The study of fidelity measures in the context of using them as a threshold criterion in the allocation of diagnostic species

Article *in* Ecologica Montenegrina · April 2020

DOI: 10.37828/em.2020.30.2

CITATIONS
0

READS
38

2 authors, including:

Igor V. Goncharenko
National Academy of Sciences of Ukraine
191 PUBLICATIONS **1,135** CITATIONS

SEE PROFILE

Ecologica Montenegrina
ISSN 2336-9744 (online) | ISSN 2337-0173 (print)

Ecologica Montenegrina 30: 28-36 (2020)
This journal is available online at: www.biotaxa.org/em
http://dx.doi.org/10.37828/em.2020.30.2

Article

The study of fidelity measures in the context of using them as a threshold criterion in the allocation of diagnostic species

IGOR GONCHARENKO[1] & OLEXANDER SENCHYLO[2]

[1]*Institute for Evolutionary Ecology, National Academy of Sciences of Ukraine, 37 Academician Lebedev Str., 03143 Kyiv, Ukraine; e-mail: goncharenko.ihor@gmail.com*; corresponding author
[2]*Taras Shevchenko National University of Kyiv, ESC «Institute of biology and medicine», 64/13 Volodymyrska Str., 01601 Kyiv, Ukraine; e-mail: senchylo@gmail.com*

Received 16 March 2020 | Accepted by *V. Pešić*: 6 April 2020 | Published online 8 April 2020.

Abstract
The paper is devoted to a comparative study of species fidelity measures from the viewpoint of using them as threshold criteria for distinguishing diagnostic species in the vegetation classification. The computational experiment was carried out on two test data sets on vegetation of the Dnieper floodplain in the forest-steppe zone of Ukraine (316 relevés) and forest vegetation of the city of Kyiv and suburbs (832 relevés). Classification of relevés was obtained using cluster analysis by the flexible-beta method on the Bray-Curtis distance matrix and the optimal number of clusters was determined using the Optimclass approach. In all fidelity calculations, classification of relevés has remained constant for comparability reasons. We compared the distribution properties of fidelity values using the most common fidelity indices, including phi-coefficient, IndVal, Ochiai index, chi-square statistic, Bruelheide's u, Fisher test, G-statistic, and TCR. Group-equalized and non-equalized modifications of fidelity indices were considered as separate measures. Each species was accounted in the cluster with the highest fidelity value. Such indicators were estimated as range and quartiles of values, evenness of distribution, variation of shares of diagnostic species with a variation of fidelity thresholds and vice versa, shares of species exceeding fidelity threshold in more than one cluster.

Key words: vegetation classification, phytosociology, species fidelity, diagnostic species, fidelity threshold.

Introduction

The concept of species fidelity is an important part of the Braun-Blanquet methodology (Braun-Blanquet, 1964; Westhoff & Van Der Maarel, 1978). Fidelity is determined by preferential concentration of the species occurrences in one syntaxon. Indices, or coefficients, of fidelity are used to statistically evaluate the degree of such “concentration” of species (Chytrý et al., 2002). This direction remains in the focus of attention of European phytocoenologists (Barkman, 1989; Botta-Dukát & Borhidi, 1999; Bruelheide, 2000; Chytrý et al., 2002; Tichy & Chytry, 2006; De Cáceres et al., 2008; Willner et al., 2009; Podani & Csányi, 2010). Very few works are devoted to this issue in American phytocoenology (Kusbach et al., 2012). Currently the works appear aimed at generalization and development of "universal" mathematical approaches to fidelity measurement (De Cáceres & Legendre, 2009). Application of statistical methods for estimation of species fidelities has acquired special importance in connection with formalization of the Braun-Blanquet method and wide spread of computer software for sorting phytocoenotic matrices (Tichý, 2002).

Statistically faithful are the species whose fidelity score exceeds the user-defined threshold (in this sense we will use the term "diagnostic species" hereinafter). Having sorted the phytocoenotic matrices, the diagnostic species appear in the upper (characterizing) part and shape a diagonal. The phytocoenotic cluster (phytocoenon, syntaxon, synoptic column, site groups), in which the species reaches the maximum fidelity value, we will call "optimal" in terms of matching the ecological amplitudes of the species and syntaxon. The total number of diagnostic species and the number of clusters that have diagnostic species are used as a criterion for determining the number of phytocoenotic clusters in the data (Tichý et al., 2010). In Juice software, which is very popular among phytocoenologists, 16 fidelity indices (Juice ver. 7.1.5) are available. Despite the considerable number of works on practical application of fidelity indices for vegetation classification, some issues remain insufficiently covered.

The purpose of the article is to perform a comparative assessment of the most common fidelity indices in terms of using them as threshold criteria. We will focus on four aspects that are relevant to the use of fidelity measures as a threshold criterion: 1) statistical parameters of the distributions of different fidelity coefficients; 2) thresholds to extract a certain percentage of diagnostic species by different fidelity coefficients; 3) sensitivity of different fidelity measures to the multiple affinity of diagnostic species to more than one cluster; 4) comparison of group-equalized and non-equalized versions of the same fidelity indices in all the foregoing respects.

Material and methods

The test data sets

Two test data sets were involved in a comparative study of fidelity measures. The first test data set consists of 316 relevés of vegetation in the Dnieper floodplain in the forest-steppe zone of Ukraine (Senchylo et al., 1997; Senchylo et al., 1998; Senchylo, 2009). The second test data set includes 832 relevés of forest semi-natural and anthropogenic vegetation of the city of Kyiv and environs (Goncharenko & Golik, 2015; Golik & Goncharenko, 2017; Goncharenko & Yatsenko, 2020). In both cases, the vegetation data itself is not the focus of our article, but was used to demonstrate the properties of the distributions of fidelity indices, which will be discussed. The main characteristics of both data sets are summarized in Table 1.

Table 1. Characteristics of two test data sets used in the current study

	Test data set 1	Test data set 2
Number of relevés	316	832
Plot size (m^2)	16–100	100
Number of species	479	334
Mean number of species per plot	16.6	23.5
Length of first DCA axis	2.54	5.76
Share of species with frequency less than 1% of number of relevés	0.39	0.35
Geographic location	the Dnieper floodplain, within the forest-steppe zone of Ukraine	Kyiv city and suburbs
Classes of vegetation	*Molinio-Arrhenatheretea, Phragmito-Magnocaricetea, Carpino-Fagetea, Salicetea purpureae*	*Carpino-Fagetea, Quercetea robori-petraeae, Robinietea, Salicetea purpureae*
Sources of data	(Senchylo et al., 1997; Senchylo et al., 1998; Senchylo, 2009)	(Goncharenko & Golik, 2015; Golik & Goncharenko, 2017; Goncharenko & Yatsenko, 2020)

The classification of relevés

Classification of relevés was carried out using the hierarchical cluster analysis by the method of "flexible beta" (Lance & Williams, 1967). The cluster analysis was applied to the Bray-Curtis distance matrix

calculated for the preliminary log-transformed data. Such data-analytical combination (DAC) is often used and well-proven in terms of the interpretability of phytocoenotic clusters (Lötter et al., 2013). Optimal number of clusters was determined by maximizing the total number of phytocoenotic clusters with diagnostic species. This approach is known as OptimClass-2 (Tichý et al., 2010). Diagnostic species were selected at the threshold of the indval index > 0.5, and the clusters with at least one diagnostic species were taken into account in OptimClass-2 approach.

Species fidelity calculations

All further fidelity calculations were made against the background of unchanged classification of relevés for comparability of fidelity scores. Species fidelities were calculated for 11 measures or coefficients (Table 2). Indices that exist in the group-equalized and non-equalized modifications were calculated separately. Only symmetric fidelity measures, the so-called joint fidelity (Chytrý et al., 2002), were taken into account in further analyzes since the traditional interpreting of species fidelity concerns primarily symmetric measures of species-to-cluster association.

Table 2. Fidelity indices used in the current study.

№	Abbreviation	Fidelity indices, sources of information
1	phi	phi-coefficient, Pearson's phi (Chytrý et al., 2002)
2	phi.g	phi-coefficient group-equalized (Tichy & Chytry, 2006)
3	indval	Dufrêne-Legendre Indicator Value Index (Dufrêne & Legendre, 1997)
4	indval.g	IndVal group-equalized (Tichy & Chytry, 2006)
5	cos	Ochiai index, cosine coefficient (De Cáceres et al., 2008)
6	cos.g	Ochiai index group-equalised (Tichy & Chytry, 2006)
7	chi	chi-square statistic (Chytrý et al., 2002)
8	u.hyp	Bruelheide's corrected u value (Bruelheide, 2000)
9	Fisher	Fisher test (Chytrý et al., 2002)
10	g	G statistic (Chytrý et al., 2002)
11	TCR	Total cover ratio (Willner et al., 2009)

At the first stage of fidelity calculations for each of two test data sets 11 matrices of species-to-cluster fidelity scores were obtained. Taking into account the large amount of these data, these intermediate matrices are not presented in the article. On the second stage of fidelity calculations, we found the maximal fidelity values (the greatest value of fidelity coefficients or minimal value of the Fisher index) for each species and took them into account for further analyzes. We have run three groups of analyses using maximal fidelity values for each species in two data sets (Table 1). The first group of analyses was aimed at investigations of statistical parameters of the distributions of different fidelity coefficients. In the second group of analyses, we repeatedly changed fidelity thresholds in order to obtain particular percentage of diagnostic species. Finally, we analyzed effects of different fidelity thresholds on multiple affinity of diagnostic species to more than one cluster.

The decrease of fidelity threshold increases affinity of a species to different clusters (Chytrý et al. 2002). Multiple affinity of a species to different clusters is a serious problem that results with ambiguous synoptic tables, especially in the case when clusters in the table are not ordered along an environmental gradient. We have changed the threshold of diagnostic species, in order to detect sensitivity of different fidelity measures to the multiple affinity problem.

All calculations were performed in the R software, ver. 3.5.3 (https://cran.r-project.org/). At different stages, packages of phytocoenotic orientation were used – vegan (Oksanen et al., 2018), labdsv (Roberts, 2016), indicspecies (De Cáceres & Jansen, 2020). Matrix calculations were carried out by means of built-in functionality of R. Cluster analysis by the flexible beta method was carried out using the cluster::agnes function at beta value = –0.25.

Results

Detecting optimal number of clusters

The optimal number of phytocoenotic clusters was defined as high as 28 for the test data set 1 and 25 for the test data set 2 (Fig. 1). On the curve of the number of clusters with diagnostic species with a growing total number of clusters at the value of abscissa k = 28, the highest value of the ordinate is reached, n = 22 clusters for the test data set 1, and the maximum of ordinate n = 13 achieved in the case of test data set 2.

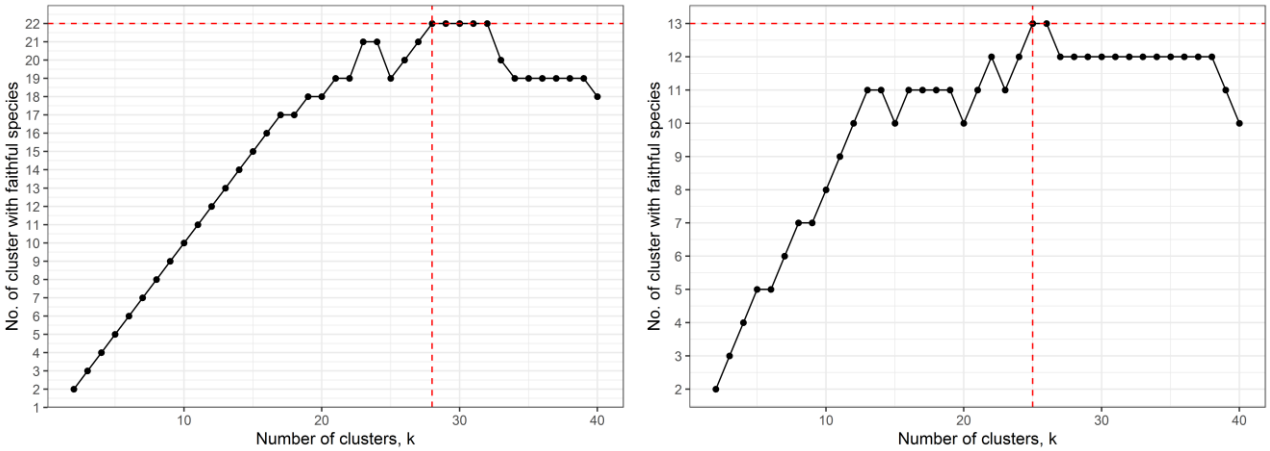

Figure 1. Detection of optimal number of clusters by the Optimclass-2 approach. From left to right: the case study of the test data set 1 and 2.

Distribution properties of fidelity values

Most fidelity indices (indval, phi, cos, TCR) are normalized, so the range of values they reach in the optimal clusters does not exceed 1 (Table 3). The others (u.hyp, g, chi) have upper (or lower limit for the Fisher index) is indefinite. Setting the thresholds according to the values specified in Table 3 results in the allocation of equal shares of species as diagnostic ones. Specifically, when setting thresholds corresponding to the third quartile (q75), then 25% of species (q50 – 50% of species, q25 – 75% and so on) of the total species number will be assigned to diagnostic ones.

Table 3. Five most used percentiles of fidelity scores considering marginal values in optimal clusters for the test data sets 1 and 2.

Quartiles of marginal fidelity values	q0	q25	q50	q75	q100
Shares of diagnostic species	1.00	0.75	0.50	0.25	0.00
Case study: test data set 1					
indval	0.160	0.277	0.366	0.495	0.983
indval.g	0.172	0.277	0.372	0.502	0.991
phi	0.139	0.262	0.339	0.466	0.954
phi.g	0.150	0.263	0.342	0.480	0.956
cos	0.156	0.277	0.356	0.481	0.956
cos.g	0.168	0.274	0.363	0.502	0.958
u.hyp	1.131	2.662	4.637	6.919	15.924
g	0.000	0.001	0.005	0.043	0.753
chi	3.276	23.380	25.667	48.029	254.392
Fisher (in reverse order)	1.17E-01	3.16E-02	1.72E-03	3.84E-06	1.42E-22
TCR	0.038	0.086	0.129	0.222	1.000

..continued on the next page

TABLE 3.

Case study: test data set 2					
indval	0.100	0.235	0.346	0.463	0.894
indval.g	0.107	0.250	0.357	0.494	0.957
phi	0.082	0.210	0.302	0.423	0.853
phi.g	0.103	0.224	0.335	0.453	0.949
cos	0.097	0.229	0.338	0.450	0.853
cos.g	0.109	0.243	0.364	0.478	0.951
u.hyp	0.000	3.379	5.582	9.533	22.528
g	0.000	0.001	0.016	0.178	3.497
chi	3.408	19.980	40.901	90.978	508.106
Fisher (in reverse order)	1.17E-01	3.50E-03	1.55E-06	2.20E-12	4.94E-55
TCR	0.012	0.063	0.135	0.207	0.890

For both data sets 1 and 2, the indices of indval, phi, cos have the quartiles which are close. This also extends to their corrected modifications – indval.g, phi.g, cos.g. Some indices (g, chi, Fisher, TCR) demonstrate asymmetric distributions. For example, the difference between the value of the third quartile (q75) and the median (q50) in case of g-index is only 0.043–0.005 = 0.038, while the difference between the third quartile (q75) and the maximum value (q100) equals 0.753 – 0.043 = 0.710, i.e. 16.5 times greater (see case of test data set 1, Table 3). For example, for the indval index the same ratio is 0.983 / 0.495 ≈ 2 times. While changing the threshold of the indval by 0.366 will increase the share of diagnostic species by next 25% portion (up to 50%), the same for the g-index is caused by shifting the threshold on a very small value (0.038).

Let's compare visually the distributions of some fidelity indices with different degree of asymmetry (Fig. 2).

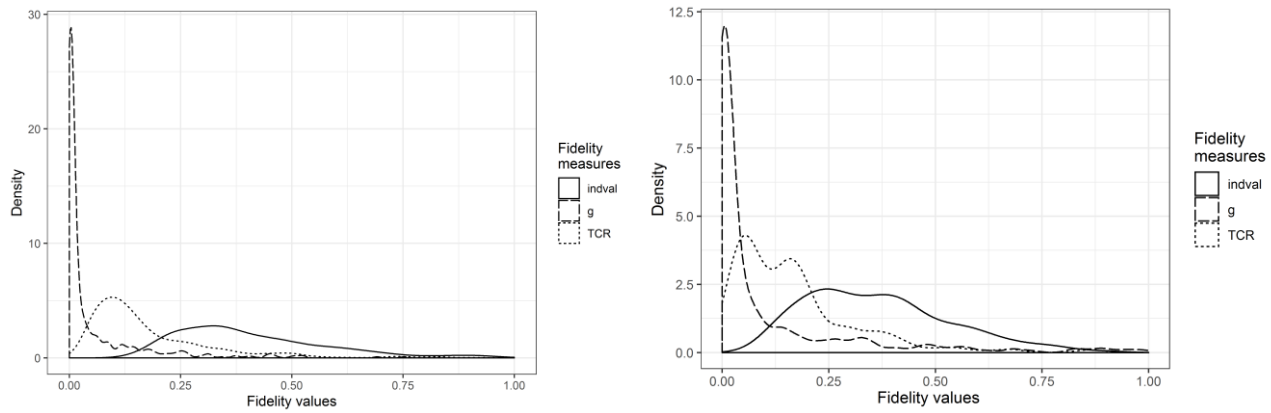

Figure 2. Representing distributions of some fidelity measures with different degree of asymmetry with kernel density estimates. From left to right: the case study of the test data set 1 and 2.

As can be seen, in both cases, left-side asymmetry is observed, but they are very different for different indices. Thus, the TCR index has the largest asymmetry among the three tested, while the indval index is characterized by a more even distribution (Fig. 2).

Thresholds of fidelity measures depending on the shares of diagnostic species

Empirical experience suggests that the fraction of diagnostic species in most cases in vegetation studies usually ranges between 10% and 40% of the total number of species. In Table 4 we chose the growth of shares of diagnostic species as a targeted indicator with the increment of 5% within the above-mentioned limits, and the thresholds of fidelity measures were dependent and have been tabulated.

Table 4. Fidelity thresholds depending on the shares of diagnostic species for the test data sets 1 and 2.

Shares of diagnostic species	indval	indval.g	phi	phi.g	cos	cos.g	u.hyp	g	chi	Fisher	TCR
Case study: test data set 1											
10%	0.63	0.67	0.62	0.63	0.62	0.63	9.68	0.17	94.03	2.61E-10	0.37
15%	0.58	0.59	0.56	0.57	0.57	0.58	8.45	0.11	71.60	1.96E-08	0.29
20%	0.54	0.57	0.51	0.53	0.51	0.53	7.85	0.07	61.81	3.09E-07	0.26
25%	0.50	0.50	0.47	0.48	0.48	0.50	6.92	0.04	48.03	3.84E-06	0.22
30%	0.47	0.47	0.44	0.45	0.45	0.46	6.38	0.03	40.82	1.60E-05	0.19
35%	0.44	0.45	0.41	0.42	0.41	0.43	5.97	0.02	35.79	4.94E-05	0.17
40%	0.41	0.41	0.39	0.40	0.40	0.41	5.62	0.01	31.72	1.68E-04	0.15
Case study: test data set 2											
10%	0.58	0.59	0.55	0.57	0.58	0.59	14.83	0.94	220.07	1.09E-19	0.36
15%	0.51	0.53	0.50	0.50	0.51	0.53	12.76	0.48	162.90	2.18E-16	0.30
20%	0.46	0.49	0.43	0.46	0.46	0.49	11.13	0.31	124.10	2.09E-14	0.24
25%	0.46	0.49	0.42	0.45	0.45	0.48	9.53	0.18	90.98	2.20E-12	0.21
30%	0.40	0.42	0.37	0.40	0.40	0.42	8.62	0.12	74.42	3.84E-11	0.19
35%	0.38	0.40	0.34	0.37	0.38	0.40	7.53	0.07	56.73	7.81E-10	0.17
40%	0.36	0.37	0.31	0.34	0.36	0.37	6.61	0.04	47.43	2.28E-08	0.17

The trends in the thresholds in Table 4 are largely a consequence of the properties of their distributions, which have been discussed. More specifically, a small change of the threshold for g-index from 0.01 to 0.17 results in as high increase as 30%, from 10% to 40%, share of diagnostic species (see case of test data set 1, Table 4). The large asymmetry of the distribution of the Fisher index is compensated by the fact that the exponent is used as the threshold, which is equivalent to the log-transformation and makes the increment step more aligned. One of the quite reasonable demands to the indices, which could be used as threshold criteria, is naturally comparability of scores and relative evenness of the incremental step, which leads to allocation of the same fraction of diagnostic species in data sets of different dimensions. The indval, phi, and cos indices satisfy this requirement. For indval index the threshold values generally range from 0.4 to 0.6 (0.41 – 0.63 for the test data set 1 and 0.36 – 0.58 for the test data set 2). For the phi index, the threshold values are slightly lower, but also similar (0.39 – 0.62 for the test data set 1 and 0.31 – 0.55 for the test data set 2). The differences of thresholds between group-equalized and non-equalized indices (indval.g vs indval, phi vs phi.g, cos.g vs cos) are not high so they do not usually require changing the thresholds when switching between group-equalized and non-equalized indices and vice versa.

Fidelity measures in context with the ability to produce non-overlapping lists of diagnostic species

Statistical allocation of diagnostic species implies choosing species whose fidelity (affinity) exceeds a certain threshold in at least one cluster. Ideally, a species should reach or exceed a threshold only in one cluster. In this case, we are talking about the "uniqueness" of the lists of diagnostic species across all clusters (synoptic columns). With decreasing of the thresholds, the number of additional peaks increases – the cases when the species fidelity reaches or exceeds the threshold in more than one cluster. However, the rate of this growth is significantly different across different indices. Because each fidelity index evaluates the relationship between a species and only one cluster, the capability to filter out species with additional lateral maximums is important.

For each of the indices the threshold has been set as indicated in Table 3, so the share of diagnostic species also varies from 10% to 40%. Table 5 provides the assessment of the non-overlap of the lists of diagnostic species across all clusters. The scores in the table below grow due to species whose fidelity values exceed the threshold in more than one cluster. Minimizing scores is preferable.

Table 5. Shares of species with fidelities exceeding a threshold in more than one synoptic column (cluster) with an increase of fractions of diagnostic species.

Shares of diagnostic species	indval.g	indval	phi	phi.g	cos	cos.g	u.hyp	g	chi	Fisher	TCR
Case study: test data set 1											
10%	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.250	0.000	0.042	0.104
15%	0.014	0.014	0.000	0.000	0.000	0.027	0.042	0.389	0.042	0.097	0.125
20%	0.031	0.031	0.010	0.021	0.031	0.042	0.031	0.458	0.031	0.156	0.120
25%	0.042	0.042	0.025	0.025	0.040	0.067	0.067	0.533	0.067	0.225	0.160
30%	0.055	0.049	0.049	0.035	0.056	0.063	0.104	0.583	0.104	0.214	0.201
35%	0.065	0.071	0.051	0.054	0.086	0.089	0.107	0.619	0.107	0.238	0.242
40%	0.073	0.070	0.057	0.047	0.093	0.093	0.115	0.635	0.115	0.260	0.255
Case study: test data set 2											
10%	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.303	0.000	0.030	0.091
15%	0.020	0.000	0.000	0.000	0.000	0.020	0.020	0.540	0.020	0.100	0.200
20%	0.015	0.030	0.030	0.015	0.030	0.015	0.045	0.642	0.045	0.104	0.224
25%	0.036	0.048	0.048	0.024	0.048	0.036	0.048	0.655	0.048	0.131	0.274
30%	0.040	0.040	0.040	0.040	0.040	0.040	0.080	0.700	0.080	0.140	0.267
35%	0.051	0.060	0.051	0.051	0.060	0.051	0.119	0.735	0.119	0.239	0.299
40%	0.112	0.052	0.090	0.090	0.052	0.112	0.187	0.784	0.172	0.299	0.261

As can be seen for both test data sets 1 and 2, at harsh thresholds with the share of diagnostic species as low as 10%, most indices (except three – g, TCR, Fisher) provide the non-overlapping lists of diagnostic species, which corresponds to zeros in Table 5. With the indval index as an example and setting loyal thresholds, which led to the share of diagnostic species of 40%, the shares of 0.070 (for the data set 1) and 0.052 (for the data set 2) of species are those with more than one maximum across synoptic columns. The good (low) scores are also possessed by the phi (0.057 for the data set 1 and 0.090 for the data set 2). Conversely, this indicator for the g-index achieves a large score (0.635 for the test data 1 and 0.784 for the test data 2). Conditionally, fidelity indices can be divided into three groups: those with high uniqueness of the lists of diagnostic species (indval, phi, cos together with their group-equalized modifications), moderate (u.hyp, chi, Fisher, TCR) and low uniqueness (g). For the indices of the first group, with the total share of diagnostic species at 40%, the share of species with additional maximums does not mostly exceed 0.1, for the second group – no more than 0.3, and for the third group greater than 0.3.

Discussion

Fidelity of species depends on homogeneity of a priori defined clusters. In this article we extracted clusters using the OptimClass methodology (Tichý et al. 2010). Alternative methods for extracting optimally homogeneous clusters involve different variants of K means clustering (Marinković et al., 2019). Additional analyses that compare fidelity of species in clusters extracted by OptimClass and K means clustering methods are still required.

All fidelity indices aim to measure the strength of the relationship between species and syntaxon. However, despite similarity of fidelity indices, not all of them are reasonable to use as a threshold criterion for allocation of diagnostic species. Indices, which have normalized values, are more intuitive in terms of choosing a threshold for them. The correction for the unequal size of clusters does not require a large change in the threshold values for group-equalized indices of phi.g, indval.g, and cos.g. For both group-equalized and non-equalized indices of indval and cos the reasonable thresholds seem to span from 0.4 to 0.6, and for the phi – just slightly less.

Evenness of distribution is also important, especially between 0.6- and 0.9-quantiles, which corresponds to the allocation of 10–40% of diagnostic species. Asymmetry of fidelity values is problematic

because it results in jump-shaped growth or reduction in the number of diagnostic species with a relatively small shift of thresholds. Because of this, the chance of overlooking the optimal value of the threshold grows. The phi, indval, cos indices satisfy most requirements while other indices are less appropriate and the g index is not recommended for using as a threshold criterion for allocation of diagnostic species. Specifically, shifting the threshold of the g index by a small value leads to an increase in the share of diagnostic species by a large value. In favour of the Fisher index is the fact that it can be applied in case of small frequencies, and the Fisher index itself is a test of significance of the null hypothesis about the relationship between species and cluster.

The use of fidelity indices for automatic sorting of species in the vegetation classification implies the formation of lists of diagnostic species that do not overlap. To avoid potential confusion that may occur around the term "differential species", it is necessary to distinguish two fundamentally different situations – when another syntaxon, where a species also occurs, is part of the same phytocoenotic table or is beyond it (i.e. not a part of data set involved in fidelity calculations). The second case is completely common in syntaxonomy, and therefore most species are not character in the strict sense, but differential with a different degree of "characticity". Nevertheless, all the fidelity indices evaluate the relationship of a species with only one cluster at a time. This is true even when they are applied to a group of clusters (De Cáceres et al., 2012), because the latter is also considered as an entire unit.

The "filtering" ability of fidelity indices to ensure the non-overlapping lists of diagnostic species is different. In this regard, the uncertainty of species-to-cluster placement increases due to the growth of probability of permutations of species optima between clusters (synoptic columns). The best marks for this indicator received by indval, phi, cos indices. In case of phi, this is also explained by the fact that the occurrence of a species outside the target cluster is counted as a negative component of correlation between species and cluster. In this sense, the term of "context-dependent" measure is used (Chytrý et al., 2002; De Cáceres & Legendre, 2009). However, there is no significant gain of phi, compared to indval or cos indices.

References

Barkman, J. (1989). Fidelity and character-species, a critical evaluation. *Vegetatio*, 85 (1–2), 105–116.

Botta-Dukát, Z. & Borhidi, A. (1999). New objective method for calculating fidelity. Example: the illyrian beechwoods. *Annali di Botanica*, 57, 73–90.

Braun-Blanquet, J. (1964). *Pflanzensoziologie. Grundzuge der Vegetationskunde*. Springer-Verlag, Wien; New York. 865 pp.

Bruelheide, H. (2000). A new measure of fidelity and its application to defining species groups. *Journal of Vegetation Science*, 11 (2), 167–178.

Chytrý, M., Tichý, L., Holt, J. & Botta-Dukát, Z. (2002). Determination of diagnostic species with statistical fidelity measures. *Journal of Vegetation Science*, 13 (1), 79–90.

De Cáceres, M., Font, X. & Oliva, F. (2008). Assessing species diagnostic value in large data sets: A comparison between phi-coefficient and Ochiai index. *Journal of Vegetation Science*, 19 (6), 779–788.

De Cáceres, M. & Jansen, F. (2020). *Indicspecies: Relationship Between Species and Groups of Sites. R Package Version: 1.7.6*. Available from: http://cran.r-project.org/web/packages/indicspecies/vignettes/indicspeciesTutorial.pdf (10.03.2020)

De Cáceres, M. & Legendre, P. (2009). Associations between species and groups of sites: indices and statistical inference. *Ecology*, 90 (12), 3566–3574.

De Cáceres, M., Legendre, P., Wiser, S.K. & Brotons, L. (2012). Using species combinations in indicator value analyses. *Methods in Ecology and Evolution*, 3 (6), 973–982. Available from: http://doi.wiley.com/10.1111/j.2041-210X.2012.00246.x (09.08.2019)

Dufrêne, M. & Legendre, P. (1997). Species assemblages and indicator species: the need for a flexible asymmetrical approach. *Ecological monographs*, 67 (3), 345–366.

Golik, G.M. & Goncharenko, I.V. (2017). Syntaxonomy, phytoindication and anthropogenic transformation of the forest vegetation of Kyiv. *Ecology and Noospherology Journal*, 28 (1–2), 49–63. Available from: https://goo.gl/EK4gJj (19.03.2020)

Goncharenko, I.V. & Golik, G.N. (2015). Classification and phytoecological evaluation of the forest-park vegetation of Kyiv. *Fitoraznoobrazie Vostochnoi Evropy*, 9 (4), 129–158. Available from: https://goo.gl/fEPJoE (19.03.2020)

Goncharenko, I.V. & Yatsenko, H.M. (2020). Phytosociological study of the forest vegetation of Kyiv urban area (Ukraine). *Hacquetia*, 19 (1), 99–126. Available from: https://ojs.zrc-sazu.si/hacquetia/article/view/7126 (17.02.2020)

Kusbach, A., Long, J.N., Van Miegroet, H. & Shultz, L.M. (2012). Fidelity and diagnostic species concepts in vegetation classification in the Rocky Mountains, northern Utah, USA. *Botany*, 90 (8), 678–693. Available from: http://www.nrcresearchpress.com/doi/10.1139/b2012-033 (17.02.2020)

Lance, G.N. & Williams, W.T. (1967). A general theory of classificatory sorting strategies: 1. Hierarchical systems. *The computer journal*, 9 (4), 373–380.

Lötter, M., Mucina, L. & Witkowski, E. (2013). The classification conundrum: species fidelity as leading criterion in search of a rigorous method to classify a complex forest data set. *Community Ecology*, 14 (1), 121–132. Available from: http://www.akademiai.com/doi/abs/10.1556/ComEc.14.2013.1.13 (22.05.2019)

Marinković, N., Karadžić, B., Pešić, V., Gligorović, B., Clemens Grosser, C., Paunović, M. Nikolić, V. Raković, M. (2019) Faunistic patterns and diversity components of leech assemblages in karst springs of Montenegro. *Knowledge and Management of Aquatic Ecosystems*, 420, 26, 1-12.

Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Szoecs, E. & Wagner, H. (2018). *Vegan: Community Ecology Package. R package version 2.4-6*. Available from: https://CRAN.R-project.org/package=vegan (21.09.2018)

Podani, J. & Csányi, B. (2010). Detecting indicator species: Some extensions of the IndVal measure. *Ecological Indicators*, 10 (6), 1119–1124. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1470160X1000052X (01.02.2020)

Roberts, D. (2016). *labdsv: Ordination and Multivariate Analysis for Ecology. R package version 1.8-0*. Available from: https://CRAN.R-project.org/package=labdsv (19.03.2020)

Senchylo, O.O. (2009). *Vegetation of the Dnieper floodplain within the forest-steppe zone*. PhD Thesis. Dissert. Cand. biol. sciences: 03.00.05 – Botany; T. Shevchenko National Kyiv Univ., Kyiv.

Senchylo, O.O., Shevchyk, V.L. & Solomakha, I.V. (1998). Vegetation of the Sobachiy Island in the Kremenchuk Reservoir. *Ukrainian phytocoenological collection, Ser. A.*, 9 (1), 21–29.

Senchylo, O.O., Shevchyk, V.L. & Solomakha, V.A. (1997). Syntaxonomy of meadow massif in the Dnieper floodplain at the upper part of the Kremenchug reservoir. *Ukrainian phytocoenological collection, Ser. A.*, 2 (7), 39–49.

Tichý, L. (2002). JUICE, software for vegetation classification. *Journal of vegetation science*, 13 (3), 451–453.

Tichy, L. & Chytry, M. (2006). Statistical determination of diagnostic species for site groups of unequal size. *Journal of Vegetation Science*, 17 (6), 809–818.

Tichý, L., Chytrý, M., Hájek, M., Talbot, S.S. & Botta-Dukát, Z. (2010). OptimClass: Using species-to-cluster fidelity to determine the optimal partition in classification of ecological communities. *Journal of Vegetation Science*, 21 (2), 287–299. Available from: http://doi.wiley.com/10.1111/j.1654-1103.2009.01143.x (09.08.2019)

Westhoff, V. & Van Der Maarel, E. (1978). The Braun-Blanquet approach. *In Classification of plant communities*, R. H. Whittaker [ed.], Dr. W. Junk, The Hague. pp. 287–399.

Willner, W., Tichý, L. & Chytrý, M. (2009). Effects of different fidelity measures and contexts on the determination of diagnostic species. *Journal of Vegetation Science*, 20 (1), 130–137. Available from: http://doi.wiley.com/10.1111/j.1654-1103.2009.05390.x (05.02.2020)