Ecological Specialization Indices for species of the Czech flora

Indexy ekologické specializace pro druhy české flóry

David Zelený[1] & Milan Chytrý[2]

[1]*Institute of Ecology and Evolutionary Biology, National Taiwan University, Roosevelt Rd. 1, 106 17 Taipei, Taiwan, e-mail: zeleny@ntu.edu.tw;* [2]*Department of Botany and Zoology, Masaryk University, Kotlářská 2, 611 37 Brno, Czech Republic, e-mail: chytry@sci.muni.cz*

Zelený D. & Chytrý M. (2019) Ecological Specialization Indices for species of the Czech flora. – Preslia 91: 93–116.

Theoretically the concept of species ecological specialization is very useful, however, practically it is often difficult to quantify due to a lack of relevant environmental data. We introduce the Ecological Specialization Index (ESI), which describes the degree of specialization of a species based on its realized niche along multiple environmental gradients and is conceptually based on the co-occurrence specialization metric theta introduced by Fridley et al. (2007). We estimated ESI for species of the Czech flora occurring in at least 10 vegetation plots stored in the Czech National Phytosociological Database. We prepared three sets of ESI values calculated from three datasets including (i) plots of all vegetation types (ESI_w, 1597 species), (ii) only plots of non-forest vegetation (ESI_{nf}, 1529 species), and (iii) only plots of forest vegetation (ESI_f, 881 species). We also provide the frequency of species in the datasets, since the reliability of the calculated ESI values increases with the species frequency. The use of these ESI values is limited to the Czech Republic, and in the case of less frequent species, the value can be influenced by sampling bias. To facilitate understanding of the ecological meaning of ESI, we related the calculated values of ESI_w to several species attributes and applied them in a case study using a local vegetation dataset from a deep river valley. We found that ESI correlates significantly with specialization metrics based on the number of phytosociological associations and habitats in which the focal species occur. The species listed in the national Red List in higher risk categories are on average more specialized than less threatened species. Neophytes tend to be significantly less specialized than archaeophytes and native species. When related to Ellenberg-type indicator values for the Czech Republic, specialists tend to be more shade-tolerant, better adapted to nutrient-poor soils and soils with either a low or high (but not intermediate) pH and to either warm or cold (but not intermediate) habitats. In a case study of herbaceous plants species in a forest understory on river valley slopes, we found that specialists tend to be confined to deeper soils on cooler north-facing slopes, to stony soils in ravine forests and sites with a denser canopy of woody species. In contrast, shallow lithic soils on eroded south-facing slopes and sites with a more open canopy tend to be dominated by generalists. The complete list of ESI values is included in an electronic appendix to this paper.

Keywords: Czech Republic, ecological amplitude, flora, generalists, realized niche breadth, specialists, theta, vascular plants, vegetation-plot database.

Introduction

Specialist species (specialists) are those that are restricted to specific habitats, adapted to using a narrow range of resources or tolerating a narrow range of environmental conditions. Conversely, generalist species (generalists) occur in many different habitats, utilize

doi: 10.23855/preslia.2019.093

a wide range of resources or are tolerant of a broad range of environmental conditions. There is a natural trade-off between being either specialist or generalist; as put by MacArthur (1972), "jack of all trades is a master of none". Although generalists can use a wide range of resources, they are not particularly good at using any of them efficiently. Generalists and specialists are relative terms: they mark the opposite extremes in a continuum of ecological specialization, but most species are somewhere between these extremes.

Ecologists predict various aspects of the ecology of a species depending on the degree of their ecological specialization. For example, specialist species are expected to be more sensitive to habitat change (both natural and human-induced) than generalists. Therefore, specialists are more susceptible to population reduction or even extinction (Walker & Preston 2006), which gives them on average a higher conservation priority than generalists. Because of their link to a narrow range of habitats, specialists are also likely to be better indicators of habitat quality than generalists (Ellenberg et al. 1991, Chytrý et al. 2018). While the concept of ecological specialization is theoretically sound and attractive, in practice the degree of specialization of a species is difficult to quantify. A commonly used method is to link species occurrences to one or several environmental gradients and quantify a species niche breadth in terms of the variation in the position of the species on this gradient (e.g. using species response curves, see Coudun & Gégout 2005, Hájková et al. 2008, Wagner et al. 2017). Unfortunately, high-quality environmental measurements for quantifying species niche breadth are usually limited to a few proxy variables that are easy to measure but do not have a direct causal effect on plant growth and other physiological processes. A further issue is that a single species can be a specialist along one gradient and a generalist along another one.

An alternative approach, which we use in this study, is to quantify the species niche breadth using indirect estimations of habitat qualities in which the species occurs. Fridley et al. (2007) introduced a method of measuring species habitat specialization, which does not require data about environmental factors at sites where the species are recorded. The method uses the pattern of co-occurrence of the focal species with other species in the community across a number of sites and requires the availability of a (possibly large) dataset of community samples (records of species co-occurring at individual sites) from a broad range of different habitats (called source dataset in this paper). It tracks the occurrence of a focal species in many samples and identifies with which species the focal species co-occurs at different sites. The species composition of these co-occurring species quantifies the local habitat conditions at each site. If we choose several samples that contain the focal species, and these samples differ considerably in species composition, it is likely that they also differ in habitat conditions. Then, a focal species that is able to grow in a wide range of habitat conditions is likely to be a generalist. Conversely, if the samples containing the focal species have similar species compositions, their habitat conditions are probably also similar, and the focal species is likely to be a specialist. Unlike measures of specialization based on estimated species amplitudes along a measured environmental variable, the co-occurrence based measure takes into consideration many environmental factors. The differences in species composition, i.e. beta diversity, among samples containing the focal species can be measured in different ways. The original metric θ proposed by Fridley et al. (2007) was based on additive beta diversity (Lande 1996) with high values for generalists and low values for specialists. Here, following the suggestion of Zelený (2009), we replace the additive measure of beta diversity by a multiplicative measure (Whittaker 1960) and introduce a more intuitive measure called Ecological Specialization Index (ESI), calculated as an inverted θ-value. ESI increases with actual increase in species specialization, and its values are real numbers in the range 0–9. Because of the way ESI is calculated, its values represent realized species niche breadth taking into account many environmental factors, approximated by the species composition and are related to the concept of realized niche as a hyper-volume in the multidimensional space defined by ecological variables, in which species can maintain a viable population (Hutchinson 1957).

The co-occurrence based index of ecological specialization critically depends on the quality and quantity of underlying compositional data in the source dataset and their geographical distribution. We used the Czech National Phytosociological Database (Chytrý & Rafajová 2003), which contains community samples from all vegetation types occurring in the Czech Republic. However, the different contexts in which the ESI values are used may require different approaches to their definition. In particular, the degree of specialization of the same species may differ between forest and non-forest habitats. For example, some species of open habitats may behave as specialists when growing in forests while behaving as generalists when growing outside forests. Therefore we calculated three sets of ESI values: across all vegetation types, for non-forest vegetation only, and for forest vegetation only.

In this paper, we describe a new dataset of Ecological Specialization Indices that we developed for species of the Czech flora and offer it for ecological analyses. To get a better insight into the ecological interpretation of the ESI values, we related them to the attributes of selected species like taxon origin (native vs alien), Red List categories and Ellenberg-type indicator values for the Czech flora. We also compared them with alternative ways of quantifying ecological specialization using compositional data, represented by the number of phytosociological associations and habitat types in which the species occur. Finally, in order to illustrate the potential application of the ESI values in vegetation studies we used a local vegetation dataset with several measured environmental variables.

Materials and methods

Vegetation data

The source dataset used in this analysis is a subset of the Czech National Phytosociological Database (Chytrý & Rafajová 2003), which at the time of data preparation (2013) contained community samples from 93,704 vegetation plots. We selected only plots with geographical coordinates (92,249 plots), removed records of non-vascular plants and juvenile individuals of woody plants and merged separate records of the same species in the tree and shrub layer to ensure that each species occurs only once in each plot. Species nomenclature follows Danihelka et al. (2012). Then we geographically stratified the database to reduce oversampling of similar vegetation types within some regions (Knollová et al. 2006). Stratification was done by assigning the plots into cells of a geographic grid of 0.75' of latitude and 1.25' of longitude (approx. 1.5 × 1.4 km) and applying an heterogeneity-constrained random (HCR) resampling procedure to choose a fixed maximum number of plots of distinct species composition from each cell (Lengyel et al.

2011). Within each grid cell, we calculated Bray-Curtis dissimilarity among all pairs of plots (using log-transformed percentage covers in species composition data), and applied the HCR resampling to select the optimal subset of plots retaining maximum mean pairwise dissimilarity among selected plots. Since the grid cells may contain plots belonging to a wide range of different vegetation types, we implemented the rule proposed by Wiser & De Cáceres (2013) that more plots are selected from grid cells where plots have higher compositional heterogeneity, with the minimum and the maximum number of selected plots 5 and 20, respectively. The resulting dataset, called the 'whole dataset' throughout this study, contains 67,453 vegetation plots and 2071 species, sampled between 1923 and 2012 by 574 different researchers. More than half of these plots (34,070) were assigned to one of 494 phytosociological associations, out of the 496 recognized in the Czech Republic according to Chytrý (2007–2013). We further divided the original dataset into a subset including only plots representing non-forest vegetation (72,719 plots) and those representing forest and scrub vegetation (19,530 plots). As forest we considered plots in which the sum of the canopy cover of the different species of trees was greater than 25%, with the exception of the tree species that form a light canopy (*Betula* sp., *Larix decidua*, *Pinus* sp., *Populus tremula* and *Taxus baccata*) for which the threshold was decreased to 15%. For each of the subsets, we also calculated the HCR resampling using the same method and parameters as in the case of the whole dataset, resulting in the 'non-forest dataset' (54,579 plots with 2034 species) and 'forest dataset' (16,397 plots with 1437 species).

Ecological Specialization Index

We developed the Ecological Specialization Index (ESI) based on the co-occurrence metric θ (Fridley et al. 2007). The θ metric for the focal species is a beta diversity of a subset of vegetation plots containing this species. Since beta diversity is high if the focal species is a generalist and low if it is a specialist, we use a simple formula to invert these values into a more intuitive ESI, with values increasing with the degree of species specialization.

As a measure of beta diversity, we used Whittaker's multiplicative beta (β_W, Whittaker 1960) calculated as $\beta_W = \gamma / \overline{\alpha}$, where γ is the cumulative number of species in a set of n plots and $\overline{\alpha}$ is the mean number of species in this set of plots. We based this choice on the study of Zelený (2009), which showed that the original algorithm of Fridley et al. (2007), using an additive beta diversity measure (Lande 1996), generates specialization values that are dependent on the size of the species pool of a community in which the focal species mostly occurs (species occurring in communities with large species pools tend to be identified as generalist more often than species occurring in communities with small species pools). Although other alternative metrics have been proposed (e.g. Manthey & Fridley 2009, Botta-Dukát 2012, Boulangeat et al. 2012), our choice of β_W is justified by extensive evaluation of different metrics using simulated data of various properties (Zelený, unpublished results).

Since β_W depends on the number of plots (n) from which it is calculated, and the size of the subset of plots containing the focal species is different for different species, β_W calculated for different species would not be comparable. Therefore, we used sample-based rarefaction to estimate the total number of species occurring on average in a subset of 10 randomly selected plots from the dataset containing the focal species (γ_{10}), and divided the estimated value by the mean number of species in individual plots ($\overline{\alpha}$). Note that this procedure is analogous to the original algorithm proposed by Fridley et al. (2007), which is based on subsampling, i.e. random selection of n subplots (n = 10 in our case) from the subset of plots containing the focal species, repeating this subsampling many times, and averaging beta diversities calculated for each subsample. Instead of subsampling with a high number of randomizations, we calculated both γ_{10} and $\overline{\alpha}$ analytically using the sample-based rarefaction equation (e.g. Ugland et al. 2003), which returns virtually identical results while requiring much less computational time.

Since β_W is known to be highly sensitive to the presence of few community samples with very different species composition in the dataset (Manthey & Fridley 2009), we followed the suggestion of Botta-Dukát (2012) and removed these plots (outliers) prior to calculating β_W. We used the distance-based algorithm of McCune & Mefford (1999) to calculate the mean of pairwise compositional dissimilarities of the focal plot to all the other plots in the subset, i.e. mean value of a single column (or row) within a symmetric pairwise dissimilarity matrix without diagonal values (zeros). Since this mean dissimilarity is relatively high in the case of outliers, we removed the plots with mean dissimilarity higher than mean dissimilarity between all pairs of plots plus two standard deviations of its variation. As a measure of dissimilarity, Botta-Dukát (2012) suggested the use of Euclidean distance applied to presence-absence data. However, we opted for Sørensen dissimilarity (one complement of Sørensen coefficient; Sørensen 1948), which is compatible with β_W used to calculate the overall beta diversity of the subset (β_W for two samples is equal to Sørensen dissimilarity between these two samples minus one; Tuomisto 2010).

Theoretically, the values of β_W are in units of the number of distinct communities with no overlap in species composition. For a subset of 10 plots, the β_W value ranges between 1 and 10 (1 if all plots have identical species composition, and 10 if these plots share no species). We inverted β_W into ESI using the formula $\text{ESI} = 10 - \beta_W$; in this way, high ESI values indicate specialists and low ESI values indicate generalists. Note that the ESI metric is continuous, theoretically ranging between 0 and 9, although the extreme values are unlikely to occur.

In summary, the steps in the calculation of ESI for the focal species are: (i) Select a subset of vegetation plots containing the focal species from the whole dataset. (ii) Remove outliers from this subset using the distance-based algorithm with Sørensen dissimilarity. (iii) Calculate the beta diversity of the subset using β_W by rarefying the dataset to 10 plots. (iv) Repeat steps 1–3 for all species with at least 10 occurrences in the whole dataset. (v) Recalculate the β_W values into ESI using the formula $\text{ESI} = 10 - \beta_W$. (vi) Repeat steps 1–5 for non-forest and forest datasets. Note that two species frequency thresholds apply: one determines the minimum number of plots with the focal species that are included in the analysis after removing potential outliers (freq_{min}), and the other is the number of plots from which β_W is calculated (n). Theoretically should apply that $\text{freq}_{\text{min}} \geq n$; in this study, both thresholds are set to 10.

Comparison of ESI with other attributes of species

To illustrate the patterns described by ESI, we compared the ESI values for individual species calculated based on the whole dataset (ESI_w) with several other attributes of these species.

Species occurrence frequency. We calculated the Spearman rank correlation between ESI and species frequency, defined as the number of vegetation plots in which the focal species is recorded within the whole dataset.

The number of phytosociological associations in which the species occurs. Some of vegetation plots used to calculate the ESI were assigned to one of 494 phytosociological associations representing nearly the complete variation in the species composition of Czech vegetation. With some simplification, each of these associations represents a different combination of habitat conditions, and the specialist species should be expected to occur in fewer associations than generalists. Since different associations are represented by different numbers of plots in the dataset, we quantified the number of associations in three alternative ways: (i) the number of associations in which the focal species occurs (A_{occ}), (ii) the number of associations in which the focal species occurs in a large proportion of the vegetation plots of that association (A_{rel}), and (iii) the number of associations in which the focal species occurs in a high absolute number of vegetation plots of that association (A_{abs}). The main difference is that A_{occ} does not consider how many plots of the given association the focal species occurs in while A_{rel} and A_{abs} do. While A_{occ} is calculated as a simple count of associations in which the focal species occurs, A_{rel} and A_{abs} are calculated using the exponential of the Shannon entropy index (Shannon 1948), also known as Shannon diversity index (Chao et al. 2014). The Shannon entropy index is calculated as $H' = -\Sigma (p_i \ln p_i)$, where p_i is the relative abundance of species i and quantifies the diversity of species in a community while considering both the number of species and the differences in their relative abundances. The Shannon diversity index uses the exponential function $e^{H'}$ to convert the units from entropy into "the number of effective species", which is equivalent to the number of common species in the case of the Shannon diversity index (Jost 2006). In this study, we replaced "the number of effective species" by "the number of effective phytosociological associations", i.e. those in which the focal species occurs more frequently, measured either by relative proportions or absolute counts. Parameter p_i represents either the normalized (i.e. divided by the sum of all values) proportion of plots of the association in which the focal species occurs (A_{rel}) or normalized absolute number of plots of the association in which the focal species occurs (A_{abs}; see Appendix 1 for equations). The logic of this calculation is that the association with a higher relative (A_{rel}) or absolute (A_{abs}) number of plots containing the focal species contributes more to the overall number of associations containing this species. We calculated Spearman's rank correlation between ESI_w and A_{occ}, A_{rel} and A_{abs}, respectively. Since common species may have a higher probability of occurring in more associations just because of their higher overall frequency, we also calculated partial Spearman's correlation between each pair of variables (i.e. ESI_w with A_{occ}, ESI_w with A_{rel} and ESI_w with A_{abs}) while controlling for the frequency of each species in the dataset ($freq_w$).

The number of habitats in which the species occurs. Similarly to the number of associations, the number of habitats in which the focal species occurs can also serve as a measure of species niche breadth (e.g. Chytrý et al. 2005, Pyšek et al. 2009). Sádlo et al. (2007) compiled a dataset of occurrence of species of the Czech flora in 88 different habitats. The assessment of the affinities of species to these habitats was first done by a statistical analysis of a large number of vegetation plots from the Czech National Phytosociological Database and then extensively revised based on the expert judgement. The preference of species for each of these 88 habitats was quantified on a scale from 1 to 4, with 1 meaning ‘species occurrence’ (species can grow in the habitat, but it is not its optimum one), 2 meaning ‘species optimum’ (the habitat or a part of it is the ecological optimum for this species), and 3 and 4 meaning species optimum combined with different degrees of dominance (high cover) in the habitat. For each species, we counted ‘the number of habitats in which the focal species occurs’ (H_{occ}, i.e. number of habitats classified as 1–4 for that species) and ‘the number of habitats in which the species has its optimum’ (H_{opt}, i.e. the number of habitats classified as 2–4). We calculated Spearman’s rank correlation coefficient between ESI_w and H_{occ} or H_{opt}, respectively, and partial Spearman’s rank correlation between each pair while controlling for species frequency in the dataset ($freq_w$).

Species origin. We classified species into archaeophytes (alien species introduced before the year 1500), neophytes (alien species introduced after that year) and native following Pyšek et al. (2012). We tested the differences in ESI values between these three categories using one-way ANOVA, followed by Tukey’s HSD tests.

Red List category. We classified species into the IUCN Red List categories (IUCN 2012) according to the national Red List (Grulich 2017). The IUCN Red List categories include: RE – regionally extinct, CR – critically endangered, EN – endangered, VU – vulnerable, NT – near threatened, LC – least concern, DD – data deficient, and NE – not evaluated. We excluded the DD category (eight species) and merged the categories LC and NE into LC, because species not included in the previous national Red List (i.e. species with the lowest risk of extinction) were not evaluated by Grulich (2017), assuming that nearly all of them would fall within the LC category. We used one-way ANOVA followed by Tukey’s HSD multiple comparison tests ($P < 0.05$) to test the differences in ESI values between the Red List categories.

Ellenberg-type indicator values for the Czech Republic. Ellenberg indicator values quantify species optima for particular environmental factors using a simple ordinal scale (Ellenberg et al. 1991). We used a dataset of Ellenberg-type indicator values adopted for the Czech Republic (Chytrý et al. 2018). This dataset also includes species not present in the original Ellenberg’s tables and contains adjusted values for some other species. Czech indicator values are based on an extensive expert revision that considered species ecological preferences within the Czech Republic and outside the country, supported by a statistical analysis of data from the Czech National Phytosociological Database (Chytrý & Rafajová 2003). The values are given on the same scale as the original Ellenberg values and are available for light, temperature, moisture, reaction, nutrients and salinity, but not for continentality (see Berg et al. 2017 for reasons).

Case study: Vltava valley forest dataset

We demonstrate the use of ESI in ecological studies on the relationship of specialist and generalist plants to environmental variables using a local dataset of forest vegetation sampled in a wide range of habitats. The dataset contains 97 vegetation plots of 10 m × 15 m recorded by the first author of this paper along transects in deep sections of the Vltava river valley in southern Bohemia, Czech Republic (Zelený & Chytrý 2007). Covers of all the vascular plants and a set of environmental variables related to topography, soil properties and canopy cover were recorded in each plot. Topographical variables included altitude (reflecting the height above the valley bottom), slope and folded aspect (deviation from 22.5°), heat load (calculated from slope and aspect using the formula of

McCune & Keon 2002) and within-plot terrain convexity in the vertical and horizontal direction. Soil variables included soil depth, soil pH and soil type including lithic soils (shallow soils on exposed rocky outcrops), stony soils (mostly in ravines), cambisols (deep, well-developed soils on gentle slopes) and fluvisols (deep soils on the river floodplain). Some plots contained more than one soil type. Cover of trees and shrubs was estimated visually as a proxy of light availability in the forest understory. Only species in the herb layer were included in the analysis. Since most of the plots were sampled in the forest and only a few in canopy openings, we used the forest Ecological Specialization Index (ESI_f) for the analysis. The relationship between species ESI_f and environmental variables was analysed using two complementary approaches: the community weighted mean approach and the fourth-corner approach. Species composition data were transformed into presences-absences prior to the analysis.

We calculated the community weighted mean (CWM) of the Ecological Specialization Index for each vegetation plot as the mean ESI_f of species in the plot (since the species composition data were transformed to presences-absences, weights of all species are the same). We related the calculated CWM values to each environmental variable using Pearson's r correlation coefficient and tested its significance by a permutation test with 49,999 permutations. Because standard tests of correlation between CWM and sample attributes (in this case environmental variables) have inflated Type I error rate and generate overly optimistic results (Peres-Neto et al. 2017), we used the 'max test' strategy introduced by ter Braak et al. (2012). The max test combines the row-based and column-based permutation tests by choosing the higher P-value as the resulting significance. We used the max test to analyse the relationship between CWM of ESI_f and each of 13 environmental variables separately, and adjusted the P-values for multiple comparisons using the false discovery rate correction (FDR; Benjamini & Hochberg 1995). Following Dray et al. (2014), we used a high number of permutations (49,999) to get enough power for the P-value correction.

The fourth-corner approach is an alternative to the CWM approach in that it relates the matrix of species attributes (here ESI_f) directly to sample attributes (environmental variables) by inflating the matrix of species composition, without calculating CWM (Legendre et al. 1997, Dray & Legendre 2008). Similar to CWM approach, there are several alternative permutation schemes to test the significance of the fourth-corner correlation, and the only one which controls for Type I error rate is also the max test strategy (ter Braak et al. 2012). As in the CWM approach, we used max test with 49,999 permutations and adjusted P-values with FDR correction. Instead of the original fourth-corner correlation, which practically cannot reach the –1 and 1 values (Legendre et al. 1997), we report the relationship based on a Chessel fourth-corner correlation, which is the original form of fourth-corner correlation rescaled into the range from –1 to 1 (see Peres-Neto et al. 2017 for details). In fact, the fourth-corner and CWM approach are numerically closely related (Peres-Neto et al. 2017). Since the CWM approach is perhaps more often used in vegetation ecology (Zelený 2018), while the fourth-corner approach is considered as more powerful (ter Braak et al. 2018), we used both methods in parallel to assess the relationship of ESI_f to environmental variables.

Software details

Data editing and plot identification by the expert system were done using JUICE software (Tichý 2002). All the other calculations and plotting were done using the R program (R Core Team 2017). Geographical stratification of the dataset by the HCR resampling method was calculated using the function hcr from the package vegclust (De Cáceres et al. 2010). ESI values were calculated using the library theta (Zelený, unpublished, https://github.com/zdealveindy/theta), which is a generalization of the original R code for θ-value calculation developed by J. Fridley (Fridley et al. 2007, Electronic Appendix 2), for calculating θ using various beta-diversity metrics. Partial Spearman's rank correlation was calculated using the function pcor in the package ppcor (Kim 2015), and Tukey's HSD post-hoc test using the function HSD.test in the package agricolae (de Mendiburu 2017). In the case study of the Vltava valley dataset, we tested the relationship between ESI_f and environmental variables using the functions test_cwm and test_fourth in the weimea package (Zelený, unpublished, https://github.com/zdealveindy/weimea). The R code for all the analyses is stored in the GitHub repository (https://github.com/zdealveindy/esi_czech).

Results

Three sets of ESI values for species with 10 or more occurrences were calculated: from the whole dataset (ESI_w, 1597 species), non-forest dataset (ESI_{nf}, 1529) and forest dataset (ESI_f, 881 species). Non-forest and forest dataset shared 829 species. For each set of ESI values, we also provide the frequency of each species in the particular dataset ($freq_w$, $freq_{nf}$ and $freq_f$). Frequencies can be used to evaluate the reliability of the calculated ESI values because the more samples the species occurs in, the more reliable is the information about its co-occurrence pattern. The minimum frequency in all the three datasets was set to 10 occurrences (the threshold to calculate beta diversity), while the maximum was 12,656 for the whole dataset (frequency of *Achillea millefolium* agg.), 12,442 for the non-forest dataset (also *Achillea millefolium* agg.) and 5090 for the forest dataset (*Oxalis acetosella*). This frequency can be used as a threshold for selecting species for analyses (e.g. only species with frequency > 50 if more reliable ESI values are required). The values of ESI for each dataset (and corresponding species frequencies) are listed in Electronic Appendix 1 at www.preslia.cz.

The ESI values calculated from the whole dataset ranged from 2.83 to 8.37 with a median of 4.97 (Fig. 1A). For both the non-forest and forest dataset, the range and median values were similar (2.68–7.54, median 4.93 for non-forest, and 3.15–7.49, median 5.08 for forest). The 10 most generalist and most specialist species (occurring in at least 50 plots) for each dataset are in Table 1, listed separately for ESI of each dataset.

Comparisons of ESI_w with other species attributes showed that this index is negatively correlated with species frequency in the dataset (Fig. 1B, Spearman's $\rho = -0.30$) and with the number of associations in which the species occurs (A_{occ}), prevails (A_{rel}) or dominates (A_{abs}; Fig. 1C). The correlation of ESI with A_{abs} is the strongest ($\rho = -0.58$), being weaker for A_{occ} ($\rho = -0.46$) and A_{rel} ($\rho = -0.45$). Partial correlation, which controls for the effect of species frequency in the dataset, returned values that are somewhat higher than those of non-partial correlations ($\rho_{part} = -0.59$ for A_{abs}, –0.59 for A_{occ} and –0.46 for A_{rel}). ESI_w is

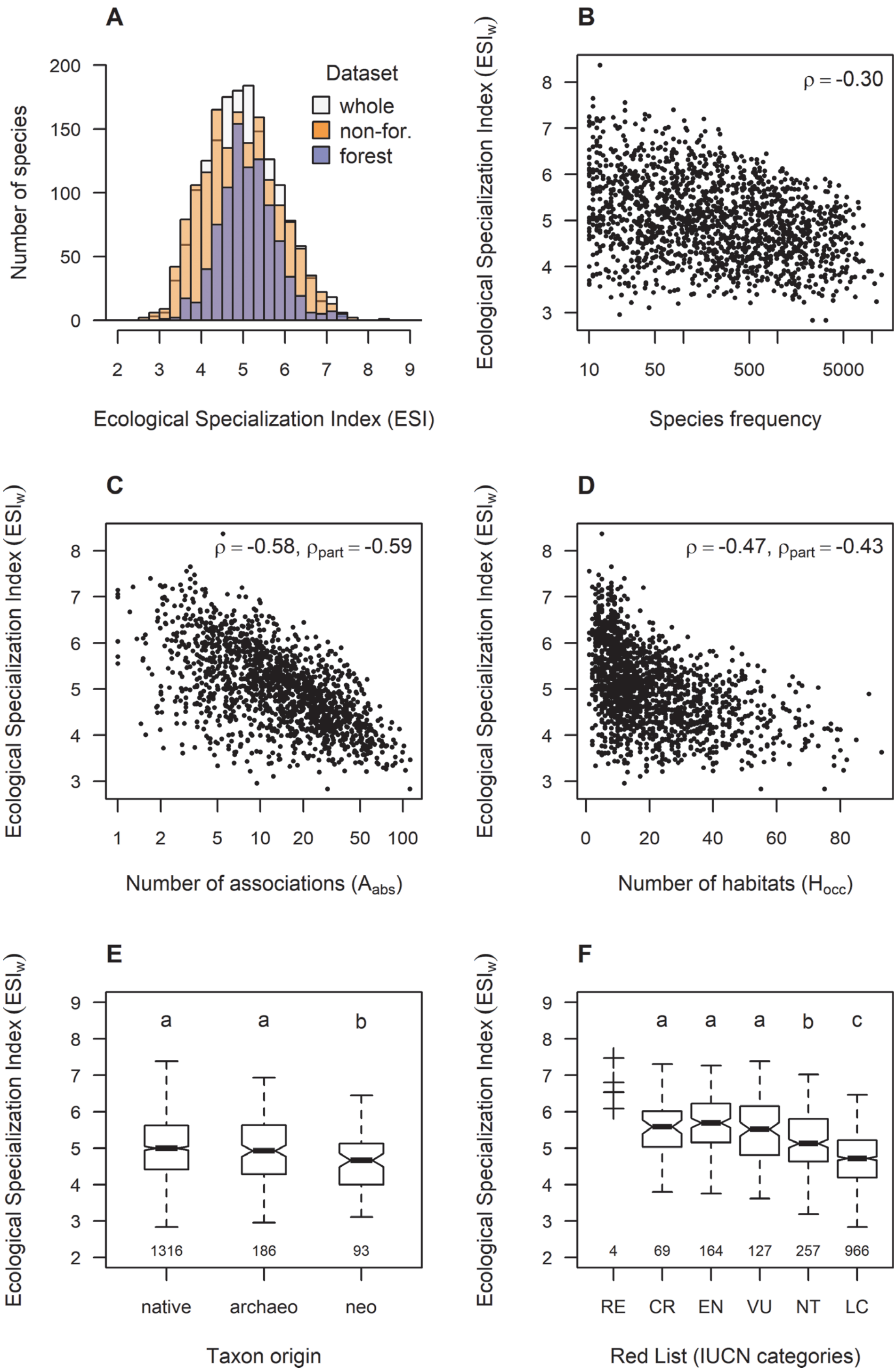
A
Number of species
200
150
100
50
0
2
3
4
5
6
7
8
9
Ecological Specialization Index (ESI)
Dataset
whole
non-for.
forest
B
Ecological Specialization Index (ESI_w)
ρ = -0.30
8
7
6
5
4
3
10
50
500
5000
Species frequency
C
Ecological Specialization Index (ESI_w)
ρ = -0.58, ρ_part = -0.59
8
7
6
5
4
3
1
2
5
10
20
50
100
Number of associations (A_abs)
D
Ecological Specialization Index (ESI_w)
ρ = -0.47, ρ_part = -0.43
8
7
6
5
4
3
0
20
40
60
80
Number of habitats (H_occ)
E
Ecological Specialization Index (ESI_w)
9
8
7
6
5
4
3
2
a
a
b
1316
186
93
native
archaeo
neo
Taxon origin
F
Ecological Specialization Index (ESI_w)
9
8
7
6
5
4
3
2
a
a
a
b
c
4
69
164
127
257
966
RE
CR
EN
VU
NT
LC
Red List (IUCN categories)

Table 1. – The list of species with the highest and lowest Ecological Specialization Index values calculated from the whole (ESI_w), non-forest (ESI_{nf}) and forest datasets (ESI_f). Ten species with the highest ESI values (specialists) and lowest ESI values (generalists) are listed for each dataset (specialists sorted by descending and generalists by ascending ESI values). The frequency of species in each dataset is given in brackets. Only species with at least 50 occurrences in a particular dataset were considered.

Whole dataset		Non-forest dataset		Forest dataset	
Species name	ESI_w ($freq_w$)	Species name	ESI_{nf} ($freq_{nf}$)	Species name	ESI_f ($freq_f$)
Top specialists					
Pinus strobus	7.40 (78)	*Empetrum nigrum* agg.	7.19 (153)	*Pinus strobus*	7.23 (78)
Empetrum nigrum agg.	7.20 (164)	*Trifolium rubens*	7.17 (101)	*Rhododendron tomentosum*	7.23 (85)
Pinus uncinata subsp. *uliginosa*	7.17 (102)	*Festuca psammophila* subsp. *dominii*	7.03 (60)	*Pinus uncinata* subsp. *uliginosa*	7.22 (112)
Chamaecytisus austriacus	7.02 (69)	*Coleanthus subtilis*	6.97 (235)	*Andromeda polifolia*	7.19 (69)
Rhododendron tomentosum	7.00 (106)	*Pulmonaria angustifolia*	6.95 (57)	*Vaccinium uliginosum*	6.78 (230)
Festuca psammophila subsp. *dominii*	7.00 (64)	*Chamaecytisus austriacus*	6.95 (71)	*Dryopteris expansa*	6.65 (59)
Coleanthus subtilis	6.97 (234)	*Veronica orchidea*	6.93 (63)	*Vaccinium oxycoccos* agg.	6.57 (192)
Trifolium rubens	6.96 (109)	*Senecio aquaticus*	6.86 (175)	*Eriophorum vaginatum*	6.56 (337)
Senecio aquaticus	6.94 (170)	*Andromeda polifolia*	6.85 (214)	*Polystichum aculeatum*	6.47 (206)
Conringia orientalis	6.94 (51)	*Elatine triandra*	6.83 (157)	*Carex pendula*	6.39 (50)
Top generalists					
Calamagrostis epigejos	2.83 (3230)	*Phragmites australis*	2.87 (2263)	*Convolvulus arvensis*	3.15 (53)
Phragmites australis	2.83 (2329)	*Pinus sylvestris*	2.89 (383)	*Elymus repens*	3.53 (186)
Verbascum densiflorum	3.19 (88)	*Betula pendula*	2.93 (520)	*Poa compressa*	3.54 (58)
Senecio viscosus	3.21 (512)	*Calamagrostis epigejos*	2.93 (2508)	*Vicia cracca*	3.54 (111)
Glyceria fluitans	3.22 (1894)	*Rubus* sect. *Rubus*	2.99 (1567)	*Chenopodium album* agg.	3.54 (78)
Urtica dioica	3.24 (11137)	*Populus tremula*	3.01 (137)	*Tanacetum vulgare*	3.64 (57)
Phalaris arundinacea	3.27 (3879)	*Poa nemoralis*	3.05 (1604)	*Cirsium arvense*	3.69 (127)
Persicaria amphibia	3.29 (1448)	*Frangula alnus*	3.06 (254)	*Poa annua*	3.69 (63)
Asplenium ruta-muraria	3.32 (351)	*Hieracium murorum*	3.07 (576)	*Pinus sylvestris*	3.71 (2530)
Rubus caesius	3.34 (1692)	*Fraxinus excelsior*	3.15 (331)	*Arrhenatherum elatius*	3.87 (449)

◀ Fig. 1. – (A) Distribution of Ecological Specialization Index values calculated using the whole, non-forest and forest datasets, respectively. (B–F) Descriptive statistics of the Ecological Specialization Index calculated from the whole dataset (ESI_w) and their relationship to various species attributes. (B) Correlation between ESI_w and the frequency of species occurrence in the whole dataset. (C) Correlation between ESI_w and the number of phytosociological associations of the national vegetation classification system (Chytrý 2007–2013) in which the focal species often occurs, calculated as an exponential of Shannon entropy index (A_{abs}). ρ – Spearman's correlation of ESI and A_{abs}, ρ_{part} – partial Spearman's correlation of ESI and A_{abs} while controlling for species frequency ($freq_w$). (D) Correlation between ESI and the number of habitats in which the species occurs (Sádlo et al. 2007); ρ – Spearman's correlation of ESI and H_{occ}, ρ_{part} – partial Spearman's correlation of ESI and H_{occ} while controlling for species frequency ($freq_w$). (E) Relationship of ESI to taxon origin (Pyšek et al. 2012); archaeo – archaeophytes, neo – neophytes. (F) Relationship of ESI to species classification in the national Red List, using the IUCN categories (Grulich 2017); RE – regionally extinct (with ESI value indicated by +), CR – critically endangered, EN – endangered, VU – vulnerable, NT – near threatened and LC – least concern. All analyses (Spearman's rank correlation in B, C and D, and ANOVA in E and F) were significant at $P < 0.001$. The letters above the barplots in E and F indicate whether Tukey's HSD post hoc comparison between categories was significant at the level of $P < 0.05$ or not. If two notches drawn on the boxes do not overlap, this is an indication that the medians of these groups differ significantly.

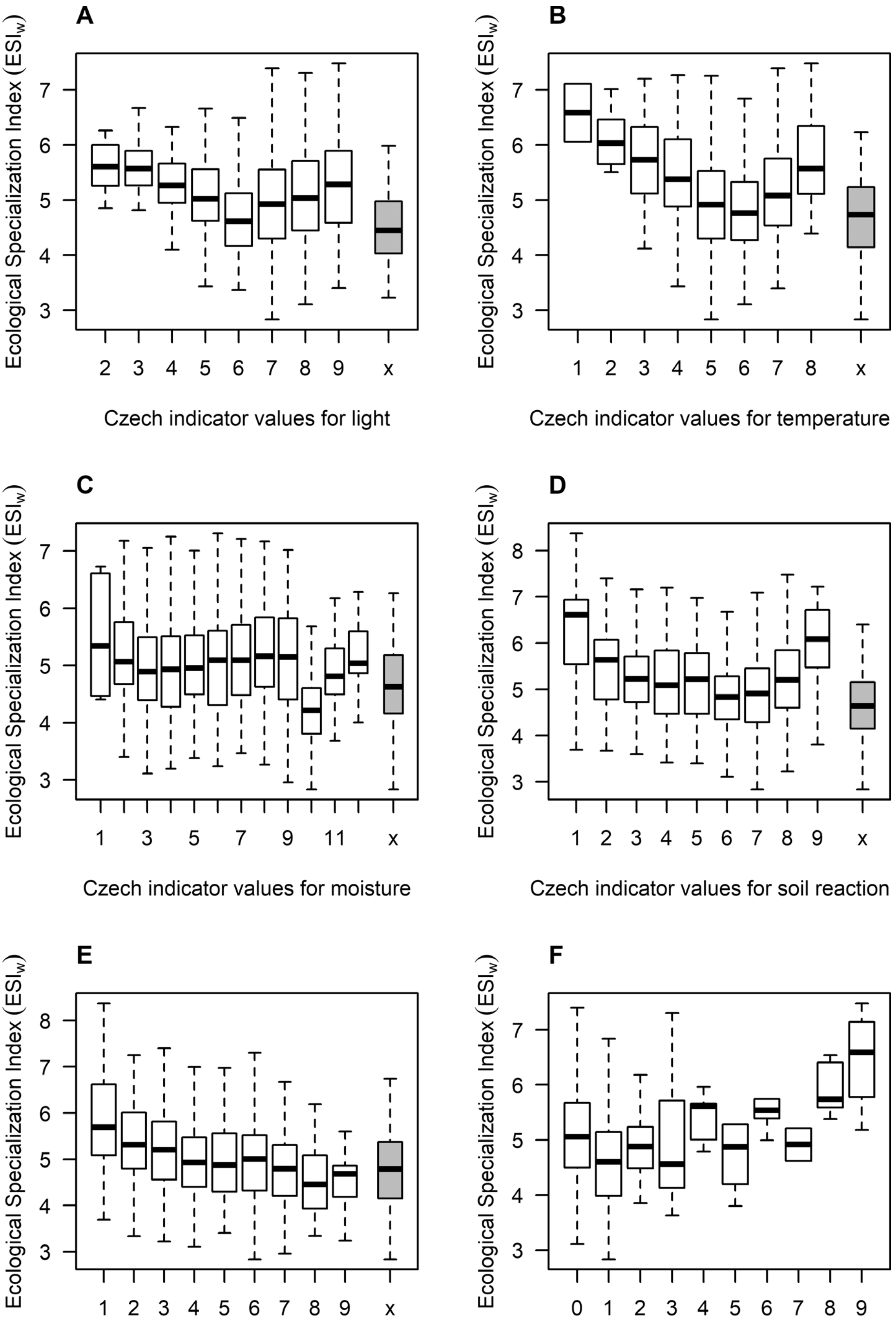

Fig. 2. – Relationships between the Ecological Specialization Index calculated from the whole dataset (ESI_w) and Ellenberg-type indicator values for the Czech flora. Species with x value are considered as generalists regarding the particular environmental gradient according to the national indicator value dataset.

also negatively correlated with the number of habitats in which a species occurs (H_{occ}, $\rho = -0.47$, Fig. 1D) or has its optimum (H_{opt}, $\rho = -0.37$), although controlling for the effect of species frequency slightly weakened these correlations ($\rho_{part} = -0.43$ for H_{occ} and –0.30 for H_{opt}).

Considering species origin, neophytes (93 species) were significantly more generalist than archaeophytes (186 species) and native species (1316 species) (Fig. 1E; global ANOVA test: $F = 9.0$, $P < 0.001$; pairwise HSD test significant at $P < 0.05$). Species classified in the IUCN Red List categories CR (critically endangered, 69 species), EN (endangered, 164 species) and VU (vulnerable, 127 species) were significantly more specialized than those classified in category NT (near threatened, 257 species). LC (least concern) species (966) were significantly less specialized than species classified in any other IUCN category (Fig. 1F; global ANOVA test, $F = 88.8$, $P < 0.001$; pairwise HSD test significant at $P < 0.05$). The four species in category RE (regionally extinct), *Gentianella germanica*, *Salicornia prostrata*, *Suaeda prostrata* and *Triglochin maritima*, were not included in the ANOVA analysis (but are included in Fig. 1F).

There are various patterns in the relationships of ESI to Ellenberg-type indicator values for the Czech flora. In the case of indicator values for light, the species adapted to shady habitats tended to be more specialized, while ESI values for light-demanding species ranged broadly (Fig 2A). In the case of temperature and soil reaction (Fig. 2B, D) there was a remarkable u-shaped response, with species adapted to the extreme conditions being more specialized than species adapted to intermediate conditions. ESI was negatively related to nutrients, with species adapted to nutrient-poor habitats being more specialized and nutrient-demanding species more generalist (Fig. 2E). In contrast, ESI did not show a clear relationship to moisture except for species with a moisture indicator value of 10 being more generalist than the others (Fig. 2C). Species tolerant of a high or extremely high salt content in the soil are more specialist than species tolerating no or low salt contents in soil (Fig. 2F). Species without assigned Ellenberg-type indicator values (indicated as ‘x’ on Fig. 2) tend to be more often generalists, with the exception of indicator values for nutrients, where their ESI_w values overlap with the range of ESI_w for more nutrient-rich soils.

The case study of forest vegetation in the Vltava valley revealed a significant ($P < 0.05$) or marginally significant ($P < 0.1$) relationship of ESI to several environmental variables (interpretation based on FDR-corrected P-values for either CWM or fourth-corner approach, Table 2). Results of the community weighted mean approach are more conservative than results of the fourth-corner approach, with only two marginally significant ($P < 0.1$) results compared to eight significant ($P < 0.05$) results of the fourth-corner approach (Table 2). More specialized species occur in cooler, less sun-exposed habitats (negative relationship with a folded aspect and heat load), on deeper soils with higher pH and in shaded conditions under a closed canopy (positive relationship with soil depth, soil pH and the cover of tree and shrub layer). Higher specialization was also recorded in communities occurring on stony soils, especially in ravine forests on the lower parts of the valley slopes (positive relationship with the occurrence of stony soils), while species with low specialization prefer lithic soils, especially on exposed convex upper slopes (negative relationship with the occurrence of lithic soils and vertical terrain convexity).

Table 2. – Relationship between Ecological Specialization Index values calculated for forest vegetation ($\mathrm{ESI_f}$) and selected environmental variables, using data from a local study of forest vegetation in the Vltava river valley. The relationship was tested by correlating community weighted mean of $\mathrm{ESI_f}$ with environmental variables (CWM approach) and by relating $\mathrm{ESI_f}$ directly to environmental variables using the fourth-corner metric (fourth-corner approach). r_{CWM} – Pearson's correlation coefficient between CWM of $\mathrm{ESI_f}$ and environmental variables; r_{Chessel} – Chessel fourth-corner correlation; P – the significance of correlation using the P_{max} test; P_{adj} – P values adjusted for multiple comparisons (false discovery rate correction). Significant ($P < 0.05$) adjusted P-values are printed in bold, marginally significant ($P < 0.1$) in italics. Species composition data were presence-absence transformed prior to the analysis.

	CWM approach			Fourth-corner approach		
Environmental variable	r_{CWM}	P	P_{adj}	r_{Chessel}	P	P_{adj}
Altitude	–0.176	0.384	0.454	–0.073	0.179	0.211
Slope	–0.188	0.245	0.319	–0.068	0.217	0.236
Folded aspect (deviation from 22.5°)	–0.519	0.005	*0.064*	–0.254	<0.001	**<0.001**
Heat load	–0.266	0.145	0.241	–0.119	0.027	**0.043**
Vertical terrain convexity	–0.383	0.052	0.112	–0.166	0.002	**0.004**
Horizontal terrain convexity	–0.201	0.187	0.270	–0.081	0.139	0.180
Lithic soil	–0.360	0.048	0.112	–0.163	0.002	**0.004**
Stony soil	0.350	0.027	0.104	0.176	0.001	**0.002**
Cambisol	0.089	0.510	0.510	0.015	0.786	0.786
Fluvisol	0.157	0.476	0.510	0.102	0.106	0.154
Soil depth	0.420	0.011	*0.070*	0.196	0.000	**0.001**
Soil pH	0.311	0.148	0.241	0.123	0.023	**0.042**
Cover of tree and shrub layer	0.369	0.032	0.104	0.186	0.001	**0.002**

Discussion

Measuring niche breadth of plant species: benefits and caveats

Experienced vegetation ecologists usually have a good empirical knowledge of the species that tend to behave more as specialists or generalists based on field observations. Here we provide an intuitive specialization index, which aims to quantify the breadths of species realized niches based on the pattern of co-occurrence of the focal species with other species in vegetation plots in which the focal species occurs (Fridley et al. 2007, Zelený 2009, Botta-Dukát 2012).

Indeed, one needs to bear in mind that any measure of species specialization is context-dependent. If specialization is quantified in terms of niche breadth based on specific environmental factors, then it is only valid in this context, since species can be a specialist along one but generalist along other gradients. The ESI values provided in this study reflect realized species niche breadth along multiple gradients, but still, they are dependent on the source dataset from which they were calculated, specifically on its spatial and compositional context, because it influences the relative importance of individual environmental variables (e.g. Siefert et al. 2012). We used a dataset that includes the range of variation in vegetation recorded throughout the Czech Republic and calculated ESI values separately for all vegetation types, non-forest vegetation and forest vegetation. The absolute and relative values of ESI would have differed if calculated from a dataset representing different spatial extent (either smaller, e.g. a certain area or locality, or larger, e.g. central Europe), or different subset of vegetation types (e.g. only broad-leaved forest vegetation as in Zelený et al. 2010 or Marinšek et al. 2015, only non-forest vegetation as in Fajmonová et al. 2013, or only synanthropic vegetation as in Šilc et al. 2014). Therefore, researchers planning to use the ESI values presented here need to consider whether the context in which they were calculated is relevant to the study they are conducting.

The dependence of ESI on the quality of the source dataset also means that the values may be negatively affected by sampling bias. If samples from some habitats are underrepresented in the dataset, then some species with a broad niche may appear as more specialized. An extreme example is *Pinus strobus*, an introduced North American species planted for timber that is invasive in sandstone areas in the northern part of the Czech Republic (Hadincová et al. 2008, Pyšek et al. 2017). The original dataset contains 121 plots with *Pinus strobus*, of which 85 were selected by geographical stratification. From these plots, most (67) were sampled in the Bohemian Switzerland National Park by a single author, while the remaining plots are mostly forestry plantations with scattered occurrence across the whole of the Czech Republic. The calculation of $\mathrm{ESI_w}$ for this species was based on 78 plots, since some were removed as outliers because their species composition differed too much. As a result, *Pinus strobus* was identified as one of the most specialized in our dataset with $\mathrm{ESI_w} = 7.40$ (considering $\mathrm{ESI_w}$ with $\mathrm{freq_w} \geq 10$).

In contrast, some of the species identified as generalists based on our calculation may actually be more specialized. This can happen, for example, if a species is a specialist of a fine-scale habitat, which is often a part of a mosaic with other habitats. For example, *Asplenium ruta-muraria*, a specialist species of calcareous rocky outcrops and walls, occurs on large cliffs with specialized chasmophytic species and on small outcrops within grasslands or forests. As a result, this species is considered a generalist based on ESI ($\mathrm{ESI_w} = 3.32$), because it co-occurs in vegetation plots with many species with different ecological requirements, which decreases its ESI value.

The ESI values calculated in this study are estimates of the size of the realized niche of species (Hutchinson 1957). They result from the interaction of three factors: (i) species fundamental niche, (ii) availability of suitable habitats for the focal species in the study area, and (iii) biotic interactions with other species. The fundamental niche is determined by species physiological limits that are a result of evolution. The actual availability of habitats in a study area influences whether a species is categorized as a generalist or a specialist, since even a generalist that is potentially able to occupy a wide range of habitats may appear to be a specialist if most of its suitable habitats do not occur in the area (Wagner et al. 2017). Biotic interactions such as competition or facilitation are also important determinants of a species realized niche, although they are difficult to measure directly. Biotic interactions depend on the composition of the regional species pool, namely the presence or absence of competitors that would narrow the realized niche or facilitators that would broaden it (Pulliam 2000). The fundamental niche of a taxonomically homogeneous species is constant across different regions, while realized niche may differ as a result of between-region differences in habitat availability and species pool composition, influencing biotic interactions. (Coudun & Gégout 2005, Hájková et al. 2008, Wasof et al. 2013, Wagner et al. 2017). Thus, the specialization values recorded in this study are specific to the Czech Republic and not directly applicable elsewhere. The regional validity of ESI values is analogous to the regional validity of species indicator values (e.g. those of

Ellenberg et al. 1991), which may need to be recalibrated when used in regions other than those for which they were initially proposed (e.g. Borhidi 1995, Hill et al. 2000, Pignatti et al. 2005, Chytrý et al. 2018).

Comparisons of the Ecological Specialization Index with other species attributes

Although we were unable to revise the calculated ESI values for each species, comparisons with other species attributes suggest that they provide ecologically meaningful information. The ESI values correlate with other measures of species specialization, such as the number of (effective) phytosociological associations or habitats in which the focal species occurs. In principle, all three metrics evaluate species specialization according to differences in the species composition of plots in which the focal species occurs, measured either as beta diversity (ESI_w, ESI_{nf}, ESI_f), the number of floristically distinct vegetation units (A_{occ}, A_{rel}, A_{abs}) or the number of ecologically distinct habitats (H_{occ}, H_{opt}). The ESI values are more tightly correlated with the association-based metrics, representing finely divided vegetation types, than with the habitat-based metrics, representing coarsely divided vegetation types. In this sense, they seem to describe realized niche breadth with finer resolution than could be obtained by simply determining the broadness of the habitats in which the species occur. This study could use a consistent and comprehensive phytosociological vegetation classification (Chytrý 2007–2013, 2017) and habitat classification that are both available for the Czech Republic (Sádlo et al. 2007, Chytrý et al. 2010). Therefore, ecological specialization of species in this country can be measured in different ways. However, since such classifications are not available in many other countries, national ESI datasets derived from national vegetation databases (Dengler et al. 2011, Chytrý et al. 2016, Bruelheide et al. 2019) provide a realistic alternative.

Species listed in the national Red List in higher risk categories are on average more specialized than less threatened species. Threatened species are often specialists of certain habitats, and their decline is caused by the decline in this habitat. The IUCN Red List criteria also consider the decline in area, extent and/or quality of habitats, in combination with the restricted extent of occurrence or small area of occupancy (IUCN 2012). An example from the Czech flora of the relationship between narrow habitat specialization and species decline is *Salicornia prostrata*, a species classified in the IUCN Red List category RE (regionally extinct; Grulich 2017). It was identified as the fourth most specialized species in this study (with $ESI_w = 7.5$ when considering species occurring in at least 10 plots, Electronic Appendix 1). This species was confined to specific inland saline habitats (Vicherek 1973), which were destroyed, and the species went extinct in the 1970s (Šumberová in Chytrý 2007). Among the 10 most specialized species identified in this study (considering ESI_w for species occurring in at least 50 plots, Table 1), seven are listed in some of the IUCN Red List categories (i.e. CR, EN, VU or NT). This does not, however, imply that species with high ESIs have higher conservation value because some of them are specialists of habitats that are widespread and not currently threatened.

When we compared the specialization values of native species and two groups of aliens, early introduced archaeophytes and more recently introduced neophytes (Pyšek et al. 2012), more neophytes appeared to be generalist than archaeophytes and native species. This pattern is consistent with the results of earlier central-European studies that show the proportion of archaeophytes within plant communities to be much more dependent on habitat than the proportion of neophytes (Chytrý et al. 2008, Lososová et al. 2012). This possibly reflects the longer residence time of archaeophytes in the areas studied (Pyšek & Jarošík 2005), during which these species have managed to colonize most of the habitats to which they are preadapted, or develop adaptations for occupying newly encountered habitats (Alexander & Edwards 2010) and establish populations at most of the suitable sites. Consequently, these species have on average similarly strong relationships to habitats as native species, which have been in the study area for much longer. In contrast, many neophytes, most of which were introduced during the last two centuries (Pyšek et al. 2012, 2017), probably have not yet established tight relationships with specific habitats and their distributions are to a large extent driven by propagule pressure (Chytrý et al. 2008). Therefore neophytes appear to be more generalist than native species and archaeophytes. Another explanation can be that species growing on a broader range of habitats in the area of their native distribution can be better adapted to various conditions in the area they invaded and thus may become more successful invaders (Kalusová et al. 2017).

Comparisons of the ESI values with Ellenberg-type indicator values were consistent with the expected pattern that species occurring near the extremes of individual environmental gradients would be more specialized than those occurring in the middle of these gradients. This is valid especially for temperature, reaction and to some extent also light. For moisture, the ESI values were distributed relatively uniformly from dry through mesic to wet habitats, except for species with the moisture indicator value of 10. This value is assigned to amphibious plants that grow in water but can also grow in drained habitats for long periods (Ellenberg et al. 1991, Chytrý et al. 2018). Therefore they co-occur both with aquatic and terrestrial plants, which is reflected in their low values of ESI. Only nutrient indicator values have a clear monotonous relationship with species niche breadth, with oligotrophic species being more specialized. This pattern probably reflects the fact that while nutrient-rich habitats are rather similar to each other, there are several different kinds of nutrient limitation at nutrient-poor sites (Tilman 1982), e.g. nitrogen limitations vs phosphorus limitation (Braakhekke & Hooftman 1999), and there can be interactions with other factors that limit the availability to plants of the nutrients present at the site (e.g. drought or high pH that limits availability of phosphorus; Tyler 2003).

Practical application and future outlook

In this paper, we provide a dataset of relative estimates of species specialization for a large proportion of the temperate flora in central Europe, with separate estimates for non-forest habitats and forests. Our tests in which we compared this dataset with other plant attributes and used it in a local study indicate that the values of the Ecological Specialization Index provide ecologically meaningful information. Still, experience of using this kind of data are currently limited and further testing in other studies, involving critical evaluation of the results, is needed.

For researchers planning to use these ESI values in their studies, we have the following practical suggestions. Of the three sets of ESI values provided in this study ($\mathrm{ESI_w}$, $\mathrm{ESI_{nf}}$ and $\mathrm{ESI_f}$), researchers are advised to choose the one that is most appropriate for the vegetation they are studying (i.e. choose $\mathrm{ESI_{nf}}$ if only studying non-forest vegetation,

ESI_f if only studying forest vegetation, and ESI_w if the study includes both forest and non-forest vegetation). Specific values of the ESI have meaning only in the context of the other values calculated using the same compositional dataset, and cannot be mixed with values calculated using a different dataset (e.g. values based on the forest dataset should not be mixed with those based on the non-forest dataset, or values based on Czech data provided here cannot be mixed with values calculated from data collected within other geographical regions). Although theoretically, ESI can reach values in the range between 0 and 9, practically this range is narrower (2.83–8.37 in the case of ESI_w) and it may be useful to set arbitrary thresholds to distinguish species that can be considered as generalists, specialists or indifferent. Here, as a rule of thumb, we propose to call species with $ESI_w < 4$ as generalists and $ESI_w > 6$ as specialists (the range of ESI_w between 4 to 6 includes 74% of the species for which there is a value of ESI_w). However, different thresholds can be used in individual studies. Although ESI can be calculated with a high precision, in practice it is perhaps not useful to report its values with a precision greater than one or two decimals. Finally, the quality of ESI values is increasing with the frequency of the species occurrence in the source dataset. Here, we provide ESIs calculated for species with at least 10 occurrences, but we encourage researchers to filter species with higher frequency (e.g. 20, 50 or even 100 occurrences) for the purpose of particular studies. There is an inherent trade-off between quantity and quality; a low frequency threshold will result in more species with low-quality ESI values, while higher frequency threshold will result in fewer species with high-quality ESI values.

So far, a national list of co-occurrence based specialization values was published only for the flora of France (Mobaied et al. 2015) based on 135,002 vegetation plots from the SOPHY database (Garbolino et al. 2012). While the study of Mobaied et al. (2015) lists specialization values calculated using three different beta-diversity metrics, we went a step further and offer an ecological interpretation of these values and illustrate their use in a local case study. Unfortunately, the use of different beta-diversity metrics and different frequency thresholds to calculate θ-values hampers the comparability of specialization values from different studies. For example, although one of the beta-diversity metrics used by Mobaied et al. (2015) was β_w which we also used in this study, the absolute θ-values are not comparable with those used in our study since the number of plots used to calculate is different (50 plots in Mobaied et al. 2015 vs 10 plots in this study). We suggest two alternative options for potential future studies publishing co-occurrence based specialization values. One option is to offer values calculated using several different beta diversity metrics (preferentially multiple beta, additive beta and pairwise or multisite Simpson metric) and several frequency thresholds to calculate the beta diversity (e.g. 10, 20 and 50 plots). An alternative option is to use a standard way of calculating and presenting the values. The approach introduced in this paper, based on calculating θ-value by multiple beta rarefied to 10 plots after removing the outliers, converting it into Ecological Specialization Index on the scale 0–9 and complementing with species weights expressed as species frequencies in the source dataset, is our proposal for a standard way of expressing co-occurrence-based specialization. ESI has an intuitive interpretation, since (unlike θ) it increases with increasing species specialization, and calculation based on rarefied multiple beta is fast even when using large source datasets. Additional availability of weights (species frequencies in the source dataset) allows for post-hoc selection of only those species with calculated ESI that occur in the source dataset with sufficient frequency (e.g. 50).

Our new dataset, as well as other similar datasets which can be developed for other areas, have multiple applications in ecological research. Habitat specificity (i.e. niche breadth) is one of three key components defining species rarity (Rabinowitz 1981), but in contrast to the other two components, geographic range size and local population size, data on niche breadth are usually missing for most species. The availability of datasets containing measures of niche breadth for large proportions of regional floras can facilitate studies exploring the relationship between different components of rarity (Slatyer et al. 2013). They can also be linked to data on plant traits (e.g. Kattge et al. 2011) to explore how individual traits relate to niche breadth (Fridley et al. 2007, Marinšek et al. 2015). They may be useful in local ecological studies in which information about species specialization is needed; this is illustrated by our local case study, in which we analysed the relationship between community-level ecological specialization and measured environmental variables using the community weighted mean and the fourth-corner approaches. Besides their use in fundamental research, these values can be used for species conservation assessment, for example as a source of information for regional Red Lists (if combined with information about habitat decline), or for predicting the response of rare species to environmental changes (e.g. Bovee et al. 2018).

Acknowledgements

We thank all those who contributed to the Czech National Phytosociological Database and its managers Dana Holubová and Ilona Knollová for making this study possible. Our thanks also go to Michal Hájek and an anonymous reviewer for valuable suggestions. This study was supported by the Czech Science Foundation (project 14-36079G, Centre of Excellence PLADIAS). Tony Dixon kindly improved our English.

Souhrn

Ačkoliv teoreticky je koncept ekologické specializace druhů velmi užitečný, v praxi obvykle chybí dostatek relevantních dat o proměnných prostředí, aby jej bylo možné numericky vyjádřit. V této studii představujeme index ekologické specializace (Ecological Specialization Index, ESI), který kvantifikuje šířku realizované niky druhů podél několika ekologických proměnných najednou. Koncept ESI vychází z indexu specializace theta (Fridley et al. 2007), který je počítán na datech o společném výskytu daného druhu s ostatními druhy v různých společenstvech. Na základě vegetačních dat z České národní fytocenologické databáze jsme vypočetli ESI pro všechny druhy, které se v databázi vyskytují alespoň v deseti fytocenologických snímcích. Připravili jsme tři seznamy ESI hodnot, spočtených na základě fytocenologických snímků zahrnujících (i) všechny vegetační typy ($\mathrm{ESI_w}$, 1597 druhů), (ii) pouze nelesní vegetační typy ($\mathrm{ESI_{nf}}$, 1529 druhů) a (iii) pouze lesní vegetační typy ($\mathrm{ESI_f}$, 881 druhů). Společně s hodnotami specializace uvádíme i frekvence výskytu jednotlivých druhů v databázi, protože kvalita vypočtené hodnoty ESI vzrůstá s frekvencí výskytu druhu v databázi. ESI hodnoty uvedené v této studii jsou použitelné pouze v rámci vegetace České republiky. Abychom ohodnotili ekologickou smysluplnost vypočtených ESI hodnot, porovnali jsme je s řadou dalších dostupných druhových vlastností a také otestovali jejich vztah k faktorům prostředí v rámci lokální případové studie. Výsledky ukazují, že ESI druhu průkazně koreluje s počty vegetačních asociací a biotopů, ve kterých se daný druh vyskytuje. Druhy zařazené do kategorií vyššího ohrožení v národním Červeném seznamu jsou v průměru více specializované než druhy zařazené v kategoriích nižšího ohrožení. Neofyty jsou v průměru méně specializované než archeofyty a původní druhy. Při srovnání s ellenbergovskými hodnotami kalibrovanými pro Českou republiku se ukázalo, že specializované druhy jsou spíše stínomilné, lépe adaptované na živinami chudé půdy, na půdy s buď nízkým, anebo vysokým (ale nikoliv středním) pH a na teplá, anebo chladná stanoviště. V rámci případové studie zaměřené na druhy bylinného patra v podrostu lesů na svazích hlubokého říčního údolí se ukázalo, že více specializované druhy se vyskytují na hlubokých půdách chladných severních svahů, kamenitých půdách na bázích

údolních svahů a stinných stanovištích s vyšší pokryvností stromového patra. Naopak mělké půdy na horních částech jižních svahů a na stanovištích s otevřeným stromovým zápojem jsou spíše osidlovány generalisty. Seznam hodnot indexu ekologické specializace pro druhy vyskytující se v České republice je dostupný v elektronické příloze této studie.

References

Alexander J. M. & Edwards P. J. (2010) Limits to the niche and range margins of alien species. – Oikos 119: 1377–1386.

Benjamini Y. & Hochberg Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. – Journal of the Royal Statistical Society, Series B (Methodological) 57: 289–300.

Berg C., Welk E. & Jäger E. J. (2017) Revising Ellenberg's indicator values for continentality based on global vascular plant species distribution. – Applied Vegetation Science 20: 482–493.

Borhidi A. (1995) Social behaviour types, the naturalness and relative ecological indicator values of the higher plants in the Hungarian flora. – Acta Botanica Hungarica 39: 97–181.

Botta-Dukát Z. (2012) Co-occurrence-based measure of species' habitat specialization: robust, unbiased estimation in saturated communities. – Journal of Vegetation Science 23: 201–207.

Boulangeat I., Lavergne S., Es J., Garraud L. & Thuiller W. (2012) Niche breadth, rarity and ecological characteristics within a regional flora spanning large environmental gradients. – Journal of Biogeography 39: 204–214.

Bovee K. M., Merriam K. E. & Gosejohan M. C. (2018) Livestock grazing affects vernal pool specialists more than habitat generalists in montane vernal pools. – Applied Vegetation Science 21: 12–20.

Braakhekke W. G. & Hooftman D. A. P. (1999) The resource balance hypothesis of plant species diversity in grassland. – Journal of Vegetation Science 10: 187–200.

Bruelheide H., Dengler J., Jiménez-Alfaro B., Purschke O., Hennekens S. M., Chytrý M., Pillar V. D., Jansen F., Kattge J., Sandel B., Aubin I., Biurrun I., Field R., Haider S., Jandt U., Lenoir J., Peet R. K., Peyre G., Sabatini F. M., Schmidt M., Schrodt F., Winter M., Aćić S., Agrillo E., Alvarez M., Ambarlı D., Angelini P., Apostolova I., Arfin Khan M. A. S., Arnst E., Attorre F., Baraloto C., Beckmann M., Berg C., Bergeron Y., Bergmeier E., Bjorkman A. D., Bondareva V., Borchardt P., Botta-Dukát Z., Boyle B., Breen A., Brisse H., Byun C., Cabido M. R., Casella L., Cayuela L., Černý T., Chepinoga V., Csiky J., Curran M., Ćušterevska R., Dajić Stevanović Z., De Bie E., Ruffray P. D., Sanctis M. D., Dimopoulos P., Dressler S., Ejrnæs R., El-Sheikh M. A. E.-R. M., Enquist B., Ewald J., Fagúndez J., Finckh M., Font X., Forey E., Fotiadis G., García-Mijangos I., de Gasper A. L., Golub V., Gutierrez A. G., Hatim M. Z., He T., Higuchi P., Holubová D., Hölzel N., Homeier J., Indreica A., Işık Gürsoy D., Jansen S., Janssen J., Jedrzejek B., Jiroušek M., Jürgens N., Kącki Z., Kavgacı A., Kearsley E., Kessler M., Knollová I., Kolomiychuk V., Korolyuk A., Kozhevnikova M., Kozub Ł., Krstonošić D., Kühl H., Kühn I., Kuzemko A., Küzmič F., Landucci F., Lee M. T., Levesley A., Li C.-F., Liu H., Lopez-Gonzalez G., Lysenko T., Macanović A., Mahdavi P., Manning P., Marcenò C., Martynenko V., Mencuccini M., Minden V., Moeslund J. E., Moretti M., Müller J. V., Munzinger J., Niinemets Ü., Nobis M., Noroozi J., Nowak A., Onyshchenko V., Overbeck G. E., Ozinga W. A., Pauchard A., Pedashenko H., Peńuelas J., Pérez-Haase A., Peterka T., Petřík P., Phillips O. L., Prokhorov V., Rašomavičius V., Revermann R., Rodwell J., Ruprecht E., Rūsiņa S., Samimi C., Schaminée J. H. J., Schmiedel U., Šibík J., Šilc U., Škvorc Ž., Smyth A., Sop T., Sopotlieva D., Sparrow B., Stančić Z., Svenning J.-C., Swacha G., Tang Z., Tsiripidis I., Turtureanu P. D., Uğurlu E., Uogintas D., Valachovič M., Vanselow K. A., Vashenyak Y., Vassilev K., Vélez-Martin E., Venanzoni R., Vibrans A. C., Violle C., Virtanen R., von Wehrden H., Wagner V., Walker D. A., Wana D., Weiher E., Wesche K., Whitfeld T., Willner W., Wiser S., Wohlgemuth T., Yamalov S., Zizka G. & Zverev A. (2019) sPlot – a new tool for global vegetation analyses. – Journal of Vegetation Science 30: 161–186.

Chao A., Gotelli N. J., Hsieh T. C., Sander E. L., Ma K. H., Colwell R. K. & Ellison A. M. (2014) Rarefaction and extrapolation with Hill numbers: a framework for sampling and estimation in species diversity studies. – Ecological Monographs 84: 45–67.

Chytrý M. (ed.) (2007–2013) Vegetace České republiky 1–4 [Vegetation of the Czech Republic 1– 4]. – Academia, Praha.

Chytrý M. (2017) Current vegetation of the Czech Republic. – In: Chytrý M., Danihelka J., Kaplan Z. & Pyšek P. (eds), Flora and vegetation of the Czech Republic, p. 229–337, Springer, Cham.

Chytrý M., Hennekens S. M., Jiménez-Alfaro B., Knollová I., Dengler J., Jansen F., Landucci F., Schaminée J. H. J., Aćić S., Agrillo E., Ambarlı D., Angelini P., Apostolova I., Attorre F., Berg C., Bergmeier E., Biurrun I., Botta-Dukát Z., Brisse H., Campos J. A., Carlón L., Čarni A., Casella L., Csiky J., Ćušterevska R., Dajić Stevanović Z., Danihelka J., De Bie E., de Ruffray P., De Sanctis M., Dickoré W. B., Dimopoulos P., Dubyna D., Dziuba T., Ejrnæs R., Ermakov N., Ewald J., Fanelli G., Fernández-González F., FitzPatrick Ú., Font X., García-Mijangos I., Gavilán García R., Golub V., Guarino R., Haveman R., Indreica A., Işık Gürsoy D., Jandt U., Janssen J. A. M., Jiroušek M., Kącki Z., Kavgacı A., Kleikamp M., Kolomiychuk V., Krstivojević Ćuk M., Krstonošić D., Kuzemko A., Lenoir J., Lysenko T., Marcenò C., Martynenko V., Michalcová D., Moeslund J. E., Onyshchenko V., Pedashenko H., Pérez-Haase A., Peterka T., Prokhorov V., Rašomavičius V., Rodríguez-Rojo M. P., Rodwell J. S., Rogova T., Ruprecht E., Rūsiņa S., Seidler G., Šibík J., Šilc U., Škvorc Ž., Sopotlieva D., Stančić Z., Svenning J.-C., Swacha G., Tsiripidis I., Turtureanu P. D., Uğurlu E., Uogintas D., Valachovič M., Vashenyak Y., Vassilev K., Venanzoni R., Virtanen R., Weekes L., Willner W., Wohlgemuth T. & Yamalov S. (2016) European Vegetation Archive (EVA): an integrated database of European vegetation plots. – Applied Vegetation Science 19: 173–180.

Chytrý M., Jarošík V., Pyšek P., Hájek O., Knollová I., Tichý L. & Danihelka J. (2008) Separating habitat invasibility by alien plants from the actual level of invasion. – Ecology 89: 1541–1553.

Chytrý M., Kučera T., Kočí M., Grulich V. & Lustyk P. (eds) (2010) Katalog biotopů České republiky [Habitat Catalogue of the Czech Republic]. Ed. 2. – Agentura ochrany přírody a krajiny ČR, Praha.

Chytrý M., Pyšek P., Tichý L., Knollová I. & Danihelka J. (2005) Invasions by alien plants in the Czech Republic: a quantitative assessment across habitats. – Preslia 77: 339–354.

Chytrý M. & Rafajová M. (2003) Czech National Phytosociological Database: basic statistics of the available vegetation-plot data. – Preslia 75: 1–15.

Chytrý M., Tichý L., Dřevojan P., Sádlo J. & Zelený D. (2018) Ellenberg-type indicator values for the Czech flora. – Preslia 90: 83–103.

Coudun C. & Gégout J.-C. (2005) Ecological behaviour of herbaceous forest species along a pH gradient: a comparison between oceanic and semicontinental regions in northern France. – Global Ecology and Biogeography 14: 263–270.

Danihelka J., Chrtek J. Jr & Kaplan Z. (2012) Checklist of vascular plants of the Czech Republic. – Preslia 84: 647–811.

De Cáceres M., Font X. & Oliva F. (2010) The management of vegetation classifications with fuzzy clustering. – Journal of Vegetation Science 21: 1138–1151.

de Mendiburu F. (2017) agricolae: statistical procedures for agricultural research. R package version 1.2-8. – URL: https://CRAN.R-project.org/package=agricolae.

Dengler J., Jansen F., Glöckler F., Peet R. K., De Cáceres M., Chytrý M., Ewald J., Oldeland J., Lopez-Gonzalez G., Finckh M., Mucina L., Rodwell J. S., Schaminée J. H. J. & Spencer N. (2011) The Global Index of Vegetation-Plot Databases (GIVD): a new resource for vegetation science. – Journal of Vegetation Science 22: 582–597.

Dray S., Choler P., Dolédec S., Peres-Neto P. R., Thuiller W., Pavoine S. & ter Braak C. J. F. (2014) Combining the fourth-corner and the RLQ methods for assessing trait responses to environmental variation. – Ecology 95: 14–21.

Dray S. & Legendre P. (2008) Testing the species traits-environment relationships: the fourth-corner problem revisited. – Ecology 89: 3400–3412.

Ellenberg H., Weber H. E., Düll R., Wirth V., Werner W. & Paulißen D. (1991) Zeigerwerte von Pflanzen in Mitteleuropa. – Scripta Geobotanica 18: 1–248.

Fajmonová Z., Zelený D., Syrovátka V., Vončina G. & Hájek M. (2013) Distribution of habitat specialists in semi-natural grasslands. – Journal of Vegetation Science 24: 616–627.

Fridley J., Vandermast D., Kuppinger D., Manthey M. & Peet R. (2007) Co-occurrence based assessment of habitat generalists and specialists: a new approach for the measurement of niche width. – Journal of Ecology 95: 707–722.

Garbolino E., De Ruffray P., Brisse H. & Grandjouan G. (2012) The phytosociological database SOPHY as the basis of plant socio-ecology and phytoclimatology in France. – Biodiversity & Ecology 4: 177–184.

Grulich V. (2017) Červený seznam cévnatých rostlin ČR [The Red List of vascular plants of the Czech Republic]. – Příroda 35: 75–132.

Hadincová V., Münzbergová Z., Wild J., Šajtar L. & Marešová J. (2008) Dispersal of invasive *Pinus strobus* in sandstone areas of the Czech Republic. – In: Tokarska-Guzik B., Brock J. H., Brundu G., Child L. E., Daehler C. C. & Pyšek P. (eds), Plant invasions: human perception, ecological impacts and management, p. 117–132, Backhuys Publishers, Leiden.

Hájková P., Hájek M., Apostolova I., Zelený D. & Dítě D. (2008) Shifts in the ecological behaviour of plant species between two distant regions: evidence from the base richness gradient in mires. – Journal of Biogeography 35: 282–294.

Hill M. O., Roy D. B., Mountford J. O. & Bunce R. G. H. (2000) Extending Ellenberg's indicator values to a new area: an algorithmic approach. – Journal of Applied Ecology 37: 3–15.

Hutchinson G. E. (1957) Concluding remarks. – Cold Spring Harbor Symposia on Quantitative Biology 22: 415–427.

IUCN (2012) IUCN Red List Categories and Criteria: Version 3.1. Second edition. – IUCN, Gland and Cambridge.

Jost L. (2006) Entropy and diversity. – Oikos 113: 363–375.

Kalusová V., Chytrý M., van Kleunen M., Mucina L., Dawson W., Essl F., Kreft H., Pergl J., Weigelt P., Winter M. & Pyšek P. (2017) Naturalization of European plants on other continents: the role of donor habitats. – Proceedings of the National Academy of Sciences of the United States of America 114: 13756–13761.

Kattge J., Díaz S., Lavorel S., Prentice I. C., Leadley P., Bönisch G., Garnier E., Westoby M., Reich P. B., Wright I. J., Cornelissen J. H. C., Violle C., Harrison S. P., Van Bodegom P. M., Reichstein M., Enquist B. J., Soudzilovskaia N. A., Ackerly D. D., Anand M., Atkin O., Bahn M., Baker T. R., Baldocchi D., Bekker R., Blanco C. C., Blonder B., Bond W. J., Bradstock R., Bunker D. E., Casanoves F., Cavender-Bares J., Chambers J. Q., Chapin F. S., Chave J., Coomes D., Cornwell W. K., Craine J. M., Dobrin B. H., Duarte L., Durka W., Elser J., Esser G., Estiarte M., Fagan W. F., Fang J., Fernández-Méndez F., Fidelis A., Finegan B., Flores O., Ford H., Frank D., Freschet G. T., Fyllas N. M., Gallagher R. V., Green W. A., Gutierrez A. G., Hickler T., Higgins S. I., Hodgson J. G., Jalili A., Jansen S., Joly C. A., Kerkhoff A. J., Kirkup D., Kitajima K., Kleyer M., Klotz S., Knops J. M. H., Kramer K., Kühn I., Kurokawa H., Laughlin D., Lee T. D., Leishman M., Lens F., Lenz T., Lewis S. L., Lloyd J., Llusià J., Louault F., Ma S., Mahecha M. D., Manning P., Massad T., Medlyn B. E., Messier J., Moles A. T., Müller S. C., Nadrowski K., Naeem S., Niinemets Ü., Nöllert S., Nüske A., Ogaya R., Oleksyn J., Onipchenko V. G., Onoda Y., Ordoñez J., Overbeck G., Ozinga W. A., Patiño S., Paula S., Pausas J. G., Peñuelas J., Phillips O. L., Pillar V., Poorter H., Poorter L., Poschlod P., Prinzing A., Proulx R., Rammig A., Reinsch S., Reu B., Sack L., Salgado-Negret B., Sardans J., Shiodera S., Shipley B., Siefert A., Sosinski E., Soussana J.-F., Swaine E., Swenson N., Thompson K., Thornton P., Waldram M., Weiher E., White M., White S., Wright S. J., Yguel B., Zaehle S., Zanne A. E. & Wirth C. (2011) TRY – a global database of plant traits. – Global Change Biology 17: 2905–2935.

Kim S. (2015) ppcor: partial and semi-partial (part) correlation. R package version 1.1. – URL: https://CRAN.R-project.org/package=ppcor.

Knollová I., Chytrý M., Tichý L. & Hájek O. (2006) Stratified resampling of phytosociological databases: some strategies for obtaining more representative data sets for classification studies. – Journal of Vegetation Science 16: 479–486.

Lande R. (1996) Statistics and partitioning of species diversity, and similarity among multiple communities. – Oikos 76: 5–13.

Legendre P., Galzin R. & Harmelin-Vivien M. L. (1997) Relating behavior to habitat: solutions to the fourth-corner problem. – Ecology 78: 547–562.

Lengyel A., Chytrý M. & Tichý L. (2011) Heterogeneity-constrained random resampling of phytosociological databases. – Journal of Vegetation Science 22: 175–183.

Lososová Z., Chytrý M., Tichý L., Danihelka J., Fajmon K., Hájek O., Kintrová K., Kühn I., Láníková D., Otýpková Z. & Řehořek V. (2012) Native and alien floras in urban habitats: a comparison across 32 cities of central Europe. – Global Ecology and Biogeography 21: 545–555.

MacArthur R. H. (1972) Geographical ecology: patterns in the distribution of species. – Harper & Row, New York.

Manthey M. & Fridley J. (2009) Beta diversity metrics and the estimation of niche width via species co-occurrence data: reply to Zeleny. – Journal of Ecology 97: 18–22.

Marinšek A., Čarni A., Šilc U. & Manthey M. (2015) What makes a plant species specialist in mixed broad-leaved deciduous forests? – Plant Ecology 216: 1469–1479.

McCune B. & Keon D. (2002) Equations for potential annual direct incident radiation and heat load. – Journal of Vegetation Science 19: 603–606.

McCune B. & Mefford M. J. (1999) PC-ORD: multivariate analysis of ecological data 4.17. User's guide. – MjM Software, Gleneden Beach, OR, US.

Mobaied S., Machon N. & Porcher E. (2015) Ecological specialization and rarity indices estimated for a large number of plant species in France. – Data in Brief 3: 165–168.

Peres-Neto P. R., Dray S. & ter Braak C. J. F. (2017) Linking trait variation to the environment: critical issues with community-weighted mean correlation resolved by the fourth-corner approach. – Ecography 40: 806–816.

Pignatti S., Menegoni P. & Pietrosanti S. (2005) Valori di bioindicazione delle piante vascolari della Flora d'Italia [Bioindicator values of vascular plants of the Flora of Italy]. – Braun-Blanquetia 39: 1–97.

Pulliam H. R. (2000) On the relationship between niche and distribution. – Ecology Letters 3: 349–361.

Pyšek P., Chytrý M., Pergl J., Sádlo J. & Wild J. (2017) Plant invasions in the Czech Republic. – In: Chytrý M., Danihelka J., Kaplan Z. & Pyšek P. (eds), Flora and vegetation of the Czech Republic, p. 339–399, Springer, Cham.

Pyšek P., Danihelka J., Sádlo J., Chrtek J. jun., Chytrý M., Jarošík V., Kaplan Z., Krahulec F., Moravcová L., Pergl J., Štajerová K. & Tichý L. (2012) Catalogue of alien plants of the Czech Republic (2nd edition): checklist update, taxonomic diversity and invasion patterns. – Preslia 84: 155–255.

Pyšek P. & Jarošík V. (2005) Residence time determines the distribution of alien plants. – In: Inderjit (ed.), Invasive plants: agricultural and ecological aspects, p. 77–96, Birkhäuser Verlag, Basel.

Pyšek P., Jarošík V., Pergl J., Randall R., Chytrý M., Kühn I., Tichý L., Danihelka J., Chrtek J. & Sádlo J. (2009) The global invasion success of Central European plants is related to distribution characteristics in their native range and species traits. – Diversity and Distributions 15: 891–903.

Rabinowitz D. (1981) Seven forms of rarity. – In: Synge H. (ed.), The biological aspects of rare plant conservation, p. 205–217, Wiley, Chichester.

R Core Team (2017) R: A language and environment for statistical computing. – R Foundation for Statistical Computing, Vienna. URL: https://www.R-project.org/.

Sádlo J., Chytrý M. & Pyšek P. (2007) Regional species pools of vascular plants in habitats of the Czech Republic. – Preslia 79: 303–321.

Shannon C. E. (1948) A mathematical theory of communication. – The Bell System Technical Journal 27: 379–423.

Siefert A., Ravenscroft C., Althoff D., Alvarez-Yépiz J., Carter B., Glennon K., Heberling M., Jo I., Pontes A., Sauer A., Willis A. & Fridley J. (2012) Scale dependence of vegetation–environment relationships: a meta-analysis of multivariate data. – Journal of Vegetation Science 23: 942–951.

Šilc U., Lososová Z. & Vrbničanin S. (2014) Weeds shift from generalist to specialist: narrowing of ecological niches along a north-south gradient. – Preslia 86: 35–46.

Slatyer R. A., Hirst M. & Sexton J. P. (2013) Niche breadth predicts geographical range size: a general ecological pattern. – Ecology Letters 16: 1104–1114.

Sørensen T. (1948) A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analysis of the vegetation on Danish commons. – Biologiske Skrifter 5: 1–34.

ter Braak C. J. F., Cormont A., & Dray S. (2012) Improved testing of species traits–environment relationships in the fourth-corner problem. – Ecology 93: 1525–1526.

ter Braak C. J. F., Peres-Neto P. R. & Dray S. (2018) Simple parametric tests for trait–environment association. – Journal of Vegetation Science 29: 801–811.

Tichý L. (2002) JUICE, software for vegetation classification. – Journal of Vegetation Science 451–453.

Tilman D. (1982) Resource competition and community structure. – Princeton University Press, Princeton, NJ.

Tuomisto H. (2010) A diversity of beta diversities: straightening up a concept gone awry. Part 1. Defining beta diversity as a function of alpha and gamma diversity. – Ecography 33: 2–22.

Tyler G. (2003) Some ecophysiological and historical approaches to species richness and calcicole/calcifuge behaviour: contribution to a debate. – Folia Geobotanica 38: 419–428.

Ugland K. I., Gray J. S. & Ellingsen K. E. (2003) The species-accumulation curve and estimation of species richness. – Journal of Animal Ecology 72: 888–897.

Vicherek J. (1973) Die Pflanzengesellschaften der Halophyten- und Subhalophytenvegetation der Tschechoslowakei. – In: Vegetace ČSSR A 5, Academia, Praha.

Wagner V., Chytrý M., Zelený D., von Wehrden H., Brinkert A., Danihelka J., Hölzel N., Jansen F., Kamp J., Lustyk P., Merunková K., Palpurina S., Preislerová Z. & Wesche K. (2017) Regional differences in soil pH niche among dry grassland plants in Eurasia. – Oikos 126: 660–670.

Walker K. J. & Preston D. C. (2006) Ecological predictors of extinction risk in the flora of lowland England, UK. – Biodiversity & Conservation 15: 1913–1942.

Wasof S., Lenoir J., Gallet-Moron E., Jamoneau A., Brunet J., Cousins S. A. O., De Frenne P., Diekmann M., Hermy M., Kolb A., Liira J., Verheyen K., Wulf M. & Decocq G. (2013) Ecological niche shifts of understorey plants along a latitudinal gradient of temperate forests in northwestern Europe. – Global Ecology and Biogeography 22: 1130–1140.

Whittaker R. H. (1960) Vegetation of the Siskiyou Mountains, Oregon and California. – Ecological Monographs 30: 279–338.

Wiser S. & De Cáceres M. (2013) Updating vegetation classifications: an example with New Zealand's woody vegetation. – Journal of Vegetation Science 24: 80–93.

Zelený D. (2009) Co-occurrence based assessment of species habitat specialization is affected by the size of species pool: reply to Fridley et al. (2007). – Journal of Ecology 97: 10–17.

Zelený D. (2018) Which results of the standard test for community-weighted mean approach are too optimistic? – Journal of Vegetation Science 29: 953–966.

Zelený D. & Chytrý M. (2007) Environmental control of vegetation pattern in deep river valleys of the Bohemian Massif. – Preslia 79: 205–222.

Zelený D., Li C.-F. & Chytrý M. (2010) Pattern of local plant species richness along a gradient of landscape topographical heterogeneity: result of spatial mass effect or environmental shift? – Ecography 33: 578–589.

Received 8 July 2018
Revision received 30 April 2019
Accepted 2 May 2019

Appendix 1. – Formulas used to calculate the number of effective phytosociological associations.

Calculation of “the number of effective phytosociological associations” in which the focal species occurs in a high relative proportion of the vegetation plots of that association (A_{rel})

$$A_{\mathrm{rel}} = e^{H'_{\mathrm{rel}}} \qquad \text{(eq 1.1)}$$

where

$$H'_{\mathrm{rel}} = \sum_{i=1}^{A} p_{(\mathrm{rel})i} \cdot \ln p_{(\mathrm{rel})i} \qquad \text{(eq 1.2)}$$

where

$$p_{(\mathrm{rel})i} = \frac{n_i}{N_i} \cdot \frac{1}{\sum_{i=1}^{A} \frac{n_i}{N_i}} \qquad \text{(eq 1.3)}$$

where $p_{(\mathrm{rel})i}$ is the normalized proportion of plots in association i in which the focal species occurs, n_i is the number of plots within association i in which the focal species occurs, N_i is the number of all plots within the association i, and A is the number of all associations in the dataset. To make sure that the sum of $p_{(\mathrm{rel})i}$ will be equal to 1, we added the normalization constant to eq 1.3.

Calculation of “the number of effective phytosociological associations” in which the focal species occurs in a high absolute number of vegetation plots of that association (A_{abs})

$$A_{\mathrm{abs}} = e^{H'_{\mathrm{abs}}} \qquad \text{(eq 2.1)}$$

where

$$H'_{\mathrm{abs}} = \sum_{i=1}^{A} p_{(\mathrm{abs})i} \cdot \ln p_{(\mathrm{abs})i} \qquad \text{(eq 2.2)}$$

where

$$p_{(\mathrm{abs})i} = n_i \cdot \frac{1}{\sum_{i=1}^{A} n_i} \qquad \text{(eq 2.3)}$$

where $p_{(\mathrm{abs})i}$ is the normalized number of plots in association i in which the focal species occurs, n_i is the number of plots within association i in which the focal species occurs, N_i is the number of all plots within the association i, and A is the number of all associations in the dataset. To make sure that the sum of $p_{(\mathrm{abs})i}$ will be equal to 1, we added the normalization constant to eq 2.3.