

RESEARCH
PAPER



Empirical evidence for the scale dependence of biotic interactions

Jonathan Belmaker^{1,*†}, Phoebe Zarnetske^{2,3†}, Mao-Ning Tuanmu⁴, Sara Zonneveld^{3,5}, Sydne Record^{2,6}, Angela Strecker⁷ and Lydia Beaudrot⁸

¹Department of Zoology, George S. Wise Faculty of Life Sciences, Tel Aviv University, Tel Aviv, 69978, Israel, ²Department of Forestry and Fisheries & Wildlife, Michigan State University, 480 Wilson Road, East Lansing, MI 48824, USA, ³School of Forestry and Environmental Studies, Yale University, 195 Prospect Street, New Haven, CT 06511, USA, ⁴Department of Ecology and Evolutionary Biology, Yale University, 165 Prospect Street, New Haven, CT 06520, USA, ⁵Biosciences, College of Life & Environmental Sciences, University of Exeter, Geoffrey Pope Building, Exeter EX4 4QD, UK, ⁶Bryn Mawr, Bryn Mawr College, 101 North Merion Avenue, PA 19010, USA, ⁷Department of Environmental Science and Management, Portland State University, Portland, OR 97201, USA, ⁸Graduate Group in Ecology and Department of Anthropology, University of California, One Shields Avenue, Davis, CA 95616, USA

*Correspondence: Jonathan Belmaker, Department of Zoology, George S. Wise Faculty of Life Sciences, Tel Aviv University, Tel Aviv, 69978, Israel.
E-mail: jbelmaker@post.tau.ac.il

†These authors have contributed equally.

ABSTRACT

Aim Although it is recognized that ecological patterns are scale dependent, the exact scales over which specific ecological processes operate are still a matter of controversy. In particular, understanding the scales over which biotic interactions operate is critical for predicting changes in species distributions in the face of the ongoing biodiversity crisis. It has been hypothesized that biotic interactions operate predominately at fine grains, yet this conjecture has received relatively little empirical scrutiny. We use US woodpeckers as a model system to assess the relative importance of biotic interactions, environmental suitability and geographic proximity to other intraspecific occurrence sites, across scales.

Location Conterminous United States.

Methods We combined species occurrence data from the North American Breeding Bird Survey (BBS) with a large pair-wise interaction matrix describing known interactions among woodpeckers and other bird species. Using a logistic mixed modelling framework we then established the relative importance of biotic interactions as predictors of woodpecker occurrences in relation to environment and geographic proximity to intraspecific occurrence sites.

Results We found that geographic proximity becomes a stronger predictor of woodpecker occurrence as grain becomes coarser, while environment is grain-invariant. As opposed to environment and geographic proximity, we found that when the focal species experienced positive biotic interactions, the importance of interactions decreased with increased grain. However, positive interactions remained important up to a grain size of entire BBS routes (*c.* 40 km), which is much coarser than the grain size used by most species distribution models. In contrast, when the focal species experienced negative interactions we did not find clear grain dependence.

Main conclusions Biotic interactions (both positive and negative) are important predictors of species occurrences. While these interactions are strongest at fine grains, they can remain important even at coarse grains, and are thus critical for predicting distributional changes in the face of the ongoing biodiversity crisis.

Keywords

Biotic interactions, Breeding Bird Survey, co-occurrence, extent, scale, grain, species distribution models.

INTRODUCTION

It is widely acknowledged that ecological processes are scale dependent (Levin, 1992; Schneider, 2001; Maurer & Taper, 2002;

McGill, 2010; Belmaker & Jetz, 2011). However, the exact scales over which specific ecological processes operate are still a matter of controversy. The current paradigm states that biotic interactions, such as competition, predation and facilitation, are

important determinants of species distributions only at small spatial scales (i.e. fine grains) where direct interspecific interactions occur (Shmida & Wilson, 1985; Whittaker *et al.*, 2001; Cavender-Bares *et al.*, 2009; McGill, 2010). At coarser grains, other processes such as dispersal and environmental associations are believed to gain prominence. Empirical evidence for the importance of biotic interactions in shaping species distributions has largely come from studies at local scales due to the difficulty of performing landscape-scale experiments and measuring the effects of biotic interactions at coarse scales. Thus, it has been a challenge to obtain data on the scale dependence of biotic interactions relative to other processes that affect species distributions.

Establishing the scale dependence of biotic interactions goes beyond the intrinsic interest of understanding the processes that structure natural assemblages. The grain size at which the influence of an interaction can no longer be detected is critical for predicting species-specific distributional changes in the face of the ongoing biodiversity crisis. If biotic interactions are only important at very fine grains we may only need species-specific attributes, such as habitat suitability or dispersal ability, to predict broad-scale changes in distribution. If, however, biotic interactions remain important across grain sizes, changes in one species may initiate cascading effects on many other species (Van der Putten *et al.*, 2010; Kissling *et al.*, 2012; Zarnetske *et al.*, 2012; Tuanmu *et al.*, 2013; Wisz *et al.*, 2013). This knowledge gap has a profound impact on the field of species distribution modelling (Guisan & Thuiller, 2005), which is now faced with uncertainty regarding how significantly biotic interactions could alter predictions and model accuracy (Elith & Leathwick, 2009; Godsoe & Harmon, 2012; Record *et al.*, 2013).

Evidence for the importance of biotic interactions at coarse grains has been accumulating (Gotelli *et al.*, 2010; Afkhami *et al.*, 2014; Araujo & Rozenfeld, 2014; de Araujo *et al.*, 2014). A study on Danish birds found signals of competition at grid resolutions of 5 and 10 km (Gotelli *et al.*, 2010). Predictions of butterfly occurrences at a resolution of 50 km were improved by incorporating the distribution of host plants (Araujo & Luoto, 2007), and predictions of owl distributions were improved by incorporating positive interactions with woodpeckers within 10 km and, to a lesser extent, 40 km (Heikkinen *et al.*, 2007). At even coarser grains the dispersal abilities of frugivores may affect the distributions of the tree species they disperse (Johnson & Webb, 1989; Beaudrot *et al.*, 2013) and over evolutionary time biotic interactions may even constrain range overlap between species (Pigot & Tobias, 2013). A recent simulation study (Araujo & Rozenfeld, 2014) showed that the scale at which biotic interactions remain important is dependent on the type of interaction under study. While positive interactions such as mutualism may remain important across grains, negative interactions such as competition were only discerned at fine grains (Araujo & Rozenfeld, 2014). Nevertheless, to date empirical analyses have been restricted to one or a very few spatial scales, thus preventing a complete evaluation of scale–response curves.

In this study, we combined species occurrence data from the North American Breeding Bird Survey (BBS) with indepen-

dently collected data on species interactions (e.g. competition, commensalism, predation). We focused on woodpeckers, a guild of birds in which most species excavate nesting cavities in dead or dying trees, because these species are known to interact in both positive and negative ways with each other and other bird species. We created a large pair-wise interaction matrix describing known interactions among woodpeckers in the United States and their interacting bird species. We then quantified the importance of positive and negative biotic interactions as predictors of species occurrences. Biotic interactions were contrasted with predictors that are typically thought to affect species distributions at broader spatial scales, including environment and geographic proximity to other intraspecific occurrences. This framework allowed us to assess how the relative importance of biotic interactions, environmental effects and geographic proximity changes across scales.

METHODS

Bird occurrence data

We used the North American Breeding Bird Survey (BBS) (USGS, 2012) to obtain breeding bird counts and locations across forested regions of the conterminous United States for the year 2006, to ensure temporal alignment with the environmental data. We focused on woodpeckers, because there is relatively good knowledge about their biotic interactions. The BBS consists of point counts during the breeding season (May and June) at 50 stops, spaced approximately every 800 m along routes approximately 39 km long. At each of the 50 stops per route, observers record the number of individuals of all species seen or heard within a 400-m radius during a 3-minute period. Some 4000 routes are repeated each year by trained volunteers, generating a unique systematic sampling of breeding birds across a broad spatial extent.

The vast majority of analyses on the BBS data are performed at the route level. However, in order to assess the influence of different predictors (e.g. biotic interactions, environment) on the occurrence of species, we included the finest stop-level resolution. Stop-level geographic locations are not published for all BBS routes, yet route locations are known, and point locations exist for the first stop of each route. For routes between 29 and 49 km long we used ARCGIS and PYTHON to generate likely locations for the remaining 49 stops, spaced equidistant along each route, beginning at the known start point. Routes that appeared interrupted or contained sections that crossed were removed manually upon visual inspection.

We restricted the BBS location data to only those stops with land-cover change from 1992 to 2006 of less than 10% to minimize the effects of land-cover change on our results. To determine this selection, we first generated buffers centred on each stop location that measured 400 m on either side of the BBS route, and 800 m along the BBS route. The 400 m buffer represents the maximum distance at which BBS observers record birds (Meehan *et al.*, 2010), while the larger 800 m distance along the route represents the potential for misplacing the loca-

tion of each stop point. Within each 400 m by 800 m buffer, we summed the 30-m National Land Cover Database (NLCD; Fry *et al.*, 2011) pixels for which land cover had changed from 1992 to 2006 within each buffered stop, and only kept those stops with buffers that contained pixels with less than 10% land-cover change over this timeframe. After filtering for land-cover change, the total number of BBS routes used in the analyses was 3126.

Altering grain and extent

We varied grain and extent as follows.

1. We changed grain by aggregating stop data within each BBS route. Grain varied by aggregating different numbers of stops, with one stop being no aggregation, then 2, 5, 10, 25 and 50 stops (all stops within each route are used). We refer to a site as the aggregated stops within a route. A species was deemed present in a site if it was seen at least once within the stops considered and absent if it did not occur at all within the stops considered. Environmental data (see below) were averaged across stops within a site. Two variations of changing grain size were examined. In the first, multiple data points were used within each route (e.g. for a grain size of 25, we used two sites within a route). This resulted in sites within routes being very close and

typically having similar environmental conditions, which may be regarded as a form of pseudoreplication. It additionally meant that sample size was smaller when the data were more aggregated (e.g. for a grain size of one we had 50 stops within a route while for a grain size of 50 we had only one stop). In the second version, we always used the first site within a route, regardless of the aggregation level, thus maintaining sample size across grains but necessarily excluding a large proportion of the data.

2. We varied extent by altering the spatial area considered, with diameters ranging from 500 km to 1000, 2000 up to 4000 km. This was done by choosing a random route and only including routes that fell within the above distances to it. Sample size always increased with increased extent.

Estimates of biotic interactions

We completed an extensive literature review to capture the interactions that other bird species may have with woodpeckers (Table 1). For example, primary cavity-nesting woodpeckers can facilitate secondary cavity nesters (Martin & Eadie, 1999; Aitken & Martin, 2007). In addition, woodpeckers interact with other birds via competition, predator–prey relationships, kleptoparasitism and brood parasitism. To identify these inter-

Table 1 The total number of bird–bird interactions experienced by woodpecker species in the conterminous United States when the woodpecker is the focal species. Interaction counts in the table are split by biotic interaction types, where the first interaction represents the effect on the focal species. For example, Positive–negative interactions indicate interactions where the focal woodpecker species has a negative effect on the interacting species, and the interacting species has a positive effect on the focal woodpecker. Negative–null and Positive–positive interactions were not found and are therefore not included. Null–positive interactions exist but were not used in this study.

Common name	Latin name	Positive–negative	Negative–positive	Positive–null	Negative–negative	Total
Acorn woodpecker	<i>Melanerpes formicivorus</i>	2	5	0	8	15
American three-toed woodpecker	<i>Picoides dorsalis</i>	0	0	1	3	4
Arizona woodpecker	<i>Picoides arizonae</i>	0	0	0	7	7
Black-backed woodpecker	<i>Picoides arcticus</i>	0	1	0	6	7
Downy woodpecker	<i>Picoides pubescens</i>	1	5	1	22	29
Gila woodpecker	<i>Melanerpes uropygialis</i>	4	1	0	4	9
Gilded flicker	<i>Colaptes chrysoides</i>	0	10	0	6	16
Golden-fronted woodpecker	<i>Melanerpes aurifrons</i>	3	1	0	3	7
Hairy woodpecker	<i>Picoides villosus</i>	0	8	1	18	27
Ladder-backed woodpecker	<i>Picoides scalaris</i>	0	0	0	2	2
Lewis's woodpecker	<i>Melanerpes lewis</i>	1	5	6	11	23
Northern flicker	<i>Colaptes auratus</i>	25	68	0	30	123
Nuttall's woodpecker	<i>Picoides nuttallii</i>	0	1	1	12	14
Pileated woodpecker	<i>Dryocopus pileatus</i>	1	6	0	13	20
Red-bellied woodpecker	<i>Melanerpes carolinus</i>	10	3	0	14	27
Red-breasted sapsucker	<i>Sphyrapicus ruber</i>	0	4	0	1	5
Red-cockaded woodpecker	<i>Picoides borealis</i>	0	4	0	11	15
Red-headed woodpecker	<i>Melanerpes erythrocephalus</i>	9	3	1	16	29
Red-naped sapsucker	<i>Sphyrapicus nuchalis</i>	0	4	0	2	6
White-headed woodpecker	<i>Picoides albolarvatus</i>	0	2	0	8	10
Williamson's sapsucker	<i>Sphyrapicus thyroideus</i>	2	2	1	9	14
Yellow-bellied sapsucker	<i>Sphyrapicus varius</i>	0	2	1	4	7
Total interactions		58	135	13	210	416

actions, we searched Birds of North America Online (Poole, 2005) for each woodpecker species in the dataset. Birds of North America is a comprehensive online resource providing life-history information, including specific sections on behaviour and interspecific interactions. Account contents are updated regularly and reference peer-reviewed journal articles. For each woodpecker, we recorded all reported interactions with any other bird species as identified by the articles containing a mention of the woodpecker species. We did not include interactions that occurred at bird feeders because these can represent unnatural situations, and we also did not include interactions related to interspecific flocking behaviour.

It is likely that not all interactions were reported for each species. This can lead to several types of biases. First, it is possible that more interactions are reported for the most common species, leading to rare species being underrepresented in the interaction matrix. However, we did not find such a pattern as the correlation between species abundance and the total number of interactions observed for a species was low ($r = 0.41$). Second, it is highly probable that rare interactions between species were not observed or reported. However, this should not influence our results as rare interactions should not strongly influence species occupancy. Finally, there might be geographic biases in reporting, based on more in-depth studies in certain locations. Although this type of bias is hard to correct, in our case we repeated our analyses across different extents and locations across the conterminous United States. Thus, geographic biases will increase the variability in our results but should not alter the major patterns. Despite these potential biases, Birds of North America is still the most comprehensive resource for North American breeding birds (Poole, 2005) and thus provides the best collection of interaction information for North American woodpeckers as a whole.

We separated the positive and negative interactions exerted on each species. For each interaction type we summed the number of positive and negative interactions experienced by each focal woodpecker species from all co-occurring species at a site to produce an estimate of total positive and total negative interactions.

Simulation studies suggest that the scale dependence of biotic interactions may change according to type of interaction (Araujo & Rozenfeld, 2014). We thus also split the effect of the interactions on the focal species into the following categories.

1. Positive effect on the focal species and negative effect on the interacting species. This category includes focal species that interact by being a brood parasite (brood parasitism), steal food from the interacting species (kleptoparasitism) or are predators (predator–prey).
2. Negative effect on the focal species and positive effect on the interacting species. This category includes focal species that are the victim of food theft (kleptoparasitism), are hosts to brood parasites (brood parasitism) or are prey species (predator–prey).
3. Positive effect on the focal species and no effect on the interacting species. This category includes focal species that benefit from commensalism (e.g. secondary cavity nesters).

4. Negative effect on the focal species and negative effect on the interacting species. This includes species that have competitive interactions.

Other interactions, such as amensalism and mutualism, were too rare to be analysed or non-existent in our literature review. A list of the total number of interactions documented for each species can be found in Table 1.

Environmental suitability

We chose a suite of environmental variables to use as predictors that are relevant to the distributions of breeding birds in North America (Meehan *et al.*, 2010; Rittenhouse *et al.*, 2010; Matthews *et al.*, 2011; Barbet-Massin & Jetz, 2014). These included climatic variables generated from PRISM (PRISM Climate Group, 2010) monthly datasets (4-km resolution), intersected at each stop point. With these data, we used the ‘biovars’ command in the R package ‘dismo’ to generate 19 bioclimatic predictors. For each buffered stop area, we also calculated the proportion of land-cover types in 2006 from the NLCD land-cover dataset (Fry *et al.*, 2011). Based on the 30-m National Biomass and Carbon Dataset for the year 2000 (NBCD2000; Kellndorfer *et al.*, 2012) we calculated the average basal area-weighted canopy height for each buffered stop area. To reflect the productivity of vegetation, we obtained the 95th percentile of the enhanced vegetation index (EVI) from a 2005–2007 time series of a 250-m 16-day MODIS composite (MOD13Q1) for each 250-m pixel. We then calculated the area-weighted average for each buffer area. We used the 95th percentile over the 3 years to capture the EVI at the peak of a growing season, but to avoid the inclusion of spuriously high values. The use of the 3-year time series increases the number of valid EVI values for the calculation and minimizes the effect of potential land-cover change. A full list of variables is presented in Table A1 in the Appendix.

For each species we estimated the probability of a site being within the suitable environment using bioclimatic envelope models, implemented through the function ‘bioclim’ in the R package ‘dismo’ (Hijmans *et al.*, 2013). The models were based on the first 10 axes of a principal components analysis (PCA) among all of the above-mentioned environmental variables, which jointly explained more than 99.9% of the total variance in environmental data. Within each grain or extent (see above) we randomly selected 80% of the BBS routes to serve as training data while the remaining 20% of the routes served as test data. The training data were used to generate predictions of environmental suitability for each of the test data points for each species. The predictions for the test data points only were used in the final analyses (logistic mixed-effects models, see below).

Geographic proximity

We performed spatial interpolation to estimate the contribution of geographic proximity to other occupied sites to the probability of species occurrence, independent of environmental suitability (Bahn & McGill, 2007). We used Gaussian kernel density

smoothers to calculate the density of occurrence for each site based on the geographic location of sites. We controlled for variation in sampling effort by dividing observed occurrence density by the density of sampling stops (Broennimann *et al.*, 2012). This ensured that sparsely sampled regions did not receive exceptionally low densities. For each species, we applied a kernel density smoother to the training data and generated predictions for the test data. This resulted in an occurrence probability for each species and testing site based on geographic proximity to other known occurrences of that species.

Analyses

We performed a logistic mixed-effects model analysis for all woodpeckers across all testing sites, using the species- and site-specific estimates obtained from the environmental suitability and geographic proximity analyses (Fig. 1). We ran a single model across all woodpecker species observed at a given grain and extent. Thus, the presence or absence of species at each site was the binary response variable with four fixed effects: (1) environmental suitability – environmental suitability according to the bioclimatic envelope model; (2) geographic proximity – occurrence probability according to the Gaussian kernel density smoother; (3) positive biotic interactions – summed interactions of all co-occurring species that exert a positive effect on the focal species; and (4) negative biotic interactions – similar to positive biotic interactions, but including only species that exert

a negative effect on the focal species. Heterogeneity among species and sites was accounted for by incorporating species- and site-specific random-effect intercepts. We compared models with random intercepts only with those with random intercepts and slopes. We found that for 80 and 58% of the models (for positive and negative interactions, respectively) random-intercept-only models were superior (in terms of the Akaike information criterion, AIC) to random-slope models. The percentage of models for which random-slope models were not statistically superior to random-intercept models (AIC differences < 2) was even higher (89 and 69% of the models, respectively). Consequently, we retained only the simpler random-intercept models throughout the study. A species was included in the models only when it had more than three presences within the sites investigated so as to remove rare species for which the random effect could not be estimated. Environmental suitability and geographic proximity were logit transformed and standardized (to a mean of zero and a standard deviation of one) to allow for comparison of fixed-effect estimates. Over local extents some analyses contained few routes, and these were excluded when the number of routes fell below 50.

We extracted parameter estimates and unconditional standard errors from the models. For each fixed effect, importance was calculated as the Akaike weight of the full model relative to the model excluding that predictor. We assessed the goodness of fit by calculating marginal and conditional R^2 (R^2_M and R^2_C , respectively), which are measures of the variance explained by

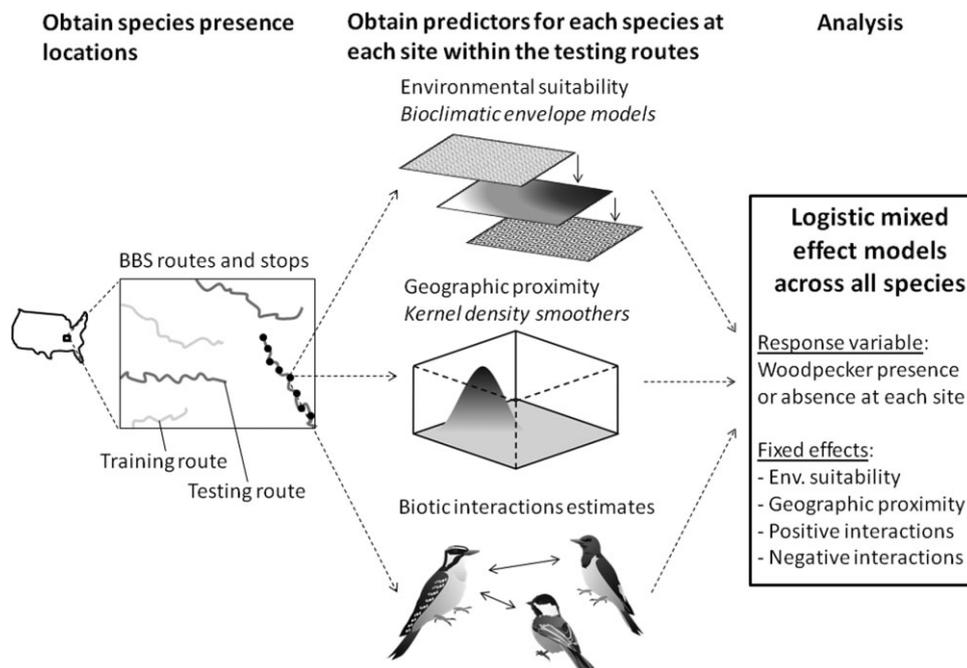


Figure 1 Schematic diagram presenting the major steps in our analyses. For each site (aggregated stops) within the Breeding Bird Survey (BBS) routes, we calculated species-specific probability of occurrence estimated from bioclimatic envelope models (environmental suitability) and geographic proximity (kernel density smoothers). We additionally calculated for each species at each site the summed number of positive and negative interactions. At the second step, these variables were combined within a mixed-effect model across woodpecker species. Bird line drawings by J. Woerner, K. Kraer, L. Van Essen and T. Saxby, IAN image library (<http://ian.umces.edu/imagelibrary/>).

the fixed effects and fixed and random effects, respectively (Nakagawa & Schielzeth, 2013). All analyses were conducted in R Version 2.15.0 (R development core team, 2012).

To ensure that the results found for biotic interactions were not an artifact related to species richness (e.g. species experience more positive interactions simply because of higher species richness) or related to reporting bias of the biotic interactions, we compared the observed patterns with a null model and calculated effect sizes (ES) as standardized deviation from the null. The null expectations were generated from 100 runs of the model. To generate the null models, we randomly shuffled interactions such that the interactions experienced by each focal species were assigned to random interacting species. For example, for a woodpecker with five interacting bird species at a site, we sampled interactions from the complete species list and randomly assigned them to one of the five interacting species. Thus, this null model did not alter the identity of co-occurring species or the total number of interactions experienced by the focal species.

BBS route identity varied across model runs, even for the same grain and extent, both because different regions were selected and because different sites were randomly designated for training and testing. We thus performed 100 runs in each

grain (aggregation of 1, 2, 5, 10, 25 and 50 stops) and extent (diameters of 500, 1000, 2000 and 4000 km) combination resulting in a total of 2000 mixed-effect models.

RESULTS

Environmental suitability and geographic proximity had strong effects on woodpecker occurrence. This was seen both in the positive coefficient estimates and the high variable importance (Fig. 2). This meant that there was a higher chance of seeing a specific woodpecker species at a site if: (1) the environment was more suitable, and (2) the site was closer to other sites in which the same species was seen. Geographic proximity showed strong grain dependence where increasing grain size resulted in increased coefficient estimates. Environment did not show similar patterns and its effect was relatively constant across grains. Generally similar results were found when only the first site within each route was included to maintain sample size across grains (Fig. S1 in Supporting Information). However, these results were more variable, reflecting the high uncertainty associated with each particular observation.

Increasing spatial extent caused a substantial increase in the variance explained by the models (Fig. S2). This result reflects

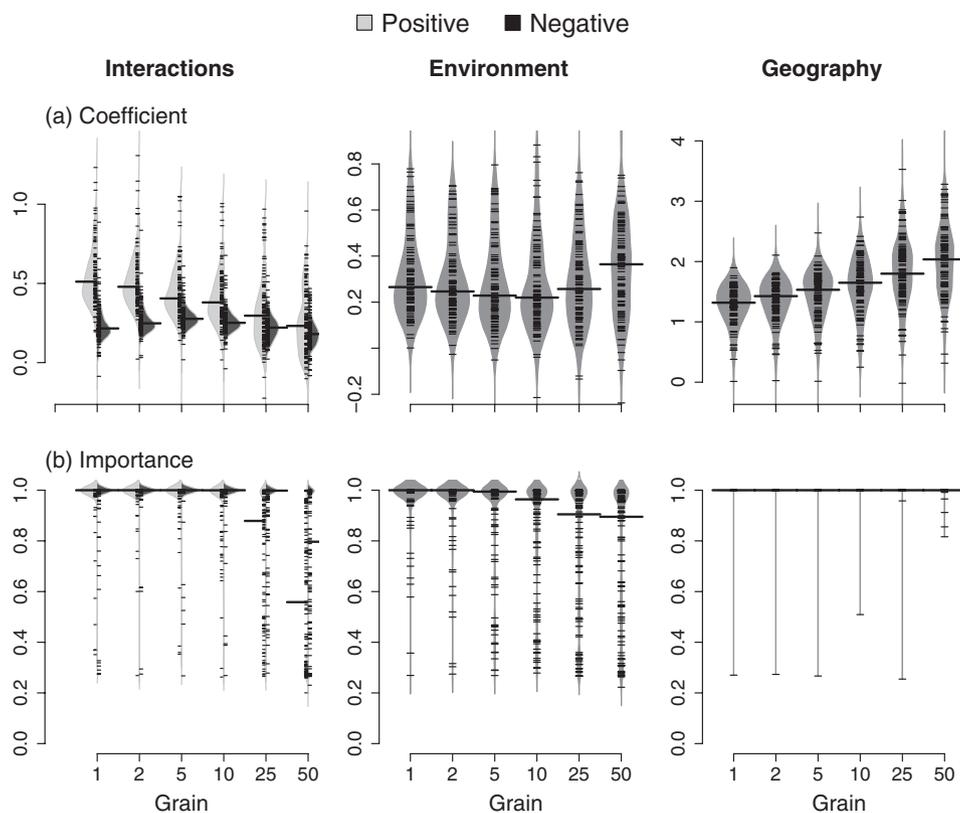


Figure 2 Violin plots depicting the scaling of the three ecological predictors (biotic interactions, environment and geographic proximity) examined across different grain sizes, measured as the level of aggregation of stops within Breeding Bird Survey routes: (a) variable coefficient; (b) importance of the variables within a full model relative to a model excluding that variable. Biotic interactions (left panels) are separated into positive (grey) and negative (black) interactions experienced by the focal species. The tick marks represent the results from each of the 100 mixed-effect models, each containing several species across all testing sites. Solid lines represent median values at each grain. The shaded area represent the result of a Gaussian kernel density smoother. Results are based on analyses from an extent of 2000 km.

the larger number of BBS routes sampled at broader extents and the associated larger environmental gradients. However, for most predictors the effect of extent on coefficient estimates was minor. Thus, when the analyses were conducted at local extents (such as extents of 500 and 1000 km) predictors maintained overall positive coefficients but with larger variability (Table A2, Fig. S3). The only predictor to show clear dependence on extent was geographic proximity, for which the estimated coefficient increased dramatically as extent broadened (Fig. S2).

Both positive and negative interactions displayed positive coefficients across scales, meaning that occurrence probability of a focal species is higher when other species which are known to interact with it were recorded at the site (Fig. 2). Positive coefficients for negatively interacting species were surprising, but can happen when the species being considered has similar habitat requirements, not captured by the environmental model, leading to net positive co-occurrence. However, at all grains and extents the coefficients for positive interactions were higher than those for negative interactions, suggesting the presence of a signal of biotic interactions. As opposed to environment and geography, positive biotic interactions showed a clear decrease in coefficient estimates with increased grain (Fig. 2). In contrast, negative interactions do not show any clear grain dependence.

To control for potential biases related to the distribution of the number of total biotic interactions across richness gradients, we calculated standardized deviations from null models as ES. Positive and negative interactions had the expected positive and negative ES (respectively), attesting to the importance of taking into account the null distribution. We found that positive interaction ES had similar patterns to the raw coefficients and were generally positive, but decreased in strength with increasing grain (Fig. 3). Nevertheless, even at the coarsest grains, ES for biotic interactions were significant for a substantially higher proportion of the models than predicted by chance alone (e.g. at a grain size of 50 and extent of 2000 km, 14% of the models are significant versus the expected 2.5%). For negative interactions we found negative ES, which declined in strength with increasing grain size (Fig. 3). However, here the ES remained non-significant across grains.

When parcelling out the biotic interactions to different interaction types we generally found few ES that were significant, probably due to the low power of the analyses. Nevertheless, we found evidence for grain dependence when the focal species experienced a positive effect and the interacting species experienced a negative effect (Fig. 4). There was no clear grain dependence with other types of interactions.

DISCUSSION

Species are altering their distributions in response to environmental changes and anthropogenic impacts. The prevailing paradigm states that at coarse grains understanding and predicting these distributional changes does not require us to account for biotic interactions among species. However, some evidence suggests that biotic interactions may be important even at coarse

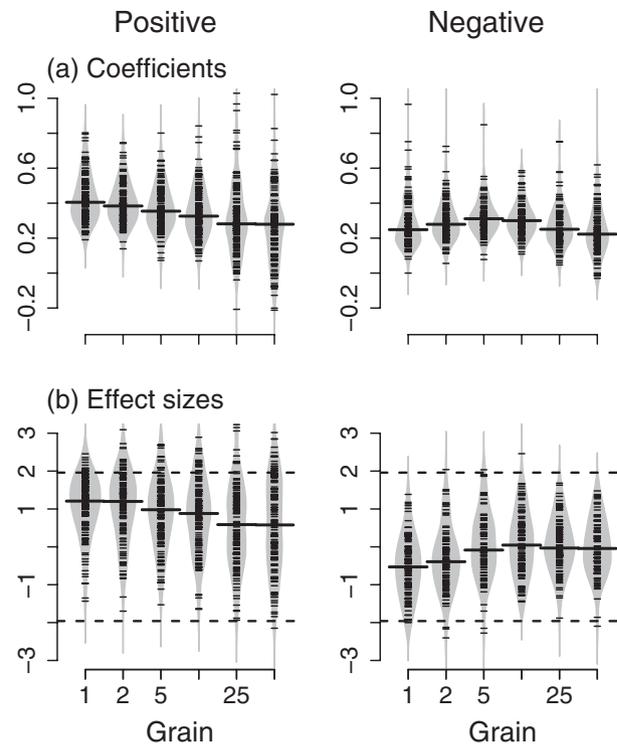


Figure 3 Violin plots depicting (a) scaling of biotic interaction coefficients, and (b) effect sizes (ES; standardized deviation from null). Dashed lines in (b) represent ES of ± 1.96 , beyond which they are statistically significant. The tick marks represent the results from each of the 100 mixed-effect models. Solid lines represent median values at each grain. The shaded area represent the result of a Gaussian kernel density smoother. Results are based on 2000 km extent.

grains (Gotelli *et al.*, 2010; Araujo & Rozenfeld, 2014; Kissling & Schleuning, 2014). Nevertheless, we still have a poor understanding of: (1) how important biotic interactions are relative to other processes, and (2) at what grain biotic interactions are too important to be ignored, for example in the context of predicting future distributional changes. In this study, we used detailed interaction information for North American woodpeckers and their interacting avian species to show that: (1) biotic interactions (both positive and negative) are important predictors of species occurrences, although typically less important than environmental variables or geographic proximity, and (2) biotic interactions remain important even at coarse grains.

We found that all three types of predictors (i.e. biotic interactions, environment and geographic proximity) were important for understanding woodpecker occurrences across scales. However, the relative importance of these predictors changed with the grain and, to a lesser degree, the extent considered. For positive interactions, although increasing grain size resulted in reduced coefficient estimates and importance of variables, interactions remained substantial even at coarse grains of entire BBS routes (c. 40 km). The significance of positive interactions was evident from the positive coefficients, high importance and generally positive effect sizes. These results show, for the first time,

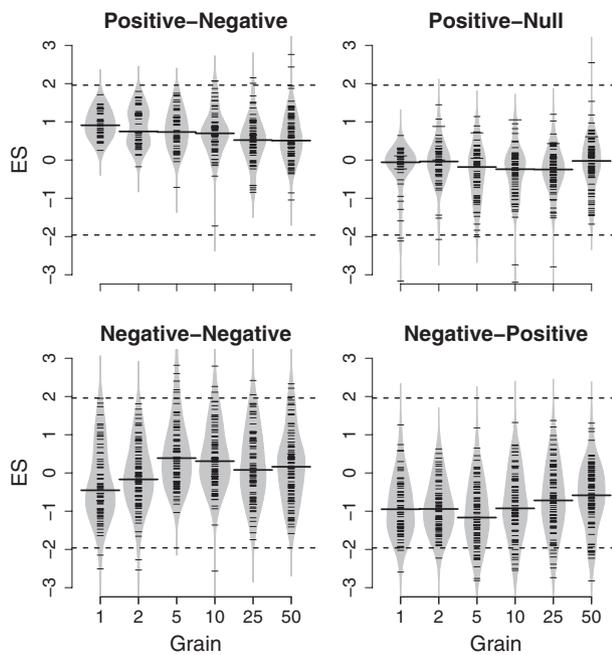


Figure 4 Violin plots depicting changes in effect sizes (ES; standardized deviation from null) across grains for different types of biotic interactions. Dashed lines represent ES of ± 1.96 , beyond which ES are statistically significant. Solid lines represent median values at each grain. The shaded area represent the result of a Gaussian kernel density smoother. Results are based on analyses from an extent of 2000 km.

that biotic interactions influence species occurrence across different scales. We suggest that fine grain analyses can always benefit from the explicit incorporation of biotic interactions, but that their importance diminishes, at least for North American woodpeckers, at grain sizes coarser than *c.* 40 km (the length of a BBS route). As roughly half of species distribution models use grains of up to 100 km² (S.R. *et al.*, unpublished data), incorporating biotic interactions across scales will be key to improving predictive performance for many different species.

We found little grain dependence in the magnitude of environmental coefficients. This is in contrast to several studies that found strong grain dependence in the strength of environment as a predictor of species distribution and richness (Guisan *et al.*, 2007; Randin *et al.*, 2009; Belmaker & Jetz, 2011). This grain independence of environmental predictors reflects our use of a large suite of environmental predictors that include climate and detailed land-cover information. Whereas climate may be important at coarse grains (Whittaker *et al.*, 2001; Belmaker & Jetz, 2011) land cover may be more influential at fine grains, resulting in overall grain invariance.

In contrast to environment, geographic proximity was markedly scale dependent. The coefficients for geographic proximity became larger as grain became coarser and as extent broadened. Thus, at grain sizes approaching entire BBS routes and extent approaching the entire continent, the presence of a species was strongly dependent on the presence of conspecifics in adjacent routes (Bahn & McGill, 2007). This pattern is likely to reflect the

cohesiveness of species ranges at coarse grains and broad extents. With finer grain, species presence becomes more patchy, reflecting the underlying heterogeneity in the environmental and biotic landscape. As a result, the statistical contribution of geographic proximity decreases.

Our approach to quantifying the importance of biotic interactions relies on observational data. Thus, when multiple species respond positively to an environmental resource we may find positive co-occurrence even if the underlying interactions between the species are negative. Indeed, the coefficients for negative interactions were positive, although we would expect to observe negative coefficients as the presence of interacting species should reduce probability of occurrence. However, the coefficients for positive interactions were always higher than the coefficients for negative interactions. If the effects of biotic interactions were entirely due to shared environmental preference we would expect to find no difference between the coefficients for positive and negative interactions. Thus, the difference between positive and negative coefficients represents the net influence of biotic interactions after accounting for shared environmental preferences of co-occurring species.

Biotic interactions exhibited mixed grain dependence. For positive interactions, both the coefficients and ES were high at fine grains and decreased with increased grain, but remained positive overall (Fig. 3). However, ES for negative interactions were slightly negative at fine grains only, and were not significant across all grains (Fig. 3). These patterns reinforce the simulation results of Araujo & Rozenfeld (2014), who found positive interactions, such as mutualism, to be important across grains but competition to be important only at fine grains. When separating the interactions into categories, we found that grain dependence was particularly apparent when the focal species benefited and the interacting species suffered (Fig. 4). In woodpeckers, such interactions mainly include kleptoparasitism, in which a woodpecker species gains or steals food originally obtained by another species and rare cases of woodpecker predation on other species. In contrast, it appears that competition (negative-negative) was generally weak and did not exhibit clear grain dependence. Together these findings suggest that emphasis should shift from focusing on competition as the prime biotic determinant of species geographical distributions to a full range of interaction types (e.g. predator-prey relationships, facilitation).

We would like to emphasize that the only way to truly estimate interaction strength is to perform manipulative experiments. Observational studies can never completely disentangle shared or divergent environmental preferences versus true biotic interactions. For example, the apparent importance of interactions at macroecological scales will be contingent on environmental data resolution. Coarse habitat designations will tend to equate negative co-occurrence patterns with competition, while at fine enough resolutions we may detect habitat partitioning as the underlying cause. Thus, the apparent importance of biotic interaction should be considered in light of the data used. For the environmental predictors commonly used in these analyses, it appears that incorporating species interactions improves

species occupancy predictions. However, a similar model using better-resolved environmental data may find a reduced signal of such biotic interactions.

Our results have important implications for applied research on species distribution models and community assembly theory. First, we show that information on biotic interactions is likely to improve species distribution models by demonstrating that qualitative data gleaned from the literature or expert knowledge provide valuable information for predicting species occurrences. Second, we find that contrary to the present ecological paradigm, biotic interactions remain important even at coarse grains. This finding shows that the inclusion or exclusion of biotic interactions should be based on ecological knowledge and cannot simply be dismissed using the coarse grain or extent of the models as a justification.

ACKNOWLEDGEMENTS

We would like to thank the numerous BBS volunteers for the bird distribution data, the Ecological Society of America for sponsoring the 2013 'Scaling-Up: Continental Scale Ecology Workshop for Early Career Scientists', its planning committee, Deborah Goldberg, Matthew Fitzpatrick, Janet Franklin, Robert Guralnick and Katherine Suding, as well as Jill Petraglia Parsons, Clifford Duke and all of the workshop participants. J.B. was supported by the Israel–German Foundation grant number I-2373-203.13/2014. P.L.Z. was supported by the Yale Climate and Energy Institute and by Michigan State University.

REFERENCES

- Afkhami, M.E., McIntyre, P.J. & Strauss, S.Y. (2014) Mutualist-mediated effects on species' range limits across large geographic scales. *Ecology Letters*, **17**, 1265–1273.
- Aitken, K.E.H. & Martin, K. (2007) The importance of excavators in hole-nesting communities: availability and use of natural tree holes in old mixed forests of western Canada. *Journal of Ornithology*, **148**, S425–S434.
- Araujo, M.B. & Luoto, M. (2007) The importance of biotic interactions for modelling species distributions under climate change. *Global Ecology and Biogeography*, **16**, 743–753.
- Araujo, M.B. & Rozenfeld, A. (2014) The geographic scaling of biotic interactions. *Ecography*, **37**, 406–415.
- de Araujo, C.B., Marcondes-Machado, L.O. & Costa, G.C. (2014) The importance of biotic interactions in species distribution models: a test of the Eltonian noise hypothesis using parrots. *Journal of Biogeography*, **41**, 513–523.
- Bahn, V. & McGill, B.J. (2007) Can niche-based distribution models outperform spatial interpolation? *Global Ecology and Biogeography*, **16**, 733–742.
- Barbet-Massin, M. & Jetz, W. (2014) A 40-year, continent-wide, multispecies assessment of relevant climate predictors for species distribution modelling. *Diversity and Distributions*, **20**, 1285–1295.
- Beaudrot, L., Rejmanek, M. & Marshall, A.J. (2013) Dispersal modes affect tropical forest assembly across trophic levels. *Ecography*, **36**, 984–993.
- Belmaker, J. & Jetz, W. (2011) Cross-scale variation in species richness–environment associations. *Global Ecology and Biogeography*, **20**, 464–474.
- Broennimann, O., Fitzpatrick, M.C., Pearman, P.B., Petitpierre, B., Pellissier, L., Yoccoz, N.G., Thuiller, W., Fortin, M.J., Randin, C., Zimmermann, N.E., Graham, C.H. & Guisan, A. (2012) Measuring ecological niche overlap from occurrence and spatial environmental data. *Global Ecology and Biogeography*, **21**, 481–497.
- Cavender-Bares, J., Kozak, K.H., Fine, P.V.A. & Kembel, S.W. (2009) The merging of community ecology and phylogenetic biology. *Ecology Letters*, **12**, 693–715.
- Elith, J. & Leathwick, J.R. (2009) Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology Evolution and Systematics*, **40**, 677–697.
- Fry, J., Xian, G.S., Jin, S., Dewitz, J., Homer, C., Yang, L., Barnes, C., Herold, N. & Wickham, J. (2011) Completion of the 2006 National Land Cover Database for the conterminous United States. *Photogrammetric Engineering and Remote Sensing*, **77**, 858–864.
- Godsoe, W. & Harmon, L.J. (2012) How do species interactions affect species distribution models? *Ecography*, **35**, 811–820.
- Gotelli, N.J., Graves, G.R. & Rahbek, C. (2010) Macroecological signals of species interactions in the Danish avifauna. *Proceedings of the National Academy of Sciences USA*, **107**, 5030–5035.
- Guisan, A. & Thuiller, W. (2005) Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, **8**, 993–1009.
- Guisan, A., Graham, C.H., Elith, J., Huettmann, F. & Distri, N.S. (2007) Sensitivity of predictive species distribution models to change in grain size. *Diversity and Distributions*, **13**, 332–340.
- Heikkinen, R.K., Luoto, M., Virkkala, R., Pearson, R.G. & Korber, J.H. (2007) Biotic interactions improve prediction of boreal bird distributions at macro-scales. *Global Ecology and Biogeography*, **16**, 754–763.
- Hijmans, R.J., Phillips, S., Leathwick, J. & Elith, J. (2013) *dismo: species distribution modeling*. R package version 0.8-5. Available at: <http://CRAN.R-project.org/package=dismo> (accessed 21 December 2012).
- Johnson, W.C. & Webb, T. (1989) The role of blue jays (*Cyanocitta cristata* L.) in the postglacial dispersal of fagaceous trees in eastern North-America. *Journal of Biogeography*, **16**, 561–571.
- Kellendorfer, J., Walker, W., LaPoint, E., Bishop, J., Cormier, T., Fiske, G., Hoppus, M., Kirsch, K. & Westfall, J. (2012) *NACP Aboveground Biomass and Carbon Baseline Data (NBCD 2000)*, USA, 2000. Available at: http://daac.ornl.gov/NACP/guides/NBCD_2000.html (accessed 17 September 2013).
- Kissling, W.D. & Schleuning, M. (2014) Multispecies interactions across trophic levels at macroscales: retrospective and future directions. *Ecography*, **38**, 346–357.
- Kissling, W.D., Dormann, C.F., Groeneveld, J., Hickler, T., Kuhn, I., McNerny, G.J., Montoya, J.M., Romermann, C., Schiffers, K., Schurr, F.M., Singer, A., Svenning, J.C., Zimmermann, N.E. & O'Hara, R.B. (2012) Towards novel approaches to model-

- ling biotic interactions in multispecies assemblages at large spatial extents. *Journal of Biogeography*, **39**, 2163–2178.
- Levin, S.A. (1992) The problem of pattern and scale in ecology. *Ecology*, **73**, 1943–1967.
- McGill, B.J. (2010) Matters of scale. *Science*, **328**, 575–576.
- Martin, K. & Eadie, J.M. (1999) Nest webs: a community-wide approach to the management and conservation of cavity-nesting forest birds. *Forest Ecology and Management*, **115**, 243–257.
- Matthews, S.N., Iverson, L.R., Prasad, A.M. & Peters, M.P. (2011) Changes in potential habitat of 147 North American breeding bird species in response to redistribution of trees and climate following predicted climate change. *Ecography*, **34**, 933–945.
- Maurer, B.A. & Taper, M.L. (2002) Connecting geographical distributions with population processes. *Ecology Letters*, **5**, 223–231.
- Meehan, T.D., Hurlbert, A.H. & Gratton, C. (2010) Bird communities in future bioenergy landscapes of the Upper Midwest. *Proceedings of the National Academy of Sciences USA*, **107**, 18533–18538.
- Nakagawa, S. & Schielzeth, H. (2013) A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, **4**, 133–142.
- Pigot, A.L. & Tobias, J.A. (2013) Species interactions constrain geographic range expansion over evolutionary time. *Ecology Letters*, **16**, 330–338.
- Poole, A. (2005) *The birds of North America online*. Cornell Laboratory of Ornithology, Ithaca, NY.
- PRISM Climate Group (2010) Oregon State University, <http://prism.oregonstate.edu>.
- R development core team (2012) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Randin, C.F., Engler, R., Normand, S., Zappa, M., Zimmermann, N.E., Pearman, P.B., Vittoz, P., Thuiller, W. & Guisan, A. (2009) Climate change and plant distribution: local models predict high-elevation persistence. *Global Change Biology*, **15**, 1557–1569.
- Record, S., Fitzpatrick, M.C., Finley, A.O., Veloz, S. & Ellison, A.M. (2013) Should species distribution models account for spatial autocorrelation? A test of model projections across eight millennia of climate change. *Global Ecology and Biogeography*, **22**, 760–771.
- Rittenhouse, C.D., Pidgeon, A.M., Albright, T.P., Culbert, P.D., Clayton, M.K., Flather, C.H., Huang, C.Q., Masek, J.G., Stewart, S.I. & Radeloff, V.C. (2010) Conservation of forest birds: evidence of a shifting baseline in community structure. *PLoS ONE*, **5**, e11938.
- Schneider, D.C. (2001) The rise of the concept of scale in ecology. *Bioscience*, **51**, 545–553.
- Shmida, A. & Wilson, M.V. (1985) Biological determinants of species-diversity. *Journal of Biogeography*, **12**, 1–20.
- Tuanmu, M.-N., Vina, A., Winkler, J.A., Li, Y., Xu, W., Ouyang, Z. & Liu, J. (2013) Climate-change impacts on understory bamboo species and giant pandas in China's Qinling Mountains. *Nature Climate Change*, **3**, 249–253.
- USGS (2012) United States Breeding Bird Survey. Available at: <http://www.pwrc.usgs.gov/bbs>.
- Van der Putten, W.H., Macel, M. & Visser, M.E. (2010) Predicting species distribution and abundance responses to climate change: why it is essential to include biotic interactions across trophic levels. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **365**, 2025–2034.
- Whittaker, R.J., Willis, K.J. & Field, R. (2001) Scale and species richness: towards a general, hierarchical theory of species diversity. *Journal of Biogeography*, **28**, 453–470.
- Wisz, M.S., Pottier, J., Kissling, W.D. *et al.* (2013) The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. *Biological Reviews*, **88**, 15–30.
- Zarnetske, P.L., Skelly, D.K. & Urban, M.C. (2012) Biotic multipliers of climate change. *Science*, **336**, 1516–1518.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's web-site.

Figure S1 Same as Fig. 1, but only taking the first (aggregated) stop within each route.

Figure S2 Changes across extent.

Figure S3 Same as Fig. 1, but using an extent of 1000 km.

BIOSKETCH

The authors are a working group initiated at the Ecological Society of America workshop: 'Scaling Up: Population and Community Ecology. A Workshop for Early Career Scientists'. The workshop focused on identifying key questions in population and community ecology that can or should be addressed at continental scales, assessing the status of existing analytical, physical, and software tools needed to address these questions, and identifying needs and capabilities for developing new tools to address continental-scale questions.

Editor: Allen Hurlbert

APPENDIX

Table A1 Environmental predictors used in the analyses.

Group	Source	Variable	Definition
Temperature	PRISM	BIO1	Annual mean temperature
	PRISM	BIO2	Mean diurnal range [mean of monthly (max. temp – min. temp)]
	PRISM	BIO3	Isothermality (BIO2/BIO7)
	PRISM	BIO4	Temperature seasonality (standard deviation)
	PRISM	BIO5	Max temperature of warmest month
	PRISM	BIO6	Min temperature of coldest month
	PRISM	BIO7	Temperature annual range (BIO5–BIO6)
	PRISM	BIO8	Mean temperature of wettest quarter
	PRISM	BIO9	Mean temperature of driest quarter
	PRISM	BIO10	Mean temperature of warmest quarter
	PRISM	BIO11	Mean temperature of coldest quarter
Precipitation	PRISM	BIO12	Annual precipitation
	PRISM	BIO13	Precipitation of wettest month
	PRISM	BIO14	Precipitation of driest month
	PRISM	BIO15	Precipitation seasonality (coefficient of variation)
	PRISM	BIO16	Precipitation of wettest quarter
	PRISM	BIO17	Precipitation of driest quarter
	PRISM	BIO18	Precipitation of warmest quarter
	PRISM	BIO19	Precipitation of coldest quarter
Land cover	NLCD	Open water	
	NLCD	Snow	
	NLCD	Developed – open space	
	NLCD	Developed – low intensity	
	NLCD	Developed – medium intensity	
	NLCD	Developed – high intensity	
	NLCD	Barren land	
	NLCD	Deciduous forest	
	NLCD	Evergreen forest	
	NLCD	Mixed forest	
	NLCD	Shrub/scrub	
	NLCD	Grassland/herbaceous	
	NLCD	Pasture/hay	
	NLCD	Cultivated cropland	
	NLCD	Woody wetlands	
	NLCD	Herbaceous wetlands	
	NBCD	avgCH_woNA	Average basal area–weighted canopy height
Productivity	MODIS	avgEVI	Average enhanced vegetation index

NLCD, National Land Cover Classification (Fry *et al.*, 2011); PRISM, climatic variables (PRISM, 2010); NBCD, National Biomass and Carbon Dataset (NBCD2000; Kellndorfer *et al.*, 2012).

Table A2 Summary statistics for the linear mixed-effect models.

Grain	Extent	R^2_M	R^2_C	Environment				Geography				Positive interactions				Negative interactions			
				Imp	Coef	SE	<i>P</i>	Imp	Coef	SE	<i>P</i>	Imp	Coef	SE	<i>P</i>	Imp	Coef	SE	<i>P</i>
1	500	0.08	0.38	0.64	0.18	0.13	0.251	0.75	0.42	0.16	0.151	0.68	0.15	35.83	0.250	0.58	0.07	0.16	0.253
2	500	0.11	0.36	0.58	0.12	112.95	0.274	0.75	0.47	0.19	0.153	0.68	0.16	38.95	0.208	0.66	0.06	0.19	0.165
5	500	0.16	0.37	0.55	0.16	3.46	0.321	0.72	0.46	0.19	0.183	0.69	0.02	54.56	0.227	0.71	0.21	0.13	0.154
10	500	0.19	0.38	0.50	0.17	3.97	0.309	0.71	0.49	0.22	0.164	0.63	0.22	27.55	0.249	0.68	0.21	0.14	0.165
25	500	0.24	0.43	0.41	0.25	74.38	0.421	0.64	0.57	0.30	0.222	0.48	0.31	0.36	0.323	0.55	0.21	0.17	0.223
50	500	0.34	0.54	0.38	0.43	119.37	0.455	0.57	0.85	0.54	0.239	0.41	0.70	90.75	0.394	0.42	0.07	0.31	0.311
1	1000	0.17	0.43	0.80	0.22	0.09	0.103	0.94	0.77	0.12	0.037	0.79	0.44	0.25	0.145	0.81	0.24	0.11	0.090
2	1000	0.19	0.43	0.75	0.24	0.11	0.117	0.95	0.84	0.14	0.029	0.80	0.47	0.23	0.115	0.85	0.24	0.11	0.068
5	1000	0.24	0.46	0.66	0.21	0.11	0.207	0.95	0.89	0.15	0.020	0.79	0.46	0.21	0.139	0.84	0.27	0.09	0.083
10	1000	0.27	0.47	0.60	0.20	0.13	0.247	0.92	0.92	0.18	0.043	0.77	0.47	0.23	0.130	0.83	0.26	0.10	0.082
25	1000	0.31	0.51	0.56	0.24	0.18	0.301	0.87	1.02	0.25	0.057	0.64	0.40	0.28	0.200	0.72	0.23	0.12	0.150
50	1000	0.36	0.55	0.52	0.24	0.25	0.313	0.86	1.16	0.37	0.069	0.57	0.46	0.36	0.229	0.57	0.21	0.15	0.250
1	2000	0.33	0.52	0.96	0.31	0.06	0.019	0.99	1.28	0.07	0.009	0.96	0.55	0.11	0.014	0.94	0.25	0.05	0.038
2	2000	0.36	0.53	0.94	0.29	0.06	0.028	0.99	1.35	0.08	0.008	0.98	0.53	0.10	0.012	0.97	0.27	0.05	0.016
5	2000	0.40	0.56	0.87	0.28	0.07	0.059	0.99	1.46	0.10	0.009	0.96	0.46	0.10	0.012	0.97	0.29	0.05	0.013
10	2000	0.44	0.59	0.82	0.27	0.08	0.074	1.00	1.60	0.13	0.001	0.93	0.40	0.11	0.023	0.96	0.27	0.05	0.015
25	2000	0.50	0.63	0.77	0.29	0.11	0.114	0.99	1.81	0.19	0.009	0.75	0.32	0.14	0.109	0.89	0.23	0.06	0.027
50	2000	0.55	0.66	0.75	0.36	0.14	0.126	1.00	2.03	0.27	0.001	0.59	0.25	0.17	0.269	0.74	0.19	0.08	0.105
1	4000	0.43	0.64	0.99	0.37	0.04	0.001	1.00	1.67	0.05	0.000	1.00	0.52	0.06	0.000	1.00	0.23	0.03	0.000
2	4000	0.46	0.65	0.99	0.35	0.05	0.002	1.00	1.75	0.06	0.000	1.00	0.50	0.05	0.000	1.00	0.24	0.03	0.000
5	4000	0.52	0.68	0.95	0.32	0.05	0.018	1.00	1.93	0.08	0.000	1.00	0.44	0.05	0.000	1.00	0.26	0.03	0.000
10	4000	0.56	0.71	0.91	0.32	0.06	0.033	1.00	2.10	0.10	0.000	0.99	0.37	0.06	0.001	1.00	0.24	0.03	0.000
25	4000	0.65	0.76	0.88	0.34	0.08	0.046	1.00	2.51	0.15	0.000	0.95	0.30	0.08	0.012	0.99	0.20	0.03	0.003
50	4000	0.71	0.79	0.87	0.36	0.10	0.050	1.00	2.87	0.22	0.000	0.82	0.28	0.10	0.063	0.94	0.18	0.04	0.020

Imp, importance; Coef, coefficient; SE, standard error.