

JUICE

**program for management,
analysis and classification
of ecological data**

2nd Edition of the Program Manual

Tichý Lubomír¹, Holt Jason² & Nejezchlebová Martina¹

2nd part



Czech Republic

2010

¹ Dept. of Botany, Masaryk Univ. Brno, Kotlarska 2, 611 37 Brno, tichy@sci.muni.cz

² Box 37, Hinsdale, MT 59241, USA, jholt@seznam.cz

Table of Contents

2.	THE FIDELITY CONCEPT.....	2
2.1.1.	Fidelity Measures	2
2.1.2.	Fidelity Measurements for Vegetation Units of Unequal Size	5
2.1.3.	Selecting Fidelity Measures and Standardization in JUICE.....	7
2.1.4.	Fidelity Tests	9
2.1.5.	Quantitative Fidelity Measures	9
2.2.	Synoptic Tables	11
2.2.1.	Synoptic Table Display	11
2.2.2.	Percentage Constancy Synoptic Table	12
2.2.3.	Categorical Synoptic Table	13
2.2.4.	Synoptic Table with Absolute Frequency ('Absolute Constancy')	14
2.2.5.	Fidelity Synoptic Table	15
2.2.6.	Synoptic Tables and Cover (Maximum, Average, Median, Modus).....	16
2.2.7.	Zlatnik's Combined Scale	16
2.2.8.	Average Cover (Barkman's Total Cover Value).....	17
2.2.9.	Combined Synoptic Tables	18
2.2.10.	Synoptic Table Export.....	19
2.2.11.	Sorting Species in Synoptic Tables	19
2.2.12.	Analysis of Synoptic Columns Combining with Exporting the Results.....	20
2.2.13.	Uniqueness of Vegetation Unit	22
2.2.13.1	Theoretical Background	23
2.2.13.2	Uniqueness in JUICE	23
2.2.14.	Average Values of Columns in Synoptic Table	24
2.2.15.	Crispness of Classification	25
2.2.15.1	Theoretical Background	25
2.2.15.2	Optimal Use of This Function in JUICE	26
2.2.16.	Comparison of Two Synoptic Tables	27
2.2.16.1	Theoretical Background	27
2.2.16.2	Comparing Two Synoptic Tables within JUICE	28
2.2.17.	Number of Species in Synoptic Table	28
2.2.18.	Merge Two nearest columns	28

2. The Fidelity Concept

Chytrý, M., Tichý, L. & Holt, J.

Statistical fidelity measurements are useful for measuring species concentration in vegetation units and comparing diagnostic values among species in a particular vegetation unit or among vegetation units for a particular species. A completely statistical approach to determination of diagnostic species by measuring fidelity, proposed by Chytrý et al. (2002), is included in JUICE. The program has 16 different fidelity measures available from the ‘Fidelity measures’ tab of the ‘Options’ window. (See Section 2.1.3.) They are for use in synoptic tables as described in Section 2.2.

2.1.1. Fidelity Measures

The fidelity measures available in JUICE can consider either binary presence/absence data or quantitative cover data. For simplicity, only four fidelity measures, using presence/absence data, will be discussed in this section. Measures that take cover into account will be discussed in Section 2.1.5. For details on other fidelity measures, refer to Chytrý et al. (2002).

Here the same notation is used as in Bruelheide (1995, 2000) and Chytrý et al. (2002):

- N ... number of relevés in the data set,
- N_p ... number of relevés in the target vegetation unit,
- n ... number of occurrences of the species in the data set,
- n_p ... number of occurrences of the species in the target vegetation unit.

(To see how these quantities are defined when quantitative cover data are taken into account, see Section 2.1.5.)

a) The **u value for a hypergeometric distribution** (Chytrý et al. 2002), further referred to as u_{hyp} , compares the observed number of occurrences of the species in the vegetation unit (n_p) with the expected number of occurrences ($\mu = n \cdot N_p / N$). The fidelity values calculated using this coefficient are not influenced by species cover or abundance (unless n and n_p are defined as in Section 2.1.5). The u value is defined as the deviation of the observed frequency of the species occurrence in the vegetation unit from the expected frequency, compared with the standard deviation:

$$u_{hyp} = (n_p - \mu) / \sigma_{hyp} . \quad (\text{Eq. 1})$$

The measure counts the number of standard deviations between the observed data and the result expected under the null hypothesis that species occurrences are independent of the target vegetation unit. Thus u_{hyp} provides a measure of the statistical significance of the observed relationship. It can take on positive or negative values with the range depending on the number of relevés in the dataset. The maximum u value, $\sqrt{N-1}$, is achieved when $N_p = n = n_p$. Fidelity rankings according to u_{hyp} will be identical to rankings according to chi squared or the phi coefficient.

b) The **phi coefficient of association** between species and vegetation units (Sokal & Rohlf 1995, Chytrý et al. 2002) is a statistical measure of association between two categories, which can be used as measure of fidelity. It can be defined as:

$$\Phi = \frac{u_{hyp}}{\sqrt{N-1}} = \frac{N \cdot n_p - n \cdot N_p}{\sqrt{n \cdot N_p \cdot (N-n) \cdot (N-N_p)}} \quad (\text{Eq. 2})$$

Traditionally, the phi coefficient considers only presence/absence, so fidelity values calculated using this coefficient are not influenced by species cover or abundance. The values of the phi coefficient range from -1 to 1 , but for convenience, they are multiplied by 100 in the program. The highest phi value of 1 is achieved if the species occurs in all relevés of the vegetation unit and is absent elsewhere. A positive value lower than 1 means that the species is absent from some relevés of the vegetation unit or present in some relevés outside the vegetation unit. A value of 0 is obtained when the relative frequency of the species in the vegetation unit equals the relative frequency in the rest of the dataset, thus indicating no relation between the target species and the target vegetation unit.

An advantage of the phi coefficient over some other statistical fidelity measures is its independence of data set size. On the other hand, the phi coefficient contains no information about statistical significance.

The phi coefficient (along with u value and χ^2) is more or less dependent on the relative size of vegetation units. If the target vegetation unit represents 10% of the total data set (Fig. 1a), Φ is relatively high for species that are not very common within the target unit, provided the species are very rare outside the target unit. However, for species that are common outside the unit, even the species with high frequency within the target vegetation unit are given low phi values. On the other hand, if the target unit makes up 50% of the total data set (Fig. 1b), species must generally have higher frequencies within the target unit in order to have high phi values; if a species has a high frequency within the target unit, a relatively high phi value can be attained even with a smaller difference in frequency between the target and other units.

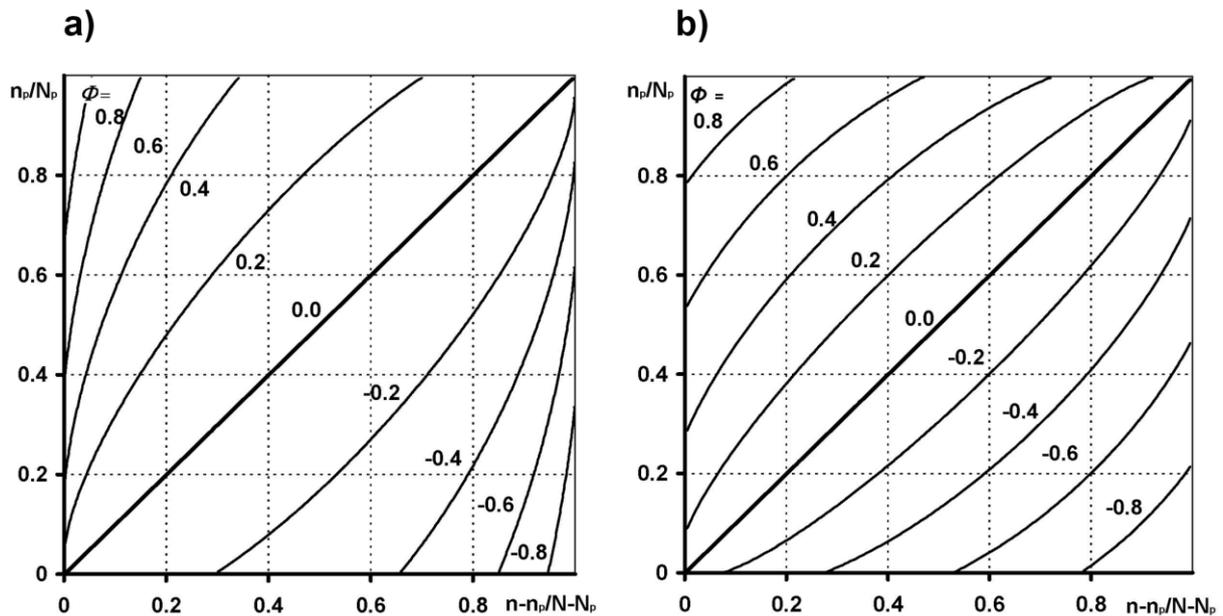


Fig. 1: Dependence of the phi coefficient on the relative frequency of species occurrences within (vertical axis) and outside (horizontal axis) the target vegetation unit, shown for vegetation units equal to (a) 10% and (b) 50% of the size of the total data set.

c) The **Dufrène-Legendre Indicator Value** (Dufrène & Legendre 1997) is defined as the product of relative species abundance and relative species frequency within the target vegetation unit. Relative species abundance is the mean abundance of the species in the target vegetation unit compared to all vegetation units in the data set; the sum of mean abundances in each

vegetation unit is used instead of the sum of actual abundances over all relevés in order to remove the effect of unequal vegetation unit size. JUICE only computes the Dufrêne-Legendre Indicator Value for presence/absence data, which means that the relative abundance is replaced by the relative frequency in the target vegetation unit divided by the sum of relative frequencies in all vegetation units of the data set. For a comparison between a target vegetation unit and the rest of the data set, the Dufrêne-Legendre Indicator Value can be expressed as:

$$IndVal = \frac{n_p/N_p}{n_p/N_p + (n - n_p)/(N - N_p)} \cdot \frac{n_p}{N_p}. \quad (\text{Eq. 3})$$

This formula can be easily extended to a comparison of vegetation units within the classified data sets by replacing the species frequency outside the target vegetation unit, $(n - n_p)/(N - N_p)$, in the denominator, by the sum of relative frequencies in all the other vegetation units of the given typology.

The Dufrêne-Legendre Indicator Value is independent of the relative size of the target vegetation unit. It implicitly standardizes the size of all vegetation units in the data set, including the target vegetation unit. All vegetation units are weighted equally, and changing N_p does not influence the resulting numerical value.

The categorical form of the Dufrêne-Legendre Indicator Value probably deviates from most researchers' intuitive expectations of the properties of a suitable fidelity measure. The next figure shows that the Dufrêne-Legendre Indicator Value gives a rather high weight to common species. It could be said, however, that the Dufrêne-Legendre Indicator Value is a fidelity measure which emphasizes the group's fidelity to the species, i.e., the value is a good indicator of how frequent the species is within the vegetation unit but it is less affected by occurrences of the species outside the vegetation unit.

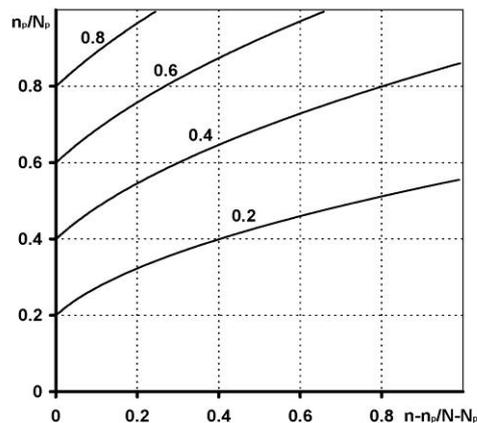


Fig. 2: Dependence of the Dufrêne-Legendre Indicator Value on the relative frequency of species occurrences within (vertical axis) and outside (horizontal axis) the target vegetation unit.

Note: The Dufrêne-Legendre Indicator Value's relationship to independence can be seen by comparing Fig. 1 (a or b) with Fig. 2. In each figure, the diagonal from the lower left to the upper right is the case of independence – the species is as frequent inside the vegetation unit as it is outside the vegetation unit. Note that this diagonal corresponds to the $\Phi = 0$ isoline in Fig.

1 but it is crossed by the isolines in Fig. 2. IndVal assigns this diagonal a value of $p/2$ (where $p = n/N$) as can be seen from where the 0.2 and 0.4 isolines cross it.

d) **Fisher's exact test** calculates $P(f(o) \geq n_p)$, the true probability of obtaining the observed number of occurrences of the species in the vegetation unit under the null hypothesis of independence. It may yield very small probability values, including those smaller than 10^{-100} , which are difficult to cope with in practical work. For this reason, $-\log_{10} P(f(o) \geq n_p)$ is a more practical quantity to use for a measure. Fisher's exact test is used either as a stand-alone fidelity measure or as a correction for calculation of missing information on statistical significance if the fidelity synoptic table with phi coefficient is displayed.

Note: Negative values of fidelity represent information which is not important in the context of most studies. Negative values are displayed only in the 'Interspecific Associations' window. (More on interspecific associations will be included in later versions of this manual.) They are hidden in synoptic tables.

2.1.2. Fidelity Measurements for Vegetation Units of Unequal Size

All the fidelity measures included in JUICE, except the Dufrene-Legendre Indicator Value, are affected by the relative size of vegetation units. Unfortunately, nearly every classification of relevés yields vegetation units of different sizes. To remove the dependence of the fidelity measures on the vegetation unit size, it is better to make a virtual standardization of the target vegetation unit (N_p) to a new value, hereafter called N_p' , which is constant for all vegetation units within the data set, without changing species frequencies within and outside each target vegetation unit, i.e., the quantities n_p/N_p and $(n - n_p)/(N - N_p)$, respectively. For example, in a data set with two vegetation units, we can set $N_p' = N/2$, i.e., equal to half of the total size of the data set, which enables us to compare the two vegetation units as if they were of the same size. Thus the resulting fidelity values are less dependent on the sampling effort and data set structure. Similarly, in a data set with four vegetation units, we can set $N_p' = N/4$. However, it is not necessary that the sum of N_p' values for all vegetation units within the data set equal N , the total size of the data set. For example, in a data set of four vegetation units, we can set $N_p' = N/2$ and compute a fidelity measure for the first vegetation unit as if this group were 50% of the total data set size and the other three groups together were also 50%. We can then compute a fidelity measure for the second vegetation unit as though it were 50% of the total data set size and so on. Such standardizations make sense in combination with the phi coefficient, but not so much with u_{hyp} or other statistical fidelity measures because data standardization violates their relationships to statistical significance.

A scheme of different standardizations of the size of vegetation units is presented in the next figure. Each line represents a data set, segments are vegetation units labelled A, B, C and D and segment length corresponds to the number of relevés in each group. The thick part of each segment represents relevés with occurrence of the given species; the thin part represents relevés where the species is absent. In figures (b) to (d), the four lines represent the four standardizations used for calculation of species fidelity to the target vegetation units A, B, C and D.

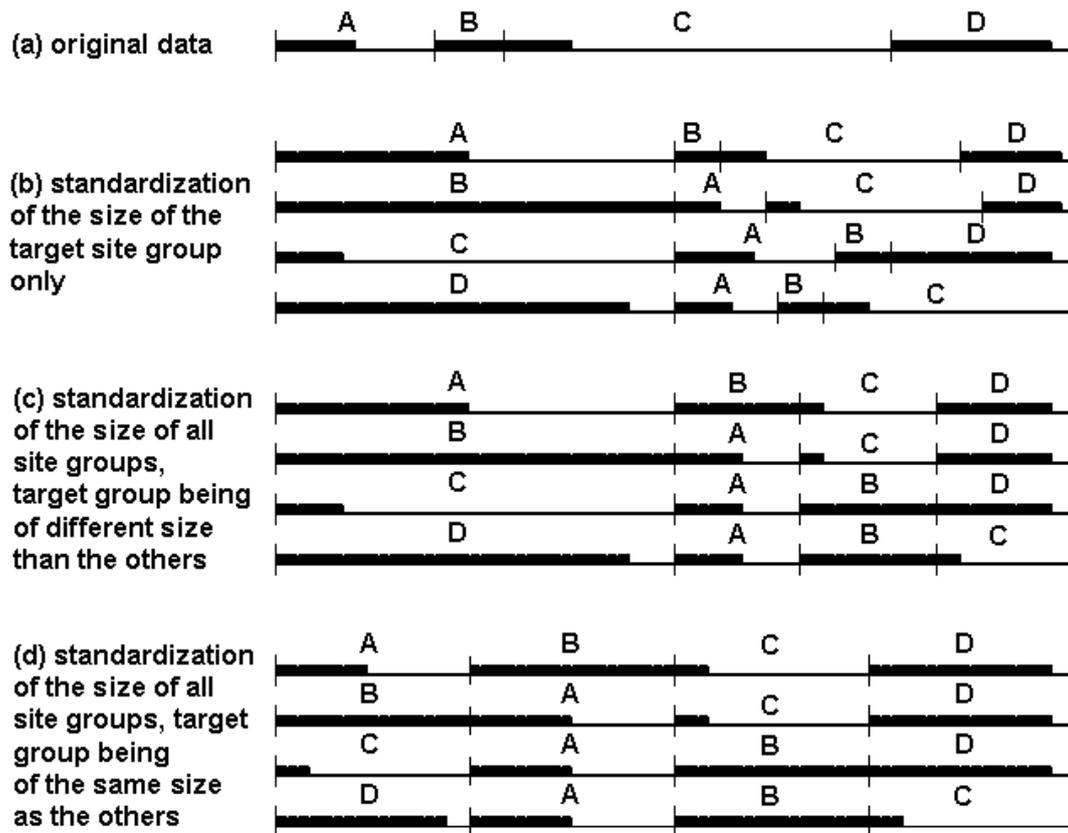


Fig. 3: A scheme of different standardizations of the size of relevé groups. Each line represents a data set, segments are groups labelled A, B, C and D and segment length corresponds to the number of relevés in each group. The thick part of each segment represents relevés that contain the given species, the thin part represents relevés where the species is absent. In figures (b) to (d), the four lines represent the four standardizations used for calculation of species fidelity to the target groups A, B, C and D.

Part (a) of the figure shows a simple artificial data set with vegetation units of unequal size. The phi coefficient depends on the size of the vegetation units. Standardizing the data set removes this dependence. After the standardizations in Fig. 3c or Fig. 3d, the fidelity measure depends only on the relative frequencies within each relevé group. The standardized size of the target vegetation unit can differ from the size of the other vegetation units (Fig. 3c) or be the same as the other vegetation units (Fig. 3d). The phi values after such standardization are entirely independent of the size of the vegetation units and can be directly compared across different vegetation units.

In some cases, however, it may not be desirable to standardize the size of all vegetation units in the data set. If diagnostic species for a few similar community types are determined using a large database, it may be useful to perform fidelity calculations with a data set that also includes relevés from other, unrelated community types (Chytrý et al. 2002). This approach finds diagnostic species that are of more general validity because they are tested against the background of other community types in the same geographical area. The additional relevés are usually treated (theoretically) as a single vegetation unit. Standardization to the size of other vegetation units would greatly and undesirably reduce its effect. However, its size can be held constant while the size of the other groups is standardized.

The phi coefficient applied to a data set with vegetation units standardized to equal size is independent of the actual differences in size of individual vegetation units. However, it depends on the standardized size of the target vegetation unit (N_p'), which may be either equal to the

size of the other vegetation units or set to any arbitrary value between 1 and $N - 1$. Setting the standardized size of the target vegetation unit (N_p') to a higher value gives a higher weight to common species and their frequency in the target vegetation unit. By contrast, setting N_p' to a lower value gives a higher weight to rare species and to the differences in species frequency within and outside the target vegetation unit. Changing the N_p' value can thus be used as a tool for modifying the properties of the phi coefficient with respect to weighting common or rare species.

The dependence of the phi coefficient on the relative size of the target vegetation unit can be tested in JUICE. See Section 2.1.3 pod and Section 2.2 on synoptic tables.

2.1.3. Selecting Fidelity Measures and Standardization in JUICE

The current fidelity measure is displayed on the Option Button on the left side of the Icon Bar. (See Section **Chyba! Nenalezen zdroj odkazů.**) Pressing this button opens the 'Fidelity measures' tab of the 'Options' window.

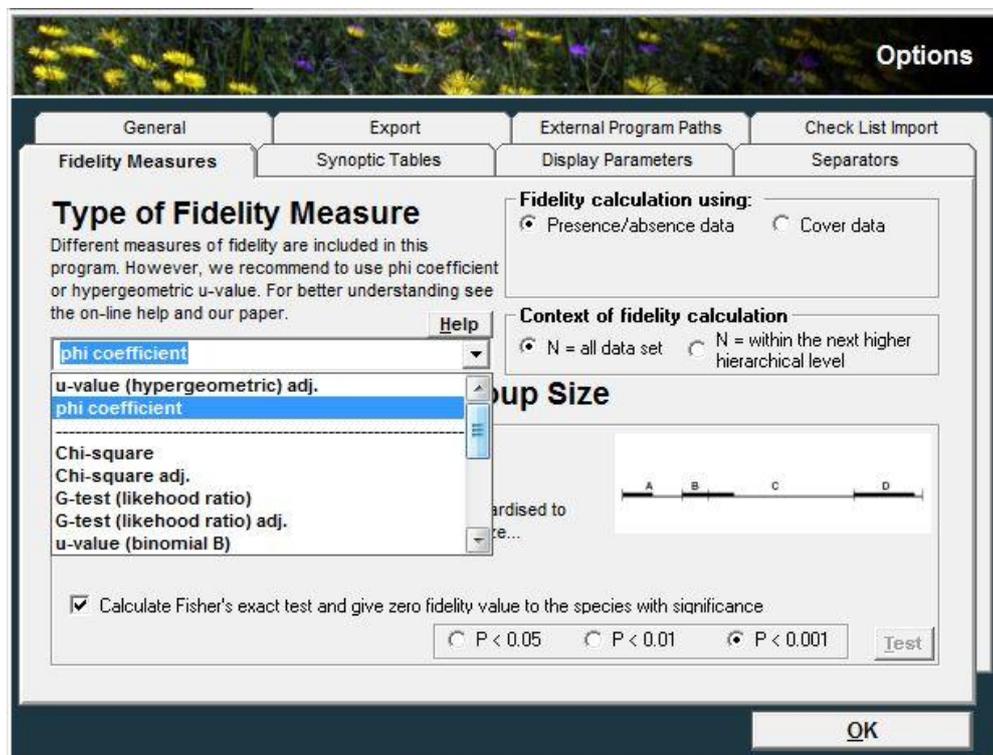


Fig. 4: Selecting a fidelity measure from the 'Fidelity measures' tab of the 'Options' window.

The user can select one of sixteen different fidelity measures, which are used to analyse presence/absence or quantitative data in association with various program functions (such as fidelity synoptic tables, interspecific associations or the COCKTAIL method).

JUICE versions 7.0.63 and higher have relevé group size standardization as discussed in Section 2.1.2 nad. The parameters of standardization can be adjusted in the 'Fidelity measures' tab of the 'Options' window.

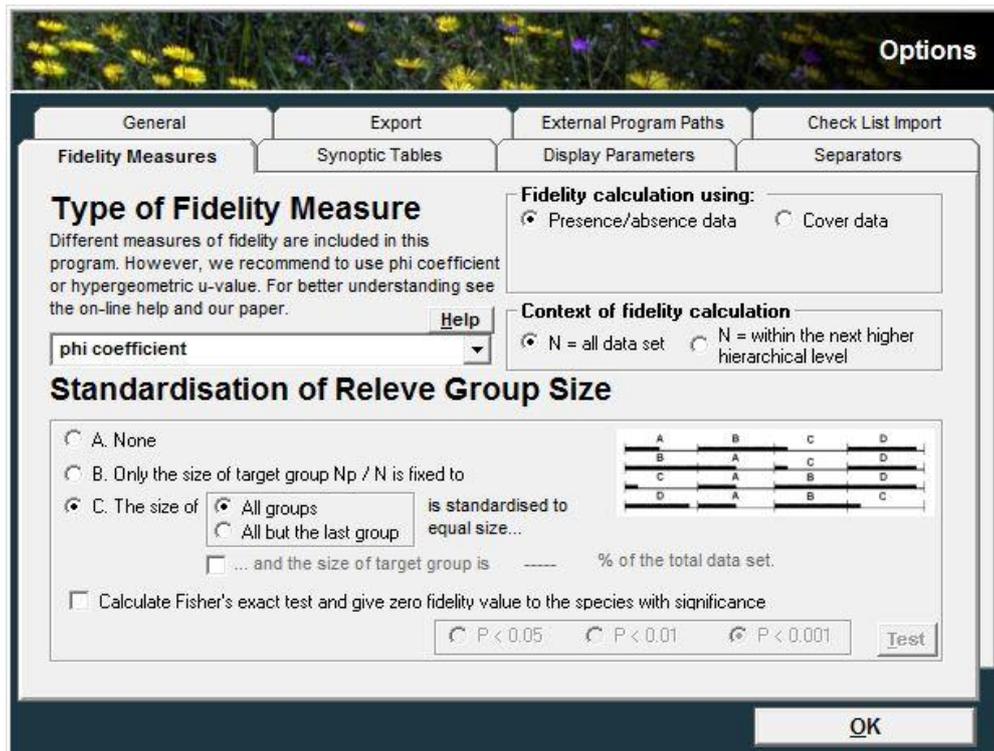


Fig. 5: Selecting a standardization method and ‘Presence/absence’ versus ‘Average cover’ from the ‘Fidelity measures’ tab of the ‘Options’ window.

Option A is to use the given data without standardization. Option B standardizes the size of the vegetation unit relative to the rest of the dataset. Option C has two steps: (1) it sets the target group to the indicated size and then (2) sets all the other groups (or all but the last) to be the same size as each other. The picture on the right side of the frame graphically represents the chosen standardization applied to a dataset divided into four vegetation units.

Warning: Because the Dufrene-Legendre Indicator Value implicitly standardizes the data set, it is only available with option C. A side-effect is that when the user scrolls through the list of measures, JUICE automatically switches to option C when ‘Dufrene and Legendre’ is highlighted. If a different standardization is desired (for a different measure) it must be re-selected manually. Similarly, highlighting ‘Dufrene and Legendre’ will cause JUICE to switch from ‘Average cover’ to ‘Presence/Absence data’.

Note 1: The data standardizations are available only for synoptic tables and related functions; thus they are only applicable to classified or partly classified datasets. All other functions using fidelity as a statistical measure (such as COCKTAIL classification and Interspecific associations) use the original data.

Note 2: You can test the standardization with real data. From a synoptic table with fidelities (see Section 2.2) open the ‘Fidelity measures’ tab of the ‘Options’ window. Changing the parameters will change the fidelity values in the synoptic table, and you will have a clear overview of what happens to your data after standardization.

Checking the Fisher’s exact test check box excludes phi coefficient fidelity values which are not statistically significant: they are assigned a fidelity of 0. The level of significance can be chosen.

2.1.4. Fidelity Tests

This function is included for users who would like to test fidelity measures with theoretical values. From the 'Help' menu select 'Fidelity Measure Test'. This opens the following window:

**Test Presence/Absence
Fidelity Measure Yourself**

N (no. of relevés in the dataset): 1000

N_p (no. of relevés in the vegetation unit): 100

n (no. of species occurrences in the dataset): 50

n_p (no. of species occurrences in the vegetation unit): 25

Fidelity value: 30.589

The value will be copied into the clipboard.

Recalculation **Close**

Fig. 6: Window for testing fidelity measures.

The fidelity measures in JUICE depend on four parameters:

- N ... number of relevés in the data set,
- N_p ... number of relevés in the target vegetation unit,
- n ... number of occurrences of the species in the data set,
- n_p ... number of occurrences of the species in the target vegetation unit.

The program uses the fidelity measure that has been selected from the 'Fidelity measures' tab of the 'Options' window. (See Section 2.1.3.) The current measure is displayed in the 'Test Fidelity Measure Yourself' window and on the Option Button on the Icon Bar.

Note 1: Standardization methods for the phi coefficient, available from the 'Fidelity measures' tab, do not affect this calculation.

Note 2: The fidelity value is copied onto the clipboard and it can be pasted into other programs.

2.1.5. Quantitative Fidelity Measures

JUICE (versions 6.3.66 and higher) can calculate fidelities not only from presence/absence data but also from quantitative cover values. For quantitative calculation, variables are defined as:

- n ... sum of all covers of the species in the data set,
- n_p ... sum of all covers of the species in the target vegetation unit.

Covers range in the interval [0,1].

The fidelity measure calculations are analogous to those described nad in Section 2.1.1.

The choice between presence/absence and quantitative fidelity measurement is available from the 'Options' window. (See Fig. 5.)

Note: Quantitative fidelity measures are only available for synoptic tables. Functions 'COCKTAIL Groups' and 'Interspecific Associations' are calculated using presence/absence data.

2.2. Synoptic Tables

Synoptic tables summarise the results of any classification of relevés. Therefore, they are one of the basic results of all phytosociological analyses and studies. They give an overview of classified vegetation units in the data set and help the researcher understand sophisticated relations among species in context with environment. Synoptic tables can contain several types of information which can be used for additional sorting of species and analysis of diagnostic, constant and dominant species. JUICE supports many different types of synoptic tables, which can be used separately or in combination.

The columns of the synoptic tables are defined by the separators in the standard table. Separators are placed or removed by holding down the Shift key and clicking on a relevé with the left mouse button. (See Section **Chyba! Nenalezen zdroj odkazů.** for more details.) To display a synoptic table, select the type of table from the ‘Synoptic Table’ menu. To return to Standard Table Display, select that type of table again.

2.2.1. Synoptic Table Display

The Synoptic Table Display has the same three parts as the Standard Table Display (species, short headers and synoptic table data – see Section **Chyba! Nenalezen zdroj odkazů.**). Mouse functions in the synoptic table are described in Section **Chyba! Nenalezen zdroj odkazů.**. Several menu items are not available in Synoptic Table Display mode.

Note: Columns in the synoptic table can be moved by clicking and dragging just as you do with relevés in Standard Table Display mode. This gives a way to move a block of relevés without using colours. Simply delimit the block with separators, switch to a synoptic table, move the column, and switch back.

Values that exceed a certain threshold are highlighted. These thresholds can be set in the ‘Synoptic tables’ tab of the ‘Options’ window. It is also possible to select the highlighting colours as well as to turn the highlighting feature on or off.

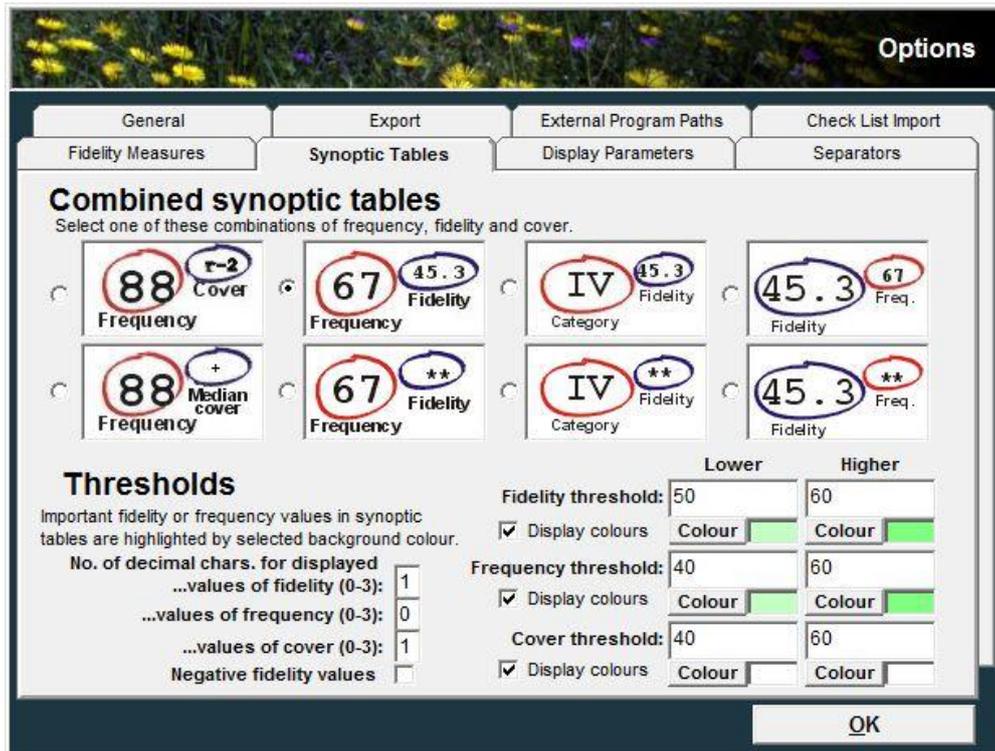


Fig. 7: The 'Synoptic tables' tab of the 'Options' window.

For more sophisticated synoptic table graphic display, two thresholds with different colours can be defined. Values greater than the 'Lower' threshold will be given one colour and values which exceed the 'Higher' threshold will be given the other colour. For single-colour highlighting, the 'Lower' threshold can be set to a value equal to or greater than the 'Higher' threshold. Colours can be completely switched off with the 'Display colours' check box.

Note: If a value is entered that causes the 'Lower' threshold to be greater than the 'Higher' threshold, the 'Higher' threshold is changed automatically. This means that it may be impossible to lower the 'Higher' threshold without first lowering the 'Lower' threshold.

2.2.2. Percentage Constancy Synoptic Table

This type of synoptic table shows percentage constancy of species for each relevé group. This function considers only presence/absence of the species without discrimination of cover values. The value of percentage constancy is rounded to the nearest natural number in the interval from 0 to 100, except that all non-zero values less than 1 are converted to the number 1.

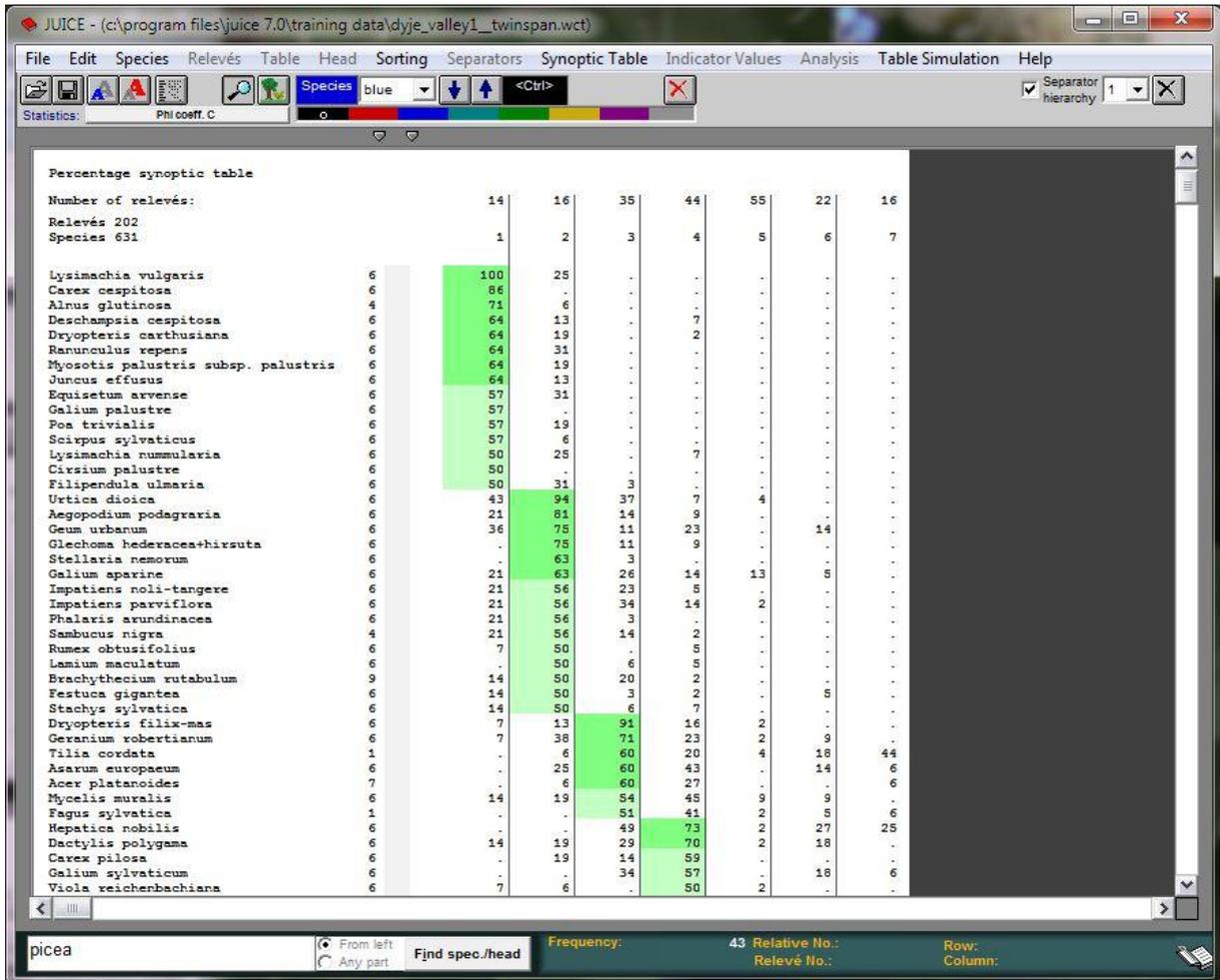


Fig. 8: Synoptic table with percentage constancy.

2.2.3. Categorical Synoptic Table

Constancy can also be displayed in a categorical form (Westhoff & Whittaker 1980). There are five constancy categories widely used in phytosociological literature: I (0 %-20 %); II (20 %-40 %); III (40 %-60 %); IV (60 %-80 %); V (80 %-100 %). Such synoptic tables are clearer and all columns are easily comparable. However, differences between categories are hard to judge. The differences between two consecutive categories could be anywhere from 1% to 39%. Moreover, species in the same category could have percentages differing by as much as 19%, so a given species may have constancy more similar to species in a different category compared to species in its own category.

Statistics: Phi coeff. C

Species: blue

Separator hierarchy: 1

Absolute frequency synoptic table

Number of relevés:	14	16	35	44	55	22	16
Relevés 202	1	2	3	4	5	6	7
Species 631	1	2	3	4	5	6	7
<i>Lysimachia vulgaris</i>	6	14	4
<i>Carex cespitosa</i>	6	12
<i>Alnus glutinosa</i>	4	10	1
<i>Deschampsia cespitosa</i>	6	9	2	.	3	.	.
<i>Juncus effusus</i>	6	9	2
<i>Myosotis palustris subsp. palustris</i>	6	9	3	.	1	.	.
<i>Dryopteris carthusiana</i>	6	9	3
<i>Ranunculus repens</i>	6	9	5
<i>Galium palustre</i>	6	8
<i>Poa trivialis</i>	6	8
<i>Equisetum arvense</i>	6	8	5
<i>Scirpus sylvaticus</i>	6	8	1
<i>Lysimachia nummularia</i>	6	7	4	.	3	.	.
<i>Filipendula ulmaria</i>	6	7	5	1	.	.	.
<i>Cirsium palustre</i>	6	7
<i>Urtica dioica</i>	6	6	15	13	3	2	.
<i>Aegopodium podagraria</i>	6	3	13	5	4	.	.
<i>Geum urbanum</i>	6	5	12	4	10	.	3
<i>Glechoma hederacea+hirsuta</i>	6	.	12	4	4	.	.
<i>Stellaria nemorum</i>	6	.	10	1	.	.	.
<i>Galium aparine</i>	6	3	10	9	6	7	1
<i>Impatiens parviflora</i>	6	3	9	12	6	1	.
<i>Phalaris arundinacea</i>	6	3	9	1	.	.	.
<i>Impatiens noli-tangere</i>	6	3	9	8	2	.	.
<i>Sambucus nigra</i>	4	3	9	5	1	.	.
<i>Lamium maculatum</i>	6	.	8	2	2	.	.
<i>Festuca gigantea</i>	6	2	8	1	1	.	1
<i>Brachythecium rutabulum</i>	9	2	8	7	1	.	.
<i>Rumex obtusifolius</i>	6	1	8	.	2	.	.
<i>Stachys sylvatica</i>	6	2	8	2	3	.	.
<i>Dryopteris filix-mas</i>	6	1	2	32	7	1	.
<i>Geranium robertianum</i>	6	1	6	25	10	1	2
<i>Acer platanoides</i>	7	.	1	21	12	.	1
<i>Tilia cordata</i>	1	.	1	21	9	2	4
<i>Asarum europaeum</i>	6	.	4	21	19	3	1
<i>Mycelis muralis</i>	6	2	3	19	20	5	2
<i>Fagus sylvatica</i>	1	.	.	18	18	1	1
<i>Hepatica nobilis</i>	6	.	.	17	32	1	6
<i>Dactylis polygama</i>	6	2	3	10	31	1	4
<i>Carex pilosa</i>	6	.	3	5	26	.	4
<i>Galium sylvaticum</i>	6	.	.	12	25	.	1
<i>Carpinus betulus</i>	7	2	.	3	22	10	5

Search: picea

From left: Any part

Find spec./head

Frequency: 43

Relative No.: Relevé No.:

Row: Column:

Fig. 10: Synoptic table with absolute constancy.

2.2.5. Fidelity Synoptic Table

This table uses the fidelity concept as described in [Section 2](#). The fidelity between the species and the relevé group is displayed. Highlighting thresholds are set in the 'Fidelity threshold' section of the 'Synoptic tables' tab of the 'Options' menu. (See Section 2.2.1.) Fidelity can be calculated either from presence/absence or from quantitative cover data. (See Section 2.1.5.)

Percentage synoptic table with fidelity (Phi coeff. C) (7 columns)

	14	16	35	44	55	22	16
Number of relevés:	14	16	35	44	55	22	16
Relevés 202							
Species 631	1	2	3	4	5	6	7
Carex cespitosa	91.5	---	---	---	---	---	---
Lysimachia vulgaris	87.6	---	---	---	---	---	---
Alnus glutinosa	78.4	---	---	---	---	---	---
Galium palustre	73.0	---	---	---	---	---	---
Juncus effusus	69.7	---	---	---	---	---	---
Scirpus sylvaticus	68.4	---	---	---	---	---	---
Cirsium palustre	67.9	---	---	---	---	---	---
Myosotis palustris subsp. palustris	66.2	---	---	---	---	---	---
Deschampsia cespitosa	65.9	---	---	---	---	---	---
Dryopteris carthusiana	65.0	---	---	---	---	---	---
Lythrum salicaria	62.6	---	---	---	---	---	---
Poa trivialis	60.8	---	---	---	---	---	---
Ranunculus repens	60.2	---	---	---	---	---	---
Caltha palustris	57.3	---	---	---	---	---	---
Ranunculus flammula	56.8	---	---	---	---	---	---
Crepis paludosa	56.8	---	---	---	---	---	---
Equisetum arvense	54.7	---	---	---	---	---	---
Mentha aquatica	50.5	---	---	---	---	---	---
Galium uliginosum	50.5	---	---	---	---	---	---
Lycopodium europaeus	50.5	---	---	---	---	---	---
Stellaria nemorum	---	74.6	---	---	---	---	---
Glechoma hederacea+hirsuta	---	73.0	---	---	---	---	---
Aegopodium podagraria	---	67.2	---	---	---	---	---
Urtica dioica	---	62.5	---	---	---	---	---
Lamium maculatum	---	60.2	---	---	---	---	---
Rumex obtusifolius	---	59.3	---	---	---	---	---
Phalaris arundinacea	---	57.2	---	---	---	---	---
Chaerophyllum aromaticum	---	56.1	---	---	---	---	---
Festuca gigantea	---	52.4	---	---	---	---	---
Sambucus nigra	---	51.2	---	---	---	---	---
Geum urbanum	---	51.1	---	---	---	---	---
Stachys sylvatica	---	51.0	---	---	---	---	---
Dryopteris filix-mas	---	---	76.9	---	---	---	---
Acer platanoides	---	---	53.4	---	---	---	---
Senecio germanicus	---	---	52.2	---	---	---	---
Acer platanoides	---	---	51.2	---	---	---	---
Actaea spicata	---	---	50.8	---	---	---	---
Viola reichenbachiana	---	---	---	57.1	---	---	---
Carex pilosa	---	---	---	55.5	---	---	---
Avenella flexuosa	---	---	---	---	61.7	---	---
Ceratodon purpureus	---	---	---	---	59.0	---	---
Hypogymnia physodes	---	---	---	---	55.8	---	---

Fig 11: Synoptic table with fidelity values.

Note: Negative fidelities are usually less useful for table explanation. Therefore, the program displays all values of zero or less as dashes.

2.2.6. Synoptic Tables and Cover (Maximum, Average, Median, Modus)

A synoptic table can also be used for presenting cover values. These are selected from the 'Cover Values' submenu of the 'Synoptic Table' menu. As an example, suppose a species has the following cover values in a group of 11 relevés (percentage number with cover code):

1(r) 1(r) 1(r) 2(+) 3(1) 4(2m) 8(2a) 8(2a) 18(2b) 38(3) 63(4).

The maximum cover for this species will be 63, the average will be 13 (rounded), the median will be 4 and the modus (the most frequent value) will be 1.

These statistics only consider relevés in which the species is present. To find the average cover over all relevés in the group, use Barkman's Total Cover (Barkman 1989) as described in Section 2.2.8 pod.

Note: Cover values in a synoptic table are always displayed as percentage numbers.

2.2.7. Zlatnik's Combined Scale

Another type of table available from the 'Cover Values' submenu is 'Zlatnik's Combined Value' (Plíva & Průša 1969). Czech foresters use a special synoptic table in which values are defined according to this table:

		Frequency			
		II	III	IV	V
A	+	1	2	4	6
b	1	3	8	13	18
u	2	9	22	37	58
n	3	15	37	62	87
d	4	21	52	87	122
.	5	27	67	112	157

Fig 12: Information window with Zlatnik’s scale.

These values combine species relative frequencies with maximum cover. To see this table in JUICE, open the ‘Synoptic Table’ menu, select ‘Sort Species In Synoptic Tables’ and press the ‘?’ button next to ‘Zlatnik’s value’.

2.2.8. Average Cover (Barkman’s Total Cover Value)

The average cover (AC) within a relevé group as defined by Barkman (1989) is:

$$AC = \text{SUM}(\% \text{ Cover}) / \text{number of relevés.}$$

Note: The ‘number of relevés’ includes those with a cover value of zero (i.e., where the species is absent).

Species	14	16	35	44	55	22	16
Molinia arundinaceae-caerulea	6	22.5	0.1	.	1.2	0.6	.
Carex cespitosa	6	19.4
Urtica dioica	6	1.7	19.2	3.2	0.3	0.1	.
Sambucus nigra	4	1.4	12.3	2.0	0.3	.	.
Agropodium podagraria	6	0.5	10.4	0.4	0.4	.	.
Hypnum cupressiforme	9	0.3	0.1	16.3	3.4	8.8	6.8
Tilia cordata	1	.	0.2	14.4	1.5	0.1	1.0
Tilia platyphyllos	1	.	.	12.3	0.9	0.2	0.7
Carex pilosa	6	.	0.3	0.4	16.2	.	.
Avenella flexuosa	6	.	.	0.1	1.3	12.4	2.4
Poa nemoralis	6	0.3	0.6	4.1	5.5	4.5	13.2
Sesleria caerulea	6	.	.	.	0.0	.	13.9
Cornus mas	4	.	.	.	2.5	2.0	12.8
Pinus sylvestris	1	0.1	.	.	0.6	9.6	0.2
Alnus glutinosa	1	73.0	.	2.0	0.3	.	.
Fagus sylvatica	1	.	65.2	23.0	25.5	0.1	0.1
Carpinus betulus	1	.	2.6	9.1	40.6	1.3	12.1
Festuca ovina subsp. ovina	6	11.1	13.8
Quercus petraea agg.	1	.	.	0.3	10.9	49.1	65.0
Brachypodium pinnatum	6	2.6
Quercus robur	1	.	.	.	1.5	1.9	1.7
Carex humilis	6	1.5	3.3
Inula ensifolia	6	7.2
Anthericum ramosum	6	.	.	.	0.0	0.7	1.4
Sorbus 'axia agg.'	4	.	.	0.1	0.0	0.5	0.1
Bupleurum falcatum	6	.	.	.	0.2	0.2	1.8
Corylus avellana	4	.	0.3	1.6	0.2	0.4	0.2
Teucrium chamaedrys	6	0.3	1.2
Vincetoxicum hirsutinavia	6	.	.	.	0.2	1.3	3.6
Forsetia tostosa	9	.	.	.	0.1	.	2.8
Euphorbia cyparissias	6	0.4	1.0
Ligustrum vulgare	4	0.1	.	.	.	0.4	2.1
Dactylis glomerat-polygama	6	0.1	1.7
Polygonatum odoratum	6	.	.	0.1	0.3	0.7	1.2
Tilia cordata	4	.	0.6	1.7	0.7	0.3	0.5
Viburnum lantana	4	.	.	.	0.2	.	0.1
Berberis vulgaris	4	.	.	.	0.0	.	0.1
Aster amellus	6	0.1
Homalothecium lutescens	9	.	.	.	0.0	.	1.6
Galium glaucum	6	0.2	0.5
Libanotis pyrenaica	6	0.2
Carex flacca	6	1.4

Fig. 13: Synoptic table with average covers.

2.2.9. Combined Synoptic Tables

Standard synoptic tables contain only one important aspect of the data set while others are not presented. Information such as frequency, fidelity and cover can be presented in multiple tables, but this takes up space and it can be difficult to see relationships between the tables. JUICE can display two values in a combined synoptic table, thus overcoming these disadvantages. Selecting ‘Combined Synoptic Table’ from the ‘Synoptic Table’ menu displays a combined synoptic table and opens the ‘Synoptic tables’ tab of the ‘Options’ window:

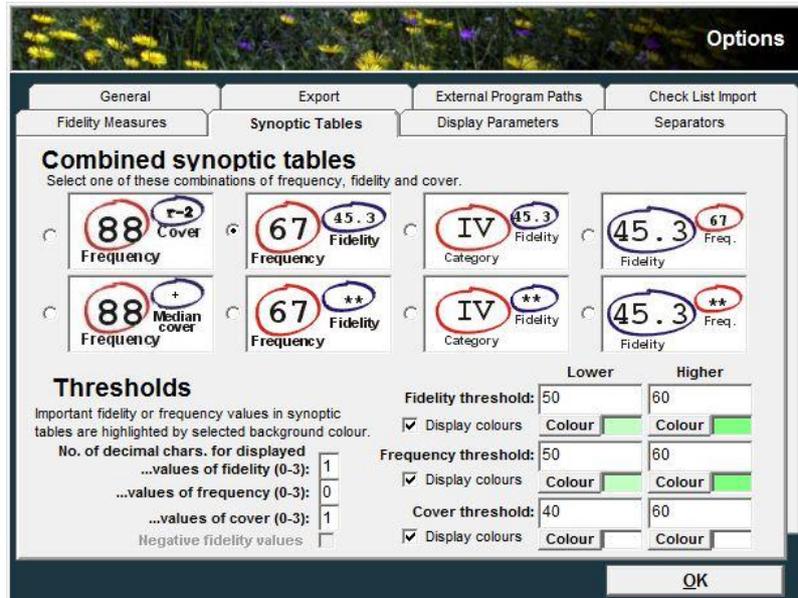


Fig. 14: Selecting a combined synoptic table from the ‘Synoptic tables’ tab of the ‘Options’ window.

From this tab, one of eight combinations of frequency, fidelity, category and cover can be displayed.

The highlighting thresholds for frequency and fidelity (see Section 2.2.1) are applied to their own columns in the table. For example, if the frequency and fidelity combination is selected, the combined table will look like this:

Synoptic table with percentage frequency and modified fidelity index {phi coefficient} (7 columns) - F

Number of relevés: 14 16 35 44 55 22

relevés 202

Species 631

	1	2	3	4	5	6
Lysimachia vulgaris	6	100 ^{87.6}	25	.	.	.
Carex cespitosa	6	86 ^{81.5}
Alnus glutinosa	4	71 ^{78.4}	6	.	.	.
Ranunculus repens	6	64 ^{60.2}	31	.	.	.
Juncus effusus	6	64 ^{59.7}	12	.	.	.
Dryopteris carthusiana	6	64 ^{55.0}	19	.	2	.
Deschampsia cespitosa	6	64 ^{55.9}	12	.	7	.
Myosotis palustris subsp. palustris	6	64 ^{56.2}	19	.	.	.
Galium palustre	6	57 ^{75.0}
Poa trivialis	6	57 ^{60.5}	19	.	.	.
Equisetum arvense	6	57 ^{54.7}	31	.	.	.
Scirpus sylvaticus	6	57 ^{68.4}	6	.	.	.
Filipendula ulmaria	6	50 ^{47.7}	31	.	.	.
Cirsium palustre	6	50 ^{57.9}	.	3	.	.
Lysimachia nummularia	6	50 ^{48.7}	25	.	7	.
Urtica dioica	6	43	94 ^{62.5}	37	7	4
Aegopodium podagraria	6	21	81 ^{67.2}	14	9	.
Geum urbanum	6	36	75 ^{51.1}	11	23	14
Glechoma hederacea+hirsuta	6	.	75 ^{75.0}	11	9	.
Stellaria nemorum	6	.	62 ^{74.6}	3	.	.
Galium aparine	6	21	62 ^{43.2}	26	14	5
Phalaris arundinacea	6	21	56 ^{57.2}	3	.	.
Impatiens noli-tangere	6	21	56 ^{47.1}	23	5	.
Impatiens parviflora	6	21	56 ^{40.3}	34	14	2
Sambucus nigra	4	21	56 ^{51.2}	14	2	.
Stachys sylvatica	6	14	50 ^{51.0}	6	7	.
Festuca gigantea	6	14	50 ^{52.4}	3	2	5
Lamium maculatum	6	.	50 ^{60.2}	6	5	.
Rumex obtusifolius	6	7	50 ^{39.3}	.	5	.
Brachythecium rutabulum	9	14	50 ^{45.7}	20	2	.
Dryopteris filix-mas	6	7	12	91 ^{76.9}	16	2

21. Quercus robur [1] Frequency: 14 Relative No.: Relevé No. Row: 52 Column: 52

Fig. 15: Combined synoptic table with frequency and fidelity.

These tables can be exported into the current RTF export file as described in Section **Chyba! Nenalezen zdroj odkazů.**

Note: In a combined synoptic table, columns of relevés cannot be moved. Other mouse functions are the same as in regular synoptic tables.

2.2.10. Synoptic Table Export

From the 'File' menu select 'Export' and 'Synoptic Table'. For more information, see Section **Chyba! Nenalezen zdroj odkazů.**

2.2.11. Sorting Species in Synoptic Tables

The 'Sort Species In Synoptic Table' function is available from the 'Sorting' menu or the 'Synoptic Table' menu. It is only active during Synoptic Table Display mode. Selecting this function opens the following window:

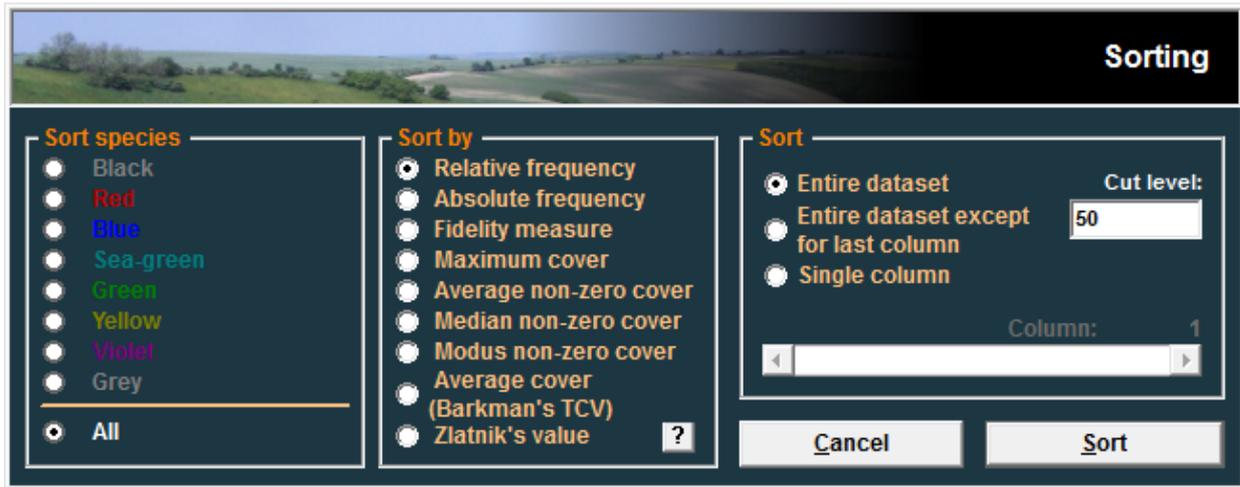


Fig. 16: The synoptic table 'Sorting' window.

Species can be sorted according to several different measures. The user has the option to sort the 'Entire dataset', the 'Entire dataset without last column' or a 'Single column'. The window contains a slide bar for selecting which column will be sorted by the 'Single Column' sort. The function can be restricted to sort only species of a certain colour.

The 'Entire dataset' sort is a cumulative sorting function:

1. When this function is called, the user is prompted to define a sorting cut-off level. The default cut-off level for relative frequency and fidelity will be the same as the 'Lower' threshold defined in the 'Synoptic tables' tab of the 'Options' window. (See Section 2.2.1.) Furthermore, changing this cut-off level for sorting changes the 'Lower' threshold for highlighting.
2. Once the cut-off level is defined, the first column is sorted.
3. Species with values higher than the cut-off level are held fixed at the top of the table. The remaining species are sorted in the second column.
4. The process repeats for each column until the entire table is sorted.

The sorting function 'Entire dataset without last column' sorts the table similarly except that the last column is not sorted.

Note: The actual sorting algorithm is more complicated than that described above. Species that have values above the cut-off level in two columns are moved below the species that exceed the cut-off level in only one column. These are followed by species that have values above the cut-off level in three columns, four columns and so on. The result will be a table with uniquely diagnostic (or constant or dominant) species in blocks at the top, followed by blocks of species which are diagnostic for multiple groups.

2.2.12. Analysis of Synoptic Columns Combining with Exporting the Results

The program can extract diagnostic, constant and dominant species from each synoptic column. This function is widely used for analysis of classification results and interpretation of vegetation clusters.

Analysis of Constancy Columns in Synoptic Table

Fidelity threshold: Frequency threshold (1-100 %): Cover threshold (1-100 %):

Diagnostic species	Constant species	Dominant species
91.5 Carex cespitosa [6]	57 Equisetum arvense [6]	93 Alnus glutinosa [1]
87.6 Lysimachia vulgaris [6]	50 Lysimachia nummularia [6]	21 Molinia arundinacea+caerulea [6]
78.4 Alnus glutinosa [4]	50 Filipendula ulmaria [6]	14 Carex cespitosa [6]
73.0 Galium palustre [6]	43 Urtica dioica [6]	7 Scirpus sylvaticus [6]
69.7 Juncus effusus [6]	43 Cardamine amara [6]	
68.4 Scirpus sylvaticus [6]	43 Caltha palustris [6]	
67.9 Cirsium palustre [6]	43 Athyrium filix-femina [6]	
66.2 Myosotis palustris subsp. palustris [6]	43 Anemone nemorosa [6]	
65.9 Deschampsia cespitosa [6]		
65.0 Dryopteris carthusiana [6]		
62.8 Alnus glutinosa [1]		
62.6 Lythrum salicaria [6]		
60.8 Poa trivialis [6]		
60.2 Ranunculus repens [6]		

Include diagnostic species
 Show references

Minimum freq. (1-100 %):

Mark in the table: Mark in the table: Mark in the table:

Select a list of species to be marked in an appropriate list window first...

No. of relevés: 0
Aver. species No.: 31.64
Aver. positive fidelity: 13.53
Sharpness: 31.00

Column: 1

Fig. 17: Window for detailed analysis of synoptic tables.

From the 'Synoptic Table' menu, select 'Analysis Of Synopt. Columns'. Initially, the three lists will appear blank. Select the column to analyse (using the slide bar at the bottom) and then press the 'Refresh' button. Three species lists will appear:

'Diagnostic species' are those with fidelity higher than the 'Lower' fidelity threshold (defined in the 'Synoptic tables' tab of the 'Options' menu – see Section 2.2.1).

'Constant species' are those with relative frequency higher than the 'Lower' frequency threshold (also defined in the 'Synoptic tables' tab). However, only constant species that are not diagnostic will be included in the list unless the 'Include diagnostic species' check box is checked.

'Dominant species' are defined as all species that have cover values higher than the cover threshold. A species appears in this list if its cover exceeds the threshold value in any relevé of the selected vegetation unit. The list of dominant species can be limited to species that exceed a certain minimum relative frequency as specified in the 'Minimum freq.' box.

Species are displayed with layer number and fidelity or frequency value. If the 'Show references' check box is checked, species that appear in more than one list will be cross-indexed to the other list or lists ('Dg' for diagnostic, 'C' for constant and 'Dm' for dominant).

Note 1: The display is not updated when a display mode check box or a threshold value is changed. To enact the desired display changes, press the 'Refresh' button.

Note 2: Changing the 'Fidelity threshold' or the 'Frequency threshold' changes the threshold throughout the program, just as though it were changed in the 'Synoptic tables' tab of the 'Options' window.

Species in the lists can be selected with the mouse or keyboard and assigned a selected colour in the table by pressing the 'Mark in the table' button.

Note: Shift+click can be used to select a block of species. In particular, to select an entire list, click on the top species and then Shift+click on the bottom species. Ctrl+click can be used to select multiple species which are not in a block or to deselect a currently selected species.

All information can be saved into the current rich text format export file. (See Section **Chyba! Nenalezen zdroj odkazů.**) Pressing the ‘Export’ button opens the following window:

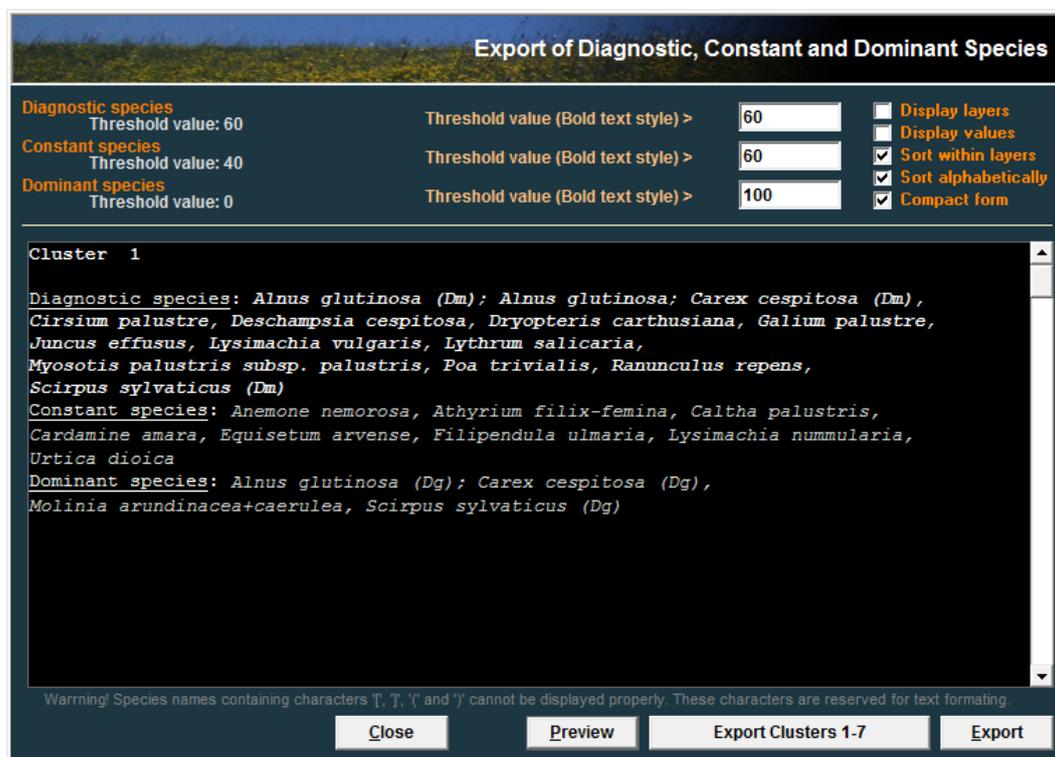


Fig. 18: Window for ‘Export Of Diagnostic, Constant And Dominant Species’.

The list can be formatted as unsorted, sorted alphabetically or within layers, with or without frequencies or fidelities, with or without layers and in compact or extended form. The ‘Lower’ threshold values are fixed – they have been defined in the previous window. The ‘Higher’ threshold values (‘Bold text style’) correspond to the values set in the ‘Synoptic tables’ tab. They may be reset to any value equal to or higher than the lower threshold. (Changing the ‘Bold text style’ value for diagnostic or constant species in this window has the same effect as changing the ‘Higher’ fidelity or frequency threshold value in the ‘Synoptic tables’ tab.) Species exceeding these values will be displayed in bold style. The sample species list in the lower part of the window will not change until the user presses the ‘Preview’ button.

The user can export information for a single cluster (column) with the ‘Export’ button. The information for the entire data set can be exported by pressing the ‘Export Clusters’ button. All data are saved into the current RTF export file. (See Section **Chyba! Nenalezen zdroj odkazů.**)

2.2.13. Uniqueness of Vegetation Unit

Uniqueness was used in the paper of Chytrý & Tichý (2003) to identify unique vegetation units in the data set. It expresses whether there are similar vegetation units of the same rank (e.g., class or alliance). A vegetation unit is unique if none of its diagnostic species has simultaneously diagnostic status in other vegetation units, while its uniqueness decreases if it shares its diagnostic species with other vegetation units.

2.2.13.1 Theoretical Background

Uniqueness is calculated in two steps. First, an Asymmetric Similarity Index between every pair of vegetation units j and k is calculated as

$$T_{jk} = \frac{\sum_i \Phi_{ij} \cdot \Phi_{ik}}{\sum_i \Phi_{ij}^2}, \quad (\text{Eq. 4})$$

where Φ_{ij} is the fidelity of species i to vegetation unit j , Φ_{ik} is the fidelity of species i to vegetation unit k and the sums only include species with $\Phi_{ij} > 0.05$ and $\Phi_{ik} > 0.05$. This index ranges in the interval $[0, \infty)$ and yields the highest values for those pairs of vegetation units in which a large proportion of diagnostic species of vegetation unit j are also diagnostic species of vegetation unit k . It is an asymmetric similarity measure because in cases of low numbers of diagnostic species in vegetation unit j and high numbers in k , where most of the diagnostic species of j are shared with k , the similarity of j to k is high while the similarity of k to j is moderate. The Similarity Index is calculated separately for all pairs of classes and all pairs of alliances (suballiances). For every pair j and k , two indices are calculated, one for the similarity of j to k and the other for the similarity of k to j .

Second, the Uniqueness Index for each vegetation unit is calculated using this formula:

$$U_j = \frac{1}{\sum_k T_{jk}}. \quad (\text{Eq. 5})$$

This Uniqueness Index is low for those vegetation units whose diagnostic species are mostly shared with other vegetation units.

2.2.13.2 Uniqueness in JUICE

The ‘Uniqueness’ function is available for synoptic tables with fidelities. After selecting ‘Uniqueness’ from the ‘Synoptic Tables’ menu, the following window will appear:

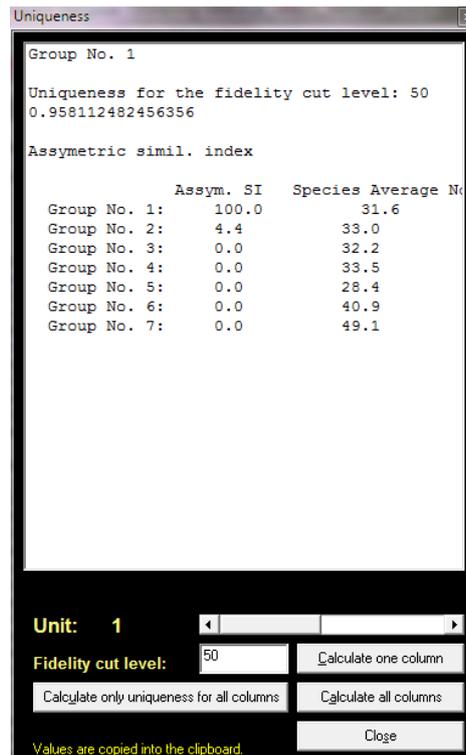


Fig. 19: Uniqueness and Asymmetric Similarity Indices calculation.

The list of species compared is limited to those with fidelities higher than the defined threshold value. This value depends on the selected fidelity measure and method of data standardization. The vegetation unit can be selected with the scroll bar. After pressing the 'Calculate one column' button, the calculated Uniqueness Index and a list of Asymmetric Similarity Indices are displayed. This information is also saved onto the clipboard.

Note: The 'Calculate only uniqueness for all columns' and 'Calculate all columns' buttons are for copying the indicated information to the clipboard (from which it can be pasted into another program). The information actually displayed in the window will be the full information for the highest-numbered relevé group.

2.2.14. Average Values of Columns in Synoptic Table

The 'Average Values Of Constancy Columns' function of the 'Synoptic Tables' menu calculates the average for each synoptic table column. The displayed values (relative frequency, fidelity, average cover etc.) are averaged and these averages are displayed in a separate window:

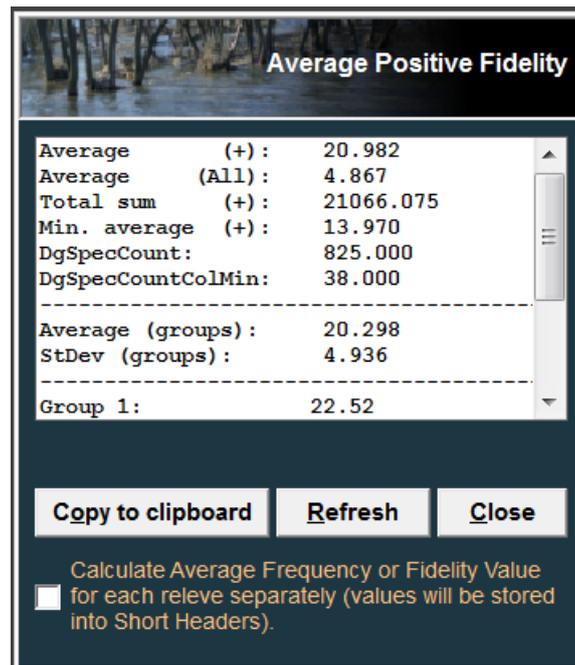


Fig. 20: Average values of constancy columns.

The results can be saved onto the clipboard for further use.

2.2.15. Crispness of Classification

Classification methods usually produce results with hierarchically distributed clusters. Sometimes it is difficult to determine an optimal number of clusters providing the highest ‘separation power’ for species. JUICE includes a function, proposed in the paper of Botta-Dukát et al. (2005), for identifying the optimal number of clusters.

2.2.15.1 Theoretical Background

The calculation is based on G statistics, which can be easily calculated for contingency tables of any size (Sokal & Rohlf 1995). Calculated for a $2 \times c$ contingency table (where c is the number of clusters), the G statistic does not measure the fidelity of a species to individual clusters, but the species’s capacity to distinguish the clusters within a given partition (the ‘separation power’ of the species). The average of separation powers is called the ‘crispness of the classification’. The higher the average separation power, the better the clusters can be distinguished by the diagnostic species, i.e., the better the classification.

The problem is that the expected separation power and consequently the crispness of classification increase with increasing number of clusters, even if relevés are assigned randomly to clusters. This effect has to be eliminated before comparing partitions with different numbers of clusters. In a random classification, separation powers have approximately a chi-squared distribution with $c-1$ degrees of freedom (Sokal & Rohlf 1995). According to the central limit theorem (e.g., Zar 1999: 76–77), if the number of species is high, which is the case in most data sets, the crispness of classification has approximately a normal distribution with expected value $c-1$ and standard deviation $\sqrt{2(c-1)/S}$, where S is the number of species.

The effect of the number of clusters is removed by subtracting this expected value and dividing the difference between the observed and expected value by the standard deviation. In this way the crispness values are standardized and can be compared among partitions with different numbers of clusters. Local maxima of crispness may indicate optimal numbers of clusters.

2.2.15.2 Optimal Use of This Function in JUICE

This function is used to its full power in conjunction with the function ‘Sort Relevés by Classifications by Others Programs’ (see Section **Chyba! Nenalezen zdroj odkazů.**):

1. Call the function ‘Sort Relevés by classification by Others Programs’ Open the resulting file produced by PC-ORD (or MULVA) or classify the data set using the ‘Cluster Analysis (PC-ORD)’ function from the ‘Analysis’ menu. The sorting window will appear automatically after the analysis. This window is always displayed on top, but the table remains active.
2. From the ‘Synoptic Table’ menu, select a table with frequency or fidelity values. Repeated selection of different numbers of clusters will automatically change the displayed synoptic table.
3. Choose a species colour and use it to select a list of species as a source for calculation of crispness of classification.
4. Call the ‘Crispness Of Classification’ function from the ‘Synoptic Table’ menu. This window, which is also always on top of the main window, will appear:

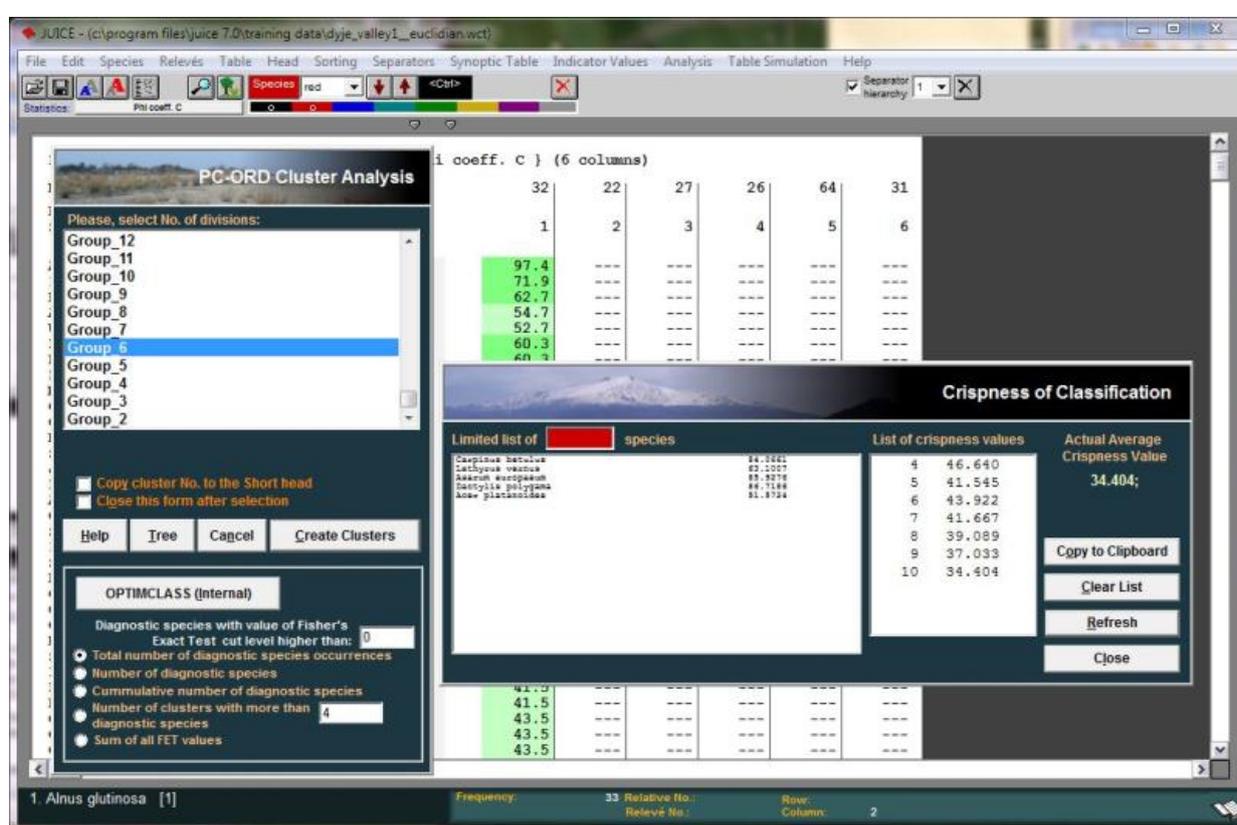


Fig. 21: The sorting function and the ‘Crispness Of Classification’ window overlaid on top of a synoptic table.

5. Press the ‘Refresh’ button. The program will calculate the appropriate value for the selected number of clusters.
6. In the ‘cluster analysis’ window, select a new number of clusters. The synoptic table will change. Press the ‘Refresh’ button in the ‘Crispness Of Classification’ window to calculate a new crispness value. (All values are saved onto the clipboard.)
7. Iterate these steps with increasing numbers of clusters.

Note 1: The program can calculate crispness of classification with a limited list of species. It is recommended, e.g., to exclude rare species (such as those occurring in fewer than 10 relevés) from this analysis.

Note 2: It is also possible to search for maximum crispness using hierarchical separators. (See Section **Chyba! Nenalezen zdroj odkazů.**) Varying separator level will change the number of columns in the synoptic table. (This can be useful, e.g., with results from TWINSpan classification.)

2.2.16. Comparison of Two Synoptic Tables

If the synoptic table is displayed in the form of percentage frequency or fidelity, it is possible to compare two synoptic tables and find their total similarity. This function was used in the paper Knollová et al. (2005) for comparison of several different classifications. The program calculates a crossing table of Euclidean distances between all pairs of synoptic columns in two synoptic tables.

Note: Calculation of fidelities can be modified by data standardization, Fisher's exact test and selection of presence/absence versus quantitative data as described in Section 2.1.3.

2.2.16.1 Theoretical Background

The distance $d(\text{TAB1}_i; \text{TAB2})$ between the i -th column of one table, TAB1, and all columns of a second table, TAB2, is calculated according to the formula:

$$d(\text{TAB1}_i; \text{TAB2}) = \frac{\min_j (ED_{ij}) \cdot (n_{\text{TAB2}} - 1)}{(\sum_j ED_{ij}) - \min_j (ED_{ij})}, \quad (\text{Eq. 6})$$

where i and j denote columns of classified resampled data sets TAB1 and TAB2, respectively; ED_{ij} is the Euclidean distance between phi coefficients in columns TAB1_i and TAB2_j ; $\min_j (ED_{ij})$ is the shortest distance between the column TAB1_i and any of the columns TAB2_j ; and n_{TAB2} is the number of columns in the table TAB2. In this way we obtain the distance between the selected column of table TAB1 and the most similar column of table TAB2, divided by the average distance between the selected column and all other columns of TAB2. The distance between the tables TAB1 and TAB2 is computed by averaging the distance values for individual columns table of TAB1:

$$d(\text{TAB1}; \text{TAB2}) = \frac{\sum_i d(\text{TAB1}_i; \text{TAB2})}{n_{\text{TAB1}}}. \quad (\text{Eq. 7})$$

This 'distance' is actually an asymmetric measure, i.e., the distances $d(\text{TAB1}; \text{TAB2})$ and $d(\text{TAB2}; \text{TAB1})$ differ. Therefore the same procedure is applied in the opposite direction:

$$d(\text{TAB2}_j; \text{TAB1}) = \frac{\min_i (ED_{ij}) \cdot (n_{\text{TAB1}} - 1)}{(\sum_i ED_{ij}) - \min_i (ED_{ij})}, \quad (\text{Eq. 8})$$

$$d(\text{TAB2}; \text{TAB1}) = \frac{\sum_j d(\text{TAB2}_j; \text{TAB1})}{n_{\text{TAB2}}}. \quad (\text{Eq. 9})$$

Finally the symmetric distance $D(\text{TAB2}; \text{TAB1})$ between tables TAB2 and TAB1 is calculated as an average of $d(\text{TAB1}; \text{TAB2})$ and $d(\text{TAB2}; \text{TAB1})$:

$$D(\text{TAB2}; \text{TAB1}) = \frac{d(\text{TAB1}; \text{TAB2}) + d(\text{TAB2}; \text{TAB1})}{2}. \quad (\text{Eq. 10})$$

This method can be used for calculating the symmetric distances between any two tables.

2.2.16.2 Comparing Two Synoptic Tables within JUICE

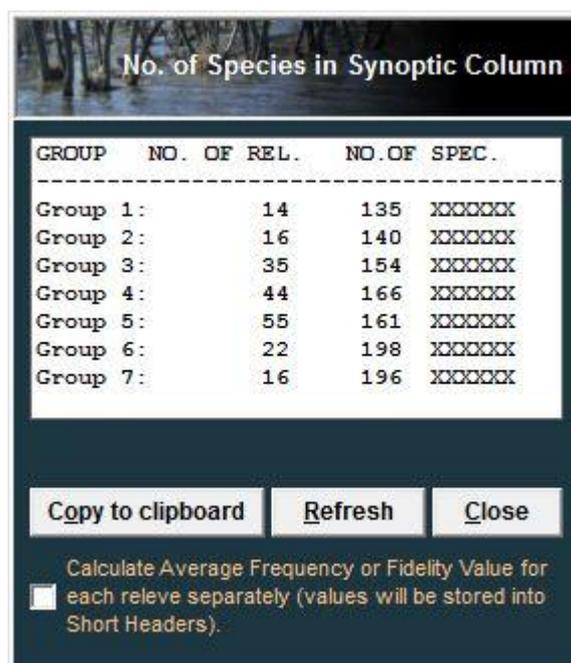
To use JUICE to compare two synoptic tables:

1. Load the first table. All records must be unique, without duplicity of either species name or layer. (Such species must be merged. See Section **Chyba! Nenalezen zdroj odkazů.**)
2. Classify the table and display it in the form of a synoptic table (using either frequency or fidelity).
3. From the 'Synoptic Table' menu, select 'Compare Two Synoptic Tables' and 'Save Image Of Synoptic Table' to save the table.
4. Load the second table and display it in the same form as the first one.
5. From the 'Synoptic Table' menu, select 'Compare Two Synoptic Tables' and 'Load Image And Compare'.

The program will display the resulting distance. A crossing table of all Euclidean distances is saved onto the clipboard.

2.2.17. Number of Species in Synoptic Table

The function 'Number of Species in Synoptic Table' in the 'Synoptic table' menu calculates a total number of species in the group of the relevés. The list of results can be copied into the clipboard.



GROUP	NO. OF REL.	NO. OF SPEC.
Group 1:	14	135
Group 2:	16	140
Group 3:	35	154
Group 4:	44	166
Group 5:	55	161
Group 6:	22	198
Group 7:	16	196

Calculate Average Frequency or Fidelity Value for each releve separately (values will be stored into Short Headers).

Fig. 22: Number of species in synoptic columns

2.2.18. Merge Two nearest columns

Sometimes user needs to find two the most similar columns and merged them. There are two possibilities: The first one calculates the similarity using an average fidelity or frequency. The second one compares Euclidian distance of columns.