Topography, energy and the global distribution of bird species richness

Richard G. Davies¹,* , C. David L. Orme², David Storch³,⁴, Valerie A. Olson⁵, Gavin H. Thomas⁶,⁷, Simon G. Ross¹, Tzung-Su Ding⁷, Pamela C. Rasmussen⁸, Peter M. Bennett⁵, Ian P. F. Owens²,⁹, Tim M. Blackburn⁶ and Kevin J. Gaston¹

¹Biodiversity and Macroecology Group, Department of Animal and Plant Sciences, University of Sheffield, Sheffield S10 2TN, UK
²Division of Biology, and ³NERC Centre for Population Biology, Imperial College London, Silwood Park, Ascot, Berkshire SL5 7PY, UK
⁴Center for Theoretical Study, Charles University, Jilská 1, 110 00 Praha 1, Czech Republic
⁵Department of Ecology, Faculty of Science, Vmičná 7, 128 44 Praha 2, Czech Republic
⁶Institute of Zoology, Zoological Society of London, Regent’s Park, London NW1 4RZ, UK
⁷School of Biosciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK
⁸School of Forestry and Resource Conservation, National Taiwan University, 1, Sec 4, Roosevelt Road, Taipei 106, Taiwan
⁹Michigan State University Museum and Department of Zoology, West Circle Drive, East Lansing, MI 48824-1045, USA

A major goal of ecology is to determine the causes of the latitudinal gradient in global distribution of species richness. Current evidence points to either energy availability or habitat heterogeneity as the most likely environmental drivers in terrestrial systems, but their relative importance is controversial in the absence of analyses of global (rather than continental or regional) extent. Here we use data on the global distribution of extant continental and continental island bird species to test the explanatory power of energy availability and habitat heterogeneity while simultaneously addressing issues of spatial resolution, spatial autocorrelation, geometric constraints upon species' range dynamics, and the impact of human populations and historical glacial ice-cover. At the finest resolution (1°), topographical variability and temperature are identified as the most important global predictors of avian species richness in multi-predictor models. Topographical variability is most important in single-predictor models, followed by productive energy. Adjusting for null expectations based on geometric constraints on species richness improves overall model fit but has negligible impact on tests of environmental predictors. Conclusions concerning the relative importance of environmental predictors of species richness cannot be extrapolated from one biogeographic realm to others or the globe. Rather a global perspective confirms the primary importance of mountain ranges in high-energy areas.

Keywords: geometric constraints; global biodiversity; habitat heterogeneity; species richness; species-energy theory; topography

1. INTRODUCTION

Understanding the determinants of the latitudinal diversity gradient, or why the species richness of the majority of taxa declines from the equator to the poles, constitutes one of the key challenges in ecology (Brown 1981; Rosenzweig 1995; Gaston 2000). Among the 30 or more environmental hypotheses proposed, two in particular emerge as contenders showing widespread empirical support. Each of these has, in turn, been regarded as having two alternative variants. The first pair falls within species-energy theory, in representing two alternative forms of available energy, ambient and productive. Ambient energy, usually represented by temperature or allied measures, is thought to influence species ranges via physiological constraints, and the effect of metabolic rates on rates of population growth and turnover (Turner et al. 1987, 1988; Currie 1991; Allen et al. 2002). Its role is also argued to extend to the influence of solar energy and temperature upon speciation rates, via the causal link between UV radiation or metabolic rates, respectively and mutation rates (Rohde 1992; Allen et al. 2002, 2006). The productive energy hypothesis predicts that the species richness of consumers is determined by the energy flowing through food webs, starting from plant productivity and biomass (Wright 1983, Wright et al. 1993), and ultimately depending on the availability of water, heat and light (Waide et al. 1999). The second pair is considered as variants of the habitat heterogeneity hypothesis, where heterogeneity is quantified either as the number of habitat types (Rosenzweig 1995; Kerr et al. 2001) or as...
the topographical variability (range in elevation) present within an area (Richerson & Lum 1980; Kerr & Packer 1997; Rahbek & Graves 2001). The mechanisms underlying each of these hypotheses are not the central concern of this study (for a complete exposition see Turner & Hawkins 2004; Evans et al. 2005). Rather, we focus on testing the relative importance of each of the representative environmental gradients in driving large-scale species richness patterns, as there is yet no consensus on this issue. In doing so, we also address four issues that are external to the proposed environmental mechanisms, but are argued to be central influences on the outcome of analytical tests.

First, it is now well established that the spatial scales at which data are sampled may have a critical influence on the outcome of investigations into the relative importance of environmental gradients (Rahbek & Graves 2001; Whittaker et al. 2001; Rahbek 2005). Past limitations on the availability and the quality of distributional and environmental data have meant that investigations have been largely restricted to individual continents or biogeographic realms. At these spatial extents, studies show the measures of energy availability to be the strongest descriptors of overall diversity, with a changeover in importance from ambient to productive energy in the transition from high to low latitudes (Hawkins et al. 2003a). Spatial patterns of both the species richness and the environmental gradients also vary with the grain size (grid cell area) of the sampling unit. Previous studies have observed that the significance of range in elevation increases with increase in grain size, while that of energy predictors tends to be more scale invariant (Rahbek & Graves 2001). The question of which are the most general global predictors of the patterns of species richness across a range of sampling resolutions has never been addressed.

Second, and more contentiously, intense discussion revolves around the contribution of non-biological processes, namely that of geometric constraints of regional domain boundaries upon species range dynamics. The latter has been argued to result in a mid-domain effect (MDE) or peak in species richness near the centre of continental land masses (Colwell & Hurlt 1994; Colwell & Lees 2000), although the measurable concordance between the predictions of such an effect and observed species richness distribution is shown in general to be weak (Zapata et al. 2003). Nevertheless, the MDE still poses an alternative null expectation to the even pattern of species richness that is implicitly the null expectation of standard regression modelling approaches to testing of the relative importance of environmental predictors. One way of accounting for mid-domain null expectations is to fit predicted species richness as a covariate within the model-building process (Jetz & Rahbek 2002). However, appropriate global mid-domain models which address the added complexity of species ranges that span multiple biogeographic domains, have only recently been developed (Storch et al. 2006).

Third, spatial autocorrelation is another key issue which has only relatively recently come to the fore in the analyses of spatially distributed ecological data. Estimates of large-scale species richness distribution, where these show similarity that is a function of the proximity of sampling locations, are typically spatially autocorrelated (hence non-independent samples) due to one or a combination of: intrinsic and/or infectious processes such as population growth and dispersal (Legendre et al. 2002); areas where false species presence or absence are recorded due to errors in distributional data (Hurlbert & White 2005); or where environmental gradients that drive species richness patterns themselves show spatial autocorrelation arising from the underlying processes that generate environment variation (Legendre et al. 2002). All of these influences may be routinely encountered in macroecological datasets. In particular, the last of them is an all but ubiquitous property of the landscape-scale habitat patches and gradients from which macroecological samples are drawn, and is equivalent to pseudoreplication in which environmental treatments are spatially segregated across an experimental grid (Hurlbert 1984; Legendre 1993). The consequence of spatial autocorrelation for comparisons of environmental hypotheses is the inflation of type I error rates and biasing of environmental parameter estimates (and their perceived relative fit), when using regression methods that assume independent model errors (Clifford et al. 1989; Cressie 1991). However, appropriate spatial modelling methods that take account of spatial autocorrelation have yet to be widely adopted in macroecological analyses.

Fourth, research into environmental determinants of overall species richness has tended to emphasize the testing of classical ecological predictors, while the other factors that may have influenced species distributional ranges, and consequent patterns of species richness observed today, have been largely omitted. Both contemporary human population density and extent of agricultural activity in an area are known to be important predictors of the numbers of threatened species (McKinney 2001; McKee et al. 2003; Scharlemann et al. 2005) using equal-area grids at resolutions comparable to 1°, 2° and 4° of latitude × longitude. We used habitat diversity (number of land-cover types) and topographic variability (elevation range) as measures of habitat heterogeneity, and mean annual temperature and the normalized difference vegetation index (NDVI) as, respectively, measures of ambient and productive energy availability. Model selection methods (Burnham & Anderson 2001) are used to determine best-fit models from all possible combinations of our four key predictors, while controlling for the role of human population density, agricultural land area and extent of glacial ice cover over the last 21 000 years. We go on to compare our highest resolution (1') models with further model specifications for that
2. MATERIAL AND METHODS

(a) Mapping

The analyses presented here are based on a database of distribution maps for 9626 extant, recognized bird species following a standard avian taxonomy (Sibley & Monroe 1990). Using a variety of published sources, breeding ranges were mapped as vectors or ‘polygons’ (for details see Orme et al. 2005, 2006). These maps were then converted to equal-area grids using a Behrmann projection at cell resolutions (at the standard parallels of 30° of N and S) of 96 486.2, 192 972.5 and 385 945.1 m, approximately equivalent to 1°, 2° and 4° geographical coordinate system grids, respectively. The global grids therefore contained 360 × 152, 180 × 76 and 90 × 38 cells, respectively, omitting the partial cells at latitudes higher than 87.13°. Species were scored as present in a grid cell if any of the available sources suggested that the breeding range fell within the cell boundaries. The overall species richness was derived by summing up all species present within each cell. Biogeographic realms were delimited using the World Wildlife Fund ecoregions map (Olson et al. 2001; Nearctic, Paleartic, Neotropical, Afrotopical, Indo-Malayan and Australasian). The final dataset used for analyses omitted grid cells falling within Oceania or Antarctica, since environmental variable data were lacking for these realms. The remaining true oceanic islands, defined as any land area located further than 200 km distance from the edge of continental shelf, were also omitted. Finally, so as to avoid bias in terms of the contribution of continental land area to the regression models, grid cells with less than 50% landcover were omitted from the final dataset.

(b) Geometric constraints models

We simulated the random dynamics of species ranges at the highest data resolution (1°) of our analyses, using the generalized spreading dye model (Storch et al. 2006) and drawing upon the complete global extent and species complement of the avian range database (Orme et al. 2005). The model assumes that species ranges are contiguous and spread from the point of origin to available neighbouring grid cells until the final number of occupied cells, hence range size, is reached (i.e. the observed distribution of range sizes in terms of number of occupied grid cells is kept). The first cell was chosen randomly, and in the subsequent steps, the species could spread to any available unoccupied cell adjacent to any already occupied, with the probability of being selected

\[ P = 1/N_{adj} \]

where \( N_{adj} \) is the number of empty adjacent cells at each respective step (i.e. \( N_{adj} \leq 8 \) for a single occupied cell but is usually higher where more than one cell is occupied).

Classical approaches to the MDE emerging from species range dynamics assume that the latter are strictly limited by domain boundaries, however defined. However, in our model of global range dynamics, it is possible for species (especially those with large ranges) to spread into areas that are bounded and smaller than the given species’ range. In such cases, when species filled the domain (island or continent) in question, it could skip to a different domain, into the cell which had the shortest distance to the last cell to be occupied within the previous domain. This simulated the rare events of long-distance dispersal, necessary for colonizing new and distant continents or islands. After the colonization event, the range dynamics continued within the new domain according to the rules described above; long-distance jumps therefore occurred only when necessary (and were in fact quite rare). Had a rule been imposed such that species’ ranges were restricted to spread only within the biogeographic realm(s) in which they were observed to occur (e.g. Storch et al. 2006), a large portion of the environmental component of the latitudinal diversity gradient would effectively be in-built (since the total species complements of each biogeographic realm, hence inter-realm differences in species numbers, would be preserved). This would have compromised our intended use of the predicted outcome of a pure mid-domain model as a covariate in environmental model building and selection. The required covariate was taken as the vector of mean values per grid cell of one hundred simulation runs of the model.

(c) Environmental data

As an alternative estimate of productive energy to NDVI, actual evapotranspiration (AET) was fitted in its place in an alternative series of global models. However, since its relative importance compared with other predictors was not qualitatively different from NDVI, these results are not presented.

Available raw data for each of the candidate environmental variables were re-projected and re-sampled to the same equal-area grid as the species richness data (for further details concerning their sources, raw resolutions and treatment, see Material and Methods in electronic supplementary material). Human population density, NDVI, AET, agricultural land area and elevation range variables were all log-transformed.

(d) Statistical analyses

We included quadratic as well as linear terms for predictors in our models in order to allow for nonlinear relationships. Where necessary, response variables were normalized by square-root- (globally at all three grid cell resolutions) and log-transformation (Indo-Malaya and Australasia at a resolution equivalent to 1°).

We applied normal error generalized least squares (GLS) modelling methods (SAS; Littell et al. 1996), fitting exponential spatial covariance structures (which were the best-fit choice among spatial covariance options). Longitudinal and latitudinal cell centroid values were used as spatial variables and all models were implemented in SAS v. 9.1.3. Spatial GLS models took account of the differences among major biogeographic realms, in the maximum geographic geographical distance or range parameter (\( \rho \)), measured in degrees, over which spatial autocorrelation in equivalent ordinary least squares (OLS) model residuals was observed to occur. This involved estimating \( \rho \) from the semi-variogram of residuals of OLS models that included the relevant combination of predictors, separately for each realm. All six estimates of \( \rho \) were then entered as spatial covariance parameters in the model, with spatial autocorrelation assumed for observations within the same realm.

We tested for collinearity among predictors through investigations of tolerance levels (Quinn & Keough 2002). These were sufficiently high in all cases at a resolution equivalent to 2° and were higher still for 1° equivalent data. However, at a resolution equivalent to 4°, low tolerances and hence some redundancy, were observed in the case of NDVI for Australasia and temperature for the Nearctic. Nevertheless, this was found to be the result of collinearity with one or other human impact factors, and comparison between full models that included and omitted human factors showed that this had negligible impact on the results and conclusions drawn.

For our highest resolution global analyses, we first ran two sets of normal errors spatial models of all combinations
of environmental predictors, in order to apply model selection methods that use Akaike’s information criterion (AIC) as a measure of overall model fit (Burnham & Anderson 2001). The two sets were a means of restricting combinations to a tractable number of models; the first set fitted all human impact predictors simultaneously in addition to the given combination of environmental predictors, while the former were excluded in the second set. Within each set, the model with the lowest AIC was selected as the best-fit model. By calculating the Akaike weight ($w$) for each model, we determined the candidate sets of models that had a combined probability of 0.95 or above that they included the best-fit model. None of the models among candidate sets contradicted the findings and conclusions drawn from the use of best-fitting models, hence only the latter results are reported. The same model selection procedures were used to determine the best-fitting non-spatial GLMs (for these results see tables 2 and 3 in electronic supplementary material); however, in all cases these were associated with higher values for $-2 \log$ likelihood of the restricted likelihood than the equivalent GLS models, indicating that the latter were a consistently more accurate description of variability in species richness (Littell et al. 1996). Estimates of variance explained cannot be derived from spatial GLS models, hence adjusted $R^2$ only for best-fit non-spatial GLM models are reported (see tables 2 and 3 in electronic supplementary material).

In order to assess the influence of geographical extent upon the relative importance of environmental predictors for our highest resolution data, we used identical model selection procedures, and both spatial and non-spatial regression methods, to determine the best-fit models for each of six major biogeographic realms (Olson et al. 2001), while simultaneously accounting for human impacts and extent of most recent glacial ice cover.

Finally, we assessed the influence of the MDE upon the relative fit of our environmental predictors, hence upon model selection outcomes, by constructing a further parallel set of models (and model selection runs), that were identical in all respects to the above model sets (both globally and within biogeographic realms), but with the additional fitting of species richness predicted from range dynamic processes based on geometric constraints (see mid-domain predictions, above) as a covariate.

3. RESULTS AND DISCUSSION

(a) Global avian species richness, sampling resolution and spatial models

Global maps of avian species richness data used in our analyses showed a consistent pattern across all three sampling resolutions (figure 1), namely higher species richness within the tropics, with peaks coinciding with major mountain chains, most notably along the Andes and the southern slopes of the Himalayas and to a lesser extent
Table 1. Best-fit global multi-predictor spatial GLS models of species richness. (All variables refer to models for which human impact predictors were additionally fitted; Environmental variables only refer to the global model for which human impact predictors were not fitted. 1°, 2° and 4° refer to equal-area spatial resolutions approximately equivalent to 1°×1°, 2°×2° and 4°×4° longitude×latitude, respectively. At 1° resolution, MDE-fit refers to the best-fit models from sets that additionally fit species richness predictions based on geometric constraints (i.e. mid-domain effect, MDE). + and − indicate significant positive and negative slopes, respectively, while * indicates significance of a categorical predictor. ++/+−/−−/−**; P<0.001; ++/+−/−***; 0.01≤P<0.01;+/−/−/−, 0.01≤P<0.05. The most significant P-values and slopes for each model are highlighted in bold. The best-fit models presented are those from the set of all combinations of the four main environmental predictors that have the lowest value for Akaike’s information criterion (AIC). Elevation range, NDVI, population density and agricultural land area were all log-transformed.)

<table>
<thead>
<tr>
<th>predictor</th>
<th>all variables</th>
<th>environmental variables only</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1°</td>
<td>1° MDE fit</td>
</tr>
<tr>
<td>elevation range</td>
<td>F_{1,13851}</td>
<td>F_{1,13849}</td>
</tr>
<tr>
<td>elevation range^2</td>
<td>107.99−−−</td>
<td>106.99−−−</td>
</tr>
<tr>
<td>habitat diversity</td>
<td>214.86++</td>
<td>214.85+++</td>
</tr>
<tr>
<td>habitat diversity^2</td>
<td>85.04++</td>
<td>86.74+++</td>
</tr>
<tr>
<td>NDVI^2</td>
<td>68.10++</td>
<td>64.30+++</td>
</tr>
<tr>
<td>temperature^2</td>
<td>4.08−</td>
<td>2.55−</td>
</tr>
<tr>
<td>temperature</td>
<td>143.56−−−</td>
<td>155.13−−−</td>
</tr>
<tr>
<td>population density</td>
<td>4.79−</td>
<td>5.04−</td>
</tr>
<tr>
<td>population density^2</td>
<td>27.71++</td>
<td>29.07+++</td>
</tr>
<tr>
<td>agricultural area^2</td>
<td>1.29−</td>
<td>2.12−</td>
</tr>
<tr>
<td>agricultural area</td>
<td>1.66−</td>
<td>2.12−</td>
</tr>
<tr>
<td>ice covered/not</td>
<td>1.84−</td>
<td>0.89−</td>
</tr>
<tr>
<td>MDE^2</td>
<td>10.895++</td>
<td>10.895++</td>
</tr>
<tr>
<td>MDE</td>
<td>71.52−−−</td>
<td>71.52−−−</td>
</tr>
<tr>
<td>AIC</td>
<td>24 521.0</td>
<td>24 392.9</td>
</tr>
</tbody>
</table>

the African Rift Valley. The best-fit multi-predictor global model accounting for glacial history and human impacts showed elevation range to be the strongest predictor of avian species richness, closely followed by temperature and then habitat diversity and productive energy (table 1). The omission of human impact variables marginally favours temperature over elevation range as the primary predictor at a resolution equivalent to 1° (table 1). The primary importance of elevation range was also supported by the relative fit of high-resolution global models in which each of the four environmental predictors was fitted separately (see table 1 in electronic supplementary material). However, in contrast to the multi-predictor models, the fit of productive energy was better than either temperature or habitat diversity.

Our high-resolution (1° equivalent) models showed that global relationships with species richness are hump-shaped (NDVI and temperature), positive-linear (habitat diversity) or U-shaped (elevation range; table 1). With elevation range, however, the relationship starts with a shallow negative phase at low to intermediate range, followed by a steeper positive phase from intermediate to high. This may reflect the transition from flatter areas (lowland plains) with moderately high levels of species richness, to areas of intermediate elevation range (including montane summits or plateaux) with lower average richness, to areas of maximal elevation range (mid-montane slopes) and species richness that encompass the greatest habitat turnover in response to elevational climate gradients. Such a relationship strongly implicates the contribution of β-diversity to maximal species richness at macroecological scales. Allied to this is the key role of latitudinal temperature gradients in determining upper altitudinal limits of habitable zones and the altitudinal breadth of seasonal temperature fluctuations, both of which influence species numbers encountered up montane slopes (Janzen 1967). The mechanisms underlying topographic and energy gradients in species diversity are complex, and the above relationships do not detract from historical hypotheses for the role of tropical areas as centres of endemism and richness resulting from high net diversification rates (Fjeldså 1994; Jetz et al. 2004).

Models based on a resolution equivalent to 2° were very similar to those obtained at the highest resolution (1° equivalent) but with productive energy entering as the third predictor variable, ahead of habitat diversity. However, at a resolution equivalent to 4°, productive energy became the primary variable followed by elevation range, habitat diversity and temperature (table 1). Hence with increase in resolution in this study, species richness became progressively uncoupled from productivity and more allied to topography and temperature.

(b) Consequences of restricting geographical extent of analyses

High-resolution data better preserve the fine-scale variation in species richness to which narrow-ranging species in particular contribute. Our highest resolution global findings concurred with the results from some previous
studies of individual continents or biogeographic realms (Rahbek & Graves 2000, 2001). However, they contrasted with a body of evidence that productive energy (net primary productivity or NDVI) is the strongest predictor of species richness at lower latitudes, with ambient energy (temperature) being the strongest predictor at higher latitudes where water is less of a limiting resource (Hawkins et al. 2003b; Turner & Hawkins 2004). There are two probable reasons for these differences. First, no previous study has considered global richness patterns, instead analysing data from one or more continents or realms separately. Such studies may downplay the importance of certain predictors if they only sample part of the overall environmental gradient in those variables. When restricting our analyses to individual biogeographic realms (defined as in Olson et al. 2001), the strongest predictors in best-fit spatial GLS models were NDVI (Afrotropics, Australasia), temperature (Nearctic, Neotropics) and elevation range (Indo-Malaya, Palaearctic; table 2). Thus, the results of models of species richness for individual realms cannot be assumed to apply either to other realms or globally.

### Spatial autocorrelation and spatial scale

The differences in findings between the analyses reported here and previous studies reflected the consequences of using models that control for spatial autocorrelation. It is becoming increasingly clear that wide-ranging species contribute disproportionately to simple models of species richness distribution, while narrow-ranging species, although more numerous, contribute far less (Jetz & Rahbek 2002; Lennon et al. 2004). This is partly because widespread species contribute to species richness in many more grid cells than narrow-ranging species. It has only recently been demonstrated that the multiple with which this disparity in representation operates increases with

---

### Table 2. Best-fit spatial multi-predictor GLS models of species richness for each biogeographic realm. (Best-fit model results include those from sets that both exclude and include the fit of mid-domain effect (MDE) predictions. See table 1 for abbreviations and methodology.)

<table>
<thead>
<tr>
<th>predictor</th>
<th>1° MDE fit</th>
<th>1° MDE fit</th>
<th>1° MDE fit</th>
<th>1° MDE fit</th>
<th>1° MDE fit</th>
<th>1° MDE fit</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Australas</td>
<td>Afrotropics</td>
<td>Indo-Malaya</td>
<td>Nearctic</td>
<td>Neotropics</td>
<td>Palaearctic</td>
</tr>
<tr>
<td>elevation range</td>
<td>30.93</td>
<td>30.75</td>
<td>93.14</td>
<td>97.40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>elevation range²</td>
<td>44.47+</td>
<td>44.58+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>habitat diversity</td>
<td>27.32+</td>
<td>31.27+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NDVI</td>
<td>33.61+</td>
<td>27.69+</td>
<td>91.17</td>
<td>95.27</td>
<td>12.29</td>
<td>12.99</td>
</tr>
<tr>
<td>NDVI²</td>
<td>23.39-</td>
<td>15.60-</td>
<td>38.63</td>
<td>37.68-</td>
<td>6.25-</td>
<td>6.46-</td>
</tr>
<tr>
<td>temperature</td>
<td>14.38+</td>
<td>3.05</td>
<td>3.56</td>
<td>33.77+</td>
<td>33.76+</td>
<td></td>
</tr>
<tr>
<td>population density</td>
<td>1.38</td>
<td>0.51</td>
<td>0.30</td>
<td>0.48</td>
<td>16.73+</td>
<td>19.22+</td>
</tr>
<tr>
<td>agricultural area</td>
<td>0.49</td>
<td>0.46</td>
<td>8.75</td>
<td>10.07-</td>
<td>0.16</td>
<td>0.19</td>
</tr>
<tr>
<td>ice covered/not</td>
<td>4.58</td>
<td>2.45</td>
<td></td>
<td></td>
<td>0.55</td>
<td>0.60</td>
</tr>
<tr>
<td>MDE</td>
<td>32.89+</td>
<td>38.58+</td>
<td>25.60-</td>
<td>3.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MDE²</td>
<td>25.61-</td>
<td></td>
<td>25.60-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AIC</td>
<td>-2382.7</td>
<td>-2382.6</td>
<td>20 380.7</td>
<td>20 339.0</td>
<td>-3796.1</td>
<td>-3778.6</td>
</tr>
<tr>
<td></td>
<td>20 411.8</td>
<td>37 430.0</td>
<td>43 740.0</td>
<td>43 701.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>20 428.4</td>
<td>37 455.1</td>
<td>43 712.4</td>
<td>43 701.2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

(c) Spatial autocorrelation and spatial scale

The differences in findings between the analyses reported here and previous studies reflected the consequences of using models that control for spatial autocorrelation. It is becoming increasingly clear that wide-ranging species contribute disproportionately to simple models of species richness distribution, while narrow-ranging species, although more numerous, contribute far less (Jetz & Rahbek 2002; Lennon et al. 2004). This is partly because widespread species contribute to species richness in many more grid cells than narrow-ranging species. It has only recently been demonstrated that the multiple with which this disparity in representation operates increases with
data resolution (Rahbek 2005). Clearly, this means that the accuracy with which the relative representation (in terms of numbers of grid cells) of narrow-ranging species to wide-ranging species increases as decrease in grain size approaches an area equivalent to the range extents of the most range-restricted species. However, the corollary that the disproportionate contribution of widespread (relative to narrow-ranging) species to spatial autocorrelation in overall species richness distribution increases with increasing resolution has so far been overlooked. This is especially relevant, where richness data are based on range maps of extent-of-occurrence which may be prone to inclusion of areas of false species presence. These influences are not trivial since it has further been demonstrated that energy gradients are strong predictors of wide-ranging species richness, but are poor predictors of the distribution of narrow-ranging species (Jetz & Rahbek 2002; Bonn et al. 2004; Ruggiero & Kitzberger 2004). Hence, as representation of wide-ranging species increases in tandem with resolution, environmental models of overall species richness are increasingly prone to find the greater importance of energy relative to other predictors (Jetz & Rahbek 2002) if they fail to account for spatial autocorrelation. This was clearly demonstrated by the steep increase in significance of NDVI with increase in resolution in best-fit global non-spatial models, outstripping that of any other predictor (see table 2 in electronic supplementary material).

The spatial model results provided a further important contrast with non-spatial model findings both from previous studies (Rahbek & Graves 2001) and the present one (see table 2 in electronic supplementary material). In non-spatial models, the strength of elevational range as a predictor relative to productive energy increased as data resolution decreased. This trend was opposite to that found in spatial models and requires explanation. An important effect of increasing grid cell size is that areas of high richness coinciding with high elevation range apparently increase in proportional extent relative to other areas (for example, compare the latitudinal width of high-richness areas coinciding with the Himalayas from resolutions equivalent to 1° to 4°, figure 1). This proportionally greater representation of high-richness/high-elevational range cells increases the strength of non-spatially modelled relationships between species richness and elevation range. However, since montane slopes tend to be narrow or otherwise limited in extent, increase in size of grid cells straddling these (and adjacent) areas are more susceptible to averaging effects upon signal-to-noise ratios. Hence, spatial models, in accounting for spatially autocorrelated errors, showed the real signal of elevation range to be weaker than that of productive energy at 4° equivalent resolution (table 1). Overall, these findings confirm that models of species richness that ignore the effects of autocorrelation cannot be assumed to approximate the results that would follow from doing so.

(d) Adjusting for geometric constraints

Unsurprisingly, at a resolution equivalent to 1°, predictions of species richness distribution based on geometric constraints upon global avian range dynamics (see figure 1 in electronic supplementary material and Storch et al. 2006) contrasted strongly with the observed patterns of species richness (figure 1a). Fitting these range dynamic predictions as a covariate in our high-resolution global spatial GLS models left our conclusions unchanged, except that in the model including human impact variables, the significance of temperature was marginally enhanced over that of elevation range (table 1). In the same model, the significance levels for the mid-domain terms were themselves greater than those for either NDVI or habitat diversity. Moreover, global models fitting mid-domain terms showed an increase in overall model fit (lower AIC) when compared with those omitting them. For our within-realm spatial GLS models, fitting of the mid-domain covariate also made no difference to the relative importance of environmental terms, but in four out of six realms resulted in an increase in the overall model fit. For Australasia, the mid-domain covariate was itself the most significant predictor ahead of all others. However, given the results for all other realms and the global model, this outcome was the exception rather than the rule. Similarly, few substantial changes in relative importance of environmental predictors were observed to result from the inclusion of mid-domain terms in our global and within-realm non-spatial model selection procedures (see table 2 in electronic supplementary material). Overall, it is noteworthy that fitting of mid-domain predictions of species richness did not substantially reduce the absolute significance of the four major environmental predictors in most of our models, and in fact enhanced this in a number of cases. Moreover, the mid-domain prediction increased overall model fit and in all but a single case was itself significant, indicating that additional variation not accounted for by the main environmental predictors was recovered. Overall, the contribution of mid-domain predictions to model fit here parallels recent studies using environmentally guided spreading dye algorithms that enforce range cohesion and find improvement in the fit of models of bird species richness over those that simply use environmental predictors (Rahbek et al. 2006; Storch et al. 2006).

(e) Relationships with human impact predictors

Despite the high proportion of bird species at risk of extinction (BirdLife International 2000), and documented relationships between levels of threat and the intensity of human impacts (McKinney 2001; McKee et al. 2003; Scharlemann et al. 2005), the inclusion of measures of such impact had limited effects on global models of avian species richness. At high resolution (1° equivalent), human population density showed a significant positive relationship with species richness. This suggests that the tendency for higher levels of human density and species richness to be favoured by similar kinds of environments (Balmford et al. 2001) overwhelms any negative effect of those densities on avian richness. The findings were similar at coarser resolutions, but with high levels of agricultural land-use and avian species richness being favoured by similar conditions. Within individual realms, relationships between species richness and human population density were also predominantly positive. However, significant negative components to these relationships at higher levels of human density (Palaearctic, Neotropics and especially Indo-Malaya) indicated that measurable human impacts on distribution are not confined to threatened species, but may also influence the overall avian richness.

4. CONCLUSIONS

This is the first study to analyse multiple possible causes of species richness patterns for a major vertebrate group at the global scale, and confirms the ultimate importance of resolution and spatial extent in analyses of the environmental determinants of large-scale species richness distribution. Moreover, our highest resolution global analyses highlight a key distinction between the emphases of spatial and non-spatial modelling methods. Non-spatial models emphasize the importance of productive energy, although this is partly a reflection of the greater geographical extent of areas of low to moderate elevational range and high productivity (e.g. across the Amazon basin). The combined influence of elevation range and temperature is the more potent driver of high richness, but areas of high elevation range are more spatially limited. Spatial models recover this signal, while non-spatial models do not. We suggest that the spatial model is more general, since it is less dependent on the contemporary prevalence of different climatic and topographic environments. Accounting for mid-domain null predictions does not alter conclusions about the relative importance of environmental predictors but may frequently improve the overall model fit.


REFERENCES


Determinants of avian species richness  R. G. Davies et al.  1197


Sibley, C. G. & Monroe, B. L.  1990 Distribution and taxonomy of the birds of the world.  New Haven, CT: Yale University Press.


