Journal of Vegetation Science **23** (2012) 387–393

FORUM

Towards consistency in vegetation classification

Miquel De Cáceres & Susan K. Wiser

Keywords
Cluster analysis; Diagnostic species; Fuzzy C-means; K-means; Physiognomy; Phytosociology; Relevé data; Supervised classification; TWINSPAN; Unsupervised classification; Vegetation databanks.; Vegetation mapping

Received 20 May 2011
Accepted 21 September 2011
Co-ordinating Editor: Michael Palmer

De Cáceres, M. (corresponding author, miquelcaceres@gmail.com): Biodiversity and Landscape Ecology Laboratory, Forest Science Center of Catalonia, Ctra antiga St Llorenç km 2, 25280, Solsona, Catalonia, Spain
Wiser, S. K. (wisers@landcareresearch.co.nz): Landcare Research, PO Box 40, Lincoln, 7640, New Zealand

Abstract

Vegetation classification is a useful tool for basic and applied research as well as for environmental management. As classification of vegetation serves many different purposes, there is no single approach to defining vegetation types. Establishing formalized standard procedures is desirable, however, because the purposes and uses of vegetation classifications are similar in different countries and regions. With the aim of promoting methodological standardization in classification across countries and vegetation scientists, this manuscript is centered on two ideas: (1) the need to explicitly distinguish between the conceptual activities involved in the definition of vegetation types (membership determination, characterization, validation and naming); and (2) the need to perform assignments of new vegetation observations to previously defined vegetation types in accordance with how these types were originally defined, a concept that we refer to as consistency in assignment. We demonstrate that our conceptual framework provides a useful tool to better understand what classification methods do. In order to manage and use classifications in a better way, vegetation scientists should produce, store and report the rules that provide consistent assignments to vegetation types.

Introduction

Vegetation types are abstract entities that delimit and name parts of the vegetation continuum to facilitate communication about them. As such, vegetation types provide a useful tool for basic and applied research, as well as for environmental management. Classification of vegetation can be based on one or multiple criteria that may include physiognomy, structure, plant functional traits, observed or potential species composition or climatic or soil conditions (e.g. UNESCO 1973; Walter 1973; Pfister & Arno 1980; Adams 1999; Dengler et al. 2008; Jennings et al. 2009). The entities to be grouped are usually vegetation stands delimited by plot boundaries, but they can also be specific vegetation strata within these boundaries or the pixels or polygons of an image. Moreover, different levels of abstraction (e.g. associations, alliances, classes, divisions or formations) and sampling/analytical approaches (i.e. size of sampling units, measurement protocols, data transformation, resemblance measures, clustering algorithms, etc.) are accepted and validly applied (Mucina 1997), and new methods continue to be proposed (e.g. De Cáceres et al. 2010; Schmidtlein et al. 2010; Tichý et al. 2011). As such, there is no universally valid approach for defining vegetation types. Rather, adopting one set of criteria over another should be based on practical considerations (Mucina 1997; Ewald 2003).

Despite the multiplicity of approaches, establishing formalized standard procedures to classify vegetation for specific needs is desirable because purposes and uses of vegetation classifications are not specific to one country or region. Rather, classification of vegetation is conducted to fulfill similar needs in different places. The exchange of vegetation information between countries, or the comparison of vegetation differences, is greatly facilitated if the same classification procedure is used in both places (e.g. Bruelheide & Chytrý 2000). Moreover, in cases where classifications of vegetation have a continental or global geographical scope (e.g. UNESCO 1973), mapping and monitoring of component vegetation patterns needs to be conducted using the same criteria everywhere (Adams 1996).

In this paper, our overall aim is to promote methodological standardization in classification among vegetation scientists and within and across countries and vegetation scientists. To pursue this objective, we propose precise definitions of very general concepts, and stress two main points: (1) the need to explicitly distinguish between the conceptual activities involved in the definition of vegetation types; and (2) the need to perform assignments of new vegetation observations to previously defined vegetation types in accordance with how these types were originally defined. We refer to this second point as *consistency* in assignments. We demonstrate that our conceptual framework provides a useful tool to better understand what current classification methods do, and enables the identification of critical issues that must be addressed if classifications of vegetation are to be managed in a better way. We believe that if vegetation scientists are more aware of the conceptual implications and long-term practical consequences of their choices, methodological standardization will more readily arise.

Before beginning our discussion, we need to establish some necessary notation. We assume that the vegetation area of interest has been surveyed in the field, or sensed by some other means, and that a set of *vegetation observations* are therefore available. The reader can think of vegetation observations as sampling units such as plot-based records (e.g. relevés) but they can have other forms (e.g. pixels of a satellite image). We use the term *vegetation type* to denote a grouping of vegetation observations at any level of abstraction, based on a set of relevant features, which we will call *vegetation attributes*. Finally, we use the term *classification scheme* to denote a set of vegetation types, which can belong to a single abstraction level or be organized into different abstraction levels (e.g. associations, alliances, divisions, formations, classes) (Dengler et al. 2008; Jennings et al. 2009).

Defining vegetation types

Membership determination, validation, characterization and naming

We suggest that four main activities regarding the definition of vegetation types should be distinguished (see Table 1). *Membership determination* is an activity related to deciding which vegetation observations belong to which types. Membership determination answers the question 'How do I group my vegetation observations?' or 'Which type does this particular observation belong to?' This contrasts to the questions answered by *characterization* ('What are the attributes of my grouping?'), *validation* ('Is this grouping acceptable for my purpose?') and *naming* ('How do we refer to this grouping?'). We illustrate these four steps taking, as an example, the floristically based definition of associations from a set of plot records. First, the membership to associations is determined by clustering plot records based on their resemblance in the multivariate space of species composition. Second, one can characterize each association using site environmental attributes (e.g. climate, soil properties, disturbance or management regimes, etc.), the geographical range occupied or its diagnostic species. Third, to decide whether the resulting types should be accepted or not, one could study their statistical robustness (e.g. Tichý et al. 2011) or examine the requirements needed for associations, following published recommendations (e.g. Willner 2006). Finally, the associations will need to be named according to some naming conventions (e.g. Weber et al. 2000).

Table 1. Activities related to the definition of vegetation types. Each of these activities can be conducted on the basis of different vegetation attributes.

Name	Activity
Membership determination	Grouping vegetation observations, or assigning a given observation to a pre-existing vegetation type or defining how the assignment should be conducted (see Table 2)
Characterization	Production of attributes that apply to a (set of) vegetation type(s)
Validation	Determining whether a given vegetation type (or the entire classification) is acceptable for a given application or not
Naming	Assigning a label to each vegetation type, according to a set of conventions

Why are these activities difficult to distinguish?

That the results of a single analysis may be used for membership determination, characterization or validation can hinder us in distinguishing these activities conceptually. For example, a linear discriminant analysis run on topographic data from a set of pre-defined vegetation types will calculate the mean and variance values of altitude, aspect and slope for each vegetation type. In addition to this characterization, the same analysis will reveal which vegetation types cannot be separated topographically and this result may lead us to reconsider the appropriateness of our classification scheme (validation). Moreover, it will provide us with a function that can be used to assign new vegetation observations to our types based on topography (membership determination). As another example of multiple activities resulting from a single procedure, OptimClass (Tichý et al. 2010) is a technique that allows many clustering alternatives of the same set of vegetation plot records to be explored (membership determination) and rates each classification result using the number of diagnostic species (validation).

The need for clarity in published classification standards

An important corollary of acknowledging the distinction between activities is that scientists aiming to promote standard conventions for vegetation classification must clearly specify what type of information is used for each activity and how the activity is conducted. In this sense, we think

Journal of Vegetation Science
Doi: 10.1111/j.1654-1103.2011.01354.x

it is fundamental that published guidelines disambiguate between *membership determination* and *validation*. In his revision of the phytosociological association concept, Willner (2006) suggested going back to Flahaut & Schröter's (1910) definition of association as 'an abstract vegetation type that has a definite floristic composition, a uniform physiognomy, and occurs in uniform habitat conditions.' A similar definition of association is provided by the US National Vegetation Classification (NVC, Jennings et al. 2009), where an association is: 'a vegetation classification unit defined on the basis of a characteristic range of species composition, diagnostic species occurrence, habitat conditions, and physiognomy.' From their use of words like 'uniform' or 'characteristic range', we conclude that these two definitions are stating necessary criteria for the validity of associations. But how are memberships determined? Jennings et al. (2009) later state that: 'for type definition, numerical multivariate analysis of the species composition is typically used to arrange the plots that span the compositional and geographic range into discrete types, as well as to show their relation to other types.' Note that the use of the words 'defined on the basis of' and 'for type definition' does not help the reader to clearly understand which activity is being described in each of these sentences. The need for clarity is even more important when the criteria for membership definition and validation vary across different levels of abstraction (e.g. Faber-Langendoen et al. 2009).

Determining the membership to vegetation types

Definition of membership statements and membership rules

We use the term *membership statement* to denote an expression specifying which vegetation observations belong to which vegetation types (e.g. 'plot records *a, b* and *c* belong to vegetation type X, whereas plot record *d* does not belong to it'). A vector containing either zeroes or ones is a numerical way to represent membership statements of a set of vegetation observations to a given type. Membership can be a continuous variable, not only binary, and be interpreted as a probability (e.g. 'plot record *e* belongs to unit X with probability 0.7') or be fuzzy (e.g. 'plot record *f* has a membership degree of 0.6 to unit Y and a membership degree of 0.4 to unit Z') (e.g. De Cáceres et al. 2009).

We use the term *membership rule* to denote a procedure that allows individual vegetation observations to be assigned to vegetation types, i.e. membership rules are classifiers. Simple examples of membership rules are: 'a plot record belongs to vegetation type Y if its altitude lies within a given altitudinal range,' or 'a plot record belongs to vegetation type Y if species A occurs among the list of species recorded.' Examples of more complex rules would be a hierarchical vegetation key (e.g. Pfister & Arno 1980; Rodwell 1991; Adams 1999), a set of species groups plus a formal logic statement combining them (Bruelheide & Flintrop 1994; Bruelheide 1997, 2000), a fuzzy membership function based on the distances of the target observation to a set of cluster prototypes (e.g. De Cáceres et al. 2009, 2010) or a trained neural network (e.g. Černá & Chytrý 2005). Moreover, two or more membership rules may be combined to create a compound rule (e.g. Kočí et al. 2003).

Activities related to membership determination

Different activities can be generally referred to as 'classification'. In the following discussion we distinguish four fundamentally different activities regarding the determination of membership to vegetation types (Table 2).

1. *Expert-based rule definition*: Vegetation scientists sometimes define membership rules using their expertise, without any explicit use of, or reference to, vegetation observations. This is often the case for global-level vegetation classifications based on climate, physiognomy and/or structure (e.g. UNESCO 1973; Adams 1999).

2. *Unsupervised classification*: Unsupervised classification (or clustering) methods allow vegetation observations to be grouped and hence produce membership statements. Clustering methods differ in the criterion they use to group

Table 2. Activities related to the determination of membership to vegetation types.

Name	Activity
Expert-based rule definition	Definition of a membership rule from expert knowledge without any explicit reference to vegetation observations
Unsupervised classification	Classification of unlabelled vegetation observations into subsets or clusters on the basis of similar vegetation attributes. The result is a set membership statements, but membership rules may also be obtained in some cases
Supervised classification	Inference of a membership rule (i.e. a classifier) from a set of training vegetation observations whose membership statements are known in advance
Assignment	Application of a membership rule to a vegetation observation to obtain a membership statement
Consistent assignment	Assignments with a rule that when applied to observations of known membership reproduces the same memberships
Indicative assignment	Assignments with a rule that when applied to observations of known membership produces similar but not the same memberships

Doi: 10.1111/j.1654-1103.2011.01354.x

observations (Jain & Dubes 1988; Legendre & Legendre 1998). Although it is not their primary aim, some clustering methods also provide membership rules.

a. *Agglomerative hierarchical clustering* – Despite their popularity for quantitative classification, hierarchical agglomerative clustering methods (e.g. complete linkage, UPGMA, Ward's method) do not allow new observations to be classified without rebuilding the entire hierarchy. Because they do not provide membership rules, we believe that using agglomerative hierarchical clustering is not a good strategy for vegetation classifications that are intended to be endurable.

b. *Divisive hierarchical clustering* – Hierarchical divisive methods (e.g. association analysis or division in ordination space) produce a membership rule in each successive division. The set of all division provides a compound rule that allows membership to be determined through all hierarchical levels.

c. *Non-hierarchical clustering* – Many non-hierarchical clustering approaches allow new observations to be classified *a posteriori*. Examples of such methods are partitioning around medoids (Kaufman & Rousseuw 1990) and the K-means/Fuzzy C-means family (MacQueen 1967; Bezdek 1981). All these methods are based on the idea that each cluster is represented by a prototype (e.g. a centroid or a medoid). Broadly speaking, the algorithm of these methods iterates two steps: (1) the assignment of observations to clusters whose prototype is closest in the multivariate attribute space; and (2) the recalculation of prototype locations. Step (1) can be taken as a membership rule for the assignment of new observations (De Cáceres et al. 2009, 2010).

3. Supervised classification: Supervised classification methods take the membership statements and the attributes of a set of 'training' vegetation observations and define a membership rule. Supervised classification provides a function to classify vegetation observations based on their affinity to a set of pre-defined units. Examples are neural networks, classification trees and discriminant analysis (e.g. Černá & Chytrý 2005; van Tongeren et al. 2008; Ejrnaes et al. 2009). The rule generated can be based on the same set of vegetation attributes that was originally used in the classification of the training observations or it can be based on a completely different set of attributes. In the latter case, however, the purpose of the analysis may be to extrapolate from the original classification exercise. For example, vegetation maps are often produced using a membership rule that, while being based on GIS or satellite image data, allows pixels or polygons to be classified into vegetation types that were originally defined using floristic or physiognomic criteria (e.g. van Etten 1998).

4. Assignments: We use the term *assignment* to denote the application of a membership rule to a vegetation observation to obtain a membership statement. Although in many cases the membership rule will come from a supervised classification method (3), it may also originate from expert definition without explicit use of data (1) or from a clustering method (2). The existence of alternative origins of a membership rule is the reason why we separated the activity of producing the rule from the act of applying it.

Consistency in vegetation classification

Consistency in assignments

We argue that assignments with a membership rule are *consistent* with a given set of membership statements if, and only if, the rule reproduces the same memberships when used to assign the same observations to which the statements refer (Table 2). For example, imagine that the initial membership statements are 'plot records *a*, *b* and *c* belong to vegetation type X' and 'plot records *d* and *e* belong to vegetation type Y.' Now imagine a membership rule that determines membership to either unit X or unit Y based on the spectrum of Raunkiaer's life forms found in the plot record. The assignment with this rule is consistent with the original statements if, and only if, the rule assigns records *a*–*c* to X and records *d*–*e* to Y. If not, we may say that the assignment is *indicative* of the membership statements, but we cannot state that the assignment is consistent with the initial classification (Table 2).

Why is consistency in assignments important?

Among others uses, vegetation classifications are expected to make the patterns of vegetation communicable (Dengler et al. 2008). New vegetation observations are continuously made and their membership to existing vegetation types needs to be known in order to map or monitor vegetation. Having multiple membership rules for the same set of vegetation types is justified because many vegetation attributes can provide valuable information about a target vegetation stand. Nevertheless, we know that membership rules based on different functions and data are unlikely to always provide the same answer. In the same way that vegetation classifications are conventions, there should be a method to determine the membership rule that is conventionally taken as the single 'true' rule. The natural way to establish this convention is to require that the assignments of vegetation observations must be done using the same criteria that were employed to create the classification. This means that the 'true' rule is the one that, when applied to observations underpinning the classification, exactly reproduces the classification membership

Journal of Vegetation Science
Doi: 10.1111/j.1654-1103.2011.01354.x

statements. Adopting this convention does not restrict the form of membership statements, which can be probabilistic or deterministic, fuzzy or crisp and point to one or many vegetation types. For a given classification scheme, however, there should be a single conventional response to the question of membership of vegetation observations, even if the answer is that the observation cannot be classified, is ambiguous or the observation belongs to more than one unit. These considerations lead us to consider the assignments with membership rules that are not consistent with the classification as *indicative* but not 'true' determinations of membership (Table 2).

Consistency is also important when comparing classifications. The relationship between two classifications is studied either by comparing their membership rules (e.g. are assignment criteria similar in classifications A and B?) or by comparing the membership statements that these rules produce when applied to the same set of vegetation observations (e.g. if many of the plot records that are assigned to type X in classification A are assigned to type Y in classification B, then there is a relationship between X and Y). These comparisons are useful in order to relate abstraction levels within a single classification scheme (e.g. to study the relationship between membership rules used for alliances and formations within the US NVC framework) or to relate two completely different classification schemes (e.g. Bruelheide & Chytrý 2000). If the rules used for these comparisons are not consistent with how the classifications were originally defined, however, then the comparison will be biased by this lack of consistency.

When do membership rules provide consistent assignments?

1. *Expert-defined rules* – When experts define membership rules without relying on vegetation observations, there is no need to evaluate consistency of assignments, because the definition of vegetation types did not produce any membership statement.

2. *Rules issued from unsupervised classification* – When clustering methods provide membership rules, the assignments using these rules are consistent with the membership statements that the method produces for the input data set. For example, if one uses Fuzzy C-means to partition a set of vegetation plot records into fuzzy clusters, in the future one will be able to use the distance to the cluster centroids to consistently assign new plot records (De Cáceres et al. 2010). Because it is one of the more popular unsupervised hierarchical divisive algorithms for vegetation classification, it is important to examine TWINSPAN (Hill 1979; Roleček et al. 2009). The TWINSPAN program can produce simple discriminant functions that, based on indicator species, allow plot records to be assigned to previously established types (e.g. Bruelheide & Chytrý 2000). Applying these rules to the original plot data, however, does not always exactly reproduce the original membership statements. Thus assignments with the rules produced by the program are not completely consistent with the way the unsupervised classification was conducted, which involved a division based on an ordination step. In order to provide consistent assignments, the TWINSPAN program should be modified to provide a membership rule based on the projection of plot records on ordination axes.

3. *Rules issued from supervised classification* – When the membership rules produced by a supervised classification method are used to assign the original observations, new membership statements are obtained. The higher the predictive performance of the supervised classification method, the closer these new membership statements will be to the original ones. Assessing the predictive accuracy of the membership rule is important to determine its validity. However, rules coming from supervised techniques do not always allow the original statements to be reproduced exactly, so consistency is not always achieved. This does not mean that supervised classification methods are not useful. They are the way to produce many rules for indicative assignments that are useful for activities such as field identification and vegetation mapping.

Final remarks: the need to report and store membership rules

In the previous section we stressed the importance of producing and using consistent membership rules. We would like to finish by reiterating the need to store and report them. Published scientific or technical reports on classification exercises commonly include the region and/or vegetation broad class of interest (e.g. forests, wetlands, grasslands, etc.), the set of vegetation observations gathered, the methods used to classify them and the characterization (e.g. synoptic tables), and perhaps validation, of the resulting types. It is rare, however, that they indicate how one would proceed to determine whether a given new plot record fits into any of the published classes. Even when user-friendly vegetation keys are published along with the description of vegetation types (e.g. Rodwell 1991), the consistency of those rules may have not been assessed. A similar situation occurs with vegetation plot databanks. Some database systems include fields that allow membership statements to be stored (Hennekens & Schaminée 2001; Dengler et al. 2011; Wiser et al. 2011), but databanks rarely store the information required to determine the membership of new observations. Databanks would benefit from being able to store explicit membership rules from published classifications, in order to permit new vegetation observations to be assigned to vegetation types.

Failing to store and report membership rules in the past has forced current assignments of new observations to be made using the membership rules created by supervised classification. Indeed, supervised classification is often used to create membership rules when the original classification was performed employing an unsupervised strategy that did not produce membership rules, or these were not recorded, or were not formally and unequivocally specified. This is frequently the case with many legacy classification schemes of the Braun-Blanquet method (Černá & Chytrý 2005; van Tongeren et al. 2008). The most important consequence of having to resort to supervised classification is the resulting loss of consistency.

For any given classification scheme, several membership rules should be made available. Among them, there should be one rule that the authors of the classification recommend for consistent assignments. It should be the duty of the authors of the classification to check that the recommended rule does indeed reproduce the original membership statements when applied to the original vegetation observations. If such a rule is complex to apply (i.e. it involves complex numerical operations), there could be a web-based service that performs the necessary operations on data uploaded by the user. In order to complement this service, other user-friendly rules should be made available for easy and fast identification when access to consistent assignment is not possible.

Acknowledgements

MDC was supported by a Beatriu de Pinós postdoctoral grant (2009 BP-B 00342) from the Catalan Agency for Management of University and Research Grants, and a Consolider Montes CSD2008-00040 project granted by the Spanish Ministry of Education and Science (MEC) and the International Mobility Fund administered by the Royal Society of New Zealand (contract SPN 10-13). SKW was supported by the New Zealand Ministry for Science and Innovation (contract C09X0916). We thank JM Fernández-Palacios, D Faber-Langendoen and an anonymous reviewer for their constructive criticisms on the content and structure of previous versions of this manuscript, to Hamish Maule and Fiona Thomson for their friendly review, and to Christine Bezar for final formatting and editing.

References

Adams, J.M. 1996. Towards a better vegetation scheme for global mapping and monitoring. *Global Ecology and Biogeography Letters* 5: 3–6.

Adams, J.M. 1999. A suggestion for an improved vegetation scheme for local and global mapping and monitoring. *Environmental Management* 23: 1–13.

Bezdek, J.C. 1981. *Pattern recognition with fuzzy objective functions*. Plenum Press, New York, NY, US.

Bruelheide, H. 1997. Using formal logic to classify vegetation. *Folia Geobotanica* 32: 41–46.

Bruelheide, H. 2000. A new measure of fidelity and its application to defining species groups. *Journal of Vegetation Science* 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. *Journal of Vegetation Science* 11: 295–306.

Bruelheide, H. & Flintrop, T. 1994. Arranging phytosociological tables by species-relevé groups. *Journal of Vegetation Science* 5: 311–316.

Černá, L. & Chytrý, M. 2005. Supervised classification of plant communities with artificial neural networks. *Journal of Vegetation Science* 16: 407–414.

De Cáceres, M., Font, X., Vicente, P. & Oliva, F. 2009. Numerical reproduction of traditional classifications and automatic vegetation identification. *Journal of Vegetation Science* 20: 620–628.

De Cáceres, M., Font, X. & Oliva, F. 2010. The management of vegetation classifications with fuzzy clustering. *Journal of Vegetation Science* 21: 1138–1151.

Dengler, J., Chytrý, M. & Ewald, J. 2008. Phytosociology. In: Jørgensen, S.E. & Fath, B.D. (eds.), *General ecology. Vol. [4] of encyclopedia of ecology*. Elsevier, Oxford, UK.

Dengler, J., Jansen, F., Glöckler, F., Chytry, M, De Cáceres, M., Ewald, J., Oldeland, J., Peet, R.K., Finckh, M., Mucina, L., Schaminnée, J. & Spencer, N. 2011. The Global Index of Vegetation-Plot Databases (GIVD): a new resource for vegetation science. *Journal of Vegetation Science* 22: 582–597.

Ejrnaes, R., Bruun, H.H., Aude, E. & Buchwald, E. 2009. Developing a classifier for the Habitats Directive grassland types in Denmark using species lists for prediction. *Applied Vegetation Science* 7: 71–80.

Etten, E.J.B. van 1998. Mapping vegetation in an arid, mountainous region of Western Australia. *Applied Vegetation Science* 1: 189–200.

Ewald, J. 2003. A critique for phytosociology. *Journal of Vegetation Science* 14: 291–296.

Faber-Langendoen, D., Tart, D.L. & Crawford, R.H. 2009. Contours of the Revised U.S. National Vegetation Classification Standard. *Bulletin of the Ecological Society of America* 90: 87–93.

Flahaut, C. & Schröter, C. 1910. Rapport sur la nomenclature phytogéographique. *Third International Botanical Congress (IBC)(Brussels)* 1: 131–164.

Hennekens, S.M. & Schaminée, J.H.J. 2001. TURBOVEG, a comprehensive data base management system for vegetation data. *Journal of Vegetation Science* 12: 589–591.

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.

Jain, A.K. & Dubes, R.C. 1988. *Algorithms for clustering data*. Prentice Hall, Englewood Cliffs, NJ, US.

Doi: 10.1111/j.1654-1103.2011.01354.x

Jennings, M.D., Faber-Langendoen, D., Loucks, O.L., Peet, R.K. & Roberts, D. 2009. Standards for associations and alliances of the U.S. National Vegetation Classification. *Ecological Monographs* 79: 173–199.

Kaufman, L. & Rousseuw, P.J. 1990. *Finding groups in data: an introduction to cluster analysis*. Wiley, New York, NY, US.

Kočí, M., Chytrý, M. & Tichý, L. 2003. Formalized reproduction of an expert-based phytosociological classification: a case study of subalpine tall-forb vegetation. *Journal of Vegetation Science* 14: 601–610.

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

MacQueen, J. 1967. Some methods for classification and analysis of multivariate observation. In: LeCam, L.M. & Neyman, J. (eds.), *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, pp. 281–297. University of California Press, Berkeley, CA, US.

Mucina, L. 1997. Classification of vegetation: past, present and future. *Journal of Vegetation Science* 8: 751–760.

Pfister, R.D. & Arno, S.F. 1980. Classifying Forest Habitat Types Based on Potential Climax Vegetation. *Forest Science* 26: 52–70.

Rodwell, J.S. 1991 et *seq. British plant communities*, Vols. 1–4. Cambridge University Press, Cambridge, UK.

Roleček, J., Tichý, L., Zelený, D. & Chytrý, M. 2009. Modified TWINSPAN classification in which the hierarchy respects cluster heterogeneity. *Journal of Vegetation Science* 20: 596–602.

Schmidtlein, S., Tichý, L., Feilhauer, H. & Faude, U. 2010. A brute-force approach to vegetation classification. *Journal of Vegetation Science* 21: 1162–1171.

Tichý, L., Chytrý, M., Hájek, M., Talbot, S.S. & Botta-Dukát, Z. 2010. OptimClass: using species-to-cluster fidelity to determine the optimal partition in classification of ecological communities. *Journal of Vegetation Science* 21: 287–299.

Tichý, L., Chytrý, M. & Smarda, P. 2011. Evaluating the stability of the classification of community data. *Ecography* (on-line version). doi: 10.1111/j.1600-0587.2010.06599.x.

Tongeren, O. van, Gremmen, N. & Hennekens, S. 2008. Assignment of relevés to pre-defined classes by supervised clustering of plant communities using a new composite index. *Journal of Vegetation Science* 19: 525–536.

UNESCO 1973. *International classification and mapping of vegetation*, pp. 1–93. UNESCO Ecology and Conservation Series 6. UNESCO, Paris, FR

Walter, H. 1973. *Vegetation of the earth*. Springer-Verlag, Berlin, DE. 318pp.

Weber, H.E., Moravec, J. & Theurillat, J.-P. 2000. International Code of Phytosociological Nomenclature. 3rd edition. *Journal of Vegetation Science* 11: 739–768.

Willner, W. 2006. The association concept revisited. *Phytocoenologia* 36: 67–76.

Wiser, S.K., Spencer, N., De Cáceres, M., Kleikamp, M., Boyle, B. & Peet, R.K. 2011. Veg-X – An exchange standard for plot-based vegetation data. *Journal of Vegetation Science* 22: 598–609.