Scale-dependent species–area relationship: Niche-based versus stochastic processes in a typical subtropical forest

Haibao Ren1 | Jens-Christian Svenning2,3 | Xiangcheng Mi1 | James A. Lutz4 | Jinxing Zhou5,6 | Keping Ma1

1State Key Laboratory of Vegetation and Environmental Change, Institute of Botany, Chinese Academy of Sciences, Beijing, China; 2Center for Biodiversity Dynamics in a Changing World (BIOCHANGE), Department of Biology, Aarhus University, Aarhus C, Denmark; 3Section for Ecoinformatics & Biodiversity, Department of Biology, Aarhus University, Aarhus C, Denmark; 4Wildland Resources Department, Utah State University, Logan, UT, USA; 5Jianshui Research Station, School of Soil and Water Conservation, Beijing Forestry University, Beijing, China and 6Engineering Research Center of Forestry Ecological Engineering, Ministry of Education, Beijing, China

Abstract

1. Determining the patterns and drivers of the small-scale species–area relationship (SAR) is crucial for improving our understanding of community assembly and biodiversity patterns. Niche-based and stochastic processes are two principal categories of mechanisms potentially driving SARs. However, their relative importance has rarely been quantified rigorously owing to scale dependence and the simplified niche volumes often used.

2. In a fully mapped, 24-hm² plot of a typical subtropical forest, we built the SARs and well-defined niche hyper-volumes of a broad range of environmental variables at scales of 10–70 m (cell sizes). We then simulated passive sampling and partitioned the variances of the SAR slopes to disentangle these two contrasting mechanisms.

3. We found that the small-scale SAR best followed a power-law relationship, consistent with large-scale SARs. The SAR slope declined with increasing scale; it was lower than expected under passive sampling at scales below 30 m and higher at larger scales. Environmental niches explained more (39%–64%) of the slope at larger scales, exceeding 50% at scales >30 m, and these niches always captured the majority of the structured slopes. Environmental position (environmental mean values) effects were steady in absolute strength across scales and explained most (98%–68%) of the niche effect, but this proportion decreased with increasing scale. The effect of environmental heterogeneity increased with spatial scales, starting to rise at the 30 m scale after controlling for environmental position. Excluding soil properties from analyses strongly reduced these niche effects, highlighting the importance of soils for structuring the small-scale SAR. There was also substantial stochasticity in the SAR slopes, which was only partially explained by passive sampling.
1 | INTRODUCTION

The species–area relationship (SAR), which describes the observed increase in species number with increased areas, is one of the oldest ‘laws’ in ecology (Rosenzweig, 1995). However, the SAR pattern and its ecological mechanisms are scale-dependent and still uncertain. Small-scale SARs are relevant to larger-scale SARs and reflect the inherent structure of communities (He & Legendre, 2002; Kunin et al., 2018). Describing the patterns of small-scale SARs and understanding the mechanisms controlling them are thus essential for community ecology and biodiversity scaling (Matthews et al., 2021). Niche-based and stochastic processes, two competing hypotheses, have been proposed to shape the SAR (Hubbell, 2001; Williams, 1964). However, there has been considerable debates regarding their relative importance and how they affect the SAR (Gewin, 2006; Gravel et al., 2006; Ning et al., 2019). So far, few if any studies have systematically examined the spatial variations in the SAR slope at and across small scales (but see Fridley et al., 2005) and quantified the actual contribution of the niche versus stochasticity to these variations.

The niche of a species is an n-dimensional hyper-volume of environmental variables (conditions and resources) that define the ecological requirements of the species to survive, grow and reproduce (Hutchinson, 1957). However, studies have rarely defined the hyper-volume sufficiently to characterize a true niche space (Cáceres et al., 2012; Legendre et al., 2009). A species niche consists of the position that the species takes in the environmental hyper-volume and the range of environmental limits that the species can tolerate, that is, the mean and variance, respectively, of environmental variables that the species utilizes (Doledec et al., 2000; Hirzel et al., 2002). Correspondingly, the mean and variance of environmental variables of a site, termed environmental position (Brunbjerg et al., 2017) and heterogeneity, respectively, measure the size of an environmental volume (Soberón, 2019). Environmental position and heterogeneity have been focal topics of studies and have sparked a long-running debate about their relative importance to biodiversity (Stevens & Carson, 2002). However, few studies have considered them together to evaluate niche effects. Accordingly, well-defined and measured niches are prerequisites to quantify precisely the niche-based effects.

Processes shaping SARs are scale dependent, which impedes our disentangling of the roles of ecological processes in shaping SARs (Chase, 2014; Turner & Tjørve, 2005). As environmental volume increases, more species’ niches are included, increasing the number of species the volume can support (Soberón, 2019). Within habitats, environments are more or less homogeneous; environmental position is thus relatively invariant with somewhat random fine-scaled fluctuations of environmental factors (i.e. within-habitat heterogeneity; Shmida & Wilson, 1985). As sampled areas increase, the environmental position does not change greatly, while the environmental fluctuation additionally begins to cover more heterogeneity between habitats. Therefore, as the spatial scale increases, the effect of environmental position would change little (Shmida & Wilson, 1985), and its relative importance would decrease owing to the increased between-habitat heterogeneity effects (Turner & Tjørve, 2005; Williams, 1964). Consequently, niche effects increase with scales due to the increased heterogeneity effects. Partitioning environmental space into the environmental position and heterogeneity helps us unravel how niche processes affect SARs.

In contrast, stochastic processes can also generate SARs. According to neutral theory, ecological drift and random dispersal determine SARs (Hubbell, 2001). At smaller spatial scales, environments are more homogeneous, species are more functionally similar, and stochastic events such as birth, death, and dispersal are more likely to occur (Chase, 2014). Specifically, local species assemblages are completely random samples from the regional species pool without dispersal limitation. This so-called passive sampling (Coleman et al., 1982; Hubbell, 2001) would completely randomize SARs in space. As spatial scales increase, many biological or ecological processes spatially cluster species’ individuals, which would weaken the effect of passive sampling. Random dispersal structured species individuals and then SARs spatially, which are spatially independent of environmental variables nonetheless (Hubbell, 2001). Furthermore, stochastic environmental fluctuation and competition for resources can generate neutrality (Holt, 2006), leading to a spatially random distribution of SARs.

4. Synthesis. Our results show that the small-scale SAR in the studied subtropical forest follows a power law, exhibits a scale shift in SAR slope at 30 m, and is strongly shaped by niche effects that are dominated by environmental position relative to heterogeneity. However, soil heterogeneity controls the increase in niche effect and the shift in the SAR slope with increasing spatial scales. Hence, edaphic factors can be responsible for scale dependence in small-scale SARs, thereby linking small-scale and large-scale SARs.

KEYWORDS
community assembly, Gutianshan forest dynamics plot, niche versus neutrality, passive sampling, power-law model, resource availability and heterogeneity, spatial effect, variance partitioning
Studies have attempted to unravel the effects of these processes on SARs (e.g. Shen et al., 2009); however, their relative importance has rarely been quantified systematically across scales. Passive sampling expects spatial randomness of SARs (Coleman et al., 1982), and the deviation of SARs from this expectation denotes the influences of niche-based and/or other stochastic processes. Environmental niches may generate spatial structures of SARs (Shmida & Wilson, 1985). Mechanisms of species dispersal may also influence SAR structure (Hubbell, 2001). However, random dispersal (Hubbell, 2001) or the dispersal–competitiveness trade-offs independent of environments (Amarasekare, 2003) should produce SARs independent of environments. Large forest dynamics plots (e.g. ≥10 h) that fully enumerate and map all trees ≥1 cm in diameter at breast height (Condit, 1998) can control for area effects, which often confound effects of other processes (Matthews et al., 2021). If highly spatially resolved data for many environmental variables are available, these large plots may allow disentangling the processes driving SARs across scales using variance partitioning techniques (Figure 1) (Legendre et al., 2009; Peres-Neto & Legendre, 2010).

We constructed a spatially explicit dataset that included sufficiently large numbers of tree individuals and environmental parameters in a typical subtropical forest with highly rugged terrains. Using this dataset, we aimed to answer the following questions: (1) What are the systematic variations in the SAR slope at and across a range of spatial scales from 100 m\(^2\) to 4900 m\(^2\)? (2) What is the relative importance of niche-based versus stochastic processes in determining the SAR across spatial scales? (3) What are the relative contributions of environmental position and heterogeneity to the SAR and total niche effect across spatial scales?

2 | METHODS

2.1 | Study area and plot

This study was conducted in the Gutianshan National Nature Reserve (29°10′–29°17′N and 118°03′–118°11′E) in the mid-subtropical zone of China. In this reserve, the mean annual temperature is 15.3°C, the mean annual precipitation is 1964 mm, and the highest elevation is 1258 m (Yu et al., 2001). Typical evergreen broad-leaved forests are widely distributed at elevations below 800 m in this region, with the dominant tree species being Castanopsis eyrei [Fagaceae] and Schima superba [Theaceae] (Yu et al., 2001).

The 24-hm\(^2\) (600 m × 400 m) Gutianshan Forest Dynamics Plot (GFDP) was established in 2005 as part of the Chinese Forest Biodiversity Monitoring Network (CForBio). The GFDP ranges from 446.3 m to 714.9 m in elevation with highly rugged terrain (elevation range of 268.6 m; Figure 2A). Vegetation is typical of lowland evergreen broad-leaved forests of the mid-subtropics. The GFDP was surveyed with total stations and divided into a grid with unit cells measuring 20 m × 20 m. At the scale of the surveyed 20 m grid, aspect ranged from 33.6° to 352°, with a mean of 182.3°, while slope ranged from 13° to 62°, with a mean of 38°. All free-standing stems ≥10 mm in diameter at breast height (DBH) were tagged, measured, mapped and identified to species. In total, 140,700 individuals were recorded, belonging to 159 species, 104 genera and 49 families. Nomenclature followed Zheng (2005).

2.2 | Species–area relationship (SAR) and spatial scale

For analysis, we divided the GFDP into grids with square cells of 10, 20, 30, 40, 50, 60 and 70 m in side length. The cell size defined the extent, that is, the scale at which we built the SAR: hereafter, the scale is given as the cell side length (L). At some scales, the GFDP was not exactly divided, and in those cases, we omitted part of the northern and/or eastern edges as needed, starting the grid from the southwest corner (the plot was aligned with the cardinal directions). For example, the 70-m cell size resulted in 40 cells, arranged as five by eight cells along the short and long axis of the GFDP, respectively, where 50 m along the northern edge and 40 m along the eastern edge were excluded. In total, 2400, 600, 260, 150, 96, 60 and 40

![Figure 1](https://example.com/figure1.png) Schematic representation of the variance in the slope of the species–area relationship (SAR). We assumed that the SAR was determined by environmental position (EP \([a + d + f + g]\)), environmental heterogeneity (EH \([b + d + e + g]\)), spatial factors (PCNMs \([c + e + f + g]\)) and passive sampling (PS \([h]\)). The environmental position and heterogeneity constitute the environmental niche \([EN \[a + b + d + e + f + g]\]) and passive sampling \([PS \[h]\]) of variation unexplained by niche effects, passive sampling and spatial factors implies other processes acting to moderate SARs. The fraction of the SAR-slope’s variance purely explained by spatial factors (PCNMs \([c]\)) could arise from species themselves, such as dispersal.

![Figure 2](https://example.com/figure2.png) Range of 268.6 m; Figure 2A). Vegetation is typical of lowland evergreen broad-leaved forests of the mid-subtropics. The GFDP was surveyed with total stations and divided into a grid with unit cells measuring 20 m × 20 m. At the scale of the surveyed 20 m grid, aspect ranged from 33.6° to 352°, with a mean of 182.3°, while slope ranged from 13° to 62°, with a mean of 38°. All free-standing stems
In building the SAR, a cell with side length $L$ was evenly divided into $n_i$ square sub-cells with the $i^{th}$ length of the sub-cell side $l_i = L / \sqrt{n_i}$, where $\sqrt{n_i}$ ranges from one to $L$ in values, and the potential values of $\sqrt{n_i}$ are: 1, 2, 3, ..., 10, 12, 14, ..., 20, 25, 30, ..., 45, 50, 60, 70. For instance, for a cell of $10 \times 10$ m, the number of sub-cells, $n_i$, is 1, 4, 9, 16, 25, 36, 49, 64, 81, 100, corresponding to the lengths of sub-cell side ($l_i$): 10, 5, 3.33, 2.5, 2, 1.67, 1.43, 1.25, 1.11 and 1 m respectively. For a given sub-cell size, we took the mean number of species over all sub-cells versus the sub-cell area in building a SAR for the cell.

Numerous formulas have been applied to fit SARs (Dengler, 2009). It has been suggested that small-scale species–area relationships are better fitted by a Gleason exponential model rather than an Arrhenius power-law model (Gleason, 1922; He & Legendre, 1996). So far, however, it has been generally accepted that the power-law model best fits and robustly describes SARs (Harte et al., 2001; Martín & Goldenfeld, 2006), including small-scale SARs (Fridley et al., 2005). The power-law model is often considered to be the simplest for comparing SARs among studies (Dengler, 2009). This study fitted species–area data with both models and selected the better one for the small-scale SAR. A power-law model is often fit in one of two ways: either in a
double-log space or nonlinearly in an arithmetic space. Fitting the power model in a double-log space results in a linear regression with a multiplicative error term, giving more weight to good fit at smaller grain sizes, reducing heteroscedasticity in the residuals, but also introducing rotational distortion (Packard et al., 2011). We implemented both methods and then selected the model with higher goodness-of-fit. The goodness-of-fitting was measured by the adjusted coefficient of determination $R^2_a$ from the model:

$$R^2_a = 1 - \frac{(n-1)\text{RSS}}{(n-k)\text{TSS}}.$$  

where $\text{RSS} = \sum_i^n (y_i - \hat{y}_i)^2$, the residual sum of squares; $\text{TSS} = \sum_i^n (y_i - \bar{y})^2$, the total sum of squares; $n$ is the number of samples or sub-side lengths; $k = 2$ (the number of parameters in the power-law model); $y_i$ is the $i$th observed SAR slope; $\hat{y}_i$ is the $i$th fitted SAR slope; and $\bar{y}$ is the mean of the observed SAR slope. Here, $R^2_a$ was computed using untransformed data. Therefore, for the double-log linear regression, $R^2_a$ was not calculated in a double-log space but the back-transformed

density (BD); soil moisture (SM); N mineralization rate from organic matter (Nmin); total C, N and P content; and the availability of N, P, K, Ca, Mg, Na, Al, Si, B, Fe, Cu, Mn and Zn; Zhang et al., 2011).

The four topographic factors were calculated from a digital elevation model with a resolution of 5 m x 5 m. Aspect, a circular variable, was converted to a southerly index ($-\cos[\text{aspect}]$) and an easterly index ($\sin[\text{aspect}]$) to meet the requirements of linear modelling (Clark et al., 1999). The 20 edaphic factors were each interpolated at the 5 m x 5 m scale (resolution) using ordinary kriging techniques from 893 soil samples, with a minimal sampling scale of 2 m throughout the GFDP (see Zhang et al., 2011 for details).

To better capture environmental heterogeneity, we proposed two additional variables for each cell: topographic roughness (TR) and elevation range (ER). Topographic roughness for a cell was calculated as the ratio of the cell's land surface area to its horizontal projected area, where the land surface area was computed as the sum of slope areas of all sub-cells of 5 m x 5 m over the entire cell. Elevation range for a cell was calculated as the maximal elevation difference within the cell.

### 2.4 Passive sampling

Passive sampling, a random process, assumes that the SAR is shaped exclusively based on random occurrences of individuals of each species in a given area (Coleman et al., 1982). The spatial distribution of each species follows a homogeneous Poisson process with a single parameter, which is the mean density of individuals per unit area over the 24-h GFDP.

We simulated this random process for each species. After one simulation for all of the species, we overlapped the distributions of all these species to determine the homogeneous Poisson assemblage and then estimated the SAR slopes for all individual cells and the corresponding variance of the slopes among the cells over the entire GFDP at the scales of 10, 20, 30, 40, 50, 60 and 70 m. We replicated the above procedure 100 times and calculated the mean (expected) variance of the SAR slope under the passive sampling assumption. The simulations were run in the R-package spatstat version 1.64–1 (Baddeley et al., 2015).

### 2.5 Spatial variables

The spatial structure of the SAR can be represented by principal coordinates of neighbour matrices (PCNMs) eigenfunctions (Legendre et al., 2009). We computed the PCNMs using principal coordinate analysis (PCoA) of a truncated geographic distance matrix among all central points of the grid cells over the entire 24-h area at each spatial scale (e.g. 600 central points at the scale of 20 m; Legendre et al., 2009; Oksanen et al., 2019). We used the PCNM eigenfunctions with positive eigenvalues as explanatory spatial variables to document the spatial structure of the SAR slopes. The numbers of effective PCNMs were 1686, 420, 182, 105, 67, 42 and 28 at the scale of 10, 20, 30, 40, 50, 60 and 70 m respectively. The eigenvalues of the PCNMs corresponded to the spatial scale;
as PCoA ranks PCNMs by decreasing eigenvalues, the PCNMs rank negatively corresponding to spatial scales (Legendre et al., 2009).

2.6 | Variation partitioning

We partitioned the variance of the SAR slope over the 24-hm² grid at a given scale (i.e. grid-cell size) into environmental position, environmental heterogeneity and spatial variables (PCNMs). The explanatory power (or effect) of environmental niches was the proportion of the variance of the SAR slope explained by environmental position and heterogeneity together (Legendre et al., 2009; Figure 1). In the present study, the variance partitioning was, in fact, partial regression since the response was a single variable—the SAR slope. The variance partition was computed using the R-package vegan version 2.5–6 (Oksanen et al., 2019). To examine the changes in the effects of the environmental niches with scales, we repeated the above computation at different spatial scales.

To capture potentially nonlinear effects of environmental niches, we used a third-degree orthogonal polynomial for each environmental variable, which expanded the environmental position matrix and the environmental-heterogeneity matrix. A forward selection was used to select the best subgroup of explanatory variables for building a parsimonious model by permutation tests, at the 5% significance level, of the increase in $R^2$ at each step. Before the variance partitioning, we applied the forward selection to these two environmental matrices, respectively, and spatial variables (PCNMs). For example, this selection procedure obtained the number of significant PCNMs: 193, 96, 34, 11, 8, 9 and 2 at the scales of 10, 20, 30, 40, 50, 60 and 70 m, respectively. We ran the forward selection using the R-package packfor version 0.0–8 (Dray et al., 2007).

3 | RESULTS

3.1 | Spatial variation in the SAR slope

The Arrhenius power-law model provided the best fit for the small-scale SARs (Figure 3); the SAR slope did vary at and across scales and deviated from expected under passive sampling (Figure 4). At

---

**FIGURE 3** Histogram of adjusted coefficients of determination ($R^2_a$) for fitting the species–area relationships (SARs) at all the scales (extent), with Arrhenius power model in both the double-log transformed space of species–area data and the arithmetic space of the original data and with Gleason model (A); the bin size is 0.01. Three examples of the SAR fitting with the three models, randomly sampled from all the fitted SARs at the scales of 10, 30 and 50 m respectively (B to D); the scales refer to the side lengths of cell squares (extent) in which SARs are fitted.
the spatial scale of 10 m, for example, the observed slopes were more widespread ($F = 3.61, p < 0.001$) and smaller ($t = -49.9, p < 0.001$) than the expected (Figure 4A–C). These observed slopes had a minimum of 0.392 on the highest ridge and a maximum of 0.819 in the lowest valley. Most (95%) of the observed slopes were between 0.485 and 0.758, following non-normality (Shapiro and Wilk’s $W = 0.995, p < 0.001$) with a mean of 0.631 (Figure 2A and 4A,C). These expected slopes ranged from 0.581 to 0.827, most (95%) of which were between 0.640 and 0.779, following normality (Shapiro and Wilk’s $W = 0.999, p = 0.51$) with a mean of 0.709 (Figure 4B,C). Both observed and expected slopes decreased at increasing scales, as did their variance. However, the observed slope was always more dispersed than the expected, and it changed from being lower than expected to being higher than the expected at the scale of 30 m (Figure 4C). The SAR slopes showed a high degree of spatial structure at all scales (Figure 5A). For example, at the scale of 10 m, the observed structure was mainly broad scale, which was different from the randomness expected under passive sampling (Kolmogorov–Smirnov test, $D = 0.58, p < 0.001$; Figure 4A,B,D).

### 3.2 Niche-based versus stochastic effects

Environmental niche accounted for 39%–64% of the variance of the SAR slopes, with increasing importance as spatial scales increased, for example, >50% at the spatial scales >30 m (Figure 5A,B). Factoring out spatial factors (PCNMs), the niche effect continued to increase with spatial scales (Figure 5A,C). The SAR slope spatially matched well with the dominant environmental variables (Figures 2 and 4A, and Figure S1). The SAR slopes fitted by the niche were spatially structured, and the structured scales and degrees were highly consistent with those of the environmental variables (Figure 6 and Appendix S1). For example, at the scale of 10 m, the SAR was steep at the sites with low elevation, weak convexity, rich nitrogen and heterogeneous phosphorus; these variables were dominant factors affecting the SAR slope (Figures 2 and 4A). The environmental variables at this scale manifested a broad-scale structure, which corresponded well to that of the SAR slope (Figure 6A–C). Elevation and phosphorus heterogeneity also dominated at other scales and were spatially congruent with the SAR slope (Figure S1). Excluding the soil variables cut down the fraction explained by the niche by 48.0% on
FIGURE 5  Proportion of the variation in the slope of the species-area relationship (SAR) explained by environmental niche factors overall (environmental position and heterogeneity together), spatial factors (PCNMs) and passive sampling, and how these change across spatial scales (see also Figure 1). The stacked proportions of the variance (adjusted $R^2$, $R^2_a$) explained by environmental niche factors overall (EN [$a + b + d + e + f + g$]), environmental position (EP [$a + d + f + g$]), environmental heterogeneity (EH [$b + d + e + g$]), spatial factors (PCNMs [$c + e + f + g$]) and passive sampling (PS [$h$]), as well as the proportion unexplained by the previous factors (UN [$i$]) (panel A). The changes across spatial scales in the proportion of the variation explained fully by the niche effects (EN), environmental position (EP) and environmental heterogeneity (EH) (panel B). The changes across spatial scales in the variation explained partially by the niche when controlling for spatial factors (PEN [$a + b + d$], partial niche effect), environmental position when controlling for environmental heterogeneity (PEP [$a + f$]), environmental heterogeneity when controlling for environmental position (PEH [$b + e$]), and jointly by environmental position and heterogeneity (PH [$d + g$]) (panel C). Niche effects (ENs) when only considering topographic variables (i.e., excluding edaphic variables), which are partitioned into the effects of environmental position (EP) and environmental heterogeneity (EH) and their joint effect (PH) (panel D). The changes across spatial scales in the proportion of the variance of SAR slope explained by passive sampling (PS) and purely by spatial factors (PC [$c$]), and in the unexplained proportion (UN) (panel E).
average to 26.3 ± 5.4% and removed the trend in the niche effect (Figure 5A,D).

As for stochasticity, passive sampling presented spatial randomness in the SAR slope (Figure 4B) and accounted for only part of the variance of the observed SAR slope unexplained by the niche. This part accounted for 10.5% to 27.7% of the total variance of the observed SAR slope, first declining rapidly and then leveling off as spatial scales increased (Figure 5A,E). Independent of the niche, pure spatial factors (PCNMs) explained 0%–20.8% of the variance of SAR slopes, with decreasing importance with increasing spatial scales (Figure 5A,E). The SAR slopes determined by pure PCNMs exhibited fine- to- broad- scaled structure, where the spatial clusters seemed to randomly scatter (Figure 6J,K, Appendix S1). Unexplained by environmental niches, PCNMs and passive sampling remained 21.6% ± 3.9% of the variation in the SAR slope, which was separate from environmental variables and was spatially random (Figure 5A,E).

### 3.3 | Environmental position versus heterogeneity effects

As a component of the environmental hyper- volume, environmental position explained a consistently large proportion (45 ± 4%) of the total variance of the SAR slope across spatial scales (Figure 5A,B), accounting for most (68%–99%) of the niche effects in a decreasing fashion with spatial scale. This steady effect decreased (although not significantly) with scales after factoring out environmental heterogeneity (Figure 5A,C). At larger scales, environmental heterogeneity explained more of the variance of the SAR slope, with or without

![Figure 6](https://example.com/figure6.png)

**Figure 6** Distribution maps of the slope of the species–area relationship (SAR) fitted by environmental niche factors overall (EN, environmental position and heterogeneity together) (A), environmental position (EP) (D), environmental heterogeneity (EH) (G) and purely spatial factors (PCNMs) independent of niche effects (J), at the scale of 10 m. The proportions of the corresponding fitted SAR slopes above explained by PCNMs individually ($R^2$) and cumulatively ($R^2_c$), showing how the fitted slopes are spatially structured (B, E, H, K). The proportions of environmental niche factors overall (C), environmental position (F) and heterogeneity (I) above explained by PCNMs individually ($R^2$) and cumulatively ($R^2_c$), showing how environmental resources and conditions are spatially structured. The PCNM rank negatively corresponds to the spatial scale, that is, the higher this rank is, the finer the spatial scale is.
controlling for environmental position (Figure 5A–C), and took more of the total niche effect, which was generally less than those taken by environmental position (t = −3.29, p = 0.004). Environmental position and heterogeneity jointly had high impacts on the SAR, which increased with spatial scale (Figure 5A,C). The heterogeneity effect was almost associated with the environmental position at small scales; it started to grow at the scale of 30 m, independent of environmental position (Figure 5A,C).

The SAR slopes were spatially correlated with variables of environmental position, for example, elevation, and with variables of environmental heterogeneity, for example, phosphorus heterogeneity (Figures 2, 4 and 6 and Figure S1). The environmental position and heterogeneity displayed spatial clusters, of which the scales and degree corresponded well to those of the SAR slope (Figure 6 and Appendix S1). The resource-richer or better-conditioned areas varied more in resources, for example, the valleys with more abundant phosphorus showed more spatial heterogeneity in phosphorous (Figure 2D,E and Figure S1). Discarding all soil factors reduced the fraction of the SAR explained by environmental position by 42.8% (down to 25.6±5.9%) and the fraction explained by environmental heterogeneity by 93.4% (down to 2.2±1.8%) and erased the scale gradient in the environmental-heterogeneity effect (Figure 5A,D).

4 | DISCUSSION

The large numbers of individual trees, species represented and measured environmental factors enabled us to establish species–area relationships (SARs) and well-defined hyper-volumes of environmental niches in spatially explicit contexts across scales. We were thus able to partition how environmental niches versus stochastic processes determined the SAR.

It is well known that the power-law model properly fits the SARs at large rather than small scales due to the different processes that drive them (He & Legendre, 1996; Rosenzweig, 1995). We found that small-scale SARs best followed the power-law model, which connects small-scale SARs to large-scale SARs. This result is consistent with Fridley et al. (2005). As the power-law represents ecological similarity across scales (Rosenzweig, 1995), we suggest that small- and large-scale SARs at least partially share similar causative processes. For building SARs in this study, individual trees were recorded once only in the quadrats where their roots were based in the 24-hm² grid-point forest system. In this system, the fact that the slope of SAR is very high at fine scales and decreases with increasing spatial scales could be purely attributable to sampling and have nothing to do with ecological processes (Williamson, 2003). However, we found that the observed SAR slopes deviated from expected under passive sampling, and both decreased with scale. These findings signal the influence of ecological processes on the SAR, highlight studies on how and why power-law slope changes across scales, and confirm that passive sampling affects the decrease in the SAR slope.

Estimating spatially explicit SARs and relevant hyper-volumes of environmental niches allowed us to rigorously quantify the importance of niche processes for the SAR across scales. We found support for the niche effect increasing with spatial scales. In our species-rich forest, the niche effect explained c. 40% of the variation in the SAR slope at the minimal scale of 10 m and >50% at scales >30 m. This result emphasizes the role of niches in shaping small-scale SARs and identifies where their effect is dominant. Reducing the environmental dimension by removing the soil factors led to a heavy underestimation of the niche effect. Chang et al. (2013) observed similar results for community composition in another subtropical forest. These findings indicate the role of well-defined niche volumes in quantifying the niche effect. Importantly, we found that the niche explained from 68% to c. 100% of the spatial structure of the SAR slope as scales increased, suggesting that the niche governs the systematic variation in the SAR slope, even at small scales.

Stochasticity was notable for the SAR even in our highly heterogeneous study plot, where niche differentiation of species is strong (Brown et al., 2013). Our results support the hypothesis that the sampling effect decreases with scales (Coleman et al., 1982), but it was generally weak. Pure spatial effects, separate from environmental niches, decreased with increasing spatial scale. These effects could stem from dispersal processes, either random dispersal limitation (Hubbell, 2001) or dispersal effects partially linked to niche processes, for example, trade-offs between dispersal limitation and competitiveness (Amarasekare, 2003) or mass effects (Shmida & Ellner, 1984). Environmental niches, PCNMs and passive sampling still left a large proportion of random variation unexplained in the SAR slope. Passive sampling is a completely random process (Coleman et al., 1982), the deviation of which indicates the importance of non-random biological and/or ecological processes. Competition, density dependence, and environmental temporal randomness likely generate stochasticity (Holt, 2006; Scheffer et al., 2018).

Deconstructing environmental volume into components of environmental position and heterogeneity is helpful to unravel how environmental niches affect SARs. Our findings support the hypothesis that the environmental-position effects are relatively steady across scales (Shmida & Wilson, 1985), while the heterogeneity effects increase with spatial scales (Williams, 1964). Environmental heterogeneity thus drives the increase in niche effects with spatial scales. The SAR slope shifted from being lower to being higher than expected at the scale where the heterogeneity effect free of the environmental position began to increase with spatial scales. This scale matching suggests that the environmental heterogeneity led to the SAR slope’s shift. Homogeneous habitats filter species (HilleRisLambers et al., 2012), which exaggerates the limitations in recruitment and decreases the slope of the SAR. Extended areas encounter more new species and steepen the SAR due to covering more environmental variations between habitats. Excluding soil factors erased the increase in heterogeneity effects and underestimated the environmental position effects to an extreme degree, which underpins the role of soil factors in driving the SAR. Soil factors are fractal at multi-scales (Burrough, 1983). These findings, therefore, provide a
potential answer as to why the power-law model generally best fits SARs.

The environmental position explained more of the niche effect than the heterogeneity, and this proportion declined with scales. These results highlight the dominance of the environmental position effect over the heterogeneity, and this dominance may gradually make way for heterogeneity effects as spatial scales continue to increase. Additionally, removing edaphic variables almost completely negates the heterogeneity effects. Scale dependence and well-defined niches are thus crucial for solving the disputes over the relevance of environmental position versus heterogeneity to the SARs (Shmida & Wilson, 1985; Turner & Tjørve, 2005). The environmental position and heterogeneity jointly had high effects on the SAR. Removing soil factors almost eradicated this joint effect, implying the dependence of resource heterogeneity on resource position (or availability). For example, phosphorous-rich valleys exhibited high phosphorous variances (Figure 2D,E). Previous experiments also show the heterogeneity dependence (Conradi et al., 2017). In addition, this joint effect increased with scales, likely because larger areas cover broader ranges of a nutrient. It is thus crucial to further explore how resource availability, heterogeneity and areas interact to shape the SAR (Ben-Hur & Kadmon, 2020).

In this study, reporting the total variance explained by environmental variables may overestimate niche effects, owing to spatial autocorrelation in the SAR slopes and environmental variables. However, this was not a problem for our gridded data since the autocorrelation was evenly spread throughout the study region and only affected regression standard errors and not coefficient estimates (Hawkins et al., 2007). If this autocorrelation is at least in part driven by the environmental variables, the variances of the SAR slopes explained by environmental variables would be appropriate to measure the niche effects (Peres-Neto & Legendre, 2010). Here, the magnitudes and scales at which environmental variables were spatially structured were consistent with those for the observed SAR slope. Also, the SAR slope fitted purely by spatial factors seemed to be randomly clustered. This observation generally agrees with the stochastic, zero-mean and spatially correlated residuals of the Cox model, which simulates aggregative variation not captured by abiotic variables (Shen et al., 2013). It is, therefore, appropriate to report niche effect as total variance explained by environmental niches, environmental-position effect as that by environmental position, and environmental-heterogeneity effect as that by environmental heterogeneity.

5 | CONCLUSIONS

Analysing individual-level data on the distribution of trees and shrubs in a 24-hm² typical subtropical forest plot in China, this study systematically quantified variation in the SAR slope at and across spatial scales (extents or cell sizes of 10 m–70m in side length) and the relative importance of environmental niches versus stochastic processes in driving these variations. The power-law model provided the best fit for the small-scale SAR, the slope of which decreased with scales and showed a scale shift at 30m relative to the expectation under passive sampling. Our results revealed that niche processes dominated the SAR at larger scales and governed its spatial structure at all scales. The effects of environmental position were steady in absolute strength across scales and dominated the niche effects over environmental heterogeneity in a decreasing fashion; the heterogeneity effects increased with increasing scale. Importantly, soil heterogeneity gave rise to the increase in niche effect with scale and the scale shift of the SAR slope.

Overall, these findings show that environmental niches strongly influence and structure the SARs across scales for trees and shrubs in a typical subtropical forest locality and do so in a scale-dependent manner. However, non-environmental-niche processes such as stochasticity also emerged as dominant in shaping SARs at <30m scales, even in this setting with its highly varied soil and terrains. Although environmental position locally controls the niche effect, soil heterogeneity drives the increased importance of the niche with scale and the shift of the SAR shape across scales, linking small-scale SARs and large-scale SARs. Future efforts should expand to other regions and larger scales to provide further detailed assessments and robust generalization of the role of environmental niches in shaping woody plant assemblages across spatial scales.

AUTHOR CONTRIBUTIONS

H.R. and K.M. conceived the ideas and designed methodology; H.R. and X.M. collected the data; H.R. analysed the data and led the writing of the manuscript, H.R., J.-C.S., X.M., J.A.L., J.Z. and K.M. interpreted the results and revised the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

ACKNOWLEDGEMENTS

We thank Dr. Emily Drummond at the University of British Columbia for her editorial assistance. We thank Dr Mingjian Yu, Dr Jianhua Chen, Mr Shengwen Chen, and Mr Teng Fang for their great contribution to data collection and species identification, Prof. Bingyang Ding for identifying cryptic species, and Dr Liwen Zhang for her contribution to soil data. We also thank the dozens of field workers for plot censuses. This study was funded by the strategic priority research program of the Chinese Academy of Sciences, grant no. XDB31030000. J.-C.S. considers this work a contribution to his VILLUM Investigator project “Biodiversity Dynamics in a Changing World” funded by VILLUM FONDEN (grant 16549). This work was carried out with the support of the Zhejiang Qianjiangyuan Forest Biodiversity National Observation and Research Station, China.

CONFLICT OF INTEREST

We declare that we have no conflict of interest.
Peres-Neto, P. R., & Legendre, P. (2010). Estimating and controlling for spatial structure in the study of ecological communities. *Global Ecology and Biogeography, 19*, 174–184.

Ren, H., Svenning, J.-C., Mi, X., Lutz, J. A., Zhou, J., & Ma, K. (2022). Data from: Scale-dependent species–area relationship: Niche-based versus stochastic processes in a typical subtropical forest. *Dryad Digital Repository*. https://doi.org/10.5061/dryad.9kd51c5kn

Rosenzweig, M. L. (1995). *Species diversity in space and time*. Cambridge University Press.

Scheffer, M., Nes, E. H.v., & Vergnon, R. (2018). Toward a unifying theory of biodiversity. *Proceeding of National Academy of Sciences of the United States of America, 115*, 639–641.

Shen, G., He, F., Waagepetersen, R., Sun, I.-F., Hao, Z., Chen, Z.-S., & Yu, M. (2013). Quantifying effects of habitat heterogeneity and other clustering processes on spatial distributions of tree species. *Ecology, 94*, 2436–2443.

Shen, G., Yu, M., Hu, X.-S., Mi, X., Ren, H., Sun, I. F., & Ma, K. (2009). Species-area relationships explained by the joint effects of dispersal limitation and habitat heterogeneity. *Ecology, 90*, 3033–3041.

Shmida, A., & Ellner, S. (1984). Coexistence of plant species with similar niches. *Vegetatio*, 58, 29–55.

Shmida, A., & Wilson, M. V. (1985). Biological determinants of species diversity. *Journal of Biogeography*, 12, 1–20.

Soberón, J. (2019). A Grinnellian niche perspective on species-area relationships. *The American Naturalist*, 194, 760–775.

Stevens, M. H. H., & Carson, W. P. (2002). Resource quantity, not resource heterogeneity, maintains plant diversity. *Ecology Letters*, 5, 420–426.

Turner, W. R., & Tjarve, E. (2005). Scale-dependence in species-area relationships. *Ecography, 28*, 721–730.

Williams, C. B. (1964). *Patterns in the balance of nature*. Academic Press.

Williamson, M. (2003). Species-area relationship at small scales in continuum vegetation. *Journal of Ecology, 91*, 904–907.

Yu, M., Hu, Z., Yu, J., Ding, B. & Fang, T. (2001) Forest vegetation types in Gutianshan natural Reserve in Zhejiang. *Journal of Zhejiang University (Agriculture and Life Science)* [In Chinese], 27, 375–380.

Zhang, L., Mi, X., Shao, H., & Ma, K. (2011). Strong plant-soil associations in a heterogeneous subtropical broad-leaved forest. *Plant and Soil*, 347, 211–220.

Zheng, C. (2005) *Keys of seed plants in Zhejiang, China*. Zhejiang Science and Technology Press [in Chinese].

**SUPPORTING INFORMATION**

Additional supporting information may be found in the online version of the article at the publisher’s website.

---

**How to cite this article:** Ren, H., Svenning, J-C, Mi, X., Lutz, J. A., Zhou, J., & Ma, K. (2022). Scale-dependent species–area relationship: Niche-based versus stochastic processes in a typical subtropical forest. *Journal of Ecology*, 110, 1883–1895. https://doi.org/10.1111/1365-2745.13924