



Contemporary environmental correlates of endemic bird areas derived from meteorological satellite sensors

D. D. P. Johnson, S. I. Hay* and D. J. Rogers

Trypanosomiasis and Land-Use in Africa (TALA) Research Group, Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK

The present-day distribution of centres of endemism is the result of an interplay between historical biogeography and contemporary environmental conditions. The relative importance of these two factors has never been established, however, for want of information on both the distributions themselves and the continental-scale measurement of environmental variables. Recently published maps of avian endemism in Africa, and the increasing availability of continental-scale surrogates of climatic conditions derived from Earth-orbiting satellites, have allowed this problem to be addressed directly. In this paper, temporal, Fourier-processed surrogate meteorological data derived from both the National Oceanic and Atmospheric Administration's series of polar-orbiting meteorological satellites and the geostationary Meteosat satellites are used within a discriminant analytical framework to describe and predict areas of bird endemism in East Africa. The technique predicts endemic bird areas (EBAs) with an accuracy of 89% (mean 85%, range 70–89%). Contemporary environmental conditions, ultimately determined by climate, therefore appear to account for a substantial fraction of the observed variation in the distribution of EBAs. On the basis of these results, several hypotheses proposed to explain the distribution of centres of avian endemism are reviewed.

Keywords: endemic bird areas; East Africa; remote sensing; NOAA-AVHRR; Meteosat-HRR; temporal Fourier processing

1. INTRODUCTION

Restricted-range or endemic bird species with limited geographical ranges and small population sizes are of great conservation concern (Stattersfield *et al.* 1998) as they are particularly vulnerable to extinction from stochastic processes (Wright & Hubble 1983). Conservation of areas containing such endemic species must be based on a clear understanding of which factors define areas of endemism in both space and time. Our understanding of these factors is still, however, far from complete.

Several alternative hypotheses have been proposed for the evolution and maintenance of centres of endemism. These involve a variety of factors, ranging from the strictly historical through to contemporary environmental. The evidence for each of these hypotheses has been reviewed by Haffer (1997) and is summarized in table 1. The relative importance of historical and contemporary factors (Endler 1982; Mayr & O'Hara 1986) has been difficult to investigate objectively, partly because historical explanations are difficult to disprove (Ball 1976; Peters 1991), but also because the relevant environmental and endemism distribution data over large areas have been lacking. One approach is to investigate to what extent the contemporary distribution of centres of endemism can be explained by recent environmental conditions (Fjelds  & Lovett 1997; Tuomisto & Ruokolainen 1997). Whereas detailed meteorological data

are not routinely available for much of the tropics, remote sensing is a powerful technique for measuring such variables consistently over continental areas (see Hay *et al.* (1996) for a review), and hence has a considerable contribution to make to ecological studies (Gaston 1996).

In this paper we use these data to investigate the distribution of endemic bird areas (EBAs), defined as areas enclosing the summed ranges of two or more overlapping 'restricted-range' bird species, each of which has a range of less than 50 000 km² (Stattersfield *et al.* 1998). This data set is, we believe, the first formal attempt to map centres of endemism for any taxonomic group on a global scale.

Recent research has shown that avian distributions can be described successfully by contemporary environmental variables. Hepinstall & Sader (1997) predicted successfully the probability of occurrence of 14 out of 23 species of land birds in Maine, USA, using data recorded by bands 4 and 5 of the Landsat Thematic Mapper (TM). Winter distributions of both sage grouse (*Centrocercus urophasianus*) in Utah (Homer *et al.* 1993) and Kirtland's warbler (*Dendroica kirtlandii*) in the Bahamas (Miller & Conroy 1990) have been predicted using vegetation classifications from Landsat TM (bands 2, 3, 4 and 5) and SPOT High Resolution Visible (HRV) (bands 1, 2 and 3) satellite images, respectively. Environmental criteria have also been used to predict the location of nesting sites of buzzards (*Buteo buteo*) in Scotland (Austin *et al.* 1996) and greater sandhill cranes (*Grus canadensis tabida*) in Minnesota (Herr & Queen 1993) using vegetation cover and landscape characteristics

*Author for correspondence (simon.hay@zoo.ox.ac.uk).

Table 1. *Hypotheses for the distribution of centres of endemism*

hypothesis	centres of endemism	isolating event	cause	period	speciation cause	speciation mode	references
palaeogeography hypothesis	regions bounded by geographical features	geographical barriers	tectonic activity	any	genetic drift	allopatric	Emsley (1965)
river hypothesis	regions bounded by rivers	rivers forming barriers	hydrological change	any	genetic drift	allopatric	Capparella (1988)
river-refuge hypothesis	forest refugia with forest-river boundaries	forest fragmentation and rivers	dry-humid cycles	Tertiary	genetic drift	allopatric	Ayres & Clutton-Brock (1992)
disturbance-vicariance hypothesis	meeting of highland and lowland	interspecific competition	cold-warm cycles	Quaternary	instability: selection	allopatric	Colinvaux (1993)
gradient hypothesis	uniform habitats	environmental (habitat) boundaries	habitat heterogeneity	any	genetic drift	parapatric	Endler (1982)
refuge hypothesis	forest refugia in forest fragments	forest fragmentation	dry-humid cycles	Tertiary	genetic drift	allopatric	Haffer (1969)
montane refuge hypothesis	montane refugia	montane forest retreat (rising)	dry-humid cycles	Tertiary	none	—	Moreau (1966); Carcasson (1964)

derived from Landsat-5 TM (bands 3, 4 and 5) and Landsat TM (bands not reported), respectively. Indeed, bird species richness *per se* has been predicted using environmental criteria, including land-use and climatic variables based on Advanced Very High Resolution Radiometer (AVHRR) data and climate (O Connor 1996). The variables they used explained 47.5% of the variance in a classification and regression tree (CART) analysis, the top-ranking predictor being the mean temperature in July.

The present paper presents the results of analyses investigating whether contemporary environmental data derived from meteorological satellites can explain the current distribution of EBAs in East Africa.

2. METHODS

(a) Summary information

This investigation used ground 'training' data from the published EBA maps of Bird Life International (Stattersfield *et al.* 1998) and multitemporal, remotely sensed surrogates for environmental variables derived from meteorological satellite sensors. Details of the processing steps required to turn the digital numbers that arise from satellite sensors into ecologically meaningful information are described below. All multitemporal satellite data were then Fourier processed to extract a set of 36 variables (table 2) that effectively describe the amplitude and timing of seasonal changes in temperature, vegetation amount and rainfall. Elevation information was also available from a digital elevation model (DEM). The boundaries of the EBAs were then overlaid on these data layers, and a set of randomly chosen points was chosen, and satellite data extracted, for both EBA and non-EBA areas. Because of the great range of environmental conditions, even within East Africa, it was considered desirable to carry out preliminary clustering of all the satellite data before further analysis. Up to five clusters each for EBA 'presence' or 'absence' were therefore defined before further processing, which involved nonlinear discriminant analysis relating the distribution data (the predicted variable, in up to five clusters) to subsets of satellite variables (the predictors, also up to five clusters). Predictor variables were selected on the basis of their ability to distinguish EBA and non-EBA areas, and the ten most important predictors in each case were then used to generate maps of East Africa showing the (posterior) probabilities of each area belonging to the EBA or non-EBA category. These maps therefore record the probabilities with which each area has environmental conditions similar to those in recorded EBAs: a close correspondence between observed EBAs and EBAs predicted on the basis of the predictor variables implies that contemporary environmental conditions are important determinants of endemic bird distributions.

(b) Endemic bird area (EBA) data

Restricted-range bird species are those with a range of less than 50 000 km². An EBA is defined as an area which encompasses the overlapping breeding ranges of restricted-range bird species, such that the complete ranges of two or more restricted-range species are entirely included within the boundary of the EBA (Stattersfield *et al.* 1998). Any disjunct distributions were summed to calculate an overall range size. To map the EBAs, all point locality records with their longitude and latitude coordinates were collated from various sources, including 3000 ornithological references, unpublished records from Bird Life's network of regional experts, and additional data collected from

the ground. The EBA boundaries were defined as the minimum polygon encompassing all records, and were subsequently modified to account for habitat requirements and altitudinal ranges of each restricted-range species. Bird Life International has identified and mapped all known EBAs for Africa, according to the above criterion (ICBP 1992; Crosby 1994; Long *et al.* 1996; Stattersfield *et al.* 1998). In this study only the EBA map for eastern Africa was used, as it was judged that this region had been most comprehensively documented, whereas the coverage for the whole continent was considerably more patchy.

(c) Digital elevation model (DEM)

Elevation data were extracted from the DEM for Africa from the Global Land Information System (GLIS) of the United States Geological Survey (USGS) (Anonymous 1996). The DEM was created by digitizing elevation data from 1:1 000 000 air navigation charts for the whole of Africa, using individual spot heights and selected contours and coastlines. These digital data were then spatially interpolated to provide a 1 km × 1 km resolution DEM product. For present purposes, the original files were then resampled to a 0.05° spatial resolution image to ensure compatibility with the resampled satellite sensor data. The accuracy of the final DEM product was evaluated by a comparison with published elevations from 207 National Oceanic and Atmospheric Administration (NOAA) National Climate Data Centre (NCDC) meteorological stations across Africa, which gave a coefficient of determination of $r^2=0.90$ ($n=207$, $p \ll 0.0001$, r.m.s.e.=145 m) (Hay 1996).

(d) Deriving environmental data from satellite sensors

Satellite sensor data for the entire Earth's surface are available from the AVHRR on board the NOAA polar-orbiting meteorological satellites (Kidwell 1995). These data were processed to produce surrogate vegetation and temperature data, as described by Hay *et al.* (1996) and references therein. Briefly, the normalized difference vegetation index (NDVI) exploits the fact that actively photosynthesizing vegetation appears darker in the visible region and brighter in the infrared region of the electromagnetic spectrum than senescing vegetation or the soil background (Sellers 1985). The NDVI is defined as

$$\text{NDVI} = \frac{(\text{channel 2} - \text{channel 1})}{(\text{channel 2} + \text{channel 1})}, \quad (1)$$

where channel 1 is in the visible wavelengths (0.58–0.68 μm) and channel 2 in the near infrared wavelengths (0.725–1.10 μm). Ten-day (decadal) images were produced by extracting information on the most cloud-free day in the period, a process called maximum value compositing (MVC) (Holben 1986). The NDVI is a specific measure of chlorophyll abundance and energy absorption (Myneni *et al.* 1995), but its use has been extended through multitemporal observations to classify major vegetation types (Tucker *et al.* 1985) and phenology (Justice *et al.* 1985) in a range of ecosystems throughout Africa.

NDVI data were obtained from the Food and Agriculture Organization's (FAO) African Real Time Environmental Monitoring using Imaging Satellites (ARTEMIS) program at 7.6 km × 7.6 km spatial resolution for the period 1982–1990. The registration of these images was checked against a georeferenced 'master' image, and corrections were made where necessary: this involved shifting images by from 0 to 3 pixels in an east–west or north–south direction, depending on the scene. The raw imagery was then corrected for satellite sensor drift in channel 1 using

Table 2. *Fourier-processed variables used in the analysis*

feature	abbreviation	description	units
altitude			
1	DEM	digital elevation model of height above mean sea level	m
vegetation			
	NDVI	normalized difference vegetation index	ratio
2	NDmean	NDVI mean	ratio
3	NDp1	NDVI phase 1	months
4	NDa1	NDVI amplitude 1	ratio
5	NDp2	NDVI phase 2	months
6	NDa2	NDVI amplitude 2	ratio
7	NDp3	NDVI phase 3	months
8	NDa3	NDVI amplitude 3	ratio
9	NDmax	NDVI maximum	ratio
10	NDmin	NDVI minimum	ratio
11	NDrange	NDVI range	ratio
rainfall			
	CCD	cold cloud duration	h
12	CCDmean	CCD mean	h
13	CCDp1	CCD phase 1	months
14	CCDa1	CCD amplitude 1	h
15	CCDp2	CCD phase 2	months
16	CCDa2	CCD amplitude 2	h
17	CCDp3	CCD phase 3	months
18	CCDa3	CCD amplitude 3	h
19	CCDmax	CCD maximum	h
20	CCDmin	CCD minimum	h
21	CCDrange	CCD range	h
temperature			
	Ch4	channel 4 (NOAA-AVHRR)	
22	Ch4mean	thermal radiance mean	K
23	Ch4p1	thermal 4 phase 1	months
24	Ch4a1	thermal 4 amplitude 1	K
25	Ch4p2	thermal 4 phase 2	months
26	Ch4a2	thermal 4 amplitude 2	K
27	Ch4p3	thermal 4 phase 3	months
28	Ch4a3	thermal 4 amplitude 3	K
29	Ch4max	thermal maximum	K
30	Ch4min	thermal minimum	K
31	Ch4range	thermal range	K
combinations			
32	Ndp – Cpd	NDVI phase – CCD phase	months
33	NDm/CDm	100 × NDVI mean/CCD mean	—
34	NDm/Ch4m	100 × NDVI mean/thermal mean	—
35	NDp1 – Ch4p1	NDVI phase 1 – thermal phase 1	months
36	NDa1/Ch4a1	100 × NDVI amplitude 1/thermal amplitude 1	—

calibration coefficients derived by Los (1993), and then MVC by selecting the highest value of the decadal pixels for each site within each month, to produce a set of monthly images for further analysis.

The NOAA-AVHRR channel 4 brightness temperature correlates with air temperature at the Earth's surface (Hay *et al.* 1996). Decadal data at 7.6 km × 7.6 km spatial resolution from the

archives of the Global Inventory Monitoring and Modelling Systems (GIMMS) group at the NASA Goddard Space Flight Center were maximum value composited for the period 1987–1992. Monthly imagery was again produced by MVC.

The proxy measure of rainfall is derived from the High Resolution Radiometer (HRR) on the geostationary Meteosat satellite series. The HRR channel 2 records the thermal infrared signal or temperature. Thick clouds with cold tops are associated with convective rainfall systems in the tropics, and have a high probability of precipitating rain when their temperature drops below a threshold value. The threshold temperature for rainfall, and the amount of rain deposited at a particular temperature, are variable. Empirical observations have quantified these variations over large parts of the Sahelian zone of Africa, to allow the production of cold cloud duration (CCD) imagery for the region that may be used as a rainfall surrogate (Snijders 1991). The CCD imagery was obtained from the FAO-ARTEMIS program as five-year monthly averages for the period 1988–1992 at a 7.6 km × 7.6 km spatial resolution.

(e) *Satellite data ordination*

The satellite data were first subjected to temporal Fourier processing: this extracts from each set of images a description of the annual, biannual, and triannual components that together comprise the observed signal. This process achieves data reduction in a biologically meaningful way, from a maximum set of 108 monthly images (the NDVI series) to a standard set of seven images: the overall mean and the amplitudes and phases (=the timing of seasonal peaks) of the three within-year cycles (Rogers & Williams 1994; Rogers *et al.* 1996; see Chatfield (1980) for the mathematical basis). In addition, the reconstructed Fourier description of the annual cycle of change in each variable was used to generate the maximum, minimum, and range of values shown during an average year.

After Fourier processing, all satellite sensor data were resampled to a 0.05° grid before further analysis. A total of 36 predictor variables was generated (table 2), each in the form of a georeferenced image file.

(f) *Data analysis*

A discriminant analytical approach (Green 1978; Williams *et al.* 1992; Rogers *et al.* 1996) was employed to describe the EBA data in East Africa, using the set of variables in table 2.

First, a random sample of 500 pixels from within and outside known EBAs was designated as the training set data. Values for all predictor variables were extracted from each image, using the training sample pixel co-ordinates, and the resulting data set was subjected to cluster analysis (using the k-means cluster algorithm of SPSS), keeping data for sites within EBAs separate from data for sites outside EBAs. One-to-five clusters were generated for each set of data, and subsequent analysis involved a user-selected, variable number of clusters to examine how partitioning the data changed the overall fit of the model. Important predictor variables were chosen during analysis in a forward, step-wise manner. The criterion for selection was that it maximized the minimum Mahalanobis distances (a variance-adjusted measure of separation in multivariate space) between any pair of presence or absence clusters during that round of variable selection. This method chooses variables that separate best the least-separable presence or absence cluster pair in the data set. Different runs of the model assumed either that the covariances of the variables were the same for every cluster (linear discriminant analysis), or that they varied according to the training set data for each cluster (nonlinear

Table 3. The accuracy of predictions according to cluster number using variable selection method of maximizing the minimum Mahalanobis distances between classes

(Combinations of cluster numbers not included are those in which sample sizes are too small for discriminant analysis.)

clusters for absence	clusters for presence	sum of clusters	% correct +ve	% false -ve	% false	A sensitivity	B specificity	A × B
1	1	2	82	13	4	0.860	0.809	0.696
1	2	3	85	10	5	0.833	0.860	0.716
1	3	4	85	13	2	0.940	0.817	0.768
1	4	5	87	10	3	0.900	0.853	0.768
2	2	4	87	8	6	0.807	0.891	0.719
2	3	5	87	10	3	0.900	0.859	0.773
2	4	6	87	10	3	0.897	0.863	0.774
3	2	5	84	9	7	0.770	0.876	0.675
3	3	6	87	10	3	0.903	0.853	0.770
3	4	7	89	8	4	0.883	0.891	0.787
4	2	6	70	0	30	0.000	0.999	0.000
4	3	7	88	8	4	0.830	0.880	0.730
4	4	8	87	9	3	0.890	0.867	0.772

discriminant analysis): obviously, the assumption involved here will affect the choice of variables by the analysis according to the selection criteria outlined above, as Mahalanobis distances are calculated from the inverses of the covariance matrices. Once the analysis had selected ten predictor variables and defined the covariance matrices, it first examined the accuracy with which the training set data could be correctly assigned to presence or absence categories: this was done through the calculation of posterior probabilities, assuming equal prior probabilities (Tatsuoka 1971; Green 1978; see Rogers *et al.* (1996) for further details and an example). Finally, the posterior probabilities were calculated for the entire region of East Africa, using the ten predictor data layers selected by the analysis, and a map was produced. For display purposes, the predictions were colour coded from red (low probability of EBA conditions) to green (high probability of EBA status).

3. RESULTS

The analysis was performed for all combinations of absence and presence clusters that gave a sufficient sample size ($n \geq 15$) for each cluster. Each of the resulting 'models' (table 3) was evaluated by examining the overall percentage correct (i.e. both presence and absence) classification of the training set, the percentage false positives (i.e. a false prediction of EBA status) and false negatives (a false prediction of non-EBA status). In addition, the sensitivity and specificity of the method were calculated (i.e. the ability to identify correctly EBA or non-EBA areas, respectively).

Single clusters for both presence and absence gave an overall accuracy of 82% correct prediction (with 13% false positives and only 4% false negatives): accuracy generally increased as cluster number increased. The best prediction was achieved using three absence and four presence clusters, giving an overall accuracy of 89%, with 7.5% false positives and 3.5% false negatives. The predicted map for this situation is shown in figure 1, the legend recording the variables selected.

Elevation was the top-ranking predictor variable in 8 of the 12 analyses shown in table 3. This is clearly associated with many of the inland areas of endemism, but not with all areas such as the Somali coastal dune system. Four of

the cluster combinations did not include the DEM as either the first or second variable, while still achieving overall predictive accuracies of 70–88% (8–10% false positives, 4–30% false negatives). The top three predictors of the best model, where elevation was not included, were channel 4 phase 2, channel 4 amplitude 3 and (NDVI phase 1 – channel 4 phase 1) (i.e. thermal and vegetation variables), although two of these were significantly, but poorly, correlated with elevation ($r=0.09$, $p<0.01$; $r=0.33$, $p<0.01$; $r=0.03$, n.s.). In addition, when the DEM was deliberately excluded as a predictor variable, the model was still able to predict EBAs with 80% accuracy (16% false positives, 5% false negatives). Without the DEM, the top predictor variables were CCD phase 2, CCD minimum and channel 4 amplitude 3 (i.e. rainfall and thermal variables). Because the DEM was excluded, the correlation coefficients between these three predictor variables and the DEM could not be calculated directly. However, where these particular variables did occur with the DEM, in other cluster combinations, two were significantly correlated with elevation ($r=-0.39$, $p<0.05$; $r=-0.05$, n.s., $r=0.33$, $p<0.01$, respectively).

4. DISCUSSION

Previously proposed hypotheses for the origin and distribution of centres of endemism (table 1) are unsatisfactory for the EBAs of East Africa. The palaeogeography hypothesis (Emsley 1965), suggesting that natural barriers delimit areas of endemism, can be rejected as no geographical structures (such as mountain ranges) form the boundary of any EBAs in the region. Similarly, the river hypothesis (Capparella 1988) and the river-refuge hypothesis (Ayles *et al.* 1992) are inapplicable, as no East African EBA is limited by the course of a large river. The disturbance–vicariance hypothesis (Colinvaux 1993) proposes that climatic cooling created a dynamic habitat periphery between lowland and upland areas, where species would become isolated owing to competition (see discussion in Haffer 1997). There is no evidence for such a process, and it can be rejected for East Africa because no cooling is evident in the Tertiary for Africa, when the principal avian

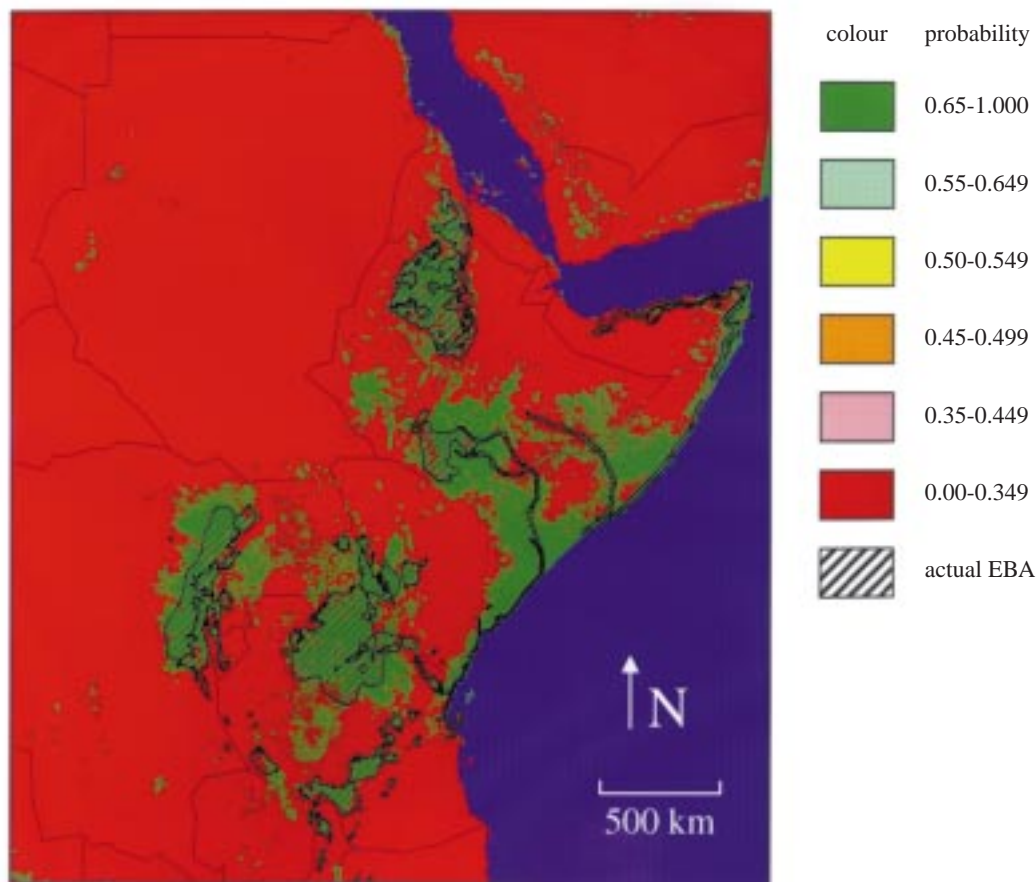


Figure 1. Results of discriminant analysis showing the prediction of endemic bird areas (EBAs) for pixels in East Africa. National borders and sea are in blue. Colour indicates the probability of EBAs, from green (high probability) to red (low probability). Hatched regions are EBAs. The ten variables included in this analysis, ranked in descending order of importance, were the DEM, Ch4a3, NDa2, Ch4a2, NDp2, Ch4p1, Ch4p2, CCDA2, Ch4mean, CCDmean (see table 2 for descriptions of variables). Comparison of predicted areas with the training data show an 89% correct overall fit, 8% false positives and 4% false negatives; sensitivity=0.88, specificity=0.89.

radiations occurred (Haffer 1997). The refugia hypothesis (Haffer 1969) applies only to lowland rainforest, and is therefore irrelevant to all East African EBAs, except perhaps for one in East Congo. The montane refuge hypothesis provides some evidence for historical factors resulting in endemism, but only in one small subset of the many areas of endemism, the montane forests. Africa has disjunct populations of more than 70 closely related bird species that co-occur in montane forests as distant as Mt Cameroon and the Albertine Rift, but not in between (Prigogine 1987), and of species co-occurring on Tanzanian mountain tops, but not in the intervening lower altitude habitat (e.g. Dinesen *et al.* 1994). This has been attributed to a lowering of montane habitat during past glacial cycles, during which time the species could have dispersed and later been trapped in high-altitude refuges when montane forest contracted. Despite debate about the mechanism (Moreau 1966; Prigogine 1987), there is considerable evidence that montane habitat was formerly connected (Carcasson 1964; Prigogine 1987; Coe 1989). Moreover, new evidence of disjunct populations of flightless insects supports such a hypothesis (Brühl 1997). Finally, the gradient hypothesis (Endler 1982), which was formulated to explain endemism biogeography on broad ecological grounds, is in agreement with data presented in Fjeldsâ (1994) and Fjeldsâ & Lovett

(1997), and has been shown in the present study to account for current EBA distribution through the use of contemporary environmental variables. Although this approach has been criticized (Mayr & O'Hara 1986), the present results support the idea that contemporary ecological factors may be more important than previously supposed in the creation and maintenance of EBAs.

Fjeldsâ *et al.* (1997) investigated whether ecoclimatic variables correlated with 'hotspots' of peak concentrations of endemic species in East Africa, which were defined as those areas that contained both neoendemics and 'phylogenetically old' species. Fjeldsâ *et al.* (1997) claim that these hotspots are found within, or on the boundary of, local reductions in ecoclimatic variability, which is measured as the summed variance of the monthly temperature (T)/vegetation index (NDVI) ratios derived from NOAA-AVHRR data. The authors show that ecoclimatic variability was lower on longitudinal transects over mountains known to be endemic hotspots than transects over mountain ranges that were not. It is not clear, however, how these control mountains were selected.

The present study shows that EBAs are environmentally distinct in their average conditions from their surrounding areas, regardless of any additional temporal environmental variation. Although Fjeldsâ *et al.* (1997) showed

ecoclimatic variability correlated with 'hotspots' of endemism, all of these were forests; the EBAs identified by Bird Life International and that we analysed are of many different habitat types. It therefore remains to be established whether climatic variability is a necessary or sufficient correlate of endemic areas in general, or of EBAs in particular.

To address questions about speciation and persistence of endemic areas, it is necessary to measure variation over a much longer time-scale than is represented in the satellite record, and to include the extremes of climatic flux. The 15 years of AVHRR data currently available are unlikely to be representative of such longer-term cycles. Nevertheless, as Fjelds  *et al.* (1997) argue, such locally moderated climatic stability may be a persistent feature that has operated on more marked climatic shifts in the past.

Centres of endemism are often high-altitude sites, but altitude *per se* is not a prerequisite, as shown by the low-altitude EBAs of the Serengeti, the Somali coastal dunes and the East Congo lowland forest, and also by some high-altitude sites that are not EBAs (e.g. the mountains of Jebel Marra, Darfur region, Sudan, or the Ahmar Mountains in north-east Ethiopia). Furthermore, high-altitude sites can contain distinct habitats, for example the contrasting Ethiopian Highlands, a dry, cool oasis for temperate-adapted species, and the Eastern Arc Mountains, which support warm, moist, montane rainforests. Both are areas of high endemism (Kingdon 1990), but for reasons of habitat, not necessarily of altitude.

In even the best model (figure 1), a considerable land area is predicted as an EBA when it is not an EBA according to the ground data. This could be due to errors in the predictions or errors in the EBA ground data, or both. For example, although endemism was falsely predicted in mountains along the Red Sea coast of Saudi Arabia, this area was formerly designated as an EBA (ICBP 1992). Systematic field surveys will resolve this issue. However, validation exercises should be seen as an iterative process where increasingly accurate training data facilitate better predictions, which in turn can be used to prioritize survey efforts. It should be noted that any false predictions should not be attributable to recent anthropogenic influences, since a criterion of EBA status is that, where known, the species must have had restricted ranges since 1800 (Stattersfield *et al.* 1998). This condition was included to avoid biases that may have arisen from diminishing species' ranges associated with habitat loss.

Accurate definitions of EBAs offer considerable potential for remote monitoring and hence management of areas of endemism. Across the globe, endemic birds represent 73% of all 'threatened' bird species. One-quarter of all birds are endemic, and all occur within less than 5% of the Earth's land area (Long *et al.* 1996). There is cause for concern as deforestation rates are particularly high near or within EBAs (Balmford & Long 1994). The improving spatial and temporal resolution of satellite sensor data (Hay 1997) can facilitate local monitoring of environmental change (especially around EBAs) for the purposes of conservation.

This study demonstrates that, in East Africa, centres of endemism can be defined objectively by their environments, but only narratively by their history. Because EBAs can be characterized by a specific subset of environmental factors,

they can be considered as ecological islands. If such sites are indeed distinct, and have been stable for a long period (for which there is some evidence (e.g. Prell *et al.* 1980; Fjelds  *et al.* 1997)), physical isolation (for instance by Pleistocene forest fragmentation) is not required to explain the distribution of restricted-range birds. Alternatively, centres of endemism may have remained climatically distinct 'migrating islands', enabling limited change to have occurred in any one geographical location. Both scenarios, however, favour the importance of environmental determinants of species' distributions.

It is of interest to note that none of the previously conceived hypotheses can account for the existence of areas of endemic areas in such widely diverse habitats and regions as the Serengeti, the Tanzanian coastal monsoon forests or the mangrove swamps. Most hypotheses are specifically directed at one habitat type, and hence fail as holistic explanations of endemism distribution and maintenance. A new theory would explain isolated populations by their occurrence in places that remained ecologically unique, at least in their environmental conditions. This suggestion automatically accommodates aspects of the more convincing historical explanations (e.g. the refugia hypothesis for the East Congo, and the montane refuge hypothesis for the montane rainforest sites) with other reasons for ecological uniqueness which set areas of endemism apart from their surroundings today (Fjelds  1994; Fjelds  *et al.* 1997; this study). Mayr & O'Hara (1986) caution against using contemporary environmental variables as explanatory of centres of endemism distributions. In contrast, this study has shown that ecological models can predict EBAs accurately, and therefore suggests that ecological factors are important, if not predominant, in explaining current EBA distributions.

We gratefully acknowledge Bird Life International and Mike Crosby for permission to use these EBA data and for help and advice. We also thank Jim Tucker of the GIMMS group at the NASA Goddard Space Flight Center and Fred Snijders of the FAO-ARTEMIS project for permission to use AVHRR (channel 4 and NDVI) and Meteosat CCD data, respectively, and William Wint for help with GIS and image processing. Richard Southwood, Malcolm Coe, Mike Packer, Jim Briskie and Clive Hambler read and commented on the manuscript, and Oliver Kr ger, Herr Walter Jetz and Susan Williams contributed to many discussions.

This publication is an output of research funded by the Department for International Development (DFID), UK, Project ZC0012. The DFID, however, can accept no responsibility for any information provided, or views expressed. D.D.P.J. also acknowledges with gratitude the support of a research masters studentship and research expenses from BBSRC while enrolled for an M.Sc. in Biology at the University of Oxford.

REFERENCES

- Austin, G. E., Thomas, C. J., Houston, D. C. & Thompson, D. B. A. 1996 Predicting the spatial distribution of buzzard *Buteo buteo* nesting areas using a Geographic Information System and remote sensing. *J. Appl. Ecol.* **33**, 1541–1550.
- Anonymous 1996 *GTOPO30 Documentation*. Universal Resource Locator: <http://edcwww.cr.usgs.gov/landdaac/gtopo30/readme.html>. Sioux Falls, USA: EROS Data Centre, Global Land Information System.
- Ayres, J. M. C. & Clutton-Brock, T. H. 1992 River boundaries and species range size in Amazonian primates. *Am. Nat.* **140**, 531–537.

- Ball, I. R. 1976 Nature and formulation of biogeographical hypotheses. *Syst. Zool.* **24**, 407–430.
- Balmford, A. & Long, A. 1994 Avian endemism and forest loss. *Nature* **372**, 623–624.
- Brühl, C. A. 1997 Flightless insects: a test case for historical relationships of African mountains. *J. Biogeogr.* **24**, 233–250.
- Capparella, A. 1988 Genetic variation in Neotropical birds: implication for the speciation process. *Act. Cong. Int. Ornith.* **19**, 1658–1664.
- Carcasson, R. H. 1964 A preliminary survey of the zoogeography of African butterflies. *E. Afr. Wildl. J.* **2**, 122–157.
- Chatfield, C. 1980 *The analysis of time series: an introduction*. London: Chapman & Hall.
- Colinvaux, P. 1993 Pleistocene biogeography and diversity in tropical forests of South America. In *Biological relationships between Africa and South America* (ed. P. Goldblatt), pp. 473–499. Yale University Press.
- Coe, M. 1989 Biogeographic affinities of the high mountains of tropical Africa. In *Quaternary and environmental research on East African mountains* (ed. W. C. Mahaney), pp. 257–278. Rotterdam: Balkema.
- Crosby, M. J. 1994 Mapping the distribution of restricted-range birds to identify global conservation priorities. In *Mapping the diversity of nature* (ed. R. I. Miller), pp. 145–154. London: Chapman & Hall.
- Dinesen, L., Lehmborg, T., Otto Svendsen, J., Hansen, L. A., & Fjeldså, J. 1994 A new genus and species of Percidine bird (Phasianidae, Percidini) from Tanzania: a relict with Indo-Malayan affinities. *Ibis* **136**, 2–11.
- Emsley, M. G. 1965 Speciation in *Heliconius* (Lepidoptera, Nymphalidae): morphology and geographic distribution. *Zoologica* **50**, 191–254.
- Endler, J. A. 1982 Problems in distinguishing historical from ecological factors in biogeography. *Am. Zool.* **22**, 441–452.
- Fjeldså, J. 1994 Geographic patterns for relict and young birds in Africa and South America and implications for conservation priorities. *Biodiv. Conserv.* **6**, 325–346.
- Fjeldså, J. & Lovett, J. C. 1997 Geographical patterns of old and young species in African forest biota: the significance of specific montane areas as evolutionary centres. *Biodiv. Conserv.* **6**, 325–346.
- Fjeldså, J., Ehrlich, D., Lambin, E. & Prins, E. 1997 Are biodiversity 'hotspots' correlated with current ecoclimatic stability? A pilot study using the NOAA-AVHRR remote sensing data. *Biodiv. Conserv.* **6**, 401–422.
- Gaston, K. J. (ed.) 1996 Species richness: measure and measurement. In *Biodiversity—a biology of numbers and difference*, pp. 77–113. Oxford: Blackwell Scientific Publications.
- Green, P. E. 1978 *Analyzing multivariate data*. Hinsdale, IL: The Dryden Press.
- Haffer, J. 1969 Speciation in Amazonian forest birds. *Science* **165**, 131–137.
- Haffer, J. 1997 Alternative models of vertebrate speciation in Amazonia: an overview. *Biodiv. Conserv.* **6**, 451–476.
- Hay, S. I. 1996 An investigation of the utility of remotely sensed meteorological satellite data for predicting the distribution and abundance of the tsetse fly (Diptera: Glossinidae). D.Phil. thesis, University of Oxford.
- Hay, S. I. 1997 Remote sensing and disease control: past, present and future. *Trans. R. Soc. Trop. Med. Hyg.* **91**, 105–106.
- Hay, S. I., Tucker, C. J., Rogers, D. J. & Packer, M. J. 1996 Remotely sensed surrogates of meteorological data for the study of the distribution and abundance of arthropod vectors of disease. *Ann. Trop. Med. Parasit.* **90**, 1–19.
- Hepinstall, J. A. & Sader, S. A. 1997 Using Bayesian statistics, Thematic Mapper satellite imagery, and breeding bird survey data to model bird species probability of occurrence in Maine. *Photogram. Eng. Remote Sensing* **63**, 1231–1237.
- Herr, A. M., & Queen, L. P. 1993 Crane habitat evaluation using GIS and remote sensing. *Photogram. Eng. Remote Sensing* **59**, 1531–1538.
- Holben, B. N. 1986 Characteristics of maximum-value composite images from temporal AVHRR data. *Int. J. Remote Sensing* **7**, 1417–1434.
- Homer, C. G., Edwards, T. C., Ramsey, R. D. & Price, K. P. 1993 Use of remote sensing methods in modelling sage grouse winter habitat. *J. Wildl. Mgmt* **57**, 78–84.
- ICBP 1992 *Putting biodiversity on the map: priority areas for global conservation*. Cambridge: International Council for Bird Preservation.
- Justice, C. O., Townshend, J. R. G., Holben, B. N., and Tucker, C. J. 1985 Analysis of the phenology of global vegetation using meteorological satellite data. *Int. J. Remote Sensing* **6**, 1271–1318.
- Kidwell, K. B. 1995 *NOAA polar orbiter data users guide (TIROS-N, NOAA-6, NOAA-7, NOAA-8, NOAA-9, NOAA-10, NOAA-11, NOAA-12, NOAA-13 and NOAA-14)*. Washington, DC: National Oceanic and Atmospheric Administration.
- Kingdon, J. 1990 *Island Africa. The evolution of Africa's rare animals and plants*. London: William Collins Sons & Co.
- Long, A. J., Crosby, M. J., Stattersfield, A. L. & Wege, D. C. 1996 Towards a global map of biodiversity: patterns in the distribution of restricted-range birds. *Global Ecol. Biogeogr. Lett.* **5**, 281–304.
- Los, S. O. 1993 Calibration adjustment of the NOAA AVHRR normalized difference vegetation index without recourse to component channel 1 and 2 data. *Int. J. Remote Sensing* **14**, 1907–1917.
- Mayr, E. & O'Hara, R. J. 1986 The biogeographic evidence supporting the Pleistocene forest refuge hypothesis. *Evolution* **40**, 55–67.
- Miller, K. V. & Conroy, M. J. 1990 SPOT satellite imagery for mapping Kirtland's warbler wintering habitat in the Bahamas. *Wildl. Soc. Bull.* **18**, 252–257.
- Moreau, R. E. 1966 *Bird faunas of Africa and its islands*. New York: Academic Press.
- Myneni, R. B., Maggion, S., Iaquinto, J., Privette, J. L., Gobron, N., Pinty, B., Kimes, D. S., Verstraete, M. M., & Williams, D. L. 1995 Optical remote sensing of vegetation: modeling, caveats, and algorithms. *Remote Sensing Environ.* **51**, 169–188.
- O'Connor, R. J., Jones, M. T., White, D., Hunsaker, C., Loveland, T., Jones, B. & Preston, E. 1996 Spatial partitioning of environmental correlates of avian biodiversity in the conterminous United States. *Biodiv. Lett.* **3**, 97–110.
- Peters, R. H. 1991 *A critique for ecology*. Cambridge University Press.
- Prell, W. L., Hutson, W. H., Williams, D. F., Bé, A. W. H., Geitzner, K. & Molino, B. 1980 Surface circulation of the Indian Ocean during the Last Glacial Maximum, approximately 18 000 yrs BP. *Quat. Res.* **14**, 309–336.
- Prigogine, A. 1987 Disjunctions of montane forest birds in the Afro-tropical region. *Bonner Zoologische Beiträge* **38**, 191–207.
- Rogers, D. J. & Williams, B. G. 1994 Tsetse distribution in Africa: seeing the wood and the trees. In *Large scale ecology and conservation biology* (ed. P. J. Edwards, R. M. May and N. R. Webb), pp. 249–273. Oxford: Blackwell Scientific Publications.
- Rogers, D. J., Hay, S. I. & Packer, M. J. 1996 Predicting the distribution of tsetse flies in West Africa using temporal Fourier processed meteorological satellite data. *Ann. Trop. Med. Parasit.* **90**, 225–241.
- Sellers, P. J. 1985 Canopy reflectance, photosynthesis and transpiration. *Int. J. Remote Sensing* **6**, 1335–1372.
- Snijders, F. L. 1991 Rainfall monitoring based on Meteosat data—a comparison of techniques applied to the Western Sahel. *Int. J. Remote Sensing* **12**, 1331–1347.

- Stattersfield, A. J., Crosby, M. J., Long, A. J. & Wege, D. C. 1998 *Endemic bird areas of the world: priorities for biodiversity conservation*. Cambridge: Bird Life International.
- Tatsuoka, M. M. 1971 *Multivariate analysis: techniques for educational and psychological research*. New York: John Wiley & Sons.
- Tucker, C. J., Townshend, J. R. G. & Goff, T. E. 1985 African land cover classification using satellite data. *Science* **227**, 369–375.
- Tuomisto, H. & Ruokolainen, K. 1997 The role of ecological knowledge in explaining biogeography and biodiversity in Amazonia. *Biodiv. Conserv.* **6**, 347–357.
- Williams, B., Rogers, D. J., Staton, G., Ripley, B. & Booth, T. 1992 Statistical modelling of georeferenced data: mapping tsetse distributions in Zimbabwe using climate and vegetation data. In *Modelling vector-borne and other parasitic diseases* (ed. B. D. Perry & J. W. Hansen), pp. 267–280. Nairobi: International Laboratory for the Research on Animal Diseases (ILRAD).
- Wright, S. J. & Hubbell, S. P. 1983 Stochastic extinction and reserve size: a focal species approach. *Oikos* **41**, 466–476.

As this paper exceeds the maximum length normally permitted, the authors have agreed to contribute to production costs.

