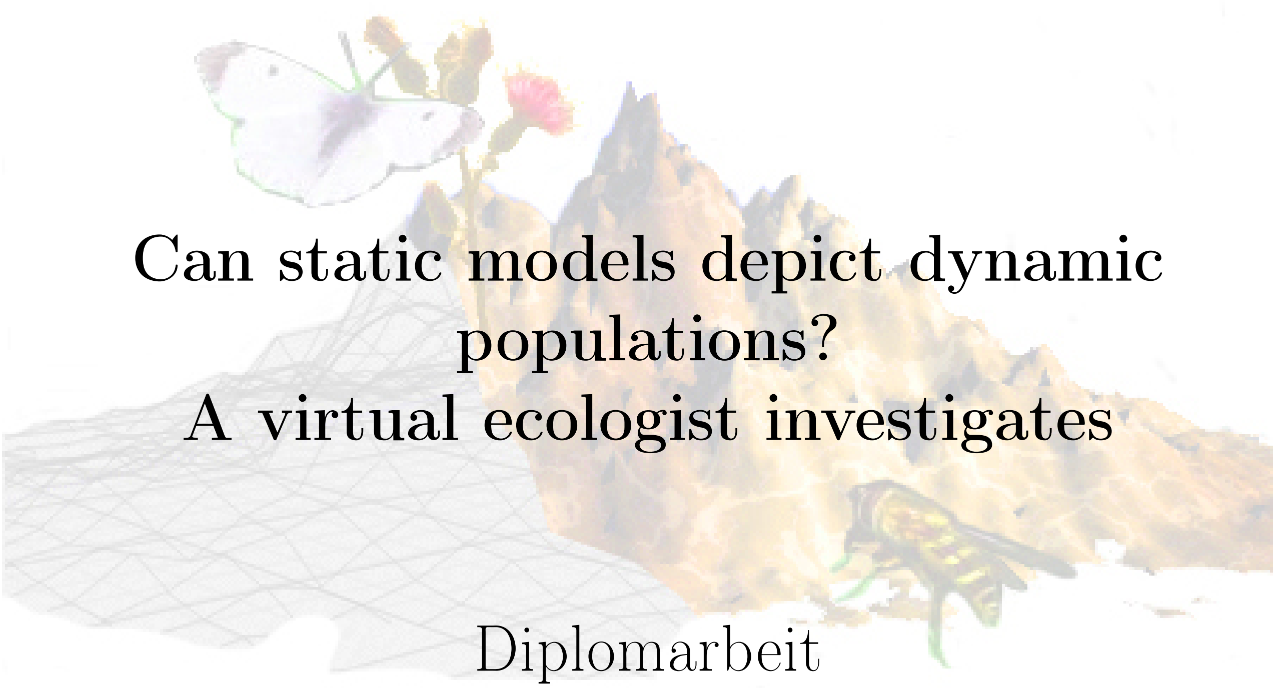
Universität Potsdam
Mathematisch-Naturwissenschaftliche Fakultät
Diplomstudiengang Geoökologie



Can static models depict dynamic populations?
A virtual ecologist investigates

Diplomarbeit

Eingereicht durch:

Damaris Zurell
Matrikelnummer: 715715

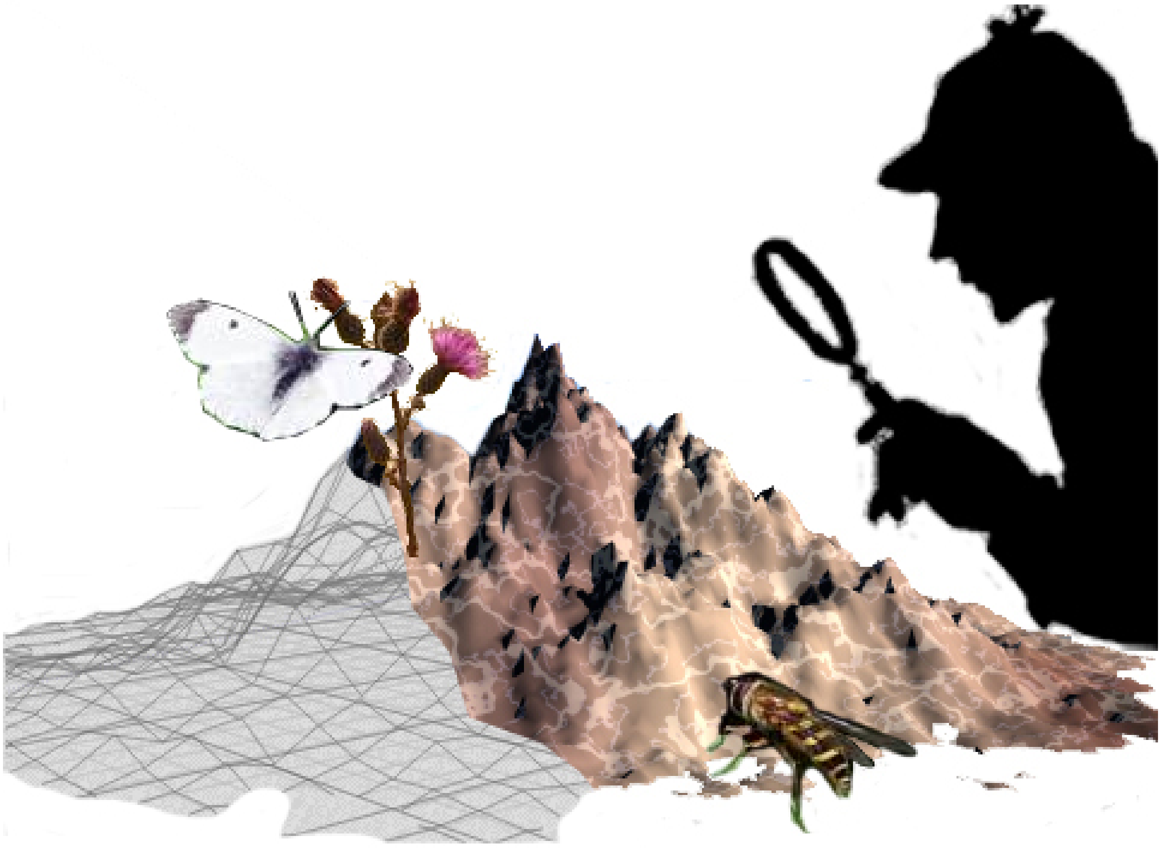
Gutachter:

Dr. Boris Schröder, Universität Potsdam
Prof. Dr. Florian Jeltsch, Universität Potsdam

Potsdam, Juni 2007

We approached the case, you remember,
with an absolutely blank mind, which is always an advantage.
We had formed no theories.
We were simply there to observe and to draw inferences from our observations.

Sherlock Holmes
The Adventure of the Cardboard Box



Contents

List of Figures

List of Tables

Abstract

As species and biodiversity are increasingly threatened by climate and land use change, there is a growing demand to predict potential changes in species distribution and composition, e. g. for conservation planning and policy making. Statistical species distribution models have been widely used in this context, though they remain subject to criticism as they implicitly assume (pseudo-) equilibrium and saturated breeding habitats, and do not explicitly incorporate demographic processes and biotic interactions. Herein I attempted to test the validity of these assumptions in a virtual dynamic system, and assess the effects of ecological processes and transient dynamics on the prediction accuracy of species distribution models. I built a spatially explicit multi-species dynamic population model, which incorporates species-specific and interspecific ecological processes, demographic and environmental stochasticity, and environmental change. A virtual ecologist sampled the population in different scenarios, and estimated species distribution models with generalised linear (GLM) and generalised additive (GAM) modelling algorithms. Spatial and spatiotemporal predictions were made and validated against simulated true species distribution. The resulting model performances were then related to the prevailing ecological processes and temporal dynamics.

I found that equilibrium state and the readiness to regain equilibrium were crucial for the validity and transferability of species distribution models. Furthermore, these effects were governed by ecological characteristics and processes, primarily stenotopy and dispersal. Prediction accuracy was also sensitive to spatial population and source-sink dynamics, with declining accuracy with an increasing use of lower-quality habitat. Above all, environmental and demographic stochasticity impeded the explanation of large proportions of deviance in species occurence. These findings were consistent throughout model algorithms. My analyses thus provide a provisional guide under which circumstances which species can be reliably predicted by species distribution models, and where interfaces to more mechanistic, process-based models are needed.

1 Introduction

Climate and land use change pose major threats to the persistence of species and biodiversity. As these are critical for the provision of ecological functions and related ecological services, spatial and temporal predictions of species distribution gain ever more importance, e.g. for deriving adequate management actions.

One way to predict potential changes in species distribution and composition are species distribution models, which have been widely used to predict species occurence in space (see Guisan & Zimmermann, 2000) and more recently also in time (Revermann *et al.*, submitted; Thuiller, 2004, 2003). These data driven models relate observed species occurence with prevailing environmental conditions. Simple species data, as presence-absence and presence-only observations, can be used to estimate the models, making them an easy to use and powerful tool for ecologists and conservationists.

But still, species distribution models remain subject to criticism as they have several shortcomings. They implicitly assume (pseudo-) equilibrium between the species and its environment, and saturated breeding habitats (Araujo & Pearson, 2005; Austin, 2002; Barry & Elith, 2006; Burgman *et al.*, 2005; Fielding & Bell, 1997; Guisan & Theurillat, 2000; Guisan & Thuiller, 2005; Guisan & Zimmermann, 2000; Pearson & Dawson, 2003). Furthermore, they explicitly incorporate neither demographic processes, nor biotic interactions such as predation and competition. I will briefly discuss how these biotic processes can promote errors in species distribution modelling.

Two critical questions arise from the assumption of pseudo-equilibrium. (1) To which degree is the species under study really in equilibrium with its environment? And (2) how long may it take to regain equilibrium, e.g. after disturbances, or to reach a new stable state, e.g. after environmental change? Following Hutchinson (1957), the species is in equilibrium with its environment when it occurs in all environmentally suitable areas and its distribution thus represents its full biotic potential. In reality, however, the species distribution may be constrained locally by history, disturbance and other temporal dynamics (Austin, 2002; Barry & Elith, 2006; Guisan & Thuiller, 2005; Pearson & Dawson, 2003).

For instance, several studies have emphasised that not all present-day species have yet reached equilibrium since the last glacial maximum (Araujo & Pearson, 2005; Svenning & Skov, 2004, 2007). When predicting species distributions under changed environmental conditions, an instantaneous realisation of a new equilibrium situation is implicitly assumed, thereby ignoring transient dynamics. Also, when the model predicts a range shift under e.g. climate change, no inferences can be made about when this process will be terminated.

The assumption of saturated breeding habitats indicates that higher quality habitat is proportionally used more frequently than lower quality habitat (Fielding & Bell, 1997). This contradicts the concepts of metapopulations and source-sink dynamics (Pulliam, 1988). According to these concepts species may be present at unsuitable sites (sink habitats) through immigration from source habitats. They may also be temporarily absent from suitable habitat due to local extinction events as a consequence of demographic and environmental stochasticity. In this way the statistical model may falsely recognise unsuitable habitat as suitable and vice versa, since it merely considers a snapshot of the dynamic situation.

Species distribution models intrinsically incorporate the effects of demographic processes, predative, competitive and other biotic interactions (Guisan *et al.*, 2002; Hijmans & Graham, 2006). The realised niche is the species response to those many complex interactions. But not all possible combinations of these factors will exist at any one place or time. Rather these interactions will vary in space and time. So actually, through taking a snapshot view of a finite space, the realised niche modelled is only valid for that particular (sub-) population at that particular time. Generalisability for the whole species and transferability in space and time must be thouroughly validated (Schröder & Richter, 1999).

In answer to these supposed pitfalls of species distribution modelling, it has been argued that the models need to become more dynamic, e.g. by linking or supplementing them with more mechanistic, process-based models, in order to better describe transient responses to a stochastically and dynamically changing environment (Austin, 2002; Burgman *et al.*, 2005; Dettki *et al.*, 2003; Guisan *et al.*, 2006; Guisan & Thuiller, 2005; Kearney & Porter, 2004;

Pearson & Dawson, 2003; Pearson *et al.*, 2006). But to date, it remains uncertain under which circumstances the incorporation of dynamic processes is really necessary. Only few studies have so far tried to systematically investigate how ecological processes and temporal dynamics influence the performance of species distribution models (but see Brotons *et al.*, 2004; McPherson & Jetz, 2007).

The objectives of this study were thus to assess the accuracy of spatiotemporal predictions of species distribution models when confronted with non-equilibrium dynamics, and to assess how species-specific ecological traits and interspecific interactions influence the effect of transient dynamics on model performance. But all these mechanisms and circumstances are hard or even impossible to capture in nature, and also transferability in space and time can not be tested easily due to the lack of data. I therefore applied a virtual species approach. This gave me the advantage of a fully known reality, where not only the true distribution of the species under study was known, but also the processes determining the observed patterns.

I built a dynamic, spatially explicit muli-species dynamic population model which incorporates species-specific traits as the ecological niche, dispersal ability and intrinsic growth rate, interspecific characteristics as competition and predation, demographic and environmental stochasticity, and environmental change. Different scenarios were calculated by systematically manipulating model properties within a certain range, species growth rates, the strength of biotic interactions, species range, dispersal strategies, and the rate of climate change.

A virtual ecologist sampled the populations at different points in time, providing me with snapshots of the dynamic situation, and calculated statistical species distribution models from the data. Models were build using generalised linear (GLM, see McCullagh & Nelder, 1989) and generalised additive modelling (GAM, see Hastie & Tibshirani, 1990) algorithms. These methods have been widely used to model and predict species occurence and have the advantage that they are implemented in most statistical software packages. Moreover, several studies have shown that the different methods lead to quite similar results (e.g. Elith, 2000; Elith *et al.*, 2006; Ferrier & Watson, 1997). I thus focussed on the comparison of parametric, model-driven GLMs and semi-parametric, data-driven GAMs, in regard to their capability to capture the species-habitat relationship, and their transferability in space and time. Both spatial and spatiotemporal predictions of the species distributions were made and evaluated against the real distribution of the virtual species.

2 Methods

2.1 Dynamic Model

The virtual world comprises an insect predator-prey system where the predator is a holometabolous insect that lives parasitically during its larval stage on an arthropod host and thereby causes the death of the host (Mills & Getz, 1996). Such host-parasitoid interactions are well suited for simple population models because they can have a much simpler structure than other enemy-victim interactions due to the tight link between trophic and reproductive aspects of the parasitoid life history (Hassell, 2000a; King & Hastings, 2003; Nicholson, 1933). As only the adult female parasitoids search for hosts and oviposit directly upon finding them, this act of parasitism closely defines the reproductive success of the parasitoids while in other predator-prey interactions several or all predator stages attack with different effectiveness (Hassell, 2000a; Schofield *et al.*, 2005).

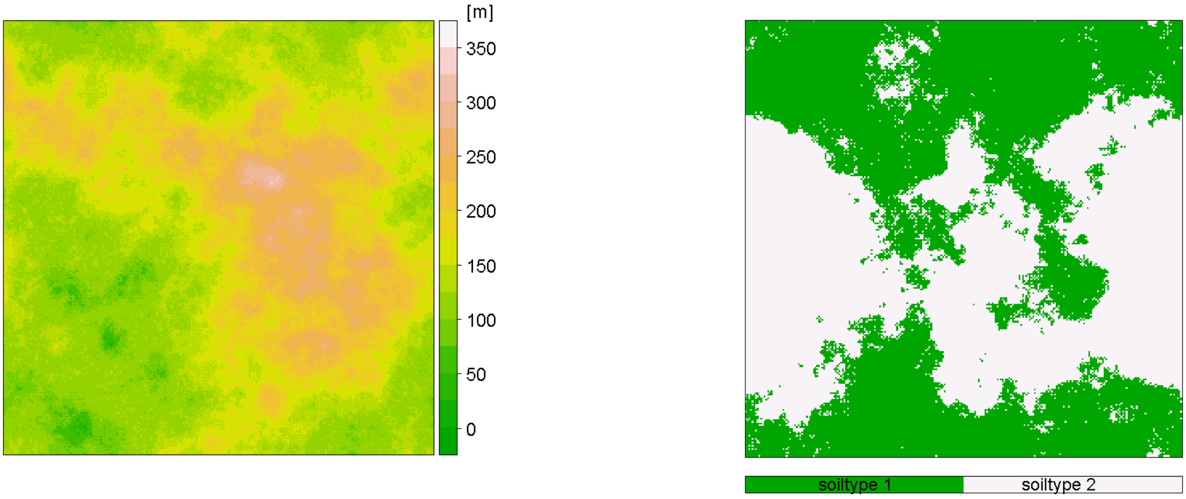
In a discrete-generation framework a perfectly synchronised parasitoid interacts with a host that has distinct generations and is univoltine. This is frequent in temperate regions where a diapause during winter is common (Hassell, 2000b; Hochberg & Holt, 1995; Mills & Getz, 1996). Movement and dispersal are limited to the adult insect stages. In analogy to real systems the model is tritrophic, i.e. the host-parasitoid interaction occurs in the presence of a host-plant which affects the presence and the abundance of the host. The model system is integrated in a dynamic environment.

A coupled-map lattice (CML) model is used to link the local (within-patch) and regional (between-patch) dynamics (Bonsall & Hassel, 2000; Comins *et al.*, 1992; Hassell *et al.*, 1991; Jones *et al.*, 1996; Wilson & Hassel, 1997). Space is represented by a two-dimensional lattice of 257×257 sites. At each site, the local host-parasitoid population dynamics are mapped annually. The populations are then connected by dispersal. I thus obtain a spatially explicit multi-species dynamic population model which may be systematically modified in several ways.

2.1.1 Environment

The two-dimensional grid is composed of different layers providing information on abiotic habitat conditions. Main input layers are topography and soil type distribution (Figure 1). Both are neutral landscapes generated using the midpoint displacement algorithm (Hargrove *et al.*, 2002; Moloney & Levin, 1996; Saupe, 1988; With, 1997). This fractal algorithm creates a three-dimensional surface, a topographical map, characterized by two parameters, the Hurst-factor H controlling the degree of spatial autocorrelation and the variance σ^2 in the displacements of points.

To generate a binary soil map a percentage of occupancy p is specified. Soil types are assigned to the fractal landscape by slicing the surface at p% of the elevational range and attributing a soil type to each slice.



(a) (b)

Figure 1: Maps of (a) elevation and (b) soil type distribution created by fractal algorithm.

Additional ecogeographical information layers are calculated. Potential soil moisture distribution is approximated by the extended topographic wetness index W (Beven & Kirkby, 1979; O'Neill *et al.*, 1997):

$$W_{ij} = \ln\left(\frac{A_{ij}}{t_{ij}\tan\beta_{ij}}\right) \tag{1}$$

where A_{ij} is the specific catchment area, t_{ij} the local soil transmissivity and β_{ij} the surface slope of the grid cell.

Temperature is interpolated by considering altitudinal differences, the temperature vertical gradient T_{lapse} and radiation (Bellasio *et al.*, 2005; Moore *et al.*, 1993; Wilson & Gallant, 2000):

$$T_{ij} = T - T_{lapse}\left(\frac{z_{ij}}{1000}\right) + C\left(S_{ij} - \frac{1}{S_{ij}}\right) \tag{2}$$

where T is the temperature averaged over the whole landscape at elevation $z = 0$ (see 2.1.5), z_{ij} the elevation of the grid cell, C a constant and S_{ij} the ratio between the insolation of the cell and the horizontal surface. The insolation is truncated to the cosine of the solar illumination angle i :

$$\cos i = \cos\theta_0\cos\beta + \sin\theta_0\sin\beta\cos(\phi_0 - A) \tag{3}$$

with θ_0 and ϕ_0 the solar zenith angle and azimuth, respectively, A is the aspect and β the surface slope (Dubayah & Loechel, 1997).

2.1.2 Within-Patch Dynamics

The difference-equation framework for my discrete-generation host-parasitoid model can be written in the generalized form:

$$\begin{aligned} N_{t+1} = Poi\left\{N^*_{t+1}\right\} &= Poi\left\{\lambda N_t g(N_t) f(P_t)\right\} \\ P_{t+1} = Poi\left\{P^*_{t+1}\right\} &= Poi\left\{c N_t[1 - f(P_t)]\right\} \end{aligned} \tag{4}$$

where N and P are the population abundances of the susceptible host stage and the searching adult female parasitoid, respectively, in generations t and $t+1$, λ is the net finite rate of increase of the host population, $g(N_t)$ the density-dependent survival of the hosts progeny, $f(P_t)$ is the proportion of hosts escaping parasitism. One minus this term therefore gives the proportion of host individuals parasitised, and c includes the average number of adult female parasitoids emerging from each host parasitised. Poi indicates that the actual population abundances at $t+1$ are drawn from a Poisson distribution with mean N^*_{t+1} and P^*_{t+1} respectively. This way the populations' reproductive success is not strictly deterministic but demographic stochasticity is accounted for.

Functional Response The function $f(P_t)$ depends on all factors affecting the level of parasitism. Nicholson (1933) and Nicholson & Bailey (1935) explored a version of model (4) in which the functional response is linear (type I functional response) and the attacks are randomly distributed amongst the host population. The fraction of hosts escaping parasitism is given by the zero term of a Poisson distribution (5) with mean aP_t where a is the area of discovery.

$$f(P_t) = exp(-aP_t) \tag{5}$$

Host Growth The intrinsic rate of increase λ is modelled temperature-dependent because in reality the metabolic rates of organisms also strongly depend on it. Above a certain threshold biological activity increases with increasing temperature until a maximum rate is reached at an optimal temperature T_{opt} and then decreases again until an upper lethal value. The relationship between physiological rates and temperature is generally humped and skewed to the left. O'Neill (1968) and O'Neill *et al.* (1972) suggested an empirical equation with empirically derived parameters often referred to as O'Neill or Oak Ridge Temperature Function (Bartell *et al.*, 1988; Richter *et al.*, 1996). Nevertheless in my model an approximation of the function's typical shape shall suffice. I therefor apply a Gumbel distribution with T_{opt} as location parameter and a scale parameter σ (Figure 2). I restrain the population's growth rate even more by introducing a simple but strong Allee effect (Allee, 1931). Below a critical population size N_{crit}, or extinction threshold, no reproduction occurs and the local population goes extinct. One might imagine the Allee effect being due to predation or mate shortage.

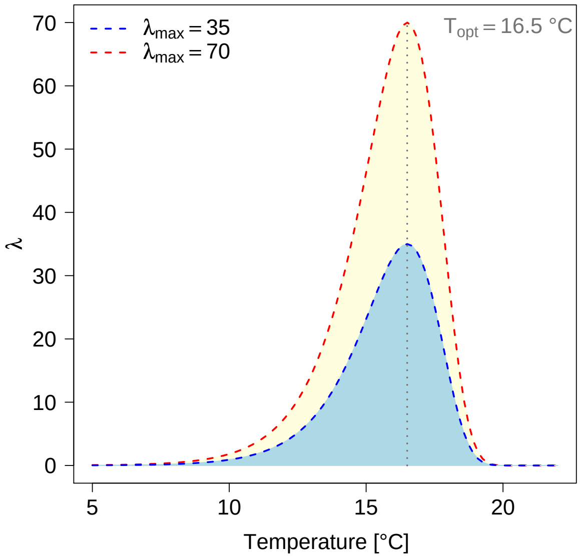


Figure 2: Temperature dependence of the host intrinsic rate of increase λ.

Host Density Dependence In my virtual world the host depends on the presence and the abundance of the host-plant. Thus even in the absence of parasitism the host-plant induces a host carrying capacity K in each lattice cell. This can be expressed via $g(N_t)$ in (4). Cook (1965); Hassell (2000b); May (1974); Moran (1950); Ricker (1954) proposed a discrete version of the logistic equation,

$$g(N_t) = exp(-sN_t) \tag{6}$$

where

$$s = \frac{\ln \lambda}{K} \tag{7}$$

The carrying capacity solely depends on the abundance of the host-plant and thus on the habitat quality of the cell. I assume a maximum carrying capacity K_{max} for the maximum possible host-plant biomass in a cell and further, that the foliage projective cover is directly proportionate to the physiological response, i.e. the maximum habitat quality yields K_{max} if no competition is present. Hence the actual carrying capacity K is given by K_{max} times the foliage projective cover F.

Function	Symbol	Value	Unit	Description
	w	100	m	cellwidth
Environment	H_{topo}	0.7	-	Hurst-factor for generating topographic map
	σ_{topo}	15	-	standard deviation in the displacements of points in topographic map
	H_{soil}	0.7	-	Hurst-factor for generating soil map
	σ_{soil}	15	-	standard deviation in the displacements of points in soil map
	$p_1 = p_2$	0.5	-	proportions of soil types in fractal soil map
Temperature	θ_0	78.25	$^\circ$	solar zenith angle
	ϕ_0	180	$^\circ$	solar azimuth
	T_{lapse}	10	Kkm^{-1}	temperature vertical gradient
	C	1	-	constant in equation (2)
Moisture	t_{s1}	30	m^2d^{-1}	transmissivity of soil type 1
	t_{s2}	5	m^2d^{-1}	transmissivity of soil type 2
Host density dependence	K_{max}	120	ind	maximum carrying capacity of a cell
	$\mu_{T,hp}$	16.5	°C	mean of host-plant's temperature utilization function
	$\mu_{W,hp}$	5.5	-	mean of host-plant's moisture utilisation function
	$\mu_{T,cp}$	12	°C	mean of competitor-plant's temperature utilisation function
	$\sigma_{T,cp}$	1	°C	standard deviation of competitor-plant's temperature utilisation function
	$\mu_{W,cp}$	10	-	mean of competitor-plant's moisture utilisation function
	$\sigma_{W,cp}$	1	-	standard deviation of competitor-plant's moisture utilisation function
	σ_F	0.1	ind	standard deviation for environmental stochasticity
Host growth	$T_{opt,\lambda}$	16.5	°C	optimal temperature for host growth rate, location parameter of Gumbel distribution
	σ_λ	1.4	°C	scale parameter of Gumbel distribution
	N_{crit}	5	ind	critical population size for Allee effect
Dispersal	β_{host}	1.5	-	host's shape parameter in equation (9)
	μ_{host}	0.75	-	fraction of local host population emigrating
	α_{para}	2	-	parasitoid's scale parameter in equation (9)
	β_{para}	1.5	-	parasitoid's shape parameter in equation (9)
	μ_{para}	0.75	-	fraction of local parasitoid population emigrating
Parasitism	c	1	ind	parameter in equation (4)

Table 1: Constants in the dynamic model

F, the physiological response of the host-plant, is determined by the prevailing environmental conditions in a cell and represents the fundamental niche of the host plant. It depends on energy and water availability, in particular on the temperature and moisture conditions of a cell, each representing a one-dimensional resource spectrum. The response of the host-plant is described by a Gaussian utilisation function with the mean being the preferred position in the spectrum and a characteristic variance (May & Mac Arthur, 1972). I have to consider that the topographic wetness index W merely represents the potential soil moisture distribution as it refers to each cell's capability of water retention relative to the whole area, while the actual amount of retained water of course depends on the actual precipitation P. Thus the same cell with one and the same wetness index may in some

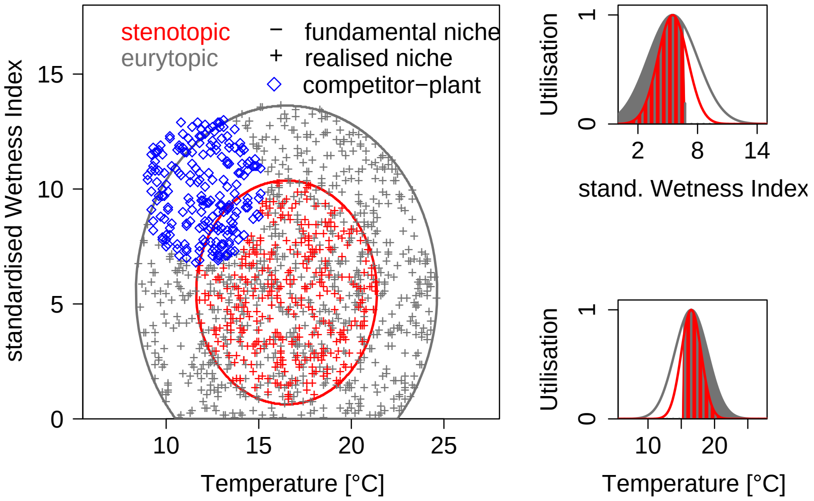


Figure 3: The fundamental and realised niche of the host-plant.

years be too moist or too dry for the plant to persist while in others it offers most suitable conditions. To account for this I standardise the topographic wetness index:

$$W_{ij,corr} = W_{ij}\left(\frac{P}{P_{base}}\right) \tag{8}$$

$W_{ij,corr}$ is the corrected or standardised wetness index, P_{base} is the base precipitation to which the plant's optimum resource utilisation refers, P is the actual precipitation of the timestep. That way in dry years (relative to the base precipitation) the cell's wetness index and accordingly the amount of retained water becomes lower and in moist years higher.

Following Liebig's law of the Minimum, the plant's physiological response F is calculated by multiplying the degree of utilisation of each resource. To incorporate environmental stochasticity, the actual physiological response is drawn from a normal distribution with mean F and a variance σ_F^2.

Competition for resources is introduced by adding a second plant species whose fundamental niche overlaps with that of the host-plant. The competitor-plant always out-competes the host-plant resulting in a narrower, realised niche of the latter (Figure 3).

If the environmental conditions are suitable, host-plant and competitor-plant are seeded instantaneously and hence are not restricted by any dispersal rules. However, to avoid too abrupt changes in the plant distribution I built in a simple memory effect such that the actual capacity of the new time step is the arithmetic mean between the calculated capacity of time step $t+1$ and the old capacity of time t. In the end host-plant and competitor-plant species each persist as a single large patchily distributed population, forming a matrix of suitable and unsuitable habitat patches for the host and indirectly the parasitoid.

2.1.3 Between-Patch Dynamics

In the between-patch dispersal phase a proportion of adult hosts and parasitoids leave their natal patch to colonise other cells in the lattice. Local dispersal is assumed, which means that the dispersers will be concentrated around the area in which they developed as juveniles according to the underlying dispersal rules. The probability p_{ij} that an individual disperses from cell i to j over the integer distance d_{ij} is described by a two-parameter Weibull distribution which allows different dispersal strategies (Söndgerath & Schröder, 2002):

$$p_{ij} = \frac{\exp(-\alpha {d_{ij}}^{\beta})}{\sum_j \exp(-\alpha {d_{ij}}^{\beta})} \tag{9}$$

where the scale parameter α determines the dispersal distance. A high value of α indicates short-range dispersal, a low one large-range dispersal. At $\alpha = 0$ the dispersers are evenly distributed throughout the lattice (global dispersal). The Weibull distribution is further defined by the shape parameter β which determines the shape of the rate function. In the case of $\beta = 1$ the Weibull distribution equals the exponential distribution and the dispersal from cell i to j is described by an exponentially decreasing probability (Johst *et al.*, 2002; Johst & Drechsler, 2003; Neubert *et al.*, 1995). The integer distance d_{ij} between cells depends on the applied neighbourhood rule, in this case an 8-cell (Moore) neighbourhood (Hogeweg, 1988).

To account for some demographic stochasticity the actual integer number of dispersing individuals is drawn from a binomial distribution with mean $p_{ij}(\mu_x X_t)$ where μ_x is the fraction of the local population X_t emigrating, which is constant over all cells (Johst &

Drechsler, 2003). Individuals dispersing to unsuitable habitats, i.e. cells without host and host-plants respectively, die. I assume cyclic boundary conditions.

2.1.4 Climate Data and Climate Change

Main climate variables in the virtual world are energy and water, in particular mean summer temperatures and mean annual precipitation. Long-term average values and typical variations for Central Europe were used to generate time series of current climate.

		Mean	Standard Deviation
Precipitation	2000	640 mm	50 mm
	2050	580 mm	50 mm
Temperature	2000	16 ℃	0.7 ℃
	2050	19 ℃	0.7 ℃

Table 2: Climate parameterisation

A generalised climate change scenario was derived for the period 2001 to 2050. Rates of change were obtained from the projected changes of the IPCC SRES scenario A2 for Central and Northern Europe (Houghton *et al.*, 2001). In this scenario temperature will increase by 3℃ in 50 years, precipitation will decrease by 10%. Parameters are shown in Table 2.

Time series were calculated through linear interpolation between current and future values taking into account interannual variations.

2.1.5 Process Simulation and Scenarios

A standard simulation lasts 120 years. At the beginning, hosts and parasitoids are randomly distributed over the suitable habitat. The model 'spins up' for 70 years, running with the long-term climatic average, ensuring a long-term equilibrium between the host population and its abiotic and biotic environment. Such equilibrium is reached when extinction and recolonisation events balance each other. Scenarios are applied over the last 50 model years, during which the population is sampled by the virtual ecologist.

Scenarios were derived in which ecological processes and temporal dynamics were systematically manipulated. Host's growth rate, the strength of biotic interaction, i.e. the level of

Function	Symbol	Values		Unit	Description
Host density dependence	$\sigma_{T,hp}$	1.5	2.5	°C	standard deviation of host-plant's temperature utilisation function
	$\sigma_{W,hp}$	1.5	2.5	-	standard deviation of host-plant's moisture utilisation function
Host growth	λ_{max}	35	70	ind	maximum intrinsic rate of increase
Dispersal	α_{host}	0.5	4	-	host's scale parameter in equation (9)
Parasitism	a	1	10	ind	area of discovery in equation (5)

Table 3: Manipulated parameters in the dynamic model

parasitism, the resource utilisation of the host-plant, the dispersal strategy, and the rate of climate change were varied within a certain range. This range depended on the boundary condition, that both the host and the parasitoid population must persist throughout the whole simulation. Once the parameter ranges were identified I took their minimum and maximum value, respectively, as input for the different scenarios (Tables 2 and 3). All possible parameter combinations were used with both constant and changing climate, resulting in 32 scenarios. The host's intrinsic rate of increase was either high or low (Figure 2), parasitoids either attacked only one host or several ones, host-plants were either stenotopic or eurytopic regarding their resource preferences (Figure 3), and finally the hosts assumed either very local or rather globalised dispersal compared to the parasitoids. I ran five replicate simulations for each scenario.

2.2 Virtual Ecologist

2.2.1 Sampling

Similar to real field surveys, a virtual ecologist sampled the host population with the same strategy as an ecologist would probably select in reality. There are several sampling techniques available. Hirzel & Guisan (2002) however, suggested regular and equal random stratified sampling the most accurate and most robust strategies. My virtual ecologist uses the latter sampling design where the sampling is stratified by those environmental predictors which are believed to have the greatest potential influence on species distribution (Hirzel & Guisan, 2002; Maggini *et al.*, 2002; Tyre *et al.*, 2001). These predictors must be chosen *a priori*. As I know that the host directly depends on the presence of the

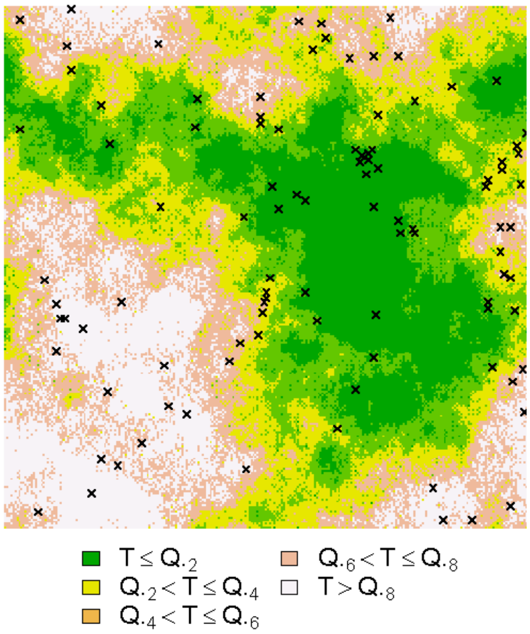


Figure 4: Example of stratified sampling design. The quantiles of the temperature distribution are taken to split the data into 5 ordinal classes.

host-plant, whose ecological niche is defined by temperature and water availability, and hosts growth rate is temperature-dependent, it would be best to use temperature and water availability as stratifying variables. But because water availability is distributed very heterogenously at a small spatial scale, solely temperature is used for stratification. The quantiles (0.2-,0.4-,0.6- and 0.8-quantiles) of the temperature distribution are taken to split the data into five ordinal classes, generating five homogenous environmental strata with respect to temperature (Figure 4).

An equal number of cells (replicates) is randomly chosen in each stratum. In a selected cell, the incidence of the species and the prevailing environmental conditions are mapped. In order to avoid spatial autocorrelation in the response variable samples are not taken in adjacent cells. Moreover, cells bordering the strata are not sampled because they may exhibit transitional conditions. The virtual ecologist makes no errors in detection, i. e. species and environmental variables are recorded exactly as given in the dynamic model.

This way the output of the statistical model can be related to the underlying demographic and stochastic processes (Tyre *et al.*, 2001).

Samples were taken in three different years, after spinup phase (year 70), after 95 years and at the end of the simulation (year 120). In the case of environmental change, I thus obtained samples of different stages of equilibria. The minimum sample size was 100. This is comparable to actual empirical studies of this type (Binzenhöfer *et al.*, 2005; Hein *et al.*, 2007). Additionally the virtual ecologist took samples of sizes 200 and 500, respectively, to measure the possible effect of sample size on the accuracy of the species distribution models. The sampled datasets were balanced with a prevalence of approximately 0.5, because this optimises accuracy of logistic models and ensures comparability between models (McPherson & Jetz, 2007; McPherson *et al.*, 2004). Sampling was repeated five times for each simulation run.

2.2.2 Species Distribution Model

Model Formulation Species occurences were coded as 0 for absences and 1 for presences. Both GLMs and GAMs were then specified using a binomial error distribution and a logistic link function. To increase robustness and generalisability we sought to find the most parsimonious model that still explained the data. Following Hosmer & Lemeshow (2000) and Harrell (2006), model selection was carried out in three steps for GLMs and in two steps for GAMs. Foremost in the GLM approach, univariate models were estimated for every predictor variable. Those variables, whose significance test yielded a p-value $>$ 0.25, were excluded from further analysis. Subsequently, the remaining and the full set of predictor variables for GLMs and GAMs, respectively, were tested for multicollinearity by calculating Spearman's rank correlation coefficient ρ_S. In accordance with Fielding & Haworth (1995), if two variables had a correlation greater than 0.7 the predictors with the bigger p-value in the univariate GLM and the smaller deviance in the univariate GAM, respectively, were excluded. For the remaining variables and for both GLM and GAM, I apllied an Akaike information criterion (AIC)-based stepwise variable selection procedure

(Akaike, 1974). Serious shortcomings have been idenitified for stepwise selection procedures (Whittingham *et al.*, 2006). I chose it nevertheless for computational speed and because the main focus of this study lies on the average model accuracy, rather than on finding the single best set of predictor variables. Furthermore I argue, that there is no single best set of predictors because the detectability of the species-habitat relationship is limited by sample size. I checked for significance of linear and quadratic terms in GLM, whereas the linear term was forced into the model each time the quadratic term was selected in the final model. In GAM the predictors were allowed as parametric linear and up to second-order polynomial terms, and as non-parametric cubic smoothing splines with up to four degrees of freedom.

Spatiotemporal Predictions For both the GLM and GAM models, spatial predictions were made for the entire region, and spatiotemporal predictions for the entire region and the respective two remainder time slices of the simulation run. Thus both past and future species distributions were predicted, depending on the year of the virtual study.

Model Evaluation The proportion of deviance in species occurence explained by the final models was quantified using the explained deviance r^2_L, an equivalent to the coefficient of determination r^2 in logistic regression (Menard, 2000).

I evaluated the models ability to discriminate between occupied and non-occupied sites, and the degree of correspondence between the estimated probabilities and the observed incidences (model's calibration).

The area under the receiver operating characteristic (ROC) curve (AUC) was used to measure model discrimination (Hanley & McNeil, 1982). AUC is independent of classification thresholds and ranges from 0 to 1, with values 0.5 for models with no predictive ability and 1 for perfectly discriminating models. Values above 0.7 indicate useful predictions according to Hosmer & Lemeshow (2000).

Model calibration was evaluated by calibration curve (Miller *et al.*, 1991), a logistic regression line fitted to the logits of the predicted probabilities and the observed incidences. A perfectly calibrated model yields a zero intercept and a unit slope (Harrell, 2006; Reineking & Schröder, 2006; Thomson *et al.*, 2007; Wintle *et al.*, 2005), departures indicate bias and consistent over- and underestimation of occupancy probabilities.

2.3 Link between Prediction Accuracy and Population Dynamics

I examined the effect of transient dynamics and ecological processes on model performance using univariate and multivariate linear regressions. As measure of prediction accuracy and thus response variable in regression analysis I used AUC, calculated for the entire data, i.e. the true species distribution. Linear models (LMs) were estimated using *niche width*, *host growth*, *level of parasitism*, *dispersal distance*, *time elapse since spinup* and *sample size* as predictor variables. In the multivariate linear regressions I allowed two-way interactions between the predictors. To ensure that interaction variables were uncorrelated with their component variables I orthogonalised the interactions by regressing the interaction variable on its two component variables and taking the residuals as new interaction variable into the multivariate LM (Burrill, n.d.). An AIC-based stepwise modelling procedure was applied to identify parsimonious models. The statistical significance of the predictors was assessed using likelihood ratio tests (Hosmer & Lemeshow, 2000). Significance levels were adjusted for multiple comparison using the Holm method (Aickin & Gensler, 1996).

Computation The dynamic model was programmed in C++. All other computations were carried out within the free software environment R 2.5.0 (R Development Core Team, 2007) using the packages Hmisc (Harrell Jr., 2007), Design (Harrell Jr., 2005), and gam (Hastie, 2006).

3 Results

3.1 Dynamic Model

Not all suitable cells were occupied by hosts, and local populations were often unsaturated. The absence of individuals from suitable habitat was due to local extinctions, caused by predation or stochastic processes, and dispersal limitation. Extinction and recolonisation events soon balanced each other, resulting in a dynamic equilibrium situation. Thus the dynamic model produced plausible spatial population dynamics on the fragmented landscape.

The different parameter combinations led to some characteristic differences in population dynamics. As one might expect, the relative proportion of occupied habitat became greater as the species was able to adapt to a wider range of environmental conditions, i. e. switched from stenotopic to eurytopic adaptation. In the scenarios with very local dispersal the subpopulations were much smaller than in case of long-distance dispersal, resulting in fine-grained distribution patterns while in the latter scenarios subpopulations were bigger, thereby producing coarser-grained patterns of occupancy. Also the distribution of subpopulations got more clumped as the hosts growth rate decreased. A higher level of parasitism caused the subpopulations to be more isolated and the turnover-rates of habitat occupancy to accelerate. This means that the time that elapsed before an occupied cell became extinct, was reduced. However, the patterns produced by host growth and parasitism rates were less pronounced. Under the climate change scenarios the potentially suitable habitat increased and the subpopulations expanded to formerly unoccupied patches. Animated graphics of the evolution of host abundances over time can be found in Appendix A.

3.2 Species Distribution Model

3.2.1 Overall Model Performance

In a first analysis I compared GLM and GAM performances and the spatial prediction accuracies achieved on the training data, and when validating against independent data of

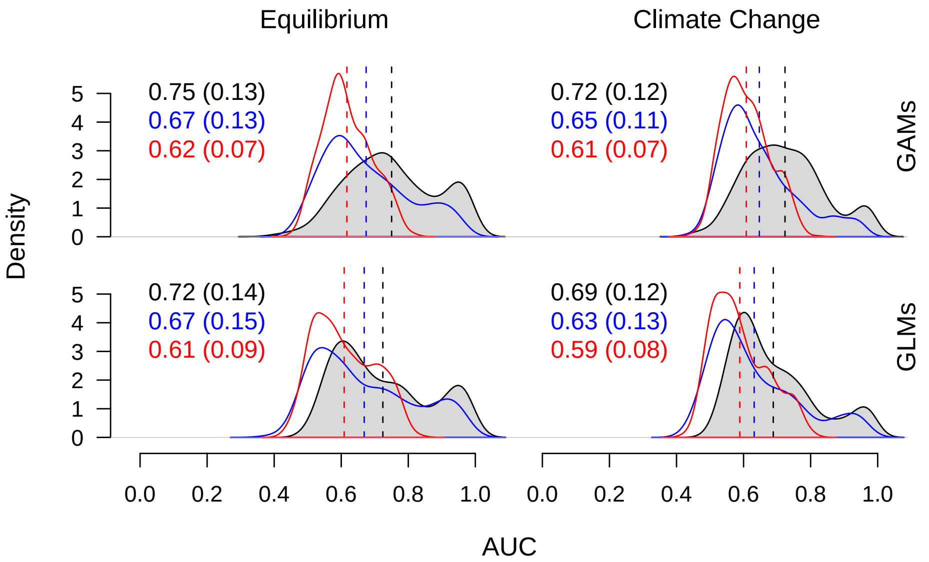


Figure 5: Distribution of spatial prediction accuracies for GLMs (bottom) and GAMs (top) under equilibrium (left) and climate change (right). Model predictions were validated against training datasets (black), independent (blue) and entire datasets (red). Vertical lines and numbers indicate mean AUC (± standard deviation), n=720.

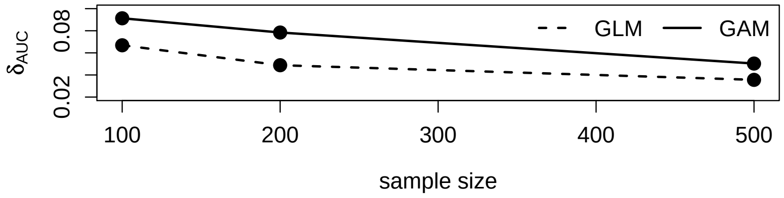


Figure 6: Differences between prediction accuracies of training and independent test data of GLMs and GAMs for different sample sizes.

the same sample size, and the entire data, respectively. Calibration statistics calculated for the independent and the entire data were also compared. I further examined the variation of prediction accuracies over time by estimating GLMs on the entire data for 100 subsequent years in an equilibrium situation and 6 replicate simulations. Spatial predictions were then made for the entire landscape and prediction accuracies were assessed in terms of AUC.

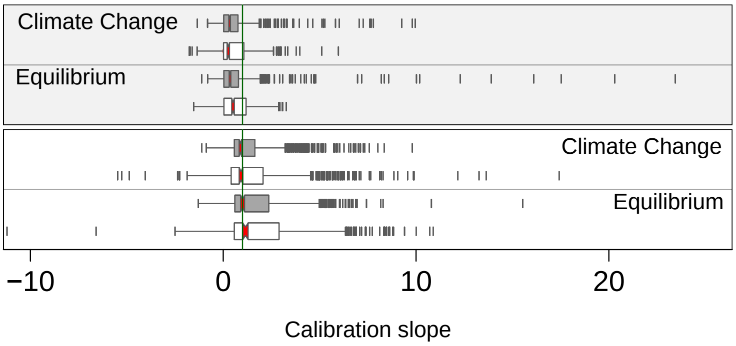


Figure 7: Boxplots of calibration slopes calculated for spatial predictions of GLMs (white boxes) and GAMs (grey boxes) on independent data of same sample size as training data (bottom, white background) and on entire data (top, grey background) under equilibrium and climate change.

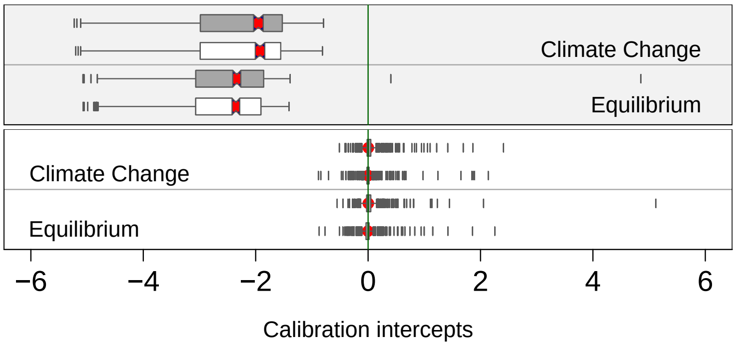


Figure 8: Boxplots of calibration intercepts calculated for spatial predictions of GLMs (white boxes) and GAMs (grey boxes) on independent data of same sample size as training data (bottom, white background) and on entire data (top, grey background) under equilibrium and climate change.

Both GLMs and GAMs had weak explanatory power on average though GAMs explained more of the deviance in the data (GLMs: mean r^2_L=0.05, maximum r^2_L=0.28; GAMs: mean r^2_L=0.09, maximum r^2_L=0.35). The accuracies achieved by spatial predictions are illustrated in Figure 5. Although the apparent model performance on the training data seemed to be good, average discrimination was poor for model predictions externally validated. Differences between apparent and externally validated accuracies were greater for GAMs than for GLMs (Figure 6), and decreased with sample size. The range of AUC values was smaller for validations on true species distribution than for validations on the sampled datasets. On average GAMs performed slightly better than GLMs, and spatial predictions under equilibrium were better than those made in climate change situations.

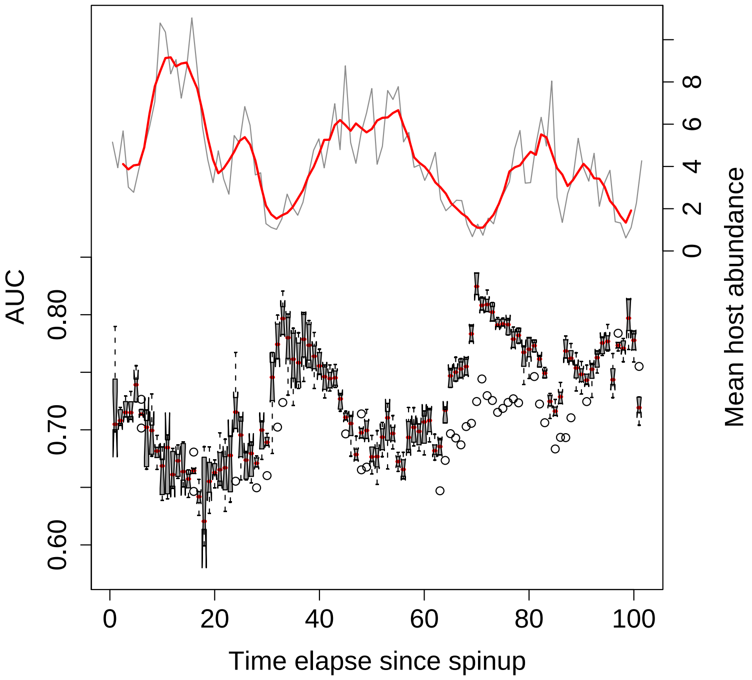


Figure 9: Variation in AUC values over time. GLMs were estimated and evaluated on entire datasets. Shown are the boxplots of AUC values, the mean host abundance (grey) and the 5-year moving average of the mean host abundance (red).

Boxplots of the calibration slopes and intercepts are shown in Figures 7 and 8. Spatial predictions made on the independent data yielded good correspondence on average between estimated occurence probabilities and observed incidences, with approximate unit median slope and zero median intercept for both model algorithms and both equilibrium and climate change situation. When predicting species distribution for the entire landscape the estimated occurence probabilities were negatively biased and consistently overestimated, with median intercepts of approximately -2.3 and -2 for both model algortihms and under equilibrium respective climate change, and with median slopes between 0.25 and 0.5 for both GLMs and GAMs and for both equilibrium and climate change.

Variation of prediction accuracies over time is illustrated in Figure 9. AUC increased when mean host abundance declined and decreased with rising host abundance. Thus the results indicate a negative correlation between prediction accuracy and mean host abundance.

3.2.2 Effects on Spatial Prediction Accuracies

The results of the multivariate LMs are listed in Table 4, univariate LMs in Table B.1. Figure 10 illustrates the proportions of explained deviances achieved by univariate LMs, i. e. relative importance of the predictors, and Figure 11 shows the effects of the different ecological traits on model accuracy. Results of the univariate LMs were similar to those of the multivariate LMs regarding the relative effects of the predictors on model performances. Consequently, only results of the multivariate LMs are reported below.

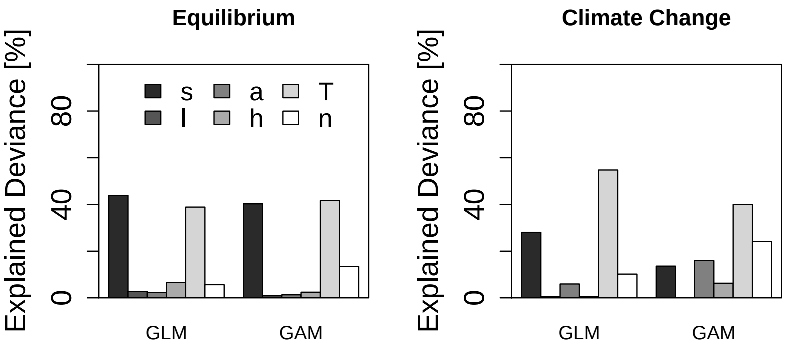


Figure 10: Proportion of explained deviance achieved by univariate LMs estimated for the spatial prediction accuracy. Predictor variables are niche width (s), host growth (l), level of parasitism (a), dispersal distance (h), time elapse since spinup (T), and sample size (n).

Equilibrium All predictors significantly influenced spatial prediction accuracy of GLMs, and only five predictors that of GAMs (Table 4). Spatial prediction accuracy increased

	GLMs		**GAMs**	
	Equilibrium	Climate Change	Equilibrium	Climate Change
intercept	**1.00 ± 0.01** $\ast\ast$	**0.87 ± 0.02** $\ast\ast$	**0.86 ± 0.01** $\ast\ast$	**0.72 ± 0.02** $\ast\ast$
niche width s	**-0.09 ± 4e-3** $\ast\ast$	**-0.05 ± 5e-3** $\ast\ast$	**-0.06 ± 3e-3** $\ast\ast$	**-0.02 ± 4e-3** $\ast\ast$
host growth l	**-7e-4 ± 1e-4** $\ast\ast$	-2e-4 ± 1e-4	-3e-4 ± 1e-4 $\ast\ast$	5e-5 ± 1e-4
level of parasitism a	**-2e-3 ± 4e-4** $\ast\ast$	**-3e-3 ± 5e-4** $\ast\ast$	**1e-3 ± 4e-4** $\ast\ast$	**3e-3 ± 5e-4** $\ast\ast$
dispersal distance h	**0.01 ± 1e-3** $\ast\ast$	2e-3 ± 1e-3	**4e-3 ± 1e-3** $\ast\ast$	**-5e-3 ± 1e-3** $\ast\ast$
time elapse T	**-2e-3** ± 9**e-5** $\ast\ast$	**-2e-3 ± 1e-4** $\ast\ast$	**-2e-3 ± 8e-5** $\ast\ast$	**-1e-3 ± 1e-4** $\ast\ast$
sample size n	**1e-4 ± 1e-5** $\ast\ast$	**1e-4 ± 1e-5** $\ast\ast$	**1e-4 ± 1e-5** $\ast\ast$	**9e-5 ± 1e-5** $\ast\ast$
$s \times l$			5e-4 ± 2e-4 $\ast$	
$s \times a$	**-3e-3 ± 8e-4** $\ast\ast$	**-4e-3 ± 1e-3** $\ast\ast$	1e-3 ± 8e-4	
$s \times h$		**0.01 ± 3e-3** $\ast\ast$	-3e-3 ± 2e-3	**6e-3 ± 2e-3** $\ast$
$s \times T$	**2e-3 ± 2e-4** $\ast\ast$	**6e-4 ± 2e-4** $\ast\ast$	**2e-3 ± 2e-4** $\ast\ast$	**9e-4 ± 2e-4** $\ast\ast$
$s \times n$	-4e-7 ± **1e-7** $\ast\ast$	**-6e-7 ± 2e-7** $\ast\ast$	-3e-7 ± 1e-7 $\ast$	**-5e-7 ± 2e-7** $\ast\ast$
$l \times a$	**-1e-4 ± 2e-5** $\ast\ast$	-5e-5 ± 3e-5	-6e-5 ± 2e-5 $\ast$	
$l \times h$	**-2e-4 ± 6e-5** $\ast\ast$		**-2e-4 ± 6e-5** $\ast\ast$	-1e-4 ± 7e-5
$l \times T$	**-2e-5 ± 5e-6** $\ast\ast$		-8e-6 ± 5e-6	
$l \times n$				
$a \times h$	**6e-4 ± 2e-4** $\ast\ast$		-6e-4 ± 2e-4 $\ast\ast$	**-1e-3 ± 3e-4** $\ast\ast$
$a \times T$	**1e-4 ± 2e-5** $\ast\ast$	4e-5 ± 3e-5	**2e-4 ± 2e-5** $\ast\ast$	**2e-4 ± 2e-5** $\ast\ast$
$a \times n$				
$h \times T$	**2e-4 ± 5e-5** $\ast\ast$	-1e-4 ± 7e-5	9e-5 ± 5e-5	**-2e-4 ± 6e-5** $\ast\ast$
$h \times n$	1e-5 ± 6e-6		1e-5 ± 6e-6 $\ast$	
$T \times n$			-8e-7 ± 5e-7	
r^2	0.71	0.43	0.61	0.36

Table 4: Effects of ecological traits and interspecific interaction on accuracy of spatial prediction of GLMs and GAMs. Given are the intercept and slopes of the multivariate linear regression model(± SE), and the explained deviance r^2. Significance is indicated by * for $p \leq 0.05$ and by ** for $p \leq 0.01$. Parameters in bold remained significant ($p \leq 0.05$) after Holm adjustment for multiple comparison.

with *sample size* and decreased with *niche width*, *dispersal distance*, and *time elapse since spinup*. GLM performance also decreased with *host growth*. *Level of parasitism* had a positive effect on the performance of GAMs and a negative on GLMs. *Niche width and time elapse since spinup* explained most of the deviance for both model algorithms (Figure 10). The most important interactions influencing prediction accuracies of both model algorithms were *time elapse since spinup* with *niche width* and *level of parasitism*, respectively.

Climate Change GLM performance was significantly influenced by four predictors, *niche width*, *level of parasitism*, *time elapse since spinup* and *sample size*. GAM performance was additionally influenced by *dispersal distance* (Table 4). But, in contrast to the equilibrium situation GAM prediction accuracy increased with *dispersal distance*. *Time elapse since spinup* explained most of the deviance in GLM and GAM performance (Figure 10). The

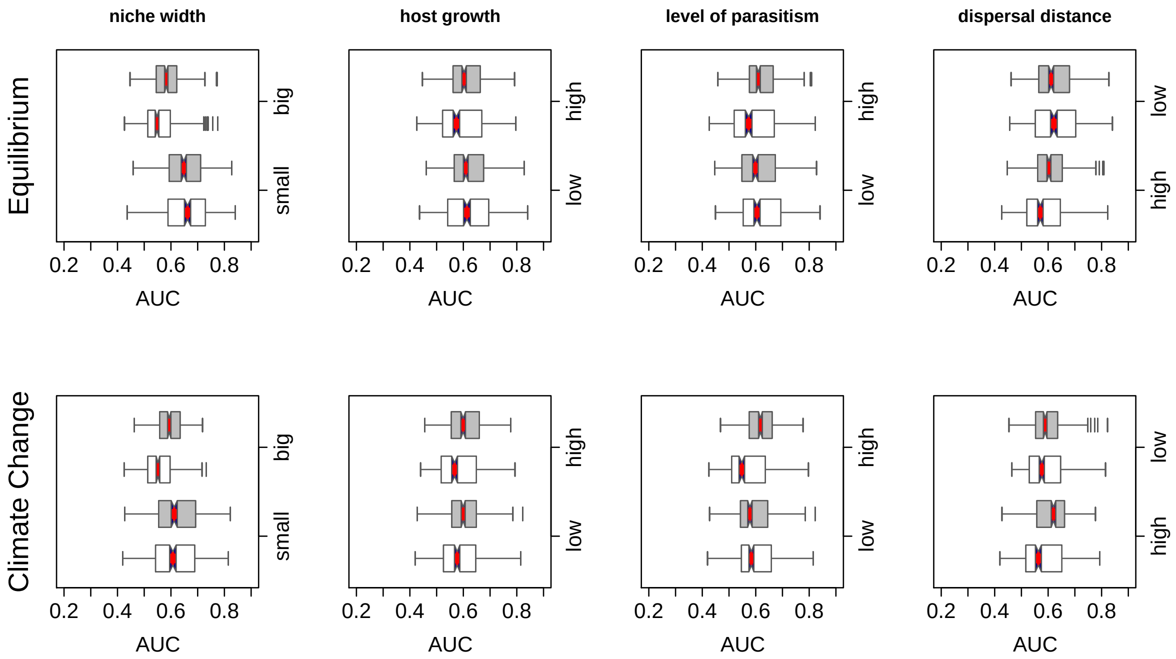


Figure 11: Effects of ecological traits on model accuracy. Shown are the boxplots of AUC values obtained from GLMs (white), and GAMs (grey) validated on entire data , for both equilibrium and climate change (n=360).

most important interactions influencing GLM performance were *niche width* with *level of parasitism* respective *dispersal distance*. In contrary, GAM prediction accuracy was impacted most strongly by the interaction between *level of parasitism* and *dispersal distance* respective *time elapse since spinup*.

Total explained deviances of the LMs were lower for GLMs and GAMs estimated under climate change than for those estimated in the equilibrium situation.

3.2.3 Effects on Spatiotemporal Prediction Accuracies

The results of the multivariate LMs are listed in Tables 5, 6 and 7, univariate LMs are shown in Table B.2. Density plots of the spatiotemporal prediction accuracies achieved by GLMs and GAMs are given in Figures 12 and 13. Figure 14 illustrates the proportion of explained deviances achieved by univariate LMs for the different spatiotemporal predictions. For convenience I introduce the abbreviations $AUC_{T.70}$, $AUC_{T.95}$ and $AUC_{T.120}$ for the spatiotemporal prediction accuracies for times T=70, T=95 and T=120, respectively. Again multivariate and univariate LMs yielded similar results and therefore only the results of the multivariate LMs are reported below.

	GLMs		**GAMs**	
	Equilibrium	Climate Change	Equilibrium	Climate Change
Projection for T=70				
intercept	**1.04 ± 0.02** $**$	**0.94 ± 0.02** $**$	**0.94 ± 0.02** $**$	**0.86 ± 0.02** $**$
niche width s	**-0.12 ± 5e-3** $**$	**-0.02 ± 5e-3** $**$	**-0.09 ± 4e-3** $**$	2e-3 ± 5e-3
host growth l	**-5e-4 ± 1e-4** $**$	-2e-4 ± 1e-4	-2e-4 ± 1e-4	9e-5 ± 1e-4
level of parasitism a	**-3e-3 ± 5e-4** $**$	**-2e-3 ± 6e-4** $**$	**-1e-3 ± 5e-4** $**$	-3e-4 ± 5e-4
dispersal distance h	**0.01 ± 1e-3** $**$	**0.02 ± 8e-3** $**$	**0.01 ± 1e-3** $**$	**0.01 ± 1e-3** $**$
time elapse T	**-2e-3 ± 1e-4** $**$	**-4e-3 ± 1e-4** $**$	**-2e-3 ± 1e-4** $**$	**-4e-3 ± 1e-4** $**$
sample size n	**1e-4 ± 1e-5** $**$	3e-5 ± 2e-5	**1e-4 ± 1e-5** $**$	3e-5 ± 7e-5 $**$
$s \times l$	6e-4 ± 3e-4 $*$	7e-4 ± 3e-4 $*$	**7e-4 ± 3e-4** $**$	**8e-4 ± 1e-5** $*$
$s \times a$	**-3e-3 ± 1e-3** $**$	-3e-3 ± 1e-3 $*$		
$s \times h$	**0.01 ± 3e-3** $**$	5e-3 ± 3e-3	7e-3 ± 3e-3 $**$	
$s \times T$	**7e-4 ± 2e-4** $**$	**4e-3 ± 3e-4** $**$	4e-4 ± 2e-4 $*$	**4e-3 ± 2e-4** $**$
$s \times n$	-4e-7 ± 2e-7 $*$	-3e-6 ± 2e-7		-3e-7 ± 2e-7
$l \times a$	-7e-5 ± 3e-5 $*$	**-9e-5 ± 3e-5** $**$		-5e-5 ± 3e-5
$l \times h$	-1e-4 ± 7e-5	-1e-4 ± 9e-5		-2e-4 ± 8e-5 $*$
$l \times T$	-1e-5 ± 6e-6			
$l \times n$				
$a \times h$	**-1e-3 ± 3e-4** $**$	**-2e-3 ± 3e-4** $**$	**-2e-3 ± 3e-4** $**$	**-2e-3 ± 3e-4** $**$
$a \times T$	**1e-4 ± 2e-5** $**$	**1e-4 ± 3e-5** $**$	**1e-4 ± 2e-5** $**$	**1e-4 ± 3e-5** $**$
$a \times n$	-4e-6 ± 3e-6			
$h \times T$	**4e-4 ± 6e-5** $**$	**5e-4 ± 7e-5** $**$	**4e-4 ± 6e-5** $**$	**6e-4 ± 7e-5** $**$
$h \times n$		2e-5 ± 9e-6 $*$	1e-5 ± 7e-6	2e-5 ± 8e-6 $*$
$T \times n$		**-5e-6 ± 8e-7** $**$		**-5e-6 ± 7e-7** $**$
r^2	0.67	0.68	0.60	0.70

Table 5: Effects of ecological traits and interspecific interaction on accuracy of spatiotemporal prediction of GLMs and GAMs for T=70. Given are the intercept and slopes of the multivariate linear regression model(± SE), and the explained deviance r^2. Significance is indicated by * for $p \leq 0.05$ and by ** for $p \leq 0.01$. Parameters in bold remained significant ($p \leq 0.05$) after Holm adjustment for multiple comparison.

Equilibrium Figure 12 shows that average accuracies were poor and very similar for all projection and model building times, with median values between 0.54 and 0.68.

$AUC_{T.70}$ for GLMs and GAMs was significantly influenced by all and five predictors, respectively (Table 5). Both GAM and GLM performance decreased with *niche width*, *dispersal distance*, *level of parasitism*, and *time elapse since spinup* and increased with *sample size*. GLM performance was additionally negatively effected by *host growth*. *Niche width* and *time elapse since spinup* explained most of the deviance in prediction accuracies for both model algorithms (Figure 14). The most important interactions for accuracies of both

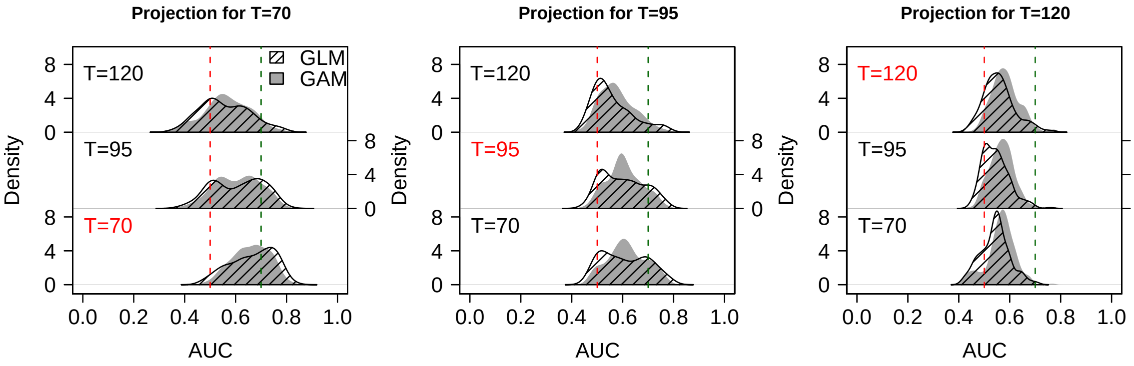


Figure 12: Distribution of AUC values for spatiotemporal predictions of GLMs and GAMs under equilibrium. For every density plot time T is given at which the models were estimated. Red T indicate spatial predictions. The dashed lines indicate AUC values the null model would achieve (red) and above which models are considered to have predictive ability (green), respectively.

models were *dispersal distance* with *level of parasitism* respective *time elapse since spinup*, and additionally *level of parasitism* with *time elapse since spinup* in the case of GAMs.

All predictors exhibit a significant effect on $AUC_{T.95}$ of both GLMs and GAMs (Table 6). Model performances decreased with *niche width*,*host growth*, *dispersal distance*, and *time elapse since spinup*, and increased with *sample size*. GLM prediction accuracy was negatively influenced by *level of parasitism*, GAM performance positively. *Niche width* explained the largest part of the deviance in the prediction accuracies of both GLMs and GAMs (Figure 14). The interaction between *niche witdh* and *time elapse since spinup* was the most important one impacting the performances of both model algorithms, followed by the interactions *level of parasitsm* with *niche width* respective *dispersal distance* in the case of GAMs.

GAM $AUC_{T.120}$ was significantly influenced by all predictors, GLM performance by all predictors except level of parasitism (Table 7). Different to the earlier time steps, *time elapse since spinup* exerted a positive impact on prediction accuracies. *Niche width* and *dispersal distance* explained most of the deviance in GLM performance, *level of parasitism* in GAM (Figure 14). The most important interactions for the prediction accuracy of GLMs were *niche width* with *time elapse since spinup* and *host growth* with *dispersal distance*,

	GLMs		**GAMs**	
	Equilibrium	Climate Change	Equilibrium	Climate Change
Projection for T=95				
intercept	**0.90 ± 0.01** $**$	**0.69 ± 0.02** $**$	**0.79 ± 0.01** $**$	**0.55 ± 0.02** $**$
niche width s	**-0.12 ± 4e-3** $**$	**-0.05 ± 5e-3** $**$	**-0.08 ± 3e-3** $**$	-7e-3 ± 4e-3
host growth l	**-7e-4 ± 1e-4** $**$	**-6e-4 ± 1e-4** $**$	**-3e-4 ± 1e-4** $**$	-3e-4 ± 1e-4 $*$
level of parasitism a	**-1e-3 ± 4e-4** $**$	-9e-4 ± 5e-4	**3e-3 ± 4e-4** $**$	**5e-3 ± 5e-4** $**$
dispersal distance h	**0.01 ± 1e-3** $**$	**0.01 ± 1e-3** $**$	**8e-3 ± 1e-3** $**$	**5e-3 ± 1e-3** $**$
time elapse T	**-1e-3 ± 9e-5** $**$	**-3e-4 ± 1e-4** $**$	**-8e-4 ± 8e-5** $**$	-5e-5 ± 1e-4
sample size n	**9e-5 ± 1e-5** $**$	**4-e5 ± 1e-5** $**$	**1e-4 ± 1e-5** $**$	**5e-5 ± 1e-5** $**$
$s \times l$	5e-4 ± 2e-4 $*$	4e-4 ± 3e-4	**8e-4 ± 2e-4** $**$	6e-4 ± 3e-4 $*$
$s \times a$			**5e-3 ± 8e-4** $**$	**6e-3 ± 1e-3** $**$
$s \times h$			**-7e-3 ± 2e-3** $**$	-5e-3 ± 3e-3 $*$
$s \times T$	**1e-3 ± 2e-4** $**$	**1e-3 ± 2e-4** $**$	**1e-3 ± 2e-4** $**$	**8e-4 ± 2e-4** $**$
$s \times n$	-2e-7 ± 1e-7		-2e-7 ± 1e-7	
$l \times a$	-6e-5 ± 2e-5 $*$	-5e-5 ± 3e-5		
$l \times h$				
$l \times T$		**2e-5 ± 7e-6** $**$		**2e-5 ± 6e-6** $**$
$l \times n$	-9e-7 ± 6e-7			
$a \times h$			**-1e-3 ± 2e-4** $**$	**-1e-3 ± 3e-4** $**$
$a \times T$		-4e-5 ± 3e-5	3e-5 ± 2e-5	
$a \times n$				
$h \times T$		**-4e-4 ± 7e-5** $**$		**-4e-4 ± 6e-5** $**$
$h \times n$			1e-5 ± 6e-6 $*$	1e-5 ± 7e-6
$T \times n$				
r^2	0.67	0.32	0.60	0.28

Table 6: Effects of ecological traits and interspecific interaction on accuracy of spatiotemporal prediction of GLMs and GAMs for T=95. Given are the intercept and slopes of the multivariate linear regression model(± SE), and the explained deviance r^2. Significance is indicated by * for $p \leq 0.05$ and by ** for $p \leq 0.01$. Parameters in bold remained significant ($p \leq 0.05$) after Holm adjustment for multiple comparison.

for GAMs *level of parasitism* with *niche width* respective *dispersal distance.*

Climate Change Figure 13 shows that prediction accuracies were generally better when the gap between model building and projection time was small. The greatest differences between the AUC values of the different model building times were found for $AUC_{T.70}$. Median prediction accuracies were poor ranging from 0.45 to 0.67.

GLM $AUC_{T.70}$ was significantly influenced by four predictors, *niche width*, *dispersal distance*, *level of parasitism*, and *time elapse since spinup*. GAM $AUC_{T.70}$ was significantly effected only by *dispersal distance* and *time elapse since spinup* (Table 5). *Time elapse since spinup* explained most of the deviance in model performances (Figure 14). The most important interactions were *time elapse since spinup* with *niche width* respective *dispersal distance.*

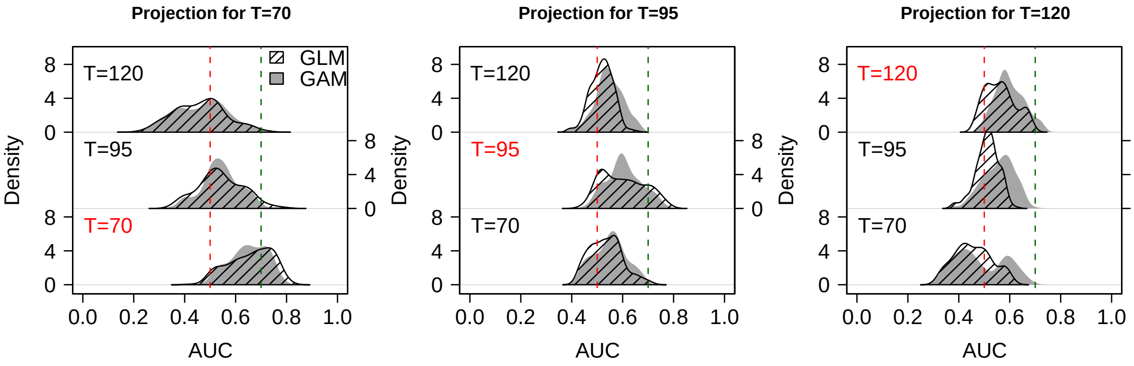


Figure 13: Distribution of AUC values for spatiotemporal predictions of GLMs and GAMs under climate change. For every density plot time T is given at which the models were estimated. Red T indicate spatial predictions. The dashed lines indicate AUC values the null model would achieve (red) and above which models are considered to have predictive ability (green), respectively.

	GLMs		**GAMs**	
	Equilibrium	Climate Change	Equilibrium	Climate Change
Projection for T=120				
intercept	**0.60** ± **0.01** ∗∗	**0.26** ± **0.01** ∗∗	**0.59** ± **0.02** ∗∗	**0.17** ± **0.01** ∗∗
niche width s	**-0.04** ± **3e-3** ∗∗	**0.02** ± **3e-3** ∗∗	**-0.05** ± **9e-3** ∗∗	**0.06** ± **3e-3** ∗∗
host growth l	**-7e-4** ± **1e-4** ∗∗	2e-5 ± 9e-5	**-2e-3** ± **3e-4** ∗∗	2e-4 ± 9e-5
level of parasitism a	-2e-4 ± 4e-4	**-2e-3** ± **3e-4** ∗∗	**5e-3** ± **3e-4** ∗∗	**7e-3** ± **3e-4** ∗∗
dispersal distance h	**0.01** ± **1e-3** ∗∗	2e-3 ± 9e-4 ∗	**6e-3** ± **8e-4** ∗∗	**-3e-3** ± **9e-4** ∗∗
time elapse T	**3e-4** ± **8e-5** ∗∗	**2e-3** ± **7e-5** ∗∗	**4e-4** ± **7e-5** ∗∗	**2e-3** ± **8e-5** ∗∗
sample size n	**5e-5** ± **1e-5** ∗∗	2e-5 ± 9e-6 ∗	**6e-5** ± **8e-6** ∗	**2e-5** ± **9e-6** ∗
$s \times l$		-4e-4 ± 2e-4 ∗	**6e-4** ± **2e-4** ∗∗	
$s \times a$		**-3e-3** ± **7e-4** ∗∗	**6e-3** ± **6e-4** ∗∗	**4e-3** ± **7e-4** ∗∗
$s \times h$		**0.01** ± **2e-3** ∗∗	-4e-3 ± 2e-3 ∗∗	**8e-3** ± **2e-3** ∗∗
$s \times T$	**6e-4** ± **2e-4** ∗∗	**-3e-3** ± **1e-4** ∗∗	**5e-4** ± **1e-4** ∗∗	**-3e-3** ± **2e-4** ∗∗
$s \times n$			-2e-7 ± 1e-7	
$l \times a$	-5e-5 ± 2e-5 ∗	-3e-5 ± 2e-5	4e-5 ± 2e-5 ∗	
$l \times h$	**-2e-4** ± **6e-5** ∗∗		**-2e-4** ± **5e-5** ∗∗	
$l \times T$	-1e-5 ± 5e-6 ∗	7e-6 ± 4e-6		7e-6 ± 4e-6
$l \times n$	-8e-7 ± 6e-7		-1e-6 ± 5e-7 ∗	
$a \times h$		-4e-4 ± 2e-4 ∗	**-1e-3** ± **2e-4** ∗∗	**-2e-3** ± **2e-4** ∗∗
$a \times T$		**-7e-5** ± **2e-5** ∗∗		
$a \times n$				
$h \times T$		**-5e-4** ± **4e-5** ∗∗	-1e-4 ± 4e-5 ∗∗	**-5e-4** ± **4e-5** ∗∗
$h \times n$	1e-5 ± 6e-6 ∗		1e-5 ± 5e-6 ∗∗	
$T \times n$	**1e-6** ± **5e-7** ∗∗	**3e-6** ± **4e-7** ∗∗	9e-7 ± 4e-7 ∗	**3e-6** ± **4e-7** ∗∗
r^2	0.38	0.71	0.49	0.78

Table 7: Effects of ecological traits and interspecific interaction on accuracy of spatiotemporal prediction of GLMs and GAMs for T=120. Given are the intercept and slopes of the parsimonious multivariate linear regression model(± SE), and the explained deviance r^2. Significance is indicated by * for $p \leq 0.05$ and by ** for $p \leq 0.01$. Parameters in bold remained significant ($p \leq 0.05$) after Holm adjustment for multiple comparison.

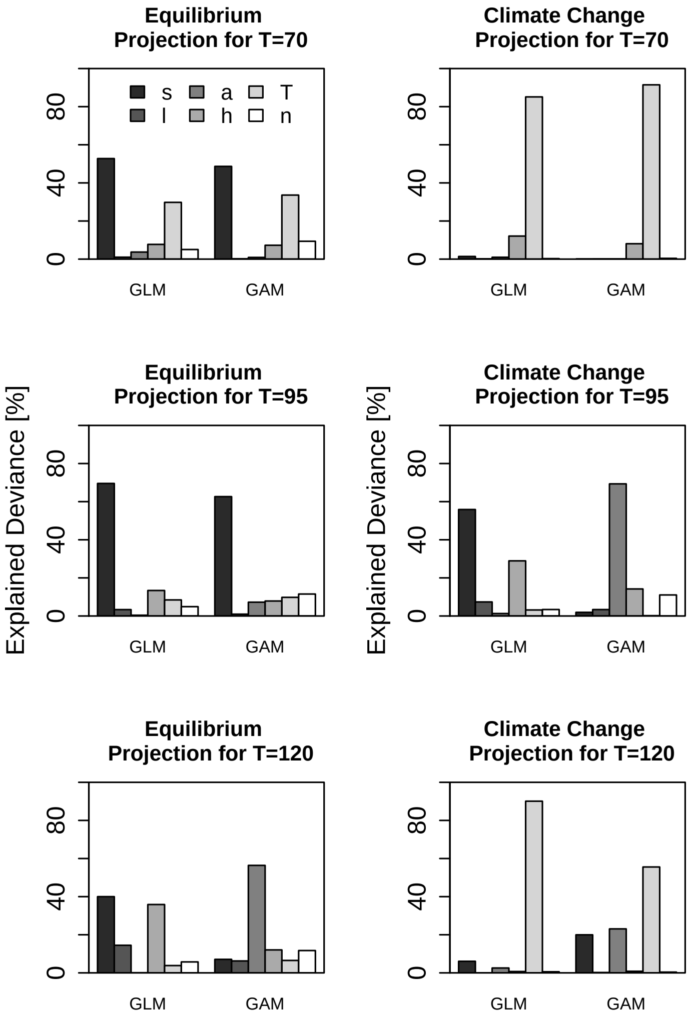


Figure 14: Proportion of explained deviance achieved by univariate LMs estimated for the spatiotemporal prediction accuracy. Predictor variables are niche width (s), host growth (l), level of parasitism (a), dispersal distance (h), time elapse since spinup (T), and sample size (n).

GLM $AUC_{T.95}$ was significantly influenced by five predictors, *niche width*, *host growth*, *dispersal distance*, *time elapse since spinup*, and *sample size*, while the accuracy of GAMs was effected by *level of parasitism*, *dispersal distance* and *sample size* (Table 6). *Niche width* explained most of the deviance in GLM performance, *level of parasitism* in GAM (Figure 14). The most important interaction influencing both GLM and GAM performance was *dispersal distance* with *time elapse since spinup*, followed by *niche width* with *time elapse* *since spinup* for GLMs and by *niche width* with *level of parasitism* for GAMs.

$AUC_{T.120}$ for GLMs was significantly impacted by three predictors, *niche width*, *level of parasitism* and *time elapse since spinup.* GAM performance was additionally influenced by *dispersal distance* and *sample size* (Table 7). Analogous to the equilibrium situation model performance was positively influenced by *time elapse since spinup*, which concurrently explained most of the deviance in both GLM and GAM performance (Figure 14). Different to all other cases *niche width* positively influenced $AUC_{T.120}$ of both GLMs and GAMs. The most important interactions were *time elapse since spinup* with *niche width* respective *dispersal distance.*

Peculiarly, explained deviances of the LMs for $AUC_{T.95}$ under climate change and for $AUC_{T.120}$ under equilibrium were much lower than for the remaining LMs (Tables 6 and 7).

4 Discussion

4.1 Dynamic Model

Hargrove *et al.* (2002) found that experts were not able to distinguish between maps of real landscapes and synthetic maps produced by the midpoint displacement algorithm. However, my study shows that this fractal algorithm is unsuitable for creating structured landscape features such as flow paths. It produces a geomorphologically young landscape with a chaotic river network, causing water availability to be distributed very heterogeneously at a small spatial scale. In respect to the dynamic population model this means that habitat quality changes remarkably over very short distances, which may reduce realism in the virtual system. Further development of the fractal algorithm is necessary. For instance it could be supplemented by an ageing component.

The dynamic model 'spins up' for 70 years to ensure a long-term equilibrium between the host population and its environment. Though the climate is in equilibrium, interannual variations and stochasticity in the system cause great variations in mean host abundances (Figure 15). However, time series analysis reveiled neither a trend nor seasonality in host abundances, implying that the host population really is in long-term equilibrium with its environment.

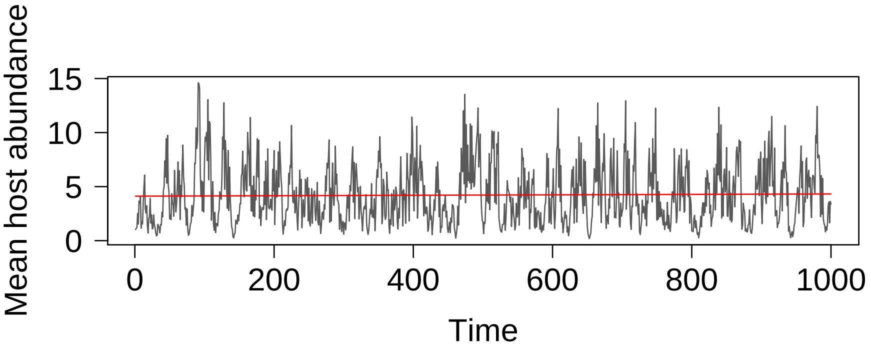


Figure 15: Long-term variation of mean host abundance (grey). The linear regression line fitted to time and mean host abundance (red) shows no trend.

Under the climate change scenarios the fraction of suitable habitat in the virtual landscape increased because sites at higher elevations became suitable due to rising temperatures, and formerly occupied habitat remained suitable within the observed time-frame. This is of course only one possibility of changing habitats under climate change. Habitats may also shift along environmental gradients or decrease.

4.2 Species Distribution Model

4.2.1 Overall Model Performance

Apparent model performance was overoptimistic, a fact already pointed out by other authors (Fielding & Bell, 1997; McPherson & Jetz, 2007; McPherson *et al.*, 2004; Reineking & Schröder, 2003). Moreover, the difference between apparent model performance and external validations is an indicator for the degree of overfitting (Stockwell & Peterson, 2002). I expected GAMs to be more overfitted than GLMs, which is supported by my findings. But surprisingly, GAM prediction accuracy and thus transferability was still better than for GLMs.

Maximum explained deviances of GLMs and GAMs were low on average indicating that average detectability of the species-habitat relationship was also low despite ideal conditions. Great proportions of residual deviances persisted due to environmental and demographic stochasticity. This is consistent with findings of Tyre *et al.* (2001).

Average model accuracies were low compared to real field studies. For instance McPherson & Jetz (2007) reported excellent average model accuracies for distribution models of 1329 bird species in southern and eastern Africa constructed with logistic regression and externally validated on withheld test data (mean AUC=0.89). Luoto *et al.* (2006) explored the distribution of 98 butterfly species on a national scale in Finland with GAMs and found acceptable average model accuracies (mean AUC=0.79) derived from cross-validation. Rudner *et al.* (2007) modelled the distribution of 52 plant species on a fine scale with GLMs and also reported excellent average accuracies based on internal validation with bootstrapping. All these studies yielded better model accuracies than found in this study regardless of spatial scale. Hence for a certain reason the performance of species distribution models is inferior in my virtual system compared to real systems. I ascribe these poor model results to the high stochasticity in the model system.

Differences between calibration statistics of external validation on sampled and entire datasets, respectively, were attributable to species prevalence. While I controlled for prevalence in the sampled datasets, the natural prevalence of the entire data was much lower leading to consistent overestimation and bias. The bias was smaller under climate change because prevalence increased with increasing species range.

The negative correlation between prediction accuracy and mean host abundance indicated that spatial population and source-sink dynamics exerted considerable influence on the performance of species distribution models. In the virtual dynamic system host abundance is proportionate to habitat quality. Hence, when mean host abundance decreases, the majority of hosts will be located in higher quality habitats. Correspondingly spatial population and source-sink dynamics become more pronounced when host abundance increases, and lower and higher quality habitat are used simultaneously. My analyses shows that in the latter case the static model was less successful in depicting the species-habitat relationship, and prediction accuracy decreased.

4.2.2 Effects on Spatial Prediction Accuracies

Equilibrium Significance of the predictor *time elapse since spinup* was attributable to spatial population dynamics. Mean host abundance increased with increasing *time elapse since spinup* and caused a decline in prediction accuracies. My results regarding *niche width* support the findings of McPherson & Jetz (2007), Segurado & Araujo (2004), and Stockwell & Peterson (2002) who reported decreasing spatial prediction accuracies with increasing range sizes. Model performance decreased with *host growth* and *dispersal distance* because both enhanced source-sink dynamics.

The positive effect of *level of parasitism* on GAMs and negative on GLMs can be related to the frequency that parasitoids were selected as predictors in the final GLMs and GAMs.

Incidence of parasitoids was kept as predictor in 69% (4992/7200) of the final GAMs and in 24% (1758/7200) of the final GLMs. When *level of parasitism* is low, parasitoids and hosts can coexist in a cell, while high *levels of parasitism* lead to complete exploitation of the hosts. Hence, in the case of low predation rates models including and excluding the incidence of parasitoids as predictor, respectively, should perform equally well, as my results support (Figure 11). However, with increasing predation and local populations going extinct, the inclusion of this predictor was crucial to distinguish between suitable but exploited and unsuitable habitat. This difference between GLM and GAM performance seems to be a methodological artefact in the model selection procedure. The first step in GLM model selection, the exclusion of predictors whose univariate model yields $p > 0.25$, caused this discrepancy. Thus in this instance, it was not the smoothing technique, but the different selection procedures making GAMs superior to GLMs.

Climate Change In contrast to the exerted influence under equilibrium GAM performance increased with *dispersal distance*, because higher *dispersal distances* promote a faster realisation of a new equilibrium situation. Also, average spatial prediction accuracy of both GLMs and GAMs was lower under climate change suggesting reduced detectability of the species-habitat relationship due to non-equilibrium dynamics. The lower explained deviances in the LMs under climate change also imply that the variance in prediction accuracies could not be explained conveniently by the ecological characteristics, *time elapse* and *sample size* alone, but had been influenced by other dynamics uncoupled from the predictors, i. e. demographic and environmental stochasticity.

4.2.3 Effects on Spatiotemporal Prediction Accuracies

Equilibrium Under equilibrium conditions spatiotemporal prediction accuracies were similar to the spatial ones. Thus by and large both GLM and GAM models were transferable in time, though the results also imply that predictions made for times nearer the model building time were generally better. This effect is attributable to spatial population and source-sink dynamics. Variations in prediction accuracies were governed by the same ecological processes as the spatial predictions.

Climate Change Under climate change the time span between model building and prediction was crucial for prediction accuracy, especially when predicting (past) species distribution for an equilibrium situation although the model was build under climate change. Related to that temporal effect were the interactions between *niche width* respective *dispersal distance* and *time elapse since spinup*, which were important in all spatiotemporal predictions under climate change. The direction of influence of *niche width* and *dispersal distance* changed with time, i. e. both had a negative effect on prediction accuracy when projection time was near model building time. These effects were reversed when the time span between model building and prediction was increased. Predictions were then better for wide-ranging and far-dispersing species, because they exhibited smaller range shifts. Explained deviances for $AUC_{T.95}$ under climate change and for $AUC_{T.120}$ under equilibrium were comparably low, i. e. great proportions of variance in the prediction accuracies were not attributable to any predictors but must have been caused by stochasticity in the dynamic system.

5 Conclusion

This virtual species approach yielded new insights on how departures from the assumptions underlying species distribution modelling, and population dynamics may affect model outcome. Throughout I tried to control for methodological drawbacks and therefore infer that my findings describe real effects on the accuracy of species distribution models.

I have shown that departures from equilibrium can impair spatiotemporal predictions, and how ecological characteristics and processes govern these effects. The equilibrium state of the species and the readiness to gain equilibrium are crucial for the validity and transferability of species distribution models. Furthermore, prediction accuracy varies with the intensity of spatial population and source-sink dynamics. And above all, great proportions of residual deviances may persist, caused by system immanent and population intrinsic stochasticity.

I see this study as a provisional guide as to when the application of species distribution models and different model algorithms are more or less reasonable. In equilibrium situations, spatial and spatiotemporal prediction accuracies increase primarily with decreasing niche width, and dispersal distance. Less pronounced spatial population and source-sink dynamics, e. g. through decreasing intrinsic growth rates, are also beneficial for prediction accuracies, as are increasing sample sizes and the inclusion of biotic interactions as predictor. Under climate change spatial predictions are more dependable for far-dispersing species. Spatiotemporal predictions under climate change are reliable only when made for short time intervals.

The great proportions of unexplained deviances and the poor prediction accuracies achieved, imply that stochasticity in the system should be accounted for. The results also strongly recommend to link or supplement the species distribution models with more mechanistic, process-based models, especially when dealing with non-equilibrium situations. Important points to consider are the strength of spatial population and source-sink dynamics and their fluctuations over time, dispersal, as well as demographic and environmental stochasticity.

It also became apparent once more that expert knowledge is inevitable because automatised modelling procedures may fail to recognise important predictors.

References

Aickin, Mikel, & Gensler, Helen. 1996. Adjusting for multiple testing when reporting research results: The Bonferroni vs. Holm Methods. *American Journal of Public Health*, **86**, 726–728.

Akaike, H. 1974. A New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control*, **19**, 716–723.

Allee, W.C. 1931. *Animal aggregations. A study in general sociology.* University of Chicago Press, Chicago.

Araujo, Miguel B., & Pearson, Richard G. 2005. Equilibrium of species distribution with climate. *Ecography*, **28**, 693–695.

Austin, M. P. 2002. Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecological Modelling*, **157**, 101–118.

Barry, Simon, & Elith, Jane. 2006. Error and uncertainty in habitat models. *Journal of Applied Ecology*, **43**, 413–423.

Bartell, S.M., Brenkert, A.L., O'Neill, R.V., & Gardner, R.H. 1988. Temporal variation in regulation of production in a pelagic food web model. *Pages 101–118 of:* Carpenter, S. R. (ed), *Complex Interactions in Lake Communities.* Springer-Verlag, NY.

Bellasio, R., Maffeis, G., Scire, J. S., Longoni, M. G., Bianconi, R., & Quaranta, N. 2005. Algorithms to account for topographic shading effects and surface temperature dependence on terrain elevation in diagnostic meteorological models. *Boundary-Layer Meteorology*, **114**, 595–614.

Beven, K.J., & Kirkby, M. J. 1979. A physically-based variable contributing area model of basin hydrology. *Hydrological Science Bulletin*, **24**, 43–69. Hydrol. Sci. Bull.

Binzenhöfer, Birgit, Schröder, Boris, Strauss, Barbara, Biedermann, Robert, & Settele, Josef. 2005. Habitat models and habitat connectivity analysis for butterflies and burnet moths - The example of *Zygaena carniolica* and *Coenonympha arcania*. *Biological Conservation*, **126**, 247–250.

Bonsall, Michael B., & Hassel, Michael P. 2000. The effects of metapopulation structure on indirect interactions in host-parasitoid assemblages. *Proc. R. Soc. Lond.*, **267**, 2207–2212.

Brotons, Lluis, Thuiller, Wilfried, Araujo, Miguel B., & Hirzel, Alexandre H. 2004. Presence-absence versus presence-only modelling methods for predicting bird habitat suitability. *Ecography*, **27**, 437–448.

Burgman, M. A., Lindenmayer, D. B., & Elith, J. 2005. Managing landscapes for conservation under uncertainty. *Ecology*, **86**, 2007–2017.

Burrill, Donald F. Modeling and interpreting interactions in multiple regression. Available at http://www.minitab.com/resources/articles/.

Comins, H. N., Hassell, M. P., & May, R. M. 1992. The spatial dynamics of host-parasitoid systems. *Journal of Animal Ecology*, **61**(3), 735–748.

Cook, L. M. 1965. Oscillation in the simple logistic growth model. *Nature*, **207**, 316.

Dettki, Holger, Löfstrand, Ronny, & Edenius, Lars. 2003. Modeling habitat suitability for Moose in coastal northern Sweden: Empirical *vs* Process-oriented approaches. *Ambio*, **32**, 549–556.

Dubayah, R., & Loechel, S. 1997. Modeling topographic solar radiation using GOES data. *American meteorological Society*, **36**, 141–154.

Elith, Jane. 2000. *Quantitative Methods in Conservation Biology.* Springer, New York. Chap. Quantitative methods for modeling species habitat: comparitive performance and an application to Australian plants, pages 39–58.

Elith, Jane, Graham, Catherine H., Anderson, Robert P., Dudik, Miroslav, Ferrier, Simon, Guisan, Antoine, Hijmans, Robert J., Huettman, Falk, Leathwick, John R., Lehmann, Anthony, Li, Jin, Lohmann, Lucia G., Loiselle, Bette A., Manion, Glenn, Moritz, Craig, Nakamura, Migual, Nakazawa, Yoshinori, Overton, Jacob McC., Peterson, A. Townsend, Phillips, Steven J., Richardson, Karen, Scachetto-Pereira, Ricardo, Schapire, Robert E., Soberon, Jorge, Williams, Stephen, Wisz, Mary S., & Zimmermann, Niklaus E. 2006. Novel methods improve prediction of species' distribution from occurence data. *Ecography*, **29**, 129–151.

Ferrier, S., & Watson, G. 1997. *An evaluation of the effectiveness of environmental surrogates and modelling techniques in predicting the distribution of biological diversity.* Tech. rept. Department of Environment Sports and Territories, Commonwealth of Australia and NSW National Parks and Wildlife Service, Canberra.

Fielding, Alan H., & Bell, John F. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, **24**, 38–49.

Fielding, Alan H., & Haworth, Paul F. 1995. Testing the generality of bird-habitat models. *Cons Biol*, **9**, 1466–1481.

Guisan, A., & Theurillat, J.-P. 2000. Equilibrium modeling of alpine plant distribution and climate change: how far can we go? *Phytocoenologia*, **30**, 353–384.

Guisan, Antoine, & Thuiller, Wilfried. 2005. Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, **8**, 993–1009.

Guisan, Antoine, & Zimmermann, Niklaus E. 2000. Predictive habitat distribution models in ecology. *Ecological Modelling*, **135**, 147–186.

Guisan, Antoine, Edwards, Thomas C., Jr., & Hastie, Trevor. 2002. Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecological Modelling*, **157**, 89–100.

Guisan, Antoine, Lehmann, Anthony, Ferrier, Simon, Austin, Mike, Overton, Jacob, MC. C., Aspinall, Richard, & Hastie, Trevor. 2006. Making better biogeographical predictions of species' distributions. *Journal of Apllied Ecology*, **43**, 386–392.

Hanley, J. A., & McNeil, B. J. 1982. The meaning and use of the area under a ROC curve. *Radiology*, **143**, 29–36.

Hargrove, W. W., Hoffmann, F. M., & Schwartz, P. M:. 2002. A fractal landscape realizer for generating synthetic maps. *Conservation Ecology*, **6**.

Harrell, Frank E. Jr. 2006. *Regression Modeling Strategies.* Secaucus, NJ, USA: Springer-Verlag New York, Inc.

Harrell Jr., Frank E. 2005. *The Design package.* Available at http://www.r-project.org/.

Harrell Jr., Frank E. 2007. *The Hmisc package.* Available at http://www.r-project.org/.

Hassell, M. P., Comins, H. N., & May, R. M. 1991. Spatial structure and chaos in insect population dynamics. *Nature*, **353**(6341), 255–258.

Hassell, Michael P. 2000a. Host-parasitoid population dynamics. *Journal of Animal Ecology*, **69**, 543–566.

Hassell, Michael P. 2000b. *The spatial and temporal dynamics of host-parasitoid interactions.* Oxford University Press Inc.

Hastie, T. J., & Tibshirani, R. J. 1990. *Generalized Additive Models.* Monographs on Statistics and Applied Probability 43. Chapman & Hall.

Hastie, Trevor. 2006. *The gam package.* Available at http://www.r-project.org/.

Hein, Silke, Binzenhöfer, Birgit, Poethke, Hans-Joachim, Biedermann, Robert, Settele, Josef, & Schröder, Boris. 2007. The generality of habitat suitabilty models: A practical test with two insect groups. *Basic and Applied Ecology*, **8**, 310–320.

Hijmans, Robert J., & Graham, Catherine H. 2006. The ability of climate envelope models to predict the effect of climate change on species distributions. *Global Change Biology*, **12**, 2272–2281.

Hirzel, Alexandre, & Guisan, Antoine. 2002. Which is the optimal sampling strategy for habitat suitability modelling. *Ecological Modelling*, **157**, 331–341.

Hochberg, Michael E., & Holt, Robert D. 1995. Refuge evolution and the population dynamics of coupled host-parasitoid associations. *Evolutionary Ecology*, **9**, 633–661.

Hogeweg, P. 1988. Cellular automata as a paradigm for ecological modeling. *Applied Mathematics and Computation*, **27**, 81–100.

Hosmer, David W., & Lemeshow, Stanley. 2000. *Applied Logistic Regression.* John Wiley & Sons, Inc.

Houghton, J.T., Ding, Y., Griggs, D. J:, Noguer, M., van der Linden, P.J., Dai, X., Maskell, K., & Johnson, C. A. (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 University Press, Cambridge, United Kingdom and New York, NY, USA.

Hutchinson, G. Evelyn. 1957. Concluding remarks, Cold Spring Harbor Symposium. *Quant. Biol.*, **22**, 415–427.

Johst, Karin, & Drechsler, Martin. 2003. Are spatially correlated or uncorrelated disturbance regimes better for the survival of species? *OIKOS*, **103**, 449–456.

Johst, Karin, Brandl, Roland, & Eber, Sabine. 2002. Metapopulation persistence in dynamic landscapes: the role of dispersal distance. *OIKOS*, **98**, 263–270.

Jones, T. H., Godfray, H. C. J., & Hassell, M. P. 1996. Relative movement patterns of a tephritid fly and its parasitoid wasps. *Oecologia*, **106**, 317–324.

Kearney, Michael, & Porter, Warren P. 2004. Mapping the fundamental niche: Physiology, climate, and the distribution of a nocturnal lizard. *Ecology*, **85**, 3119–3131.

King, Aaron A., & Hastings, Alan. 2003. Spatial mechanisms for coexistence of species sharing a common natural enemy. *Theoretical Population Biology*, **64**, 431–438.

Luoto, M., Heikkinen, R. K., Pöyry, J., & Saarinen, K. 2006. Determinants of the biogeographical distribution of butterflies in boreal regions. *Journal of Biogeography*, **33**, 1764–1778.

Maggini, R., Guisan, A., & Cherix, D. 2002. A stratified approach for modeling the distribution of a threatened ant species in the Swiss National Park. *Biodiversity and Conservation*, **11**, 2177–2141.

May, Robert M. 1974. *Stability and Complexity in model ecosystems.* Princeton University Press, Princeton, NJ.

May, Robert M., & Mac Arthur, Robert H. 1972. Niche overlap as a function of environmental variability. *PNAS*, **69**, 1109–1113.

McCullagh, P., & Nelder, John A. 1989. *Generalized linear models.* 2nd edn. Monographs on statistics and applied probability, vol. 37. Chapman and Hall.

McPherson, Jana M., & Jetz, Walter. 2007. Effects of species' ecology on the accuracy of distribution models. *Ecography.*

McPherson, Jana M., Jetz, Walter, & Rogers, David J. 2004. The effects of species' range shifts on the accuracy of distribution models: ecological phenomenon or statistical artefact? *Journal of Apllied Ecology*, **41**, 811–823.

Menard, Scott. 2000. Coefficients of determination for multiple logistic regression analysis. *The American Statistician*, **54**, 17–24.

Miller, M. E., Hui, S. L., & Tierney, W. M. 1991. Validation techniques for logistic regression models. *Stat. Med.*, **10**, 1213–26.

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

Moloney, K. A., & Levin, S. A. 1996. The effects of disturbance architecture on landscape-level population dynamics. *Ecology*, **77**, 375–394.

Moore, I. D., Norton, T. W., & Williams, J. E. 1993. Modeling environmental heterogeneity in forested landscapes. *Journal of Hydrology*, **150**, 717–747.

Moran, P. A. P. 1950. Some remarks on animal population dynamics. *Biometriks*, **6**, 250–258.

Neubert, M. G., Kot, M., & Lewis, M. A. 1995. Dispersal and pattern formation in a discrete-time predator-prey model. *Theor. Popul. Biol.*, **48**, 7–43.

Nicholson, A. J. 1933. The balance of animal populations. *Journal of Animal Ecology*, **2**, 131–178.

Nicholson, A. J., & Bailey, V. A. 1935. The balance of animal populations. *Proceedings of the Zoological Society of London*, **3**, 551–598.

O'Neill, Michael P., Schmidt, John C., Dobrowolski, James P., Hawkins, Charles P., & Neale, Christopher M. U. 1997. Identifying sites for Riparian wetland restoration: application of a model to the upper Arkansas river basin. *Restoration Ecology*, **5**, 85–102.

O'Neill, R. V. 1968. Population energetics of a milipede. Narceus Americanus (Beavois). *Ecology*, **49**, 803–809.

O'Neill, R. V., Oldstein, R.A., Shugart, H.H., & Makin, J.B. 1972. Terrestrial ecosystem energy model. *In: Eastern deciduous forest biome memo report.*

Pearson, Richard G., & Dawson, Terence P. 2003. Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Global Ecology and Biogeography*, **12**, 361–371.

Pearson, Richard G., Thuiller, Wilfried, Araujo, Miguel B., Martinez-Meyer, Enrique, Brotons, Lluis, McClean, Colin, Miles, Lera, Segurado, Pedro, Dawson, Terence P., & Lees, David C. 2006. Model-based uncertainty in species range prediction. *Journal of Biogeography*, **33**, 1704–1711.

R Development Core Team. 2007. *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Available at http://www.r-project.org/.

Pulliam, H. R. 1988. Sources, sinks and population regulation. *The American Naturalist*, **132**, 652–661.

Reineking, Björn, & Schröder, Boris. 2003. *Gene, Bits und Ökosysteme.* P. Lang Verlag Frankfurt/M. Chap. Computer-intensive Methods in the Analysis of Species-habitat Relationships, pages 165–172.

Reineking, Björn, & Schröder, Boris. 2006. Constrain to perform: Regularization of habitat models. *Ecological Modelling*, **193**, 675–690.

Revermann, R., Zbinden, N., Schmid, H., Marti, Ch., Spaar, R., & Schröder, B. submitted. Suitable habitat for ptarmigan (*Lagopus muta helvetica* Thienemann 1829) in the Swiss Alps and its response to rapid climate change in the 21st century - a multi-scale approach.

Richter, Otto, Diekkrüger, Bernd, & Nörtesheuser, Peter. 1996. *Environmental Fate Modelling of Pesticides: From the Laboratory to the Field Scale.* VCH.

Ricker, W. E. 1954. Stock and recruitment. *Journal of the Fish Research Board of Canada*, **11**, 559–623.

Rudner, M., Biedermann, R., Schröder, B., & Kleyer, M. 2007. Integrated Grid Based Ecological and Economic (INGRID) landscape model - A tool to support landscape management decisions. *Environmental Modelling and Software*, **22**, 177–187.

Saupe, Dietmar. 1988. Algorithms for random fractals. *Pages 71–113 of: The Science of Fractal Images.* New York, NY, USA: Springer-Verlag New York, Inc.

Schofield, Peter G., Chaplain, Mark A. J., & Hubbard, Stephen F. 2005. Dynamic heterogeneous spatio-temporal pattern formation in host-parasitoid systems with synchronised generations. *J. Math. Biol.*, **50**, 559–583.

Schröder, Boris, & Richter, Otto. 1999. Are habitat models transferable in space and time? *Zeitschrift für Ökologie und Naturschutz*, **8**, 195–205.

Segurado, Pedro, & Araujo, Miguel B. 2004. An evaluation of methods for modelling species distributions. *Journal of Biogeography*, **31**, 1555–1568.

Söndgerath, Dagmar, & Schröder, Boris. 2002. Population dynamics and habitat connectivity affecting the spatial spread of populations - a simulation study. *Landscape Ecology*, **17**, 57–70.

Stockwell, David R. B., & Peterson, A. Townsend. 2002. Effects of sample size on accuracy of species distribution models. *Ecological Modelling*, **148**, 1–13.

Svenning, Jens-Christian, & Skov, Flemming. 2004. Limited filling of the potential range in European tree species. *Ecology Letters*, **7**, 565–573.

Svenning, Jens-Christian, & Skov, Flemming. 2007. Could the tree diversity pattern in Europe be generated by postglacial dispersal limitation? *Ecology Letters*, **10**, 453–460.

Thomson, James R., Mac Nally, Ralph, Fleishman, Erica, & Horrocks, Greg. 2007. Predicting bird species distributions in reconstructed landscapes. *Conservation Biology*, **21**, 752–766.

Thuiller, Wilfried. 2003. BIOMOD - optimizing predictions of species distributions and projecting potential future shifts under global change. *Global Change Biology*, **9**, 1353–1362.

Thuiller, Wilfried. 2004. Patterns and uncertainties of species' range shifts under climate change. *Global Change Biology*, **10**, 2020–2027.

Tyre, Andrew J., Possingham, Hugh P., & Lindenmayer, David B. 2001. Inferring process from pattern: can territory occupancy provide information about life history parameters? *Ecological Applications*, **11**(6), 1722–1737.

Whittingham, Mark J., Stephens, Philip A., Bradbury, Richard B., & Freckleton, Robert P. 2006. Why do we still use stepwise modelling in ecology and behaviour? *Journal of Animal Ecology*, **75**, 1182–1189.

Wilson, . H. B., & Hassel, M. P. 1997. Host-parasitoid spatial models: the interplay of demographic stochasticity and dynamics. *Proc. R. Soc. Lond.*, **264**, 1189–1195.

Wilson, John P., & Gallant, John C. (eds). 2000. *Terrain Analysis: Principles and Applications*. John Wiley & Sons.

Wintle, Brendan A., Elith, Jane, & Potts, Joanne M. 2005. Fauna habitat modelling and mapping: A review and case study in the Lower Hunter Central Coast region of NSW. *Austral Ecology*, **30**, 719–738.

With, Kimberly A. 1997. The application of neutral landscape models in conservation biology. *Conservation Biology*, **11**, 1069–1080.

A Evolution of Population Abundances

Animated graphics of the evolution of population abundances over time can be found on the enclosed data DVD.

B Univariate LMs

	Equilibrium			**Climate Change**		
Predictor	Intercept	Parameter	r^2	Intercept	Parameter	r^2
GLMs						
niche width	0.80 ± 0.01	**-0.09 ± 6e-3** $\ast\ast$	0.27	0.70 ± 0.01	**-0.05 ± 6e-3** $\ast\ast$	0.11
hosts growth rate	0.64 ± 0.01	**-7e-4 ± 2e-4** $\ast\ast$	0.02	0.60 ± 0.01	-2e-4 ± 2e-4	2e-3
level of parasitism	0.62 ± 5e-3	**-2e-3 ± 7e-4** $\ast\ast$	0.01	0.60 ± 5e-3	**-3e-3 ± 7e-4** $\ast\ast$	0.02
dispersal distance	0.59 ± 5e-3	**0.01 ± 2e-3** $\ast\ast$	0.04	0.58 ± 5e-3	2e-3 ± 2e-3	2e-3
time elapse	0.81 ± 0.01	**-2e-3 ± 1e-4** $\ast\ast$	0.24	0.77 ± 0.01	**-2e-3 ± 1e-4** $\ast\ast$	0.21
sample size	0.58 ± 6e-3	**1e-4 ± 2e-5** $\ast\ast$	0.03	0.56 ± 6e-3	**1e-4 ± 2e-5** $\ast\ast$	0.04
GAMs						
niche width	0.74 ± 0.01	**-0.06 ± 5e-3** $\ast\ast$	0.18	0.66 ± 0.01	**-0.02 ± 5e-3** $\ast\ast$	0.03
hosts growth rate	0.63 ± 9e-3	-3e-4 ± 2e-4	4e-3	0.61 ± 8e-3	5e-5 ± 1e-4	1e-4
level of parasitism	0.61 ± 4e-3	1e-3 ± 6e-4 $\ast$	6e-3	0.59 ± 4e-3	**3e-3 ± 6e-4** $\ast\ast$	0.04
dispersal distance	0.61 ± 4e-3	**4e-3 ± 2e-3** $\ast\ast$	0.01	0.62 ± 4e-3	**-5e-3 ± 1e-3** $\ast\ast$	0.01
time elapse	0.760 ± 0.01	**-2e-3 ± 1e-4** $\ast\ast$	0.18	0.70 ± 0.01	**-1e-3 ± 1e-4** $\ast\ast$	0.09
sample size	0.59 ± 5e-3	**1e-4 ± 2e-5** $\ast\ast$	0.06	0.58 ± 5e-3	**9e-5 ± 1e-5** $\ast\ast$	0.05

Table B.1: Effects of ecological traits and interspecific interaction on accuracy of spatial prediction of GLMs and GAMs estimated on sampled and validated against entire data. Given are the intercept and slope of the univariate linear regressions (± SE), and the explained deviance r^2. Significance is indicated by * for $p \leq 0.05$ and by ** for $p \leq 0.01$. Parameters in bold remained significant ($p \leq 0.05$) after Holm adjustment for multiple comparison.

	Equilibrium			**Climate Change**		
Predictor	Intercept	Parameter	r^2	Intercept	Parameter	r^2
Projection for T=70						
GLMs						
niche width	0.85 ± 0.01	**-0.12 ± 6e-3** $**$	0.32	0.59 ± 0.02	-0.02 ± 9e-3 $*$	7e-3
host growth rate	0.64 ± 0.01	**-5e-4 ± 2e-4** $*$	6e-3	0.56 ± 0.01	-2e-4 ± 3e-4	9e-4
level of parasitism	0.63 ± 6e-3	**-3e-3 ± 9e-4** $**$	0.02	0.56 ± 7e-3	-2e-3 ± 1e-3	5e-3
dispersal distance	0.58 ± 6e-3	**0.01 ± 2e-3** $**$	0.05	0.51 ± 7e-3	**0.02 ± 3e-3** $**$	0.06
time elapse	0.82 ± 0.02	**-2e-3 ± 2e-4** $**$	0.18	0.93 ± 0.02	**-4e-3 ± 2e-4** $**$	0.43
sample size	0.58 ± 7e-3	**1e-4 ± 2e-5** $**$	0.03	0.55 ± 8e-3	3e-5 ± 3e-5	1e-3
GAMs						
niche width	0.80 ± 0.01	**-0.09 ± 6e-3** $**$	0.25	0.55 ± 0.02	2e-3 ± 9e-3	6e-5
host growth rate	0.62 ± 0.01	-2e-4 ± 2e-4	1e-3	0.55 ± 0.01	9e-5 ± 2e-4	2e-4
level of parasitism	0.62 ± 5e-3	-1e-3 ± 8e-4	5e-3	0.55 ± 7e-3	-3e-4 ± 9e-4	2e-4
dispersal distance	0.59 ± 5e-3	**0.01 ± 2e-3** $**$	0.04	0.52 ± 7e-3	**0.01 ± 2e-3** $**$	0.04
time elapse	0.79 ± 0.01	**-2e-3 ± 2e-4** $**$	0.17	0.91 ± 0.02	**-4e-3 ± 2e-4** $**$	0.45
sample size	0.58 ± 6e-3	**1e-4 ± 2e-5** $**$	0.05	0.55 ± 8e-3	3e-5 ± 2e-5	2e-3
Projection for T=95						
GLMs						
niche width	0.82 ± 0.01	**-0.12 ± 5e-3** $**$	0.44	0.66 ± 0.01	**-0.05 ± 5e-3** $**$	0.13
host growth rate	0.63 ± 0.01	**-7e-4 ± 2e-4** $**$	0.02	0.58 ± 9e-3	**-6e-4 ± 2e-4** $**$	0.02
level of parasitism	0.59 ± 5e-3	-1e-3 ± 7e-4	2e-3	0.56 ± 4e-3	-9e-4 ± 6e-4	3e-3
dispersal distance	0.56 ± 5e-3	**0.01 ± 2e-3** $**$	0.08	0.53 ± 4e-3	**0.01 ± 2e-3** $**$	0.07
time elapse	0.68 ± 0.02	**-1e-3 ± 2e-4** $**$	0.05	0.58 ± 0.01	**-3e-4 ± 1e-4** $*$	8e-3
sample size	0.57 ± 6e-3	**9e-5 ± 2e-5** $**$	0.03	0.54 ± 5e-3	**4e-5 ± 2e-5** $*$	8e-3
GAMs						
niche width	0.76 ± 9e-3	**-0.08 ± 4e-3** $**$	0.32	0.58 ± 0.01	-7e-3 ± 5e-3	2e-3
host growth rate	0.62 ± 8e-3	-3e-4 ± 2e-4	5e-3	0.58 ± 8e-3	-3e-4 ± 1e-4	4e-3
level of parasitism	0.58 ± 4e-3	**3e-3 ± 6e-4** $**$	0.04	0.54 ± 4e-3	**5e-3 ± 5e-4** $**$	0.09
dispersal distance	0.58 ± 4e-3	**8e-3 ± 2e-3** $**$	0.04	0.56 ± 4e-3	**5e-3 ± 1e-3** $**$	0.02
time elapse	0.68 ± 0.01	**-8e-4 ± 1e-4** $**$	0.05	0.57 ± 0.01	-5e-5 ± 1e-4	2e-4
sample size	0.57 ± 5e-3	**1e-4 ± 2e-5** $**$	0.06	0.56 ± 5e-3	**5e-5 ± 1e-5** $**$	0.01
Projection for T=120						
GLMs						
niche width	0.63 ± 8e-3	**-0.04 ± 4e-3** $**$	0.13	0.47 ± 0.01	**0.02 ± 6e-3** $**$	0.02
host growth rate	0.59 ± 7e-3	**-7e-4 ± 1e-4** $**$	0.05	0.51 ± 9e-3	2e-5 ± 2e-4	2e-5
level of parasitism	0.55 ± 3e-3	-2e-4 ± 5e-4	3e-4	0.52 ± 4e-3	**-2e-3 ± 6e-4** $**$	0.01
dispersal distance	0.52 ± 3e-3	**0.01 ± 1e-3** $**$	0.12	0.51 ± 5e-3	2e-3 ± 2e-3	3e-3
time elapse	0.52 ± 0.01	**3e-4 ± 1e-4** $**$	0.01	0.31 ± 0.01	**2e-3 ± 1e-4** $**$	0.34
sample size	0.54 ± 4e-3	**5e-5 ± 1e-5** $**$	0.02	0.51 ± 5e-3	2e-5 ± 2e-5	2e-3
GAMs						
niche width	0.60 ± 8e-3	**-0.02 ± 4e-3** $**$	0.02	0.43 ± 0.01	**0.06 ± 6e-3** $**$	0.10
host growth rate	0.59 ± 6e-3	**-4e-4 ± 1e-4** $**$	0.02	0.54 ± 0.01	2e-4 ± 2e-4	1e-3
level of parasitism	0.54 ± 3e-3	**5e-3 ± 4e-4** $**$	0.18	0.51 ± 5e-3	**7e-3 ± 7e-4** $**$	0.12
dispersal distance	0.56 ± 3e-3	**6e-3 ± 1e-3** $**$	0.04	0.55 ± 5e-3	-3e-3 ± 2e-3	4e-3
time elapse	0.54 ± 9e-3	**4e-4 ± 1e-4** $**$	0.02	0.33 ± 0.01	**2e-3 ± 1e-4** $**$	0.27
sample size	0.56 ± 4e-3	**6e-5 ± 1e-5** $**$	0.04	0.54 ± 6e-3	2e-5 ± 2e-5	2e-3

Table B.2: Effects of ecological traits and interspecific interaction on accuracy of spatiotemporal prediction of GLMs and GAMs. Given are the intercept and slope of the univariate linear regressions (± SE), and the explained deviance r^2. Significance is indicated by * for $p \leq 0.05$ and by ** for $p \leq 0.01$. Parameters in bold remained significant ($p \leq 0.05$) after Holm adjustment for multiple comparison.

Zusammenfassung

Klima- und Landnutzungswandel bergen zunehmende Gefahren für Arten und Biodiversität. Die Vorhersage möglicher Auswirkungen dieser Veränderungen gewinnt daher immer mehr an Relevanz, z. B. für die Naturschutzplanung. Statistische Habitateignungsmodelle wurden in diesem Zusammenhang häufig eingesetzt. Dennoch wird deren Verwendung oftmals kritisiert aufgrund ihrer impliziten Annahmen eines (Quasi-) Gleichgewichtes und gesättigter Bruthabitate. Außerdem werden demographische Prozesse und biotische Interaktionen nicht explizit in den Habitateignungsmodellen berücksichtigt. In dieser Studie habe ich versucht, die Gültigkeit der zugrunde liegenden Annahmen sowie die Effekte ökologischer Prozesse und transienter Dynamiken auf die Vorhersagegüte von Habitateignungsmodellen zu überprüfen, und zwar mithilfe eines virtuellen dynamischen Systems. Dazu habe ich ein räumlich explizites Multispezies populationsdynamisches Modell entwickelt, welches art- und interspezifische ökologische Prozesse, demographische und Umweltstochastizität sowie Umweltwandel berücksichtigt. Ein virtueller Ökologe hat in verschiedenen Szenarien Stichproben erhoben und mithilfe verallgemeinerter linearer (GLMs) und verallgemeinerter additiver Modelle (GAMs) Habitateignungsmodelle geschätzt. Mit diesen wurden räumliche und raumzeitliche Vorhersagen der Artverbreitung erstellt und gegen die simulierte wahre Verbreitung validiert. Die resultierenden Vorhersagegüten wurden dann in Beziehung gesetzt zu den vorherrschenden ökologischen Prozessen und zeitlichen Dynamiken.

Die Ergebnisse zeigen, dass der Gleichgewichtszustand und die Bereitwilligkeit, in einen Gleichgewichtszustand zurückzukehren, entscheidend sind für die Gültigkeit und Übertragbarkeit von Habitateignungsmodellen. Außerdem werden diese Effekte durch ökologische Charakteristika und Prozesse beeinflusst. Entscheidend sind dabei vor allem die ökologische Potenz der Arten und ihr Ausbreitungsverhalten. Vorhersagegüten reagierten darüber hinaus empfindlich auf räumliche Populations- sowie Quellen-Senken-Dynamiken. So sanken sie mit zunehmender Nutzung von Habitaten geringerer Qualität. Demographische und Umweltstochastizität erschweren außerdem die Erklärung großer Anteile an Varianz im

Vorkommen der Arten. Diese Ergebnisse waren konsistent für die verschiedenen Modellalgorithmen. Meine Analysen stellen somit eine vorläufige Orientierungshilfe dar, unter welchen Umständen die Verbreitung von Arten zuverlässig durch Habitateigungsmodelle vorhergesagt werden kann und wo Schnittstellen zu mechanistischen, prozessbasierten Modellen vonnöten sind.

Danksagung

An dieser Stelle möchte ich allen danken, die direkt oder indirekt zum Gelingen dieser Arbeit beigetragen haben.

Zuallererst möchte ich Dr. Boris Schröder und Prof. Dr. Florian Jeltsch danken für die Bereitstellung des Diplomarbeitsthemas, ihre Betreuung und ihr Vertrauen in meine Fähigkeiten bei der Anfertigung dieser Arbeit. Dr. Boris Schröder sei besonders dafür gedankt, dass er sich trotz engen Zeitplans immer ausführlich Zeit genommen hat, stets mit mir Ideen und Probleme ausgiebig diskutiert hat, nicht nur in allen fachlichen Fragen, sondern auch den kleinen und großen Sorgen, die eine solche Arbeit begleiten.

Besonders bedanken möchte ich mich auch bei Andreas Bauer, dessen technische Unterstützung, Verständnis und Hilfe bei der Bereitstellung von Rechenressourcen, beträchtlich zum Gelingen dieser Arbeit (und meiner persönlichen Aufmunterung) beigetragen haben.

Weiterhin gebührt mein Dank meinen Freunden und meiner Familie, die mich unterstützt und ermutigt haben, wo es nötig war. Ohne die in allen Lebenslagen umfangreiche Unterstützung meiner Eltern, Marianne und Günter Zurell, wäre diese Arbeit und der (hoffentlich erfolgreiche) Abschluss meines Studiums nicht möglich gewesen. Und zuletzt, jedoch für mich an erster Stelle, danke ich meinem Freund, Andreas Passing, dass er mir in den letzten Jahren immer hilfreich zur Seite gestanden und mich immer wieder neu aufgebaut und motiviert hat, vor allem jedoch für seine Ruhe und Geduld.

Eidesstattliche Erklärung

Ich versichere hiermit an Eides Statt, dass ich die von mir eingereichte Diplomarbeit selbständig verfasst und ausschließlich die angegebenen Hilfsmittel benutzt habe.

Die Arbeit wurde bisher in gleicher oder ähnlicher Form keiner anderen Prüfungsbehörde vorgelegt und auch nicht veröffentlicht.

Ort, Datum

Unterschrift