Static species distribution models in dynamically changing systems: how good can predictions really be?

Damaris Zurell, Florian Jeltsch, Carsten F. Dormann and Boris Schröder

D. Zurell (Damaris.Zurell@uni-potsdam.de) Univ. of Potsdam, Inst. of Geoecology, Karl-Liebknecht-Str. 24/25, DE-14476 Postdam, Germany. – F. Jeltsch, Univ. of Potsdam, Inst. for Biochemistry and Biology, Maulbeerallee 2, DE-14469 Potsdam, Germany. – C. F. Dormann, Helmholtz Centre for Environmental Research UFZ, Dept of Computational Landscape Ecology, Permoserstr. 15, DE-04318 Leipzig, Germany. – B. Schröder, Univ. of Potsdam, Inst. of Geoecology, Karl-Liebknecht–Str. 24/25, DE-14476 Potsdam, Germany and Leibniz-Centre for Agricultural Landscape Research (ZALF) e.V., Soil Landscape Modelling, Eberswalder Str. 84, DE-15374 Müncheberg, Germany.

It is widely acknowledged that species respond to climate change by range shifts. Robust predictions of such changes in species' distributions are pivotal for conservation planning and policy making, and are thus major challenges in ecological research. Statistical species distribution models (SDMs) have been widely applied in this context, though they remain subject to criticism as they implicitly assume equilibrium, and incorporate neither dispersal, demographic processes nor biotic interactions explicitly. In this study, the effects of transient dynamics and ecological properties and processes on the prediction accuracy of SDMs for climate change projections were tested. A spatially explicit multi-species dynamic population model was built, incorporating species-specific and interspecific ecological processes, environmental stochasticity and climate change. Species distributions were sampled in different scenarios, and SDMs were estimated by applying generalised linear models (GLMs) and boosted regression trees (BRTs). Resulting model performances were related to prevailing ecological processes and temporal dynamics.

SDM performance varied for different range dynamics. Prediction accuracies decreased when abrupt range shifts occurred as species were outpaced by the rate of climate change, and increased again when a new equilibrium situation was realised. When ranges contracted, prediction accuracies increased as the absences were predicted well. Far-dispersing species were faster in tracking climate change, and were predicted more accurately by SDMs than short-dispersing species. BRTs mostly outperformed GLMs. The presence of a predator, and the inclusion of its incidence as an environmental predictor, made BRTs and GLMs perform similarly.

Results are discussed in light of other studies dealing with effects of ecological traits and processes on SDM performance. Perspectives are given on further advancements of SDMs and for possible interfaces with more mechanistic approaches in order to improve predictions under environmental change.

Among the expected consequences of the ongoing climate change are shifts in species' geographic ranges, range expansions and contractions. Robust prediction of these distributional changes are a prerequisite for dynamic and sustainable conservation strategies, and thus constitute a major challenge in present-day ecological research (Guisan and Thuiller 2005, Vaughan and Ormerod 2005).

Statistical species distribution models (SDMs) have been widely used to project species range shifts, and to derive extinction risks for different climate change scenarios (Bakkenes et al. 2002, Midgley et al. 2002, Thomas et al. 2004, Thuiller 2004). These data-driven models relate field observations to environmental predictor variables. They provide an easy-to-use and potentially powerful tool for ecologists and conservationists because simple spatial incidence data can be used to derive the statistical models (Scott et al. 2002). Despite these merits, SDMs also show

particular limitations regarding climate change projections (Pearson and Dawson 2003, Guisan and Thuiller 2005, Araújo and Rahbek 2006, Dormann 2007). Foremost, they assume equilibrium between the species and its environment, and aim at predicting a new equilibrium state when extrapolating. Thus, transient dynamics are essentially ignored when projecting into the future. Furthermore, SDMs do not explicitly incorporate demographic processes and biotic interactions, and only few attempts have been made yet to narrow uncertainties due to dispersal assumptions (Midgley et al. 2006). All these issues may lead to substantial uncertainties in climate change projections (Dormann et al. 2008). To date, there is little knowledge of the consequences as the accuracy of SDMs in predicting future species distributions is not easy to evaluate (Thuiller 2004, Araújo et al. 2005). This is mainly because the events we aim to project have not yet occurred, and the future

species' geographic ranges are therefore unknown (but see Araújo et al. (2008) for an application for predicting current distributions from historical data).

One way to overcome the data limitations are artificial or virtual experiments (Berger et al. 1999, Austin et al. 2006, Schröder and Seppelt 2006). Such approaches have the further advantages of allowing us perfect knowledge and control over the underlying processes. Previous studies employing virtual experiments and SDMs have tested optimal sampling strategies (Hirzel and Guisan 2002, Reese et al. 2005), compared the performance of different statistical methods, model selection strategies, or threshold criteria for binary predictions (Hirzel et al. 2001, Reineking and Schröder 2006, Jiménez-Valverde and Lobo 2007), and assessed how good SDMs were at identifying "source" habitats (Tyre et al. 2001), the effect of diverse occurrence-environment relationships (Austin et al. 2006, Meynard and Quinn 2007) or the use of favourability functions (Albert and Thuiller 2008). To our knowledge, it has not been tested yet how transient dynamics and ecological properties and processes affect SDM accuracy when projecting into the future.

In this study we propose a virtual experiment to test SDM prediction accuracy under ongoing climate change by developing species distribution models on data from a complex, dynamic population model, which is used to model transient responses of a species to climate change. We therefore built a dynamic, spatially explicit multi-species population model which incorporated species-specific ecological properties and processes such as the ecological niche, dispersal ability and intrinsic growth rate, interspecific interactions such as competition and predation, environmental stochasticity, and climate change. Different scenarios were developed by systematically manipulating model properties. For each modelling scenario SDMs were estimated by applying two different SDM methods, generalised linear models (GLMs) and boosted regression trees (BRTs). Projected species distributions by SDMs were compared to simulated "true" species distributions by the dynamic population model focusing on the following questions: 1) do transient dynamics lead to a decrease in projection accuracy under climate change? 2) Are modern, flexible statistical modelling techniques (represented in our study by BRT) more capable of projecting future species ranges than long-established, parametric methods (represented here by GLM)? 3) Are these effects confounded by differing ecological properties and processes?

Methods

Dynamic population model

The purpose of the dynamic population model was to mimic scenarios of real-world situations with a complex virtual world containing a focal species characterised by species-specific properties and processes, and influenced by predation, environmental stochasticity, and climate change. To achieve this, a discrete-generation host-parasitoid system was set in a real environment. Such host-parasitoid systems are well suited for simple population models because they can have a much simpler structure than many other enemy-victim interactions due to the tight link between trophic and reproductive aspects of the parasitoid life history (Nicholson 1933, Hassell 2000, King and Hastings 2003).

The host was the focal species, a holometabolous insect with distinct generations which is henceforth referred to as butterfly. A highly synchronised parasitoid parasitised the butterfly during its larval stage, and thereby caused the death of the butterfly (Mills and Getz 1996). Movement and dispersal were limited to the adult insect stages. In analogy to real systems the model was tritrophic, i.e. the butterfly-parasitoid interaction depended on a host plant, henceforth called plant, which affected the presence and abundance of the butterfly. A coupled-map lattice model was used to link the local and regional dynamics (Hassell et al. 1991, Comins et al. 1992, Bonsall and Hassel 2000). In each cell of a two-dimensional lattice the local butterfly-parasitoid population dynamics were mapped annually. The populations were then connected by dispersal. We thus obtained a spatially explicit multi-species dynamic population model which allowed systematic modifications in several ways.

Structure, scales and scheduling

Space was represented by a two-dimensional lattice of 148×113 sites with a cell size of 1×1 km. Absorbing boundary conditions were assumed, representing an open system where butterflies and parasitoids were able to leave the habitat, but not to (re-)enter it. One time step represented one year and the whole simulation covered a period of 150 yr.

During initialisation, the environmental factors elevation and potential moisture were attributed to each cell in the lattice, and the climatic factors temperature and precipitation were assigned to each time step. Butterflies and parasitoids were randomly distributed over the suitable habitat. During simulation, each time step was characterised by four processes (Fig. 1). Foremost, climate state and habitat state were updated. Temperature and moisture were assigned to each cell by climatological downscaling. The host plant foliage projective cover was determined by the prevalent temperature and moisture regime, and induced a carrying capacity K for butterflies in each lattice cell. Resource competition at plant level was introduced by adding a second plant species, the competitor, whose fundamental niche overlapped with that of the host plant (Fig. 1), and which did not serve as a host plant for the butterfly. At the beginning of each time step, butterflies and parasitoids dispersed throughout the lattice. After colonisation, butterflies reproduced and the offspring could then be parasitised by female parasitoids. Detailed descriptions of the modelled processes are contained in Supplementary material Text S1, parameter values are given in Table 1.

Input

The environmental data employed in the simulations were based on real environmental data derived from a digital terrain model developed by the Swiss Federal Statistical Office, GEOSTAT. The coordinates of the 148×113 km grid are 607 000–754 000 m E/152 000–264 000 m N (Swiss grid CH1903). Environmental input layers were
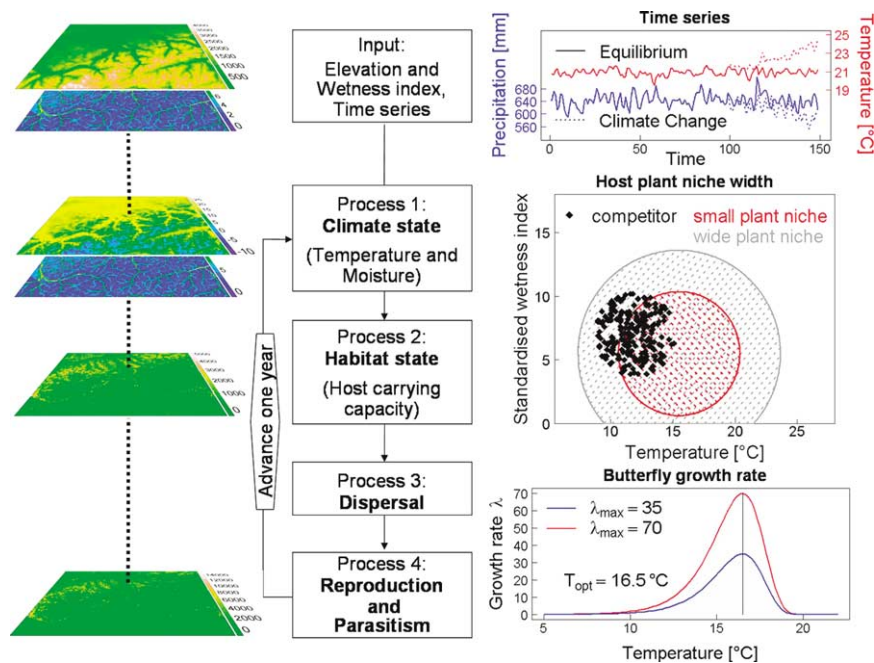
Figure 1. Process scheduling of the dynamic model is illustrated in the flowchart in the centre. Each simulation started with the input of elevation and potential moisture distribution (top left) as well as temperature and precipitation time series (top right). After time initiation, four processes were carried out within each time step: additional ecogeographical information layers were calculated, i.e. actual temperature and moisture distribution (Process 1) as well as local host carrying capacities (Process 2). Then dispersal of hosts and parasitoids was simulated (Process 3) and reproduction and parasitism took place (Process 4). The figure in the centre right depicts the fundamental and realised niche of the host plant. Temperature dependency of host growth rate is shown in the bottom right figure.

elevation, drainage area above each cell, slopes and aspects. Altitude in the study region ranged from 269 to 3854 m a.s.l. Main climate variables in the virtual world were energy and water, in particular mean summer temperatures and mean summer precipitation. Time series were calculated manually by drawing for each time step a random number from a normal distribution with a characteristic mean (temperature $T = 21°C$ at 0 m a.s.l.; precipitation $Prec = 640$ mm) and standard deviation ($\sigma_T = 0.5°C$; $\sigma_{Prec} = 30$ mm). A generalised climate change scenario was derived for the period 2001 to 2050. Rates of change approximately followed the projected changes of the IPCC SRES scenario A2 for central and northern Europe (Houghton et al. 2001). Mean temperature was increased by 3°C in 50 yr, mean precipitation decreased by 50 mm; mean values within the 50 yr were linearly interpolated.

Scenarios

A standard simulation run lasted 150 yr. The model "spinned up" for 100 yr, running with average climate, thereby ensuring a long-term equilibrium between butterfly

Table 1. Constants in the process-based dynamic model. Respective equations are contained in Supplementary material Text S1.

Function	Symbol	Value	Unit	Description
Grid	w	1	km	Cell width
Climate state	T_{lapse}	7	°C km^{-1}	Adiabatic gradient
Habitat state	K_{max}	5000	ind	Maximum carrying capacity of a cell
	$\mu_{T,plant}$	15.5	°C	Mean of host plant's temperature utilisation function
	$\mu_{W,plant}$	5.5	–	Mean of host plant's moisture utilisation function
	$\mu_{T,comp}$	12	°C	Mean of competitor's temperature utilisation function
	$\sigma_{T,comp}$	1	°C	Standard deviation of competitor's temperature utilisation function
	$\mu_{W,comp}$	7	–	Mean of competitor's moisture utilisation function
	$\sigma_{W,comp}$	1	–	Standard deviation of competitor's moisture utilisation function
	σ_F	0.008		Standard deviation for environmental stochasticity
Dispersal	$\beta_{butterfly}$	1.5	–	Butterfly's shape parameter in eq. 4
	$\mu_{butterfly}$	0.75	–	Fraction of local butterfly population emigrating
	$\alpha_{parasitoid}$	2	–	Parasitoid's scale parameter in eq. 4
	$\beta_{parasitoid}$	1.5	–	Parasitoid's shape parameter in eq. 4
	$\mu_{parasitoid}$	0.75	–	Fraction of local parasitoid population emigrating
Reproduction and parasitism	$T_{opt,\lambda}$	16.5	°C	Optimal temperature for butterfly growth rate, location parameter of Gumbel distribution
	σ_λ	1.4	°C	Scale parameter of Gumbel distribution
	N_{crit}	50	ind	Critical population size for Allee effect
	c	1	ind	Parameter in eq. 5

population and environment which was usually reached after five years. Scenarios were applied over the last 50 model years.

Scenarios were derived in which ecological properties and processes as well as climate were systematically manipulated in a factorial simulation experiment of five factors with two levels each (Table 2). Climate change and parasitism were either turned on or off, the latter by adjusting the area of discovery, i.e. the attack rate by the parasitoids. Butterfly growth rate was either low or high through manipulation of the maximum growth rate λ_{max}. Plant niche width was either narrow or wide, which was achieved by varying the standard deviation of the plant's resource utilisation functions. Finally, butterfly dispersal distance was either short or long through manipulation of the scale parameter α in eq. 4. A value of $\alpha = 4$ resulted in a maximum dispersal distance of one cell, $\alpha = 0.5$ in a maximum dispersal distance of five cells, while the parasitoids dispersed with a fixed maximum dispersal distance of two cells ($\alpha = 2$). All possible parameter combinations resulted in $2^5 = 32$ scenarios. For each scenario we ran 10 replicate simulations.

Sampling by virtual ecologist

Similar to real field studies, a virtual ecologist sampled the butterfly population with the same strategy as an ecologist might choose in reality. A random stratified sampling strategy was applied (Hirzel and Guisan 2002), with temperature as the stratifying variable. The quantiles (0.2-, 0.4-, 0.6- and 0.8-quantiles) of the temperature distribution were taken to split the data into five ordinal classes, generating five homogenous environmental strata with respect to temperature. An equal number of cells (replicates) were chosen in each stratum. In a selected cell, the incidence of the butterfly and the prevailing environmental conditions were recorded. In order to reduce spatial autocorrelation in the response variable, samples were not taken in adjacent cells. The virtual ecologist made no errors in detection, i.e. butterfly occurrence and environmental conditions were recorded exactly as given in the dynamic model. This way, the performance of the statistical model could be directly related to the underlying demographic and stochastic processes (Tyre et al. 2001). Training data were sampled directly after "spin up", i.e. after 100 model years, with a sample size of 1000. For each simulation run, five training data sets were sampled to capture the variability introduced by snapshot data; SDM accuracies for these five training data sets were averaged later. Independent (test) data sets (sample size = 1000) were sampled in the year 100, and in every subsequent 10 yr until year 150.

Statistical modelling

Model formulation

Butterfly occurrences were analysed by generalised linear models (GLMs) and boosted regression trees (BRTs) using a binomial error distribution and a logistic link function. GLMs have been traditionally used in species distribution modelling, and fit parametric terms. BRTs were developed within the machine-learning community, and are an ensemble-prediction method combining regression trees and boosting. They are very flexible, but at the same time resistant to overfitting, and are able to automatically model complex interactions between predictor variables (Ridgeway 1999, Leathwick et al. 2006, Thuiller et al. 2006, Elith et al. 2008).

Before the application of GLMs and BRTs, the predictor variables were tested for multicollinearity by calculating Spearman's rank correlation coefficient ρ_S. Following Fielding and Haworth (1995), if two variables had a correlation $\rho_S > 0.7$ the predictor with less ecological importance in respect to butterfly occurrence was removed. GLMs were estimated for the remaining predictor variables by applying an Akaike information criterion (AIC)-based stepwise variable selection procedure (Akaike 1974, Harrell 2001). According to the ecological knowledge we had (from the "virtual" reality), we included linear (e.g. host plant cover) and quadratic terms (e.g. temperature and soil moisture) in the GLMs as well as an interaction term between temperature and soil moisture. The linear term was forced into the model each time the quadratic term or the interaction term was selected in the final model. BRTs were estimated with a tree complexity of 2, a bag fraction of 0.75 and a learning rate of 0.005 which ensured that the models were fitted with at least 1000 trees (cf. Elith et al. 2008).

Extrapolation in space and time

The resulting SDMs were used to make predictions to independent (test) data sets (sample size = 1000) for the year 100, in which the models were fitted, and for every subsequent 10 yr until year 150 where the simulation ended. Thus, SDMs were fitted under average climate where the butterfly population was in long-term equilibrium with its environment, and, in the case of climate change, the occurrence of butterflies was projected to differing time slices under gradually ongoing climate change.

Model transferability: validation

For each prediction in space and in time, three different measures of accuracy were calculated in order to get a multi-facetted view of how good the SDM predictions were. The proportion of deviance explained by the final models was

Table 2. Range of parameter values in the process-based dynamic model. Please note that low values of $\alpha_{butterfly}$ correspond to long dispersal distances and high values to short dispersal distances. Respective equations are contained in Supplementary material Text S1.

Function	Symbol	Values		Unit	Description
Habitat state	$\sigma_{T/W,plant}$	1.5	2.5	°C	Standard deviations of host plant's temperature and moisture utilisation functions
Dispersal	$\alpha_{butterfly}$	0.5	4	–	Butterfly's scale parameter in eq. 4
Reproduction and parasitism	λ_{max}	35	70	ind	Maximum butterfly population growth rate
	a	0	0.01	ind	Area of discovery in eq. 6

quantified by the explained deviance r_L^2, a logistic regression equivalent to the coefficient of determination r^2 (Menard 2000). We derived the deviance by applying eq. 1.10 in Hosmer and Lemeshow (2000). The models' ability to discriminate between occupied and non-occupied sites was assessed by calculating AUC, the area under the ROC-curve (Fielding and Bell 1997). AUC is independent of classification thresholds and typically assumes values between 0.5 and 1, with AUC = 0.5 for models with predictive ability no better than the null model and AUC = 1 for perfectly discriminating models. Values of AUC > 0.7 indicate useful predictions according to Hosmer and Lemeshow (2000). Additionally, the calibration curve was calculated, to determine the agreement between observations and predicted values, i.e. the goodness-of-fit (Pearce and Ferrier 2000). It was derived from a logistic regression of the observed values on the logit of the predicted values (Harrell 2001, Reineking and Schröder 2006). A perfectly calibrated model would exhibit a calibration curve with a slope of one and an intercept of zero. Departures from these values indicate bias and spread, respectively, in the predicted values.

Results

Dynamic population model

Not all suitable cells were occupied by butterflies. This was a consequence of local extinctions caused by predation or stochastic processes, or due to dispersal limitation. The relative proportion of occupied habitat, the prevalence, became greater when the plants had wider niches resulting in more available habitat for the butterflies, and when the butterflies were able to travel farther distances which allowed them to spread throughout more of their geographical niche and to reach even more distant patches (Fig. 2). Prevalence was also greater for scenarios without parasitism, as extinction then only occurred due to stochastic processes
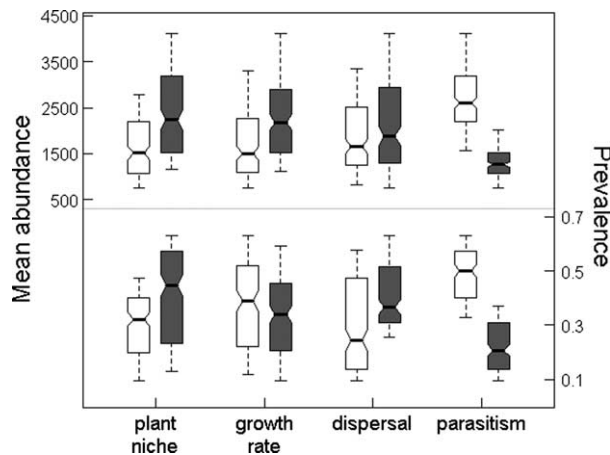
Figure 2. Effects of ecological properties on mean abundances and prevalences of butterflies in the virtual world in year 100 when butterflies were in long-term equilibrium with their environment. Open and filled boxes indicate low and high values for a given ecological parameter, respectively (Table 2). Sample size is n = 160.

in the environment. Although high butterfly growth rates caused higher butterfly abundances, prevalence was not affected, and did not differ considerably between scenarios of high and low butterfly growth rates.

In all scenarios under climate change the butterfly population moved southwards, not gradually, but rather in distinct steps which corresponded to distinct steps in the temperature trajectory (Fig. 3). After 115 yr, the butterflies started to shift their geographical range southwards which was accompanied by a small range contraction. Then after 135 yr, distinct range contractions took place accompanied by ongoing southwards movements. The northernmost patches which formerly supported intermediate relative butterfly abundances became unoccupied after year 140.

Statistical models

Under average climate, both BRTs and GLMs achieved high prediction accuracies. According to the rules of thumb given by Hosmer and Lemeshow (2000) the average discrimination ability could be considered as outstanding with (mean and median) AUC > 0.9. Furthermore, models were transferable in space without noticeable loss in predictive power (Fig. 4a). Under climate change, the distinct steps in the range dynamics were reflected in the prediction accuracies achieved for the different time slices (Fig. 4b). The range shift after year 115 caused only a slight
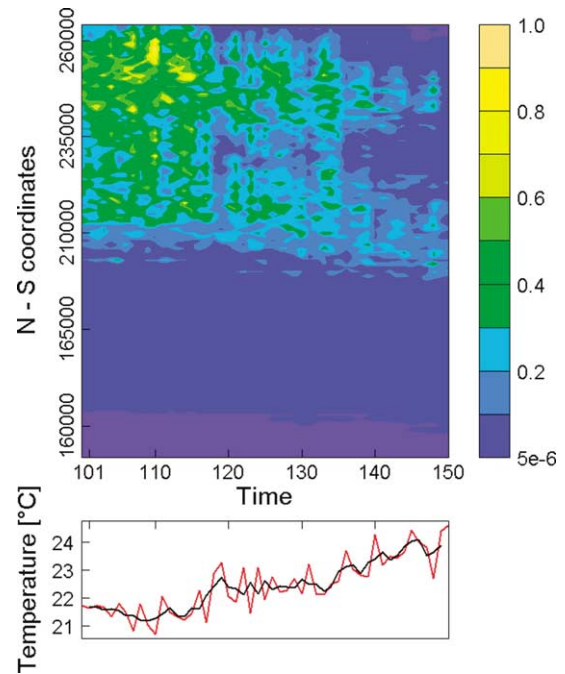
Figure 3. Range dynamics under climate change. The shift in relative butterfly abundances in north (N) -south (S) direction is depicted in the top panel. It is the average butterfly abundance of all (16) scenarios and replicate runs under climate change, and in each 1-km wide N-S transect in years 100 to 150, when climate change took place. The bottom panel shows the temperature trajectory under climate change for the years 100 to 150 (red) and the 3-yr moving average (black). Range shifts took place after year 115, after year 135 distinct range contractions occurred. These steps coincided with steps in the temperature trajectory.
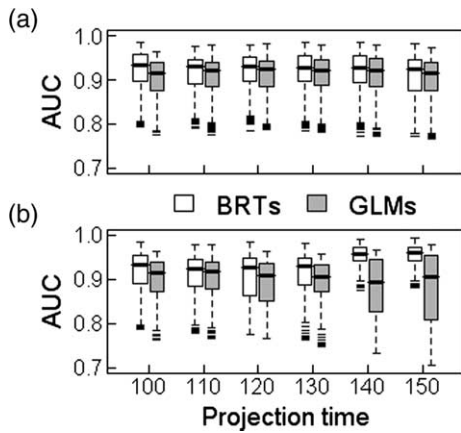
Figure 4. Boxplots of AUC values for scenarios under average climate (a) and climate change (b). SDMs were fitted in year 100; predictions were made on independent data for year 100 and every subsequent 10 yr. The AUC values achieved for BRTs and GLMs, respectively, for the different climate regimes in the respective years are depicted in the boxplots; n = 160.

difference, namely a wider range of prediction accuracies for the year 120 while for the year 130 the pattern seen in Fig. 4a was retrieved. Thus, abrupt range shifts caused a loss in predictive power in some scenarios, but, after a small time lag, predictive power was resumed. After year 140 which corresponded to the distinct range contractions, there was a profound difference between the prediction accuracies achieved by BRTs and GLMs. While the mean discriminatory power of BRTs even increased and the range decreased, the opposite was true for prediction accuracies of GLMs with a decrease in mean performance and a much wider range. Still, all scenarios yielded AUC > 0.7 indicating that for all scenarios both BRTs and GLMs were able to make useful predictions.

The effects of different ecological properties on the mean prediction accuracies achieved under climate change are shown in Fig. 5. Accuracies for the year 100 correspond to accuracies the SDMs would achieve under average climate. When the butterflies were not influenced by a parasitoid-interaction (Fig. 5a–c) differing butterfly population growth rates caused the only considerable effect on prediction accuracies, with higher AUCs for lower growth rates. There were trends that long butterfly dispersal distances and wide plant niche widths led to higher prediction accuracies. BRT prediction accuracies slightly decreased for year 120, except for scenarios with far butterfly dispersal distances where mean AUCs remained constant, and then increased and achieved even higher values in the year 150 than at the time of model estimation. GLM prediction accuracies exhibited the same pattern up to the year 130 but than strongly decreased for the years 140 and 150 where the distinct range contractions occurred. The exception were the scenarios with wide plant niche widths which only showed slight decreases in prediction accuracies under range contractions. Both BRT and GLM prediction accuracies for the year 120 only decreased for scenarios with short butterfly dispersal distances and there was this aforementioned time lag until predictive power was retrieved. Thus, butterflies with short

dispersal distances did not track the range shift instantaneously but with some time lag.

When parasitoids attacked the butterflies, the resulting prediction accuracies for the different time slices looked completely different (Fig. 5d–f). There were only minimal differences between BRTs and GLMs, and when the BRT prediction accuracies increased with range contractions, those of GLMs did as well. The virtual ecologist had perfect knowledge of the occurrence of parasitoids at all times. In the presence of a parasitoid-interaction the inclusion of this parasitoid incidence as predictor in the SDMs put GLMs on a par with BRTs. In contrast to scenarios without parasitism, butterfly population growth rates now did not show a remarkable effect on prediction accuracies anymore, and the effects of plant niche widths and butterfly dispersal distances were reversed. Butterfly dispersal ability exhibited the most pronounced effect with much higher prediction accuracies for short dispersal distances up to the year 130. When range contractions occurred this effect was smaller but still noticeable.

While for the year 150 under climate change even the lowest prediction accuracies achieved by GLMs still indicated useful predictions, they had alarmingly low explanatory power in some scenarios (Fig. 6). Without any parasitoid-interaction BRTs explained at least 50% of the deviance and achieved even values up to 90%. GLMs, on the other hand, did not even explain 20% of the deviance in five out of eight cases. Only in scenarios with wide plant niche widths and low butterfly population growth rates and in the scenario with high plant niche width, high butterfly population growth rate and far butterfly dispersal distance did GLMs have explanatory power above 40 up to 60% explained deviance. Explained deviance achieved intermediate values between 30 and 70% in scenarios with parasitism assumed, and were similar for BRTs and GLMs.
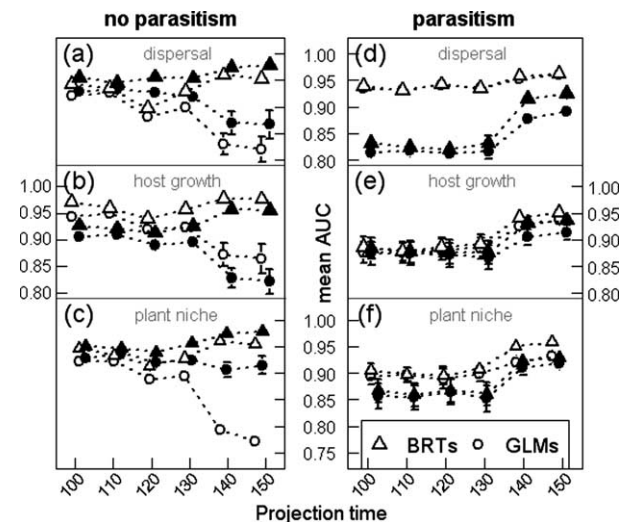

Figure 5. Effects of ecological properties on mean prediction accuracies (AUCs) of BRTs and GLMs achieved under climate change for all time slices, and for cases without parasitism (a–c) and with parasitism (d–f), respectively. Open and filled symbols indicate low and high values for a given ecological parameter, respectively (Table 2). Error bars indicate 95%-confidence intervals. Sample size is n = 40.

Calibration statistics allowed us to judge the bias and spread in the predicted probabilities of occurrence compared to observed occurrences. The calibration curves obtained for BRTs and GLMs under climate change and without parasitism are shown in Fig. 7. Both BRTs and GLMs slightly underestimated the probability of butterfly occurrence in the year 120 when the butterflies started to shift their geographical range. This was true for all scenarios regardless of dispersal ability, indicating that the butterflies persisted at the trailing edge of the range shift for some time whereas the SDMs predicted unsuitable habitat. When range contractions took place, the probability of occurrence was overestimated, only slightly by BRTs but consistently by GLMs which showed a strong bias in their predictions (Fig. 8). Overall, predictions made by BRTs fitted the observations well even under ongoing climate change. Bias and spread in the predictions differed for the different scenarios, i.e. for different ecological properties (Fig. 8). For instance for the year 150, short butterfly dispersal distances and wide plant niche widths caused BRTs to be slightly biased towards overestimating the probability of butterfly occurrence while low butterfly population growth rates caused no bias, and all other scenarios caused bias towards underestimating the probabilities of occurrence. All scenarios resulted in calibration slopes greater than one indicating that higher predicted values were underestimating the occurrence of butterflies while lower predicted values were overestimating the occurrence of butterflies. Nevertheless, for BRTs differences to the ideal calibration curve were small and in no way alarming. The strong bias in GLM predictions led to consistent overestimation of butterfly occurrence for all scenarios (Fig. 8a). Here, differing calibration slopes only determined whether this overestimation was more extreme for very low or very high predicted values (Fig. 8b). Overall, due to consistent overestimation
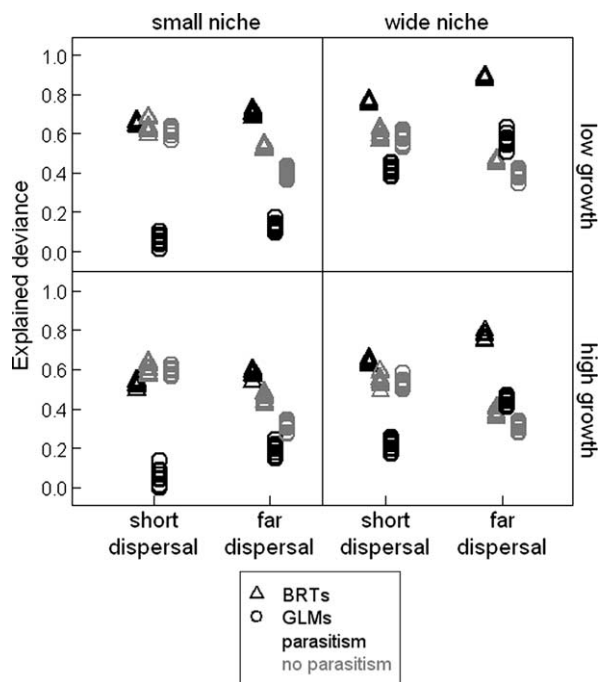
GLMs were not able to accurately predict the absences under range contraction which, in contrast, was the major strength of the BRTs.

Discussion

Prediction accuracies under climate change

The main outcome of this study was that the ways in which species respond to climate change lead to quite different projection accuracies achieved by SDMs. As one might expect, prediction accuracies initially decreased when species started to shift their geographic range due to climate change. The calibration statistics helped to explain how this mismatch between simulated true species distribution and forecasts made by SDMs was determined on one hand by the dispersal ability of the species and on the other hand by the ability of the species to endure, at least for some time, suboptimal conditions at the trailing edge of the range shift where habitat became gradually unsuitable (Morin and Thuiller 2009). Both mechanisms led to a time lag after the range shift where the predictive performance of SDMs was decreased, although the mechanisms at the trailing edge were of minor importance. The uncertainty in prediction accuracies introduced by the unlimited dispersal assumption strongly depended on the dispersal ability of the species
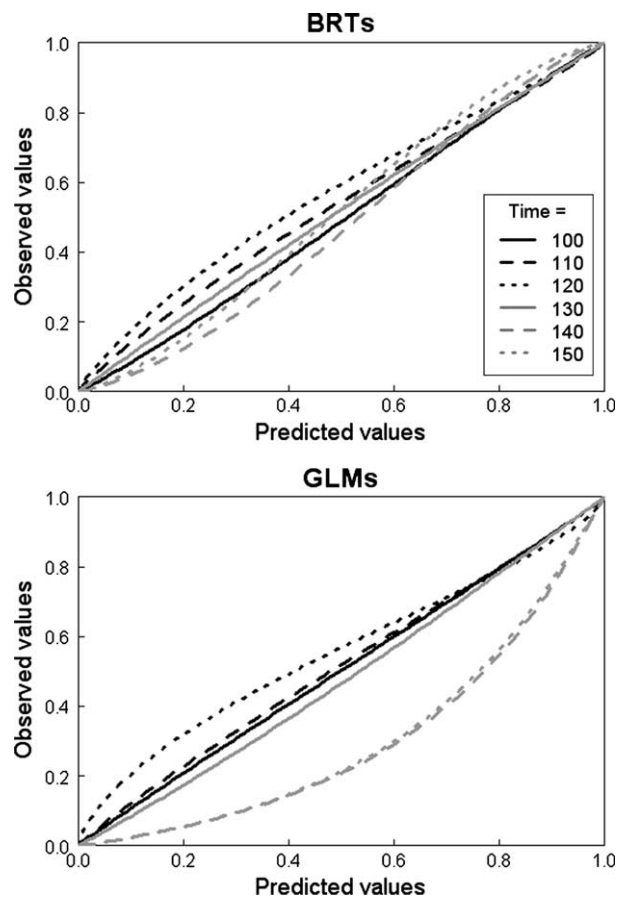

Figure 7. The mean calibration curves achieved by BRTs and GLMs under climate change without parasitism for all time slices. Sample sizes are n = 80.

Figure 6. Effects of ecological properties on explained deviance achieved by BRTs and GLMs under climate change for year 150.
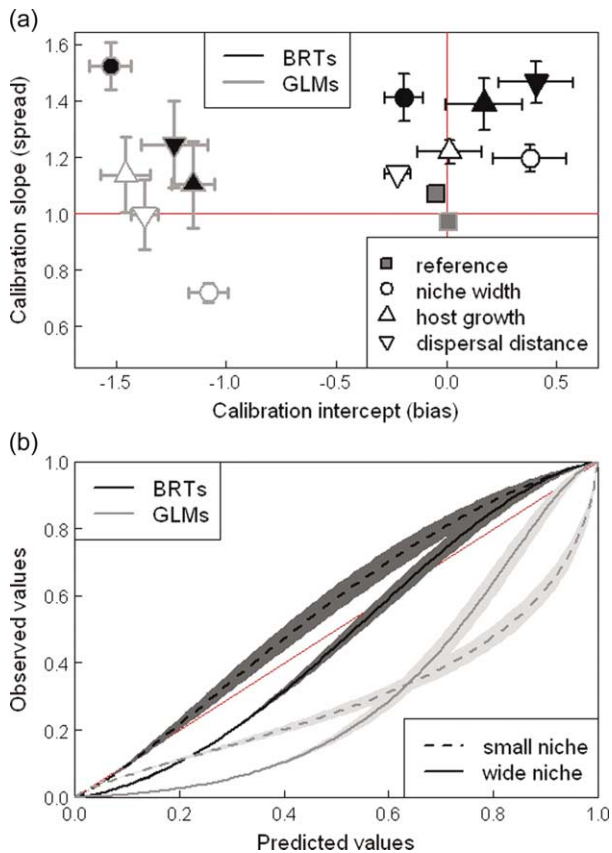
Figure 8. Effects of ecological properties on calibration statistics of BRTs and GLMs under climate change without parasitism for the year 150. (a) Open and filled symbols indicate low and high magnitudes, respectively (Table 2). The reference is the fit in year 100. Error bars (a) and shading lines (b) present the 95%-confidence interval. Sample sizes are n = 40, respectively n = 80 for the reference (a).

(Thuiller et al. 2005, Midgley et al. 2006). As soon as the species' potential geographic range remained static long enough for the species to fill its entire range, i.e. as a new equilibrium situation was realised, prediction accuracies increased again to values the SDMs would achieve under average climate. This also implied that rates and intensity of climate change are of utmost importance for the predictive performance of SDMs. If the potential geographic range of the species were to shift continuously, or if climate change were accelerated for periods of time, this would cause greater discrepancies between predicted and true species occurrences. Somewhat counterintuitively at first sight, the predictive performance of SDMs increased when species geographic ranges contracted. At second sight however, we saw that this happened because the absences were predicted more accurately. Thus, range contractions and consequently refugia could be modelled quite accurately by SDMs, and thus may allow identification of core areas for nature conservation. The results also highlight the benefit and importance of using different performance measures for SDMs. Models performing well in terms of predicting climate-change induced distributional shifts from a discrimination perspective (measured using the popular AUC statistic) may nevertheless perform poorly in terms of

calibration, with significant implications for estimates of extinction risk and colonisation success.

Model comparison

As one might have expected, BRTs outperformed GLMs both when making predictions under both average climate and climate change. The high flexibility of BRTs and their ability to model thresholds in species' occurrence made them superior to GLMs. Surprisingly, in the presence of a parasitoid interacting with the butterfly this effect vanished and BRTs and GLMs were equivalent, highlighting the importance to include biotic interactions as predictors. It was striking that the mean discriminatory power of BRTs increased when range contractions took place while the mean performance of GLMs decreased. But can we generalise these results and, thus, can we expect BRTs to generally better perform under climate change than GLMs? The answer is no, because attention should be paid to the circumstances where the models were estimated. As BRTs model thresholds in species' occurrences and extrapolate beyond the parameter range by predicting the mean response of the parameter region closest to the newly encountered parameter space, the prediction accuracy strongly depends on the extent to which the recorded occurrences correspond to the entire niche of the species. Ideally, the full range of a species should be used for estimating SDMs because then the probability is reduced to extrapolate to environmental conditions the species has never encountered before (Thuiller et al. 2004, Dormann 2007), examples are given by Pearson et al. (2002) and del Barrio et al. (2006). If this is met, then BRTs will probably always outperform parametric methods as GLMs, simply because they allow a highly non-linear threshold-like fitting of presence-absence data, rather than relying on monotone approximations such as the logistic curves of binomial GLMs.

Effects of ecological properties and processes

Several studies reported decreasing spatial prediction accuracies with increasing range sizes and niche breadth, i.e. more accurate predictions could be made for specialists than for generalists (Pearce et al. 2001, Stockwell and Peterson 2002, Segurado and Araújo 2004, Seoane et al. 2005, Hernandez et al. 2006, Brotons et al. 2007, McPherson and Jetz 2007). This however, is not beyond controversy as e.g. Garrison and Lupo (2002) reported better model performances for species with larger range sizes. We encountered both effects: when butterflies were interacting with a parasitoid, specialists were modelled more accurately. In contrast, when no parasitoid-interaction was present there was no considerable effect for spatial predictions but under ongoing climate change the distributions of generalists were modelled more accurately than of specialists. There is no easy biological explanation for this. To explain why specialists can be modelled more accurately than generalists Stockwell and Peterson (2002) suggested that widespread species may show local ecological adaptations. Modelling all these subpopulations together would effectively overestimate the species' niche, and therefore reduce model

performance. However, our dynamic model did not incorporate any local adaptation and we may thus rule out this explanation. A likely explanation why in the presence of a parasitoid-interaction SDMs performed better for specialists than for generalists is that the relationship between butterfly and parasitoid occurrence might be noisier for wide-ranging butterflies. Without parasitism a remarkable difference between model performances for generalists and specialists only occurred when the species started to shift their geographic range and in the case of GLMs increased noticeably when range contractions occurred. A reason might be that suitable habitat became more isolated with ongoing range shifts and contractions, and that this isolating effect was severer for specialists. Thus, butterflies were not able to reach all suitable habitat patches, this effect being more pronounced for narrow-ranging than for wide-ranging butterflies.

Without parasitism lower butterfly population growth rates led to higher predictive performances of both BRTs and GLMs under average climate and for range shifts. There was no considerable effect when range contractions occurred and when the butterflies were interacting with parasitoids. Higher growth rates caused higher abundances and because of local dispersal also higher prevalences, but only in particular regions where the temperature was near the optimum growing temperature. This may have resulted in biased habitat selection patterns and therefore response surfaces, obscuring the true species-habitat relationship. These findings are in contrast to the results of Seoane et al. (2005) who obtained better models for species that can reach high densities. They, however, predicted abundances instead of probabilities of occurrence, and for accurate predictions of species abundances other ecological factors may be of importance.

Under average climate and without parasitism, model performances did not differ between short and far-dispersing butterflies. This is consistent with results found by Garrison and Lupo (2002) and Stockwell and Peterson (2002). Pearce et al. (2001) on the other hand reported poorer prediction accuracies for mobile species though this effect was not significant. As mentioned before, when the species started to shift their geographic range due to climate change short-dispersing butterflies were initially outpaced by climate change. When butterflies were interacting with parasitoids prediction accuracies were much better for short-dispersing butterflies. This is an effect of parasitoids' dispersal ability. When parasitoids were worse dispersers than the butterflies then the relationship between butterfly and parasitoid occurrences was much noisier. Thus, not only the dispersal ability of the focal species is of importance but also the dispersal behaviour of species it is interacting with.

Limitations and extensions

Creating a virtual world is fraught with difficulties. Our spatially explicit tritrophic system and companion virtual ecologist represent only one possible implementation, but it enabled us to manipulate important biotic interaction and dispersal effects on SDM accuracy. One caveat of our study is the assumption that our virtual ecologist acts flawlessly and under optimum conditions: the virtual species is detected perfectly and the spatial samples are complete in coverage and instantaneous in time. In studies of real data ecologists have to deal with false negatives in the data, incomplete coverage of environmental predictors, more complex species-habitat relationships, etc., the effects of which clearly remain to be explored by future extensions of our approach.

Additionally, several other effects can be explored with this approach, which lie outside experimental manipulation in the real world. First the effect of other types of ecological processes can be investigated, such as spatial dependency of biotic interactions (i.e. some taking place only at the edge of a species distribution) or changing biotic interactions under environmental change, the effects of changing disturbance regimes or local ecological adaptation. Secondly, several statistical and sampling issues can be addressed, such as the usefulness of proxies (e.g. NDVI as surrogate for host plant abundance); effect of missing important variables (e.g. omitting incidence of parasitoids from SDMs); or density-dependence in detection probability of the focal species (a problem that underlies the development of efficient survey designs).

Perspectives and research needs in species distribution modelling

Several steps must be taken in order to improve predictions of species distributions under scenarios of environmental change. Distributional patterns of species in space and time are determined by environmental variability, and processes acting at specific spatial scales and times may be crucial for the occurrence of a species. Climate change may even increase variability in the environment both spatially and temporally, and thus this variability must be accounted for in species distribution modelling. The importance of spatial scale and hierarchical structure in ecological processes has long been recognised (Mackey and Lindenmayer 2001), but only very few studies have so far explicitly dealt with multiple spatial scales in a hierarchical manner (Graf et al. 2005, Diez and Pulliam 2007, McMahon and Diez 2007, Albert et al. 2008). Ignoring hierarchical structure in processes may be fallacious and result in erroneous projections of future species distributions under environmental change (Davis et al. 1998, Diez and Pulliam 2007, Dormann 2007). In real world studies, the underlying processes are rarely known, making, in our opinion, consideration of hierarchy in scales and processes indispensable, for example by means of multilevel modelling (Gelman and Hill 2007). Our results showed that temporal patterns of climate change and transient dynamics greatly affect SDM prediction accuracies, and other studies have suggested that temporal variability increases the probability of population extinction (cf. Thuiller et al. 2008). Midgley et al. (2006) showed that even a simple ''time-slice'' approach may reveal transient range dynamics which are obscured by ''one-step'' projections as commonly applied when projecting future species ranges by SDMs. We recommend the use of ''time-continuous'' approaches with discrete time steps in which step lengths should be as small as possible or should at least be determined by

temporal patterns and rates of expected environmental change (cf. Schröder et al. 2008 for an example on landscape-scale). Overlaying the resulting habitat suitability maps of the different time steps or years may allow the identification of core areas within a species' range (cf. Osborne and Suarez-Seoane 2007) and thus core areas for nature conservation.

Our study showed that the performance of SDMs predicting species which experienced range shifts strongly depended on two processes: dispersal at the leading edge and extinction or persistence at the trailing edge of the range shift. The incorporation of these processes into species distribution modelling is thus of major importance. Several strategies have already been tested to incorporate animal dispersal or animal dispersed pollen and seeds into SDMs, including the simple assumption of maximum dispersal rates (Midgley et al. 2006), and connectivity analyses (Graf et al. 2005, del Barrio et al. 2006). More complex approaches mechanistically modelling dispersal rely on dispersal kernels (del Barrio et al. 2006) or individual-based models (Graf et al. 2005). Applying such mechanistic approaches to range projections under climate change additionally requires integrated modelling of local population dynamics. The choice which approach to use depends on the specific aims of the study in question. However, for the prediction of species' responses to climate change an integration of dispersal and local population dynamics in a mechanistic manner seems promising (del Barrio et al. 2006, Keith et al. 2008), on the one hand because knowledge of local population dynamics is needed for the prediction of persistence at the trailing edge and local extinctions within the species' range. On the other hand, incorporation of local population dynamics may help to predict species' responses to, for example, seasonally asymmetric climate change or extreme events. Explicit consideration of dynamic species' responses may provide the basis for dynamic and integrated conservation strategies.

Biotic interactions must be included in SDMs. But then, future projections of species' distribution then also require knowledge and thus prediction of the distribution of interacting organisms (Hawkins and Porter 2003, Schweiger et al. 2008). In cases where the link between predator and prey is as tight as in our virtual world, where the parasitoid only foraged on the butterflies and was directly dependent on them, the temporal distribution of the predator may be approximated by a simple dispersal model simulating movement between patches exhibiting a probability of prey occurrence above a certain threshold. Also, the distribution of lower trophic levels on which the focal species forages must be known or predicted at all projection times introducing further uncertainties in range projections. There is no simple solution to that, and for now we must be content with making projections under specific assumptions regarding the distribution of any interacting organisms.

Conclusion

Different range dynamics may lead to quite different prediction accuracies of SDMs under climate change. A species' ability to track climate change, determined by dispersal ability and the rate of change, is decisive for SDM

performance when species shift their geographic range. Range contractions may be predicted quite accurately as the absences are predicted well. Flexible methods as BRTs will probably always outperform parametric methods due to their ability to fit non-monotone relationships. The study demonstrates the benefits and capabilities of integrating dynamic and statistical modelling approaches in different ways. On the one hand, dynamic population modelling as virtual playground for testing statistical models allows the extensive exploration of specific questions. On the other hand, the integration of dynamic processes into species distribution modelling may help to improve predictions of species distributions under environmental change. Here, the present study provides valuable insights which processes are of relevance when species undergo transient dynamics and should hence be incorporated in species distribution models.

References

Akaike, H. 1974. A new look at the statistical model identification. – IEEE T. Autom. Contr. 19: 716–723.

Albert, C. H. and Thuiller, W. 2008. Favourability functions versus probability of presence: advantages and misuses. – Ecography 31: 417–422.

Albert, C. H. et al. 2008. Land-use change and subalpine tree dynamics: colonization of *Larix decidua* in French subalpine grasslands. – J. Appl. Ecol. 45: 659–669.

Araújo, M. B. and Rahbek, C. 2006. How does climate change affect biodiversity? – Science 313: 1396–1397.

Araújo, M. B. et al. 2005. Validation of species-climate impact models under climate change. – Global Change Biol. 11: 1504–1513.

Araújo, M. B. et al. 2008. Quaternary climate changes explain diversity among reptiles and amphibians. – Ecography 31: 8–15.

Austin, M. P. et al. 2006. Evaluation of statistical models used for predicting plant species distributions: role of artificial data and theory. – Ecol. Model. 199: 197–216.

Bakkenes, M. et al. 2002. Assessing effects of forecasted climate change on the diversity and distribution of European higher plants for 2050. – Global Change Biol. 8: 390–407.

Berger, U. et al. 1999. Virtual biologists observe virtual grasshoppers: an assessment of different mobility parameters for the analysis of movement patterns. – Ecol. Model. 115: 119–127.

Bonsall, M. B. and Hassel, M. P. 2000. The effects of metapopulation structure on indirect interactions in host-parasitoid assemblages. – Proc. R. Soc. B 267: 2207–2212.

Brotons, L. et al. 2007. Updating bird species distribution at large spatial scales: applications of habitat modelling to data from long-term monitoring programs. – Divers. Distrib. 13: 276–288.

Comins, H. N. et al. 1992. The spatial dynamics of host-parasitoid systems. – J. Anim. Ecol. 61: 735–748.

Davis, A. J. et al. 1998. Making mistakes when predicting shifts in species range in response to global warming. – Nature 391: 783–786.

del Barrio, G. et al. 2006. Integrating multiple modelling approaches to predict the potential impacts of climate change on species' distributions in contrasting regions: comparison and implications for policy. – Environ. Sci. Policy 9: 129–147.

Diez, J. M. and Pulliam, H. R. 2007. Hierarchical analysis of species distributions and abundance across environmental gradients. – Ecology 88: 3144–3152.

Dormann, C. F. 2007. Promising the future? Global change projections of species distributions. – Basic Appl. Ecol. 8: 387–397.

Dormann, C. F. et al. 2008. Components of uncertainty in species distribution analysis: a case study of the great grey shrike *Lanius excubitor* L. – Ecology 89: 3371–3386.

Elith, J. et al. 2008. A working guide to boosted regression trees. – J. Anim. Ecol. 77: 802–813.

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

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

Garrison, B. A. and Lupo, T. 2002. Accuracy of bird range maps based on habitat maps and habitat relationship models. – In: Scott, J. M. et al. (eds), Predicting species occurrences – issues of accuracy and scale. Island Press, pp. 367–375.

Gelman, A. and Hill, J. 2007. Data analysis using regression and multilevel/hierarchical models. – Cambridge Univ. Press.

Graf, R. F. et al. 2005. The importance of spatial scale in habitat models: capercaillie in the Swiss Alps. – Landscape Ecol. 20: 703–717.

Guisan, A. and Thuiller, W. 2005. Predicting species distribution: offering more than simple habitat models. – Ecol. Lett. 8: 993–1009.

Harrell, J. F. E. 2001. Regression modeling strategies. – Springer.

Hassell, M. P. 2000. Host-parasitoid population dynamics. – J. Anim. Ecol. 69: 543–566.

Hassell, M. P. et al. 1991. Spatial structure and chaos in insect population dynamics. – Nature 353: 255–258.

Hawkins, B. A. and Porter, E. E. 2003. Does herbivore diversity depend on plant diversity? The case of California butterflies. – Am. Nat. 161: 40–49.

Hernandez, P. A. et al. 2006. The effect of sample size and species characteristics on performance of different species distribution modeling methods. – Ecography 29: 773–785.

Hirzel, A. and Guisan, A. 2002. Which is the optimal sampling strategy for habitat suitability modelling. – Ecol. Model. 157: 331–341.

Hirzel, A. H. et al. 2001. Assessing habitat-suitability models with a virtual species. – Ecol. Model. 145: 111–121.

Hosmer, D. W. and Lemeshow, S. 2000. Applied logistic regression. – Wiley.

Houghton, J. T. et al. (eds) 2001. Climate change 2001: the scientific basis. – Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge Univ. Press.

Jiménez-Valverde, A. and Lobo, J. M. 2007. Threshold criteria for conversion of probability of species presence to either–or presence–absence. – Acta Oecol. 31: 361–369.

Keith, D. A. et al. 2008. Predicting extinction risks under climate change: coupling stochastic population models with dynamic bioclimatic habitat models. – Biol. Lett. 4: 560–563.

King, A. A. and Hastings, A. 2003. Spatial mechanisms for coexistence of species sharing a common natural enemy. – Theor. Popul. Biol. 64: 431–438.

Leathwick, J. R. et al. 2006. Variation in demersal fish species richness in the oceans surrounding New Zealand: an analysis using boosted regression trees. – Mar. Ecol. Prog. Ser. 312: 267–281.

Mackey, B. G. and Lindenmayer, D. B. 2001. Towards a hierarchical framework for modelling the spatial distribution of animals. – J. Biogeogr. 28: 1147–1166.

McMahon, S. M. and Diez, J. M. 2007. Scales of association: hierarchical linear models and the measurement of ecological systems. – Ecol. Lett. 10: 437–452.

McPherson, J. M. and Jetz, W. 2007. Effects of species' ecology on the accuracy of distribution models. – Ecography 30: 135–151.

Menard, S. 2000. Coefficients of determination for multiple logistic regression analysis. – Am. Stat. 54: 17–24.

Meynard, C. N. and Quinn, J. F. 2007. Predicting species distributions: a critical comparison of the most common statistical models using artificial species. – J. Biogeogr. 34: 1455–1469.

Midgley, G. F. et al. 2002. Assessing the vulnerability of species richness to anthropogenic climate change in a biodiversity hotspot. – Global Ecol. Biogeogr. 11: 445–451.

Midgley, G. F. et al. 2006. Migration rate limitations on climate change-induced range shifts in Cape Proteaceae. – Divers. Distrib. 12: 555–562.

Mills, N. J. and Getz, W. M. 1996. Modelling the biological control of insect pests: a review of host-parasitoid models. – Ecol. Model. 92: 121–143.

Morin, X. and Thuiller, W. 2009. Species' range shift under climate change – reducing predictions' uncertainties through the comparison of niche and process-based approaches. – Ecology, in press.

Nicholson, A. J. 1933. The balance of animal populations. – J. Anim. Ecol. 2: 131–178.

Osborne, P. E. and Suarez-Seoane, S. 2007. Identifying core areas in a species' range using temporal suitability analysis: an example using little bustards *Tetrax tetrax* L. in Spain. – Biodivers. Conserv. 16: 3505–3518.

Pearce, J. and Ferrier, S. 2000. An evaluation of alternative algorithms for fitting species distribution models using logistic regression. – Ecol. Model. 128: 127–147.

Pearce, J. et al. 2001. An evaluation of the predictive performance of distributional models for flora and fauna in north-east New South Wales. – J. Environ. Manage. 62: 171–184.

Pearson, R. G. and Dawson, T. P. 2003. Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? – Global Ecol. Biogeogr. 12: 361–371.

Pearson, R. G. et al. 2002. SPECIES: a spatial evaluation of climate impact on the envelope of species. – Ecol. Model. 154: 289–300.

Reese, G. C. et al. 2005. Factors affecting species distribution predictions: a simulation modeling experiment. – Ecol. Appl. 15: 554–564.

Reineking, B. and Schröder, B. 2006. Constrain to perform: regularization of habitat models. – Ecol. Model. 193: 675–690.

Ridgeway, G. 1999. The state of boosting. – Comput. Sci. Stat. 31: 172–181.

Schröder, B. and Seppelt, R. 2006. Analysis of pattern-process-interactions based on landscape models – overview, general concepts, methodological issues. – Ecol. Model. 199: 505–516.

Schröder, B. et al. 2008. A landscape model for quantifying the trade-off between conservation needs and economic constraints in the management of a semi-natural grassland community. – Biol. Conserv. 141: 719–732.

Schweiger, O. et al. 2008. Climate change can cause spatial mismatch of trophically interacting species. – Ecology 89: 3472–3479.

Scott, J. M. et al. (eds) 2002. Predicting species occurences – issues of accuracy and scale. – Island Press.

Segurado, P. and Araújo, M. B. 2004. An evaluation of methods for modelling species distributions. – J. Biogeogr. 31: 1555–1568.

Seoane, J. et al. 2005. Species-specific traits associated to prediction errors in bird habitat suitability modelling. – Ecol. Model. 185: 299–308.

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

Thomas, C. D. et al. 2004. Extinction risk from climate change. – Nature 427: 145–148.

Thuiller, W. 2004. Patterns and uncertainties of species' range shifts under climate change. – Global Change Biol. 10: 2020–2027.

Thuiller, W. et al. 2004. Effects of restricting environmental range of data to project current and future species distributions. – Ecography 27: 165–172.

Thuiller, W. et al. 2005. Climate change threats to plant diversity in Europe. – Proc. Nat. Acad. Sci. USA 102: 8245–8250.

Thuiller, W. et al. 2006. Predicting patterns of plant species richness in megadiverse South Africa. – Ecography 29: 733–744.

Thuiller, W. et al. 2008. Predicting global change impacts on plant species' distributions: future challenges. – Perspect. Plant Ecol. Evol. Syst. 9: 137–152.

Tyre, A. J. et al. 2001. Inferring process from pattern: can territory occupancy provide information about life history parameters? – Ecol. Appl. 11: 1722–1737.

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

Download the Supplementary material as file E5810 from <www.oikos.ekol.lu.se/appendix>.