How well does presence-only-based species distribution modelling predict assemblage diversity? A case study of the Tenerife flora

Silvia C. Aranda and Jorge M. Lobo

S. C. Aranda, Dept Biodiversidad y Biología Evolutiva, Museo Nacional de Ciencias Naturales (CSIC), C/ José Gutiérrez Abascal, 2, ES-28006, Madrid, Spain, and Univ. dos Açores, Dept de Ciências Agrárias, CITA-A (Azorean Biodiversity Group), Terra-Chã, PT-9700-851 Angra do Heroísmo, Terceira, Açores, Portugal. – J. M. Lobo (mcnj117@mncn.csic.es), Dept Biodiversidad y Biología Evolutiva, Museo Nacional de Ciencias Naturales (CSIC), C/José Gutiérre Abascal, 2, ES-28006, Madrid, Spain.

Prediction error is considered an important problem in species distribution models. To address this issue, we here examined the accuracy of overlays of presence-only-based models for many individual species in representing patterns of assemblage diversity. For this purpose, we used a database of 977 160 records of seed plant occurrences on an intensively surveyed, species-rich island (Tenerife, Canary Islands) for modelling the distribution of all its 841 native plant species individually. The modelling was done using Maxent, one of the best-performing presence-only modelling techniques, using various thresholds to convert the estimated suitability values into predicted presence or absence. Distribution models for each individual species were overlaid to predict species richness and composition, which were then compared to the observed values for well-surveyed grid cells. We found high levels of compositional error, when the best performing suitability threshold for predicting species richness was applied. Our best prediction had a mean species richness error of 24% and a mean compositional error of 60% relative to the observed values for the well-surveyed cells; >50% of all species were included erroneously in >25% of the well-surveyed cells. Hence, large quantities of data are not necessarily enough to obtain reliable predictions of assemblage diversity, limiting the usefulness of this methodology in conservation planning.

Despite the compilation of massive amounts of data on the distributions of organisms, distributional knowledge is frequently incomplete and biased, particularly for hyper-diverse groups and regions (the Wallacean shortfall; Whittaker et al. 2005). Species distribution models have been suggested as a valuable strategy to overcome this deficiency for basic and applied purposes (Guisan and Zimmermann 2000, Soberón and Peterson 2005, Araújo and Guisan 2006). However, the quality of these models depends on the biological information on which they are based (Jiménez-Valverde et al. 2008), which is often incomplete and biased (Hortal et al. 2007, 2008). This source of uncertainty is one of the critical factors that limit the utility of distribution model results for conservation purposes (Rondinini et al. 2006).

The different modelling techniques used to predict species distributions must be viewed as correlative or forecasting simulations sensu Legendre and Legendre (1998, p. 493), in which little a priori information is available on the underlying processes that actually shape species distributions. That is, these models can arguably be considered mere interpolation procedures, the results of which depend on the quality of the data used for model calibration, including biases in collection effort across spatial and environmental gradients in the study region

(Hortal and Lobo 2006, Hortal et al. 2008). If environmentally unique areas in the region do not have enough distributional data, model predictions will have a higher probability of being erroneous (Kadmon et al. 2004, Hortal et al. 2007). In this case, species distributions are mapped by extrapolating beyond the ranges of the environmental gradient covered by the training data set, and should thus be treated with caution because they are likely to depend on the technique used to parameterize the model (Randin et al. 2006, Jiménez-Valverde et al. 2008). Apart from general incompleteness, distributional data and locations of well-surveyed areas are frequently spatially and environmentally biased: while some regions have been surveyed repeatedly thanks to sociological or accessibility factors, many others remain poorly surveyed (Dennis et al. 1999, Dennis and Thomas 2000, Lobo et al. 2007, Soberón et al. 2007). A previous study that aimed at predicting the distribution of seed plant species richness across Tenerife Island (Canary Islands) used data from well-surveyed localities to demonstrate that both data incompleteness and bias hinder estimating a reliable picture of biodiversity patterns, even if the data come from an apparently comprehensive compilation of all the available taxonomic and distributional information (Hortal et al. 2007).

Estimating species richness patterns by means of a predictive modelling using the available information does not always provide accurate results even when using high-quality data. Here, we assess whether accurate predictions of assemblage diversity (species richness and composition) can be achieved by overlaying results of individual presence-only-based species distribution models. We focus on presence-only data as this is often the only kind of available data on species distributions. To this end, we used an apparently exhaustive database on plant species distributions on Tenerife and Maxent, one of the best-performing species distribution modelling techniques for presence-only data (Elith et al. 2006). Briefly, we developed distribution models for each species and then overlaid the resulting maps to predict species richness and composition (a ''modelling-then-classification'' procedure, Ferrier et al. 2002a, b). We then compared the predicted richness and composition estimates to observed values from well-surveyed sites. We also provide guidance on how to minimize errors generated by standard applications of species distribution modelling methods for deriving assemblage diversity patterns from multiple individual-species predictions.

Methods

Area of study

The Canary Islands archipelago (27°37′–29°24′N, 13°23′–18°8′W) is comprised of seven main islands. Tenerife is the largest, with a surface area of 2034 km^2 (Fig. 1) and a mild climate year around (Fernández-Palacios and Martín 2002). Within the archipelago, the flora of Tenerife is especially rich, having a total of ca 1349 species of seed plants, hereof 318 endemic. The complex geological history of Tenerife helps to explain its astonishing floristic diversity; although the current configuration of the island dates from 1 million years ago, it originated by the merging of three ''palaeo-islands'' that emerged during the late Tertiary and early Quaternary (Marrero 2004).

Biological data

Biological information was obtained from BIOTA-Canarias, a database of presences of all species in the Canary Islands <www.gobiernodecanarias.org/cmayot/medioambiente/medionatural/biodiversidad/especies/bancodatos/index.html>. This database has been used previously in biogeographic and conservation studies (Emerson and Kolm 2005, Hortal et al. 2007). Its taxonomic and distributional information provides the best data source on the biodiversity of the Canary Islands (Izquierdo et al. 2001, 2005). In the specific case of Tenerife Island, a huge amount of data on seed plant species has been gathered (Santos 2001), compiled by botanists since Linnaeus and with diverse, systematized surveys during the 20th century. We used the 977 160 species occurrence records for the 841 native seed plant species that were available for Tenerife at the time of data extraction (29 November 2003). Each occurrence record referred to one of the ~500 ×500 m UTM grid cells (n = 8469 cells; see Hortal et al. 2007 for further details).

Well-surveyed grid-cells (n = 264; Fig. 1) were selected by plotting the relationships between the number of database records and the number of species in every cell for each of island's geo-environmental regions (see Hortal et al. 2007 for further details). The number of database records per cell was used as a surrogate of sampling effort (Lobo 2008a). Cells with ≥50 database records and ≥3 database records per species were considered well-surveyed.

Environmental variables

It is always difficult to know which environmental variables are most important in delimiting species distributions, and even more so when vast numbers of species are considered simultaneously. For this reason, we first chose a large suite of environmental predictors, which we then subsequently reduced in number by ordination procedures. Thirty climatic variables were selected: mean monthly temperature (12 months), annual mean temperature, minimum and maximum annual temperatures, mean monthly precipitation (12 months), total annual precipitation, and precipitation of the driest and wettest quarters. Additionally, five topographic variables were also selected: mean altitude, altitude range, slope, aspect and diversity of slopes. All the climatic variables were provided by the Spanish Inst. Nacional de Meteorología, while topographic variables were derived from a digital elevation model with a resolution of 100 ×100 m. Topographic variables were managed using the IDRISI Kilimanjaro software (Clark Labs 2003). Slope and aspect diversity were both calculated using the Shannon-Wiener index of the relative frequency of each class in all the twenty-five pixels of 100 ×100 m contained in every 500 ×500 m grid-cell. Altitude range was calculated as the difference between the highest and lowest elevations in the 500 ×500 m grid-cells.

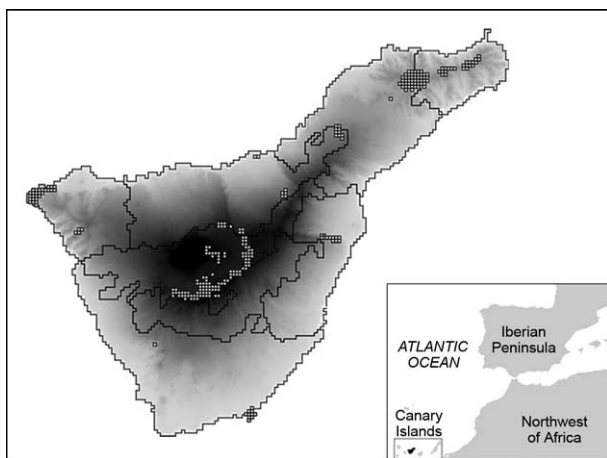All of these continuous variables were standardized to zero mean and one standard deviation to eliminate

Figure 1. Elevation map of the Tenerife Island showing 1) the well-surveyed 500 ×500 m grid-cells (n = 264) obtained from the plant BIOTA database, and 2) the eight main homogeneous geo-environmental regions described by Hortal et al. (2007). The inset shows the geographic location of Tenerife in the Canary Islands archipelago and the neighbouring continental areas.

measurement-scale effects. To select variables that better represent the environment of the region we used Jolliffe's principal component method (Rencher 2002). First, a Principal Component Analysis (PCA) was carried out with all of the continuous variables, and five non-correlated factors with eigenvalues ≥ 1 were obtained that explained 91.0% of the climatic and topographic variation across Tenerife. We subsequently selected the variables with the highest loadings on each of these factors (provided that they were > 0.8) as predictors. These variables included mean October temperature, mean annual precipitation, altitude range, June precipitation, and aspect diversity. Slope diversity and December precipitation were also included as predictors because they were not significantly correlated with any of the five PCA factors.

Apart from these continuous predictors, qualitative variables related to soil type (alfisols, entisols, inceptisols, vertisols, and utilisols), geological substrate (basalts, ancient basalts, and sediments), and aspect (N, NE, E, SE, S, SW, W and NW) were also incorporated in the modelling procedure to further restrict the predicted distribution generated by the exclusive use of climatic related predictors. The incorporation of these factors is necessary to obtain distributional representations closer to the realized distribution of species (Jiménez-Valverde et al. 2008). Digital maps of soil type and geology were provided by the Consejería de Política Territorial y Medio Ambiente of the Government of the Canary Islands. The continuous aspect values were converted into eight binary dummy variables of 45° width to avoid the problem of circular data (Hosmer and Lemeshow 1989). In total, sixteen binomial qualitative variables were also used as predictors (presence as 1, absence as 0).

Model building

We used the maximum entropy species distribution modelling approach (Maxent) to estimate the realized distribution of each plant species on Tenerife (Phillips et al. 2004, 2006). This machine-learning method is considered one of the best species distribution modelling techniques for presence-only data (Elith et al. 2006). Briefly, Maxent uses a deterministic algorithm to find an optimal probability distribution based on a set of environmental constraints (see details in Phillips et al. 2004, 2006, Phillips and Dudík 2008). Only linear and quadratic terms of the predictors were included in the modelling process in order to avoid overfitting. A preliminary analysis using the default Maxent features for the modelling process, produced similar, though slightly less overpredicted, distributions for a random set of 30 species (not shown), indicating that our finding were robust to the choice of allowable response shapes. Relative suitability for a species was estimated by the cumulative suitability output from Maxent for each grid cell. The cumulative output is computed as the sum of probability values for that cell and all the others with equal or lower probability multiplied by 100 to produce a continuous measure of favourability of presence that varies from 0 to 100. All models were run with Maximum Entropy Species Distribution Modelling v. 2.3 (<www.cs.princeton.edu/~schapire/maxent/>).

Statistical analyses

We assessed whether accurate predictions of species richness and composition could be assembled by overlaying the Maxent results using three different approaches for estimating assemblage diversity. First, we estimated species richness by summing the raw suitabilities estimated per grid cell for each species. This approach assumes that the summed suitabilities estimate how favourable conditions are for the presence of a high number of species. Contrary to the two next methods, this approach does not provide an estimate of species composition. As an alternative to the first approach, assemblage diversity was instead estimated by first obtaining binary species distribution maps by transforming the continuous suitability values for each species into presence-absence predictions, and then overlaying all these individual-species models. This methodology was implemented in two ways. Since suitability scores range from 0 to 100, in one approach 21 decision thresholds were selected at intervals of 5 to derive the respective presence-absence maps for each species, recognizing as presences the cells with suitability scores equal to or higher than 1, 5, ..., 95 and 100. As an alternative approach, the conversion into presence-absence predictions was done using the lowest predicted value associated with an observed presence record was used (hereafter Minimum Training Presence or MTP) as the threshold. This threshold guarantees that all presences are predicted as suitable (Pearson et al. 2007).

We estimated the accuracy of the species richness predictions for all these approaches by comparing observed richness in the well-surveyed grid cells to the number of species predicted present by each approach. Predicted and observed species richness scores were compared by calculating the non-parametric Spearman rank correlation (r_s) between the two variables in the well-surveyed cells.

Since cells with similar observed and predicted species richness could differ in their species composition, we finally also compared predicted and observed turnover in species composition in the well-surveyed cells, deriving the former directly from the overlain predicted presence-absence species distribution maps. We note that all assessments of predictions of species composition or species distributions only involved the two approaches that generated presence-absence predictions for each species. Mantel correlation tests (R) were used to assess similarity in the predicted and observed turnover in species composition (Sokal 1979). The Mantel test measures the correlation (R) between two distances or similarity matrices. Significance testing is based on a Monte Carlo permutational method that accommodates the lack of independence between site pairs. We calculated the species composition distance matrices by computing the Simpson β-diversity index (β_{sim}) for all pairs of well-surveyed cells, first based on the observed assemblages, and then also based on each the 22 sets of predicted assemblages obtained by the different thresholds mentioned above. β-diversity is defined as the difference in species composition between sites (Whittaker 1972), and β_{sim} was selected to measure it because it is independent of richness differences (Koleff et al. 2003). Hence, Mantel tests were performed between the observed and each of the 22 predicted species composition distance matrices. PAST

v. 1.68 free software was used for these computations (Hammer et al. 2001).

The threshold method generating the highest correlations between the observed and predicted species richness and composition scores was selected for further analyses. The difference between predicted and observed species richness scores (species richness errors) was calculated for all well-surveyed cells. Likewise, the difference in species composition between observed and predicted maps (species composition errors) was also computed as the total number of false positives and false negatives in each well-surveyed cell. Both error measures were quantified relative to the observed richness by dividing by the number of observed species. The two error measures were then regressed separately against the formerly mentioned environmental and spatial variables using General Linear Models to examine if the errors have an environmental or spatial structure. All variables were submitted to a backward selection procedure to estimate their explanatory capacity. Quadratic and cubic functions of each variable were considered, to account for curvilinear relationships. The errors were also regressed against the third-degree polynomial of grid cell longitude and latitude to detect spatial patterns (Trend Surface Analysis, Legendre and Legendre 1998). STATISTICA v.7 was used for all of these computations (StatSoft 2004).

Besides estimating species richness and species composition errors by cell, we also estimated errors on a species by species basis. We calculated the number and percentage of well-surveyed cells for which each species was erroneously predicted.

Results

Comparison between observed and predicted scores

The intercept of the linear relationship between predicted and observed values can be considered as a measure of overprediction in species-poor cells: around 6 species per grid-cell for raw suitability scores and 153 species for the MTP criterion (Fig. 2). Both slopes differed statistically from zero. In the case of summed suitabilities the slope was below unity (0.76 ± 0.02 SD, $t = 32.82$, $p < 0.0001$), indicating underprediction in the richest cells. In contrast, the slope was higher than unity (1.51 ± 0.06, $t = 23.89$, $p < 0.0001$) for MTP-based predictions, indicating overprediction in these same cells. Nevertheless, the correlation between the observed and MTP-based predicted richness scores was relatively high (Spearman $r_s = 0.82$, $n = 264$, $p < 0.0001$), as was the correlation between observed and MTP-based predicted turnover in species composition (Mantel $R = 0.67$; $n = 264$, $p < 0.0001$).

Considering the approach with 21 fixed suitability thresholds, the results obtained by using a least restrictive threshold (presence assigned when suitability > 1) were similar to the MTP-based results (Fig. 3). At the opposite extreme, the most restrictive thresholds (presence assigned when suitability > 95 or $= 100$) generated poor predictions of both species richness and composition. Between these two extreme situations, a wide range of thresholds provided relatively accurate predictions (Fig. 3), but particularly so at
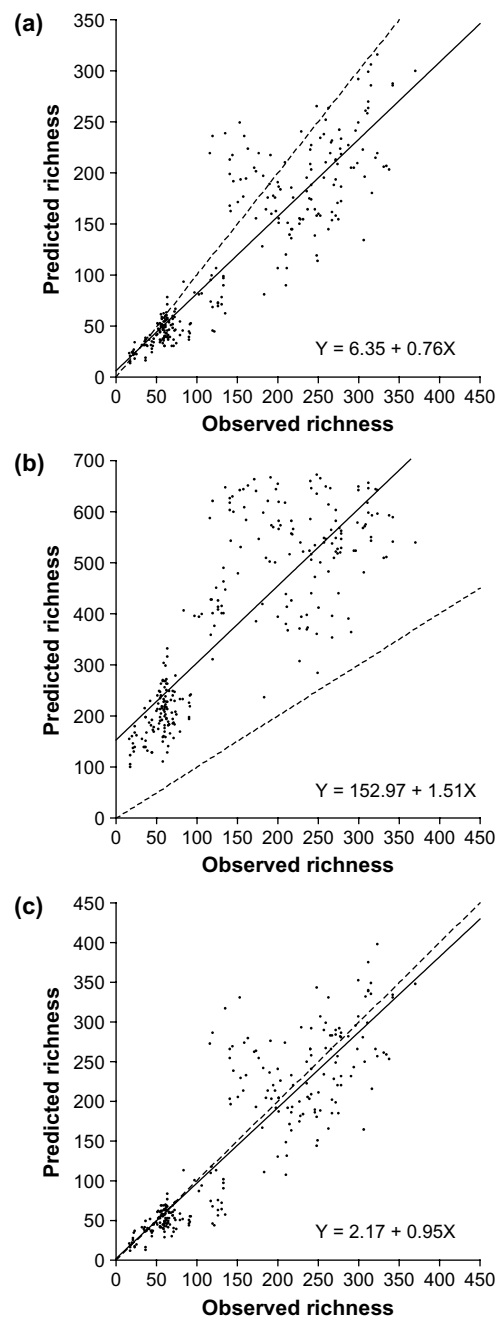

Figure 2. Plot of observed and predicted richness derived from overlapping (a) individual distribution models using the sum of raw individual suitability values, (b) the lowest predicted value associated with an observed presence record (MTP) as threshold criteria to cut continuous suitability values, or (c) using the threshold suitability value that generates the highest species richness and composition correlations (suitability = 35; see text). Continuous line represents the linear regression equation, while broken lines indicate x = y.

thresholds of 30 and 35 (equal correlations for both thresholds; richness, $r_s = 0.88$, $p < 0.0001$; compositional turnover, $R = 0.93$, $p < 0.0001$; $n = 264$). Using a suitability score of 35 as threshold, the linear relationship between predicted and observed species richness had an intercept not significantly different from zero (2.17 ± 5.24; $t = 0.41$, $p = 0.7$) and a significant slope only slightly lower
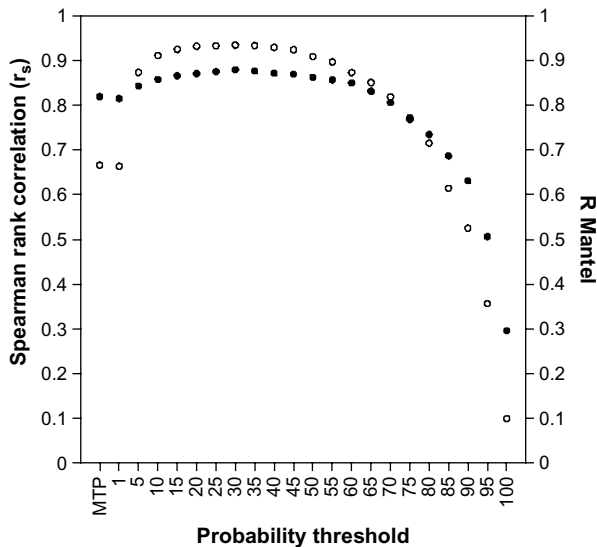
Figure 3. Spearman rank correlation values between observed and predicted species richness distribution (dark circles) and Mantel test correlations between observed and predicted turnover in species composition (open circles) for the different suitability thresholds used for converting the continuous suitability values into binary species presence-absence predictions. MTP (Minimum Training Presence threshold) is the lowest suitability value associated with an observed presence record.

than one (0.95 ± 0.03; $t = 32.02$, $p < 0.0001$), showing in this case that the richest cells were only slightly under-estimated (Fig. 2).

Examination of model errors

Richness maps created using the threshold of 35 showed an absolute mean species richness error of 23.6% (SD $= 22.1$; minimum $= 0.0\%$, maximum $= 138.3\%$) with 20.2% of the variability in these errors being explained by the environmental variables ($F_{4, 259} = 17.64$, $p < 0.0001$). The quadratic function of slope diversity provided the best univariate relationship between model errors and environmental variables ($R^2_{adj} = 16.2\%$); the highest errors in species richness were found at both low and high slope diversity scores. Errors were also spatially structured; geographic variables explained 33.2% of total variability ($F_{5, 258} = 27.15$, $p < 0.0001$), largely by a cubic function of longitude ($R^2_{adj} = 29.0\%$), showing that eastern areas were generally overpredicted.

Errors were even higher for species composition: an average of 59.7% (± 21.9 SD) of species predicted in a cell were mistakenly classified as present or absent. Errors were always $\geq 29.6\%$ in any cell. The mean number of false positives (species predicted present when they are in fact absent) was 39.6 (± 39.7 SD; minimum $= 2$, maximum $= 193$), while the mean number of false negatives (species predicted absent when they are really present) was 44.7 (± 33.1 SD; minimum $= 5$, maximum $= 164$). Errors in species composition did not occur randomly: environmental variables accounted for 28.9% of total variability ($F_{7, 256} = 16.27$, $p < 0.0001$), with presence of inceptisols as the most relevant predictor ($R^2_{adj} = 15.5\%$): inceptisol

cells had higher compositional errors ($\sim 75\%$ of species incorrectly predicted). Spatial variables explained 14.6% of the variability in composition errors ($F_{2, 261} = 23.41$, $p < 0.0001$), with the highest error values in cells with high overprediction in the northeastern part of the island.

Considering the prediction errors in the well-surveyed cells on species-by-species basis, each of the 841 species were on average erroneously omitted or included in 14.1 (± 15.4 SD) cells, an average error level per species of 89%. In total, 432 species (51.4% of the total) were erroneously included or excluded in $>\frac{1}{4}$ of the number of cells in which they have been recorded.

Discussion

Two main strategies have been proposed to obtain useful data on distributions of assemblages for conservation (Ferrier et al. 2002a, b, Ferrier and Guisan 2006): "classification-then-modelling" and "modelling-then-classification". While the former implies predicting composite biodiversity variables such as species richness, rarity, or composition from partial data directly, the latter involves overlaying many single-species distribution models. Here, we examine the possibilities of obtaining reasonable geographic representations of assemblage diversity (species richness and composition) by overlaying individual model predictions, using an intensely surveyed island (Tenerife) in the Canary Archipelago as a study case. A former study based on the same data that applied the "classification-then-modelling" approach showed that even comprehensive databases of species occurrences might not have adequate geographic and environmental coverage to accomplish unbiased and reliable species richness predictions (Hortal et al. 2007). Our findings in the present study show that overlaying single-species models may provide a relatively reliable prediction of the geographical variation in species richness, but with a high level of compositional errors, the percentage of species erroneously predicted present or absent from a cell averaging 60% of the predicted number of species for the threshold achieving the best richness prediction.

When attempting to predict species assemblage diversity by overlaying single-species distribution model predictions, the accuracy of not only the species richness predictions, but the species composition predictions should therefore be evaluated. Even though the richness of a given cell can be predicted accurately, its composition may be quite different from the actual one, with important implications for both ecology and conservation. Our results show that, on average, 40 species were erroneously added to each cell, while another almost 45 species were erroneously removed. Since the mean number of species in the well-surveyed cells was 147 (± 97 SD), there were strong differences between the observed and predicted species composition. This mismatch may result from the high resolution of the cells (500×500 m), resulting in including or excluding species that are actually present in neighbouring, environmentally similar grid-cells. We examined whether composition errors diminished when the resolution was coarsened by analyzing each of the eight previously identified geo-environmental regions of Tenerife (Fig. 1; Hortal et al. 2007). In this case,

model predictions generated an average species richness error of 16%, with omission errors not greater than 12% (mean = 4.2%), but commission errors exceeded 50% (mean = 46.1%; Table 1). However, each species was erroneously omitted or included in 1.2 regions ($\pm$1.3 SD), for an average error level of 41% according to the number of regions in which the species were observed. In total, 349 species (41.5% of total) were erroneously included or excluded in >¼ of the regions in which they were observed. However, interestingly, most of these over-predictions in a region involve species erroneously included in low numbers of cells; the mean number of cells overestimated for all the species wrongly predicted in a region ranged from 5.5 ($\pm$9.7 SD) in zone 4 (around 1% of regional area) to 31.9 ($\pm$49.3 SD) in zone 2 (around 7% of area). These results indicate that compositional errors are considerable, both locally and regionally, but that the error rate is likely to diminish drastically when model predictions are examined at larger study area extents, relative to the species range sizes.

Why do such high levels of compositional error occur in the present study? One major source of uncertainty potentially responsible for these compositional errors is probably the difficulty of including the effect of dispersal limitation and historical factors in species distribution models (Svenning et al. 2008). Overprediction is characteristic of most distribution predictions even when these are aimed at obtaining reliable representations of the realized distribution (Fielding and Haworth 1995, Araújo and Williams 2000, Feria and Peterson 2002, Stockwell and Peterson 2002, Loiselle et al. 2003, Brotons et al. 2004, Segurado and Araújo 2004, Graham and Hijmans 2006, Stockman et al. 2006). This probably happens because most modelling approaches have only taken current environmental conditions into account as range-limiting factors, while ignoring other factors that may constrain species' distributions, such as dispersal limitation, historical factors, or biotic interactions. Without using adequate predictors

Table 1. Species richness (SR) and species composition errors for each of the eight geo-environmental regions of Tenerife (Fig. 1) identified by Hortal et al. (2007). Species richness errors (i.e. predicted minus observed values) were computed in percentage of the observed number of species in each region. Composition errors were calculated in the percentage of the number of species observed in each region. Omission error is the percentage of species actually present in each region that were predicted absent by the models, while commission error is the percentage of species erroneously predicted present in each region, i.e. despite being absent. Overall performance is the correct classification rate (Fielding and Bell 1997); that is, the number of well-predicted presences and absences over the total number species considered in the confusion matrix.

			Composition errors (%)		
Regions	Area (km^2)	SR errors (%)	Overall performance	Omission error	Commission error
1	330	10.7	90.0	0.9	47.0
2	122	17.4	86.1	1.0	46.8
3	436	15.8	78.1	10.7	37.8
4	261	20.8	80.1	7.5	35.2
5	224	9.9	80.6	11.6	30.5
6	230	16.0	80.5	6.6	45.5
7	345	21.6	82.4	1.1	64.1
8	170	19.2	83.5	1.4	61.6

and reliable absence data, model results inevitably tend to reflect potential distributions of species (Soberón 2007, Lobo 2008b, Chefaoui and Lobo 2008, Jiménez-Valverde et al. 2008). Overall, most species distribution predictions should be considered situated somewhere intermediately along a gradient between potential and realized distributions (Jiménez-Valverde et al. 2008, Lobo 2008b). This fact limits the utility of species distribution models in systematic conservation planning. Nevertheless, these considerations are only relevant with respect to commission errors (over-predictions), and as discussed above omission errors were a more prominent type of compositional error in the present study.

The finding that omission errors were even more frequent than commission errors is an important observation for understanding the compositional errors. Hence, restricting overpredictions could probably reduce the rate of false presences, but often only at the expense of increasing even more the number of species erroneously predicted absent from each cell. The latter would be the case if overpredictions were reduced by allowing more complex species–environment relationships. We believe that an important cause of joint relevance of commission and omission errors is the low quality of the data used to calibrate the models. In spite of the large number of occurrence records gathered, historical data collection in Tenerife has been biased geographically (Hortal et al. 2007), which could compromise the representativeness of the distributions estimated from these data. Hence, the use of extensive data such as that used in this study (977 160 database records) does not seem to guarantee accurate predictions. Although we use an average of 1162 database records per species, 84.2% of cells have $\leq$2 database records per species, and 99.5% of cells have $\leq$5 records. Also, 33.4% of grid-cells have $\leq$50 database records. Hence, wide areas remain poorly surveyed, as shown by the high correlation between the observed species richness values and the number of database records (r = 0.935; Hortal et al. 2007). Furthermore and importantly, the available biological information is not distributed homogeneously across the environmental spectrum of the island; only 2.4% of the cells can be considered as well-surveyed and these cells are spatially and environmentally biased (Hortal et al. 2007). Hence, most cells have unknown levels of false absences that approaches based on overlaying single-species distribution models seem not to be able to predict adequately.

The predictions of species distribution models should be assessed by comparing model predictions with independent distributional data (see Fielding and Bell 1997 and Vaughan and Ormerod 2005 for reviews of evaluation methods). However, an appropriate validation of the realized distribution requires reliable and accurate absence information. In the case of profile methods (sensu Guisan and Zimmermann 2000) like Maxent, the lack of reliable absence information hinders the assessment of model reliability. Here, we propose a new and easy-to-use approach to estimating model performance when many individual distribution predictions are generated automatically and simultaneously using profile methods. Comparisons of species richness and composition based on overlaying predicted individual species distributions with

data from well-surveyed sites may be a useful strategy to assess model performance. Model errors can be examined by comparing assemblage diversity predictions derived from multiple individual species models to observed assemblage data from well-surveyed localities, as done in the present study. Such error assessments should not only evaluate the magnitude, but also the location of errors along the spatial and environmental gradients considered (Fielding and Bell 1997, Fielding 2002). Our analyses show that the variability of model errors can be explained relatively well by environmental and spatial variables, for both species richness and composition. The spatial and environmental aggregation of model errors may be caused by the biased structure of the raw data used to build and validate them (Hortal et al. 2007), but also by the difficulty of including predictors capable of accounting for the effects of dispersal limitation. Species distribution models devoted to revealing the realized distributions of species need to be carried out using biological information representing a good sample of the species' presences: reliable data well distributed across the spectrum of environmental and spatial conditions present in the region analyzed (Hortal and Lobo 2006). Furthermore, such species distribution models should also incorporate predictors capable of accounting for factors that exclude a species from unoccupied, but environmentally favourable areas, in combination with accurate absence information for such localities (Peterson 2006, Jiménez-Valverde et al. 2008) and/or presence-only data without strong geographic or environmental biases (De Marco et al. 2008). Unless properly accounted for, these shortcomings reduce the usefulness of ecological niche modelling techniques for conservation purposes seriously (Rondinini et al. 2006). Further research in this field should be directed towards developing tools that are able to handle the often inevitable bias and scarceness of species occurrence data, as well as to select the minimum biological information needed for generating reliable results.

References

Araújo, M. B. and Williams, P. H. 2000. Selecting areas for species persistence using occurrence data. – Biol. Conserv. 96: 331–345.

Araújo, M. B. and Guisan, G. 2006. Five (or so) challenges for species distribution modelling. – J. Biogeogr. 33: 1677–1688.

Brotons, L. W. et al. 2004. Presence-absence versus presence-only modelling methods for predicting bird habitat suitability. – Ecography 27: 437–448.

Chefaoui, R. M. and Lobo, J. M. 2008. Assessing the effects of pseudo-absences on predictive distribution model performance. – Ecol. Model. 210: 478–486.

Clark Labs 2003. Idrisi Kilimanjaro. GIS software package. – Clark Labs, Worcester, MA.

De Marco, P. et al. 2008. Spatial analysis improves species distribution modelling during range expansion. – Biol. Lett. 4: 577–580.

Dennis, R. L. H. and Thomas, C. D. 2000. Bias in butterfly distribution maps: the influence of hot spots and recorder's home range. – J. Insect Conserv. 4: 73–77.

Dennis, R. L. H. et al. 1999. Bias in butterfly distribution maps: the effects of sampling effort. – J. Insect Conserv. 3: 33–42.

Elith, J. et al. 2006. Novel methods improve prediction of species' distributions from occurrence data. – Ecography 29: 129–151.

Emerson, B. C. and Kolm, N. 2005. Species diversity can drive speciation. – Nature 434: 1015–1017.

Feria, T. P. and Peterson, A. T. 2002. Prediction of bird community composition based on point-occurrence data and inferential algorithms: a valuable tool in biodiversity assessments. – Divers. Distrib. 8: 49–56.

Fernández-Palacios, J. M. and Martín, J. L. 2002. Naturaleza de las Islas Canarias, 2nd ed. – Publicaciones Turquesa, Madrid.

Ferrier, S. and Guisan, A. 2006. Spatial modelling of biodiversity at the community level. – J. Appl. Ecol. 43: 393–404.

Ferrier, S. et al. 2002a. Extended statistical approaches to modelling spatial pattern in biodiversity in northeast New South Wales. I. Species-level modelling. – Biodivers. Conserv. 11: 2275–2307.

Ferrier, S. et al. 2002b. Extended statistical approaches to modelling spatial pattern in biodiversity in northeast New South Wales. II. Community-level modelling. – Biodivers. Conserv. 11: 2309–2338.

Fielding, A. H. 2002. What are the appropriate characteristics of an accuracy measure? – In: Scott, J. M. et al. (eds), Predicting species occurrences: issues of accuracy and scale. Island Press, pp. 271–280.

Fielding, A. H. and Haworth, P. F. 1995. Testing the generality of bird-habitat models. – Conserv. Biol. 9: 1466–1481.

Fielding, A. H. and Bell, J. F. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. – Environ. Conserv. 24: 38–49.

Graham, C. H. and Hijmans, R. J. 2006. A comparison of methods for mapping species ranges and species richness. – Global Ecol. Biogeogr. 15: 578–587.

Guisan, A. and Zimmermann, N. E. 2000. Predictive habitat distribution models in ecology. – Ecol. Model. 135: 147–186.

Hammer, O. et al. 2001. PAST: palaeontological statistics software package for education and data analysis. – Palaeontol. Electron. 4: 1–9.

Hortal, J. and Lobo, J. M. 2006. Towards a synecological framework for systematic conservation planning. – Biodivers. Inform. 3: 16–45.

Hortal, J. et al. 2007. Limitations of biodiversity databases: case study on seed-plant diversity in Tenerife, Canary Islands. – Conserv. Biol. 21: 853–863.

Hortal, J. et al. 2008. Historical bias in biodiversity inventories affects the observed environmental niche of the species. – Oikos 117: 847–858.

Hosmer, D. W. and Lemeshow, S. 1989. Applied logistic regression, 2nd ed. – Wiley.

Izquierdo, I. et al. 2001. Lista de especies silvestres de Canarias (hongos, plantas y animales terrestres). – Gobierno de Canarias, Consejería de Política Territorial y Medio Ambiente.

Izquierdo, I. et al. 2005. Lista de especies silvestres de Canarias (hongos, plantas y animales terrestres), 2nd ed. – Gobierno de Canarias, Consejería de Política Territorial y Medio Ambiente.

Jiménez-Valverde, A. et al. 2008. Not as good as they seem: the importance of concepts in species distribution modelling. – Divers. Distrib. 14: 885–890.

Kadmon, R. et al. 2004. Effect of roadside bias on the accuracy of predictive maps produced by bioclimatic models. – Ecol. Appl. 14: 401–413.

Koleff, P. et al. 2003. Measuring beta diversity for presence-absence data. – J. Anim. Ecol. 72: 367–382.

Legendre, P. and Legendre, L. 1998. Numerical ecology. – Elsevier.

Lobo, J. M. 2008a. Database records as a surrogate for sampling effort provide higher species richness estimations. – Biodivers. Conserv. 17: 873–881.

Lobo, J. M. 2008b. More complex distribution models or more representative data? – Biodivers. Inform. 5: 14–19.

Lobo, J. M. et al. 2007. How does the knowledge about the spatial distribution of Iberian dung beetle species accumulate over time? – Divers. Distrib. 13: 772–780.

Loiselle, B. A. et al. 2003. Avoiding pitfalls of using species distribution models in conservation planning. – Conserv. Biol. 17: 1591–1600.

Marrero, A. 2004. Procesos evolutivos en plantas insulares, el caso de Canarias. – In: Fernández-Palacios, J. M. and Morici, C. (eds), Ecología insular/Island ecology. Asociación Española de Ecología Terrestre and Cabildo Insular de La Palma, Spain, pp. 305–356.

Pearson, R. G. et al. 2007. Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. – J. Biogeogr. 34: 102–117.

Peterson, T. A. 2006. Uses and requirements of ecological niche models and related distributional models. – Biodivers. Inform. 3: 59–72.

Phillips, S. J. and Dudík, M. 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. – Ecography 31: 161–175.

Phillips, S. J. et al. 2004. A maximum entropy approach to species distribution modeling. – In: Proceedings of the 21st International Conference on Machine Learning. ACM Press, pp. 665–662.

Phillips, S. J. et al. 2006. Maximum entropy modelling of species geographic distributions. – Ecol. Model. 190: 231–259.

Randin, C. F. et al. 2006. Are niche-based species distribution models transferable in space? – J. Biogeogr. 33: 1689–1703.

Rencher, A. C. 2002. Methods of multivariate analysis, 2nd ed. – Wiley.

Rondinini, C. et al. 2006. Tradeoffs of different types of species occurrence data for use in systematic conservation planning. – Ecol. Lett. 9: 1136–1145.

Santos, A. 2001. Flora Vascular Nativa. – In: Fernández-Palacios, J. M. and Martín Esquivel, J. L. (eds), Naturaleza de las Islas Canarias. Ecología y Conservación. Publicaciones Turquesa, Madrid, pp. 185–192.

Segurado, P. and Araújo, M. B. 2004. An evaluation of methods for modelling species distributions. – J. Biogeogr. 31: 1555–1569.

Soberón, J. 2007. Grinnellian and Eltonian niches and geographic distributions of species. – Ecol. Lett. 10: 1115–1123.

Soberón, J. and Peterson, A. T. 2005. Interpretation of models of fundamental ecological niches and species' distributional areas. – Biodivers. Inform. 2: 1–10.

Soberón, J. et al. 2007. Assessing completeness of biodiversity databases at different spatial scales. – Ecography 30: 152–160.

Sokal, R. R. 1979. Testing statistical significance of geographic variation patterns. – Syst. Zool. 28: 627–632.

StatSoft 2004 STATISTICA (data analysis software system), version 7. – <www.statsoft.com>.

Stockman, A. K. et al. 2006. An evaluation of a GARP model as an approach to predicting the spatial distribution of non-vagile invertebrate species. – Divers. Distrib. 12: 81–89.

Stockwell, D. R. B. and Peterson, A. T. 2002. Effects of sample size on accuracy of species distribution models. – Ecol. Model. 148: 1–13.

Svenning, J.-C. et al. 2008. Postglacial dispersal limitation of widespread forest plant species in nemoral Europe. – Ecography 31: 316–326.

Vaughan, V. and Ormerod, S. J. 2005. The continuing challenges of testing species distribution models. – J. Appl. Ecol. 42: 720–730.

Whittaker, R. H. 1972. Evolution and measurement of species diversity. – Taxon 21: 213–251.

Whittaker, R. J. et al. 2005. Conservation biogeography: assessment and prospect. – Divers. Distrib. 11: 3–23.