Preslia, Praha, 75: 1–15, 2003

Czech National Phytosociological Database: basic statistics of the available vegetation-plot data

Česká národní fytocenologická databáze: základní statistika dostupných snímkových dat

Milan Chytrý & Marie Rafajová

Department of Botany, Masaryk University, Kotlářská 2, CZ-611 37 Brno, Czech Republic, e-mail: chytry@sci.muni.cz

Chytrý M. & Rafajová M. (2003): Czech National Phytosociological Database: basic statistics of the available vegetation-plot data. – Preslia 75: 1–15.

The vegetation relevés stored electronically in the Czech National Phytosociological Database are reviewed. The database was established in 1996, with the central database located in the Department of Botany, Masaryk University, Brno (www.sci.muni.cz/botany/database.htm). On 15 November 2002 this central database contained 54,310 relevés from the Czech Republic, collected by 332 authors between 1922–2002. Ca. 54% of the relevés were taken from published papers or monographs, 21% from theses and the rest from various unpublished reports and field-books. These relevés include 1,259,008 records of individual plant species. Territorial coverage of the country by the relévés is irregular as the areas with attractive natural or semi-natural vegetation are more intensively sampled, with gaps in coverage of less attractive or poorly accessible areas. Most relevés are of broad-leaved deciduous forests (*Querco-Fagetea*), meadows (*Molinio-Arrhenatheretea*), dry grasslands (*Festuco-Brometea*), and marsh grasslands (*Phragmito-Magnocaricetea*). The quality of the data is discussed, such as researcher bias, preferential selection of sampling sites, spatial autocorrelation and missing values for some data elements.

Keywords: Bioinformatics, Czech Republic, plant community, phytosociology, relevé, vegetation sampling, survey

Introduction

Modern information technologies have facilitated the development of electronic databases, which include previously scattered information on various aspects of biodiversity (Michener & Brunt 2000). Diversity of plant communities at the level of landscapes has been traditionally documented by relevés, i.e. species lists with simple estimates of cover-abundance, recorded in plots ranging in size between 10^{-1} and 10^2 m^2. Several institutions worldwide have compiled phytosociological databases that consist of electronic archives of relevés from particular regions or national territories (Wohlgemuth 1992, Brisse et al. 1995, Ewald 1995, 2001, Font & Ninot 1995, Schaminée & Hennekens 1995, Schaminée & Stortelder 1996, Mucina et al. 2000a; see also www.vegbank.org). Since the mid 1990s, this was facilitated by the wide availability of Turboveg database management software (Hennekens & Schaminée 2001; see also www.synbiosys.alterra.nl/turboveg/).

Ewald (2001) estimated there are approximately one million relevés already stored electronically in phytosociological databases worldwide, of which ca. 750,000 originate from central and western Europe, a region with a strong tradition of vegetation sampling for classification purposes (Braun-Blanquet 1964, Westhoff & van der Maarel 1973). Undoubtedly, there are still more relevés in papers, reports and field protocols that have not yet been computerized (Rodwell 1995). For example, the databases for the Netherlands and France contain ca. 350,000 and 137,000 relevés, respectively (Ewald 2001, see also Brisse et al. 1995, Schaminée & Stortelder 1996). In contrast, there are very few computerized relevés for countries of northern, eastern, and southern Europe.

Phytosociological databases are most often designed for vegetation classification and as a source of information on the spatial distribution of vegetation or habitat diversity, as required by the nature conservation agencies or forest management institutions. However, this data can be used for a variety of other purposes, such as determining changes in vegetation, defining major environmental gradients, estimating species niches, calibrating indicator values for species, mapping the distribution of individual species, and modelling potential distribution of species and plant communities.

The purpose of this paper is to review the data in the Czech National Phytosociological Database, which consists of relevés from the Czech Republic.

History of the Czech National Phytosociological Database

The idea to set up vegetation database first appeared at the Institute of Botany, Průhonice, towards the end of the 1970s. Associated activities during the 1980s included compilation of a checklist of species of central European vascular plants, bryophytes and lichens (Neuhäuslová & Kolbek 1982), publication of a bibliography of Czechoslovak phytosociological studies (Bibliographia 1983–1992), and the development of the computer program Edifyt for editing phytosociological tables by Eduard Brabec at the end of the 1980s. However, this database was not developed further in the 1990s.

Interest in the project was rekindled in the mid 1990s due to the support of the newly established international working group: European Vegetation Survey (Mucina et al. 1993, Rodwell et al. 1995, 1997). In 1995–1998 Professor John S. Rodwell from Lancaster University, UK, received funding for the project "Safeguarding the Biodiversity of Eastern Europe" under the Darwin Initiative funded by the British Government. As a part of this project, a training course on vegetation databanking was organized for vegetation scientists from five central and eastern European countries, including the Czech Republic. This course was held at Lancaster in February 1996, under the supervision of Prof. John S. Rodwell and Julian Dring, in close co-operation with Stephan M. Hennekens (Wageningen, The Netherlands), the author of the computer program Turboveg. In February 1997, a training course on the use of the Turboveg program was held at the Masaryk University, Brno, for vegetation scientists from the Czech Republic and Slovakia. As a result, the Turboveg program has been adopted for storing phytosociological relevés in the Czech Republic and Slovakia (Chytrý 1996, Valachovič 1996) and widely used by many institutions and individuals.

The Czech National Phytosociological Database was established in early 1996 (Chytrý 1997), after a new species checklist of vascular plants, bryophytes and lichens of the Czech Republic, Slovakia, Austria, and Hungary was prepared in co-operation with Prof. Harald Niklfeld, Dr. Walter Gutermann and Prof. Ladislav Mucina from the University of Vienna, and Drs. Ivan Jarolímek and Milan Valachovič from the Institute of Botany in Bratislava. The common checklist was adopted to facilitate the exchange of data among the four countries. Vascular plant nomenclature is that used in the revised and updated checklist of Ehrendorfer (1973) and the nomenclature for bryophytes and lichens is that of Frey et al. (1995) and Pišút et al. (1993), respectively.

The organizational structure of the Czech National Phytosociological Database follows the model of several local databases, managed by individuals or research groups at different institutions with the central database in the Department of Botany, Masaryk University, Brno, where data from the local databases, papers and other documents are kept. Besides data collecting and storing, the central database also provides relevés upon request to participating researchers for the purposes of various projects. Funding from the Grant Agency of the Czech Republic supported the appointment of Marie Rafajová as the full-time administrator of the central database since 1999. This made it possible to start with a more systematic computerization of relevés and apply quality assurance procedures to the growing body of data. Updated information on the database is available in English on the web site of the Department of Botany, Masaryk University, at www.sci.muni.cz/botany/database.htm, and more comprehensive information in Czech, including technical guidelines for managers of local databases at www.sci.muni.cz/botany/dbase_cz.htm.

Development of the database is accompanied by research into the structure of large phytosociological data sets (Chytrý 2001), testing new formalized methods of vegetation classification and data analysis (Bruelheide & Chytrý 2000, Chytrý et al. 2002) and development of the computer program Juice (Tichý 2002), which includes several tools for numerical classification and analysis of phytosociological data. This program is freely available on the web site www.sci.muni.cz/botany/juice.htm.

Recently, these developments stimulated the initiation of a parallel project: the Database of Forest Typology at the Forest Management Institute (ÚHÚL). In cooperation with the Department of Botany, Masaryk University, computerization of the data from field protocols of the forest site-type research started there in 2001, using the Turboveg program and the same standards as the Czech National Phytosociological Database.

Basic statistical figures

All the figures given in this section refer to the state of the central database on 15 November 2002 and do not include information from the database of the Forest Management Institute. To this date, the central database contained 54,310 relevés from the Czech Republic, of which 54% were from published papers or monographs (Table 1). Most relevés taken from journals came from Preslia (3240), Folia Geobotanica (et Phytotaxonomica) (2295), Muzeum a současnost (1907), Rozpravy Československé akademie věd (1626), Studie ČSAV (1381), Bohemia centralis (1298), Feddes Repertorium (1048), Folia Musei rerum naturalium Bohemiae occidentalis (993) and Opera corcontica (957). About 21% were unpublished relevés in master, doctoral and "Candidate of Science" theses, mainly of workers who studied in the Departments of Botany of Charles University, Prague, Masaryk University, Brno, Palacký University, Olomouc, and the Institute of Botany of the Academy of Sciences of the Czech Republic, Průhonice. About 6% of the relevés were from unpublished surveys of nature reserves carried out by the Agency for Nature Conservation and Landscape Protection of the Czech Republic. The quality of these relevés, however, varies considerably. Remaining data came from various unpublished sources. Relevés that have been repeatedly presented in different publications or manuscripts were included

Table 1. – Number and bibliographic sources of the relevés included in the Czech National Phytosociological Database and the Database of Forest Typology of the Forest Management Institute. Relevés from unpublished manuscripts (theses, reports) that were published later, are included only under publications.

	Number of bibliographic references included in the database	Number of relevés included in the database	Estimated number of existing relevés not yet included in the database
Czech National Phytosociological Database			
Published papers and monographs	896	29,120	4,000
Theses	119	11,635	3,000
Nature reserve survey reports	292	3,426	1,000
Unpublished relevés	–	10,129	12,000
Total	1,307	54,310	20,000
Database of Forest Typology	–	14,130	45,000

only once in the database; in Table 1 these relevés are attributed to the oldest publication. The relevés in the database were recorded by 332 authors, although many of these authors are responsible for only one or a few relevés (Table 2). The number of relevés still to be included in the database is estimated at approximately 20,000 (Table 1).

The database reflects the history of vegetation research in the Czech Republic. The oldest relevés in the database, which are perhaps the oldest relevés from the Czech Republic, were recorded by Alois Zlatník in 1922 (Zlatník 1928a, 1928b). The proportion of relevés from 1920 to 1950 is rather low. For example, 95% of the relevés were made after 1950 and 51% after 1980 (Fig. 1). Geographical coverage of the Czech Republic has differed in time (Fig. 2). Before 1950 (Fig. 2A), few researchers prevailed, in particular J. Klika, R. Mikyška, J. Šmarda and A. Zlatník, who focused on the most attractive areas of the colline belt in Bohemia, namely the České středohoří (Bohemian Central Range) and the Bohemian Karst, the famous dry grassland sites of southern Moravia (Pavlovské vrchy Hills, Mohelno) and the mountainous areas of the Krkonoše, Hrubý Jeseník, and Žďárské vrchy. Between 1951–1980 (Fig. 2B) many of the gaps in territorial coverage were filled; still, distinct gaps remained, e.g. in the Českomoravská vrchovina (Bohemian-Moravian Highlands) and central Moravia. The fieldwork in 1981–2002 (Fig. 2C) was mainly done in areas already studied, with emphasis on the less explored parts of the Moravian Carpathians. The cumulative territorial coverage achieved during the entire history of phytosociological research (Fig. 2D) is fairly good, however, there are still some gaps, mainly in the area west and southwest of Plzeň, in the middle Vltava area, in the area between the Sázava river and the foothills of the Železné hory Mts, in some parts of the Bohemian-Moravian Highlands, in the upland areas of northern Moravia (Nízký Jeseník and Oderské vrchy) and in the Ostrava city agglomeration. In contrast, the highest numbers of relevés are from areas with a high diversity of natural and semi-natural vegetation, particularly in the České středohoří Mts, Bohemian Karst and the Berounka river valley, Kokořín area, Český ráj area, Ještědský hřbet Mts, Krkonoše Mts, Hrubý Jeseník Mts, Šumava Mts, Třeboň basin, middle Labe valley, Železné hory Mts, Žďárské vrchy Mts, river valleys of southwestern Moravia, Pavlovské vrchy Hills and adjacent areas, Bílé Karpaty Mts, Hostýnské vrchy Mts and Moravskoslezské Beskydy Mts. Most relevés were recorded between 200–600 m a. s. l., which are the most common altitudes encountered in the Czech Republic (Fig. 3).

Table 2. – Twenty-five most frequent authors of the relevés in the database and the number of relevés.

Balátová-Tuláčková E.	2261	Otýpková Z.	794
Rydlo J.	1917	Šumberová K.	780
Blažková D.	1876	Kropáč Z.	730
Vicherek J.	1578	Kučera T.	725
Neuhäusl R. & Neuhäuslová Z.	1573	Tlusták V.	720
Sofron J.	1358	Sádlo J.	718
Chytrý M.	1229	Grüll F.	710
Toman M.	1127	Hédl R.	679
Klika J.	1037	Pešout P.	673
Kopecký K.	1000	Pyšek A.	661
Kolbek J.	966	Sýkora T.	644
Albrecht J.	879	Novák J.	573
Jirásek J.	864		

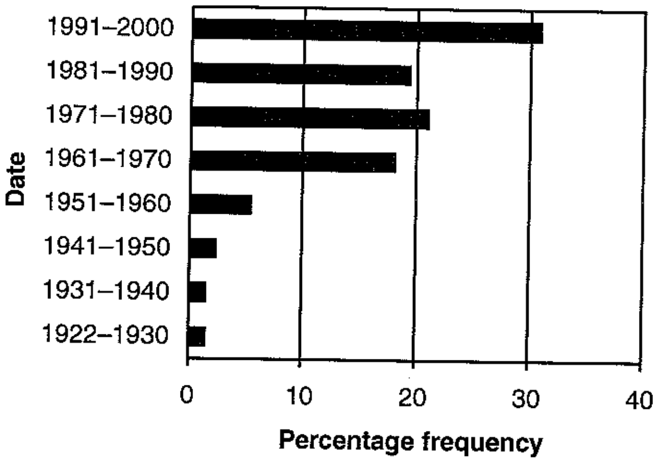
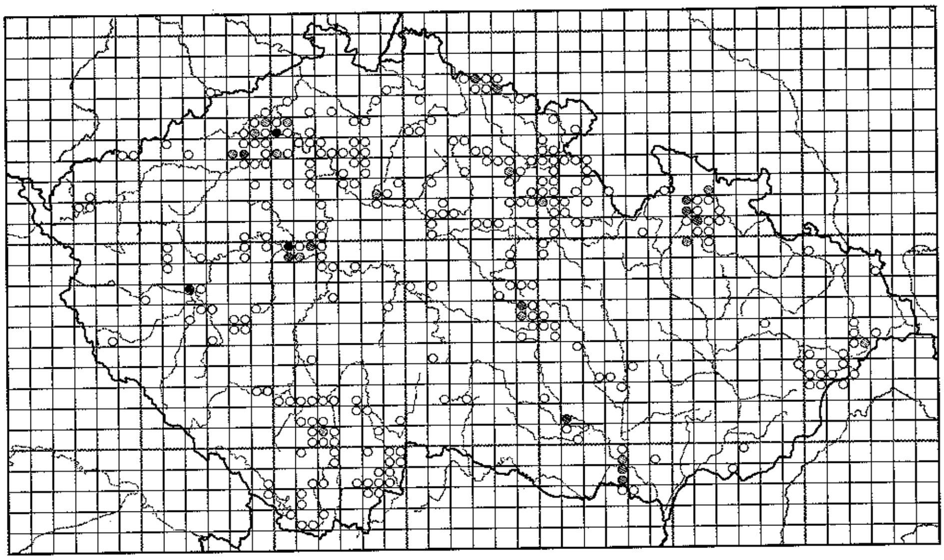


Fig. 1. The proportion of relevés made in each decade of the 20th century. Missing dates for relevés were substituted by publication dates.

However, a comparison of the number of relevés recorded at different altitudes with the total land area found shows that sampling was positively biased to the higher altitudes (> 700 m), while the middle altitudes, especially those between 400–600 m, have been under-sampled.

Assignment of the relevés to syntaxa as indicated by relevé authors (Table 3) roughly reflects the frequency of vegetation types in the Czech Republic; however, this frequency estimate may be biased by the attractiveness of individual vegetation types to researchers. Most relevés are of natural forests and semi-natural grasslands, whereas those of anthropogenous vegetation types are rarer. It is of interest to identify vegetation classes that are strongly under-represented in the database. There are less than 30 relevés of extinct types of halophytic vegetation (*Thero-Suaedetea* and *Thero-Salicornietea*), classes with fragmentary development in the Czech Republic (*Salicetea herbaceae* and *Crypsietea aculeatae*), the class *Charetea fragilis,* which is rarely sampled due to the difficulty of identifying stoneworts, and

A. 1922–1950



B. 1951–1980

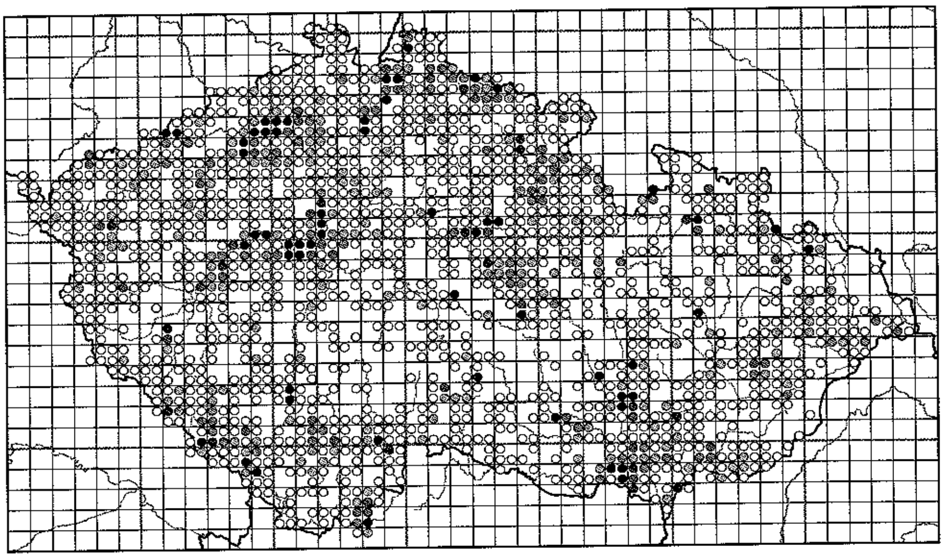
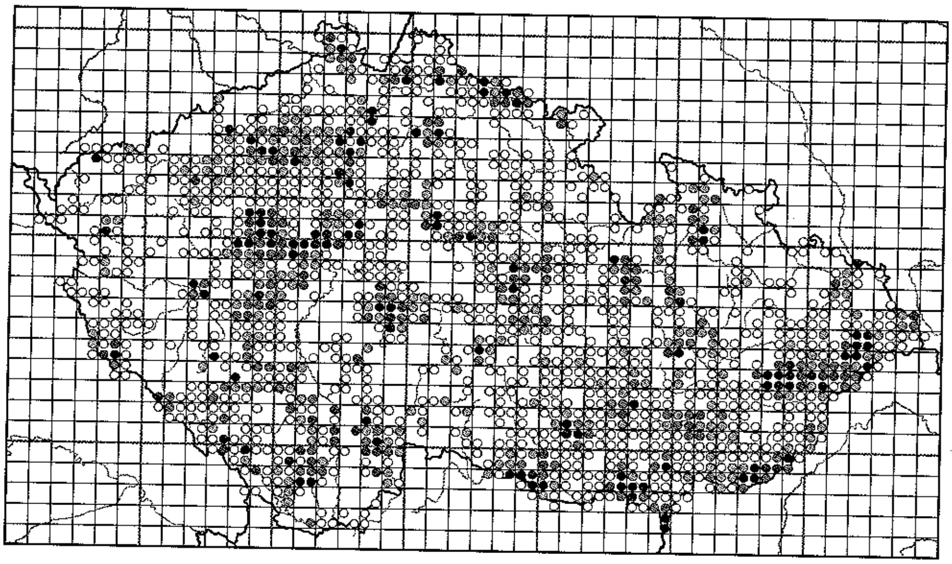


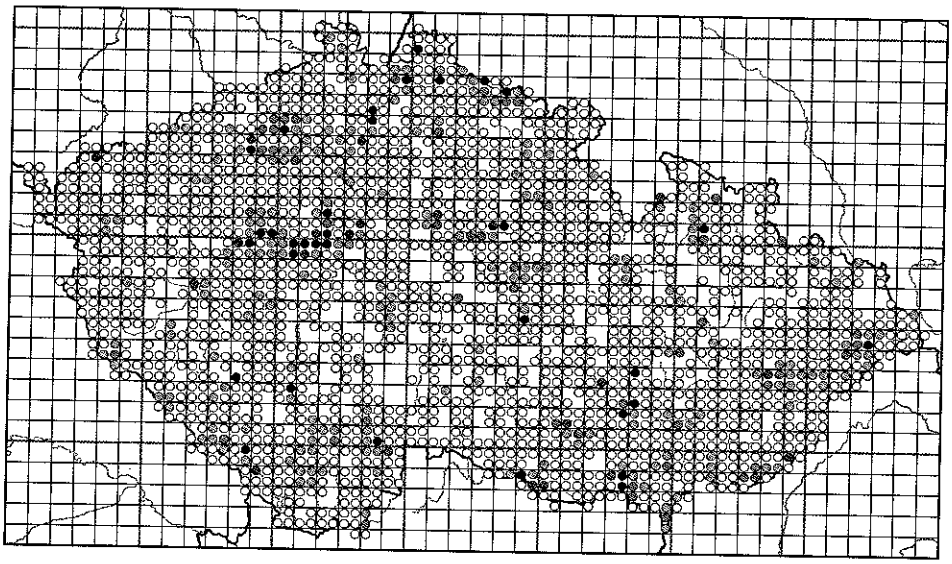
Fig. 2. Distribution of relevés for different periods of time. Missing dates for relevés were substituted by publication dates. Dark colours indicated areas with more relevés.

the class *Parietarietea*, whose syntaxonomic identity is doubtful in the Czech Republic. It is striking that 79 out of the 548 associations reported by Moravec et al. (1995) for the Czech Republic are not documented in the database and 28 of them are represented by a single relevé (Table 4). Some of these associations may be valuable for recognizing and naming

C. 1981–2002



D. 1922–2002



certain vegetation phenomena, which are poorly documented by relevés. Some others, however, are obviously not being recognized by Czech field botanists: instead of having a practical value, they perhaps make the standard syntaxonomic system used in this country unnecessarily complex. Thus, phytosociological databases can help identify syntaxa that are in need of syntaxonomic revision.

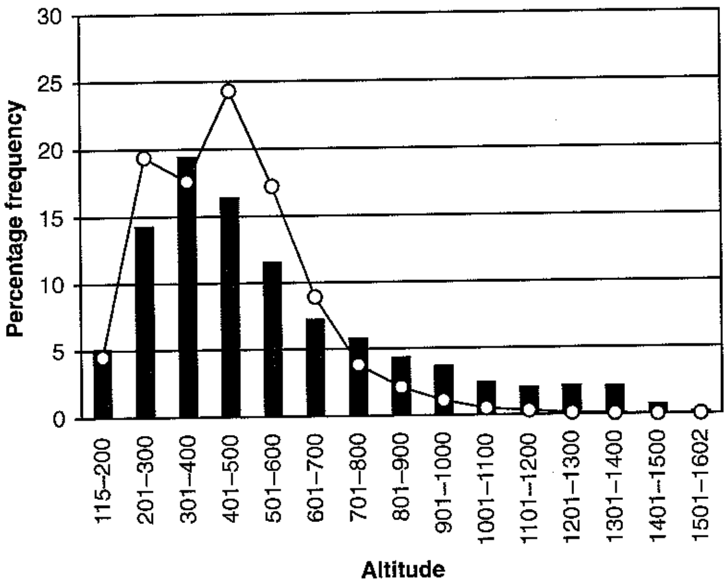


Fig. 3. Altitudinal distribution of relevés. Bars – percentage of the relevés recorded for each 100 m interval of altitude; open circles: percentage of land area at each 100 m interval of altitude. Only relevés with an indication of altitude are included.

The database is also a valuable source of floristic information. The average number of species per relevé is 23, making up a total of 1,259,008 floristic records, which include 1,177,013 records of vascular plants and 81,995 records of bryophytes and lichens. The most frequent species, including aggregate species, are those of mesic grasslands, native species with an expansive tendency (apophytes) and some native trees. The list of the most frequent species in the database, presented in Table 5, provides a rough estimate of the most frequent species in the Czech flora at the scale of vegetation plots.

The Database of Forest Typology of the Forest Management Institute (ÚHÚL) contained 14,130 relevés on 20 November 2002. It is estimated that the number of relevés available in the field protocols of this institute yet to be included in the database is 45,000 (Table 1), however, about 20–30% of them are repeated relevés in previously sampled plots (V. Zouhar, personal communication). The oldest relevés are from 1950; most were recorded between 1950–1970 (ca. 70–80%) and between 1990–2002 (ca. 20%). The relevés of the Forest Management Institute rather regularly cover the area of forest in the Czech Republic and the location of individual plots is precisely described by reference to forest sectors. They often contain excellent descriptions of habitat characteristics, but the quality of species recording is variable.

Data quality

Relevés in the phytosociological databases are collected from different sources which vary in sampling procedures and quality. Serious but hardly avoidable errors are introduced by overlooking or misidentifying some species in the field. Observer bias in estimating cover

Table 3. – Fifteen most frequent classes, alliances and associations in the database, and the number of relevés.

A) Classes			
Querco-Fagetea	9221	*Nardo-Callunetea*	1528
Molinio-Arrhenatheretea	7792	*Plantaginetea majoris*	1511
Festuco-Brometea	5755	*Secalietea*	1340
Phragmito-Magnocaricetea	4544	*Potametea*	1225
Vaccinio-Piceetea	1914	*Mulgedio-Aconitetea*	890
Scheuchzerio-Caricetea fuscae	1862	*Sedo-Scleranthetea*	805
Galio-Urticetea	1837	*Oxycocco-Sphagnetea*	795
Chenopodietea	1564		
B) Alliances			
Calthion	3322	*Caricion gracilis*	1321
Fagion	2631	*Luzulo-Fagion*	1254
Festucion valesiacae	1717	*Piceion excelsae*	1094
Arrhenatherion	1643	*Tilio-Acerion*	1027
Bromion erecti	1623	*Quercion pubescenti-petraeae*	917
Carpinion	1590	*Molinion*	867
Alnion incanae	1369	*Polygonion avicularis*	777
Phragmition communis	1329		
C) Associations			
Melampyro nemorosi-Carpinetum	881	*Calamagrostio villosae-Piceetum*	413
Dentario enneaphylli-Fagetum	845	*Angelico-Cirsietum oleracei*	379
Arrhenatheretum elatioris	679	*Phragmitetum communis*	375
Caricetum gracilis	635	*Glycerietum maximae*	314
Cirsietum rivularis	506	*Luzulo albidae-Quercetum*	289
Scirpetum sylvatici	469	*Carici pilosae-Carpinetum*	278
Aceri-Carpinetum	467	*Calamagrostio villosae-Fagetum*	276
Luzulo-Fagetum	451		

may also be serious (Lepš & Hadincová 1992, Klimeš et al. 2001). In addition, different authors use different taxonomic concepts of many species, subspecies or aggregate species. Therefore, species or subspecies often need to be transformed to species sensu lato or aggregated species in order to equalize taxonomic concepts in the data sets compiled from relevés by different authors. Such transformations result in a loss of information, but are necessary in order to avoid biases in the analyses.

Another issue strongly related to data quality is the preferential selection of relevé sites, which means that the authors tend to select stands that fit their a priori idea of a particular vegetation type and avoid other stands. Consequently, phytosociological databases contain biased samples of vegetation diversity in certain areas. In a large database that contains relevés made by many authors who had different preferences for selection of the site, this may result in a mere noise rather than a systematic bias. However, the paradigms accepted by most researchers in a given geographical area and time may cause a serious systematic bias (Frey 1995). A clear bias in phytosociological databases is the under-representation of species-poor vegetation stands. Although species-poor stands may be common, researchers often tend to avoid them, believing that they are difficult to classify (Wołek 1997). In the Czech National Phytosociological Database, the bias towards species-rich stands is indicated by the non-increasing or unimodal species-area curves for some vegetation types (Chytrý 2001).

Table 4. – The associations included in the list of Czech syntaxa (Moravec et al. 1995) but not represented or only documented by a single relevé in the database. Associations documented by a single relevé are indicated by an asterisk.

*Asplenieteta trichomanis: Biscutello-Asplenietum septentrionalis, Asplenietum septentrionali-adianti-nigri, Saxifrago paniculatae-Agrostietum alpinae**, *Hedysaro hedysaroidis-Molinietum**;

Parietarietea: Corydalidetum luteae;

*Juncetea trifidi: Carici rigidae-Juncetum trifidi**;

Salicetea herbaceae: Polytrichetum sexangularis;

Charetea fragilis: Charetum braunii, Charetum asperae, Charetum globularis, Charetum canescentis;

*Potametea: Nupharetum pumilae**, *Nymphoidetum peltatae**, *Potametum colorati, Groenlandietum densae**, *Parvopotamo-Zannichellietum pedicellatae, Najadetum minoris, Sparganio-Potametum pectinati, Batrachio trichophylli-Callitrichetum cophocarpae, Batrachietum rionii**;

Isoëto-Littorelletea: Isoëtetum echinosporae, Isoëtes lacustris comm., *Sparganium angustifolium* comm., *Pilularietum globuliferae;*

Utricularietea: Scorpidio-Utricularietum minoris;

*Isoëto-Nanojuncetea: Lindernio-Eleocharitetum ovatae, Centunculo-Anthoceretum punctati**, *Junco tenageiae-Radioletum linoidis, Samolo-Cyperetum fusci;*

*Phragmito-Magnocaricetea: Typhetum laxmannii, Schoenoplectetum tabernaemontani**, *Butomo-Alismatetum lanceolati, Helosciadietum, Catabrosetum aquaticae;*

*Montio-Cardaminetea: Caltho minoris-Philonotidetum seriatae**, *Cardaminetum amarae, Allio sibirici-Cratoneuretum filicini**;

*Scheuchzerio-Caricetea fuscae: Eleocharitetum pauciflorae, Drepanoclado revolventis-Caricetum lasiocarpae, Amblystegio scorpioidis-Caricetum limosae, Amblystegio stellati-Caricetum dioicae**, *Drepanoclado revolventis-Caricetum diandrae, Carici limosae-Sphagnetum contorti, Carici filiformis-Sphagnetum apiculati**, *Carici chordorrhizae-Sphagnetum apiculati;*

*Molinio-Arrhenatheretea: Meo athamantici-Cirsietum heterophylli, Cirsio heterophylli-Alchemilletum acutilobae, Trifolio repentis-Veronicetum filiformis, Veronico longifoliae-Filipenduletum**, *Cnidio-Violetum elatioris, Lysimachio-Filipenduletum picbaueri**, *Stachyo palustris-Thalictretum flavae**;

*Sedo-Scleranthetea: Airetum praecocis, Arabidopsietum thalianae, Hypno tamariscini-Festucetum duriusculae, Diantho deltoidis-Armerietum, Thymo angustifolii-Festucetum ovinae, Saxifrago tridactylitae-Poëtum compressae**, *Poo badensis-Allietum montani, Sempervivetum soboliferi;*

Festuco-Brometea: Minuartio setaceae-Thymetum angustifolii, Adonido-Brachypodietum pinnati;

*Trifolio-Geranietea sanguinei: Geranio-Anemonetum sylvestris**, *Vicietum sylvaticae**;

*Rhamno-Prunetea: Roso gallicae-Prunetum**, *Prunetum mahaleb;*

Salicetea purpureae: Agrostio-Salicetum purpureae;

*Erico-Pinetea: Cytiso-Pinetum**;

Vaccinio-Piceetea: Cladonio rangiferinae-Pinetum sylvestris;

Robinietea: Balloto nigrae-Robinietum;

Epilobietea angustifolii: Arctietum nemorosi, Avenello-Molinietum caeruleae, Calamagrostio villosae-Franguletum;

Bidentetea tripartitae: Pulicario vulgaris-Bidentetum;

*Chenopodietea: Atriplici-Chenopodietum crassifolii, Puccinellio-Chenopodietum glauci, Malvetum pusillae, Chenopodietum muralis, Matricario-Anthemidetum cotulae, Brometum sterilis**, *Lepidietum drabae, Atriplex rosea* comm., *Amarantho-Fumarietum, Rorippo-Chenopodietum polyspermi, Setarietum viridi-verticillatae, Panico sanguinalis-Eragrostietum minoris, Hibisco-Eragrostietum, Digitario-Portulacetum**, *Eragrostio poaeoidis-Panicetum capillaris;*

*Artemisietea vulgaris: Xanthietum spinosi, Lappulo-Cynoglossetum**;

Galio-Urticetea: Aristolochio-Cucubaletum bacciferi, Convolvulo-Epilobietum hirsuti, Carduus crispus comm.*, *Petasitetum officinali-glabrati, Torilidetum japonicae, Cephalarietum pilosae, Artemisio-Melilotetum albae, Rumici obtusifolii-Chenopodietum, Urtico-Heracleetum mantegazziani;*

*Agropyretea repentis: Cynodontetum dactyli, Poëtum pratensi-compressae**;

*Plantaginetea majoris: Rumici crispi-Agrostietum stoloniferae**, *Rumici crispi-Agropyretum, Blysmo-Juncetum compressi**, *Cynodonto-Plantaginetum majoris;*

Secalietea: Scleranthо annui-Arnoseridetum minimae.

Table 5. – Forty most frequent species or species aggregates in the database and the number of occurrences. Double or triple records in different vegetation layers of a single relevé are treated as a single occurrence.

Achillea millefolium agg.	11660	*Arrhenatherum elatius*	5596
Poa pratensis (incl. *P. angustifolia*)	8509	*Fagus sylvatica*	5587
Urtica dioica	8083	*Poa nemoralis*	5542
Taraxacum sect. *Ruderalia*	8080	*Alopecurus pratensis*	5240
Deschampsia cespitosa	7343	*Agrostis capillaris*	5179
Festuca rubra agg.	7191	*Sorbus aucuparia*	5170
Ranunculus repens	6706	*Acer pseudoplatanus*	5096
Ranunculus acris	6683	*Potentilla erecta*	5002
Plantago lanceolata	6659	*Elymus repens*	4985
Veronica chamaedrys	6570	*Lathyrus pratensis*	4942
Rumex acetosa	6528	*Holcus lanatus*	4792
Picea abies	6510	*Rubus idaeus*	4587
Oxalis acetosella	6304	*Cirsium arvense*	4564
Dactylis glomerata	6282	*Hypericum perforatum*	4442
Anthoxanthum odoratum	6175	*Sanguisorba officinalis*	4357
Vaccinium myrtillus	5973	*Trifolium repens*	4342
Avenella flexuosa	5951	*Lychnis flos-cuculi*	4336
Senecio ovatus	5895	*Lotus corniculatus*	4302
Euphorbia cyparissias	5858	*Aegopodium podagraria*	4165
Poa trivialis	5686	*Galium aparine*	4139

In addition to the effect of preferential sampling on the scale of landscape, phytosociological data in the Czech Republic are also affected by the irregular distribution of sites within the country (Fig. 2). The bulk of the data originates from small areas which have been frequently sampled, while there are few or no data from other areas, where the vegetation type in question may be rather common. For example, there are 42,945 relevés in the database that have both an indication of syntaxon and a precise location by geographical coordinates. If we permit no more than one relevé of each syntaxon per grid square of 0.5 longitudinal × 0.3 latitudinal minute (ca. 0.6 × 0.55 km), there are 15,987 superfluous relevés. Such a clumped distribution of relevés raises the problem of spatial autocorrelation (Fortin et al. 1989) with samples from close sites inherently similar. Consequently, analyses of the data from the database tend to be biased towards local peculiarities of the over-sampled areas, and classifications may reflect discontinuities in the data rather than discontinuities in the field. These trends can be weakened, but not completely removed, by stratified selection of relevés prior to analyses, such as the one mentioned above.

A striking problem is the lack of values for some variables. Mucina et al. (2000b) attempted to standardize the sampling procedure by proposing obligatory and optional data elements associated with relevés. Of these obligatory data elements, the relevés in the Czech National Phytosociological Database frequently lack field-book number, total cover of all vegetation layers, heights of the vegetation layers, and for aquatic vegetation the separate cover of emergent, natant and submerged plants. Many relevés do not describe the habitat characteristics. The accuracy of the location varies from an exact location measured by a Global Positioning System as in some recent relevés to an indication of a broad geographic area such as a mountain range. Geographic site description of most relevés permits the identification of the location to between 50–500 m, which is sufficient for syntheses on national and regional scales. For these relevés the Greenwich coordinates were added a posteriori; currently 45,269 relevés, i.e. 83.4% of the database, have coordinates,

Table 6. – The percentages of missing values of different variables in the relevés in the database. Percentages of missing cover values are calculated relative to the number of relevés that include a particular layer.

Variable	Percentage of missing values
Year of sampling	15.4%
Plot size	15.8%
Altitude	38.0%
Tree layer cover	19.0%
Shrub layer cover	22.2%
Herb layer cover	21.3%
Cryptogam layer cover	18.0%
At least one of the above values is missing	55.7%

and the rest are difficult to locate with this accuracy. Relevés without any indication of locality are not included in the database. Table 6 shows the proportions of missing values for the other obligatory data elements proposed by Mucina et al. (2000b). It is striking that the number of relevés containing all the obligatory data elements is rather low (see the bottom row of the table).

Although the poor data quality limits the kinds of analyses that can be done, there is a need to make a full use of the existing data for rapid assessment in biodiversity conservation and landscape planning. In some cases, improvements can be achieved by stratified relevé selection from the database and subsequent standardization of the data set. However, future vegetation surveys should pay more attention to an a priori stratification of field sampling procedures according to statistically well founded plans (Austin & Heyligers 1989).

Future outlook

In the near future, missing relevés from publications and theses will be computerized. The inclusion of this new data is unlikely to substantially change the basic structure of the database (e.g. proportion of individual syntaxa, proportion of relevés from different periods, territorial coverage). After that, the database will grow mainly due to incorporation of newly recorded relevés.

The major project using the database is the synthetic vegetation classification of the Czech Republic, which started on 1 January 2002 (project GAČR 206/02/0957). In addition, the database is available for non-commercial use by the scientific community in the Czech Republic and abroad. It will also be linked to existing databases in other European countries as a part of the current initiative SynBioSys Europe (www.synbiosys.alterra.nl/eu/), which was established by the working group European Vegetation Survey in October 2001. This initiative, managed by a steering committee headed by Dr. Joop H. J. Schaminée (Alterra Green World Research, Wageningen, the Netherlands), is directed towards the development of an information system for evaluation and management of the biodiversity of plant species, vegetation types and landscapes in Europe. This initiative will not only contribute Czech data for synthetic studies at the European level, but will also provide a new framework for the assessment of the vegetation and landscape diversity of the Czech Republic.

Souhrn

Od 90. let 20. století probíhá v několika institucích v Evropě i jinde ve světě shromažďování velkých souborů fytocenologických snímků a jejich převádění do elektronické podoby. Jen ve střední a západní Evropě existuje v současné době nejméně 750 000 elektronických snímků. Tato data jsou nenahraditelným zdrojem informací o diverzitě vegetace a flóry v měřítku krajiny.

Česká národní fytocenologická databáze byla založena v roce 1996 díky podpoře mezinárodní pracovní skupiny Evropský přehled vegetace. Centrální databáze je spravována na katedře botaniky PřF MU v Brně. Kromě centrální databáze existují dílčí databáze jednotlivců a pracovních skupin v jiných institucích, odkud jsou dílčí data předávána do centrální databáze a naopak centrální databáze poskytuje partnerům výběry dat pro konkrétní projekty. Průběžně aktualizované organizační a technické informace včetně softwarové podpory jsou dostupné na internetové adrese www.sci.muni.cz/botany/dbase_cz.htm.

K 15. listopadu 2002 obsahovala centrální databáze celkem 54 310 fytocenologických snímků od 322 autorů z let 1922–2002 (tab. 1, 2, obr. 1). Největší část snímků pocházela z publikovaných monografií a článků, zejména v časopisech Preslia a Folia geobotanica et phytotaxonomica. Další snímky pak byly převzaty z diplomových a disertačních prací, inventarizačních průzkumů chráněných území a také z nepublikovaných terénních zápisníků. Odhadem lze v České republice předpokládat existenci dalších asi 20.000 snímků dosud nezahrnutých v databázi, ovšem publikovaných nebo dostupných v rukopisech uložených ve veřejných knihovnách je patrně jen méně než polovina z tohoto odhadu. Rozmístění snímků na území České republiky je dosti nerovnoměrné (obr. 2). Více jich pochází z botanicky atraktivních oblastí s velkou diverzitou přirozené a polopřirozené vegetace, zatímco z některých území údaje zcela chybějí. Většina snímků pochází z nadmořských výšek 200–600 m (obr. 3).

Syntaxonomické spektrum snímků obsažených v databázi (tab. 3) částečně odráží relativní hojnost jednotlivých vegetačních typů v České republice, větší množství snímků některých tříd je ale spíše jen důsledkem jejich atraktivity pro badatele. Většina snímků pochází z přirozené lesní nebo polopřirozené travinné vegetace. Méně než 30 snímků je k dispozici pro vymizelé typy halofilní vegetace (třídy *Thero-Suaedetea* a *Thero-Salicornietea*) a pro fragmentárně vyvinuté nebo z různých důvodů problematické třídy *Parietarietea, Salicetea herbaceae, Charetea fragilis* a *Crypsietea aculeatae*. Z asociací, které uvádějí Moravec et al. (1995), není 79 dokumentováno ani jedním a 28 pouze jedním snímkem (tab. 4). Některé z těchto asociací mohou být užitečné pro pojmenování určitých nápadných částí vegetačního kontinua, ale jsou zatím málo doloženy snímky. Mnohé další asociace z tohoto seznamu by si však zasloužily syntaxonomickou revizi, protože patrně nemají praktický význam a spíše přispívají ke zbytečné složitosti u nás používaného systému vegetačních jednotek.

Snímky dosud zadané do databáze obsahují celkem 1 177 013 jednotlivých údajů o výskytu druhů cévnatých rostlin a 81 995 údajů o výskytu mechorostů a lišejníků. Srovnání frekvence zastoupení různých druhů v databázi dává hrubou představu o nejhojnějších druzích v české krajině (tab. 5).

Kromě České národní fytocenologické databáze existuje v České republice také Databáze typologie lesů, která je od roku 2001 vytvářena v Ústavu pro hospodářskou úpravu lesů (ÚHÚL), rovněž s využitím programu Turboveg. V současné době je v této databázi uloženo celkem 14 130 typologických zápisů, přičemž se předpokládá, že celkový počet dalších zápisů existujících v terénních protokolech ústavu dosahuje až 45 000, i když pravděpodobně asi čtvrtina z nich jsou opakované zápisy na dříve dokumentovaných plochách.

Problémem každé fytocenologické databáze je nevyrovnaná kvalita dat. Nejzávažnější jsou chyby vzniklé při záznamu druhů v terénu, využití dat je však ztíženo také rozmanitostí taxonomických koncepcí druhů, které uplatňují různí autoři snímků. Dalším problémem je preferenční sběr dat, kdy badatelé vybírají ke snímkování jen určité porosty podle apriorních schémat, zatímco jiné, zejména druhově chudé porosty, zpravidla přehlížejí. Analýzy dat mohou být také zkreslovány tím, že některé vybrané lokality byly dokumentovány velmi intenzívně, zatímco jiná, často rozsáhlá území, nebyla snímkována vůbec. Problémy vznikají také kvůli málo přesným lokalizacím nebo chybějícím údajům z hlaviček snímků (tab. 6). Vzhledem k těmto omezením musí být data pro každou analýzu pečlivě vybírána a zpravidla standardizována, obvykle za cenu ztráty části informací.

Česká národní fytocenologická databáze bude v nejbližší době doplněna o zbývající fytocenologické snímky z publikací a veřejně dostupných rukopisů a poté bude růst hlavně díky nově zaznamenaným snímkům. Bude použita pro začínající syntetický projekt klasifikace vegetace České republiky a současně bude k dispozici pro různé jiné vědecké projekty. Bude propojena s ostatními fytocenologickými databázemi existujícími v Evropě v rámci mezinárodního projektu evropského informačního systému o vegetaci a krajině „SynBioSys Europe" (www.synbiosys.alterra.nl/eu/).

Acknowledgements

Our thanks are due to all those colleagues who helped us in different ways to organize the basic structure of the database, particularly J. Dring, W. Gutermann, H. Härtel, S.M. Hennekens, I. Jarolímek, T. Kučera, J. Moravec, L. Mucina, Z. Neuhäuslová, H. Niklfeld, J.S. Rodwell, L. Tichý, M. Valachovič and J. Wild, and to many others who contributed data or helped with the electronic recording of the data: D. Blažková, L. Bravencová, D. Cigánek, J. Čáp, N. Čeplová, T. Černý, J. Danihelka, M. Derková, M. Duchoslav, H. Dušková, K. Dvořáčková, V. Eltsova, M. Fabšičová, T. Grulich, M. Hájek, P. Hájková, P. Hanáková, M. Havlová, H. Havránková, E. Hettenbergerová, A. Hoffmann, V. Horáková, J. Chlapek, J. Jirásek, J. Juřička, K. Karimová, P. Karlík, L. Klimeš, M. Kočí, M. Koloničná, F. Krahulec, Z. Kropáč, T. Kučera, K. Kučerová, P. Kuneš, I. Kuželová, Z. Lososová, P. Lustyk, J. Němec, Z. Neuhäuslová, J. Novák, I. Ostrý, Z. Otýpková, M. Paloudová, P. Pavlíčková, J. Pekárová, P. Petřík, M. Pokorný, P. Pyšek, J. Rozehnal, J. Sádlo, K. Šumberová, L. Tichý, J. Tkačíková, J. Vicherek, J. Vojta, T. Vymyslický, P. Vysloužil, and E. Vysloužilová. Special thanks are due to R. Hédl who computerized ca. 3500 relevés of forest vegetation and to the team of P. Pyšek, including I. Ostrý, J. Pergl and I. Koukolíková, who computerized ca. 3000 relevés of synanthropic vegetation. V. Zouhar gave us information on the Database of Forest Typology and P. Pyšek provided GIS-based data on the land areas found at different altitudes. We appreciated the comments on an earlier version of this paper by T. Kučera and L. Tichý. The database was developed and this paper was prepared with funding from the projects GAČR 206/99/1523 and MSM 143100010.

References

Austin M.P. & Heyligers P.C. (1989): Vegetation survey design for conservation: Gradsect sampling of forests in North-eastern New South Wales. – Biol. Conserv. 50: 13–32.

Bibliographia (1983–1992): Bibliographia syntaxonomica čechoslovaca ad annum 1970. Vols. 1–20. – Botanický ústav ČSAV, Průhonice.

Braun-Blanquet J. (1964): Pflanzensoziologie. Grundzüge der Vegetationskunde. Ed. 3. – Springer, Berlin, Wien, New York.

Brisse H., de Ruffray P., Grandjouan G. & Hoff M. (1995): The Phytosociological Database "SOPHY" Part 1: Calibration of indicator plants, Part II: Socio-ecological classification of the relevés. – Ann. Bot. (Roma) 53: 177–223.

Bruelheide H. & Chytrý M. (2000): Towards unification of national vegetation classifications: a comparison of two methods for analysis of large data sets. – J. Veg. Sci. 11: 295–306.

Chytrý M. (1996): Databázový systém pro projekt přehledu vegetace České republiky (Database system for the project of the vegetation survey of the Czech Republic). – Zpr. Čes. Bot. Společ. 31: 193–200.

Chytrý M. (1997): Česká národní fytocenologická databáze: počáteční stav a perspektivy (Czech National Phytosociological Database: initial state and perspectives). – Zpr. Čes. Bot. Společ., Mater. 15: 27–40.

Chytrý M. (2001): Phytosociological data give biased estimates of species richness. – J. Veg. Sci. 12: 439–444.

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

Ehrendorfer F. (ed.) (1973): Liste der Gefäßpflanzen Mitteleuropas. Ed. 2. – G. Fischer, Stuttgart.

Ewald J. (1995): Eine vegetationskundliche Datenbank bayerischer Bergwälder. – Hoppea 56: 453–465.

Ewald J. (2001): Der Beitrag pflanzensoziologischer Datenbanken zur vegetationsökologischen Forschung. – Ber. R.-Tüxen-Ges. 13: 53–69.

Font X. & Ninot J.-M. (1995): A regional project for drawing up inventories of flora and vegetation in Catalonia (Spain). – Ann. Bot. (Roma) 53: 99–105.

Fortin J.-M., Drapeau P. & Legendre P. (1989): Spatial autocorrelation and sampling design in plant ecology. – Vegetatio 83: 209–222.

Frey H.-U. (1995): Waldgesellschaften und Waldstandorte im St. Galler Berggebiet. – Veröff. Geobot. Inst. ETH, Stift. Rübel 126: 1–280.

Frey W., Frahm J.P., Fischer E. & Lobin W. (1995): Die Moos- und Farnpflanzen Europas. – G. Fischer, Stuttgart.

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

Klimeš L., Dančák M., Hájek M., Jongepierová I. & Kučera T. (2001): Scale-dependent biases in species counts in a grassland. – J. Veg. Sci. 12: 699–704.

Lepš J. & Hadincová V. (1992): How reliable are our vegetation analyses? – J. Veg. Sci. 3: 119–124.

Michener W.K. & Brunt J.W. (2000): Ecological data: design, management and processing. – Blackwell Science, Oxford.

Moravec J., Balátová-Tuláčková E., Blažková D., Hadač E., Hejný S., Husák Š., Jeník J., Kolbek J., Krahulec F., Kropáč Z., Neuhäusl R., Rybníček K., Řehořek V. & Vicherek J. (1995): Rostlinná společenstva České republiky a jejich ohrožení (Red list of plant communities of the Czech Republic and their endangerment). Ed. 2. – Severočes. Přír., suppl. 1995: 1–206.

Mucina L., Bredenkamp G.J., Hoare D.B. & McDonald D.J. (2000a): A national vegetation database for South Africa. – S. Afr. J. Sci. 96: 497–498.

Mucina L., Rodwell J. S., Schaminée J. H. J. & Dierschke H. (1993): European Vegetation Survey: current state of some national programmes. – J. Veg. Sci. 4: 429–439.

Mucina L., Schaminée J. H. J. & Rodwell J. S. (2000b): Common data standards for recording relevés in field survey for vegetation classification. – J. Veg. Sci. 11: 769–772.

Neuhäuslová Z. & Kolbek J. (eds.) (1982): Seznam vyšších rostlin, mechorostů a lišejníků střední Evropy užitých v bance geobotanických dat BÚ ČSAV (A list of higher plants, bryophytes and lichens of Central Europe used in the Bank of Geobotanical Data in the Botanical Institute of Czechoslovak Academy of Sciences). – Botanický ústav ČSAV, Průhonice.

Pišút I., Lackovičová A. & Lisická E. (1993): Súpis lišajníkov Slovenska (A list of lichens of Slovakia). – Biológia 48, Suppl. 1: 53–98.

Rodwell J. S. (1995): The European Vegetation Survey questionnaire: an overview of phytosociological data, vegetation survey programmes and databases in Europe. – Ann. Bot. (Roma) 53: 87–98.

Rodwell J. S., Mucina L., Pignatti S., Schaminée J. H. J. & Chytrý M. (1997): European Vegetation Survey: the context of the case studies. – Folia Geobot. Phytotax. 32: 113–115.

Rodwell J. S., Pignatti S., Mucina L. & Schaminée J. H. J. (1995): European Vegetation Survey: update on progress. – J. Veg. Sci. 6: 759–762.

Schaminée J. H. J. & Hennekens S. M. (1995): Update of the installation of Turboveg in Europe. – Ann. Bot. (Roma) 53: 159–161.

Schaminée J. H. J. & Stortelder A. H. F. (1996): Recent developments in phytosociology. – Acta Bot. Neerl. 45: 443–459.

Tichý L. (2002): JUICE, software for vegetation classification. – J. Veg. Sci. 13: 451–453.

Valachovič M. (1996): TURBOVEG – celoeurópsky databázový program pre fytocenologické dáta (TURBOVEG – a pan-European database program for phytosociological data). – Bull. Slov. Bot. Spoloč. 18: 189.

Westhoff V. & van der Maarel E. (1973): The Braun-Blanquet approach. – In: Whittaker R. H. (ed.), Ordination and classification of plant communities, W. Junk, The Hague, p. 617–737.

Wohlgemuth T. (1992): Die vegetationskundliche Datenbank. – Schweiz. Z. Forstwes. 143: 22–36.

Wołek J. (1997): Species co-occurrence patterns in pleustonic plant communities (class *Lemnetea*). Are there assembly rules governing pleustonic community assembly? – Fragm. Flor. Geobot., Suppl. 5: 3–100.

Zlatník A. (1928a): Études écologiques et sociologiques sur le *Sesleria coerulea* et le *Seslerion calcariae* en Tchécoslovaquie. – Rozpr. Král. Čes. Společ. Nauk, cl. math.-natur., 8/1: 1–116.

Zlatník A. (1928b): Apercu de la végétation des Krkonoše (Riesengebirge). – Preslia 7: 94–152.

Received 18 January 2002
Revision received 3 December 2002
Accepted 7 January 2003