

# Consequences of spatial autocorrelation for niche-based models

P. SEGURADO,\* M. B. ARAÚJO† and W. E. KUNIN‡

\*Macroecology and Conservation Unit, University of Évora, Estrada dos Leões, 7000–730 Évora, Portugal;

†Department of Biodiversity and Evolutionary Biology, National Museum of Natural Sciences, CSIC, CIJosé Gutiérrez Abascal, 2, 28006 Madrid, Spain; and ‡Earth and Biosphere Institute, IICB, University of Leeds, Leeds LS2 9JT, UK

## Summary

1. Spatial autocorrelation is an important source of bias in most spatial analyses. We explored the bias introduced by spatial autocorrelation on the explanatory and predictive power of species' distribution models, and make recommendations for dealing with the problem.

2. Analyses were based on the distribution of two species of freshwater turtle and two virtual species with simulated spatial structures within two equally sized areas located on the Iberian Peninsula. Sequential permutations of environmental variables were used to generate predictor variables that retained the spatial structure of the original variables. Univariate models of species' distributions using generalized linear models (GLM), generalized additive models (GAM) and classification tree analysis (CTA) were fitted for each variable permutation. Variation of accuracy measures with spatial autocorrelation of the original predictor variables, as measured by Moran's  $I$ , was analysed and compared between models. The effects of systematic subsampling of the data set and the inclusion of a contagion term to deal with spatial autocorrelation in models were assessed with projections made with GLM, as it was with this method that estimates of significance based on randomizations were obtained.

3. Spatial autocorrelation was shown to represent a serious problem for niche-based species' distribution models. Significance values were found to be inflated up to 90-fold.

4. In general, GAM and CTA performed better than GLM, although all three methods were vulnerable to the effects of spatial autocorrelation.

5. The procedures utilized to reduce the effects of spatial autocorrelation had varying degrees of success. Subsampling was partially effective in avoiding the inflation effect, whereas the inclusion of a contagion term fully eliminated or even overcompensated for this effect. Direct estimation of probability using variable simulations was effective, yet seemed to show some residual spatial autocorrelation effects.

6. *Synthesis and applications.* Given the expected inflation in the estimates of significance when analysing spatially autocorrelated variables, these need to be adjusted. The reliability and value of niche-based distribution models for management and other applied ecology purposes can be improved if certain techniques and procedures, such as the null model approach recommended in this study, are implemented during the model-building process.

*Key-words:* AUC, classification tree analysis, freshwater turtles, generalized additive models, generalized linear models, Iberian Peninsula, likelihood ratio statistics, model validation, toroidal shifts

*Journal of Applied Ecology* (2006) **43**, 433–444

doi: 10.1111/j.1365-2664.2006.01162.x

## Introduction

Niche-based models are familiar tools used to explain and predict species' spatial distributions (Walker 1990; Pereira & Itami 1991; Austin *et al.* 1996; Manel *et al.* 1999). Recently, niche models have received increased attention, in part because of the need to predict species' range shifts under future climate-change scenarios (Guisan & Theurillat 2000; Midgley *et al.* 2002; Peterson *et al.* 2002; Thomas *et al.* 2004; Thuiller *et al.* 2005; Araújo, Thuiller & Pearson 2006). However, there are a number of unresolved methodological issues requiring further enquiry. One example is the problem of non-independence between data used for calibration of models and that used for validation (Araújo *et al.* 2005a). Non-independence is often the result of using spatially autocorrelated data to calibrate and validate the models, and one of the consequences is that the perceived ability of models to make realistic predictions in space (Randin *et al.* 2006) and time (Araújo *et al.* 2005a) may be inflated. This problem may be greater than previously anticipated, as illustrated by studies showing high levels of intermodel variability in projections of species' range shifts under climate change scenarios (Thuiller 2004; Thuiller *et al.* 2004; Araújo *et al.* 2005b; Araújo, Thuiller & Pearson 2006). We have addressed the biases in model predictions that arise from using different procedures of model adjustment and validation under varying levels of spatial dependencies among predictor and response variables.

Sample size is a crucial parameter in the outcome of classical hypothesis testing as it determines the necessary degrees of freedom for pattern detection. In spatial analyses the simple count of sample units is not always an adequate estimator of effective sample size. For example, if the values of a variable depend on the distance between sample points, a set of closely spaced observations effectively provides less information than the same number of observations more widely separated in space. Such spatial dependency between values is termed spatial autocorrelation (SA; Cliff & Ord 1973) and its causes and consequences have been the focus of much research (Legendre 1993; Koenig & Knops 1998; Lennon 2000; Dale & Fortin 2002; Fortin & Payette 2002; Legendre *et al.* 2002). SA leads to an overestimation of the effective sample size (leading to pseudoreplication), inflating the statistical significance of measured spatial relationships and consequently increasing the likelihood of type I errors (false positives). There is a serious possibility that previous analyses that used correlative approaches might be flawed because of 'red herrings' generated by SA (Lennon 2000), with both the estimated predictive power and the choice of variables being seriously biased. Because of the potential importance of such biases, many methods have been developed to help account for SA within models, including a priori procedures at the level of sampling design (Harrison 1997; Legendre *et al.* 2002; Legendre *et al.* 2004), modifications at the level of model adjustment

(Keitt *et al.* 2002; Lichstein *et al.* 2002) and a posteriori procedures, such as the use of correction factors, to improve statistical accuracy of models (Dutilleul 1993; Legendre *et al.* 2002).

Niche-based models use several alternative techniques to summarize relationships between species occurrences and environmental variation (Guisan & Zimmermann 2000; Segurado & Araújo 2004), usually in the context of spatial and temporal predictions. Other designations for this family of models can be found in the literature, such as habitat models (Guisan & Zimmermann 2000), species distribution models (Olden, Jackson & Peres-Neto 2002), bioclimatic envelope models (Pearson & Dawson 2003) and presence/absence models (Fielding & Bell 1997). Although authors acknowledge the importance of SA, they often disregard or minimize the extent to which the presence of spatially autocorrelated data might affect the explanatory power and predictive accuracy of models (i.e. the ability in which models calibrated in one set represent observations in an independent set; for discussion see Araújo *et al.* 2005a; Randin *et al.* 2006). This could be a serious shortcoming in models as species' occurrences tend to be aggregated at most spatial scales, and the more aggregated species' occurrences are, the more likely it is that environmental variables will show some explanatory power simply because of the fact that environmental conditions tend to be more similar at neighbouring sites. Indeed, the strength of the correlation between variables has been shown to be increasingly more pronounced as SA grows stronger, whereas unbiased correlations are produced when at least one variable exhibits no SA (Lennon 2000).

Model generalization (e.g. variable selection in stepwise logistic regression, pruning of classification trees and stopping rules in artificial neural networks) are common procedures to avoid overadjustments to calibration data and are designed to increase the predictive power of models (Franklin 1998; Pearce & Ferrier 2000; Thuiller 2003). In the case of regression-based techniques this procedure implies that an assessment of the explanatory power of variables is made. The problem is that the inflation of explanatory power for spatially autocorrelated variables makes them, a priori, disproportionately likely to be selected in the final models. This is made at the expense of selecting potentially more important variables with lower SA. Therefore, variable selection procedures can be an additional source of bias in model fitting.

Patterns of species' distributions may be spatially autocorrelated because of contagious population dynamics and historical factors, but they may also be the result of spatial structure among environmental predictors (Storch *et al.* 2003). In fact, species and the environment may share spatial structure because of the effect of spatially structured environmental predictors and non-environmental contingencies that may or may not be related amongst them (Borcard, Legendre & Drapeau 1992). If part of the spatial structure in the species' data is shared by the environmental data, knowing the relative weight of each item that contributes to the observed spatial structure

is an important challenge when testing causal hypotheses to data (Borcard, Legendre & Drapeau 1992; Storch *et al.* 2003). A common procedure to cancel the effect of spatial structure of species' occurrences is to incorporate a term for SA into the analysis (Smith 1994; Augustin, Muggleston & Buckland 1996; Araújo & Williams 2000; Keitt *et al.* 2002; Segurado & Araújo 2004), usually a measure of contagion that encompasses the effect of spatial neighbourhood in the statistical test.

An alternative procedure to avoid pseudoreplication is to subsample the original species' distribution data, usually by adopting a systematic scheme that constrains observations to be spaced far enough from each other (Gates *et al.* 1994; Brito, Crespo & Paulo 1999; Guisan & Theurillat 2000). This method has the disadvantage of not using all the available information and thereby artificially limiting sample size, a procedure that may have serious consequences for the predictive performance of models (Araújo *et al.* 2005b).

The effect of SA on correlation and linear regression significance values has been tested elsewhere using artificially generated variables with known spatial structures (Lennon 2000). In the context of niche-based modelling the extent to which SA in the response and predictor variables influences model performance is poorly known. In particular, the effect of SA in the validation data set used to estimate predictive power of models has never been assessed. In this study the effects of SA in species' distribution models were quantified using a null model approach with test variables generated from randomizations of spatially structured environmental data. First, the overall effect of SA on measures of model performance was assessed. Secondly, we explored whether different modelling techniques, differing in the response functions used, were equally sensitive to the effect of SA. We also quantified the effect of spatial autocorrelation in the predictive ability

of covariates entering models of species' distributions using both resubstitution and data set partitioning (Olden, Jackson & Peres-Neto 2002; Araújo *et al.* 2005a). Finally, the effectiveness of different approaches to reduce undesirable effects of SA was evaluated.

## Materials and methods

### DATA

We used the distributions of two species of freshwater turtles, the Mediterranean pond turtle *Mauremys leprosa* (Schweiger, 1812) and the European pond turtle *Emys orbicularis* (Linnaeus, 1758), in two equally sized rectangular areas on the Iberian Peninsula (Fig. 1). In each rectangular area, data on species' distributions were located in  $66 \times 21$  universal transverse mercator (UTM)  $10 \times 10$ -km grid cells. The main criteria for delimitation of the two rectangular areas were to maximize (i) the geographical extent of the rectangles and (ii) the geographical distance between them, in order to ensure the greatest feasible amount of information while retaining high levels of spatial independence between rectangles. Distribution data were compiled from four different sources: the updated atlas of the Portuguese herpetofauna (Godinho *et al.* 1999), the atlas and Red Data Book of the amphibian and reptiles of Spain (Keller & Andreu 2002; da Silva 2002), the UNIBA database (Alentejo's Biodiversity Database Unit; www.cea.uevora.pt/umc, 2006) and P. Segurado (unpublished data). The spatial structure of the two species' distributions in the study region is distinct: *E. orbicularis* is found in fewer grid cells and it is widely scattered compared with *M. leprosa*, which has more clumped distributions. There are also regional differences in occupancy patterns: in the western rectangle occurrences of *M. leprosa* are more densely distributed and

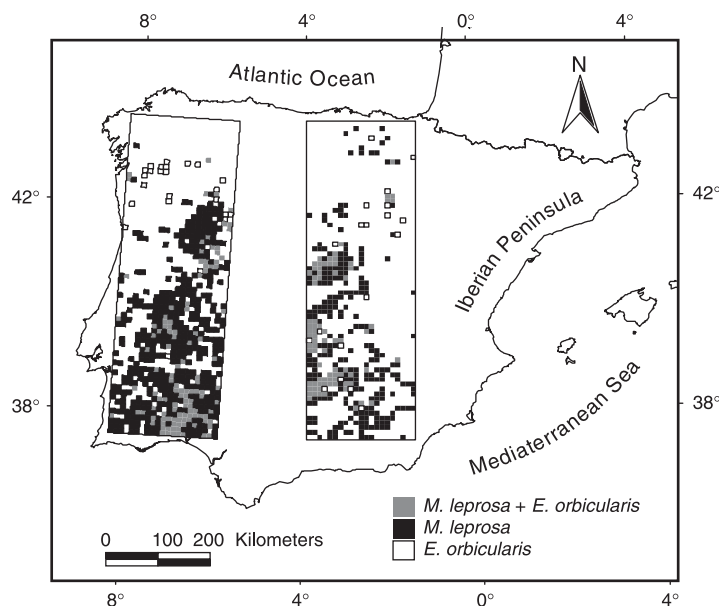


Fig. 1. Location of the rectangular areas used in the analysis (10-km linear resolution).

the occurrences of *E. orbicularis* are slightly more scattered than in the eastern rectangle. Simulated distributions of two species with distinctive spatial structures were also generated from the *M. leprosa* database creating a random distribution and a clumped distribution. The random distribution was obtained by randomly assigning the position of species' presences, while the clumped distribution was obtained by joining all species' presences into a single contiguous block. The species' distributions, as well as the simulated distributions, were used as response variables in the analyses.

Environmental data included climatic and topographic information (Table 1) resampled at the same grid resolution as the species' occurrence data. There is a danger that species might respond indirectly to topography, which would limit the models' predictive powers. However, topographic variables have the potential to summarize important surrogate predictor variables, such as habitat availability for freshwater turtles, that are not captured by the available variables. For example, both species show a preference for still and slow-moving water habitats and therefore their distributions might respond directly to slope. Climate data included 11 variables compiled from point data with 10-minute resolution (New, Hulme & Jones 2000). A randomly generated predictor variable showing negligible SA was also included in the analyses. The two rectangular areas did not differ considerably in their environmental range, although the western rectangle included a slightly wider gradient range for some variables.

## PATTERN GENERATION

The effect of SA was evaluated by generating simulated patterns with known and fixed spatial structures. There are two main categories of such pattern generation: (i) fully synthetic patterns, generated purely from mathematical principles such as the method based on the inverse discrete Fourier transform (Lennon 2000); and (ii) patterns generated from real data using restricted or sequential permutations of real patterns (Fortin & Jacquez 2000) or using more elaborate approaches such as the random patterns implemented by Roxburgh & Chesson (1998) and the patch model proposed by Watkins & Wilson (1992).

In this study we employed sequential permutations of environmental variables based on toroidal shifts (Palmer & Van der Maarel 1995; Fortin & Jacquez 2000; Dale & Fortin 2002; Fortin & Payette 2002; Storch *et al.* 2003) to generate patterns from the original environmental variables. With this randomization technique, coordinates of the original variable are moved by a common random factor in every geographical direction; cells that are shifted beyond one side of the range of real coordinates are moved to the opposite side of the range. This randomization assures that the main spatial structure is maintained. This technique is more straightforward, easier to implement and computationally more efficient than alternative methods. Its main drawback is that it can create unrealistic environmental patterns with abrupt orthogonal lines originated by the

**Table 1.** Variables included in the analysis, Moran's *I*-values for the original variables and mean Moran's *I*-values for the 1000 toroidal permutations (SD values inside parentheses)

Variable	Eastern area		Western area	
	Original Moran's <i>I</i>	Mean Moran's <i>I</i>	Original Moran's <i>I</i>	Mean Moran's <i>I</i>
<i>Mauremys leprosa</i>	0.416	–	0.543	–
<i>Emys orbicularis</i>	0.386	–	0.264	–
Species with clumped distribution	0.959	–	0.975	–
Species with a random distribution	–0.016	–	0.009	–
Predictor variables				
Altitude, mean (Alt)	0.817	0.764 (0.012)	0.882	0.819 (0.019)
Altitude, minimum (Altmin)	0.843	0.796 (0.013)	0.847	0.775 (0.024)
Altitude, maximum (Altmax)	0.784	0.739 (0.011)	0.889	0.837 (0.015)
Slope, mean (Slope)	0.733	0.729 (0.011)	0.832	0.787 (0.012)
Hill shade, mean (Hill)	0.292	0.284 (0.007)	0.181	0.171 (0.007)
Hill shade, minimum (Hillmin)	0.657	0.658 (0.008)	0.752	0.716 (0.013)
Hill shade, maximum (Hillmax)	0.626	0.623 (0.006)	0.716	0.681 (0.012)
Mean annual temperature (Tann)	0.973	0.943 (0.016)	0.981	0.937 (0.013)
Mean temperature of the coldest month (Mtc)	0.945	0.929 (0.014)	0.971	0.909 (0.015)
Mean temperature of the warmer month (Mtw)	0.987	0.952 (0.016)	0.981	0.943 (0.014)
Mean annual growing degree days (Gdd)	0.975	0.944 (0.017)	0.983	0.938 (0.012)
Mean annual global net radiation (Rann)	0.957	0.906 (0.022)	0.958	0.861 (0.016)
Mean annual evapotranspiration/potential evapotranspiration (A2P)	0.950	0.915 (0.017)	0.975	0.914 (0.011)
Mean annual precipitation sum (Pann)	0.901	0.877 (0.023)	0.968	0.908 (0.023)
Mean winter precipitation sum (Pwin)	0.895	0.890 (0.021)	0.966	0.898 (0.025)
Mean summer precipitation sum (Psum)	0.924	0.879 (0.020)	0.971	0.932 (0.016)
Mean spring precipitation sum (Pspr)	0.916	0.872 (0.018)	0.966	0.914 (0.020)
Mean autumn precipitation sum (Paut)	0.888	0.879 (0.021)	0.968	0.900 (0.024)
Random environment	0.001	0.000 (0.003)	–0.018	–0.014 (0.003)

shifted edges. We assumed that this feature would not have an effect in the analysis because the existence of linear edges in species' distributions is unlikely. As toroidal shifts can be too liberal (Fortin & Jacquez 2000), an image reflection and a 180-degree rotation were initially performed for each variable. This procedure ensured that even toroids involving only small shifts would differ substantially from the original pattern, thus making the pattern simulation more conservative.

Moran's  $I$  statistics was used to estimate general patterns of spatial dependency of variables. In order to evaluate the degree to which the original spatial structure was maintained after randomizations, 1000 toroidal shifts were run and for each permutation the Moran's  $I$  was calculated. The distribution of the resulting values was compared with the Moran's  $I$ -value of the original variables.

#### ASSESSING THE EFFECT OF SPATIAL AUTOCORRELATION

Three modelling techniques, differing in their ability to model complex response shapes, were used to relate species' distributions to each of the 1000 toroidal shifts of the environmental variables: (i) generalized linear models, (ii) generalized additive models and (iii) classification tree analysis. Generalized linear models (GLM; McCullagh & Nelder 1983) are generalizations of the classical linear regression allowing error distributions other than the normal distribution; here a binomial error distribution was assumed (logistic regression). Generalized additive models (GAM; Hastie & Tibshirani 1990) are semi-parametric forms of GLM that use smooth functions instead of the usual regression coefficients. GAM were fitted using cubic splines as the smooth function and assuming a binomial error distribution. Classification tree analysis (CTA; Clark & Pregibon 1992) is a non-parametric technique that is based on recursive partitions of the dimensional space defined by the predictor variables into groups that are as homogeneous as possible for the response variable. We used a recursive algorithm that successively splits the data into binary branches by choosing the splits that cause the maximum reduction of the residual deviance.

For each permutation, the model classification accuracy was measured by calculating the receiver operational characteristic (ROC) curve and summing the area under that curve (AUC; Fielding & Bell 1997). The AUC assesses whether model predictions differ from that expected by chance, varying from 0.5 (random classification) to 1 (perfect classification). GLM performance was also measured using the likelihood ratio test statistics (LRS; Hosmer & Lemeshow 1989), which correspond to the reduction of model residual deviance in relation to null model deviance.

Models were calibrated on both rectangles and model accuracy was measured using the whole calibration set (i.e. resubstitution). Accuracy (AUC) of GLM was also

measured by application to the second rectangular area, which was interpreted as providing an independent validation. Explanatory power was expressed by measures of model accuracy using the calibration set, while predictive power was expressed by measures of model accuracy using the validation set.

The overall effect of SA was assessed exploring the relationship between the 95th percentile and the standard deviation of measures of accuracy for each null pattern set with Moran's  $I$  of the original predictor variables. The variation of the 95th percentile was analysed because it represents a common threshold in most statistical hypothesis testing. This parameter is expected to be inflated by SA.

In GLM, the LRS is assumed to be chi-square distributed and therefore it is also possible, for each run of 1000 permutations, to calculate the number of LRS tests on the calibration set that are found to have values above the expected number according to a chosen type I significance level. The ratio between this number and the expected number of significant tests according to the significance level (e.g. 50 out of 1000 for  $P = 0.05$ ) expresses the type I inflation ratio (Lennon 2000). Here, a significance level of 0.01 was considered in order to compare with other results found in the literature.

All data analyses were performed with 5-PLUS 2000 (Statistical Sciences 1999) using the default functions for model adjustments. Random toroidal shifts were performed using a modified function from the Splancs library of 5-PLUS (Rowlingson & Diggle 1993).

#### DEALING WITH SPATIAL AUTOCORRELATION

Two methods that address the effects of SA were compared using toroidal shifts and the procedures described above. The first method consisted in subsampling the original data set by eliminating cells in a systematic manner. All cells with even coordinates were eliminated from the original data sets. The second method included the incorporation of an autocovariate term accounting for the SA of observations. Autologistic models use contagion as an autocovariate term in the logistic regression equation. The measure of contagion was based on a two-order neighbourhood as the weighted average of the number of occupied grid cells among a set of  $k_a$  neighbours of a central grid cell  $y_a$ , so that:

$$\text{Contagion} = \left( \frac{\sum_{b=1}^{k_a} w_{ab} y_b}{\sum_{b=1}^{k_a} w_{ab}} \right) \quad \text{eqn 1}$$

where the weight given to the grid cell  $y_b$  is  $w_{ab} = 1/d_{ab}$ , and  $d_{ab}$  is the distance between grid cells  $y_a$  and  $y_b$ . Two orders of neighbours, assigning a weight of  $d = 1$  to the first-order and a weight of  $d = 2$  to the second-order neighbours, were used. Neighbours in the first order were the eight adjacent cells touching the central cell

along the edges and at the corners within a rectangular grid. The second-order neighbours were the next group of cells concentric to the first order with 16 grid cells.

These two methods were assessed for their ability to produce unbiased estimates of the explanatory power as measured by LRS. The contribution of contagion terms was removed from the LRS estimates, in order to assess exclusively the contribution of each variable to the explanatory power. This assessment was based on the distribution of *M. leprosa* only. The predictive power of models was not assessed using the autologistic approach because the spatial autocovariate term is a function of the calibration data and cannot be predicted for independent validation data.

Finally, corrected univariate variable significances were estimated for distribution models of *M. leprosa* and *E. orbicularis*. A Monte Carlo simulation approach was used to compare the AUC computed using the original predictor variables with the test statistics generated from 1000 toroidal shifts. Significance (probability of rejecting a true null hypothesis) was defined as the fraction of the AUC statistics of the 1000 simulated variables that fell above the AUC statistics of the original variable.

## Results

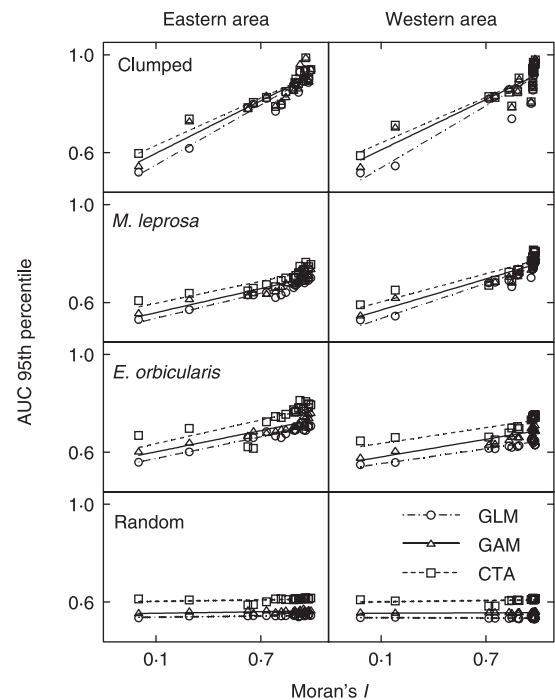
### SPATIAL AUTOCORRELATION OF VARIABLES

Spatial dependency is stronger for *M. leprosa* than for *E. orbicularis*, as shown by the overall SA of variables measured by the Moran's *I* index (Table 1). In the western rectangle the SA of *M. leprosa* occurrence is more pronounced while that of *E. orbicularis* is less pronounced in the western than in the eastern rectangle. Among environmental variables, the topographical descriptors are generally less spatially autocorrelated than climatic variables. Hill shade variables have the lowest Moran's *I*-values. For most environmental variables, with the exception for mean hill shade and mean temperature of the warmest month, the western rectangle shows slightly higher SA values than the eastern rectangle.

The mean Moran's *I*-values for each set of 1000 toroidal shifts of the environmental variables tend to be slightly lower than the original Moran's *I*, probably because of the effect of the shifted edges, which can slightly disturb the spatial structure. However, toroidal displacement did not shift significantly the relative order of SA between variables. On the other hand, the variability of values among the 1000 simulations, as measured by the standard deviation, was reasonably low, representing a small fraction of the mean values (SD/mean ranging from 0.010 to 0.042), which means that spatial structure was maintained among permutations (Table 1).

### EFFECT ON THE EXPLANATORY POWER OF VARIABLES

Amongst sets of 1000 toroidal permutations the 95th percentile of the model accuracy measures increased as

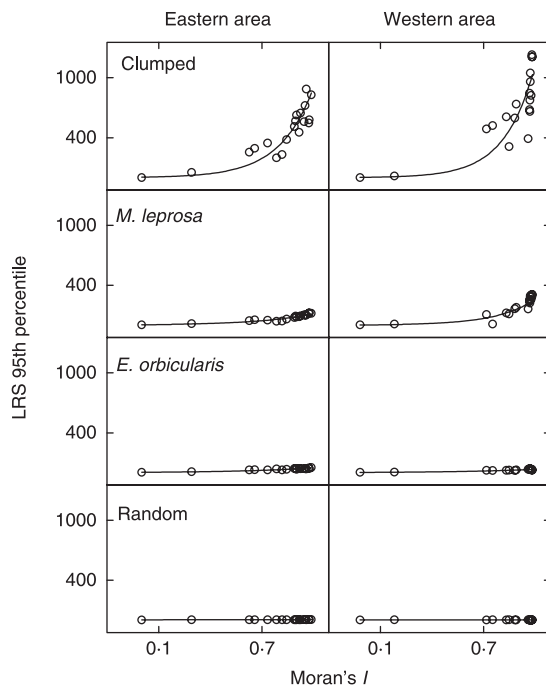


**Fig. 2.** Variation of the 95th percentile of AUC with Moran's *I* of environmental variables using three modelling techniques (GLM, GAM and CTA), four distributions (*M. leprosa*, *E. orbicularis*, a simulated distribution with totally clumped occurrences and a random distribution) and 1000 simulated surfaces (lines represent linear fits).

the SA of the predictor variables increased (Fig. 2). This effect was more pronounced for species' distributions that were more spatially autocorrelated and was eliminated in a simulated random species distribution. The 95th percentile of the AUC distribution of the 1000 toroidal shifts per variable increased in an approximately linear fashion with SA for each modelling approach (Fig. 2). In the western section of the studied rectangle the variation of AUC values with SA of predictor variables was more pronounced than in the eastern rectangle for *M. leprosa* and less pronounced for *E. orbicularis* (Fig. 2).

There were slight differences on the effect of SA for GAM and CTA compared with GLM, as shown by the variation of the 95th percentile of AUC values with Moran's *I* of the environmental variables (Fig. 2). The variation of AUC 95th percentiles was less marked for GAM and CTA than for GLM models; this was particularly evident for species with more autocorrelated distributions. This weaker variation was because of the better performances of GAM and CTA compared with GLM models for species' distributions and predictor variables with reduced SA while showing comparable performances for those with higher SA (Fig. 2). CTA tended to perform better than GAM, especially for the less autocorrelated distributions. These results were consistent between rectangles.

The 95th percentile of the LRS distribution of GLM models using variable permutations was more sensitive



**Fig. 3.** Relationship between spatial autocorrelation and the 95th percentiles of LRS statistics for GLM, regressing each simulated surface with presence/absence data of *M. leprosa*, *E. orbicularis*, a simulated distribution with totally clumped occurrences and a random distribution.

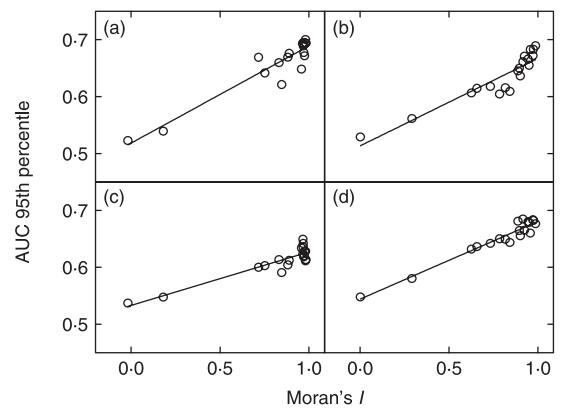
than AUC to the increase of Moran's *I*. This parameter tended to increase in an exponential fashion with SA (Fig. 3).

#### EFFECT ON THE PREDICTIVE POWER OF VARIABLES

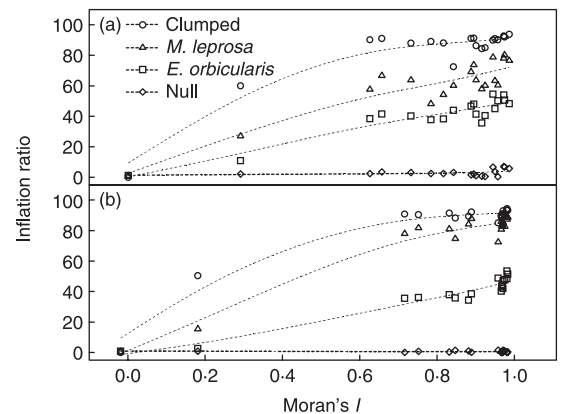
The effect of the increase in Moran's *I* using an independent validation data set to evaluate the predictive power of GLM models was as strong as using the calibration set. The 95th percentile of deviance also tended to increase with SA in a linear form (Fig. 4). This trend was more marked for models of *M. leprosa* calibrated in the eastern rectangle and for models of *E. orbicularis* calibrated in the western rectangle.

#### INFLATION RATIOS

Inflation ratio increased with Moran's *I*-values (Fig. 5) of the environmental variables, especially at lower Moran's *I*-values. Thus there was a substantial inflation of the apparent predictive power of analyses using variables with even a modest amount of SA. Predictor variable significance could be inflated by a factor up to 90 for the clumped distribution. At higher Moran's *I*-values, inflation ratios tended to stabilize (Fig. 5). Inflation ratios also increased where species' distributions displayed increased SA; indeed, in the absence of SA in the distribution (Fig. 5, Random distributions) even highly autocorrelated environments did not cause



**Fig. 4.** Model validation: relationship between Moran's *I* for the validation data set and the 95th percentiles of *M. leprosa* model's AUC, for each set of 1000 permutations of the original variables. Lines represent linear fits. (a) *M. leprosa*, eastern rectangle model validated with the western rectangle data set; (b) *M. leprosa*, western rectangle model validated with the eastern rectangle data set; (c) *E. orbicularis*, eastern rectangle model validated with the western rectangle data set; (d) *E. orbicularis*, western rectangle model validated with the eastern rectangle data set.



**Fig. 5.** Variation of the inflation ratio (number of times a significant test was found in relation to the expected number of significant tests, according to the significance level adopted, in this case  $P < 0.01$ , with the Moran's *I*-values of the original variables). Dashed lines are spline fits. (a) Eastern rectangle; (b) western rectangle.

inflated predictive power estimates. Overall, the higher the level of SA in both species' distributions and environmental variables used, the higher the inflation ratios. Consequently, it was unsurprising to find that inflation ratios for *M. leprosa* were more pronounced in the western rectangle (mean inflation ratio of 75.3 vs. 60.2 for the eastern rectangle) while those for *E. orbicularis* were more pronounced in the eastern rectangle (mean inflation ratio of 40.4 vs. 38.3 for the eastern rectangle). The observation that random distributions did not suffer variations in the inflation ratio supported the interpretation that it was SA rather than alternative unmeasured factors that caused inflation of the significance estimates in the models.

DEALING WITH SPATIAL  
AUTOCORRELATION

Comparison of two methods for dealing with SA within models showed that including an autocovariate term in the regression was more effective than systematically subsampling the area (Fig. 6). Indeed, the autologistic procedure may slightly overcompensate. For example, in the eastern rectangle there was a slight decrease in the location measure of the LRS test significances for an increase of Moran's  $I$  (Fig. 6a). The subsampling procedure was only partially effective in avoiding the inflation effect of model performance because of SA. The increase of LRS test significance with an increase of Moran's  $I$  was less pronounced using this procedure (Fig. 6; note that the axes have different ranges).

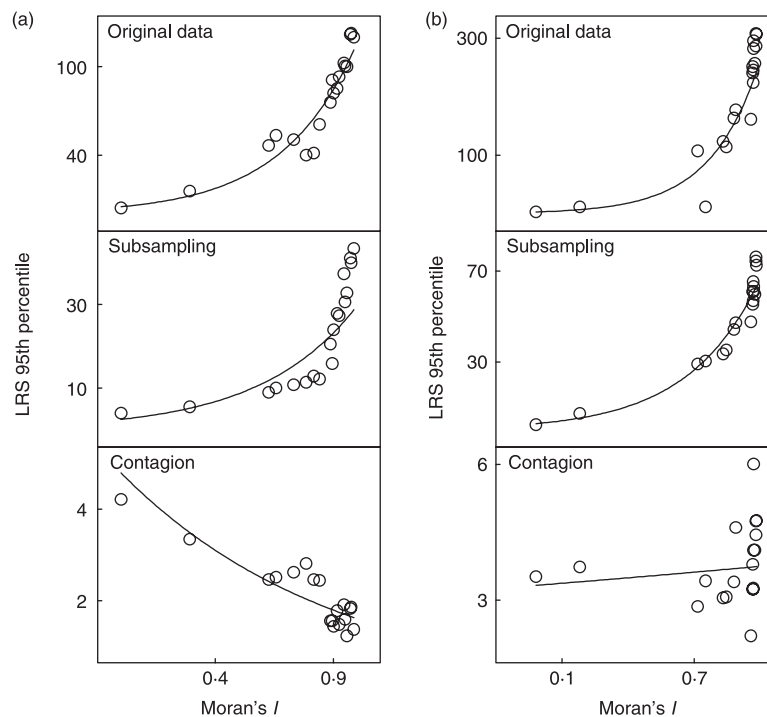
When surface permutations were used to produce Monte Carlo significance values of AUC accounting for SA, no obvious relationship with SA was observed, which suggested that unbiased estimates were produced. However, the most significant variables had consistently higher Moran's  $I$ -values, especially for *M. leprosa* (Table 2). Variables tended to have greater explanatory power for *M. leprosa* than for *E. orbicularis* and in the western rectangle there were more variables with significant ( $P < 0.05$ ) effects (e.g. for GLM 16 significant variables for *M. leprosa* and 13 for *E. orbicularis*) than there were in the eastern rectangle (11 significant variables for *M. leprosa* and two for *E. orbicularis*). The majority of significant variables for *M. leprosa* distribution were

common to the three modelling approaches. In the eastern rectangle nine significant variables were common and two uncommon, while in the western rectangle 15 significant variables were common and three uncommon. For *M. leprosa* seven significant variables were simultaneously common to both rectangles and the three modelling approaches. For *E. orbicularis* in the eastern rectangle, only one significant variable was common to the three modelling approaches and three variables were uncommon, while in the western rectangle five significant variables were common and eight uncommon. No significant variables were simultaneously common to both rectangles and the three modelling approaches for *E. orbicularis*.

## Discussion

In this study we used a simple and straightforward method of pattern simulation to explore the effect of SA on niche-based models. The results reinforce the idea that conclusions from niche-based models could be compromised because of the autocorrelated nature of both predictor variables and species' occurrences (Lennon 2000; Hampe 2004).

When plotted against a measure of SA (Moran's  $I$ ), model performance showed clear trends. For example, we found that significant log-likelihood ratio tests in GLM need to be exponentially larger as SA of predictor variables are stronger. This means that even slight changes in the degree of SA have a strong effect on the



**Fig. 6.** Relationship between spatial autocorrelation and the 95th percentiles of the LRS tests generated by regressing each simulated surface with presence/absence data of *M. leprosa*. Comparison between results using the original *M. leprosa* data on presence/absence and two different procedures usually adopted to avoid effects of spatial autocorrelation: subsampling of the original data set and forcing the inclusion of a contagion term in the model. Lines represent exponential fits. (a) Eastern rectangle; (b) western rectangle.



**Table 2.** Significance of each predictor variable using a Monte Carlo simulation approach (\* $P < 0.05$ ; \*\* $P < 0.01$ ). Variables are displayed in a increasing order of autocorrelation as measured by Moran's

Variables	<i>Mauremys leprosa</i>			<i>Emys orbicularis</i>		
	GLM	GAM	TREE	GLM	GAM	TREE
(a) Eastern rectangle						
Random	0.753	0.596	0.936	0.582	0.986	0.977
Hill	0.013*	0.201	0.283	0.004**	0.011*	0.044*
Hillmax	0.717	0.928	0.937	0.283	0.418	0.028*
Hillmin	0.959	0.869	0.922	0.997	0.369	0.003**
Slope	0.988	0.852	0.795	0.613	0.412	0.552
Altmax	0.153	0.313	0.319	0.454	0.773	0.847
Alt	0.017*	0.059	0.077	0.160	0.419	0.548
Altmin	0.018*	< 0.001	0.012*	0.164	0.137	0.389
Paut	0.957	0.357	0.351	0.982	0.157	0.072
Pwin	0.940	0.412	0.557	0.988	0.122	0.273
Pann	0.185	0.418	0.366	0.724	0.698	0.791
Pspr	< 0.001	< 0.001	< 0.001	0.134	0.320	0.440
Psum	< 0.001	< 0.001	< 0.001	0.077	0.289	0.358
Mtc	< 0.001	< 0.001	< 0.001	0.107	0.368	0.638
A2P	0.005**	0.018*	0.048*	0.235	0.678	0.436
Rann	< 0.001	< 0.001	0.008*	0.140	0.790	0.460
Tann	< 0.001	0.003**	< 0.001	0.091	0.173	0.340
Gdd	< 0.001	< 0.001	< 0.001	0.090	0.149	0.289
Mtw	< 0.001	0.001**	0.013*	0.042*	0.191	0.421
(b) Western rectangle						
Random	0.426	0.383	0.204	0.882	0.982	0.982
Hill	0.013*	0.022*	0.049*	0.410	0.091	0.551
Hillmax	0.036*	0.033*	0.023*	0.044*	0.035*	0.242
Hillmin	0.013*	0.008**	0.008**	0.004**	0.021*	0.019*
Slope	0.015*	0.020*	0.041*	0.034*	0.070	0.033*
Altmin	0.109	0.050	0.048*	0.551	0.840	0.857
Alt	0.068	0.041*	0.062	0.285	0.333	0.289
Altmax	0.028*	0.017*	0.009**	0.146	0.159	0.132
Rann	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
Pwin	0.001**	< 0.001	< 0.001	0.011*	0.006**	0.018*
Pspr	0.002**	0.002**	0.010*	0.008**	0.024*	0.010*
Paut	0.009**	< 0.001	< 0.001	0.040*	0.027*	0.037*
Pann	0.007**	0.007**	< 0.001	0.012*	0.053	0.113
Psum	0.005**	0.006**	0.002**	0.020*	0.231	0.289
Mtc	0.045*	0.110	0.078	0.252	0.459	0.568
A2P	0.005**	0.001**	< 0.001	0.006**	0.071	0.038*
Tann	0.006**	0.004**	< 0.001	0.012*	0.033*	0.078
Mtw	< 0.001	< 0.001	< 0.001	< 0.001	0.060	0.001**
Gdd	0.002**	0.001**	0.012*	0.006**	0.029*	0.169

probability of a predictor variable being chosen as significant. The same trend was observed with measures of models' predictive accuracy (here AUC), although this measure was less sensitive to SA on both the response and predictor variables.

Results also suggest that sensitivities of model accuracy (AUC) to SA using either validation or calibration data sets are nearly identical. Model performances are commonly measured using either resubstitution or by splitting data into calibration and validation sets (Olden, Jackson & Peres-Neto 2002; Araújo *et al.* 2005a). The last approach does not fully avoid an overestimation of model accuracy because two sources of bias may arise: (i) SA between the calibration and validation data sets and (ii) SA within data sets. In fact, even when using more complex approaches such as cross-validation, bootstrapping and jack-knifing (Guisan & Zimmermann

2000), most procedures involve splitting the original data set into calibration and validation subsets using random assignments. These two subsets have inevitably some degree of non-independence because of SA (Araújo *et al.* 2005a). In this study we used two data sets that were spatially separated in order to avoid an overestimation of accuracy because of a lack of independence between data sets. This allowed an assessment of the role of SA within each data set when predicting distributions in other regions, i.e. its effect on model transferability (Randin *et al.* in press). Despite differences in the environments of two rectangles, results suggest that the models' transferability depends greatly on the SA of the environmental variables.

As demonstrated before with correlation and linear regression analysis (Lennon 2000), SA of covariates inflates their statistical significance in predictive models.

In the present study, the maximum inflation ratio was 94 for the clumped distribution, i.e. variables were found to be significant 94 times more often than expected by chance. Inflation ratios reached 89 for *M. leprosa* and 54 for *E. orbicularis*. These inflation values are similar to those estimated by Lennon (2000) using synthetically generated spatial patterns.

The two a priori procedures used to minimize the effect of SA partially cancelled (systematic subsampling) or fully cancelled (inclusion of the contagion term) the effect of SA. However, it should be noted that testing for relationships between environmental variables and species' occurrences after forcing the inclusion of a contagion term might represent a problem. This procedure cancels the SA of species' distributions without differentiating environmental from demographic/historical contributions. By cancelling the environmental contribution to the SA, models tend to underestimate the importance of environmental variables that co-vary with species' occurrences (Araújo & Williams 2000). Variables that enter in models with an autocovariate term are most likely to explain the non-autocorrelated aspects of patterns, which may represent only a minor fraction of total variance. Moreover, such a model would most probably lack predictive power outside the calibration set or under future climate-change scenarios. Indeed, models that incorporate a term for SA cannot be extrapolated to regions where no occurrence data are available. Hence, the inclusion of an autocovariate term is only possible when projections are made within the geographical range of the calibration data set.

When sample size is not a limiting factor, subsampling of the original data matrix is a possibility. Even though it does not completely eliminate the inflation effect, it reduces it substantially. This can be done by simply selecting samples in a systematic manner as done here, or by using geostatistical tools such as variogram and correlogram plots (Maurer 1994; Diniz-Filho, Bini & Hawkins 2003), to analyse the overall pattern of spatial dependency and to help establish a minimum distance between samples that will reduce SA at a given amount (Catry *et al.* 2003). None the less, as populations and environmental variables tend to be autocorrelated at all scales, it seems likely that the spacing out of samples will never fully eliminate SA effects.

Another type of procedure is to account a posteriori for SA using a Monte Carlo approach, such as that used here, to estimate variable significance based on null spatial patterns. However, even when *P*-values are estimated using this approach, variables that exhibit higher SA tend to be more significant. Nevertheless, these are likely to be the better candidates as actual causative factors of species' distributions; they are more likely to contribute to the autocorrelated nature of species' patterns of occurrence.

The use of semi-parametric modelling techniques, such as GAM, and non-parametric techniques, such as classification trees, slightly reduces the effect of SA in our analyses. These techniques place fewer constraints

on the shape of species' responses to their environments, and stronger adjustments to predictions are consequently produced in comparison with parametric techniques, such as GLM. Our results suggest that this is particularly evident for response or environmental variables that display only moderate levels of SA. Thus these techniques are slightly less sensitive to variations in the variables' patterns of SA than traditional parametric approaches such as GLM.

The results described above, and thus the recommendations given for dealing with SA, apply only to univariate modelling. When there is more than one candidate variable to explain a species' distribution, the assessment of the effect of SA on models requires a more complex and thorough approach. Indeed, the impact of SA on multivariate model building is an urgent issue for future investigation. The first step of multivariate modelling is variable selection. Methods based on stepwise variable selection are still widely used by ecologists although it has been demonstrated that automated model building procedures can result in the selection of a subset of predictor variables with no direct effect on the response variable (Derksen & Keselman 1992). Nevertheless, even if SA inflates variable significance this does not mean that the final model configuration will exclusively include the most autocorrelated variables. If an autocorrelated variable is included in the model, it may explain a substantial fraction of the SA in the species' pattern of occurrence. If that is the case the remaining variables may then be related to the less autocorrelated aspects of the distribution, probably driven by factors that act at finer spatial scales (Diniz-Filho, Bini & Hawkins 2003).

Some recommendations for the variable selection process can, however, be drawn from our results. Variable selection based on simulated patterns is not cost effective. We recommend, instead, before model adjustment running univariate tests based on null patterns such as the procedure described in this study. Strong autocorrelated variables that loose explanatory power when their significance is adjusted for SA should be handled with special care because they could inflate the models' significance. An alternative would be simply to exclude such variables from the analysis. Employing methods based on the recent information-theoretic paradigm (Burnham & Anderson 2002; but see Stephens *et al.* 2005) during model building, such as variable selection procedures using the Akaike information criterion (Akaike 1974), is always preferable because it does not fully rely on significance thresholds. For modelling species' distributions that show a high degree of SA it might be adequate to use semi-parametric or non-parametric techniques, as well as avoiding the use of statistics such as LRS, which are bound to be more sensitive to SA than other accuracy measures (e.g. AUC).

The value of niche-based distribution models for planning and management purposes greatly depends on their ability to overcome different sources of biases that are inherent with biological data. SA is just one

among other sources of bias, yet probably the most challenging one. It is our belief that simple procedures such as the ones discussed in this study would help to enhance ecological reliability of models and therefore to increase their applied value during the decision-making process.

### Acknowledgements

We thank Stephen Roxburgh, Ana Bio and Wilfried Thuiller for their comments on the manuscript. Pedro Segurado was funded by FCT (SFRH/BD/8493/2002). This research is partly funded by the EC Integrated FP6 ALARM (GOCE-CT-2003-506675) project.

### References

- Akaike, H. (1974) A new look at statistical model identification. *IEEE Transactions on Automated Control*, **19**, 716–722.
- Araújo, M.B. & Williams, P.H. (2000) Selecting areas for species persistence using occurrence data. *Biological Conservation*, **96**, 331–345.
- Araújo, M.B., Pearson, R.G., Thuiller, W. & Erhard, M. (2005a) Validation of species–climate impact models under climate change. *Global Change Biology*, **11**, 1504–1513.
- Araújo, M.B., Thuiller, W. & Pearson, R.G. (2006) Climate warming and the decline of amphibians and reptiles in Europe. *Journal of Biogeography* (in press).
- Araújo, M.B., Whittaker, R.J., Ladle, R.J. & Erhard, M. (2005b) Reducing uncertainty in projections of extinction risk from climate change. *Global Ecology and Biogeography*, **14**, 529–538.
- Augustin, N.H., Muggleston, M.A. & Buckland, S.T. (1996) An autologistic model for the spatial distribution of wildlife. *Journal of Applied Ecology*, **33**, 339–347.
- Austin, G.E., Thomas, C.J., Houston, D.C. & Thompson, B.A. (1996) Predicting the spatial distribution of buzzard *Buteo buteo* nesting areas using geographical information systems and remote sensing. *Journal of Applied Ecology*, **33**, 1541–1550.
- Borcard, D., Legendre, P. & Drapeau, P. (1992) Partiaing out the spatial component of ecological variation. *Ecology*, **73**, 1045–1055.
- Brito, J.C., Crespo, E.G. & Paulo, O.S. (1999) Modelling wildlife distribution: logistic multiple regression vs overlap analysis. *Ecography*, **22**, 151–260.
- Burnham, K.P. & Anderson, D.R. (2002) *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, 2nd edn. Springer-Verlag, New York, NY.
- Catry, P., Campos, A., Segurado, P., Silva, M. & Strange, I. (2003) Population census and nesting habitat selection of thin-billed prion *Pachyptila belcheri* on New Island, Falkland Islands. *Polar Biology*, **26**, 202–207.
- Clark, L.A. & Pregibon, D. (1992) Tree-based models. *Statistical Models in S* (eds J.M. Chambers & T.J. Hastie), pp. 377–419. Chapman & Hall, New York, NY.
- Cliff, A.D. & Ord, J.K. (1973) *Spatial Autocorrelation*. Pion Limited, London, UK.
- Dale, M.R.T. & Fortin, M.-J. (2002) Spatial autocorrelation and statistical tests in ecology. *Écoscience*, **9**, 162–167.
- Derksen, S. & Keselman, H.J. (1992) Backward, forward and stepwise automated subset selection algorithms: frequency of obtaining authentic and noise variables. *British Journal of Mathematical and Statistical Psychology*, **45**, 265–282.
- Diniz-Filho, J.A.F., Bini, L.M. & Hawkins, B.A. (2003) Spatial autocorrelation and red herrings in geographical ecology. *Global Ecology and Biogeography*, **12**, 53–64.

- Dutilleul, P. (1993) Modifying the *t* test for assessing the correlation between two spatial processes. *Biometrics*, **49**, 305–314.
- Fielding, A.H. & Bell, J.F. (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, **24**, 38–49.
- Fortin, M.-J. & Jacquez, G.M. (2000) Randomization tests and spatially autocorrelated data. *Bulletin of the Ecological Society of America*, **81**, 201–206.
- Fortin, M.-J. & Payette, S. (2002) How to test the significance of the relation between spatially autocorrelated data at the landscape scale: a case study using fire and forest maps. *Écoscience*, **9**, 213–218.
- Franklin, J. (1998) Predicting the distribution of shrub species in southern California from climate and terrain-derived variables. *Journal of Vegetation Science*, **9**, 733–748.
- Gates, S., Gibbons, D.W., Lack, P.C. & Fuller, R.J. (1994) Declining farmland bird species: modelling geographical patterns of abundance in Britain. *Large-Scale Ecology and Conservation Biology* (eds P.J. Edwards, R.M. May & N.R. Webb), pp. 153–177. Blackwell Science, London, UK.
- Godinho, R., Teixeira, A., Rebelo, R., Segurado, P., Loureiro, A., Álvares, F., Gomes, N., Cardoso, P., Camilo-Alves, C. & Brito, J.C. (1999) Atlas of the continental Portuguese herpetofauna: an assemblage of published and new data. *Revista Española de Herpetología*, **13**, 61–82.
- Guisan, A. & Theurillat, J.P. (2000) Assessing alpine plant vulnerability to climate change: a modelling perspective. *Integrated Assessment*, **1**, 307–320.
- 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.
- Harrison, S. (1997) How natural habitat patchiness affects the distribution of diversity in Californian serpentine chaparral. *Ecology*, **78**, 1898–1906.
- Hastie, T. & Tibshirani, R. (1990) *Generalised Additive Models*. Chapman & Hall, London, UK.
- Hosmer, D.W. Jr & Lemeshow, S. (1989) *Applied Logistic Regression*. John Wiley & Sons, New York, NY.
- Keitt, T.H., Bjørnstad, O.N., Dixon, P.M. & Citron-Pousty, S. (2002) Accounting for spatial pattern when modeling organism–environment interactions. *Ecography*, **25**, 616–625.
- Keller, C. & Andreu, A.C. (2002) *Emys orbicularis*. *Atlas y Libro Rojo de Los Anfíbios y Reptiles de España* (eds J.M. Pleguezelos, R. Marquez & M. Lizana), pp. 137–140. Dirección General de la Conservación de la Naturaleza, Madrid, Spain.
- Koenig, W.D. & Knops, M.H. (1998) Testing for spatial autocorrelation in ecological studies. *Ecography*, **21**, 423–429.
- Legendre, P. (1993) Spatial autocorrelation: trouble or new paradigm? *Ecology*, **74**, 1659–1673.
- Legendre, P., Dale, M.R.T., Fortin, M.-J., Casgrain, P. & Gurevitch, J. (2004) Effects of spatial structures on the results of field experiments. *Ecology*, **85**, 3202–3214.
- Legendre, P., Dale, M.R.T., Fortin, M.-J., Casgrain, P., Gurevitch, J., Hohn, M. & Myers, K. (2002) The consequence of spatial structure for the design and analysis of ecological field surveys. *Ecography*, **25**, 601–615.
- Lennon, J.J. (2000) Red-shifts and red herrings in geographical ecology. *Ecography*, **23**, 101–113.
- Lichstein, J.W., Simons, T.R., Shiner, S.A. & Franzreb, K.E. (2002) Spatial autocorrelation and autoregressive models in ecology. *Ecological Monographs*, **72**, 445–463.
- McCullagh, P. & Nelder, J.A. (1983) *Generalized Linear Models*. Chapman & Hall, London, UK.
- 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.
- Maurer, B.A. (1994) *Geographical Population Analysis: Tools for the Analysis of Biodiversity*. Blackwell Science, Oxford, UK.
- Midgley, G.F., Hannah, L., Millar, D., Rutherford, M.C. & Powrie, L.W. (2002) Assessing the vulnerability of species richness to anthropogenic climate change in a biodiversity hotspot. *Global Ecology and Biogeography*, **11**, 445–451.
- New, M., Hulme, M. & Jones, P.D. (2000) Representing twentieth century space time climate variability. II. Development of 1901–96 monthly grids of terrestrial surface climate. *Journal of Climate*, **13**, 2217–2238.
- Olden, J.D., Jackson, D.A. & Peres-Neto, P.R. (2002) Predictive models of fish species distributions: a note on proper validation and chance predictions. *Transactions of the American Fisheries Society*, **131**, 329–336.
- Palmer, M.W. & Van der Maarel, E. (1995) Variance in species richness, species association, and niche limitation. *Oikos*, **73**, 203–213.
- Pearce, J. & Ferrier, S. (2000) An evaluation of alternative algorithms for fitting species distribution models using logistic regression. *Ecological Modelling*, **128**, 127–147.
- 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.
- Pereira, J.M.C. & Itami, R.M. (1991) GIS-based habitat modeling using logistic multiple regression: a study of the Mt Graham red squirrel. *Photogrammetric Engineering and Remote Sensing*, **57**, 1475–1486.
- Peterson, A.T., Ortega-Huerta, M.A., Bartley, J., Sanchez-Cordero, V., Soberon, J., Buddemeier, R.H. & Stockwell, D.R.B. (2002) Future projections for Mexican faunas under global climate change scenarios. *Nature*, **416**, 626–629.
- Randin, C.F., Dirnböck, T., Dullinger, S., Zimmermann, N.E., Zappa, M. & Guisan, A. (2006) Are niche-based species distribution models transferable in space? *Journal of Biogeography* (in press).
- Rowlingson, B. & Diggle, P.J. (1993) Splanx: spatial point pattern analysis code in S-PLUS. *Computers and Geosciences*, **19**, 627–655.
- Roxburgh, S.H. & Chesson, P. (1998) A new method for detecting species associations with spatially autocorrelated data. *Ecology*, **79**, 2180–2192.
- Segurado, P. & Araújo, M.B. (2004) An evaluation of methods for modelling species distributions. *Journal of Biogeography*, **31**, 1555–1568.
- da Silva, E., (2002) *Mauremys leprosa* (Schweiger, 1812). *Atlas y Libro Rojo de Los Anfibios y Reptiles de España* (eds J.M. Pleguezelos, R. Marquez & M. Lizana), pp. 143–146. Dirección General de la Conservación de la Naturaleza, Madrid, Spain.
- Smith, P.A. (1994) Autocorrelation in logistic regression modelling of species distribution. *Global Ecology and Biogeography Letters*, **4**, 47–61.
- Statistical Sciences (1999) *S-Plus Version 2000 for Windows*. Mathsoft Inc., Seattle, WA.
- Stephens, P.A., Buskirk, S.W., Hayward, G.D. & Martínez del Rio, C. (2005) Information theory and hypothesis testing: a call for pluralism. *Journal of Applied Ecology*, **42**, 4–12.
- Storch, D., Konvička, M., Beneš, J., Martinková, J. & Gaston, K.J. (2003) Distribution patterns in butterflies and birds of the Czech Republic: separating effects of habitat and geographic position. *Journal of Biogeography*, **30**, 1195–1205.
- Thomas, C.D., Cameron, A., Green, R.E., Bakkenes, M., Beaumont, L.J., Collingham, Y., 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.J., Ortega-Huerta, M.A., Peterson, A.T., Phillips, O. & Williams, S.E. (2004) Extinction risk from climate change. *Nature*, **427**, 145–148.
- Thuiller, W. (2003) Optimizing predictions of species distributions and projecting potential future shifts under global change. *Global Change Biology*, **9**, 1353–1362.
- Thuiller, W. (2004) Patterns and uncertainties of species' range shifts under climate change. *Global Change Biology*, **10**, 2220–2227.
- Thuiller, W., Araújo, M.B., Pearson, R.G., Whittaker, R.J., Brotons, L. & Lavorel, S. (2004) Uncertainty in predictions of extinction risk. *Nature*, **430**, 33.
- Thuiller, W., Lavorel, S., Araújo, M.B., Sykes, M.T. & Prentice, I.C. (2005) Climate change threats to plant diversity in Europe. *Proceedings of the National Academy of Sciences, USA*, **102**, 8245–8250.
- Walker, P.A. (1990) Modelling wildlife distributions using a geographic information system: kangaroos in relation to climate. *Journal of Biogeography*, **17**, 279–289.
- Watkins, A.J. & Wilson, J.B. (1992) Fine-scale community structure of lawns. *Journal of Ecology*, **80**, 15–24.

Received 24 October 2005; final copy received 26 January 2006  
Editor: Rob Freckleton