ORIGINAL ARTICLE

Predicting species distributions across the Amazonian and Andean regions using remote sensing data

Wolfgang Buermann[1]*, Sassan Saatchi[1,2], Thomas B. Smith[1,3], Brian R. Zutta[1,3], Jaime A. Chaves[1,3], Borja Milá[1,3] and Catherine H. Graham[4]

[1]*Center for Tropical Research, UCLA Institute of Environment, University of California, Los Angeles, CA, USA,* [2]*Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, USA,* [3]*Department of Ecology and Evolutionary Biology, University of California, Los Angeles, CA, USA,* [4]*Department of Ecology and Evolution, Stony Brook University, Stony Brook, NY, USA*

*Correspondence: Wolfgang Buermann, Center for Tropical Research, UCLA Institute of Environment, University of California, Los Angeles, La Kretz Hall, Suite 300, 619 Charles E. Young Dr. East, Los Angeles, CA 90095-1496, USA.
E-mail: buermann@ucla.edu

ABSTRACT

Aim We explore the utility of newly available optical and microwave remote sensing data from the Moderate Resolution Imaging Spectroradiometer (MODIS) and QuikSCAT (QSCAT) instruments for species distribution modelling at regional to continental scales. Using eight Neotropical species from three taxonomic groups, we assess the extent to which remote sensing data can improve predictions of their geographic distributions. For two bird species, we investigate the specific contributions of different types of remote sensing variables to the predictions and model accuracy at the regional scale, where the benefits of the MODIS and QSCAT satellite data are expected to be most significant.

Location South America, with a focus on the tropical and subtropical Andes and the Amazon Basin.

Methods Potential geographic distributions of eight species, namely two birds, two mammals and four trees, were modelled with the MAXENT algorithm at 1-km resolution over the South American continent using climatic and remote sensing data separately and combined. For each species and model scenario, we assess model performance by testing the agreement between observed and simulated distributions across all thresholds and, in the case of the two focal bird species, at selected thresholds.

Results Quantitative performance tests showed that models built with remote sensing and climatic layers in isolation performed well in predicting species distributions, suggesting that each of these data sets contains useful information. However, predictions created with a combination of remote sensing and climatic layers generally resulted in the best model performance across the three taxonomic groups. In Ecuador, the inclusion of remote sensing data was critical in resolving the known geographically isolated populations of the two focal bird species along the steep Amazonian–Andean elevational gradients. Within remote sensing subsets, microwave-based data were more important than optical data in the predictions of the two bird species.

Main conclusions Our results suggest that the newly available remote sensing data (MODIS and QSCAT) have considerable utility in modelling the contemporary geographical distributions of species at both regional and continental scales and in predicting range shifts as a result of large-scale land-use change.

Keywords

Conservation biogeography, ecological niche characterization, microwave remote sensing, MODIS, optical remote sensing, QSCAT, South America, spatial scale, species distribution modelling.

INTRODUCTION

A wide range of applications in biogeography, including evolutionary (Peterson, 2001; Hugall *et al.*, 2002; Graham *et al.*, 2004, 2006), ecological (Anderson *et al.*, 2002), invasive species (Peterson & Robins, 2003), conservation (Godown & Peterson, 2000; Sánchez-Cordero & Martínez-Meyer, 2000) and climate change studies (Berry *et al.*, 2002; Peterson *et al.*, 2002; Thomas *et al.*, 2004; but see also Buckley & Roughgarden, 2004; Thuiller *et al.*, 2004; Araújo *et al.*, 2006), rely on information concerning the geographical distribution of species and their environmental requirements. In species distribution modelling, which is often used in these studies, two sets of input data are required: occurrence data and spatial information on the environmental characteristics of the species habitats. Whereas significant work has been done on the complexity and robustness of the statistical methodology (e.g. Elith *et al.*, 2006) as well as on the quality and quantity of occurrence data (e.g. McPherson *et al.*, 2004; Segurado & Araújo, 2004; Hernandez *et al.*, 2006), relatively little attention has been paid to the environmental data sets used in ecological modelling. Here we explore the use of an array of new satellite data, which have seldom been used to predict species distributions and are often poorly understood by modellers.

The availability of satellite observations of the Earth's surface in the past three decades has heralded a new era in the ecological modelling of species habitats. Direct measurements of the Earth's hydrological and biophysical characteristics, its geological features and its climate from space have provided new data layers with spatial and temporal resolutions relevant to landscape-scale habitat characteristics and ecological processes (Gould, 2000; Nagendra, 2001; Turner *et al.*, 2003). The optical portion of the electromagnetic spectrum acquired from space-borne passive sensors, such as the Advanced Very High Resolution Radiometer (AVHRR), is most commonly used to infer vegetation characteristics. Green leaves harbour pigments, notably chlorophyll, that absorb strongly at blue and red wavelengths, whereas a lack of such absorption in the near-infrared results in strong scattering. A popular measure that captures this unique spectral response is the Normalized Difference Vegetation Index (NDVI); sometimes referred to as a measure of 'greenness', it is computed through the normalized difference in surface reflectances at near-infrared and red wavelengths. NDVI, or similar vegetation indices, are among the few satellite variables that have been used extensively in ecological niche modelling (e.g. Fuentes *et al.*, 2001; Osborne *et al.*, 2001; Luoto *et al.*, 2002; Zinner *et al.*, 2002; Parra *et al.*, 2004; Roura-Pascual *et al.*, 2006).

Through the recent launch of the NASA satellites Terra and Aqua with the Moderate Resolution Imaging Spectroradiometer (MODIS) on board, global optical measurements of vegetation at finer spatial resolution (250–500 m) with high revisit times (1–2 days) and improved data quality have become widely available (Justice *et al.*, 1998). In particular, the use of narrower spectral bands from the visible to middle infrared and advances in atmospheric corrections in the MODIS era have minimized the impact of water vapour and aerosol effects, which were problematic in previous satellite missions (Vermote *et al.*, 1997). Furthermore, improvements in retrieval techniques that translate the remotely sensed surface reflectances to more interpretable biophysical parameters have led to a number of advanced MODIS vegetation products. Among these, and applied in this study, are the leaf area index (LAI), defined as the one-sided projected green leaf area per unit ground area (Knyazikhin *et al.*, 1998; Myneni *et al.*, 2002), and the vegetation continuous field (Hansen *et al.*, 2002) products as measures of vegetation density, seasonality and percentage of tree cover. Although the application of optical satellite data over areas with persistent cloud cover remains a challenge, the frequent revisit time of MODIS at high spatial resolution allows circumvention of this basic physical limitation to some extent through appropriate spatial and temporal compositing.

In addition, microwave remote sensing data from active space-borne scatterometers [e.g. SeaWinds on board of NASA's QuikSCAT (QSCAT)] at relatively high spatial (2.25 km) and temporal (3 days) resolutions with global coverage are also becoming widely available. Microwave radiation probes the dielectric properties of land and, thus, is particularly sensitive to surface moisture. Furthermore, microwave radiation is sensitive to surface roughness, and hence its interpretation is somewhat more complex than that of optical data. However, recent studies have shown that the intensity of the radar backscatter over vegetated regions relates to biomass density and forest structure (Saatchi *et al.*, 2000, 2007; Long *et al.*, 2001). In addition, microwave measurements are less sensitive to cloud cover and penetrate deeper into the canopy, leading to a better discrimination of distinct forest architectures over densely forested regions. Taken together, the newly available optical and microwave remote sensing data layers provide additional information on surface parameters, such as vegetation density and seasonality as well as canopy moisture and roughness, with great potential to describe the ecology and geographical distributions of species.

In this study, we tested the utility of newly available remote sensing data for distribution modelling from regional to continental scales for a suite of species with different ecological characteristics. At these scales, the remote sensing data act as surrogates for land-cover variables and, hence, are closely related to climate variations (Egbert *et al.*, 2002). In total, we modelled the distributions of eight Neotropical species, namely two birds, two mammals and four trees, over the South American continent at 1-km spatial resolution. Because the usefulness of remote sensing data in distribution modelling may be dependent on a species' ecology (Hernandez *et al.*, 2006; McPherson & Jetz, 2007), these species were chosen to capture a variety of ecological traits such as range size and habitat specificity. Specifically, for widespread species in relatively homogeneous climates, such as throughout the Amazon basin, we expect remote sensing data to be particularly useful in providing additional constraints on the species ecological niches. In contrast, for Andean species that reside in

narrow climate regimes, the benefit of remote sensing data may be more limited. Furthermore, the utility of remote sensing data is likely to be scale-dependent, increasing in importance from continental scales where a broad species range is depicted to regional and landscape scales where vegetation and landscape features need to be incorporated.

Models were developed using bioclimatic data from the WorldClim data base (Hijmans *et al.*, 2005) and remote sensing data in isolation and combined, and tested by analysing the area under the receiver operating characteristic curve (AUC). For two bird species, the wedge-billed woodcreeper (*Glyphorynchus spirurus*) and the speckled hummingbird (*Adelomyia melanogenys*), with the former showing a very widespread distribution throughout the rainforests of the Amazon basin and Andean foothills and the latter a very narrow distribution in the Andean montane forests, we provide a detailed analysis on model performance. This includes omission tests at selected thresholds, evaluations of model predictive power as well as of the contributions of various environmental variables to the predictions, and visual inspections of the predicted maps. Given that the greatest benefit of the newly available optical and microwave remote sensing layers is expected at the regional level, we also provide an in-depth analysis of the predictions for the two focal bird species in Ecuador, a country with highly variable microclimates and ecological gradients along the steep slopes of the Andes.

DATA AND METHODS

Species and environmental data

Study species

For this study, point localities for eight Neotropical species were obtained. The species analysed were two common South American birds that were selected for further detailed analysis, the wedge-billed woodcreeper (*Glyphorynchus spirurus*) and the speckled hummingbird (*Adelomyia melanogenys*), as well as two Neotropical mammals and four Amazonian trees, which have been used in recent distribution modelling studies (Phillips *et al.*, 2006; Saatchi *et al.*, 2008). In the following, brief descriptions of ecological characteristics are given for each of the selected species, with more detail for the two focal bird species. Additional information about the six non-focal species is provided in the respective references.

Glyphorynchus spirurus is a widespread Neotropical bird species (length, 130–160 mm; mass, 10.5–21 g), which is frequent in the understorey of lowland and submontane tropical forests throughout the Amazon basin and Andean foothills (Del Hoyo *et al.*, 2003). There are 13 subspecies of *G. spirurus* known in the region, with ranges spanning from the southern Amazon basin to south-eastern Mexico, and along the Atlantic coastal forests of Brazil (Del Hoyo *et al.*, 2003).

Adelomyia melanogenys, one of the smaller species of South American hummingbirds (length, 85–90 mm; mass, 4.2–4.9 g), is found in wet to humid montane forests along the Andean slopes at elevations ranging from 1000 to 3000 m (Fjeldså & Krabbe, 1990; Del Hoyo *et al.*, 1999). Following the orientation of the Andes, this species occupies an impressive north-to-south range from northern Venezuela to north-western Argentina. Seven subspecies of *A. melanogenys* are recognized, occupying various regions of the northern Andes (Del Hoyo *et al.*, 1999).

The localities of four Amazonian tree species were taken from a recent distribution modelling study on Amazon tree diversity (Saatchi *et al.*, 2008). The species analysed were jacareuba (*Calophyllum brasiliensis*; 88 point localities), andiroba (*Carapa guianensis*; 73), balatá (*Manilkara bidentata*; 127), and ucuúba (*Virola surinamensis*; 113). The four trees differ markedly in habitat specificity and range size (Saatchi *et al.*, 2008). Amongst them, *C. brasiliensis* shows the widest distribution throughout northern South America and is found in the rain forests of the Amazon basin and the Andean lowlands up to elevations of 1500 m, as well as in the woodlands of south-eastern Brazil. *Carapa guianensis*, on the other hand, has the most specific habitat requirements and is found along streams, in the periodically inundated swamp forests, and in upland forests along the rivers of the Amazon Basin, and at elevations ranging from shorelines to 1200 m. *Manilkara bidentata* is quite common throughout the rain forests of the Amazon basin and along the coastal regions of northern Brazil and Guyana, but is restricted to lower elevations. *Virola surinamensis* is found in swampy, fertile and periodically inundated riverbanks, in Amazonian várzea forests, and in degraded and secondary forests. Compared with *M. bidentata*, its range is more restricted to the northern Amazon basin.

From Phillips *et al.* (2006), we obtained point localities for two Neotropical mammals: the brown-throated three-toed sloth, *Bradypus variegatus* (77 point localities), and a small montane murid rodent, *Microryzomys minutus* (85). *Bradypus variegatus* is widely distributed in the deciduous, evergreen and montane forests of the Amazon basin and in the vicinity of the tropical and subtropical Andes. Its geographic range extends from Honduras to northern Argentina. *Microryzomys minutus,* in contrast, shows a much narrower distribution throughout the tropical and subtropical Andean wet montane forests, with elevational ranges from 1000 to 4000 m and a geographic extent from Venezuela to Bolivia.

In the case of the two focal bird species, we obtained precise point-locality data by sampling each bird species using mist nets during multiple field surveys throughout Ecuador between 1999 and 2006 (22 sites for *G. spirurus*, 26 for *A. melanogenys*). In addition, we acquired occurrence data for these two bird species from other sources (see Table S1 in Supplementary Material) leading to a total of 78 unique georeferenced locations for *G. spirurus* and 180 for *A. melanogenys* (Fig. 1). In compiling this data set, we screened the records to include only georeferenced points obtained from recent surveys (within the last 10 years), thereby ensuring a broadly overlapping time frame with the acquisition of the

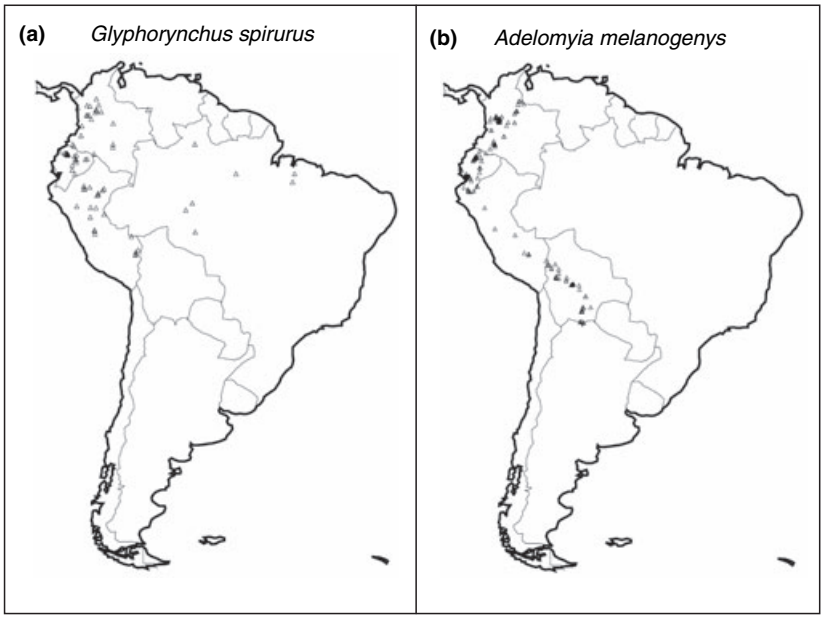

Figure 1 Point localities for (a) the wedge-billed woodcreeper (*Glyphorynchus spirurus*; 78 records) and (b) the speckled hummingbird (*Adelomyia melanogenys*; 170 records) utilized in this study.

remote sensing data. The modelling study was concentrated on the distribution of each of the two bird species pooled across subspecies because of the limited number of occurrences.

Climate data

A series of bioclimatic metrics were obtained from WorldClim version 1.4 (Hijmans *et al.*, 2005). These metrics are derived from monthly temperature and rainfall climatologies and represent biologically meaningful variables for characterizing species range (Nix, 1986). Eleven temperature and eight precipitation metrics were used, expressing spatial variations in annual means, seasonality and extreme or limiting climatic factors. The monthly climatologies were developed using long time series of a global network of weather stations from various sources, such as the Global Historical Climatology Network (GHCN), the United Nations Food and Agricultural Organization (FAO), the World Meteorological Organization (WMO), the International Center for Tropical Agriculture (CIAT), and additional country-based station networks. The station data were interpolated to monthly climate surfaces at 1-km spatial resolution using a thin-plate smoothing spline algorithm with latitude, longitude, and elevation as independent variables (Hijmans *et al.*, 2005).

Remote sensing data

Remote sensing data obtained from a variety of satellite sensors provided measurements of environmental variables directly related to species habitat characteristics (Turner *et al.*, 2003). We compiled a large data set covering a diverse range of surface parameters, namely vegetation density and seasonality, moisture, roughness, and topography.

To quantify spatial and temporal vegetation patterns, we used the MODIS 8-day LAI product (Myneni *et al.*, 2002) derived from atmospherically corrected MODIS surface reflectances (Vermote *et al.*, 1997) over the 5-year period 2000–2004. The accuracy of the MODIS LAI product has been evaluated against ground measurements of leaf area in a host of vegetation types across the globe (see online supporting information in Myneni *et al.*, 2007), and was generally found to be exact within LAI values of roughly 0.6. For each year, we created monthly composites by averaging the 8-day LAI product. Even though the MODIS algorithm is equipped with improved cloud masking (Platnick *et al.*, 2003), effects from subpixel cloudiness on LAI estimates over areas with persistent cloud cover, such as within the Amazon basin and along the Amazonian–Andean elevational gradients, were occasionally observed. To reduce these effects along with any natural interannual variability present in the data, we created monthly climatologies by averaging the 5 years of data (2000–2004). The climatological monthly composites were then used to generate five metrics: annual maximum (Fig. 2a), minimum, mean, standard deviation, and range (difference of maximum and minimum). These LAI metrics provide spatial information on net primary productivity and vegetation seasonality.

We also included the MODIS-derived vegetation continuous field (VCF) product as a measure of the percentage of tree canopy cover within each 500-m pixel (Hansen *et al.*, 2002). This data set was produced from time-series composites of MODIS data of the year 2001. Preliminary validation results over the conterminous United States (Hansen *et al.*, 2002) suggest that the VCF product can reliably separate more open (e.g. shrublands, savannas), deforested and, to some extent, fragmented areas from those of closed forests (Fig. 2b). For this study, the 500-m native tree cover data were aggregated to 1 km.

In addition to these optical remote sensing data, we included microwave QSCAT data available in 3-day composites at 2.25-km resolution (Long *et al.*, 2001). The 3-day data

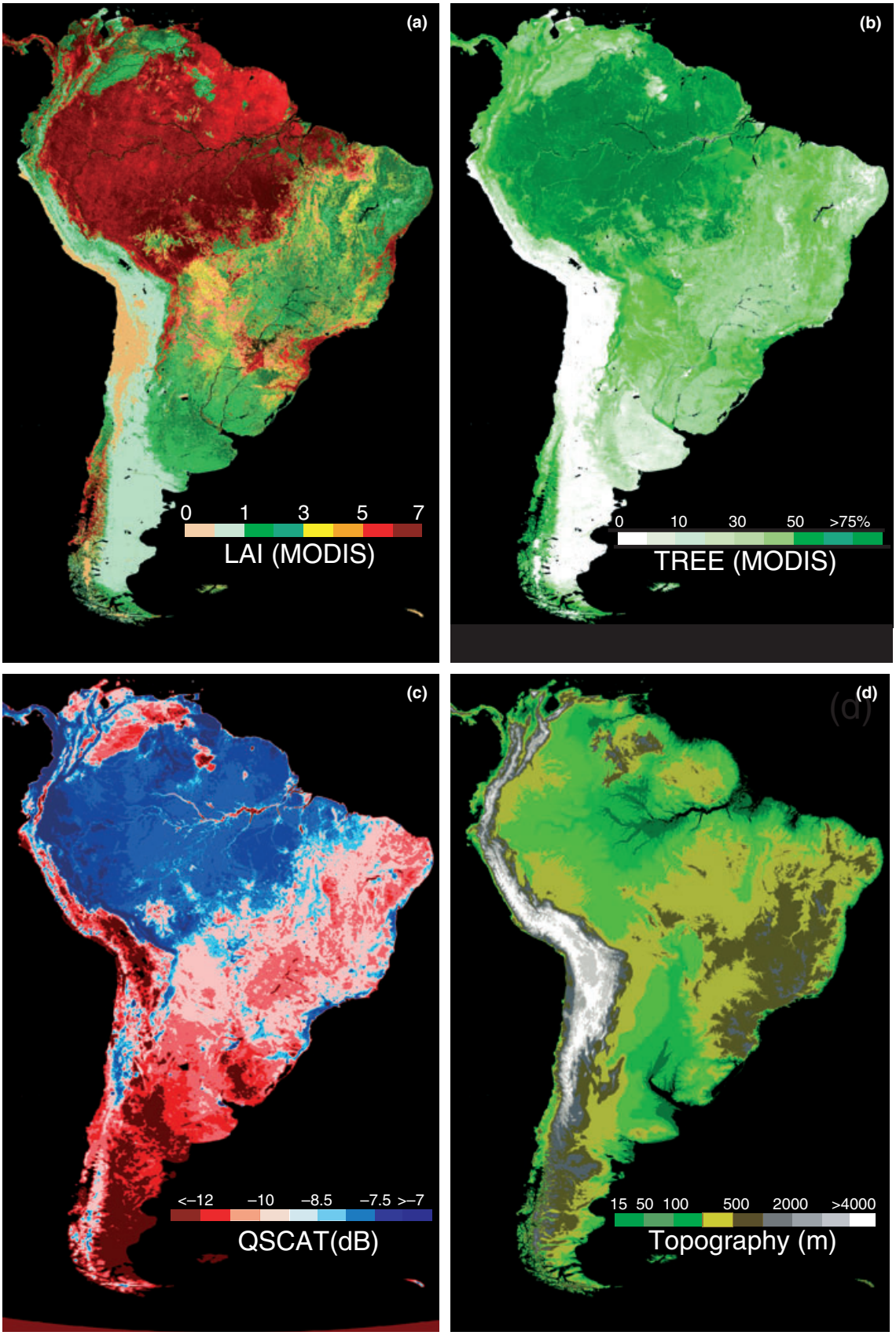

Figure 2 A selection of the remote sensing data layers used in this study. The panels depict (a) MODIS LAI (leaf area index) annual maximum, (b) MODIS percentage tree cover, (c) QSCAT annual mean, and (d) mean elevation from SRTM.

of the year 2001 with complete data coverage were used to create average monthly composites at 1-km resolution and then further processed to produce four metrics that included annual mean and standard deviation of radar backscatter with horizontal and vertical polarizations. QSCAT backscatter measurements are calibrated with high accuracy [*c.* 0.1 decibel (dB); Tsai *et al.*, 1999], and the long-term averages ensure the reduction of any possible high-frequency noise from compositing the data or from atmospheric disturbances (e.g. rain events) while preserving information on the backscatter variability. The QSCAT radar measurements, at wavelengths of *c.* 2 cm, are sensitive to surface canopy roughness, surface canopy moisture, and other seasonal attributes, such as deciduousness of vegetation (Frolking *et al.*, 2006). For low-density vegetation cover, such as woodlands, shrublands, and grassland savannas, measurements at various polarizations correlate positively with vegetation biomass and surface moisture (Saatchi *et al.*, 2007). For areas covered with dense forests, the QSCAT data are sensitive to large-scale variations in canopy roughness and moisture conditions (Fig. 2c).

Finally, we used the Shuttle Radar Topography Mission (SRTM) digital elevation data, aggregated from the native 90-m resolution to 1 km (Fig. 2d). In addition to mean elevation, the standard deviation based on the 90-m data was included as an indicator of surface ruggedness.

Overall, 10 continuous remote sensing data layers (5 LAI, 4 QSCAT, 1 VCF) and the two elevation data layers (mean and standard deviation) representing various vegetation and landscape features were included in this analysis. In Table 1, we present a brief overview of the original remote sensing products/data from which the various metrics were derived, along with corresponding ecological interpretations, native spatial resolution and orbiting frequency.

Data reduction

To facilitate interpretation of the environmental information and how it relates to a species' habitat, we reduced the total of 31 data layers (19 WorldClim + 10 remote sensing + 2 SRTM) to a set of less correlated variables. Covariance within the subsets of climate and remotely sensed variables was estimated by computing cross-correlation matrixes based both on the birds' point localities and on 1000 points randomly drawn from the northern half of South America in order to capture both covariance at local levels close to the bird vicinities and more general covariance over larger spatial scales. Various criteria were used to decide which layers of correlated pairs to retain for further analysis (with Pearson's correlations of the order of 0.9 or larger). These included keeping layers that are more commonly used in distribution modelling (WorldClim), that exhibit larger contrast/variance over the study area (QSCAT), and that have the highest data quality (LAI).

For the WorldClim metrics, significant correlations among a number of the original 19 data layers were observed, and a subset of nine bioclimatic variables was selected for further analysis, namely *annual mean temperature*, *mean diurnal temperature range*, *temperature seasonality*, *minimum temperature of coldest month*, *maximum temperature of warmest month*, *annual mean rainfall*, *rainfall seasonality*, *rainfall of coldest quarter* and *rainfall of warmest quarter*. High correlations were also evident among the five LAI metrics and the four QSCAT metrics that were produced for this study, and, based on the criteria given above, we selected the *LAI annual maximum* (Fig. 2a) and *LAI annual range* layers as well as the *QSCAT annual mean* (Fig. 2c) and *QSCAT seasonality* layers with horizontal polarization for further analysis.

As a result, the final reduced data set used in this study converged to a total of 16 environmental layers: nine climate (five temperature and four rainfall), two LAI, two QSCAT, two topography (mean and standard deviation), and one tree-cover layer. It should be noted that the model algorithm (MAXENT) used in this study is largely robust to covariance among variables, and that data reduction was performed mainly to improve interpretation.

Modelling approach

MAXENT

We used the MAXENT algorithm (version 2.1), which has been very recently introduced for the modelling of species distributions (Phillips *et al.*, 2006). MAXENT is a general-purpose algorithm that generates predictions or inferences from an incomplete set of information. The MAXENT approach is based on a probabilistic framework. The main assumption is that the incomplete empirical probability distribution (which is based

Table 1 Overview of the remote sensing data sets used in this study. For each remote sensing data layer, native spatial and temporal resolutions as well as an ecological interpretation are given along with the respective model scenarios in which these data layers were included. Model scenario acronyms refer to model runs with remote sensing and elevation data (RSE), climate and elevation data (CLIME), and climate, remote sensing and elevation data combined (CLIMERS).

Data record	Instrument	Ecological variable	Resolution	Scenario
Leaf area index (LAI)	MODIS	Vegetation density, seasonality and net primary productivity	1 km and 8 day	RSE, CLIMERS
Percentage tree cover	MODIS	Forest cover and heterogeneity	500 m	RSE, CLIMERS
Scatterometer-backscatter	QSCAT	Surface moisture and roughness (forest structure), seasonality	2.25 km and 3 day	RSE, CLIMERS
Digital elevation model	SRTM	Topography and ruggedness	90 m	RSE, CLIME CLIMERS

on the species occurrences) can be approximated by a probability distribution of maximum entropy (the MAXENT distribution) subject to certain environmental constraints, and that this distribution approximates a species' potential geographic distribution (Phillips *et al.*, 2006).

Like most maximum-likelihood estimation approaches, the MAXENT algorithm *a priori* assumes a uniform distribution and performs a number of iterations in which the weights associated with the environmental variables, or functions thereof, are adjusted to maximize the average probability of the point localities (also known as the average sample likelihood), expressed as the training gain (Phillips, 2006). These weights are then used to compute the MAXENT distribution over the entire geographic space. Consequently, this distribution expresses the suitability of each grid cell as a function of the environmental variables for that grid cell. A high value of the function (in units of cumulative probability) for a particular grid cell indicates that the grid cell is predicted to have suitable conditions for the species in question (Phillips, 2006).

MAXENT has several characteristics that make it highly suitable for the task of modelling species ranges (Phillips *et al.*, 2006). These include a deterministic framework; the ability to run with presence-only point occurrences; a high performance with few point localities (Hernandez *et al.*, 2006); better computing efficiency, enabling the use of large-scale high-resolution data layers; continuous output from least to most suitable conditions; and the ability to model complex responses through a number of distinct feature classes (e.g. functions of environmental variables). As a consequence, in a recent large model intercomparison project with 15 other algorithms, MAXENT's performance was generally rated among the highest (Elith *et al.*, 2006).

There are several aspects of MAXENT (2.1) that support the interpretation of the model results. For example, MAXENT has a built-in jackknife option through which the importance of individual environmental data layers can be estimated. It also provides response curves showing how the prediction depends on a particular environmental variable (Phillips, 2006). For all model runs in this study, we used the MAXENT default settings for regularization and in selecting the feature classes (functions of environmental variables). These include linear, quadratic, product, threshold and hinge features, depending on the number of point localities (Phillips, 2006).

Scenarios and quantitative analysis

To evaluate the merit of the set of newly available satellite data in species distribution modelling, we ran MAXENT with climate and satellite layers in isolation and combined. Three scenarios were evaluated: (1) MAXENT runs with climate and elevation data only (CLIME); these include the five temperature, four precipitation and two topography layers from the final reduced subset (see above), (2) remote sensing and elevation data only (RSE), with two QSCAT, two LAI, one tree-cover and two topography layers from the reduced subset, and (3) climate, elevation and remote sensing data combined (CLIMERS) using all 16 layers of the reduced subset (see *Data reduction*). For visualization, predicted maps were generated with the three predictor sets CLIME, RSE and CLIMERS using all available point localities of a given species. The corresponding maps for the six non-focal species (two mammals and four trees) are provided in Figs S1–S6 in the Supplementary Material.

The spatial accuracy of the predictions was tested following Phillips *et al.* (2006). For all eight species we created 10 random data partitions with 60% of the point localities assigned for training and 40% for testing, and ran the three scenarios with each of these 10 data partitions. Model performance was then tested across all thresholds (threshold-independent) and, in the case of the two focal bird species, at fixed thresholds (threshold-dependent).

In the threshold-dependent case, we evaluated extrinsic omission rates, defined as the fraction of test localities that fall into pixels outside the predicted area, both at constant proportional predicted area, defined as the fraction of all the pixels that are predicted as suitable for a species, and at the 'balance' threshold. Whereas the former allows a direct comparison of the extrinsic omission rates among the three model scenarios, the latter balances two measures of quality of a binary prediction, namely intrinsic (training) omission and proportional predicted area (Phillips, 2006). For all model scenarios and data partitions, we tested whether test points fell into areas predicted present more often than expected at random by applying a one-tailed binomial test on the extrinsic omission rate and proportional predicted area (Anderson *et al.*, 2002). In the constant proportional predicted area case, a two-tailed Wilcoxon signed rank test was utilized to evaluate statistical significance in differences in overall extrinsic omission rates between the CLIME and RSE scenarios. A one-tailed Wilcoxon signed rank test was utilized to test if decreases in overall extrinsic omission rates in the CLIMERS scenarios relative to those in the CLIME and RSE scenarios were statistically significant.

In the threshold-independent evaluation, we tested the performance of the MAXENT scenarios using receiver operating characteristic (ROC) analysis and pseudo-absence in place of absence localities (Phillips *et al.*, 2006). In this case, the AUC is typically used as a measure of model performance. We employed a ties-corrected one-tailed Mann–Whitney *U*-test (AccuROC; Vida, 1993) to test if a particular prediction in terms of AUC was significantly better than random. A two-tailed Wilcoxon signed rank test was used to test whether differences in overall AUC in the CLIME and RSE scenarios were significant. Finally, a one-tailed Wilcoxon signed rank test was also used to assess whether increases in overall AUC in the CLIMERS scenarios relative to those in the CLIME and RSE scenarios were statistically significant at levels $P < 0.005$, $P < 0.05$ and $P < 0.2$.

RESULTS

Quantitative results

Threshold-independent tests

For all eight Neotropical species in this study, AUC values generated using the extrinsic test data were highly statistically significant ($P < 0.001$, Mann–Whitney *U*-test) for all model scenarios and data partitions. This suggests that each predictor set alone contains useful information for modelling the species' distributions. AUC tests on model performance indicated that for six species scenarios with climate and elevation variables (CLIME) performed better than scenarios with remote sensing and elevation variables (RSE), whereas for two species these two scenarios performed equally well (Table 2). For seven species, however, scenarios with climate, remote sensing and elevation variables combined (CLIMERS) performed better than the corresponding scenarios with a subset of the predictor variables (Table 2). These results are suggestive of a generally improved discrimination of suitable vs. unsuitable habitats across taxonomic groups when all environmental data layers are used in combination.

To test whether increases in model performance in the CLIMERS scenarios were not simply an artifact of the larger number of environmental data layers used, we performed additional model runs with the full set of data layers within the climate subset (19 WorldClim and 2 elevation), the remote sensing subset (10 remote sensing + 2 elevation) and with all environmental layers (19 WorldClim + 10 remote sensing + 2 elevation). Model performance for scenarios with reduced and full sets of variables within the various predictor sets were generally stable, documenting that adding more variables with no significant new information (see also *Data reduction*) does not lead to improved model performance.

Threshold-dependent tests

For the two focal bird species in this study, model performance was also evaluated at two selected thresholds: the optimized 'balance' threshold and at constant proportional predicted area. In the three model scenarios, RSE, CLIME and CLIMERS, with 10 random data partitions, the 'balance' thresholds ranged from 0.5 to 6.0 (in units of cumulative probabilities) for *G. spirurus*, and from 1.5 to 4.5 for *A. melanogenys*. For *G. spirurus*, the binary maps corresponding to these threshold values were broadly consistent with its known distribution (Del Hoyo *et al.*, 2003), but for *A. melanogenys* these threshold values were somewhat too low, resulting in overprediction (Del Hoyo *et al.*, 1999). However, to facilitate statistical analysis in a consistent manner, we evaluated model performance at the 'balance' threshold for both bird species. At the 'balance' threshold, the binomial test on extrinsic omission rate and predicted area was highly significant ($P < 0.001$, one-tailed) for all scenarios and data partitions.

A more detailed look at extrinsic omission and predicted area showed that the combination of climate, remote sensing and elevation data resulted in more accurate and spatially more explicit predictions for both bird species (Table 3). For *G. spirurus*, the CLIMERS scenario achieved the lowest overall extrinsic omission rates with a relatively low overall predicted area. For *A. melanogenys*, the situation is somewhat different. Both the CLIME and RSE scenarios showed extrinsic omission rates lower than those in CLIMERS, but these lower omission rates were only realized at the cost of a notably larger overall predicted area. Independent evidence of overprediction in the

Table 2 Comparison of model performance across all thresholds for eight Neotropical species produced in the model scenarios RSE, CLIME, and CLIMERS (see Table 1). For each scenario, median AUC and range (brackets) calculated on extrinsic test data and based on 10 random data partitions are shown.

	Birds		**Mammals**	
	G. spirurus	*A. melanogenys*	*B. variegatus*	*M. minutus*
RSE	0.877 (0.861–0.896)	0.975 (0.970–0.980)	0.834 (0.790–0.873)	0.976 (0.971–0.981)
CLIME	**0.889** (0.865–0.909)	0.975 (0.956–0.983)	0.823 (0.779–0.875)	**0.985** (0.984–0.988)
CLIMERS	**0.898** (0.880–0.918)	**0.978** (0.971–0.984)	**0.888** (0.826–0.914)	0.984 (0.982–0.986)
	Trees			
	C. brasiliensis	*V. surinamensis*	*C. guianensis*	*M. bidentata*
RSE	0.736 (0.704–0.802)	0.805 (0.754–0.874)	0.849 (0.811–0.885)	0.837 (0.812–0.863)
CLIME	*0.759* (0.715–0.780)	**0.848** (0.799–0.881)	**0.900** (0.881–0.927)	**0.888** (0.867–0.909)
CLIMERS	**0.789** (0.755–0.829)	*0.851* (0.808–0.897)	*0.905* (0.869–0.927)	**0.900** (0.886–0.917)

To assess statistical significance in the comparison of model performance between the CLIME and RSE scenarios, a two-tailed Wilcoxon signed rank test was used, and median AUC values in underlined-bold ($P < 0.005$), bold ($P < 0.05$) and italics ($P < 0.2$) are indicative of a statistically significant better model performance in the corresponding scenario. A one-tailed Wilcoxon signed rank test was used to test whether increases in AUCs in the CLIMERS scenario relative to the CLIME and the RSE scenarios were statistically significant. In this case, median AUCs in the CLIMERS scenarios are only marked in underlined-bold ($P < 0.005$), bold ($P < 0.05$) and italics ($P < 0.2$) when differences for both pairwise comparisons (RSE-CLIMERS and CLIME-CLIMERS) are below the respective significance level.

Table 3 Comparison of model performance at selected thresholds for *G. spirurus* and *A. melanogenys*. The table shows (a) proportional predicted area and extrinsic omission rates at the 'balance' threshold, and (b) extrinsic omission rates at constant proportional predicted area produced in the model scenarios RSE, CLIME and CLIMERS. For each scenario, the median and range (brackets; only decimal shown) values based on 10 random data partitions are shown.

	G. spirurus		*A. melanogenys*	
	Area	Omission	Area	Omission
(a) 'Balance' threshold				
RSE	0.367 (200–419)	0.061 (032–129)	0.141 (102–171)	0.022 (014–042)
CLIME	0.458 (340–561)	0.071 (000–161)	0.123 (098–177)	0.029 (014–042)
CLIMERS	0.382 (317–450)	0.058 (000–097)	0.084 (068–118)	0.039 (014–056)
(b) Constant proportional predicted area				
RSE		0.126 (065–186)		0.144 (097–194)
CLIME		0.124 (065–194)		0.121 (097–194)
CLIMERS		0.109 (065–143)		0.089 (056–120)

For (b), the constant proportional predicted areas are 0.295 (5% threshold) for *G. spirurus* and 0.038 (10% threshold) for *A. melanogenys*, and stem from the CLIMERS scenarios with all point localities used.

RSE and CLIME scenarios for *A. melanogenys* comes from visible inspection at the continental scale, where large areas were predicted outside its known range (see below).

Because this omission test is highly sensitive to the proportional predicted area (Anderson *et al.*, 2003), we also compared extrinsic omission rates across the three scenarios at a constant proportional predicted area close to the observed ranges of the two bird species (Table 3). For *G. spirurus,* extrinsic omission rates in the CLIME and RSE scenarios were similar. For *A. melanogenys*, a smaller overall extrinsic omission rate in CLIME relative to RSE was evident, although it was not statistically significant ($P = 0.232$; two-tailed Wilcoxon signed rank test). Based on one-tailed Wilcoxon signed rank tests, the decreases in overall extrinsic omission rates from CLIME to CLIMERS ($P = 0.200$ for *G. spirurus* and $P = 0.010$ for *A. melanogenys*) as well as from RSE to CLIMERS ($P = 0.077$ for *G. spirurus* and $P = 0.002$ for *A. melanogenys*) were generally statistically significant.

Model predictive power, variable importance and response curves

To diagnose variations among the three scenarios in model predictive power for the two focal bird species in this study, we monitored the likelihood of the training and test point localities (training and test gain) under the MAXENT probability distribution (Phillips, 2006). For both species, this exercise indicated a stepwise increase in training and test gains from the RSE to CLIME to CLIMERS scenarios, respectively. In detail, the overall training/test gains (based on 10 random data partitions) increased for *G. spirurus* from 1.45/1.18 (RSE) to 1.53/1.41 (CLIME) to 1.85/1.64 (CLIMERS), and for *A. melanogenys* from 2.89/2.85 (RSE) to 3.00/2.90 (CLIME) to 3.34/2.99 (CLIMERS). These findings demonstrate the complementary information content of the remote sensing and climate data layers in describing the habitat characteristics of each species, resulting in a model with the highest predictive power.

The absolute and relative importance of individual environmental variables as predictors of the distributions of the two birds can be estimated through the training gains when the variable of interest is used in isolation and excluded from the whole set of variables in the MAXENT runs (Fig. 3). For *G. spirurus*, this test indicated that the data layers with the most useful information by themselves are the *QSCAT annual mean* followed by *rainfall of the warmest quarter*. The latter also appears to have the most information that is not shared with the other variables. For *A. melanogenys*, the layers with the most important information by themselves are the two *topography* layers (mean and standard deviation) followed by the *maximum temperature of the warmest month* and *temperature seasonality*, whereas *mean topography* and *QSCAT annual mean* seem to harbour a significant portion of information that is not contained in the other variables. *QSCAT annual mean* was also generally the most important remote sensing variable in the predictions of the additional six non-focal species (see Fig. S7). To test whether differences in native spatial resolutions of the environmental variables (see Table 1) introduce biases in comparisons of their predictive ability, we repeated the analysis with all environmental variables coarsened to the native 2.25-km QSCAT resolution. In this comparison, we found no significant changes in variable importance, suggesting that differences in native data grain size among variables play a minor role.

MAXENT also computes response curves showing how the predictions depend on the variables, which greatly facilitates the interpretation of a species ecological niche and its defining or limiting environmental factors. The responses of the most important environmental variables in the predictions for the two focal bird species in this study generally agree well with the corresponding sample histograms (Fig. 4). For *G. spirurus*, the ecological niche appears to be defined by *QSCAT annual mean* values of > -8.0 dB and *rainfall of warmest quarter* between roughly 500 and 1200 mm (Fig. 4a,b). Such environmental conditions are characteristic of dense humid forests, where the species is known to occur (Ridgely & Tudor, 1994; Del Hoyo *et al.*, 2003). *Adelomyia melanogenys*, on the other hand, appears to prefer mid-elevations and variable terrain typically extending from 1000 to 3000 m a.s.l. (Fig. 4c,d). In addition, this hummingbird species is adapted to a relatively narrow

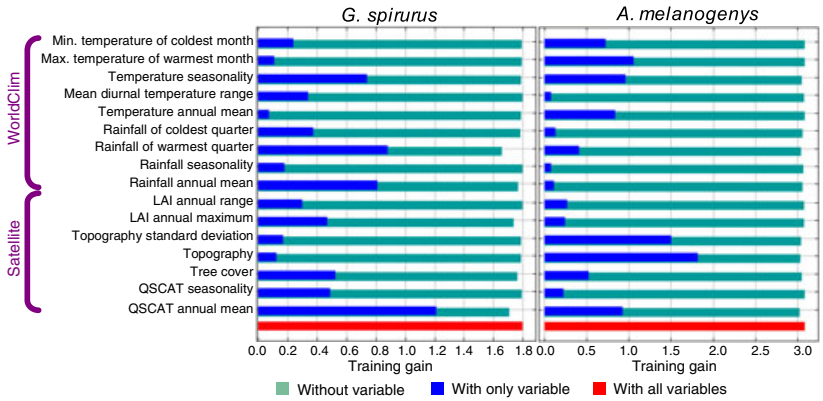

Figure 3 Results of a jackknife test of variable importance for *G. spirurus* and *A. melanogenys* in the CLIMERS scenario using all point localities. Bars represent training gain, a measure of the average likelihood of the point localities under the MAXENT probability distribution, when a particular variable is used in isolation, excluded, and when all variables are used in the predictions of the distributions of the two birds. Variables with highest gains when used in isolation contain the most useful information by themselves, whereas variables that lead to large decreases in gains when omitted from the predictor set contain information that is not present in the other variables (see also Phillips, 2006). Results of jackknife tests of variable importance are provided for the six non-focal species in Fig. S7.

maximum temperature of the warmest month range of 18–32°C, but can tolerate a wide range of *temperature seasonality* (Fig. 4e,f). A closer inspection suggests that this extended range in the response to *temperature seasonality* results from local adaptations among the populations of *A. melanogenys* in the Andes. In comparison with the tropical populations (diamonds, Fig. 4f), which reside in regions of low *temperature seasonality* (< 10°C), the extratropical populations at the southern margin (crosses, Fig. 4f) of the species distribution can tolerate much wider seasonal swings in temperature of up to 30°C.

Visual interpretations

Continental scale

The output of the MAXENT model consists of a continuous range with values ranging from 0 to 100 (in units of cumulative probabilities) indicating the least suitable to the most suitable conditions for the taxa under consideration (Phillips *et al.*, 2006). A visual inspection of the predicted potential geographic distributions of the two focal bird species in this study showed a broad agreement with their known distributions (Ridgely & Tudor, 1994; Del Hoyo *et al.*, 1999, 2003) when low thresholds of *c.* 2.5–5% (*G. spirurus*) and 10% (*A. melanogenys*) were applied (Fig. 5). The good match between known ranges and predicted distributions also suggests that the sampling was adequate for predicting the current distribution of the two bird species.

Predictions for *G. spirurus* covered a vast area in the northern half of the Southern American continent that closely followed the spatial extent of rain forest biomes, including those in the Amazon basin and the Andes lowlands (Fig. 5a–c). In addition, MAXENT correctly predicted the known disjunct population of *G. spirurus* in the Brazilian Atlantic Forest (Del Hoyo *et al.*, 2003). In contrast, *A. melanogenys* occupied a very restricted region along the western and eastern flanks of the Andes, and its north-to-south extent from northern Venezuela to northern Argentina was correctly identified (Fig. 5d-f).

A closer look at the predicted potential geographic distributions for *G. spirurus* and *A. melanogenys* produced with the three MAXENT scenarios revealed some distinct characteristics. For both bird species, all three scenarios agreed well in their spatial predictions of more suitable conditions (> 20% cumulative threshold). The CLIMERS scenario agreed best with the known ranges of the two bird species (Fig. 5c,f), whereas the CLIME and RSE scenarios predicted extensive but not coinciding areas outside the known ranges of each bird (Figs 5a,b and 5d,e for *G. spirurus* and *A. melanogenys*, respectively). Similar patterns of overpredictions were also evident in the CLIME and RSE scenarios for the additional two mammal and four tree species that were used in this study (see Figs S1–S6).

Ecuadorian Andes

For Ecuador, predictions for the two focal bird species with remote sensing data layers (RSE and CLIMERS) at plausible cumulative thresholds of 2.5–5% (*G. spirurus*) and 10% (*A. melanogenys*) were more in agreement with their known ranges (Fig. 6). In the case of *G. spirurus*, these predictions showed more fine-grained features than in the CLIME scenarios, in particular in the Central Coast region and the western Andean foothills (Fig. 6b–d), where large-scale continued deforestation has led to fragmented landscapes and habitat loss (Sierra *et al.*, 2002). For *A. melanogenys*, RSE and CLIMERS predictions showed narrower distributions along both the coast and the western Andean montane cloud forests than those in the CLIME scenario (Fig. 6b–d), again matching more consistently human-altered landscape patterns (Sierra

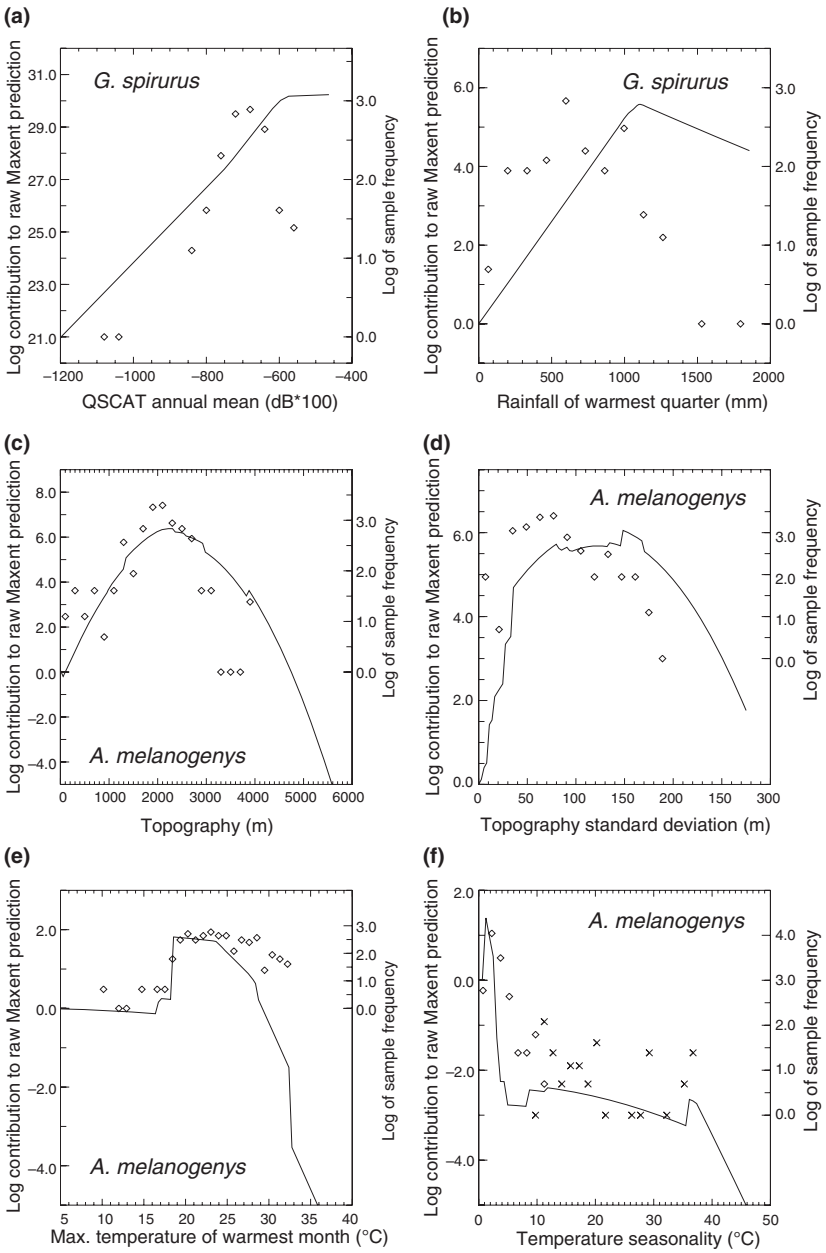

Figure 4 MAXENT response curves (lines) and sample histograms (diamonds) of a selected set of environmental variables for (a,b) *G. spirurus* and (c–f) *A. melanogenys*. The response curves illustrate how the contribution to the raw prediction depends on a particular environmental variable (Phillips, 2006). Because the MAXENT model is an exponential model, the probability assigned to a pixel is proportional to the exponential of an additive combination of the variables. The response curves above show the contribution to this exponent (*y*-axis) as a function of a particular environmental variable (*x*-axis). Maximum values in the response curves correspond to the highest predicted suitability. The response curves were derived from MAXENT runs using all point localities and the respective environmental variable in isolation, and, thus, do not include interactions with other environmental variables. The sample histograms show the number of point localities (frequency) that fall in a certain interval of the environmental variable under consideration. In (f), sample frequencies that correspond to Bolivian point localities are plotted as crosses.

et al., 2002). However, it must be noted that deforestation patterns at scales smaller than one kilometre were not captured with the remote sensing data used in this study. By contrast, the predictions with climate and elevation data (CLIME) did not reveal the disjunct populations of the two bird species along the eastern and western Andes, and in the case of *A. melanogenys* along the Chongón Colonche cordillera (Figs 6a,b). Smoother distributions in the CLIME scenarios were anticipated, as climatic gradients are less sharp than landscape gradients, in particular in regions with large-scale anthropogenic modifications of the landscape. In addition, the large uncertainties in the interpolated climate surfaces in mountainous and poorly sampled areas (Hijmans *et al.*, 2005), such as in the Ecuadorian Andes, severely limit their application over these areas.

DISCUSSION

Quantitative tests on model success with various predictor sets showed that models built with remote sensing and climatic data layers in isolation performed well in predicting the distributions of eight Neotropical species, suggesting that each of these data sets contains useful information. Models with climate and elevation data generally performed better across the three taxonomic groups than those using remote sensing and elevation data, which appears to be related to a tendency for increased overprediction with the latter predictor sets (see below). Overall, however, predictions created with a combination of remote sensing, climate and elevational variables generally performed best across all species. Effects of a species' ecology on the usefulness of remote sensing data in

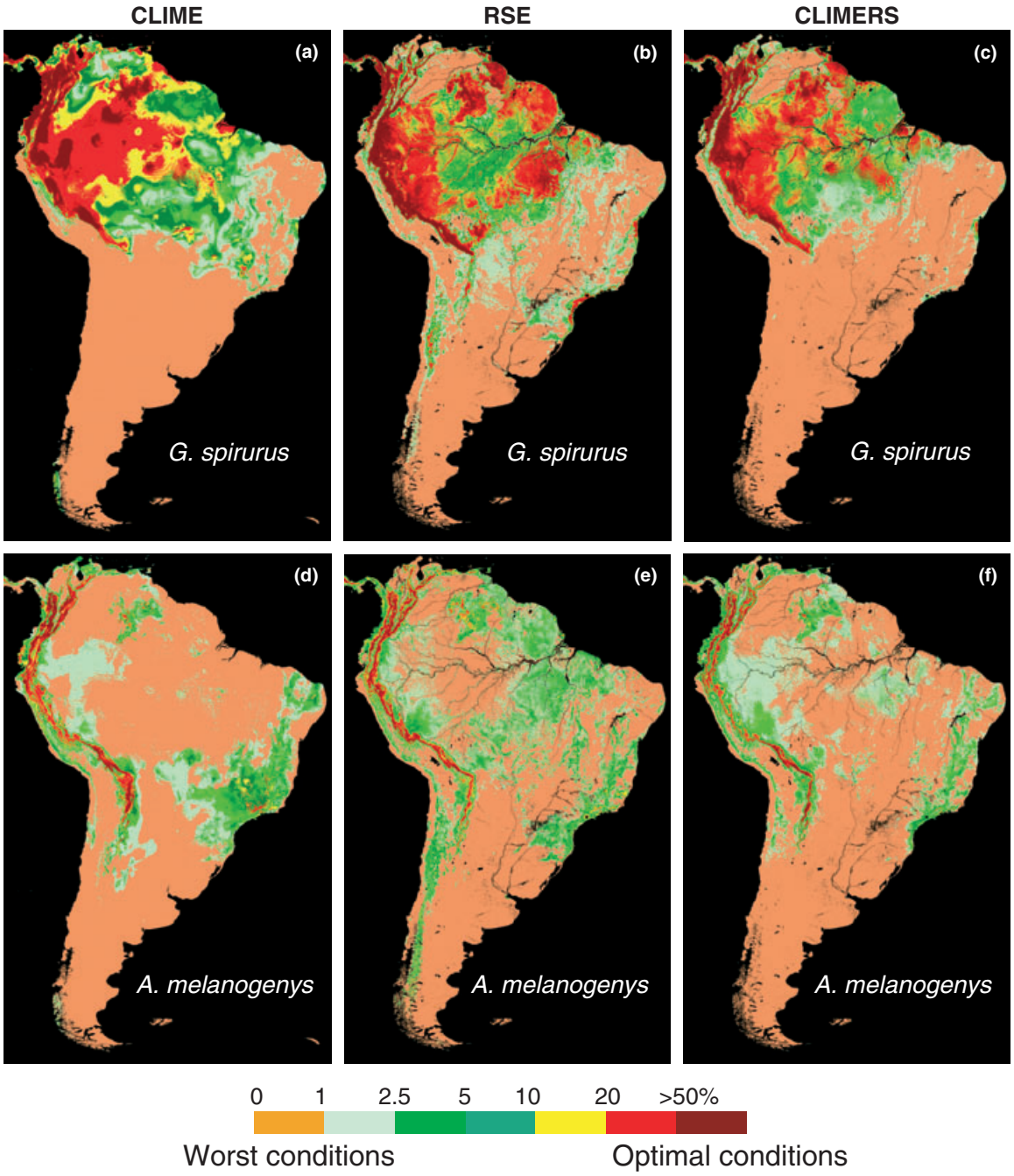

Figure 5 Predicted potential geographic distributions of the two focal bird species produced with the three predictor sets CLIME, RSE and CLIMERS using all point localities. The panels show the MAXENT predictions for *G. spirurus* and *A. melanogenys* in the (a,d) CLIME, (b,e) RSE and (c,f) CLIMERS scenarios, respectively. Corresponding MAXENT predictions with the three predictor sets are also provided for the additional two mammal and four tree species in supplementary Figs S1–S6.

distributional modelling were discernible. The addition of remote sensing data as predictor variables led to the largest improvement in model performance for species with relatively large range sizes and low habitat specificities (e.g. *C. brasiliensis* and *B. variegatus*). In this case, the inclusion of remote sensing data apparently leads to a much more constrained modelled ecological niche. This is probably because there is little variation in climate across the Amazon basin, so that these variables alone do not provide much predictive power. Remote sensing data reveal differences in forest structure in this homogeneous climate space, thereby providing valuable additional information to refine species distributions.

In contrast, for the two species with the most restricted range sizes and highest habitat sensitivities (*A. melanogenys* and *Mi. minutus*), the addition of remote sensing data yielded only marginal (*A. melanogenys*) or no (*M. minutus*)

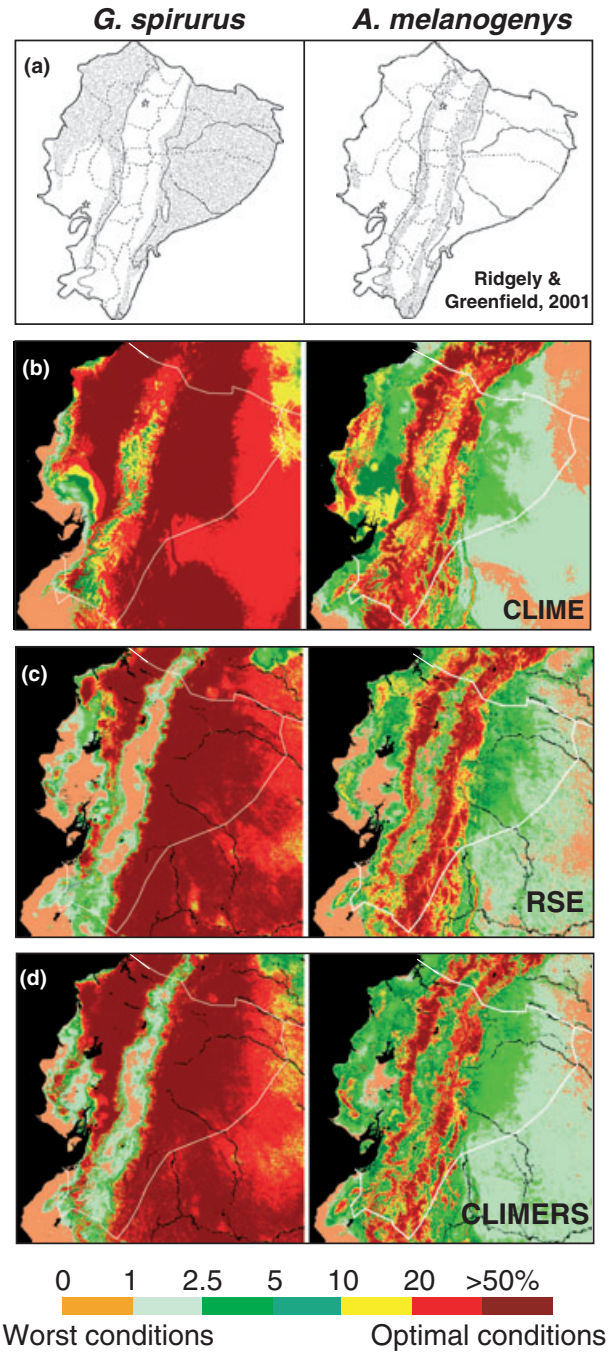

Figure 6 (a) Range maps of the two focal bird species in Ecuador and corresponding close-ups of the predicted potential geographic distributions in the three model scenarios (b) CLIME, (c) RSE and (d) CLIMERS. The range maps for *G. spirurus* and *A. melanogenys* were reproduced from Ridgely & Greenfield (2001). The predictions for *G. spirurus* (b–d *left panels*) and *A. melanogenys* (b–d *right panels*) are close-ups of the corresponding MAXENT predictions in Fig. 5.

improvements in model success. However, because of the overwhelming importance of elevational variables (see Figs 3 and S7) in the predictions of these two species, inferences concerning the merits of remote sensing data are limited.

Our findings emphasize the complementary information content in the climatic and remote sensing layers, leading generally to a more constrained modelled relationship of the ecological niche and more accurate spatial predictions. Similar results were obtained by Parra *et al.* (2004), who showed that a combination of climate and NDVI data sets improved predicted distributions of Andean bird species, although less significantly than in this study.

At continental scales, potential geographic distributions for the two focal bird species and for the additional two mammal and four tree species (see Figs S1–S6) created with only a subset of the available environmental information showed substantial prediction of area outside their observed ranges. Across the eight species, overprediction was generally most pronounced in model scenarios with remote sensing and elevation data. These patterns are not unexpected, as regions far distant from a species' point localities may reveal similar surface reflectance signatures in remotely sensed spectral space, even if they exist in different climate regimes. Hence, results of continental-scale species distribution modelling studies that rely entirely on remotely sensed measurements (e.g. Roura-Pascual *et al.*, 2006) should be interpreted cautiously. In the scenarios with climate data, overprediction may be a result of not capturing important characteristics of these species' ecological niches. Alternatively, overprediction in these scenarios may be a result of not capturing all limiting bioclimatic parameters with the required accuracy, or of artifacts resulting from interpolating climate within sparse station networks. In this context, it was reassuring that, when the full suite of the available environmental information was used, the corresponding predictions for the two focal bird species and the six additional species showed the least amount of overprediction and generally coincided best with their observed distributions.

Along the Amazonian–Andean elevational gradients, the two focal bird species occupy marginally overlapping habitats. *Glyphorynchus spirurus* is more frequent in the lowland and foothill tropical rain forests, where it forages chiefly for arthropods on the trunks of trees. *Adelomyia melanogenys*, in contrast, is found in the wet to humid montane forests at mid-elevations, where it feeds primarily on the nectar of flowering shrubs and epiphytes. With respect to variable importance and response curves, the results for the two bird species are consistent with these habitat characteristics. The geographic distribution of *G. spirurus* is broadly defined by precipitation patterns, with significant amounts of rainfall even in the warmest season typical for humid rain forests. In contrast, *A. melanogenys* has generally adapted to a flora in a relatively narrow hydro-thermal regime in the mid-elevations of the northern central Andes. For both forest-bird species, microwave-based QSCAT measurements of surface canopy moisture and roughness were important in the predictions. In comparison to optical measurements and derived variables (LAI, tree cover), the higher significance of QSCAT measurements could be the result of backscatter sensitivity to canopy roughness. In other words, optical measurements gradually lose their sensitivity to leaf area because of signal

saturation, whereas QSCAT backscatter is less saturated and can resolve the differences that may exist for dense forests with approximately similar LAI but different canopy structure. In comparison to precipitation, QSCAT measurements may also better capture the true moisture signature of the wet to humid montane forested habitats of *A. melanogenys*, as the 'horizontal precipitation' received at these locations in the form of mist and clouds off the ocean in addition to real rain contribute significantly to the overall moisture supply (Still *et al.*, 1999).

The Ecuador case study provided promising insights regarding the utility of the newly available satellite data at the regional scale. Because this study was conducted at 1-km spatial resolution close to the native or effective resolution of the remote sensing data layers, their full potential as surrogates for capturing microclimatic and land cover variations could be exploited. At these scales, the accuracy of the interpolated climate surfaces is more sensitive to the station network density, in particular over regions with highly variable terrains, such as the Ecuadorian Andes (Hijmans *et al.*, 2005). Predictions that incorporated remote sensing data could resolve the spatially isolated populations of the two focal bird species in the vicinity of the Ecuadorian Andes much more clearly. In addition, the inclusion of remote sensing data layers led to a sharper delineation of the predicted areas and better exclusion of areas that have suffered large-scale land-use and deforestation impacts.

In conclusion, the findings of this study suggest that the landscape, vegetation, and ecosystem attributes derived from, or present in, the applied remote sensing data contribute significantly to defining habitat characteristics, natural barriers and gradients that may exist even within similar climatic conditions. We expect that comparable results can be obtained using other species and algorithms. From the standpoint of conserving biodiversity, these results have significant implications. More accurate mapping of species distributions in biodiversity hotspots, such as along the steep Amazonian–Andean elevational gradients (Myers *et al.*, 2000; Bush, 2002), will improve the classification of areas with high habitat suitability that are threatened by both human land-use and climate change.

ACKNOWLEDGEMENTS

This work was performed at the UCLA Center for Tropical Research, the Jet Propulsion Laboratory, California Institute of Technology, and Stony Brook University with grant support from the National Aeronautic and Space Administration (IDS/03-0169-0347). Support for C.H.G. also came from a NASA New Investigator Award (34829). We would especially like to thank Woody Turner for his support and encouragement, and Ana Paula Giorgi for her assistance in obtaining point localities for the study species. We also would like to thank the subject editor and two anonymous referees for their comments, through which the manuscript has considerably improved.

REFERENCES

Anderson, R.P., Gomez-Laverde, M. & Peterson, A.T. (2002) Geographical distributions of spiny pocket mice in South America: insights from predictive models. *Global Ecology and Biogeography*, **11,** 131–141.

Anderson, R.P., Lew, D. & Peterson, A.T. (2003) Evaluating predictive models of species' distributions: criteria for selecting optimal models. *Ecological Modelling*, **162,** 211–232.

Araújo, M.B., Thuiller, W. & Pearson, R.G. (2006) Climate warming and the decline of amphibians and reptiles in Europe. *Journal of Biogeography*, **33,** 1712–1728.

Berry, P.M., Dawson, T.P., Harrison, P.A. & Pearson, R.G. (2002) Modelling potential impacts of climate change on the bioclimatic envelope of species in Britain and Ireland. *Global Ecology and Biogeography*, **11,** 453–462.

Buckley, L.B. & Roughgarden, J. (2004) Biodiversity conservation: effects of changes in climate and land use. *Nature*, **430,** doi:10.1038/nature02717.

Bush, M.B. (2002) Distributional change and conservation on the Andean flank: a palaeoecological perspective. *Global Ecology and Biogeography*, **11,** 463–473.

Del Hoyo, J., Elliott, A. & Sargatal, J. (1999) *Handbook of the birds of the world, vol. 5, barn-owls to hummingbirds*. Lynx Edicions, Barcelona.

Del Hoyo, J., Elliott, A. & Sargatal, J. (2003) *Handbook of the birds of the world, vol. 8, broadbills to tapaculos*. Lynx Edicions, Barcelona.

Egbert, S.L., Martínez-Meyer, E., Ortega-Huerta, M. & Peterson, A.T. (2002) Use of datasets derived from time-series AVHRR imagery as surrogates for land cover maps in predicting species' distributions. *IGARSS 2002: IEEE International Geoscience and Remote Sensing Symposium and 24th Canadian Symposium on Remote Sensing. Volumes I–VI. Proceedings—Remote Sensing: Integrating Our View of the Planet*, pp. 2337–2339. Institute of Electrical and Electronics Engineers, New York.

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

Fjeldså, J. & Krabbe, N. (1990) *Birds of the High Andes*. Zoological Museum and Apollo Books, Stenstrup, Denmark.

Frolking, S., Milliman, T., McDonald, K., Kimball, J., Zhao, M. & Fahnestock, M. (2006) Evaluation of the SeaWinds scatterometer for regional monitoring of vegetation phenology. *Journal of Geophysical Research*, **111,** doi: 10.29/2005JD006588.

Fuentes, M.V., Malone, J.B. & Mas-Coma, S. (2001) Validation of a mapping and prediction model for human fasciolosis transmission in Andean very high altitude endemic areas using remote sensing data. *Acta Tropica*, **79,** 87–95.

Godown, M.E. & Peterson, A.T. (2000) Preliminary distributional analysis of US endangered bird species. *Biodiversity and Conservation*, **9,** 1313–1322.

Gould, W. (2000) Remote sensing of vegetation, plant species richness, and regional biodiversity hotspots. *Ecological Applications*, **10,** 1861–1870.

Graham, C.H., Ron, S.R., Santos, J.C, Schneider, C.J. & Moritz, C. (2004) Integrating phylogenetics and environmental niche models to explore speciation mechanisms in Dendrobatid frogs. *Evolution*, **58,** 1781–1793.

Graham, C.H., Moritz, C. & Williams, S.E. (2006) Habitat history improves prediction of biodiversity in rainforest fauna. *Proceedings of the National Academy of Sciences USA*, **103,** 632–636.

Hansen, M.C., DeFries, R.S., Townshend, J.R.G., Sohlberg, R.A., Dimiceli, C. & Carroll, M. (2002) Towards an operational MODIS continuous field of percent tree cover algorithm: Examples using AVHRR and MODIS data. *Remote Sensing of the Environment*, **83,** 303–319.

Hernandez, P.A., Graham, C.H., Master, L.L. & Albert, D.L. (2006) The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography*, **29,** 773–785.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high-resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25,** 1965–1978.

Hugall, A., Moritz, C., Moussalli, A. & Stanisic, J. (2002) Reconciling paleodistribution models and comparative phylogeography in the Wet Tropics rainforest land snail *Gnarosophia bellendenkerensis* (Brazier 1875). *Proceedings of the National Academy of Sciences USA*, **99,** 6112–6117.

Justice, C.O., Vermote, E., Townshend, J.R.G., Defries, R., Roy, D.P., Hall, D.K., Salomonson, V.V., Privette, J.P., Riggs, G., Strahler, A., Lucht, W., Myneni, R.B., Knyazikhin, Y., Running, S.W., Nemani, R.R., Wan, Z., Huete, A., Leeuwen, W., Wolfe, R.E., Giglio, L., Muller, J.P., Lewis, P. & Barnsley, M.J. (1998) The moderate resolution imaging spectroradiometer (MODIS): land remote sensing for global research. *IEEE Transactions of Geosciences and Remote Sensing*, **36,** 1228–1249.

Knyazikhin, Y., Martonchik, J.V., Myneni, R.B., Diner, D.J. & Running, S.W. (1998) Synergistic algorithm for estimating vegetation canopy leaf area index and fraction of absorbed photosynthetically active radiation from MODIS and MISR data. *Journal of Geophysical Research*, **103,** 32257–32274.

Long, D.G., Drinkwater, M.R., Holt, B., Saatchi, S. & Bertoia, C. (2001) Global ice and land climate studies using scatterometer image data. *EOS Transactions AGU*, **82,** 503.

Luoto, M., Kuussaari, M. & Toivonen, T. (2002) Modelling butterfly distribution based on remote sensing data. *Journal of Biogeography*, **29,** 1027–1037.

McPherson, J.M. & Jetz, W. (2007) Effects of species' ecology on the accuracy of distribution models. *Ecography*, **30,** 135–151.

McPherson, J.M., Jetz, W. & Rogers, D.J. (2004) The effects of species' range sizes on the accuracy of distribution models: ecological phenomenon or statistical artifact? *Journal of Applied Ecology*, **41,** 811–823.

Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B. & Kent, J. (2000) Biodiversity hotspots for conservation priorities. *Nature*, **403,** 853–858.

Myneni, R.B., Hoffman, S., Knyazikhin, Y., Privette, J.L., Glassy, J., Tian, Y., Wang, Y., Song, X., Zhang, Y., Smith, G.R., Lotsch, A., Friedl, M., Morisette, J.T., Votava, P., Nemani, R.R. & Running, S.W. (2002) Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data. *Remote Sensing of the Environment*, **83,** 214–231.

Myneni, R.B., Yang, W., Nemani, R.R., Huete, A.R., Dickinson, R.E., Knyazikhin, Y., Didan, K., Fu, R., Robinson, I., Juarez, N., Saatch, S.S., Hashimoto, H., Ichii, K., Shabanov, N.V., Tan, B., Ratana, P., Privette, J.L., Morisette, J.T., Vermote, E.F., Roy, D.P., Wolfe, R.E., Friedl, M.A., Running, S.W., Votava, P., El-Saleous, N., Devadiga, S., Su, Y. & Salomonson, V.V (2007) Large seasonal changes in leaf area of Amazon rainforests. *Proceedings of the National Academy of Sciences USA*, **104,** 4820–4823.

Nagendra, H. (2001) Using remote sensing to assess biodiversity. *International Journal of Remote Sensing*, **22,** 2377–2400.

Nix, H. (1986) A biogeographic analysis of Australian elapid snakes. *Atlas of Elapid Snakes of Australia* (ed. by R. Longmore), pp. 4–15. Australian Government Publishing Service, Canberra.

Osborne, P.E., Alonso, J.C. & Bryant, R.G. (2001) Modelling landscape-scale habitat use using GIS and remote sensing: a case study with great bustards. *Journal of Applied Ecology*, **38,** 458–471.

Parra, J.L., Graham, C.H. & Freile, J.F. (2004) Evaluating alternative data sets for ecological niche models of birds in the Andes. *Ecography*, **27,** 350–360.

Peterson, A.T. (2001) Predicting species' geographic distributions based on ecological niche modeling. *Condor*, **103,** 599–605.

Peterson, A.T. & Robins, C.R. (2003) Using ecological-niche modeling to predict barred owl invasions with implications for spotted owl conservation. *Conservation Biology*, **17,** 1161–1165.

Peterson, A.T., Ortega-Huerta, M.A., Bartley, J., Sánchez-Cordero, V., Soberón, J., Buddemeier, R.W. & Stockwell, D.R.B. (2002) Future projections for Mexican faunas under global climate change scenarios. *Nature*, **416,** 626–629.

Phillips, S. (2006) *A brief tutorial on Maxent*. AT & T Research. Available at: http://www.cs.princeton.edu/~schapire/maxent/tutorial/tutorial.doc.

Phillips, S., Anderson, R.P. & Schapire, R.E. (2006) Maximum entropy modelling of species geographic distributions. *Ecological Modelling*, **190,** 231–259.

Platnick, S., King, M.D., Ackerman, S.A., Menzel, W.P., Baum, B.A., Riedi, J.C. & Frey, R.A. (2003) The MODIS cloud products: algorithms and examples from Terra. *IEEE Transactions on Geoscience and Remote Sensing*, **41,** 459–473.

Ridgely, R.S. & Greenfield, P.J. (2001) *The birds of Ecuador*. Cornell University Press, Ithaca, NY.

Ridgely, R.S. & Tudor, G. (1994) *The birds of South America. vol. II: the suboscine passerines*. University of Texas Press, Austin.

Roura-Pascual, N., Suarez, A.V, McNyset, K., Gómez, C., Pons, P., Touyama, Y., Wild, A., Gascon, F. & Peterson, A.T. (2006) Niche differentiation and fine-scale projections for Argentine ants based on remotely sensed data. *Ecological Applications*, **16,** 1832–1841.

Saatchi, S., Nelson, B., Podest, E. & Holt, J. (2000) Mapping land cover types in the Amazon Basin using 1 km JERS-1 mosaic. *International Journal of Remote Sensing*, **21,** 1201–1234.

Saatchi, S.S, Houghton, R.A., dos Santos, R.C., Avala, F., Soares, J.V. & Yu, Y. (2007) Distribution of aboveground live biomass in the Amazon basin. *Global Change Biology*, **13,** 816–837.

Saatchi, S., Buermann, W., ter Steege, H., Mori, S. & Smith, T. (2008) Modelling distribution of Amazonian tree species and diversity using remote sensing measurements. *Remote Sensing of the Environment*, doi:10.1016/j.rse.2008.01.008.

Sánchez-Cordero, V. & Martínez-Meyer, E. (2000) Museum specimen data predict crop damage by tropical rodents. *Proceedings of the National Academy of Sciences USA*, **97,** 7074–7077.

Segurado, P. & Araújo, M.B. (2004) An evaluation of methods for modelling species distributions. *Journal of Biogeography*, **31,** 1555–1568.

Sierra, R., Campos, F. & Chamberlin, J. (2002) Assessing biodiversity conservation priorities: ecosystem risk and representativeness in continental Ecuador. *Landscape and Urban Planning*, **59,** 95–110.

Still, C.J., Foster, P.J. & Schneider, S.H. (1999) Simulating the effects of climate change on tropical montane cloud forests. *Nature*, **398,** 608–610.

Thomas, C.D., Cameron, A., Green, R.E., Bakkenes, M., Beaumont, L.J., Collingham, Y.C., Erasmus, B.F.N., de Siqueira, M.F., Grainger, A., Hannah, L., Hughes, L., Huntley, B., van Jaarsveld, A.S., Midgley, G.F., Miles, L., Ortega-Huerta, M.A., Peterson, A.T., Phillips, O.L. & Williams, S.E. (2004) Extinction risk from climate change. *Nature*, **427,** 145–148.

Thuiller, W., Araújo, M.B., Pearson, R.G., Whittaker, R.J., Brotons, L. & Lavorel, S. (2004) Biodiversity conservation: Uncertainty in predictions of extinction risk. *Nature*, **430,** doi:10.1038/nature02716.

Tsai, W.Y., Graf, J.E., Winn, C., Huddleston, J.N., Dunbar, S., Freilich, M.H., Wentz, F.J., Long, D.G. & Jones, W.L. (1999) Post-launch sensor verification and calibration of the NASA scatterometer. *IEEE Transactions on Geoscience and Remote Sensing*, **37,** 1517–1542.

Turner, W., Spector, S., Gardiner, N., Fladeland, M., Sterling, E. & Steininger, M. (2003) Remote sensing for biodiversity science and conservation. *Trends in Ecology & Evolution*, **18,** 306–314.

Vermote, E.F., El Saleous, N.Z., Justice, C.O., Kaufman, Y.J., Privette, J., Remer, L., Roger, J.C. & Tanre, D. (1997) Atmospheric correction visible to middle infrared EOS-MODIS data over land surface, background, operational algorithm and validation. *Journal of Geophysical Research*, **102,** 17131–17141.

Vida, S. (1993) A computer program for non-parametric receiver operating characteristic analysis. *Computational Methods and Programs in Biomedicine*, **40,** 95–101.

Zinner, D., Peláez, F. & Tokler, F. (2002) Distribution and habitat of grivet monkeys (*Cercopithecus aethiops aethiops*) in eastern and central Eritrea. *African Journal of Ecology*, **40,** 151–158.

SUPPLEMENTARY MATERIAL

The following supplementary material is available for this article online:

Table S1 Source information on occurence data for the two focal bird species.

Figure S1 Point localities for *B. variegatus* and its predicted potential geographic distribution with the three predictor sets CLIME, RSE, and CLIMERS.

Figure S2 Point localities for *M. minutus* and its predicted potential geographic distribution with the three predictor sets CLIME, RSE, and CLIMERS.

Figure S3 Point localities for *C. brasiliensis* and its predicted potential geographic distributions with the three predictor sets CLIME, RSE, and CLIMERS.

Figure S4 Point localities for *C. guianensis* and its predicted potential geographic distribution with the three predictor sets CLIME, RSE, and CLIMERS.

Figure S5 Point localities for *M. bidentata* and its predicted potential geographic distribution with the three predictor sets CLIME, RSE, and CLIMERS.

Figure S6 Point localities for *V. surinamensis* and its predicted potential geographic distribution with the three predictor sets CLIME, RSE, and CLIMERS.

Figure S7 Results of a jackknife test of variable importance for the six non-focal species, two mammals (*B. variegatus* and *M. minutus*) and four trees (*C. brasiliensis*, *C. guianensis*, *M. bidentata* and *V. surinamensis*), in the CLIMERS scenario using all point localities.

This material is available as part of the online article from: http://www.blackwell-synergy.com/doi/abs/10.1111/j.1365-2699.2007.01858.x

Please note: Blackwell Publishing is 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

Wolfgang Buermann has a PhD from Boston University. Currently, he is a post-doctoral fellow in the Center for Tropical Research at the University of California, Los Angeles. His interests involve assessing the impact of land use and climate change on ecosystem functioning and species distribution, and assessing the merit of remote sensing data in ecological modelling.

Sassan Saatchi has a PhD from George Washington University. He is a senior researcher at the Jet Propulsion Laboratory, California, where he studies the physics of remote sensing and develops techniques for measuring vegetation and surface characteristics from space.

Thomas B. Smith has a PhD from the University of California, Berkeley. He is a professor in the Department of Ecology and Evolutionary Biology at the University of California, Los Angeles (UCLA), and is also the director of the Center for Tropical Research hosted at UCLA's Institute of the Environment. His research interests include topics in evolutionary ecology, speciation, and the conservation of vertebrates, especially in the tropics.

Editor: Walter Jetz