Received: 24 April 2022 | Revised: 25 August 2022 | Accepted: 6 October 2022

DOI: 10.1002/ecy.3917

ECOLOGY
ECOLOGICAL SOCIETY OF AMERICA

ARTICLE

On the shape and origins of the freshwater species–area relationship

Sophia I. Passy[1] | **Joseph L. Mruzek**[2] | **William R. Budnick**[3] | **Thibault Leboucher**[4] | **Aurélien Jamoneau**[5] | **Jonathan M. Chase**[6] | **Janne Soininen**[7] | **Eric R. Sokol**[8] | **Juliette Tison-Rosebery**[5] | **Annika Vilmi**[9] | **Jianjun Wang**[10] | **Chad A. Larson**[11]

[1]Department of Biology, University of Texas at Arlington, Arlington, Texas, USA

[2]Forestry and Environmental Conservation Department, Clemson University, Clemson, South Carolina, USA

[3]Department of Fisheries and Wildlife, Michigan State University, East Lansing, Michigan, USA

[4]Laboratory for Continental Environments, National Scientific Research Center, University of Lorraine, Metz, France

[5]INRAE, EABX, Cestas, France

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

[7]Department of Geosciences and Geography, University of Helsinki, Helsinki, Finland

[8]National Ecological Observatory Network (NEON), Battelle, Boulder, Colorado, USA

[9]Freshwater Centre, Finnish Environment Institute, Oulu, Finland

[10]State Key Laboratory of Lake Science and Environment, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing, China

[11]Washington State Department of Ecology, Environmental Assessment Program, Lacey, Washington, USA

Correspondence
Sophia I. Passy
Email: sophia.passy@uta.edu

Handling Editor: Helmut Hillebrand

Abstract

The species–area relationship (SAR) has over a 150-year-long history in ecology, but how its shape and origins vary across scales and organisms remains incompletely understood. This is the first subcontinental freshwater study to examine both these properties of the SAR in a spatially explicit way across major organismal groups (diatoms, insects, and fish) that differ in body size and dispersal capacity. First, to describe the SAR shape, we evaluated the fit of three commonly used models, logarithmic, power, and Michaelis–Menten. Second, we proposed a hierarchical framework to explain the variability in the SAR shape, captured by the parameters of the SAR model. According to this framework, scale and species group were the top predictors of the SAR shape, climatic factors (heterogeneity and median conditions) represented the second predictor level, and metacommunity properties (intraspecific spatial aggregation, γ-diversity, and species abundance distribution) the third predictor level. We calculated the SAR as a sample-based rarefaction curve using 60 streams within landscape windows (scales) in the United States, ranging from 160,000 to 6,760,000 km^2. First, we found that all models provided good fits ($R^2 \geq 0.93$), but the frequency of the best-fitting model was strongly dependent on organism, scale, and metacommunity properties. The Michaelis–Menten model was most common in fish, at the largest scales, and at the highest levels of intraspecific spatial aggregation. The power model was most frequent in diatoms and insects, at smaller scales, and in metacommunities with the lowest evenness. The logarithmic model fit best exclusively at the smallest scales and in species-poor metacommunities, primarily fish. Second, we tested our framework with the parameters of the most broadly used SAR model, the log–log form of the power model, using a structural equation model. This model supported our framework and revealed that the SAR slope was best predicted by scale- and organism-dependent metacommunity properties, particularly spatial aggregation, whereas the intercept responded most strongly to species group and γ-diversity. Future research

Ecology. 2023;104:e3917.
https://doi.org/10.1002/ecy.3917

should investigate from the perspective of our framework how shifts in metacommunity properties due to climate change may alter the SAR.

KEYWORDS
biodiversity, climate, diatoms, environmental heterogeneity, fish, insects, intraspecific spatial aggregation, metacommunity, scale, skewness, species abundance distribution, stream

INTRODUCTION

The species–area relationship (SAR) is an almost universal ecological phenomenon that describes how species richness increases with area or number of sampling units (Drakare et al., 2006; Lomolino, 2000; Matthews et al., 2021; Rosenzweig, 1995; Scheiner et al., 2011). It has a broad application in ecology for upscaling biodiversity from local to regional and larger scales, which are targets of conservation policies (Kunin et al., 2018), detection of biodiversity hotspots (Fattorini, 2021; Guilhaumon et al., 2008; Veech, 2000), and estimation of extinction with habitat loss (Chaudhary & Mooers, 2018; Fattorini et al., 2021; Pimm & Raven, 2000). Herein, we focus on SARs of noncontiguous habitats on the mainland (hereafter simply SARs), rather than island SARs, and describe how species richness changes with an increasing number of samples (individual stream reaches), rather than the size of an island or other isolate.

There are two major directions in the SAR research. The first is concerned with the form of the best-fitting SAR model, which is most often linear (logarithmic), concave (power), or saturating (Michaelis–Menten) when richness is expressed as a function of log-area (Figure 1; DeMalach et al., 2019; He & Legendre, 1996). However, this research has generally been limited to terrestrial systems (DeMalach et al., 2019; Dengler et al., 2020; Matthews et al., 2016; Triantis et al., 2012), leaving open important questions about the freshwater SAR, namely, how fast richness increases with area, whether it saturates, and if the model form depends on species group, scale, and the environment, similarly to terrestrial systems.

The second line of SAR research, also predominantly terrestrial, investigates what factors control the SAR shape, which is generally described by the parameters of the log–log form of the power SAR model because they have a biological interpretation. These factors include scale, organismal biology, and habitat characteristics (Drakare et al., 2006; Matthews et al., 2016, 2019; Rosenzweig, 1995; Triantis et al., 2012). They may influence the SAR shape directly or indirectly through metacommunity properties, including intraspecific spatial aggregation, regional species pool, and species abundance distribution (SAD) (Chase et al., 2019; Chase & Knight, 2013; DeMalach et al., 2019; May, 1975; Preston, 1962). For example, higher spatial aggregation (individuals of a species clustered together), larger species pools, and more uneven SADs lead to a steeper richness increase with area.

Both the SAR shape and its underlying mechanisms depend on the spatial scale at which the SAR is measured, including grain and extent (Palmer & White, 1994). As the scale increases, the SAR shape becomes saturating (He & Legendre, 1996; Matthews et al., 2016; Triantis et al., 2012). The relative importance of the mechanisms creating the SAR also shifts with dispersal, operating at small to intermediate scales, habitat diversity, primarily at intermediate scales, and speciation, mostly at the largest scales (Lomolino, 2000; Triantis et al., 2012; Turner & Tjørve, 2005).

The SAR and its drivers also depend on organismal biology. Weaker dispersers, which tend to be larger bodied, have more aggregated spatial distributions, more symmetrical SADs (Borda-de-Água et al., 2017), and steeper SARs (Drakare et al., 2006). However, how scale and species group affect the SAR relative to other factors is not well understood, particularly in freshwaters, where more SAR research is needed (Soininen & Teittinen, 2019). This is problematic not only for our fundamental understanding of the SAR but also from a conservation point of view. Streams and rivers are some of the most threatened ecosystems worldwide (Tickner et al., 2020; Wiens, 2016), and basic knowledge of biodiversity scaling is essential for determining what factors will require preservation or mitigation and at what scale, particularly under climate change. Therefore, the goal of this study was to determine how scale (spatial extent) and species groups differing in body size and dispersal capacity, including diatoms, insects, and fish, control the shape of the SAR. We hypothesized that the SAR shape would be defined by scale and species group, viewed here as exogenous factors, through their effect on hierarchical predictors of the SAR, including climatic and metacommunity factors (Figure 2).

Scale can affect metacommunity factors and the SAR indirectly through environmental factors. Aquatic

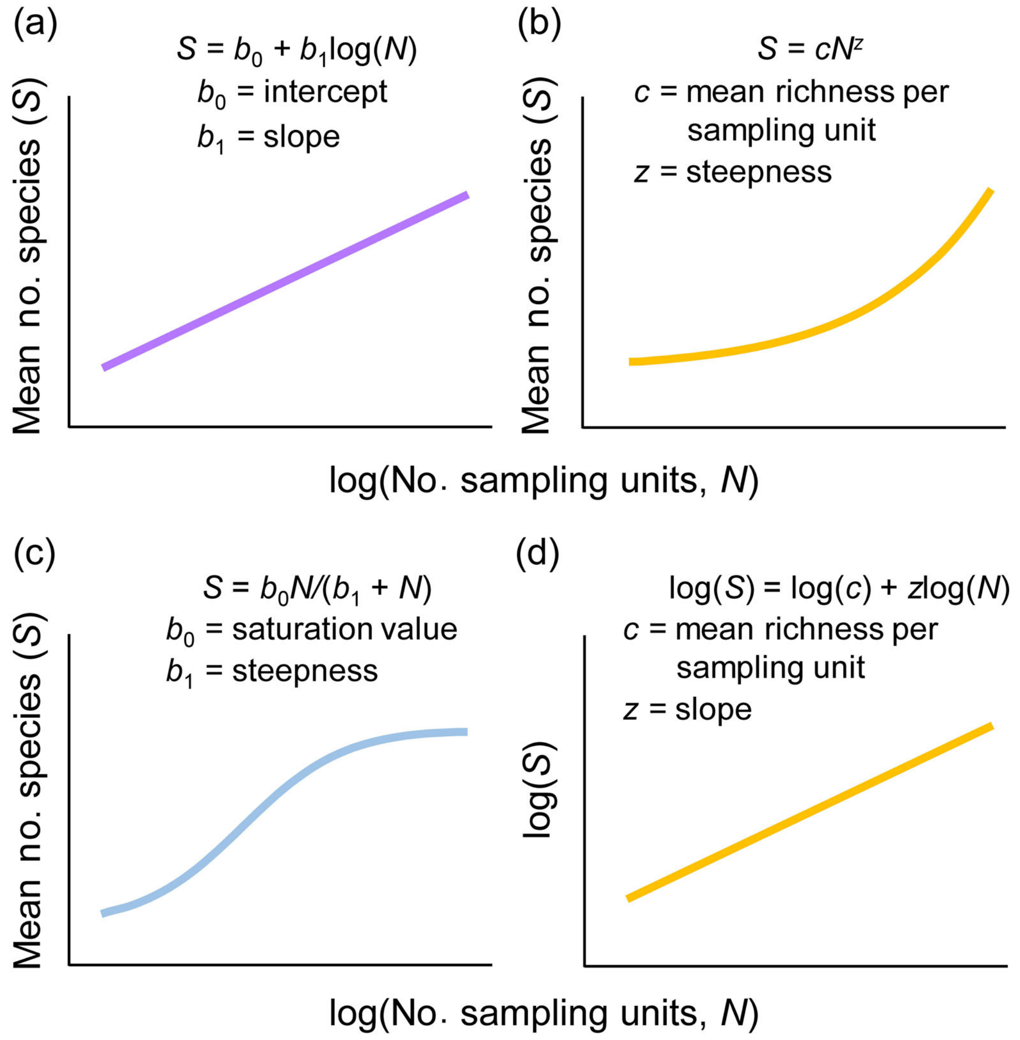

FIGURE 1 Common models of the species–area relationship and interpretation of their parameters (b_0, b_1, c, and z): (a) logarithmic, (b) power, (c) Michaelis–Menten, and (d) log–log form of the power model.

communities are influenced by a variety of environmental factors, but our research here was focused on climate (heterogeneity and medians) because at large scales climate outperformed local water physicochemistry and land use in determining the distributions of diatoms, insects, and fish in US streams (Pound et al., 2021). Similarly, our preliminary research showed that local physicochemistry and land use did not add substantially to the variance in the SAR already explained by climate (data not shown). Higher climatic heterogeneity at larger scales is expected to generate more aggregated conspecific distributions (from species sorting in their preferred habitats) and higher regional species richness, also known as γ-diversity (due to greater habitat diversity), both of which can contribute to a steeper SAR (Chase & Knight, 2013; Turner & Tjørve, 2005). To account for any scale-dependent trends in the environment with an impact on the metacommunity factors and the SAR, we also included median values of climatic variables in our model (Figure 2).

Scale can also affect metacommunity factors directly, for example, by controlling the veil line of the SAD and the rates of dispersal. This is shown in our conceptual model as a direct link between scale and metacommunity factors (Figure 2). First, as scale increases, rarer species in the left tail of the abundance distribution are progressively unveiled, and the SAD becomes more even and symmetric (Connolly et al., 2005; de Lima et al., 2020; Morlon et al., 2009). However, such a trend may emerge if the effect of scale is not separated from the effect of sample size, that is, when there are more samples at larger scales and, consequently, fewer rare species (de Lima et al., 2020). To control the influence of sample size on the SAD (and the SAR), our scale-explicit framework was tested with samples and individuals standardized across scales (Figure 3).

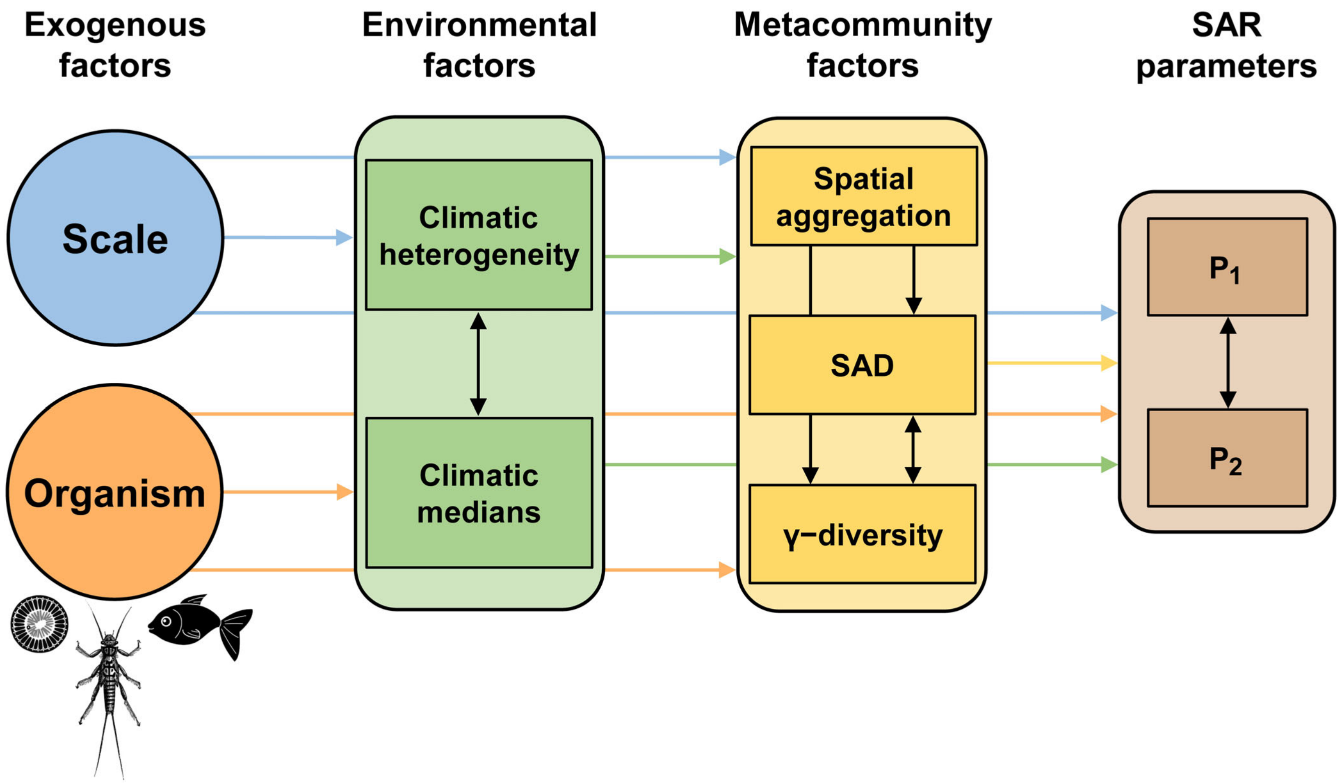

FIGURE 2 Conceptual framework depicting the pathways of scale and organismal control of parameters (P_1 and P_2) of the species–area relationship (SAR) through environmental and metacommunity factors. Circles: exogenous variables; rectangles: endogenous variables. Organism includes diatoms, insects, and fish. Because the three organismal groups were not always sampled from the same streams, the path between organism and environmental factors tests for differences in climatic conditions among these groups. Single-headed arrows = causal pathways, double-headed arrows = correlations. SAD, species abundance distribution.

Second, scale determines the rates of dispersal and, along with this, metacommunity properties. High dispersal at smaller scales can decrease γ-diversity via intraspecific disaggregation and subsequent homogenization and competitive exclusion (Ben-Hur & Kadmon, 2020; Chave et al., 2002; Matthiessen et al., 2010; Mouquet & Loreau, 2003). Conversely, dispersal limitation at larger scales may increase intraspecific spatial aggregation and, along with this, γ-diversity and evenness, given that greater aggregation causes individuals to be more strongly controlled by intraspecific than interspecific competition (Chave et al., 2002; Ives, 1991). In the proposed framework, spatial aggregation is viewed as a predictor of both γ-diversity and the SAD, which may be correlated, as shown for species richness and evenness (Soininen et al., 2012). Finally, metacommunity factors are predicted to vary across organismal groups as a result of differences in dispersal capacity.

The objectives of this research were to examine the following issues along regional to subcontinental scales and across major stream organisms: (i) the models that best fit the SAR, (ii) the factors that discriminate the best SAR model fits, and (iii) the factors underlying the variability of the SAR parameters from the perspective of our hierarchical framework.

MATERIALS AND METHODS

Data sets

Data were collected between 1993 and 2019 during the warmer months (May–September) from streams in all major watersheds in the United States by the National Water-Quality Assessment (NAWQA) Program of the US Geological Survey and the National Rivers and Streams Assessment (NRSA) of the US Environmental Protection Agency. However, the most intensive sampling occurred within a comparatively short period of time, in 2008 and 2009 for diatoms and insects and in 2008, 2009, 2013, and 2014 for fish. Diatoms and insects were sampled from a predefined area of substrate in 2278 and 2270 distinct localities, respectively. Fish were collected by electrofishing and seines from 2296 distinct localities. Diatoms and fish were identified to species and insects to genus. Given that the number of individuals per sample varied

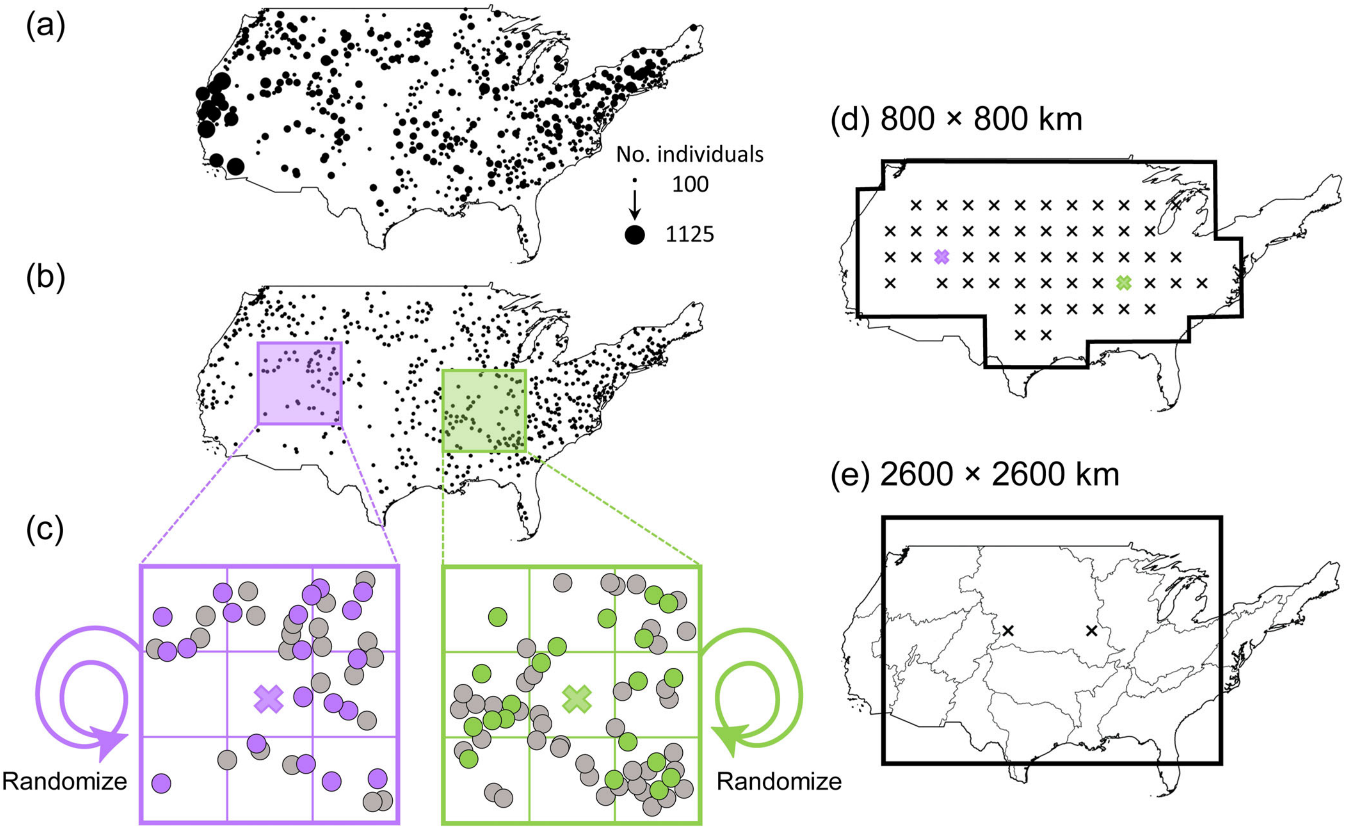

FIGURE 3 Analytical and spatial approaches using a subset of the insect data as an example. (a) Distribution of samples shown as circles with size proportional to the number of counted individuals. (b) Prior to all analyses, individuals were sampled down to 100 in each sample. Two 800 × 800-km windows are shown. (c) To ensure that a window adequately represents its respective scale, it was subdivided into nine subwindows and used in further analyses only if seven out of nine of the subwindows contained samples. A predefined number of samples was randomly pulled (shown in purple and green circles in the respective windows) *n* times, where *n* was proportional to the number of samples in the window. For example, the green window had 1.5 times more samples, and it was subsampled 1.5 times more than the purple window. Crosses indicate the center points of the windows. (d, e) Maps showing center points of windows of increasing size for two sample scales. The black polygons outline the area covered by all windows. In (d) the center points of the windows shown in (c) are color coded. The subdivision of the United States into hydrologic regions is given in (e).

and this could bias the richness estimates and the SAD, each sample was sampled down once to 400 individuals for diatoms and 100 individuals for insects and fish (Figure 3a,b). For each stream locality, bioclimatic data on temperature and precipitation averages, minima, maxima, ranges, seasonality, and extremes (19 variables) were retrieved from the WorldClim database (Hijmans et al., 2005).

Spatial design

Scale is defined in terms of grain and extent. Here, grain was kept constant and represented a single stream site, whereas extent was widely varied (Figure 3). To test for the effect of scale on the SAR and its drivers, we implemented a recently developed landscape window approach (Leboucher et al., 2019). This method allows subdivision of the study area into progressively larger square windows and, thus, the examination of a given pattern or relationship in a spatially explicit way. In the United States, successive windows increased by 200 km per side and measured from 400 × 400 km to 2600 × 2600 km, comprising 12 different scales (Figure 3d,e). To ensure that a window adequately represented its respective scale, each window was subdivided into nine subwindows and used in further analyses only if seven of nine of the subwindows contained samples. At each scale, the center points of the windows were spaced by one-third of the width of the window such that adjacent windows had some overlap. For example, the centers of the 800 × 800-km windows were spaced by 266.7 km (Figure 3d). The overlap was necessary to ensure maximal coverage of the study area. For instance, two overlapping 2600 × 2600-km windows covered most of the United States (Figure 3e). Conversely, two abutting (nonoverlapping) windows of the same size did not fit within the boundaries of the United States and failed to meet our seven-out-of-nine criterion. In this case, one of the windows had to be discarded, leaving a large area of the United States unsampled. To counterbalance potential pseudoreplication, that is, some of the samples in two adjacent windows are the same due to window overlap, we carried out randomizations.

Larger windows contained more samples, and differences in sample size can bias the estimates of the SAD and the SAR via sampling effects. Even within the same scale, windows varied in number of samples (Figure 3c). To account for this variability, we did the following. First, at all scales, the same number of samples (60) was randomly pulled without replacement from each window to measure environmental heterogeneity and median environmental conditions and to calculate biotic metrics. A sample size of 60 was large enough to estimate the SAR and allowed us to maximize the number of valid windows, meeting the seven-out-of-nine criterion. Second, in each window, the number of random pulls of 60 samples was in proportion to the number of samples in this window (Figure 3c). There were 1000 such randomizations at each scale, or 12,000 randomizations in total. Thus, all scales had the same number of samples, and all samples had the same number of individuals for a given organismal group, giving rise to metacommunities of the exact same number of individuals at all scales, that is, 24,000,000 diatoms (1000 random pulls × 60 samples × 400 individuals) and 6,000,000 insects and fish (1000 random pulls × 60 samples × 100 individuals).

Calculating and modeling the SAR

There are different ways of measuring the SAR, and here we adopted a Type IIIB species–area curve sensu Scheiner (2003), based on noncontiguous habitats (here stream sites) within a region. In essence, this is a sample-based rarefaction curve (Gotelli & Colwell, 2001), but in the literature, it is often referred to as a species accumulation curve, even though the two curves are not identical. For the generation of this curve, we used the "exact" method of Ugland et al. (2003) in the "specaccum" function in the vegan R package (Oksanen et al., 2019) to calculate mean species richness for all combinations of 1, 2, ..., 60 stream sites in each random pull at each of the 12 scales, without respect to the spatial position of these sites. Then mean species richness (S) was expressed as a function of the number of sampled sites and fitted with logarithmic, power, and Michaelis–Menten models (Figure 1a–c). The best-fitting model was selected based on corrected Akaike information criterion (AIC_c). Smaller AIC_c values indicate better fit, and absolute differences in AIC_c ($\Delta\text{AIC}_\text{c}$) > 2 indicate strong support for the better model. We also calculated the $\log_{10}$-$\log_{10}$ form of the power model (Figure 1d). The fit of this model was not compared to the fits of the other three models because such comparisons are not meaningful if S is log-transformed in some models but not transformed in others (Dengler et al., 2020).

Drivers of the SAR

According to our framework (Figure 2), the environmental drivers of the SAR include climatic medians and heterogeneity. Using a matrix of random pulls × centered and standardized median values for all climatic variables, we carried out a principal component analysis (PCA) for the data set containing all three species groups as well as for the data sets of individual species groups. PCA reduces the number of predictors and eliminates collinearity among them by generating synthetic orthogonal principal components (PCs). In all cases, only the first two PCs were significant (with eigenvalues exceeding those randomly generated by a broken-stick model) and used in further analyses. For each random pull of 60 samples, we calculated the median Euclidean distance of the centered and standardized climatic variables as a metric of climatic heterogeneity.

The metacommunity drivers of the SAR included intraspecific spatial aggregation, total richness (γ-diversity), and the SAD, which were calculated for each species group in each random pull from a window. A global dispersion parameter k across all taxa in a window was calculated using maximum-likelihood methods (Xu et al., 2015). Specifically, parameter k was obtained after maximizing the following log-likelihood function:

$$l=\sum_{i=1}^{\gamma}\left[o_i\log(p_i)+(m-o_i)\log(1-p_i)\right],$$

where γ is the total number of species in the random pull from a window, o_i is the number of sites occupied by species i, m is the total number of sites (60 in our case), and p_i is the predicted occupancy of species i from the negative binomial distribution. The predicted occupancy was calculated using the following equation:

$$p_i=1-\left(1+\frac{n_i}{mk}\right)^{-k},$$

where n_i is the abundance of species i. Ecologically, parameter k is interpreted as an intraspecific aggregation metric because it accounts for the degree of clumping in the spatial distribution of individuals among samples (White & Bennetts, 1996). As smaller values of parameter

k signify greater aggregation, for a more straightforward interpretation we obtained the reciprocal, $1/k$, which increases with aggregation.

The SAD was evaluated by two metrics—skewness, measuring the symmetry in species abundances about the mean, and Pielou's J, measuring the equitability in species abundances. Negative skewness indicates excess of rare species, whereas positive skewness shows excess of common species compared to the normal distribution. Skewness was calculated using the R package moments (Komsta & Novometsky, 2015).

Statistical analyses

Comparing metacommunities and species groups

Metacommunities best fit with logarithmic, power, and Michaelis–Menten models were compared with Cliff's $|d|$ statistic in terms of climatic and metacommunity factors. Using this statistic, we also compared species groups with respect to climatic and metacommunity factors and the SAR model parameters. Cliff's $|d|$ measures the probability that randomly selected values from two groups will be different and is unaffected by sample size. We adopted a $|d|$ of 0.33 as a threshold above which differences between two groups were considered significant (Jamoneau et al., 2018 and references therein).

Testing the factors controlling the best SAR model fit (logarithmic, power, Michaelis–Menten)

First, the frequency of the best SAR model fit (logarithmic, power, and Michaelis–Menten) was examined as a function of scale and species group. The response of this frequency to scale was modeled with a linear regression. Then we calculated the frequency of each of the three models when they were identified as best in each of the 12 scales for each species group. The frequencies of the three best SAR models across scales (response variables) were compared among species groups (factor) (12 scales $\times$ 3 species groups $= 36$) with multivariate ANOVA (MANOVA) and univariate F-tests, followed by Tukey's multiple comparison tests. Second, we tested which variables (scale, species group, climatic factors, or metacommunity properties) contributed the most to the separation of the best SAR models with discriminant function analysis (DFA). Linear DFA was carried out using the best SAR model as a grouping factor (logarithmic, power, and Michaelis–Menten) and the following as predictors: scale, species (diatoms and fish, represented by dummy variables, given that only two dummy variables were needed to represent three categories), climatic variables, and metacommunity properties. All variables that varied significantly among the three groups (Cliff's $|d| > 0.33$) were included in the DFA. However, since there were differences in the covariance matrices among the three groups, we proceeded with a quadratic DFA because it does not require equality of covariance matrices. Because many of the variables in the DFA model were correlated, we performed a backward selection to eliminate redundant variables, whereby the variable with the lowest scores on the two discriminant functions was removed. The backward selection continued until the Jackknife classification accuracy dropped. Regressions, MANOVA, and DFA were performed using SYSTAT 13.

Testing the factors controlling the SAR model parameters

The drivers of the SAR parameters in all species groups were tested with a structural equation model (SEM) after $\log_{10}$-transformation of scale and ln-transformation of γ-diversity. The SEM was generated with the R package lavaan (Rosseel, 2012). In the SEM, exogenous variables were scale and species group (coded as dummy variables with diatoms being the base level), whereas endogenous variables included climatic factors (climatic heterogeneity and PCA scores), metacommunity properties (aggregation parameter $1/k$, γ-diversity, and skewness), and the parameters of the SAR model. In addition to the overall SEM, SEMs were also run at a group level to test for among-group differences in causal paths. The SEM fit was evaluated by the root mean square error of approximation (RMSEA) statistic. Redundant pathways, which did not add to the explained variance, were eliminated. This was done by sequentially removing variables with low z-values until the RMSEA statistic became non-significant. The standardized total effect sizes (direct + indirect) of each predictor of the SAR parameters in the overall SEM were calculated using the standardized path coefficients.

RESULTS

Climatic variability

The first two climatic PCA axes using data on all three species groups explained, respectively, 52.8% and 27.0% of the variance in the data. The first axis was positively correlated with temperature and precipitation but negatively correlated with temperature range and precipitation seasonality. The second axis was primarily defined by temperature isothermality and seasonality, which correlated in opposite ways with the axis. The PCA scores did not differ significantly across species groups (Cliff's $|d| = 0.01–0.31$), nor did the median distances in climatic conditions (henceforth climatic heterogeneity) (Cliff's $|d| = 0.02–0.05$), indicating that the three species groups were sampled from similar climatic conditions. PCAs at a species group level identified the same major gradients.

Best-fitting SAR models and their drivers

All three SAR models provided very good fits with a minimum R^2 of 0.93 (logarithmic), 0.95 (power), and 0.98 (Michaelis–Menten) and a maximum R^2 approaching 1.00 in all models. Based on $\Delta\mathrm{AIC}_c$, Michaelis–Menten was the best SAR model in 78.0% of the random pulls and the power model in 17.1% of random pulls, whereas SARs best fit by the logarithmic model were comparatively rare, that is, 1.7% of the random pulls. SARs that were equally fit by two or three models were also rare (0.003%–3.0%) and are not discussed further.

Linear regressions of best model frequency against scale revealed that the frequency of the power model decreased, whereas the frequency of the Michaelis–Menten model increased with scale (Figure 4a). The logarithmic model frequency also declined with scale but was not modeled because only four scales exhibited logarithmic SARs. Then we compared the frequency of each model across species groups (Figure 4b). The frequency of the logarithmic model did not vary among groups ($p = 0.18$), whereas the frequency of the power and Michaelis–Menten models depended on species group, as confirmed by MANOVA (Wilks' $\lambda = 0.32$, $p < 0.001$) and univariate F-tests ($p < 0.005$). Tukey's post hoc tests further showed that Michaelis–Menten was significantly more frequently the best-fitting model in fish compared to diatoms ($p < 0.005$), whereas fish and insects did not differ significantly from each other ($p = 0.08$), nor did diatoms and insects ($p = 0.40$). Conversely, the power model had significantly lower frequency in fish than in diatoms and insects ($0.001 < p < 0.01$), which did not differ significantly from each other ($p = 0.40$).

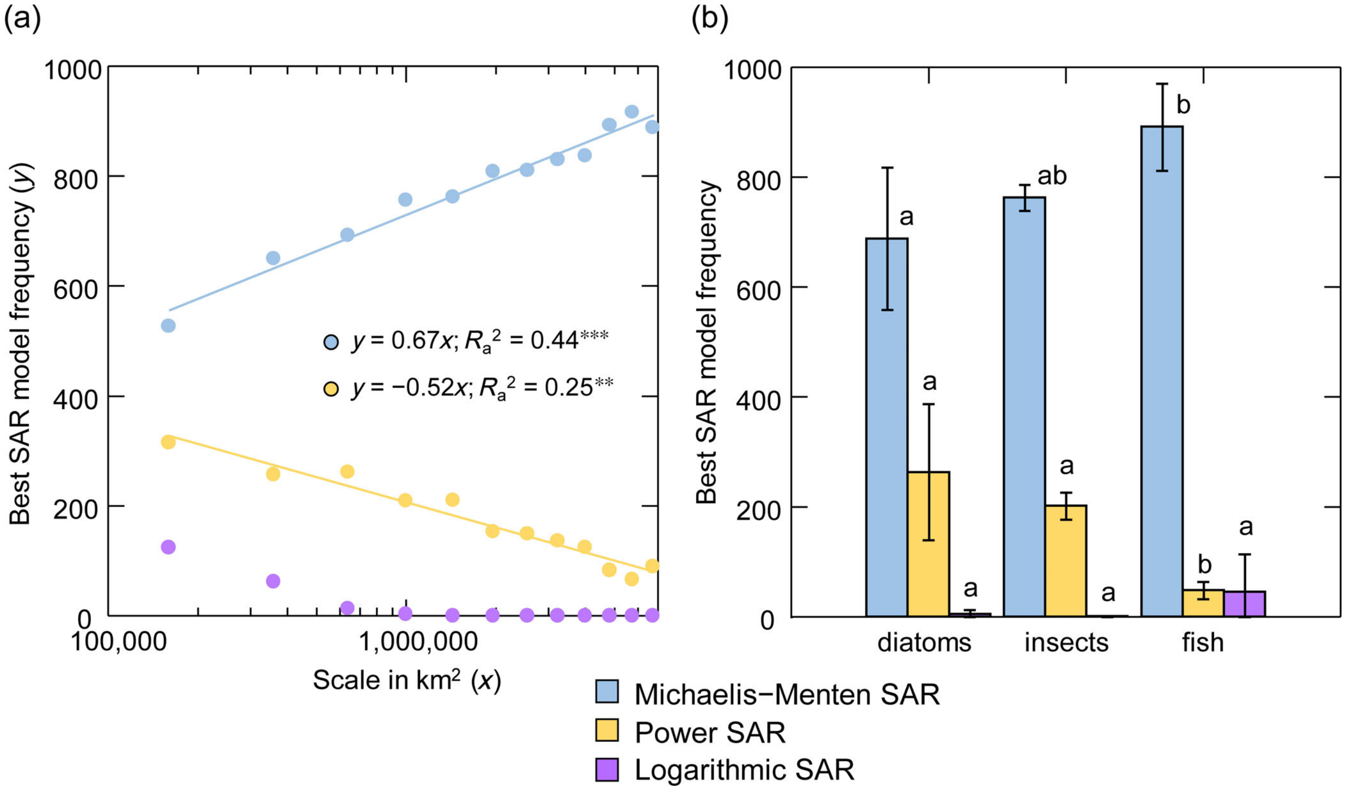

FIGURE 4 Frequency of best-fitting models for species–area relationship (SAR) across (a) scales and (b) species groups. In (a), the best SAR model frequency was calculated for each of three SAR models, Michaelis–Menten, power, and logarithmic, across all species groups and scales. This frequency was regressed against scale for the Michaelis–Menten model and the power model, and the standardized regression coefficients and adjusted R^2 (R_a^2) are provided. $^{**}p < 0.01$ and $^{***}p < 0.001$. In (b), the best model frequency was calculated for each scale and species group (i.e., 12 scales per species group), and then species groups were compared with ANOVA and a Tukey's post hoc test. Different letters in (b) indicate significant differences among groups ($n = 36$, $p < 0.05$).

Quadratic DFA identified five variables as having the strongest discriminative power with respect to the best SAR model (logarithmic, power, and Michaelis–Menten), including ln(spatial aggregation), ln(γ-diversity), skewness, PC2, and diatoms, and classified correctly 87% of the 34,858 cases following a Jackknife procedure (Figure 5). The logarithmic and Michaelis–Menten models were correctly classified in 93% and 95% of the cases, respectively, whereas the power model was correctly classified in only 51% of the cases. The first discriminant function was most strongly negatively correlated with ln(γ-diversity) and most strongly positively correlated with diatoms and skewness. The second discriminant function was most strongly negatively correlated with skewness and ln(spatial aggregation).

Testing the conceptual framework

Although the Michaelis–Menten was most frequently the best-fitting model, the log–log model with a minimum R^2 of 0.92 and a maximum R^2 of 0.99 also provided a very good fit of the SAR. Given that the parameters of the log–log SAR model are most broadly investigated and have biological meaning, we proceeded with further analyses of the SAR using these parameters. Thus, slope z reflects the rate of richness increase with the increase in number of stream localities, and intercept log(c) is the log of the mean richness at a stream locality (Figure 1d). Notably, slope z of the log–log model and the steepness parameter b_1 of the Michaelis–Menten model had a 97% correlation and intercept log(c) of the log–log model and saturation parameter b_0 of the Michaelis–Menten model had a 92% correlation. Therefore, results of the analyses of the parameters of the log–log model apply to the parameters of the Michaelis–Menten model, which was most frequently the best model.

Predicting climatic and metacommunity factors

Our overall SEM, testing the hierarchical effects on the SAR parameters (Figure 6a), had a very good fit

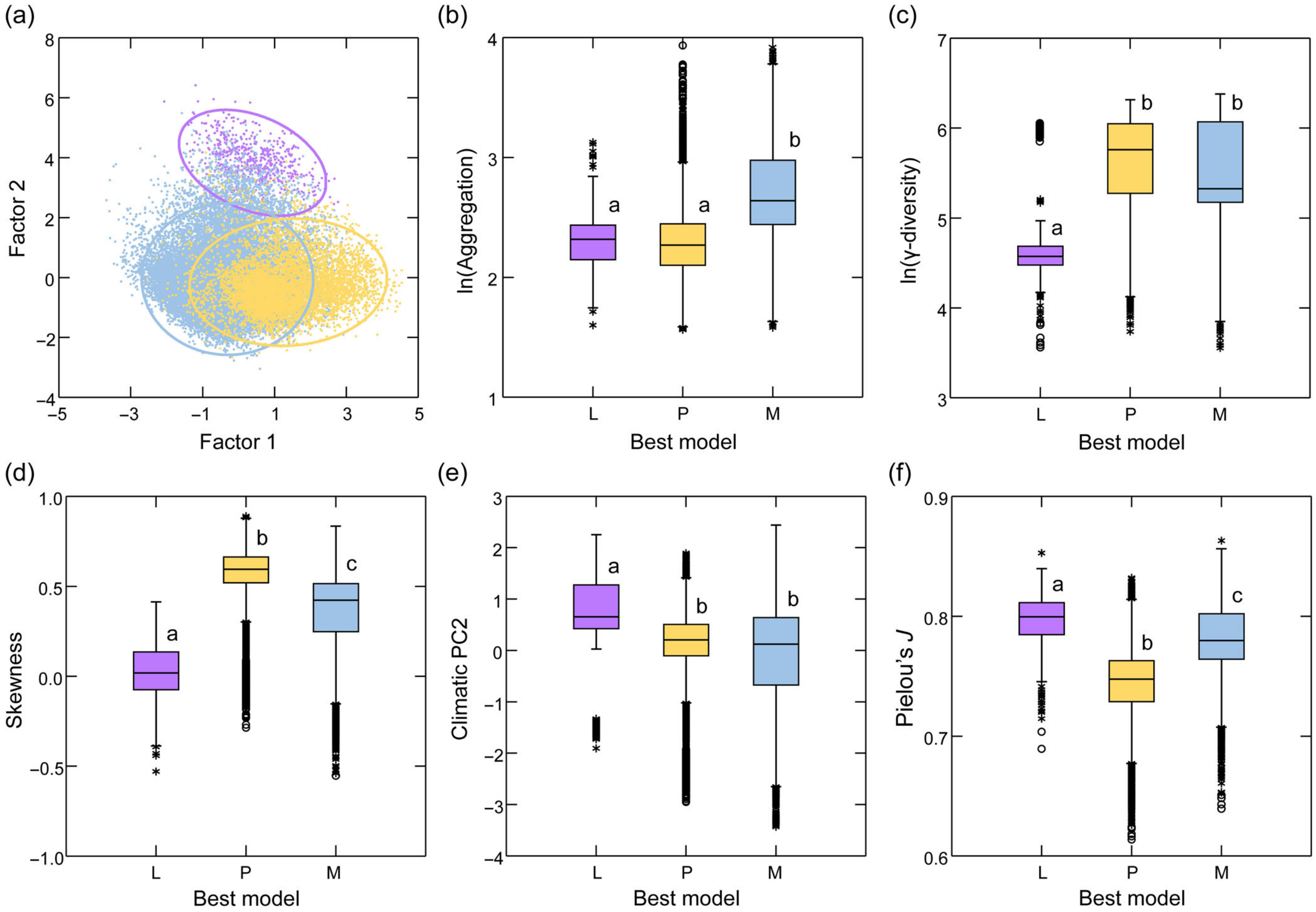

FIGURE 5 (a) Discriminant function analysis (DFA) of the best models of the species–area relationship (logarithmic, L, in purple; power, P, in yellow; and Michaelis–Menten, M, in blue) with 95% confidence ellipses. Box plots of variables included in the DFA (b–e) and Pielou's evenness (J) (f). In (b–f) different letters indicate significant differences (Cliff's $|d| > 0.33$). Asterisks indicate values outside the inner fences, while open circles denote values outside the outer fences.

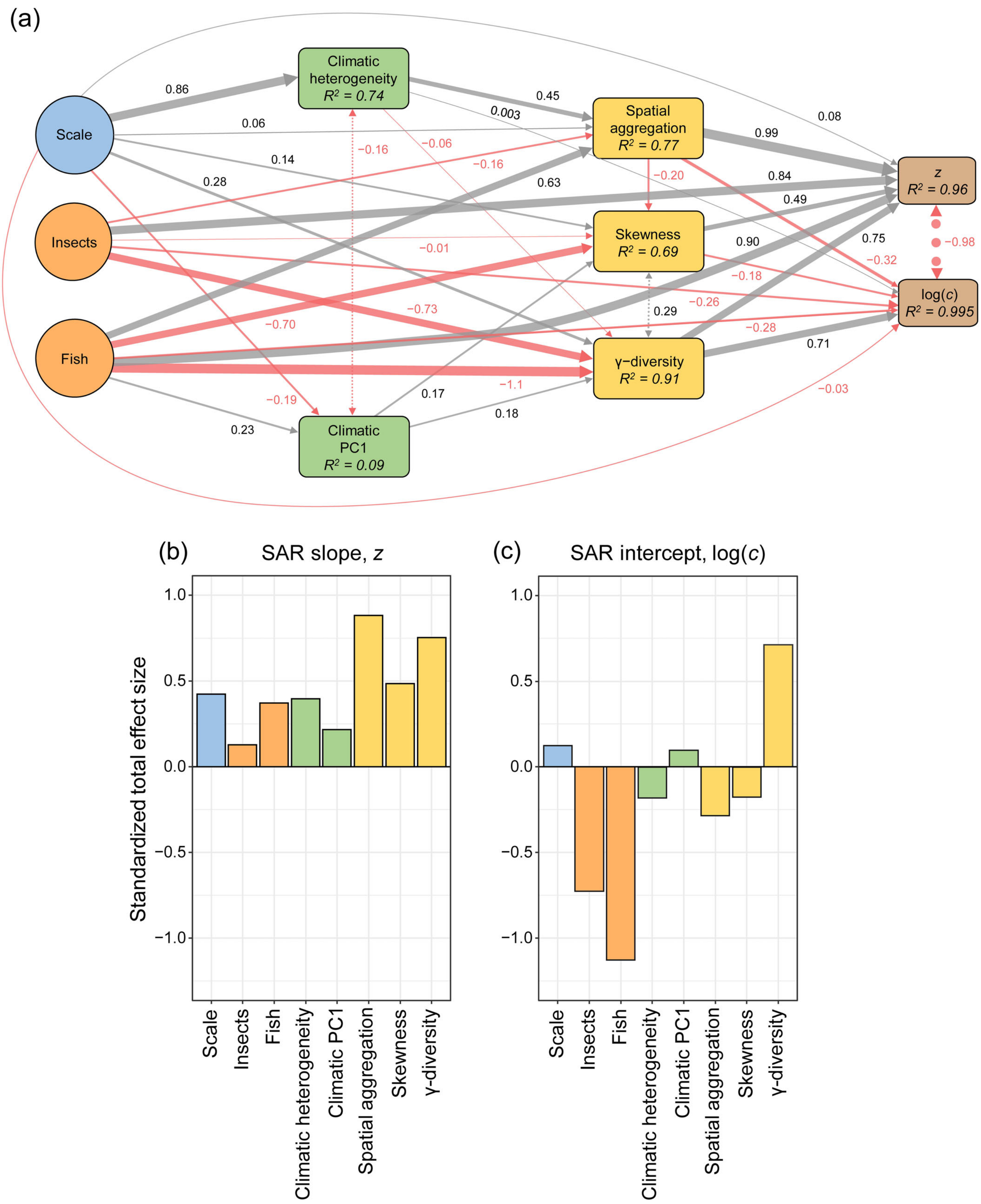

FIGURE 6 (a) Structural equation model testing the proposed framework for the pathways of control of slope, *z*, and intercept, log(*c*), of the log–log species–area relationship (SAR). Scale was $\log_{10}$-transformed and γ-diversity was ln-transformed. Single-headed solid arrows = causal pathways, double-headed dotted arrows = correlations with, respectively, standardized regression coefficients and correlation coefficients. Gray arrows = positive relationships, red arrows = negative relationships. R^2 is given for each endogenous variable. Standardized total effect sizes of scale, organism, climatic, and metacommunity factors for *z* (b) and log(*c*) (c) of the log–log SAR, derived from the structural equation model.

(RMSEA = 0.05, $p = 0.07$). From the climatic factors, heterogeneity was well predicted by scale ($R^2 = 0.74$), whereas PC1 had a very weak response to its predictors, scale and fish ($R^2 = 0.09$). Although fish entered the model, suggesting some differences in PC1 among species groups, these differences were in fact minimal, given the nonsignificant Cliff's $|d|$ results for PC1. The metacommunity variables, spatial aggregation, γ-diversity, and skewness, were well predicted by the exogenous and climatic factors, having an R^2 of 0.77, 0.91, and 0.69, respectively. Spatial aggregation was most strongly and positively affected by fish, followed by climatic heterogeneity, that is, the highest aggregation was observed in fish (Figure 7a) and in heterogeneous conditions. Species groups were the primary predictors of γ-diversity, followed by scale. As to be expected, γ-diversity was higher at large scales but lower in fish and insects relative to diatoms (Figure 7b). Skewness was most strongly and negatively controlled by fish, followed by spatial aggregation. Skewness was significantly lower in fish than in the other two groups (Figure 7c) and decreased with aggregation.

SEMs at a group level also had very good fits (RMSEA = 0.003–0.021, $p = 1.00$) and complemented the results from the overall SEM of the combined data (Appendix S1: Table S1). The group-level SEMs revealed that spatial aggregation was driven by climatic heterogeneity, similarly to the overall SEM, but also by scale and climatic medians. Although insects and/or fish were the strongest predictors of γ-diversity and skewness in the overall SEM, spatial aggregation, scale, and climatic medians were most influential at the group level.

Predicting the parameters of the log–log SAR

According to the overall SEM, the strongest and positive predictors of the SAR slope were, in descending order, spatial aggregation, fish, insects, and γ-diversity (Appendix S1: Table S1). The SAR intercept responded most strongly and positively to γ-diversity. Group-level SEMs indicated that spatial aggregation remained the strongest predictor of the SAR slope, whereas γ-diversity and

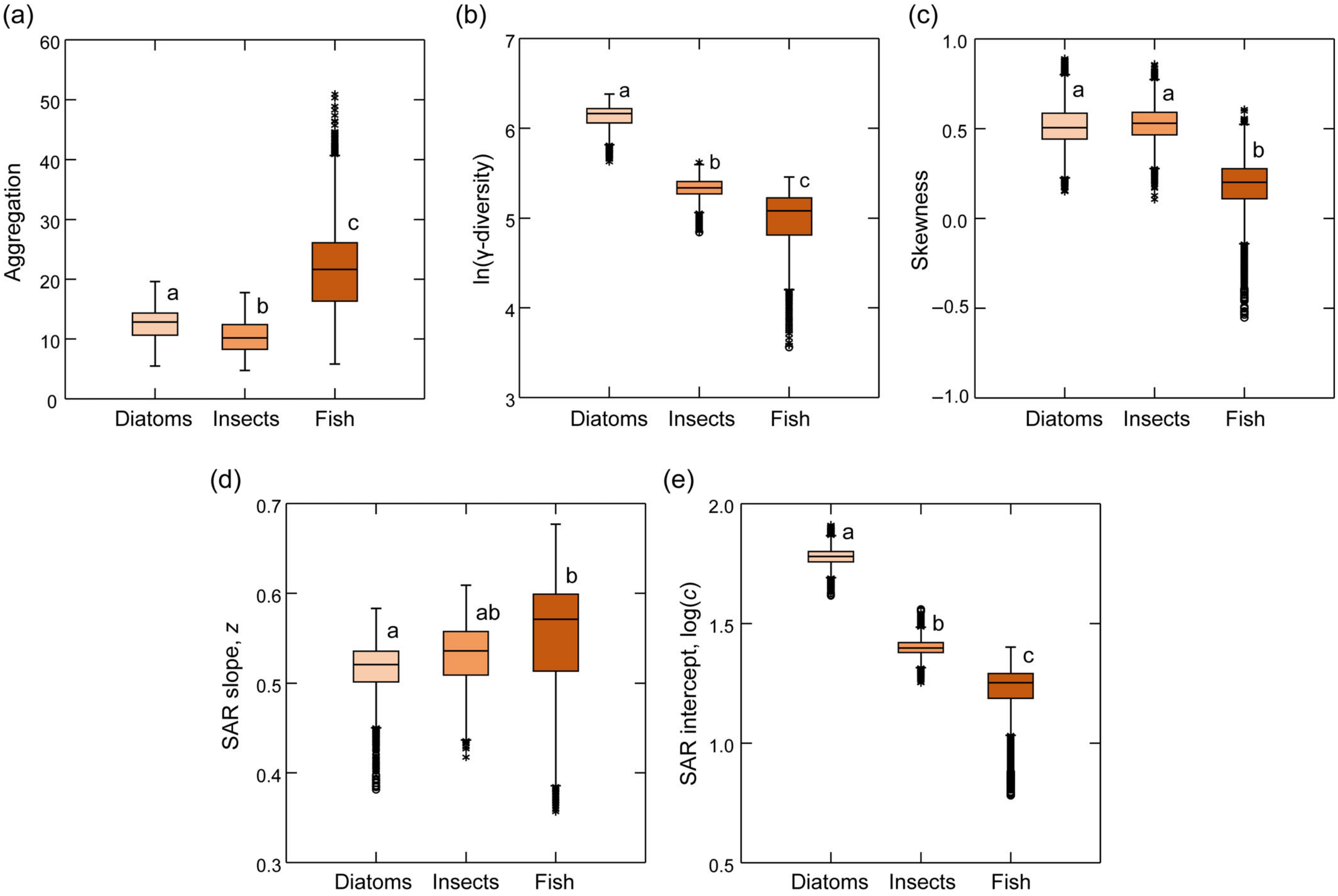

FIGURE 7 (a–e) Box plots of selected variables from the structural equation model in Figure 6a across species groups. In each panel, different letters indicate significant differences (Cliff's $|d| > 0.33$). Asterisks indicate values outside the inner fences, while open circles denote values outside the outer fences.

spatial aggregation most strongly constrained the SAR intercept (Appendix S1: Table S1). The fact that most of the variance in the SAR parameters (88%–99.5%) was explained in the two types of SEMs suggests that all important predictors were included.

Calculation of the standardized total effect sizes of all predictors of the SAR parameters in the overall SEM revealed that metacommunity factors (spatial aggregation, γ-diversity, and skewness) had the strongest effects on the slope, followed by scale, climatic heterogeneity, and fish (Figure 6b). The intercept, on the other hand, was primarily determined by species group (fish and insects), followed by γ-diversity (Figure 6c).

DISCUSSION

In this novel spatially explicit subcontinental study on the freshwater SAR, we determined the best model fit, the factors underlying the SAR shape, and the pathways of abiotic and biotic control of the parameters of the log–log model. First, we demonstrated that the best SAR model was most frequently Michaelis–Menten, but the frequency of all models (Michaelis–Menten, power, and logarithmic) varied significantly across species groups and/or scales. Second, we revealed that all three SAR models provided good fits and were comparatively well discriminated by metacommunity properties, species group, and climate. Third, we showed that the parameters of the log–log SAR were products of complex interactions among hierarchical factors. A comparison of the standardized total effect sizes of these factors indicated that metacommunity properties (spatial aggregation, γ-diversity, and skewness) had the strongest total effects on the slope, whereas species group (fish and insects) and γ-diversity had the strongest total effects on the intercept. In the following sections, we will discuss these findings in more detail.

Best-fitting SAR models and their drivers

The questions of what model provides the best fit for the SAR and under what circumstances, that is, at what scales, species groups, and characteristics of the metacommunity, have been central to ecology (Connor & McCoy, 1979; Dengler et al., 2020; He & Legendre, 1996; May, 1975; Preston, 1962; Tjørve et al., 2021). The species accumulation curve is predicted to be saturating (Dengler et al., 2020), although a logarithmic best fit has also been reported (DeMalach et al., 2019). In this study, the Michaelis–Menten model most often provided the best fit, but saturation was not observed, given that the saturation constant b_0 fell outside the range of the data. This is consistent with results from island SARs, where asymptotes were rarely detected (Triantis et al., 2012).

The interest in the shape of the SAR comes from its ability to capture the rate of species increase with area and its sensitivity to environmental and biological drivers. Thus, He and Legendre (1996) reported a transition from logarithmic to power to saturating SARs with the increase in scale and explained it as follows. At small scales, the logarithmic model, which has a slower richness increase with area compared to the other two models, is more likely because spatial autocorrelation of environmental factors and species composition results in a weak increase of species with area. At intermediate scales, the power model predominates because increasing heterogeneity produces a steeper richness increase with area. At the largest scales, the encounter of new habitats decreases, heterogeneity tapers off, and species richness saturates, which is best captured by an asymptotic model, such as the Michaelis–Menten model. Similar to He and Legendre (1996), we detected logarithmic SARs exclusively at small scales, power SARs, predominantly at smaller scales, and Michaelis–Menten SARs most frequently at the largest scales. However, the SAR best model frequencies further depended on metacommunity properties, species group, and the environment.

Both the logarithmic and the power SARs were observed in metacommunities with lower levels of intraspecific spatial aggregation. However, these metacommunities exhibited diverging symmetry and evenness of the regional abundance distribution and γ-diversity, and generally belonged to different species groups. Evenness (Pielou's *J*) was not included in the DFA model owing to the high correlation with variables already in the model (i.e., skewness, Pearson $r = -0.74$ and spatial aggregation, Pearson $r = 0.66$, $p < 0.001$ for both correlations). Nevertheless, since evenness had a strong discriminative power and its effect was transmitted by other variables, it is discussed here. Logarithmic SARs described species-poor and mostly fish metacommunities with significantly lower and sometimes negative skewness and significantly higher evenness. Power SARs, in contrast, described species-rich and predominantly diatom and insect metacommunities with significantly higher and generally positive skewness and significantly lower evenness than metacommunities best fit by the logarithmic or Michaelis–Menten model. Excess of rare species (negative skewness) would result in low richness in small areas, whereas high evenness would give rise to a slow richness increase toward the maximum richness of the metacommunity (Chase & Knight, 2013; He & Legendre, 2002; Tjørve et al., 2008). Because the logarithmic model predicts lower richness in small areas and is linear when richness is expressed as a function of log-area (Figure 1; Appendix S1: Figure S1), it performed better than the power model in metacommunities with higher rarity and evenness. Conversely, an excess of common species (positive skewness) and low evenness produce, respectively, high richness in small areas and a steep richness increase toward the maximum richness, best fit here by the power model. This model predicts higher richness in small areas and is concave, increasing steeply in large areas when richness is expressed as a function of log-area (Figure 1; Appendix S1: Figure S1). These results are partially consistent with a study on the species accumulation curve in plants where the power model was also detected predominantly in metacommunities with the lowest evenness (DeMalach et al., 2019); however, here the logarithmic rather than the Michaelis–Menten model was associated with the highest evenness. Our results further indicated that species dispersal ability might be a strong determinant of the SAR shape. Good dispersers have high skewness (Borda-de-Água et al., 2017), and here diatoms and insects, with significantly higher skewness than fish (Figure 7c), displayed significantly higher frequency of the power model (Figure 4b).

An additional limit on the richness increase with area was imposed by γ-diversity. Significantly lower γ-diversity in fish compared to the other two groups (Figure 7b) would require a smaller area to detect most species and can thus explain why the sigmoid Michaelis–Menten SAR was most common in fish. In contrast, significantly higher γ-diversity in diatoms and insects would explain the higher incidence of power SARs compared to fish. Increased performance of the power model in groups or data sets with higher mean richness was reported in plants (Dengler et al., 2020), implying that the effect of richness on the power SAR frequency may be more general.

We observed the greatest frequency of the Michaelis–Menten model at large scales, which were the most climatically heterogeneous. This is inconsistent with predictions for power SARs at high heterogeneity and intermediate scales (He & Legendre, 1996). We explained this discrepancy with the effect of spatial aggregation, which was highest at high climatic heterogeneity and large scales. Spatial aggregation, particularly of rare species, causes the SARs to become sigmoid (Tjørve et al., 2008). Indeed, SARs best fit here by the sigmoid Michaelis–Menten model were associated with significantly higher aggregation than SARs best fit by logarithmic and power models (Figure 5b). Additionally, metacommunities with Michaelis–Menten best fits had significantly more rare species (lower skewness) compared to metacommunities best fit by the power model (Figure 5d). Thus, greater aggregation in metacommunities with more rare species gave rise to Michaelis–Menten rather than power fits at large and climatically heterogeneous scales. Notably, the environment also had an impact on the form of the SAR—environments with higher temperature seasonality (high PC2 scores, Figure 5e) harbored metacommunities with logarithmic SARs. Although the shape of the SAR has been examined with respect to scale, organismal group, environment, and metacommunity properties (Connor & McCoy, 1979; DeMalach et al., 2019; Dengler et al., 2020; He & Legendre, 1996), this investigation is unique in addressing simultaneously all these sources of SAR variability and demonstrating the importance of species' dispersal capacity and scale-dependent metacommunity and environmental factors.

Testing the conceptual framework

Predicting climatic and metacommunity factors

All climatic and metacommunity factors except climatic medians (PC1 and PC2) were well predicted by our overall and group-level SEMs. Climatic medians were an exception because they were weakly correlated with scale and did not differ appreciably among the three species groups, which were sampled from the same regions. Climatic heterogeneity, on the other hand, sharply increased with scale, as is to be expected across the range of scales in this investigation.

High climatic heterogeneity resulted in high spatial aggregation in the overall and group-level SEMs, which is to be anticipated considering that all three study groups encompassed species with distinct thermal niches (Pound et al., 2021). However, climatic heterogeneity alone was not sufficient to describe spatial aggregation, implying the operation of factors other than climate. These factors were most likely organism-, dispersal-, and environment-related. Specifically, fish—the most dispersal-limited group in our study (Beisner et al., 2006; Shurin et al., 2009)—had significantly greater spatial aggregation than diatoms and insects and the strongest positive control over aggregation in the overall SEM. In diatom and fish SEMs, scale, climatic heterogeneity, and climatic medians had similar contributions, whereas in insects, climatic medians were the primary driver of spatial aggregation. Spatial aggregation increased with temperature and precipitation in diatoms but decreased in insects, whereas in fish, spatial aggregation increased with isothermality. Research on forests has shown that both habitat heterogeneity and limited dispersal work in concert to generate spatially aggregated species distributions (McFadden et al., 2019; Shen et al., 2009). To our knowledge, there is no such research in freshwaters, but our analysis implies that climatic heterogeneity and limited dispersal, together with climatic conditions, determine the spatial distributional patterns of species and genera in aquatic habitats and, thus, have broader importance across ecosystems.

In the overall SEM, γ-diversity was largely determined by species group—fish and insects had the strongest effects (negative). Notably, γ-diversity differed significantly among the three groups—it was the lowest in fish but the highest in diatoms. Contrary to our expectation for a positive response of γ-diversity to climatic heterogeneity and spatial aggregation, γ-diversity showed little sensitivity to climatic heterogeneity and did not depend on aggregation in our overall model. This was because the heterogeneity and aggregation effects on γ-diversity were largely subsumed by organismal group, given that within-group correlations of γ-diversity with climatic heterogeneity and aggregation showed modest to strong positive responses (fish: Pearson $r = 0.17–0.21$; insects: Pearson $r = 0.66–0.72$; and diatoms: Pearson $r = 0.56–0.81$, $p < 0.001$ for all correlations). Group-level SEMs confirmed that spatial aggregation was indeed among the strongest positive predictors of γ-diversity across groups, as anticipated. In these models, climatic heterogeneity did not emerge as a strong determinant of γ-diversity, most likely because its effect was subsumed by scale.

We expected more even and symmetric SADs at large scales, consistent with earlier investigations (Connolly et al., 2005; Morlon et al., 2009). The SAD did become more even with scale, which correlated positively with Pielou's J (Pearson $r = 0.37$, $p < 0.001$). The response of skewness to scale, on the other hand, was very weakly positive in our overall SEM due to divergent scale responses among groups, that is, negative in diatoms (Pearson $r = -0.40$, $p < 0.001$) but positive in insects and fish (Pearson $r = 0.17–0.26$, $p < 0.001$). For trees, beetles, and birds, the influence of scale on the SAD symmetry was due primarily to sampling effects, whereby more samples at large scales reduced the number of rare species (de Lima et al., 2020). Here, the number of samples was constant across scales, but group-specific responses caused scale to exert only a negligible overall influence on the SAD symmetry. De Lima et al. (2020) showed that spatial aggregation was the second most important determinant of SAD symmetry after sampling effects. Here, spatial aggregation had a negative effect on skewness in the overall SEM, but this effect was much stronger in the group-level SEMs. In the overall SEM, skewness was mostly defined by fish, which had higher symmetry, possibly due to weaker dispersal, leading to a lower frequency of common species and occasional preponderance of rare species (negative skewness).

Predicting the parameters of the log–log SAR

The SAR is a product of multiple mechanisms, operating at different scales, including sampling effects, environmental heterogeneity, dispersal, and evolutionary independence (Connor & McCoy, 1979; Rosenzweig, 1995; Scheiner et al., 2011; Turner & Tjørve, 2005). Additionally, the SAR varies across organismal groups, but investigations of multiple organisms inhabiting the same region are quite rare, despite the advantage of being unbiased by differences among taxa in the environment and geologic history (Fattorini et al., 2017). Here we examined the drivers of the SAR in diatoms, insects, and fish from all major US watersheds across regional to subcontinental scales after controlling for sampling effects. Our SEMs provided evidence for the proposed conceptual framework and further identified the most influential pathways of SAR control, which differed between the slope and the intercept.

Spatial aggregation had the strongest total and direct effects on the SAR slope, and, consistently with expectations (Scheiner et al., 2011), the direct effect was positive. Thus, as the scale and climatic heterogeneity increased, conspecific aggregation into optimal climatic niches resulted in steeper SAR slopes in the overall and group-level SEMs. Tjørve et al. (2008) reported that unevenness had a stronger impact on the shape of the SAR than spatial aggregation, while DeMalach et al. (2019) observed comparable influences. In our case, spatial aggregation had a stronger effect on the SAR slope than skewness across SEMs. Spatial aggregation had a negative effect on the SAR intercept, which is also to be expected, as more clumped distributions of individuals within a species produce lower community richness at local scales (Chase & Knight, 2013).

The size of the species pool is an important driver of the SAR (Chase & Knight, 2013; Matthews et al., 2019), strongly influencing how fast and high the SAR rises, for example, due to metacommunity processes such as dispersal rates (Leibold & Chase, 2017; O'Sullivan et al., 2019). However, the effect of γ-diversity on the SAR relative to the influence of climate, spatial aggregation, and the SAD is not well understood. Here we showed that γ-diversity, a proxy of the species pool, had the second greatest total effect on the SAR slope followed by skewness in the overall SEM. Metacommunities with high γ-diversity had high skewness, both of which resulted in steep SAR slopes. Although common species reduce the SAR slope, in this investigation, high skewness (excess of common species) was associated with low evenness of the SAD, which elevates the SAR slope. γ-diversity also had strong direct and total effects (both positive) on the SAR intercept, revealing that metacommunities with greater regional species pools had higher average local richness. At a species group level, γ-diversity was a less important predictor of the SAR slope but remained the strongest or the second strongest predictor of the SAR intercept.

The relationship of the SAR slope with scale is not straightforward as negative, positive, and unimodal responses have been noted. Triantis et al. (2012) reported that as scale increased, more factors contributed to the SAR on islands, causing the slope to decrease. Conversely, as scale increases and the area begins to cover different biogeographic provinces without a common species pool, the SAR slope may increase (Rosenzweig, 1995). A study on British flora documented a unimodal response, with low SAR slopes at small scales due to species interactions, high slopes at intermediate scales due to habitat heterogeneity, and again low slopes at large scales due to low species turnover (Crawley & Harral, 2001). In our study of regional to subcontinental areas, we saw strong positive correlations of the SAR slope with scale across the three groups (Pearson $r = 0.56–0.86$, $p < 0.001$). However, our overall and group-level SEMs revealed that scale constrained the slope only weakly positively, whereas much of the scale effects were indirect, that is, through climatic heterogeneity and γ-diversity.

The scale effect on the slope could be attributed to some extent to evolutionary independence. Evolutionary independence is an important positive driver of the SAR slope at the largest scales, where inclusion of different provinces with distinct evolutionary histories and species pools generates the steepest SAR slopes and the greatest overall richness (Rosenzweig, 1995; Turner & Tjørve, 2005). To test whether evolutionary independence could have contributed to the SAR slope, we calculated the percentage of species unique to the 14 hydrologic regions enclosed in the two largest windows (2600 × 2600 km; Figure 3e). We observed 0%–10% (4% average) unique diatoms, 0%–16% (4%–5% average) unique insects, and 0%–48% (14%–15% average) unique fish. This suggests that at large scales, the encounter of species unique to a hydrologic region may have influenced both the SAR slope and γ-diversity. Scale also constrained the SAR intercept, but its effect was indirect, mostly through γ-diversity. As is to be expected, greater species pools at large scales were responsible for higher intercepts.

Finally, the SAR was dependent on species group in our overall SEM. Consistent with earlier research showing that the SAR slope increased with body size (Azovsky, 2002; Drakare et al., 2006), we observed that fish exhibited significantly steeper SAR slopes than diatoms, whereas insects had a SAR slope intermediate between diatoms and fish (Figure 7d). In the overall SEM, fish and insects had a strong, positive, and direct influence on the SAR slope as well as an indirect influence through spatial aggregation, skewness, and/or γ-diversity. Fish and insects also had a weakly negative direct effect on the SAR intercept but a strongly negative total effect due to their influence on γ-diversity. Thus, species-poor, dispersal-limited, and spatially aggregated fish had SARs with higher slopes but lower intercepts, whereas species-rich diatoms with comparatively weak dispersal limitation and lower level of aggregation showed the opposite trend. A decrease in the SAR intercept from autotrophs to invertebrates and finally to vertebrates (Triantis et al., 2012), and from diatoms to macroinvertebrates (Azovsky, 2002), as seen here (Figure 7e), has been explained with the ability of autotrophs and smaller organisms to pack more species per unit area.

In conclusion, we examined the best model fits and underlying mechanisms of the SAR across scales and major species groups for the first time in streams. Metacommunity properties were among the strongest predictors of the SAR shape, as shown previously (Chase & Knight, 2013; DeMalach et al., 2019; He & Legendre, 2002; Tjørve et al., 2008), but we further traced their origins to species group, scale, and climate, broadening our understanding of the hierarchical drivers of the SAR. Our results have important implications for climate change research. Altered species/genus distributions and richness of stream diatoms, insects, and fish projected with climate change (Pound et al., 2021) may have profound impacts on the spatial aggregation, γ-diversity, and species abundance distribution and, consequently, on the SAR shape. Therefore, future studies should explore how climate change may affect the scaling of biodiversity in streams from the perspective of our framework.

AUTHOR CONTRIBUTIONS

Sophia I. Passy conceived the ideas, designed the research, and wrote the paper with contributions from all co-authors. Joseph L. Mruzek, William R. Budnick, Chad A. Larson, Sophia I. Passy, Thibault Leboucher, and Aurélien Jamoneau analyzed the data, and Chad A. Larson compiled the data.

ACKNOWLEDGMENTS

We thank Sébastien Boutry for help with the diatom harmonization, Alethea Passy-Crow for creating artwork, and two anonymous reviewers for insightful comments, which improved our work.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

Data (Passy et al., 2022a) are available in Dryad at https://doi.org/10.5061/dryad.4tmpg4fdq. R code (Passy et al., 2022b) is available in Zenodo at https://doi.org/10.5281/zenodo.7196013.

ORCID

Sophia I. Passy https://orcid.org/0000-0002-8230-9380
Joseph L. Mruzek https://orcid.org/0000-0002-8067-3824
William R. Budnick https://orcid.org/0000-0001-9288-6782
Thibault Leboucher https://orcid.org/0000-0003-0353-8896
Aurélien Jamoneau https://orcid.org/0000-0003-1181-2269
Jonathan M. Chase https://orcid.org/0000-0001-5580-4303
Janne Soininen https://orcid.org/0000-0002-8583-3137
Eric R. Sokol https://orcid.org/0000-0001-5923-0917
Juliette Tison-Rosebery https://orcid.org/0000-0003-1090-7020
Annika Vilmi https://orcid.org/0000-0001-5694-5802
Jianjun Wang https://orcid.org/0000-0001-7039-7136
Chad A. Larson https://orcid.org/0000-0002-0329-8979

REFERENCES

Azovsky, A. I. 2002. "Size-Dependent Species-Area Relationships in Benthos: Is the World more Diverse for Microbes?" *Ecography* 25: 273–82.

Beisner, B. E., P. R. Peres, E. S. Lindström, A. Barnett, and M. L. Longhi. 2006. "The Role of Environmental and Spatial Processes in Structuring Lake Communities from Bacteria to Fish." *Ecology* 87: 2985–91.

Ben-Hur, E., and R. Kadmon. 2020. "Disentangling the Mechanisms Underlying the Species-Area Relationship: A Mesocosm Experiment with Annual Plants." *Journal of Ecology* 108: 2376–89.

Borda-de-Água, L., R. J. Whittaker, P. Cardoso, F. Rigal, A. M. C. Santos, I. R. Amorim, A. Parmakelis, K. A. Triantis, H. M. Pereira, and P. A. V. Borges. 2017. "Dispersal Ability Determines the Scaling Properties of Species Abundance Distributions: A Case Study Using Arthropods from the Azores." *Scientific Reports* 7: 3899.

Chase, J. M., L. Gooriah, F. May, W. A. Ryberg, M. S. Schuler, D. Craven, and T. M. Knight. 2019. "A Framework for Disentangling Ecological Mechanisms Underlying the Island Species-Area Relationship." *Frontiers of Biogeography* 11: 1–17.

Chase, J. M., and T. M. Knight. 2013. "Scale-Dependent Effect Sizes of Ecological Drivers on Biodiversity: Why Standardised Sampling Is Not Enough." *Ecology Letters* 16: 17–26.

Chaudhary, A., and A. O. Mooers. 2018. "Terrestrial Vertebrate Biodiversity Loss under Future Global Land Use Change Scenarios." *Sustainability* 10: 2764.

Chave, J., H. C. Muller-Landau, and S. A. Levin. 2002. "Comparing Classical Community Models: Theoretical Consequences for Patterns of Diversity." *American Naturalist* 159: 1–23.

Connolly, S. R., T. P. Hughes, D. R. Bellwood, and R. H. Karlson. 2005. "Community Structure of Corals and Reef Fishes at Multiple Scales." *Science* 309: 1363–5.

Connor, E. F., and E. O. McCoy. 1979. "The Statistics and Biology of the Species-Area Relationships." *American Naturalist* 113: 791–833.

Crawley, M. J., and J. E. Harral. 2001. "Scale Dependence in Plant Biodiversity." *Science* 291: 864–8.

de Lima, R. A. F., P. A. Condé, C. Banks-Leite, R. C. Campos, M. I. M. Hernández, R. R. Rodrigues, and P. I. Prado. 2020. "Disentangling the Effects of Sampling Scale and Size on the Shape of Species Abundance Distributions." *PLoS One* 15: e0238854.

DeMalach, N., H. Saiz, E. Zaady, and F. T. Maestre. 2019. "Plant Species-Area Relationships Are Determined by Evenness, Cover and Aggregation in Drylands Worldwide." *Global Ecology and Biogeography* 28: 290–9.

Dengler, J., T. J. Matthews, M. J. Steinbauer, S. Wolfrum, S. Boch, A. Chiarucci, T. Conradi, et al. 2020. "Species-Area Relationships in Continuous Vegetation: Evidence from Palaearctic Grasslands." *Journal of Biogeography* 47: 72–86.

Drakare, S., J. J. Lennon, and H. Hillebrand. 2006. "The Imprint of the Geographical, Evolutionary and Ecological Context on Species-Area Relationships." *Ecology Letters* 9: 215–27.

Fattorini, S. 2021. "The Identification of Biodiversity Hotspots Using the Species-Area Relationship." In *The Species-Area Relationship: Theory and Application*, edited by T. J. Matthews, K. A. Triantis, and R. J. Whittaker, 321–44. Cambridge: Cambridge University Press.

Fattorini, S., P. A. V. Borges, L. Dapporto, and G. Strona. 2017. "What Can the Parameters of the Species-Area Relationship (SAR) Tell Us? Insights from Mediterranean Islands." *Journal of Biogeography* 44: 1018–28.

Fattorini, S., W. Ulrich, and T. J. Matthews. 2021. "Using the Species-Area Relationship to Predict Extinctions Resulting from Habitat Loss." In *The Species-Area Relationship: Theory and Application*, edited by T. J. Matthews, K. A. Triantis, and R. J. Whittaker, 345–67. Cambridge: Cambridge University Press.

Gotelli, N. J., and R. K. Colwell. 2001. "Quantifying Biodiversity: Procedures and Pitfalls in the Measurement and Comparison of Species Richness." *Ecology Letters* 4: 379–91.

Guilhaumon, F., O. Gimenez, K. J. Gaston, and D. Mouillot. 2008. "Taxonomic and Regional Uncertainty in Species-Area Relationships and the Identification of Richness Hotspots." *Proceedings of the National Academy of Sciences of the United States of America* 105: 15458–63.

He, F., and P. Legendre. 1996. "On Species-Area Relations." *American Naturalist* 148: 719–37.

He, F., and P. Legendre. 2002. "Species Diversity Patterns Derived from Species-Area Models." *Ecology* 83: 1185–98.

Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis. 2005. "Very High Resolution Interpolated Climate Surfaces for Global Land Areas." *International Journal of Climatology* 25: 1965–78.

Ives, A. R. 1991. "Aggregation and Coexistence in a Carrion Fly Community." *Ecological Monographs* 61: 75–94.

Jamoneau, A., S. I. Passy, J. Soininen, T. Leboucher, and J. Tison-Rosebery. 2018. "Beta Diversity of Diatom Species and Ecological Guilds: Response to Environmental and Spatial Mechanisms along the Stream Watercourse." *Freshwater Biology* 63: 62–73.

Komsta, L., and F. Novometsky. 2015. "Moments: Moments, Cumulants, Skewness, Kurtosis and Related Tests." R Package Version 0.14. https://cran.r-project.org/package=moments.

Kunin, W. E., J. Harte, F. He, C. Hui, R. T. Jobe, A. Ostling, C. Polce, et al. 2018. "Upscaling Biodiversity: Estimating the Species-Area Relationship from Small Samples." *Ecological Monographs* 88: 170–87.

Leboucher, T., W. R. Budnick, S. I. Passy, S. Boutry, A. Jamoneau, J. Soininen, W. Vyverman, and J. Tison-Rosebery. 2019. "Diatom Beta-Diversity in Streams Increases with Spatial Scale and Decreases with Nutrient Enrichment across Regional to Sub-Continental Scales." *Journal of Biogeography* 46: 734–44.

Leibold, M. A., and J. M. Chase. 2017. *Metacommunity Ecology*. Princeton, NJ: Princeton University Press.

Lomolino, M. V. 2000. "Ecology's most General, Yet Protean Pattern: The Species-Area Relationship." *Journal of Biogeography* 27: 17–26.

Matthews, T. J., F. Guilhaumon, K. A. Triantis, M. K. Borregaard, and R. J. Whittaker. 2016. "On the Form of Species-Area Relationships in Habitat Islands and True Islands." *Global Ecology and Biogeography* 25: 847–58.

Matthews, T. J., F. Rigal, K. A. Triantis, and R. J. Whittaker. 2019. "A Global Model of Island Species-Area Relationships." *Proceedings of the National Academy of Sciences of the United States of America* 116: 12337–42.

Matthews, T. J., K. A. Triantis, and R. J. Whittaker, eds. 2021. *The Species-Area Relationship: Theory and Application*. Cambridge: Cambridge University Press.

Matthiessen, B., E. Mielke, and U. Sommer. 2010. "Dispersal Decreases Diversity in Heterogeneous Metacommunities by Enhancing Regional Competition." *Ecology* 91: 2022–33.

May, R. M. 1975. "Patterns of Species Abundance and Diversity." In *Ecology and Evolution of Communities*, edited by M. L. Cody and J. M. Diamond, 81–120. Cambridge, MA: Harvard University Press.

McFadden, I. R., M. K. Bartlett, T. Wiegand, B. L. Turner, L. Sack, R. Valencia, and N. J. B. Kraft. 2019. "Disentangling the Functional Trait Correlates of Spatial Aggregation in Tropical Forest Trees." *Ecology* 100: e02591.

Morlon, H., E. P. White, R. S. Etienne, J. L. Green, A. Ostling, D. Alonso, B. J. Enquist, et al. 2009. "Taking Species Abundance Distributions beyond Individuals." *Ecology Letters* 12: 488–501.

Mouquet, N., and M. Loreau. 2003. "Community Patterns in Source-Sink Metacommunities." *American Naturalist* 162: 544–57.

Oksanen, J., F. G. Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlinn, P. R. Minchin, et al. 2019. "Vegan: Community Ecology Package." R Package Version 2.5-4. https://cran.r-project.org/package=vegan.

O'Sullivan, J. D., R. J. Knell, and A. G. Rossberg. 2019. "Metacommunity-Scale Biodiversity Regulation and the Self-Organised Emergence of Macroecological Patterns." *Ecology Letters* 22: 1428–38.

Palmer, M. W., and P. S. White. 1994. "Scale Dependence and the Species-Area Relationship." *American Naturalist* 144: 717–40.

Passy, S., J. Mruzek, W. Budnick, T. Leboucher, A. Jamoneau, J. Chase, J. Soininen, et al. 2022a. "Data from: On the Shape and Origins of the Freshwater Species-Area Relationship." Dryad, Data Set. https://doi.org/10.5061/dryad.4tmpg4fdq.

Passy, S., J. Mruzek, W. Budnick, T. Leboucher, A. Jamoneau, J. Chase, J. Soininen, et al. 2022b. "Data from: On the Shape and Origins of the Freshwater Species-Area Relationship." Zenodo, Software. https://doi.org/10.5281/zenodo.7196013.

Pimm, S. L., and P. Raven. 2000. "Extinction by Numbers." *Nature* 403: 843–5.

Pound, K. L., C. A. Larson, and S. I. Passy. 2021. "Current Distributions and Future Climate-Driven Changes in Diatoms, Insects and Fish in US Streams." *Global Ecology and Biogeography* 30: 63–78.

Preston, F. W. 1962. "Canonical Distribution of Commonness and Rarity: Part I." *Ecology* 43: 185–215.

Rosenzweig, M. L. 1995. *Species Diversity in Space and Time*. Cambridge: Cambridge University Press.

Rosseel, Y. 2012. "Lavaan: An R Package for Structural Equation Modeling." *Journal of Statistical Software* 48: 1–36.

Scheiner, S. M. 2003. "Six Types of Species-Area Curves." *Global Ecology and Biogeography* 12: 441–7.

Scheiner, S. M., A. Chiarucci, G. A. Fox, M. R. Helmus, D. J. McGlinn, and M. R. Willig. 2011. "The Underpinnings of the Relationship of Species Richness with Space and Time." *Ecological Monographs* 81: 195–213.

Shen, G., M. Yu, X.-S. Hu, X. Mi, H. Ren, I.-F. Sun, and K. Ma. 2009. "Species-Area Relationships Explained by the Joint Effects of Dispersal Limitation and Habitat Heterogeneity." *Ecology* 90: 3033–41.

Shurin, J. B., K. Cottenie, and H. Hillebrand. 2009. "Spatial Autocorrelation and Dispersal Limitation in Freshwater Organisms." *Oecologia* 159: 151–9.

Soininen, J., S. Passy, and H. Hillebrand. 2012. "The Relationship between Species Richness and Evenness: A Meta-Analysis of Studies across Aquatic Ecosystems." *Oecologia* 169: 803–9.

Soininen, J., and A. Teittinen. 2019. "Fifteen Important Questions in the Spatial Ecology of Diatoms." *Freshwater Biology* 64: 2071–83.

Tickner, D., J. J. Opperman, R. Abell, M. Acreman, A. H. Arthington, S. E. Bunn, S. J. Cooke, et al. 2020. "Bending the Curve of Global Freshwater Biodiversity Loss: An Emergency Recovery Plan." *BioScience* 70: 330–42.

Tjørve, E., W. E. Kunin, C. Polce, and K. M. C. Tjørve. 2008. "Species-Area Relationship: Separating the Effects of Species Abundance and Spatial Distribution." *Journal of Ecology* 96: 1141–51.

Tjørve, E., K. M. C. Tjørve, E. Šizlingová, and A. Šizling. 2021. "Determinants of the Shape of Species-Area Curves." In *The*

Species-Area Relationship: Theory and Application, edited by T. J. Matthews, K. A. Triantis, and R. J. Whittaker, 78–106. Cambridge: Cambridge University Press.

Triantis, K. A., F. Guilhaumon, and R. J. Whittaker. 2012. “The Island Species-Area Relationship: Biology and Statistics.” *Journal of Biogeography* 39: 215–31.

Turner, W. R., and E. Tjørve. 2005. “Scale-Dependence in Species-Area Relationships.” *Ecography* 28: 721–30.

Ugland, K. I., J. S. Gray, and K. E. Ellingsen. 2003. “The Species-Accumulation Curve and Estimation of Species Richness.” *Journal of Animal Ecology* 72: 888–97.

Veech, J. A. 2000. “Choice of Species-Area Function Affects Identification of Hotspots.” *Conservation Biology* 14: 140–7.

White, G. C., and R. E. Bennetts. 1996. “Analysis of Frequency Count Data Using the Negative Binomial Distribution.” *Ecology* 77: 2549–57.

Wiens, J. J. 2016. “Climate-Related Local Extinctions Are Already Widespread among Plant and Animal Species.” *PLoS Biology* 14: e2001104.

Xu, W., G. Chen, C. Liu, and K. Ma. 2015. “Latitudinal Differences in Species Abundance Distributions, Rather than Spatial Aggregation, Explain Beta-Diversity along Latitudinal Gradients.” *Global Ecology and Biogeography* 24: 1170–80.

SUPPORTING INFORMATION

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

How to cite this article: Passy, Sophia I., Joseph L. Mruzek, William R. Budnick, Thibault Leboucher, Aurélien Jamoneau, Jonathan M. Chase, Janne Soininen, et al. 2023. “On the Shape and Origins of the Freshwater Species–Area Relationship.” *Ecology* 104(3): e3917. https://doi.org/10.1002/ecy.3917