

Modelling plant species richness using functional groups

K. Steinmann^{a,*}, H.P. Linder^b, N.E. Zimmermann^a

^a Land Use Dynamics Unit, Swiss Federal Research Institute WSL, Zürcherstrasse 111, CH-8903 Birmensdorf, Switzerland

^b Institute of Systematic Botany, Zollikerstrasse 107, CH-8008 Zürich, Switzerland

ARTICLE INFO

Article history:

Received 26 February 2008

Received in revised form

23 December 2008

Accepted 14 January 2009

Available online 21 February 2009

Keywords:

Clustering species

Generalized linear models

Meso-scale

Morphological traits

Niche width

ABSTRACT

Conservation biologists increasingly rely on spatial predictive models of biodiversity to support decision-making. Therefore, highly accurate and ecologically meaningful models are required at relatively broad spatial scales. While statistical techniques have been optimized to improve model accuracy, less focus has been given to the question: How does the autecology of a single species affect model quality? We compare a direct modelling approach versus a cumulative modelling approach for predicting plant species richness, where the latter gives more weight to the ecology of functional species groups. In the direct modelling approach, species richness is predicted by a single model calibrated for all species. In the cumulative modelling approach, the species were partitioned into functional groups, with each group calibrated separately and species richness of each group was cumulated to predict total species richness. We hypothesized that model accuracy depends on the ecology of individual species and that the cumulative modelling approach would predict species richness more accurately. The predictors explained plant species richness by ca. 25%. However, depending on the functional group the deviance explained varied from 3 to 67%. While both modelling approaches performed equally well, the models of the different functional groups highly varied in their quality and their spatial richness pattern. This variability helps to improve our understanding on how plant functional groups respond to ecological gradients.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

Predictive modelling of species richness relies on niche theory and gradient analysis (Austin, 2002). The fundamental niche is the environmental range a species population can persistently occupy without immigration (Hutchinson, 1957). The realized niche of a species additionally includes biotic interactions such as, e.g. competitive exclusion (Hutchinson, 1957). According to the continuum concept the species composition gradually changes along an environmental gradient as the realized niches of species populations overlap (Austin, 1985). The response of species along environmental gradients is often assumed to follow a Gaussian shape with identical width and height but individually distributed optima (Whittaker, 1956; Gauch and Whittaker, 1972; ter Braak, 1985; but see Abrams, 1995 and Austin, 2002). However, Austin and Smith (1989) pointed out that physiological processes and interactions between species may lead to skewed, bimodal or more complex response curves. As a species' niche response is the result of multiple reactions to varying environmental gradients, it is unlikely to exclusively take a multivariate Gaussian shape (Minchin, 1989;

Austin et al., 1994). Woodward and Kelly (1997) showed that even a simple temperature gradient can result in highly skewed maximum photosynthetic rates for plants in different biomes. It is now accepted that the physiological responses of species to environmental factors are skewed (Oksanen and Minchin, 2002). In addition, plant species with optima close to the end of a gradient seem to have narrower niche widths, compared to species with growth optima near the mid-point of the same gradient (Thuiller et al., 2004). Hence, according to its physiological and morphological traits, each species has its own niche, with its characteristic shape and optimum along a given environmental gradient. This expression of the realized niche is often calibrated in statistical models for predicting species spatial occurrence or richness patterns (Guisan and Zimmermann, 2000). While there is a wide range of studies concerning the optimization of different statistical techniques and comparing their quality (see Guisan et al., 2002; Thuiller et al., 2003; Rushton et al., 2004; Austin et al., 2006), comparably few studies attempt to optimize the selection of model predictors (but see Mac Nally, 2000; Austin et al., 2006). The question whether or to what degree the autecology of single species might affect the model quality, has been neglected even more. In the present study, we assume that the width, position and shape of realized niches of single species will influence the predictive power of species richness models.

* Corresponding author. Tel.: +41 1 739 25 07; fax: +41 1 739 22 15.
E-mail address: Katharina.Steinmann@wsl.ch (K. Steinmann).

Conceptually, species richness can be modelled at the community level or at the individual level (Ferrier and Guisan, 2006). In the first case species richness is predicted by a set of environmental parameters calibrated by one single model (direct approach). In the second case, individual species are modelled separately as a function of environmental variables. Distribution maps of the individual species are stacked to produce a species distribution map (complete cumulative approach).

The first method allows for rapid analyses of very large numbers of species, and integrates the interaction between species. The second method shows its strength by accounting for individualistic species responses (for a review, see Ferrier and Guisan, 2006). Thus the first method is straightforward, but may show deficiencies as pooling of all species with their different realized niches may result in information loss (Minchin, 1989). For example two environmentally different sites may share the same number of species, yet the species composition at these two sites may differ greatly (Sarr et al., 2005). In one habitat, species may primarily be limited by temperature, in the other habitat species may be constrained by precipitation (Broennimann et al., 2006). An overall richness model may therefore include (or exclude) both parameters, hence the predicted species richness response may be fuzzy. In contrast, predicting the spatial distribution of each species separately might result in more precise cumulative predictions of species richness. However, depending on the targeted ecosystem and the spatial scale, the total number of species can be huge (>1000 species), while presence data of single species may be scarce. Overlaying simulated presence absence maps for individual species may result in an accumulation of errors when calculating species richness, and thus lower the strength of the full cumulative model.

A compromise between the two described approaches involves dividing all species into functional groups consisting of similar eco-physiology, predicting the species richness within each group separately and then cumulating richness values group-wise to calculate total species richness (i.e. a partially cumulative approach). This approach may reduce information loss and minimize the confidence interval associated with model accuracy. The challenge of the intermediate approach is to find appropriate criteria to define functional groups representing similar eco-physiological behaviour (Smith, 1997). Adaptations to specific environmental constraints require specific traits, which are reflected in morphological, physiological or life history characters of each single species (Parkhurst and Loucks, 1972). Therefore, morphological as well as life history traits seem suitable to determine the niche of a specific plant species (Lavergne et al., 2004) and are often used to divide plant species into functional groups.

The principal goal of our study was to determine whether the modelling of plant species richness using multiple functional groups is more accurate than modelling the total plant species richness. The second goal was to evaluate whether the results differ depending on the level of functional differentiation. As such, functional groups were built on a coarse level (4 life form groups) and on a finer level (40 groups based on morphological traits). As more than 70% of the species used were perennial herbs, only the latter group was partitioned into smaller groups using a clustering method applied on morphological characters. In order to evaluate the ecological relevance of the clusters, the habitat affinity of the species of each cluster was tested. By this, we aimed at testing the following two hypotheses: (1) The accuracy of plant species richness models can be improved by using functional groups modelled separately and then aggregated a posteriori. (2) The accuracy of plant species richness models using a functional group approach depends on the average niche width, with higher accuracy expected for groups with narrow niches.

2. Materials and methods

2.1. Study site

The study was carried out in Switzerland (~6–10.5°N and ~46–47.5°E). The climate is generally humid temperate with rather mild winters and moderately warm summers. Annual mean temperature ranges from ca. 1.5 °C (in high alpine valleys) to 11.5 °C (in the southern part of Switzerland). The mean annual precipitation sums vary between 600 mm and 1700 mm.

2.2. Response variables

The species richness data were derived from vegetation recordings at 784 sites in Switzerland. The data originate from the biodiversity monitoring program of the Swiss federal office of environment (Plattner et al., 2004). Each record contains all vascular plant species growing in a circular plot of 10 m² (radius = 1.78 m). The sampling sites are organised in a regular grid with mesh size of 6 km × 4 km. The initial point of the grid was generated randomly. Urban areas and inappropriate growing conditions such as lakes and glaciers were excluded. Therefore, only 421 out of the 784 sites were kept for further analyses.

2.3. Predictors

For the predictive modelling of species richness, we used a suite of parameter maps to represent climatic (12 maps), topographic (5 maps) and edaphic (4 maps) conditions in combination with a coarse habitat classification (8 classes) as environmental predictors (see Table S1 in the electronic appendix). Most of the climatic parameters were available as monthly variables. The climatic parameters were generated according to Zimmermann and Kienast (1999) using a digital elevation model (DEM) with a spatial resolution of 25 m and data from meteorological measurements. Climate variables represented monthly Normals of the period 1961–1990. Degree days of the growing season are based on a threshold of 0 °C. Site water balance was calculated as the monthly difference of precipitation and potential evapotranspiration over the water holding capacity of the soil (see Maggini et al., 2006 for details). The topographic wetness index expresses the lateral water flow. It was calculated according to Beven and Kirkby (1979). The topographic position is a measure of convexity of the terrain and was calculated from the DEM directly. As a surrogate of potential stand productivity two edaphic layers were used: (1) coarse fragment content and (2) nutrient availability. Finally, the presence of limestone bedrock was derived from the digital geotechnical map of Switzerland (1:200,000, de Quervain et al., 1963–1967). While all continuous data were available at a 25 m resolution, the categorical data had a resolution of 100 m.

2.4. Regression modelling and model testing

To avoid model calibration bias, the predictive parameters were first checked for collinearity. After discarding collinear parameters according to the method described in Belsley (1991), 25 predictors out of the original 95 were kept for modelling species richness pattern. The plot of species richness showed a negative binomial distribution, thus we used generalized linear models with a negative binomial model family. In two cases the group of modelled species was too small (0 or 1 species per plot), thus a binomial model family was used. We employed a stepwise forward variable selection to build models. The number of parameters entering the final model was set to minimize the Bayesian information criterion (BIC), a comparably strict selection algorithm (Schwarz, 1978). Continuous predictors were allowed to enter the model as linear and

quadratic terms. In order to test for spatial autocorrelation sample variograms were computed from the model residuals. Model calibration was calculated from the deviance explained (D^2) and model performance was evaluated from the mean absolute error (MAE) and from the prediction bias (mean of the error). All three parameters were calculated as the mean statistic originating from a 10-fold cross-validation. GIS layers of each predictor were used for spatial predictions of species richness throughout Switzerland. Areas outside the calibrated predictor space were excluded from the predictions (*safe predictions*, Guisan et al., 2006). Maps were generated for all calibrated models.

The two modelling approaches (partially cumulative versus direct prediction) were compared by testing the MAE of both approaches for significant differences (paired Wilcoxon test).

We did not test a complete cumulative model because 94% of our species occurred in fewer than 50 plots, which is below a threshold that allows sound model calibration (Manel et al., 2001).

All statistical analyses were performed in the statistical software environment R (R Development Core Team, 2006).

2.5. Clustering species into functional groups

For the different models, the total list of plant species was divided into four functional groups (annuals and biannuals, perennial herbs, shrubs and trees). The information for this classification of life form was taken from Aeschimann et al. (2004). Perennial herb species were further classified into 40 subgroups according to a set of morphological characters. The characters included rosette type, lateral growth, shoot metamorphosis and leaf anatomy with 3, 4, 7 and 6 classes, respectively (see Table 1 for definitions). We used Klotz et al. (2002) for defining the morphological characteristics and we compiled the characteristics for the species from Hess et al. (1976) and Klotz et al. (2002). We applied a cluster analysis by partitioning around medoids (Kaufmann and Rousseeuw, 1990) to divide the perennial plant group into subgroups. We then used silhouette coefficients (*sc*: defined as the maximal average silhouette width for the entire data set) to judge the appropriateness of the cluster structure. A *sc* > 0.5 indicates a reasonable structure in the data (Kaufmann and Rousseeuw, 1990). Hence the number of morphological species groups (*k*) built was set so that *k* was minimized while *sc* was not allowed to drop below a value of 0.5. This procedure resulted in 40 subgroups.

2.6. Habitat affinity of the functional groups

To test whether clustering based on morphological characteristics resulted in ecologically meaningful groups, we applied a correlation analysis for habitat affinity. We assigned the species to nine habitat types (Table 2) according to Aeschimann et al. (2004). The different proportions of species per habitat type (considering all species) served as null model. For each cluster group, the species proportions per habitat type were compared against the null model using Spearman's rank correlation coefficient. Low correlations (with associated high

Table 1
Plant traits used for clustering the perennial herbs into functional groups.

Rosette type	Lateral growth	Shoot metamorphosis	Leaf anatomy
fr: full rosette	Cushion	b: bulbill	hel: helomorph
hr: half rosette	Standard	n: standard shoot	hyd: hydromorph
nr: no rosette	Turf	p: pleiocorm	hyg: hygromorph
	Tussock	rh: rhizome	m: mesomorph
		rht: rhizome and tiller building	scl: scleromorph
		rp: rhizome pleiocorm	suc: succulent
		rt: root tuber	

Table 2
Description of the habitat types used to test group affinities.

Code	Habitat type
1	Water courses
2	Nitrophilous vegetation
3	Scree, gravel and rocky area
4	Creeks
5	Swamps
6	Meadows, pasture, turfs, snow beds
7	Dwarf-shrub heathland, tall forb meadow
8	Scrubland <i>sensu lato</i>
9	Forest

P-values) indicate that the two species lists are not equally distributed across the habitats. Thus, it indicates that the respective cluster group contains species, more associated to a specific habitat and hence the group reflects ecological idiosyncrasies.

2.7. Niche width and model quality

The variance of the normalized predictor values among all observations per group was used as a measurement of niche width. The niche width of a single cluster group was then defined as the mean variance along the predictor gradients of the observed presence data points of the respective group. Only predictors that remained in the respective models were used to calculate the average species niche width per group. As the mean absolute errors increased with the maximal observed species richness, they were expressed relative to the maximal observed richness number per group. Finally the niche widths were correlated with the relative MAEs using Pearson's product-moment correlation.

3. Results

3.1. Model performance

Clustering the 789 perennial species from four morphological traits resulted in 40 distinct groups (Table S2 in the electronic appendix). On average each group contained 17 species (the smallest group contained 4 species, the largest group 67 species). All 25 predictors (which remained after testing for collinearity) were significant in at least one of the 45 models (see Table S3 in the electronic appendix). Computing the variograms of the model residuals showed that the distances between sampling points were large enough to avoid spatial autocorrelation. Direct prediction of the total species richness resulted in a D^2 of 0.28 (Table 3), whereas predicting functional group richness at the plot level revealed D^2 values between 0.06 (biannual plants) and 0.68 (trees), with an average of 0.31.

The model which predicted the total species richness directly showed a weak negative bias (−0.05, i.e. a marginal underes-

Table 3
Mean absolute error (MAE) and explained deviance (D^2) of the models.

Group	MAE	D^2
All species	7.9	0.28
Trees	0.75	0.68
Shrubs	1.44	0.17
Perennial herbs	7.02	0.34
(Bi-)annuals	0.73	0.06
Sum of 4 groups	7.97	
Sum of 40 groups	6.69	
Summary statistics of the 40 functional groups		
Mean	0.36	0.23
Variance	0.06	0.01

timation) and an MAE of 7.9 species. The accumulation of the modelled species richness from four functional groups resulted in a small positive bias (0.08) and a similar MAE of 7.97 species. The direct richness model of the perennial group showed a positive bias of 0.08 and an MAE of 7.02, whereas the cumulated predictions of perennial species richness from the 40 functional groups reduced the bias to 0.02 and the MAE to 6.69. All 40 perennial subgroup models were unbiased and had an MAE <1.2 species.

3.2. Comparison of the two modelling approaches

The sum of the predicted species richness of the four different life forms resulted in very similar spatial pattern compared to the directly modelled species richness (Fig. 1a and b). We found a high similarity between the predicted perennial plant richness and the sum of all 40 richness models of perennial herbs clustered from functional traits (Fig. 1c and d). Since roughly 75% of the recorded species were perennial herbs, it is not surprising that

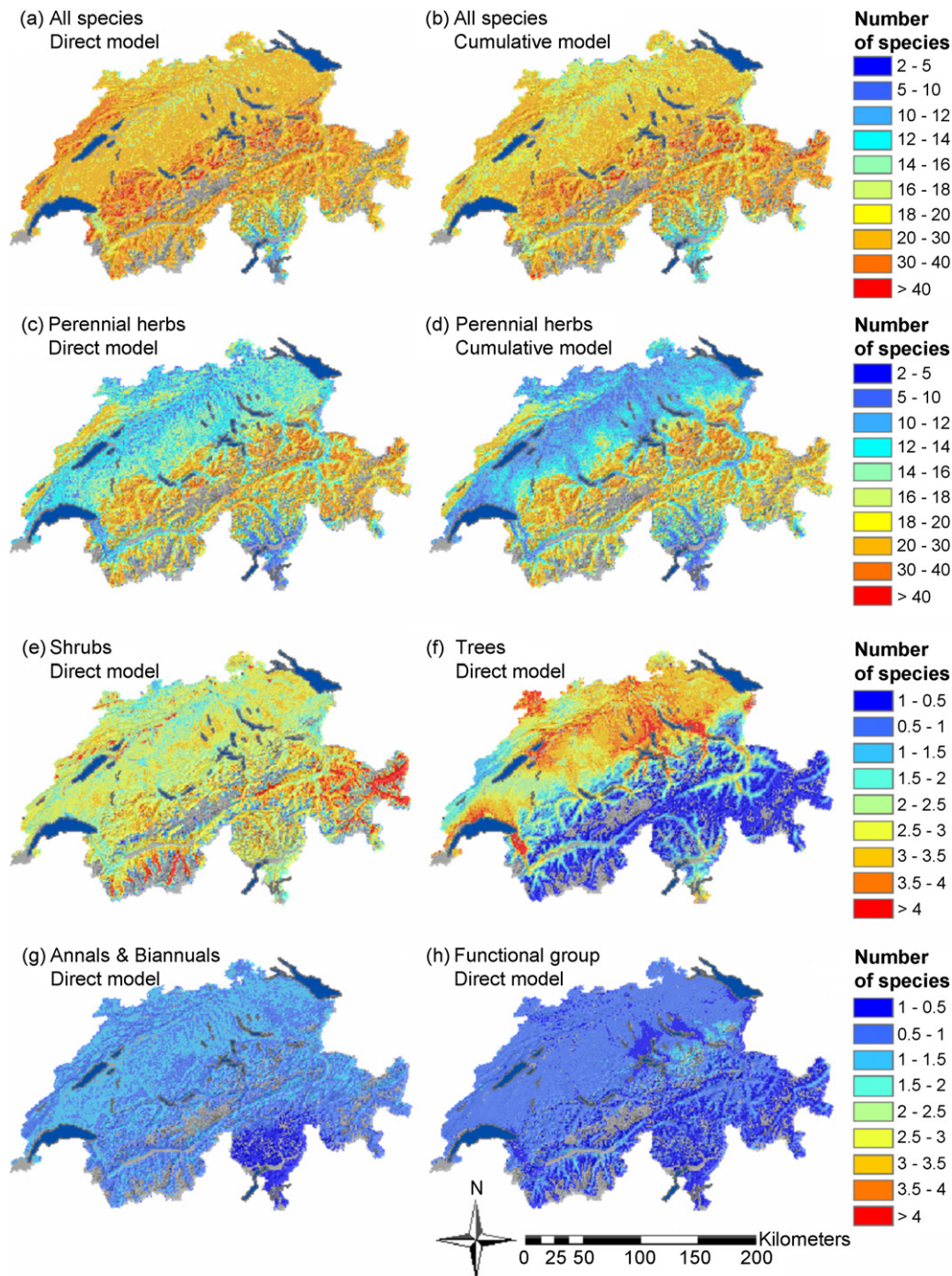


Fig. 1. Spatial distribution of plant species richness. The maps represent the following models: (a) direct modelling of species richness; (b) cumulative modelling of species richness from four life forms; (c) direct modelling of species richness of perennial herbs; (d) cumulative modelling of species richness from 40 functional groups of perennial herbs; (e) direct modelling of shrub species richness; (f) direct modelling of tree species richness; (g) direct modelling of annual and biannual species richness; (h) direct modelling of a functional group (the group with plants of the type *Urtica dioica* is given as an example). Grey regions represent areas with climatic parameters exceeding the range of values used for calibration (safe predictions). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

the richness pattern of this group shares a high similarity with the pattern of the total species richness. However, the spatial patterns of the different functional groups varied considerably (Fig. 1e–h and Fig. S1 in the electronic appendix).

The comparison of the model residuals for total plot species richness revealed no significant difference between the predictions of the two modelling approaches ($P=0.32$ for absolute residual values). Similarly, the comparison of the perennial model residuals did not reveal significant differences ($P=0.08$ for absolute residuals). It indicates that the richness models from functional groups did not significantly increase prediction accuracies.

3.3. Niche width, habitat affinity and model evaluation

The realized niches, defined by the mean variance of the normalized predictors entering the model for a specific functional group, varied between 0.3 (*Crassulaceae* species, group c29) and 1.8 (species of the type *Potentilla reptans*, group c3). The model fit could not be explained by the average niche width per perennial group ($r=0.08$, $P=0.63$).

Ten out of the forty perennial groups consisted of species clearly belonging to a specific habitat type (see Table S2 in the electronic appendix). However, groups that significantly reflect specific habitats did not necessarily result in better models (a *t*-test comparing the explained deviances of the two groups was insignificant; $P=0.43$).

4. Discussion

Both the selected predictors and the model quality varied considerably between functional groups. In line with other studies (Sarr et al., 2005; Broennimann et al., 2006) the model for tree species richness performed best. This may be explained by the dominance of this life form, which is highly competitive after establishment. In suitable habitats, the detection probability for Trees – and therefore the prediction certainty – is higher compared to small herbs as the sensitivity for trees is higher than for small herbs. In contrast, short lived-species showed rather poor model fit, which was also found by Broennimann et al. (2006). Since (bi-)annual plants require open space to regenerate, they are found less often under dense vegetation, which reduces their detection probability and consequently their model quality. Such plants belong to the “rural” plant type (according to Collins et al., 1993), a plant type that has been identified as being difficult to predict (Edwards et al., 2005; Zimmermann et al., 2007). Within the perennial species groups, wetland species were best predicted. Even though their habitat is restricted, they show a wide geographic distribution and high local abundance. Thus, they belong to predictable habitat specialists (Rabinowitz et al., 1986). Yet we did not find a general agreement between niche width and model fit. We believe that this is the main reason why the two different model approaches tested here did not yield differing model accuracies for the total species richness, and only marginal differences for the perennial species richness models. The influence of both the ecological information loss (mixture of niches) and the cumulated uncertainty (many group models accumulated) might cancel each other out when modelling species richness at a plot level where randomness is high, a conclusion that was also found by Guisan and Theurillat (2000) when comparing individual species distribution versus species richness models.

Effects of geographic range size and varying population densities of single species may have strongly influenced the different models (Venier et al., 1999; Stockwell and Peterson, 2002; McPherson et al., 2004). By grouping the perennial herbs according to their lateral growth and shoot metamorphosis, we differentiated between clonal species and species reproducing sexually. As shown by Kolb

et al. (2006) the local abundance of a species is related to such life histories. The lateral spread of clonal species results in a high ramet density within a given habitat patch. Therefore, groups containing clonal species are expected to result in higher model accuracy, compared to groups with similar habitat width and lower local abundance. Hence, independently of niche width, we generated models of differing quality.

In summary, we found no clear improvements for species richness modelling when splitting species into functional groups. The direct modelling of plot species richness performed equally well. However, in line with Bruun et al. (2006) our analysis showed a clearly differing response of species groups to environmental gradients. In particular, the spatial patterns of the 40 classes of the perennial herbs differed highly from each other.

Even though our hypothesis of improving predictive species richness models by considering the autecology of the species was not substantially supported by the data, it is not necessarily falsified either for the following reasons: (1) The sampling unit of 10 m² is small and does not reflect the local flora of a site well enough. Thus, random effects and non-equilibrium processes (Huston, 1979, 1994) introduce noise and impede the detection of ecological patterns, which depend on the spatial resolution (Willis and Whitaker, 2002; Rahbek, 2005; Luoto et al., 2007). To circumvent this species package problem, Crawley and Hurrall (2001) suggested that areas >100 m² should be used to characterize alpha diversity. (2) In the data set used, more than 50% of all sampling sites were covered by forests. On the other hand the presence of dwarf-shrubs, boulder and gravel sites, mires, peat land or riverside vegetation had low frequencies (<1% each). Therefore, we had incomplete coverage of ecological gradients and some species were not well represented in the data set. This resulted in less precise models for groups containing several such species (Venier et al., 1999; Edwards et al., 2005). A stratified random sampling design may overcome this type of deficit (Tillé, 2001). (3) The selected plant traits may be irrelevant for partitioning of species into ecological meaningful groups. Yet, since some groups attained high model qualities, they seem to show a clear ecological preference. Also, as one fourth of the functional groups showed a significant habitat affiliation, the selected plant traits seem to carry ecological information.

Although the individual errors of the class models averaged to comparable model quality at the overall species richness level, the individual models allow us to make better-informed predictions of subgroups at the landscape scale. This clearly improves our understanding of how functional groups are distributed spatially, and how these patterns and their associated species richness may alter in the future under scenarios of global changes. We see this as the most important reason for distinguishing functional groups for such future projections.

Acknowledgements

This research was funded by the 6th Framework Programme of the European Union (Contract Number GOCE-CT-2003-505376). We thank the Swiss Federal Agency SAEFL and Hintermann and Weber Consultancy for access to the data from the biodiversity monitoring programme.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ecolmodel.2009.01.006.

References

- Abrams, P.A., 1995. Monotonic or unimodal diversity-productivity gradients: what does competition theory predict? *Ecology* 76 (7), 2019–2027.

- Aeschmann, D., Lauber, K., Moser, D.M., Theurillat, J.-P., 2004. *Flora Alpina*. Haupt Verlag, Bern.
- Austin, M.P., 1985. Continuum concept, ordination methods, and niche theory. *Annu. Rev. Ecol. Syst.* 16, 39–61.
- Austin, M.P., 2002. Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecol. Model.* 157, 101–118.
- Austin, M.P., Smith, T.M., 1989. A new model for the continuum concept. *Vegetatio* 83, 35–47.
- Austin, M.P., Nicholls, A.O., Doherty, M.D., Meyers, J.A., 1994. Determining species response functions to an environmental gradient by means of a β -function. *J. Veg. Sci.* 5, 215–228.
- Austin, M.P., Belbin, L., Meyers, J.A., Doherty, M.D., Luoto, M., 2006. Evaluation of statistical models used for predicting plant species distributions: role of artificial data and theory. *Ecol. Model.* 199, 197–216.
- Belsley, D.A., 1991. *Conditioning Diagnostics*. Wiley & Sons, New York, 40 pp.
- Beven, K.J., Kirkby, M.J., 1979. A physically based, variable contributing area model of basin hydrology. *Hydrol. Sci. Bull.* 24, 43–69.
- Broennimann, O., Thuiller, W., Hughes, G., Midgley, G.F., Alkemade, J.M.R., Guisan, A., 2006. Do geographic distribution, niche property and life form explain plant's vulnerability to global change? *Glob. Chang. Biol.* 12, 1079–1093.
- Bruun, H.H., Moen, J., Virtanen, R., Grytnes, J.-A., Oksanen, L., Angerbjörn, A., 2006. Effects of altitude and topography on species richness of vascular plants, bryophytes and lichens in alpine communities. *J. Veg. Sci.* 17, 37–46.
- Collins, S.L., Glenn, S.M., Roberts, D.W., 1993. The hierarchical continuum concept. *J. Veg. Sci.* 4, 149–156.
- Crawley, M.J., Harral, J.E., 2001. Scale dependence in plant biodiversity. *Science* 291, 864–868.
- Edwards, T.C., Cutler, D.R., Zimmermann, N.E., Geiser, L., Algeria, J., 2005. Model-based stratifications for enhancing the detection of rare ecological events. *Ecology* 86 (5), 1081–1090.
- Ferrier, S., Guisan, A., 2006. Spatial modelling of biodiversity at the community level. *J. Appl. Ecol.* 43, 393–404.
- Gauch, H.G., Whittaker, R.H., 1972. Coenocline simulation. *Ecology* 53 (3), 446–451.
- Guisan, A., Theurillat, J.-P., 2000. Equilibrium modelling of alpine plant distribution: how far can we go? *Phytocoenologia* 30, 353–384.
- Guisan, A., Zimmermann, N.E., 2000. Predictive habitat distribution models in ecology. *Ecol. Model.* 135, 147–186.
- Guisan, A., Edwards, T.C., Hastie, T., 2002. Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecol. Model.* 157, 89–100.
- Guisan, A., Broennimann, O., Engler, R., Vust, M., Yoccoz, N.G., Lehmann, A., Zimmermann, N.E., 2006. Using niche-based models to improve the sampling of rare species. *Conserv. Biol.* 20 (2), 501–511.
- Hess, H.E., Landolt, E., Hirzel, R., 1976. *Flora der Schweiz und angrenzender Gebiete*. Birkhäuser, Heidelberg.
- Huston, M., 1979. A general hypothesis of species diversity. *Am. Nat.* 113 (1), 81–101.
- Huston, M., 1994. *Biological Diversity. The Coexistence of Species on Changing Landscapes*. Cambridge University Press, New York.
- Hutchinson, G., 1957. Concluding remarks. *Cold Spring Harb. Symp. Quant. Biol.* 22, 415–427.
- Kaufmann, L., Rousseeuw, P.J., 1990. *Finding Groups in Data. An Introduction to Cluster Analysis*. Wiley & Sons, New York.
- Klotz, S., Kühn, I., Durka, W., 2002. *Biolfior – Eine Datenbank mit biologisch-ökologischen Merkmalen zur Flora von Deutschland*. Schriftenreihe für Vegetationskunde, 38.
- Kolb, A., Barsch, F., Diekmann, M., 2006. Determinants of local abundance and range size in forest vascular plants. *Glob. Ecol. Biogeogr.* 15, 237–247.
- Lavergne, S., Thompson, J.D., Garnier, E., Debussche, M., 2004. The biology and ecology of narrow endemic and widespread plants: a comparative study of trait variation in 20 congeneric pairs. *Oikos* 107, 505–518.
- Luoto, M., Virkkala, R., Heikkinen, R.K., 2007. The role of land cover in bioclimatic models depends on spatial resolution. *Glob. Ecol. Biogeogr.* 16, 34–42.
- Mac Nally, R., 2000. Regression and model-building in conservation biology, biogeography and ecology: the distinction between - and reconciliation of 'predictive' and 'explanatory' models. *Biodivers. Conserv.* 9, 655–671.
- Maggini, R., Lehmann, A., Zimmermann, N.E., Guisan, A., 2006. Improving generalized regression analysis for the spatial prediction of forest communities. *J. Biogeogr.* 33, 1729–1749.
- Manel, S., Williams, H.C., Ormerod, S.J., 2001. Evaluating presence-absence models in ecology: the need to account for prevalence. *J. Appl. Ecol.* 38, 921–931.
- McPherson, J.M., Jetz, W., Rogers, D.J., 2004. The effects of species' range sizes on the accuracy of distribution models: ecological phenomenon or statistical artefact? *J. Appl. Ecol.* 41, 811–823.
- Minchin, P.R., 1989. Montane vegetation of the Mt. Field massif, Tasmania: a test of some hypotheses about properties of community patterns. *Vegetatio* 83, 97–110.
- Oksanen, J., Minchin, P.R., 2002. Continuum theory revisited: what shape are species responses along ecological gradients? *Ecol. Model.* 157, 119–129.
- Parkhurst, D.F., Loucks, O.L., 1972. Optimal leaf size in relation to environment. *J. Ecol.* 60 (2), 505–537.
- Plattner, M., Birrer, S., Weber, D., 2004. Data quality in monitoring plant species richness in Switzerland. *Commun. Ecol.* 5 (1), 135–143.
- de Quervain, F., Hofmänner, F., Jenny, V., Köppl, V., Frey, D., 1963–1967. *Geotechnische Karte der Schweiz 1:200 000*. Kümmerli & Frei, Bern.
- R Development Core Team, 2006. *R: A Language and Environment for Statistical Computing*. R version 2.4.0.
- Rabinowitz, D., Cairns, S., Dillon, T., 1986. Seven forms of rarity and their frequency in the flora of the British Isles. In: Soulé, M.E. (Ed.), *Conservation Biology. The Science of Scarcity and Diversity*. Sinauer, Massachusetts, pp. 182–204.
- Rahbek, C., 2005. The role of spatial scale and the perception of large-scale species-richness patterns. *Ecol. Lett.* 8, 224–239.
- Rushton, S., Ormerod, S., Kerby, G., 2004. New paradigms for modelling species distribution? *J. Appl. Ecol.* 41, 193–200.
- Sarr, D., Hibbs, D., Huston, M.A., 2005. A hierarchical perspective of plant diversity. *Q. Rev. Biol.* 80 (2), 187–212.
- Schwarz, G., 1978. Estimating the dimension of a model. *Ann. Stat.* 6 (2), 461–464.
- Smith, T.M., 1997. Examining the consequences of classifying species into functional types: a simulation model analysis. In: Smith, T.M., Shugart, H.H., Woodward, F.I. (Eds.), *Plant Functional Types*. Cambridge University Press, Cambridge, pp. 319–340.
- Stockwell, D.R.B., Peterson, A.T., 2002. Effects of sample size on accuracy of species distribution models. *Ecol. Model.* 148, 1–13.
- ter Braak, C.J.F., 1985. Correspondence analysis of incidence and abundance data: properties in terms of a unimodal response model. *Biometrics* 41 (4), 859–873.
- Thuiller, W., Araujo, M.B., Lavorel, S., 2003. Generalized models vs. classification tree analysis: predicting spatial distributions of plant species at different scales. *J. Veg. Sci.* 14, 669–680.
- Thuiller, W., Lavorel, S., Midgley, G., Lavergne, S., Rebelo, T., 2004. Relating plant traits and species distributions along bioclimatic gradients for 88 *Leucadendron* taxa. *Ecology* 85 (6), 1688–1699.
- Tillé, Y., 2001. *Théorie des Sondages*. Dunod, Paris.
- Venier, L.A., McKenney, D.W., Wang, Y., McKee, J., 1999. Models of large-scale breeding-bird distribution as a function of macro-climate in Ontario, Canada. *J. Biogeogr.* 26, 315–328.
- Whittaker, R.H., 1956. *Vegetation of the Great Smoky Mountains*. *Ecol. Monogr.* 26 (1), 1–80.
- Willis, J.J., Whittaker, R.J., 2002. Species diversity: scale matters. *Science* 295, 1245–1248.
- Woodward, F.I., Kelly, C., 1997. Plant functional types: towards a definition by environmental constraints. In: Smith, T.M., Shugart, H.H., Woodward, F.I. (Eds.), *Plant Functional Types*. Cambridge University Press, Cambridge, United Kingdom, pp. 47–65.
- Zimmermann, N.E., Kienast, F., 1999. Predictive mapping of alpine grasslands in Switzerland: species versus community approach. *J. Veg. Sci.* 10, 469–482.
- Zimmermann, N.E., Edwards, T.C., Moisen, G.G., Frescino, T.S., Blackard, A., 2007. Remote sensing-based predictors improve distribution models of rare, early successional and broadleaf tree species in Utah. *J. Appl. Ecol.* 44, 1057–1067.