Folia Geobot (2013) 48:437–447
DOI 10.1007/s12224-013-9152-6

Does Ordinal Cover Estimation Offer Reliable Quality Data Structures in Vegetation Ecological Studies?

Carlo Ricotta · Enrico Feoli

Abstract Using visual estimation of species cover in ordinal interval classes may reduce costs in vegetation studies. In phytosociology, species cover within plots is usually estimated according to the well-known Braun-Blanquet scale and ordinal data from this scale are usually treated using common exploratory analysis tools that are adequate for ratio-scale variables only. This paper addresses whether the visual estimation of ordinal cover data and the treatment of these data with multivariate procedures tailored for ratio-scale data would lead to a significant loss of information with respect to the use of more accurate methods of data collection and analysis. To answer these questions we used three data sets sampled by different authors in different sites of Tuscany (central Italy) in which the species cover is measured with the point quadrat method. For each data set we used a Mantel test to compare the dissimilarity matrices obtained from the original point-quadrat cover data with those obtained from the corresponding ordinal interval classes. The results suggest that the ordinal data are suitable to represent the plot-to-plot dissimilarity structure of all data sets in a reasonable way and that in using such data there is no need to apply dissimilarity coefficients specifically tailored for ordinal scales.

Keywords Braun-Blanquet scale · Dissimilarity coefficients · Mantel test · Ordinal transformation · Species cover abundance

> *Heretics are those who pursue*
> *orthodoxy too vigorously*
>
> Leland Wilkinson (in Hand 1996)

C. Ricotta (✉)
Department of Environmental Biology, University of Rome 'La Sapienza', Piazzale Aldo Moro 5, 00185 Rome, Italy
e-mail: carlo.ricotta@uniroma1.it

E. Feoli
Department of Life Sciences, University of Trieste, Via E. Weiss 2, 34127 Trieste, Italy
e-mail: feoli@units.it

Introduction

Data on plant species abundances may be collected in several ways and treated using many methods (either numerical or not) depending on the aim of the study and the required precision. Obviously, the more precise the sampling method is, the higher the costs are in terms of time (and money). In phytosociology, the species cover within plots (called relevés) is usually estimated according to an ordinal alphanumeric scale first proposed by Braun-Blanquet (1932; hereafter BB scale).

According to the original proposal of Braun-Blanquet (see e.g., Braun-Blanquet 1932), an alphanumerical code composed of both numbers and non-numbers (r, +, 1, 2, 3, 4 and 5) is assigned to each species based on a visual estimation of the percentage of its cover in a given plot. Clearly, the presence of non-numbers automatically complicates any numerical manipulation (Podani 2006); therefore, several methods have been proposed to convert the alphanumerical values of the BB scale to some other scale expressed by natural numbers (revised in van der Maarel 1979). The most frequently used transformation is the one proposed by van der Maarel (1979); it consists in converting the scores of the BB alphanumerical scale to ranks or ‘ordinal transform values’ (OTV) ranging from 1 to 9, where the score 2 on the original BB scale is divided into 3 scores corresponding to finer abundance/dominance classes (see Table 1).

A well-known effect of ordinal transformation is a sort of ‘downweighting’ of species with high cover. Nonetheless the precision level obtained with the BB scale is considered adequate for the information it provides and the knowledge it has generated on terrestrial vegetation systems (Dengler et al. 2011; Feoli et al. 2011; Schaminée et al. 2011). Due to its operational simplicity, the BB scale and its ordinal transformation has also been frequently used for estimating species cover in studies that have little to do with syntaxonomy and phytosociology. For example, recently Ricotta and Moretti (2010) used a BB-type scale to describe vegetation plots sampled in a chestnut forest in southern Switzerland for quantifying the degree of the species functional differentiation with respect to different fire regimes, while Samaritani et al.

Table 1 Cover percentage approximations of the Braun-Blanquet cover scale used in this study. Based on species cover, the original BB score 2 was subsequently divided into three OTVs corresponding to finer abundance/dominance classes (see van der Maarel 1979, 2007)

BB original scale	Ordinal Transform Values (OTVs)	Cover interval (%)
r	1	<0.5
+	2	0.5–1
1	3	1–2.5
2	4	2.5–5
2	5	5–12.5
2	6	12.5–25
3	7	25–50
4	8	50–75
5	9	>75

(2011) used OTV data in a study aimed at understanding organic carbon dynamics in soils of riparian vegetation.

However, in spite of the widespread success of the BB scale and its ordinal transformation, its use is justified only if the results obtained are equivalent to those obtained with more accurate methods for estimating species cover (see e.g., Moore and Chapman 1986). In addition, regardless of the aim of the study, a number of authors (e.g., Dale 1989; Podani 2006) argue that the OTVs of the BB scale cannot be properly treated using usual exploratory data analysis tools, like multivariate ordination and clustering methods, because these methods are generally adequate for ratio-scale variables only, while the OTVs are ordinal ones; as such, they need to be manipulated only by multivariate procedures applicable to ordinal scales.

Podani (2006; see also Podani 2005) has recently emphasized this point. As observed by Podani (2006), most numerical syntaxonomy studies analyzing BB-type data disregard this constraint and are based on multivariate procedures like ordination or clustering methods that use arithmetic operations as addition or division that are adequate for ratio-scale variables only. That is, for scales that possess a natural zero and for which both intervals and ratios between values are meaningful. On the contrary, for ordinal scales the values used for the measurement are interpreted only in terms of their arrangement in a given order, for example, from least to most abundant. Therefore, differences between states and their ratios are not interpreted (Anderberg 1973; Dale 1989). "On the ordinal scale only the relations = and < are meaningful, so that only the ordering of values would convey information. In other words, the difference between 1 and 2 is not necessarily as that between 3 and 4, 1+2 is not equal to 3 and 1/2 is not the same as 2/4" (Podani 2006).

In this framework, Ricotta and Avena (2006) argued that the common use of conventional multivariate methods to analyze OTVs is not due to 'mathematical ignorance', but rather to the common observation that conventional multivariate methods reproduce the researcher's intuitive classification/ordination scheme at least as good as multivariate methods explicitly developed for ordinal data. This paper aims thus to explore the effects of ordinal data transformation and the performance of metric *vs* ordinal dissimilarity measures on the structure of a plot-to-plot resemblance matrix. To do this, using six well-known ratio-scale dissimilarity coefficients we first calculated the resemblance matrices of three sets of vegetation plots in which the species cover is measured accurately with the point-quadrat method. Next we compared these matrices with the matrices obtained applying both ratio-scale and ordinal dissimilarity measures on the OTVs. The idea behind this approach is that the results of ordination and classification methods depend always on the structure of the resemblance matrices (Feoli and Orloci 1991; Podani 2007). Accordingly, to be relevant, the resemblance matrices calculated from the OTVs need to correspond as closely as possible to the resemblance matrices calculated from the non-transformed percentage cover data, which we consider the 'ground truth' of the vegetation systems under study.

Data

We analyzed three sets of plots sampled by a random design by different authors in different sites in Tuscany (central Italy): A) 35 1-m-square plots sampled in spring 1994 in a species-poor garrigue on serpentine soils south of Siena (Chiarucci et al.

1998). Total number of species is 31. B) 40 1-m square plots sampled in late June – early July 2000 in a calcareous grassland on Mt. Labbro (132 species; Maccherini et al. 2007). C) 56 1-m square plots sampled in May 2005 in shrub and herbaceous badland vegetation in the Lucciola Bella Natural Reserve (90 species; Marignani et al. 2007). In all data sets the cover of each vascular plant species is measured on a metric scale with the point-quadrat method (Moore and Chapman 1986) with a density of 441 pins/m^2 for set A and 100 pins/m^2 for set B and C. Species present in a plot but not touched by any pins were recorded with an arbitrary cover of 0.1 %.

All matrices can be considered very 'sparse' matrices (i.e., containing many zeros) with different degrees of species occupancy. Matrix A has the highest rate of occupied cells in the matrix, or occupancy ratio (C=0.44), together with the highest floristic homogeneity among plots calculated as the average Jaccard similarity value for presence/absence data (J=0.48; see Table **2**). Matrix C has both the lowest occupancy ratio (C=0.08) and the lowest floristic homogeneity (J=0.10), as well as the lowest number of species per plot (mean species per plot=10). However, Matrix C has the highest value of mean species cover (8 %). As the occupancy ratio of a presence/absence matrix of species × plots is a direct measure of average species rarity, not surprisingly, this measure is tightly related to floristic homogeneity.

Methods

Using six well-known and widely-used ratio-scale dissimilarity coefficients (e.g., Faith et al. 1987), we constructed six pairwise dissimilarity matrices for each set of plots. These matrices are considered a reliable reference for the vegetation systems under study. Let X = $\{x_{ij}\}$ denote the raw data matrix with n species (rows) and m plots (columns). Let further x_{i+} and x_{+j} represent row i and column j of X, respectively. These dissimilarity coefficients are:

$$\text{Euclidean distance ED} = \sqrt{\sum\nolimits_i \left(x_{ij} - x_{ik}\right)^2} \tag{1}$$

$$\text{Manhattan distance (or city-block metric) } CB = \sum\nolimits_i \left|x_{ij} - x_{ik}\right| \tag{2}$$

Table 2 Summary statistics of the data matrices used in this study (species-poor garrigue south of Siena, Mt. Labbro and Lucciola Bella, respectively). s/p – mean species number per plot; C – rate of occupied cells in the matrix

	Matrix A	Matrix B	Matrix C
Nr. of species	31	132	127
Nr. of plots	35	40	56
s/p	15.52	21.20	9.89
Occupancy rate (C)	0.44	0.16	0.08
Mean species cover	1.07	4.09	8.10
Mean Jaccard similarity among plots (J)	0.48	0.22	0.10

The Euclidean and the Manhattan distances are special cases of a more general parametric family of distance coefficients, the Minkowski distance $\mathrm{MD}_\alpha = \left(\sum_i \left|x_{ij} - x_{ij}\right|^\alpha\right)^{1/\alpha}$ where $\alpha \geq 1$; for $\alpha = 1$, we have the Manhattan distance, while for $\alpha = 2$, we have the Euclidean distance. Because for the calculation of the Euclidean distance the differences are squared before pooled, ED is more sensitive to large differences than CB.

$$\text{Chord distance CH} = \sqrt{2\left(1 - \frac{\sum_i x_{ij}x_{ik}}{\sqrt{\sum_i {x_{ij}}^2 \sum_i {x_{ik}}^2}}\right)} \tag{3}$$

CH is obtained by normalizing the vectors pointing to the plots to unit length and computing the Euclidean distance afterwards.

$$\text{Bray-Curtis index BC} = \frac{\sum_\mathrm{i} \left|\mathrm{x_{ij}} - \mathrm{x_{ik}}\right|}{\sum_\mathrm{i} \left(\mathrm{x_{ij}} + \mathrm{x_{ik}}\right)} \tag{4}$$

$$\text{Canberra metric CM} = \sum\nolimits_i \frac{\left|x_{ij} - x_{ik}\right|}{\left(x_{ij} + x_{ik}\right)} \tag{5}$$

CM is obtained by standardizing the absolute differences in the Manhattan distance with the sum of the two values (Podani 2000) and differs from the Bray-Curtis index, because in BC there is a separate summation for the absolute differences in the numerator and for the sum of values in the numerator.

$$\text{Marczewski-Steinhaus coefficient MS} = \frac{\sum_\mathrm{i} \left|\mathrm{x_{ij}} - \mathrm{x_{ik}}\right|}{\sum_\mathrm{i} \max\left\{\mathrm{x_{ij}}, \mathrm{x_{ik}}\right\}} \tag{6}$$

The cover values of all species in the raw matrices were then converted to OTVs according to Table **1**. Next, for each ordinal data matrix, we calculated a pairwise dissimilarity matrix using the six dissimilarity coefficients listed above and four dissimilarity matrices with the following coefficients explicitly developed for ordinal scales (see Podani 2000):

$$\text{Kendall's tie-adjusted tau } \tau = 1 - \frac{2(\mathrm{a} - \mathrm{b})}{\sqrt{\left[\mathrm{n}(\mathrm{n}-1) - 2\mathrm{T_j}\right]\left[\mathrm{n}(\mathrm{n}-1) - 2\mathrm{T_k}\right]}} \tag{7}$$

where n is the number of species, a is the number of pairs of species i and h ordered identically for plots j and k ($x_{ij} > x_{ik}$ and $x_{hj} > x_{hk}$), b is the number of pairs of species that are ordered reversely in j and k ($x_{ij} > x_{ik}$ and $x_{hj} < x_{hk}$). T_j and T_k are the numbers of tied species pairs in plots j and k, respectively ($x_{ij} = x_{hj}$ and $x_{ik} = x_{hk}$).

$$\text{Goodman and Kruskal's gamma } \gamma = 1 - (a - b)/(a + b) \tag{8}$$

$$\text{Relative rank difference } RD = \frac{1}{n} \sum\nolimits_i \frac{\left|r_{ij} - r_{ik}\right|}{\max\{r_i\} - \min\{r_i\}} \tag{9}$$

where r_{ij} are the values of the BB ordinal scale of Table **1** after conversion to ranks (ranging from 1 to 9).

$$\text{Podani discordance PD} = 1 - 2(a - b + c - d)/[n(n-1)] \tag{10}$$

This measure was proposed by Podani (1997) for all cases in which presence/absence is also meaningful, like phytosociological data with OTVs. In (Eq. 10) n, a and b are defined as above, while c is the number of pairs of species tied in both j and k, corresponding to joint presence or joint absence, as in the examples given below in which the rows represent the species and the columns represent the plots:

$$\begin{matrix} 1 & 1 & \textit{or} & 1 & 2 & \textit{or} & 0 & 0 \\ 1 & 1 & & 1 & 2 & & 0 & 0 \end{matrix}$$

That is, such pairs of species increase the similarity (decrease the dissimilarity) of the objects. In turn, d is the number of all pairs of species that are tied at least for one of the plots being compared so that either one, two or three scores are zero. The following examples will clarify this point:

$$\begin{matrix} 1 & 0 & \textit{or} & 1 & 1 & \textit{or} & 1 & 0 & \textit{or} & 0 & 1 \\ 0 & 0 & & 1 & 0 & & 1 & 0 & & 0 & 3 \end{matrix}$$

These pairs of variables indicate contradiction of the plots at least in presence/absence relations and will contribute to increased dissimilarity.

For comparing the reference matrices with the corresponding matrices obtained after transforming the point-quadrat cover to OTVs we performed a Mantel test (Mantel 1967). The application of Mantel's test in community ecology has a long history and was used mainly for the analysis of correlation between environmental and biological data (Wildi 1992; Feoli and Orloci 2011). It is calculated as the Pearson correlation coefficient between the pairwise elements of two matrices. As these elements are not independent, the significance of the Mantel correlation is usually tested by randomly permuting the order of the elements within one matrix keeping the other matrix unchanged. P-values are then computed as the proportion of permutation-derived values that are as extreme or more extreme than the actual matrix correlation coefficient. In this paper, the Mantel test was calculated with the program MATEDIT (Burba et al. 2008) fixing the number of randomizations at 9,999.

Results

For all sets of plots (A–C) the results of the Mantel correlations between the reference matrices and the dissimilarity matrices computed from the OTVs are shown in Tables **3**, **4** and **5**. In the first column of Tables **3**, **4** and **5** we compare resemblance matrices that are obtained with the same ratio-scale dissimilarity coefficient, but different data (transformed *vs* untransformed), thus showing the effects of ordinal data transformation on the structure of matrix dissimilarity. By contrast, in columns 2–5, we compare resemblance matrices that are obtained with different dissimilarity coefficients and different data, thus showing the combined effects of using OTVs and ordinal dissimilarity coefficients at the same time. Note that in this paper we did not

Table 3 Coefficients of linear matrix correlation (R) between each of the 6 dissimilarity matrices computed from the point-quadrat cover data of the species-poor garrigue south of Siena (data set A) and the dissimilarity matrices obtained from applying the same ratio-scale dissimilarity of Column 2 together with the four ordinal dissimilarity measures to the corresponding ordinal transform values (OTVs). All coefficients are significant at $P<0.05$ (two-tailed test)

		BB-type data (OTVs)					
		Ratio-scale dissimilarity	τ	γ	RD	PD	Mean
Point-quadrat cover data	ED	0.566	0.247	0.273	0.274	0.171	0.306
	CB	0.696	0.425	0.471	0.508	0.328	0.486
	CH	0.812	0.508	0.510	0.384	0.367	0.516
	BC	0.829	0.621	0.609	0.493	0.488	0.608
	CM	0.963	0.810	0.807	0.927	0.884	0.878
	MS	0.826	0.618	0.604	0.488	0.495	0.606
	Mean	0.782	0.538	0.546	0.512	0.456	0.567

address the problem of comparing resemblance matrices computed from OTVs with different ratio-scale and/or ordinal dissimilarity coefficients. For instance, all these matrices are basically surrogates for resemblance matrices calculated from non-transformed percentage cover data, and none of them can be assumed as a reliable reference for the actual dissimilarity patterns among plots.

As expected, due to the monotonic relationship between the non-transformed cover values (Cv) and the corresponding OTVs (i.e., $\mathrm{OTV} \approx 1.415 \ln \mathrm{Cv} + 2$; van der Maarel 2007), the dissimilarity matrices obtained applying the ratio-scale coefficients to the OTVs are highly correlated with the reference matrices that are calculated with the same ratio-scale coefficient on the original cover data. In contrast, the values of matrix correlation between the reference matrices and the matrices calculated with the ordinal coefficients on the OTVs are generally

Table 4 Coefficients of linear matrix correlation (R) between each of the 6 dissimilarity matrices computed from the point-quadrat cover data of the Mt. Labbro data (data set B) and the dissimilarity matrices obtained from applying the same ratio-scale dissimilarity of Column 2 together with the four ordinal dissimilarity measures to the corresponding ordinal transform values (OTVs). n.s. – not significant at $P<0.05$ (two-tailed test)

		BB-type data (OTVs)					
		Ratio-scale dissimilarity	τ	γ	RD	PD	Mean
Point-quadrat cover data	ED	0.232	0.474	0.370	$-0.172^{\text{n.s.}}$	$0.003^{\text{n.s.}}$	0.181
	CB	0.381	0.612	0.497	0.190	0.319	0.400
	CH	0.823	0.579	0.507	0.314	0.401	0.525
	BC	0.880	0.714	0.603	0.331	0.457	0.597
	CM	0.989	0.442	0.403	0.984	0.942	0.752
	MS	0.836	0.642	0.542	0.372	0.471	0.573
	Mean	0.690	0.577	0.487	0.337	0.432	0.505

Table 5 Coefficients of linear matrix correlation (R) between each of the 6 dissimilarity matrices computed from the point-quadrat cover data of the Lucciola Bella data (data set C) and the dissimilarity matrices obtained from applying the same ratio-scale dissimilarity of Column 2 together with the four ordinal dissimilarity measures to the corresponding ordinal transform values (OTVs). n.s. – not significant at $P<0.05$ (two-tailed test)

		BB-type data (OTVs)					
		Ratio-scale dissimilarity	τ	γ	RD	PD	Mean
Point-quadrat cover data	ED	0.382	0.499	0.533	$-0.073^{\text{n.s.}}$	$-0.048^{\text{n.s.}}$	0.259
	CB	0.578	0.585	0.482	0.408	0.425	0.496
	CH	0.817	0.634	0.576	$0.095^{\text{n.s.}}$	0.157	0.456
	BC	0.876	0.740	0.668	$0.101^{\text{n.s.}}$	0.171	0.511
	CM	0.997	0.306	$0.073^{\text{n.s.}}$	0.992	0.986	0.671
	MS	0.854	0.707	0.619	0.121	0.186	0.497
	Mean	0.751	0.579	0.492	0.274	0.313	0.482

lower, showing that in most cases the dissimilarity measures tailored for OTVs are less adequate for preserving the actual dissimilarity structure of the three sets of plots.

Looking at the average correlation values between the resemblance matrices obtained with the point-quadrat data and the OTVs for the three data sets, we observe a regular decrease from data set A to data set C as a function of the decreasing occupancy rate of the matrices (i.e., of increasing sparseness; see Table 2). For instance, the average value of the matrix correlation coefficient $\overline{R}$ equals 0.567 for matrix A, 0.505 for matrix B and 0.482 for matrix C, meaning that differences among the dissimilarity matrices calculated with different dissimilarity coefficients may increase as a function of the floristic heterogeneity of the data set.

The dissimilarity matrices calculated with the Canberra metric (CM) on the point-quadrat data show the highest correlation values with the dissimilarity matrices calculated with OTVs. This is because the CM gives high weight to the species presences irrespective of their cover values. This is evident when we consider the average correlation of the CM matrices computed from the original data with the matrices calculated with coefficients tailored for OTVs and the mean species cover values in the three data sets. The average correlation for the first set is 0.878, while mean species cover is 1.07. For the second set average correlation is 0.752 and mean species cover is 4.09. Finally, for the third set average correlation is 0.671 and mean species cover is 8.10. Not surprisingly, the lowest correlation values are found for the Euclidean distance (ED) because ED is particularly critical for sparse matrices and with uneven species cover (Orloci 1978; Legendre and Legendre 1998); in fact, it may happen that plots with no species in common may have a lower dissimilarity than plots with several species in common. Low correlation values are also found for the Manhattan distance (CB) that usually yields results that are similar to those of the Euclidean distance; it is only different in that it is less sensitive to large differences in species cover.

Discussion

In this paper we explored the effects of ordinal data transformation and the performance of metric *vs* ordinal dissimilarity measures in preserving the dissimilarity structure of three sets of vegetation plots that we considered representative of three different vegetation systems in terms of floristic homogeneity and rate of matrix occupancy (see Table **2**). Overall, whenever an accurate estimation of plant species abundances is considered important for our conclusions, it is prudent to measure abundance as accurately as possible and to manipulate it with adequate dissimilarity coefficients. Nonetheless, in spite of the loss of importance of the more abundant species due to the ordinal data transformation, the dissimilarities obtained from applying the ratio-scale coefficients CH, BC, CM and MS to the OTVs are closely related to the dissimilarities calculated with the same coefficients on the original cover data. In this view, van der Maarel (2007) notes that in spite of being of ordinal nature, the OTVs come close enough to a metric scale, and that the results of the usual ratio-scale multivariate methods can be accepted as a meaningful approximation of reality.

To the contrary the dissimilarity matrices calculated with ordinal coefficients on OTVs are generally less correlated to the reference matrices. This outcome is not surprising, as these matrices have been obtained using different dissimilarity coefficients and a different data scale with respect to the reference matrices. Therefore, in spite of their mathematical admissibility, dissimilarity matrices calculated with ordinal coefficients on OTVs were less informative on the reference dissimilarity structure than less admissible dissimilarity matrices calculated with ratio-scale coefficients on OTVs.

Of course, the question of whether a result is meaningful for the scientist is by far not limited to OTVs. For instance, Podani (2006) notes that "mathematical admissibility, meaningfulness of dissimilarities, consistency and universality" should prevail over the human factor in any scientific discipline. However, as already said, the use of dissimilarity coefficients based on the arithmetical operations of sum, difference product and division should not be considered unsuitable for BB-type data. To justify this viewpoint, Ricotta and Avena (2006) proposed the following staircase metaphor: imagine a building with a staircase made of steps of different height. If we ask: "How many steps should I climb to reach the third floor?" we assume that the number of steps is the relevant piece of information, irrespective of their height. As a result, if we focus our attention on the number of steps, the difference between 1 step and 2 steps is equal to the difference between 3 steps and 4 steps, 1 step + 2 steps is equal to 3 steps, and 1/2 is the same as 2/4.

From a historical perspective, it was Stevens (1946) that first presented a hierarchy of data scales (nominal, ordinal, interval and ratio scale) based on invariance of their meaning under different classes of transformations. Five years later, Stevens (1951) went a step forward by associating specific statistical procedures to the measurement scales for which they were admissible. This apparently simple and uncontroversial taxonomy of measurement scales and statistical procedures was quickly adopted in measurement theory to derive an axiomatic approach to measurement problems (Luce et al. 1990; Hand 1996). According to this axiomatic approach, if we assume that *i*) we know what attributes to measure, *ii*) we know what questions we shall ask about these attributes, and *iii*) we have assigned numbers that preserve the relevant features and relationships of these attributes, then we can derive axiomatic results from such assumptions that tell us the scale type to be used together with the statistical procedures that are appropriate for the selected scale type (Velleman and Wilkinson 1993; Hand 1996; Podani 2007).

Nonetheless, as observed by Velleman and Wilkinson (1993), even if we restrict our perspective to the above axiomatic approach without considering different theories of measurement (see Hand 1996), the security that it implies is illusory because the underlying assumptions are usually false. An example may help clarifying this point. Contrary to Stevens' approach, Lord (1953) argues that the choice of a permissible statistical method depends on the question it is designed to answer, and not on the measurement scale. Accordingly, he imagined that a professor who owned the university football jersey number concession was accused to assign unusually low numbers to the freshman class. Here, Velleman and Wilkinson (1993) note: "Although Lord's professor protests that football numbers are only nominal-scale values, the statistician he consults is happy to add them up, square them, compute their mean, and perform other operations needed for the application of Tchebycheff's inequality (avoiding reference to normality) to test the accusation that the numbers were too low". Likewise, if we analyze OTVs with multivariate procedures applicable to ratio scales only, we are basically asking how many ranks are there between, say, OTV 2 and 5, implicitly assuming that the number of ranks is the relevant piece of information. Whether this question is meaningful or not depends more on what kind of attributes researchers wish to highlight with plot data, rather than on the original representation scale used for measuring species cover. To conclude, we fully agree with Podani (2007) that the results of any analysis cannot be better than the quality of the input data, no matter how we manipulate them. Nonetheless, paraphrasing Paul Velleman (in Hand 1996, p. 486): "[plant ecologists] may prefer the freedom to do something 'impermissible' and to judge a posteriori whether it was warranted".

References

Anderberg MR (1973) *Cluster analysis for applications*. Academic Press, New York

Braun-Blanquet J (1932) *Plant sociology: the study of plant communities*. McGraw-Hill, New York

Burba N, Feoli E, Malaroda M (2008) MATEDIT: a software tool to integrate information in decision making processes. In Neves R, Baretta JW, Mateus M (eds) *Perspectives on integrated coastal zone management in South America*. IST PRESS, Lisboa, pp 123–127

Chiarucci A, Maccherini S, Bonini I, De Dominicis V (1998) Effects of nutrient addition on species diversity and ground cover of "serpentine" vegetation. *Pl Biosyst* 132:143–150

Dale MB (1989) Dissimilarity for partially ranked data and its application to cover-abundance data. *Vegetatio* 82:1–11

Dengler J, Ewald J, Kuhn I, Peet R (2011) Ecoinformatics and global change – an overdue liason. *J Veg Sci* 22:577–581

Hand DJ (1996) Statistics and the theory of measurement. *J Roy Stat Soc A* 159:445–492

Faith D, Minchin PR, Belbin R (1987) Compositional dissimilarity as a robust measure of ecological distance. *Vegetatio* 69:57–68

Feoli E, Ganis P, Venanzoni R, Zuccarello V (2011) Toward a framework of integrated knowledge of terrestrial vegetation system: The role of databases of phytosociological relevés. *Pl Biosyst* 145:74–84

Feoli E, Orloci L (1991) *Computer assisted vegetation analysis*. Kluwer Academic Publishers, Dordrecht

Feoli E, Orloci L (2011) Can similarity theory contribute to the development of a general theory of the plant community? *Community Ecol* 12:135–141

Legendre P, Legendre L (1998) *Numerical ecology.* Elsevier, Amsterdam

Lord FM (1953) On the statistical treatment of football numbers. *Amer Psychol* 8:750–751

Luce LD, Krantz DH, Suppes P, Tversky A (1990) *Foundations of measurement*, Vol 3. Academic Press, New York

Maccherini S, Marignani M, Castagnini P, van den Brink PJ (2007) Multivariate analysis of the response of overgrown semi-natural calcareous grasslands to restorative shrub cutting. *Basic Appl Ecol* 8:332–342

Mantel N (1967) The detection of disease clustering and a generalized regression approach. *Cancer Res* 27:209–220

Marignani M, Del Vico E, Maccherini S (2007) Spatial scale and sampling size affect the concordance between remotely sensed information and plant community discrimination in restoration monitoring. *Biodivers & Conservation* 16:3851–3861

Moore PD, Chapman SB (1986) *Methods in plant ecology.* Blackwell Scientific Publications, Oxford

Orloci L (1978) *Multivariate analysis in vegetation research*. Dr W Junk Publishers, The Hague

Podani J (1997) A measure of discordance for partially ranked data when presence/absence is also meaningful. *Coenoses* 12:127–130

Podani J (2000) *Introduction to the exploration of multivariate biological data*. Backhuys, Leiden

Podani J (2005) Multivariate exploratory analysis of ordinal data in ecology: Pitfalls, problems and solutions. *J Veg Sci* 16:497–510

Podani J (2006) Braun-Blanquet's legacy and data analysis in vegetation science. *J Veg Sci* 17:113–117

Podani J (2007) Spatial confusion or clarity? Reply to Ricotta and Avena. *J Veg Sci* 18:921–924

Ricotta C, Avena G (2006) On the evaluation of ordinal data with conventional multivariate procedures. *J Veg Sci* 17:839–842

Ricotta C, Moretti M (2010) Assessing the functional turnover of species assemblages with tailored dissimilarity matrices. *Oikos* 119:1089–1098

Schaminée JHJ, Janssen JAM, Hennekens SM, Ozinga WA (2011) Large vegetation databases and information systems: New instruments for ecological research, nature conservation and policy making. *Pl Biosyst* 145:85–90

Samaritani E, Shrestha J, Fournier B, Frossard E, Gillet F, Guenat C, Niklaus PA, Tockner K, Mitchell EAD, Luster J (2011) Heterogeneity of soil carbon pools and fluxes in a channelized and a restored floodplain section (Thur River, Switzerland). *Hydrol Earth Syst Sci* 8:1059–1091

Stevens SS (1946) On the theory of scales of measurement. *Science* 103:677–680

Stevens SS (1951) Mathematics, measurement and psychophysics. In Stevens SS (ed) *Handbook of experimental psychology.* Wiley, New York, pp 1–49

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

van der Maarel E (2007) Transformation of cover-abundance values for appropriate numerical treatment – Alternatives to the proposals by Podani. *J Veg Sci* 18:767–770

Velleman PF, Wilkinson L (1993) Nominal, ordinal, interval, and ratio typologies are misleading. *Amer Statist* 47:65–72

Wildi O (1992) On the use of Mantel's statistics and flexible shortest path adjustement in the analysis of ecological gradients. *Coenoses* 7:91–101

Received: 19 December 2011 / Revised: 4 October 2012 / Accepted: 8 October 2012 /
Published online: 15 February 2013