of between-plot variance (models 1–15 in Table 2). Energy availability had a positive effect and heterogeneity of abiotic conditions generally a negative effect on the OAR slope (Table 2). The OAR slope tended to be shallower in areas with higher variability of topographic factors, especially convexity and mean elevation,

TABLE 2 The results of linear mixed-effect models of occupancy–area relationship (OAR) slope (z) fitted over the grains of 1–100 m for all 3,546 OARs, against plot- and species-level variables

| Model No | Plot Level | | | | | | | | | | | | Species Level | | | | | R_m^2(%) | R_{pc}^2(%) | AICc |
| | Intc | I. Energy availability | II. Heterogeneity of abiotic condition | | | | | | | III. Biotic condition | | Tree size | | Coarse-grain occupancy | | | | | |
		ET	ER	SCN	SME	SSL	FCN	FME	FSL	MID	S	ts	ts²	cgo	cgo²			
1	***	*			-*											3.1		−3198
2	***	*	-*													3.2		−3198
3	***				-*					-**						3.6		−3199
4	***	-*	-*							-**						3.7		−3200
5	***						*		-**							3.8		−3201
6	***	*		-*						-**						3.6		−3201
7	***	**		-**												3.7		−3202
8	***								-**	-**						3.9		−3203
9	***			-*												1.6		−3204
10	***						*		-**		**					7.5		−3204
11	***									-*						1.2		−3205
12	***			-*					-*							2.4		−3206
13	***			-*							**					7.7		−3207
14	***								-**		**					7.4		−3209
15	***										**					7.1		−3211
16	•											-**	***				8.9	−3381
17														-***	***	10.5	10.7	−3485
18												-*	***	***	***	11.0	13.8	−3519
19	*										***	-*	***	-***	***	12.0	15.3	−3517

Note. Models 1–15 test for the effects on OAR slope of plot-level variables, models 16–18 for the effects of species-level variables, and model 19 examines the power of both sets to predict OAR slope. For confidence intervals of coefficients for each predictor, see the Supporting Information (Table S2).

Abbreviations: AICc = second-order Akaike information criterion; cgo = coarse-grain occupancy (proportional occupancy at the grain of 100 × 100 m); ER = elevational range; ET = actual evapotranspiration; FCN = fractal dimension of convexity; FME = fractal dimension of mean elevation; FSL = fractal dimension of slope; MID = mean number of individuals per species; R_m^2 = a marginal coefficient of determination, the proportion of total variance explained by the fixed effects alone (i.e., variance explained within plots for species-specific variables); R_{pc}^2 = the proportion of total variance explained by the fixed effects and random slope effects (excluding random intercept effects), following (Johnson, 2014; Nakagawa & Schielzeth, 2013); S = species richness; SCN = standard deviation of convexity; SME = standard deviation of mean elevation; SSL = standard deviation of slope; ts = tree size (the maximum diameter at breast height). The significance code "(***)" represents significance at level .001, "(**)" at level .01, "(*)" at level .05, and "(•)" at level .1. The sign "−" before the significance code represents a negative effect on OAR slope; no sign represents a positive effect.
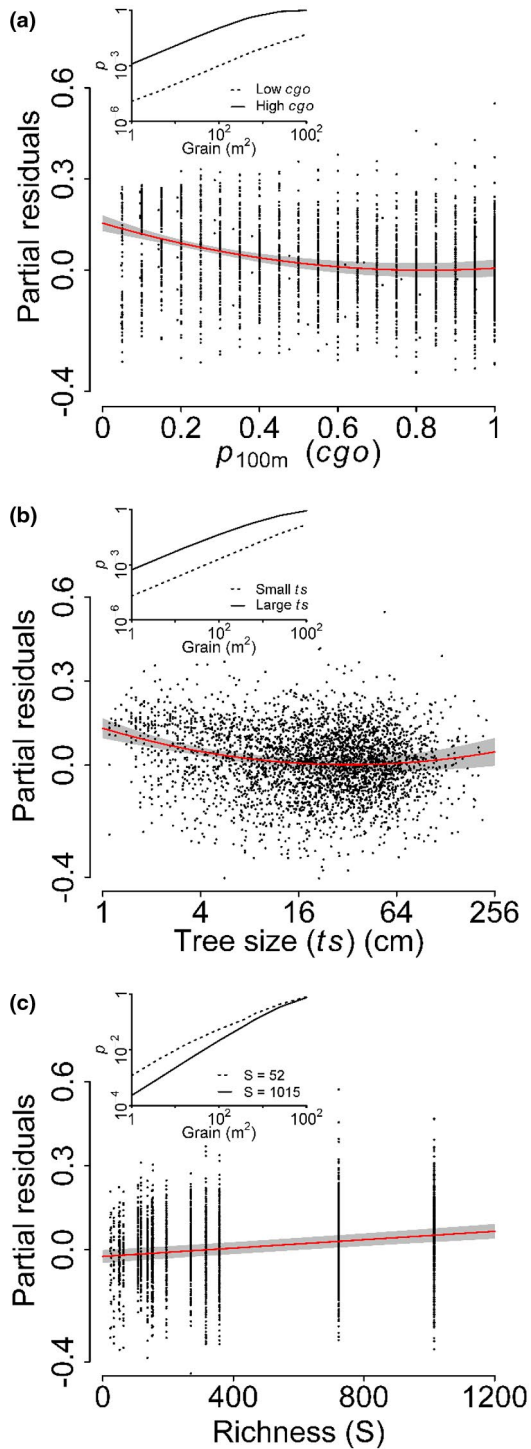
FIGURE 4 Partial residual effect on occupancy–area slope (z) of coarse-grain occupancy (cgo) at the grain of 100 m × 100 m (a), tree size (b) and plot species richness (c), each conditional on the other two predictors, with 95% confidence intervals (grey shade), based on model 19 in Table 2. The insets illustrate the differences in occupancy–area relationship in extreme conditions of explanatory variables. The low species richness in the inset is for the Dongling plot and the high one for the Yasuni plot. The ranges for the other two variables in the insets are subsets of the entire dataset in which their values are smaller than their 2.5% quantiles and larger than their 97.5% quantiles. The y axes of the insets refer to median occupancy (p) for relevant subsets at each grain [Colour figure can be viewed at wileyonlinelibrary.com]

although the slope variability was not significant. The OAR slope was also affected by the spatial structure of environmental conditions and was steeper when convexity patterns showed a more random distribution (higher fractal dimension). However, the spatial structure of topographic slope showed the opposite effect. The fractal dimension of topographic slope was positively correlated with the variability of itself ($r = .67$, $p < .001$) and convexity ($r = .69$, $p < .001$). The OAR slope tended to be steeper in the plots with more species and with a lower mean number of individuals per species. Species richness, which was positively correlated with energy availability ($r = .61$, $p = .01$) and negatively with the mean number of individuals per species ($r = −.60$, $p = .01$), accounted for 7.1% of total variance in the OAR slope but much more than other plot-level factors (model 15 in Table 2).

4 | DISCUSSION

For thousands of tree species across forests worldwide, we identified a pronounced variation in the shape of the OAR and thus the scaling from coarse- to fine-grain occupancy. Although between-plot differences in the frequency distribution of OAR slopes existed, and although we still identified a significant latitudinal gradient in OAR slopes, the variation occurred predominantly at the species level. The OAR slopes varied most strongly among species within plots rather than among plots worldwide. We found that both species traits and plot-specific environmental conditions were able to explain some of the variation in OARs, suggesting that there is scope for generalizing the OAR beyond local measurements. But again, the predictive power of these covariates was limited.

In contrast to the work on the OAR to date, which has focused on single regions, our study examined the relationship over a global (if incompletely filled) extent. Earlier work has shown that the OAR slope varies across spatial grains (Hartley et al., 2004), whereas we found a distinct variation with spatial extent. We found that 302 (out of 3,157) species that occurred in multiple plots in different regions varied greatly in their OAR slope, and there was limited consistency among plots (Figure 3). This highlights the importance of local adaptation in addition to stochastic or unmeasured drivers of local OARs. This region and plot dependence of the OAR slopes also highlights the importance of extent in OAR assessments, because the OAR slope might change substantially with an increasing extent of study. The capture of regional or site-associated species traits (e.g., in tree diameter) might be crucial, rather than mean species traits, to achieve strong predictive fits.

Our predictions (Table 1) regarding the significance and directionality of putative drivers of the OAR were largely confirmed. Large tree size was associated with the traits favouring seed dispersal, such as height of seed release (Díaz et al., 2016), and with seeds of a large size, which larger, more wide-ranging animals disperse more favourably (Seidler & Plotkin, 2006). We found somewhat steeper OAR slopes in smaller-diameter trees, supporting the

alternative proposal of larger-diameter trees potentially being better competitors at smaller scales. But this decline in OAR slopes with tree diameter was nonlinear, and medium to large trees showed weak associations with slope.

Coarse-grain occupancy had a negative effect on OAR slope, confirming our expectations and previous reports (Conlisk, Conlisk, Enquist, Thompson, & Harte, 2009; Wilson, Thomas, Fox, Roy, & Kunin, 2004). The well-known fact that species with large populations tend to have wide-ranging occupancy (Borregaard & Rahbek, 2010) underpins the negative effect. In contrast, most randomly distributed species have low abundances (He, Legendre, & LaFrankie, 1997). These rare species have high large-grain occupancy and low fine-grain occupancy, and thus rapidly increase in occupancy with grains, which might depress the decrease in OAR slope with coarse-grain occupancy (i.e., levelling off the relationship). Coarse-grain occupancy alone provided a reasonably strong prediction of the OAR, thus highlighting the potential, even over the large extents analysed, for limited coarser-grain survey data to help inform fine-scale occupancy.

The relatively limited variation in OAR slope between plots worldwide (compared with among species) implies that plot-wide attributes are less important in predicting the OAR slope. This is because the spatial distribution of species, which determines OARs, depends on not only environmental conditions but also on species-specific adaptation and sensitivity to them by its fundamental niche and traits (McGill, Enquist, Weiher, & Westoby, 2006). Even so, biotic and abiotic variables at the plot level significantly shaped OARs. Species richness explained most of the between-plot variation in OAR slope. This might be attributable to the strong association between the richness and the abundance structure and fine-grain spatial aggregation of species in communities (He & Legendre, 2002). Mean abundance per species had a negative effect on the OAR slope, which became insignificant after controlling for species richness. This result supports the hypothesis that OAR slope declines as the mean abundance increases, and reflects the close correlation between the mean abundance and species richness (He & Legendre, 2002).

The increasing trend of OAR slope towards the tropics can be attributed to species richness and energy availability, both of which decrease with latitude (Evans et al., 2005). We found that energy availability had a weak but positive effect on the occupancy–area slope. This could be the result of intraspecific competition positively correlated with energy (LaManna et al., 2017), which decreases mean abundance per species and fine-grain spatial aggregation (He & Legendre, 2002). However, species richness and energy availability are strongly correlated. It is thus an open question why species richness explained most of the among-plot variation in occupancy–area slope, whereas available energy explained very little.

We found that elevation range and variability of topographic variables within plots exhibited negative effects on the OAR slope, and spatially structured topographic factors suppressed the OAR slope in general. These results support the conclusion that rugged topography creates patchier habitats and thus more spatial aggregation of

species at the fine grains (Cáceres et al., 2012; Legendre et al., 2009). In contrast, spatially structured terrain slope steepened the increase in occupancy with grains. This is because the fractal dimension of terrain slope was positively correlated with the variability of topographic slope and convexity; that is, a spatially structured topographic slope implies simple terrain and a lack of habitat heterogeneity.

With increasing species spatial and trait data there is a potentially exciting possibility of a model-based prediction of fine-scale occurrences of species, or even their abundances, over large extents (Jetz et al., 2016; Kattge et al., 2011). A more generalized and predictable model of the OAR might have the potential to support the threat status assessment for species (Azaele et al., 2012). Here, we have shown that the coarser-grain survey data combined with species trait data and environmental information indeed have the potential to improve our OAR predictions for thousands of species. However, although statistically significant, the detected effects and predictive ability attained remained limited. Obviously, the large geographical dispersion of plots used in the present study (at least, outside China) represented a highly challenging test for the statistical generalizability of OARs. We measured OARs at the plot level. When extending spatial extent and grains, therefore, the grain dependence of OARs (Hartley et al., 2004) might imply a shift in the drivers of OARs with grains. For future work that is able to draw on a larger set of sites, we suggest considering spatial distance as an additional covariate and extending grains, together with the exploration of more regionally parameterized OAR models and richer trait data.

4.1 | Conclusions

By synthesizing worldwide forest plot data, our study built on a growing array of fine-scale occurrence data that allow the estimation of OARs. We asked whether species traits or environmental drivers predictably affected observed OARs and whether these effects were sufficiently strong to enable a robust prediction of unobserved OARs and fine-scale occupancies for different regions or even species. Our analyses uncovered increasing OAR steepness towards the tropics and confirmed key predictors of the OAR. But over a global extent and over 3,500 tree species, the variation in OAR slopes is substantial and not fully captured by traits or environmental conditions. Richer regional data together with remotely sensed information and richer trait data might soon offer additional opportunities for a more generalized, cross-scale prediction of species occupancy, with many uses in ecology and conservation.

DATA ACCESSIBILITY

Data for plots in the ForestGEO network are available via the online portal at: https://www.forestgeo.si.edu. The Smithsonian ForestGEO network participants and the CForBio participants are committed to making the plot data widely available. Data are always available to anyone for the purposes of checking or recreating any published analysis in _Global Ecology and Biogeography_ or any other publication. Data are available for new and different analyses based on the plot-specific data access conditions listed at the Smithsonian ForestGEO portal (https://www.forestgeo.si.edu).

ORCID

Haibao Ren (iD) https://orcid.org/0000-0002-8955-301X

Petr Keil (iD) https://orcid.org/0000-0003-3017-1858

Keping Ma (iD) https://orcid.org/0000-0001-9112-5340

James Lutz (iD) https://orcid.org/0000-0002-2560-0710

REFERENCES

Azaele, S., Cornell, S. J., & Kunin, W. E. (2012). Downscaling species occupancy from coarse spatial scales. _Ecological Applications_, _22_, 1004–1014. https://doi.org/10.1890/11-0536.1

Bartoń, K. (2015). _MuMIn: Multi-model inference. R package version 1.14.0_. Rerived from https://CRAN.Rproject.org/package=MuMIn

Barwell, L. J., Azaele, S., Kunin, W. E., & Isaac, N. J. B. (2014). Can coarse-grain patterns in insect atlas data predict local occupancy? _Diversity and Distributions_, _20_, 895–907.

Borregaard, M. K., & Rahbek, C. (2010). Causality of the relationship between geographic distribution and species abundance. _The Quarterly Review of Biology_, _85_, 3–25. https://doi.org/10.1086/650265

Breheny, P., & Burchett, W. (2015). _visreg: Visualization of regression models. R package version 2.2-0_. Rerived from https://CRAN.R-project.org/package=visreg

Chesson, P. (2003). Understanding the role of environmental variation in population and community dynamics – Introduction. _Theoretical Population Biology_, _64_, 253–254. https://doi.org/10.1016/j.tpb.2003.06.002

Condit, R. S. (1998). _Tropical forest census plots: Methods and results from Barro Colorado Island, Panama and a comparison with other plots_. Berlin: Springer-Verlag.

Condit, R., Engelbrecht, B. M. J., Pino, D., Pérez, R., & Turner, B. L. (2013). Species distributions in response to individual soil nutrients and seasonal drought across a community of tropical trees. _Proceedings of the National Academy of Sciences USA_, _110_, 5064–5068. https://doi.org/10.1073/pnas.1218042110

Conlisk, E., Conlisk, J., Enquist, B., Thompson, J., & Harte, J. (2009). Improved abundance prediction from presence–absence data. _Global Ecology and Biogeography_, _18_, 1–10. https://doi.org/10.1111/j.1466-8238.2008.00427.x

Constantine, A. G., & Hall, P. (1994). Characterizing surface smoothness via estimation of effective fractal dimension. _Journal of the Royal Statistical Society. Series B_, _56_, 97–113.

Davies, R. G., Orme, C. D. L., Storch, D., Olson, V. A., Thomas, G. H., Ross, S. G., … Gaston, K. J. (2007). Topography, energy and the global distribution of bird species richness. _Proceeding of the Royal Society B: Biological Sciences_, _274_, 1189–1197. https://doi.org/10.1098/rspb.2006.0061

De Cáceres, M., Legendre, P., Valencia, R., Cao, M., Chang, L.-W., Chuyong, G., … He, F. (2012). The variation of tree beta diversity across a global network of forest plots. _Global Ecology and Biogeography_, _21_, 1191–1202. https://doi.org/10.1111/j.1466-8238.2012.00770.x

Díaz, S., Kattge, J., Cornelissen, J. H. C., Wright, I. J., Lavorel, S., Dray, S., … Gorné, L. D. (2016). The global spectrum of plant form and function. _Nature_, _529_, 167–171. https://doi.org/10.1038/nature16489

Evans, K. L., Warren, P. H., & Gaston, K. J. (2005). Species-energy relationships at the macroecological scale: A review of the mechanisms. _Biological Review_, _80_, 1–25. https://doi.org/10.1017/S1464793104006517

Halley, J. M., Hartley, S., Kallimanis, A. S., Kunin, W. E., Lennon, J. J., & Sgardelis, S. P. (2004). Uses and abuses of fractal methodology in ecology. _Ecology Letters_, _7_, 254–271. https://doi.org/10.1111/j.1461-0248.2004.00568.x

Harms, K. E., Condit, R., Hubbell, S. P., & Foster, R. B. (2001). Habitat associations of trees and shrubs in a 50-ha Neotropical forest plot. _Journal of Ecology_, _89_, 947–959. https://doi.org/10.1111/j.1365-2745.2001.00615.x

Harte, J., Conlisk, E., Ostling, A., Green, J. L., & Smith, A. B. (2005). A theory of spatial structure in ecological communities at multiple spatial scales. _Ecological Monographs_, _75_, 179–197. https://doi.org/10.1890/04-1388

Hartley, S., & Kunin, W. E. (2003). Scale dependency of rarity, extinction risk, and conservation priority. _Conservation Biology_, _17_, 1559–1570. https://doi.org/10.1111/j.1523-1739.2003.00015.x

Hartley, S., Kunin, W. E., Lennon, J. J., & Pocock, M. J. O. (2004). Coherence and discontinuity in the scaling of species' distribution patterns. _Proceedings of the Royal Society B: Biological Sciences_, _271_, 81–88.

He, F., & Condit, R. (2007). The distribution of species: Occupancy, scale, and rarity. In D. Storch , P. A. Marquet , & J. H. Brown (Eds.), _Scaling biodiversity_ (pp. 32–50). Cambridge: Cambridge University Press.

He, F., & Legendre, P. (2002). Species diversity patterns derived from species–area models. _Ecology_, _83_, 1185–1198.

He, F., Legendre, P., & LaFrankie, J. V. (1997). Distribution patterns of tree species in a Malaysian tropical rain forest. _Journal of Vegetation Science_, _8_, 105–114. https://doi.org/10.2307/3237248

Hui, C., McGeoch, M. A., Reyers, B., le Roux, P. C., Greve, M., & Chown, S. L. (2009). Extrapolating population size from the occupancy–abundance relationship and the scaling pattern of occupancy. _Ecological Applications_, _19_, 2038–2048. https://doi.org/10.1890/08-2236.1

Hurlbert, A. H., & Jetz, W. (2007). Species richness, hotspots, and the scale dependence of range maps in ecology and conservation. _Proceedings of the National Academy of Sciences USA_, _104_, 13384–13389. https://doi.org/10.1073/pnas.0704469104

IUCN. (2010). Guidelines for using the IUCN Red List categories and criteria. Version 8.1. Prepared by the Standards and Petitions Subcommittee in March 2010. Retrieved from http://intranet.iucn.org/webfiles/doc/SSC/RedList/RedListGuidelines.pdf

Jetz, W., Cavender-Bares, J., Pavlick, R., Schimel, D., Davis, F. W., Asner, G. P., … Ustin, S. L. (2016). Monitoring plant functional diversity from space. _Nature Plants_, _2_, 16024. https://doi.org/10.1038/nplants.2016.24

Jetz, W., McPherson, J. M., & Guralnick, R. P. (2012). Integrating biodiversity distribution knowledge: Toward a global map of life. _Trends in Ecology and Evolution_, _27_, 151–159. https://doi.org/10.1016/j.tree.2011.09.007

Johnson, D. J., Beaulieu, W. T., Bever, J. D., & Clay, K. (2012). Conspecific negative density dependence and forest diversity. _Science_, _336_, 904–907.

Johnson, P. C. D. (2014). Extension of Nakagawa & Schielzeth's R^2_{GLMM} to random slopes models. _Methods in Ecology and Evolution_, _5_, 944–946.

Kattge, J., Díaz, S., Lavorel, S., Prentice, I. C., Leadley, P., Bönisch, G., ... Wirth, C. (2011). TRY – a global database of plant traits. *Global Change Biology*, 17, 2905–2935. https://doi.org/10.1111/j.1365-2486.2011.02451.x

Kunin, W. E. (1998). Extrapolating species abundance across spatial scales. *Science*, 281, 1513–1515. https://doi.org/10.1126/science.281.5382.1513

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2015). *lmerTest: Tests in linear mixed effects models. R package version 2.0-29*. Retrieved from https://CRAN.Rproject.org/package=lmerTest

LaManna, J. A., Mangan, S. A., Alonso, A., Bourg, N. A., Brockelman, W. Y., Bunyavejchewin, S., ... Myers, J. A. (2017). Plant diversity increases with the strength of negative density dependence at the global scale. *Science*, 356, 1389–1392. https://doi.org/10.1126/science.aam5678

Legendre, P. (1993). Spatial autocorrelation: Trouble or new paradigm? *Ecology*, 74, 1659–1673. https://doi.org/10.2307/1939924

Legendre, P., Mi, X., Ren, H., Ma, K., Yu, M., Sun, I.-F., & He, F. (2009). Partitioning beta diversity in a subtropical broad-leaved forest of China. *Ecology*, 90, 663–674. https://doi.org/10.1890/07-1880.1

Lutz, J. A., Furniss, T. J., Johnson, D. J., Davies, S. J., Allen, D., Alonso, A., ... Zimmerman, J. K. (2018). Global importance of large-diameter trees. *Global Ecology and Biogeography*, 27, 849–864. https://doi.org/10.1111/geb.12747

McGill, B. J., Enquist, B. J., Weiher, E., & Westoby, M. (2006). Rebuilding community ecology from functional traits. *Trends in Ecology and Evolution*, 21, 178–185. https://doi.org/10.1016/j.tree.2006.02.002

Mu, Q., Zhao, M., & Running, S. W. (2011). Improvements to a MODIS global terrestrial evapotranspiration algorithm. *Remote Sensing of Environment*, 115, 1781–1800. https://doi.org/10.1016/j.rse.2011.02.019

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, 4, 133–142.

Packard, G. C., Birchard, G. F., & Boardman, T. J. (2011). Fitting statistical models in bivariate allometry. *Biological Review*, 86, 549–563. https://doi.org/10.1111/j.1469-185X.2010.00160.x

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed effect models in S and S-PLUS*. New York, NY: Springer.

Pocock, M. J. O., Hartely, S., Telfer, M. G., Preston, C. D., & Kunin, W. E. (2006). Ecological correlates of range structure in rare and scarce British plants. *Journal of Ecology*, 94, 581–596. https://doi.org/10.1111/j.1365-2745.2006.01123.x.

Rüger, N., Wirth, C., Wright, S. J., & Condit, R. (2012). Functional traits explain light and size response of growth rates in tropical tree species. *Ecology*, 93, 2626–2636. https://doi.org/10.1890/12-0622.1

Schemske, D. W., Mittelbach, G. G., Cornell, H. V., Sobel, J. M., & Roy, K. (2009). Is there a latitudinal gradient in the importance of biotic interactions? *Annual Review of Ecology and Systematics*, 40, 245–269.

Schlather, M., Malinowski, A., Oesting, M., Boecker, D., Strokorb, K., Engelke, S., ...Team, R. C. (2017). RandomFields: Simulation and analysis of random fields. R package version 3.1.5.0. Retrieved from https://CRAN.Rproject.org/package=RandomFields

Seidler, T. G., & Plotkin, J. B. (2006). Seed dispersal and spatial pattern in tropical trees. *PLoS Biology*, 4, e344. https://doi.org/10.1371/journal.pbio.0040344

Stein, A., Gerstner, K., & Kreft, H. (2014). Environmental heterogeneity as a universal driver of species richness across taxa, biomes and spatial scales. *Ecology Letters*, 17, 866–880. https://doi.org/10.1111/ele.12277

Thomson, F. J., Moles, A. T., Auld, T. D., & Kingsford, R. T. (2011). Seed dispersal distance is more strongly correlated with plant height than with seed mass. *Journal of Ecology*, 99, 1299–1307. https://doi.org/10.1111/j.1365-2745.2011.01867.x

Westoby, M. (1998). A leaf-height-seed (LHS) plant ecology strategy scheme. *Plant and Soil*, 199, 213–227.

Westoby, M., & Wright, I. J. (2006). Land-plant ecology on the basis of functional traits. *Trends in Ecology and Evolution*, 21, 261–268. https://doi.org/10.1016/j.tree.2006.02.004

Wilson, R. J., Thomas, C. D., Fox, R., Roy, D. B., & Kunin, W. E. (2004). Spatial patterns in species distributions reveal biodiversity change. *Nature*, 432, 393–396. https://doi.org/10.1038/nature03031

Wright, J. S. (2002). Plant diversity in tropical forests: A review of mechanisms of species coexistence. *Oecologia*, 130, 1–14. https://doi.org/10.1007/s004420100809

Zuur, A. F., Ieno, E. N., & Elphick, C. S. (2010). A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution*, 1, 3–14. https://doi.org/10.1111/j.2041-210X.2009.00001.x

BIOSKETCH

HAIBAO REN is interested in the scaling of biodiversity and species distributions across space and time. He is also interested in interpreting species coexistence in tree communities. This study resulted from a collaborative international research initiative for the comparison of occupancy scaling for tree species worldwide.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.