Preslia 84: 925–937, 2012

Identification of suitable unoccupied habitats: direct versus an indirect approach

Identifikace vhodných neobsazených stanovišť: srovnání přímých a nepřímých metod

Lucie Hemrová & Zuzana Münzbergová

Department of Botany, Faculty of Science, Charles University, CZ-128 01 Prague, Czech Republic, and Institute of Botany, Academy of Sciences of the Czech Republic, CZ-252 43 Průhonice, Czech Republic, e-mail: hemrova@natur.cuni.cz, zuzmun@natur.cuni.cz

Hemrová L. & Münzbergová Z. (2012): Identification of suitable unoccupied habitats: direct versus an indirect approach. – Preslia 84: 925–937.

Identification of habitats suitable for a species is a key task when studying landscape dynamics. Direct (cultivation experiments) or indirect (species distribution models) methods are employed to identify both suitable but currently unoccupied habitats and habitats that are unsuitable yet occupied (remnant populations). Although both direct and indirect approaches have been used increasingly in recent years the predictive power of cultivation experiments and of models based on different types of data have not been compared. In the present study, we compare the predictive power of distribution models for a short-lived herbaceous plant, *Jasione montana*. Our models are based on the environmental characteristics of the habitats, on the species composition of the habitats, or on both of these types of data. The predictions of the different models were compared (using AUC values) with the results of our cultivation experiment. We found that the models based on the species composition of the habitats performed better than the model based only on environmental characteristics. The models also differed significantly in the unoccupied habitats they identified as suitable. The most accurate was the model based on both environmental characteristics and species composition. This model also significantly explained both the presence/absence and abundance of *J. montana* individuals in the cultivation experiment. Nevertheless, the variation in occurrence of *J. montana* in the cultivation experiment explained by this model was below 50%. We therefore assume that the predictions of this model, in spite of the high AUC values, were inaccurate for at least some habitats. The results of this study are species and landscape specific, so they cannot be generalized. Our study, however, demonstrates that assembling data on both environmental characteristics and species composition of habitats is likely to be useful for predicting habitat suitability at a landscape scale. This study also demonstrates that a high AUC value is not a guarantee that a model's prediction is reliable because a cultivation experiment may provide different results. When identifying habitats that are suitable for a species (e.g. for the purpose of a metapopulation study), the results should be subjected to a sensitivity analysis.

Keywords: Beals index, community composition, distribution patterns, dry acidic grasslands, Ellenberg indicator values, patch occupancy, seed addition experiment

Introduction

Identification of habitats suitable for a species is an important task when studying its landscape dynamics (e.g. Ehrlén & Eriksson 2000, Münzbergová 2004, Guisan & Thuiller 2005, Herben et al. 2006). By identifying suitable habitats, it is possible to recognize habitats occupied due to extinction debt (Eriksson 1996). Analogically, we can identify habitats that are unoccupied because of the dispersal limitations of species (e.g. Ehrlén & Eriksson 2000, Turnbull et al. 2000, Münzbergová 2004).

Habitats suitable for a plant species can be recognized using a wide range of direct and indirect methods. The most common direct approach involves introduction of a species by sowing its seeds in occupied and unoccupied habitats (e.g. Ehrlén & Eriksson 2000, Turnbull et al. 2000, Münzbergová 2004, Ehrlén et al. 2006, Mildén et al. 2006, Moore & Elmendorf 2006). While this approach is considered useful in many cases it has several limitations. Firstly, habitat unsuitability can be proved only in the later stages of plant development (e.g. Turnbull et al. 2000, Ehrlén et al. 2006). Sowing experiments should therefore be continued until the plants have matured and reproduced, which may take up to several decades in the case of a perennial herbaceous plant (Ehrlén & Lehtilä 2002). Secondly, due to demographic stochasticity, a high initial number of seeds should be used in every sowing experiment. Thirdly, the success of sowing experiments depends on the year of establishment due to among-year variation in the species' survival (e.g. caused by weather conditions). Among-year variation may thus result in misleading results, so sowing experiments should be carried out repeatedly. Fourthly, sowing experiments cannot be done using invasive or endangered species. Finally, sowing experiments are time-consuming and require a lot of fieldwork. In many cases, these disadvantages result in indirect approaches being favoured over direct ones.

Indirect methods predict habitat suitability using the characteristics of both the habitats occupied and unoccupied by the species (e.g. Bastin & Thomas 1999, Thuiller et al. 2003, Münzbergová & Herben 2004, Liu et al. 2005, Moore & Elmendorf 2006, Wichmann et al. 2008). Compared to the direct approach, the indirect one requires less field work because a field survey is done only once. In addition, a lot of data on habitat characteristics can be obtained from various maps. One clear disadvantage of indirect methods is that the resulting predictions can depend on the frequency of species' occurrence (e.g. Gu & Swihart 2004, Segurado & Araújo 2004). Nevertheless, this problem can be solved, as suggested for example by Osawa et al. (2011). Another problem is that the species are assumed to occur primarily in suitable habitats. This assumption can result in imprecise predictions for species forming many remnant populations. The occurrence of remnant populations is more likely in long-lived species and species with extensive clonal growth (Eriksson 1996). Using a short-lived species with rapid responses to habitat deterioration can thus avoid this problem. And finally, the quality of a model depends on the right selection of parameters. This problem can be partly overcome by using the species composition of habitats to assess their suitability, since recorded species may reflect the same habitat characteristics that are also important for the target species (e.g. Münzbergová 2004, Münzbergová & Herben 2004, Guisan & Thuiller 2005, Ehrlén et al. 2006, Mildén et al. 2006, Moore & Elmendorf 2006). Species composition can, however, result from dispersal limitation, not habitat limitation, which limits the utility of the method. Similarly, the use of species composition is restricted in many species with remnant populations at the sites studied.

Species distribution models are usually developed using a range of abiotic and biotic environmental predictors. The predictors are usually obtained from maps or databases for larger study areas and from field data collected for smaller study areas (i.e. up to ca 100 km^2). Abiotic predictors include parameters characterizing soil and geology (e.g. Ehrlén & Eriksson 2000, Thuiller et al. 2003, Guisan et al. 2007, Wichmann et al. 2008), habitat configuration (Bastin & Thomas 1999), topography and bioclimatic conditions (e.g. Thuiller et al. 2003, Liu et al. 2005, Osawa et al. 2011, Rupprecht et al. 2011), habitat history and land use (Bastin & Thomas 1999, Rouget et al. 2001). The biotic predictors include parameters of particular species or vegetation such as occurrence, cover, height or leaf type (e.g. Münzbergová 2004, Guisan et al. 2007). In contrast to these predictors, detailed information on species composition of habitats has not often been exploited for making predictions about species occurrence (but see e.g. Dupré & Diekmann 1998, Münzbergová & Herben 2004, Tájek et al. 2011, Hemrová et al. 2012). It has been suggested that parameters derived from species composition might be better predictors of species occurrence than environmental predictors (e.g. Dupré & Diekmann 1998, Münzbergová 2004, Ehrlén et al. 2006, Elmendorf & Moore 2008). A comparison of the performance of models based (i) only on species composition, (ii) only on environmental parameters and (iii) both on species composition and environmental parameters with each other and with results of cultivation experiments is crucial for determining which method is best at predicting the suitability of habitats for species. Although all these approaches have been used repeatedly, no such comparison has been done.

In the present study, we therefore predict the presence of a species using indirect methods (i)–(iii) and compare the model predictions with the results of a cultivation experiment. We ask the following questions: (i) What is the influence of characteristics derived from species composition on model performance? (ii) Do models using different habitat characteristics detect the same suitable but unoccupied habitats? (iii) What is the correspondence between habitat suitability predicted using different habitat characteristics and the results of a cultivation experiment?

To answer these questions, we decided to estimate habitat suitability for the short-lived herbaceous plant *Jasione montana*. By choosing a short-lived species, we reduced both the time necessary for following the cultivation experiment and the risk of encountering remnant populations. For a set of dry acidic grassland habitats in a particular area, we collected data on abiotic conditions and species composition. We developed models predicting habitat suitability using data on abiotic conditions, on species composition, or on their combination. In several habitats, we also established a cultivation experiment. The results of the cultivation experiment were subsequently compared with the results of the models predicting habitat suitability.

Methods

Species and area studied

Jasione montana L. (*Campanulaceae*) is a moderately frequent, short-lived, monocarpic, rosette forming herbaceous plant inhabiting dry grasslands in most of Europe (Tutin 1976). According to Parnell (1985) individual plants of *J. montana* can produce up to 4335 seeds. As its seeds are small and light [0.6 mm, 19.4 μg (Kleyer et al. 2008), V_{term} = 0.7 $m \cdot s^{-1}$ (L. Hemrová, unpublished data)], it is likely they are dispersed by wind. Experiments in which the seed was attached to animal fur (roe deer, wild boar and brown hare), however, indicate the seed could be dispersed via exozoochory (L. Hemrová, unpublished data). Sixty percent of the seeds is viable based on a mix of 1200 seeds from three populations germinated in a growth chamber under the following conditions: 12 hours light and 20°C, 12 hours dark and 10°C (L. Hemrová, unpublished data). Seeds mature from September to October, and some of them start to germinate shortly afterwards. This species has a persistent seed bank (Thompson et al. 1997) but most of its mature seeds germinate the following spring or summer (L. Hemrová, unpublished data). *Jasione montana* has a scattered distribution in the study area. Most of the *J. montana* populations contained less than 50 flowering individuals and only one contained more than 100.

The study area (square and about 3 km^2; NW corner: 49°21'38"N, 13°39'24"E; SE corner 49°20'47"N, 13°41'04"E) is situated in a fragmented agricultural landscape in southwestern Bohemia in the Czech Republic. Dry acidic grasslands hosting *J. montana*, primary *Hyperico perforati-Scleranthion perennis* and *Koelerio-Phleion phleoidis* (Chytrý et al. 2001) are typical of the landscape. The grasslands are surrounded by agricultural fields or partly by shrubs, forests or paths. In the study area, *J. montana* is restricted to these grasslands.

Data collection

We identified the locations of all the dry acidic grasslands in the study area (N = 122) in 2005–2006. These habitats contained similar vegetation and were surrounded by visually obvious barriers such as forests, shrubs, fields or paths. To increase the sample size, we added to the dataset six habitats hosting *J. montana* that were outside the study area (≤ 7 km). In total, we surveyed 128 dry acidic grassland habitats (ranging from one to 3260 m^2), 26 of which hosted *J. montana*. Habitat occupancy was detected based on presence of *J. montana* anytime between 2005 and 2010.

To assess habitat suitability for *J. montana*, we recorded both abiotic conditions and species composition in each habitat in 2005–2006. Abiotic parameters were chosen to represent commonly assembled habitat characteristics such as slope, potential direct solar irradiation (PDSI) or soil properties (e.g. Rouget et al. 2001, Thuiller et al. 2003, Wichmann et al. 2008). Data on nine recorded abiotic parameters (Table 1) included: area (logarithm), slope, PDSI for the 21st day of the month from January to June (calculated using data on slope and aspect of given habitat; only the two least correlated were finally used), contact with a path (expressed as a percentage of habitat boundary with a path), presence of deep soil (> 10 cm), shallow soil (< 5 cm), gravel and stones in a habitat (Table 1). Other habitat characteristics were calculated using species composition.

One phytosociological relevé of 9, 4 or 1 m^2, depending on the area of the habitat, was randomly placed and recorded in each habitat. Based on the species composition, we first calculated Beals index. Beals index expresses the probability of a species' presence in a habitat based on the number of joint occurrences with other species in a dataset (Beals 1984, Münzbergová & Herben 2004). We then calculated mean Ellenberg indicator values for light, temperature, moisture, nutrients, soil reaction and continentality (Ellenberg 1988, Table 1) for each habitat.

In late October 2007, we established a cultivation experiment in 21 habitats situated in the study area, three of which hosted several flowering individuals of *J. montana*. This proportion of occupied habitats corresponded to the habitat occupancy in the study area (i.e. 20 out of 122 habitats were occupied). We selected the habitats to cover the whole range of habitat quality in the area. Twelve plots (30 × 30 cm each) arranged in a block (4 × 3 plots) were placed > 1 m from the closest flowering *J. montana* individual to reduce the number of seedlings originating from natural seed rain and < 3 m from the recorded phytosociological relevé. Before sowing, we mechanically disturbed six plots by removing

Table 1. – List of the habitat characteristics used to develop predictive models of habitat suitability for *Jasione montana* using (i) abiotic parameters, (ii) Beals index, (iii) Ellenberg indicator values and (iv) a combination of habitat characteristics. R^2 and effect (+/–) is shown for characteristics included in the model (values significant at P = 0.05 are highlighted in bold), n indicates characteristics not included in the model, x indicates characteristics excluded from the test.

Habitat characteristics	Type	Models			
		Abiot	Beals	Ellenberg	Full
Abiotic parameters					
Area	continous	+0.030	x	x	n
Slope	continous	+0.029	x	x	+0.023
PDSI					
March	continous	n	x	x	n
June	continous	n	x	x	n
Contact with path	continous	**+0.037**	x	x	**+0.080**
Deep soil	binomial	n	x	x	–0.020
Shallow soil	binomial	n	x	x	n
Gravel	binomial	+0.023	x	x	n
Stones	binomial	n	x	x	+0.024
Beals index	continous	x	**+0.433**	x	**+0.456**
Ellenberg indicator values					
Light	continous	x	x	+0.021	n
Moisture	continous	x	x	n	n
Soil reaction	continous	x	x	n	n
Nutrients	continous	x	x	**–0.044**	n
Temperature	continous	x	x	**–0.055**	n
Continentality	continous	x	x	**–0.033**	n

all above-ground vegetation. Three seed sources (originating from three external populations) were used in each block sown with seed, but the seeds were not mixed. Specifically, 300 seeds from one particular source were sown in one disturbed and 300 in one undisturbed plot in a block, so that six plots in a block were sown (3 sources × 2 plots). The remaining six plots were used as controls. The numbers of *J. montana* individuals in the cultivation experiment were recorded each September from 2008 to 2011.

Data analysis

To estimate habitat suitability, we tested for the effects of (i) abiotic parameters, (ii) Ellenberg indicator values and (iii) the combination of all habitat characteristics (including Beals index) on the presence of *J. montana*. Beals index was not used in a separate test, as its values directly represent the probability of the presence of *J. montana*. We used a generalized linear model (GLM) with a binomial distribution of the dependent variable (presence/absence of *J. montana*) in these tests. To simplify the models (see correlation matrix in Electronic Appendix 1), we used a step-wise bidirectional regression starting with the maximal model. We used the Akaike information criterion (AIC, Crawley 2002) to identify those habitat characteristics that best explained the presence of *J. montana*. We then calculated the probability of *J. montana* presence in individual habitats (Crawley 2002) for each developed model. In total, we had four models predicting *J. montana* presence: Beals index (model Beals), a model based on abiotic parameters (model Abiot), a model based on Ellenberg indicator values (model Ellenberg) and a model based on the combination of all habitat characteristics (model Full). To identify the habitats suitable for *J. montana*, we used the cut-off probability value in which the sum of sensitivity (i.e. the ratio of correctly predicted presences) and specificity (i.e. the ratio of correctly predicted absence) was maximized (Liu et al. 2005).

The performances of each model was evaluated by both calculating the area under the receiver operating characteristic curve (AUC, Swets 1988) and by comparing the suitability estimates with the results of the cultivation experiment. The receiver operating characteristic curve is obtained by plotting sensitivity vs (1–specificity) for various probability thresholds. The area under the curve may range from 0 for an inverse model, through 0.5 for a random model, to 1 for a perfect model. First, we tested the presence and abundance of *J. montana* individuals in the cultivation experiment using a GLM with a binomial and Poisson distribution, respectively. Predicted probabilities of a model based on habitat characteristics were used as an independent variable. Since no natural regeneration of *J. montana* was observed in the control plots, we tested both the total number of individuals and the presence of any individual in a sown block in (i) 2008, (ii) 2009 and (iii) 2010 and (iv), the sum of flowering individuals in a sown block from 2008 to 2011, and the presence of any flowering individual in a sown block anytime between 2008 to 2011. Results from 2011 could not be tested, as only a few individuals were recorded then in this experiment. We could not perform these tests for the plots with different disturbance regimes separately because only a few individuals were found in undisturbed plots. We, however, tested for the effect of disturbance regime on the number of *J. montana* individuals in the cultivation experiment. We used a GLM with Poisson distribution to do this. A manipulative disturbance regime (0/1) was used as an independent variable. As the dependent variable, we used the total number of individuals in the plots of a given regime in each block sown with seed in (i) 2008, (ii) 2009 and (iii) 2010, and (iv) the sum of flowering individuals in the plots of a given regime in each block from 2008 to 2011; the habitat code was used as a covariate. All analyses were performed in R 2.14.1. (R Development Core Team 2011).

Results

The Abiot model included four of the nine abiotic parameters tested: area, slope, contact with path and presence of gravel (df error = 123, $P = 0.01$, $R^2 = 0.11$, Table 1). The Ellenberg model included four of the six indicator values: light, nutrients, temperature and continentality (df error = 123, $P < 0.001$, $R^2 = 0.17$, Table 1). The Full model included five of the 16 test variables: slope, contact with path, presence of deep soil, presence of stones and Beals index (df error = 122, $P < 0.001$, $R^2 = 0.54$, Table 1).

The Beals model identified the least and the Ellenberg model most of the habitats suitable for but unoccupied by *J. montana* (Table 2, Fig. 1). Sixty-four percent of the suitable unoccupied habitats identified by the Beals model, 43% identified by the Abiot model and 29% identified by the Ellenberg model were also identified as suitable yet unoccupied by the Full model. AUC values largely differed among the models, the lowest being for the Abiot model (Table 2).

Table 2. – Number of suitable habitats that are not occupied by *Jasione montana*, AUC and correspondence of model predictions with the presence and number of *J. montana* individuals in the cultivation experiment in different years. R^2 are shown only for significant tests ($P < 0.05$; df error = 19); n.s. indicates non-significant tests.

Model	Number of habitats	AUC	*J. montana* number				*J. montana* presence			
			2008	2009	2010	Flowering 2008–2011	2008	2009	2010	Flowering 2008–2011
Abiot	15	0.72	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
Beals	12	0.88	n.s.	n.s.	n.s.	n.s.	**0.14**	n.s.	**0.42**	0.46
Ellenberg	18	0.79	n.s.	n.s.	**0.28**	0.29	n.s.	n.s.	n.s.	n.s.
Full	14	0.93	n.s.	n.s.	n.s.	**0.31**	**0.40**	n.s.	**0.22**	0.32

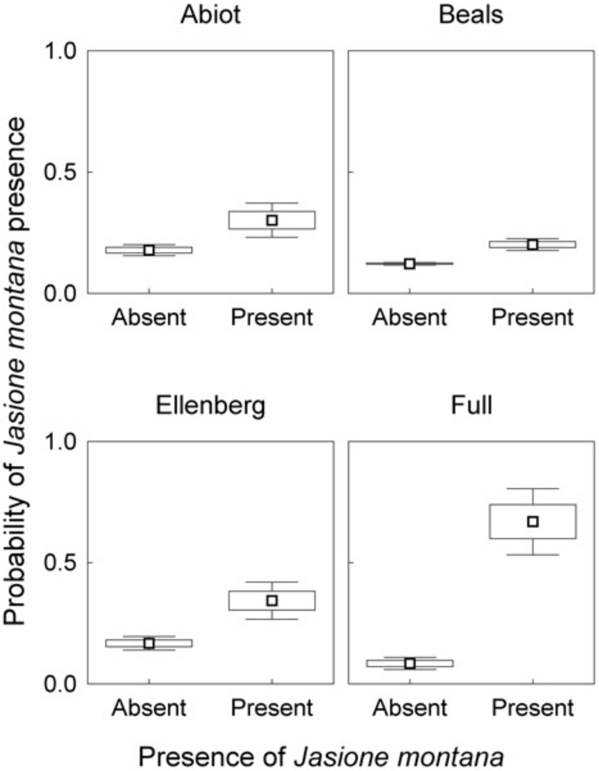

Fig. 1. – Probability of the presence of *Jasione montana* in individual habitats predicted by the models: Abiot (using abiotic parameters), Beals (Beals index values), Ellenberg (using Ellenberg indicator values) and Full (using a combination of habitat characteristics). Probabilities are shown for habitats both with and without recent presence of *J. montana* separately. Mean, SE and 1.96 SE are shown.

Jasione montana was successfully recruited in 13 of 21 habitats and survived until flowering in five habitats (including all three occupied habitats) of those sown with seed of this species. We recorded 102 vegetative individuals and no flowering individual of *J. montana* in 2008. Thirty two percent of the individuals survived from autumn 2008 to autumn 2009. We recorded 96 vegetative and one flowering individual in 2009, 29 vegetative and 15 flowering individuals in 2010 and 7 vegetative and 7 flowering individuals in 2011.

We found no significant relationship between the predictions of the Abiot model and both the presence and number of *J. montana* individuals in the cultivation experiment (Table 2). The Beals model significantly predicted the presence of *J. montana* individuals in the experiment in 2008 and 2010 and the presence of flowering individuals in this experiment (Table 2). The Ellenberg model significantly predicted the number of *J. montana* individuals in the experiment in 2010 and the number of flowering individuals in the experiment. Only the Full model significantly predicted both the presence and number of *J. montana* plants in the experiment. Specifically, the Full model significantly predicted the presence of *J. montana* individuals in 2008, in 2010 and the presence and the number of flowering individuals in the experiment (Table 2).

We also found a significant ($P < 0.05$) positive effect of disturbance on *J. montana* abundance in the cultivation experiment in 2008 (df error = 20, $P < 0.001$, $R^2 = 0.28$), 2009 (df error = 20, $P < 0.001$, $R^2 = 0.28$) and 2010 (df error = 20, $P = 0.005$, $R^2 = 0.03$, Fig. 2). Disturbance also had a significant positive effect on the total number of flowering *J. montana* individuals in the cultivation experiment (df error = 20, $P < 0.001$, $R^2 = 0.15$, Fig. 2).

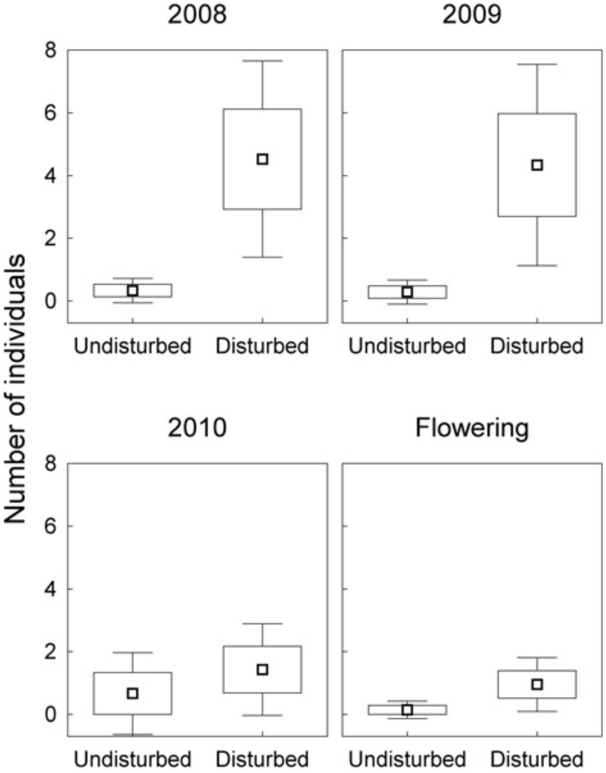

Fig. 2. – Number of individuals of *Jasione montana* in 2008, 2009 and 2010, and that of flowering individuals in 2008–2011 recorded in undisturbed and disturbed plots in the cultivation experiment. Mean, SE and 1.96 SE are shown.

Discussion

Models predicting habitat suitability

The importance of abiotic habitat characteristics for species occurrence is discussed in many previous studies (e.g. Bastin & Thomas 1999, Dupré & Ehrlén 2002, Thuiller et al. 2003). It has also been suggested that parameters derived from species composition such as Beals index or Ellenberg indicator values might be better predictors of species occurrence than abiotic habitat characteristics (e.g. Dupré & Diekmann 1998, Münzbergová 2004, Ehrlén et al. 2006, Elmendorf & Moore 2008). In this study, we found that habitat characteristics included in our models strongly influenced their predictions. Based on the classification of model performance suggested by Swets (1988), no model had a low accuracy ($0.5 < AUC < 0.7$), the models Abiot, Ellenberg and Beals were indicated as useful for some applications ($0.7 < AUC < 0.9$) and the Full model was very accurate ($AUC > 0.9$). Similarly, our comparison of the models' predictions with the results of the cultivation experiment revealed a low ability of the Abiot model to predict both the presence and abundance of *J. montana* individuals in the cultivation experiment. The Beals and Ellenberg models significantly predicted the presence and abundance of *J. montana* individuals in the cultivation experiment, respectively. Only the Full model significantly predicted both the presence and abundance of *J. montana* individuals in the cultivation experiment. This finding corroborates the results of previous studies by Tájek et al. (2011) and Hemrová et al. (2012) that identified the combination of abiotic characteristics and parameters derived from species composition as the best predictors of the presence of species in their studies. Our results indicate that it is difficult to preselect habitat characteristics that are important for species presence and demonstrate the potential advantage of parameters derived from species composition. We also found that habitat characteristics commonly used in species distribution models can prove insufficient for predicting species distribution. By contrast, species composition can reflect not only the current habitat quality and configuration but also past or present land use, which has repeatedly been shown to be important for many species (e.g. Lindborg & Eriksson 2004, Chýlová & Münzbergová 2008, Johansson et al. 2008). Parameters derived from species composition should therefore be included among habitat characteristics commonly used in species distribution models. Their different importance for the distribution of different species has to be kept in mind, however. Of the parameters derived from species composition, we highly recommend using Beals index because of its small number of degrees of freedom along with its high explanatory power (e.g. Münzbergová 2004, Ehrlén et al. 2006, Moore & Elmendorf 2006, Hemrová et al. 2012). Beals index can be calculated from both phytosociological relevés and species lists. As recording species composition can be difficult in large areas, the possibility of using the existing databases and past records of species occurrences to compute Beals index should be considered; however, the predictive power of this approach should be estimated when applied to large areas.

Landscape dynamics of Jasione montana

The Full model, which was identified as the best predictor of *J. montana* presence, identified 56% of the occupancy of suitable habitats by *J. montana* in the area. A similar degree of dispersal limitation of *J. montana* was also confirmed by the cultivation experiment (60% of suitable habitats occupied). It is suggested that habitat occupancy relates to the number of available habitats in the landscape (e.g. Husband & Barrett 1996, Bastin & Thomas 1999) and also to species life-history traits (e.g. Dupré & Ehrlén 2002, Jacquemyn et al. 2003, Verheyen et al. 2004, Tremlová & Münzbergová 2007). During the course of our study, we recorded a high percentage mortality of *J. montana* in the cultivation experiment. This experiment revealed that disturbance positively affected *J. montana* recruitment and survival. Similarly, the Abiot model and the Full model that included contact of habitats with paths, i.e. the parameter representing disturbance of the habitats. The importance of disturbance for species recruitment is in agreement with the findings of Parnell (1985), who observed *J. montana* in disturbed sites and of Weigelt et al. (2007) who identified *J. montana* as a poor competitor.

Reliability of models and the cultivation experiment

The Full model accurately predicted the presence of *J. montana* individuals in the landscape. This model explained more than half of the variation in *J. montana* presence in the landscape but only a third of the variation in both the abundance and presence of individuals of *J. montana* in the cultivation experiment. The discrepancy between the ability of the model to predict this species' presence in the landscape and the results of the cultivation experiment may have several explanations. Firstly, it could relate to partial suitability of the habitats, i.e. some of the plots were located in less suitable parts of the habitats. We suppose, however, that the predicted habitat suitability was reliable for the cultivation plots, since the blocks sown with seed were placed close to the phytosociological relevés and thus were representative of the habitat quality. Secondly, we could have sown an insufficient number of seeds, so the cultivation experiment consequently provided misleading results due to demographic stochasticity. This hypothesis is supported by the small number of flowering individuals recorded in this experiment (23 individuals in total in three occupied and two unoccupied habitats). We suppose, however, that *J. montana* flowered in all the suitable habitats included in this experiment because flowering *J. montana* were recorded in all the blocks sown with seed placed in habitats occupied by *J. montana*. We therefore assume that the results of the cultivation experiment were a good indicator of habitat suitability. We also suppose, however, that the model predictions for some habitats suffered from inaccuracy in spite of their high AUC values. When habitats suitable for a species are identified, for example, for the purpose of a metapopulation study, a sensitivity analysis of the model determination of habitat suitability should be performed.

We found that long-term rather than short-term results of the cultivation experiment were significantly explained by habitat suitability. Seedling abundance could not be used as good measure of habitat unsuitability but seedling survival and their subsequent reproduction can. This finding corroborates studies suggesting that long-term rather than short-term results of sowing experiments reflect habitat suitability (e.g. Turnbull et al. 2000, Münzbergová 2004, Ehrlén et al. 2006). Finally, *J. montana* is supposed to be an annual or biennial species (Tutin 1976). We therefore expected that a short time would be needed for *J. montana* to grow into flowering individuals. Unexpectedly, most flowering individuals were recorded in the third and fourth year of the experiment. This shows that a cultivation experiment can turn out to be more time-consuming than initially thought (Ehrlén et al. 2006), even for a short-lived species.

See http://www.preslia.cz for Electronic Appendix 1

Acknowledgements

We thank A. Lampei Bucharová, P. Dostál, T. Dostálek, V. Hadincová, T. Herben, J. Knappová, M. Macek, P. Petřík, H. Skálová, J. Vojta and the anonymous referees for their comments on previous versions of this paper, T. Herben for providing the program for calculation of PDSI and O. Beran for his help in the field. Tony Dixon kindly improved our English. This study was supported by grant no. 110010 from the Grant Agency of Charles University and grant no. P504/10/0456 from the Grant Agency of the Czech Republic. It was also partly supported by the long-term research development project no. RVO 67985939 from the Academy of Sciences of the Czech Republic.

Souhrn

Identifikace stanovišť vhodných pro daný druh je jednou z klíčových otázek při studiu dynamiky druhů na úrovni krajiny. Pro identifikaci vhodných stanovišť se využívá přímých (výsevový pokus) nebo nepřímých (predikce výskytu druhu pomocí modelů) metod. Obě tyto metody dokážou určit, která stanoviště jsou vhodná, ale prozatím neobsazená daným druhem, a naopak která stanoviště jsou nevhodná, ale druhem stále ještě obsazená. Přímé i nepřímé metody určení vhodných stanovišť se v posledních letech používají stále častěji, ale spolehlivost výsevového pokusu a modelů založených na různých typech dat nebyla zatím porovnána. V naší práci porovnáváme spolehlivost modelů využívajících buď jen charakteristiky prostředí, nebo jen druhové složení stanovišť, nebo oba zmíněné typy dat. Modely byly vytvořeny pro krátkověkou bylinu *Jasione montana* L. Spolehlivost modelů byla porovnávána pomocí hodnot AUC a výsledků výsevového pokusu. Zjistili jsme, že modely, které zahrnovaly parametry odvozené z druhového složení stanovišť, predikovaly výskyt *J. montana* v krajině lépe, než model založený pouze na charakteristikách prostředí. Modely se také značně lišily v určení vhodných neobsazených stanovišť. Výskyt *J. montana* v krajině nejlépe predikoval model zahrnující jak parametry odvozené z druhového složení, tak charakteristiky prostředí. Tento model také jako jediný vysvětlil jak přítomnost, tak i počet jedinců *J. montana* ve výsevovém pokusu. Protože však vysvětlená variabilita ve výskytu *J. montana* ve výsevovém pokusu činila méně než 50 %, došli jsme k závěru, že predikce tohoto modelu byla pro některá stanoviště nepřesná. Jsme si vědomi toho, že naše závěry o spolehlivosti modelů využívajících různé typy dat jsou těžko zobecnitelné na jiné druhy a jiné typy krajin, nicméně naše studie ukázala, že současné využití charakteristik prostředí a druhového složení stanovišť může být přínosné při určování jejich vhodnosti v krajinném měřítku. Jak charakteristiky prostředí, tak i druhové složení totiž mohou být důležitým ukazatelem vhodnosti stanovišť pro daný druh. Z naší studie také vyplynulo, že vysoké hodnoty AUC nejsou zárukou naprosté spolehlivosti modelu, neboť výsevový pokus může poskytnout odlišné výsledky. Pokud jsou tedy vhodná stanoviště identifikována pro účely navazující metapopulační studie, je třeba v takové studii zhodnotit vliv určení vhodných stanovišť na její výsledek.

References

Bastin L. & Thomas C. D. (1999): The distribution of plant species in urban vegetation fragments. – Landsc. Ecol. 14: 493–507.

Beals E. W. (1984): Bray-Curtis ordination: an effective strategy for analysis of multivariate ecological data. – Adv. Ecol. Res. 14: 1–55.

Chýlová T. & Münzbergová Z. (2008): Past land use co-determines the present distribution of dry grassland plant species. – Preslia 80: 183–198.

Chytrý M., Kučera T. & Kočí M. (eds) (2001): Katalog biotopů České republiky [Habitat catalogue of the Czech Republic]. – AOPK ČR, Praha.

Crawley M. J. (2002): Statistical computing: an introduction to data analysis using S-Plus. – Wiley, Chichester.

Dupré C. & Diekmann M. (1998): Prediction of occurrence of vascular plants in deciduous forests of south Sweden by means of Ellenberg indicator values. – Appl. Veg. Sci. 1: 139–150.

Dupré C. & Ehrlén J. (2002): Habitat configuration, species traits and plant distributions. – J. Ecol. 90: 796–805.

Ehrlén J. & Eriksson O. (2000): Dispersal limitation and patch occupancy in forest herbs. – Ecology 81: 1667–1674.

Ehrlén J. & Lehtilä K. (2002): How perennial are perennial plants? – Oikos 98: 308–322.

Ehrlén J., Münzbergová Z., Diekmann M. & Eriksson O. (2006): Long-term assessment of seed limitation in plants: results from an 11-year experiment. – J. Ecol. 94: 1224–1232.

Ellenberg H. (1988): Vegetation ecology of Central Europe. – Cambridge University Press, Cambridge.

Elmendorf S. C. & Moore K. A. (2008): Use of community-composition data to predict the fecundity and abundance of species. – Cons. Biol. 22: 1523–1532.

Eriksson O. (1996): Regional dynamics of plants: a review of evidence for remnant, source-sink and metapopulations. – Oikos 77: 248–258.

Gu W. D. & Swihart R. K. (2004): Absent or undetected? Effects of non-detection of species occurrence on wildlife-habitat models. – Biol. Cons. 116: 195–203.

Guisan A. & Thuiller W. (2005): Predicting species distribution: offering more than simple habitat models. – Ecol. Lett. 10: 435–435.

Guisan A., Zimmermann N. E., Elith J., Graham C. H., Phillips S. & Peterson A. T. (2007): What matters for predicting the occurrences of trees: techniques, data, or species' characteristics? – Ecol. Monogr. 77: 615–630.

Hemrová L., Červenková Z. & Münzbergová Z. (2012): The effects of large herbivores on the landscape dynamics of a perennial herb. – Ann. Bot. (in press, doi: 10.1093/aob/mcs064).

Herben T., Münzbergová Z., Mildén M., Ehrlén J., Cousins S. A. O. & Eriksson O. (2006): Long-term spatial dynamics of *Succisa pratensis* in a changing rural landscape: linking dynamical modelling with historical maps. – J. Ecol. 94: 131–143.

Husband B. C. & Barrett S. C. H. (1996): A metapopulation perspective in plant population biology. – J. Ecol. 84: 461–469.

Jacquemyn H., Butaye J. & Hermy M. (2003): Influence of environmental and spatial variables on regional distribution of forest plant species in a fragmented and changing landscape. – Ecography 26: 768–776.

Johansson L. J., Hall K., Prentice H. C., Ihse M., Reitalu T., Sykes M. T. & Kindström M. (2008): Semi-natural grassland continuity, long-term land-use change and plant species richness in an agricultural landscape on Oland, Sweden. – Landsc. Urban. Plan. 84: 200–211.

Kleyer M., Bekker R. M., Knevel I. C., Bakker J. P., Thompson K., Sonnenschein M., Poschlod P., van Groenendael J. M., Klimeš L., Klimešová J., Klotz S., Rusch G. M., Hermy M., Adriaens D., Boedeltje G., Bossuyt B., Dannemann A., Endels P., Götzenberger L., Hodgson J. G., Jackel A-K., Kühn I., Kunzmann D., Ozinga W. A., Römermann C., Stadler M., Schlegelmilch J., Steendam H. J., Tackenberg O., Wilmann B., Cornelissen J. H. C., Eriksson O., Garnier E. & Peco B. (2008): The LEDA Traitbase: a database of life-history traits of Northwest European flora. – J. Ecol. 96: 1266–1274.

Lindborg R. & Eriksson O. (2004): Historical landscape connectivity affects present plant species diversity. – Ecology 85: 1840–1845.

Liu C., Berry P. M., Dawson S. & Pearson R. G. (2005): Selecting thresholds of occurrence in the prediction of species distributions. – Ecography 28: 385–393.

Mildén M., Münzbergová Z., Herben T. & Ehrlén J. (2006): Metapopulation dynamics of a perennial plant, *Succisa pratensis*, in an agricultural landscape. – Ecol. Model. 199: 464–475.

Moore K. A. & Elmendorf S. C. (2006): Propagule vs. niche limitation: untangling the mechanisms behind plant species' distributions. – Ecol. Lett. 9: 797–804.

Münzbergová Z. (2004): Effect of spatial scale on factors limiting species distributions in dry grassland fragments. – J. Ecol. 92: 854–867.

Münzbergová Z. & Herben T. (2004): Identification of suitable unoccupied habitats in metapopulation studies using co-occurrence of species. – Oikos 105: 408–414.

Osawa T., Mitsuhashi H., Uematsu Y. & Ushimaru A. (2011): Bagging GLM: improved generalized linear model for the analysis of zero-inflated data. – Ecol. Inform. 6: 270–275.

Parnell J. A. N. (1985): Biological flora of the British Isles: *Jasione montana* L. – J. Ecol. 73: 341–358.

R Development Core Team (2011): R: a language and environment for statistical computing. – R Foundation for Statistical Computing, Vienna, URL: http://www.R-project.org.

Rouget M., Richardson D. M., Lavorel S., Vayreda J., Gracia C. & Milton S. J. (2001): Determinants of distribution of six *Pinus* species in Catalonia, Spain. – J. Veg. Sci. 12: 491–502.

Rupprecht F., Oldeland J. & Finckh M. (2011): Modelling potential distribution of the threatened tree species *Juniperus oxycendrus*: how to evaluate the predictions of different modelling approaches? – J. Veg. Sci. 22: 647–659.

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

Swets J. A. (1988): Measuring the accuracy of diagnostic systems. – Science 240: 1285–1293.

Tájek P., Bucharová A. & Münzbergová Z. (2011): Limitation of distribution of two rare ferns in fragmented landscape. – Acta Oecol. 37: 495–502.

Thompson K., Bakker J. P. & Bekker R M. (1997): The soil seed banks of North West Europe: methodology, density and longevity. – Cambridge University Press, Cambridge.

Thuiller W., Vayreda J., Pino J., Sabate S., Lavorel S. & Gracia C. (2003): Large-scale environmental corelates of forest tree distributions in Catalonia (NE Spain). – Glob. Ecol. Biogeogr. 12: 313–325.

Tremlová K. & Münzbergová Z. (2007): Importance of species traits for species distribution in fragmented landscapes. – Ecology 88: 965–977.

Turnbull L. A., Crawley M. J. & Rees M. (2000): Are plant populations seed-limited? A review of seed sowing experiments. – Oikos 88: 225–238.

Tutin T. G. (1976): *Jasione*. – In: Tutin T. G., Heywood V. H., Burges N. A., Valentine D. H., Walters S. M. & Webb D. A. (eds), Flora Europaea 4: 100–102, Cambridge University Press, Cambridge.

Verheyen K., Vellend M., Van Calster H., Peterken G. & Hermy M. (2004): Metapopulation dynamics in changing landscapes: a new spatially realistic model for forest plants. – Ecology 85: 3302–3312.

Weigelt A., Schumacher J., Walther T., Bartelheimer M., Steinlein T. & Beyschlag W. (2007): Identifying mechanisms of competition in multi-species communities. – J. Ecol. 95: 53–64.

Wichmann M. C., Alexander M. J., Hails R. S. & Bullock J. M. (2008): Historical distribution and regional dynamics of two *Brassica* species. – Ecography 31: 673–684.

Received 15 March 2012
Revision received 18 June 2012
Accepted 27 June 2012