ORIGINAL ARTICLE

Using potential distributions to explore determinants of Western Palaearctic migratory songbird species richness in sub-Saharan Africa

M. S. Wisz*, B. A. Walther and C. Rahbek

Center for Macroecology, Zoological Museum at the University of Copenhagen, DK-2100, Copenhagen Ø, Denmark

ABSTRACT

Aim Although the breeding ranges of most Western Palaearctic migratory passerines are well documented in Europe, their overwintering ranges and patterns of species richness in Africa remain poorly understood. To illustrate potential patterns of species richness despite severely limited data, we extrapolated species ranges from a new and unique data bank of locality records that documents overwintering locations of these birds in Africa.

Location Sub-Saharan Africa.

Methods We predicted potential geographical distributions of 60 species of passerine birds based on overwintering records using bioclimatic models. We then combined these predictions to estimate potential species richness and explored response shapes using spatial linear regression. We also evaluated the evidence for a mid-domain effect using a one-dimensional null model.

Results Spatial linear regression analyses of the species richness pattern revealed non-linear relationships to seasonality in precipitation, minimum net primary productivity, minimum average temperature, habitat heterogeneity, percentage of tree cover, distance from the Sahara Desert and inter-annual variability in net primary productivity. The explanatory power of these variables decreased with geographic range size. The one-dimensional null model of species richness based on distance from the Sahara Desert did not show evidence of a mid-domain effect.

Main conclusions Distributions of migrants seem generally strongly determined by distance from the Sahara Desert working in concert with climatic effects, but this cannot adequately explain richness patterns of species with small ranges in Africa, many of which are of substantial conservation concern.

Keywords

Africa, bioclimatic models, contemporary climate, geographic range size, passerine birds, species richness, Western Palaearctic migrants.

*Correspondence: Mary S. Wisz, Department of Arctic Environment, National Environmental Research Institute, PO Box 358, DK-4000 Roskilde, Denmark.
E-mail: msw@dmu.dk

INTRODUCTION

Every year, millions of songbirds embark on a spectacular journey from their European breeding grounds to their wintering ranges in sub-Saharan Africa. These continent-spanning movements result in a seasonal explosion of local species richness in places that are otherwise thought to be relatively depauperate in resident species (Leisler, 1992; Jones, 1995; Hockey, 2000). Over the years, many compelling hypotheses have emerged to explain why some places are inundated annually by overwintering birds, while others are not. However, the paucity of high-quality species locality data has made it difficult to study patterns of migrant species richness and the factors that may explain the seasonal influx of bird diversity to Africa.

One long-standing hypothesis relates to the distance travelled by migratory birds. Theory predicts that a migrant should return to the breeding grounds as early as possible to compete

successfully for the best breeding territory (Kokko, 1999). Any increased cost from distance migrated should be offset by benefits gained for long-term survival and reproduction (Drent *et al.*, 2003). Some studies have shown a decreasing north–south gradient of Western Palaearctic migrant species richness within sub-Saharan Africa (Underhill *et al.*, 1992; Hockey, 2000). Thus, throughout the Afro-tropics, Western Palaearctic migrant diversity may be expected to be highest in the most northern over-wintering locations.

Recent research also points to the importance of seasonality as a determinant of migrant species richness (Hurlbert & Haskell, 2003). In particular, seasonality in rainfall, associated primarily with shifts of the Intertropical Convergence Zone, appears to trigger the movements of many Palaearctic migrants in Africa (Moreau, 1972; Curry-Lindahl, 1981; Lack, 1986; Pearson & Lack, 1992; Jones, 1995). As rainfall increases, plant productivity and insect abundance increase sharply, providing insectivorous birds with a rich food supply that presumably is not exhausted by the resident bird population (Jones, 1995). In view of this, seasonality and migrant species diversity may be positively correlated.

Yet another hypothesis relates to available energy. The species energy theory (SET) proposes that species richness should increase with some measure of ambient available energy, because greater resource availability may translate into larger population sizes for species, lower extinction risk, and an accumulation in species richness over time (Hurlbert & Haskell, 2003; Willig *et al.*, 2003). Numerous empirical studies have revealed a linear or hump-shaped relationship between species richness and net primary productivity, a commonly used surrogate for available energy (e.g. Waide *et al.*, 1999; Balmford *et al.*, 2001; Jetz & Rahbek, 2002). Some similar relationship might help to describe patterns of migrant species richness.

Another hypothesis relating to energy suggests that areas with unpredictable inter-annual resource availability may act as a selective force against resident birds, leaving more resources available to migrants in productive years (Alerstam & Enckell, 1979). Consequently, places with very low interannual variability in productivity may show relatively low levels of migrant richness in comparison with locations with intermediate to high levels.

Many species may be unable to allocate sufficient energy to growth and reproduction in cold locations, and these places may provide opportunities for migratory birds to exploit during warmer months (Hurlbert & Haskell, 2003). Accordingly, in an analysis of country checklists, Hockey (2000) found that the temperature of the coldest month was positively correlated with the percentage of resident species in Africa. However, some empirical studies have revealed a hump-shaped relationship to temperature, because few species can utilize locations that are extremely cold or hot throughout the year (Willig *et al.*, 2003).

An historical hypothesis posits that, after the emergence of the Sahara in the late Pleistocene, only savanna birds could maintain their migratory routes to the Palaearctic (Mönkkönen *et al.*, 1992). Indeed, very few Western Palaearctic migrants are known from forests or desert (Morel & Morel, 1978; Jones, 1995). Hence, a hump-shaped relationship may exist between Palaearctic migrant species richness and forest cover, because few migrants are expected to occur in treeless habitats (e.g. deserts) and continuous forest (e.g. rain forest), whereas more should occur in savanna-like habitats.

The preference for woodland and savanna habitats among Palaearctic migrants may also allow them to exploit areas where continuous forest habitat has been modified recently by humans (Leisler, 1992), and various observers have noted that migrants moved into previously forested areas after degradation or clearance in Africa (Morel & Morel, 1992). Many migrants seem to have a preference for edge habitats (Lack, 1986), and if migrants are indeed found in areas characterized by a patchy network of forest and more open habitats, it can be predicted that habitat heterogeneity (i.e. the number of distinct habitats within a specified area) will have a positive correlation with migrant richness.

The overall pattern of species richness of resident birds in Africa is driven by geographically widespread species (Jetz & Rahbek, 2002). Moreover, the role of many explanatory variables has been shown to change dramatically with decreasing range size, possibly because narrowly distributed species respond to finer-scale habitat features than do wider-ranging species, and narrow-ranging species might also be more sensitive to topographic dispersal barriers (Jetz & Rahbek, 2002). One might expect not to see the same dramatic changes among migrant species with different range sizes, as migrants are certainly not restricted by dispersal ability to the same extent as narrow-ranging resident birds.

In this paper, we evaluate quantitatively the weight of evidence supporting these hypotheses in a spatial analysis of the determinants of Western Palaearctic migrant songbird species richness in Africa. Owing to the paucity of data on the African wintering ranges of these birds, we used a new data base of point-locality records in conjunction with bioclimatic modelling tools to predict the geographical distributions of 60 individual species of Western Palaearctic migrant songbirds, and combined these into a predictive map of species richness.

METHODS

Data acquisition

We used a point-locality data base compiled from published and unpublished records from over 100 individuals, organizations, and ringing schemes within Europe and Africa (Walther & Rahbek, 2002). Although the data base currently contains information for almost 300 migrant species, we excluded species with fewer than 25 individual records dated between November and February, recorded to a spatial precision of 5 km. We also excluded species that do not regularly migrate to the sub-Saharan region, and any species for which identification problems exist between migratory Western Palaearctic and resident African subspecies (*Anthus*

novaeseelandiae, *Lanius meridionalis*, *L. excubitor*, and *Saxicola torquata*). In total, we analysed 60 species of migrant passerines based on 6941 dated point-locality records representing 2327 unique localities at 5-km resolution (Appendix S1).

Environmental predictors

All of the following predictors except distance from the Sahara Desert were used to predict the distributions of species. All predictors, including distance from the Sahara Desert, were used in the analysis of the determinants of species richness. Where the environmental data were provided at a finer spatial resolution, we resampled them to a common spatial resolution of 0.05° using ArcInfo 8.0.

Distance from the Sahara Desert

To estimate the relative distance travelled from the boundary of the Sahara Desert, we calculated latitudinal distance from 22° N latitude (the approximate southern limit of the Sahara) to each point-locality record.

Percentage forest cover

We used percentage tree-cover estimates from the Moderate Resolution Imaging Spectroradiometer (MODIS) satellite 500 m Global Vegetation Continuous Fields data set, which provides percentage tree-cover estimates from satellite imagery collected from 31 October 2000 to 9 December 2001 (Hansen *et al.*, 2003).

Seasonality of precipitation

We used the coefficient of variation of monthly average precipitation as a measure of seasonality of precipitation. We derived this layer from the Centre for Resource and Environmental Studies (CRES) African climate data set (Hutchinson *et al.*, 1995). This continent-wide data compilation is based on interpolated climatic observations from 6051 meteorological stations and represents standard mean conditions for the period 1920–80.

Minimum net primary productivity

To estimate minimum net primary productivity we used a simple classification of NDVI (Normalized Difference Vegetation Index) from AVHRR (Advanced Very High Resolution Radiometer) satellites. The data are based on 10-day NDVI maximum-value composite images with stretched NDVI values in the range 0–255 from 1982 to 1999 that were originally provided by USGS/USAID. We worked from mean NDVI values of monthly maximum-value composite images, processed by Clark Labs (Worcester, MA, USA), which excluded cloud-covered pixels. We calculated the average minimum monthly NDVI value for the period 1982–1999. Thus, for each 5-km pixel across sub-Saharan Africa, we calculated the average of all the Januarys from 1982 to 1999, then all the Februarys, and so on, and then identified the minimum average value across all months.

Inter-annual variability in net primary productivity (IAV-NDVI)

In order to estimate inter-annual variability in net primary productivity, we calculated the coefficient of variation of the mean annual NDVI estimates of the full 18-year period (1982–99) from the Clark Labs data set. Thus, we calculated average NDVI values across all months for each year, and then computed the coefficient of variation across all years.

Average temperature of the coldest month

We calculated the average temperature of the coldest month from the monthly minimum-temperature layers provided in the CRES data set. The CRES data set provides average minimum and average maximum monthly temperatures interpolated from approximately 1500 meteorological stations for the period 1920–80. We kept the data at their original spatial resolution of 0.05°.

Habitat heterogeneity

We calculated habitat heterogeneity using the 1-km-resolution Maryland Global Land Cover Map. This map is based on AVHRR, and categorizes the landscape into 13 broad categories (evergreen needleleaf forest, evergreeen broadleaf forest, deciduous needleleaf forest, deciduous broadleaf forest, mixed forest, woodland, wooded grassland, closed shrubland, open shrubland, grassland, cropland, bare ground, and urban and built up) (Hansen *et al.*, 2000). We calculated the number of land-cover categories within each 0.05° pixel (range 1–8) using ArcInfo 8.0.

Creating species distribution maps using bioclimatic envelopes

It was not possible to examine species richness patterns using sparse locality records alone, and it was inappropriate to aggregate them to coarse-resolution grids. Consequently, we modelled each species distribution using bioclimatic envelopes (Busby, 1991) and the environmental layers described above. Our bioclimatic envelope models, given a sample of locality records, predicted the species to be absent from any location outside its minimum or maximum observed range of values for any of the predictor variables, and present anywhere else.

Bioclimatic models often overpredict species distributions (Graham *et al.*, 2004). Consequently, we clipped our predictions using a map of all ecoregion polygons (Burgess *et al.*, 2004) that intersect with a given species' range as determined by 1° resolution range maps for that species documented in the Copenhagen data set (Burgess *et al.*, 1998; Brooks *et al.*, 2001).

Calculating species richness

After predicting the potential distribution of each species, we summed the individual species maps to yield five different maps of potential species richness across sub-Saharan Africa using ArcInfo 8.0. The first map corresponds to total species richness for all migrants. The remaining four maps were calculated for four groups of species based on their membership of each of the four geographic range-size quartiles, ranging from the most narrowly distributed 25% of species to the most geographically widespread 25% of species.

Exploring richness patterns with generalized linear models and spatial linear regression models

For each of the five species richness maps (all species and the four quartiles), we randomly sampled the values of the environmental predictors and the values of potential species richness from 10,000 grid cells, excluding all cells with a species richness count of zero. To identify the most parsimonious model that characterizes the African regions used by migrants, we built multiple Poisson regression models predicting the number of migrant species per grid cell using information-theoretic approaches (Burnham & Anderson, 1998, 2002). We then used spatial linear regression models to obtain parameter estimates and to calculate the proportion of variation explained by the global model, as well as each predictor variable.

Information-theoretic methods and generalized linear models

In order to explore determinants of migrant species richness using model averaging, we built 2186 generalized linear models (GLMs) with Poisson error structure, thus representing all possible combinations and subsets of the predictors with and without their quadratic expressions. The global model (i.e. the most complex model) was composed of these 14 predictors (one linear plus one quadratic expression of each of the seven environmental variables) plus an intercept, while the simplest model included only the intercept. Because parabolic relationships are rarely selected in model selection exercises (Burnham & Anderson, 1998, 2002), we did not allow any quadratic terms (x^2) in any model without also including the corresponding linear term (x) in order to reduce the number of models and thus model selection uncertainty.

Overdispersion occurs when the residual deviance is greater than the residual degrees of freedom, and results from the underlying errors being greater than assumed (i.e. not equal to the mean) (Crawley, 2002). If efforts are not made to account for this, standard errors tend to be too small, and the significance of model terms is overestimated. Our data were overdispersed because our residual degrees of freedom exceeded our residual deviance by a factor of nearly 2. To account for this phenomenon while evaluating the weight of evidence for each of these 2,186 competing models, we used QAIC, a modification of the Akaike information criteria (AIC) (Akaike, 1973) for over-dispersed count data (Burnham & Anderson,

1998, 2002). Within a set of candidate models, models with relatively low QAIC values are the most parsimonious and strike a balance between bias and variance of model predictions. QAIC is a measure of the relative Kullback–Leibler information lost in using candidate model i to approximate truth j.

$$QAIC = \frac{-2\log_e(\ell(\theta|\text{data}))}{c} + 2K,$$

where $\log_e(\ell(\theta|\text{data}))$ is the value of the maximized log-likelihood over the estimated parameters given the data and the model, c is the overdispersion parameter, and K is the number of parameters in candidate model i. The overdispersion parameter c can be estimated from the goodness-of-fit chi-square statistic (χ^2) of the global model and its degrees of freedom:

$$c = \frac{\chi^2}{\text{d.f.}}.$$

Spatial linear regression models (SLMs)

In order to evaluate and control for spatial autocorrelation, we estimated the proportion of variation explained by the global model using SLMs (Cressie, 1993), and report parameter estimates, confidence intervals, and proportion of variation explained from these models only.

We defined as spatial neighbours the eight cells immediately surrounding a cell (the queen's case), and then estimated the spatial autocorrelation in our total species richness map, and in our range-size richness maps using Moran's I statistic, a weighted correlation coefficient used to detect departures from spatial randomness (Moran, 1950). Values range from -1 to $+1$ for this statistic. Positive values indicate clustering, values equal to zero indicate spatial randomness, and negative values indicate contrasting patterns in neighbouring cells (Goodchild, 1986).

Spatial autoregressive modelling (SAR) has been increasingly used in ecology in recent years (e.g. Keitt et al., 2002; Tognelli & Kelt, 2004). We used the SAR modelling approach in S-plus with the following covariance structure: $S = [(I-rhoN)^{-t}W^{-1}(I-rhoN)]^{-1}sigma^2$, where S is the covariance matrix, I is the identity matrix, rho and $sigma$ are scalar parameters to be estimated in the spatial regression model, N is the weighted neighbour matrix, t indicates a transposed matrix, and W denotes a diagonal matrix of weights (Ripley, 1981; Cressie, 1993). We estimated the percentage variation in richness explained by calculating the coefficient of determination r^2 of the fitted spatial regression model to migrant richness for the global model, as well as the variation in richness explained by each linear predictor expressed as $x + x^2$.

We performed model selection on the full data set, and repeated the analysis on four subsets of the data set representing each of the four geographic range-size quartiles. Because the global model was supported by the majority of the evidence in model selection (see Results), we performed spatial

regression on the global model for the full data set and each geographic range size quartile.

Distance from the Sahara Desert

We also studied the relationship of species richness to distance from the Sahara using a one-dimensional null model that assumed peak diversity occurs at intermediate distances from the Sahara Desert. We implemented this test using the program RANGE MODEL, a Monte Carlo simulation tool for visualizing and assessing one-dimensional geometric constraints on species richness using the method recommended for empirical range frequency distributions with random midpoints (Colwell, 2000). To implement this method in RANGE MODEL, we calculated the midpoint of the distance between a given species' most southern predicted occurrence and 22° north latitude (approximately the southern margin of the Sahara Desert). RANGE MODEL then simulated 10,000 random midpoints for the empirical ranges in each quartile and predicted a species richness estimate for each 1° bin of distance. We then compared these simulated species richness estimates with the species richness estimates derived from our bioclimatic envelope predictions using the coefficient of determination, r^2.

RESULTS

Continental pattern of migrant richness

Our bioclimatic envelopes successfully generated maps of the potential distributions of all 60 species (see Appendix S1 in Supplementary Material). After partitioning species into quartiles based on potential geographic range size, the 15 geographically most widespread species (fourth quartile) covered between 8,901,100 km^2 (*Acrocephalus arundiaceus*) and 15,976,675 km^2 (*Hirundo rustica*). The third quartile covered between 4,418,575 km^2 (*Motacilla alba*) and 8,609,700 km^2 (*Sylvia atricapilla*). The second quartile covered between 1,236,125 km^2 (*Sylvia hortensis*) and 4,286,550 km^2 (*Acrocephalus palustis*). The most narrowly distributed quartile ranged from 2,900 km^2 (*Locustella luscinioises*) to 1,049,825 km^2 (*Luscinia svecica*). Although there is a longitudinal gradient in rainfall and a latitudinal gradient in temperature on most continents, the climate gradient from the Sahara and southwards to the equatorial zone is special because precipitation increases toward the equator while mean annual temperature decreases. Our predicted richness patterns relate to these gradients in a complex manner. Within the major vegetation zones described by White (1983), potential species richness peaked in the Sahel Acacia Wooded Grassland and Deciduous Bushland (up to 46 species in a single grid cell) and in the Eastern and Western Sudanian Woodlands (Fig. 1a). The dense Guinea–Congolian Forest had low species richness, with no more than four migratory species predicted. In contrast, species richness was as high as 20 species within more open patches of Lowland Guinea–Congolian Forest, even

when completely surrounded by forests where as few as four species are predicted. The Zanzibar–Inhambane East African Coastal Mosaic Forests and the Coastal Woodlands, Grasslands and Shrublands of Somalia had very low species richness. Richness declined from the northern Sahel through the East African Highlands towards the south. South of the Guinea–Congolian Forests, richness declined sharply moving from the Zambezian Miombo Woodlands and Dry, Deciduous Forest and Secondary Grasslands to the east towards the Southern Mosaic of Lowland Guinea–Congolian rain forest and Secondary Grasslands and Wetter Zambezain Miombo Woodlands farther west.

Most of the locality records originated from well-sampled regions, for example parts of Uganda and southern Kenya. Very few locality records were available for the Guinea–Congolian Forests or the Sahel Acaica Wooded Grasslands and Deciduous Bushlands of Sudan. While low species richness was predicted for the former, some of the highest species richness was predicted for the latter (Fig. 1b).

Effect of range size on continental pattern of migrant richness

Of the total of 12,599,085 *c*. 5-km records comprising the total migrant species richness occurring in 688,661 grid cells in our study area, 55.4% were from the 15 widest ranging species (fourth quartile), while the third, second, and first quartiles accounted for only 30.3% 12.1% and 2.2% of grid cells, respectively (Appendix S1). Species richness for the first quartile explained very little of the variance of all species ($r^2 = 0.15$), but explained substantially more of the variance moving through the more geographically widespread quartiles ($r^2 = 0.70$ for the second quartile, $r^2 = 0.86$ for the third quartile, and $r^2 = 0.79$ for the fourth quartile).

Species richness was much lower per grid cell for narrowly distributed species than for geographically widespread ones, indicating that the narrow-ranging species were not highly clustered together. However, there were some notable peaks in richness of narrowly distributed species in the Senegal Delta, within the Niger Inundation Zone (a major entry point for many migrants crossing the Sahara), and in the Eastern Highlands (Fig. 2). These zones also exhibited high richness for second-quartile species, whose distributions also extended as a discontinuous northern band across the Sahelian Acaccia Wooded Grassland and Deciduous Bushland, and in parts of the Ethiopian Highlands, with some richness reaching farther south into South Africa. Some species in the third quartile occupied the Wetter Zambezian Miombo Woodlands south of the Guinea–Congolian Forests. The fourth quartile had high species richness nearly everywhere except in the Guinea–Congolian Forests, Somalia, Ethiopia's High Montane Vegetation, and arid or cool regions of Southern Africa (Fig. 2d). Very different relationships existed between the predictors and species richness estimates for narrowly distributed and geographically widespread species, as indicated by the changing sign of the slope describing the relationship between
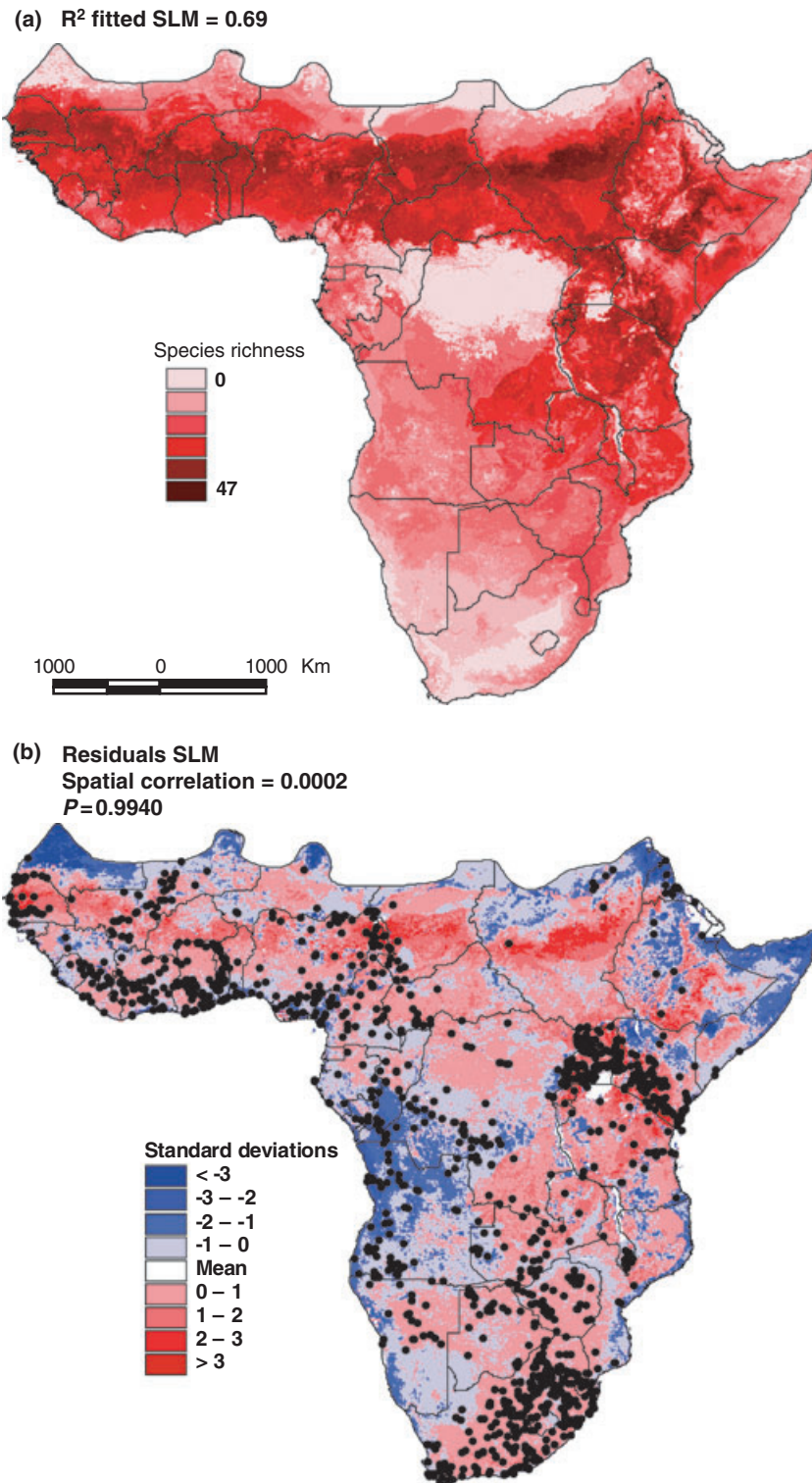
Figure 1 (a) Species richness pattern of 60 species of Western Palaearctic migrant passerines in sub-Saharan Africa based on potential distributions. Equal-interval classification is shown, with colour ramps indicating minimum (light, top of legend) to maximum (dark, bottom of legend) species richness. R^2 is the fit of the global spatial linear regression model (SLM) to this pattern. (b) The geographical pattern of model residuals for the global spatial autoregressive SLM, as presented in Table 1. Standard deviation classification ranges from dark cyan (< -3 SD) to dark red (> +3SD). Points represent all sampling localities used in our analyses. The spatial correlation in the residuals (estimated by Moran's I) is very near zero, indicating that we have successfully corrected for spatial autocorrelation ($P = 0.994$ for the null hypothesis of no spatial autocorrelation) in the sample of 10,000 records used to build the model.
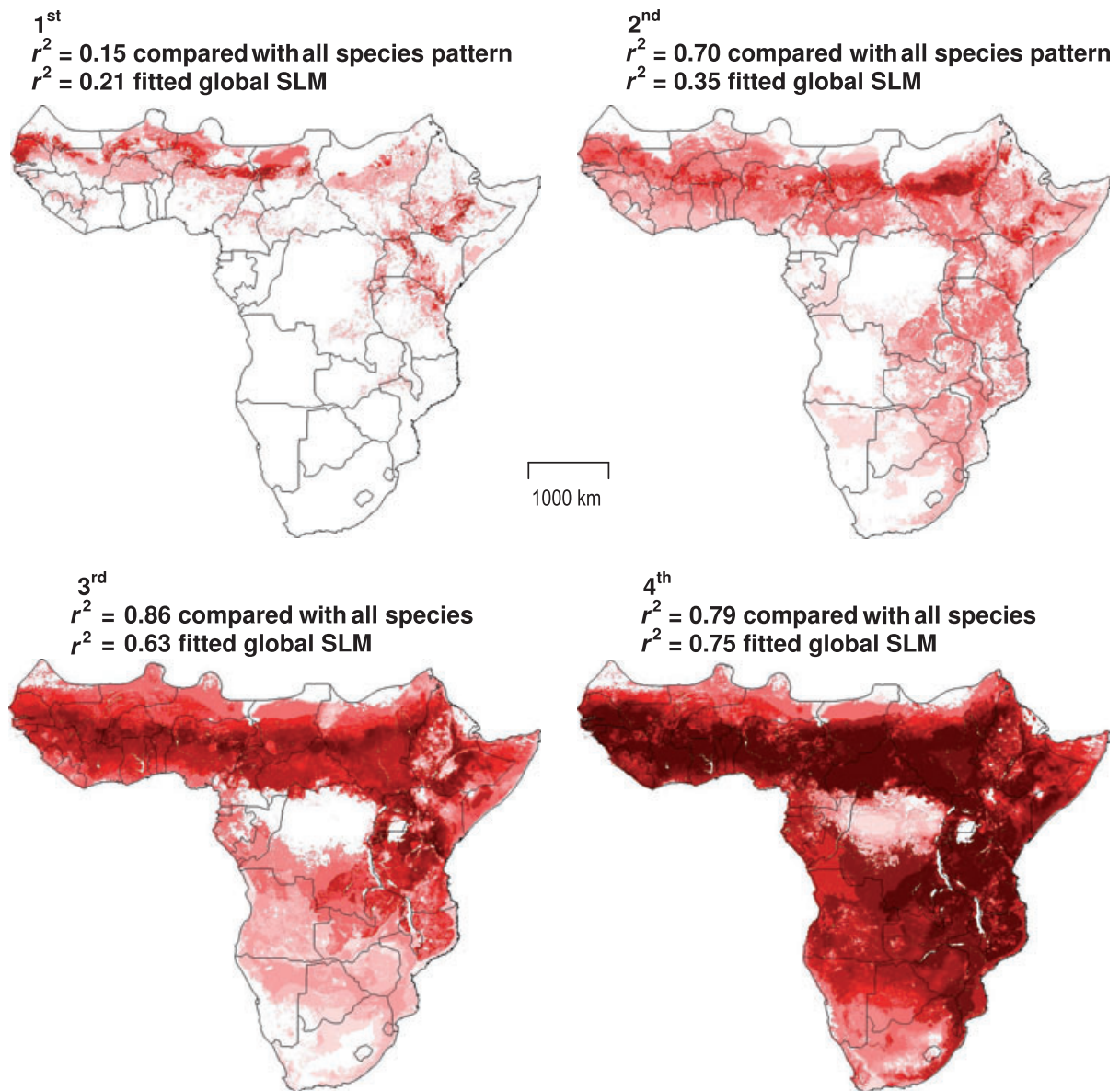
Figure 2 The geographical pattern of potential species richness for Western Palaearctic migrant passerines in Africa by range-size quartiles. Quartiles range from first (narrowest ranging 25%) to fourth (widest ranging 25%) of the 60 species overall and the empirical potential richness pattern. Equal-interval classification is shown, with light colours indicating minimum and dark colours indicating maximum species richness. r^2 compared to the species pattern is the coefficient of determination between the species richness pattern for each quartile and that of all species. r^2 for the fitted global SLM is the coefficient of determination between the spatial linear regression model and the empirical potential richness pattern.

species richness and minimum temperature, seasonality of precipitation, and inter-annual variability in NDVI (IAV-NDVI) (Table 1). Richness for the first to third quartiles was highest within a band 10° south of the Sahara, but gradually decreased with increasing distance. Richness for widespread species and all species was highest between 20° S and 10° N, but then declined nearer the Sahara, while overall richness peaked at a distance of 10° measured from the Sahara (c. 10° N latitude). The migrant richness pattern for all species exhibited positive spatial autocorrelation (Moran's $I = 0.864$, $P < 0.001$); that is, if one grid cell had high species richness, so would its neighbouring cells.

Results from GLM model selection and from SLMs

Our data supported the inclusion of each of our predictor variables and in most cases all of their quadratic terms in the GLM. For all species and for the second, third and fourth quartiles, essentially all the weight of evidence supported the global model ($w_i = 1$), which included an intercept, all seven linear predictors, and their quadratic expressions (Table 1). In the first quartile, 66% of the weight of evidence supported the global model ($w_i = 0.66$), while 33% supported the 14-term model that excluded the quadratic term for distance from the Sahara ($w_i = 0.33$). Across the range-size quartiles and all

Table 1 Spatial autoregressive linear regression results for seven linear predictors and their quadratic expressions of species richness of Western Palaearctic migrant passerines in Africa (all species), and richness of species partitioned into range-size quartiles from the narrowest ranging 25% (first quartile) to the widest ranging 25% (fourth quartile) of species. Predictors are habitat heterogeneity (HabHet), distance from Sahara (Distance), minimum average temperature (MinTemp), habitat heterogeneity squared (HabHet^2), minimum monthly average annual NDVI (MinNDVI), inter-annual variability in NDVI (IAV-NDVI), seasonality of precipitation (PPTSeason), minimum average annual NDVI (MinNDVI), minimum average temperature (MinTemp), etc. w_i is the Akaike weight for the global generalised linear model, indicating substantial weight of evidence that the five global models are the most parsimonious of 2,186 possible combinations of the linear predictors. The low spatial autocorrelation of errors (all Errors Moran's $I < 0.001$) in each resulting spatial linear regression model (SLM) indicates that the effect of spatial autocorrelation has been successfully removed by the SLM. Values in bold indicate parameter estimates whose confidence interval includes zero, indicating a lack of reliability in estimating the sign of the relationship for that particular regression coefficient

| | Parameter estimates ±95% CI for $n = 10,000$ | | | | | | | | | | t ($P > 0.01$ in bold) | | | | |
| | First quartile | | Second quartile | | Third quartile | | Fourth quartile | | All species | | First quartile | Second quartile | Third quartile | Fourth quartile | All species |
Variable	β	±95% CI	β	±95% CI	β	±95% CI	β	±95% CI	β	±95% CI					
Intercept	1.87	0.03	3.38	0.05	6.54	0.07	10.25	0.06	18.83	0.32	135.03	131.74	186.84	355.12	115.79
HabHet	-2.89	2.04	-5.64	2.85	-10.55	3.52	10.27	4.74	**-0.30**	**8.72**	-2.78	-3.88	-5.88	4.25	**-0.07**
HabHet^2	-9.67	1.61	-10.87	2.41	-23.02	2.86	-25.41	4.20	-46.49	7.26	-11.78	-8.84	-15.75	-11.86	-12.54
IAVNDVI	12.20	3.32	4.84	4.01	9.75	4.87	14.30	5.66	17.45	12.53	7.19	2.37	3.92	4.95	2.73
IAVNDVI^2	-8.52	2.20	-3.70	3.06	**1.57**	**3.99**	-10.02	4.82	**1.00**	**8.85**	-7.60	-2.38	**0.77**	-4.08	**0.22**
Distance	13.25	5.84	87.73	8.20	250.65	11.00	63.24	8.78	0.30	0.02	4.45	20.97	44.67	14.11	24.96
Distance^2	**-0.99**	**4.64**	42.67	6.91	15.20	9.43	-48.97	7.72	< 0.01	<<0.01	**-0.42**	12.10	3.16	-12.43	**0.56**
MinNDVI	7.05	5.30	**-2.07**	**6.28**	14.34	8.08	25.24	9.27	-30.64	22.62	2.61	**-0.65**	3.48	5.34	-2.66
MinNDVI^2	-12.93	2.90	-23.45	3.85	-52.37	4.81	-87.21	6.20	-165.53	13.23	-8.75	-11.93	-21.35	-27.59	-24.53
MinTemp	-12.76	2.85	26.16	7.62	70.78	9.86	125.26	8.88	273.34	27.72	-8.78	6.73	14.07	27.64	19.33
MinTemp^2	-24.11	2.75	-52.66	5.53	-75.18	7.04	-83.08	5.97	-195.83	18.29	-17.17	-18.67	-20.94	-27.26	-20.99
PPTSeason	10.95	7.16	-34.49	7.64	-127.33	10.02	-61.27	8.55	-257.09	26.82	3.00	-8.85	-24.91	-14.05	-18.79
PPTSeason^2	25.19	3.02	-75.80	5.56	-86.99	7.42	-158.37	6.08	-337.74	19.84	16.36	-26.72	-22.99	-51.03	-33.37
Tree cover	-22.54	4.02	-22.23	5.81	-56.78	7.58	-87.48	9.18	-257.85	21.90	-10.99	-7.50	-14.68	-18.69	-23.08
Tree cover^2	**1.02**	**2.72**	-15.93	3.83	-44.63	5.03	-170.62	5.95	-262.51	15.24	**0.74**	-8.16	-17.39	-56.20	-33.77
Log-likelihood	-4415		-4927		-5144		-5315		-5994						
w_i(poisson GLM)	0.66		≈ 1		≈1		≈1		≈1						
r^2 fitted SLM	0.21		0.35		0.63		0.75		0.69						
Errors Moran's I (MC)	-0.04		-0.05		-0.05		-0.02		-0.03						

species, we saw that the parameter estimates had different magnitudes, and occasionally different directions (i.e. signs) for many of the predictor variables, indicating that the relationship to each predictor variable was not the same for all groups of species. The direction was maintained across all quartiles and for all species for the intercept, the quadratic term for habitat heterogeneity, minimum temperature, the main effects of inter-annual variability in NDVI (IAV-NDVI), distance from the Sahara, and tree cover, but not unequivocally for the other terms. Confidence intervals included zero for some of our parameter estimates in these global spatial linear regression models, indicating uncertainty in our estimates of the direction of the relationship between species richness and the predictor. The only terms with such uncertainty are quadratic terms for tree cover and distance from the Sahara in the first quartile, minimum NDVI in the second quartile, the quadratic term for IAV-NDVI in the third quartile, and the quadratic terms for habitat heterogeneity and IAV-NDVI for all 60 species together.

In order to examine how the explanatory power of the global model differed for each geographic range-size quartile, we compared r^2 of the fitted values from each global SLM with the species richness response used to build each model. r^2 increased dramatically from only 0.21 for the first quartile to 0.75 for the fourth quartile. The global model explained 69% of the variance in total species richness. There was no spatial pattern in the sample residuals for these models, and values of

Errors Moran's I were close to zero, indicating essentially no remaining spatial autocorrelation (Lennon, 2000; Diniz-Filho et al., 2003).

In order to study the shape of the relationship between species richness and each predictor for each range-size category, we plotted response curves for each individual predictor. In general, relationships to predictor variables were visually more pronounced for more geographically widespread quartiles than for more narrowly distributed ones. Clear, non-linear relationships were found for most of the predictor variables for all 60 species, third, and fourth range-size quartiles. However, these relationships were less pronounced for the first and second quartiles, and difficult to discern from these plots (Fig. 3), despite the fact that non-linear relationships are suggested for all variables by model selection based on AIC.

Distance from the Sahara Desert

In order to investigate whether distance from the Sahara Desert could predict the species richness of migrants in sub-Saharan Africa, we compared one-dimensional null-model predictions generated by Range Model with our bioclimatic envelope predictions (see Methods). A marked simulated richness peak did not emerge for any quartile, or for all species (Fig. 4), indicating that the null model does not explain migrant species richness (see Discussion). Although
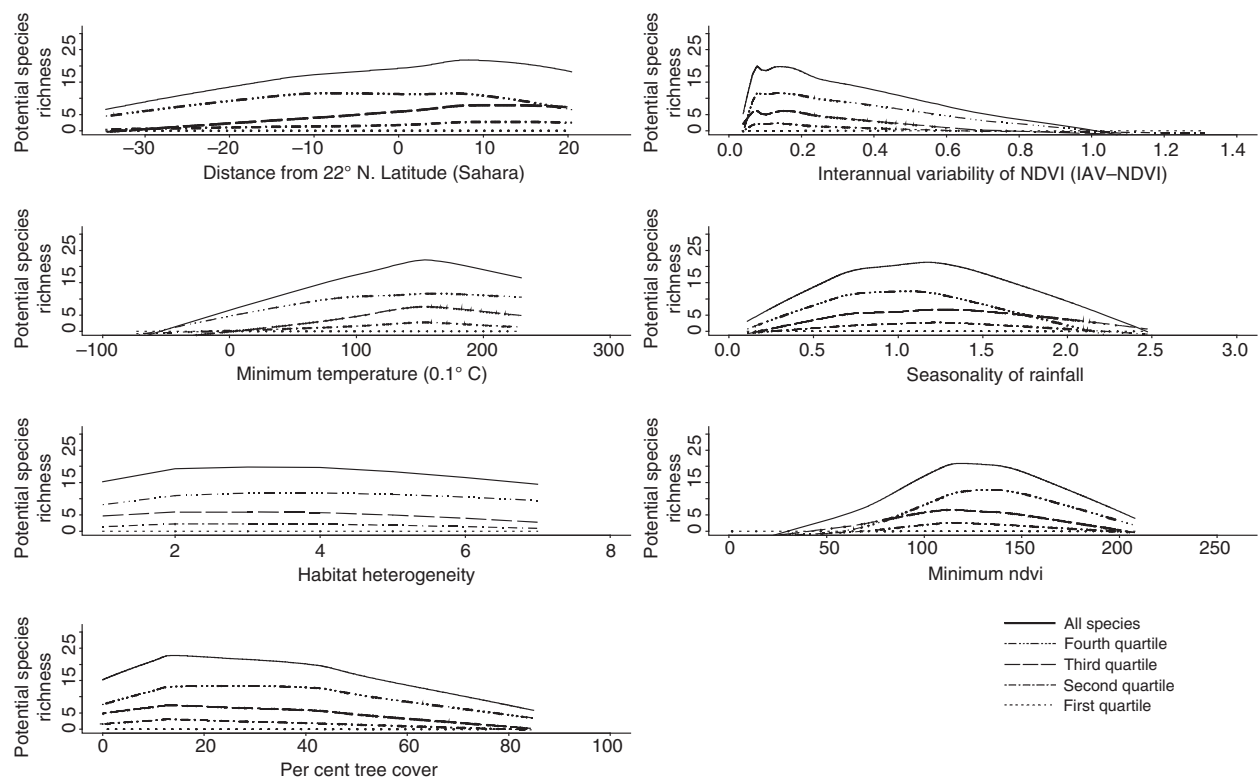

Figure 3 Fitted values of predicted potential species richness (based on the selected spatial linear regression model) as a function of each linear predictor.
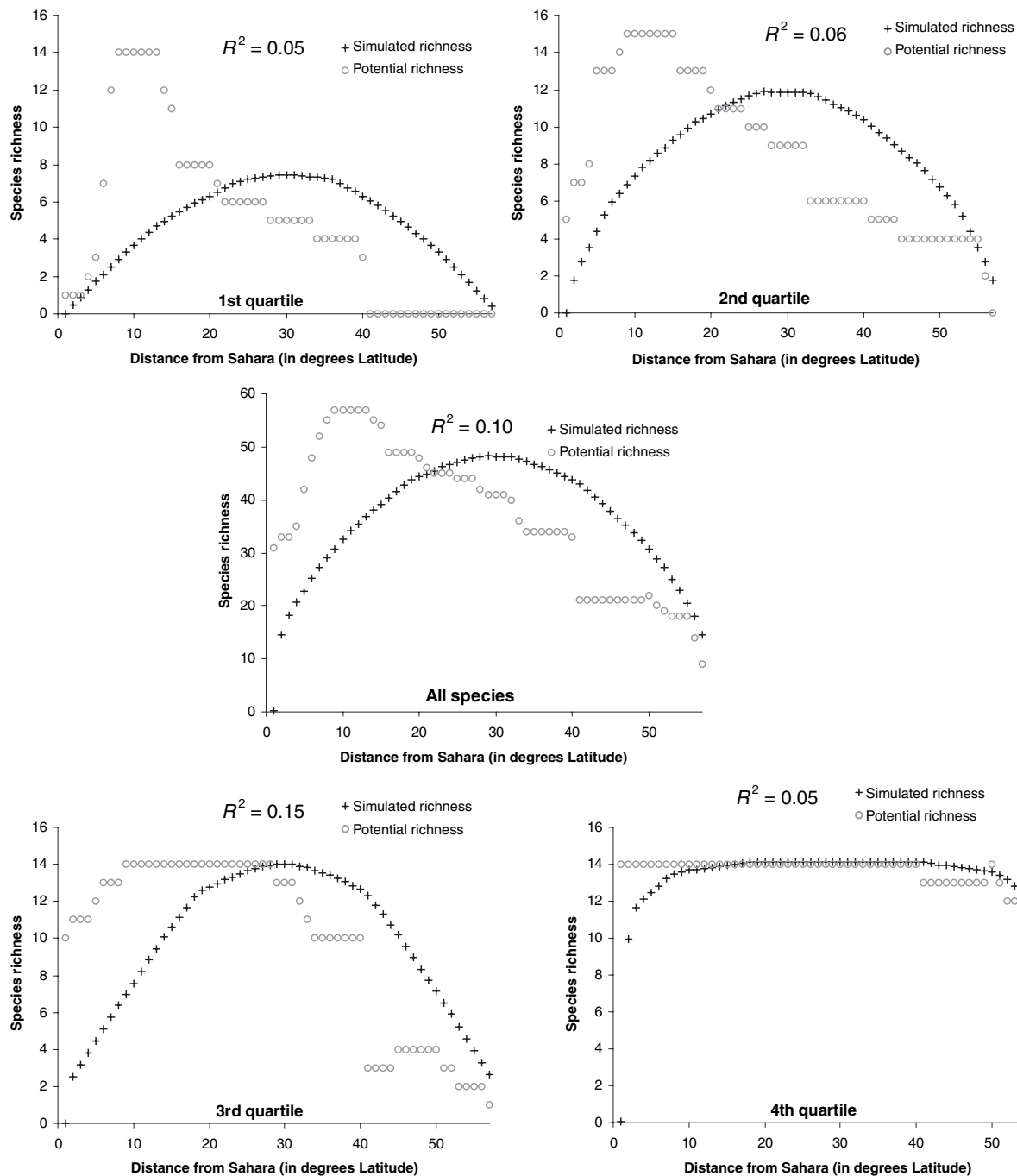
Figure 4 Species richness patterns for each range-size quartile as a function of distance from the Sahara (in degrees of latitude). Simulated patterns are those predicted by the one-dimensional null model based on Monte Carlo simulations of 10,000 random midpoints for the ranges of distance derived from the 15 potential species distributions from each quartile. Potential richness is derived from bioclimatic models for the most narrowly distributed 25% of species (first quartile), to the most widely distributed 25% of species (fourth quartile). R^2 = coefficient of determination between simulated and potential richness.

the null model predicted the highest values of species richness in the fourth quartile for intermediate values of distance, it could not predict the marked peak in richness near the Sahara. Species richness in each quartile was always higher closer to the Sahara than could be explained by the null model alone.

DISCUSSION

Our results suggest that the potential distributions of these species are determined by distance from the Sahara Desert working in concert with climatic effects (Table 1). Moreover, richness peaked close to the Sahara within the Sahel and

Northern and Eastern Savanna Zones. Thus, this pattern did not support the one-dimensional null model, which predicted that species richness should be concentrated at intermediate latitudes between the Southern margin of the Sahara and the Cape. On the contrary, we found that such latitudes are actually quite low in species richness in the first to third quartiles, and that richness is always higher in the Savanna and Sahel zones just south of the Sahara than predicted by the one-dimensional null model (Figs 2a & 4).

Measures of energy, water, or water-energy have been shown to explain richness better than any other climatic or non-climatic variable (Hawkins et al., 2003). Although we were not able to establish the relative explanatory power of our predictor variables, our results were consistent with the predictions of species energy theory. We found that net primary productivity and temperature bore hump-shaped relationships to species richness, although we found that narrowly distributed species lacked this relationship (Fig. 3). Our hump-shaped relationship to net primary productivity was also supported in the resident avifauna (Jetz & Rahbek, 2002). Although species energy theory predicts that migrant richness should be highest wherever there are seasonal pulses in surplus resource foods that cannot be consumed by the resident avifauna, we found that richness was lowest at both extremes of seasonality and peaked at intermediate values (Fig. 3). We attribute the low potential richness in the highly seasonal but sparsely vegetated Sahel zones closest to the Sahara to unsuitability in other important variables, for example insufficient tree cover or high temperatures. Such an interpretation is also consistent with recently suggested modifications to species energy theory that propose that more combinations of physiological parameters (and thus species) can survive under warm and wet than cold or dry conditions (Currie et al., 2004). Our species richness estimates were low in places where inter-annual variability in NDVI (IAV-NDVI) was extremely high. These high values usually corresponded to coastal areas where marine and freshwater communities experience high inter-annual fluctuations in the amount of shoreline vegetation.

Percentage tree cover bore a hump-shaped relationship to species richness for geographically widespread species, supporting the hypothesis that migrants are typically not found in densely forested or barren habitats, but occur in savanna-type habitats with intermediate forest cover. This result may be somewhat confounded by the low number of records from the forested Congo basin, partly resulting from either low sampling effort (see Curry-Lindahl, 1981) or the difficulty of detecting canopy-dwelling insectivores (Walther, 2003). However, whenever surveys were carried out in tree canopies, only the canopy-dwelling golden oriole (Oriolus oriolus) and aerial-feeding species such as swallows were recorded (Pearson & Lack, 1992).

Although migrants are excellent dispersers, some migrants occupy rather small wintering ranges in Africa compared with others. As Jetz & Rahbek (2002) found for resident birds, geographically widespread migrants drive the overall continental pattern of species richness. Similarly, we found that the predictor variables selected based on the hypotheses presented in the literature explain richness patterns in geographically widespread species better than those in narrowly distributed ones, also supporting some of the findings of Jetz & Rahbek (2002). Furthermore, confirming the pattern for resident birds, net primary productivity, temperature, and habitat heterogeneity explain some of the species richness pattern for migrants. Although we found that the one-dimensional null model does not explain the distribution of migrants well at all, we found that other variables (distance from the Sahara, seasonality in precipitation, and percentage forest cover) play important roles.

Our results highlight the need for new spatial environmental predictors that better reflect the habitat needs of some species, particularly those with small ranges. We expect that the distribution of small-range species might be determined by fine-scale environmental features that are not reflected in our spatial predictor variables. Although currently unavailable, detailed electronic data sets mapping aquatic vegetation zones might help us better predict some of the first-quartile species known to be associated with reed and marsh habitats, such as the threatened Basra reed warbler, Acrocephalus griseldis, and aquatic warbler, Acrocephalus paludicola (BirdLife-International, 2000).

Bioclimatic models are sensitive to the number of locality records used because they can only interpolate between the extremes of a species range in environmental space (Kadmon et al., 2003). It is therefore possible that species that are poorly sampled at the margins of their range could be predicted to have larger geographic range sizes with a more complete data set. However, this does not, for example, imply that narrowly distributed species would be outside the Sahel or Northern and Eastern savanna zones, because no first-quartile species has ever been recorded from outside these areas, except as vagrants (Dowsett & Forbes-Watson, 1993). Ideally, bioclimatic models should be validated with a reliable, independent data set of presences and absences (Guisan & Zimmermann, 2000), but this clearly was not possible for this small data set based on opportunistically collected records. Our locality records were compiled over 3 years from a wide variety of sources, and additional data simply were not available at an appropriate spatial resolution for validating these models. As biodiversity data sets become more complete (Graham et al., 2004), we anticipate that this will become possible in the future. Nonetheless, we expect that these predictions offer the most conservative spatial extrapolation of the overwintering distributions of these species that can be made with the data available. Further work should quantitatively explore how uneven sampling effort could affect richness patterns, with and without bioclimatic modelling. Because the data for each species in our sample are characterized by sampling biases, an ideal, although ambitious, approach might begin with completely simulated species distributions (e.g. Minchin, 1987), backed by a real example using strategically sampled species.

Other algorithms can result in different predictions for each species and could yield different predictions of species richness. Further work with our data could include predictions from other algorithms and the application of consensus methods such as those proposed by Thuiller (2004). Moreover, novel modelling algorithms recently introduced to ecology (e.g. those rooted in maximum entropy or improved genetic algorithms) could be considered, as these have been shown to yield robust predictions (Elith *et al.*, 2006).

Although the variables used in modelling the distribution of each individual species were also used to investigate overall patterns of species richness, this is not a circular approach. First, all our bioclimatic envelope predictions were clipped using their extent of occurrence and habitat polygons. Second, the overall pattern of species richness is a composite of many individually generated and clipped distributions and should therefore be somewhat independent of the variables that helped to generate the individual distributions. Nevertheless, further work should investigate to what degree estimates of correlations might be upwardly biased by our method.

Based on our comparison of candidate models using QAIC, the global model including all explanatory variables was the one most strongly supported by the data. Any other model that lacked any one of the variables received less support or essentially no support. However, because all the weight of evidence was on the global model, it is not possible for us to rank the importance of the predictor variables using information-theoretic approaches (Burnham & Anderson, 1998, 2002). Thus, our results point towards the importance of all of our predictors for explaining the distribution of migrant species richness.

Our analysis of the potential distributions of migrant songbirds highlighted locations where their richness fluctuates most throughout the year, and allowed us to investigate important determinants of migrant species richness. Although we could not validate predictions of the potential distributions of individual species, our approach allowed us to address hypotheses that otherwise could not be explored with the limited data available. While the determinants examined in this paper deserve further consideration, they are alone insufficient to explain the distributions of narrowly distributed species. Moreover, further work, including field research, is needed to explore how resident birds and migrants share the landscape, and to see if other migrant birds, including intra-African migrants, exhibit similar patterns.

ACKNOWLEDGEMENTS

REFERENCES

Akaike, H. (1973) Information theory as an extension of the maximum likelihood principle. *Second International Symposium on Information Theory* (ed. by B.N. Petrov and F. Csaki), pp. 267–281. Akademiai Kiado, Budapest.

Alerstam, T. & Enckell P. (1979) Unpredictable habitats and evolution of bird migration. *Oikos*, **33**, 228–232.

Balmford, A., Moore, J.L., Brooks, T., Burgess, N., Hansen, L.A., Williams, P. & Rahbek, C. (2001) Conservation conflicts across Africa. *Science*, **291**, 2616–2619.

BirdLife-International (2000) *Threatened birds of the world*. Lynx Edicions and BirdLife International, Cambridge, UK.

Brooks, T., Balmford, A., Burgess, N., Fjeldså, J., Hansen, L.A., Moore, J., Rahbek, C. & Williams, P., (2001) Towards a blueprint for conservation in Africa. *Bioscience*, **51**, 613–624.

Burgess, N., Fjeldså, J., Rahbek, T.C. (1998) Mapping the distributions of Afrotropical vertebrate groups. *Species*, **30**, 16–17.

Burgess, N.D., D'amico Hales, J., Underwood, E. & Dinerstein, E. (2004) *Terrestrial ecoregions of Africa and Madagascar: a conservation assessment*. Island Press, Washington, DC.

Burnham, K.P. & Anderson, D.R. (1998, 2002) *Model selection and multi model inference: a practical information-theoretic approach*. Springer Verlag, New York, NY.

Busby, J.R. (1991) BIOCLIM – a bioclimatic analysis and prediction system. *Nature conservation: cost effective biological surveys and data analysis* (ed. by C.R. Margules and M.P. Austin), pp. 64–68. CSIRO, Canberra.

Colwell, R.K. (2000) *RangeModel A Monte Carlo simulation tool for assessing geometric constraints on species richness*. User's Guide and application published at: http://viceroy.eeb.uconn.edu/rangemodel.

Crawley, M.J. (2002) *Statistical computing: an introduction to data analysis using S-plus*. John Wiley & Sons, Chichester.

Cressie, N.A.C. (1993) *Statistics for spatial data*. John Wiley & Sons, New York, NY.

Currie, D.J., Mittelbach, G.G., Cornell, H.V., Field, R., Guégan, J.-F., Hawkins, B.A., Kaufman, D.M., Kerr, J.T., Oberdorff, T., O'Brien, E. & Turner, J.R.G. (2004) Predictions and tests of climate-based hypotheses of broad-scale variation in taxonomic richness. *Ecology Letters*, **7**, 1121–1134.

Curry-Lindahl, K. (1981) *Bird migration in Africa*. Academic Press, London.

Diniz-Filho, J.A.F., Bini, L.M. & Hawkins, B.A. (2003) Spatial autocorrelation and red herrings in geographical ecology. *Global Ecology and Biogeography*, **12**, 53–64.

Dowsett, R.J. & Forbes-Watson, A.D. (1993) *Checklist of birds of the Afrotropical and Malagasy regions. Volume 1: species limits and distributions*. Tauraco Press, Liège.

Drent, R., Both, C., Green, M., Madsen, J. & Piersmu, T. (2003) Pay-offs and penalties of competing migratory schedules. *Oikos*, **103**, 274–292.

Elith, J., Graham, C.H., Anderson, R.P., Dudík, M., Ferrier, S., Guisan, A., Hijmans, R.J., Huettmann, F., Leathwick, J.R., Lehmann, A., Li, J., Lohmann, L., Loiselle, B.A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J.M., Peterson, A.T., Phillips, S.J., Richardson, K., Scachetti-Pereira, R., Schapire, R.E., Soberón, J., Williams, S., Wisz, M.S. & Zimmermann, N.E. (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, **29**, 129–151.

Goodchild, M.F., 1986, *Spatial autocorrelation, Concepts and Techniques in Modern Geography 47*. Geo Books, Norwich.

Graham, C.H., Ferrier, S., Huettman, F., Mortiz, C. & Peterson, A.T. (2004) New developments in museum-based informatics and applications in biodiversity analysis. *Trends in Ecology and Evolution*, **19**, 497–503.

Guisan, A. & Zimmermann, N.E. (2000) Predictive habitat distribution models in ecology. *Ecological Modelling*, **135**, 147–186.

Hansen, M., DeFries, R., Townshend, J.R.G. & Sohlberg, R. (2000) Global land cover classification at 1 km resolution using a decision tree classifier. *International Journal of Remote Sensing*, **21**, 1331–1365.

Hansen, M., DeFries, R., Townshend, J.R., Carroll, M., Dimiceli, C. & Sohlberg, R. (2003) *500m, MODIS Vegetation Continuous Fields*. The Gobal Land Cover Facility, College Park, Maryland. Dataset published at http//glcf.umiacs.umd.edu/data/modis/vcf/.

Hawkins, B.A., Field, R., Cornell, H.V., Currie, D.J., Guégan, J.F., Kaufman, D.M., Kerr, J.T., Mittelbach, G.G., Oberdorff, T., O'Brien, E.M., Porter, E.E. & Turner, J.R.G. (2003) Energy, water, and broad-scale geographic patterns of species richness. *Ecology*, **84**, 3105–3117.

Hockey, P.A.R. (2000) Patterns and correlates of bird migrations in sub-Saharan Africa. *Emu*, **100**, 401–417.

Hurlbert, A.H. & Haskell, J.P. (2003) The effect of energy and seasonality on avian species richness and community composition. *The American Naturalist*, **161**, 83–97.

Hutchinson, M.F., Nix, H.A., McMahon, J.P. & Ord, K.D. (1995) *A topographic and climate database for Africa*. Centre for Resource and Environmental Studies (CRES), Canberra, Australia.

Jetz, W. & Rahbek, C. (2002) Geographic range size and determinants of avian species richness. *Science*, **297**, 1548–1551.

Jones, P.J. (1995) Migration strategies of Palearctic passerines in Africa. *Israel Journal of Zoology*, **41**, 393–506.

Kadmon, R., Farber, O. & Danin, A. (2003) A systematic analysis of factors affecting the performance of climatic envelope models. *Ecological Applications*, **13**, 853–867.

Keitt, T.H., Bjørnstad, O.N., Dixon, P.M. & Citron-Pousty, S. (2002) Accounting for spatial pattern when modelling organism–environment interactions. *Ecography*, **25**, 616–625.

Kokko, H. (1999) Competition for early arrival in migratory birds. *Journal of Animal Ecology*, **68**, 940–950.

Lack, P.C. (1986) Ecological correlates of migrants and residents in a tropical African savanna. *Ardea*, **74**, 111–119.

Leisler, B. (1992) Habitat selection and coexistence of migrants and Afrotropical residents. *Ibis*, **134** (Suppl.), 77–82.

Lennon, J.J. (2000) Red-shifts and red herrings in geographical ecology. *Ecography*, **23**, 101–113.

Minchin, P.R. (1987) Simulation of multidimensional community patterns: towards a comprehensive model. *Vegetatio*, **71**, 145–156.

Mönkkönen, M., Helle, P. & Welsh, D. (1992) Perspectives on Palearctic and Nearctic bird migrations; comparisons and overview of life-history and ecology of migrant passerines. *Ibis*, **134** (Suppl.), 7–13.

Moran, P.A.P. (1950) Notes on continuous stochastic phenomena. *Biometrika*, **37**, 17–23.

Moreau, R.E. (1972) *The Palaearctic–African bird migration systems*. Academic Press, London.

Morel, G.J. & Morel, M.Y. (1978) Recherches écologiques sur une savane sahélienne du Ferlo septentrional, Sénégal. Etude d'une communauté avienne. *Cahier l'ORSTOM Series Biologie*, **13**, 3–24.

Morel, G.J. & Morel, M.Y. (1992) Habitat use by Palaearctic migrant passerine birds in West Africa. *Ibis*, **134** (Suppl.), 83–88.

Pearson, D.J. & Lack, P.C. (1992) Migration patterns and habitat use by passerine and near-passerine migrant birds in eastern Africa. *Ibis*, **134** (Suppl.), 89–98.

Ripley, B.D. (1981) *Spatial statistics*. Wiley, New York, NY.

Thuiller, W. (2004) Patterns and uncertainties of species' range shifts under climate change. *Global Change Biology*, **10**, 2020–2027.

Tognelli, M.F. & Kelt, D.A. (2004) Analysis of determinants of mammalian species richness in South America using spatial autoregressive models. *Ecography*, **27**, 427–436.

Underhill, L.G., Pryz-Jones, R.P., Harrison, J.A. & Martinez, P. (1992) Seasonal patterns of occurrence of Palaearctic migrants in southern Africa using atlas data. *Ibis*, **134** (Suppl.), 99–108.

Waide, R.B., Willig, M.R., Steiner, C.F., Mittelbach, G., Gough, L., Dodson, S.I., Juday, G.P. & Parmenter, R. (1999) The relationship between productivity and species richness. *Annual Review of Ecology and Systematics*, **30**, 257–300.

Walther, B.A. (2003) Why canopy access is essential to understand canopy birds: four examples from the Surumoni Crane Project. *Ornithologica Neotropica*, **15**, 41–52.

Walther B.A. & Rahbek C. (2002) Where do Palearctic migratory birds overwinter in Africa? *Dansk Ornitologisk Forenings Tidsskrift*, **96,** 4–8.

White, F. (1983) *The vegetation of Africa: a descriptive memoir to accompany the UNESCO/AETFAT/UNSO vegetation map of Africa.* UNESCO, Paris.

Willig, M.R., Kaufman, D.M. & Stevens, R.D. (2003) Latitudinal gradients of biodiversity: Pattern, process, scale, and synthesis. *Annual Review of Ecology and Systematics*, **34,** 273–309.

SUPPLEMENTARY MATERIAL

The following supplementary material is available for this article:

Appendix S1 Potential distributions of 60 species of Western Palaearctic migrant songbirds in Africa.

This material is available as part of the online article from: http://www.blackwell-synergy.com/doi/abs/10.1111/j.1365-2699.2006.01661.x (this link will take you to the article abstract).

Please note: Blackwell Publishing are not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

BIOSKETCHES

Mary Wisz is a spatial modeller in the Department of Arctic Environment at the National Environmental Research Institute in Denmark. Her current interests lie in spatial patterns of species' occurrence and abundance, the dynamics of species' habitat selection and habitat use in marine and terrestrial ecosystems, and ecological modelling techniques.

Bruno Walther is a post-doctoral zoologist at Stellenbosch University, South Africa. He works on the biogeography, ecology, and conservation of Palaearctic migratory birds in Africa. He is also interested in species richness estimation methods, and host–parasite and predator–prey interactions.

Carsten Rahbek is a professor at the University of Copenhagen, Denmark, and the director of the Center for Macroecology. His main research interests are patterns of species distribution, species range sizes, species assemblages and species richness, and what determines such patterns. His recent focus has been on the role of scale and the conceptual formulation and practical design of null and predictive models that allow direct testing of hypotheses related to patterns of diversity.

Editor: Ole Vetaas