Journal of Vegetation Science 14: 601-610, 2003
© IAVS; Opulus Press Uppsala.

Formalized reproduction of an expert-based phytosociological classification: A case study of subalpine tall-forb vegetation

Kočí, Martin[1,2]; Chytrý, Milan[1,*] & Tichý, Lubomír[1,3]

[1]Department of Botany, Masaryk University, Kotlářská 2, CZ-611 37 Brno, Czech Republic; [2] E-mail koci@sci.muni.cz;
[3]tichy@sci.muni.cz; [*]Corresponding author; Fax +420541211214; E-mail chytry@sci.muni.cz

Abstract. Delimitation of vegetation units in phytosociology is traditionally based on expert knowledge. Applications of expert-based classifications are often inconsistent because criteria for assigning relevés to vegetation units are seldom given explicitly. Still, there is, e.g. in nature conservation, an increasing need for a consistent application of vegetation classification using computer expert systems for unit identification.

We propose a procedure for formalized reproduction of an expert-based vegetation classification, which is applicable to large phytosociological data sets. This procedure combines Bruelheide's Cocktail method with a similarity-based assignment of relevés to constancy columns of a vegetation table. As a test of this method we attempt to reproduce the expert-based phytosociological classification of subalpine tall-forb vegetation of the Czech Republic which has been made by combination of expert judgement and stepwise numerical classification of 718 relevés by TWINSPAN. Applying the Cocktail method to a geographically stratified data set of 21 794 relevés of all Czech vegetation types, we defined groups of species with the statistical tendency of joint occurrences in vegetation. Combinations of 12 of these species groups by logical operators AND, OR and AND NOT yielded formal definitions of 14 of 16 associations which had been accepted in the expert-based classification. Application of these formal definitions to the original data set of 718 relevés resulted in an assignment of 376 relevés to the associations. This assignment agreed well with the original expert-based classification. Relevés that remained unassigned because they had not met the requirements of any of the formal definitions, were subsequently assigned to the associations by calculating similarity to relevé groups that had already been assigned to the associations. A new index, based on frequency and fidelity, was proposed for calculating similarity. The agreement with the expert-based classification achieved by the formal definitions was still improved after applying the similarity-based assignment. Results indicate that the expert-based classification can be successfully formalized and converted into a computer expert system.

Keywords: Braun-Blanquet approach; Cocktail; Czech Republic; Expert system; Matching; *Mulgedio-Aconitetea*; Species group; Vegetation survey.

Nomenclature: Kubát et al. (2002).

Abbreviations: *FDI* = Fidelity Index; *FFI* = Frequency-Fidelity Index; *FQI* = Frequency Index.

Introduction

In the last decade, phytosociological classification of vegetation (Westhoff & van der Maarel 1978) has been increasingly applied in nature conservation and landscape planning (Ostermann 1998). Nature managers are in urgent need of consistent systems of vegetation classification that would be valid over large areas and that would provide unequivocal criteria for assignment of vegetation stands to the classification units. An ideal tool for nature conservation authorities would be computer expert systems (Noble 1987) for identification of vegetation units.

Traditional expert-based vegetation classifications, which are widely accepted in many countries, suffer from several inconsistencies. Vegetation units accepted in these classifications have usually been defined by many different researchers who have used variable and in most cases not explicitly stated classification criteria. Frequently there are considerable overlaps in delimitation of vegetation units by different researchers. Extensive national revisions of vegetation units (e.g. Oberdorfer 1977-1992; Mucina et al. 1993; Schaminée et al. 1995-1999; Valachovič 1995 et seq.; Dierschke 1996 et seq.) eliminated these inconsistencies only incompletely.

The advent of numerical classification methods since the 1960s (Mucina & van der Maarel 1989) and the availability of large electronic databases of vegetation relevés since the 1990s (Ewald 2001; Hennekens & Schaminée 2001) made it possible to develop highly formalized, consistent and repeatable classifications as an alternative to the traditional expert-based approaches. However, even sophisticated combinations of several formalized methods (e.g. Wildi 1989) so far have not received universal acceptance especially in those European countries where phytosociological tradition is firmly established. In our opinion, the main reason is that the results of commonly used agglomerative or divisive classification algorithms strongly depend on the geographical or ecological extent and stratification of the data sets (Bruelheide & Chytrý 2000). Confronted with multiple solutions of numerical classifications, many researchers often take a conservative approach and rather stick to the units of the traditional expert-based classification. An increasingly common practice seems to be combining numerical classification with subsequent interpretation of clusters selected subjectively across different hierarchical levels, merging selected clusters so that they better correspond to traditional units, or even manually re-assigning selected relevés among clusters (Hennekens et al. 1995; Bergmeier 2002; Krestov & Nakamura 2002; Willner 2002). This practice of *post hoc* manual re-arrangement of the numerical classification results indicates that formalization of the traditional expert-based classification by cluster analysis and related unsupervised methods has mostly failed so far.

An alternative to the commonly used numerical classification algorithms is the Cocktail method proposed by Bruelheide (1995, 2000). This method produces formalized definitions of vegetation units by providing unequivocal criteria for assignment of relevés to these units. At the same time it appears to be able to delimit vegetation units in a similar way to the traditional expert-based classification (Bruelheide & Chytrý 2000), but with the elimination of the latter’s inherent inconsistencies. An important difference between the Cocktail method and the numerical classification algorithms is that the Cocktail method does not assign all the relevés in the data set to vegetation units. It preferably defines vegetation units in those parts of the vegetation continuum, where several species with rather narrow ecological or geographical ranges meet, while those parts of the vegetation continuum which contain only widespread generalist species are often not assigned to any vegetation unit by the Cocktail method.

However, non-assignment of some relevés to the vegetation units may become a problem in some applications of vegetation classification, notably in vegetation mapping. Therefore it seems to be advantageous to apply the Cocktail method in combination with numerical procedures that assign relevés to vegetation units by calculating similarity between the relevés and constancy columns of vegetation tables (Hill 1989; Dodd et al. 1994). If these procedures are run in large phytosociological data sets, diagnostic species of vegetation units can be formally defined (Chytrý et al. 2002) and performance of the similarity calculations can be enhanced by positive weighting of diagnostic species.

The aim of the present paper is to test the ability of the Cocktail method, combined with a newly designed procedure of similarity-based assignment of relevés to vegetation units, to reproduce an expert-based vegetation classification in a formal way. As a test data set, we use the subalpine tall-forb and deciduous scrub vegetation of the Czech Republic, previously classified at the level of associations by expert knowledge.

Material and Methods

Material

We took the classification of the subalpine tall-forb and deciduous scrub vegetation of the class *Mulgedio-Aconitetea* in the Czech Republic (Kočí 2001) as an example of the traditional expert-based classification. The data set used for creating this classification consisted of 718 relevés of subalpine tall-forb vegetation, with species cover estimated on the Braun-Blanquet or Domin scale (Westhoff & van der Maarel 1978). This classification was largely based on expert knowledge, being a compromise between the field experience of the author and different local classifications published in earlier literature. The classification was aided by numerical divisive algorithm of the TWINSPAN program (Hill 1979), which was used in several successive runs. Several TWINSPAN end-groups were either merged or further divided according to the subjective opinion of the author. Assignment of each of the relevés to the end-groups was checked manually and some relevés were eventually moved to groups other than those suggested by TWINSPAN. In the end, each of the 718 relevés was assigned to one of 16 recognized associations.

The Cocktail classification, which was used to formally reproduce the expert-based classification by Kočí (2001), was performed with a data set of 21 794 relevés, containing all vegetation types of the Czech Republic. This data set was taken from the Czech National Phytosociological Database (Chytrý & Rafajová 2003), using a geographically stratified selection that made it possible to avoid a great influence of the over-sampled areas on the results. In this selection, we took only one relevé of each syntaxon per grid square of 1.25 longitudinal × 0.75 latitudinal minute (ca. 1.5 km × 1.4 km). The assignment to syntaxa at the level of association (or alliance) was according to the original assignments by the relevé authors. If two or more relevés of the same association were encountered in the same grid square, selection priority was given to the relevés with recorded cryptogams and to newer relevés. If this selection still yielded more than one relevé in the grid square, one of them was selected at random. Due to the stratified selection, several relevés contained in the above-mentioned data set of 718 relevés were not included in this data set.

Cocktail classification

The Cocktail method (Bruelheide 1995, 2000) was designed to mimic the classification procedure of the Braun-Blanquet approach. Cocktail classification is basically created by expert knowledge and not by an unsupervised algorithm of a computer program. An expert makes subjective choices during the classification process while the program suggests possible solutions and ensures that particular steps in the classification process are applied consistently throughout the data set. Delimitations of the resulting vegetation units are explicitly formally described, which means that also new relevés that were not present in the original data set can be unequivocally classified as belonging or not belonging to the particular vegetation unit.

The Cocktail procedure starts with defining groups of species that tend to occur together in relevés of a large database. Using a large database that covers a broad spectrum of different habitats and a large geographical area is important for obtaining species groups of more general validity. Species of the same group usually have similar habitat requirements and phytogeographical affinities. Cocktail species groups correspond to the concept of sociological species groups (Doing 1969) and often they are closely related to the groups of diagnostic species for particular vegetation units as recognized in phytosociological literature. Extraction of each group starts with one or a few species preselected by the researcher. Other species with the most similar distribution across the relevés of the database are added stepwise to this starting species or species group. In our case, co-occurrence tendency of species was measured by the phi coefficient of association (Sokal & Rohlf 1995; Chytrý et al. 2002).

Unlike Bruelheide (1995, 2000) who used a fully automated process of species group optimization, we allowed for more manual control with the aim to arrive at ecologically more coherent species groups. After selecting a starting group of two or three species, we calculated the phi coefficient of association between each species in the data set and the group of relevés that contained the starting species group. Of the species not belonging to the species group, we usually chose the one with the highest Φ value and included it in the group as its next member. In a few cases we included the species with the second or third highest Φ value, particularly if the species with the highest Φ value was already included in another species group or had several times more or less occurrences in the data set than the species already included in the species group. This solution was chosen because groups of species with large differences in occurrence frequency would not be ecologically coherent: their species might have roughly identical ecological optima but much more frequent species would usually have broader ecological ranges. After including the new species in the species group, we re-defined the group of relevés and recalculated the phi coefficient for all species in the data set and the new group of relevés. If the species group disintegrated after this step, i.e. some of the species not included in the species group had a higher Φ value than some of the species included, the group was rejected. By contrast, if the species belonging to the group had the highest Φ values, the group was either accepted or further optimized by including additional species. The optimization process was stopped if any of the candidate species for addition in the next step either caused group disintegration or substantially changed the ecological coherence of the group.

To consider a species group as being contained in a relevé, not all the species of the group need to be present. Bruelheide (1995, 2000) defined the minimally required number of species of the group as the intersection of expected and observed cumulative distribution functions for relevés having 0 to k species, k being the number of all species included in the group. However, our pilot studies showed that this criterion tends to yield a low minimum number of species if the group consists of species that are rare in the data set and a high minimum number of species if the group mainly includes common species. Our data set of 21 794 relevés included many different vegetation types, which made subalpine tall-forb species relatively rare; then the calculated minimum number of species was two for most groups. By contrast, in our data set of 718 relevés, which contained only subalpine tall-forb vegetation, tall forbs were relatively common and the calculated minimum number of species increased for several groups. This indicates that the minimum criterion derived from the cumulative distribution functions is strongly dependent on the data set structure, which complicates the transfer of species groups between different data sets. Therefore we employed a simpler criterion, taking half of the species of the group as the minimum number, e.g. at least two out of four or three out of five.

After defining several species groups, the Cocktail method creates definitions of vegetation units by combinations of species groups using logical operators such as *AND*, *OR* or *AND NOT* (Bruelheide 1997). For example, a relevé is assigned to vegetation type *X* if it contains species group *A* and at least one of the species groups *B* or *C*, while at the same time the species group *D* is absent. As our aim was reproduction of an expert-based classification of Kočí (2001), we combined species groups in such a way as to arrive close to an understanding of associations in that classification. However, pure combinations of species groups were not sufficient to reproduce most of the expert-based associations. Therefore in some Cocktail definitions of associations we used dominance of single species in combination with occurrence of species groups. A threshold cover value of either 25% or 50% was used as dominance criterium.

Similarity-based assignment of relevés to vegetation units

Some Cocktail definitions of associations created in this study marginally overlapped, i.e. a few relevés could be assigned to more than one association. These relevés and the relevés unassigned to any of the associations were subsequently assigned to one of the associations by calculating their similarity to the constancy columns of a synoptic table created from the relevés which were unequivocally assigned to the associations by Cocktail definitions.

For the similarity-based assignment of relevés to vegetation units we designed a new index, called Frequency-Fidelity Index (*FFI*), which is a combination of Frequency Index (*FQI*) and Fidelity Index (*FDI*). The Frequency Index is derived from the measure of 'compositional satisfaction' proposed by Hill (1989), modified so as to deal with percentage frequencies of species occurrence rather than with constancy classes:

$$FQI = \Sigma_{i \in \mathrm{R}} FQ_i \,/\, \Sigma_{i \in \mathrm{C}} FQ_i \qquad (1)$$

FQ_i is the frequency (constancy) of species i in a constancy column of a synoptic table. Species present in the relevé are indicated as $i \in \mathrm{R}$ and species present in the constancy column as $i \in \mathrm{C}$. In the numerator, frequencies are summed over all species of the constancy column that are also present in the relevé considered, while in the denominator the sum is calculated over all species of the constancy column.

The Frequency Index satisfactorily measures the similarity of relevés to constancy columns in terms of species composition; however, it gives the same weight to diagnostic (specialist) species, i.e. those with a distinct concentration of occurrence in a certain vegetation unit, and generalist species, i.e. those occurring in most vegetation units of a synoptic table. Thus, for a constancy column of a vegetation unit consisting of an equal proportion of generalist and diagnostic species, a relevé containing none of the diagnostic species but sharing all the generalist species will have the same *FQI* value as a relevé sharing all the diagnostic species but lacking the generalists. As this feature may be disadvantageous in phytosociological applications, we introduce an alternative index, called Fidelity Index (FDI), which is based on fidelity, i.e. concentration of species occurrence in a given vegetation unit:

$$FDI = \Sigma_{i \in \mathrm{R}} FD_i \,/\, \Sigma_{i \in \mathrm{C}} FD_i \qquad (2)$$

This index is calculated in an identical way as the Frequency Index but uses a fidelity measure (*FD*) instead of frequency (Chytrý et al. 2002). In the calculations, we considered only positive fidelity values, which clearly indicate one particular vegetation unit. By contrast, negative fidelity values of a particular species are often very similar for several vegetation units. E.g., a specialist species occurring in a single vegetation unit will have nearly identical values of negative fidelity in all other vegetation units. The fidelity measure (*FD*) used in this paper was the phi coefficient (Sokal & Rohlf 1995; Chytrý et al. 2002).

A major disadvantage of the Fidelity Index is that it poorly discriminates between the relevés composed exclusively of generalist species shared with the constancy column on one hand, and the relevés composed of totally different species than the constancy column on the other hand. Both of these two types of relevés yield a *FDI* value close or equal to zero. Therefore we combined the Frequency Index and Fidelity Index into a single Frequency-Fidelity Index (*FFI*), which retains the advantages and lacks the disadvantages of both:

$$FFI = (FQI + FDI) / 2 \qquad (3)$$

Comparison of expert-based and formalized classification

To compare the expert-based and formalized classification, we applied the Cocktail definitions of associations to the data set of 718 relevés originally used for the expert-based classification by Kočí (2001). We assigned the relevés that met these definitions to the associations. Subsequently we calculated the Frequency-Fidelity Index for the relevés that had been assigned to more than one association by the Cocktail method and assigned them to that of the candidate associations for which the highest *FFI* value was yielded. Then we applied the same procedure to the relevés that had not been assigned to any of the associations by the Cocktail definitions.

The assignment of relevés to the associations in the expert-based classification was first compared with their assignment by the Cocktail definitions only, and second, with the combined assignment by the Cocktail definitions and similarity calculations. All analyses in this paper were performed with the JUICE program (Tichý 2002; www.sci.muni.cz/botany/juice.htm).

Results

Using the Cocktail method, we defined several species groups, of which 12 were used in the formal definitions of associations (Table 1). Some of these groups have a rather narrow ecological range and they more or less correspond to the groups of diagnostic species for associations or alliances within the class *Mulgedio-Aconitetea* (e.g. *Aconitum plicatum* Group, *Betula carpatica* Group, *Bupleurum vapincense* Group, and *Laserpitium archangelica* Group). Some other groups have a wide ecological range, with an optimum beyond the subalpine tall-forb vegetation, e.g. the *Mercurialis perennis* Group in beech forests and the *Trientalis europaea* Group in spruce forests. Φ values for the species of the groups in Table 1 are proportional to the group quality, higher values indicating that the species of the group have a stronger co-occurrence tendency.

Of 16 associations of the expert-based classification (Kočí 2001; App. 1), 14 were reproduced by the Cocktail definitions (Table 2). In contrast with the original classification, association 3, *Sileno-Calamagrostietum* was merged with association 2, *Crepido-Calamagrostietum,* because the former was a species-poor variant of the latter. Association 9, *Piceo-Salicetum* was abandoned since it was a rare community known only from two sites and mainly defined by its structure rather than by floristic composition.

When the Cocktail definitions of associations were applied to the original data set of 718 relevés used by Kočí (2001), 376 relevés were assigned to the associations. Of these 376 relevés, 15 were simultaneously assigned to two associations; these relevés were subsequently assigned to one of the two suggested associations by similarity calculations. Assignments by the Cocktail definitions were in good agreement with the original expert-based classification (Table 3), except for the associations with few or poor diagnostic species (4, *Bistorto-Deschampsietum*, 5, *Violo-Deschampsietum,* and 12, *Trollio-Geranietum*). Some relevés originally classified to association 8, *Pado-Sorbetum* were assigned to the structurally and floristically similar association 7, *Salici-Betuletum* by the Cocktail definitions.

This good agreement was even improved, particularly for associations 4 and 5, after similarity-based assignment of the so far unclassified relevés to the constancy columns consisting of the relevés that had been assigned by the Cocktail definitions (Table 4). Cocktail definition of the combined association 2-3 mainly captured the relevés of association 2, *Crepido-Calamagrostietum*, which had been defined in the expert-based classification by the presence of several specialist species. Relevés of the expert-based association 3, *Sileno-Calamagrostietum*, which had a similar floristic composition as association 2 but lacked the specialist species, were satisfactorily assigned to the combined association 2-3 only by similarity calculations.

Table 1. Species groups used for the Cocktail classification of subalpine tall-forb vegetation of the Czech Republic. Species of each group are ranked by decreasing fidelity to the group, calculated as the phi coefficient in a data set of 21 794 relevés of all vegetation types of the Czech Republic. The range of Φ values (multiplied by 100) between the first and the last species is given, higher values indicating more distinct groups. The number of relevés in which the group occurs is the number of relevés containing at least half of the group's species.

Species group	**$\Phi \times 100$**	**No. of relevés**	**Species** (ranked by decreasing Φ value)
Aconitum plicatum	64 - 38	47	*Aconitum plicatum, Viola biflora, Epilobium alpestre, Carduus personata*
Betula carpatica	70 - 36	14	*Salix silesiaca, Ribes petraeum, Betula carpatica, Prunus padus* ssp. *borealis*
Bupleurum vapincense	67 - 43	5	*Thymus pulcherrimus* ssp. *sudeticus, Pleurospermum austriacum, Bupleurum longifolium* ssp. *vapincense, Thesium alpinum*
Cardamine amara	55 - 34	300	*Cardamine amara, Chrysosplenium alternifolium, Chaerophyllum hirsutum, Crepis paludosa, Chrysosplenium oppositifolium*
Geranium sylvaticum	65 - 35	215	*Geranium sylvaticum, Cardaminopsis halleri, Phyteuma spicatum, Silene dioica, Crepis mollis, Cirsium heterophyllum*
Laserpitium archangelica	67 - 28	6	*Laserpitium archangelica, Delphinium elatum, Campanula latifolia, Stachys alpina, Scrophularia scopolii, Aconitum lycoctonum* ssp. *lycoctonum*
Ligusticum mutellina	63 - 57	25	*Campanula barbata, Ligusticum mutellina, Viola lutea* ssp. *sudetica, Avenula planiculmis*
Mercurialis perennis	78 - 71	1688	*Mercurialis perennis, Galeobdolon luteum* s. lat., *Galium odoratum, Dryopteris filix-mas*
Petasites albus	62 - 45	312	*Petasites albus, Stellaria nemorum, Cicerbita alpina, Thalictrum aquilegiifolium*
Trientalis europaea	71 - 60	410	*Trientalis europaea, Homogyne alpina, Calamagrostis villosa*
Vaccinium myrtillus	79 - 55	1739	*Vaccinium myrtillus, Avenella flexuosa, Polytrichum formosum, Dicranum scoparium*
Veratrum lobelianum	57 - 32	108	*Rumex arifolius, Adenostyles alliariae, Athyrium distentifolium, Veratrum album* ssp. *lobelianum, Gentiana asclepiadea, Ranunculus platanifolius*

Table 2. Associations recognized in the expert-based classification (Kočí 2001) and their Cocktail definitions in terms of occurrence of species groups and dominance of species. In the Cocktail definitions, species groups are printed in bold and dominant species in normal letters. Brief descriptions of associations are found in App. 1

Association	**Cocktail definition**
1. *Sphagno compacti-Molinietum caeruleae*	***Trientalis europaea* Group** AND *Molinia caerulea* cover > 50%
2–3. *Crepido-Calamagrostietum villosae* and *Sileno vulgaris-Calamagrostietum villosae*	***Veratrum lobelianum* Group** AND ***Vaccinium myrtillus* Group** AND ***Trientalis europaea* Group** AND *Calamagrostis villosa* cover > 25%
4. *Bistorto-Deschampsietum alpicolae*	(***Trientalis europaea* Group** OR ***Vaccinium myrtillus* Group**) AND *Deschampsia cespitosa* cover > 50% AND NOT (***Ligusticum mutellina* Group** OR *Salix lapponum* cover > 25%)
5. *Violo sudeticae-Deschampsietum cespitosae*	***Ligusticum mutellina* Group** AND *Deschampsia cespitosa* cover > 25%
6. *Bupleuro-Calamagrostietum arundinaceae*	***Bupleurum vapincense* Group**
7. *Salici silesiacae-Betuletum carpaticae*	***Betula carpatica* Group** AND NOT (*Prunus padus* ssp. *borealis* cover > 25% OR *Sorbus aucuparia* cover > 25%)
8. *Pado-Sorbetum*	***Betula carpatica* Group** AND (*Prunus padus* ssp. *borealis* cover > 25% OR *Sorbus aucuparia* cover > 25%)
9. *Piceo-Salicetum silesiacae*	This association was not reproduced with the Cocktail method.
10. *Ranunculo platanifolii-Adenostyletum alliariae*	(***Aconitum plicatum* Group** OR ***Veratrum lobelianum* Group**) AND *Adenostyles alliariae* cover > 25% AND NOT ***Laserpitium archangelica* Group**
11. *Salicetum lapponum*	(***Trientalis europaea* Group** OR ***Veratrum lobelianum* Group** OR ***Aconitum plicatum* Group**) AND *Salix lapponum* cover > 25%
12. *Trollio altissimi-Geranietum sylvatici*	(***Aconitum plicatum* Group** AND ***Geranium sylvaticum* Group**) AND NOT (***Bupleurum vapincense* Group** OR ***Laserpitium archangelica* Group** OR ***Petasites albus* Group**)
13. *Laserpitio-Dactylidetum glomeratae*	***Laserpitium archangelica* Group**
14. *Chaerophyllo-Cicerbitetum alpinae*	(***Cardamine amara* Group** OR ***Petasites albus* Group**) AND *Cicerbita alpina* cover > 25%
15. *Daphno mezerei-Dryopteridetum filicis-maris*	(***Mercurialis perennis* Group** OR ***Veratrum lobelianum* Group**) AND *Dryopteris filix-mas* cover > 25%
16. *Adenostylo-Athyrietum alpestris*	(***Trientalis europaea* Group** OR ***Veratrum lobelianum* Group**) AND *Athyrium distentifolium* cover > 25%

Table 3. Relationships between the associations of the expert-based classification and of the formalized classification. Only relevés assigned by the Cocktail definitions are considered in the formally defined associations. The associations of the two classifications are compared by the phi coefficient (multiplied by 100); negative Φ values are not shown and values higher than 50 are printed in bold. Numbers of associations are the same as in Table 2.

Formally defined associations		**Expert-based associations recognized by Kočí (2001)**															
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
	No. of relevés	*34*	*84*	*120*	*36*	*41*	*55*	*15*	*13*	*4*	*59*	*14*	*15*	*23*	*53*	*37*	*115*
1	*27*	**78**	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2-3	*51*	-	**54**	5	-	-	-	-	-	-	-	-	-	-	-	-	-
4	*9*	-	-	-	26	19	-	-	-	-	-	-	-	-	-	-	-
5	*12*	-	-	-	7	34	-	-	-	-	-	-	13	-	-	-	-
6	*30*	-	-	-	-	-	**70**	-	-	-	-	-	-	1	-	-	-
7	*19*	-	-	-	-	-	-	**77**	30	-	-	-	-	-	-	-	-
8	*6*	-	1	-	-	-	-	-	**56**	-	-	-	-	-	-	-	-
10	*31*	-	-	-	-	-	-	-	-	-	**69**	-	2	-	-	-	-
11	*13*	-	-	-	-	-	-	-	-	-	-	**96**	-	-	-	-	-
12	*8*	-	-	-	-	3	-	-	-	-	-	-	36	21	-	-	-
13	*10*	-	-	-	-	-	6	-	-	-	-	-	-	**52**	-	-	-
14	*29*	-	-	-	-	-	-	-	-	-	7	-	-	-	**60**	-	-
15	*27*	-	-	-	-	-	-	-	-	-	-	-	-	-	-	**75**	-
16	*104*	-	-	-	-	-	-	-	-	8	-	-	-	-	-	-	**82**

Discussion

The Cocktail method, combined with the similarity-based assignment of relevés to vegetation units, satisfactorily reproduced the expert-based classification of subalpine tall-forb vegetation of the Czech Republic (Kočí 2001) at the level of associations. We formulated unequivocal assignment criteria for 14 of 16 associations recognized by Kočí (2001) (Table 2). The remaining two associations of the original classification could also be formally defined, but we decided to abandon them as they lacked positive differentiation and were highly similar to some other associations. Tables 3 and 4 show generally good agreement between the expert-based classification and the formalized classification proposed in this paper. A poor agreement was found between some expert-based associations composed mainly of generalist species and their corresponding Cocktail definitions; however, this was substantially improved by the subsequent similarity-based assignment of relevés to the associations. These results show that the expert-based vegetation classification can be successfully reproduced by the formalized methods and converted to the computer expert systems for identification of vegetation units (Noble 1987). Similar results can also be achieved by other methods such as neural networks (Ejrnaes et al. 2002). However, the absolute agreement will rarely be achieved since the expert-based classifications contain various inconsistencies and consider also non-floristic classification criteria such as vegetation structure, chorology, stand history, position in successional seres, and abiotic site factors (Westhoff 1967; Feoli 1984; Pignatti et al. 1995).

Our ability to reproduce narrowly conceived vegetation units as associations was due to involving the dominance of individual species as a classification criterion, in addition to the presence of species groups. Previous applications of the Cocktail method showed that if only species groups without other classification criteria were used (Bruelheide 1995; Jandt 1999, 2000; Bruelheide & Chytrý 2000), it was mostly possible to define alliances or broad groups of associations but not the associations. On the other hand, studies employing dominance or presence of individual species in addition to the species groups (Pflume 1999; Täuber 2000) were quite successful in reproducing the traditional associations. These results reflect the fact that most associations traditionally recognized in Central Europe are defined not only by species presence/absence but also by quantitative proportions among species.

Cocktail definitions of vegetation units may overlap to some degree, which results in an inconvenient property that some relevés may be assigned to more than one vegetation unit. These overlaps are larger the broader the definitions of vegetation units are. Therefore it is often necessary to create narrow definitions including only the most typical, core relevés of vegetation units. Still, even with narrow definitions some relevés may be assigned to more than one vegetation unit. Jandt (1999), Pflume (1999) and Täuber (2000) fixed this issue by introducing a hierarchy of vegetation units. They started with the unit that was subjectively given the highest

Table 4. Relationships between the associations of the expert-based classification and of the formalized classification. Both the relevés assigned directly by the Cocktail definitions and relevés assigned by subsequent similarity calculations are considered in the formally defined associations. The associations of the two classifications are compared by the phi coefficient (multiplied by 100); negative *Φ* values are not shown and values higher than 50 are printed in bold. Numbers of associations are the same as in Table 2.

Formally defined associations		**Expert-based associations recognized by Koćí (2001)**															
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
	No. of relevés	*34*	*84*	*120*	*36*	*41*	*55*	*15*	*13*	*4*	*59*	*14*	*15*	*23*	*53*	*37*	*115*
1	*64*	**69**	-	8	11	-	-	-	-	-	-	-	-	-	-	-	-
2-3	*184*	-	**55**	46	-	-	2	-	-	-	-	-	-	-	-	-	-
4	*25*	-	-	-	**55**	18	-	-	-	-	-	-	-	-	-	-	-
5	*25*	-	-	-	3	**58**	-	-	-	-	-	-	13	-	-	-	-
6	*31*	-	-	-	-	-	**71**	-	-	-	-	-	-	-	-	-	-
7	*22*	-	-	-	-	-	-	**77**	28	-	-	-	-	-	-	-	-
8	*8*	-	1	-	-	-	-	-	**58**	17	-	-	-	-	-	-	-
10	*58*	-	-	-	-	-	-	-	-	-	**81**	-	10	1	-	-	-
11	*14*	-	-	-	-	-	-	-	-	-	-	**100**	-	-	-	-	-
12	*28*	-	-	-	-	26	-	-	-	-	-	-	32	37	-	-	-
13	*11*	-	-	-	-	-	5	-	-	-	-	-	-	**56**	-	-	-
14	*54*	-	-	-	-	-	-	-	-	-	1	-	-	-	**91**	-	-
15	*38*	-	-	-	-	-	3	-	-	-	-	-	-	-	-	**82**	-
16	*156*	-	-	-	-	-	-	-	-	10	-	-	-	-	-	-	**79**

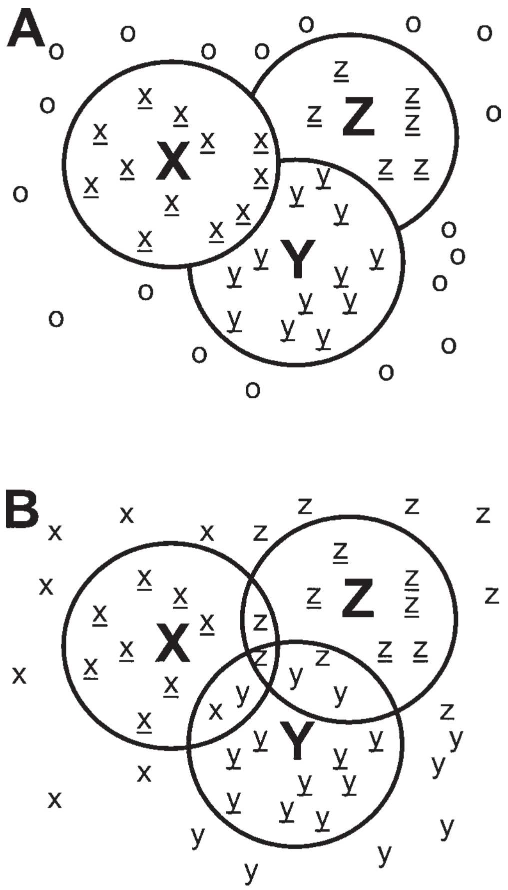

Fig. 1. Scheme of the Cocktail classification of relevés into three vegetation units, followed by a similarity-based assignment of unclassified relevés to these units. Circles denote the boundaries of the three vegetation units X, Y and Z. Small letters x, y, z and o are symbols for relevés distributed in a vegetation continuum, different letters denoting their assignment to the vegetation units X, Y, Z and non-assignment, respectively. Underlined letters indicate the relevés assigned to the vegetation units by Cocktail definitions, non-underlined letters are the relevés assigned by calculating their similarity to the vegetation units. **A.** Hierarchical approach of the earlier studies, where vegetation unit X is defined first and has the highest priority, while Z has the lowest priority; **B.** Non-hierarchical approach used in this paper, with relevés that occurred in the overlapping parts of the Cocktail definitions of vegetation units being assigned using a similarity criterion.

priority and each of the next units included the definitions of all the previously defined units in the negative part of its definition (Fig. 1A). Such an approach, however, is unduly influenced by vegetation units from which the classification starts. The new procedure proposed in this paper solves the overlap issue by a similarity-based assignment of the equivocally classified relevés to one of the possible vegetation units (Fig. 1B). The advantage of this procedure is that definitions of vegetation units are independent of the earlier defined units.

An important feature of the Cocktail method is that some relevés which do not meet any of the association definitions are not classified. Most of the unclassified relevés are from vegetation stands that lack specialist species, which could be used as diagnostic species of associations. This feature is similar to the traditional Braun-Blanquet approach, in which the researchers preferably sample vegetation stands with specialized species and often neglect stands composed mainly of generalists. This makes the Cocktail method ideally suited for classification of phytosociological data which are biased towards vegetation stands with specialist species. In traditional phytosociology, the stands without specialist species are usually referred to as impoverished, atypical, initial, degraded, transitional, trunk or basal communities. Kopecký & Hejný (1978) proposed to assign these stands directly to higher syntaxa, such as alliances, orders or classes. Although their approach has been followed only in studies of synanthropic vegetation (Kopecký et al. 1995), it is applicable to any vegetation type. Using the Cocktail method, definitions of higher syntaxa can be formed in a similar way as definitions of associations, and relevés unassigned to the associations can be assigned directly to the higher syntaxa.

In vegetation mapping and other applications of vegetation classification, however, it might be a disadvantage if some patches of a single stand were assigned to an association and others directly to higher syntaxa. We propose a solution by calculating similarity between associations and the relevés not belonging to the associations according to Cocktail definitions. Then, users of the classification will distinguish three categories of relevés or vegetation stands (Fig. 1), including those (1) belonging to the association, i.e. corresponding to its Cocktail definition; (2) not belonging to the association but related to it, i.e. not corresponding to the Cocktail definition but being similar to the relevés corresponding to this definition; and (3) not belonging and not related to the association. If necessary, the second category can be merged with the first for the purposes of vegetation mapping. The distinction between the second and third category can be made in two ways. The first option is defining a similarity threshold that provides a criterion for relevés to be considered as either related or unrelated to the association. The second option, used in this paper, is calculation of similarities between the relevé and each of predefined set of associations, and subsequent assignment of the relevé to the most similar association.

Acknowledgements. We thank Helge Bruelheide, Jason Holt, Ute Jandt, Onno van Tongeren and two anonymous referees for their stimulating ideas and comments on a previous version of the manuscript. Maruška Rafajová managed our vegetation database. This research was supported by the grants GAČR 206/99/1523 and MSM 143100010.

References

Bergmeier, E. 2002. Plant communities and habitat differentiation in the Mediterranean coniferous woodlands of Mt. Parnon (Greece). *Folia Geobot.* 37: 309-331.

Bruelheide, H. 1995. Die Grünlandgesellschaften des Harzes und ihre Standortsbedingungen. Mit einem Beitrag zum Gliederungsprinzip auf der Basis von statistisch ermittelten Artengruppen. *Diss. Bot.* 244: 1-338.

Bruelheide, H. 1997. Using formal logic to classify vegetation. *Folia Geobot. Phytotax.* 32: 41-46.

Bruelheide, H. 2000. A new measure of fidelity and its application to defining species groups. *J. Veg. Sci.* 11: 167-178.

Bruelheide, H. & Chytrý, M. 2000. Towards unification of the national vegetation classifications: A comparison of two methods for the analysis of large data sets. *J. Veg. Sci.* 11: 295-306.

Chytrý, M. & Rafajová, M. 2003. Czech National Phyto-Sociological Database: basic statistics of the available vegetation-plot data. *Preslia* 75: 1-15.

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

Dierschke, H. (ed.) 1996. *Synopsis der Pflanzengesellschaften Deutschlands.* Floristisch-soziologische Arbeitsgemeinschaft & Reinhold-Tüxen Gesellschaft, Göttingen, DE.

Dodd, M.E., Silvertown, J., McConvay, K., Potts, J. & Crawley, M. 1994. Application of the British National Vegetation Classification to the communities of the Park Grass Experiment through time. *Folia Geobot. Phytotax.* 29: 321-334.

Doing, H. 1969. Sociological species groups. *Acta Bot. Neerl.* 18: 398-400.

Ejrnaes, R., Aude, E., Nygaard, B. & Münier, B. 2002. Prediction of habitat quality using ordination and neural networks. *Ecol. Appl.* 12: 1180-1187.

Ewald, J. 2001. Der Beitrag pflanzensoziologischer Datenbanken zur vegetationsökologischen Forschung. *Ber. Reinhold-Tüxen-Ges.* 13: 53-69.

Feoli, E. 1984. Some aspects of classification and ordination of vegetation data in perspective. *Stud. Geobot.* 4: 7-21.

Hennekens, S.M. & Schaminée, J.H.J. 2001. TURBOVEG, a comprehensive data base management system for vegetation data. *J. Veg. Sci.* 12: 589-591.

Hennekens, S.M., van der Maarel, E. & Stortelder, A.H.F. 1995. Numerieke methoden [Numerical methods]. In: Schaminée, J.H.J., Stortelder, A.H.F. & Westhoff, V. (eds.) *De vegetatie van Nederland. Deel 1. Inleiding tot de plantensociologie – grondslagen, methoden en toepassingen* [The vegetation of the Netherlands. Vol. 1. Introduction to phytosociology - fundamentals, methods and applications], pp. 151-162, Opulus Press, Uppsala, SE. (In Dutch.)

Hill, M.O. 1979. *TWINSPAN. A FORTRAN program for arranging multivariate data in an ordered two-way table by classification of the individuals and attributes.* Ecology & Systematics, Cornell University, Ithaca, NY, US.

Hill, M.O 1989. Computerized matching of relevés and association tables, with an application to the British National Vegetation Classification. *Vegetatio* 83: 187-194.

Jandt, U. 1999. Kalkmagerrasen am Südharzrand und im Kyffhäuser. Gliederung im überregionalen Kontext, Verbreitung, Standortsverhältnisse und Flora. *Diss. Bot.* 322: 1-246.

Jandt, U. 2000. Application of the species-group method to the data base of calcareous grasslands in Germany. In: White, P.S., Mucina, L. & Lepš, J. (eds.) *Vegetation science in retrospect and perspective*, pp. 115-119, Opulus Press, Uppsala, SE.

Kočí, M. 2001. Subalpine tall-forb vegetation (*Mulgedio-Aconitetea*) in the Czech Republic: syntaxonomical revision. *Preslia* 73: 289-331.

Kopecký, K. & Hejný, S. 1978. Die Anwendung einer deduktiven Methode syntaxonomischer Klassifikation bei der Bearbeitung der strassenbegleitenden Pflanzengesellschaften Nordostböhmens. *Vegetatio* 36: 43-51.

Kopecký, K., Dostálek, J. & Frantík, T. 1995. The use of the deductive method of syntaxonomic classification in the system of vegetational units of the Braun-Blanquet approach. *Vegetatio* 17: 95-112.

Krestov, P.V. & Nakamura, Y. 2002. Phytosociological study of the *Picea jezoensis* forests of the Far East. *Folia Geobot.* 37: 441-473.

Kubát, K., Hrouda, L., Chrtek jun., J., Kaplan, Z., Kirschner, J. & Štěpánek, J. (eds.) 2002. *Klíč ke květeně České republiky* [*Key to the flora of the Czech Republic*]. Academia, Praha, CZ. (In Czech.)

Mucina, L. & van der Maarel, E. 1989. Twenty years of numerical syntaxonomy. *Vegetatio* 81: 1-15.

Mucina, L., Grabherr, G., Ellmauer, T. & Wallnöfer, S. 1993. *Die Pflanzengesellschaften Österreichs. Teil I-III.* G. Fischer, Jena, DE.

Noble, I.R. 1987. The role of expert systems in vegetation science. *Vegetatio* 69: 115-121.

Oberdorfer, E. (ed.) 1977-1992. *Süddeutsche Pflanzengesellschaften*. Fischer, Jena, DE.

Ostermann, O.P. 1998. The need for management of nature conservation sites designated under Natura 2000. *J. Appl. Ecol.* 35: 968-973.

Pflume, S. 1999. *Laubwaldgesellschaften im Harz. Gliederung, Ökologie, Verbreitung.* M. Galunder-Verlag, Wiehl, DE.

Pignatti, S., Oberdorfer, E., Schaminée, J.H.J. & Westhoff, V. 1995. On the concept of vegetation class in phytosociology. *J. Veg. Sci.* 6: 143-152.

Schaminée, J.H.J. 1995–1999. *De vegetatie van Nederland* [*The vegetation of the Netherlands*]. Opulus Press, Uppsala, SE. (In Dutch.)

Sokal, R.R. & Rohlf, F.J. 1995. *Biometry.* 3rd ed. Freeman, New York, NY, US.

Täuber, T. 2000. *Zwergbinsen-Gesellschaften (Isoëto-Nano-*

juncetea) in Niedersachsen – Verbreitung, Gliederung, Dynamik, Keimungsbedingungen der Arten und Schutzkonzepte. Cuvillier, Göttingen, DE.

Tichý, L. 2002. JUICE, software for vegetation classification. *J. Veg. Sci.* 13: 451-453.

Valachovič, M., Otahelová, H., Stanová, V. & Maglocký, Š. 1995. *Rastlinné spoločenstvá Slovenska 1. Pionierska vegetácia* [*Plant communities of Slovakia 1. Pioneer vegetation*]. Veda, Bratislava, SK. (In Slovak.)

Westhoff, V. 1967. Problems and use of structure in the classification of vegetation. The diagnostic evaluation of structure in the Braun-Blanquet system. *Acta Bot. Neerl.* 15: 495-511.

Westhoff, V. & van der Maarel, E. 1978. The Braun-Blanquet approach. In: Whittaker, R.H. (ed.) *Classification of plant communities*, pp. 289-399. W. Junk, The Hague, NL.

Wildi, O. 1989. A new numerical solution to traditional phytosociological tabular classification. *Vegetatio* 81: 95-106.

Willner, W. 2002. Syntaxonomische Revision der südmitteleuropäischen Buchenwälder. *Phytocoenologia* 32: 337-453.

Received 15 February 2002;
Revision received 2 February 2003;
Accepted 13 March 2003.
Coordinating Editor: H. Bruelheide.

App. 1. Brief description of the associations recognized in the expert-based classification (Kočí 2001).

1. *Sphagno compacti-Molinietum caeruleae* Wagnerová ex Berciková 1976 – species-poor chionophilous and hygrophilous community dominated by *Molinia caerulea*, which occupies mostly leeward edges of cirques

2–3. *Crepido-Calamagrostietum villosae* (Zlatník 1925) Jeník 1961 and *Sileno vulgaris-Calamagrostietum villosae* Jeník et al. 1980 – closed grasslands dominated by *Calamagrostis villosa*, confined to leeward sites with dry, deep and nutrient-rich brown alpine soils; the former association includes species-rich, the latter species-poor stands

4. *Bistorto-Deschampsietum alpicolae* (Zlatník 1928) Burešová 1976 – species-poor community of *Deschampsia cespitosa* and *Polygonum bistorta* occupying shallow, wet depressions with long-lasting snow cover on the ridges and moderate slopes above the timberline

5. *Violo sudeticae-Deschampsietum cespitosae* (Jeník et al. 1980) Kočí 2001 – *Deschampsia cespitosa* and *Poa chaixii* dominated community, occupying depressions near springs and shaded places around the timberline, on soils with permanently percolating ground water

6. *Bupleuro-Calamagrostietum arundinaceae* (Zlatník 1928) Jeník 1961 – species-rich community dominated by *Calamagrostis arundinacea*, confined to steep, dry and sunny places, mostly located at the bases of slopes of cirques and on avalanche paths

7. *Salici silesiacae-Betuletum carpaticae* Rejmánek et al. 1971 – subalpine open shrubberies of *Betula carpatica* and *Salix silesiaca* with tall forbs, occurring on steep slopes with sliding snow and infrequent avalanches

8. *Pado-Sorbetum* W. Matuszkiewicz et A. Matuszkiewicz 1975 – subalpine open shrubberies of *Sorbus aucuparia* subsp. *glabrata* with herb layer dominated by tall forbs, confined to the bottoms of cirques and to moist ravines on shallow, stony and acidic soils

9. *Piceo-Salicetum silesiacae* Rejmánek et al. 1971 – open shrubberies of *Salix silesiaca* and *Picea abies* with ferns and woodland species in herb layer, influenced by snow accumulation and spring floods, occurring along submontane rivers

10. *Ranunculo platanifolii-Adenostyletum alliariae* (Krajina 1933) Dúbravcová et Hadač ex Kočí 2001 – species-rich community dominanted by *Adenostyles alliariae*, confined to moderate slopes, in the surroundings of springs and streams, in shaded places and wet depressions around the timberline

11. *Salicetum lapponum* Zlatník 1928 – subalpine low-willow shrubberies occupying shallow wet depressions, surroundings of springs and mires, and the upper edges of cirques with permanently moist, shallow, often peaty and acidic soils

12. *Trollio altissimi-Geranietum sylvatici* Jeník et al. 1980 – species-rich tall-forb community confined to the surrounding of streams and springs with moist soils, occurring at their upper edges of cirques outside the avalanche tracks

13. *Laserpitio-Dactylidetum glomeratae* Jeník et al. 1980 – species-rich tall-forb community of the bottoms of cirques, confined to deep and moist soils, rich in nutrients supplied by avalanches, aeolic sedimentation, and percolating ground water

14. *Chaerophyllo-Cicerbitetum alpinae* (Kästner 1938) Sýkora et Hadač 1984 – community dominated by *Petasites albus* and *Cicerbita alpina*, confined to shaded and wet surroundings of streams and springs and the bottoms of V-shaped valleys and canyons in the montane and supramontane belts

15. *Daphno mezerei-Dryopteridetum filicis-maris* Sýkora et Štursa 1973 – species-rich community dominated by *Dryopteris filix-mas*, occupying mostly dry and warm screes and scree cones with shallow soil at the bases of steep slopes in cirques, covered by thick snow accumulations in winter

16. *Adenostylo-Athyrietum alpestris* (Zlatník 1928) Jeník 1961 – species-poor, chionophilous community dominated by *Athyrium distentifolium*, occurring on deep soils on shaded wet places around the timberline, in places with a thick snow cover in winter