ECOGRAPHY

Research

Evaluating the ecological realism of plant species distribution models with ecological indicator values

Marjon Hellegers, Wim A. Ozinga, Arjen van Hinsberg, Mark A. J. Huijbregts, Stephan M. Hennekens, Joop H. J. Schaminée, Jürgen Dengler and Aafke M. Schipper

M. Hellegers (https://orcid.org/0000-0002-8696-4921) ✉ (marjon.hellegers@pbl.nl), A. van Hinsberg (https://orcid.org/0000-0002-4452-9892) and A. M. Schipper (https://orcid.org/0000-0002-5667-0893), PBL Netherlands Environmental Assessment Agency, the Hague, the Netherlands. – W. A. Ozinga (https://orcid.org/0000-0002-6369-7859), S. M. Hennekens (https://orcid.org/0000-0003-1221-0323) and J. H. J. Schaminée (https://orcid.org/0000-0002-0416-3742), Wageningen Environmental Research, Wageningen, the Netherlands. – M. A. J. Huijbregts (https://orcid.org/0000-0002-7037-680X), MH and AMS, Dept of Environmental Science, Radboud Univ., Nijmegen, the Netherlands. – J. Dengler (https://orcid.org/0000-0003-3221-660X), Vegetation Ecology Group, Inst. of Natural Resource Sciences (IUNR), Zurich Univ. of Applied Sciences (ZHAW), Wädenswil, Switzerland, and Plant Ecology Group, Bayreuth Center of Ecology and Environmental Research (BayCEER), Univ. of Bayreuth, Bayreuth, Germany, and German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany.

Species distribution models (SDMs) are routinely applied to assess current as well as future species distributions, for example to assess impacts of future environmental change on biodiversity or to underpin conservation planning. It has been repeatedly emphasized that SDMs should be evaluated based not only on their goodness of fit to the data, but also on the realism of the modeled ecological responses. However, possibilities for the latter are hampered by limited knowledge on the true responses as well as a lack of quantitative evaluation methods. Here we compared modeled niche optima obtained from European-scale SDMs of 1476 terrestrial vascular plant species with empirical ecological indicator values indicating the preferences of plant species for key environmental conditions. For each plant species we first fitted an ensemble SDM including three modeling techniques (GLM, GAM and BRT) and extracted niche optima for climate, soil, land use and nitrogen deposition variables with a large explanatory power for the occurrence of that species. We then compared these SDM-derived niche optima with the ecological indicator values by means of bivariate correlation analysis. We found weak to moderate correlations in the expected direction between the SDM-derived niche optima and ecological indicator values. The strongest correlation occurred between the modeled optima for growing degree days and the ecological indicator values for temperature. Correlations were weaker for SDM-derived niche optima with a more distal relationship to ecological indicator values (notably precipitation and soil moisture). Further, correlations were consistently highest for BRT, followed by GLM and GAM. Our method gives insight into the ecological realism of modeled niche optima and projected core habitats and can be used to improve SDMs by making a more informed selection of environmental variables and modeling techniques.

Keywords: BioScore, ecological niche model, Ellenberg indicator values, model selection, scale dependency, validation

Introduction

Global biodiversity is currently declining at an unusually high rate (Butchart et al. 2010, Barnosky et al. 2011, Tittensor et al. 2014). This brings about a clear demand for quantitative models able to project changes in biodiversity in response to anthropogenic pressures as well as policy measures designed to counteract the decline (Pereira et al. 2010, Harfoot et al. 2014). Species distribution models (SDMs) are quantitative relationships between the occurrence of a species and a set of environmental factors (Elith and Leathwick 2009). These models help to reveal and understand the habitat preferences of a species and can be used to project its distribution as a function of changing environmental conditions, which in turn provides vital information in order to underpin biodiversity policy (Araújo et al. 2011, Visconti et al. 2016, Morán-Ordóñez et al. 2017).

SDMs have been applied to assess the impact of changes in environmental conditions, such as climate and land use, on a wide range of taxonomic groups, including plants, insects and vertebrates (Thuiller et al. 2004, Settele et al. 2008, Morán-Ordóñez et al. 2017). Confidence in projections of SDMs is typically achieved by means of testing against left-out data (Araújo et al. 2011, Dubuis et al. 2011, Mitchell et al. 2017). However, validation against left-out data may provide overly optimistic estimates of the models' predictive ability due to a lack of independence of the testing data set (Heikkinen et al. 2012, Rapacciuolo et al. 2012, Bahn and McGill 2013). Therefore, it has been repeatedly emphasized that SDMs should be evaluated not only based on their goodness of fit to data, but also based on their ecological realism (Austin et al. 2006, Elith and Leathwick 2009, Merow et al. 2014). The selection of an ecologically relevant set of environmental variables is of crucial importance (Fourcade et al. 2017), in particular whether these are direct or indirect predictors of species occurrence and whether they are appropriate in relation to the grain size, extent and species of concern (Pearson and Dawson 2003, Austin 2007, Elith and Leathwick 2009). For example, it has been suggested that species' distributions at large scales are primarily defined by climatic factors, whereas land use and soil properties become increasingly dominant at smaller extent and finer grain (Pearson and Dawson 2003). After the SDMs have been fitted, the ecological realism of the modeled species' responses to the selected predictors should be evaluated (Austin et al. 2006). However, possibilities are limited by a lack of knowledge on the true responses and a lack of quantitative approaches to evaluate the modeled responses (Austin et al. 2006, Austin 2007).

In this study, we use independent data on species niche optima to evaluate the ecological realism of SDMs. To that end, we evaluate to what extent SDMs of 1476 vascular plant species in Europe adequately capture niche optima based on a comparison with independent ecological indicator values. This increases our understanding of the ability of SDMs to accurately predict core habitat areas, which is key in studies aiming to project impacts of environmental change and in conservation planning. We used the modeled response curves from the SDMs to extract niche optima for various relevant habitat variables representative of climate (minimum temperature, annual precipitation, growing degree days), pollution (nitrogen deposition), soil properties (carbon, silt and clay content, coarse fragments, pH and cation exchange capacity) and land use or land cover (arable land, pastures, heterogeneous agriculture, scrubs, forests, open spaces and inland wetlands). For each environmental variable and species, we extracted the niche optimum from the SDM (hereafter, modeled indicator value; MIV) as the value of that variable corresponding with the highest probability of occurrence. We then compared the MIVs with independent empirical ecological indicator values (EIVs) indicating the preferences of the plant species for various environmental conditions, expressed as realized niche optima on ordinal scales (Landolt 1977, Ellenberg et al. 1991, 2001). EIVs have been derived for many vascular plant species occurring in Europe, based on field observations of species co-occurrence patterns, in-situ measured environmental factors and occasional experiments, and are available for various key habitat factors (temperature, continentality, soil moisture, soil nutrient availability, soil acidity, light availability and salinity) (Hill et al. 1999, Ellenberg et al. 2001, Pignatti 2005).

Upfront, we expected to find relationships between MIVs and EIVs for pairs of variables representative of the same type of site conditions that are key to plants. More specifically, we expected the MIVs for growing degree days, annual precipitation, nitrogen deposition and soil pH to be positively related to the EIVs for temperature, soil moisture, soil nutrient availability and soil acidity, respectively. Similarly, we expected negative relationships between the MIVs of mean temperature of the coldest month and forest cover and the EIVs of continentality and light availability, respectively. Because of the extent (Europe) and resolution (1 km) of our SDMs, we expected that the relationships would be stronger for the climate- and landscape-related EIVs (temperature, continentality and light) than for the soil-related EIVs (soil moisture, soil nutrient availability and soil acidity), as soil properties are expected to become important for species distributions mostly at extents below 10 km and fine-grained resolution (Pearson and Dawson 2003). Further, we expected stronger relationships if the environmental variable in the SDM was more similar or more directly related to the EIV (Austin 2007, Petitpierre et al. 2017). From this perspective, we expected stronger relationships between the MIVs for growing degree days, soil pH and forest cover and the EIVs for temperature, soil acidity and light, respectively, than for the other three pairs of MIVs and EIVs. Overall, we therefore expected the strongest relationships for growing degree days with temperature and forest cover with light (i.e. representative scale and direct relationships), and the weakest relationships for annual precipitation with soil moisture and nitrogen deposition with soil nutrients (i.e. small-scale variables and more distal relationships).

Material and methods

Species

In order to limit computation time, we first retrieved a representative selection of terrestrial vascular plant species in Europe. To that end, we selected species characteristic of distinct, terrestrial EUNIS habitat types that are included in the European Red List of Habitats (Janssen et al. 2016), and 10 additional anthropogenic habitat types derived from the EuroVegChecklist (Supplementary material Appendix 1 Table A1; Mucina et al. 2016). Lists of characteristic species are available for the EUNIS habitat types of forests (37 types), heathlands, shrub and tundra (38 types) and grasslands (50 types; Schaminée et al. 2013, 2014, 2015, 2016). To identify species characteristic of the remaining 66 EUNIS habitat types of coastal habitats, mires and bogs and sparsely vegetated habitats and of the 10 anthropogenic habitat types, we used species lists reported by Hennekens et al. (2017). The selection procedure yielded a total of 4541 vascular plant species (Supplementary material Appendix 1 Text section A1).

Vegetation plots

We retrieved a total of 1 310 202 vegetation plots from the European Vegetation Archive (EVA; Chytrý et al. 2016). An overview of datasets included is provided in the Supplementary material Appendix 1 Table A2. We excluded plots according to the following criteria.

1) plots located in 'inland waters', 'marine waters' and 'maritime wetlands', according to the Corine Land Cover map (Hazeu et al. 2008).

2) plots recorded before 1990, to minimize temporal differences between the plots and the environmental variables (see the section on environmental variables).

3) plots with a known spatial uncertainty larger than 1 km.

4) plots with missing values for one or more environmental variables.

With these selection criteria we retained 45% of the vegetation plots retrieved from EVA (= 599 038 plots) for further analysis (Supplementary material Appendix 1 Fig. A1). As we extracted the data in 2017, plots added after 2016 were not included.

Ecological indicator values

We used a dataset of ecological indicator values (EIVs) of temperature, continentality, soil moisture, soil nutrient availability, soil acidity, light and salinity compiled by Louette et al. (2010), which was in turn based on the lists with ecological indicator values of plant species for central European (Ellenberg et al. 2001), Switzerland (Landolt 1977), eastern Germany (Frank et al. 1990), Great Britain (Hill et al. 1999) and Italy (Pignatti 2005). Louette et al. (2010) harmonized the taxon names by adopting the names from the SynBioSys Taxon Database, which are also used in

EVA (Chytrý et al. 2016). The scale used by Landolt (1977), which runs from 1 to 5, was linearly extrapolated to the general scale of 1 to 9 and 1 to 12 for moisture (Louette et al. 2010). The extended range of 1 to 12 that Pignatti (2005) used for the indicator values for light and temperature was retained, to reflect the wider range of climatic conditions than in the areas represented by the other lists, especially in the warmer southern part of Italy, where solar radiation is more direct (Louette et al. 2010). Per species and environmental condition, the EIV is given as an integer that represents the optimum of the species on the corresponding ordinal scale. For species with different values of EIV in multiple lists, we calculated a mean EIV across the lists.

Environmental variables

We selected a set of environmental variables covering climate, soil, land cover and pollution (nitrogen deposition) (Table 1). For climate, we selected a set of variables which are known to pose physiological limitations on species distributions in Europe by imposing energy or water deficiency (Whittaker et al. 2007, Araújo et al. 2011). These are mean temperature of the coldest month, total annual precipitation, annual growing degree days and a moisture index (Araújo et al. 2011). We calculated the moisture index as the sum of the monthly differences between precipitation and potential evapotranspiration (Heikkinen et al. 2010). To calculate monthly potential evapotranspiration (PET), we followed Skov and Svenning (2004):

$$PET = 58.93 \times T(above\,0°C)/12 \qquad (1)$$

where T(above 0°C) represents the monthly mean temperature values above 0°C. We obtained the climatic variables from monthly mean temperature and precipitation values for the time period of 1979–2013 retrieved from the CHELSA data set (Karger et al. 2017). Soil variables included organic carbon, silt, sand and clay content, bulk density, volume of coarse fragments, pH and cation exchange capacity, which we retrieved from the SoilGrids dataset (Hengl et al. 2017). We selected values from the top 5 cm of the soil. Land cover was represented by the non-urban land cover types (classification level 2) of the Corine land cover map for the year 2000 (Hazeu et al. 2008). We aggregated the land cover class of permanent crops, where less than 4000 vegetation plots were recorded, with the class of arable land (Table 1). Finally, we used nitrogen deposition in the year 2013 as a proxy for nitrogen input in the soil (Fagerli et al. 2015).

We resampled all continuous environmental variables to a 1 km resolution using the mean value, based on the ETRS89 Lambert Azimuthal Equal-Area projection. For the land cover variables we quantified the fraction of each type within the 1 km grid cell. In order to minimize collinearity while keeping both energy and water-related climate variable, we selected only variables with a variance inflation factor (VIF) below 5 for inclusion in the SDMs. This led to

Table 1. Environmental variables included in the models.

Variable	Unit	Data source	Original resolution	Reference
Mean temperature of the coldest month	°C	Chelsa	30 arc sec	Karger et al. (2017)
Total annual precipitation	mm	Chelsa	30 arc sec	Karger et al. (2017)
Annual growing degree days	°C	Chelsa	30 arc sec	Karger et al. (2017)
Moisture index	mm	Chelsa	30 arc sec	Karger et al. (2017)
Arable land and permanent crops (CLC-codes 211–223)	ha ha^{-1}	Corine land cover	100 m	Hazeu et al. (2008)
Pastures (CLC-code 231)	ha ha^{-1}	Corine land cover	100 m	Hazeu et al. (2008)
Heterogeneous agricultural areas (CLC-code 241–244)	ha ha^{-1}	Corine land cover	100 m	Hazeu et al. (2008)
Scrub and/or herbaceous vegetation associations (CLC-code 321–324)	ha ha^{-1}	Corine land cover	100 m	Hazeu et al. (2008)
Forests (CLC-code 311–313)	ha ha^{-1}	Corine land cover	100 m	Hazeu et al. (2008)
Open spaces with little or no vegetation (CLC-code 331–335)	ha ha^{-1}	Corine land cover	100 m	Hazeu et al. (2008)
Inland wetlands (CLC-code 411–412)	ha ha^{-1}	Corine land cover	100 m	Hazeu et al. (2008)
Organic carbon content	g kg^{-1}	Soil Grids	1 km	Hengl et al. (2017)
Clay content	weight %	Soil Grids	1 km	Hengl et al. (2017)
Silt content	weight %	Soil Grids	1 km	Hengl et al. (2017)
Sand content	weight %	Soil Grids	1 km	Hengl et al. (2017)
Bulk density (fine earth)	kg m^{-3}	Soil Grids	1 km	Hengl et al. (2017)
Coarse fragments volumetric	volumetric %	Soil Grids	1 km	Hengl et al. (2017)
Cation exchange capacity of soil (CEC)	cmolc kg^{-1}	Soil Grids	1 km	Hengl et al. (2017)
Soil pH×10 in H2O	pH	Soil Grids	1 km	Hengl et al. (2017)
Nitrogen deposition	mg m^{-2}	EMEP	0.1 degrees	Fagerli et al. (2015)

the omission of the moisture index, sand content and bulk density. The first was positively correlated with total annual precipitation. Sand content had a negative correlation with clay and silt content. Bulk density was negatively correlated to the cation exchange capacity and carbon content. We used the remaining set of 17 variables to fit the SDMs and to retrieve modeled indicator values. We reduced the skewness of the distributions of total annual precipitation, annual growing degree days and nitrogen deposition by means of a square root transformation.

Fitting the SDMs

For each species we fitted an SDM based on three modeling techniques (generalized linear model, GLM; generalized additive model, GAM; boosted regression trees, BRT) that are frequently used and have proven to perform well (Rapacciuolo et al. 2012). Moreover, the range of complexity of the modeling techniques enabled us to capture simple linear or unimodal as well as skewed and more complex responses (Merow et al. 2014). We fitted the SDMs with the BIOMOD2 package using default settings (ver. 3.3–7; Thuiller et al. 2016). In order to reduce spatial bias and pseudo-replication, we randomly selected only one plot where the species was present per 1 km grid cell. We then selected species with at least 100 presence points, thus ensuring about five presences per environmental variable after setting aside 20% of the observations for model evaluation (17 variables×5 observations/0.8 = 106

observations, which we rounded down to 100). This left us with 1500 species for which we had enough presence points as well as EIVs available. Because of the large number of available vegetation plots (599 038 plots), we made a selection of absences for the model fitting, in order to limit calculation time. For each species we sampled absences from plots where the species was not recorded. We sampled the absences randomly, such that for each species there was not more than one absence record per 1 km grid cell, equal to the selection of the presence records. For the GLM and GAM we used at least 10 000 absence records and for the BRT 1000 (Barbet-Massin et al. 2012). When the number of presence records was larger than either 1000 or 10 000, we sampled a number of absence records equal to the number of presences. Presences and absences were given an equal weight in the model fitting. We calibrated the models using a single random sample of 80% of the presences and absences and evaluated the models against the remaining 20% of the data. Then, we build for each species an ensemble model based on the three modeling techniques, each weighted based on the cross-validated TSS value. We discarded species if one or more of the modeling techniques did not converge, leaving 1481 of the 1500 species. We further discarded species with an ensemble model with true skill statistic (TSS) values < 0.3 or area under the receiver operating characteristic curve (AUC) values < 0.7 (Araújo et al. 2011). With these thresholds five more species were left out (Supplementary material Appendix 1 Table A3), leaving 1476 species for the final analysis.

Retrieving modeled indicator values

We retrieved species-specific response curves for each variable and modeling technique according to the Evaluation Strip method proposed by Elith et al. (2005), with the response.plot2 tool from the BIOMOD2 package (ver. 3.3–7; Thuiller et al. 2016). Response curves were obtained by varying the variable of interest across its range while fixing the other variables at their mean values across the plots where the species of concern was observed. We then calculated species-specific ensemble response curves for each variable, as averages of the response curves of the individual modeling techniques each weighted by the respective cross-validated TSS value, similar to the method to calculate the ensemble models. We further retrieved species-specific variable importance values for each variable and modeling technique, using the get_variables_importance tool from the BIOMOD2 package (ver. 3.3–7; Thuiller et al. 2016). The variable importance per modeling technique was calculated as 1 minus the Pearson's correlation between the predictions of the full model and the predictions of the model where the variable of concern was randomized. We calculated the ensemble variable importance as weighted average of the variable importance values of the individual modeling techniques, with the respective cross-validated TSS value as weight. Finally, for variables with an ensemble variable importance > 0.05, we extracted the modeled niche indicator values of each species (MIVs) from the ensemble response curve as the value of the environmental variable corresponding to the maximum probability of occurrence. In case of multiple values corresponding to the maximum occurrence probability, we calculated their median.

Data analysis

We assessed pairwise correlations between the MIVs and the EIVs by means of Spearman rank correlation analyses. To test the sensitivity of the correlations to the threshold of the ensemble variable importance, which was used to select the variables for which MIVs were extracted, we repeated the analysis with an ensemble variable importance threshold of 0.1 instead of 0.05. To test whether the results were robust to differences in EIVs for species with different values in multiple lists, we repeated the analysis based on the minimum and maximum EIV values for species across lists. To assess possible differences among modeling techniques, we repeated the analysis with MIVs retrieved from the response curves specific to each modeling techniques. We performed all modeling and data analyses in the R environment (ver. 3.5.2; <www.R-project.org>).

Results

The most important variables in the SDMs were annual growing degree days and mean temperature of the coldest month, followed by some of the soil variables (including pH and
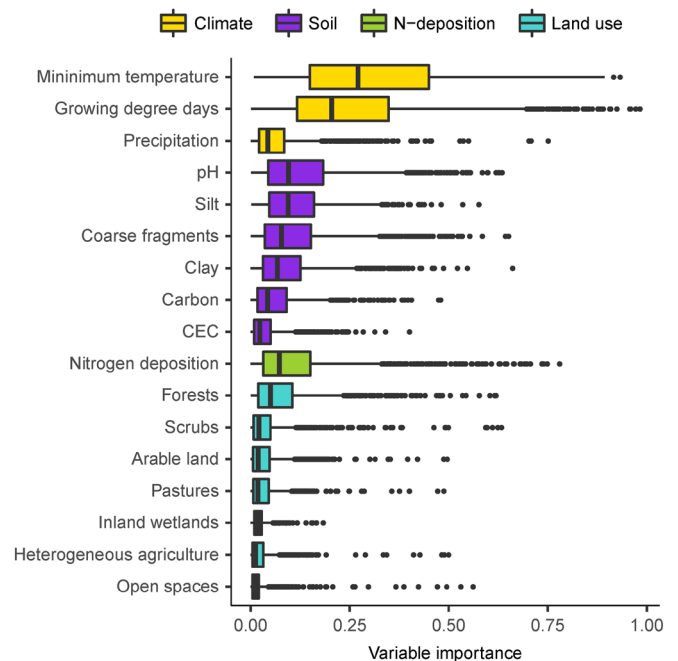

Figure 1. Boxplots of the ensemble variable importance of all variables used to fit the SDMs for all 1476 species. The boxes show the medians, the 25 and 75 percentiles and the whiskers represent 1.5 times the inter-quartile distance.

silt content) and nitrogen deposition (Fig. 1). Precipitation was clearly less important than the other climate variables. Within the land cover and land use variables, forest cover generally had the highest importance.

With regard to the directions of the relationships between MIVs and EIVs, our expectations were all confirmed, i.e. we found negative relationships of forest cover (MIV) with light (EIV) and minimum temperature (MIV) with continentality (EIV), and positive relationships for the other four pairs of MIVs and EIVs (Fig. 2). As expected, we found the strongest correlation between the MIV of annual growing degree days and the EIV of temperature ($\rho = 0.53$, $p < 0.0001$, $n = 1380$; Fig. 2a) and the weakest correlation between the MIV of total annual precipitation and the EIV of soil moisture ($\rho = 0.10$, $p = 0.0086$, $n = 656$; Fig. 2c). Correlations for the other four pairs were all between $\rho = 0.3$ and 0.4 (absolute values).

The correlations were slightly stronger when only species were included for which the variable had an importance above 0.1. This was most obvious for the correlation between forest cover (MIV) and light (EIV), which increased from $\rho = -0.39$ to -0.56 (Supplementary material Appendix 1 Fig. A2). Results were robust to the use of different EIVs for species occurring in multiple lists (Supplementary material Appendix 1 Fig. A3, A4). Among the three modeling techniques, the strongest correlations between MIVs and EIVs were obtained for BRT, followed by GLM and lastly GAM (Supplementary material Appendix 1 Fig. A5–A7).
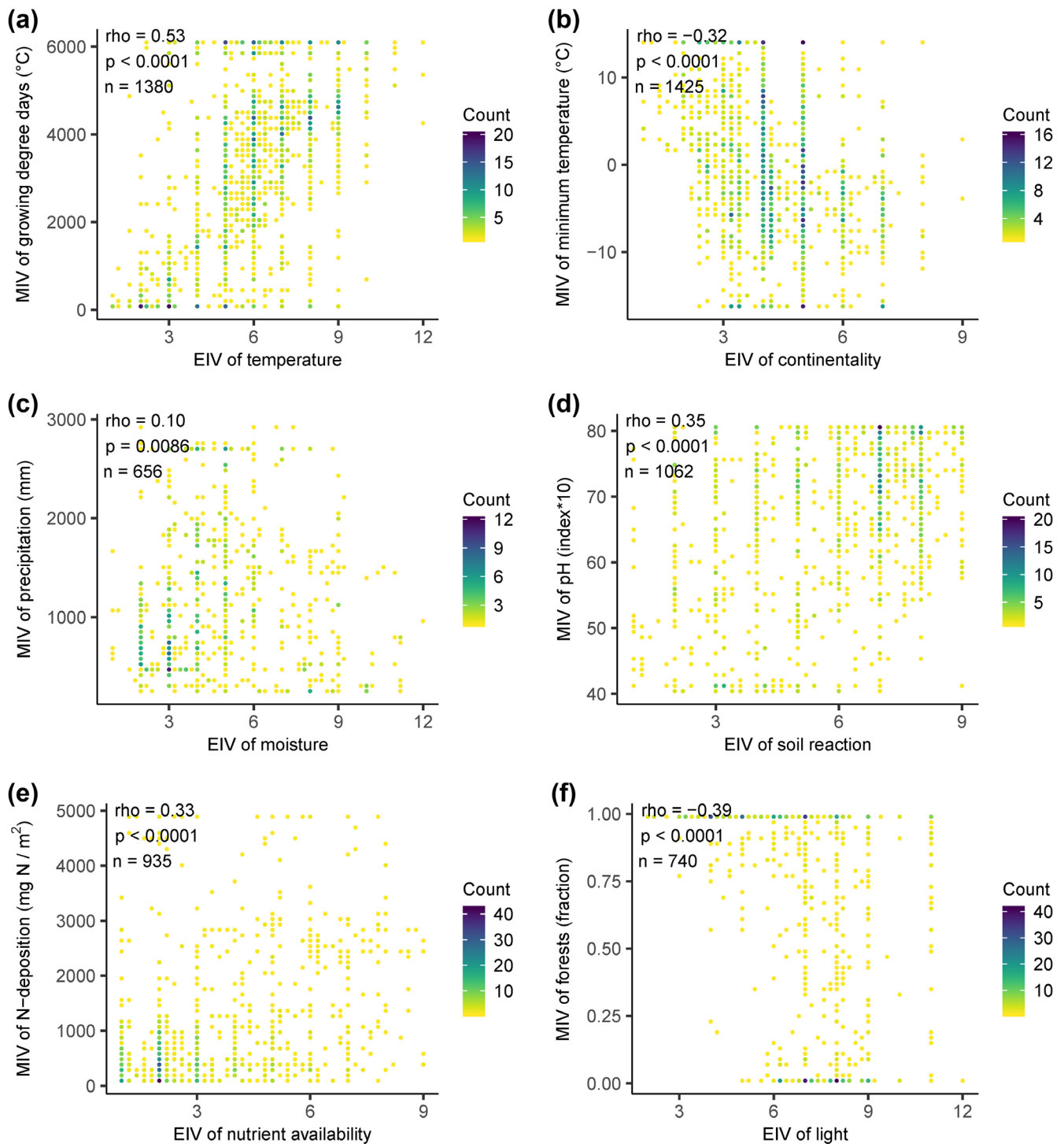
Figure 2. Bivariate correlations between species ecological indicator values (EIV) and modeled indicator values (MIV). n refers to the number of species included in the bivariate correlation and is equal to the number of species for which the variable of concern had an ensemble variable importance > 0.05.

Discussion

In the present study we assessed the ecological realism of European-scale SDMs of 1476 vascular plant species by comparing niche optima estimated by the models (modeled indicator values; MIVs) with ecological indicator values (EIVs) representative of niche optima. Although several studies have criticized EIVs, for example for the supposedly non-systematic derivation inferred from field experience, the review by Diekmann (2003) revealed that multiple studies

have shown that EIVs correlate well with measured site conditions. This indicates that EIVs can be reliably used to verify niche optima retrieved from SDMs.

Overall, the niche optima retrieved from the SDMs showed weak to moderate agreement with the independent data, whereby the directions of the six hypothesized relationships between MIVs and EIVs were all confirmed. As expected, we found the strongest correlation between the MIV of growing degree days and the EIV of temperature (Fig. 2), reflecting that temperature is one of the most important factors driving the distribution of species at large scales (Thuiller et al. 2005). The low correlation between the MIV of total annual precipitation and the EIV of soil moisture ($\rho = 0.10$) may reflect that soil moisture in wet ecosystems is largely determined by hydrological factors other than precipitation, notably the groundwater level. In ecosystems where the groundwater level is far below the root zone, soil moisture is more dependent on precipitation (Walter 1985). Indeed, we found an increase in correlation from $\rho = 0.10$ to 0.33 between the MIVs of precipitation and the EIVs of soil moisture when considering only species of dry to moist soils (Supplementary material Appendix 1 Table A4). Similarly, the relatively weak correlation between the MIV of nitrogen deposition and the EIV of soil nutrient availability ($\rho = 0.33$) reflects that soil nutrient availability is also influenced by other factors, such as management, soil and land cover type (Kooijman et al. 1998, Nissinen and Hari 1998). In addition, the EIV of soil nutrient availability is also affected by other nutrients, such as phosphorus (Ellenberg et al. 2001).

Against our expectations, we did not find the relationships between MIVs and EIVs to be stronger for climate- and landscape-related (i.e. large-scale) variables than for the soil-related (i.e. small-scale) variables. This might reflect complex multi-scale patterns of variation in the respective environmental variables. For example, soil variables such as pH can vary strongly within meters (Campbell et al. 1989, Lechowicz and Bell 1991). This fine-grain heterogeneity was neglected in the SDMs, which were based on environmental variable values averaged to a 1 km resolution. On the other hand, there are also large-scale gradients in soil acidity (Azevedo et al. 2013), which may be adequately captured by the SDMs. Similarly, temperature-related variables show not only large-scale patterns (Pearson and Dawson 2003), but also local heterogeneity (Scherrer and Körner 2011). This fine-grain heterogeneity was not included in our SDMs, which may partly explain why the correlations with the temperature-related EIVs were only weak to moderate.

We found various correlations between MIVs and EIVs that we had not expected upfront, due to remaining correlations between environmental variables after the VIF-based variable selection. For example, the MIVs of several soil variables showed weak to moderate correlations with the EIV of temperature (Table 2). This can be explained by similarities in the large-scale gradients of these soil variables and climatic conditions in Europe, due to the formation of the current soil by past climatic conditions (Hengl et al. 2017). Thus, the niche optima

of the soil variables retrieved from the SDMs partly reflect species' response to soil conditions and partly to climate variables.

We found systematic differences in the correlations between the MIVs and EIVs between the different models in our ensemble. Although the BRT retrieved niche optima for the smallest numbers of species (i.e. the least cases where the variable importance was larger than the threshold), the optima were ecologically more realistic than for the GLM and the GAM. This is confirmed by Elith and Graham (2009), who also found that BRT described species' responses better than GLM. Moreover, several studies show that BRT has a higher predictive accuracy than GLM and GAM (Elith and Graham 2009, Marmion et al. 2009). Together these results indicate that BRT is more suitable to project species' core habitat than GAM and GLM. However, BRT can have poor predictive accuracy in case of temporal or spatial bias in the data (Bell and Schlaepfer 2016). This could lead to reduced predictability outside of the core habitat, even when MIVs for niche optima show high correlations with corresponding EIVs.

Our evaluation method can be used to improve the selection of environmental variables and evaluation of SDMs. First, models should be selected based on a good fit on an independent test set of observations, for example based on TSS-values. Subsequently, our method can be used to identify and discard environmental variables which have a low variable importance as well as weak correlations with the EIV, such as total annual precipitation (Fig. 1, 2). Furthermore, our method can be used to identify alternatives for variables with a high variable importance and an ecologically realistic but relatively weak correlation with the EIV, such as nitrogen deposition and pH. For plant SDMs, a possible improvement could be to retrieve site-specific soil variables from process-based models, such as VSD+ (Bonten et al. 2016). However, often better data sources will not be available. In that case it is important to be aware that projected core habitats might not fully capture optimal conditions of the variable of concern, either because fine-grain heterogeneity is ignored, e.g. for pH, or due to the complex relationship of indirect variables with direct variables, e.g. for nitrogen deposition (Pearson and Dawson 2003, Austin 2007).

The novel evaluation approach as presented in this study provides insights into the ecological realism of species niche optima as captured by SDMs. Hence, our methodology can be used for a more informed selection of modeling techniques and ecologically relevant environmental variables, which is particularly relevant to prevent spurious model relationships and consequently false extrapolations (Merow et al. 2014, Fourcade et al. 2017). Moreover, this evaluation method gives insight into the ecological realism of projected core habitats, and consequently can affect the underpinning of conservation planning. Our approach requires an independent source of information on species' habitat preferences. Next to EIVs, indicator values derived from other sources, such as laboratory or field experiments, and for taxonomic groups other than terrestrial vascular plants, can be considered. Further,

Table 2. Spearman rank correlations between the ecological indicator values (EIVs) and the modeled indicator values (MIVs). The value in the first column refers to the number of species included in the bivariate correlation and is equal to the number of species with an ensemble variable importance above 0.05. Bold numbers indicate p-values above 0.3 or below −0.3; p-values are in italics. A critical value of $p<0.05$ corresponds to $p<0.0004$ based on a Bonferroni correction for multiple comparisons.

MIVs	EIVs						
	Light	Soil moisture	Soil acidity	Soil nutrient availability	Salinity	Temperature	Conti-nentality
Minimum temperature (1425)	0.17 *<0.0001*	−0.14 *<0.0001*	−0.19 *<0.0001*	−0.07 *0.0108*	0.23 *<0.0001*	**0.43** *<0.0001*	**−0.32** *<0.0001*
Growing degree days (1380)	−0.03 *0.3511*	−0.28 *<0.0001*	0.10 *0.0002*	0.09 *0.0015*	−0.02 *0.4833*	**0.53** *<0.0001*	0.14 *<0.0001*
Precipitation (656)	−0.17 *<0.0001*	0.10 *0.0086*	0.01 *0.8137*	−0.09 *0.0156*	−0.22 *<0.0001*	**−0.44** *<0.0001*	−0.25 *<0.0001*
pH (1062)	0.24 *<0.0001*	**−0.39** *<0.0001*	**0.35** *<0.0001*	−0.04 *0.2277*	0.15 *<0.0001*	**0.37** *<0.0001*	0.23 *<0.0001*
Silt (1078)	−0.28 *<0.0001*	0.22 *<0.0001*	0.24 *<0.0001*	0.26 *<0.0001*	−0.21 *<0.0001*	−0.24 *<0.0001*	−0.01 *0.7113*
Coarse fragments (983)	−0.04 *0.2537*	**−0.32** *<0.0001*	0.03 *0.3764*	−0.20 *<0.0001*	**−0.32** *<0.0001*	−0.03 *0.2817*	−0.12 *0.0002*
Clay (884)	0.06 *0.0635*	**−0.33** *<0.0001*	0.29 *<0.0001*	0.00 *0.9781*	0.01 *0.8521*	**0.35** *<0.0001*	0.14 *<0.0001*
Carbon (632)	−0.03 *0.3799*	**0.32** *<0.0001*	−0.01 *0.8143*	0.10 *0.0120*	0.04 *0.3523*	**−0.34** *<0.0001*	−0.02 *0.6490*
Cation exchange capacity (372)	−0.13 *0.0107*	0.20 *0.0001*	0.18 *0.0005*	−0.05 *0.3622*	−0.17 *0.0011*	**−0.46** *<0.0001*	0.12 *0.0153*
Nitrogen deposition (935)	−0.29 *<0.0001*	0.29 *<0.0001*	0.11 *0.0010*	**0.33** *<0.0001*	0.02 *0.5017*	−0.23 *<0.0001*	−0.06 *0.0887*
Forests (740)	**−0.39** *<0.0001*	−0.07 *0.0410*	0.04 *0.2771*	−0.11 *0.0021*	**−0.37** *<0.0001*	−0.07 *0.0504*	−0.09 *0.0148*
Scrubs (369)	0.24 *<0.0001*	**−0.34** *<0.0001*	−0.17 *0.0009*	**−0.47** *<0.0001*	−0.27 *<0.0001*	0.14 *0.0070*	0.00 *0.9532*
Arable land (359)	−0.09 *0.0889*	0.10 *0.0515*	−0.06 *0.2223*	0.04 *0.4227*	−0.13 *0.0168*	0.02 *0.6746*	0.07 *0.1890*
Pastures (329)	−0.15 *0.0072*	**0.36** *<0.0001*	−0.09 *0.0916*	0.11 *0.0523*	−0.05 *0.3536*	−0.05 *0.3583*	0.03 *0.6416*
Inland wetlands (45)	0.15 *0.3057*	**0.71** *<0.0001*	**−0.43** *0.0017*	**−0.63** *<0.0001*	**−0.39** *0.0049*	**−0.51** *0.0002*	0.01 *0.9584*
Heterogeneous agriculture (203)	−0.28 *<0.0001*	0.20 *0.0036*	−0.12 *0.0839*	0.08 *0.2435*	−0.17 *0.0124*	0.19 *0.0057*	−0.09 *0.2014*
Open spaces (78)	**0.40** *0.0002*	−0.10 *0.3734*	−0.03 *0.7860*	**−0.38** *0.0004*	−0.20 *0.0694*	−0.23 *0.0362*	0.00 *0.9650*

the approach may be extended to niche characteristics other than optima, notably upper and lower tolerance limits or shapes of response curves, depending on the availability of suitable data. For example, the extensive availability of lab test data on thermal tolerance of particularly ectotherms (Bennett et al. 2018) provides a promising opportunity to validate niche width of these species as captured by SDMs. This would allow for a more in-depth analysis of the transferability of SDMs, as transferability depends not only on optima but also on the niche width and the shape of the response curve.

References

Araújo, M. B. et al. 2011. Climate change threatens European conservation areas. – Ecol. Lett. 14: 484–492.

Austin, M. P. 2007. Species distribution models and ecological theory: a critical assessment and some possible new approaches. – Ecol. Model. 200: 1–19.

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

Azevedo, L. B. et al. 2013. Global assessment of the effects of terrestrial acidification on plant species richness. – Environ. Pollut. 174: 10–15.

Bahn, V. and McGill, B. J. 2013. Testing the predictive performance of distribution models. – Oikos 122: 321–331.

Barbet-Massin, M. et al. 2012. Selecting pseudo-absences for species distribution models: how, where and how many? – Methods Ecol. Evol. 3: 327–338.

Barnosky, A. D. et al. 2011. Has the Earth's sixth mass extinction already arrived? – Nature 471: 51–57.

Bell, D. M. and Schlaepfer, D. R. 2016. On the dangers of model complexity without ecological justification in species distribution modeling. – Ecol. Model. 330: 50–59.

Bennett, J. M. et al. 2018. GlobTherm, a global database on thermal tolerances for aquatic and terrestrial organisms. – Sci. Data 5: 180022.

Bonten, L. T. C. et al. 2016. A model to calculate effects of atmospheric deposition on soil acidification, eutrophication and carbon sequestration. – Environ. Model. Softw. 79: 75–84.

Butchart, S. H. M. et al. 2010. Global biodiversity: indicators of recent declines. – Science 328: 1164–1168.

Campbell, D. J. et al. 1989. The soil solution chemistry of some Oxfordshire soils: temporal and spatial variability. – J. Soil Sci. 40: 321–339.

Chytrý, M. et al. 2016. European Vegetation Archive (EVA): an integrated database of European vegetation plots. – Appl. Veg. Sci. 19: 173–180.

Diekmann, M. 2003. Species indicator values as an important tool in applied plant ecology – a review. – Basic Appl. Ecol. 4: 493–506.

Dubuis, A. et al. 2011. Predicting spatial patterns of plant species richness: a comparison of direct macroecological and species stacking modelling approaches. – Divers. Distrib. 17: 1122–1131.

Elith, J. and Graham, C. H. 2009. Do they? How do they? WHY do they differ? On finding reasons for differing performances of species distribution models. – Ecography 32: 66–77.

Elith, J. and Leathwick, J. R. 2009. Species distribution models: ecological explanation and prediction across space and time. – Annu. Rev. Ecol. Evol. Syst. 40: 677–697.

Elith, J. et al. 2005. The evaluation strip: a new and robust method for plotting predicted responses from species distribution models. – Ecol. Model. 186: 280–289.

Ellenberg, H. et al. 1991. Zeigerwerte von Pflanzen in Mitteleuropa, 2nd ed. – Scripta Geobotanica, Verlag Erich Goltze KG.

Ellenberg, H. et al. 2001. Zeigerwerte von Pflanzen in Mitteleuropa. 3. durchgesehene Auflage. – Scr. Geobot. 18: 1–262.

Fagerli, H. et al. 2015. Transboundary particulate matter, photo-oxidants, acidifying and eutrophying components. – EMEP Status Report 1/2015, Norwegian Meteorological Inst.

Fourcade, Y. et al. 2017. Paintings predict the distribution of species, or the challenge of selecting environmental predictors and evaluation statistics. – Global Ecol. Biogeogr. 27: 245–256.

Frank, D. et al. 1990. Biologisch-ökologische Daten zur Flora der DDR. – Martin-Luther-Univ. Halle-Wittenberg.

Harfoot, M. B. J. et al. 2014. Integrated assessment models for ecologists: the present and the future. – Global Ecol. Biogeogr. 23: 124–143.

Hazeu, G. W. et al. 2008. D01.02.01 Land cover maps for environmental modeling at multiple scales. – Alterra and WSL, Wageningen, the Netherlands.

Heikkinen, R. K. et al. 2010. Assessing the vulnerability of European butterflies to climate change using multiple criteria. – Biodivers. Conserv. 19: 695–723.

Heikkinen, R. K. et al. 2012. Does the interpolation accuracy of species distribution models come at the expense of transferability? – Ecography 35: 276–288.

Hengl, T. et al. 2017. SoilGrids250m: global gridded soil information based on machine learning. – PLoS One 12: e0169748.

Hennekens, S. M. et al. 2017. BioScore 3 – plants. Background and preprocessing of distribution data. – Statutory Research Tasks Unit for Nature and the Environment (WOT Natuur and Milieu), WOt-technical report 106.

Hill, M. O. et al. 1999. Ellenberg's indicator values for British plants. ECOFACT vol. 2 technical annex. – Inst. of Terrestrial Ecology, Huntingdon, UK.

Janssen, J. A. M. et al. 2016. European red list of habitats – part 2. Terrestrial and freshwater habitats. – European Commission.

Karger, D. N. et al. 2017. Climatologies at high resolution for the earth's land surface areas. – Sci. Data 4: 170122.

Kooijman, A. M. et al. 1998. Nutrient limitations and their implications on the effects of atmospheric deposition in coastal dunes; lime-poor and lime-rich sites in the Netherlands. – J. Ecol. 86: 511–526.

Landolt, E. 1977. Ökologische Zeigerwerte der Schweizer Flora. – Veröffentlichungen des Geobotanischen Inst. der ETH, Stiftung Rübel 64: 1–208.

Lechowicz, M. J. and Bell, G. 1991. The ecology and genetics of fitness in forest plants. II. Microspatial heterogeneity of the edaphic environment. – J. Ecol. 79: 687–696.

Louette, G. et al. 2010. BioScore–cost-effective assessment of policy impact on biodiversity using species sensitivity scores. – J. Nat. Conserv. 18: 142–148.

Marmion, M. et al. 2009. Evaluation of consensus methods in predictive species distribution modelling. – Divers. Distrib. 15: 59–69.

Merow, C. et al. 2014. What do we gain from simplicity versus complexity in species distribution models? – Ecography 37: 1267–1281.

Mitchell, P. J. et al. 2017. Sensitivity of fine-scale species distribution models to locational uncertainty in occurrence data across multiple sample sizes. – Methods Ecol. Evol. 8: 12–21.

Morán-Ordóñez, A. et al. 2017. Evaluating 318 continental-scale species distribution models over a 60-year prediction horizon: what factors influence the reliability of predictions? – Global Ecol. Biogeogr. 26: 371–384.

Mucina, L. et al. 2016. Vegetation of Europe: hierarchical floristic classification system of vascular plant, bryophyte, lichen, and algal communities. – Appl. Veg. Sci. 19: 3–264.

Nissinen, A. and Hari, P. 1998. Effects of nitrogen deposition on tree growth and soil nutrients in boreal Scots pine stands. – Environ. Pollut. 102: 61–68.

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

Pereira, H. M. et al. 2010. Scenarios for global biodiversity in the 21st century. – Science 330: 1496–1501.

Petitpierre, B. et al. 2017. Selecting predictors to maximize the transferability of species distribution models: lessons from cross-continental plant invasions. – Global Ecol. Biogeogr. 26: 275–287.

Pignatti, S. 2005. Valori di bioindicazione delle piante vascolari della Flora d'Italia (Bioindicator values of vascular plants of the flora of Italy). – Braun-Blanquet. 39: 3–95.

Rapacciuolo, G. et al. 2012. Climatic associations of British species distributions show good transferability in time but low predictive accuracy for range change. – PLoS One 7: e40212.

Schaminée, J. H. J. et al. 2013. Review of EUNIS forest habitat classification. – European Environmental Agency (EEA/NSV/13/005), Copenhagen.

Schaminée, J. H. J. et al. 2014. Vegetation analysis and distribution maps for EUNIS habitats. – European Environmental Agency (EEA/NSV/14/006), Copenhagen.

Schaminée, J. H. J. et al. 2015. Review of grassland habitats and development of distribution maps of heathland, scrub and tundra habitats of EUNIS habitats classification. – European Environmental Agency (EEA/NSV/15/005), Copenhagen.

Schaminée, J. H. J. et al. 2016. Development of distribution maps of grassland habitats of EUNIS habitat classification. – European Environmental Agency (EEA/NSS/16/005), Copenhagen.

Scherrer, D. and Körner, C. 2011. Topographically controlled thermal-habitat differentiation buffers alpine plant diversity against climate warming. – J. Biogeogr. 38: 406–416.

Settele, J. et al. 2008. Climatic risk atlas of European butterflies. – BioRisk 1: 1–712.

Skov, F. and Svenning, J.-C. 2004. Potential impact of climate change on the distribution of forest herbs in Europe. – Ecography 27: 366–380.

Thuiller, W. et al. 2004. Do we need land-cover data to model species distributions in Europe? – J. Biogeogr. 31: 353–361.

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

Thuiller, W. et al. 2016. biomod2: ensemble platform for species distribution modeling. – R package ver. 3.3-7.1, <https://CRAN.R-project.org/package=biomod2>.

Tittensor, D. P. et al. 2014. A mid-term analysis of progress toward international biodiversity targets. – Science 346: 241–244.

Visconti, P. et al. 2016. Projecting global biodiversity indicators under future development scenarios. – Conserv. Lett. 9: 5–13.

Walter, H. 1985. Vegetation of the Earth and ecological systems of the geo-biosphere. – Springer.

Whittaker, R. J. et al. 2007. Geographic gradients of species richness: a test of the water-energy conjecture of Hawkins et al. (2003) using European data for five taxa. – Global Ecol. Biogeogr. 16: 76–89.