Ecological Indicators 118 (2020) 106761

Contents lists available at ScienceDirect

Ecological Indicators

journal homepage: www.elsevier.com/locate/ecolind

From abundance-based to functional-based indicator species

Carlo Ricotta[a,*], Alicia T.R. Acosta[b], Marco Caccianiga[c], Bruno E.L. Cerabolini[d], Sandrine Godefroid[e,f,g], Marta Carboni[b]

[a] Department of Environmental Biology, University of Rome 'La Sapienza', Piazzale Aldo Moro 5, 00185 Rome, Italy
[b] Department of Sciences, University 'Roma Tre', Viale Marconi 446, 00146 Rome, Italy
[c] Department of Biosciences, University of Milano, Via G. Celoria 26, 20133 Milano, Italy
[d] Department of Biotechnology and Life Sciences, University of Insubria, Via J.H. Dunant 3, 21100 Varese, Italy
[e] Research Department, Meise Botanic Garden, Nieuwelaan 38, 1860 Meise, Belgium
[f] Service Général de L'Enseignement Supérieur et de la Recherche Scientifique, Fédération Wallonie-Bruxelles, rue A. Lavallée 1, 1080 Brussels, Belgium
[g] Laboratory of Plant Ecology and Biogeochemistry, Université Libre de Bruxelles, CP 244, Boulevard du Triomphe, 1050 Brussels, Belgium

ARTICLE INFO

Keywords:
Functional abundance
Functional centroid
Fuzzy sets
Permutation methods
Species occurrences

ABSTRACT

Indicator species with high fidelity to a-priori defined groups of sites are a relevant tool to ecologically characterize plant or animal assemblages. The identification of indicator or diagnostic species is usually performed by summarizing the species abundances within each group of sites. Species with high concentration in a given group of sites are considered diagnostic of that particular group. Among the methods proposed for the determination of indicator species, only very few have considered the species functional traits. This is quite surprising, as species influence ecosystem processes via their traits. Therefore, the species functional traits should give a much better ecological characterization of a group of sites than the species abundances. The aim of this paper is thus to use the species functional characteristics to improve their diagnostic value. These characteristics include the species functional traits and all species-level indicators of environmental association. The proposed method consists of combining the species abundances and their functional characteristics into a single composite index, which can be interpreted as the species fuzzy degree of compatibility with each group of sites. The interpretation of this index in terms of fuzzy set theory allows to introduce a high degree of flexibility in the computation of the species diagnostic values. To show the behavior of the proposed index, two worked examples with data on Alpine vegetation in northern Italy and urban alien species in the city of Brussels (Belgium) are used.

1. Introduction

Indicator species with high fidelity to a particular group of sites are an important tool for the characterization of plant or animal assemblages (Tichý and Chytrý 2006). The detection of indicator or diagnostic species is typically performed by calculating the species occurrences or abundances in distinct groups of sites (or plots, relevés, sampling units, etc.). Species with high abundance concentration in a given group of sites compared to the other groups are considered diagnostic of that specific group (Tichý and Chytrý 2006; De Cáceres et al. 2010). The grouping may be based on compositional or environmental differences among sites or on other distinctive properties, such as different successional stages, land use types, or different levels of controlled experimental designs (Dufrêne and Legendre 1997; De Cáceres et al. 2012).

Many different approaches have been proposed to identify diagnostic species, (e.g. Dufrêne and Legendre 1997; Chytrý et al. 2002; Podani and Csányi 2010). The most widely used method for indicator species analysis is due to Dufrêne and Legendre (1997) who introduced a species-specific composite index called IndVal (Indicator Value). Given a community composition matrix containing the presence/absence or the abundance values x_{jn} of S species ($j = 1, 2, \ldots, S$) in N plots ($n = 1, 2 \ldots, N$), let the number of plots in group k ($k = 1, 2, \ldots, G$) be N_k. IndVal is the product of two terms. The first term, which is called *specificity* (A_{jk} in the notation of Dufrêne and Legendre), is the ratio of the mean abundance of species j in group k and the sum of means of the same species over all groups. Specificity thus measures the probability that a given plot n that contains species j belongs to a target group k. Maximum specificity ($A_{jk} = 1$) is obtained if species j appears only in group k irrespective of its abundance. Minimum specificity ($A_{jk} = 0$) is

* Corresponding author.
E-mail address: carlo.ricotta@uniroma1.it (C. Ricotta).

https://doi.org/10.1016/j.ecolind.2020.106761
Received 21 April 2020; Received in revised form 12 July 2020; Accepted 23 July 2020

obtained if species j is not contained in k.

The second component of IndVal is called *fidelity* (B_{jk}) and represents the number of presences of species j in group k compared to the total number of plots N_k in that group, thus measuring the probability that a given plot n that belongs to a target group k contains species j. Maximum fidelity ($B_{jk} = 1$) is obtained if species j occurs in all plots of group k, whereas minimum fidelity ($B_{jk} = 0$) is obtained if species j is not contained in k. For details, see Dufrêne and Legendre (1997).

Specificity and fidelity are then multiplied and scaled in the range [0, 100] to express the indicator value of species j for group k in terms of percentages:

$$Indval_{jk} = A_{jk} \times B_{jk} \times 100 \quad (1)$$

Here, we assume that each of the S species is found at least in one plot, and that all N plots contain at least one species.

To determine the significance of the association of species j with group k we next have to compare the actual value of $IndVal_{jk}$ with a null distribution obtained by randomly reassigning the abundances of j among the N plots. This operation generates a distribution of the test statistic under the null hypothesis that the occurrence of species j in a given plot n is due to chance alone (De Cáceres et al. 2012).

Although indicator species analysis is generally used to ecologically characterize different groups of sites, to the best of our knowledge, only Ricotta et al. (2015) have proposed to include the functional traits of species, in addition to their occurrence and abundance, for the determination of indicator species. This is in spite of the fact that functional traits are routinely used in community ecology studies to inform on processes of community assembly (Wright et al. 2004; McGill et al. 2006; Adler et al. 2013; Kraft et al. 2015; Díaz et al. 2016), community responses to environmental change (Voigt et al. 2007; Moretti and Legg 2009; Moles et al. 2014), and effects on ecosystem functioning (Lavorel and Garnier 2002; Garnier et al. 2004; Díaz et al. 2007). Indeed, it is well known that functional traits inform on the ecological strategies of species and, in turn, species influence ecosystem processes via their traits (Mason and de Bello 2013). Hence, the dominant functional traits in plant or animal assemblages should give a much better ecological characterization of a group of sites, in terms of the local environmental conditions and ecosystem functioning, than the mere occurrence of species.

Ricotta et al. (2015) proposed a two-step procedure to include the species functional traits in the evaluation of their diagnostic values. First, the indicator species that best characterize a given group of plots are identified with the usual statistical tools based either on the species incidence or abundance data. Next, the functional association between the abundance-based indicator species and the target group of plots is tested by calculating the functional distance between the indicator species and the functional centroids of all plots in that group. A species is considered diagnostic of a given group if its mean functional distance from the plot centroids of the target group is significantly lower than expected.

According to this approach, functional indicator analysis is limited to those species that are considered diagnostic of a given group of plots in terms of traditional abundance-based methods. The main reason invoked by Ricotta et al. (2015) for this restriction is that indicator species are generally used for the *a-posteriori* ecological typification of one or more groups of plots. Therefore, "to consider a species as functionally diagnostic of a particular habitat, the species should possess a reasonable chance of being detected in the field" (Ricotta et al. 2015). On the other hand, this two-step procedure prevents the species without significant diagnostic capacity in terms of species abundances from being tested for their functional relevance. However, it may happen that a species that is not diagnostic in terms of species abundances alone shows instead a strong functional association with a target group of plots.

The aim of this paper is thus to develop a general method to include any measurable species characteristic *sensu* Garnier et al. (2017) for the characterization of their diagnostic value. These characteristics may include the species traits and all species-level indicators of environmental association, such as Ellenberg's indicator values (Ellenberg et al. 1991), or Grime's primary adaptive strategies (Grime 1977; Grime and Pierce 2012). Two worked examples on Alpine vegetation in northern Italy and on urban alien species in the city of Brussels (Belgium) are used to show the behavior of the proposed method.

2. Methods

A promising line of attack to include the species characteristics into indicator species analysis may consist in combining the species abundances and their functional or environmental association with a given group of plots into a single composite index, in a similar way to that of Dufrêne and Legendre (1997). Let x_{jn} be the presence/absence (0/1) or the abundance value of species j in plot n and d_{jk} be the functional dissimilarity between species j and the functional centroid of all plots in group k. If d_{jk} is measured in the range [0, 1], this quantity can be interpreted as the fuzzy degree of functional distinctness of species j with respect to group k. Consequently, the similarity s_{jk} between species j and the functional centroid of group k can be simply calculated as the complement of d_{jk}: $s_{jk} = 1 - d_{jk}$. The functional association ϕ_{jk} between species j and group k may then be expressed as the mean abundance of species j in all plots belonging to group k multiplied by s_{jk}:

$$\phi_{jk} = \frac{\sum_{n \in k} x_{jn} \times s_{jk}}{N_k} \quad (2)$$

where s_{jk} represents the fuzzy degree of functional compatibility between species j and group k, such that the product $\pi_{jn} = x_{jn} \times s_{jk}$ may be interpreted as the 'functional abundance' of species j in plot n (i.e., the fraction of abundance of species j that is functionally compatible with plot n) and ϕ_{jk} as the mean 'functional abundance' of species j in group k. Note that for presence and absence data $0 \leq \phi_{jk} \leq 1$. In this case, maximum functional association between species j and group k is obtained if j occurs in all plots of group k with $s_{jk} = 1$. The significance of the association of a given species j with a target group k can then be tested with the usual permutation methods by randomly reassigning the functional abundances of j among the N plots.

3. Worked examples

3.1. Alpine vegetation on a glacier foreland in northern Italy

We used data on plant communities sampled along a primary succession on the foreland of the Rutor Glacier (northern Italy). The data were sampled by Caccianiga et al. (2006). The original data set can be found in Ricotta et al. (2016: Appendix S2) and contains the abundances of 45 Alpine species collected in 59 vegetation plots of approximately 25 m^2. All species abundances were measured with a five-point ordinal scale transformed to ranks. Based on the age of the moraine deposits, the plots were classified into three successional stages: early succession (17 plots), mid succession (32 plots), and late succession (10 plots); for additional details, see Caccianiga et al. (2006).

Six quantitative traits, which provide a good representation of the species global spectrum of form and function (Díaz et al. 2016) were selected: canopy height (CH; mm), leaf dry matter content (LDMC; %), leaf dry weight (LDW; mg), specific leaf area (SLA; $mm^2 \times mg^{-1}$), leaf nitrogen content (LNC; %), and leaf carbon content (LCC; %). All data are freely available in Caccianiga et al. (2006, pp. 16-17).

To explore the behavior of the selected traits along the primary succession, we performed a principal component analysis (PCA) on the multivariate functional centroids of each plot, defined as the mean of all trait values in each plot weighted by the total abundance of each species in that plot (Garnier et al. 2004). Before calculations, all traits were

standardized to zero mean and unit standard deviation.

To assess the functional association of each species with the three successional stages, we first calculated the multivariate functional centroid of each stage (i.e., the mean of all trait values in each stage weighted by the total abundance of each species in that stage). We then computed the Euclidean distance between the trait values of each species and the functional centroid of each successional stage. These distances were then locally rescaled to the unit range by dividing each species-to-centroid distance by the maximum value in the dataset. Finally, we used the species-to-centroid similarities s_{jk} (i.e. the complement of the scaled species-to-centroid distances) to calculate the functional association ϕ_{jk} of all 45 species with each of the three successional stages.

To evaluate whether the functional association of each species with the three successional stages was significantly higher than expected by chance, we permuted the functional abundance values of each species π_{jn} among all 59 plots. The null hypothesis is that there is no difference in the value of ϕ_{jk} among the successional stages. P-values of positive functional association between a given species and each group of plots were then calculated as the proportion of permutation-derived values of ϕ_{jk} that were as high or higher than the actual value (999 permutations, two-tailed test). All calculations were done with a new R script available in Appendix 1.

Note that unlike in Ricotta at al. 2015, we calculated the functional abundances of each species π_{jn} based on the mean similarity from the group centroids and not from single plots. This is because the functional centroid of a given group of plots is likely a better indicator of the overall functioning of that group compared to the centroids of the single plots in that group. Those who think it is better to base the analysis on the centroids of single plots simply have to replace in Eq. (2) the functional similarity s_{jk} between species j and the functional centroid of group k, with the functional similarity s_{jn} between species j and the functional centroid of plot n.

We also performed a traditional indicator species analysis to test whether the species were significantly associated with any of the three successional stages based only on presence/absence scores or on abundance data. In this case, an indicator or diagnostic species is defined as a species that is more common in a given group of plots than expected by chance alone. Hence, for species incidence or abundance data, to assess whether a species is diagnostic of a given successional stage, the total abundance (occurrence) of that species in each stage was compared with a random model in which the species abundance (occurrence) values are randomly permuted within all plots (999 permutations; two-tailed test), thus simulating the null condition whereby all plots have the same probability to host each species, irrespective of their ecological preferences (De Cáceres and Legendre 2009; Ricotta et al. 2015).

3.2. Urban alien flora of Brussels

Lososová et al. (2011) and Godefroid and Ricotta (2018) showed that different urban land use types host different assemblages of alien plant species. Since alien species may represent a major threat to ecosystems and biodiversity, analyzing their ecological preferences is of crucial importance, particularly for urban areas where they represent a relevant portion of the local flora. Therefore, we analyzed the environmental association of alien plant species for five urban land uses in the city of Brussels.

The urban area of Brussels (161 km^2; 1.2 million inhabitants) was divided in a grid of square cells of 1 km^2. Within each cell, all spontaneous species of the vascular flora were sampled between 1992 and 1994. Each species was then classified as alien or native according to Pyšek et al. (2004). Based on the dominant land use type, each cell was associated to one of the following classes: densely built-up urban areas (UD), open built-up urban areas (UO), urban forests (FOR), industrial areas (IND), and agricultural areas (AGR). This land use classification was considered suitable for indicator species analysis in the city of Brussels (see Godefroid and Ricotta 2018). Only the 159 grid cells that are included in the administrative limits of the city for at least 75% of their area were used in this study.

To explore the response of plant species to urban soil and climatic conditions, we used the Ellenberg indicator values (EIVs; Ellenberg et al. 1991) for soil nutrient availability (N), soil reaction (R), soil moisture (F) and light (L). EIVs have been widely used in vegetation science for the assessment of the species ecological niches (Diekmann 2003). For all EIVs, each species is given a value on a 9-point ordinal scale based on expert knowledge, field observations, and partly on direct measurements denoting the position at which plants reach peak abundance along environmental gradients (Godefroid and Dana 2007, Bartelheimer and Poschold 2016). We used the Ellenberg indicator values re-calibrated for the British Isles by Hill et al. (1999) because they are bioclimatically closer to the study area compared to the original values estimated for Central European conditions. First, the multivariate EIV centroid of each land use class was calculated as the mean of the four Ellenberg indicator values of all species (native and alien) present in that land use class (Ellenberg et al. 1991). Next, we calculated the Euclidean distance between the indicator values of each alien species and the centroids of each land use type. Only neophyte species introduced after AD1500 with available EIVs (88 species) were considered in this study. Like in the previous example, the resulting species-to-centroid distances were linearly rescaled to the unit range by dividing each distance by the maximum value in the dataset. By combining the species presence/absence scores with the scaled species-to-centroid similarities s_{jk}, we then calculated the functional association ϕ_{jk} of the alien species with the urban land uses. Finally, using indicator species analysis (999 permutations; two-tailed test), we identified the alien species that best characterize each land use type in terms of both presence/absence scores and functional association values.

4. Results

4.1. Alpine vegetation

The principal component analysis of the 59 vegetation plots sampled on the foreland of the Rutor Glacier (Fig. 1) shows that the three successional stages are functionally well distinct in ordination space. Along the primary succession, a significant increase of leaf dry matter and leaf carbon content is observed together with a decrease of specific leaf area and leaf nitrogen content.

Using traditional indicator species analysis on presence/absence scores, we found 24 species showing significant association ($p < 0.05$, two-tailed test) with one of the three successional stages (Table 1). If the same analysis is performed on species abundances, the number of indicator species increased to 28. Finally, if the indicator species analysis is performed on the functional association ϕ_{jk} of the Alpine species with the three successional stages (thus considering both the species abundances *and* their functioning), the number of indicator species increased to 31.

Apart from *Veronica bellidioides*, all species that were identified as diagnostic of one or more successional stages in terms of presence/absence scores are also diagnostic of the same successional stages in terms of species abundances (see Table 1). Likewise, all abundance-based diagnostic species represent a subset of the functionally-based diagnostic species. Therefore, traditional indicator species analysis and functional indicator species analysis are not in contrast with each other. For some species, occurrence-based or distance-based analysis is powerful enough to highlight their diagnostic value. For some other species, we also need to consider their functional traits. In this sense, at least in our case, functional indicator species analysis enables to highlight the diagnostic value of a larger pool of species.

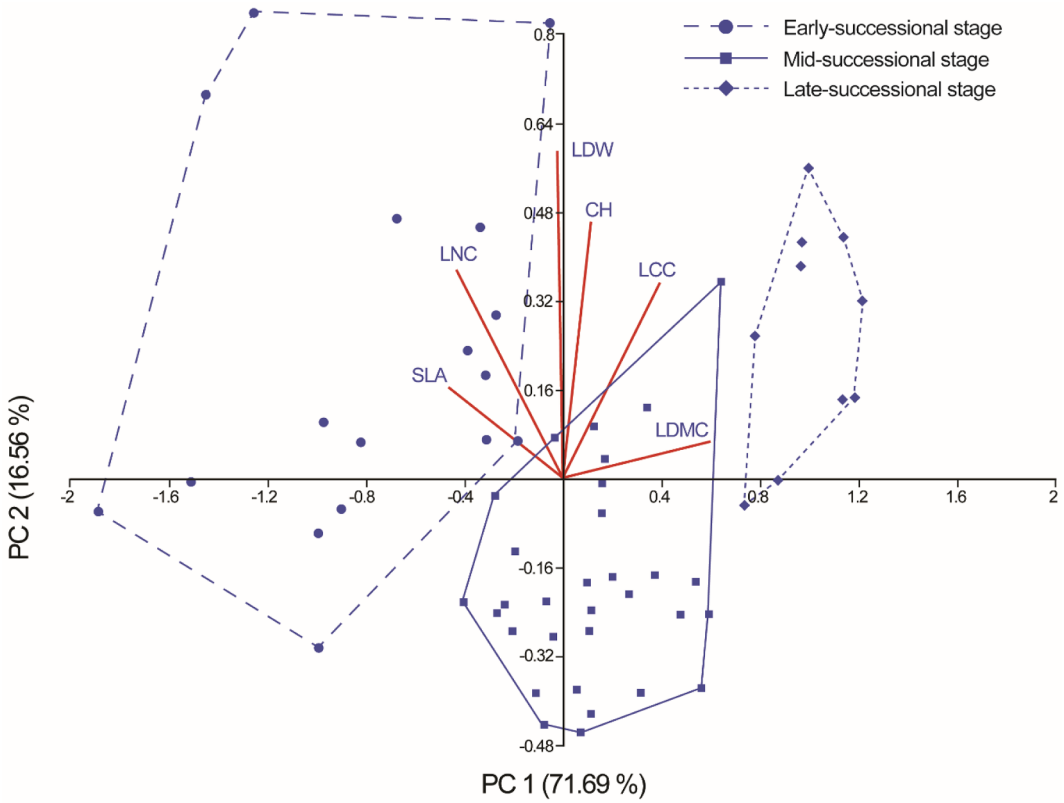

Fig. 1. Biplot of the principal component analysis of the 59 plots of Alpine vegetation with the convex hulls of the three stages identified along the primary succession. The amount of variance explained by the first two axes is shown in brackets. CH = canopy height, LDMC = leaf dry matter content, LDW = leaf dry weight, SLA = specific leaf area, LNC = leaf nitrogen content, LCC = leaf carbon content.

4.2. Urban alien flora of Brussels

Godefroid and Ricotta (2018) highlighted a significant relationship between land use composition and EIVs, such that 1 km^2 cells with similar land use composition tend to be similar also in terms of Ellenberg indicator values. Compared to less human impacted land uses, densely urbanized areas have on average lower EIVs for soil moisture and higher EIVs for light, soil reaction (pH) and nitrogen availability.

Traditional indicator species analysis on presence/absence scores identified 24 alien species showing significant association ($p < 0.05$, two-tailed test) with certain land use classes (Table 2). Therefore, it seems that a relevant portion of alien species has rather narrow ecological requirements allowing them to differentiate between more or less impacted land use types (for details, see Godefroid and Ricotta 2018). The largest number of diagnostic species is associated to open built-up areas (13 species), whereas for presence/absence scores, only one single species, *Sisymbrium altissimum*, is diagnostic of the industrial areas.

If the indicator species analysis is performed on the functional association ϕ_{jk} of the alien species with the urban land uses (taking into account both the species presence/absence scores *and* their EIVs), the number of species that are significantly associated with at least one land use class increases to 35. As in the previous example, all species that were diagnostic of one or more land use class in terms of presence/absence scores are also diagnostic of the same land use classes in terms of EIVs.

5. Discussion and conclusions

Plant functional traits have been increasingly used to explore patterns of co-occurrence in plant communities (e.g. McGill et al. 2006, Adler et al. 2013, Nathan et al. 2015), but this approach has been overlooked in indicator species analyses. Although diagnostic species are generally used as ecological indicators of community or habitat types (De Cáceres et al. 2010), most of the methods used for their determination do not take into account their functional traits. In this paper we thus developed a method for indicator species analysis which combines the species abundances with their functional traits.

Tested on an urban alien flora, this method has demonstrated its effectiveness. First, by allowing more indicator species to be detected (35 instead of 24 with the traditional method). Second, because the 11 newly identified indicators make ecological sense. Among these, there are for example several species escaped from gardens, such as *Asparagus officinalis*, *Aster lanceolatum* or *Syringa vulgaris*, logically identified as indicators of open built-up areas. Also, species that are typically recorded in highly urbanized areas (e.g. wasteland and pavements), like *Coronopus didymus* and *Hirschfeldia incana*, were correctly detected as indicators of densely built-up areas, while the classical approach was unable to identify them as such.

Likewise, the early-successional stage of the Rutor glacier foreland is comprised of fast-growing species with strongly ruderal characteristics *sensu* Grime (1977), which highlight the influence of disturbance on pioneer communities. By contrast, the mid and late-successional stages are progressively characterized by an increasing number of stress-tolerator species that are adapted to substantial periods of low temperature (Caccianiga et al. 2006). As shown in Fig. 1, this shift from ruderal to stress-tolerator communities along the primary succession correlates strongly with changes in resource-use traits, such as increasing leaf dry matter content (LDMC) and decreasing specific leaf area (SLA). In this view, the three newly identified indicators, *Agrostis rupestris*, *Luzula spicata* and *Veronica bellidioides* are all stress-tolerator species with high LDMC and low SLA (see Caccianiga et al. 2006, Table 2). Therefore, while the usual abundance-based approach was unable to identify them as diagnostic, their association with the mid-

Table 1
Alpine species with significant diagnostic power for the three successional stages on the Rutor glacier foreland in terms of presence/absence scores (P/A), species abundances (AB) and functional association values ϕ_{jk} ($p < 0.05$, 999 randomizations, two-tailed test). The functional association values are obtained by combining the species abundances with their functional traits. Nomenclature according to Pignatti (1982).

Indicator Species	Early Succession			Mid Succession			Late Succession		
	P/A	AB	ϕ_{jk}	P/A	AB	ϕ_{jk}	P/A	AB	ϕ_{jk}
Achillea moschata Wulfen				x	x	x			
Adenostyles leucophylla (Willd.) Rchb.	x	x	x						
Agrostis rupestris All.						x			
Anthoxanthum odoratum L.				x	x	x			
Arabis alpina L.	x	x	x						
Avenula versicolor (Vill.) M. Laínz							x	x	x
Carex curvula All.							x	x	x
Carex sempervirens Vill.							x	x	x
Cerastium uniflorum Clairv.	x	x	x						
Erigeron uniflorus L.				x	x	x			
Festuca halleri All.					x	x			
Gnaphalium supinum L.	x	x	x						
Hieracium angustifolium Hoppe				x	x	x			
Homogyne alpina (L.) Cass.							x	x	x
Juncus trifidus L.							x	x	x
Leucanthemopsis alpina (L.) Heywood	x	x	x						
Linaria alpina (L.) Mill.	x	x	x						
Luzula lutea (All.) DC.							x	x	x
Luzula spicata (L.) DC.						x			
Minuartia recurva (All.) Schinz & Thell.				x	x	x			
Myosotis alpestris F. W. Schmidt				x	x	x			
Oxyria digyna (L.) Hill	x	x	x						
Phleum rhaeticum (Humphries) Rauschert		x	x						
Poa alpina L.				x	x	x			
Saxifraga aizoides L.	x	x	x						
Saxifraga bryoides L.					x	x			
Silene acaulis (L.) Jacq.				x	x	x			
Trifolium badium Schreb.		x	x						
Trifolium pallescens Schreb.				x	x	x			
Tussilago farfara L.	x	x	x						
Veronica bellidioides L.							x		x

successional stage (*Agrostis rupestris* and *Luzula spicata*) and the late-successional stage (*Veronica bellidioides*) is supported by their functional characteristics.

Like in traditional indicator species analysis which randomizes the species occurrences or abundances among plots, functional indicator species analysis randomizes the functional abundances π_{jn} of species among plots. As a result, there is no ecological inconsistency between the abundance-based and functional-based methods for identifying diagnostic species. If the functional characteristics of a target species are strongly correlated with specific ecosystem properties such that the species is exclusive or highly preferential of a given community, occurrence-based analysis is powerful enough to highlight its diagnostic value. Some more ubiquitous species show significant differences among communities in terms of their abundances, while for a third group of species, a functional fine-tuning to the specific biotic and abiotic conditions of their habitat is needed to highlight their diagnostic value. Therefore, by applying a 'cascade' of different methods, we can identify step by step groups of species with different diagnostic power, thus improving the flexibility of the species association with a target group of plots.

Like for any other ecological indicator that is based on the analysis of the functional traits of species, the most relevant decisions to be taken include which traits to use, the choice of the dissimilarity coefficient and how to calculate the functional centroid (Anderson 2006, Lavorel et al. 2008). Note that a relevant aspect of the proposed method is that for the calculation of the multivariate functional centroid of a target group of plots, abundant species contribute more than rare species. The biological justification for this differential weighting of rare and abundant species is that the extent to which plant species affect a wide variety of ecosystem functions, such as carbon balance or nutrient dynamics, is largely determined by their abundance. This 'mass-ratio' effect is determined by the observation that, at least for plants, a larger body mass implies major contribution to syntheses, resource fluxes and degradative processes (see Grime 1998). Alternative methods to calculate the functional centroid of a target group of plots may consist in weighting all species equally, irrespective of their abundance, or in iteratively excluding the target species from the calculation of the multivariate centroid by means of leave-one out methods.

Finally note that, from a more technical viewpoint, the interpretation of the functional dissimilarity d_{jk} as a fuzzy degree of functional distinctness between a given species j and a target group of plots k opens the way for introducing a high degree of flexibility in the computation of the species functional abundances π_{jn}.

Fuzzy set theory has been already described elsewhere (Klir and Yuan 1995; Klir and Wierman 1999), and the reader is referred to these papers for details. For our purposes, it is sufficient to observe that if we define the quantity d_{jk} in terms of fuzzy set theory, then functional similarity becomes its fuzzy complement $s_{jk} = C(d_{jk})$. Based on this fuzzy characterization of the relationship between d_{jk} and s_{jk}, we can use a wide range of generalized complement operators for defining a biologically meaningful measure of fuzzy compatibility between a given species and a target group of plots.

Besides the standard fuzzy complement operation $C(d_{jk}) = 1 - d_{jk}$ there exists a broad class of functions $C(d_{jk})$: $[0, 1] \rightarrow [0, 1]$ that allow to assign a value $C(d_{jk})$ in the range $[0, 1]$ to each dissimilarity value d_{jk}. These functions need to conform to a set of requirements that make them suitable as fuzzy generalizations of the standard complement operation (see Klir and Wierman 1999). The resulting values can then be used to calculate the species functional abundances in each plot as $\pi_{jn} = x_{jn} \times C(d_{jk})$. For example, a fuzzy complement operation that might be used in the context of indicator species analysis is the threshold function (Klir and Yuan 1995). For a threshold value t in the

Table 2
Alien species with significant diagnostic power for the selected urban land use classes of Brussels in terms of presence/absence scores (P/A) and functional association values ϕ_{jk} ($p < 0.05$, 999 randomizations, two-tailed test). UD = densely built-up urban areas, UO = open built-up urban areas, FOR = urban forests, IND = industrial areas, AGR = agricultural areas. Nomenclature according to Lambinon et al. (2012).

Indicator Species	UD		UO		FOR		IND		AGR	
	P/A	ϕ_{jk}	P/A	ϕ_{jk}	P/A	ϕ_{jk}	P/A	ϕ_{jk}	P/A	ϕ_{jk}
Acer platanoides L.			x	x						
Acer pseudoplatanus L.				x		x				
Allium schoenoprasum L.			x	x						
Alnus incana (L.) Moench					x	x				
Asparagus officinalis L.				x						
Aster lanceolatus Willd.				x						
Barbarea intermedia Boreau				x						
Buddleja davidii Franch.	x	x						x		
Castanea sativa Mill.			x	x	x	x				
Coronopus didymus (L.) Smith		x								
Cymbalaria muralis P. Gaertn., B. Mey. et Scherb.			x	x						
Fallopia japonica (Houtt.) Ronse Decraene				x						
Fallopia sachalinensis (F. Schmidt Petrop.) Ronse Decraene			x	x						
Galinsoga quadriradiata Ruiz et Pav.	x	x								
Hirschfeldia incana (L.) Lagrèze-Fossat		x								
Impatiens glandulifera Royle				x						x
Impatiens parviflora DC.					x	x				
Juncus tenuis Willd.					x	x				
Ligustrum ovalifolium Hassk.									x	x
Mahonia aquifolium (Pursh) Nutt.	x	x								
Matricaria discoidea DC.	x	x								
Phalaris canariensis L.	x	x								
Pseudofumaria lutea (L.) Borkh.			x	x						
Robinia pseudoacacia L.			x	x						
Saponaria officinalis L.								x		
Sisymbrium altissimum L.							x	x		
Solidago canadensis L.			x	x						
Solidago gigantea Ait.			x	x						
Symphoricarpos albus (L.) S.F. Blake			x	x						x
Syringa vulgaris L.				x						
Taxus baccata L.			x	x		x				
Trifolium hybridum L.									x	x
Veronica filiformis Smith			x	x						
Veronica persica Poiret			x	x					x	x
Vinca major L.										x

range [0, 1], this function is defined as:

$$C(d_{jk}) = \begin{cases} 1 & for \quad d_{jk} < t \\ 0 & for \quad d_{jk} \geqslant t \end{cases} \tag{3}$$

According to Eq. (3), the species occurrences or abundances in a given group of plots are considered only if their functional dissimilarity to the functional centroid of that group of plots is below an a-priori defined threshold.

Two additional complement operators that are commonly used in multivariate analysis for computing a measure of similarity in the range [0, 1] from the corresponding distances are: $C(d_{jk}) = \sqrt{1 - d_{jk}}$ or $C(d_{jk}) = \sqrt{1 - d_{jk}^2}$ (see Legendre and Legendre 2012). All these functions enable the practitioner to change the sensitivity of the complement operator to high or low values of d_{jk} according to his/her specific requirements. For additional aspects on the ecological applications of fuzzy complement functions, see Ricotta (2008).

Overall, we see the flexibility associated to fuzzy complement operators as a great potential advantage for ecologists as it allows to compute relevant aspects of the species diagnostic power from different viewpoints and motivations. Therefore, we hope that the use of functional indicator species analysis will help improve the ecological characterization of plant and animal assemblages.

Author contributions

C.R. conceived the idea; All authors developed the methodology and analyzed the data; M.CAC. and B.C. collected the data for Italy and S.G. for Belgium; C.R. took the lead in writing the main text and M.CAR. in writing the R script. All authors revised the manuscript critically and approved the final version.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ecolind.2020.106761.

References

Adler, P.B., Fajardo, A., Kleinhesselink, A.R., Kraft, N.J.B., 2013. Trait-based tests of coexistence mechanisms. Ecol. Lett. 16, 1294–1306.
Anderson, M.J., 2006. Distance-based tests for homogeneity of multivariate dispersions. Biometrics 62, 245–253.
Bartelheimer, M., Poschold, P., 2016. Functional characterizations of Ellenberg indicator values – a review on ecophysiological determinants. Funct. Ecol. 30, 506–516.
Caccianiga, M., Luzzaro, A., Pierce, S., Ceriani, R.M., Cerabolini, B., 2006. The functional basis of a primary succession resolved by CSR classification. Oikos 112, 10–20.
Chytrý, M., Tichý, L., Holt, J., Botta-Dukát, Z., 2002. Determination of diagnostic species with statistical fidelity measures. J. Veg. Sci. 13, 79–90.
Díaz, S., Kattge, J., Cornelissen, J.H.C., Wright, I.J., Lavorel, S., et al., 2016. The global spectrum of plant form and function. Nature 529, 167–171.
Díaz, S., Lavorel, S., de Bello, F., Quétier, F., Grigulis, K., Robson, T.M., 2007.

Incorporating plant functional diversity effects in ecosystem service assessments. Proc. Natl. Acad. Sci. 104, 20684–20689.
De Cáceres, M., Legendre, P., 2009. Associations between species and groups of sites: indices and statistical inference. Ecology 90, 3566–3574.
De Cáceres, M., Legendre, P., Moretti, M., 2010. Improving indicator species analysis by combining groups of sites. Oikos 119, 1674–1684.
De Cáceres, M., Legendre, P., Wiser, S.K., Brotons, L., 2012. Using species combinations in indicator value analyses. Methods Ecol. Evol. 3, 973–982.
Diekmann, M., 2003. Species indicator values as an important tool in applied plant ecology - a review. Basic Appl. Ecol. 4, 493–506.
Dufrêne, M., Legendre, P., 1997. Species assemblages and indicator species: the need for a flexible asymmetrical approach. Ecol. Monogr. 67, 345–366.
Ellenberg, H., Weber, H.E., Düll, R., Wirth, V., Werner, W., Paulissen, D., 1991. Zeigerwerte von Pflanzen inMitteleuropa. Scripta Geobotanica 18, 1–48.
Garnier, E., Cortez, J., Billes, G., Navas, M.L., Roumet, C., Debussche, M., Laurent, G., Blanchard, A., Aubry, D., Bellmann, A., Neill, C., Toussaint, J.P., 2004. Plant functional markers capture ecosystem properties during secondary succession. Ecology 85, 2630–2637.
Garnier, E., Stahl, U., Laporte, M.-A., Kattge, J., Mougenot, I., et al., 2017. Towards a thesaurus of plant characteristics: an ecological contribution. J. Ecol. 105, 298–309.
Godefroid, S., Dana, E.D., 2007. Can Ellenberg's indicator values for Mediterranean plants be used outside their region of definition? J. Biogeogr. 34, 62–68.
Godefroid, S., Ricotta, C., 2018. Alien plant species do have a clear preference for different land uses within urban environments. Urban Ecosystems 21, 1189–1198.
Grime, J.P., 1977. Evidence for the existence of three primary strategies in plants and its relevance to ecological and evolutionary theory. Am. Nat. 111, 1169–1194.
Grime, J.P., 1998. Benefits of plant diversity to ecosystems: immediate, filter and founder effects. J. Ecol. 86, 902–910.
Grime, J.P., Pierce, S., 2012. The evolutionary strategies that shape ecosystems. Wiley-Blackwell, Chichester, UK.
Hill, M.O., Mountford, J.O., Roy, D.B., Bunce, R.G.H. (1999) Ellenberg's indicator values for British plants. ECOFACT Vol. 2, Technical Annex. Institute of Terrestrial Ecology, Huntingdon, UK.
Klir, G., Yuan, B., 1995. Fuzzy Sets and Fuzzy Logic: Theory and Applications. Prentice Hall, Upper Saddle River, New Jersey.
Klir, G.J., Wierman, M.J., 1999. Uncertainty-Based Information. Physica-Verlag, Heidelberg.
Kraft, N.J.B., Godoy, O., Levine, J.M., 2015. Plant functional traits and the multidimensional nature of species coexistence. Proceed. Nat. Acad. Sci. USA 12, 797–802.
Lambinon, J., Delvosalle, L., Duvigneaud, J. (2012) Nouvelle Flore de la Belgique, du Grand-Duché de Luxembourg, du Nord de la France et des Régions voisines. Jardin botanique national de Belgique, Meise.
Lavorel, S., Grigulis, K., McIntyre, S., Williams, N.S.G., Garden, D., Dorrough, J., Berman, S., Quétier, F., Thébault, A., Bonis, A., 2008. Assessing functional diversity in the field – methodology matters!. Funct. Ecol. 22, 134–147.
Lavorel, S., Garnier, E., 2002. Predicting changes in community composition and ecosystem functioning from plant traits: revisiting the Holy Grail. Funct. Ecol. 16, 545–556.
Legendre, P., Legendre, L., 2012. Numerical Ecology. Elsevier, Amsterdam.
Lososová, Z., Horsák, M., Chytrý, M., Čejka, T., Danihelka, J., Fajmon, K., Hájek, O., Juřičková, L., Kintrová, K., Láníková, D., Otýpková, Z., Řehořek, V., Tichý, L., 2011. Diversity of Central European urban biota: effects of human-made habitat types on plants and land snails. J. Biogeogr. 38, 1152–1163.
Mason, N.W.H., de Bello, F., 2013. Functional diversity: a tool for answering challenging ecological questions. J. Veg. Sci. 24, 777–780.
McGill, B.J., Enquist, B.J., Weiher, E., Westoby, M., 2006. Rebuilding community ecology from functional traits. Trends Ecol. Evol. 21, 178–185.
Moles, A.T., Perkins, S.E., Laffan, S.W., Flores-Moreno, H., Awasthy, M., et al., 2014. Which is a better predictor of plant traits: temperature or precipitation? J. Veg. Sci. 25, 1167–1180.
Moretti, M., Legg, C., 2009. Combining plant and animal traits to assess community functional responses to disturbance. Ecography 32, 299–309.
Pignatti, S., 1982. Flora d'Italia. Edagricole, Bologna.
Podani, J., Csányi, B., 2010. Detecting indicator species: some extensions of the IndVal measure. Ecol. Ind. 10, 1119–1124.
Pyšek, P., Richardson, D.M., Rejmánek, M., Webster, G., Williamson, M., Kirschner, J., 2004. Alien plants in checklists and floras: towards better communication between taxonomists and ecologists. Taxon 53, 131–143.
Ricotta, C., Carboni, M., Acosta, A.T.R., 2015. Let the concept of indicator species be functional!. J. Veg. Sci. 26, 839–847.
Ricotta, C., de Bello, F., Moretti, M., Caccianiga, M., Cerabolini, B., Pavoine, S., 2016. Measuring the functional redundancy of biological communities: a quantitative guide. Methods Ecol. Evol. 7, 1386–1395.
Ricotta, C., 2008. Computing additive -diversity from presence and absence scores: A critique and alternative parameters. Theor. Popul Biol. 73, 244–249.
Tichý, L., Chytrý, M., 2006. Statistical determination of diagnostic species for site groups of unequal size. J. Veg. Sci. 17, 809–818.
Voigt, W., Perner, J., Hefin Jones, T., 2007. Using functional groups to investigate community response to environmental changes: two grassland case studies. Glob. Change Biol. 13, 1710–1721.
Wright, I.J., Reich, P.B., Westoby, M., Ackerly, D.D., Baruch, Z., et al., 2004. The worldwide leaf economics spectrum. Nature 428, 821–827.

```
#########################################################################
################
#Function Phi.Val to calculate Functional Indicator Value of species for
groups of plots.
#
#########################################################################
################

#Disclaimer:
#users of this code are cautioned that, while due care has been taken and
it is believed accurate, it has not been rigorously tested and its use and
results are solely the responsibilities of the user.  #Description:
#given a matrix of N species abundances × M plots, together with a factor
with group belongings for M plots, and either a matrix of T traits x N
species or a matrix of T Community weighted Means x M plots, this function
calculates the Functional Indicator Value of species for groups of plots
as described in the main text, and assess the significance of the
association with the groups based on randomizations.   #Dependencies:
#FD, vegan, picante, abind

#Usage:
#    Phi.Val(samp,    groups,    traits,    spxplot.phi=NULL,    nrep=99,
dist.method="euclidean", null.model = "richness")

#Arguments:
#samp: Community data matrix
#groups: factor with group belonging for each plot in samp
#traits: Traits x Species matrix
#spxplot.phi: functional association/abundance phi of species in plots
(optional, for greater flexibility in calculating phis)
#null.model: Null model to use (see ?randomizeMatrix). Default = "richness"
(this randomizes the within rows, so functional associations are randomized
among plots/groups)
#dist.method= Distance method to calculate species-to-plot multivariate
functional distance (see ?vegdist). Default= "euclidean"
#nrep: Number of randomizations

#Returns:
#A 3-dimensional array in which rows are species, columns relate to the
mean functional association phi to each group (column1=observed values,
column 2=mean of randomized values, column3=p-value), on the third
dimension are the different groups of plots.

#########################################################################
########
require(FD)            #functcomp
require(vegan)         #vegdist
require(picante)       #randomize matrix
require(abind)


Phi.Val <- function(samp, groups, traits, spxplot.phi=NULL, nrep=99,
dist.method="euclidean", null.model = "richness"){

      # Calculate functional association phi ------------------------

      if(is.null(spxplot.phi)) {

          #calculate abundance by group
          group.abund<-sapply(split(samp,groups),colSums)
```

```
          #calculate cwms of group
          cwms.group<-
sapply(traits,function(x){colSums(x*group.abund)/colSums(group.abund)})

          #standardize traits and cwms
          all<-rbind(traits,cwms.group)
          all.st<-decostand(all, "standardize")

          #calculate species to group multivariate distances
          all.dist<-as.matrix(vegdist(all.st, method=dist.method))
            #method="euclidean" default
          spxgroup.dist<-
all.dist[row.names(traits),row.names(cwms.group)]       #extract
distances to group centroids
          spxgroup.d <- spxgroup.dist/max(spxgroup.dist)
            #devide by max value to obtain 0<d<1

          #Obtain species similarities to group
          spxgroup.s <- 1-spxgroup.d

          #Calculate functional association phi to plots
            spxplot.s <- spxgroup.s[,groups]
          spxplot.phi <- spxplot.s * t(samp)

      }

    # Observed values of average phi per group ----------------------

    spxgroup.phi<-t(apply(spxplot.phi,1,tapply,groups,mean))

    # Randomize values   ------------------------------------------

    #simulated means per group
    simus                              <-                              array(NA,
dim=c(nrow(spxgroup.phi),ncol(spxgroup.phi),nrep+1),
dimnames=list(rownames(spxgroup.phi),colnames(spxgroup.phi),     c("obs",
1:nrep)))
    simus[,,1] <- spxgroup.phi

    for (n in 1:nrep+1){
      tmp<-randomizeMatrix(spxplot.phi, null.model = null.model)
      simus[,,n]<-t(apply(tmp,1,tapply,groups,mean, na.rm=T))
        }

    #calculate mean of simulated values, p-values
    rand.means <- apply(X = simus[,,-1], MARGIN = 1:2, FUN = mean, na.rm =
TRUE)
      p.values   <-   apply(X   =   simus,   MARGIN   =   1:2,   function(o)
mean(o[names(o)=="obs"] <= o, na.rm = TRUE))

    # put it together
    out.tmp<-
abind(obs=spxgroup.phi,rand.means=rand.means,p.values=p.values,along=3)
    out<-aperm(out.tmp, c(1,3,2))

    return(out)
}
```