Global Ecology and Biogeography, (Global Ecol. Biogeogr.) (2015) **24**, 1225–1235

A Journal of Macroecology

Global Ecology and Biogeography

RESEARCH PAPER

Bird diversity and environmental heterogeneity in North America: a test of the area–heterogeneity trade-off

Rachel Chocron[1], Curtis H. Flather[2] and Ronen Kadmon[1]*

[1]*Department of Ecology, Evolution and Behavior, Institute of Life Sciences, Hebrew University of Jerusalem, Givat-Ram, Jerusalem 91904, Israel,* [2]*USDA Forest Service, Rocky Mountain Research Station, Fort Collins, CO 80526, USA*

*Correspondence: Ronen Kadmon, Department of Ecology, Evolution and Behavior, Institute of Life Sciences, Hebrew University of Jerusalem, Givat-Ram, Jerusalem 91904, Israel.
E-mail: kadmon@mail.huji.ac.il

ABSTRACT

Aim Deterministic niche theory predicts that increasing environmental heterogeneity increases species richness. In contrast, a recent stochastic model suggests that heterogeneity has a unimodal effect on species richness since high levels of heterogeneity reduce the effective area available per species, thereby increasing the likelihood of stochastic extinction (the 'area–heterogeneity trade-off'). We tested these contrasting predictions using data on bird distributions in North America.

Location North America.

Methods The effect of heterogeneity on species richness was tested using simultaneous autoregressive regression models based on two measures of heterogeneity (elevational range and land-cover richness) each quantified at two scales (400 m, 5 km), three measures of species richness (observed, corrected for incomplete detection, and corrected for regional richness) and three variable selection methods [forced entry, Akaike information criterion (AIC)-based and a null-model approach]. Covariates included precipitation, temperature, elevation and latitude. For all variables, both linear and quadratic terms were included in the analyses.

Results Overall, heterogeneity had a weak effect on species richness and the contribution of the quadratic term of heterogeneity to the explained variance was very small (< 1%). Nevertheless, in all 36 models, the coefficients of both the linear and quadratic terms of heterogeneity were statistically significant and the estimated inflection point was within the range of the data, as predicted by the area–heterogeneity trade-off. Moreover, in 30 out of the 36 models, support for a unimodal effect of heterogeneity on species richness based on information theoretic criteria was strong ($\Delta AIC > 10$), and in 22 of those 30 models the null hypothesis of a monotonically positive relationship could be rejected at the 0.05% significance level.

Main conclusions Patterns of bird richness in North America were predominantly consistent with the predictions of the area–heterogeneity trade-off. Future attempts to understand the mechanisms affecting species diversity should pay more attention to the potential consequences of this fundamental trade-off.

Keywords

BBS, community ecology, elevational range, habitat diversity, land-cover richness, niche theory, species richness.

INTRODUCTION

A fundamental notion in ecology is that species richness increases with increasing heterogeneity in environmental conditions (hereafter environmental heterogeneity). This generalization is supported by numerous studies (Stein *et al.*, 2014) and is attributed to the fact that a more heterogeneous environment provides suitable conditions for a larger number of species with different ecological requirements (Rosenzweig, 1995).

DOI: 10.1111/geb.12353
http://wileyonlinelibrary.com/journal/geb

In contrast to this widely accepted notion, a recent model integrating the main elements of niche theory (environmental heterogeneity and niche partitioning) and island biogeography theory (stochastic colonization and extinction processes) suggests that environmental heterogeneity has a unimodal, rather than a positive, effect on species richness (Kadmon & Allouche, 2007). According to this model, environmental heterogeneity has two opposing effects on species richness: it increases opportunities for niche partitioning (as predicted by classical niche theory; Hutchinson, 1957) but at the same time reduces the amount of *suitable* area available for individual species, thereby increasing the probability of stochastic extinctions (as predicted by island biogeography theory; MacArthur & Wilson, 1967). The trade-off between these positive and negative effects (hereafter, the area–heterogeneity trade-off, AHTO) should lead to a general unimodal, rather than a positive, heterogeneity–diversity relationship.

Hortal *et al.* (2009) questioned the model proposed by Kadmon & Allouche (2007), arguing that its prediction that species richness may decrease with increasing heterogeneity contradicts empirical evidence. They further argued that this prediction stems from a crucial assumption of the model – namely, that each species is able to become established in only one type of habitat. Based on simulations of an alternative model and empirical analysis of 24 insular systems they concluded that '. . . species richness increases monotonically with increased habitat diversity and never decreases' (Hortal *et al.*, 2009, E213).

Since Hortal *et al.* (2009) published their paper, considerable evidence has accumulated against their conclusion. First, growing evidence indicates that negative heterogeneity–diversity relationships do occur in natural communities much more often than could be attributed to chance (see Tamme *et al.*, 2010; Seiferling *et al.*, 2014; Stein *et al.*, 2014 for recent meta-analyses). Although none of these meta-analyses has explicitly tested for unimodal responses, Tamme *et al.* (2010) found that negative heterogeneity–diversity relationships were more frequent at small spatial scales, and Seiferling *et al.* (2014) found that negative relationships were particularly frequent in systems representing intermediate levels of anthropogenic 'footprints' which they interpreted as the most heterogeneous systems. Both conclusions are fully consistent with the predictions of the AHTO. The claim that the unimodal response predicted by Kadmon & Allouche (2007) stems from simplifying assumptions of their model is also contradicted by recent studies showing that similar predictions are obtained from different models (Allouche *et al.*, 2012; Laanisto *et al.*, 2013; de Souza Júnior *et al.*, 2014). These theoretical findings suggest that the prediction of unimodal heterogeneity–diversity relationship is model-independent, and can be generalized to a wider range of scenarios than those originally proposed by Kadmon & Allouche (2007).

Unfortunately, explicit tests of the AHTO are far from trivial. First, detecting a significant unimodal response of species richness to environmental heterogeneity requires data representing a wide range of heterogeneity *within the observation units* since the expected relationship (positive, negative or unimodal) is conditioned on the range of heterogeneity captured among observation units (see Seiferling *et al.*, 2014). Second, the spatial scale at which the data are collected (the grain size of the observation unit) should match the scale at which stochastic extinctions can be detected, which is usually unknown. Theoretical analyses (Allouche *et al.*, 2012) show that data collected over relatively large spatial scales are not expected to show the decreasing phase of the heterogeneity–diversity relationship because probabilities of stochastic extinction decrease with increasing scale. However, small observation units may also be inappropriate for testing the AHTO because variation in species richness at small spatial scales may reflect behavioural rather than demographic processes (MacArthur, 1965; Recher, 1969). Thus, although unimodal heterogeneity–diversity relationships have previously been documented at small spatial scales (Bar-Massada & Wood, 2014), empirical tests of the AHTO should focus on intermediate spatial scales, where local colonization–extinction processes rather than behavioural or biogeographical processes determine species diversity.

The temporal scale over which the data are collected may also influence the functional form of the heterogeneity–diversity relationship. Theoretical analyses suggest (Allouche *et al.*, 2012) and empirical data confirm (Hurlbert & White, 2005) that pooling of occurrence data over time introduces positive bias in estimates of species richness and may convert negative heterogeneity–diversity relationships into positive ones. For this reason, analyses of heterogeneity–diversity relationships based on data obtained from herbaria (Jimenez *et al.*, 2009), museum collections (Rahbek & Graves, 2001), atlases based on long-term observations (Fløjgaard *et al.*, 2011), regional checklists (Ricklefs & Lovette, 1999), range maps (Davies *et al.*, 2007), niche models (Lennon *et al.*, 2000) or databases mixing biodiversity data from different sources (GBIF; Finch *et al.*, 2008), are not useful for testing the AHTO. Thus, an optimal dataset for testing the AHTO should be based on direct observations taken at a large number of sites of intermediate grain size, that represent a wide range of environmental heterogeneity and are sampled simultaneously over a short time interval to reduce bias due to temporal turnover. As in any analysis, a large geographical extent may strengthen the results by confirming that the observed pattern is not an artefact of a particular location or a specific combination of environmental conditions.

The North American Breeding Bird Survey (BBS; Sauer *et al.*, 2014) is a unique data source that satisfies all of the above considerations. The survey is carried out simultaneously in thousands of sites every year and combines a large geographical extent (the North American continent), an intermediate grain size (roadside routes 39.4-km long), a short sampling period (the peak of the nesting season) and a standardized sampling protocol. Previous analyses of BBS data have shown that local extinctions are common at the route scale (Boulinier *et al.*, 1998) and that the dynamics of individual species has a strong stochastic component (Boucher-Lalonde *et al.*, 2014; Kalyuzhny *et al.*, 2014). We therefore expected the BBS to be an appropriate source of empirical evidence for testing the AHTO hypothesis.

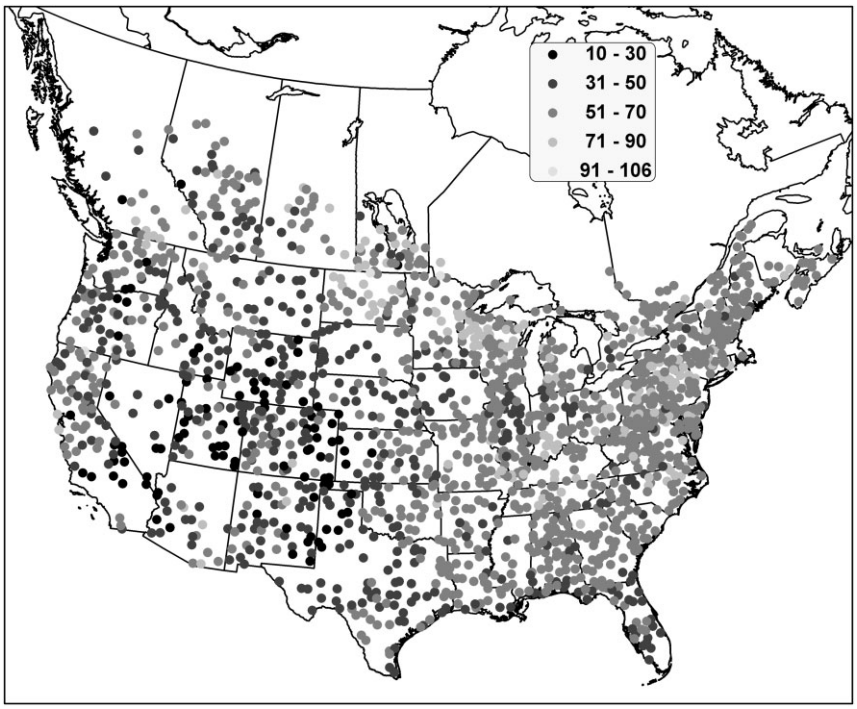

Figure 1 Map of the 1783 North American Breeding Bird Survey routes included in the analysis. Each route is shaded according to its value of observed species richness (mean number of species per year between 1996 and 2006).

We tested our hypothesis by analysing the relationship between the mean number of species per survey route and two measures of environmental heterogeneity: elevational range (maximum minus minimum elevation) and land-cover richness (number of different land-cover types). Elevational range is a complex environmental gradient encompassing variation in both abiotic and biotic components of the environment. Land-cover richness reflects both natural and human land-use effects, and may capture aspects of habitat conditions not accounted for by elevational range.

Overall, our study had two main objectives. The first was to test whether patterns of breeding bird richness in North America exhibit the unimodal heterogeneity–diversity relationship predicted by the AHTO. The second was to evaluate if the empirical results were robust to the manner by which the data were analysed, with particular emphasis on (1) the measure used to quantify heterogeneity, (2) the scale at which heterogeneity was quantified, (3) the manner by which species richness was estimated (with or without different kinds of standardization), and (4) the manner by which the statistical model was specified. The latter issue is particularly important in statistical tests of factors affecting species richness, because such factors almost always show spatial correlation which can influence variable selection and their qualitative effects on the response variable (Olden & Jackson, 2000).

METHODS

Selection of BBS data for the analysis

We accumulated annual bird observation records over those years when each BBS route was run within an 11-year window (1996 to 2006) centred on the date of the imagery we used to determine land-cover heterogeneity (2001). Because annual BBS bird counts completed on a single morning during the breeding season can be highly variable (see Donovan & Flather, 2002), we estimated an annual average within the 11-year period to control for annual variation attributable to factors other than environmental conditions near the route (e.g. the idiosyncrasies of bird behaviour on any given day). A temporal window of this length would be problematic if land-cover heterogeneity changed substantially over this period. However, Rittenhouse *et al.* (2012) found little land-cover transition (*c.* 3%) over a 10-year period using similar satellite-based imagery. A total of 1783 routes distributed over the conterminous United States and southern Canada (Fig. 1) met our selection criteria (see Appendix S1 in Supporting Information).

Calculation of environmental heterogeneity

The distance within which birds are sampled according to the BBS protocol is limited to 400 m from the survey route. We therefore used a buffer of 400 m around a route as a reasonable approximation for the environment faced by birds detected along the route. However, birds sampled within this distance probably experience a larger area throughout their life-cycle track (Tittler *et al.*, 2009). We used the allometric model relating natal dispersal to body mass presented by Sutherland *et al.* (2000, Table 2) to determine a buffer zone that may better reflect the scale of population processes that integrate the effects of habitat heterogeneity on the number of species that could be detected on a particular BBS route. We used only body mass data (from Dunning, 2007) that were measured in North America during the breeding season. For each species, we calculated the

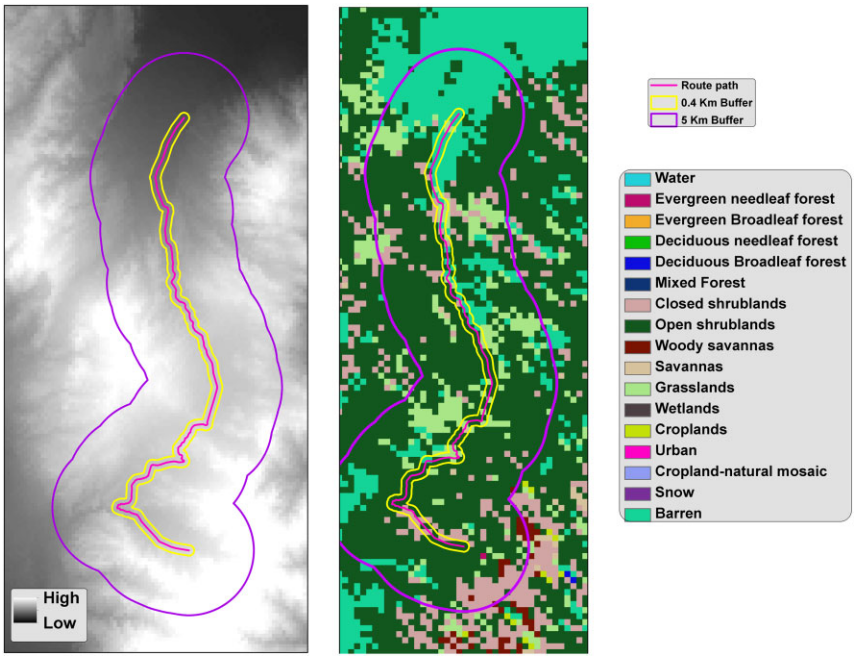

Figure 2 Calculation of environmental heterogeneity around a North American Breeding Bird Survey route. Two buffers of constant distance were calculated around each route: 400-m wide and 5-km wide. Elevational range was calculated using a digital elevation model of 250-m resolution, as maximum elevation minus minimum elevation within the buffer (left panel). Land-cover richness was calculated using a land-cover imagery of 500-m resolution, as the number of land-cover classes within the buffer (right panel).

median natal dispersal distance using the allometric model and then averaged the median distances among all species we detected.

The resulting value (4.64 km) was rounded to 5 km, which was used as an additional buffer size for the analysis. Within each buffer (400 m and 5 km) we quantified the degree of environmental heterogeneity using two different indices: elevational range and land-cover richness (Fig. 2). Elevational range was determined from digital elevation data with a spatial resolution of 0.25 km. Land-cover richness was determined using a land-cover map with 17 land-cover classes and spatial resolution of 0.5 km based on MODIS imagery data from 2001. Although the calculation of land-cover richness could be influenced by map resolution, testing this effect was beyond the scope of our analysis. (see Appendix S2 for details on the calculation of the explanatory variables).

It should be noted that values of both elevational range and land-cover richness were positively, but weakly, correlated over the two scales (Table S1). Still, quantifying the magnitude of heterogeneity at the two scales was important since previous analyses of BBS data found that the size of the buffer may influence the relationship between landscape properties and species richness (e.g. Mayer & Cameron, 2003). Similarly, the use of both elevational range and land-cover richness as measures of heterogeneity was important, since both measures have been used extensively in previous studies and have often shown different effects on species richness (e.g. Moreno-Rueda & Pizarro, 2009).

Calculation of other environmental variables

In addition to environmental heterogeneity we determined for each route three key variables that have previously been found to be important in determining species richness by influencing the inherent capacity of different environments to support different kinds of species (Appendix S2): mean annual precipitation (van Rensburg *et al.*, 2002), mean summer temperature (Lennon *et al.*, 2000) and mean elevation (Hurlbert & White, 2005). We also determined geographic coordinates of the centre of each route in order to control for possible broad-scale spatial variation in species richness through trend-surface modelling (quadratic) as recommended by Legendre & Legendre (2012). However, preliminary analyses showed that the Y coordinate was strongly correlated with summer temperature ($R = -0.838$). We therefore used only the X coordinate in our analyses. Pearson correlation coefficients among all independent variables used in our analyses are provided in Table S1.

Calculation of species richness

For each route, we tallied the number of species detected within a given year along the route. The resulting annual richness values were averaged over the years for which appropriate data were available for each route. We call this measure 'observed species richness'.

Raw BBS counts are known to be downwardly biased (Nichols *et al.*, 1998). This bias might be particularly problematic if detection probabilities vary between different habitat types or different levels of environmental heterogeneity. We used capture–recapture models for closed populations based on the jackknife estimators as proposed by Burnham & Overton (1979) to correct for such bias. We corrected the raw species counts obtained for each route and year using the COMDYN software (Hines *et al.*, 1999) and then averaged the estimated values over the relevant years. We call the resulting value 'species richness corrected for incomplete detection'.

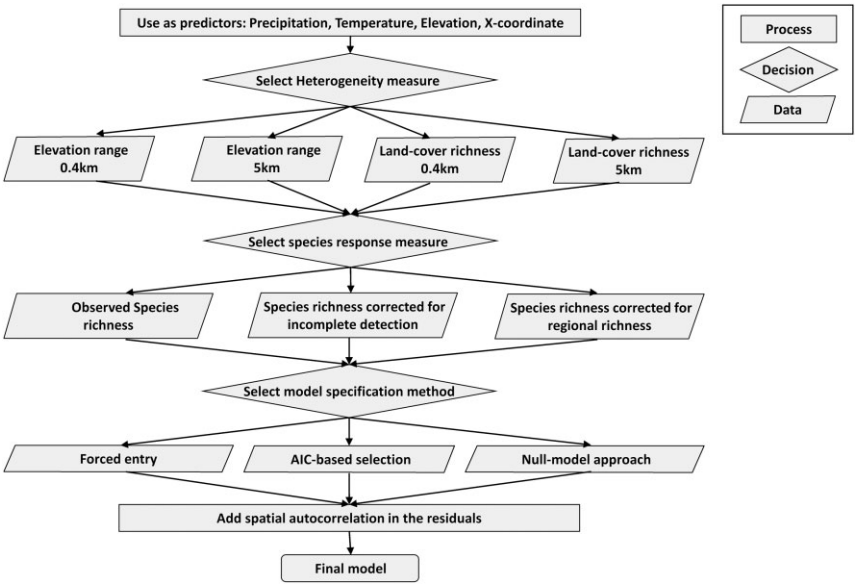

Figure 3 A flow chart of the procedure used to construct models testing the effects of elevational range and land-cover richness on species richness. This procedure resulted in 36 models (two measures of heterogeneity × two spatial scales × three species response variables × three variable selection methods).

Another potential confounding factor that needs to be addressed when trying to explain variation in local species richness is the regional species pool size (Cam *et al.*, 2000). In order to control for pool size variation we standardized the estimate of species richness obtained for each route by the number of species recorded in the respective Bird Conservation Region (BCR). BCRs are ecologically distinct geographic units with similar bird communities, habitats and management, that have previously been shown to provide reasonable strata for BBS-based analyses (Sauer *et al.*, 2003). We determined the species pool of each BCR based on all the species recorded in the respective BBS routes between 1996 and 2006. We then divided the observed species richness of each route by its BCR richness, and transformed the resulting proportion using arcsine transformation. We term the resulting measure 'species richness corrected for regional richness'.

Statistical analysis

We used multiple linear regressions to test the effect of environmental heterogeneity (elevational range within a 400-m buffer, land-cover richness within a 400-m buffer, elevational range within a 5-km buffer and land-cover richness within a 5-km buffer) on each measure of species richness (observed species richness, species richness corrected for incomplete detection and species richness corrected for regional richness). Each of these 12 tests was performed using three methods of variable selection in order to evaluate the robustness of the results:

1. Forced entry: all independent variables (the relevant measure of heterogeneity, mean annual precipitation, mean summer temperature, mean elevation and the *X* coordinate) as well as the squared term of each variable are entered into the model simultaneously (a total of 10 predictor variables per model).

2. Akaike information criterion (AIC)-based selection: construction of linear regression models with all possible combinations of the 10 predictor variables, ranking the models according to the AIC and selecting the model with the lowest AIC value.

3. A null model approach: building linear regression models with all possible combinations of the independent variables and their squared terms except for environmental heterogeneity, and selecting the model with the lowest AIC value as a null model. Then adding the relevant measure of environmental heterogeneity and its squared term to the null model.

Figure 3 presents a flow chart of the statistical analysis. Each statistical model consisted of three steps: selection of a candidate measure of heterogeneity (in addition to the other independent variables), selection of a species response measure and selection of a statistical method. This procedure resulted in 36 statistical models.

After building the models we tested whether their residuals were spatially autocorrelated using Moran's *I* correlograms. In all models the residuals were spatially autocorrelated, with Moran's *I*-values exceeding 0.1 up to distances of 200 km. We therefore re-estimated all models using simultaneous autoregressive regression (SAR). The SAR models were performed using binary connectivity matrices with a threshold distance of 200 km based on the empirical correlograms. All the results presented in this paper are for SAR models.

For each final model, the effect of each independent variable was defined by checking the significance and sign of the regression coefficients. If the coefficient of the squared term was statistically significant ($P < 0.05$) and negative and the inflection point was inside the range of the data, the effect was defined as unimodal. If the coefficient of the squared term was significant and negative and the inflection point was outside this range, the effect was defined as positive decelerating (inflection point to the right of the range) or accelerating (inflection point to the left of the range). If the coefficient of the squared term was significant and positive and the inflection point was inside the range of the data, the effect was defined as U-shaped. If the coefficient of the squared term was significant and positive, and the inflection

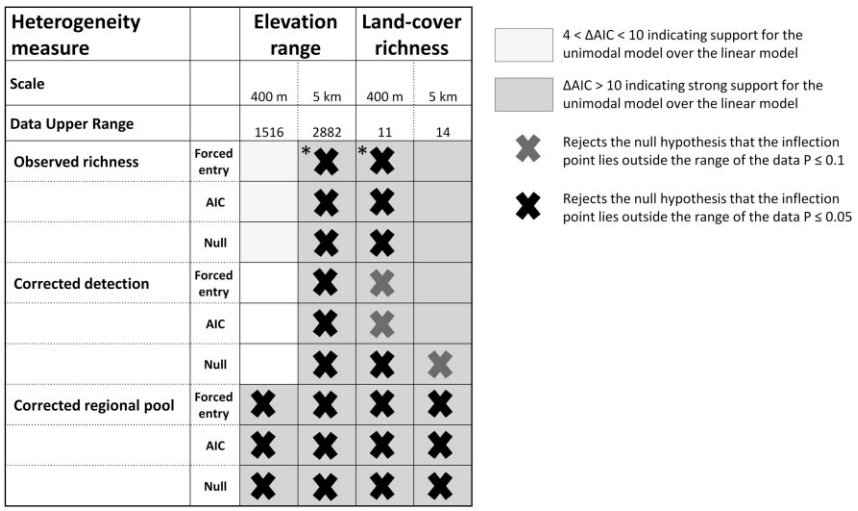

Figure 4 Summary of the 36 models used for testing the effect of environmental heterogeneity on bird richness in North America. Shading indicates the level of support for a unimodal effect over a linear effect based on information theoretic criteria (note that across all 36 models ΔAIC was > 2 and the inflection point estimate was within the range of the data). Crosses indicate the significance level at which the null hypothesis that the inflection point lies outside the range of the data (i.e. that the response is monotonically positive) can be rejected based on its confidence interval. Confidence intervals were calculated using the conservative approach proposed by Mandel (2013). An asterisk (*) indicates the models shown in Fig. 5.

point was outside the range of the data, the effect was defined as negative decelerating (inflection point to the right of the range) or positive accelerating (inflection point to the left of the range). If only the linear coefficient was significant and positive the effect was defined as positive linear, and if it was negative the effect was defined as negative linear.

Differences in goodness of fit among the various models were evaluated based on their AIC value and the Nagelkerke (1991) pseudo R^2 index (termed here R^2_N). In cases where the effect of heterogeneity was classified as significantly unimodal, we also calculated the difference in AIC between the relevant model and a nested model from which the squared term of heterogeneity was removed. A difference of at least two is interpreted as a support for the conclusion that the unimodal model better fits the data (Burnham & Anderson, 2002).

In order to be most conservative, in cases where the effect of heterogeneity was significantly unimodal *and* the inflection point estimate was inside the range of the data, *and* the difference in AIC was > 2, we also calculated one-side 90% and 95% confidence intervals to the inflection point and tested the hypothesis that the upper range of the data (elevational range or land-cover richness) is beyond this range. Confidence intervals were determined using the method proposed by Mandel (2013) since this method does not require assuming normality and is more conservative than the standard Delta method (Mandel, 2013). One-sided confidence intervals were used because our hypothesis was that the upper range of the data is *higher* than the inflection point (i.e. a one-tailed hypothesis, where the null hypothesis is that the upper range of the data is equal to, or lower than, the inflection point). All statistical analyses were carried out using R (R Development Core Team, 2013).

RESULTS

The results obtained for the 36 models were highly consistent: for all models, the linear term of environmental heterogeneity was positive, the quadratic term was negative, both terms were highly significant and the increase in AIC value when the quadratic term was removed from the model was greater than two (Table S1). Moreover, in 30 out of the 36 models, the evidence based on information theoretic criteria revealed strong (ΔAIC > 10) support for the unimodal specification; and in 22 of those instances the null hypothesis that the inflection point is above the upper limit of the data (i.e. that the actual relationship is monotonically positive) could be rejected at the 0.05% level (Fig. 4). These overall results indicate that in the majority of cases variation in species richness among BBS routes was better fitted by the unimodal response predicted by the AHTO than the positive response predicted by classical niche theory. It should be noted, however, that the effects of both measures of heterogeneity on species richness were very weak and noisy, and the contribution of the quadratic term of heterogeneity to the explained variance was < 1% in all cases (see Fig. 5 for typical examples).

There was also a strong consistency in the effects of the various properties of the models on their explanatory power. First, models in which richness was corrected for differences in regional species richness always showed a much better fit to the data than corresponding models focusing on observed richness or richness corrected for incomplete detection (Fig. 6a, Table S2). This result was consistent under all methods of variable selection, for both measures of heterogeneity, and for both scales (Tables S2, S3). These models also showed the largest increase in AIC values when the squared term of heterogeneity was removed from the model (Fig. 6b, Table S2), and in all of them, the upper limit of the data was significantly above the 95% confidence limit of the inflection point (Figs 4 & 6c, Table S2).

Second, models in which the data were corrected for incomplete detection always showed a lower explanatory power than models focusing on the observed richness (Fig. 6a, Table S2). These models also showed the lowest increase in AIC values

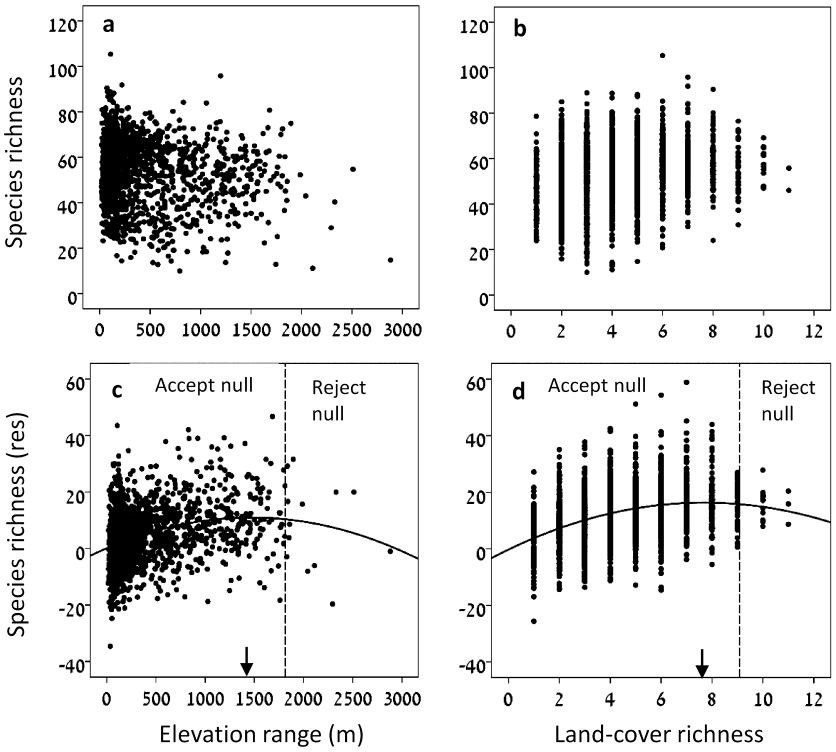

Figure 5 Examples of relationships between species richness and the two measures of environmental heterogeneity before (a, b) and after (c, d) correcting for the effects of all other variables. In plots (a) and (b) the response variable is observed richness. In plots (c) and (d) the response variable is the partial residuals of observed richness, as obtained from the models that maximized ΔAIC in analyses based on elevational range (c) and land-cover richness (d) as measures of heterogeneity (indicated by asterisks in Fig. 4). Note that plotting the partial residuals of species richness as the response variable allows us to visualize the response predicted by the model, the inflection point estimate (arrow) and the upper limit of the one-side 95% confidence level calculated for the inflection point (the vertical dashed line). For both models, the upper limit of the data falls above the 95% confidence interval, implying that the null hypothesis of a monotonically positive response can be rejected at the 0.05% significance level.

when the squared term of heterogeneity was removed from the model (Fig. 6b, Table S2). Third, models based on land-cover richness always performed better (in terms of both R^2_N and AIC values) than those based on elevational range (Fig. 6a, Tables S2 & S3). These models also showed a larger increase in AIC values when the quadratic term of heterogeneity was removed from the model (Fig. 6b, Table S2). The differences between the two measures of heterogeneity were consistent for all methods of variable selection, all measures of species richness and both spatial scales, indicating that land-cover richness is a better predictor of bird richness than elevational range in our study system.

When both elevational range and land-cover richness were entered simultaneously as predictors to the models and the analysis was repeated using all possible combinations of response variables, variable selection methods and scale (Fig. 3), the best model in terms of both AIC and R^2_N values was the one in which both measures of heterogeneity were incorporated within a buffer of 400 m, species richness was corrected for regional richness and the explanatory variables were selected using the null-model approach (Table 1). In this final model the effects of both land-cover richness and elevational range were significantly unimodal (Table 1) and removing their squared term increased the AIC value of the model by 35.9 and 6.6, respectively.

DISCUSSION

Our analysis, encompassing 1783 BBS routes representing a large part of North America, indicated that environmental heterogeneity has a rather weak though statistically significant unimodal effect on species richness, as predicted by the AHTO (Kadmon & Allouche, 2007). This result was highly robust to the method of variable selection, the manner by which species richness was standardized, the measure of heterogeneity and the scale at which the magnitude of heterogeneity was quantified. Quantitative differences among the various models were also highly consistent: other things being equal, models in which species richness was corrected for differences in regional richness provided the best fit to the data, and land-cover richness was a better predictor of species richness than elevational range. It should be noted, however, that the overall effect of environmental heterogeneity on bird richness was rather weak and explained an extremely small portion of the observed variation. In this respect our findings are similar to previous studies showing that environmental heterogeneity is a very weak and usually insignificant predictor of continental-scale variation in richness of BBS routes (see Table S4 for a detailed comparison). However, except for one study, none of the previous analyses had incorporated the quadratic term of environmental heterogeneity as a potential predictor in the model.

Our finding that environmental heterogeneity had a significant unimodal effect on bird richness in the majority of the 36 model cases we examined is notable given that studies focusing on larger scales have usually detected significant positive effects of topographic and/or land-cover heterogeneity on bird richness (Stein *et al.*, 2014). We attribute this difference to two mechanisms. First, since larger areas support larger populations, the likelihood of extinction (the probability that all individuals of a certain species would die or emigrate within a given time

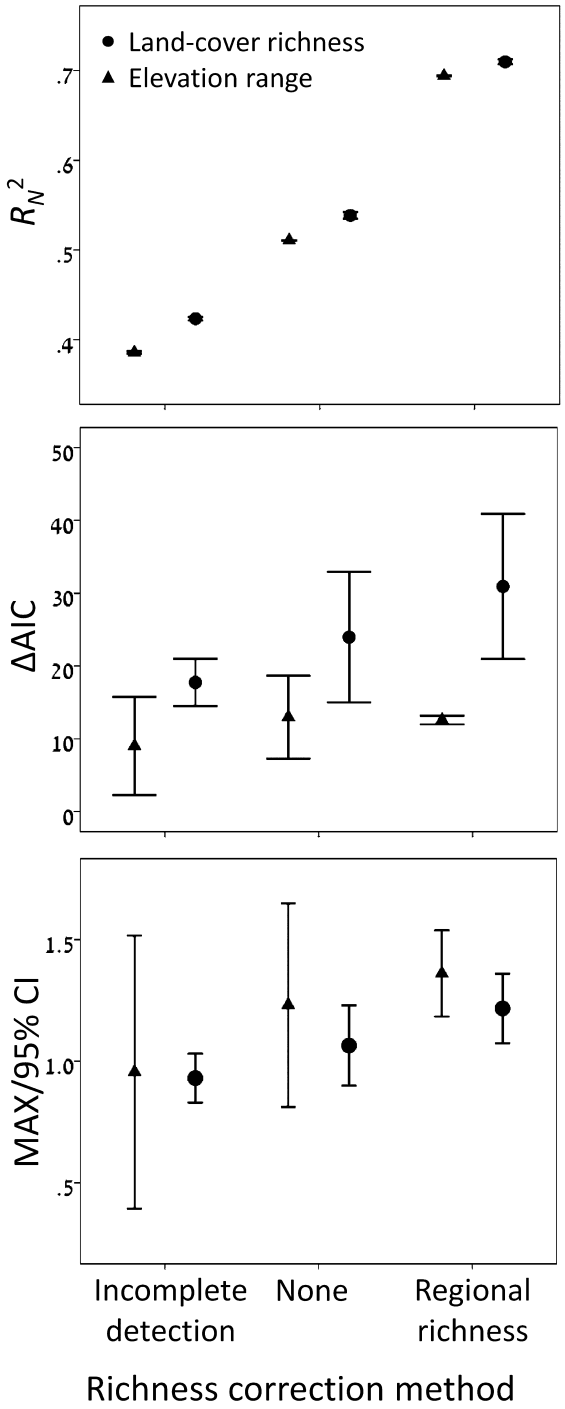

Figure 6 Effect of the measure of species richness (corrected for incomplete detection, corrected for regional richness, or no correction) and the measure of heterogeneity (elevational range versus land-cover richness) on (top) the portion of overall variance accounted for by the model; (middle) the difference in AIC between a model including the squared term of heterogeneity and a model from which the squared term was removed; and (bottom) the ratio between the upper limit of the data (maximum elevational range or maximum land-cover richness at the relevant scale) and the upper limit of 95% confidence level calculated for the inflection point. See Tables S2 &S3 for full results of all models.

interval) decreases with increasing scale. It is therefore unlikely that stochastic extinctions would balance the positive effect of heterogeneity at large spatial scales (Allouche *et al.*, 2012). Second, studies focusing on large spatial scales are usually based on distribution data obtained from the scientific literature, atlases, checklists or range maps, rather than direct observations (Stein *et al.*, 2014). Such secondary data sources contain information collected over long time spans, during which colonization events are accumulated but extinction events are ignored. As a result, analyses based on such data are not expected to show a decrease of species richness at high levels of heterogeneity. Models of the AHTO are consistent with these hypotheses, and show that increasing the spatial and/or temporal scale of the analysis may shift negative heterogeneity–diversity relationships into positive ones (Kadmon & Allouche, 2007; Allouche *et al.*, 2012).

Two recent analyses of BBS data provide direct support for this scale dependence. Veech & Crist (2007) analysed the effect of spatial scale on patterns of species diversity and found that spatial aggregation of BBS routes shifts the relationship between elevational range and species richness from negative to positive. Hurlbert & White (2005) analysed the effect of temporal scale on patterns of species diversity and found that temporal aggregation of data collected at the route scale leads to a similar shift. Importantly, while previous studies have attributed this scale dependence to statistical processes (an increase in the range of environmental heterogeneity with increasing spatial scale, or an increase in detection probability with increasing temporal scale) we attribute these patterns to a decrease in the likelihood of stochastic extinctions (in the case of spatial aggregation) or bias caused by accumulating colonization events while ignoring extinction events (in the case of temporal aggregation). The latter conclusion is further supported by a previous analysis of BBS data showing that, although route-scale richness increases monotonically with temporal aggregation, incomplete sampling ceases to limit the cumulative number of species after 2–4 years (White, 2004). This finding confirms that aggregating data over time-scales longer than 4 years induces a positive bias to the actual number of coexisting species.

Our conclusion that patterns of species richness at the scale of BBS routes are influenced by stochastic extinctions is further supported by previous studies showing that stochastic processes are important in determining the dynamics of bird populations at this scale (Keitt & Stanley, 1998; Kalyuzhny *et al.*, 2014). Moreover, recent analyses of factors affecting extinction rates of bird populations at the scale of BBS routes demonstrate that extinction probabilities increase with increasing environmental heterogeneity (Stegen *et al.*, 2013) and are negatively correlated with population size (Boucher-Lalonde *et al.*, 2014). These overall findings strengthen support for the hypothesis that the unimodal heterogeneity–diversity relationship observed in this study is driven by the AHTO.

The observed unimodal response of North American birds to elevational range resembles the pattern obtained by Allouche *et al.* (2012) for breeding birds in Catalonia. The unimodal pattern obtained for Catalonia was subject to two comments,

Table 1 Results of the best simultaneous autoregressive regression model (lowest AIC and highest R^2_N) for the effect of all variables examined in this study on bird richness in North American Breeding Bird Survey routes.

	Coefficient	SE	*z*-value	*P*
Intercept	2.19×10^{-1}	3.85×10^{-2}	5.7	1.3×10^{-8}
Mean annual precipitation	3.26×10^{-1}	2.53×10^{-2}	12.9	$<1.0 \times 10^{-15}$
Mean annual precipitation^2	-1.44×10^{-1}	1.29×10^{-2}	−11.1	$<1.0 \times 10^{-15}$
Mean summer temperature	1.50×10^{-2}	3.54×10^{-3}	4.2	2.2×10^{-5}
Mean summer temperature^2	-3.92×10^{-4}	9.12×10^{-5}	−4.3	1.7×10^{-5}
X coordinate	1.21×10^{-3}	2.76×10^{-4}	4.4	1.1×10^{-5}
X coordinate^2	-1.50×10^{-4}	1.70×10^{-5}	−8.8	$<1.0 \times 10^{-15}$
Elevation	-9.03×10^{-2}	1.27×10^{-2}	−7.1	1.2×10^{-12}
Elevation^2	1.66×10^{-2}	4.50×10^{-3}	3.7	2.3×10^{-4}
Elevational range 400 m	1.02×10^{-1}	2.23×10^{-2}	4.6	4.6×10^{-6}
Elevational range 400 m^2	-5.04×10^{-2}	1.72×10^{-2}	−2.9	3.4×10^{-3}
Land-cover richness 400	2.51×10^{-2}	2.84×10^{-3}	8.8	$<1.0 \times 10^{-15}$
Land-cover richness 400 m^2	-1.72×10^{-3}	2.78×10^{-4}	−6.2	6.0×10^{-10}

Response variable, species richness corrected for regional richness; method of variable selection, null model approach (forced entry approach provided similar results); R^2_N of the model, 0.72.

one attributing it to correlation between elevational range and mean elevation (Hortal *et al.*, 2013) and the other attributing it to sampling bias and non-random distribution of environmental conditions and species pools along the heterogeneity gradient (Carnicer *et al.*, 2013). Although further analyses of the Catalonian data rejected these alternative explanations (Allouche *et al.*, 2013a,b), these comments have pointed to the importance of controlling for alternative mechanisms in empirical tests of the AHTO. In this study we controlled for spatial variation in mean elevation by incorporating its linear and quadratic effects in all statistical models. We also controlled for variation in regional species richness, other key factors that are known to affect large-scale patterns of species richness (precipitation and temperature) and local spatial autocorrelation. Sampling bias is probably not a significant factor in our analysis since BBS data are obtained using a standardized sampling protocol. It should also be noted that the two measures of heterogeneity were very weakly correlated (elevational range explained less than 5% of the variation in land-cover richness) and both had unimodal effects on species richness in the best model. Other aspects of the patterns observed in this study (the superiority of land-cover richness as a predictor of richness, the improvement obtained by correcting for regional richness) were also highly robust. Thus, although we cannot totally rule out alternative explanations, it is unlikely that all of these responses were statistical artefacts.

Summary

Our analysis of breeding bird distribution in North America shows that both elevational range and land-cover richness, two of the most widely used measures of environmental heterogeneity (Stein *et al.*, 2014), exhibit statistically significant and highly robust unimodal effects on species richness. These findings are consistent with previous evidence for strong stochasticity (Keitt & Stanley, 1998; Kalyuzhny *et al.*, 2014), frequent extinctions (Boulinier *et al.*, 1998) and positive effects of environmental heterogeneity on extinction probability (Stegen *et al.*, 2013) in BBS routes. All of these patterns are consistent with the predictions of the AHTO. We conclude that future attempts to understand the mechanisms affecting the diversity of ecological communities should pay more attention to the potential consequences of this fundamental trade-off.

ACKNOWLEDGEMENTS

We thank the USGS Patuxent Research Center for facilitating the data used in this analysis, and all the volunteers of the BBS for their extraordinary work. We also thank two anonymous referee, V. Devictor and D. Currie for valuable comments on the manuscript. This work was supported by the Israeli Ministry of Science and Technology TASHTIOT Program and by the Israeli Science Foundation grants no. 454/11 and 1026/11.

REFERENCES

Allouche, O., Kalyuzhny, M., Moreno-Rueda, G., Pizarro, M. & Kadmon, R. (2012) Area–heterogeneity tradeoff and the diversity of ecological communities. *Proceedings of the National Academy of Sciences USA*, **109**, 17495–17500.

Allouche, O., Kalyuzhny, M., Moreno-Rueda, G., Pizarro, M. & Kadmon, R. (2013a) Reply to Carnicer et al.: environmental heterogeneity reduces breeding bird richness in Catalonia by increasing extinction rates. *Proceedings of the National Academy of Sciences USA*, **110**, E2861–E2862.

Allouche, O., Kalyuzhny, M., Moreno-Rueda, G., Pizarro, M. & Kadmon, R. (2013b) Reply to Hortal et al.: patterns of bird distribution in Spain support the area-heterogeneity tradeoff. *Proceedings of the National Academy of Sciences USA*, **110**, E2151–E2152.

Bar-Massada, A. & Wood, E. (2014) The richness–heterogeneity relationship differs between heterogeneity measures within and among habitats. *Ecography*, **37**, 528–535.

Boucher-Lalonde, V., Theriault, F.L. & Currie, D.J. (2014) Can climate explain interannual local extinctions among bird species? *Journal of Biogeography*, **41**, 443–451.

Boulinier, T., Nichols, J.D., Hines, J.E., Sauer, J.R., Flather, C.H. & Pollock, K.H. (1998) Higher temporal variability of forest breeding bird communities in fragmented landscapes. *Proceedings of the National Academy of Science USA*, **95**, 7497–7501.

Burnham, K.P. & Anderson, D.R. (2002) *Model selection and multimodel inference: a practical information-theoretical approach*, 2nd edn. Springer, New York.

Burnham, K.P. & Overton, W.S. (1979) Robust estimation of population-size when capture probabilities vary among animals. *Ecology*, **60**, 927–936.

Cam, E., Nichols, J.D., Sauer, J.R., Hines, J.E. & Flather, C.H. (2000) Relative species richness and community completeness: birds and urbanization in the Mid-Atlantic states. *Ecological Applications*, **10**, 1196–1210.

Carnicer, J., Brotons, L., Herrando, S. & Sol, D. (2013) Improved empirical tests of area-heterogeneity tradeoffs. *Proceedings of the National Academy of Sciences USA*, **110**, E2858–E2860.

Davies, R.G., Orme, C.D.L., Storch, D., Olson, V.A., Thomas, G.H., Ross, S.G., Ding, T.-S., Rasmussen, P.C., Bennett, P.M., Owens, I.P.F., Blackburn, T.M. & Gaston, K.J. (2007) Topography, energy and the global distribution of bird species richness. *Proceedings of the Royal Society B: Biological Sciences*, **274**, 1189–1197.

de Souza Júnior, M.B., Ferreira, F.F. & de Oliveira, V.M. (2014) Effects of the spatial heterogeneity on the diversity of ecosystems with resource competition. *Physica A – Statistical Mechanics and its Applications*, **393**, 312–319.

Donovan, T.M. & Flather, C.H. (2002) Relationships among North American songbird trends, habitat fragmentation, and landscape occupancy. *Ecological Applications*, **12**, 364–374.

Dunning, J.B. (2007) *CRC handbook of avian body masses*, 2nd edn. CRC Press, Boca Raton, FL.

Finch, O.D., Blick, T. & Schuldt, A. (2008) Macroecological patterns of spider species richness across Europe. *Biodiversity and Conservation*, **17**, 2849–2868.

Fløjgaard, C., Normand, S., Skov, F. & Svenning, J.-C. (2011) Deconstructing the mammal species richness pattern in Europe – towards an understanding of the relative importance of climate, biogeographic history, habitat heterogeneity and humans. *Global Ecology and Biogeography*, **20**, 218–230.

Hines, J.E., Boulinier, T., Nichols, J.D., Sauer, J.R. & Pollock, K.H. (1999) COMDYN: software to study the dynamics of animal communities using a capture-recapture approach. *Bird Study*, **46**, S209–S217.

Hortal, J., Triantis, K.A., Meiri, S., Thebault, E. & Sfenthourakis, S. (2009) Island species richness increases with habitat diversity. *The American Naturalist*, **174**, E205–E217.

Hortal, J., Carrascal, L.M., Triantis, K.A., Thebaulte, E., Meiri, S. & Sfenthourakis, S. (2013) Species richness can decrease with altitude but not with habitat diversity. *Proceedings of the National Academy of Sciences USA*, **110**, E2149–E2150.

Hurlbert, A.H. & White, E.P. (2005) Disparity between range map- and survey-based analyses of species richness: patterns, processes and implications. *Ecology Letters*, **8**, 319–327.

Hutchinson, G.E. (1957) Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology*, **22**, 415–427.

Jimenez, I., Distler, T. & Jorgensen, P.M. (2009) Estimated plant richness pattern across northwest South America provides similar support for the species-energy and spatial heterogeneity hypotheses. *Ecography*, **32**, 433–448.

Kadmon, R. & Allouche, O. (2007) Integrating the effects of area, isolation, and habitat heterogeneity on species diversity: a unification of island biogeography and niche theory. *The American Naturalist*, **170**, 443–454.

Kalyuzhny, M., Seri, E., Chocron, R., Flather, C.H., Kadmon, R. & Shnerb, N. (2014) Niche versus neutrality: a dynamical analysis. *The American Naturalist*, **184**, 439–446.

Keitt, T.H. & Stanley, H.E. (1998) Dynamics of North American breeding bird populations. *Nature*, **393**, 257–260.

Laanisto, L., Tamme, R., Hiiesalu, I., Szava-Kovats, R., Gazol, A. & Paertel, M. (2013) Microfragmentation concept explains non-positive environmental heterogeneity–diversity relationships. *Oecologia*, **171**, 217–226.

Legendre, P. & Legendre, F.J. (2012) *Numerical ecology*. Elsevier, Amsterdam.

Lennon, J.J., Greenwood, J.J.D. & Turner, J.R.G. (2000) Bird diversity and environmental gradients in Britain: a test of the species–energy hypothesis. *Journal of Animal Ecology*, **69**, 581–598.

MacArthur, R.H. (1965) Patterns of species diversity. *Biological Reviews*, **40**, 510–533.

MacArthur, R.H. & Wilson, E.O. (1967) *The theory of island biogeography*. Princeton University Press, Princeton, NJ.

Mandel, M. (2013) Simulation-based confidence intervals for functions with complicated derivatives. *American Statistician*, **67**, 76–81.

Mayer, A.L. & Cameron, G.N. (2003) Landscape characteristics, spatial extent, and breeding bird diversity in Ohio, USA. *Diversity and Distributions*, **9**, 297–311.

Moreno-Rueda, G. & Pizarro, M. (2009) Relative influence of habitat heterogeneity, climate, human disturbance, and spatial structure on vertebrate species richness in Spain. *Ecological Research*, **24**, 335–344.

Nagelkerke, N.J.D. (1991) A note on a general definition of the coefficient of determination. *Biometrika*, **78**, 691–692.

Nichols, J.D., Boulinier, T., Hines, J.E., Pollock, K.H. & Sauer, J.R. (1998) Estimating rates of local species extinction, colonization, and turnover in animal communities. *Ecological Applications*, **8**, 1213–1225.

Olden, J.D. & Jackson, D.A. (2000) Torturing data for the sake of generality: how valid are our regression models? *Ecoscience*, **7**, 501–510.

R Development Core Team (2013) *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Available at: http://www.R-project.org.

Rahbek, C. & Graves, G.R. (2001) Multiscale assessment of patterns of avian species richness. *Proceedings of the National Academy of Sciences USA*, **98**, 4534–4539.

Recher, H.F. (1969) Bird diversity and habitat diversity in Australia and North America. *The American Naturalist*, **103**, 75–80.

Ricklefs, R.E. & Lovette, I.J. (1999) The roles of island area *per se* and habitat diversity in the species-area relationships of four Lesser Antillean faunal groups. *Journal of Animal Ecology*, **68**, 1142–1160.

Rittenhouse, C.D., Pidgeon, A.M., Albright, T.P., Culbert, P.D., Clayton, M.K., Flather, C.H., Masek, J.G. & Radeloff, V.C. (2012) Land-cover change and avian diversity in the conterminous United States. *Conservation Biology*, **26**, 821–829.

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

Sauer, J.R., Fallon, J.E. & Johnson, R. (2003) Use of North American Breeding Bird Survey data to estimate population change for bird conservation regions. *Journal of Wildlife Management*, **67**, 372–389.

Sauer, J.R., Hines, J.E., Fallon, J.E., Pardieck, K.L., Ziolkowski, D.J., Jr & Link, W.A. (2014) *The North American Breeding Bird Survey, results and analysis 1966–2012*. Version 02.19.2014. USGS Patuxent Wildlife Research Center, Laurel, MD.

Seiferling, I., Proulx, R. & Wirth, C. (2014) Disentangling the environmental heterogeneity–species diversity relationship along a gradient of human footprint. *Ecology*, **95**, 2084–2095.

Stegen, J.C., Freestone, A.L., Crist, T.O., Anderson, M.J., Chase, J.M., Comita, L.S., Cornell, H.V., Davies, K.F., Harrison, S.P., Hurlbert, A.H., Inouye, B.D., Kraft, N.J.B., Myers, J.A., Sanders, N.J., Swenson, N.G. & Vellend, M. (2013) Stochastic and deterministic drivers of spatial and temporal turnover in breeding bird communities. *Global Ecology and Biogeography*, **22**, 202–212.

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

Sutherland, G.D., Harestad, A.S., Price, K. & Lertzman, K.P. (2000) Scaling of natal dispersal distances in terrestrial birds and mammals. *Conservation Ecology*, **4**, 1–56.

Tamme, R., Hiiesalu, I., Laanisto, L., Szava-Kovats, R. & Partel, M. (2010) Environmental heterogeneity, species diversity and co-existence at different spatial scales. *Journal of Vegetation Science*, **21**, 796–801.

Tittler, R., Villard, M.-A. & Fahrig, L. (2009) How far do songbirds disperse? *Ecography*, **32**, 1051–1061.

van Rensburg, B.J., Chown, S.L. & Gaston, K.J. (2002) Species richness, environmental correlates, and spatial scale: a test using South African birds. *The American Naturalist*, **159**, 566–577.

Veech, J.A. & Crist, T.O. (2007) Habitat and climate heterogeneity maintain beta-diversity of birds among landscapes within ecoregions. *Global Ecology and Biogeography*, **16**, 650–656.

White, E.P. (2004) Two-phase species-time relationships in North American land birds. *Ecology Letters*, **7**, 329–336.

SUPPORTING INFORMATION

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

Appendix S1 Selection of North American Breeding Bird Survey routes for the analysis.

Appendix S2 Calculation of independent variables used in the analysis.

Figure S1 Relationships between species richness and variables used as predictors in the models.

Table S1 Pearson correlation coefficients between the variables used as predictors of species richness in the models.

Table S2 A summary of 36 simultaneous autoregressive regression models testing the effects of elevational range and land-cover richness on species richness.

Table S3 Full results of 36 simultaneous autoregressive regression models testing the effects of elevational range and land-cover richness on species richness.

Table S4 Characteristics of previous analyses of North American Breeding Bird Survey data testing the effects of elevational range and/or land-cover richness on species richness at the route scale.

BIOSKETCHES

Rachel Chocron conducted this work as a MSc student at the Department of Ecology, Evolution and Behavior, The Hebrew University of Jerusalem.

Curtis H. Flather is a research ecologist with the US Forest Service, Rocky Mountain Research Station. His research is focused on understanding the response of biodiversity to changing climate, land use, natural disturbance and land management activities to support resource planning activities within the agency.

Ronen Kadmon is a plant ecologist at the Hebrew University of Jerusalem. His research focuses on the mechanisms that determine the dynamics and structure of ecological communities, with special emphasis on the mechanisms that determine species diversity and its variation over space and time.

Editor: Vincent Devictor