

# A comparative evaluation of presence-only methods for modelling species distribution

Asaf Tsoar\*, Omri Allouche, Ofer Steinitz, Dotan Rotem and Ronen Kadmon

Department of Evolution, Systematics and Ecology, Institute of Life Sciences, The Hebrew University of Jerusalem, Campus Edmond Safra, Jerusalem 91904, Israel

## ABSTRACT

In spite of increasing application of presence-only models in ecology and conservation and the growing number of such models, little is known about the relative performance of different modelling methods, and some of the leading models (e.g. GARP and ENFA) have never been compared with one another. Here we compare the performance of six presence-only models that have been selected to represent an increasing level of model complexity [BIOCLIM, HABITAT, Mahalanobis distance (MD), DOMAIN, ENFA, and GARP] using data on the distribution of 42 species of land snails, nesting birds, and insectivorous bats in Israel. The models were calibrated using data from museum collections and observation databases, and their predictions were evaluated using Cohen's Kappa based on field data collected in a standardized sampling design covering most parts of Israel. Predictive accuracy varied between modelling methods with GARP and MD showing the highest accuracy, BIOCLIM and ENFA showing the lowest accuracy, and HABITAT and DOMAIN showing intermediate accuracy levels. Yet, differences between the various models were relatively small except for GARP and MD that were significantly more accurate than BIOCLIM and ENFA. In spite of large differences among species in prevalence and niche width, neither prevalence nor niche width interacted with the modelling method in determining predictive accuracy. However, species with relatively narrow niches were modelled more accurately than species with wider niches. Differences among species in predictive accuracy were highly consistent over all modelling methods, indicating the need for a better understanding of the ecological and geographical factors that influence the performance of species distribution models.

## Keywords

Bats, BIOCLIM, birds, distribution range, DOMAIN, ecological niche models, ENFA, GARP, HABITAT, habitat suitability, Kappa, Mahalanobis distance, predictive maps, prevalence, snails, tolerance.

\*Correspondence: Asaf Tsoar, Department of Evolution, Systematics and Ecology, Institute of Life Sciences, The Hebrew University of Jerusalem, Campus Edmond Safra, Jerusalem 91904, Israel.  
Tel.: 972 26586080; Fax: 972 26584655;  
E-mail: tsoarasaf@pob.huji.ac.il

## INTRODUCTION

Knowledge about the geographical distribution of species is crucial for conservation and management of biodiversity (Margules & Pressey, 2000). Yet, for most regions and most taxa, detailed data on species distribution are usually not available and collecting such data is costly and labour intensive (Prendergast *et al.*, 1999; Bowker, 2000; Ottaviani *et al.*, 2004). Consequently, ecologists and conservation scientists increasingly rely on predictive models as a means for estimating patterns of species distribution and informing conservation strategies (Peterson & Robins, 2003; Araujo *et al.*, 2004; Ortega-Huerta & Peterson, 2004; Sánchez-Cordero *et al.*, 2005). The number of methods available for modelling patterns of species distribution is immense (Guisan & Zimmermann, 2000; Scott *et al.*, 2002; Guisan & Thuiller, 2005),

and evaluating the relative performance of different methods remains a continuous challenge in ecology and conservation biology (Loiselle *et al.*, 2003; Thuiller, 2003; Ottaviani *et al.*, 2004; Vaughan & Ormerod, 2005; Elith *et al.*, 2006; Pearson *et al.*, 2006).

In general, two categories of methods for species distribution modelling can be distinguished: methods that need presence-absence data for model construction and methods that use presence-only data as a basis for generating predictions. Presence-only data differ from presence-absence data in that they indicate locations where the target organism was observed to occur, but cannot be used to define locations where the organism does not occur (Dettmers & Bart, 1999). Some modelling techniques use 'pseudo-absence' data for model construction (e.g. Stockwell & Peters, 1999; Zaniwski *et al.*, 2002; Engler *et al.*, 2004), but these

are still classified as presence-only methods because there is no real use of absence data in the construction of the model.

Presence-only records of species distribution, particularly those obtained from herbaria and museum collections, provide the oldest and most reliable documentation available on the distribution of plant and animal species (Ponder *et al.*, 2001; Williams *et al.*, 2002). Presence-only information is also much more available and requires much less collection efforts than presence-absence information. Recent progress in biodiversity informatics (Bisby, 2000) and the development of extensive web services and databases of biodiversity data (e.g. GBIF, MaNIS, FishNet, HerpNet) have contributed significantly to the accessibility of such data to the public. Yet, despite the numerous applications of presence-only models, relatively few studies have compared the relative performance of different methods (Elith & Burgman, 2002; Ferrier *et al.*, 2002; Farber & Kadmon, 2003; Loiselle *et al.*, 2003; Segurado & Araujo, 2004; Elith *et al.*, 2006). This lack of data is in contrast to methods based on presence-absence data or models using pseudo-absences such as generalized linear model (GLM) and generalized additive model (GAM), whose properties have been investigated much more intensively (e.g. Manel *et al.*, 1999; Pearce & Ferrier, 2000; Thuiller *et al.*, 2003; Thuiller, 2003). An exceptional case is a recent evaluation of 16 different methods carried out by Elith *et al.* (2006) based on distribution data of 226 species from six regions of the world. In the study, recently developed methods that have rarely been applied to modelling species distributions consistently outperformed more established methods.

In this study, we compare the performance of some of the most common methods of presence-only distribution models using data on the distribution of snails, birds, and bats in Israel. Specifically, we compare six different methods: BIOCLIM, HABITAT, DOMAIN, Mahalanobis distance (hereafter MD), Ecological Niche Factor Analysis (ENFA), and Genetic Algorithm for Rule-set Prediction (GARP). To our knowledge, this is the first study to compare some of the leading presence-only methods with one another (e.g. ENFA and GARP). Considering previous evidence that differences in predictive accuracy among models may depend on the prevalence (proportion of the validation sites in which the species was found) and niche width of the modelled species (Pearce & Ferrier, 2000; Brotons *et al.*, 2004; Segurado & Araujo, 2004; Elith *et al.*, 2006), we also quantified the prevalence and niche width of each species and tested whether differences among species in these properties interact with differences in the modelling method in determining the accuracy of model predictions. Such knowledge could assist in interpreting differences in predictive accuracy between models and may allow a more educated selection of modelling techniques.

## METHODS

### Environmental variables

Three climatic variables were chosen for constructing the models: mean annual rainfall, mean daily temperature of the hottest month (August), and mean minimum temperature of the coldest

month (January). These variables were chosen because they showed high correlations with other climatic variables in the study area but relatively low correlations among them (Steinitz *et al.*, 2005). Together, these variables capture the main climatic gradients of Israel (Kadmon & Danin, 1997; Kurtzman & Kadmon, 1999) and previous studies have shown that they are important determinants of distribution ranges of land snails (Heller, 1988; Kadmon & Heller, 1998; Steinitz *et al.*, 2005), birds (Shirihai, 1996; Steinitz *et al.*, 2005), and bats (Yom-Tov & Werner, 1996; Yom-Tov & Kadmon, 1998) in this region. Many previous applications of species distribution models have used similar indices as predictors of distribution patterns (Box *et al.*, 1993; Carpenter *et al.*, 1993; Shao & Halpin, 1995; Eeley *et al.*, 1999) and it seems that such a combination of rainfall and temperature variables effectively represents correlates of physiological tolerance (Martinez-Meyer, 2005). Preliminary analyses have indicated that adding additional climatic and topographical variables (up to a total of 23 variables) did not improve the accuracy of model predictions. A map of mean annual rainfall of the study area was constructed using data from 475 rainfall stations (Kadmon & Heller, 1998). Maps of mean daily temperature of the hottest month and mean minimum temperature of the coldest month were produced using the data collected in 38 climatic stations (Kurtzman & Kadmon, 1999). For the purpose of this study, all maps were rescaled into spatial resolution of 1 km<sup>2</sup>.

### Calibration data

The data used to construct the models included 7340 georeferenced records of land snails (17 species, 2123 records), nesting land birds (16 species, 4688 records), and bats (9 species, 529 records) in Israel. The snail data were obtained from the Mollusk Collection of The Hebrew University of Jerusalem. Records of bird distribution were obtained from the Zoological Collections of Tel Aviv University and the Database Unit of the Israel Nature and Parks Authority. Data on bat distribution were obtained from the Mammalian Center of the Society for the Protection of Nature in Israel, the Zoological Collections of Tel Aviv University, the Harrison Institute, and unpublished observations of individual researchers (M. Dor, Y. Barak, Y. Carmel, R. Feldman, and C. Korine). Each record was checked for its geographical coordinates and records positioned within any of the validation sites (see below) were removed from the calibration data set.

### Validation data

A field sampling was designed to collect a standardized set of presence-absence data for model validation. The size of the validation sites was adjusted to fit the spatial resolution of the predictive maps generated by the models (1 km<sup>2</sup>). The geographical distribution of the sites was determined using a stratified random sampling procedure with the aid of a geographical information system in order to represent the main geographical and climatic gradients of Israel. A total of 27 sites of 1 km<sup>2</sup> were selected for field sampling and of these, land snails were sampled in all sites, birds in 21 sites, and bats in 19 sites. Although this

sampling project required extensive logistic efforts, we believe that a larger sample size could have increased the power of our statistical tests.

A detailed description of the sampling design and the methods used for sampling snails and birds can be found elsewhere (Steinitz *et al.*, 2005). Bats were sampled using a combination of mist nets (a total of 50 m × 2.5 m) and three types of echolocation monitoring systems: frequency division bat detectors with data loggers (ANABAT II, Titley Electronics, NSW, Australia); time expansion bat detector (Pettersson D-980, Pettersson Elektronik, Uppsala, Sweden); and heterodyne bat detectors (Pettersson D-200, Pettersson Elektronik). Previous studies have shown that this combination of monitoring and capturing techniques maximizes the efficiency of bat sampling (O'Farrell & Gannon, 1999). Each site was visited twice, once during the spring and again during the summer. Sampling started 1 h before sunset and continued until half an hour after sunrise. The echolocation data were identified using a reference library that was created based on data collected in preliminary surveys and relevant literature.

## Models

Predictive maps were produced based on species occurrence data using the methods BIOCLIM (Busby, 1986; Nix, 1986), HABITAT (Walker & Cocks, 1991), MD (Farber & Kadmon, 2003), DOMAIN (Carpenter *et al.*, 1993), ENFA (Hirzel *et al.*, 2002), and GARP (Stockwell & Peters, 1999). All of these methods are based on the concept of the ecological niche (Hutchinson, 1957). In each method, some rules or mathematical algorithms are used to define the ecological niche of the species based on the distribution of the species records in the multidimensional environmental space. Once the species niche is defined, its projection into the geographical space produces a predictive map.

BIOCLIM defines the ecological niche of a species as the bounding hyper-box that encloses all the records of the species in the climatic space (Busby, 1991). Thus, it creates a rectilinear 'envelope' in the environmental space, defined by the most extreme (minimum and maximum) records of the species on each environmental variable. To reduce the sensitivity of model predictions to outliers, the species records are sorted along each variable and only the records that lie within a certain percentile range of the data are used for model construction. In this study we applied a percentile range of 95% (disregarding 2.5% of the values on each side).

HABITAT defines the ecological niche as the convex hull of the species records in the environmental space (Walker & Cocks, 1991). It differs from BIOCLIM in its ability to better adjust the boundaries of the environmental envelope to the distribution of the species records in the climatic space. As with BIOCLIM, however, only the outer records are used to determine the boundaries of the ecological niche.

The MD method ranks potential sites by their Mahalanobis distance to a vector expressing the mean environmental conditions of all the records in the environmental space. A certain distance threshold is then used to define the boundaries of the ecological niche. This algorithm produces an elliptic envelope that explicitly

accounts for possible correlations between the environmental variables (Farber & Kadmon, 2003).

DOMAIN uses a point-to-point similarity metric (based on the Gower distance) to assign a value of habitat suitability to each potential site based on its proximity in the environmental space to the closest (most similar) occurrence location (Carpenter *et al.*, 1993). A threshold value of suitability can then be selected to determine the boundaries of the ecological niche. Note that in contrast to all previous methods, environmental envelopes defined by DOMAIN are not necessarily continuous in the environmental space.

ENFA calculates a measure of habitat suitability based on the analysis of marginality (how the species mean differs from the global mean) and environmental tolerance (how the species variance compares to the global variance). A threshold of suitability value can then be applied to determine the boundaries of the ecological niche (Hirzel *et al.*, 2002).

GARP is a genetic algorithm that produces sets of rules that delineate ecological niches in an artificial-intelligence-based approach (Stockwell & Peters, 1999). Occurrence points are resampled randomly to create training and test data sets and the algorithm works in an iterative process of rule selection, evaluation, testing, and incorporation or rejection. Predictive accuracy is evaluated at each step based on the test presence data and a set of 'pseudo-absence' points, and the change in predictive accuracy from one iteration to the next is used to evaluate whether or not a particular rule should be incorporated into the model.

## Model application

All models except GARP were implemented within the MATLAB environment. ENFA was simulated in MATLAB to produce results equal to those calculated by BIOMAPPER (Hirzel *et al.*, 2004). We used the Medians algorithm, which proved accurate in most situations and has good generalization power (BIOMAPPER manual). Box-Cox transformation of the environmental variables produced slightly poorer results, and was therefore not used.

GARP predictions were generated using Desktop GARP (Scachetti-Pereira, 2001). For each species we produced 100 prediction maps, using 50% of the calibration data set, with a convergence limit of 0.01 and a maximum of 1000 iterations per model. Ten best subsets were chosen using a soft extrinsic measure of 20% omission and 50% commission. A potential site was included in the environmental envelope if six or more of these 10 maps predicted presence. This threshold was determined based on preliminary analyses that were designed to identify the threshold that maximizes the average value of Kappa (see below).

To enable a standardized evaluation of all models, all predictions were expressed as binary (presence-absence) predictive maps. Threshold values were applied to transform predictions generated by MD, ENFA, and DOMAIN to binary predictions. For each model, the threshold that maximized the average Kappa for all species was selected. Thresholds used were 5.5 for MD, 0.4 for ENFA, and 0.97 for DOMAIN. The percentile range used in BIOCLIM (95%) was also selected to maximize the average value of Kappa.

## Accuracy assessment

Predictions of each model were compared to the validation data set to form a confusion matrix, from which Cohen's Kappa (Cohen, 1960) was calculated. The Kappa statistic defines the accuracy of prediction, relative to the accuracy that might have resulted by chance alone. It ranges from  $-1$  to  $+1$ , where  $+1$  indicates perfect agreement between predictions and observations and values of 0 or less indicate agreement no better than random classification. Following previous evidence indicating that low prevalence may introduce bias to estimates of accuracy based on Kappa (McPherson *et al.*, 2004), only species with four or more presences in the validation data set were included in the analysis.

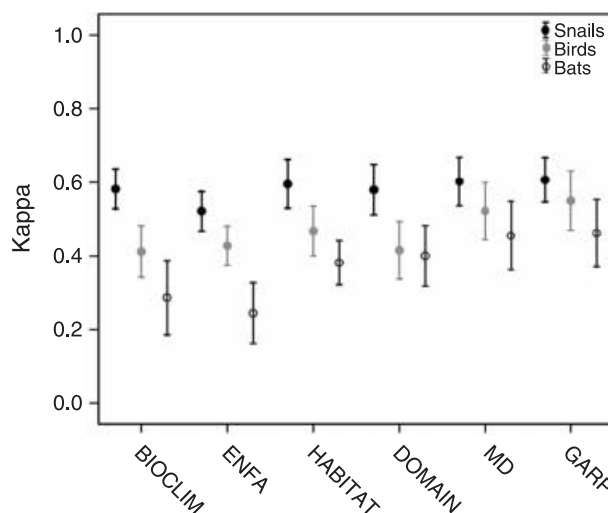
## Statistical analysis

The combined effects of modelling method and taxonomic group on Kappa were tested using repeated measures analysis of variance with modelling method as a within-subject factor and taxonomic group as a between-subject factor. Analyses were performed with and without prevalence (the proportion of the validation sites in which the species was found, McPherson *et al.*, 2004) and tolerance (a measure of niche width ranging from 0 to 1, Hirzel *et al.*, 2004) as covariates. Significance levels (*P*-values) were calculated using exact (multivariate) tests because sphericity could be assumed based on Mauchly's test.

Further analyses were performed to test for correlations between values of Kappa obtained from the various modelling methods. Since high correlations can be achieved even if two methods produce different predictions, we also compared the predictive maps produced for each pair of models at the level of individual sites. Such comparisons were performed separately for each species using the Simple Matching coefficient (Hubalek, 1982). This measure counts the number of sites predicted to have the same status (presence or absence) by two models and divides it by the total number of validation sites. The average proportion of validation sites that were identically classified by any two methods was used as a measure for classification agreement.

## RESULTS

Differences in Kappa among the three taxa were highly consistent: snails showed the highest scores, birds showed intermediate scores, and bats showed the lowest scores in all models (Fig. 1). Variation in Kappa among modelling methods was less consistent but a superiority of GARP and MD over all other models and a low performance of BIOCLIM and ENFA could be detected, particularly for birds and bats (Fig. 1). Repeated measures analysis testing the combined effects of modelling method and taxonomic group on Kappa indicated that the differences among models were highly significant, those among taxa were nearly significant, and the interaction between models and taxa was not significant (Table 1). Similar results were obtained when tolerance and prevalence were included as covariates in the analysis (Table 1). Neither tolerance nor prevalence interacted with the modelling methods, indicating that differences among models



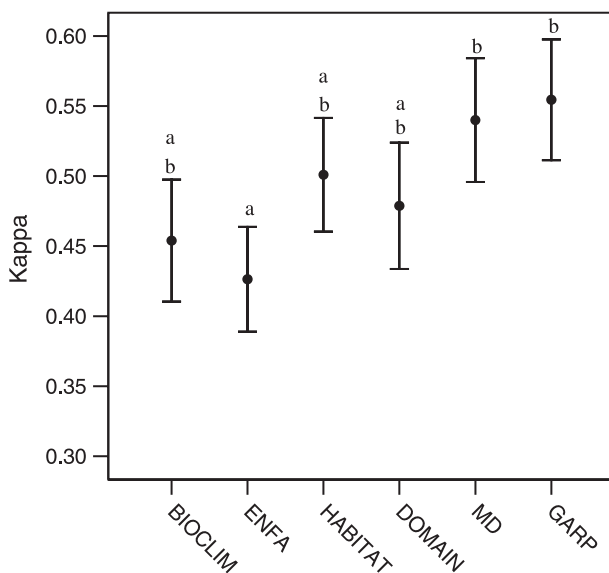
**Figure 1** Effects of modelling method (BIOCLIM, HABITAT, GARP, ENFA, DOMAIN, and MD) and taxonomic group (snails, birds, and bats) on Kappa. Error bars represent mean  $\pm$  1 standard error.

**Table 1** Results (*P*-values) of repeated measures analyses testing the effects of modelling methods (BIOCLIM, HABITAT, MD, DOMAIN, ENFA, GARP) and taxonomic group (snails, birds, bats) on Cohen's Kappa. Analyses were performed with and without tolerance and prevalence as covariates. Statistically significant *P*-values are in bold.

	Without covariates	With covariates
Source of variation		
Model	<b>0.002</b>	<b>0.007</b>
Taxonomic group	0.073	<b>0.048</b>
Model $\times$ taxonomic group	0.679	0.414
Prevalence		0.191
Tolerance		<b>&lt; 0.001</b>
Model $\times$ prevalence		0.165
Model $\times$ tolerance		0.153

were robust to variation in the distribution properties of the species as quantified in this study (Table 1). The results further show that tolerance had a highly significant negative effect on Kappa and that the effect of prevalence was not statistically significant (Table 1).

When data from all taxa were pooled (Fig. 2), GARP and MD showed the highest Kappa scores (average  $\pm$  SD =  $0.554 \pm 0.280$  and  $0.540 \pm 0.286$ , respectively), HABITAT and DOMAIN showed intermediate scores ( $0.501 \pm 0.263$  and  $0.479 \pm 0.292$ , respectively), and BIOCLIM and ENFA showed the lowest scores ( $0.454 \pm 0.282$  and  $0.426 \pm 0.242$ , respectively). Pairwise comparisons revealed statistically significant differences between GARP and BIOCLIM ( $P = 0.008$ ), MD and BIOCLIM ( $P = 0.011$ ), GARP and ENFA ( $P = 0.002$ ), and MD and ENFA ( $P < 0.001$ ). After a conservative Bonferroni correction for multiple (15) comparisons (since the same set of species was used to evaluate each modelling technique), only the latter two differences remain statistically significant.



**Figure 2** Differences in Kappa among modelling methods (BIOCLIM, HABITAT, GARP, ENFA, DOMAIN, and MD) when data for all taxa (snails, birds, and bats) are pooled. Error bars represent mean  $\pm$  1 standard error. Models sharing the same letters do not differ from each other significantly ( $P > 0.05$  following Bonferroni corrections for multiple comparisons).

Kappa scores of predictive maps generated for individual species were positively and significantly correlated over all modelling methods (Table 2). The classification agreement between different methods was higher than 0.7 in all cases except for the GARP method that showed much lower agreement with all other methods (Table 3).

## DISCUSSION

Pearce & Ferrier (2000) suggested that algorithms used to fit species distribution models can be ranked according to their 'function complexity'. They compared different GLM and GAM models and found that increasing model complexity was not necessarily associated with better predictive power. The models examined in our study also differ considerably from each other in their complexity. BIOCLIM can be considered as the simplest model: it assumes a rectilinear environmental envelope, it cannot deal with correlations or interactions between the environmental factors, and only the outer records along each environmental variable are used to define the boundaries of the ecological niche. HABITAT is a more flexible model that better fits the environmental envelope to the distribution of the records in the environmental space, but as with BIOCLIM, only the outer records are taken into account in determining the ecological niche. DOMAIN differs from BIOCLIM and HABITAT in its ability to cope with discontinuity of the species records in the environmental space. Its main limitation is that, for each potential site, only a single record (the nearest neighbour at the environmental space) is used to determine its suitability to the modelled species. ENFA and MD take into account the distribution of all the records in

**Table 2** Pearson correlation coefficients between Kappa scores of predictive maps produced for 42 species of snails, birds, and bats, using six modelling methods (BIOCLIM, HABITAT, MD, DOMAIN, ENFA, and GARP). Asterisks indicate levels of statistical significance (\*\* $P < 0.01$ ; \*\*\* $P < 0.001$ ).

	HABITAT	MD	DOMAIN	ENFA	GARP
BIOCLIM	0.628***	0.731***	0.665***	0.630***	0.658***
HABITAT		0.726***	0.554***	0.449**	0.679***
MD			0.709***	0.757***	0.775***
DOMAIN				0.546***	0.667***
ENFA					0.555***

**Table 3** Average proportions of classification agreement between predictive maps produced for 42 species of snails, birds, and bats, using six modelling methods (BIOCLIM, HABITAT, MD, DOMAIN, ENFA, GARP).

	HABITAT	MD	DOMAIN	ENFA	GARP
BIOCLIM	0.790	0.793	0.731	0.825	0.561
HABITAT		0.826	0.790	0.748	0.425
MD			0.781	0.766	0.453
DOMAIN				0.722	0.456
ENFA					0.428

the environmental space and create elliptic envelopes that are consistent with the assumption of unimodal responses to environmental gradients. MD has the advantage that it is calculated based on the covariance matrix of the environmental variables and, therefore, more directly reflects patterns of correlations between the environmental factors (Farber & Kadmon, 2003). GARP is the most flexible algorithm among the methods examined in this study and has the advantage of an inherent mechanism for iteratively evaluating and improving the rules used for generating the predictive maps. The results of our analysis indicate that the six modelling methods can be categorized into three groups based on their Kappa scores: GARP and MD that were consistently the most accurate models, BIOCLIM and ENFA that showed the lowest accuracy, and HABITAT and DOMAIN that were characterized by intermediate levels of accuracy. Thus, except for the relatively poor performance obtained for ENFA, our results are consistent with the hypothesis that increasing model complexity may contribute to predictive accuracy. Elith *et al.* (2006) analysed a wider spectrum of modelling methods and reached a similar conclusion, although their analysis was based on a different approach.

Similarly to previous studies (e.g. Manel *et al.*, 2001; Guisan & Hofer, 2003; Petit *et al.*, 2003; Rouget *et al.*, 2004), we found considerable variation in predictive accuracy among individual species. For example, the values of Kappa obtained for GARP and MD, the most accurate models in this study, ranged from  $-0.07$  to  $1.00$  and from  $-0.16$  to  $0.93$ , respectively. Kappa scores obtained for individual species were positively and significantly

correlated across models (Table 2), indicating that the differences in predictive accuracy among species were relatively robust to variation in the modelling methods. This conclusion is further supported by the lack of significant interactions between the effects of modelling method and species characteristics.

Our results show that tolerance had a negative effect and prevalence had no effect on predictive accuracy. The observed negative effect of tolerance on predictive accuracy is in agreement with previous studies, showing that distribution ranges of species with restricted ecological niches can be modelled with higher accuracy than those of more generalist species (Pearce *et al.*, 2001; Hepinstall *et al.*, 2002; Stockwell & Peterson, 2002; Kadmon *et al.*, 2003; Berg *et al.*, 2004; Brotons *et al.*, 2004; Segurado & Araujo, 2004). Brotons *et al.* (2004) proposed that species inhabiting a wide range of habitats in a certain area might not be limited by any of the measured predictive factors at the scale at which the models are fitted and may therefore show low predictive accuracy. Another possible explanation for the negative relationship between tolerance and predictive accuracy is that widespread species show local or regional differences in ecological characteristics as a result of local adaptations (Stockwell & Peterson, 2002). Pooling such ecologically variable populations in a single model can result in a weak predictive accuracy.

Previous evidence concerning the effect of prevalence on predictive accuracy is less consistent. Some studies documented positive effects (Boone & Krohn, 1999; Manel *et al.*, 2001; Kadmon *et al.*, 2003; Berg *et al.*, 2004), other studies documented negative effects (Pearce *et al.*, 2001; Stockwell & Peterson, 2002; Guisan & Hofer, 2003; Rouget *et al.*, 2004; Segurado & Araujo, 2004), in some studies no relationship was found (Brotons *et al.*, 2004), and in others the effect of prevalence varied among different biological groups (Pearce & Ferrier, 2000). In our analysis the effect of prevalence on Kappa was not statistically significant but it should be taken into account that we excluded species with relatively low prevalence (< 15%) from our analysis. This restriction was adopted following previous results indicating that low prevalence may lead to statistical bias in estimates of accuracy based on the Kappa statistic (McPherson *et al.*, 2004).

Several studies have argued that limited dispersal may prevent species to occur in areas with suitable ecological conditions and should therefore reduce the accuracy of predictive maps produced by species distribution models (Pearson & Dawson, 2003; Peterson, 2003; Hampe, 2004). Based on this hypothesis, one would expect that predictive maps produced for snails would be less accurate than those produced for birds and bats, which are characterized by much higher dispersal ability. Our results do not support this prediction. Actually, among the three taxa examined in this study, snails showed the highest values of Kappa for all modelling algorithms (Fig. 1). One possible explanation for this result is that snails were characterized by a lower tolerance (average  $\pm$  SD =  $0.52 \pm 0.18$ ) than either birds ( $0.74 \pm 0.14$ ) or bats ( $0.60 \pm 0.26$ ). However, the differences in Kappa among taxa remained statistically significant even after controlling for the effect of tolerance (Table 1). Another potential explanation is that intensive dispersal may blur niche relationships (and thus, reduce the predictive power of distribution models) by causing

species to occur in unsuitable areas, either temporarily, or for long time periods as sink populations (Pulliam, 2000). Such 'over dispersal' may deteriorate, rather than enhance, the predictive power of distribution models. A related mechanism by which high dispersal ability may deteriorate predictions of species distribution models is the greater ability of mobile species to respond to local, short-term fluctuations in the environment. Such short-term adaptive responses reduce the correlation between distribution patterns and indices expressing long-term climatic conditions, and therefore, deteriorate the predictive power of models based on mean values of rainfall and temperature. Differences in home range size may have further contributed to the observed differences in Kappa because species with small range size (like snails) are more likely to exhibit correlations with local environmental conditions than species characterized by large range size (e.g. birds and bats).

The relatively low values of predictive accuracy obtained for bats suggest that factors other than climate are important in determining distribution patterns of this group at the spatial scales examined in this study. One factor that is known to have considerable effects on patterns of bat distribution and was not included in our models is lithology, which influences the availability of caves used as roosting sites (Altringham, 1996). Anthropogenic disturbances are another factor that had substantial effects on patterns of bat distribution in Israel during the last decades (Shalmon, 2002). The main sources of such disturbances are cave visitations, secondary poisoning, and fumigation of caves against fruit bats that were considered to be agricultural pests. Yom-Tov & Mendelsohn (1988) noted that bats are particularly sensitive to anthropogenic disturbances and estimated that most bats of the Mediterranean region of Israel do not occupy their potential distribution ranges because of such disturbances.

## CONCLUSIONS

In spite of considerable differences in the complexity of the modelling algorithm, the six models examined in this study showed relatively small (though statistically significant) differences in predictive accuracy. On the other hand, our results show that distribution properties of the species may have considerable effects on predictive accuracy. These findings suggest that properties of the species may have a greater impact on predictive accuracy than differences in modelling techniques. Similar results were reported in studies comparing other modelling techniques (e.g. Manel *et al.*, 1999; Elith & Burgman, 2002; Thuiller *et al.*, 2003; Berg *et al.*, 2004). We therefore recommend that future studies should devote more efforts to the identification of ecological and geographical factors that influence the accuracy of species distribution models (e.g. Segurado & Araujo, 2004). We also recommend that, in addition to developing more sophisticated modelling algorithms, future studies should incorporate factors such as disturbance, dispersal limitation, and biotic interactions, which are known to affect patterns of species distribution but are usually ignored in distribution modelling. Several studies have already taken this approach (Leathwick & Austin, 2001; Anderson *et al.*, 2002; Thomas *et al.*, 2004).

## ACKNOWLEDGEMENTS

We thank H-U. Schnitzler and A. Denzinger for providing us the echolocation monitoring equipment and for their technical support, M. B. Fenton for valuable advice and field assistance, A. Ben-Nun for assistance with GIS issues, O. Farber for his help in programming, and A. T. Peterson and S. Ferrier for valuable comments on the manuscript. The research was partially supported by the Israel Nature and Parks Authority and The Israel Science Foundation (grant no. 545/03).

## REFERENCES

- Altringham, J.D. (1996) *Bats biology and behaviour*. Oxford University Press, Oxford.
- Anderson, R.P., Peterson, A.T. & Gomez-Laverde, M. (2002) Using niche-based GIS modeling to test geographic predictions of competitive exclusion and competitive release in South American pocket mice. *Oikos*, **98**, 3–16.
- Araujo, M.B., Cabeza, M., Thuiller, W., Hannah, L. & Williams, P.H. (2004) Would climate change drive species out of reserves? An assessment of existing reserve-selection methods. *Global Change Biology*, **10**, 1618–1626.
- Berg, A., Gardenfors, U. & von Proschwitz, T. (2004) Logistic regression models for predicting occurrence of terrestrial molluscs in southern Sweden — importance of environmental data quality and model complexity. *Ecography*, **27**, 83–93.
- Bisby, F.A. (2000) The quiet revolution: biodiversity informatics and the internet. *Science*, **289**, 2309–2312.
- Boone, R.B. & Krohn, W.B. (1999) Modeling the occurrence of bird species: Are the errors predictable? *Ecological Applications*, **9**, 835–848.
- Bowker, G.C. (2000) Mapping biodiversity. *International Journal of Geographic Information Systems*, **14**, 739–754.
- Box, E.O., Crumpacker, D.W. & Hardin, E.D. (1993) A climatic model for location of plant species in Florida, USA. *Journal of Biogeography*, **20**, 629–644.
- Brotons, L., Thuiller, W., Araujo, M.B. & Hirzel, A.H. (2004) Presence-absence versus presence-only modelling methods for predicting bird habitat suitability. *Ecography*, **27**, 437–448.
- Busby, J.R. (1986) A biogeoclimatic analysis of *Nothofagus cunninghamii* (Hook.) Oerst. in southeastern Australia. *Australian Journal of Ecology*, **11**, 1–7.
- Busby, J.R. (1991) BIOCLIM — A Bioclimate Analysis and Prediction System. Nature Conservation: Cost effective biological surveys and data analysis. *Nature conservation* (ed. by C.R. Margules and M.P. Austin), pp. 64–68. CSIRO, Melbourne.
- Carpenter, G., Gillison, A.N. & Winter, J. (1993) DOMAIN — A flexible modeling procedure for mapping potential distributions of plants and animals. *Biodiversity and Conservation*, **2**, 667–680.
- Cohen, J. (1960) A coefficient of agreement of nominal scales. *Educational and Psychological Measurement*, **20**, 37–46.
- Dettmers, R. & Bart, J. (1999) A GIS modeling method applied to predicting forest songbird habitat. *Ecological Applications*, **9**, 152–163.
- Eeley, H.A.C., Lawes, M.J. & Piper, S.E. (1999) The influence of climate change on the distribution of indigenous forest in KwaZulu-Natal, South Africa. *Journal of Biogeography*, **26**, 595–617.
- Elith, J. & Burgman, M.A. (2002) Predictions and their validation: rare plants in the central highlands, Victoria, Australia. *Predicting species occurrences: issues of accuracy and scale* (ed. by J.M. Scott, P.J. Heglund, M.L. Morrison, J.B. Haufler, M.G. Raphael, W.A. Wall and F.B. Samson), pp. 303–313. Island Press, Washington, DC.
- Elith, J., Graham, C.H., Anderson, R.P., Dudik, 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.C., Peterson, A.T., Phillips, S.J., Richardson, K.S., Scachetti-Pereira, R., Schapire, R.E., Soberon, J., Williams, S., Wisz, M.S. & Zimmermann, N.E. (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, **29**, 129–151.
- 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.
- Farber, O. & Kadmon, R. (2003) Assessment of alternative approaches for bioclimatic modeling with special emphasis on the Mahalanobis distance. *Ecological Modelling*, **160**, 115–130.
- Ferrier, S., Watson, G., Pearce, J. & Drielsma, M. (2002) Extended statistical approaches to modelling spatial pattern in biodiversity in north-east New South Wales. I. Species-level modelling. *Biodiversity and Conservation*, **11**, 2275–2307.
- Guisan, A. & Hofer, U. (2003) Predicting reptile distributions at the mesoscale: relation to climate and topography. *Journal of Biogeography*, **30**, 1233–1243.
- Guisan, A. & Thuiller, W. (2005) Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, **8**, 993–1009.
- Guisan, A. & Zimmermann, N.E. (2000) Predictive habitat distribution models in ecology. *Ecological Modelling*, **135**, 147–186.
- Hampe, A. (2004) Bioclimate envelope models: what they detect and what they hide. *Global Ecology and Biogeography*, **13**, 469–471.
- Heller, J. (1988) The biogeography of the land snails of Israel. *The zoogeography of Israel* (ed. by Y. Yom-Tov and E. Tchernov), pp. 325–353. Kluwer Academic Publishers, Dordrecht, the Netherlands.
- Hepinstall, J.W., Krohn, W.B. & Sader, S.A. (2002) Effects of niche width on the performance and agreement of avian habitat models. *Predicting species occurrences: issues of accuracy and scale* (ed. by J.M. Scott, P.J. Heglund, M.L. Morrison, J.B. Haufler, M.G. Raphael, W.A. Wall and F.B. Samson), pp. 593–606. Island Press, Washington, DC.
- Hirzel, A.H., Hausser, J., Chessel, D. & Perrin, N. (2002) Ecological niche factor analysis: How to compute habitat-suitability maps without absence data? *Ecology*, **83**, 2027–2036.
- Hirzel, A.H., Hausser, J. & Perrin, N. (2004) *Biomapper 3.1*.

- Division of Conservation Biology, University of Bern, Bern, Switzerland. URL: <http://www.unil.ch/biomapper>.
- Hubalek, Z. (1982) Coefficients of association and similarity based on binary (presence-absence) data: an evaluation. *Biological Reviews*, **57**, 669–689.
- Hutchinson, G.E. (1957) Concluding remarks. *Cold Spring Harbor Symposium on Quantitative Biology*, **22**, 415–427.
- Kadmon, R. & Danin, A. (1997) Floristic variation in Israel: a GIS analysis. *Flora*, **192**, 341–345.
- Kadmon, R., Farber, O. & Danin, A. (2003) A systematic analysis of factors affecting the performance of climatic envelope models. *Ecological Applications*, **13**, 853–867.
- Kadmon, R. & Heller, J. (1998) Modelling faunal responses to climatic gradients with GIS: land snails as a case study. *Journal of Biogeography*, **25**, 527–539.
- Kurtzman, D. & Kadmon, R. (1999) Mapping of temperature variables in Israel: a comparison of different interpolation methods. *Climate Research*, **13**, 33–43.
- Leathwick, J.R. & Austin, M.P. (2001) Competitive interactions between tree species in New Zealand's old-growth indigenous forests. *Ecology*, **82**, 2560–2573.
- Loiselle, B.A., Howell, C.A., Graham, C.H., Goerck, J.M., Brooks, T., Smith, K.G. & Williams, P.H. (2003) Avoiding pitfalls of using species distribution models in conservation planning. *Conservation Biology*, **17**, 1591–1600.
- Manel, S., Dias, J.M., Buckton, S.T. & Ormerod, S.J. (1999) Alternative methods for predicting species distribution: an illustration with Himalayan river birds. *Journal of Applied Ecology*, **36**, 734–747.
- Manel, S., Williams, H.C. & Ormerod, S.J. (2001) Evaluating presence-absence models in ecology: the need to account for prevalence. *Journal of Applied Ecology*, **38**, 921–931.
- Margules, C.R. & Pressey, R.L. (2000) Systematic conservation planning. *Nature*, **405**, 243–253.
- Martinez-Meyer, E. (2005) Climate change and biodiversity: some considerations in forecasting shifts in species' potential distributions. *Biodiversity Informatics*, **2**, 42–55.
- McPherson, J.M., Jetz, W. & Rogers, D.J. (2004) The effects of species' range sizes on the accuracy of distribution models: ecological phenomenon or statistical artefact? *Journal of Applied Ecology*, **41**, 811–823.
- Nix, H.A. (1986) A biogeographic analysis of Australian elapid snakes. *Atlas of Elapid Snakes of Australia*, pp. 4–15. Australian Government Publications Service, Canberra.
- O'Farrell, M.J. & Gannon, W.L. (1999) A comparison of acoustic versus capture techniques for the inventory of bats. *Journal of Mammalogy*, **80**, 24–30.
- Ortega-Huerta, M.A. & Peterson, A.T. (2004) Modelling spatial patterns of biodiversity for conservation prioritization in North-eastern Mexico. *Diversity and Distribution*, **10**, 39–54.
- Ottaviani, D., Lasinio, G.J. & Boitani, L. (2004) Two statistical methods to validate habitat suitability models using presence-only data. *Ecological Modelling*, **179**, 417–443.
- Pearce, J. & Ferrier, S. (2000) Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling*, **133**, 225–245.
- Pearce, J., Ferrier, S. & Scotts, D. (2001) An evaluation of the predictive performance of distributional models for flora and fauna in north-east New South Wales. *Journal of Environmental Management*, **62**, 171–184.
- 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.
- Pearson, R.G., Thuiller, W., Araujo, M.B., Martinez-Meyer, E., Brotons, L., McClean, C., Miles, L., Segurado, P., Dawson, T.P. & Lees, D.C. (2006) Model-based uncertainty in species range prediction. *Journal of Biogeography*, **33**, 1704–1711.
- Peterson, A.T. (2003) Predicting the geography of species' invasions via ecological niche modeling. *Quarterly Review of Biology*, **78**, 419–433.
- Peterson, A.T. & Robins, C.R. (2003) Using ecological-niche modeling to predict barred owl invasions with implications for spotted owl conservation. *Conservation Biology*, **17**, 1161–1165.
- Petit, S., Chamberlain, D., Haysom, K., Pywell, R., Vickery, J., Warman, L., Allen, D. & Firbank, L. (2003) Knowledge-based models for predicting species occurrence in arable conditions. *Ecography*, **26**, 626–640.
- Ponder, W.F., Carter, G.A., Flemons, P. & Chapman, R.R. (2001) Evaluation of museum collection data for use in biodiversity assessment. *Conservation Biology*, **15**, 648–657.
- Prendergast, J.R., Quinn, R.M. & Lawton, J.H. (1999) The gaps between theory and practice in selecting nature reserves. *Conservation Biology*, **13**, 484–492.
- Pulliam, H.R. (2000) On the relationship between niche and distribution. *Ecology Letters*, **3**, 349–361.
- Rouget, M., Richardson, D.M., Nel, J.L., Le Maitre, D.C., Egoh, B. & Mgid, T. (2004) Mapping the potential ranges of major plant invaders in South Africa, Lesotho and Swaziland using climatic suitability. *Diversity and Distributions*, **10**, 475–484.
- Sánchez-Cordero, V., Cirelli, V., Munguía, M. & Sarkar, S. (2005) Place prioritization for biodiversity content using species ecological niche modeling. *Biodiversity Informatics*, **2**, 11–23.
- Scachetti-Pereira, R. (2001) *Desktop GARP, version 1.1.4*. The University of Kansas Biodiversity Research Center, Lawrence, Kansas.
- Scott, J.M., Heglund, P.J., Morrison, M.L., Haufler, J.B., Raphael, M.G., Wall, W.A. & Samson, F.B., eds (2002) *Predicting species occurrences: issues of scale and accuracy*. Island Press, Washington, DC.
- Segurado, P. & Araujo, M.B. (2004) An evaluation of methods for modelling species distributions. *Journal of Biogeography*, **31**, 1555–1568.
- Shalmon, B. (2002) Mammals. *Endangered species in Israel, red list of threatened animals vertebrates* (ed. by A. Dolev and A. Perevolotsky), pp. 201–282. Nature and Parks Authority and the Society for the Preservation of Nature, Jerusalem. (in Hebrew).
- Shao, G.F. & Halpin, P.N. (1995) Climatic controls of eastern north American coastal tree and shrub distributions. *Journal of Biogeography*, **22**, 1083–1089.



- Shirihai, H. (1996) *The birds of Israel*. Academic Press, London, UK.
- Steinitz, O., Heller, J., Tsoar, A., Rotem, D. & Kadmon, R. (2005) Predicting regional patterns of similarity in species composition for conservation planning. *Conservation Biology*, **19**, 1978–1988.
- Stockwell, D. & Peters, D. (1999) The GARP modelling system: problems and solutions to automated spatial prediction. *International Journal of Geographical Information Science*, **13**, 143–158.
- Stockwell, D.R.B. & Peterson, A.T. (2002) Effects of sample size on accuracy of species distribution models. *Ecological Modelling*, **148**, 1–13.
- 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.
- Thuiller, W., Araujo, M.B. & Lavorel, S. (2003) Generalized models vs. classification tree analysis: Predicting spatial distributions of plant species at different scales. *Journal of Vegetation Science*, **14**, 669–680.
- Vaughan, I.P. & Ormerod, S.J. (2005) The continuing challenges of testing species distribution models. *Journal of Applied Ecology*, **42**, 720–730.
- Walker, P.A. & Cocks, K.D. (1991) HABITAT: a procedure for modelling a disjoint environmental envelope for a plant or animal species. *Global Ecology and Biogeography Letters*, **1**, 108–118.
- Williams, P.H., Margules, C.R. & Hilbert, D.W. (2002) Data requirements and data sources for biodiversity priority area selection. *Journal of Biosciences*, **27**, 327–338.
- Yom-Tov, Y. & Kadmon, R. (1998) Analysis of the distribution of insectivorous bats in Israel. *Diversity and Distributions*, **4**, 63–70.
- Yom-Tov, Y. & Mendelssohn, H. (1988) Changes of the distribution and abundance of vertebrates during the 20th century in Israel. *The zoogeography of Israel* (ed. by Y. Yom-Tov and E. Tchernov), pp. 515–548. Kluwer Academic Publishers, Dordrecht.
- Yom-Tov, Y. & Werner, Y.L. (1996) Environmental correlates of geographical distribution of terrestrial vertebrates in Israel. *Israel Journal of Zoology*, **42**, 307–315.
- Zaniewski, A.E., Lehmann, A. & Overton, J.M.C. (2002) Predicting species spatial distributions using presence-only data: a case study of native New Zealand ferns. *Ecological Modelling*, **157**, 261–280.