

How fine is fine-scale? Questioning the use of fine-scale bioclimatic data in species distribution models used for forecasting abundance patterns in butterflies

KATHARINA J. FILZ¹, THOMAS SCHMITT¹ and JAN O. ENGLER^{1,2}

¹Biogeography Department, Faculty of Regional and Environmental Sciences, Trier University, Universitätsring 15, D-54296 Trier, Germany; e-mails: kfilz@yahoo.de; thsh@uni-trier.de

²Zoological Research Museum Alexander Koenig, Adenauerallee 160, D-53113 Bonn, Germany; e-mail: j.engler.zfmk@uni-bonn.de

Key words. Invertebrates, spatial abundance, environmental suitability, environmental niche model, MAXENT

Abstract. The use of species distribution models (SDMs) to predict the spatial occurrence and abundance of species in relation to environmental predictors has been debated in terms of species' ecology and biogeography. The predictive power of these models is well recognized for vertebrates, but has not yet been tested for invertebrates. In this study, we aim to assess the use of SDMs for predicting local abundances of invertebrates at a macroscale level. A maximum entropy algorithm was used to build SDMs based on occurrence records of 61 species of butterflies and bioclimatic information with a 30 arc second resolution. Predictions of habitat suitability were correlated with butterfly abundance data derived from independently conducted field surveys in order to check for a relationship between the predictions of the model and local abundances. Even though the model accurately described the current distributions of the species in the study area at a macroscale, the observed occurrences of the species (i.e. presence/absence) recorded by the field surveys differed significantly from the model's predictions for the corresponding grid cells. Moreover, there was no correlation between observed abundance and the model's predictions for most species of butterflies. We conclude that the spatial abundance of butterflies cannot be predicted from environmental suitability modelled at a resolution as large as in this study. Using the finest scale bioclimatic information currently available (i.e. 30 arc seconds) it is not adequate to predict species abundances as structural and ecological factors as well as climatic patterns acting at a smaller scale are key determinants of the occurrence and abundance of invertebrates. Therefore, future studies have to account for the role of the resolution in environmental predictors when assessments of spatial abundances via SDMs will be conducted.

INTRODUCTION

Environmental conditions and population processes determine the spatial distribution of species and hence biodiversity patterns over geographic ranges (Gaston, 2003). In this context, species distribution models (SDMs) are commonly used to predict the potential distribution of species with regard to their ecological niches (e.g. Franklin, 2009; Soberon & Nakamura, 2009; Araújo & Peterson, 2012). Based on multiple climatic and environmental variables recorded over the known distribution, these models aim to predict spatial patterns of environmental suitability and are used to infer the likelihood of species' occurrence across a given geographic range. In recent times, SDMs have emerged as powerful tools for addressing important topics in ecology (Ficetola et al., 2007; Rödder et al., 2009), evolution (Habel et al., 2010, 2011) and conservation (Rodríguez et al., 2007; Araújo et al., 2011). Despite the well recognized relationship between environmental suitability and local abundance of species, only a few authors have tried to deduce spatial abundance patterns of species from SDMs (Pearce & Ferrier, 2001; Nielsen et al., 2005; Pearce & Boyce, 2006; VanDerWal et al., 2009). The basic idea behind this assumption is that environmental suitability predicted by a SDM for a given location can be used as an indicator of species' abundance as it indicates how well the physical

and ecological constraints of species are met. If this is the case then the species will be abundant at locations with high environmental suitability and vice versa. Hence, models that predict environmental suitability based on occurrence data might also provide information on spatial variation in abundance.

It is much more difficult to obtain data on the abundance of species within their ranges than data on their occurrence. Species often vary greatly in abundance in time and in the different habitats within their ranges (Murphy et al., 2006). Even closely related species differ greatly in their mean population densities in the same range (Bink, 1992). Furthermore, the ease with which individuals can be detected strongly depends on season, time of day, weather conditions, habitat type, faunal activity and the skill of the observer. Therefore, detailed data on abundance are available for only a few species at a local scale and area-wide information is commonly lacking because collecting such data for a large number of species requires a lot labour and is consequently expensive. Assuming a positive correlation between habitat quality and species abundance within a region, such information can be used to target field surveys over large geographical areas. Moreover, modelling species abundances could provide information that could be used to develop regional conservation programs. However, their effective implementation at large scales require that there

is a good correlation between “high-quality” habitats and species abundance. As relative abundance is likely to be a good indicator of population viability and reflect factors such as reproductive success, carrying capacity and susceptibility of populations to extinction (Keller et al., 1986; Hobbs & Hanley, 1990), deducing abundance from environmental suitability and extrapolating the results across a region might be a useful tool. The use of SDMs to predict species specific abundance patterns might be a feasible way of determining spatially explicit assessments of species’ abundances, which, in turn, could be used to optimise specific conservation efforts.

Since SDMs have been only used to predict spatial patterns in abundance of vertebrates (cf. VanDerWal et al., 2009; Brambilla & Ficetola, 2012), we aim to apply this approach to invertebrates for the first time, using central European species of butterflies as the model system. Occurrence records for 61 species were used to develop species distribution models using the maximum entropy algorithm MAXENT (Phillips et al., 2006). Predictions of environmental suitability were then correlated with abundance data obtained from standardized field surveys in order to investigate whether it is possible to predict variations in the spatial abundance of species of butterflies using SDMs.

The aim of this study was to determine the suitability of the data on bioclimatic information at the finest resolution (30 arc seconds) currently available in the Worldclim database (Hijmans et al., 2005; www.worldclim.org), which is the standard source of variables for modelling species distributions world-wide (i.e. Hijmans et al., 2005), as demonstrated by a citation index of 1,432 (ISI Web of Science query, 6-6-2012). However, even if the resolution is quite fine on the macroscale (grid cell resolution equates approximately 1 km² along the equator), we hypothesize that it might be difficult or even impossible at this level of resolution to use SDMs to predict accurately abundance patterns. If so, then this study will demonstrate the limitations of the current method of assessing abundance data using the bioclimatic datasets currently available.

METHODS

Species abundance data were obtained from field surveys carried out at 14 locations in south-western Germany in 2010 and 2011 using a standardized transect method developed for the British Monitoring Scheme (Pollard & Yates, 1993). Butterflies were counted when walking along transects, which varied in length between 123 and 1430 m. Each butterfly seen within 5 m ahead and 2.5 m on each side of the observer was counted. Individuals were either identified and counted without capturing them or caught using a butterfly net for closer determination. All locations were surveyed several times each year (monthly between April and October) to avoid misinterpretation of extreme or zero abundances that can arise from single surveys and seasonal variations in the numbers of certain butterfly species. The means of the species counts for each month were transformed to numbers per 1000 m of transect and summed to give an annual value taking into account the differences in transect length. The transformed monthly counts of species were summed to give a single annual value for each transect.

Species distribution modelling (SDM) requires environmental and species occurrence data. Occurrence of butterflies was obtained from intensive field surveys at 148 locations across the study area and from a GBIF query (<http://www.gbif.org>). We selected a region between 50.5°N and 48.9°N and 5.8°E and 8.2°E, which includes an area in south-western Germany and adjacent regions in Luxembourg and France (Fig. 1). The long history of human settlement and different land-uses in this region resulted in a complex landscape matrix, which ensured the survival of a diversity of species including a larger number of butterfly taxa. The landscape encompasses a mosaic consisting of residential areas, arable fields, vineyards, meadows, forests and semi-natural calcareous grassland. The latter especially is a favourable habitat for many butterfly species and rare and endangered taxa (Wenzel et al., 2006). Therefore, butterflies have been well studied in this area.

Information on the occurrence of a total of 61 butterfly species, for which there was information on their abundance available, was used. The mean number of occurrence records per species was 48, ranging from 5 to 126 records for a single species (Table 1). Even if modelling algorithms have a high potential error rate if the information on occurrence is limited to just a few locations (Hernandez et al., 2006; Wisz et al., 2008), all species were included, because the information on occurrence was mainly limited by the rarity of a given species. Since most of the rare butterflies in the study area are habitat specialists, we assumed they have a high preference for a specific niche. Thus, rare species with a very limited range of environmental tolerance can also result in accurate SDMs, even if information on their occurrence is scarce (sensu Hernandez et al., 2006; de Siqueira et al., 2009).

Bioclimatic information with a spatial resolution of 30 arc seconds was obtained from the Worldclim database (Version 1.4, <http://www.worldclim.org>; Hijmans et al., 2005). Nineteen Bioclim variables, all of which are assumed to strongly influence the occurrence and abundance of butterfly species, were checked for multi-colinearity by conducting pairwise Pearson correlations. High inter-correlations between predictor variables might inflate the performance of SDMs (Heikkinen et al., 2006), when redundant information was used for calculating the climatic niche of a species. If $r^2 > 0.75$, we therefore selected only one of these strongly inter-correlated variables. In these cases we preferred those with a higher relevance to butterfly biology (i.e. extremes rather than means) as extremes seem to limit butterfly distributions in a more direct way than means. The final data set included eight variables: “isothermality” (bio3), “temperature seasonality” (bio4), “maximum temperature of warmest month” (bio5), “minimum temperature of coldest month” (bio6), “mean temperature of wettest quarter” (bio8), “mean temperature of driest quarter” (bio9), “precipitation of driest month” (bio14) and “precipitation seasonality” (bio15).

We used MAXENT 3.3.3k (Phillips et al., 2006; Elith et al., 2011; available through: <http://www.cs.princeton.edu/~shapire/maxent>), a machine-learning algorithm following the principles of maximum entropy, for species distribution modelling. MAXENT models potential species distributions based on environmental predictors (i.e. the above mentioned eight bioclimatic variables in this case) and presence-only data (Elith et al., 2011). In doing so, this algorithm frequently outperforms other methods (e.g. Elith et al., 2006). In addition, and even more important in the context of this study, MAXENT is the best of the available algorithms when there are few species records (Hernandez et al., 2006; Pearson et al., 2007; Wisz et al., 2008). This is particularly important when modelling the distribution of the rare species of butterflies included in this survey.

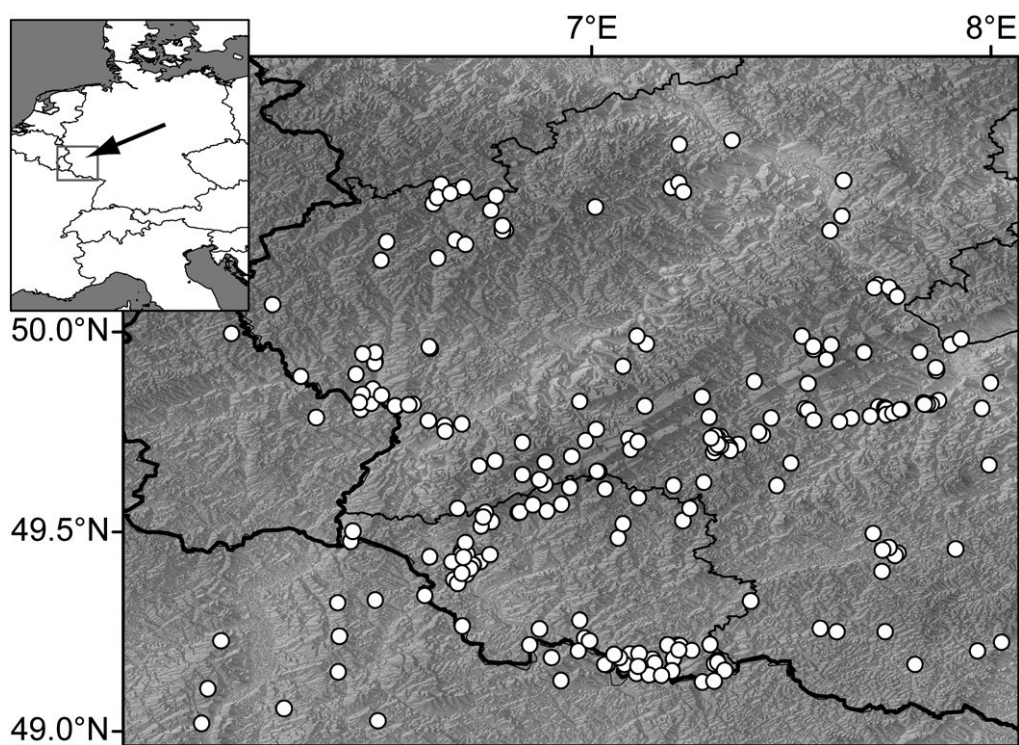


Fig. 1. Map showing the location of the study area in central Europe. Circles indicate the occurrence locations for which there was information on butterfly species available.

MAXENT allows for the calculation of the “area under the receiver operation characteristic curve” (AUC) in order to test the predictive outcome of SDMs (Phillips et al., 2006). AUC values range from 0.5 for models with no predictive ability to 1.0 for those giving perfect predictions (Swets, 1988) and can be used to assess the ability of the model to distinguish species records from background data (Phillips et al., 2006). Models were computed with 30% of the records randomly omitted as test points from the model during training in 100 iterations, in order to assess the internal consistency of the model (Phillips et al., 2006). Subsequently, the average of all 100 models automatically computed by MAXENT was used in further analyses. The logistic output of MAXENT is a continuous map interpreted as the potential distribution of the species studied in the area of interest based on the predicted environmental suitability from 0 (unsuitable conditions) to 1 (optimal conditions). We used a non-fixed threshold as recommended by Liu et al. (2005) and set the minimum training presence prediction value as presence/absence threshold.

We compared the expected occurrence (represented by the presence/absence prediction in the corresponding grid cell as derived from the SDM) with the observed occurrence (i.e. species presence/absence along each transect) for each species and tested the general deviance across all species using the χ^2 -test. In addition, spearman rank correlations between the abundance information for each species for each transect and the corresponding predictions derived from the SDMs for the respective grid cells (i.e. the predicted environmental suitability for the respective butterfly species ranging from 0 to 1) were obtained. Calculations were conducted in R 2.15 (R development core team, 2012).

RESULTS

The distribution of butterfly species in the study area was patchy and the number of occurrence records varied

considerably between species and habitats. A total of 65 species were recorded during transect walks. The highest number of species ($n = 59$) was recorded in calcareous grasslands, most of which were protected areas. Fifty three of the 65 species of butterfly were recorded in vineyard fallows in cultivated landscapes. The frequency of detection of a species mostly depended on its rarity in this region, i.e. the number of records was positively correlated with the mean number of individuals per transect along which the species in question was present ($r^2 = 0.38$). The most frequent species was the Meadow Brown (*Maniola jurtina*, $n = 858$ individuals). More than 100 individuals were recorded of 15 species, 51–100 of eight species, 11–50 of 21 species and ten or fewer individuals of 21 species. The abundance data for all species recorded along all transects is given in Table 1.

The AUC values of the model’s performance averaged across the set of 61 butterfly species analyzed (i.e. 4 of the 65 species were not modelled, due to insufficient occurrence records) was 0.785 (sd: 0.057). Therefore, the model can be considered as “useful” for predicting the local presence or absence of a species according to the classification scheme adapted from Swets (1988) and modified by Araújo et al. (2005). The lowest AUC value for a species was that for *T. betulae* (AUC = 0.56, $n = 12$), i.e. it is not possible to reliably predict the distribution of this species in the study area, and highest for *L. dispar* (AUC = 0.93, $n = 6$), i.e. the prediction of the distribution of this species in the study area is reliable. There is no linear relationship between the number of occurrences and the performance of the model ($r^2 = 0.003$, $p = 0.67$), thus the distribution of common and widespread

TABLE 1. Summary table for 61 species of butterflies for which abundance data was collected in the field and SDMs conducted (n locations – number of occurrence records used in the development of the model; AUC – Area Under the receiver operation characteristic Curve; sum abundance – recorded abundances of species during field surveys; % congruence – percentage of congruence between expected and observed presence of species, r – spearman rank correlation coefficient; p – p -value of spearman rank correlation).

Species	n locations	Test AUC	Sum abundance	% congruence	r	p
<i>A. aglaja</i>	32	0.72	59	78.6	0.27	0.354
<i>A. cardamines</i>	40	0.78	60	92.9	–0.12	0.674
<i>A. crataegi</i>	25	0.80	22	42.9	0.30	0.298
<i>A. hyperanthus</i>	63	0.78	326	100.0	0.45	0.104
<i>A. ilia</i>	5	0.78	1	85.7	0.10	0.726
<i>A. iris</i>	18	0.68	5	28.6	0.36	0.208
<i>A. levana</i>	53	0.80	25	78.6	–0.31	0.283
<i>A. paphia</i>	52	0.79	22	71.4	0.03	0.928
<i>B. dia</i>	46	0.81	29	28.6	–0.15	0.612
<i>B. ino</i>	17	0.66	2	14.3	–0.10	0.730
<i>B. selene</i>	26	0.86	2	57.1	0.42	0.133
<i>C. alceae</i>	21	0.82	3	50.0	0.46	0.097
<i>C. alfacariensis</i>	41	0.81	2	92.9	0.45	0.109
<i>C. arcania</i>	51	0.78	139	92.9	0.72	0.003
<i>C. argiolus</i>	30	0.84	17	71.4	0.19	0.517
<i>C. hyale</i>	64	0.79	45	71.4	–0.25	0.397
<i>C. minimus</i>	33	0.78	81	50.0	0.36	0.206
<i>C. palaemon</i>	32	0.82	1	7.1	–0.45	0.109
<i>C. pamphilus</i>	126	0.76	272	100.0	0.57	0.033
<i>C. rubi</i>	54	0.84	54	78.6	0.62	0.018
<i>C. semiargus</i>	57	0.76	61	85.7	–0.08	0.776
<i>E. medusa</i>	40	0.82	2	7.1	0.45	0.109
<i>E. tages</i>	56	0.80	7	14.3	–0.21	0.478
<i>G. rhamni</i>	73	0.80	35	85.7	0.68	0.007
<i>I. lathonia</i>	58	0.76	30	85.7	–0.63	0.017
<i>L. camilla</i>	16	0.84	6	78.6	0.53	0.049
<i>L. dispar</i>	6	0.93	2	64.3	–0.17	0.571
<i>L. maera</i>	7	0.63	1	92.9	0.45	0.109
<i>L. megera</i>	78	0.78	36	71.4	–0.60	0.023
<i>L. phlaeas</i>	42	0.79	17	71.4	0.04	0.897
<i>L. sinapis</i>	71	0.76	118	100.0	0.03	0.911
<i>L. tityrus</i>	54	0.79	10	42.9	0.30	0.303
<i>M. athalia</i>	23	0.78	48	64.3	0.25	0.392
<i>M. aurelia</i>	24	0.77	5	21.4	0.07	0.800
<i>M. cinxia</i>	15	0.82	13	57.1	–0.13	0.646
<i>M. diamina</i>	21	0.89	3	85.7	0.24	0.414
<i>M. galathea</i>	93	0.77	805	100.0	0.17	0.557
<i>M. jurtina</i>	123	0.78	857	100.0	–0.01	0.976
<i>N. c-album</i>	28	0.81	13	64.3	0.23	0.432
<i>N. io</i>	74	0.78	58	92.9	–0.01	0.976
<i>N. polychloros</i>	18	0.76	10	50.0	–0.25	0.380
<i>N. quercus</i>	13	0.90	2	85.7	0.25	0.382
<i>N. urticae</i>	88	0.78	79	100.0	0.17	0.568
<i>O. sylvanus</i>	43	0.77	120	78.6	–0.26	0.364
<i>P. aegeria</i>	44	0.80	17	57.1	0.26	0.376
<i>P. bellargus</i>	33	0.86	1	92.9	–0.24	0.407
<i>P. brassicae</i>	69	0.76	56	92.9	–0.21	0.467
<i>P. coridon</i>	68	0.79	335	28.6	0.05	0.856
<i>P. icarus</i>	119	0.78	226	100.0	–0.02	0.958
<i>P. machaon</i>	56	0.75	21	85.7	0.19	0.526
<i>P. malvae</i>	43	0.83	24	85.7	0.38	0.186
<i>P. napi</i>	98	0.77	103	100.0	–0.18	0.542
<i>P. rapae</i>	91	0.76	512	100.0	–0.34	0.233
<i>S. pruni</i>	23	0.76	16	78.6	0.13	0.664
<i>S. sertorius</i>	31	0.81	3	21.4	0.03	0.909
<i>T. acteon</i>	55	0.79	2	21.4	0.36	0.205
<i>T. betulae</i>	12	0.56	12	57.1	0.30	0.297
<i>T. lineola</i>	51	0.76	419	100.0	0.06	0.834
<i>T. sylvestris</i>	59	0.76	302	85.7	0.20	0.502
<i>V. atalanta</i>	46	0.76	16	78.6	–0.06	0.848
<i>V. cardui</i>	68	0.77	12	57.1	0.33	0.247

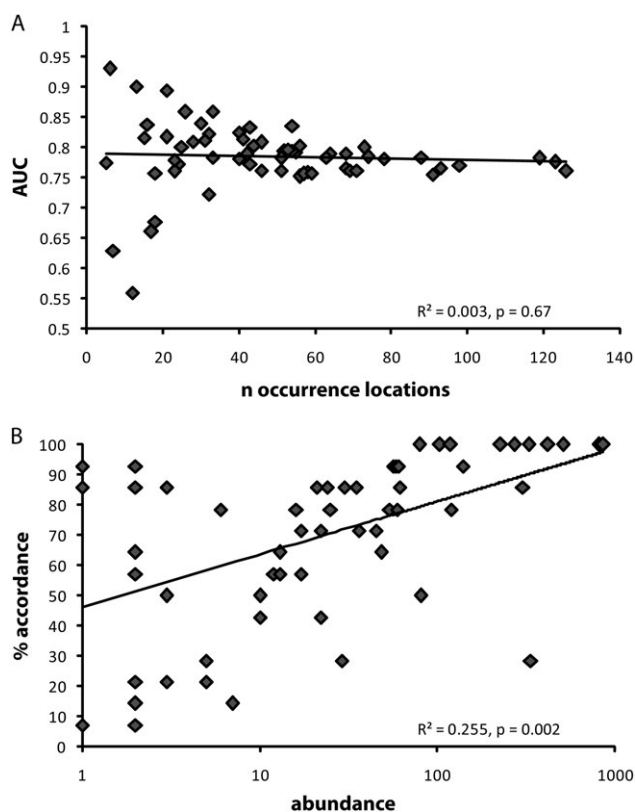


Fig. 2. The relationships between (A) the AUC values (Area Under the receiver operation characteristic Curve, AUC) and number of occurrences and (B) % accuracy (expected / observed presence) and observed abundance.

butterfly species, on average, were not better predicted by the model than local and rare taxa. However, the variance of the predictive power was high for species for which there were few records of occurrence (Fig. 2), meaning that the degree of uncertainty of the model's predictions for a species increased as the numbers of records of occurrence of this species decreased.

The observed butterfly species occurrence (i.e. presence/absence) along the 14 transects differed significantly from the model's predictions for the corresponding grid cells ($\chi^2 = 47.08$, $df = 1$, $p < 0.0001$). Thus, the model was not able to predict the presence or absence of butterfly species along transects. The percentage of congruence between expected (i.e. modelled) and observed presence was highest for common species that were abundant in the area (e.g. *M. jurtina*, *P. rapae*, *P. icarus* and *C. pamphilus*) and decreased from species that were abundant to those that were scarce (Fig 2b, $r^2 = 0.255$, $p = 0.002$). However, the occurrence of a few rather rare species (e.g. *H. comma*, *A. iris*, *M. diamina* and *M. aurelia*) was well predicted, as the model predicted mostly species absence. For most species there was no correlation between abundance and the model's prediction tested using spearman rank correlation. However, for five species (i.e. *C. arcania*, *C. pamphilus*, *C. rubi*, *G. rhamni* and *L. camilla*) there are significant correlations with the amount of explained variance ranging from 28 to 52% (Table 1). Thus, the model can be used to predict the local

abundance of these species. On the other hand, the correlation was significantly negative for *I. lathonia* and *L. megera*, and therefore it is not possible to use this model for predicting the local abundance of these two species.

DISCUSSION

The use of bioclimatic data for modelling the distributions of species has recently become more common in ecological research (e.g. Franklin, 2009). Some studies even link predictions derived from SDMs with other ecological patterns like productivity (Brambilla & Ficetola, 2012) or species abundance (Pearce & Ferrier, 2001; VanDerWal, 2009; Huntley et al., 2012). However, there are no such studies on invertebrates or on the use of state-of-the-art variable sets for such analyses. Therefore, this study aimed to determine whether species distribution models (SDMs) based on fine-scale bioclimatic variables can be used to forecast species specific abundance patterns for butterflies in a heterogeneous landscape.

In general, both the simple presence/absence predictions and observations for specific sites and more specifically the forecast of species abundance and that of the prediction for the corresponding grid cells where transects were located differed greatly. Consequently, our data support, with a few exceptions, the contention that species abundance of butterflies cannot be predicted using models based on bioclimatic variables recorded at this scale in central Europe. Even a reliable prediction of presence or absence of a species was impossible. Thus, it is likely that fine-scale bioclimatic data with a resolution of 30 arc seconds is too coarse, in terms of representing landscape features that might be responsible for species occurrence (Brambilla et al., 2009; Cord & Rödder, 2011). Moreover, depending on the geographical extent of the study, the use of variables at this resolution might blur model predictions. Thus, local factors might have masked the effects of the climatic variables used.

It is evident, however, that the AUC values, which are measures of model quality, indicated a good fit in most cases (Table 1, Fig 2a). The use of this statistic for discriminating between models has been often criticized in the past (e.g. Lobo et al., 2008; Jimenez-Valverde, 2011), but because of the lack of an alternative measure (Baldwin, 2009, but see Hijmans, 2012) it is still widely used in niche modelling studies. In this study, the AUC values did not reflect the presence/absence situation of butterfly species at the study sites investigated, because the factors included in the model do not reflect the finer scale variation (see below) that is important in determining the distribution of the butterflies. Therefore, we recommend careful and critical use of AUC values when they alone are used to determine the goodness of fit of a model's predictions with reality.

Although mesoclimatic conditions strongly influence regional distribution patterns, their importance at finer (i.e. local) scales is considerably reduced as other factors become the key determinants of the occurrence and abundance of species. Structural and ecological parameters are especially important at the micro-scale, as butterflies

often need highly complex habitat conditions for oviposition, larval development, growth of larval food-plants and imagoes' nectar-plants (e.g. Settele et al., 1999; Asher et al., 2001). These conditions are often strongly influenced by the nature of the soil and the use of the land by humans (Schmitt & Rákósy, 2007; Dover & Settele, 2009). It is not possible to include these parameters in SDMs as such fine scale data is not available for large areas (but see Cord & Rödder, 2011; Brambilla & Ficetola, 2012; Pfeifer et al., 2012). In this context, it is also necessary to emphasize the importance of the microclimatic conditions that are markedly influenced by human activities, such as the construction of traditional stone walls, ecologically rich waysides and slopes, hedgerows and small eroded patches. All of these provide suitable conditions for butterflies at places not predicted by the SDMs because they considerably increase temperature and modulate humidity at a very local scale of some few square metres. This is far below the spatial resolution of the climatic data currently used in model construction so that it is necessary to determine the abundance of butterflies and other invertebrates with complex habitat requirements by means of time- and labour-intensive field work, and this likely to continue to be the case for the foreseeable future.

ACKNOWLEDGEMENTS. Permissions to work and collect butterflies in protected areas in Rhineland-Palatinate were granted by the Struktur- und Genehmigungsdirektion Nord (Koblenz, Germany). We thank M. Weitzel for assisting with field surveys in 2011. Financial support was provided for KJF by the Friedrich Ebert Foundation and the Ministry for Environment, Agriculture, Viticulture, Food and Forests Rhineland-Palatinate, and for JOE by the German Federal Environmental Foundation. We thank M. Konvicka, M. Brambilla, A.F.G. Dixon and an anonymous referee for helpful comments on an earlier draft of the manuscript and language corrections.

REFERENCES

- ARAÚJO M.B. & PETERSON A.T. 2012: Uses and misuses of bioclimatic envelope modelling. — *Ecology* **93**: 1527–1539.
- ARAÚJO M.B., PEARSON R.G., THUILLER W. & ERHARD M. 2005: Validation of species-climate impact models under climate change. — *Glob. Change Biol.* **11**: 1504–1513.
- ARAÚJO M.B., ALAGADOR D., CABEZA M., NOGUÉS-BRAVO D. & THUILLER W. 2011: Climate change threatens European conservation areas. — *Ecol. Lett.* **14**: 484–492.
- BALDWIN R.A. 2009: Use of maximum entropy modeling in wildlife research. — *Entropy* **11**: 854–866.
- BINK F.A. 1992: *Ecologische Atlas van de Dagvlinders van Noordwest-Europa*. Schuyt & Co. Uitgevers en Importeurs, Haarlem, 512 pp.
- BRAMBILLA M. & FICETOLA G.F. 2012: Species distribution models as a tool to estimate reproductive parameters: a case study with a passerine bird species. — *J. Anim. Ecol.* **81**: 781–787.
- BRAMBILLA M., CASALE F., BERGERO V., CROVETTO G.M., FALCO R., NEGRI I., SICCARDI P. & BOGLIANI G. 2009: GIS-models work well, but are not enough: Habitat preferences of *Lanius collurio* at multiple levels and conservation implications. — *Biol. Conserv.* **142**: 2033–2042.
- CORD A. & RÖDDER D. 2011: Inclusion of habitat availability in species distribution models through multi-temporal remote-sensing data? — *Ecol. Appl.* **21**: 3285–3298.
- DOVER J. & SETTELE J. 2009: The influences of landscape structure on butterfly distribution and movement: a review. — *J. Insect Conserv.* **13**: 3–27.
- ELITH J., GRAHAM C.H., ANDERSON R.P., DUDÍK M., FERRIER S., GUISAN A., HUMANS R.J., HUETTMANN F., LEATHWICK J.R., LEHMANN A., LI J., LOHMANN L.G., LOISELLE B.A., MANION G., MORITZ C., NAKAMURA M., NAKAZAWA Y., OVERTON J.M.M., PETERSON A.T., PHILLIPS S.J., RICHARDSON K., SCACHETTI-PEREIRA R., SHAPIRE R.E., SOBERÓN J., WILLIAMS S., WISZ M.S. & ZIMMERMANN N.E. 2006: Novel methods improve prediction of species' distributions from occurrence data. — *Ecography* **29**: 129–151.
- ELITH J., PHILLIPS S.J., HASTIE T., DUDÍK M., CHEE Y.E. & YATES C.J. 2011: A statistical explanation of MaxEnt for ecologists. — *Divers. Distrib.* **17**: 43–57.
- FICETOLA G.F., THUILLER W. & MIAUD C. 2007: Prediction and validation of the potential global distribution of a problematic alien invasive species – the American bullfrog. — *Divers. Distrib.* **13**: 476–485.
- FRANKLIN J. 2009: *Mapping Species Distributions: Spatial Inference and Prediction*. Cambridge University Press, Cambridge, 338 pp.
- GASTON K.J. 2003: *The Structure and Dynamics of Geographic Ranges*. Oxford University Press, Oxford, 280 pp.
- HABEL J.C., RÖDDER D., STEFANO S., MEYER M. & SCHMITT T. 2010: Strong genetic cohesiveness between Italy and North Africa in four butterfly species. — *Biol. J. Linn. Soc.* **99**: 818–830.
- HABEL J.C., RÖDDER D., SCHMITT T. & NÈVE G. 2011: Global warming will affect the genetic diversity and uniqueness of *Lycaena helle* populations. — *Glob. Change Biol.* **17**: 194–205.
- HEIKKINEN R.K., LUOTO M., ARAÚJO M.B., VIRKKALA R., THUILLER W. & SYKES M.T. 2006: Methods and uncertainties in bioclimatic envelope modelling under climate change. — *Prog. Phys. Geog.* **30**: 751–777.
- HERNANDEZ P.A., GRAHAM C.H., MASTER L.L. & ALBERT D.L. 2006: The effect of sample size and species characteristics on performance of different species distribution modelling methods. — *Ecography* **29**: 773–785.
- HIJMANS R.J. 2012: Cross-validation of species distribution models: removing spatial sorting bias and calibration with a null model. — *Ecology* **93**: 679–688.
- HIJMANS R.J., CAMERON S.E., PARRA J.L., JONES P.G. & JARVIS A. 2005: Very high resolution interpolated climate surfaces for global land areas. — *Int. J. Climatol.* **25**: 1965–1978.
- HOBBS N.T. & HANLEY T.A. 1990: Habitat evaluation: do use/availability data reflect carrying capacity? — *J. Wildlife Manag.* **54**: 515–522.
- HUNTLEY B., ALTWEGG R., BARNARD P., COLLINGHAM Y.C. & HOLE D.G. 2012: Modelling relationships between abundance patterns and climate. — *Global Ecol. Biogeogr.* **21**: 668–681.
- JIMÉNEZ-VALVERDE A. 2012: Insights into the area under the receiver operating characteristic curve (AUC) as a discrimination measure in species distribution modelling. — *Glob. Ecol. Biogeogr.* **21**: 498–507.
- KELLER C.J., BROWN J.D. & KARR J.R. 1986: What is habitat suitability and how should it be measured? In Verner J., Morrison M.L. & Ralph C.J. (eds): *Wildlife 2000: Modelling Wildlife-habitat Relationships of Terrestrial Vertebrates*. University of Wisconsin Press, Madison, pp. 476–488.

- LIU C., BERRY P.M., DAWSON T.P. & PEARSON R.G. 2005: Selecting thresholds of occurrence in the prediction of species distributions. — *Ecography* **28**: 385–393.
- LOBO J.M., JIMÉNEZ-VALVERDE A. & REAL R. 2008: AUC: a misleading measure of the performance of predictive distribution models. — *Glob. Ecol. Biogeogr.* **17**: 145–151.
- MURPHY H.T., VANDERWAL J. & LOVETT-DOUST J. 2006: Distribution of abundance across the range in eastern North American trees. — *Global Ecol. Biogeogr.* **15**: 63–71.
- NIELSEN S.E., JOHNSON C.J., HEARD D.C. & BOYCE M.S. 2005: Can models of presence-absence be used to scale abundance? two case studies considering extremes in life history. — *Ecography* **28**: 197–208.
- PEARCE J.L. & BOYCE M.S. 2006: Modelling distribution and abundance with presence only-data. — *J. Appl. Ecol.* **43**: 405–412.
- PEARCE J. & FERRIER S. 2001: The practical value of modelling relative abundance of species for regional conservation planning: a case study. — *Biol. Conserv.* **98**: 33–43.
- PEARSON R.G., RAXWORTHY C.J., NAKAMURA M. & PETERSON A.T. 2007: Predicting species distributions from small numbers of occurrence records, a test case using cryptic geckos in Madagascar. — *J. Biogeogr.* **34**: 102–117.
- PFEIFER M., DISNEY M., QUAIPE T. & MARCHANT R. 2012: Terrestrial ecosystems from space: a review of earth observation products for macroecology applications. — *Global Ecol. Biogeogr.* **21**: 603–624.
- PHILLIPS S.J., ANDERSON R.P. & SCHAPIRE R.E. 2006: Maximum entropy modeling of species geographic distributions. — *Ecol. Model.* **190**: 231–259.
- POLLARD E. & YATES T.J. 1993: *Monitoring Butterflies for Ecology and Conservation. The British Butterfly Monitoring Scheme*. Chapman & Hall, London, 292 pp.
- R DEVELOPMENT CORE TEAM 2012: *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>
- RÖDDER D., SCHMIDTLEIN S., VEITH M. & LÖTTERS S. 2009: Alien invasive slider turtle in unpredicted habitat: A matter of niche shift or of predictors studied? — *PlosOne* **4**(11): e7843.
- RODRIGUEZ J.P., BROTONS L., BUSTAMANTE L. & SEOANA J. 2007: The application of predictive modeling of species distribution to biodiversity conservation. — *Divers. Distrib.* **13**: 243–251.
- SCHMITT T. & RÁKOSY L. 2007: Changes of traditional agrarian landscape and their conservation implications: a case study of butterflies in Romania. — *Divers. Distrib.* **13**: 855–862.
- SETTELE J., FELDMANN R. & REINHARDT R. 1999: *Die Tagfalter Deutschlands-Ein Handbuch für Freilandökologen, Umweltplaner und Naturschützer*. Ulmer, Stuttgart, 452 pp.
- SIQUEIRA M.F. DE, DURIGAN G., DE MARCO JR. P. & PETERSON A. T. 2009: Something from nothing: Using landscape similarity and ecological niche modeling to find rare plant species. — *J. Nat. Conserv.* **17**: 25–32.
- SOBERON J. & NAKAMURA M. 2009: Niches and distributional areas: Concepts, methods, and assumptions. — *Proc. Natl. Acad. Sci. U.S.A.* **106**: 19644–19650.
- SWETS K. 1988: Measuring the accuracy of diagnostic systems. — *Science* **240**: 1285–1293.
- VANDERWAL J., SHOO L.P., JOHNSON C.N. & WILLIAMS S.E. 2009: Abundance and the environmental niche: Environmental suitability estimated from niche models predicts the upper limit of local abundance. — *Am. Nat.* **174**: 282–291.
- WENZEL M., SCHMITT T., WEITZEL M. & SEITZ A. 2006: The severe decline of butterflies on western calcareous grasslands during the last 30 years: A conservation problem. — *Biol. Conserv.* **128**: 542–552.
- WISZ M.S., HUMANS R.J., PETERSON A.T., GRAHAM C.H., GUISAN A. & NPSDW GROUP 2008: Effects of sample size on the performance of species distribution models. — *Divers. Distrib.* **14**: 763–773.

Received June 11, 2012; revised and accepted October 15, 2012