



Applying climatically associated species pools to the modelling of compositional change in tropical montane forests

Duncan J. Golicher^{1*}, Luis Cayuela², J. Rob M. Alkemade³,
Mario González-Espinosa¹ and Neptalí Ramírez-Marcial¹

¹Departamento de Ecología y Sistemática Terrestre, El Colegio de la Frontera Sur, Carretera Panamericana y Periférico Sur s/n, C.P. 29290, San Cristóbal de Las Casas, Chiapas, Mexico, ²Departamento de Ecología, Universidad de Alcalá, Carretera de Barcelona km. 33,600, C.P. 28871 Alcalá de Henares, Madrid, Spain, ³Netherlands Environmental Assessment Agency, Antonie van Leeuwenhoeklaan 9, 3721 MA, Bilthoven, The Netherlands

ABSTRACT

Aim Predictive species distribution modelling is a useful tool for extracting the maximum amount of information from biological collections and floristic inventories. However, in many tropical regions records are only available from a small number of sites. This can limit the application of predictive modelling, particularly in the case of rare and endangered species. We aim to address this problem by developing a methodology for defining and mapping species pools associated with climatic variables in order to investigate potential species turnover and regional species loss under climate change scenarios combined with anthropogenic disturbance.

Location The study covered an area of 6800 km² in the highlands of Chiapas, southern Mexico.

Methods We derived climatically associated species pools from floristic inventory data using multivariate analysis combined with spatially explicit discriminant analysis. We then produced predictive maps of the distribution of tree species pools using data derived from 451 inventory plots. After validating the predictive power of potential distributions against an independent historical data set consisting of 3105 botanical collections, we investigated potential changes in the distribution of tree species resulting from forest disturbance and climate change.

Results Two species pools, associated with moist and cool climatic conditions, were identified as being particularly threatened by both climate change and ongoing anthropogenic disturbance. A change in climate consistent with low-emission scenarios of general circulation models was shown to be sufficient to cause major changes in equilibrium forest composition within 50 years. The same species pools were also found to be suffering the fastest current rates of deforestation and internal forest disturbance. Disturbance and deforestation, in combination with climate change, threaten the regional distributions of five tree species listed as endangered by the IUCN. These include the endemic species *Magnolia sharpii* Miranda and *Wimmeria montana* Lundell. Eleven vulnerable species and 34 species requiring late successional conditions for their regeneration could also be threatened.

Main conclusions Climatically associated species pools can be derived from floristic inventory data available for tropical regions using methods based on multivariate analysis even when data limitations prevent effective application of individual species modelling. Potential consequences of climate change and anthropogenic disturbance on the species diversity of montane tropical forests in our study region are clearly demonstrated by the method.

Keywords

Chiapas, climate change, climatic envelopes, conservation, disturbance, diversity, species distribution modelling.

*Correspondence: Duncan J. Golicher, Departamento de Ecología y Sistemática Terrestre, El Colegio de la Frontera Sur, Carretera Panamericana y Periférico Sur s/n, C.P. 29290, San Cristóbal de Las Casas, Chiapas, Mexico.
E-mail: dgoliche@ecosur.mx

INTRODUCTION

In recent years considerable progress has been made in making information from biological collections and floristic inventories available in digital form. These data can make a key contribution to the shaping of conservation policy (Funk & Richardson, 2002). To do this, the maximum possible amount of relevant information must be extracted from them and combined with other information sources (Margules & Austin, 1994). Potential distributions for species can be inferred from such data using predictive distribution modelling. The most popular strategy has been to model distributions of individual species one at a time. However, the types of models used in a study can have dramatic effects on predicted range shifts and extinction rates (Elith *et al.*, 2006; Lawler *et al.*, 2006). It can be difficult to parameterize complex niche models for each species in tropical regions where the availability of data is still comparatively limited.

Spatial modelling of biodiversity at the community level can have significant benefits for applications involving large numbers of species, particularly when a sizeable proportion of these species is rarely recorded in the data set (Ferrier & Guisan, 2006). This approach is particularly promising for tropical regions where knowledge of empirical distributions of organisms is still limited. A variety of techniques aim to find the potential geographical distribution of organisms using computer software, a set of explanatory variables and a set of statistical models (Guisan & Zimmermann, 2000; Ferrier & Guisan, 2006). The resulting output is valuable to decision makers wishing to identify areas of high or low sensitivity with respect to change in key variables. The most important predictor variables selected are often climatic in nature. It is therefore natural to regard predictive modelling as a suitable tool for evaluating the impact of climate change or habitat loss on biological diversity (Kienast *et al.*, 1998; Iverson *et al.*, 1999; Bakkenes *et al.*, 2002; Pearson & Dawson, 2003; Thomas *et al.*, 2004).

In this study we develop and test a methodology for defining and mapping pools of tree species in relation to climatic predictors in order to evaluate the impact of anthropogenic disturbance and future climate change scenarios on forest diversity. Because rare tropical tree species can have too few recorded occurrences to allow their potential distribution to be accurately modelled on a species by species basis, we first ordinate and classify biological survey data, then model the resulting groups in relation to climatic variables. Although this is a common practice in community ecology, we have decided to avoid the term 'community model' (Ferrier & Guisan, 2006) as it implies that species co-occur in close enough proximity to allow biological interaction between them. We therefore propose the use of the term climatically associated species pools (CASPs) to define lists of species that are known to occur within the bounds of complex multidimensional climatic envelopes which, when mapped onto geographical space, do not necessarily form contiguous regions. Within the area covered by a CASP, finer-scale environmental variability and biotic interactions are also assumed to be acting. These determine the realized spatial and environmental proximity of species co-occurrences. In other words, CASPs are potential

species pools that are sampled from during the process of forming ecological communities at a local scale. Non-climatic variables, such as human disturbance or soil type, act as filters to reduce the actual number of species that may be found at any single site within a CASP. Microclimatic effects that cannot be accurately mapped at the scale at which the CASPs are defined may also be included in a set of site-specific filters later used in downscaling.

We illustrate the applicability of this approach with an example involving tree species data derived from 451 inventory plots in the highlands of Chiapas, Mexico. Suitable tools for biodiversity modelling are a priority for this biologically important region in which tree species diversity is known to be highly threatened by ongoing disturbance, deforestation and landscape-scale fragmentation (Ramírez-Marcial *et al.*, 2001; Cayuela *et al.*, 2006a,b). Regional species distribution is still not well known, and no study to date has attempted to predict the potential effect of climate change on the region's forests. We had three complementary objectives when designing our analytical framework: (1) to define baselines of species richness – an explicit requirement in order to parameterize models that predict the global consequences of human activities on ecosystems; (2) to evaluate the impacts of anthropogenic disturbance and assess potential threats to species by using remote sensing data upon a stratified CASP-based map; and (3) to assess the potential impact of climate change on tree diversity. Recent work suggests that tropical montane regions may be among the most affected by climate change (Still *et al.*, 1999; Pounds & Puschendorf, 2004). However, comparatively little fine-scale information exists on the effects of climate change in tropical montane regions, such as our study area, possibly because of difficulties in defining suitable and available inputs to models (Lips, 1998; Pounds *et al.*, 1999).

METHODS

Study site

The highlands of Chiapas is a biologically diverse region extending over 11,000 km² which includes 30% of about 9000 vascular plant species of the flora of Chiapas (Breedlove, 1986). Our study area consisted of a subset of this region limited to land over 1400 m a.s.l., resulting in the total area of 6800 km² shown in Fig. 1. The maximum elevation is 2901 m. Poor roads and abrupt topography, combined with political and social factors, complicate access to much of the area.

The soils include thin lithic rendzina and rather infertile chromic luvisols. The high elevation results in a cool climate, given the tropical latitude, with mean temperatures ranging from 14 to 17 °C. It is also comparatively humid with annual rainfall between 1000 and 1800 mm. Most rain falls between May and November with a 5–6 month dry season. This climate can be further subdivided along the elevation and rainfall gradients that we used in our analysis.

Several forest types have been identified in the region including oak, pine–oak, pine and evergreen cloud forests (Miranda, 1952; Rzedowski, 1978; Breedlove, 1981; González-Espinosa *et al.*, 1991). Traditional agricultural practices have produced a mosaic

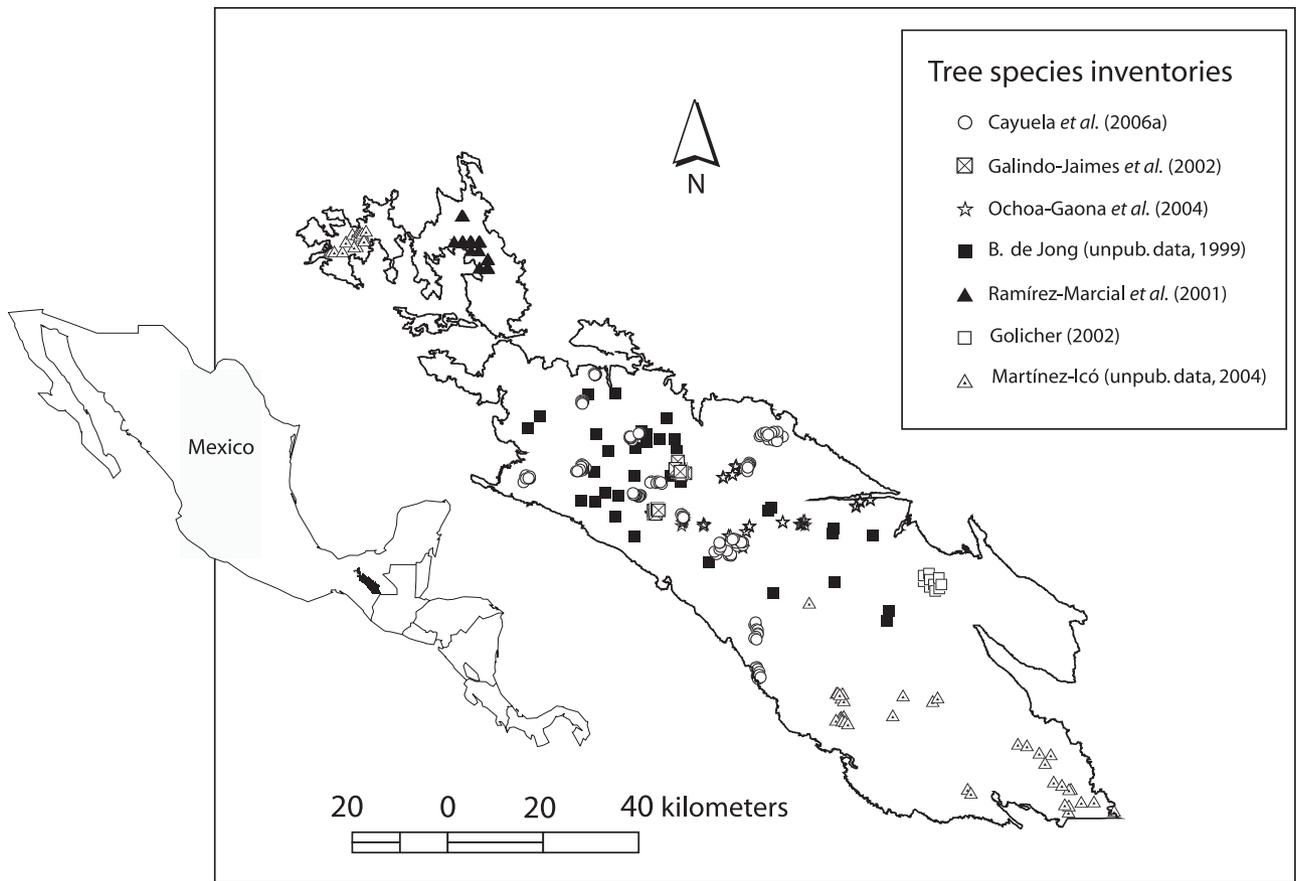


Figure 1 The study region in the highlands of Chiapas (land over 1400 m), showing the location of the tree species inventory data used for modelling. Data were collected by B. de Jong (unpublished data, 1999), Ramírez-Marcial *et al.* (2001), Galindo-Jaimes *et al.* (2002), Golicher (2002), Ochoa-Gaona *et al.* (2004), M. Martínez-Icó (unpublished data, 2004) and Cayuela *et al.* (2006a).

landscape of forest fragments embedded in a matrix of secondary vegetation and crop fields (Ramírez-Marcial *et al.*, 2001; Galindo-Jaimes *et al.*, 2002). Changes in land use, particularly in the last three decades, have accelerated deforestation and disturbance (Ochoa-Gaona & González-Espinosa, 2000; Cayuela *et al.*, 2006b). Almost all forest in the region has arisen from natural regeneration following long-term anthropogenic disturbance at a range of scales, with intact undisturbed mature forest being restricted to a few inaccessible fragments. Of the approximately 3000 km² of forest, only two small areas are under any form of protection: the national park of Lagunas de Montebello (60.22 km²) in the south-east, and the private reserve of Huitepec (1.36 km²) in the centre of the region. There are only a few, small, localized examples of forest plantations, making up most of the forested area that can be detected through remote sensing as potential habitat for native trees.

Data sources

Data on the occurrence of tree species were taken from a large set of floristic inventories compiled over a period of several years (Fig. 1). This resulted in a total of 451 plots at 71 separate localities within which a total of 4448 separate occurrences had been

recorded for 108 tree species. Unlike data from botanical collections, lack of occurrence could be interpreted as true absence. Only presence/absence was used, as a variety of methods and plot sizes had been used in collecting the data. The size of inventory plots ranged from 0.05 ha (500 m²) to 0.56 ha (5600 m²). Plots tended to be clustered within forest fragments. In a few cases, the distance between plots was as little as 100 m, although most were separated by more than 500 m. A range of differing plot sizes and spatially autocorrelated data are common when diverse studies are pooled. Thus, data of this type provide some challenges for robust statistical inference. This motivated the development of an appropriate methodology. Statistical analysis of ordination axes should a priori be less vulnerable to distortions produced by varying plot sizes than modelling on a species by species basis, because plots sharing some common species will receive similar scores even if small plot size reduces the probability of occurrence of rare species. Furthermore the pattern of ordination across varying plot sizes is more robust when presence/absence data are used (Otýpková & Chytrý, 2006).

We developed 36 raster coverages representing long-term mean minimum temperature, mean maximum temperature and mean rainfall for each month using 50 years of climate data obtained from the Comisión Nacional del Agua (CONAGUA).

The method of data collection used by CONAGUA is described by Golicher *et al.* (2006b). We used universal kriging to extrapolate these data onto a 1-km grid to produce raster coverages for monthly rainfall. In the case of temperature we combined the climate station data with a model of elevational lapse rates calculated for each month in order to convert the digital elevation model (downloaded from the Instituto Nacional de Estadística, Geografía e Informática website <http://mapserver.inegi.gob.mx/geografia/espanol/prodyserv/actualizacion/mde/descripcion.cfm>) into a temperature map. Temperature variability was modelled at a fine scale (30 m × 30 m pixel) and rainfall was resampled to the same scale. While this approach produced more local detail in the climatic maps than is available from global coverages (e.g. WORLDCLIM; Hijmans *et al.*, 2005), we point out that the potential effects of very fine-scale geomorphology on humidity, mist deposition and frost occurrence could not be included in the models at this stage. Because the overall climate of the region is comparatively homogeneous and highly elevation dependent, derived bioclimatic variables are very highly correlated.

Statistical methods for determining CASPs

We conducted most of our analysis using the R statistical language (R Development Core Team, 2005). The first step in the analysis was to identify groups of sites with similar species composition and similar climates. To achieve this we used canonical correspondence analysis (CCA) as implemented in the R package 'vegan' (Oksanen *et al.*, 2005).

We plotted site scores on the first two axes and then interpreted the results with respect to climate. We did this by generating smooth surfaces for climatic variables by fitting thin plate splines using general additive models (Wood, 2003) and interpolating the fitted values on the constrained ordination diagram (Oksanen *et al.*, 2005).

We then produced groups of sites that lay together on the first two ordination axes using K-means clustering (Forgy, 1965; Hartigan & Wong, 1979). The use of contours on the ordination diagram allowed us to identify and describe the climate envelope associated with the species list. A CASP was then defined quite simply as the list of species that were found in these clusters of similar sites. Note that this did not lead to mutually exclusive species lists as would have occurred if we clustered species on the ordination diagram.

A challenge we faced was deciding how many CASPs to produce before mapping the results onto geographical space. Errors in this choice are comparable to those resulting from over- and underfitting a generalized linear or rule-based species model. In the case of single-species models, overfitting suggests that a species is only found in a region that has a combination of variables coinciding almost exactly with those sites where the species has already been recorded. This also occurs if we use too many CASPs. In contrast, underfitting suggests that the species can be found over a broader range of environmental variation than that in the sites where it has been recorded. This may occur if very few CASPs are used. Initially we proposed the use of between three

and nine CASPs. For each of these groups, we looked at the difference in mean ranks of Bray–Curtis dissimilarity between groups compared with within-group dissimilarity using the ANOSIM R statistic (Clarke, 1993). The aim was to maximize between-group dissimilarity. This analysis was combined with a visual inspection of the ordination diagram and expert judgement.

The final step in our analysis was to project the CASPs onto space. We achieved this using flexible discriminant analysis as implemented in the R package 'mda' (Leisch *et al.*, 2005). A weighted Kappa statistic was used to test the accuracy of within-sample discrimination. The fitted model object was then used to project the results. In order to produce future scenarios we altered the input to the discriminant model to represent projected climates as outlined below.

Remotely sensed data

We used satellite imagery in order to look at the current distribution of forest and changes in forest cover over a 10-year period (1990–2000). We also used these data to determine the internal forest disturbance within the areas covered by each of the CASPs. The images consisted of a subset from three Landsat Thematic Mapper (TM) and Enhanced Thematic Mapper (ETM+) scenes with a resolution of 30 m (path 21 row 48, path 21 row 49, path 22 row 48) for 1990 and 2000, respectively. These images were geometrically, atmospherically and topographically corrected, as described by Cayuela *et al.* (2006b).

We calculated the normalized difference vegetation index (NDVI) from the red and near-infrared bands of satellite images for 1990 and 2000 (Tucker, 1979). The NDVI is sensitive to photosynthetically active biomass, and is correlated with leaf area index and net primary productivity (Cramer *et al.*, 1999). We then performed a supervised classification of forested and non-forested areas for each CASP in 1990 and 2000. Cross-validation with a previous classification of part of the study area using the same data (Cayuela *et al.*, 2006b) showed a high degree of consistency. Based on the resulting land-cover maps, we classified the forest within each CASP into open, closed or dense using NDVI- and CASP-based stratification. Within each of the CASPs, we standardized values of NDVI by subtracting the mean and dividing by the standard deviation using only forested pixels. Values below –1 SD were interpreted as open forest, between –1 and 1 as closed forest, and above 1 as dense forest. All these calculations were performed with Idrisi 14.0.

Climate change scenarios

Climate scenarios are plausible representations of the future that are consistent with assumptions about future emissions of greenhouse gases and other pollutants. There is a great deal of uncertainty surrounding the assumptions that determine the range of possible scenarios included in the IPCC SRES fourth assessment report (IPCC, 2007). These uncertainties are reflected in the results from general circulation models (GCMs), such as HadCM3 which combine emissions scenarios with assumptions regarding the physical basis of climate change.

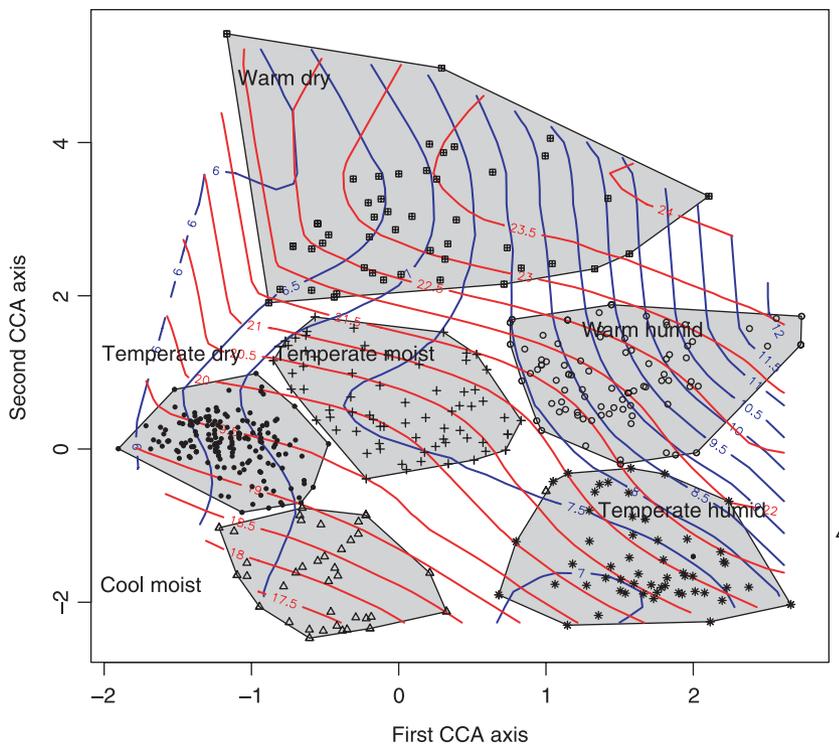


Figure 2 Canonical correspondence analysis plot displaying the K-means grouping of sites into six climatically associated species pools (CASPs). Red contours show November mean maximum temperature. Blue lines show June mean daily rainfall.

GCMs have a typical spatial resolution of $2.5^{\circ} \times 3.75^{\circ}$ (latitude by longitude). In our region this represents a surface spatial resolution of about $417 \text{ km} \times 278 \text{ km}$, which is much too coarse to be used directly in our models. We therefore downscaled three plausible future climates that we developed through investigating the results from three separate GCM runs that used the low-emission scenarios suggested by the IPCC. These were ECHAM4 of the Deutsches Klimarechenzentrum DKRZ in Germany, CGCM2 of the Canadian Centre for Climate Modeling and Analysis in Canada and HadCM3 of the Hadley Centre for Climate Prediction and Research in the UK. All these global models can be visualized at <http://lw.ipcc-data.org/cgi-bin/ddcvis/gcmcf>. Based on these models, we decided that an increase in mean temperature of around 1°C and a variation in mean annual rainfall of around 360 mm (either positive or negative) was plausible by 2050. Warming was downscaled using a locally derived elevational lapse rate of 5.1°C per 1000 m , which we obtained from fitting linear models to data from 180 climate stations. We interpreted changes in rainfall as change of approximately 20% either above or below present values, rather than an additive change over the surface. This led us to three credible scenarios: (1) warming of 1°C with no change in rainfall; (2) warming of 1°C with a 20% increase in rainfall; and (3) warming of 1°C with a 20% decrease in rainfall.

Model validation

We tested the predictions of the models against an independent data source provided by a data base of botanical collections (González-Espinosa *et al.*, 2004, 2005). A total of 3105 records were found within the study area for the tree species of interest.

We assumed this to be a sufficiently representative sample of the species pool associated with each CASP. Therefore, non-collected species were accepted as absences. Within each CASP, predicted and observed presences/absences were compared by means of contingency tables for all species that were present in both data sets. The Kappa index of agreement (KIA) was calculated for each CASP. A mean Kappa value was then obtained by weighting the individual Kappa statistics by the proportional area occupied by each CASP. We also used generalized linear models (logistic regression) and generalized additive models in order to produce predictive maps for all species having over five occurrences in the data set. The results obtained were matched against the maps produced by shading the CASPs in which the species occurred.

RESULTS

After analysing eigenvalues for a range of CCA models of differing complexities we selected a simple model that included winter rainfall, summer rainfall and a single temperature layer displaying the greatest variation (November maximum temperature). Figure 2 summarizes the results of the CCA. We divided the sites into six groups based on the use of the ANOSIM R statistic, visual inspection of the ordination diagram and expert judgement. The use of contours instead of the traditional arrows to represent vectors helped us to identify more clearly the gradients associated with each group of sites. Temperature and rainfall were almost independently related to each of the axes, the first axis being interpretable as a moisture gradient and the second axis as a temperature gradient. We noted that for all the models there was a slight increase in the eigenvalues of unconstrained axes when climate constraints were relaxed. We interpreted this to be due to

Table 1 Physical and biological characteristics of the six CASPs defined for the highlands of Chiapas, Mexico, based on 451 field inventories. From the original list of 125 tree species, only those that were present in more than one site were selected for ordination and determination of climatically associated species pools (CASPs) ($n = 108$), whereas the remaining were added subsequently as unique species ($n = 17$).

| CASP | Definition | Area (km ²) | Field inventories | Genera | Total species | Unique species | Threatened species | Successional status | | |
|------|-----------------|-------------------------|-------------------|--------|---------------|----------------|--------------------|---------------------|--------------|------|
| | | | | | | | | Early | Intermediate | Late |
| 1 | Temperate humid | 97 | 54 | 46 | 74 | 5 | 19 | 17 | 29 | 28 |
| 2 | Cool moist | 178 | 38 | 35 | 56 | 5 | 10 | 20 | 16 | 20 |
| 3 | Warm humid | 1234 | 89 | 40 | 67 | 6 | 18 | 27 | 28 | 12 |
| 4 | Temperate moist | 1822 | 69 | 48 | 79 | 15 | 17 | 33 | 27 | 19 |
| 5 | Temperate dry | 1498 | 173 | 41 | 67 | 10 | 13 | 27 | 21 | 19 |
| 6 | Warm dry | 1835 | 49 | 27 | 48 | 10 | 8 | 28 | 16 | 4 |

the additional effects of anthropogenic disturbance, soil types and management of the forests, which, if included in the model, would lead to slightly better ordination.

The predictors used in the discriminant analysis were November maximum temperature, June rainfall and January rainfall. The addition of two rainfall layers led to slightly better discrimination. The within-sample confusion matrix was inspected in order to verify the accuracy of discrimination. The weighted Kappa statistic (0.65) showed adequate within-sample discrimination given the large number of groups.

The information included in Table 1 demonstrates the special conservation value of CASP 1. A temperate humid forest type draws its composition from this species list. The associated species pool has the highest number of species (74) and the highest number of threatened species (19) and late successional species (28). This CASP is restricted to a potential area of less than 100 km², the smallest area of all the CASPs. CASP 2, which is associated with cooler but slightly drier conditions at higher elevation, is also of conservation concern due to its low potential area and high number of species that are vulnerable to disturbance. The detailed species by species information is included in Appendix S1 (see Supplementary Material).

Deforestation over a period of 10 years and forest disturbance, as measured by standardized NDVI within each CASP, are shown in Fig. 3. All areas except that predicted for CASP 4 have been subjected to recent deforestation. From a conservation perspective, the large loss in area for CASP 1 is of greatest immediate concern as it holds many vulnerable species. Our analysis suggests that the area covered by forest that draws its compositional components from this particular species pool has declined from 8100 hectares to just 4900 hectares in the period between 1990 and 2000. This is a loss of 40%. Dense forest in the same region also declined sharply. Our compositional analysis of CASP 2 also suggests that any further loss of potential area for this species pool must be considered highly worrying from a conservation perspective.

Figure 4 shows that the CASPs that are currently most threatened by deforestation and disturbance are also most likely to be affected by climate change. Again, CASP 1 is the most vulnerable, being predicted to lose area under any of the credible scenarios. Species turnover between CASP 3 and CASP 1 is extremely high, with the potential for 28 species to be lost from the fragmented

species-rich forests on mountain ridges in the north of the study area. Among the species that could be lost are *Abies guatemalensis* Rehder, *Clethra chiapensis* L.M. González (before *Clethra macrophylla*) and *Weinmannia pinnata* L. CASP 2 is also extremely vulnerable to any change that involves warming without increased rainfall. The relatively species-poor CASP 5, which replaces CASP 2, lacks *Cyathea fulva* (Mar. & Gal.) Fée, *Hedyosmum mexicanum* Cordemoy, *Meliosma dentata* (Liebmann) Urban, *Persea liebmannii* Mez, *Photinia microcarpa* Standley, *Podocarpus matudai* Lundell, *Quercus acutifolia* Née, *Quercus lancifolia* Cham. & Schldl., *Myrsine myricoides* (Schltdl.) Lundell and *Weinmannia pinnata* L. While some of these species are present in CASP 4, the potential area for this species list is also reduced to a small remnant in the south-west of the region.

The current distributions of all six CASPs were an acceptably accurate match to the distribution of the botanical collections, as assessed using the Kappa statistic (Table 2). Following the relative assessment scale proposed by Monserud & Leemans (1992), we found a good level of agreement between predicted and observed species pools in four out of six cases (Kappa between 0.55 and 0.7) and a very good level of agreement in two out of six cases (Kappa between 0.7 and 0.85). Overall, species presence was better predicted than species absence (Table 2). There was no correlation between the Kappa statistic and the area occupied by CASPs (Spearman's rank correlation = 0.2, $P = 0.714$). The mean weighted Kappa was 0.643, which indicates an overall good level of agreement between predicted and observed values (Monserud & Leemans, 1992). This analysis probably underestimates the predictive power of the models as we suspect that many discrepancies between predicted distribution and the positions recorded for the botanical collections can be attributed to inaccurate georeferencing of the collections themselves. The validation exercise may thus act in both directions as it may be necessary to re-evaluate the distribution of some species as suggested by collection data obtained before the availability of GPS. Comparisons of the results with collection data and individual species models fitted using logistic regression also validated our approach. For all but five of 108 species the CASP approach produced maps which matched the pattern of the distribution of collection points at least as well as those produced by logistic regression. In most cases the CASP approach produced a better match than

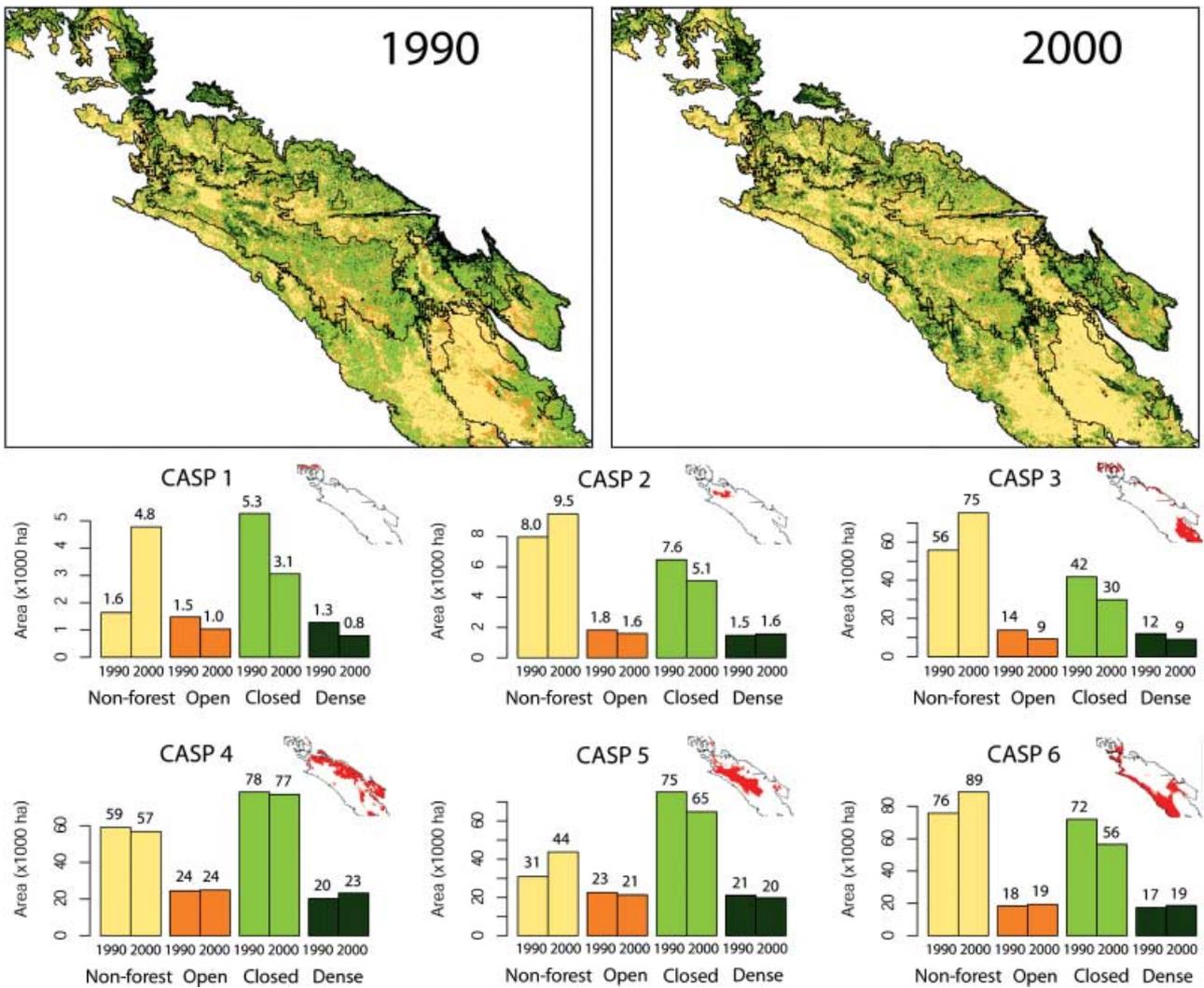


Figure 3 Land-cover map showing forested and non-forested areas in 1990 and 2000 (above). Forest was classified into open, closed and dense based on values of standardized NDVI within each CASP. Values of standardized NDVI below -1 were interpreted as open forest, between -1 and 1 as closed forest, and above 1 as dense forest. Bar plots (below) show the estimated area of the non-forest and each of the successional forest categories within each climatically associated species pool (CASPs).

logistic regression, and for rarer species logistic regression failed completely.

DISCUSSION

The applied value of CASPs

CASPs are empirical tools rather than theoretically derived models. CASPs are data driven and can change as new data become available. They can also be adapted to different spatial scales. For example, a detailed study along a single clearly defined transect through a climatic gradient could be used to produce a set of CASPs. A regional or global-scale study could be used to produce a different set of CASPs. At larger scales a more complex set of derived bioclimatic variables would be appropriate in order to include seasonal effects on temperature and precipitation. Differing sets of CASPs with alternative spatial resolutions can be

used to address different questions. As the occurrence of any single species is not necessarily restricted to a single CASP, the approach can be used to predict shifts in species composition that do not assume complete replacement of one mutually exclusive set of species by another under climate change scenarios. Other ecological and biogeographical approaches also apply a range of ecological concepts to the question of delimiting geographical boundaries to species distribution, such as ecoregions (Olson & Dinerstein, 1998, 2002; Kier *et al.*, 2005). These approaches are particularly valuable for setting conservation priorities. However, unlike CASPs, they are typically derived from expert judgement alone and may be only weakly coupled to empirical data sets (Jepson & Whittaker, 2002).

Although the CASP approach converts a continuum into discrete entities, our underlying assumption is that species do tend to show individualistic responses along climatic gradients. This is reflected both in our use of ordination along gradients as

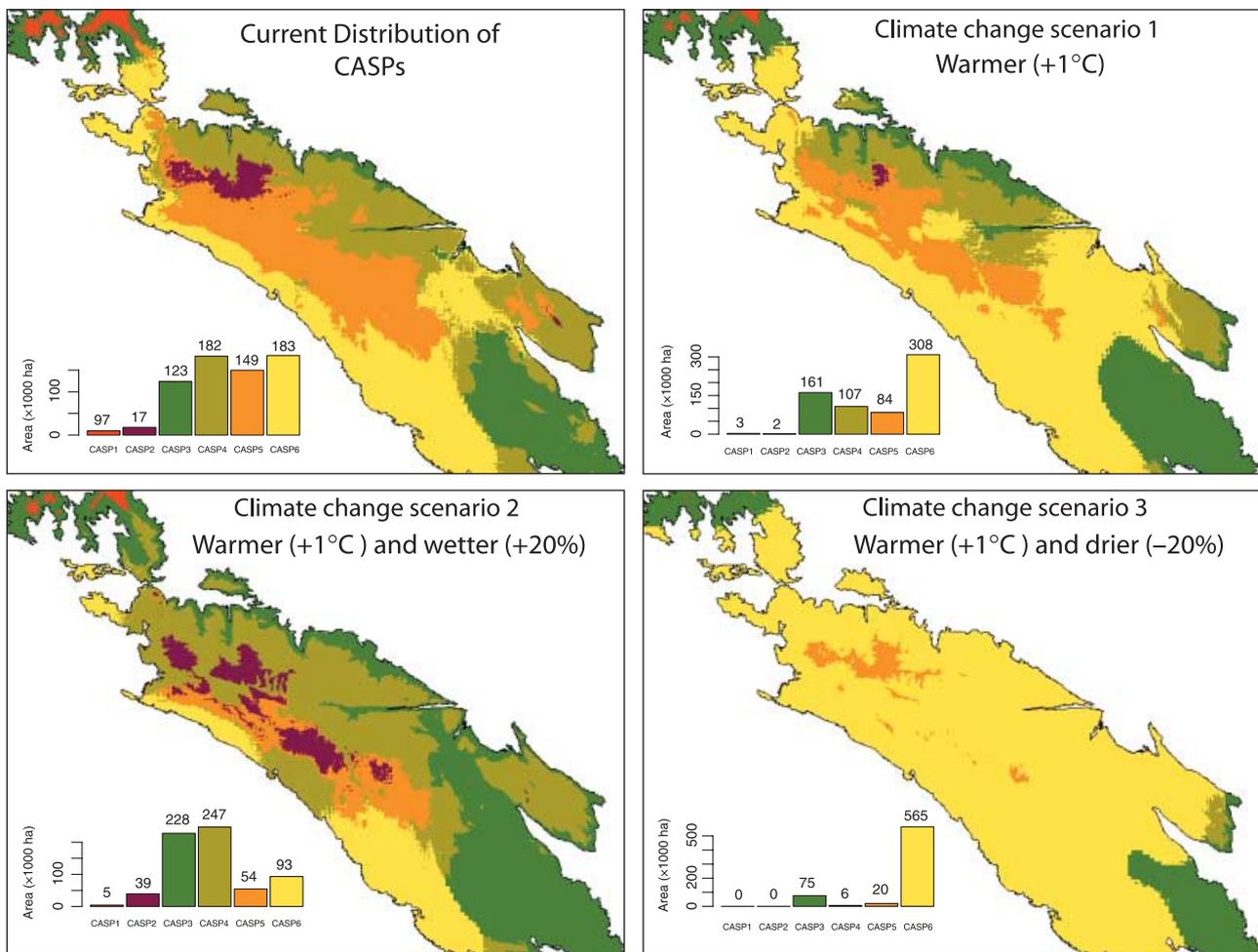


Figure 4 Predicted spatial extent of the areas in which species composition is drawn from each of the six climatically associated species pools (CASPs) under the present climate and three scenarios for 2050 based on the results of global circulation models.

a step in CASP modelling and the insistence that different sets of CASPs can, and indeed should, be defined and used for different purposes. We assume that any species grouping, including a CASP, is artificial. The species pool concept is a convenient device for structuring and communicating information regarding co-occurrence (Zobel *et al.*, 1998). Expert judgement is essential for evaluating the utility and completeness of any species pool. We use an automated, model-based approach to producing CASPs and their associated distribution maps in order to ensure transparency and repeatability. Criticism and evaluation of a CASP-based model with respect to stated aims is part of the modelling process.

The acceptable validation of the approach reinforced our belief in the utility of taking a pragmatic, although data-driven, approach to modelling species distributions. As pointed out by Guisan *et al.* (2006), 'data hungry' models such as the genetic algorithm for rule-set prediction (GARP), generalized additive models (GAMs), classification and regression trees (CARTs) and generalized linear models (GLMs) cannot be reliably fitted for species with fewer than 20 occurrences. In such a situation it has been suggested that climatic envelopes should be combined with

approaches that allow rare species to effectively 'borrow strength' from commoner species (Ferrier *et al.*, 2002; Edwards *et al.*, 2005). This is the basis of our approach. To date, relatively few successful applications of the individual species modelling approach have been published for rare and endangered plant species. Lack of data, spatial inaccuracy and lack of valid absences are the main reasons identified for this shortcoming (Engler *et al.*, 2004).

CASPs can be also used for the purpose of stratification. We demonstrated the applied value of CASPs in combination with remote sensing data to evaluate the impacts of human disturbance and assess potential threats to species changes. Previous work has pointed out the intrinsic challenges in separating climatic effects on tree species from those due to anthropogenic disturbance (Cayuela *et al.*, 2006a). This poses a major problem when using NDVI or any other vegetation index as an indicator of human disturbance, as rates of forest regrowth and the structural characteristics of the different successional forest stages may vary across regions due to differences in climatic or edaphic factors, thus rendering most relationships among spectral properties and forest conditions site-specific (Stomes & Hargrove, 2000; Vieira *et al.*, 2003). CASPs are a particularly useful tool in this context,

Table 2 Contingency tables and estimated Kappa statistic for agreement of predicted and observed species pools in the area covered by each climatically associated species pool (CASPs).

| Predicted CASP | | Botanical collections | | | Kappa |
|----------------|----------|-----------------------|---------|-------|-------|
| | | Presence | Absence | Total | |
| CASP 1 | Presence | 46 | 27 | 73 | 0.603 |
| | Absence | 6 | 29 | 35 | |
| | Total | 52 | 56 | 108 | |
| CASP 2 | Presence | 60 | 12 | 72 | 0.706 |
| | Absence | 20 | 16 | 36 | |
| | Total | 80 | 28 | 108 | |
| CASP 3 | Presence | 56 | 7 | 63 | 0.646 |
| | Absence | 25 | 20 | 45 | |
| | Total | 81 | 27 | 108 | |
| CASP 4 | Presence | 44 | 4 | 48 | 0.703 |
| | Absence | 36 | 24 | 60 | |
| | Total | 80 | 28 | 108 | |
| CASP 5 | Presence | 36 | 18 | 54 | 0.556 |
| | Absence | 12 | 42 | 54 | |
| | Total | 48 | 60 | 108 | |
| CASP 6 | Presence | 54 | 8 | 62 | 0.613 |
| | Absence | 23 | 23 | 46 | |
| | Total | 77 | 31 | 108 | |

as they allow us to hold for the climatic effect. Not only does this help to avoid confusing disturbed forest with a naturally drier more open forest type, but it also permits the effects of disturbance on species composition to be analysed. Both the synthetic results shown here and the more detailed information provided in Appendix S1 proved to be useful tools for analysing conservation priorities.

In addition, CASPs are a suitable source of input to some statistical models for comparing species diversity between sites, which require as input prior knowledge of the number of species expected to be found in a wider area in order to overcome problems associated with differing sample sizes (O'Hara, 2005; Golicher *et al.*, 2006a).

Our study was commissioned within a broader initiative aimed at providing input to the GLOBIO model (Alkemade *et al.*, 2006). GLOBIO is designed to improve understanding of the impact of human expansion on different ecosystems and regions through monitoring infrastructure expansion and climate change at a global level. Although further work is required to integrate results at this scale, the approach appears to be a promising means of providing synthetic input for large-scale models.

Caveats and limitations

Different species undoubtedly respond idiosyncratically to climate change, causing vegetation communities to disassemble as individual species shift their ranges in different directions (Thuiller, 2003; Thomas *et al.*, 2004). However, it should be noted that, because CASPs do not assume exclusivity of membership, they allow for some flexibility in this respect. Furthermore

actual plant communities are assumed to be compiled from subsets of the climatically associated pool of species.

At this stage we do not have integrated knowledge of tree migration rates into the climate change scenarios. Thus this part of the study can also be criticized on the grounds that migration from areas outside our study area is likely to occur. There are inevitable gaps in our knowledge that prevent accurate prediction of the immediate results of climate change on the distribution of tree species. However, we feel that this does not diminish the value of predictive models of potential distribution. Their value may well lie in an ability to clearly identify knowledge gaps in a changing world. Migration rates of trees, the time needed for vegetation to reach equilibrium or even quasi-equilibrium with climate and the ability of individual species or communities to adapt are all as yet unknown. Individuals of long-lived tree species may survive for many years or even centuries outside the limits of their true climatically determined potential distribution. In this case regeneration, rather than survival, is prevented by more extreme climatic conditions. When this occurs, human activities that remove reproductive trees from a landscape will act to accelerate the rate of change (Ramírez-Marcial *et al.*, 2005). If previous disturbance has led to a fragmented landscape over which the dispersal of some tree species is slow or impossible, then a predicted discrepancy between observed forest composition and predicted forest composition as represented by a potential distribution model suggests an unusually high rate of species loss through human activity. This may result in a loss of the ability of native forests to provide essential ecological services.

Implications for conservation

The implications of this case study for the purposes of regional conservation are clear. The remaining forests within the cool humid area predicted to be suitable for CASP 1 and CASP 2 should receive special conservation status. Deforestation and disturbance pose an immediate and ongoing threat to the area suitable for occupation by species in these lists. Any further disturbance of this area is expected to result in rapid species loss.

The longer-term implications of climate change present particular challenges for conservation. One of the most notable results is that the areas currently potentially occupied by the species in CASP 2 and CASP 4 would undergo major species turnover under a moderately warmer and drier climate change scenario. Although most of the species have ranges that extend beyond the study region, *Magnolia sharpii* and *Wimmeria montana* are believed, on current evidence, to be narrowly endemic. Both could potentially become extinct if conservation measures are not taken. Even species with a broader range would be likely to be threatened by climate change as a similar reduction in their potential occupied area is likely elsewhere. Most of the cool moist conditions in southern Mexico and Central America are restricted to small areas on and around mountain peaks and ridges.

We do not consider fatalism in the face of inevitable loss to be a suitable response to these results. Appendix S1 shows that 34

out of 42 species suggested to be threatened at a regional scale prefer late successional conditions. Species turnover is almost certainly accelerated by anthropogenic disturbance. Therefore we recommend that refuge areas with cool, moist microclimates should be specifically targeted for conservation in order to slow regional species loss. Directed conservation could potentially ameliorate the short- to medium-term effects of climate change on regional species richness. The large area occupied by CASP 4 shares many species with CASPs 1 and 2 and is slightly less vulnerable to climate change. We suggest that attention should be given to the task of locating and protecting the remaining species-rich forests within this area.

Modelling potential distributions is particularly important for conservation in our study region, which is not unusual in having a large proportion of the biodiversity outside protected areas. In our data set only 58 species of the total of 108 were found in the 42 plots within the two protected areas. Information on the distribution and abundance of species outside these reserves is patchy at best. Many of the forests of the region are not easily accessible. This is a common situation in the tropics. The CASP methodology was found to be a useful tool for addressing this challenge and to help guide targeted conservation efforts at a regional scale.

ACKNOWLEDGEMENTS

The analytical component of this work was financed by the Netherlands Environmental Assessment Agency (MNP) in the context of the Biodiversity International project (E/555050). A substantial part of the data collection was financed by the European Commission BIOCORES project INCO contract ICA4-CT-2001-10095. L.C. was partly financed by the European Commission ALFA PROJECT (contract II-0411-FA-FCD-FI-FC). We are grateful to Miguel Martínez Icó for his work in identifying tree species both in the field and the herbarium.

REFERENCES

- Alkemade, R., Bakkenes, M., Bobbink, R., Miles, L., Nellemann, C., Simons, H. & Tekelenburg, T. (2006) GLOBIO 3: Framework for the assessment of global terrestrial biodiversity. MNP (2006). *Integrated modelling of global environmental change. An overview of IMAGE 2.4* (ed. by A.F. Bouwman, T. Kram and K. Klein Goldewijk), pp. 171–186. MNP, Bilthoven, The Netherlands.
- Bakkenes, M., Alkemade, J.R.M., Ihle, F., Leemans, R. & Latour, J.B. (2002) Assessing effects of forecasted climate change on the diversity and distribution of European higher plants for 2050. *Global Change Biology*, **8**, 390–407.
- Breedlove, D.E. (1981) *Flora of Chiapas. Part 1: introduction to the Flora of Chiapas*. California Academy of Sciences, San Francisco, CA.
- Breedlove, D.E. (1986) *Listados florísticos de México IV. Flora de Chiapas*. Universidad Nacional Autónoma de México, México, DF.
- Cayuela, L., Golicher, J.D., Rey Benayas, J.M., González-Espinosa, M. & Ramírez-Marcial, N. (2006a) Fragmentation, disturbance and tree diversity conservation in tropical montane forests. *Journal of Applied Ecology*, **43**, 1172–1181.
- Cayuela, L., Rey Benayas, J.M. & Echeverría, C. (2006b) Clearance and fragmentation of tropical montane forests in the Highlands of Chiapas, Mexico (1975–2000). *Forest Ecology and Management*, **226**, 208–218.
- Clarke, K.R. (1993) Non-parametric multivariate analysis of changes in community structure. *Australian Journal of Ecology*, **18**, 117–143.
- Cramer, W., Kicklighter, D.W., Bondeau, A., Moore, B., Churkina, C., Nemry, B., Ruimy, A. & Schloss, A.L. (1999) Comparing global models of terrestrial net primary productivity (NPP): overview and key results. *Global Change Biology*, **5**, 1–15.
- DOF (Diario Oficial de la Federación) (2002) *Norma Oficial Mexicana NOM-059-ECOL-2001. Protección ambiental-Especies nativas de México de flora y fauna silvestres-Categorías de riesgo*, Mexico City.
- Edwards, T.C., Jr, Cuttler, R., Zimmermann, N.E., Geiser, L. & Alegria, J. (2005) Model-based stratifications for enhancing the detection of rare ecological events: Lichens as a case study. *Ecology*, **86**, 1081–1090.
- Elith, J., Graham, C.H., Anderson, R.P., Dudík, M., Ferrier, S., Guisan, A., Hijmans, R.J., Huettmann, F., Leathwick, J.R., Lehmann, A., Li, J., Lohmann, L.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., Soberón, J., Williams, S., Wisz, M.S. & Zimmermann, N.E. (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, **29**, 129–151.
- Engler, R., Guisan, A. & Rechsteiner, L. (2004) An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. *Journal of Applied Ecology*, **41**, 263–274.
- Ferrier, S. & Guisan, A. (2006) Spatial modelling of biodiversity at the community level. *Journal of Applied Ecology*, **43**, 393–404.
- Ferrier, S., Drielsma, M., Manion, G. & Watson, G. (2002) Extended statistical approaches to modelling spatial pattern in biodiversity in north-east New South Wales II. Community-level modelling. *Biodiversity and Conservation*, **11**, 2309–2338.
- Forgy, E.W. (1965) Cluster analysis of multivariate data: efficiency vs. interpretability of classifications. *Biometrics*, **21**, 768–769.
- Funk, V.A. & Richardson, K.S. (2002) Systematic data in biodiversity studies: use it or lose it. *Systematic Biology*, **51**, 303–316.
- Galindo-Jaimes, L., González-Espinosa, M., Quintana-Ascencio, P.F. & García-Barrios, L. (2002) Tree composition and structure in disturbed stands with varying dominance by *Pinus* spp. in the Highlands of Chiapas, Mexico. *Plant Ecology*, **162**, 259–272.

- Golicher, D.J. (2002) *The dynamics of disturbed Mexican Pine Oak forest: a modelling approach*. PhD thesis, University of Edinburgh.
- Golicher, D.J., O'Hara, R.B., Ruiz-Montoya, L. & Cayuela, L. (2006a) Lifting a veil on biodiversity: a Bayesian approach to fitting relative abundance models. *Ecological Applications*, **16**, 202–212.
- Golicher, D.J., Ramírez-Marcial, N. & Levy-Tacher, S. (2006b) Correlations between precipitation patterns in Southern Mexico and the El Niño sea surface temperature index. *Interciencia*, **31**, 80–86.
- González-Espinosa, M., Quintana-Ascencio, P.F., Ramírez-Marcial, N. & Gaytán-Guzmán, P. (1991) Secondary succession in disturbed *Pinus-Quercus* forest in the highlands of Chiapas, Mexico. *Journal of Vegetation Science*, **2**, 351–360.
- González-Espinosa, M., Rey-Benayas, J.M., Ramírez-Marcial, N., Huston, M.A. & Golicher, D. (2004) Tree diversity in the northern Neotropics: regional patterns in highly diverse Chiapas, Mexico. *Ecography*, **27**, 741–756.
- González-Espinosa, M., Ramírez-Marcial, N., Méndez-Dewar, G., Galindo-Jaimes, L. & Golicher, J.D. (2005) Riqueza de especies de árboles en Chiapas: variación espacial y dimensiones ambientales asociadas al nivel regional. *Diversidad Biológica en Chiapas* (ed. by M. González-Espinosa, N. Ramírez-Marcial and L. Ruiz-Montoya), pp. 81–125. Plaza y Valdés, México, DF.
- Guisan, A. & Zimmermann, N.E. (2000) Predictive habitat distribution models in ecology. *Ecological Modelling*, **135**, 147–186.
- Guisan, A., Broennimann, O., Engler, R., Vust, M., Yoccoz, N.G., Lehmann, A. & Zimmermann, N.E. (2006) Using niche-based models to improve the sampling of rare species. *Conservation Biology*, **20**, 501–511.
- Hartigan, J.A. & Wong, M.A. (1979) A K-means clustering algorithm. *Applied Statistics*, **28**, 100–108.
- 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.
- IPCC (Intergovernmental Panel on Climate Change) (2007) *Climate change 2007: the physical science basis*. Cambridge University Press, Cambridge.
- IUCN (2007) *2007 IUCN Red List of Threatened Species*. <http://www.iucnredlist.org>.
- Iverson, L.R., Prasad, A. & Schwartz, M.W. (1999) Modelling potential future individual tree-species distributions in the eastern United States under a climate change scenario: a case study with *Pinus virginiana*. *Ecological Modelling*, **115**, 77–93.
- Jepson, P. & Whittaker, R.J. (2002) Ecoregions in context: a critique with special reference to Indonesia. *Conservation Biology*, **16**, 42–57.
- Kienast, F., Wildi, O. & Brzeziecki, B. (1998) Potential impacts of climate change on species richness in mountain forests – an ecological risk assessment. *Biological Conservation*, **83**, 291–305.
- Kier, G., Mutke, J., Dinerstein, E., Ricketts, T.H., Küper, W., Kreft, H. & Barthlott, W. (2005) Global patterns of plant diversity and floristic knowledge. *Journal of Biogeography*, **32**, 1107–1116.
- Lawler, J.J., White, D., Neilson, R.P. & Blaustein, A.R. (2006) Predicting climate-induced range shifts: model differences and model reliability. *Global Change Biology*, **8**, 1568–1584.
- Leisch, F., Hornik, K. & Ripley, B.D. (2005) *MDA: mixture and flexible discriminant analysis*, R package version 0.3-0. www.r-project.org.
- Lips, K.R. (1998) Decline of a tropical montane amphibian fauna. *Conservation Biology*, **12**, 106–117.
- Margules, C.R. & Austin, M.P. (1994) Biological models for monitoring species decline: the construction and use of data bases. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **344**, 69–75.
- Miranda, F. (1952) *La vegetación de Chiapas*, Part 1. Gobierno del Estado de Chiapas, Tuxtla Gutiérrez, Chiapas, México.
- Monserud, R.A. & Leemans, R. (1992) Comparing global vegetation maps with the Kappa statistics. *Ecological Modelling*, **62**, 275–293.
- Ochoa-Gaona, S. & González-Espinosa, M. (2000) Land use and deforestation in the highlands of Chiapas, Mexico. *Applied Geography*, **20**, 17–42.
- Ochoa-Gaona, S., González-Espinosa, M., Meave, J.A. & Sorani-Dal Bon, V. (2004) Effects of forest fragmentation on the woody flora of the highlands of Chiapas, Mexico. *Biodiversity and Conservation*, **13**, 867–884.
- O'Hara, R.B. (2005) Species richness estimators: how many species can dance on the head of a pin? *Journal of Animal Ecology*, **74**, 375–386.
- Oksanen, J., Kindt, R. & O'Hara, R.B. (2005) *vegan: community ecology*, R package version 1.6-9 (available at: <http://cc.oulu.fi/~jarioksa/>).
- Olson, D.M. & Dinerstein, E. (1998) The Global 200: a representation approach to conserving the Earth's distinctive ecoregions. *Conservation Biology*, **12**, 502–515.
- Olson, D.M. & Dinerstein, E. (2002) The Global 200: priority ecoregions for global conservation. *Annals of the Missouri Botanical Garden*, **89**, 199–224.
- Otýpková, Z. & Chytrý, M. (2006) Effects of plot size on the ordination of vegetation samples. *Journal of Vegetation Science*, **17**, 465–472.
- Pearson, R.G. & Dawson, T.P. (2003) Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Global Ecology and Biogeography*, **12**, 361–371.
- Pounds, J.A. & Puschendorf, R. (2004) Clouded futures. *Nature*, **427**, 107–109.
- Pounds, J.A., Fogden, M.P.L. & Campbell, J.H. (1999) Biological responses to climate change on a tropical mountain. *Nature*, **398**, 611–615.
- R Development Core Team (2005) *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna (<http://www.R-project.org>).
- Ramírez-Marcial, N., González-Espinosa, M. & Williams-Linera, G. (2001) Anthropogenic disturbance and tree diversity in

- Montane Rain Forests in Chiapas, Mexico. *Forest Ecology and Management*, **154**, 311–326.
- Ramírez-Marcial, N., Camacho-Cruz, A. & González-Espinosa, M. (2005) Potencial florístico para la restauración de bosques en Los Altos y Montañas del Norte de Chiapas. *Diversidad Biológica en Chiapas* (ed. by M. González-Espinosa, N. Ramírez-Marcial and L. Ruiz-Montoya), pp. 325–369. Plaza y Valdés, México, DF.
- Rzedowski, J. (1978) *Vegetación de México*. Limusa, México, D.F.
- Still, C.J., Forest, P.N. & Schneider, S.H. (1999) Simulating the effects of climate change on tropical montane cloud forests. *Nature*, **398**, 608–610.
- Stomes, D.M. & Hargrove, W.W. (2000) Potential NDVI as a baseline for monitoring ecosystem functioning. *International Journal of Remote Sensing*, **21**, 401–407.
- 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. (2003) BIOMOD – optimizing predictions of species distributions and projecting potential future shifts under global change. *Global Change Biology*, **9**, 1353–1362.
- Tucker, C.J. (1979) Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, **8**, 127–150.
- Vieira, I.C.G., de Almeida, A.S., Davidson, E.A., Stone, T.A., de Carvalho, C.J.R. & Guerrero, J.B. (2003) Classifying successional forests using Landsat spectral properties and ecological characteristics in eastern Amazonia. *Remote Sensing of the Environment*, **87**, 470–481.
- Wood, S.N. (2003) Thin plate regression splines. *Journal of the Royal Statistical Society, Series B*, **65**, 95–114.
- Zobel, M., van der Maarel, E. & Dupré, C. (1998) Species pool: the concept, its determination and significance for community restoration. *Applied Vegetation Science*, **1**, 55–66.

BIOSKETCHES

Duncan J. Golicher is a researcher at El Colegio de la Frontera Sur, Mexico. His work focuses on forest dynamics and landscape-level modelling, with a particular emphasis on tropical montane forests.

Luis Cayuela is a post-doctoral researcher at Alcalá de Henares University, Spain. His current research focuses on the effects of forest loss and fragmentation on species diversity, forest restoration and conservation.

J. Robert M. Alkemade is a researcher at the Netherlands Environmental assessment Agency, his main interest is in forecasting future changes in biodiversity driven by environmental change and policies.

Mario González-Espinosa and **Neptalí Ramírez-Marcial** research at El Colegio de la Frontera Sur on plant ecology, diversity patterns, restoration of native forests and conservation management.

Editor: Erica Fleishman

SUPPLEMENTARY MATERIAL

The following supplementary material is available for this article:

Appendix S1 Tree species included in each of the six Climatically Associated Species Pools defined for the Highlands of Chiapas, Mexico.

This material is available as part of the online article from: <http://www.blackwell-synergy.com/doi/abs/10.1111/j.1466-8238.2007.00362.x>
(This link will take you to the article abstract).

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.