Annals of Forest Science (2021) 78: 13
https://doi.org/10.1007/s13595-020-01017-0

RESEARCH PAPER

The assignment of relevés to pre-existing vegetation units: a comparison of approaches using species fidelity

Hamed Asadi[1,2] · Omid Esmailzadeh[1] · Miquel De Cáceres[3] · Seyed Mohsen Hosseini[1]

Received: 26 November 2019 / Accepted: 12 November 2020 / Published online: 1 February 2021

Abstract

• ***Key message*** **Total fidelity value index can be used for the assignment of new relevés to existing vegetation units and it can be used to refine classifications derived from unsupervised clustering.**

• ***Context*** Diagnostic species is an important concept in vegetation classification. Apart from its usefulness to characterize species niche preferences, the diagnostic species concept is used in vegetation classification: (1) for the assignment of new relevés to the vegetation units of an existing classification; (2) to refine vegetation classifications by reassigning relevés that sustain the definition of vegetation units.

• ***Aims*** The main aims were to evaluate the relative predictive performance of different statistical fidelity measures for the reassignment of relevés to existing vegetation units, and in which cases reassignments improve the quality of the original classification.

• ***Methods*** We took the classifications produced by three commonly used unsupervised classification methods, and all relevés were reassigned to the closest vegetation unit according to the total fidelity value index (*TFVI*), where fidelity value had been calculated using one of eight distinct statistical measures, and according to the frequency-positive fidelity index (*FPFI*). Classifications obtained after relevé reassignments were compared to the initial ones using the Adjusted Rand Index. The quality of all classification solutions, including the initial ones, was evaluated using thirteen different evaluator statistics.

• ***Results*** The predictive performance of *IndVal* was the best among all eight fidelity indices in the TFVI framework, and also outperformed FPFI. The TFVI framework based on group-equalized fidelity indices produced better results than other assignment rules in terms of the chosen evaluator statistics. Re-assignments based on *IndVal*, *r*, or FPFI produced classifications with the best quality, when combining the results of all evaluators.

• ***Conclusion*** We conclude that TFVI based on *IndVal* and *r* has the best quality for assigning of new relevés to existing vegetation units, and it also could be used to refine classifications derived from unsupervised clustering. Consequently, our results reiterate that TFVI, which is new in vegetation sciences, can be a good alternative for FPFI, as the most commonly used in the assignment of vegetation plots (relevés), to predefined vegetation types in large datasets.

Keywords Classification validity · Fidelity values · Indicator value · Hyrcanian forests · Phi coefficient · Phytosociology · Similarity indices · Vegetation classification

Handling Editor: Laurent Bergès

Contribution of the co-authors HA and OE originally formulated the idea, HA, OE and MD developed methodology, HA and OE conducted fieldwork, HA, OE and MD performed statistical analyses, and HA, OE, MD and SYD wrote the manuscript.

✉ Omid Esmailzadeh
oesmailzadeh@modares.ac.ir

Extended author information available on the last page of the article

Abbreviations

FPFI	Frequency-positive fidelity index
TFVI	Total fidelity value index
TPFI	Total phi fidelity Index
TIVI	Total indicator value index
TWINSPAN	Two-way indicator species analysis
PAM	Partitioning around medoids
r	Correlation indices of fidelity
IndVal	Indicator value indices
ARI	Adjusted rand index
PARTANA	Partition analysis
PBC	Point biserial correlation

$IndVal_{ind}$	Indicator value indices for species abundance
$INdVal_{pa}$	Indicator value indices for presence-absence data
$r_{ø}$	Correlation indices based on presence-absence data
r_{ind}	Correlation indices based on abundance data; indices with superscript *g* are considered as a group-equalized and without g are considered as a non-equalized

1 Introduction

Diagnostic species is an important concept in vegetation classification (Whittaker 1962; Westhoff and van der Maarel 1973). It refers to those species that, due to their niche preferences, concentrate their occurrence or abundance in a single vegetation unit or in a few vegetation units (Dengler et al. 2008). Under the Braun-Blanquet approach to vegetation classification, the degree of concentration is usually called fidelity. There are several statistical techniques for the determination of fidelity and diagnostic species. Among them, those that assess the strength of association between species and groups of relevés (i.e., vegetation plot records) using correlation or indicator value indices are the most widely used (Chytrý et al. 2002; De Cáceres and Legendre 2009; De Cáceres et al. 2012).

Apart from its usefulness to characterize species niche preferences, the diagnostic species concept is used in vegetation classification: (1) for the assignment of new relevés to the vegetation units of an existing classification; (2) to refine vegetation classifications by reassigning relevés that sustain the definition of vegetation units (Tichý 2005; Dai et al. 2006). To conduct these tasks, phytosociologists traditionally (re)assigned relevés informally using the diagnostic species concept. However, nowadays fidelity values can be formally incorporated into the calculation of the numerical similarity between the relevé to be assigned and each of the vegetation units (van Tongeren et al. 2008). Similarity indices are often used to assign relevés to vegetation units based on complete species composition. For example, Hill (1989) described a method based on a modified Czekanowski coefficient of similarity between observed and expected numbers of species in each constancy class. Two assumptions underpin the incorporation of fidelity values into the calculation of the similarity between relevés and vegetation units: (a) some species may be more informative than others for the assignment of relevés to vegetation units; (b) since diagnostic species are indicators of their preferred habitats (De Cáceres et al. 2012), fidelity values are a reasonable way of defining species weights in the calculation of similarity to vegetation units(Chytr and Tichý. 2003). Following this rationale, Tichý (2005) developed four similarity indices using the *phi* coefficient as a fidelity measure. Among them, the *frequency-positive fidelity index* (*FPFI*) has been often used for the assignment of relevés that were misclassified or classified to more than one groups in supervised classifications (Boublík et al. 2007; Douda 2008; Boublík 2010; Janišová et al. 2010; Svitková and Šibík 2013; Landucci et al. 2013; Rodríguez-Rojo et al. 2014; Chytry and Tichy 2018; Maciejewski et al. 2020). Another approach employing the diagnostic species concept in a similarity index was proposed by Dai et al. (2006), who suggested using the *total indicator value index* (*TIVI*) to test the validity of a TWINSPAN classification and to refine the initial classification by reassigning relevés. While Dai et al. (2006) based relevé assignments on fidelity values calculated using the indicator value (*IndVal*) index (Dufrêne and Legendre 1997), Esmailzadeh and Asadi (2014) recently suggested replacing IndVal with the phi coefficient, calling the resulting approach *total phi fidelity index* (*TPFI*). Gégout and Coudun (2012) also developed the *fidelity index* of relevés (FI) to reassign relevés according to fidelity values, but in contrast to previous indices FI is defined as the sum of phi values, regardless of species abundance or frequency.

Here, we propose that *TIVI* and *TPFI* can be unified in a single assignment framework that can be called the *total fidelity value index* (*TFVI*). For each relevé and vegetation unit, *TFVI* is defined as the sum, across all species occurring in the relevé, of the species fidelity values for the vegetation unit multiplied by the species abundance values in the relevé. Each relevé is then assigned to the vegetation unit for which *TFVI* is highest. Several alternatives can be used as a choice of fidelity measure in the *TFVI* framework (Tichý and Chytrý 2006; De Cáceres and Legendre 2009), and it is unknown how this choice may affect the results. Therefore, in this paper, we compare the performance of eight different statistical fidelity measures for their use within the *TFVI* framework. Because relevé assignments are commonly based on the *FPFI*, we also include this assignment rule in our evaluation. We take vegetation data from the Hyrcanian Box tree (*Buxus hyrcana* Pojark.) and common yew (*Taxus baccata* L.) forests and the result of three frequently-used unsupervised methods as initial vegetation classification. We evaluate the nine different assignment rules: (a) in terms of the predictive performance; (b) in terms of the quality of the vegetation units after re-assignment and also in terms of several evaluator indices. Evaluating predictive performance (a) implies the assumption that the number of groups is known and the original classification is the "truth" to be reproduced by the assignment rule. In contrast, evaluating the quality of classification after reassignment (b) assumes that the original classification may be improved. We conduct this evaluation without fixing the number of groups, but we penalize those cases when the classification obtained after reassignment reduces the number of relevés sustaining the definition of some vegetation units to the point of compromising their statistical validity.

2 Methods

2.1 Study area

The Hyrcanian region is a narrow green belt covering an area of 1.9 million ha in northern Iran and 20,000 ha in the Republic of Azerbaijan. In Iran, this region is located between the Caspian Sea and the northern foothills of Alborz mountains in three northern provinces: Gilan (western part), Mazandaran (middle part), and Gorgan (eastern part). Unlike most of Iran, the Hyrcanian region is relatively humid, with an average annual rainfall that ranges between 530 mm in the east and 1350 mm in the west (with an occasional record over 2000 mm in some locations). Rainfall mostly occurs during late fall, winter, and spring. The average annual temperature varies from 15 °C in the west to 17.5 °C in the east. The average temperature of the warmest month ranges from 28 to 35 °C while that of the coldest month is between 1.5 and 4 °C (Sagheb-Talebi et al. 2014). Brown soils (i.e., calcareous, forest acidic, podzolic, and non-podzolic soils) are the most important soil type, comprising approximately 90% of the region (Habibi Kasseb 1992). Hyrcanian forests are dominated by combinations of oriental beech (*Fagus orientalis* Lipsky), Caucasian oak (*Quercus castaneifolia* C.A.Mey.), hornbeam (*Carpinus betulus* L.), and Persian ironwood (*Parrotia persica* C.A.Mey.), and depending on the site *Acer velutinum* Boiss., *Tilia rubra* DC., *Fraxinus excelsior* L., *Alnus subcordata* C.A.Mey., and *B. hyrcana* (Marvie Mohadjer 2005).

In Hyrcanian forests, plant communities of *B. hyrcana* are the remnants of broad masses that formerly occupied lowlands (along with *Q. castaneifolia*, *C. betulus*, and *P. persica*) and steep slopes of wet valleys (along with *F. orientalis*) but are now restricted to limited areas. They are characterized by a low species richness and variability in species composition and are adapted to grow on sites within a range of edaphic conditions, as long as they are exposed to the adequate air moisture in the southern parts of the Caspian Sea, Northern Iran (Asadi et al. 2011; Esmailzadeh et al. 2014; Soleymainpour and Esmailzadeh 2015). The broad ecological niche and the high sociability of *B. hyrcana* makes the composition of box tree stands in Hyrcanian forests to be highly variable across the study area. Box trees co-occur with *Zelkova carpinifolia* (Pall.) Dippel and *Celtis australis* L. as drought-tolerant tree species in the eastern lowland Hyrcanian forests to *Pterocarya fraxinifolia* (Lam.) Spach and *Populus caspica* (Bornm.) Bornm. as hygrophilous tree species in the western part of Hyrcanian lowland forests. Along elevation gradients, box trees co-occur with *Q. castaneifolia* at low elevations and *F. orientalis* in highland forests (up to 1700 m). Box tree stands are distributed in a wide range of geographical slopes: in steep, north-oriented slopes, they are accompanied by *T. baccata*, *Prunus laurocerasus* L., and *Danae racemosa* (L.) Moench as a hygrophilous indicator species of Hyrcanian highland forests. However, in slopes with a little lower air humidity, they are accompanied by *T. rubra*, *Acer cappadicicum* Gled., and *Sorbus torminalis* (L.) Crantz.

The Hyrcanian *T. baccata* communities' dataset was also included as the second dataset for reinforcing the significance of the results. *T. baccata* is the only coniferous species that is able to distribute in the main Hyrcanian forest types. *T. baccata* is often individually scattered or consisting small populations in humid sites (i.e., with high air humidity, but not humid soils) of northern steep slopes as well as hillslopes of northern valleys throughout Hyrcanian mountainous forests (Sagheb-Talebi et al. 2014). In much more humid sites of Hyrcanian forests, it forms pure and mixed dense large populations (Sagheb-Talebi et al. 2014).

2.2 Sampling

Habitats containing *B. hyrcana* forests were searched from the Cheshmeh-Bolbol protected area (west of Golestan Province) to Lire-sar in the west of Mazandaran province (Fig. 1). These habitats are located from 50 m a.s.l in Sisangan protected area to 1750 m a.s.l in Farim, the highest altitude of *B. hyrcana* in Hyrcanian forests. After this initial survey, vegetation was sampled using the Braun-Blanquet relevé method in summer 2010 to 2014. Vegetation plots (with an area 400 m^2) were conducted in stands considered representative to provide an appropriate representation of the variability of *B. hyrcana* forests (Mueller-Dombois and Ellenberg, 1974). To include any possible change in vegetation indicating variations in habitat conditions, while considering the principle of a representative stand, we defined transects which were 400-m stretches systematically set along the altitudinal gradient and relevés were conducted whenever floristically or environmental (especially in topographical features) alteration was perceived. In each vegetation plot, all vascular plant species were recorded, and their percentage cover was visually assessed using a modification of the ordinal van der Maarel (1979) cover-abundance scale (0, absent; 1, 0–1%; 2, 1–2.5%; 3, 2.5–5%; 4, 5–12.5%; 5, 12.5–25%; 6, 25–50%; 7, 50–75%; 8, 75–100%). Cover-abundance class values were replaced by the mean cover of each cover class. The resulting dataset (referred here as the *B. hyrcana* dataset) included 484 relevés and 157 species.

Also 408 relevés in habitats containing *T. baccata* in central and eastern of the Hyrcanian forests in summer 2015, 2017, and 2018 were sampled. These forests were distributed from Siah-roudbar Valleys in the east of Golestan province to Mazga in the west of Mazandaran province (Fig. 1). These habitats are located from 1000 m a.s.l in Gazoo to 2000 m a.s.l in Afrathakhteh. This dataset is referred to here as the *T. baccata* dataset.

2.3 Initial classifications

We used three unsupervised classification methods to classify the compositional structure of the both *B. hyrcana* and *T. baccata* datasets into vegetation units: (1) modified TWINSPAN (Roleček et al. 2009), (2) partitioning around medoids (PAM, Kauffman and Rousseeuw 1990), and (3) k-means (Mac Queen 1967). Modified TWINSPAN is a hierarchical divisive method that combines the classical TWINSPAN algorithm (two-way indicator species analysis; Hill 1979) with an analysis of heterogeneity of the clusters prior to each division. Unlike the original version, modified TWINSPAN does not enforce a dichotomy of classification but instead, at each step, divides only the most heterogeneous cluster of the previous hierarchical level. Thus, the application of modified TWINSPAN results in vegetation units of similar internal heterogeneity (Luther-Mosebach et al. 2012). We applied total inertia (i.e., the sum of all eigenvalues in correspondence analysis) as measure of cluster heterogeneity (Roleček et al. 2009) and pseudo-species cut levels were set to 0%, 1%, 2.5%, 5%, 12.5%, 25%, 50%, 75%, and 100%. K-means and PAM are commonly used non-hierarchical clustering methods (Legendre and Legendre, 2012; Tichý et al. 2014). Both of them require the number of clusters and the initial group members to be specified by the user. The main difference between k-means and PAM is that in the former each cluster is represented by its centroid, a multivariate mean, whereas in the latter, each cluster is represented by its medoid, the cluster member that has the minimum sum of distances to all the other members of the cluster. Both methods have an objective error function that is progressively minimized by iteratively reassigning objects (i.e., relevés) to their nearest cluster center (centroid or medoid). Iterations terminate when no further reassignments are possible. We ran k-means and PAM starting from 100 random initial configurations to avoid local minima of the error function. The floristic resemblance between pairs of relevés was assessed using the Hellinger distance, which can be emulated by transforming relevé data prior to calculation of Euclidean distances (Legendre and Gallagher 2001; De

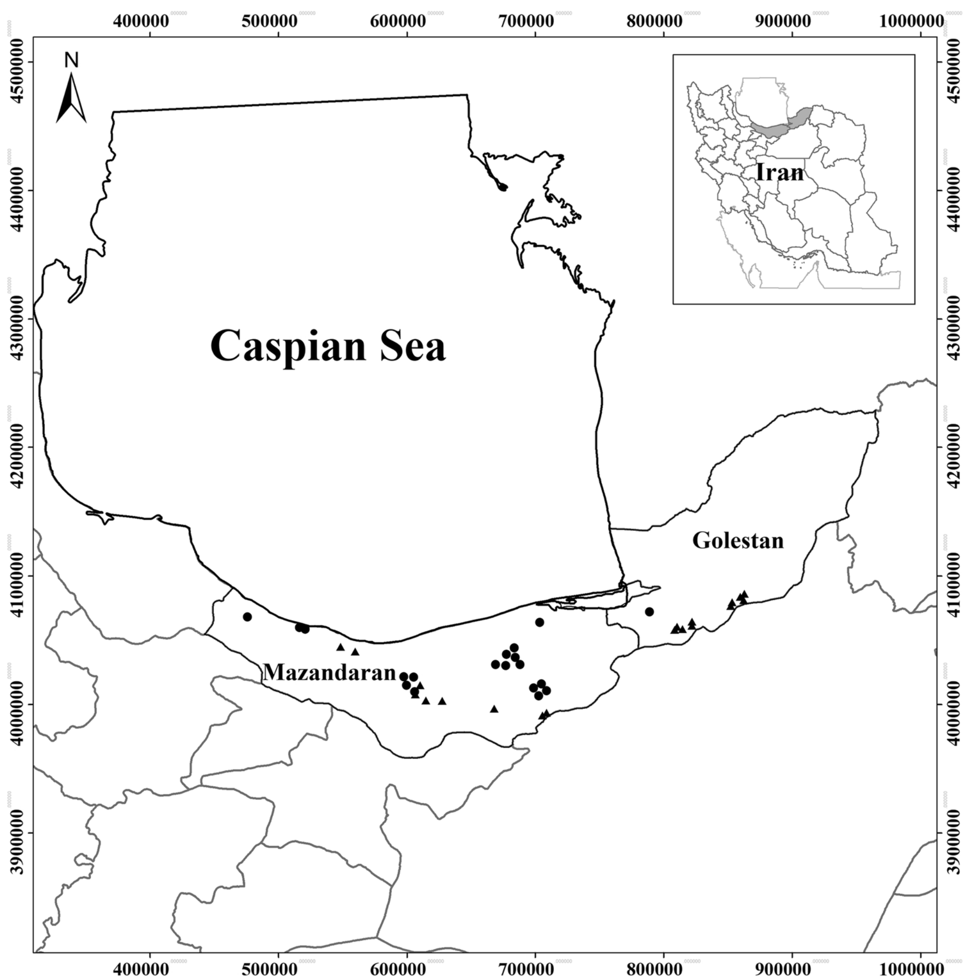


Fig. 1 Map of the two Iranian provinces where Hyrcanian forests were sampled. Dots indicate sampling locations corresponding to the *B. hyrcana* (circles) and *T. Baccata* (triangles) dataset

Cáceres et al. 2008). All classification methods were executed using the JUICE software (Tichy 2002) based on both vegetation datasets.

We used the OptimClass procedure (Tichy et al. 2010) to determine the optimum number of clusters (Appendix 1) in both *B. hyrcana* and *T. baccata* datasets. Specifically, we searched for the partition with the largest number of faithful species across all clusters. Faithful species were determined based on the p-value of the Fisher's exact test as a measure of fidelity (Tichy et al. 2010). OptimClass was run on 40 classification algorithms with five distance measures (i.e., Euclidean, relative Euclidean, correlation, chi-square, and relative Sorensen) and eight methods of group linkage (i.e., Flexible Beta, McQuitty's, Ward's, centroid, group average, median, nearest and farthest neighbor methods) based on square-root transformed cover percentage. Across all classification methods, the relationship between the number of faithful species and the number of clusters (groups) revealed that classification by 18 group numbers in *B. hyrcana* forests and 17 group numbers in *T. baccata* forests presented the most number of faithful species and it was thereafter considered as optimal group number in. For the evaluation of predictive performance, we took the partitions obtained by TWINSPAN, PAM, and k-means with 18 and 17 groups as the initial classification to be recovered in numbers in *B. hyrcana* and *T. baccata* datasets, respectively. For the evaluation of the quality of classification after reassignment, however, we kept the partitions generated by TWINSPAN, PAM, and k-means for 2, 3, …, 18 groups in *B. hyrcana* datasets and 2, 3, …, 18 groups in *T. baccata* datasets as initial classifications to be refined.

2.4 Assignment rules

The $TFVI_{ig}$ value for a given relevé i and vegetation group g is defined as:

$$TFVI_{ig} = \sum_{j=1}^{s_t} C_{ij} \times FV_{gj} \tag{1}$$

where FV_{gj} is the fidelity (indicator) value of species j in group g and C_{ij} is the cover value of species j in relevé i. Once the $TFVI_{ig}$ value is calculated for all vegetation units, the assignment rule consists of assigning relevé i to the group corresponding to the highest $TFVI_{ig}$ value.

There are many alternative indices to define the fidelity value of species (Tichý and Chytrý 2006; De Cáceres and Legendre 2009). To determine how important was the choice of a fidelity index for assignments in the TFVI framework, we compared eight different alternatives, differing in the general approach (i.e., correlation vs. indicator value indices, or IndVal, indices), in the way differences in group size are dealt (i.e., non-equalized vs. group-equalized) and in whether species abundance values are taken into account (Table 1, taken from De Cáceres and Legendre 2009). For correlation indices, we only used species with positive fidelity values in the calculation of TFVI. Calculations were

Table 1 Non-equalized and group-equalized versions of the correlation indices (r) and the indicator value indices ($IndVal$)

Indices	Non-equalized	Group-equalized
Correlation (r)	$r_\phi = \frac{N \times n_p - n \times N_p}{\sqrt{(N \times n - n^2) \times (N \times N - N_p^2)}}$	$r_\phi^g = \frac{N \times n_p^g - n^g \times N_p^g}{\sqrt{(N \times n^g - n^{g2}) \times (N \times N_p^g - N_p^{g2})}}$
	$r_{ind} = \frac{N \times a_p - a \times N_p}{\sqrt{(N \times c \times a - a^2) \times (N \times N_p - N_p^2)}}$	$r_{ind}^g = \frac{N \times a_p^g - a^g \times N_p^g}{\sqrt{(N \times c \times a^g - a^{g2}) \times (N \times N_p^g - N_p^{g2})}}$
Indicator value ($IndVal$)	$IndVal_{pa} = A_{pa} \times B_{pa} = \frac{n_p}{n} \times \frac{n_p}{N_p}$	$IndVal_{pa}^g = A_{pa}^g \times B_{pa} = \frac{n_p / N_p}{\sum_{k=1}^{k} n_k / N_k} \times \frac{n_p}{N_p}$
	$IndVal_{ind} = A_{ind} \times B_{pa} = \frac{a_p}{a} \times \frac{n_p}{N_p}$	$IndVal_{ind}^g = A_{ind}^g \times B_{pa} = \frac{a_p / N_p}{\sum_{k=1}^{k} a_k / N_k} \times \frac{n_p}{N_p}$

Notes: $r_\emptyset$ and r_{ind} are classified as correlation indices which the first one, phi coefficient, is calculated based on presence–absence data and the second one, correlation index for Individual-based, is derived from abundance data. $IndVal_{\mathrm{ind}}$ is indicator value indices for species abundance data, and $IndVal_{\mathrm{pa}}$ is indicator value indices for presence–absence data. We follow mathematical formulae of the indices used in De Cáceres and Legendre (2009)

N total number of relevés, N_{p} number of relevés belonging to the target group, n number of occurrences of the species among all relevés, n_{p} number of occurrences of the species within the target group, a_{p} sum of the abundance values of the species within the target group, a sum of the abundance values of the species over all relevés, c constant representing the total number of individuals or the total abundance per relevé. symbols are used in the Group-equalized indicator value indices, K number of groups, N_{k} number of relevés belonging to the kth group, n_{k} number of occurrences of the species in the kth group, a_{k} sum of the abundance values of the species in the kth group performed using the R statistical language and the "indicspecies" package, version 1.7.5 (De Cáceres and Legendre 2009). Because it is frequently used for relevé assignments, we also considered *FPFI* (Tichý 2005) as an additional assignment rule to be evaluated. *FPFI* is a combination of the *frequency index* (*FQI*) and *positive fidelity index* (*PFDI*) (Eq. 1). The FQI_{ig} (Eq. 2), $PFDI_{ig}$ (Eq. 3), and $FPFI_{ig}$ (Eq. 4) value for a given relevé i and vegetation group g are defined as:

$$FQI_{ig} = 100 \times \left(\sum_{j \in R} FQ_{gj} \Big/ \sum_{j \in C} FQ_{gj} \right) \tag{2}$$

$$PFDI_{ig} = 100 \times \left(\sum_{j \in R} FD_{gj} \Big/ \sum_{j \in C} FD_{gj} \right) \tag{3}$$

$$FPFIig = 100 \times (FQIig + PFDIig)/2 \tag{4}$$

where FQ_{gj} is the frequency (constancy) value of species j in group g. Species present in the relevé are indicated as $j \in R$ and species present in constancy column as $j \in C$. FD_{gj} is positive fidelity value (phi coefficient) for species j to a g vegetation unit. All assignment algorithms were done based on the results of every third initial classification methods in both *B. hyrcana* and *T. baccata* datasets separately.

2.5 Evaluation of predictive performance

Before using it for the assignment of new relevés, it is important to evaluate the predictive performance of any given assignment rule (i.e., to evaluate to which degree a "known" classification can be reproduced). We used each of the three 18-group initial classifications (produced by modified TWINSPAN, k-means, and PAM) for the reassignment of the 484 relevés according to TFVI and FPFI. Reassignment of each target relevé was done after excluding it from the calculation of fidelity values. Note that this does not make the assignment rule completely independent of the relevé to be assigned, since it still influences how the original classification was obtained. Nevertheless, this effect is very difficult to remove and we think it may be small in most cases. In total, 27 new classifications were obtained being the result of reassigning relevés of each initial classification using each of the nine assignment rules (i.e., TFVI with eight fidelity measures plus FPFI) in each datasets.

Assuming that the initial classification was the "known" classification solution, we calculated the adjusted Rand index (ARI) (Hubert and Arabie 1985) to assess the degree of agreement between the initial classification and the classification obtained after reassignment. ARI values are bounded between 0 and 1, with a value of 1 meaning perfect agreement between the initial classification and the TFVI result and a value of 0 meaning that agreement was not any better than what would be obtained by chance. ARI was calculated using the R statistical language in the "*mclust*" package, version 5.1 (Fraley et al. 2012). Ranks of ARI values were used to compare the performance of different assignment rules across the three classifications in each datasets.

2.5.1 Evaluation of the quality of vegetation classifications

For the evaluation of the quality of vegetation units, we tested the nine assignment rules on the partitions generated by the three classification methods for each number of groups between 2 and 18 in *B. hyrcana* datasets and 2 and 17 in *T. baccata* datasets. This resulted in 459 new classifications (3 classification methods × 9 assignment rules × 17 clustering levels) in *B. hyrcana* datasets and 432 new classifications (3 classification methods × 9 assignment rules × 16 clustering levels) in *T. baccata* datasets. In some cases, reassignments led to a reduction in the number of members of some vegetation units. The minimum group size for accepting a group as 'statistically valid' was conventionally set to three relevés, and assignment rules leading to a reduction in the number of valid vegetation units were penalized (see below). The 459 classifications for *B. hyrcana* and the 432 classifications for *T. baccata* were evaluated were evaluated with eight internal classification evaluators, most of which have been tested and reviewed in the literature (Aho, Roberts and Weaver 2008; Roberts 2015). The eight evaluators consist of five geometric and three non-geometric measures, and are summarized in Table 2 (see discussion about the circularity in the choice of evaluators in Sect. 4). Among non-geometric evaluators, we applied (1) *Morisita's index of niche overlap* (Horn 1966) that evaluates classification effectiveness concerning species distributions, (2) *ISAMIC-indicator species analysis to minimize intermediate constancy* (Roberts 2010) measures the constancy (either high or low) of species within groups irrespective of how many groups that species occurs in (Roberts 2015), and (3) *ISA-indicator species analysis* (Aho et al. 2008) is derived from the indicator value (*IndVal*) of species (Dufrene and Legendre 1997). The *IndVal* has long been the most popular measure to assess species importance in community classifications (Podani and Csanyi 2010) as it is one of the most widely used goodness of clustering index (Roberts 2015). High *ISA* values indicating high fidelity and abundance of species within groups (Aho et al. 2008). *ISA* is presented in two modes: average p value and number of significant indicator ($\alpha = 0.05$), but we used the last one. p Values for *ISA* was calculated with Monte-Carlo procedures.

For the geometric evaluators, we considered five indices: (4) *C-index* (Hubert and Levin 1976), (5) *PBC-point*

Table 2 Summary of classification solution evaluators in this paper

Evaluator	Optimality criteria*
Non-geometric evaluator	
Morisita's index of niche overlap (Horn 1966)	This index represents a measure of group overlap for a particular clustering solution. High proportional occurrence of species within a single group causes to niche overlap is decreased. It means minimal niche overlap indicates optimal solutions, so we used 1- *Morisita* instead of *Morisita* in the analysis.
ISAMIC-indicator species analysis to minimize intermediate constancy (Robert 2010)	This index is a measure of consistent presence or absence of species in groups and it is bounded 0-1. The higher values the better.
ISA (number of significant indicators) (Aho et al. 2008)	High ISA values indicate high fidelity and abundance of species within clusters. P-values for ISA-values calculated with Monte-Carlo procedures.
Geometric evaluators	
C-index (Hubert and Levin 1976)	This index shows the ratio of within to between group distances. This index is confined to interval 0–1 and minimum C-index scores is its optimality. So 1- (*C-index*) was considered instead of *C-index* in the analysis.
PARTANA ratio (Robert 2005)	High *PARTANA* value implies Low within group dissimilarity and high dissimilarity of relevés within groups to relevés outside of groups. So higher *PARTANA* values were considered as optimal clustering solutions.
Point biserial correlation (*PBC*) (Brogden, 1949)	The higher *PBC* values were considered as optimal clustering solutions
ASW-average silhouette width (Rousseeuw 1987)	High ASW indicates samples within clusters are compositionally similar, and dissimilar to nearest neighbor samples outside clusters. This index is confined to interval 0–1 and a high value indicates optimal solutions.
ANOSIM-analysis of similarities(Clarke 1993)	This index uses the rank of dissimilarity values in between and within groups. This index is confined to interval − 1, 0, + 1 and a high value indicates optimal solutions. Value 0 indicating completely random grouping.

*Equations and descriptions for all evaluators included in Appendix 2

biserial correlation (Brogden 1949), (6) PARTANA ratio (Roberts 2005), (7) *ASW-average silhouette width* (Rousseeuw 1987), and (8) *ANOSIM-analysis of similarity* (Clarke 1993). Geometric indices evaluate classification effectiveness based on the relationship of pairwise dissimilarities within- and between-groups. For the calculation of geometric evaluators, Euclidean distance was used to create the required distance matrices based on presence-absence and abundance data (after Hellinger's transformation) as two types of species data. We used both presence-absence and abundance data to avoid biasing our evaluation towards reassignments made with either incidence-based or abundance-based fidelity measures. Finally, we used thirteen classification evaluators. All evaluators were run in R with "plant.ecol" package, version 0.4-1 (Aho, K. 2017. https://sites.google.com/a/isu.edu/aho) for *Morisita*, *C-index*, and *PBC*; "labdsv" (Robert and Robert 2016. https://cran.r-project.org/web/packages/optpart/index.html) for ISAMIC; "optpart" (Robert 2010. https://cran.r-project.org/web/packages/optpart/index.html) for PARTANA; "cluster" (Maechler et al. 2013. https://cran.r-project.org/web/packages/cluster/index.html) for Silhouette; "indicspecies" (De Cáceres et. al 2016. https://cran.r-project.org/web/packages/indicspecies/index.html) for ISA, and "vegan" (Oksanen et al. 2013. https://cran.r-project.org/web/packages/vegan/index.html) for ANOSIM indices.

A numerical value was obtained for each evaluator and each classification solution. The reduction in the number of groups can lead to an inflation of the evaluator statistic (because the vegetation units that loose relevé members during reassignment have lower quality). Therefore, we penalized the fact that some assignment rules led to some vegetation units having less than three relevés. In these cases, we decreased the value of the evaluator multiplying the ratio between the number of "statistically valid" groups and the number of initial groups (this proportion hereafter called "correction ratio"). After possibly modifying the evaluator values, we averaged the values of each evaluator across the partitions of 2 to 18 groups in *B. hyrcana* datasets and 2 and 17 in *T. baccata* datasets. For each classification method, the ten evaluator values, corresponding to the initial classification and the nine assignment rules, were ranked from best (1) to worst (10). To combine the results of the different evaluators, we calculated median ranks (across the thirteen evaluator statistics) for each classification. Ranks of these medians were used to compare the different assignment rules globally. All synthesizing process was summarized in a flowchart (Fig. 2). All analyses were separately done based on two available datasets.

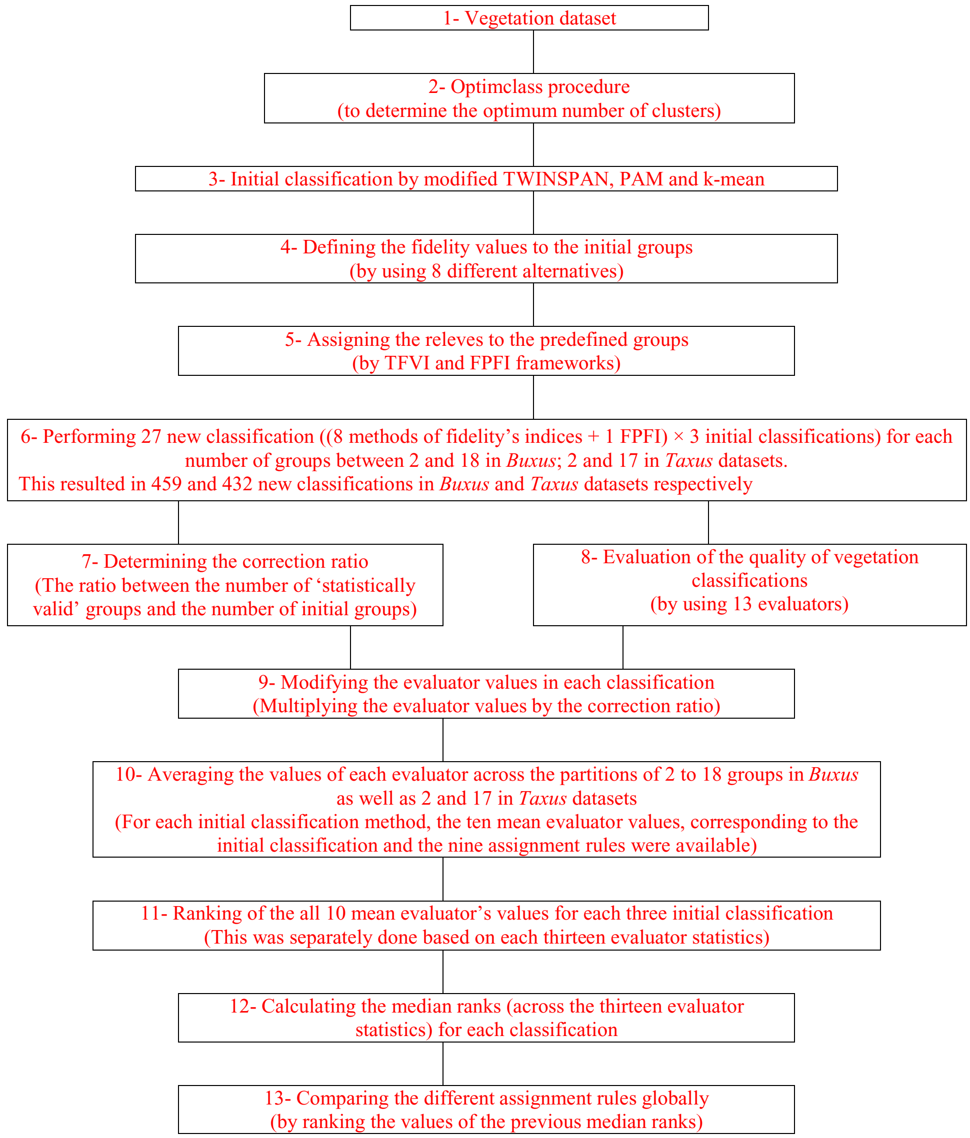


Fig. 2 Flowchart of all synthesizing process

3 Results

3.1 Evaluation of predictive performance

All assignment rules produced some distortion of the three initial classifications. The predictive performance of the TFVI framework varied strongly depending on the fidelity measure, whereas that of the *FPFI* assignment rule was rather low especially in the *T. baccata* dataset. *Indval* obtained the best rank in its performance to reproduce all three initial classifications in both *B. hyrcana* and *T. baccata* dataset and, hence, the best median rank. The order of the remaining assignment rules in terms of performance was *IndVal*, *Indval*, *r*, *r*, *r*, *r*, FPFI, and *IndVal*, respectively, in the *B. hyrcana* dataset (Table 3, section A). However, *Indval* had the best predictive performance in the *T. baccata* dataset, but with a different order in the remaining assignment order, little has been changed including: *r*, *r*, *IndVal*, *IndVal*, *Indval*, *r*, *r*, and FPFI (Table 3, section B).

3.2 Quality of the classification after reassignment

The average (from 2 to 18 groups in *B. hyrcana* and from 2 to 17 groups in *T. baccata* dataset) values of quality evaluators are shown in Table 4 (averages after correcting for the decrease in the proportion of valid groups are underlined). These are shown in two sections A and B corresponded to *B. hyrcana* and *T. baccata* dataset, respectively. Among the eight fidelity measures, only the reassignment with *Indval* did not change the number of 'statistically valid' groups for partitions generated by any of the three classification methods in the *B. hyrcana* dataset. Using other indices in

Table 3 Adjusted Rand index (ARI) values for all cluster solutions in *B. hyrcana* (section A) and *T. baccata* (section B) datasets. Ranks (1 to 9) are indicated in parentheses

Classification solutions	Modified TWINSPAN	K-means	PAM	Median of ranks
(Section A)				
FPFI	0.26 (6)	0.37 (7)	0.20 (9)	7 (7.5)
$IndVal_{ind}$	0.40 (5)	0.43 (2)	0.38 (3)	3 (2.5)
$IndVal^{g}_{ind}$	0.51 (1.5)	0.60 (1)	0.44 (1)	1 (1)
$IndVal_{pa}$	0.18 (9)	0.38 (6)	0.27 (8)	8 (9)
$IndVal^{g}_{pa}$	0.48 (3)	0.23 (9)	0.40 (2)	3 (2.5)
r_{ind}	0.25 (7.5)	0.39 (4.5)	0.34 (5.5)	5.5 (6)
r^{g}_{ind}	0.25 (7.5)	0.39 (4.5)	0.36 (4)	4.5 (5)
r_{ϕ}	0.51 (1.5)	0.36 (8)	0.32 (7)	7 (7.5)
r^{g}_{ϕ}	0.45 (4)	0.40 (3)	0.34 (5.5)	4 (4)
(Section B)				
FPFI	0.39 (9)	0.59 (6)	0.39 (9)	9 (9)
$IndVal_{ind}$	0.69 (3)	0.75 (4)	0.44 (8)	4 (4)
$IndVal^{g}_{ind}$	0.72 (1)	0.77 (1)	0.67 (1)	1 (1)
$IndVal_{pa}$	0.47 (6)	0.61 (5)	0.53 (6)	6 (5)
$IndVal^{g}_{pa}$	0.46 (7)	0.48 (9)	0.64 (3)	7 (7)
r_{ind}	0.42 (8)	0.57 (7)	0.57 (5)	7 (7)
r^{g}_{ind}	0.52 (5)	0.53 (8)	0.52 (7)	7 (7)
r_{ϕ}	0.69 (3)	0.76 (3)	0.65 (2)	3 (3)
r^{g}_{ϕ}	0.71 (2)	0.76 (2)	0.64 (3)	2 (2)

the *TFVI* framework or using the *FPFI* assignment rule in *B. hyrcana* dataset led to a decrease in the number of valid groups for at least one initial classification. But the changes in the number of statistically valid groups, which were derived by nine algorithms, were relatively low in the *T. baccata* dataset and reassignment with *Indval* as well as *r* did not change the number of groups in all three initial classification methods.

Boxplot of the ranks of the thirteen evaluators, except ISA, in ten clustering solutions for *B. hyrcana* dataset indicate that at least one of the nine reassignment methods resulted in an improvement in the value of the evaluators compared to the initial classification (Fig. 3). In this dataset and from the point of view of abundance-based evaluators except for 1-*Morisita* and *ISAMIC*, *TFVI* based on *Indval* was the best solution. The *TFVI* based on *r* and *r* assignment rules obtained the first rank for *Morisita* and *ISAMIC* indices, respectively. However, in terms of incidence-based evaluators (i.e., 1-*C.index*, *PARTANA*, and *ANOSIM*), *TFVI* based on *r* assignment was ranked first, followed by the *FPFI* (i.e., *Sil and PBC*). The initial solution was ranked first for the ISA evaluator only. In this relation, boxplot of the ranks related to all evaluators in ten clustering solution for *T. baccata* dataset indicate that at least one of the nine reassignment methods improve the value of the evaluators compared to the initial (Fig. 4). From the point of view of abundance-based evaluators in *T. baccata* dataset except for *Morisita*, *PBC* and *ISAMIC*, *TFVI* based on *Indval* was the best solution. The *TFVI* based on *r*, *IndVal* and *r* assignment rules obtained the first rank for *Morisita*, *PBC* and *ISAMIC* indices, respectively. However, in terms of incidence-based evaluators (i.e., *PARTANA* and *ANOSIM*), *TFVI* based on *r* assignment was ranked first, followed by the *IndVal* (i.e., 1-*C.index* and *PBC*). But *TFVI* based on *Indval* acquired the first rank in term of incidence-based *Sil* evaluator. Based on *T. baccata* dataset, the initial solution was not ranked first at all.

Calculating median ranks revealed that assignment rules sometimes led to classifications with the better overall quality compared to the initial classification in both datasets (Table 5). In the *B. hyrcana* dataset, in terms of incidence-based evaluators, an improvement of the quality of classification was obtained by *TFVI* in combination with *r* and *FPFI*, whereas in terms of abundance-based evaluators, only reassignments using *TFVI* based on *Indval* led to classifications of better quality than the initial classification. This process in the *T. baccata* dataset also revealed that the *TFVI* algorithm based on *r* led to the best refinement of initial classification. While in terms of abundance-based evaluators, *Indval*, *IndVal*, and *FPFI* were ranked first, while *r* gained the second importance.

In the overall comparison of assignment rules (including both incidence-based and abundance-based evaluators) in both vegetation datasets, we found that TFVI based on *Indval*, *r* indices, and *FPFI* performed best in terms of evaluation

Table 4 Average values of each evaluator of the initial classification and that obtained with each assignment rule. Averages were calculated across the different number of groups 2 to 18 in *B. hyrcana* (section A) and 2 to 17 in *T. baccata* (section B) datasets, before (not underlined) and after (underlined) correcting for the ratio of "statistically valid" Appendix 3 groups. Ranks (1 to 10) are calculated from the corrected averages (i.e., from italicized values). Incidence-based evaluators are specified in the table with braces

Section (A)	Initial		FPFI		$IndVal_{ind}$		$Indval^{g}_{ind}$		$IndVal_{pa}$		$Indval^{g}_{pa}$		r_{ind}		r^{g}_{ind}		$r_{\varnothing}$		$r^{g}_{\varnothing}$	
Modified TWINSPAN classification solutions																				
1-C-index	0.74	2	0.682	5	0.841	9	0.812	1	0.569	10	0.668	8	0.695	4	0.683	7	0.725	3	0.729	6
	0.74		0.655		0.572		0.812		0.257		0.581		0.660		0.621		0.667		0.634	
$(1\text{-C.index})_p$	0.59	4	0.662	1	0.698	8	0.598	3	0.511	10	0.534	9	0.605	6	0.557	7	0.651	2	0.661	5
	0.59		0.635		0.474		0.598		0.204		0.465		0.574		0.507		0.599		0.575	
1-Morisita	0.44	8	0.480	4	0.380	9	0.450	6	0.224	10	0.443	7	0.484	2	0.482	3	0.496	1	0.469	5
	0.44		0.480		0.380		0450		0.224		0.443		0.484		0.482		0.496		0.496	
PBC	0.29	3	0.237	8	0.502	2	0.373	1	0.143	10	0.245	9	0.280	5	0.273	7	0.292	4	0.301	6
	0.29		0.228		0.341		0.373		0.057		0.213		0.266		0.248		0.269		0.262	
$(PBC)_p$	0.17	7	0.250	1	0.346	2	0.177	6	0.073	10	0.118	9	0.206	5	0.150	8	0.244	4	0.260	3
	0.17		0.240		0.235		0.177		0.029		0.102		0.196		0.136		0.224		0.226	
PARTANA	1.38	2	1.294	4	1.504	9	1.479	1	1.112	10	1.268	8	1.303	5	1.285	7	1.353	3	1.356	6
	1.38		1.243		1.023		1.479		0.444		1.103		1.238		1.169		1.245		1.180	
$(PARTANA)_p$	1.23	3	1.330	1	1.359	9	1.236	2	1.060	10	1.135	8	1.234	5	1.163	7	1.311	4	1.325	6
	1.23		1.277		0.924		1.236		0.424		0.987		1.172		1.058		1.206		1.153	
Sil	0.02	4	− 0.076	8	0.056	2	0.050	1	− 0.122	6	0.042	3	− 0.099	10	− 0.087	9	− 0.047	5	− 0.056	7
	0.02		*− 0.073*		0.038		0.050		*− 0.049*		0.036		*− 0.094*		*− 0.079*		*− 0.043*		*− 0.049*	
$(Sil)_p$	0.01	2	0.069	1	− 0.020	3	− 0.015	4	− 0.127	9	− 0.041	6	− 0.051	8	− 0.109	10	− 0.027	5	− 0.044	7
	0.01		0.066		*− 0.014*		*− 0.015*		*− 0.050*		*− 0.035*		*− 0.047*		*− 0.099*		*− 0.024*		*− 0.037*	
ISAMIC	0.28	1	0.280	3	0.286	9	0.282	2	0.292	10	0.289	7	0.282	4	0.286	6	0.288	5	0.287	8
	0.28		0.268		0.194		0.282		0.117		0.251		0.268		0.260		0.265		0.250	
ISA(*n*)	124	1	118	3	111	9	123	2	104	10	115	8	119	4	110	7	120	5	126	6
	124		113		75		123		42		100		113		100		110		110	
ANOSIM	0.55	2	0.438	7	0.715	4	0.674	1	0.162	10	0.388	9	0.459	6	0.433	8	0.529	3	0.535	5
	0.55		0.420		0.486		0.674		0.065		0.338		0.436		0.394		0.487		0.466	
(ANOSIM)*p*	0.32	5.5	0.467	1	0.475	5.5	0.339	4	0.093	10	0.199	9	0.318	7	0.222	8	0.432	2	0.449	3
	0.32		0.448		0.323		0.339		0.037		0.173		0.302		0.202		0.397		0.390	
K-means classification solutions																				
1-C-index	0.80	2	0.732	6	0.765	8	0.831	1	0.673	10	0.6787	9	0.754	3	0.753	4	0.715	7	0.745	5
	0.80		0.732		0.696		0.831		0.397		0.576		0.754		0.753		0.715		0.745	
(1-C-index)*p*	0.74	3	0.777	2	0.717	8	0.730	5	0.635	10	0.667	9	0.718	6	0.707	7	0.744	4	0.781	1
	0.74		0.777		0.652		0.730		0.375		0.567		0.718		0.707		0.744		0.781	
1-Morisita	0.5	4.5	0.53	3	0.49	9	0.5	4.5	0.59	10	0.54	8	0.49	6.5	0.49	6.5	0.55	1.5	0.55	1.5
	0.50		0.530		0.445		0.500		0.348		0.459		0.490		0.490		0.550		0.550	
PBC	0.32	2	0.303	4	0.352	3	0.360	1	0.269	10	0.220	9	0.299	7	0.302	5	0.262	8	0.299	6
	0.32		0.303		0.320		0.360		0.159		0.187		0.299		0.302		0.262		0.299	

Table 4 (continued)

Section (A)	Initial		FPFI		$IndVal_{ind}$		$Indval^{g}_{ind}$		$IndVal_{pa}$		$Indval^{g}_{pa}$		r_{ind}		r^{g}_{ind}		$r_{\varnothing}$		$r^{g}_{\varnothing}$	
(PBC)*p*	0.30	5	0.393	1	0.333	4	0.295	7	0.257	10	0.242	9	0.296	6	0.288	8	0.327	3	0.375	2
	0.30		0.393		0.303		0.295		0.151		0.206		0.296		0.288		0.327		0.375	
PARTANA	1.48	2	1.361	6	1.405	8	1.516	1	1.268	10	1.294	9	1.402	3	1.399	4	1.345	7	1.384	5
	1.48		1.361		1.278		1.516		0.748		1.100		1.402		1.399		1.345		1.384	
(PARTANA)*p*	1.48	3	1.503	2	1.405	8	1.446	5	1.271	10	1.345	9	1.422	6	1.405	7	1.461	4	1.516	1
	1.48		1.503		1.279		*1.446*		0.749		1.143		1.422		1.405		1.461		1.516	
Sil	0.06	2	− 0.026	6	0.029	3	0.095	1	− 0.159	10	− 0.081	8	− 0.015	4	− 0.015	5	− 0.07	9	− 0.049	7
	0.06		*− 0.026*		0.027		*0.095*		*− 0.094*		*− 0.069*		*− 0.015*		*− 0.015*		*− 0.07*		*− 0.049*	
(Sil)*p*	0.05	2	0.063	1	0.001	8	0.019	4	− 0.037	10	− 0.020	9	0.004	7	0.005	6	0.007	5	0.019	3
	0.05		0.063		0.001		*0.019*		*− 0.022*		*− 0.017*		0.004		0.005		0.007		0.019	
ISAMIC	0.28	4	0.279	5	0.280	8	0.282	2.5	0.288	10	0.283	9	0.276	7	0.277	6	0.282	2.5	0.283	1
	0.28		0.279		0.255		*0.282*		0.170		0.241		0.276		0.277		0.282		0.283	
ISA(*n*)	134	1	130	3	124	8	132	2	119	10	129	9	125	4	124	5.5	123	7	124	5.5
	134		130		113		*132*		70		110		125		124		123		124	
ANOSIM	0.70	2	0.538	6	0.587	7	0.737	1	0.396	10	0.434	9	0.603	3	0.601	4	0.518	8	0.578	5
	0.70		0.538		0.534		*0.737*		0.234		0.368		0.603		0.601		0.518		0.578	
(ANOSIM)*p*	0.64	3	0.675	2	0.563	8	0.601	5	0.349	10	0.479	9	0.560	6	0.536	7	0.618	4	0.678	1
	0.64		0.675		0.512		*0.601*		0.206		0.407		0.560		0.536		0.618		0.678	
PAM classification solutions																				
1-C-index	0.78	2	0.618	9	0.781	7	0.795	1	0.628	10	0.677	8	0.760	4	0.761	3	0.741	6	0.753	5
	0.78		0.618		0.718		0.795		0.433		0.622		0.760		0.761		0.741		0.753	
(1-C-index)*p*	0.73	3	0.640	7	0.683	8	0.721	4	0.554	10	0.652	9	0.697	5	0.672	6	0.793	2	0.801	1
	0.73		0.640		0.629		0.721		0.382		0.600		0.697		0.672		0.793		0.801	
1-Morisita	0.52	4.5	0.510	6	0.520	8	0.480	7	0.580	10	0.510	9	0.520	4.5	0.530	3	0.600	1	0.580	2
	0.52		0.510		0.478		0.480		0.400		0.469		0.520		0.530		0.600		0.580	
PBC	0.30	7	0.174	9	0.364	3	0.335	2	0.204	10	0.217	8	0.304	6	0.312	5	0.322	4	0.340	1
	0.30		0.174		0.335		0.335		0.141		0.199		0.304		0.312		0.322		0.340	
(PBC)*p*	0.29	4	0.236	8	0.288	6	0.294	3	0.146	10	0.225	9	0.272	5	0.254	7	0.421	2	0.436	1
	0.29		0.236		0.265		0.294		0.101		0.207		0.272		0.254		0.421		0.436	
PARTANA	1.45	2	1.203	8	1.428	7	1.461	1	1.210	10	1.293	9	1.410	3	1.408	4	1.372	6	1.387	5
	1.45		1.203		1.314		1.461		0.835		1.189		1.410		1.408		1.372		1.387	
(PARTANA)*p*	1.46	3	1.295	7	1.357	8	1.428	4	1.158	10	1.323	9	1.390	5	1.350	6	1.523	2	1.534	1
	1.46		1.295		1.248		1.428		0.799		1.217		1.390		1.350		1.523		1.534	
Sil	0.02	2	− 0.099	10	0.020	3	0.033	1	− 0.084	6	− 0.081	8	− 0.034	4	− 0.036	5	− 0.081	9	− 0.069	7
	0.02		*− 0.099*		0.018		0.033		*− 0.058*		*− 0.074*		*− 0.034*		− 0.036		− 0.081		*− 0.069*	

Table 4 (continued)

Section (A)	Initial		FPFI		$IndVal_{ind}$		$Indval^{g}_{ind}$		$IndVal_{pa}$		$Indval^{g}_{pa}$		r_{ind}		r^{g}_{ind}		$r_{\varnothing}$		$r^{g}_{\varnothing}$	
(Sil)*p*	0.06	2	− 0.0023	3	− 0.017	4	0.107	1	− 0.049	10	− 0.025	9	− 0.0219	7	− 0.0225	8	− 0.019	5	− 0.021	6
	0.06		− 0.0023		*− 0.016*		0.107		*− 0.033*		*− 0.023*		− 0.0219		− 0.0225		− 0.019		*− 0.021*	
ISAMIC	0.28	3	0.280	5	0.284	8	0.279	7	0.289	10	0.283	9	0.280	6	0.282	4	0.285	2	0.286	1
	0.28		0.280		0.261		0.279		0.199		0.261		0.280		0.282		0.285		0.286	
ISA(*n*)	133	1	117	8	121	9	132	2	118	10	130	7	121	5.5	122	4	121	5.5	123	3
	133		117		111		132		81		120		121		122		121		123	
ANOSIM	0.65	2	0.305	9	0.626	6	0.671	1	0.311	10	0.430	8	0.618	3	0.613	4	0.565	7	0.591	5
	0.65		0.305		0.576		0.671		0.214		0.395		0.618		0.613		0.565		0.591	
(ANOSIM)*p*	0.61	3	0.427	8	0.497	7	0.588	4	0.218	10	0.463	9	0.533	5	0.480	6	0.695	2	0.712	1
Section (B)	**Initial**		**FPFI**		$IndVal_{ind}$		$IndVal_{ind}{}^{g}$		$IndVal_{pa}$		$IndVal_{pa}{}^{g}$		r_{ind}		$r_{ind}{}^{g}$		$r_{\varnothing}$		$r_{\varnothing}{}^{g}$	
Modified TWINSPAN classification solutions																				
1-C-index	0.68	4	0.717	1	0.625	8	0.702	3	0.516	9	0.508	10	0.681	6	0.643	7	0.703	2	0.679	5
	0.68		0.717		0.589		0.702		0.516		0.478		0.641		0.605		0.703		0.679	
(1-C-index)*p*	0.57	8	0.596	6	0.645	4	0.583	7	0.516	9	0.510	10	0.739	1	0.683	2	0.628	3	0.600	5
	0.57		0.596		0.607		0.583		0.516		0.480		0.695		0.643		0.628		0.600	
1-Morisita	0.35	9	0.365	8	0.435	3	0.346	10	0.404	4	0.402	7	0.450	2	0.451	1	0.392	6	0.401	5
	0.35		0.365		0.409		0.346		0.404		0.378		0.424		0.424		0.392		0.401	
PBC	0.24	6	0.300	1	0.222	8	0.260	4	0.062	9	0.057	10	0.298	2	0.246	7	0.261	3	0.251	5
	0.24		0.300		0.209		0.260		0.062		0.053		0.281		0.231		0.261		0.251	
(PBC)*p*	0.21	8	0.237	6	0.284	4	0.221	7	0.086	9	0.079	10	0.404	1	0.344	2	0.281	3	0.256	5
	0.21		0.237		0.267		0.221		0.086		0.074		0.380		0.323		0.281		0.256	
PARTANA	1.61	5	1.889	1	1.517	8	1.793	2	1.087	9	1.077	10	1.693	6	1.555	7	1.697	4	1.702	3
	1.61		1.889		1.428		1.793		1.087		1.014		1.593		1.463		1.697		1.702	
(PARTANA)*p*	1.62	4	1.619	5	1.541	8	1.648	3	1.248	9	1.253	10	1.631	7	1.634	6	1.688	1	1.649	2
	1.62		1.619		1.450		1.648		1.248		1.179		1.535		1.538		1.688		1.649	
Sil	− 0.082	5	0.016	1	− 0.198	10	0.016	2	− 0.107	7	− 0.102	6	− 0.180	9	− 0.148	8	− 0.060	4	− 0.007	3
	− 0.082		0.016		*− 0.187*		0.016		*− 0.107*		*− 0.096*		*− 0.169*		*− 0.139*		*− 0.060*		*− 0.007*	
(Sil)*p*	0.03	4	0.035	5	0.010	7	0.052	2	0.000	9	0.008	8	− 0.012	10	0.018	6	0.039	3	0.057	1
	0.03		0.035		0.010		0.052		0.000		0.008		− 0.012		0.017		0.039		0.057	
ISAMIC	0.22	4	0.225	5	0.221	10	0.223	6	0.226	3	0.226	9	0.227	7	0.226	8	0.228	2	0.228	1
	0.22		0.225		0.208		0.223		0.226		0.213		0.214		0.213		0.228		0.228	
ISA(*n*)	68	5	78	4	83	3	91	1	52	7	53	8	37	9	32	10	60	6	84	2
	68		78		78		91		52		50		35		30		60		84	
ANOSIM	0.53	5	0.725	1	0.433	8	0.682	2	0.080	9	0.074	10	0.523	6	0.459	7	0.591	4	0.621	3
	0.53		0.725		0.408		0.682		0.080		0.069		0.493		0.432		0.591		0.621	

Table 4 (continued)

Section (B)	Initial		FPFI		$IndVal_{ind}$		$IndVal_{ind}^{g}$		$IndVal_{pa}$		$IndVal_{pa}^{g}$		r_{ind}		r_{ind}^{g}		$r_{\emptyset}$		$r_{\emptyset}^{g}$	
(ANOSIM)*p*	0.68	4	0.675	5	0.551	8	0.697	2	0.137	9	0.139	10	0.641	9	0.642	6	0.739	1	0.687	3
	0.68		0.675		0.518		0.697		0.137		0.131		0.603		0.604		0.739		0.687	
K-means classification solutions																				
1-C-index	0.70	6	0.729	5	0.793	3	0.735	4	0.780	1	0.646	10	0.630	8	0.613	9	0.695	7	0.748	2
	0.70		0.729		0.746		0.735		0.780		0.532		0.630		0.613		0.695		0.748	
(1-C-index)*p*	0.60	8	0.624	7	0.685	4	0.639	5	0.730	1	0.701	10	0.632	6	0.659	3	0.599	9	0.674	2
	0.60		0.624		0.645		0.639		0.730		0.578		0.632		0.659		0.599		0.674	
1-Morisita	0.44	5	0.383	8	0.462	6	0.383	9	0.447	4	0.452	10	0.458	2	0.463	1	0.448	3	0.410	7
	0.44		0.383		0.435		0.383		0.447		0.372		0.458		0.463		0.448		0.410	
PBC	0.27	6	0.293	5	0.484	2	0.315	3	0.496	1	0.244	10	0.218	8	0.212	9	0.269	7	0.302	4
	0.27		0.293		0.455		0.315		0.496		0.201		0.218		0.212		0.269		0.302	
(PBC)*p*	0.25	9	0.256	8	0.350	2	0.277	6	0.396	1	0.345	5	0.274	7	0.306	4	0.243	10	0.312	3
	0.25		0.256		0.329		0.277		0.396		0.284		0.274		0.306		0.243		0.312	
PARTANA	1.77	6	1.901	5	2.185	2	1.975	3	2.100	1	1.617	10	1.561	8	1.465	9	1.751	7	1.956	4
	1.77		1.901		2.057		1.975		2.100		1.332		1.561		1.465		1.751		1.956	
(PARTANA)*p*	1.66	2	1.642	5	1.563	9	1.649	4	1.560	8	1.630	10	1.586	6	1.579	7	1.659	3	1.795	1
	1.66		1.642		1.471		1.649		1.560		1.342		1.586		1.579		1.659		1.795	
Sil	− 0.075	7	0.020	3	0.069	1	0.019	4	− 0.013	5	− 0.145	8	− 0.162	9	− 0.189	10	− 0.056	6	0.030	2
	− 0.075		0.020		0.065		0.019		− 0.013		− 0.120		− 0.162		− 0.189		− 0.056		0.030	
(Sil)*p*	0.04	2	0.038	4	0.024	7	0.037	5	0.021	8	0.044	6	0.003	9	− 0.002	10	0.042	3	0.079	1
	0.04		0.038		0.022		0.037		0.021		0.036		0.003		*− 0.002*		0.042		0.079	
PAM classification solutions																				
ISAMIC	0.22	1	0.224	7	0.227	9	0.223	8	0.224	6	0.224	10	0.226	4	0.228	3	0.228	2	0.225	5
	0.22		0.224		0.213		0.223		0.224		0.185		0.226		0.228		0.228		0.225	
ISA(*n*)	66	4	74	3	49	6	80	2	37	9	55	7	39	8	36	10	55	5	88	1
	66		74		46		80		37		45		39		36		55		88	
ANOSIM	0.62	6	0.727	4	0.800	3	0.769	1	0.724	5	0.530	9	0.506	8	0.434	10	0.612	7	0.761	2
	0.62		0.727		0.753		0.769		0.724		0.437		0.506		0.434		0.612		0.761	
(ANOSIM)*p*	0.70	2	0.687	5	0.625	9	0.695	4	0.608	8	0.674	10	0.636	6	0.629	7	0.697	3	0.809	1
	0.70		0.687		0.588		0.695		0.608		0.555		0.636		0.629		0.697		0.809	
1-C-index	0.70	4	0.727	3	0.811	2	0.800	1	0.657	7	0.719	8	0.636	9	0.620	10	0.675	5	0.672	6
	0.70		0.727		0.764		0.800		0.657		0.634		0.599		0.547		0.675		0.672	
(1-C-index)*p*	0.61	8	0.613	7	0.681	6	0.714	1	0.689	2	0.638	10	0.696	5	0.693	9	0.664	3	0.659	4
	0.61		0.613		0.641		0.714		0.689		0.563		0.655		0.612		0.664		0.659	
1-Morisita	0.38	7	0.359	9	0.386	8	0.416	4	0.443	1	0.372	10	0.426	6	0.464	5	0.442	2	0.433	3
	0.38		0.359		0.364		0.416		0.443		0.328		0.401		0.409		0.442		0.433	

Table 4 (continued)

Section (B)	Initial		FPFI		$IndVal_{ind}$		$IndVal_{ind}^{g}$		$IndVal_{pa}$		$IndVal_{pa}^{g}$		r_{ind}		r_{ind}^{g}		$r_{\emptyset}$		$r_{\emptyset}^{g}$	
PBC	0.26	5	0.308	3	0.545	2	0.549	1	0.269	4	0.296	6	0.222	9	0.208	10	0.246	7	0.241	8
	0.26		0.308		0.512		0.549		0.269		0.261		0.209		0.183		0.246		0.241	
(PBC)*p*	0.25	8	0.250	9	0.342	4	0.377	1	0.341	2	0.273	10	0.350	3	0.349	7	0.316	5	0.312	6
	0.25		0.250		0.322		0.377		0.341		0.240		0.330		0.308		0.316		0.312	
PARTANA	1.83	4	1.943	3	2.226	2	2.212	1	1.633	8	1.903	5	1.563	9	1.494	10	1.645	6	1.634	7
	1.83		1.943		2.095		2.212		1.633		1.679		1.471		1.318		1.645		1.634	
(PARTANA)*p*	1.66	3	1.576	4	1.480	10	1.480	7	1.528	6	1.618	8	1.645	5	1.604	9	1.677	2	1.681	1
	1.66		1.576		1.393		1.480		1.528		1.428		1.548		1.416		1.677		1.681	
Sil	0.003	4	0.028	2	0.019	3	0.060	1	– 0.161	10	– 0.025	5	– 0.107	6	– 0.153	9	– 0.129	8	– 0.126	7
	0.003		0.028		0.018		0.060		*– 0.161*		*– 0.022*		*– 0.101*		*– 0.135*		*– 0.129*		*– 0.126*	
(Sil)*p*	0.05	1	0.028	2	0.004	9	0.002	10	0.007	8	0.031	3	0.021	5	0.010	7	0.010	6	0.024	4
	0.05		0.028		0.004		0.002		0.007		0.027		0.020		0.008		0.010		0.024	
ISAMIC	0.22	3	0.223	4	0.223	8	0.221	6	0.222	5	0.222	10	0.224	7	0.227	9	0.230	1	0.228	2
	0.22		0.223		0.210		0.221		0.222		0.196		0.211		0.200		0.230		0.228	
ISA(*n*)	77	1	76	2	62	6	46	9	70	3	78	4	62	6	40	10	48	8	68	5
	77		76		58		46		70		69		58		35		48		68	
ANOSIM	0.68	4	0.739	1	0.771	3	0.734	2	0.509	8	0.721	5	0.525	9	0.462	10	0.566	6	0.562	7
	0.68		0.739		0.725		0.734		0.509		0.636		0.494		0.407		0.566		0.562	
(ANOSIM)*p*	0.70	3	0.633	3	0.537	10	0.535	9	0.581	7	0.674	6	0.685	4	0.649	8	0.712	2	0.717	1
	0.70		0.633		0.505		0.535		0.581		0.595		0.645		0.572		0.712		0.717	

Notes: Because minimum C-index and Morisita scores indicate optimal classifications, they presented in minus 1 type for having same trend with other indices. Then, high scores of evaluators indicate optimal classifications

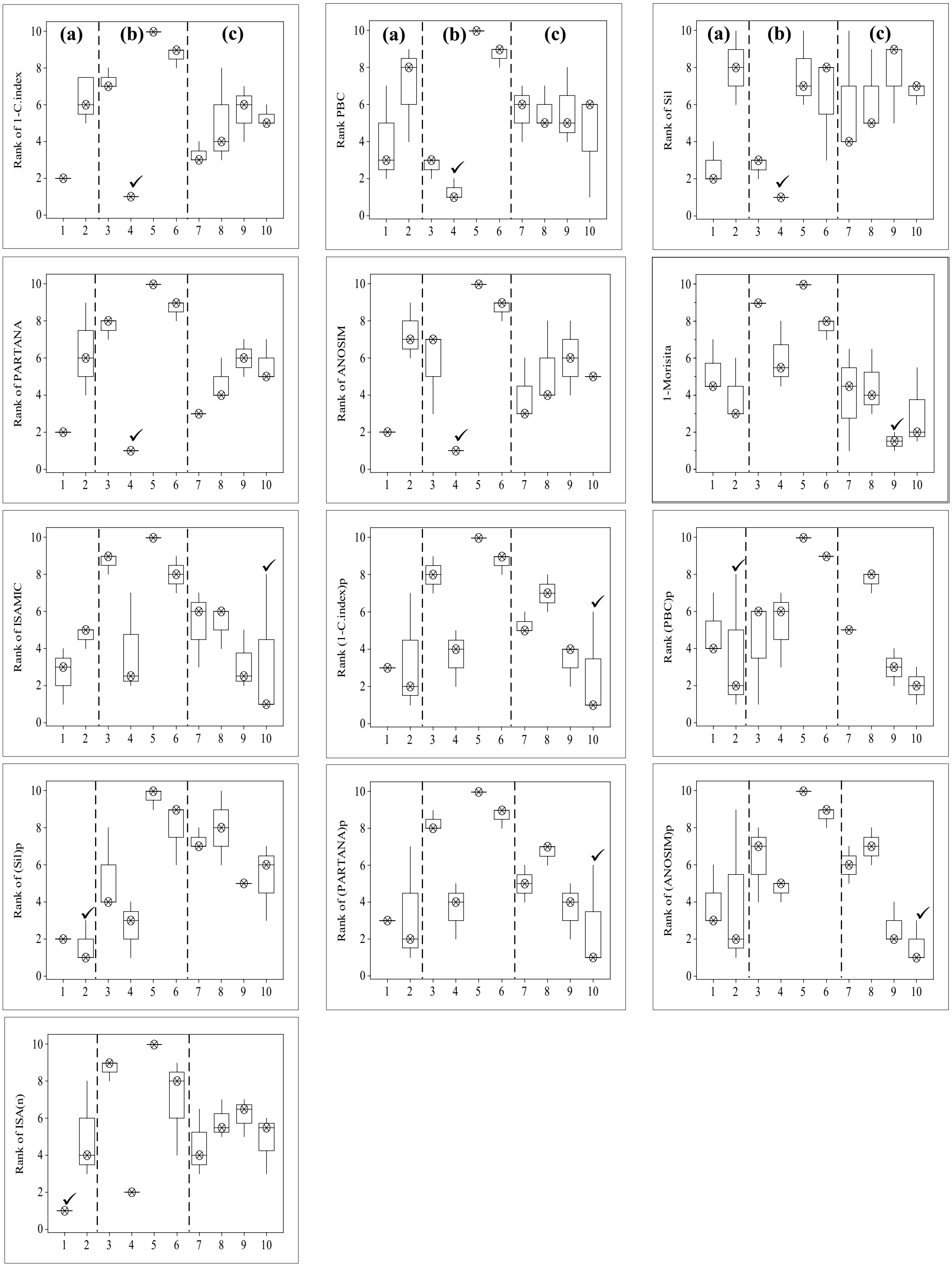


Fig. 3 Boxplot of the evaluator ranks across ten cluster solutions in the *B. hyrcana* dataset. Boxplots were drawn based on 51 values per box (3 initial classifications × 17 clustering levels): 1 initial algorithm, 2 FPFI and 3, 4, 5, 6, 7, 8, 9, and 10 are *IndVal*, 4 *Indval*, 5 *IndVal*, 6 *Indval*, 7 *r*, 8 *r*, 9 *r* based, 10 *r* based TFVI models. These are partitioned into three definite parts including (**a**) initial with *FPFI* solutions, (**b**) *IndVal*-based solutions, and (**c**) *phi*-based solutions. The box extends from the first quartile to the third quartile. The crossed line in the center of the box is the median. Each boxplot is based on three ranks, corresponding to the three initial classifications. The best solution (lowest average rank) is indicated using a checkmark symbol

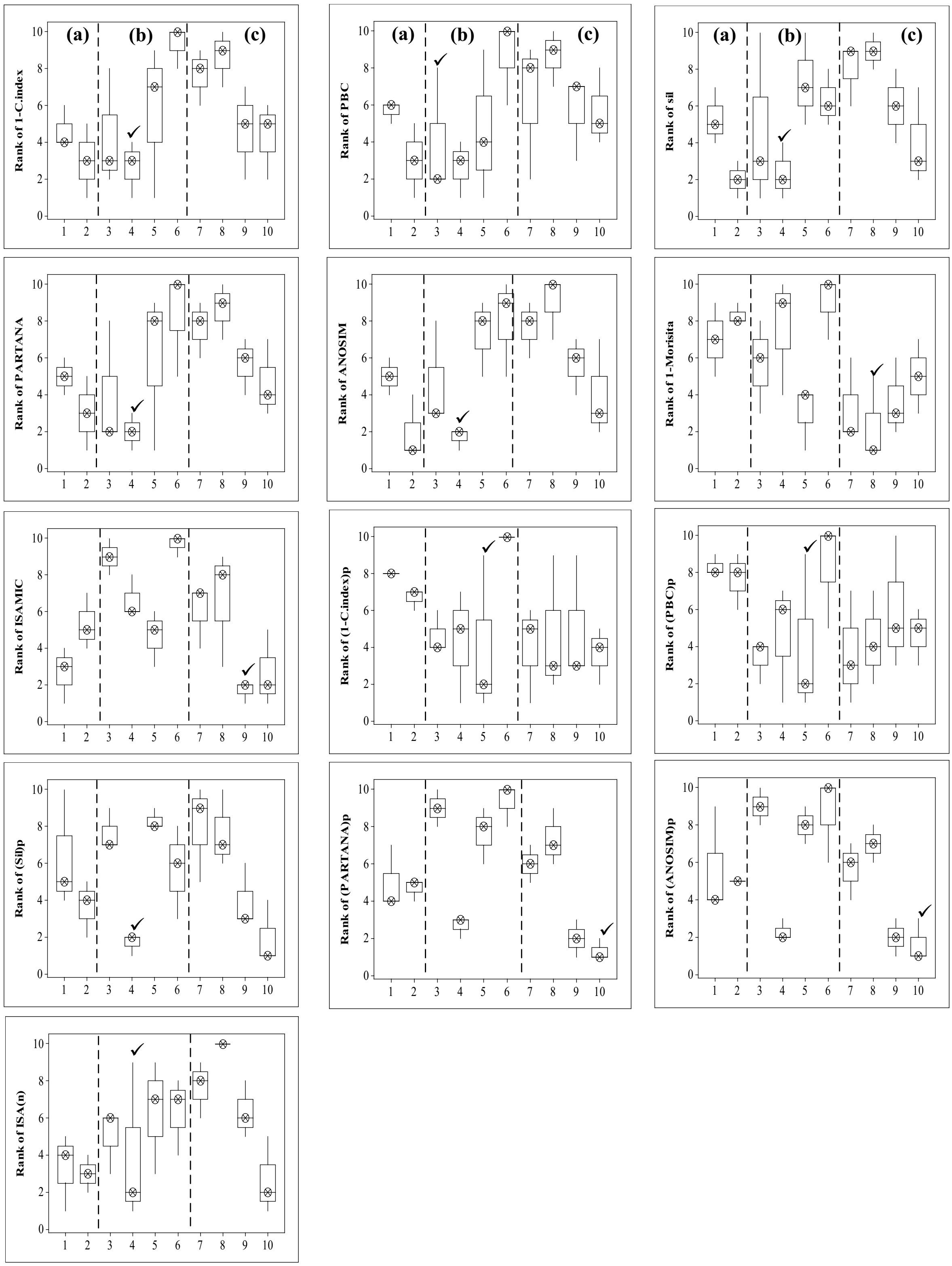


Fig. 4 Boxplot of the evaluator ranks across ten cluster solutions in the *T. baccata* dataset: 1 initial algorithm, 2 FPFI and 3, 4, 5, 6, 7, 8, 9, and 10 are *IndVal*, 4 *Indval*, 5 *IndVal*, 6 *Indval*, 7 *r*, 8 *r*, 9 *r* based, 10 *r* based TFVI models. These are partitioned in three definite parts including (**a**) initial with *FPFI* solutions, (**b**) *IndVal*-based solutions, and (**c**) *phi*-based solutions. The box extends from the first quartile to the third quartile. The crossed line in the center of the box is the median. Each boxplot is based on three ranks, corresponding to the three initial classifications. The best solution (lowest average rank) is indicated using a checkmark symbol

Table 5 Median ranks for incidence-based, for abundance-based evaluators and all evaluators in *B. hyrcana* (section A) and *T. baccata* (section B) dataset. Ranks of median ranks are shown in parentheses

Classification solutions	Incidence-based evaluators		Abundance-based evaluators		All evaluators	
	Median	Individual rank	Median	Individual rank	Median	Global rank
Section A						
Initial	3	(3)	2	(2)	3	(3)
FPFI	2	(2)	5.5	(6)	3	(3)
$IndVal_{ind}$	7	(7.5)	8	(8.5)	8	(8)
$IndVal^{g}_{ind}$	4	(4.5)	1	(1)	2	(1)
$IndVal_{pa}$	10	(10)	10	(10)	10	(10)
$IndVal^{g}_{pa}$	9	(9)	8	(8.5)	9	(9)
r_{ind}	6	(6)	4.25	(3)	5	(6)
r^{g}_{ind}	7	(7.5)	5	(4.5)	6	(7)
r_{ϕ}	4	(4.5)	5.75	(7)	4	(5)
r^{g}_{ϕ}	1	(1)	5	(4.5)	3	(3)
Section B						
Initial	4	(4)	5	(5)	4.75	(5)
FPFI	5	(5)	2	(2)	3.5	(3)
$IndVal_{ind}$	7	(7.5)	2	(2)	4.75	(5)
$IndVal^{g}_{ind}$	2	(2.5)	2	(2)	2.25	(1)
$IndVal_{pa}$	9	(9)	6	(6.5)	7.75	(8)
$IndVal^{g}_{pa}$	10	(10)	10	(10)	10	(10)
r_{ind}	6	(6)	8	(8)	7	(7)
r^{g}_{ind}	7	(7.5)	9	(9)	8.25	(9)
r_{ϕ}	2	(2.5)	6	(6.5)	4. 5	(5)
r^{g}_{ϕ}	1	(1)	3	(4)	2. 5	(2)

statistics, followed by the initial classifications and finally the other TFVI algorithm (Table 5). From the point of view incidence-based evaluators, *TFVI* based on *r* had the highest quality while the *Indval* as abundance-based indicator value has the best rank based on abundance-based evaluator. Results also indicated that FPFI was a reliable assignment index, achieving the second rank after *TFVI* based on *r* when considering incidence-based evaluators and similar rank as *TFVI* based on *r* when considering all evaluators in the *B. hyrcana* dataset. FPFI also obtained the third rank after *Indval* and *r* by considering all evaluators in *T. baccata* dataset.

4 Discussion

Finding efficient, simple and precise rules for the assignment of new or misclassified relevés to existing vegetation units is an important topic in vegetation science (Bruelheide 1997; Černá and Chytrý 2005; Tichý 2005; Dai et al. 2006; van Tongeren et al. 2008). In parallel, vegetation scientists often recommend refining the results produced using unsupervised classification methods before accepting vegetation units (Wiser and De Cáceres 2013; Tichý et al. 2014). These two tasks can be conducted employing species fidelity data.

Tichý (2005) compared several similarity indices for the assignment of relevés to the vegetation units using simulated data. Among them, Tichý (2005) recommended *FPFI* (which combines frequency information with fidelity values) for the assignment of relevés to preexisting vegetation units. Following a similar approach, Dai et al. (2006) and Esmailzadeh and Asadi (2014) developed *TFVI* and *TPFI*, respectively, based on fidelity and the cover percentage of each species. Esmailzadeh and Asadi (2014) concluded that *TPFI* using a group-equalized phi fidelity index could be used as an approach to improve TWINSPAN results. In this paper we generalized *TFVI* and *TPFI* into a single framework, which we called *TFVI*, for the assignment of relevés to existing vegetation units based on fidelity values and the cover percentage of each species. We sought to determine the most suitable fidelity measure to be used in the *TFVI* framework. For this purpose, we took the vegetation units derived from a *B. hyrcana* as well as *T. baccata* datasets and tested the performance of the *FPFI* assignment rule and the *TFVI* framework using eight different fidelity measures. Despite only testing assignment rules on a single (but real) dataset, we obtained some interesting findings, which we describe in the following paragraphs.

We found that assignments using the *TFVI* framework often had higher predictive performance than assignments using *FPFI*. The reason for this result may be related to the usage of cover percentage (i.e., species abundance data) as a weighting criterion instead of species frequency (i.e., presence/absence data). Assuming that a high percentage cover of a species implies more favorable environmental conditions than its frequency, the weighting of fidelity values by cover percentage causes the influence of each species to be related to the availability of favorable conditions. In the TFVI framework, species with high cover percentage as well as high fidelity for the target unit will be more influential in assignments.

We also found that using group-equalized fidelity indices led to better results in the *TFVI* framework, both in terms of predictive performance and quality of the resulting classification, compared to the use of non-equalized fidelity indices. Diagnostic values analysis using non-equalized indices is biased towards common species (e.g., Chytrý et al. 2002; Tichý and Chytrý 2006). Group equalization allows assessing diagnostic value independently of the size of the data set and of the size of the target site group, resulting in a better treatment of species rarity in fidelity calculations (Tichý and Chytrý 2006). In the case of the phi coefficient, another advantage of group equalization is that for each species, the order of its relative frequencies within different vegetation units is the same as the order of its fidelities to those vegetation units (Tichý and Chytrý 2006). Our results emphasize the importance of group size equalization not only for diagnostic value calculations but also for assignments of relevés based on fidelity values.

TFVI assignments based on *Indval* preserved the initial number of groups regardless of the method used to produce the initial classification (i.e., modified TWINSPAN, k-means or PAM). Hence, *TFVI* based on *Indval* can be considered superior to other assignment rules in the sense that it does not produce strong alterations of the vegetation concepts in the original (unsupervised) classification. In addition, classifications obtained using assignments based on *Indval* resulted in the highest predictive performance. If assigning new relevés to existing vegetation units is the only usage of the assignment rule, *Indval* should be recommended because of its higher predictive power. However, if the assignment rule is applied for the refinement of a classification, other fidelity measures may also be suitable, because in this case having a good predictive performance may not be as important as improving the quality of the classification.

The choice of an evaluator index often implies a bias in the evaluation towards classification procedures that better match the concepts considered important in the conception of the evaluator index. Our quality evaluation results also showed that the ranking of evaluators is influenced by the type of vegetation data (i.e. incidence or abundance- based). Actually, in our analysis non-geometric evaluators (i.e., *ISAMIC* and *Morisita*) as well as incidence-based evaluators indicated that classifications obtained using correlation measures followed by FPFI were better than *IndVal*-based reassignments. Since the calculation of non-geometric and incidence-based evaluators is also based on frequency values, it can be said that the results of these evaluators were biased towards classification solutions obtained using phi coefficients such as *TFVI* based on *r* and FPFI. A similar bias could be argued towards abundance-based site-group association measures when using abundance-based evaluators. We found *Indval* to perform very well with abundance-based evaluators but this was not the case for other abundance-based measures such as *IndVal*.

In terms of the overall quality of the resulting classification, our results indicate that the *TFVI* framework works better when the chosen fidelity measure is either *Indval* or *r*. Indeed, comparisons based on all evaluators indicated that *TFVI* based on $Indval_{ind}^{g}$ was the first option, followed by *r* and *FPFI*. There are two differences between *Indval* and *r*. One is that *Indval* uses abundance data for the calculation of fidelity values, whereas *r* does not (but remember that the *TFVI* framework uses species abundance values for assignments regardless of the fidelity measure). The second difference is that *Indval* does not take into account species absences values outside the target site group. The fact that absences outside the target site group contribute to the strength of association in r_{φ}^{g} suggests a potential overestimation of the fidelity value (De Caceres et al. 2008). In this sense, De Cáceres and Legendre (2009) mentioned that indicator value indices have the advantage, compared to correlation indices, of being less dependent on the context of fidelity determination.

5 Conclusion

While the results of our analysis based on two vegetation datasets in the Hyrcanian forests points towards a slight preference of *Indval* over *r* for relevé assignments, we acknowledge that additional studies are necessary, using both simulated and real datasets, before more conclusive recommendations can be made in favor of one or another. Given its good performance in the context of the *TFVI* framework, one could ask whether *Indval*, being based on species abundances, should also be preferred over phi fidelity indices in the *FPFI* framework too. Additional work is also needed to test this hypothesis. It may well be the case that the preference for one site-group association measure or another depends on both the intended usage of the assignment rule (i.e., for assigning new relevés to an existing classification vs. refining an initial classification) and on the kind of vegetation considered (e.g., forests vs. grasslands, or species-poor vs. species-rich vegetation).

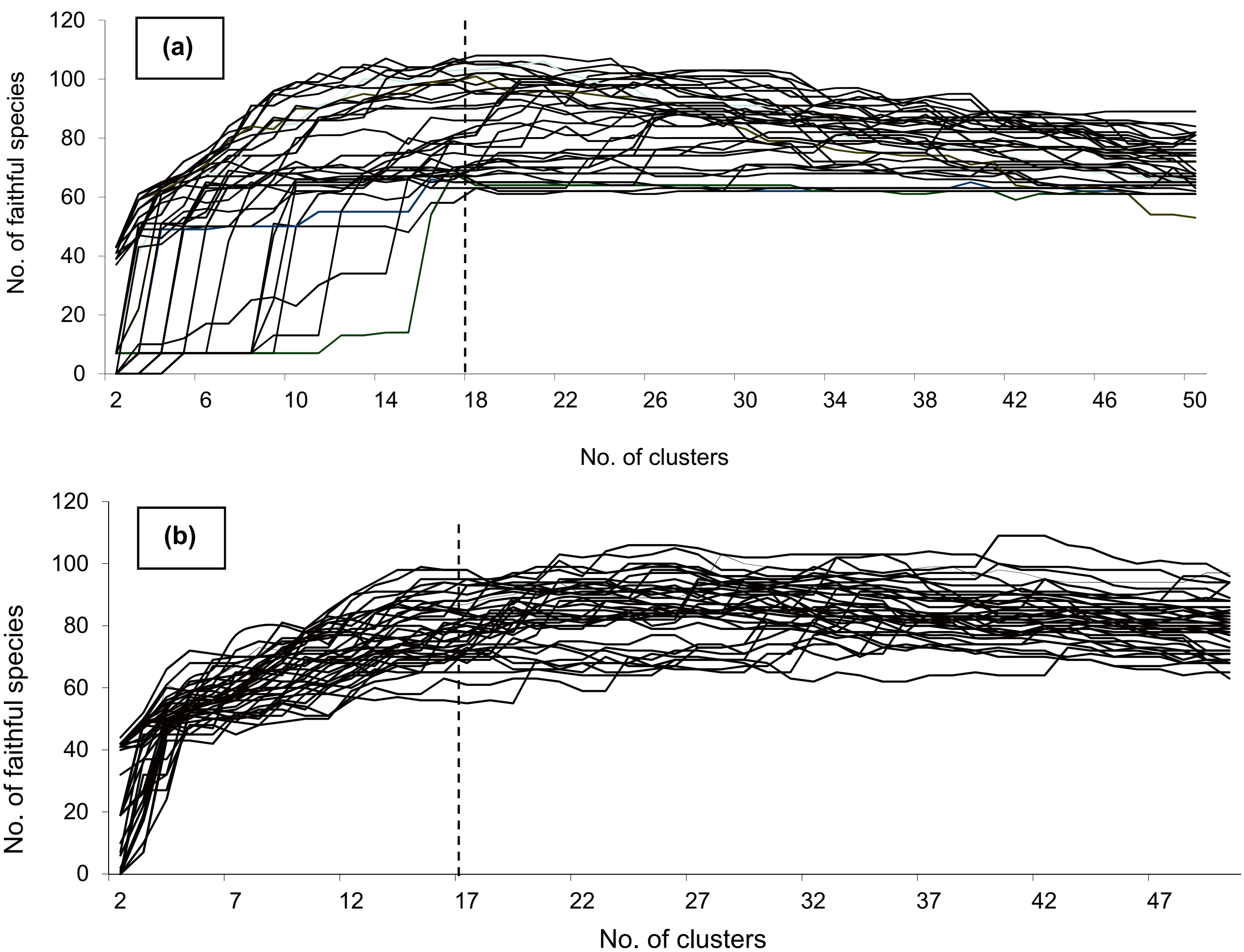
Appendix 1

Result of OptimClass 1 analysis applied to the 40 clustering solution using Fisher's exact test with cut level of 10–8 for species to be considered as faithfull. Each curve represents the number of faithfull species from one cluster analysis defined by a uniqe combination of cover transformation, distance measures and group linkage. (a) and (b) are related to *B. hyrcana* and *T. baccata* dataset, respectively.

Appendix 2

Equations and descriptions for all evaluators

Morisita's Index of niche overlap with the aim of explaining of floristically niche overlap for a particular classification solution (Aho et al 2008) was used to evaluate the separability of vegetation units. This is calculated using proportional occurrence of each species in a group with respect to all species in that group (Eq. 5). C_H scores were calculated for all pairwise niche overlap measures between groups. For g groups: [(g2 – g)/2 = no. of comparisons]. The mean of these pairwise relationships was calculated to represent a measure of cluster overlap for a particular classification solution. Minimal niche overlap indicates optimal solutions.



$$C_H = \frac{2\sum_{i=1}^{n} p_{ij} \times p_{ik}}{\sum_{i=1}^{n} p_{ij}^2 + \sum_{i=1}^{n} p_{ik}^2} \tag{5}$$

C_H = adapted Morisita index of niche overlap; p_{ij} = proportional occurrence of species i with respect to all species in group j; p_{ik} = proportional occurrence of species i with respect to all species in cluster k; n = total number of species.

ISAMIC-indicator species analysis to minimize intermediate constancy (Roberts 2010) is an analysis that measures the constancy (either high or low) of species within clusters irrespective of how many clusters a species occurs in (Robert 2015). It calculates the degree to which species are either always present or always absent within clusters (Robert 2010). The index is calculated as Eq. 6.

$$IS = \frac{\sum_{j=1}^{n} \left(\left(2 \sum_{g=1}^{k} \left| C_{jg} - 0.5 \right| \right) / k \right)}{n} \tag{6}$$

IS = ISAMIC score; n = the number of species; k = is the number of groups; c_{jg} = is the constancy of species j in group g. The statistic is bounded 0–1 and higher values are better.

ISA-indicator species analysis (Aho et al. 2008) is calculated by *IndVal* index (Dufrene and Legendre 1997). The IndVal index combines species fidelity (the proportion of sites in which species j is present within group g) with species specificity to groups (is the ratio of the mean abundance of species j in group g and the sum of means of the same species over all groups) to calculate a combined index (Podni and Csanyi 2010). Fidelity is calculated as relative frequency (Eq. 7), and its range is 0–1. The minimum obtained when the species is absent from group k and maximum resulting when the species occurs in every site in that group. Specificity is obtained as Eq. 8. These two terms are multiplied and then scaled to 100 to express the indicator value of species j in group g in terms of percentage (Eq. 9). By the way, each species is assigned an indicator value for every group and the largest value is tested for significance with Monte Carlo procedure, resulting in p values. Finally, the number of significant indicators at $\alpha = 0.05$ was derived from this test and expressed as ISA of each clustering solutions.

$$B_{jg} = \sum_{g=1}^{k} \frac{np_{jg}}{Np_j} \tag{7}$$

$$A_{jg} = \frac{\sum_{i=1}^{n_g} a_{jig} / n_g}{\sum_{g=1}^{k} \sum_{i=1}^{n_g} a_{jig} / n_g} \tag{8}$$

$$IndVal_{jg} = 100 \times A_{jg} \times B_{jg} \tag{9}$$

B_{jk} = species fidelity; np_{jg} = number of occurrences of species j in the group g; Np_j = number of occurrences of species j in all group; a_{jig} = abundance of species j in sample unit i of group g; n_g = number of sample units in group g.

C-index (Hubert and Levin, 1976) is defined as Eq. 10, where dw is the sum of within cluster similarity for all clusters. If p is the number of pairs of relevés in the same vegetation units, pairs of samples for which both samples are located in the same units, max (d_w) and min (d_w) are the sum of the p- pairs of relevés with largest and smallest distances respectively. This index is confined to interval 0–1 and minimum C-index scores were considered as optimal clustering solutions of TFVI models (Aho et al. 2008).

$$C_{index} = \left(\frac{d_w - \min(d_w)}{\max(d_w) - \min(d_w)} \right) \tag{10}$$

PARTANA ratio index (Eq. 11) calculates the ratio of the mean within-cluster similarity to the mean among-cluster similarity (Roberts 2015) so in this research it was used as a goodness of vegetation units which are classified by TFVI models. High *PARTANA* value implies low within group dissimilarity and high dissimilarity of relevés within groups to relevés outside of groups (Aho et al. 2008) so higher *PARTANA* values, were considered as optimal clustering solutions of TFVI models.

$$P = \frac{\sum_{z=1}^{C} \sum_{\substack{i=1 \\ i \in z}}^{N-1} \sum_{\substack{j=i+1 \\ j \in 1}}^{N} S_{ij} \Big/ \sum_{z=1}^{C} \left(n_z^2 - n_z \right) / 2}{\sum_{i=1}^{N-1} \sum_{\substack{j=i+1 \\ \omega_i \neq \omega_j}}^{N} S_{ij} \Big/ \sum_{z=1}^{C-1} \sum_{k=z+1}^{C} n_z.n_k} \tag{11}$$

where i and j are relevé, C is the number of vegetation units, N is total number of relevé, n_k is the number of relevé in unit k, $i\varepsilon z$ indicates that relevé i is a member of unit z, ω is membership, and $\omega_i \neq \omega_j$ indicates that relevé i and j are not members of the same units. n_z is the number of relevé in the z unit ($z = 1, 2, \ldots, C$). S_{ij} is the similarity of two relevé i and j. $S_{ij} = ((1\text{-}dij)/\max dij)$ and dij is maximum of all possible pairwise Hellinger Euclidean distances.

Point biserial correlation (PBC) is a correlation measure between a continuous variable (i.e., distance measure) and a binary variable (i.e., a variable whose values are 0 or 1). PBC was defined as Pearson correlation of D and B matrixes (Eq. 12). D is distance data matrix of all possible pairwise Hellinger Euclidean distances and B is a symmetric matrix of ones and zeroes with the same dimensions as D; 0 = relevés in the same vegetation units, 1 = relevés in different vegetation units. The higher PBC values were considered as optimal clustering solutions of TFVI models.

$$PBC = corr(D, B) \tag{12}$$

Silhouette coefficient (*SW*) combine ideas of both cohesion (measures how closely related are objects in a cluster) and separation (measure how distinct or well separated a cluster is from other clusters), but for individual points as well as clusters. Cohesion is the within- cluster mean distance $a(i)$ as the average dissimilarity between object i and all other objects in the cluster to which i belongs (K for instance) (Eq. 13). Separation is the between- cluster minimum distance $b(i)$ as the minimum average dissimilarity to the instances of each cluster that are different to K (Eq. 14). The silhouette coefficient is calculated for each object i based on its cohesion and separation values as Eq. 15. The average of all output values for each object is called the average silhouette width (*ASW*) which is the final result and is in the [− 1, 1] range. A high value indicates good quality clusters (Guerra et al. 2012).

$$a(i) = \frac{1}{n_k - 1} \sum_{\substack{g \\ i \neq i' \\ i,i' \in g}}^{k} d(p_i, p_{i'}) \tag{13}$$

$$b(i) = \min \left[\frac{1}{n_{k'}} \sum_{\substack{i,j=1 \\ i \in k \\ j \in k'}}^{n_k, n_{k'}} d(p_i, p_j) \right] \tag{14}$$

$$s(i) = \frac{b(i) - a(i)}{\max(b(i), a(i))} \tag{15}$$

Analysis of similarities (*ANOSIM*) as a test of the significance of the groups that had been defined a priori (Clarke 1993). It provides a way to test statistically whether there is a significant difference between two or more groups of sampling units (Oksanenen 2013). This method uses the results of dissimilarity matrix which is derived by a suitable distance index; however, the rank order of dissimilarity values is used. If two groups of sampling units are really different in their species composition, then compositional dissimilarities between the groups ought to be greater than those within the groups. The *ANOSIM* statistic (R) is based on the difference of mean ranks between groups (r_B) and within groups (r_W) which is showed by Eq. 16. R is bounded in the [− 1, o, 1], which $R = 1$ indicates that all the most similar samples are within the same groups. $R = 0$ occurs if the high and low similarities are perfectly mixed and bear no relationship to the group. A value of − 1 indicates that the most similar samples are all outside of the groups. The null hypothesis is therefore that there are no differences between the members of the various groups. The statistical significant is done by Monte-Carlo permuting test. If the value of R is significant, you can conclude that there is evidence that the samples within groups are more similar than would be expected by random chance.

$$R = \frac{\overline{r}_B - \overline{r}_W}{n(n-1)/4} \tag{16}$$

$\overline{r}_B$ and $\overline{r}_W$ are the mean of the ranked similarity between groups and within groups, respectively, and n is the total number of samples (objects).

Appendix 3

The average correction ratios of group's penalty for the evaluator's values

Classification solutions	Modified TWINSPAN	K-means	PAM
B. hyrcana dataset			
FPFI	0.96	1.00	1.00
IndVal	0.68	0.91	0.92
Indval	1.00	1.00	1.00
IndVal	0.40	0.59	0.69
Indval	0.87	0.85	0.92
r	0.95	1.00	1.00
r	0.91	1.00	1.00
r	0.92	1.00	1.00
r	0.87	1.00	1.00
T. baccata dataset			
FPFI	1.00	1.00	1.00
IndVal	0.94	0.94	0.94
Indval	1.00	1.00	1.00
IndVal	1.00	1.00	1.00
Indval	0.94	0.82	0.88
r	0.94	1.00	0.94
r	0.94	1.00	0.88
r	1.00	1.00	1.00
r	1.00	1.00	1.00

Acknowledgments We are grateful to Erwin Dreyer, Laurent Bergès, and two anonymous reviewers for their valuable comments on the former version of the manuscript, and Saleh Sekhavat and Abbas Ahmadi for their tireless help with field sampling of *B. hyrcana* forests and laboratory work. Also, we would like to thank Pari Karami, Meysam Soufi, and Mohammad-Reza Abbaszadeh for providing *T. bacata* dataset.

Funding This study is financially supported by the Tarbiat Modares University.

Data availability The datasets generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval The authors declare that they obtained the approval of Research Council of Natural Resources Faculty in Tarbiat Modares University and in all correspondence with permission from Iranian Forest, Range and Watershed Organization ethics committee for conducting the study in Hyrcanian forests.

References

Aho K (2017) Plant Ecology: Computational tools for plant ecology. R package version 0.5-1. https://sites.google.com/a/isu.edu/aho/

Aho K, Roberts DW, Weaver T (2008) Using geometric and non-geometric internal evaluators to compare eight vegetation classification methods. J Veg Sci 19(4):549–562

Asadi H, Hosseini SM, Esmailzadeh O, Ahmadi A (2011) Flora, Life form and chorological study of Box tree (*Buxus hyrcana* Pojark.) sites in Khybus protected forest Mazandaran. J. Plant Biol 3:27–40 (**[In Persian]**)

Boublík K (2010) Formalized classification of the vegetation of *Abies alba*-dominated forests in the Czech Republic. Biologia 65:822–831

Boublík K, Petrik P, Sadlo J, Hedl R, Willner W, Cerny T, Kolbek J (2007) Calcicolous beech forests and related vegetation in the Czech Republic: a comparison of formalized classifications. Preslia 79(2):141–161

Brogden HE (1949) A new coefficient: application to biserial correlation and to estimation of selective efficiency. Psychometrika 14:169–182

Bruelheide H (1997) Using formal logic to classify vegetation. Folia Geobotanicaet Phytotaxonomica 32:41–46

Černá L, Chytrý M (2005) Supervised classification of plant communities with artificial neural networks. J Veg Sci 16:407–414

Chytrý M, Tichý L (2003) Diagnostic, constant and dominant species of vegetation classes and alliances of the Czech Republic: a statistical revision. Masaryk University, Brno, CZ

Chytrý M, Tichý L (2018) National vegetation classification of the Czech Republic: a summary of the approach. Phytocoenologia 48:121–131

Chytrý M, Tichý L, Holt J, Botta-Dukát Z (2002) Determination of diagnostic species with statistical fidelity measures. J Veg Sci 13:79–90

Clarke KR (1993) Non-parametric multivariate analyses of changes in community structure. Aust J Ecol 18:117–143

Dai X, Page B, Duffy KJ (2006) Indicator value analysis as a group prediction technique in community classification. S Afr J Bot 72:589–596

De Cáceres M (2016) How to use the "indicspecies" package. R package (ver. 17.1)

De Cáceres M, Font X, Oliva F (2008) Assessing species diagnostic value in large data sets: a comparison between phi-coefficient and Ochiai index. J Veg Sci 19:779–788

De Cáceres M, Legendre P, Wiser SK, Brotons L (2012) Using species combinations in indicator value analyses. Methods Ecol Evol 3:973–982

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

Dengler J, Chytrý M, Ewald J (2008) Phytosociology. In: Jørgensen SE, Fath BD (eds) Encyclopedia of ecology. Elsevier, Amsterdam, The Netherlands, pp 2767–2779

Douda J (2008) Formalized classification of the vegetation of alder carr and floodplain forests in the Czech Republic. Preslia 80:199–224

Dufrêne M, Legendre P (1997) Species assemblages and indicator species: the need for a flexible asymmetrical approach. Ecol Monogr 67:345–366

Esmailzadeh O, Asadi H (2014) Total phi fidelity index (TPFI) as a new algorithm in plant communities analysis. Iran J Forest 6:215–232 (**[In Persian]**)

Esmailzadeh O, Asadi H, Ahmadi A (2014) Phytosociology of Khybus Protected Area. J Wood Forest Sci Technol 19:1–20 (**[In Persian]**)

Fraley C, Raftery AE, Murphy TB, Scrucca L (2012) *mclust* Version 4 for R: Normal mixture modeling for model-based clustering, classification, and density estimation. Technical Report No. 597, Department of Statistics, University of Washington.

Gégout JC, Coudun C (2012) The right relevé in the right vegetation unit: a new typicality index to reproduce expert judgement with an automatic classification programme. J Veg Sci 23(1):24–32

Guerra L, Robles V, Bielza C, Larrañaga P (2012) A comparison of clustering quality indices using outliers and noise. Intell Data Anal 16(4):703–715

Habibi KH (1992) Fundaments of forest soil science. Tehran University Press, Tehran, IR (**[In Persian]**)

Hill MO (1979) TWINSPAN: a FORTRAN program for arranging multivariate data in an ordered two-way table by classification of the individuals and attributes. Cornell University, Ithaca, NY, Section of Ecology and Systematics

Hill MO (1989) Computerized matching of relevés and association tables, with an application to the British National Vegetation Classification. Vegetatio 83:187–194

Horn HS (1966) Measurement of 'overlap' in comparative ecological studies. Am Nat 100:419–424

Hubert L, Arabie P (1985) Comparing partitions. J Classif 2:193–218

Hubert LJ, Levin JR (1976) A general framework forassessing categorical clustering in free recall. Psychology Bulletin 83:1072–1080

Janišová M, Uhliarová E, Ružičková H, Jandt U, Becker T, Dengler J (2010) Expert system-based classification of semi-natural grasslands in submontane and montane regions of central Slovakia. Tuexenia 30:375–422

Kaufman L, Rousseeuw PJ (1990) Finding groups in data: an introduction to cluster analysis. Wiley, New York, US

Landucci F, Gigante D, Venanzoni R, Chytrý M (2013) Wetland vegetation of the class Phragmito-Magno-Caricetea in central Italy. Phytocoenologia 43:67–102

Legendre P, Gallagher ED (2001) Ecologically meaningful transformations for ordination of species data. Oecologia 129:271–280

Legendre P, Legendre L (2012) Numerical ecology, 3rd edn. Elsevier, Amsterdam, NL

Luther-Mosebach J, Dengler J, Schmiedel U, Röwer IU, Labitzky T, Gröngröft A (2012) A first formal classification of the Hardeveld vegetation in Namaqualand, South Africa. Appl Veg Sci 15:401–431

Maciejewski L, Pinto PE, Wurpillot S, Drapier J, Cadet S, Muller S, Agou P, Renaux B, Gégout JC (2020) Vegetation unit assignments: phytosociology experts and classification programs show similar performance but low convergence. Appl Veg Sci. https://doi.org/10.1111/avsc.12516

Mac Queen J (1967) Some methods for classification and analysis of multivariate observations. In: Cam, L.M.L. & Neyman, J. (eds.) Proceedings of the fifth Berkeley symposium on mathematical

statistics and probability, pp. 281–297. University of California Press, Berkeley, US.
Maechler M, Rousseeuw P, Struyf A, Hubert M, Hornik K, Studer M (2013) Package 'cluster'. Dosegljivo na
Marvie Mohadjer MR (2005) Silviculture. University of Tehran Press, Tehran, IR (**[In Persian]**)
Mueller-Dombois D, Ellenberg H (1974) Aims and methods of vegetation ecology. John Wiley, New York, US
Podani J, Csányi B (2010) Detecting indicator species: some extensions of the IndVal measure. Ecol Indic 10(6):1119–1124
Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'hara, RB. Oksanen, MJ (2013) Package 'vegan'. Community ecology package, version 2(9), 1–295
Roberts D (2005) Vegetation classification in R, for labdsv ver. 1.1–1, vegetation ecology package. http://www.cran.r-project.org
Roberts DW (2010) Optpart: optimal partitioning of similarity relations. R package version, 2-0
Roberts DW (2015) Vegetation classification by two new iterative reallocation optimization algorithms. Plant Ecol 216(5):741–758
Roberts DW, Roberts MDW, (2016) Package 'labdsv'. Ordination and Multivariate, 775
Rodríguez-Rojo MP, Fernández-González F, Tichý L, Chytrý M (2014) Vegetation diversity of mesic grasslands (*Arrhenatheretalia*) in the Iberian Peninsula. Appl Veg Sci 17:780–796
Roleček J, Tichý L, Zelený D, Chytrý M (2009) Modified TWINSPAN classification in which the hierarchy respects cluster heterogeneity. J Veg Sci 20:596–602
Rousseeuw PJ (1987) Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J Comput Appl Math 20:53–65
Sagheb-Talebi K, Sajedi T, Pourhashemi M (2014) Forests of Iran: A Treasure from the Past, a Hope for the Future. Springer, Berlin, GE
Soleymanipour SS, Esmailzadeh O (2015) Flora, life form and chorology of Box trees (*Buxus hyrcana*) habitats in forests of the Farim area of Sari. Taxonomy and Biosystematics 7:39–56 (**[In Persian]**)
Svitková I, Šibík J (2013) An expert-based classification of high-altitude arctic-alpine vegetation of the class Caricirupestris-Kobresietea Ohba 1974: achievements and obstacles. Plant Biosystems 147:315–327
Tichý L (2002) JUICE, software for vegetation classification. J Veg Sci 13:451–453
Tichý L (2005) New similarity indices for the assignment of relevés to the vegetation units of an existing phytosociological classification. Plant Ecol 179:67–72
Tichý L, Chytrý M (2006) Statistical determination of diagnostic species for site groups of unequal size. J Veg Sci 17:809–818
Tichý L, Chytrý M, Botta-Dukát Z (2014) Semi-supervised classification of vegetation: preserving the good old units and searching for new ones. J Veg Sci 25:1504–1512
Tichý L, Chytrý M, Hájek M, Talbot SS, Botta-Dukát Z (2010) OptimClass: using species-to-cluster fidelity to determine the optimal partition in classification of ecological communities. J Veg Sci 21(2):287–299
Van der Maarel E (1979) Multivariate methods in phytosociology, with reference to the Netherlands. Junk The Hague, NL
van Tongeren O, Gremmen N, Hennekens S (2008) Assignment of relevés to pre-defined classes by supervised clustering of plant communities using a new composite index. J Veg Sci 19:525–536
Westhoff V, van der Maarel E (1973) The Braun-Blanquet approach. In: Whittaker RH (ed) Ordination and classification of plant communities. W. Junk, The Hague, NL, pp 617–737
Whittaker RH (1962) Classification of natural communities. The Botanical Review 28:1–239
Wiser SK, De Cáceres M (2013) Updating vegetation classifications: an example with New Zealand's woody vegetation. J Veg Sci 24:80–93

Authors and Affiliations

Hamed Asadi[1,2] · Omid Esmailzadeh[1] · Miquel De Cáceres[3] · Seyed Mohsen Hosseini[1]

Hamed Asadi
h.asadi@sanru.ac.ir

Miquel De Cáceres
miquelcaceres@gmail.com

Seyed Mohsen Hosseini
hosseini@modares.ac.ir

[1] Department of Forestry, School of Natural Resources & Marine Sciences, Tarbiat Modares University, P.O. Box, 46417-76489 Noor, Iran

[2] Department of Forestry, College of Natural Resources, University of Agricultural Science and Natural Resources of Sari, P.O. Box 578, Sari, Iran

[3] Forest Sciences Center of Catalonia (CTFC), Ctra. antiga St Llorenc km 2, 25280 Solsona, Spain