Journal of Vegetation Science 16: 407-414, 2005

Supervised classification of plant communities with artificial neural networks

Černá, Lenka[1,2] & Chytrý, Milan[1*]

[1]*Department of Botany, Masaryk University, Kotlářská 2, CZ-611 37 Brno, Czech Republic;*
[2]*E-mail lela2@seznam.cz; *Corresponding author; Fax +420 541211214; E-mail chytry@sci.muni.cz*

Abstract

Questions: Are artificial neural networks useful for the automatic assignment of species composition records from vegetation plots to *a priori* established classes (vegetation units)? Is the assignment more accurate (1) if the classes are defined by numerical classification rather than by expert-based classification; (2) if the training data set is selected to include plots that are richer in diagnostic species of particular classes?

Material: Species composition records (relevés) from 4186 plots of Czech grasslands.

Methods: Plots were classified into 11 phytosociological alliances (expert classification) and into 11 clusters derived from numerical cluster analysis. Some plots were used for training the classifiers, which were the multi-layer perceptrons (MLP; a type of artificial neural network). Other plots were used for testing the performance of these classifiers. Plots used for training were selected (1) randomly; (2) according to higher representation of diagnostic species of particular classes.

Results: Different MLP classifiers correctly classified 77-83% of plots to the classes of the expert classification and 70-78% to the classes of the numerical classification. The better result in the former case was mainly due to two classes in the expert classification, which were well recognized by the classifiers and at the same time contained a large proportion of the plots of the entire data set. Correct classification of the plots belonging to these large classes resulted in a good overall performance of the classifiers. After training with randomly chosen plots, the classifiers produced better results than after training with plots that contained more diagnostic species. This indicates that the biased selection of the training plots disables the classifiers to recognize the entire variation within the classes and results in errors when new plots are to be classified.

Conclusions: MLP is suitable for assigning vegetation plots to already established classes. Unlike some other methods of supervised classification, it performs well even in communities that are poor in diagnostic species. However, the method does not provide clear assignment keys that could be used for class identification in field surveys. It is therefore more appropriate in applications that aim at a reliable class assignment rather than understanding the assignment rules.

Keywords: Cluster analysis; Grassland; Multi-layer perceptron; Phytosociological data; Predictive habitat modelling; Vegetation survey.

Abbreviations: MLP = Multi-layer perceptron.

Introduction

Community ecologists and vegetation scientists routinely use a range of numerical methods for classifying species-by-sites matrices (e.g. Legendre & Legendre 1998) in order to establish community types. Ejrnæs et al. (2004) emphasised an important dichotomy between unsupervised and supervised classifications, which is commonly applied in studies of vegetation patterns based on remote sensing (Ripley 1996), but is rarely used in the context of species-based community classifications. Unsupervised classifications include both agglomerative methods of cluster analysis and divisive methods such as TWINSPAN. These methods require a minimum input from the user and produce classes by searching for patterns in the analysed data set, without considering any external information. When classification is repeated after the addition of new data to the previously classified data set, serious changes may appear, including shifts in the cluster membership of the previously classified sites. This is a disadvantage, because every new classification exercise produces a new classification system that is difficult to compare both with the established standards of existing national or international classifications and with the results of other classification exercises (Bruelheide & Chytrý 2000). In addition, unsupervised classification methods do not include procedures for new sampling units to be assigned to the previously established classes.

By contrast, supervised classification methods are learning an established classification from a training data set, which contains predictor variables measured in each sampling unit and *a priori* class assignments of the sampling units. In this way a classifier is developed which can be used to assign new sampling units to classes. There are several methods suitable for the supervised classification of community composition data, e.g. quadratic discriminant analysis (Ejrnæs et al. 2004), multinomial log-linear regression, classification trees and artificial neural networks (Ripley 1996). In this paper, we will focus on the latter.

Artificial neural networks (Ripley 1996; Lek & Guégan 1999) are computational modelling tools inspired by the structure of the human brain. They learn from experience and recognize complex patterns, predict class membership or values of different variables. One advantage of artificial neural networks is their non-parametric nature, which makes them appropriate for the analysis of nearly any kind of data irrespective of their statistical properties. Artificial neural networks have frequently been reported as giving a more accurate prediction than other supervised methods (Cairns 2001; Liu et al. 2003), but at the expense of the interpretability of the results. They represent a black-box approach which hides the underlying prediction process.

Artificial neural networks consist of a number of units called neurons, which are arranged in layers. The simplest structure consists of an input layer and an output layer of neurons, but usually also one or more hidden layers are placed between the two (Fig. 1). The neurons are interconnected by coefficients called weights, which are successively modified when the network is in operation. After feeding values from the input data set, each neuron passes its given value to the connections leading out from it, and on each connection the value is multiplied by the weight associated with that connection. Each neuron in the next layer then receives a value which is the sum of the values produced by the connections leading to it, performs a simple computation using a predefined function, and delivers the value to the neurons in the next layer (Ripley 1996).

Currently the most popular types of artificial neural networks include the Kohonen network (also called 'self-organizing feature map'; Kohonen 1982) and the multi-layer perceptron (MLP; Rumelhart et al. 1986). The Kohonen network is an unsupervised method which identifies clusters in data; it can be applied to the analysis of species composition data in a similar way as cluster analysis or ordination (Chon et al. 1996; Foody 1999a; Brosse et al. 2001; Giraudel & Lek 2001). By contrast, MLP uses supervised learning and creates a classifier by fitting output values (responses) to input data (predictors) in the training data set. Subsequently, the trained classifier may be used to predict the output values for new cases that were not contained in the training data set.

In the context of vegetation classification, MLP has mainly been used for predictive modelling of the spatial distribution of vegetation or land-cover classes from remote sensing data (Paola & Schowengerdt 1995; Foody 1996; Zhang et al. 1997; Cairns 2001). It is rarely applied to community ecology, where it is potentially suitable for predictive modelling of e.g. species richness (Guégan et al. 1998), biomass or productivity (Lae et al. 1999), performance of dominant species (Tan & Smeins 1996) or conservation priority of habitats (Ejrnæs et al. 2002). MLP can be trained to recognize *a priori* classes from either environmental variables measured in each sampling site (Hirlbert & Ostendorf 2001), a combination of environmental variables and species composition (Liu et al. 2003; Zhang et al. 2004), or pure species composition (Ejrnæs et al. 2002).

The objective of this paper is to test the efficiency of the multi-layer perceptron as a tool for the supervised classification of species composition data in vegetation science. We will focus on comparing its performance when the *a priori* classification is based either on expert knowledge or cluster analysis, and when training is done either with a set of randomly selected sampling units or with a set of units that are considered as the best representatives of particular classes.

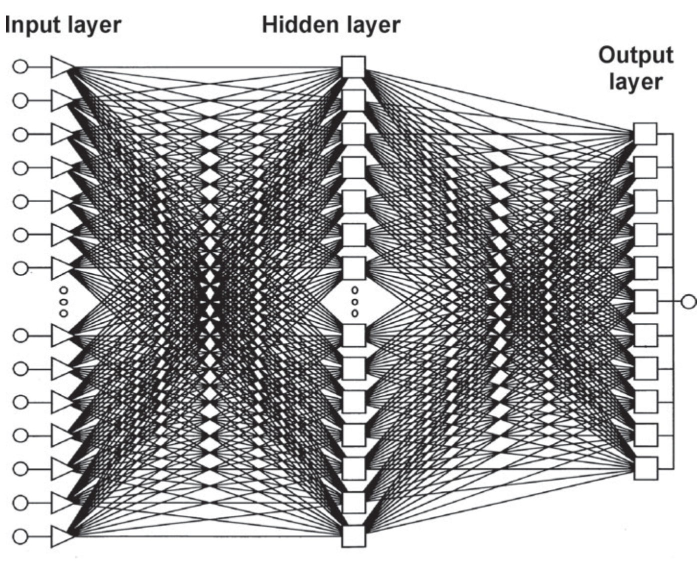

Fig. 1. Schematic architecture of the MLP network with one hidden layer, used to predict the assignment of vegetation plots to one of 11 classes. Input data are cover values of species; the number of neurons in the input layer (triangles) is equal or lower than the number of species in the data set. Output is the plot assignment to one of 11 classes.

Material and Methods

Data sets and a priori classifications

The basic data set used for developing MLP classifiers included 4186 plot records of species composition of semi-natural grasslands from the Czech Republic, taken from the Czech National Phytosociological Database (Chytrý & Rafajová 2003). As plots were irregularly distributed across the national territory, the selection of the data set was based on a geographically stratified resampling which deleted some randomly selected plots in oversampled areas. Plots < 4 m^2 and > 100 m^2 were excluded. To reduce noise in the data, the records of cryptogams, juvenile trees and species with less than five occurrences were deleted. After this reduction, 598 species were included in the analyses. Species cover values were transformed to the nine-degree ordinal scale (van der Maarel 1979).

We used two *a priori* classifications that were subsequently used to train and test the MLP classifier. The first classification ('*expert classification*') was a traditional phytosociological classification, entirely based on expert knowledge. We subdivided the data set of vegetation plots into 11 phytosociological alliances of the standard national vegetation classification (Moravec et al. 1995), using the alliance assignments given by the authors of the individual plot records. The ecologically and floristically closely related alliances *Cnidion* and *Veronico-Lysimachion*, which were represented by few plots, were merged. The second classification ('*numerical classification*') was prepared from the same data set using cluster analysis in the PC-ORD 4 program (McCune & Mefford 1999), with Euclidean

Table 1. Phytosociological and ecological interpretation of classes in expert and numerical classification, numbers of plots in the individual classes and subsets, and performances of the best MLP classifiers in each of the two variants of the selection of the training data set (i.e., either random or by diagnostic species). The best classifiers shown in this table are those marked by * in Table 2.

					Random selection				Selection by diagnostic species			
		No. of plots in the training data set	No. of plots in the test data set	Total no. of plots	No. of test plots assigned to the group by the MLP classifier	No. of test plots correctly assigned to the group by the MLP classifier	Sensitivity	Positive predictive power	No. of test plots assigned to the group by the MLP classifier	No. of test plots correctly assigned to the group by the MLP classifier	Sensitivity	Positive predictive power
Expert classification												
1	*Arrhenatherion* – mesic meadows	138	275	825	273	235	85	86	244	220	80	90
2	*Polygono-Trisetion* – montane mesic meadows	20	40	120	37	23	58	62	42	28	70	67
3	*Cynosurion* – mesic pastures	25	49	148	53	38	78	72	48	31	63	65
4	*Alopecurion* – subatlantic lowland wet meadows	45	89	268	84	55	62	65	138	65	73	47
5	*Calthion* – submontane wet meadows	328	655	1965	670	616	94	92	627	574	88	92
6	*Cnidion* – subcontinental lowland wet meadows	9	17	52	16	14	82	88	12	11	65	92
7	*Molinion* – meadows of wet, nutrient poor sites	70	140	420	136	91	65	67	156	95	68	61
8	*Nardion* – subalpine *Nardus* grasslands	9	18	54	14	9	50	64	15	11	61	73
9	*Violion* – submontane *Nardus* grasslands	36	72	216	67	49	68	73	60	42	58	70
10	*Nardo-Juncion* – wet *Nardus* grasslands	8	8	24	10	4	50	40	12	8	100	67
11	*Nardo-Agrostion* – montane *Nardus* grasslands	16	31	94	34	23	74	68	40	25	81	63
	Total	704	1394	4186	1394	1157	83	83	1394	1110	80	80
Numerical classification												
1	Mesic to moderately dry meadows (*Arrhenatherion*)	104	209	627	191	155	74	81	182	149	71	82
2	Submontane *Nardus* grasslands (*Violion*)	90	180	540	206	149	83	72	177	129	72	73
3	Montane heathlands and *Nardus* grasslands (*Nardetalia*)	44	87	263	86	74	85	86	86	68	78	79
4	Various types of mesic to wet meadows	84	168	505	158	114	68	72	186	124	74	67
5	Wet meadows on base-poor soils (*Calthion*)	60	120	360	108	81	68	75	103	79	66	77
6	Intermittently wet meadows on nutrient-poor soils (*Molinion*)	31	62	186	57	51	82	89	67	52	84	78
7	Various types of wet meadows	113	225	676	246	165	73	67	253	149	66	59
8	Montane meadows (*Polygono-Trisetion*)	15	29	89	31	27	93	87	31	26	90	84
9	*Cirsium rivulare* wet meadows (*Calthion*)	31	63	188	60	53	84	88	58	42	67	72
10	*Scirpus sylvaticus* wet meadows (*Calthion*)	52	104	313	117	94	90	80	123	91	88	74
11	*Filipendula ulmaria* wet meadows (*Calthion*)	73	147	439	134	121	82	90	128	110	75	86
	Total	697	1394	4186	1394	1084	78	78	1394	1019	73	73

distance and Ward clustering method. We accepted 11 clusters from the resulting dendrogram, in order to use the same number of classes as in the expert classification. Ecological and phytosociological interpretations of classes of both classifications are presented in Table 1 and details of species composition in Apps. 1 and 2.

From each of these two classified data sets we selected the *training data set* and the *selection data set*, which were further used for the development of the MLP classifier (see below), and the *test data set*, which was used to evaluate model performance. We used two alternative procedures for dividing plots of the basic data sets into training, selection and test data sets. The first procedure was the random selection of plots. In the second procedure we selected as training data the most typical plots of particular classes, in order to test the hypothesis that training with typical plots would improve the quality of prediction. As a criterion for the selection of typical plots we took the number of diagnostic species of the given class occurring in each plot. Diagnostic species for each class were determined by calculating the phi coefficient of association (Chytrý et al. 2002) between each species and each class. This coefficient ranges from –1 to +1, higher values meaning that the species is more associated with the given class and can therefore be considered as a diagnostic species of the class. To avoid the dependence of the phi coefficient on the relative size of the classes within the data set, we re-calculated this coefficient for the case of equal size of all classes, which was set to 10% of the entire data set size. This procedure enabled direct comparisons of the phi coefficients between the classes that contained unequal numbers of plots. The phi coefficients were calculated using the JUICE 6.3 program (Tichý 2002). We assigned diagnostic status to those species that exceeded the subjectively selected value of $\Phi = 0.25$. For each class, we first randomly selected 1/3 of its plots and assigned them to the test data set. Then we ranked the remaining plots within each class by decreasing number of diagnostic species, and excluded half of these plots, the one which was poorer in diagnostic species. The other half was further randomly divided into two halves, and one of them was used for the training data set and the other for the selection data set. Thus, the division of plots within each class followed the ratio 1 : 1 : 2 : 2 in turn for the training, selection and test data sets and excluded plots. For the sake of comparability, the same division ratio was used for the analyses based on the random selection. In one small class of the expert classification, where the training data set would contain only four plots after this division, we moved another four plots of this class from the selection data set to the training data set.

MLP classifiers for supervised classification

Supervised classification was performed with the multi-layer perceptron (MLP; Rumelhart et al. 1986) in the STATISTICA 7.0 program (www.statsoft.com). This artificial neural network comprises one input layer of neurons, at least one hidden layer and one output layer (Fig. 1). The maximum number of neurons it contains is determined by the number of input variables; in our case, the input layer could contain up to 598 neurons, each corresponding to one species in the data set. Each vector of input values contained species cover values in a particular plot. The hidden layers encode and organize the information received from the input layer and deliver it to the output layer. The output layer contains as many neurons as there are classes in the *a priori* classification (11 in our case). During the training process, data vectors (plots) with a known class membership are submitted to the network and the output values are compared with the correct class membership. Errors identified in these comparisons are used for iterative adjustment of the weights on each connection of the network until the pre-defined error-function value decreases below a certain threshold.

One major problem of artificial neural networks is the risk of over-learning (over-fitting), especially in larger networks with more complex underlying functions. Over-learning occurs when the network is trained to minimize the error on the training data set, but at the same time looses its ability to generalize and recognize newly encountered cases. We prevented over-learning by using an independent data set, called *selection data set*, in the process of network training. The network was trained on the training data set and the error on this data set naturally dropped as the training process proceeded. At the same time, the error was measured on the selection data set. Over-learning was indicated by ceasing of dropping or, indeed, by rising of the error on the selection data set, with simultaneously continued dropping of the error on the training data set. If this situation occurred, the training process was stopped.

Another problem of artificial neural networks is the possible convergence of underlying functions to local minima. To avoid this problem, we used a two-phase training, with initial application of the back-propagation algorithm, which is less prone to stick in local minima, followed by the conjugate gradient descent algorithm. We also tested, for each model, five MLP networks with different architectures, including those with one or two hidden layers, with different number of input variables (some species with lower capacity to discriminate classes were excluded) and with different number of training epochs. Selection of the network architectures was made using the *Intelligent Problem Solver* module in the STATISTICA program.

Evaluation of the classifiers

The MLP classifications of the test data sets were compared with the *a priori* classifications of the same data sets using the concepts of sensitivity and positive predictive power (Fielding & Bell 1997). Sensitivity is the probability of correct classification, i.e. the proportion of sampling units that belong to a particular class and have been correctly assigned to this class by the classifier. Positive predictive power is the probability that a sampling unit belongs to a particular class if the classifier assigns it to this class. Let us use the following contingency table to compare the numbers of sampling units that are correctly and incorrectly classified with respect to class i, where i is 1, 2, ..., n, and n is the number of classes in the given classification:

No. of sampling units ...	actually belonging to class i	actually not belonging to class i
classified to class i	a_i	b_i
not classified to class i	c_i	d_i

Using this notation we calculated the sensitivity of the classifier for each class i as $S_i = a_i / (a_i + c_i)$, the positive predictive power for each class i as $\mathrm{PPP}_i = a_i / (a_i + b_i)$ and the overall sensitivity for the whole classification as $S = \Sigma a_i / N$, where N is the total number of plots. The values of these variables are given in percentages throughout this paper.

Results

The MLP classifiers for the expert classification correctly classified 81-83% plots of the test data set when trained with randomly selected plots and 77-80% plots when trained with plots containing a high proportion of diagnostic species of particular classes (Table 2). These values were significantly different from each other and significantly higher than the values for the classifier for numerical classification (ANOVA, $P < 0.05$). Plot assignments to the classes of the *a priori* classification and the MLP assignments by the best classifier (i.e. that with the highest sensitivity on the test data set) within each classification type and training data set variant are compared in Table 1 and, in more detail, in Apps. 3 and 4. In the best classifier for expert classification, based on the random selection of the test data set, sensitivity for individual classes ranged between 50-94%. Poorest sensitivity and positive predictive power occurred in those classes that were represented by few plots in the training data set, while the highest values of both of these measures were reached in the large class 5 which contained almost half of the plots of the entire data set. When the classifier was trained with plots rich in diagnostic species, sensitivity for individual classes was between 58-100%, however, the overall sensitivity was lower than after training with the randomly selected plots due to a lower sensitivity of the

Table 2. Basic details of architecture and performance of the MLP classifiers tested. More details on the classifiers with the highest sensitivity on the test data set within each variant (i.e. 5, 10, 12 and 19, marked with asterisks) are reported in Table 1.

Classifier	No. of neurons			Sensitivity		
	Input layer	Hidden layer 1	Hidden layer 2	Training data set	Selection data set	Test data set
Expert classification, random selection of the training data set						
1	311	73	67	99.6	80.7	81.2
2	378	87	73	100.0	79.7	81.2
3	385	98	0	98.7	81.8	80.9
4	215	85	0	97.6	81.6	82.1
5 *	419	97	0	100.0	82.6	83.0
Expert classification, selection of the training data set by diagnostic species						
6	260	98	0	100.0	90.8	79.6
7	313	92	85	100.0	91.4	78.3
8	314	75	68	100.0	92.2	77.0
9	84	55	0	99.6	92.5	78.4
10 *	421	94	0	100.0	93.5	79.6
Numerical classification, random selection of the training data set						
11	450	100	0	100.0	79.2	76.7
12 *	425	96	0	100.0	77.0	77.8
13	312	70	64	96.1	75.0	73.6
14	388	88	82	99.7	75.6	76.1
15	251	78	0	100.0	76.0	74.1
Numerical classification, selection of the training data set by diagnostic species						
16	203	72	68	99.3	91.0	70.3
17	391	97	0	100.0	91.5	72.5
18	372	77	75	100.0	89.8	71.8
19 *	135	72	0	99.6	90.2	73.1
20	392	89	0	100.0	91.0	71.9

classifier for the large class 5 in this case.

The MLP classifiers for the numerical classification correctly classified 74-78% of plots of the test data set when trained with randomly selected plots and 70-73% of plots when trained with the plots that were rich in diagnostic species (Table 2). These values were significantly different from each other and lower than those for the expert classification (ANOVA, $P < 0.05$). Sensitivity for individual classes ranged between 68-93% after training with the randomly selected plots and between 66-90% after training with plots rich in diagnostic species (Table 1). The differences in both sensitivity and the positive predictive power among the classes were lower for the numerical classification than for the expert classification.

In both expert and numerical classification, selection of the training data set based on the higher representation of diagnostic species resulted in a higher sensitivity on the training and selection data sets, but the crucial parameter for the evaluation of the classifier performance, i.e. sensitivity on the test data set, was lower than in the case of random selection of the training data set (Table 2).

Discussion

The MLP classifiers were able to classify correctly 77-83% of plots to the classes of the expert classification and 70-78% to the classes of numerical classification (Table 1). The better results in the case of the expert classification may be surprising, considering that the classification was made subjectively by different researchers with differing opinions and experiences. Not all classes of the expert classification were equally well defined either. For example, the numerical classification indicated a low degree of differentiation of *Nardus* grasslands by including them all in two classes (2 and 3), while the expert classification divided them into four classes (8-11). Inconsistencies inherent to the expert classification resulted in the greater variation in sensitivity of the classifiers among classes, which ranged between 50-100% for expert classifications and 66-93% for numerical classifications. However, the higher overall sensitivity of the classifiers for expert classification does not mean that numerical classification is inherently worse or more difficult to reproduce with the MLP classifier. Both this study and our previous pilot studies indicate that in the classified data sets in which one class contains a high proportion of the plots of the entire data set, the overall sensitivity mainly depends on the classifier's ability to recognize successfully this particular class. In our case, there were two large classes in the expert classification (1 and 5), which together included 68% of the plots in the entire data set. These classes corresponded to the alliances *Arrhenatherion* and *Calthion*, respectively, which are generally accepted as being well-defined alliances of the traditional phytosociological classification of Central European grasslands (Ellenberg 1996; Havlová et al. 2004; Botta-Dukát et al. 2005). The classifiers for expert classification were able to recognize these large classes successfully, both in terms of sensitivity and positive predictive power, and this resulted in higher overall values of these measures for the expert classification. On the other hand, there was a rather heterogeneous but large class 7 in the numerical classification, which was poorly recognized by the classifiers. Therefore the overall sensitivity and positive predictive power of the numerical classification were decreased. We presume that without the effects of these large classes, the overall performance of the classifiers would not considerably differ between expert and numerical classification. From this point of view, there seems to be no fundamental difference in the ability of the MLP classifier to reproduce the presumably less consistent expert classification and the more consistent numerical classification.

In our study, we only considered the unequivocal assignment of each plot to a single class by the classifier. However, neural networks can produce a fuzzy assignment, giving the probabilities of class membership for each plot. Instead of assigning each plot to the most probable class, as we did, it would also be possible to evaluate the classifier's performance by taking membership probabilities for more than one class for each plot. In that case, the sensitivity of the classifiers would most probably be higher than our conservative estimate of 70-83%, because in the cases of misclassification, the membership probabilities for the correct class were usually the second highest. Given this fact and considering that many misclassifications might have resulted from inconsistencies in the *a priori* classifications, we conclude that MLP may be a successful technique for the supervised classification of species-by-sites matrices.

In both cases, expert and numerical classification, worse results were achieved when the classifiers were trained with plots 'typical' of the particular classes (containing several diagnostic species) rather than with randomly selected plots (Table 1), especially with the classifiers for expert classification where this difference was significant. Although the sensitivity on training- and selection data sets was slightly higher with the use of plots rich in diagnostic species, indicating that classes within this subset of plots were easier to discriminate, it was lower on the test data set, which contained plots both poor and rich in diagnostic species. Thus, our original hypothesis that training with typical plots would result in a higher proportion of correct assignments was not supported. Due to removal of the less typical plots from the training data set the classifier did not learn to recognize such plots and tended to misclassify them in the test data set. A similar result was presented by Foody (1999b) in the context of the remote sensing data classification. In that case, the neural network trained with a set of border patterns performed better than one trained with a set of patterns drawn from the cores of the classes.

This result points to the ability of the MLP classifier of recognizing class membership also for plots that are poor in diagnostic species. This is an advantage against the expert classification of traditional phytosociology, which is predominantly based on diagnostic species, and also against the formalized supervised methods derived from similar principles. Such methods include, for example: (1) the indicator ordination option in TWINSPAN, which provides indicator species and thresholds for plot assignment to the classes (Hill 1979); (2) calculations of similarity coefficients between the species composition of individual plots and species frequencies within classes of *a priori* classifications (Hill 1989; Kočí et al. 2003); (3) the COCKTAIL method, which assigns plots to classes on the basis of the occurrence of species from pre-defined sociological species groups (Bruelheide 2000; Bruelheide & Chytrý 2000; Kočí et al. 2003). However, the black-box approach of the MLP classifiers does not provide sufficient information on the underlying assignment process, which could be used for developing simple keys for the identification of vegetation types in field mapping.

In our application of MLP we directly used the species-by-sites matrix as input data. Ejrnæs et al. (2002) applied an artificial neural network in a similar context (predicting the degree of habitat naturalness from species composition) but used a different, two-step approach. First they subjected the species-by-sites matrix to ordination and then used site scores on the first two or three ordination axes as input data for the neural network. With this approach, supervised classification can be done through the passive ordination of new sites and their subsequent assignment to the appropriate class by the neural network. Our approach is more straightforward, using a single step, but with large data sets it can be computationally more demanding. Future studies should compare both approaches, especially with respect to the possible effect of noise contained in the species-by-sites matrices on the one hand and the potential loss of information due to using only the higher ordination axes on the other hand. The latter issue can possibly have negative influence on the classifier's performance especially when it is assumed to recognize higher numbers of classes.

Acknowledgements. Our thanks are due to Lubomír Tichý and Václav Veleba for programming support and Rasmus Ejrnæs, Marie Forbelská, Jiří Jarkovský, Jan Roleček and Onno van Tongeren for helpful comments. This study was supported through projects GA ČR 206/02/0957 and MSM 0021622416.

References

Botta-Dukát, Z. Chytrý, M., Hájková, P. & Havlová, M. 2005. Vegetation of lowland wet meadows along a climatic continentality gradient in Central Europe. *Preslia* 77: 89-111.

Brosse, S., Giraudel, J.L. & Lek, S. 2001. Utilisation of non-supervised neural networks and principal component analysis to study fish assemblages. *Ecol. Model*. 146: 159-166.

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 national vegetation classifications: A comparison of two methods for analysis of large data sets. *J. Veg. Sci*. 11: 295-306.

Cairns, D.M. 2001. A comparison of methods for predicting vegetation type. *Plant Ecol*. 156: 3-18.

Chon, T.S., Park, Y.S., Moon, K.H. & Cha, E.Y. 1996. Patternizing communities by using an artificial neural network. *Ecol. Model*. 90: 69-78.

Chytrý, M. & Rafajová, M. 2003. Czech National Phytosociological 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.

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

Ejrnæs, R., Bruun, H.H., Aude, E. & Buchwald, E. 2004. Developing a classifier for the Habitats Directive grassland types in Denmark using species lists for prediction. *Appl. Veg. Sci*. 7: 71-80.

Ellenberg, H. 1996. *Vegetation Mitteleuropas mit den Alpen*. 5th ed. Ulmer, Stuttgart, DE.

Fielding, A.H. & Bell, J.F. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environ. Conserv*. 24: 38-49.

Foody, G.M. 1996. Fuzzy modelling of vegetation from remotely sensed imagery. *Ecol. Model*. 85: 3-12.

Foody, G.M. 1999a. Applications of the self-organising feature map network in community data analysis. *Ecol. Model*. 120: 97-109.

Foody, G.M. 1999b. The significance of border training patterns in classification by a feedforward neural network using back propagation learning. *Int. J. Remote Sens*. 20: 3549-3562.

Giraudel, J.L. & Lek, S. 2001. A comparison of self-organizing map algorithm and some conventional statistical methods for ecological community ordination. *Ecol. Model*. 146: 329-339.

Guégan, J.F., Lek, S. & Oberdorff, T. 1998. Energy availability and habitat heterogeneity predict global riverine fish diversity. *Nature* 391: 382-384.

Havlová, M., Chytrý, M. & Tichý, L. 2004. Diversity of hay meadows in the Czech Republic: major types and environmental gradients. *Phytocoenologia* 34: 551-567.

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*. 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.

Hirlbert, D.W. & Ostendorf, B. 2001. The utility of artificial neural networks for modelling the distribution of vegetation in past, present and future climates. *Ecol. Model.* 146: 311-327.

Kočí, M., Chytrý, M. & Tichý, L. 2003. Formalized reproduction of an expert-based phytosociological classification: A case study of subalpine tall-forb vegetation. *J. Veg. Sci.* 14: 601-610.

Kohonen, T. 198. Self-organised formation of topologically correct feature maps. *Biol. Cybern.* 43: 59-69.

Lae, R., Lek, S. & Moreau, J. 1999. Predicting fish yield of African lakes using neural networks. *Ecol. Model.* 120: 325-335.

Legendre, P. & Legendre, L. 1998. *Numerical ecology*. 2nd ed. Elsevier, Amsterdam, NL.

Lek, S. & Guégan, J.F. 1999. Artificial neural networks as a tool in ecological modelling, an introduction. *Ecol. Model.* 120: 65-73.

Liu, C.M., Zhang, L.J., Davis, C.J., Solomon, D.S., Brann, T.B. & Caldwell, L.E. 2003. Comparison of neural networks and statistical methods in classification of ecological habitats using FIA data. *For. Sci.* 49: 619-631.

McCune, B. & Mefford, M.J. 1999. *PC-ORD. Multivariate analysis of ecological data. Version 4.* MjM Software Design, Gleneden Beach, OR, US.

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. Red list of plant communities of the Czech Republic and their endangerment. 2nd ed. *Severočes. Přír. Suppl.* 1995: 1-206.

Paola, J.D. & Schowengerdt, R.A. 1995. A review and analysis of backpropagation neural networks for classification of remotely-sensed multispectral imagery. *Int. J. Remote Sens.* 16: 3033-3058.

Ripley, B.D. 1996. *Pattern recognition and neural networks*. Cambridge University Press, Cambridge, UK.

Rumelhart, D.E., Hinton, G.E. & Wiliams, R.J. 1986. Learning representations by back-propagating errors. *Nature* 323: 533-536.

Tan, S.S. & Smeins, F.E. 1996. Predicting grassland community changes with an artificial neural network model. *Ecol. Model.* 84: 91-97.

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

van der Maarel, E. 1979. Transformation of cover-abundance values in phytosociology and its effects on community similarity. *Vegetatio* 39: 97-114.

Zhang, L.J., Liu, C.M., Davis, C.J., Solomon, D.S., Brann, T.B. & Caldwell, L.E. 2004. Fuzzy classification of ecological habitats from FIA data. *For. Sci.* 50: 117-127.

Zhang, X.H., Li, C. & Yuan, Y.L. 1997. Application of neural networks to identifying vegetation types from satellite images. *AI Appl.* 11: 99-106.

Received 5 July 2004;
Accepted 14 June 2005.
Co-ordinating Editor: H. Bruelheide.

For Apps. 1-4, see JVS/AVS Electronic Archives; www.opuluspress.se

App. 1. Synoptic table of the percentage species frequencies for classes of expert classification. Diagnostic species of particular classes (grey shading) are those with a fidelity of $\Phi \geq 0.25$, using this coefficient adjusted for the equal size of classes. They are ranked by decreasing Φ value. Species names follow Kubát et al. (2002, *Klíč ke květeně České republiky*, Academia, Praha) and Frey et al. (1995, *Die Moos- und Farnpflanzen Europas,* G. Fischer, Stuttgart).

Class no.	1	2	3	4	5	6	7	8	9	10	11
No. of plots	825	120	148	268	1965	52	420	54	216	24	94
Diagnostic species for individual classes											
Arrhenatherum elatius	77	12	24	12	5	13	12	.	13	.	.
Dactylis glomerata	80	61	58	32	14	13	21	.	15	.	4
Knautia arvensis agg.	48	17	24	5	1	.	10	.	34	4	3
Trisetum flavescens	67	58	34	20	10	2	19	.	9	.	6
Galium mollugo agg.	67	28	22	29	12	27	14	.	17	8	4
Crepis biennis	27	1	11	6	1	2	2	.	.	.	.
Campanula patula	56	43	28	18	5	8	20	6	19	.	16
Plantago lanceolata	81	50	76	28	13	58	55	11	61	8	21
Leontodon hispidus	54	53	42	6	3	17	17	20	40	17	31
Plantago media	29	1	32	3	1	.	8	.	6	.	.
Taraxacum sect. *Ruderalia*	71	43	68	56	14	67	28	.	9	4	1
Tragopogon orientalis	14	.	1	.	.	.	1	.	1	.	.
Leucanthemum vulgare agg.	67	45	48	27	11	25	44	9	40	12	17
Daucus carota	25	2	21	3	1	.	5	.	4	.	.
Securigera varia	16	.	13	.	.	.	.	.	3	.	.
Achillea millefolium agg.	87	86	89	40	25	37	58	17	68	12	59
Salvia pratensis	13	.	8	.	.	2	1	.	.	.	.
Trifolium dubium	29	1	17	9	3	10	10	.	6	.	.
Ranunculus bulbosus	14	.	16	.	.	2	.	.	2	.	.
Pastinaca sativa	16	.	2	4	1	.	1	.	.	.	.
Veronica chamaedrys	75	76	49	36	27	17	38	13	46	4	37
Heracleum sphondylium	44	41	9	26	13	.	5	.	6	.	2
Geranium pratense	33	2	9	31	6	6	7	.	.	.	.
Crepis mollis	4	42	.	3	9	.	9	6	7	12	24
Alchemilla vulgaris agg.	46	88	61	35	37	.	45	17	38	4	48
Lolium perenne	6	.	42	1	.	.	.	.	.	.	.
Ononis spinosa	1	.	20	.	.	.	2	.	1	.	.
Poa annua	.	.	16	.	.	.	.	.	.	.	1
Cynosurus cristatus	18	19	55	12	8	.	14	2	7	.	1
Leontodon autumnalis	9	18	38	11	1	10	7	2	7	.	7
Euphorbia cyparissias	8	.	26	.	.	.	.	.	10	.	.
Phleum pratense	11	24	40	9	5	.	6	.	8	.	2
Senecio jacobaea	3	1	17	.	.	.	.	.	3	.	.
Brachypodium pinnatum	4	.	20	.	1	.	2	.	1	.	.
Trifolium repens	53	51	68	27	13	8	26	13	20	4	18
Carlina vulgaris	2	.	12	.	.	.	.	.	2	.	.
Fragaria vesca	7	2	20	.	.	.	1	2	7	.	1
Euphrasia rostkoviana	6	13	23	2	.	.	5	9	8	4	5
Bellis perennis	13	12	30	9	3	.	6	.	.	.	.
Agrimonia eupatoria	5	.	15	.	.	.	1	.	.	.	.
Scirpus sylvaticus	1	.	1	16	60	2	12	.	2	.	.
Caltha palustris	1	.	1	12	54	17	10	.	.	4	.
Juncus effusus	1	.	1	16	45	2	19	2	12	17	.
Crepis paludosa	.	7	.	1	28	.	7	2	5	.	2
Angelica sylvestris	4	9	.	21	51	2	28	.	10	8	3
Filipendula ulmaria	5	2	2	29	52	25	32	.	1	.	.
Equisetum palustre	3	.	.	15	37	.	16	.	2	4	.
Myosotis palustris agg.	4	25	6	31	55	13	33	.	7	8	4
Galium palustre agg.	1	.	1	15	32	37	10	.	.	12	.
Galium uliginosum	2	5	3	19	51	2	44	4	17	4	.
Chaerophyllum hirsutum	1	11	.	2	17	.	.	2	.	.	2
Equisetum fluviatile	.	.	.	1	13	.	2	.	.	.	.
Cirsium rivulare	1	2	.	1	17	.	5	.	.	.	.
Cardamine amara	.	.	.	.	7	.	.	.	.	.	.
Cnidium dubium	.	.	.	.	.	58	3	.	.	.	.
Potentilla reptans	10	.	11	3	2	69	4	.	1	.	.
Carex praecox	1	.	.	1	.	44	4	11	2	12	.

App. 1, cont.

Class no.	1	2	3	4	5	6	7	8	9	10	11
No. of plots	825	120	148	268	1965	52	420	54	216	24	94
Phalaris arundinacea	1	.	.	22	5	60	2	.	.	.	.
Gratiola officinalis	.	.	.	.	.	31	.	.	.	.	.
Rumex crispus	3	1	16	13	4	50	2	.	.	.	.
Pulegium vulgare	.	.	.	.	.	23	.	.	.	.	.
Inula britannica	.	.	1	.	.	23	.	.	.	.	.
Odontites vernus	1	.	1	1	.	25	.	.	.	.	.
Cirsium arvense	12	.	17	16	5	56	8	.	1	.	.
Poa palustris	2	.	2	14	8	48	2	.	.	.	.
Cerastium dubium	.	.	.	.	.	21	.	.	.	.	.
Allium angulosum	.	.	.	.	.	21	1	.	.	.	.
Symphytum officinale	8	.	1	22	7	50	5	.	.	.	.
Pseudolysimachion maritimum	.	.	.	1	.	21	2	.	.	.	.
Aster novi-belgii agg.	.	.	.	.	.	17	.	.	.	.	.
Scutellaria hastifolia	.	.	.	.	.	17	.	.	.	.	.
Lythrum virgatum	.	.	.	.	.	17	.	.	.	.	.
Rorippa sylvestris	.	.	.	1	.	17	.	.	.	.	.
Lathyrus palustris	.	.	.	.	.	17	.	.	.	.	.
Carex acuta	1	.	.	16	13	48	5	.	.	.	.
Carex vulpina	.	.	1	14	2	29	4	.	.	.	.
Carex melanostachya	.	.	.	.	.	15	.	.	.	.	.
Potentilla anserina	2	.	8	3	2	27	2	.	.	.	.
Elytrigia repens	12	2	9	16	3	38	3	.	1	.	.
Eleocharis palustris	.	.	.	1	2	19	.	.	.	.	.
Iris pseudacorus	.	.	.	1	1	15	.	.	.	.	.
Silaum silaus	1	.	.	1	.	17	4	.	.	.	.
Rorippa austriaca	.	.	.	.	.	10	.	.	.	.	.
Lysimachia nummularia	20	10	7	38	26	62	17	.	1	.	.
Viola pumila	.	.	.	.	.	10	2	.	.	.	.
Ranunculus repens	22	12	37	67	40	75	23	.	5	.	.
Thalictrum flavum	.	.	.	1	.	10	1	.	.	.	.
Lotus tenuis	.	.	.	.	.	8	.	.	.	.	.
Myosotis ramosissima	4	.	.	.	.	13	2	.	.	.	.
Vicia tetrasperma	9	.	5	1	.	19	1	.	4	.	.
Molinia caerulea agg.	1	.	1	4	10	.	59	7	14	33	.
Succisa pratensis	2	.	1	8	8	.	58	.	26	12	.
Galium boreale	5	1	.	8	3	25	39	.	5	.	.
Betonica officinalis	6	.	9	3	1	6	23	.	3	.	.
Sanguisorba officinalis	31	7	13	69	42	37	82	.	25	33	.
Homogyne alpina	.	1	.	.	.	.	.	44	1	4	5
Hieracium alpinum	.	.	.	.	.	.	.	22	.	.	1
Hypochaeris uniflora	.	.	.	.	.	.	.	20	.	.	1
Festuca supina	.	.	.	.	.	.	.	19	.	.	2
Veratrum album ssp. *album*	.	2	.	.	1	.	.	19	.	.	3
Ranunculus platanifolius	.	4	.	.	.	.	.	17	.	.	7
Rhinanthus pulcher	1	1	1	.	.	.	.	17	.	.	3
Trientalis europaea	.	.	.	.	.	.	.	15	.	8	.
Ligusticum mutellina	.	.	.	.	.	.	.	11	.	.	.
Hieracium prenanthoides	.	.	.	.	.	.	.	11	.	.	1
Melampyrum pratense	.	.	.	.	.	.	.	13	3	4	.
Maianthemum bifolium	.	2	.	.	.	.	.	11	.	.	4
Melampyrum sylvaticum	.	.	.	.	.	.	.	11	.	.	5
Pseudorchis albida	.	1	.	.	.	.	.	9	.	.	1
Luzula sylvatica	.	.	.	.	.	.	.	9	.	.	.
Thesium alpinum	.	.	.	.	.	.	.	9	.	.	2
Polygonatum verticillatum	.	2	.	.	.	.	.	9	.	.	2
Polygala vulgaris	8	8	18	.	.	.	8	11	44	21	7
Festuca ovina	5	.	3	.	1	.	19	.	43	17	.
Viola canina	6	3	5	1	1	.	13	9	40	.	6
Dianthus deltoides	8	5	5	.	.	.	2	6	28	.	.
Galium pumilum agg.	8	10	5	1	1	.	3	2	31	21	5
Hieracium pilosella	10	9	18	.	.	.	2	11	33	4	24
Rumex acetosella	5	3	7	1	.	.	.	4	21	.	4
Briza media	31	26	41	10	20	.	50	17	64	38	16

App. 1, cont.

Class no.	1	2	3	4	5	6	7	8	9	10	11
No. of plots	825	120	148	268	1965	52	420	54	216	24	94
Juncus squarrosus	.	.	.	.	.	.	.	2	1	92	.
Vaccinium uliginosum	.	.	.	.	.	.	.	9	4	33	.
Hieracium laevigatum	1	6	.	.	.	.	.	9	4	.	23
Hieracium iseranum	.	5	.	.	.	.	.	.	.	.	13
Gnaphalium sylvaticum	1	8	1	.	.	.	.	7	6	.	17
Luzula campestris agg.	44	48	30	16	22	2	58	57	70	62	74
Pimpinella saxifraga	39	14	39	2	.	2	15	2	43	4	1
Lotus corniculatus	49	13	54	12	3	17	25	7	40	17	3
Trifolium pratense	63	36	71	28	13	40	33	7	23	.	10
Cardaminopsis halleri	2	62	1	1	2	.	.	6	3	.	40
Geranium sylvaticum	2	66	1	.	5	.	.	7	2	.	43
Phyteuma spicatum	3	56	1	.	1	.	.	13	4	.	44
Silene dioica	.	43	.	1	1	.	.	17	1	.	38
Hypericum maculatum	16	81	15	8	12	.	17	28	37	12	72
Bistorta major	5	68	1	12	33	.	15	59	10	25	78
Plantago major agg.	5	.	28	9	1	27	3	.	.	.	1
Thymus pulegioides	17	3	31	.	.	.	5	7	39	.	6
Serratula tinctoria	1	.	.	1	1	42	15	.	.	.	.
Solidago virgaurea	.	8	1	.	1	.	.	54	3	.	28
Luzula luzuloides	1	15	2	1	1	.	.	39	5	.	37
Viola lutea ssp. *sudetica*	.	6	.	.	.	.	.	26	.	.	15
Vaccinium vitis-idaea	.	.	.	.	.	.	2	28	8	21	3
Gentiana asclepiadea	.	2	.	.	.	.	.	24	.	.	15
Calamagrostis villosa	.	2	.	.	.	.	.	22	3	17	4
Crepis conyzifolia	.	8	.	.	.	.	.	17	.	.	17
Arnica montana	.	2	.	.	.	.	1	19	11	17	4
Galium saxatile	.	10	.	.	.	.	.	19	9	12	34
Danthonia decumbens	2	.	7	1	1	.	10	13	39	29	4
Veronica officinalis	5	22	7	.	1	.	4	20	31	12	43
Pedicularis sylvatica	.	.	1	.	.	.	1	2	13	21	.
Potentilla aurea	.	27	.	.	.	.	.	41	1	.	55
Campanula rotundifolia agg.	21	49	7	1	1	.	7	6	48	8	40
Agrostis capillaris	35	87	64	16	21	.	36	56	75	54	81
Poa chaixii	.	21	.	.	1	.	.	22	2	.	27
Phleum rhaeticum	.	17	.	.	.	.	.	19	.	.	26
Silene vulgaris	7	25	1	.	.	.	.	35	3	.	59
Campanula bohemica	.	14	.	.	.	.	.	24	.	.	36
Vaccinium myrtillus	.	4	1	.	.	.	2	72	18	33	47
Hieracium lachenalii	1	11	.	.	.	.	.	31	6	25	35
Calluna vulgaris	.	.	.	.	.	.	8	28	35	67	4
Carex pilulifera	2	12	1	1	1	.	6	31	37	12	53
Potentilla erecta	9	27	22	9	23	.	61	81	74	75	59
Avenella flexuosa	1	13	.	.	1	.	.	78	25	58	82
Nardus stricta	5	24	7	3	8	.	36	100	78	100	85
Other species with frequency > 5%											
Rumex acetosa	71	88	43	75	64	37	68	15	49	12	59
Ranunculus acris	66	85	53	69	60	44	79	17	50	38	52
Festuca rubra agg.	72	88	68	45	50	6	71	59	71	38	95
Alopecurus pratensis	51	72	27	85	60	85	48	2	8	.	13
Anthoxanthum odoratum	60	77	64	39	36	10	64	74	64	50	82
Holcus lanatus	51	4	36	65	46	6	70	2	47	21	.
Poa pratensis agg.	74	43	51	68	36	69	51	4	20	.	10
Lathyrus pratensis	47	26	17	53	53	50	55	.	9	17	3
Deschampsia cespitosa	18	49	23	70	50	21	69	28	39	33	46
Poa trivialis	28	32	30	62	61	48	24	.	1	.	.
Lychnis flos-cuculi	24	17	14	63	50	62	60	.	23	17	4
Festuca pratensis	57	15	62	55	31	56	49	.	8	.	.
Cerastium holosteoides	64	41	59	51	22	29	40	6	22	.	6
Ranunculus auricomus agg.	12	23	7	43	48	50	52	2	12	4	2
Cardamine pratensis	17	18	10	47	40	56	43	.	14	8	10
Cirsium palustre	4	2	3	17	46	.	50	.	23	33	.
Carex nigra	2	.	2	23	41	.	39	13	18	62	14

App. 1-4. Internet supplement to: Černá, L. & Chytrý, M. 2005.
Supervised classification of plant communities with artificial neural networks.
J. Veg. Sci. 16: 407-414.

OPULUS PRESS

App. 1, cont.

Class no.	1	2	3	4	5	6	7	8	9	10	11
No. of plots	825	120	148	268	1965	52	420	54	216	24	94
Vicia cracca	43	38	18	25	21	15	29	2	19	8	4
Carex panicea	1	.	4	13	36	.	56	2	26	50	.
Prunella vulgaris	28	25	40	17	16	27	35	4	21	.	3
Centaurea jacea	36	2	34	20	8	38	49	7	31	12	.
Stellaria graminea	28	30	27	18	15	12	24	6	31	4	5
Avenula pubescens	35	7	10	8	13	2	32	.	28	.	.
Cirsium oleraceum	6	2	1	19	28	.	9	.	.	.	.
Lysimachia vulgaris	1	.	1	10	27	10	21	.	5	.	.
Equisetum arvense	18	.	4	14	19	17	11	.	6	.	.
Carex hirta	8	.	6	25	18	15	18	.	5	.	.
Ajuga reptans	12	22	7	9	17	2	17	4	7	.	11
Agrostis stolonifera	5	2	9	15	22	17	12	.	4	4	.
Carex pallescens	4	6	11	12	15	.	30	7	25	8	5
Juncus conglomeratus	1	.	1	9	22	.	21	.	10	8	.
Lotus uliginosus	2	.	3	12	20	.	20	.	2	.	.
Valeriana dioica	1	1	.	3	20	.	22	2	6	4	.
Carex ovalis	3	9	8	24	13	.	24	2	15	17	11
Galium verum agg.	28	1	20	12	5	10	20	.	18	8	.
Rhinanthus minor	20	21	15	7	5	2	18	4	19	4	1
Glechoma hederacea	19	2	6	25	7	31	7	.	.	.	.
Agrostis canina	.	.	.	8	16	.	14	.	6	29	.
Colchicum autumnale	12	.	2	12	9	23	19	.	.	.	.
Pimpinella major	21	16	5	12	5	4	12	4	3	.	1
Anemone nemorosa	4	12	1	5	11	.	16	7	15	.	15
Cirsium canum	8	.	.	17	11	25	14	.	.	.	.
Juncus filiformis	.	2	.	6	16	.	6	6	1	29	14
Urtica dioica	5	2	7	13	14	13	1	.	1	.	.
Viola palustris	.	2	1	1	16	.	5	.	6	12	.
Trifolium hybridum	6	2	15	25	7	31	13	2	1	.	.
Vicia sepium	19	22	7	11	5	4	2	.	2	.	1
Holcus mollis	6	24	4	3	9	.	4	6	12	.	26
Aegopodium podagraria	13	20	2	11	8	.	2	.	2	.	.
Anthriscus sylvestris	20	12	8	15	5	.	.	.	2	.	.
Dactylorhiza majalis	1	1	.	2	13	.	13	.	6	4	.
Carex brizoides	1	1	1	6	13	.	5	.	2	4	.
Mentha arvensis	2	.	4	7	11	4	8	.	1	.	.
Lythrum salicaria	.	.	.	9	11	17	8	.	.	.	.
Achillea ptarmica	1	.	1	4	9	.	13	.	4	8	.
Geranium palustre	1	.	.	5	11	.	3	.	.	.	.
Epilobium palustre	.	.	.	2	12	2	4	.	.	.	.
Trollius altissimus	1	2	.	4	8	.	15	4	3	8	.
Geum rivale	.	3	.	1	11	.	4	.	.	4	.
Saxifraga granulata	16	3	4	7	2	.	10	.	2	.	.
Hypericum perforatum	18	.	18	1	1	.	2	.	19	.	.
Juncus articulatus	.	.	1	5	10	.	5	.	3	4	.
Carex echinata	.	.	1	1	9	.	8	.	3	21	.
Selinum carvifolia	2	.	1	9	4	4	21	.	4	.	.
Primula elatior	2	21	.	.	8	.	3	.	1	.	2
Eriophorum angustifolium	.	.	.	.	10	.	5	.	1	8	.
Rumex obtusifolius	3	8	8	18	6	8	.	.	.	.	1

App. 2. Synoptic table of percentage species frequencies for classes of numerical classification. Diagnostic species of particular classes (grey shading) are those with fidelity $\Phi \geq 0.25$, using this coefficient adjusted for the equal size of classes. They are ranked by decreasing Φ value. Nomenclature as in App. 1.

Class no.	1	2	3	4	5	6	7	8	9	10	11
No. of plots	627	540	263	505	360	186	676	89	188	313	439
Diagnostic species for individual classes											
Arrhenatherum elatius	80	39	2	9	1	6	11	4	7	3	6
Dactylis glomerata	85	61	2	23	6	15	17	25	28	7	15
Geranium pratense	47	5	.	9	1	3	13	.	7	3	7
Crepis biennis	32	9	.	3	.	2	2	.	2	1	1
Galium mollugo agg.	70	44	5	14	3	10	16	10	28	9	12
Taraxacum sect. *Ruderalia*	78	50	3	48	12	7	27	2	20	8	7
Trisetum flavescens	65	54	3	29	8	6	8	24	15	1	4
Plantago media	33	18	.	3	1	9	1	.	1	.	2
Pastinaca sativa	20	1	.	1	.	1	1	.	3	.	3
Securigera varia	19	7	.	.	.	1	.	.	.	.	.
Daucus carota	28	13	.	3	.	7	1	.	5	1	3
Tragopogon orientalis	16	5	.	.	.	1	.	.	1	.	.
Convolvulus arvensis	15	3	.	.	.	.	2	.	.	.	.
Lolium perenne	14	4	.	1	.	.	.	.	.	.	1
Heracleum sphondylium	47	29	2	15	9	3	13	28	10	5	14
Salvia pratensis	15	6	.	.	.	3	.	.	.	.	.
Festuca rupicola	15	4	.	.	.	2	2	.	.	.	.
Medicago lupulina	15	5	.	1	.	4	1	.	2	.	.
Pimpinella saxifraga	27	53	8	14	.	5	2	.	1	.	.
Campanula rotundifolia agg.	11	44	27	6	1	3	1	9	.	.	.
Leontodon hispidus	44	64	20	20	3	11	2	4	7	1	1
Thymus pulegioides	13	34	7	3	.	3	.	.	.	.	.
Carlina acaulis	3	20	8	.	.	1	.	.	.	.	.
Hieracium pilosella	5	26	17	2	.	2	.	.	.	.	.
Lotus corniculatus	45	51	10	23	3	13	4	1	5	2	4
Polygala vulgaris	3	24	22	6	1	5	.	.	1	.	.
Leucanthemum vulgare agg.	58	67	17	55	12	19	10	9	26	3	3
Dianthus deltoides	4	17	5	2	.	2	.	.	.	.	.
Trifolium repens	47	61	10	47	21	8	12	2	13	4	2
Viola canina	1	22	11	12	1	6	1	1	1	.	.
Galium pumilum agg.	4	19	9	3	3	2	.	6	.	.	.
Avenella flexuosa	.	6	60	1	3	2	2	.	.	.	.
Nardus stricta	1	24	91	30	28	27	4	10	2	1	.
Vaccinium myrtillus	.	3	44	1	1	4	1	.	.	.	.
Potentilla aurea	.	5	30	.	1	.	.	1	.	.	.
Carex pilulifera	.	10	42	8	4	5	.	2	.	.	.
Galium saxatile	.	2	25	.	.	.	1	.	.	.	.
Campanula bohemica	.	2	21	.	.	.	.	.	.	.	.
Luzula luzuloides	.	5	25	.	1	.	.	6	.	.	.
Solidago virgaurea	1	2	23	.	1	.	.	.	2	1	.
Silene vulgaris	7	8	31	.	1	.	.	4	.	.	.
Hieracium lachenalii	.	6	21	.	1	1	.	.	.	.	.
Calluna vulgaris	.	5	25	2	1	15	1	.	.	.	.
Veronica officinalis	2	18	32	6	1	1	.	.	2	.	.
Poa chaixii	.	3	18	.	1	.	1	16	.	.	.
Phleum rhaeticum	.	2	14	.	1	.	.	.	.	.	.
Gentiana asclepiadea	.	.	11	.	1	.	.	1	.	.	.
Homogyne alpina	.	.	11	.	1	.	.	.	.	.	.
Crepis conyzifolia	.	1	11	.	.	.	.	.	.	.	.
Hieracium laevigatum	.	2	13	.	1	1	.	1	.	.	.
Viola lutea ssp. *sudetica*	.	1	11	.	1	.	.	.	.	.	.
Vaccinium vitis-idaea	.	1	13	1	.	4	.	.	.	.	.
Silene dioica	1	8	19	2	1	.	1	8	.	.	1
Phyteuma spicatum	1	16	22	1	2	.	.	15	1	.	.
Arnica montana	.	1	11	1	1	2	.	.	.	.	.
Holcus lanatus	45	43	18	92	69	46	43	10	66	27	20
Sanguisorba officinalis	32	19	11	85	44	72	41	38	37	35	40
Carex pallescens	2	9	16	40	25	11	7	37	31	6	1
Lychnis flos-cuculi	24	22	12	80	78	33	51	46	63	36	18

App. 2, cont.

Class no.	1	2	3	4	5	6	7	8	9	10	11
No. of plots	627	540	263	505	360	186	676	89	188	313	439
Juncus filiformis	.	.	7	7	52	5	8	22	1	13	2
Carex nigra	1	1	21	41	90	32	23	25	55	38	17
Agrostis canina	.	.	5	10	47	18	11	8	21	9	2
Cirsium palustre	1	7	16	42	82	50	32	35	11	48	34
Valeriana dioica	.	1	2	12	50	32	15	17	17	10	6
Viola palustris	.	.	5	3	37	8	9	22	1	21	8
Carex echinata	.	.	3	4	26	9	4	2	18	8	.
Myosotis palustris agg.	2	10	4	47	75	15	39	72	69	67	30
Carex canescens	.	.	1	2	19	1	3	.	5	8	.
Juncus conglomeratus	.	1	9	19	43	27	12	13	40	11	10
Epilobium palustre	.	.	.	2	25	7	8	4	6	14	7
Cardamine pratensis	14	18	8	65	68	18	44	47	20	27	15
Ranunculus auricomus agg.	12	14	6	63	71	31	43	67	23	30	39
Molinia caerulea agg.	1	4	10	18	16	96	7	1	7	2	13
Carex davalliana	.	.	.	1	2	15	.	.	1	.	1
Cirsium heterophyllum	.	3	7	1	6	.	1	99	.	1	4
Bistorta major	6	16	50	27	54	9	21	79	5	20	29
Crepis mollis	1	14	11	14	19	8	3	44	4	.	4
Geranium sylvaticum	.	17	19	2	5	.	3	37	.	2	4
Holcus mollis	2	13	19	3	14	5	8	39	2	8	5
Phyteuma nigrum	1	3	1	5	5	2	.	17	.	.	.
Cirsium rivulare	1	.	.	5	3	5	6	.	94	19	7
Cruciata glabra	1	7	.	1	1	3	2	.	46	3	2
Eriophorum latifolium	.	.	.	.	2	4	.	.	27	.	.
Carex flava	.	.	2	1	4	9	1	.	29	2	.
Dactylorhiza majalis	.	1	3	9	20	15	9	.	42	4	4
Epipactis palustris	.	.	.	.	.	1	.	.	14	1	.
Valeriana simplicifolia	.	.	.	.	1	.	1	.	17	2	1
Juncus articulatus	.	.	3	5	16	9	5	.	32	7	3
Juncus inflexus	.	.	.	1	.	2	3	.	22	9	4
Hypericum tetrapterum	.	1	.	1	1	1	2	.	17	3	.
Carex flacca	.	1	1	3	2	10	1	.	20	2	.
Mentha longifolia	.	.	.	.	.	.	4	.	24	14	10
Eupatorium cannabinum	.	.	.	.	.	1	1	.	13	4	3
Primula elatior	.	7	1	4	4	4	7	10	26	3	7
Lysimachia nemorum	.	.	.	.	1	.	2	1	13	2	1
Mentha arvensis	3	1	1	10	6	5	9	1	30	19	3
Scirpus sylvaticus	1	.	3	21	48	20	37	22	61	99	59
Galium palustre agg.	.	1	.	8	44	12	27	8	35	56	20
Juncus effusus	.	1	5	26	58	21	31	33	46	63	31
Epilobium obscurum	.	.	.	.	8	.	5	.	.	18	5
Filipendula ulmaria	7	2	1	33	39	33	39	29	35	33	90
Geranium palustre	1	.	.	2	2	4	9	.	6	3	32
Knautia arvensis agg.	46	43	7	8	1	4	2	3	1	.	1
Plantago lanceolata	75	83	26	64	19	26	15	10	22	4	3
Achillea millefolium agg.	86	90	48	63	34	38	24	69	30	12	10
Campanula patula	50	51	8	25	3	5	8	15	10	3	1
Agrostis capillaris	22	75	68	39	46	22	16	88	9	8	5
Potentilla erecta	4	31	74	39	47	70	12	44	49	9	7
Hypericum maculatum	7	42	51	15	16	13	6	66	13	7	8
Succisa pratensis	2	5	14	36	22	49	5	3	14	2	3
Galium uliginosum	1	4	12	42	79	48	35	74	13	42	43
Eriophorum angustifolium	.	.	1	1	29	11	4	1	27	3	1
Caltha palustris	1	.	2	15	64	11	39	15	58	64	53
Carex panicea	1	4	15	44	70	64	23	20	66	20	9
Other species with frequency > 5%											
Rumex acetosa	65	72	46	91	88	39	59	85	74	55	30
Ranunculus acris	60	73	41	93	89	61	55	78	72	40	28
Festuca rubra agg.	59	90	76	82	89	46	40	85	72	24	15
Alopecurus pratensis	52	43	11	76	59	19	65	85	29	58	61
Anthoxanthum odoratum	48	79	70	78	77	35	26	30	57	18	3

App. 1-4. Internet supplement to: Černá, L. & Chytrý, M. 2005.
Supervised classification of plant communities with artificial neural networks.
J. Veg. Sci. 16: 407-414.

OPULUS PRESS

App. 2, cont.

Class no.	1	2	3	4	5	6	7	8	9	10	11
No. of plots	627	540	263	505	360	186	676	89	188	313	439
Poa pratensis agg.	79	53	8	69	51	28	47	29	30	19	20
Lathyrus pratensis	47	31	3	67	46	36	53	47	63	36	54
Deschampsia cespitosa	18	24	45	81	76	51	46	64	43	35	36
Poa trivialis	32	20	.	51	64	12	60	48	51	67	54
Alchemilla vulgaris agg.	35	70	33	67	47	22	26	58	65	17	22
Veronica chamaedrys	70	76	29	48	41	19	24	71	30	13	10
Festuca pratensis	64	35	2	69	38	19	36	7	55	15	16
Cerastium holosteoides	64	57	6	65	26	9	25	3	42	10	5
Luzula campestris agg.	29	62	66	61	60	32	11	31	21	5	2
Ranunculus repens	30	15	2	44	47	12	52	13	48	44	20
Angelica sylvestris	6	6	6	33	62	30	37	51	46	45	45
Trifolium pratense	56	61	7	50	19	12	13	3	24	5	4
Briza media	20	51	35	44	46	37	10	27	41	4	3
Vicia cracca	36	40	6	31	23	21	18	46	36	10	18
Lysimachia nummularia	27	8	.	28	15	5	33	.	51	27	21
Prunella vulgaris	28	30	8	34	23	28	14	7	45	7	4
Equisetum palustre	3	1	1	16	30	26	30	4	52	40	40
Centaurea jacea	35	32	11	40	10	36	11	1	24	1	4
Stellaria graminea	20	38	11	35	16	11	13	29	15	13	8
Avenula pubescens	26	29	7	38	21	17	12	22	7	1	7
Cirsium oleraceum	11	1	1	13	11	15	32	.	20	18	42
Lysimachia vulgaris	1	1	4	9	17	31	19	4	26	34	43
Equisetum arvense	21	6	3	13	12	12	16	6	38	18	23
Crepis paludosa	.	2	5	8	37	8	19	35	41	18	24
Carex hirta	7	5	2	29	14	10	20	2	35	18	11
Ajuga reptans	10	16	6	22	30	10	12	17	28	8	5
Agrostis stolonifera	8	1	3	16	38	13	16	9	32	19	12
Lotus uliginosus	1	2	2	19	31	25	17	11	12	18	12
Cynosurus cristatus	14	30	.	28	16	3	4	1	13	3	1
Carex ovalis	2	5	13	34	36	10	11	18	4	7	1
Galium verum agg.	28	21	1	14	2	23	10	.	3	.	5
Rhinanthus minor	11	28	3	24	10	5	2	7	4	3	.
Glechoma hederacea agg.	24	5	.	11	2	1	16	.	11	8	6
Colchicum autumnale	17	3	.	11	2	16	12	.	30	3	10
Pimpinella major	27	9	2	12	3	11	6	22	2	2	5
Trifolium dubium	24	22	.	16	2	2	4	.	5	1	.
Anemone nemorosa	2	9	13	16	18	14	5	25	4	3	11
Cirsium canum	11	1	.	15	4	17	17	.	12	4	9
Urtica dioica	7	2	.	1	1	2	16	7	3	23	27
Chaerophyllum hirsutum	1	3	1	2	10	.	13	33	15	15	23
Phleum pratense agg.	14	18	3	9	6	4	5	9	12	4	3
Trifolium hybridum	7	3	1	23	7	1	14	1	10	5	4
Vicia sepium	24	11	1	5	3	1	6	8	7	2	6
Carex acuta	1	.	.	5	8	4	21	2	4	12	20
Aegopodium podagraria	12	11	.	6	4	3	9	9	6	5	15
Anthriscus sylvestris	22	11	.	4	3	.	8	2	2	3	8
Cirsium arvense	16	4	1	7	2	8	16	.	1	5	8
Galium boreale	7	2	3	23	2	28	7	.	3	.	5
Carex brizoides	.	1	2	7	14	6	11	27	1	7	18
Symphytum officinale	13	1	.	4	.	5	17	.	8	6	9
Lythrum salicaria	1	.	.	2	3	15	13	.	17	15	17
Bellis perennis	14	13	.	16	2	2	1	.	6	1	2
Achillea ptarmica	.	3	2	11	15	11	9	19	.	5	6
Equisetum fluviatile	.	.	.	1	16	3	9	3	8	18	15
Leontodon autumnalis	10	17	3	8	1	4	3	1	4	.	2
Trollius altissimus	.	1	3	15	11	14	6	1	2	1	10
Festuca ovina	3	17	7	12	3	18	1	2	.	.	.
Geum rivale	.	1	.	4	14	5	9	15	18	4	9
Elytrigia repens	17	3	.	2	1	.	12	1	3	3	4
Saxifraga granulata	8	15	1	15	3	2	1	1	.	.	.
Hypericum perforatum	18	17	.	2	.	2	2	2	1	.	.
Poa palustris	2	1	.	4	4	1	13	2	21	7	7
Selinum carvifolia	3	2	2	12	6	20	4	.	9	1	5
Rumex obtusifolius	6	2	.	3	3	1	13	2	3	10	3
Potentilla reptans	13	5	.	2	.	5	8	.	6	3	2

App. 3. Comparison of the class membership of the plots in the expert classification and the class assignment of these plots by the MLP classifier in the test data set. The MLP classifier was trained with (A) randomly selected plots and (B) with plots containing high numbers of the class diagnostic species. Results of the classifier with the highest sensitivity on the test data set within each variant are shown. The numbers in the table are plot counts.

	Class no.	1	2	3	4	5	6	7	8	9	10	11	Σ
(A) Class assignment by the MLP classifier trained with randomly selected plots													
Classes of a priori expert classification	1	235	5	10	7	8	0	9	0	1	0	0	275
	2	6	23	0	1	3	0	0	0	2	0	5	40
	3	9	1	38	0	1	0	0	0	0	0	0	49
	4	7	2	1	55	17	0	7	0	0	0	0	89
	5	3	2	1	8	616	2	21	0	1	0	1	655
	6	1	0	0	1	1	14	0	0	0	0	0	17
	7	4	1	0	12	22	0	91	0	5	5	0	140
	8	0	0	1	0	0	0	0	9	4	0	4	18
	9	7	1	2	0	2	0	8	1	49	1	1	72
	10	0	0	0	0	0	0	0	1	3	4	0	8
	11	1	2	0	0	0	0	0	3	2	0	23	31
	Σ	273	37	53	84	670	16	136	14	67	10	34	1394
(B) Class assignment by the MLP classifier trained with plots rich in diagnostic species													
Classes of a priori expert classification	1	220	8	9	9	10	0	12	0	7	0	0	275
	2	1	28	1	1	4	0	0	0	2	0	3	40
	3	11	1	31	3	1	0	1	0	1	0	0	49
	4	1	1	0	65	11	0	11	0	0	0	0	89
	5	4	1	0	46	574	0	28	0	1	0	1	655
	6	1	0	0	2	1	11	2	0	0	0	0	17
	7	3	0	2	11	22	1	95	0	5	1	0	140
	8	0	0	1	0	0	0	1	11	1	0	4	18
	9	3	2	4	0	2	0	6	3	42	3	7	72
	10	0	0	0	0	0	0	0	0	0	8	0	8
	11	0	1	0	1	2	0	0	1	1	0	25	31
	Σ	244	42	48	138	627	12	156	15	60	12	40	1394

App. 4. Comparison of the class membership of the plots in the numerical classification and the class assignment of these plots by the MLP classifier in the test data set. The MLP classifier was trained with (A) randomly selected plots and (B) with plots containing high numbers of the class diagnostic species. Results of the classifier with the highest sensitivity on the test data set within each variant are shown. The numbers in the table are plot counts.

	Class no.	1	2	3	4	5	6	7	8	9	10	11	Σ
(A) Class assignment by the MLP classifier trained with randomly selected plots													
Classes of a priori numerical classification	1	155	36	1	8	0	0	9	0	0	0	0	209
	2	18	149	2	6	0	0	4	1	0	0	0	180
	3	0	6	74	2	0	1	3	0	1	0	0	87
	4	8	10	4	114	8	2	22	0	0	0	0	168
	5	0	1	4	9	81	0	18	0	2	3	2	120
	6	1	2	0	5	2	51	1	0	0	0	0	62
	7	9	1	1	10	13	1	165	0	3	13	9	225
	8	0	1	0	0	1	0	0	27	0	0	0	29
	9	0	0	0	1	0	0	3	0	53	6	0	63
	10	0	0	0	1	3	0	4	0	0	94	2	104
	11	0	0	0	2	0	2	17	3	1	1	121	147
	Σ	191	206	86	158	108	57	246	31	60	117	134	1394
(B) Class assignment by the MLP classifier trained with plots rich in diagnostic species													
Classes of a priori numerical classification	1	149	27	0	7	0	0	26	0	0	0	0	209
	2	27	129	10	8	0	1	4	1	0	0	0	180
	3	0	6	68	8	2	2	0	0	1	0	0	87
	4	2	13	0	124	4	4	21	0	0	0	0	168
	5	0	0	2	12	79	2	16	1	2	3	3	120
	6	1	1	1	5	2	52	0	0	0	0	0	62
	7	0	1	5	17	13	4	149	2	7	17	10	225
	8	0	0	0	1	0	0	0	26	0	0	2	29
	9	1	0	0	0	0	1	8	0	42	10	1	63
	10	0	0	0	0	2	0	4	0	5	91	2	104
	11	2	0	0	4	1	1	25	1	1	2	110	147
	Σ	182	177	86	186	103	67	253	31	58	123	128	1394