Diversity and Distributions, (Diversity Distrib.) (2007) **13**, 361–371

BIODIVERSITY RESEARCH

Reconstructing community relationships: the impact of sampling error, ordination approach, and gradient length

Claire N. Hirst and Donald A. Jackson*

Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, Ontario, Canada

*Correspondence: Donald A. Jackson, Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, ON, Canada M5S 3G5. Tel.: (416) 978-0976; E-mail: jackson@zoo.utoronto.ca

ABSTRACT

Effectively summarizing complex community relationships is an important feature in studies such as biodiversity, global change, and invasion ecology. The reliability of such community summaries depends on the degree of sampling variability that is present in the data, the structure of the data, and the choice of ordination method, but the relative importance of these factors is not understood. We compared the validity of results from different ordination methods by applying five levels of sampling error to a simulated coenoplane model at two gradient lengths using two types of data (abundance and presence–absence). The multivariate methods we compared were correspondence analysis (CA), detrended correspondence analysis (DCA), non-metric multidimensional scaling (NMDS), principal component analysis (PCA) and principal coordinates analysis (PCoA). Our results showed CA and PCA using presence–absence data were the most successful methods regardless of sampling error and gradient length, closely followed by the other methods using presence–absence data. With abundance data, PCA and CA were the most successful approaches with the short and long gradients, respectively. Approaches based on PCoA and NMDS using abundance data did not perform well regardless of the choice of distance measure used in the analysis. Both of these methods, along with the PCA using abundance data, were strongly affected by the longer gradient, leading to more distorted results.

Keywords

Community ecology, gradient length, multivariate statistics, ordination, sampling error.

INTRODUCTION

Communities are complex systems that are not easily described qualitatively or quantitatively — there is a broad array of biotic, abiotic, and historical factors that can structure an ecological community (Jackson *et al.*, 2001). Accordingly, ecologists use multivariate ordination methods to create databased, quantitative summaries of the large complex patterns of species distribution and abundance within communities, and to connect these patterns to environmental conditions (Zimmer *et al.*, 2000; Kenkel *et al.*, 2002; Bowman *et al.*, 2006). In addition to providing basic understanding about ecological systems, these summaries are valuable because detailed knowledge of community structure is necessary to restore the integrity and diversity of a community that has been damaged by human interference and to monitor changes in a given community over time (e.g. Kremen, 1992).

The structure of a community is affected by how many species are present, their relative abundance, and how broadly each component species is distributed along environmental gradients. These differences in structure among sites provide the basic information used in community analysis. The reliability of the results from ordination and clustering methods can be greatly affected by the properties of the observations collected (Gauch *et al.*, 1977; Oksanen & Minchin, 1997). Differences in species composition among sampling locations are quantified by various measures of gradient length. Increases in gradient length cause multivariate ordination methods to perform poorly and potentially to produce results that are, at best, difficult to interpret with very long gradients (De'ath, 1999; Tamas *et al.*, 2001). As a result, ecologists using multivariate summaries of community data need to consider how gradient length affects the usefulness of a particular ordination method.

Most ecological studies rely on data sampled from the field to draw some conclusion; this means that sample data must be both meaningful and accurate. One substantial concern is the type of data that is collected; accurate abundance data are more informative than presence–absence data, but if sample data are distorted by sampling biases or errors, this more detailed level of measurement may be misleading rather than informative (Jackson & Harvey, 1997). There are a variety of problems inherent to sampling

DOI: 10.1111/j.1472-4642.2007.00307.x
www.blackwellpublishing.com/ddi

ecological populations and communities. One of these problems is rarity. Rare species are those that occur at low prevalence or in low numbers, or both, making them difficult to sample accurately, if at all. Often, rare species add 'noise' to community summaries, and as a result, ecologists have debated extensively on whether or not rare species comprise valuable ecological information, and whether it is justifiable to remove rare species from a data set to obtain more representative results about the general community patterns (Courtemanch, 1996; Brazner & Beals, 1997; Resh *et al.*, 2005). The issue is made more difficult because the particular definition of rarity that a researcher employs is arbitrary. For plant communities, the concept of rarity is obscured by additional problems (McCoy & Mushinsky, 1992), including at what point clonal plants should be counted as individuals (similar problems exist for colonial animals such as coral). Other problems include differences arising from measuring the number of individuals vs. biomass, and for animals, how the differences among life stages should be considered. The relevance of rare species to a study will also depend on the researcher's objectives; research aimed at conserving rare species must first be able to detect them (Green & Young, 1993).

Another sampling problem is bias resulting from the gear type that is used in sampling, or the habitat or season in which sampling is done. For example, sampling methodology often differs among studies of benthic invertebrates, and biased results may arise if a researcher samples only a single habitat (Bradley & Ormerod, 2002) or uses a particular gear (Kerans *et al.*, 1992). The accuracy and precision of sample data may be improved by repeated sampling but this is time-consuming and, because many different factors can induce noise — for example, sampling limitations, species vagility, and the researcher's own bias — variability among replicates can be quite high (Gauch, 1982). The effect of these various sampling problems is that it is frequently difficult to establish a reliable estimate of abundance, and that all sample data include various sources of error (Ostermiller & Hawkins, 2004).

We are therefore confronted with various sources of error and complications in being able to adequately summarize patterns of species composition and resemblance of sampling locations to one another. We have the problem associated with: (1) how representative our samples are (i.e. errors associated in the sampling of a particular device); (2) choices in the type of sampling methodology or device which favours the sampling of particular species in some habitats relative other species or habitats; (3) intrinsic variability in each species distribution; and finally (4) methods of data analysis that affect our ability to effectively summarize patterns in species composition and site resemblance. Our study focuses primarily on the combination 2–3 through our ability to adequately assess the 'actual' abundance of each species within a community across a series of site and the interaction of this combination with (4) given the numerous choices in resemblance measure and ordination technique available to researchers.

Many ecologists use multivariate ordination methods to summarize sample data and to describe patterns of species distribution. The best known of these methods is principal component analysis (PCA), which summarizes major linear patterns of covariation into a few axes. PCA provides a low-dimensional summary of high-dimensional data through the use of a covariance or correlation matrix to summarize patterns of covariation among variables. It provides an effective method when there are linear relationships between variables (see Peres-Neto *et al.*, 2003). PCA is essentially a specialized case of another technique, principal coordinates analysis (PCoA), which is similar to PCA in its goal but permits the use of a greater breadth of resemblance measures (see Jackson *et al.*, 1989; Legendre & Legendre, 1998). Such choice provides a careful user the option of selecting a measure that emphasizes data attributes that are of interest (e.g. relative species abundance vs. absolute abundance). Correspondence analysis (CA) doubly standardizes data by row and column totals and calculates the degree of association based on the chi-squared distance. While many of the other methods work with specific types of data, or employ different measures of resemblance depending on the type of data involved, CA has been shown to be effective with data ranging from presence–absence through abundance and has also been shown to be effective with compositional (i.e. proportional or percentage) data that lead to difficulties with many of the other approaches (Jackson, 1997). PCA, PCoA, and CA are all vulnerable to a particular mathematical artefact, in which the initial ordination axes may be distorted into arches or horseshoes (ter Braak & Prentice, 1988), particularly when species turnover is high. Detrended correspondence analysis (DCA) was developed as an attempt to correct this arch effect by 'detrending': dividing the arch into segments and then centring each segment on the second axis by subtracting the mean. However, the form of detrending and number of segments chosen can lead to very different solutions (Jackson & Somers, 1991; Legendre & Legendre, 1998). Non-metric multidimensional scaling (NMDS) is a non-parametric approach that creates a graphical summary (e.g. a two- or three-dimensional graphical solution) of the original relationships in the data by systematically rearranging the distance matrix from this plot until the distances between its elements are ranked as similarly as possible to the distance matrix based on the original species data. The number of rearrangements (iterations) that is made is an arbitrary choice based on a variety of criteria, and the initial configuration of the artificial matrix is known to affect the final result frequently (Fasham, 1977) and is either based on using a random initial configuration or on a configuration based on some of the results from some other ordination methods, e.g. PCoA. Readers interested in additional details on these methods are referred to Jongman *et al.* (1987), Legendre & Legendre (1998), Manly (2005), or other recent monographs on community analysis and multivariate statistical methods.

Combining these ideas, it appears that the usefulness of a specific community analysis and its summary will depend on the reliability of the data that were used, the structure of the data, and the choice of ordination method. While the impact of gradient length has been examined, it is not clear how the combination of changes in sampling error and gradient length will affect different multivariate ordination methods. In this study, we conducted a comparative analysis to examine how changes in gradient length and sampling error influence the reliability of ordination methods in correctly summarizing the general relationships. The specific

objectives were (1) to visually and quantitatively compare how well different ordination methods recreated a simulated community at different levels of sampling variability along relatively long and short gradients, and (2) to assess how sensitive each ordination method was to increasing sampling error.

Simulated data sets have been used previously to quantitatively and qualitatively compare ordination methods (e.g. Karadzic & Popovic, 1993; McCune, 1997). While simulated models have certain drawbacks, and may not precisely reflect ecological reality (Jackson, 1993a; Oksanen & Minchin, 2002), they are nonetheless a useful tool that facilitates an objective evaluation of the effects of sampling error on resulting ordination and community interpretation, a topic that has not been examined in detail before.

METHODS

We generated an initial data set in COMPAS by selecting the default values that were provided by this program (see Minchin, 1987 for details related to COMPAS), so that the species response functions had log-random distributions of modal abundance and uniform random distributions of modal coordinates on each gradient. Our data set included randomly generated beta response curves, and abundance values for 40 species across a two-dimensional space, which simulated the plane of available habitat. A 10×10 grid was used to draw sets of species from the entire environmental space, and in total there were 100 sets of species (Fig. 1). This first grid was called *full gradient.* The rate of turnover for this grid was 3.10×2.72 half-change units. Note that the specifications for this data set were selected after McCune (1994).

A second 10×10 grid of 100 sets of species was constructed from the central region of the first grid. This second grid was called *half gradient*. The sampling intensity on the half-gradient grid was effectively doubled because it was half the size of the full-gradient grid but still included 100 observations; that is, the first grid had 10 sampling points per gradient, evenly placed at positions between 0 and 100 and the second had grid had 10 sampling points per gradient, evenly placed between 38.88 and 88.89. The abundance values for each species on the two gradients were calculated by interpolation using the formula from Minchin (1987). The parameters required by this formula were provided by COMPAS for the full-gradient grid, and we reapplied these same values for the sampling positions on the half-gradient grid. Only 24 species were present in the half-gradient grid; the remaining 16 species had distributions that were too peripheral to be included in the half gradient. All of the treatments that follow within the Methods section were applied to both grids (see Table 1 for a summary of treatments).

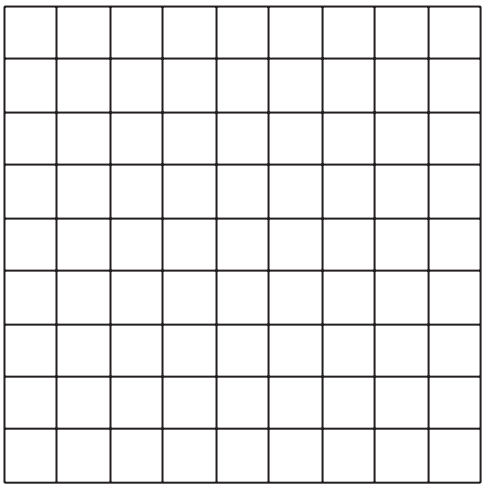

Figure 1 The original coenoplane model, before ordination and application of sampling variability. The sampling points used for comparisons are based on the intersection points of lines.

In order to examine the effect of sampling variability, sample errors were added to species' abundances as random, normally distributed departures from the original values. The error levels we used were 5%, 10%, 25%, and 50%. The errors were calculated as a number drawn from a normal distribution $N(0,s)$, where s is the standard deviation for a given species across all sampling points multiplied by the particular level of error. We added the random error to each species' original abundance value at each sampling point to generate separate data matrices for species' abundances along the grid at five levels of sampling variability (i.e. 0%, 5%, 10%, 25%, and 50%). Negative abundance values that were derived were replaced with zeros. The estimation of the percentage of error added to the data was calculated on the data prior to the replacement of negative values with zero values, which accurately reflects the patterns of abundance one would encounter, but error values do not map directly to the data sets after replacements with zero values.

For each data matrix, five ordination methods (and some variations on certain methods; see Table 1) were applied to test how effectively each technique would summarize the pattern in the original data (i.e. either the full or half gradient 10×10 grids). The ordinations, with their respective measures of resemblance,

Table 1 Description of the ordination methods. Each method, with its variations, was carried out on full and half gradient grids, at error levels of 0%, 5%, 10%, 25%, and 50%

Method	Data type	Resemblance measure selected	Initial configuration
Correspondence analysis	Abundance, presence–absence	N/A	N/A
Detrended correspondence analysis	Abundance, presence–absence	N/A	N/A
Non-metric multidimensional scaling	Abundance, presence–absence	Chord distance, Euclidean distance	Random, PCoA solution
Principal component analysis	Abundance, presence–absence	Correlation	N/A
Principal coordinates analysis	Abundance, presence–absence	Chord distance, Euclidean distance	N/A

were as follows: PCA = principal component analysis using a correlation matrix, $PCoA_C$ and $PCoA_E$ = principal coordinates analysis using chord distance and Euclidean distance, CA = correspondence analysis using abundance data, DCA = detrended correspondence analysis, and $NMDS_C$ and $NMDS_E$, which were based on selecting the best solution obtained from 10 random starting configurations. Given literature debates regarding the relative merits of random vs. metric solutions as starting configurations, we also used starting configurations based on a PCoA of Euclidean distance as it always provides lower stress levels than obtained from a starting configuration based on a PCoA of chord distances. These solutions are denoted as $NMDS_P$. We chose the default values, which included STRESS2 as the stress coefficient (see Rohlf, 1997), monotone regression, a minimum stress of 0.001, and a maximum ratio of stresses falling between 0.999 and 1. The ordination that provided the minimum level of stress was selected as the best representation of the method. Each method was conducted using both abundance data and presence–absence data; presence–absence data are indicated by the subscript 'pa', e.g. CA_{pa}.

We chose chord distance, also called stand normalization, because it has been shown to be a more effective standardization measure than double standardization or no standardization, for recreating a coenoplane model (Kenkel & Orlóci, 1986). We chose Euclidean distance as an alternate measure to contrast against chord distance given that the former emphasizes absolute abundances, whereas the latter emphasizes species relative abundances. All ordinations were performed using the NTSYSPC 02.02H statistical package (Rohlf, 1997), with the exception of DCA, which used CANOCO for Windows 4.02 (ter Braak, 1986).

To examine differences in the ordination techniques and the effects of different levels of sampling error, the 80 ordinations (5 matrices × 8 ordination methods per grid × 2 data types; see Table 1) were compared to the original grid, and pairwise to each other, using Procrustes analysis (Jackson, 1995; Peres-Neto & Jackson, 2001) which provides numerical value for distortion between a reference and a target matrix (the m^2 and residual sum of squared values). The m^2 values from the pairwise comparisons were ordinated by PCoA in NTSYS-PC, and a minimum spanning tree was superimposed in order to reveal relationships of relative similarity between the ordinations (Jackson, 1993b). We restricted the comparison to using two-dimensional solutions from each ordination given the two-dimensional nature of the coenoplane and graphical presentations of the results. However it is possible, and likely, that some relevant patterns in species composition are expressed on axes beyond the second axes for some of the methods considered (e.g. this is well known to contribute to the horseshoe phenomenon in PCA).

RESULTS

Effect of sampling error and gradient length on ordination

As sampling error increased, distortion as measured by m^2 values from Procrustes analysis tended to be greater when error was added, but random in the pattern of whether it increased or not relative to the degree of error added (Table 2), or alternatively the distortion increased with increasing error (Table 3). The original grid of sampling points on the coenoplane is a two-dimensional grid (Fig. 1). Figures 2–5 permit visual comparison of the coenoplane model with the ordination results from different methods at 0%, 10%, and 25% sampling error, for the full gradient and half gradient using abundance and presence–absence data, respectively. For the full gradient, at all levels of sampling and for both types of data, CA_{pa} consistently returned the lowest m^2 values out of the ordination methods examined, and provided the least distorted multivariate summary of the original sampling grid. Overall, CA_{pa} and other methods based on presence–absence data, followed by CA and DCA based on abundance data, were the best methods for reproducing the original grid structure: these methods returned small m^2 values and were robust to sampling error within the range of 0% to 50% error. Of all the methods we examined, the $PCoA_E$ and the $NMDS_P$ returned the most distorted results at the full-gradient length, closely followed by the other PCoA and NMDS solutions, with the exception of $NMDS_E$.

For the half-gradient grid at all levels of error, PCA_{pa} and the other ordinations based on presence–absence data, followed by PCA based on abundance data, returned the least distorted results of all the methods. The m^2 values from these methods

Table 2 Procrustean m^2 values for the eight ordination methods compared to the full-gradient grids, at five levels of sampling error (0%, 5%, 10%, 25%, and 50%), using both abundance and presence–absence data. High m^2 values indicate greater distortion (maximum = 1). m^2 values are derived from Procrustean comparisons of the first two axes from each ordination solution with the original grids

	Level of sampling error				
	0%	5%	10%	25%	50%
m^2 values — abundance data					
CA	0.159	0.169	0.146	0.155	0.151
DCA	0.161	0.149	0.161	0.140	0.137
$NMDS_C$	0.768	0.783	0.790	0.444	0.799
$NMDS_E$	0.454	0.431	0.456	0.448	0.611
$NMDS_P$	0.793	0.794	0.792	0.793	0.819
PCA	0.478	0.486	0.481	0.447	0.488
$PCoA_C$	0.720	0.736	0.796	0.799	0.740
$PCoA_E$	0.797	0.799	0.798	0.818	0.819
m^2 values — presence–absence data					
CA_{pa}	0.045	0.049	0.046	0.043	0.051
DCA_{pa}	0.058	0.050	0.051	0.044	0.048
$NMDS_{Cpa}$	0.087	0.088	0.093	0.089	0.089
$NMDS_{Epa}$	0.087	0.088	0.094	0.094	0.093
$NMDS_{Ppa}$	0.087	0.088	0.094	0.094	0.093
PCA_{pa}	0.064	0.073	0.077	0.093	0.092
$PCoA_{Cpa}$	0.080	0.087	0.096	0.103	0.104
$PCoA_{Epa}$	0.087	0.097	0.111	0.128	0.128

Table 3 Procrustean m^2 values for the eight ordination methods compared to the half-gradient grids, at five levels of sampling error (0%, 5%, 10%, 25%, and 50%), using both abundance and presence–absence data. High m^2 values indicate greater distortion (maximum = 1). m^2 values are derived from Procrustean comparisons of the first two axes from each ordination solution with the original grids

	Level of sampling error				
	0%	5%	10%	25%	50%
m^2 values — abundance data					
CA	0.464	0.463	0.468	0.479	0.456
DCA	0.273	0.271	0.318	0.364	0.290
$NMDS_C$	0.195	0.198	0.205	0.198	0.247
$NMDS_E$	0.251	0.250	0.255	0.277	0.290
$NMDS_P$	0.254	0.255	0.263	0.284	0.354
PCA	0.140	0.138	0.143	0.151	0.147
$PCoA_C$	0.320	0.319	0.345	0.346	0.338
$PCoA_E$	0.464	0.463	0.467	0.471	0.474
m^2 values — presence–absence data					
CA_{pa}	0.095	0.102	0.097	0.110	0.138
DCA_{pa}	0.079	0.084	0.084	0.101	0.112
$NMDS_{Cpa}$	0.108	0.113	0.130	0.145	0.140
$NMDS_{Epa}$	0.097	0.105	0.118	0.131	0.127
$NMDS_{Ppa}$	0.097	0.105	0.119	0.131	0.127
PCA_{pa}	0.077	0.085	0.093	0.098	0.110
$PCoA_{Cpa}$	0.095	0.109	0.120	0.140	0.139
$PCoA_{Epa}$	0.079	0.093	0.104	0.121	0.114

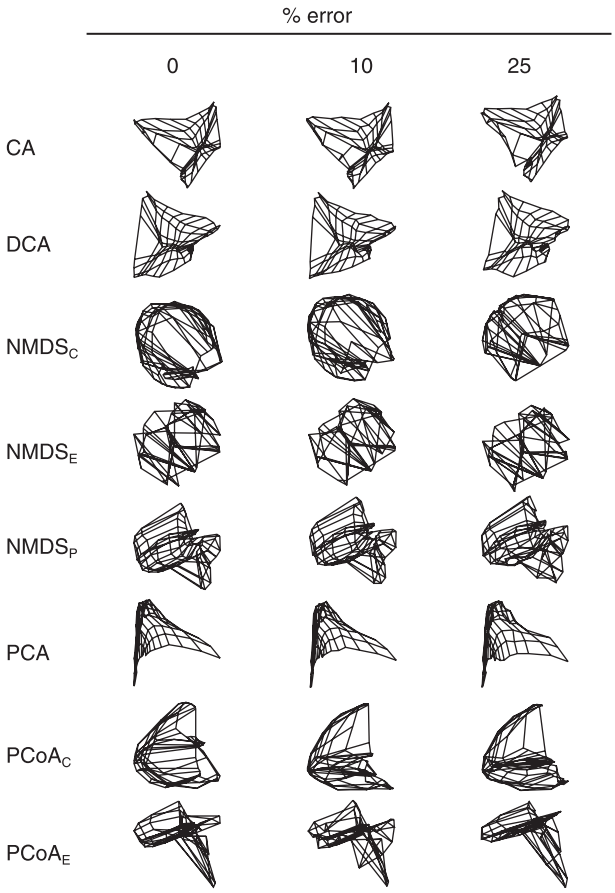

Figure 2 Graphical representations of ordination results for the full gradient grid at 0%, 10%, and 25% error, using abundance data. Lines represent the same lines as in coenoplane grid (Fig. 1), but the degree of distortion relative to the regular grid provides a measure of the ability of the ordination method to recover the original data structure. CA, correspondence analysis; DCA, detrended correspondence analysis; NMDS, non-metric multidimensional scaling; PCA, principal component analysis; PCoA, principal coordinates analysis.

were quite consistent at all levels of sampling error up to and including 50% error. CA showed a greater degree of distortion with the half gradient than with the full gradient, and exhibited m^2 values of approximately 0.45 at all levels of sampling error with abundance data, typically much greater than those resulting from other methods. NMDS ordinations based on abundance data become much less distorted at the half-gradient length, and exhibited the lowest m^2 values of any method except those based on PCA. NMDS ordinations on the half-gradient grid were relatively robust to sampling errors ranging up to 50%. Note that NMDS solutions give similar m^2 values for random and PCoA-ordinated initial configurations, and that the m^2 values for $NMDS_{Epa}$ and $NMDS_{Ppa}$ are nearly identical.

Overall, the results could be placed into two groups based on the relative degree of distortion on the shorter gradient vs. the longer gradient where non-linearity is more likely to be a relevant factor. Using abundance data, the results based on the CA and DCA showed less distortion on the longer gradient than on the shorter gradient, whereas all other results showed greater distortion with the longer gradient (Tables 2 & 3). One general feature of both the short- and the long-gradient solutions was that the m^2 values were smaller for solutions based on species presence–absence data rather than the relative abundance data and this was a consistent feature across the different ordination solutions and levels of induced sampling error.

Relationships among ordination methods

Figure 6 summarizes the similarity of the ordination solutions for the full gradient at all levels of sampling error and for both abundance and presence–absence data. In Fig. 6, groupings were located based on similarity, using a PCoA of m^2 values from Procrustean pairwise comparisons of ordination scores and their positions on the original grid.

For the full-gradient graph (Fig. 6), there are three main observations to be made. First, ordinations based on abundance data tend to fall out according to method for all levels of sampling error. PCoA and NMDS ordinations tended to fall out together, and the particular groups that were formed ($PCoA_E$ and $NMDS_P$ at the left upper corner, $NMDS_C$ and $PCoA_C$ in the lower left corner) share common resemblance measures within each group in general. The exception is $NMDS_E$, which fell out on the lower right of the graph when all levels of error were considered (Fig. 6), likely because of its relatively smaller m^2 values with the

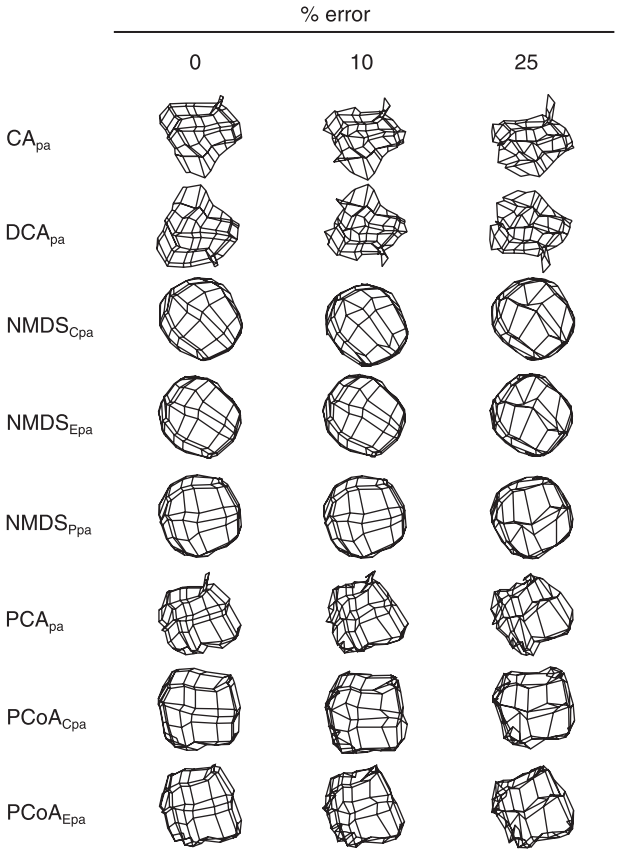

Figure 3 Graphical representations of ordination results for the full gradient grid at 0%, 10%, and 25% error, using presence–absence data. Lines represent the same lines as in coenoplane grid (Fig. 1), but the degree of distortion relative to the regular grid provides a measure of the ability of the ordination method to recover the original data structure.

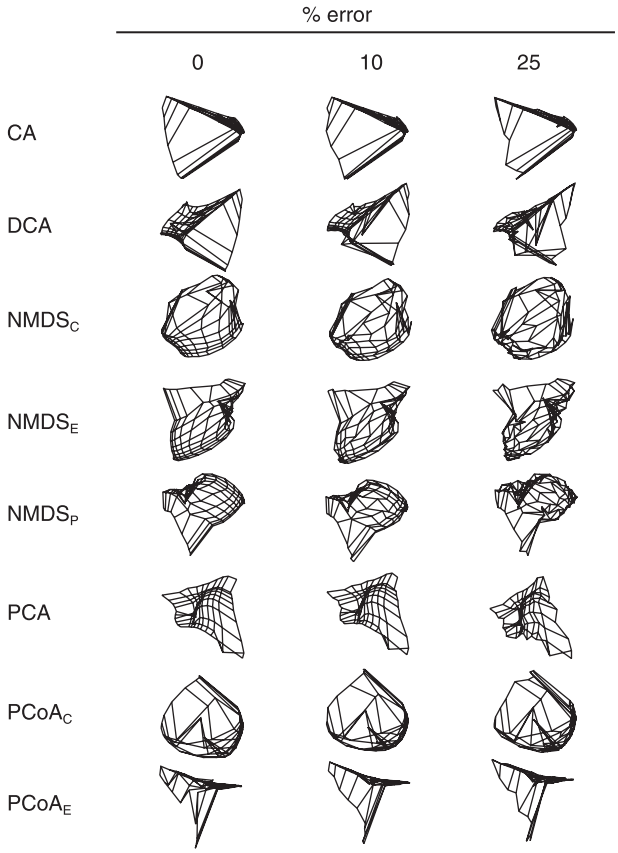

Figure 4 Graphical representations of ordination results for the half gradient grid at 0%, 10%, and 25% error, using abundance data. Lines represent the same lines as in coenoplane grid, but the degree of distortion relative to the regular grid provides a measure of the ability of the ordination method to recover the original data structure.

original grid compared to NMDS$_C$ and NMDS$_P$. From the ordinations based on abundance data, CA and the DCA cluster fell closest to the original grid. All PCA results are more closely related to those from the CA and DCA results than they are to the PCoA or NMDS solutions. Second, all ordinations based on presence–absence data are positioned closely to each other and to the original grid, and fall out according to method, as well as the level of error (see Fig. 6 inset for the more detailed relationships). Overall, the NMDS$_{pa}$ methods tend to fall out together and, as for abundance data, NMDSP$_{pa}$ and NMDS$_{Epa}$ are positioned particularly closely to each other. Out of the methods examined, CA$_{pa}$ and DCA$_{pa}$ are closest to the original grid. Lastly, it is useful to look at the eigenvalues for Figure 6: the total variation explained by the plot is high (> 85%) and the structure of the minimum spanning tree (MST) indicates that plot is an effective summary of the existing relationships.

For the half-gradient graph (Fig. 7), ordinations once again tended to fall out by method. DCA and CA are much more poorly related to the half-gradient grid than to the full-gradient grid (Fig. 6), and PCA is more closely related to the true arrangement of observations. The three variations of NMDS also fall out relatively closely on the second axis; NMDS$_E$ and NMDS$_P$ produced nearly identical ordinations at every level of sampling error, indicating that the choice of initial configuration has a much smaller effect than the choice of similarity measure. Once again, ordinations based on presence–absence data fall out much closer to the original grid than ordinations based on abundance data, and nearest-neighbour relationships fall out by the method as well as the level of error. The half-gradient grid is most closely related to PCA$_{pa}$. DCA$_{pa}$ and CA$_{pa}$ tend to fall out together across error levels and are reasonably close to the half-gradient grid as well; NMDS$_{pa}$ and PCoA$_{pa}$ are close to the original grid only at 0% error. The eigenvalues for Figure 7 are high (~84%), indicating that this figure explains a great deal of the total variation, and provides an effective overall summary of the existing relationships, but the minimum spanning tree shows that some of the multivariate similarity is not correctly displayed for the PCA solutions based on abundance (i.e. there is a crossing of the MST lines).

DISCUSSION

It is a frustrating necessity that researchers constructing ordinations based on species abundances (or any other analytical solution) must assume, with perhaps little justification or consideration,

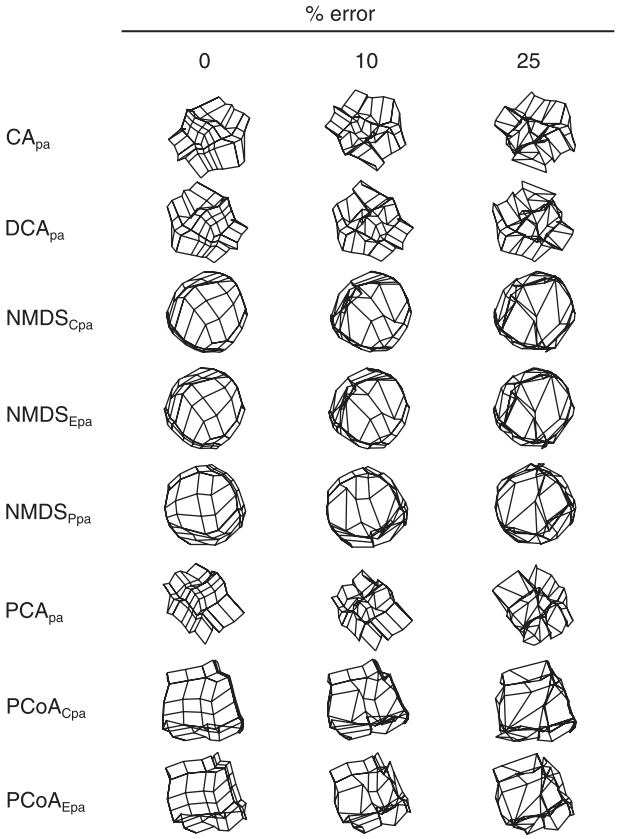

Figure 5 Graphical representations of ordination results for the half gradient grid at 0%, 10%, and 25% error, using presence–absence data. Lines represent the same lines as in coenoplane grid, but the degree of distortion relative to the regular grid provides a measure of the ability of the ordination method to recover the original data structure.

that their data correctly summarize ecological relationships. Sampling error from a variety of sources may skew the accuracy or precision of abundance values, and the degree of this sampling error will likely be unknown. To address this problem, researchers require ordination techniques that are robust to 'reasonable' amounts of sampling error present, or at least within a known range of error. Of the methods that we examined, we found that CA$_{pa}$, DCA$_{pa}$, and PCA$_{pa}$ returned the least distorted results at both gradient lengths, and at all levels of sampling error. With abundance data, results were specific to gradient length; for example, we found that PCA is a particularly good choice when a gradient is short and intensively sampled. This is because PCA assumes linear relationships among the variables, and for the short gradient, the sampling pattern is more likely to be linear (Ominski *et al.*, 1999). The use of PCA is often faulted for use with ecological data, which frequently contain non-linear relationships and many zeros (Kessel & Whittaker, 1976; Fasham, 1977; Legendre & Legendre, 1998). However, our results show that PCA may provide useful and informative results with a short gradient and therefore gradient length must also be a consideration in the analysis when abundance data are used. Given that the best results were obtained using presence–absence data, it is perhaps surprising that a measure, i.e. Pearson correlation coefficient, known to be vulnerable to non-linearities and zero values performed so well. We caution against the general acceptance of using presence–absence data in PCA given that there are logical reasons why one would not want to provide linear summaries of binary data, but clearly additional research in this result is warranted. One promising avenue of research relates to the use of the Hellinger transformation in PCA. Legendre & Gallagher (2001) found it to provide good reconstruction of distances in one-dimensional simulated communities and further research is warranted for community studies.

Using abundance data, CA and DCA returned the most accurate solutions for the long gradient when sampling error was up to 50% of the mean species abundance. The relatively improved performance of CA and relatively poor performance of PCA on the longer gradient make sense: CA has been found to outperform PCA at high levels of beta diversity (Gauch *et al.*, 1977; Jongman *et al.*, 1987; Legendre & Legendre, 1998) because it is more tolerant of non-linear changes in species abundances (Noy-Meir *et al.*, 1975). On the half gradient both CA and DCA of abundance data performed poorly, although DCA returned better results than CA, but analyses that used presence–absence data provided superior performance. The improvement of DCA over the CA is likely a combination of both rescaling and 'detrending' following segmentation, because the DCA grid may partly resolve an arch effect in the CA grid (see Fig. 4). It is also possible that the poorer performance of CA on the half gradient plot results from a particular feature of this method, in which rare species and sites with few species may be disproportionately emphasized. DCA performed well at both gradient lengths, and was fairly robust at a broad range of sampling errors. However, as there is little theoretical justification for using DCA (Legendre & Legendre, 1998), the choice of segment number in DCA is arbitrary (Digby & Kempton, 1987) and affects the stability of the ordination solution (Jackson & Somers, 1991), and the ordination results do not demonstrate reduced distortion when compared against other methods; DCA should be used with caution or avoided in ecological analyses. The variability in DCA solutions has not been considered here (although well documented in Jackson & Somers, 1991) and such variability in community ordinations simply further complicates the selection of a useful multivariate summary. As these problems are likely to apply to DCA based on presence–absence data as well, this method is not recommended. While many of these findings regarding DCA have been reported previously, inexplicably the method still remains frequently used.

With abundance data, PCoA performed badly overall. It is possible that Euclidean distance and chord distance were not relevant similarity measures for this data set; however, this calls into question the reliability of PCoA for general use given that these two metrics represent very different classes of resemblance measures, i.e. those based on measures incorporating information about species absolute abundance and those based on measures incorporating information regarding relative abundance. PCoA$_{C}$ produced less distorted results than PCoA$_{E}$ at both gradient lengths and all levels of sampling error. This result suggests that a researcher employing PCoA should select a resemblance measure with care (see Legendre & Gallagher, 2001).

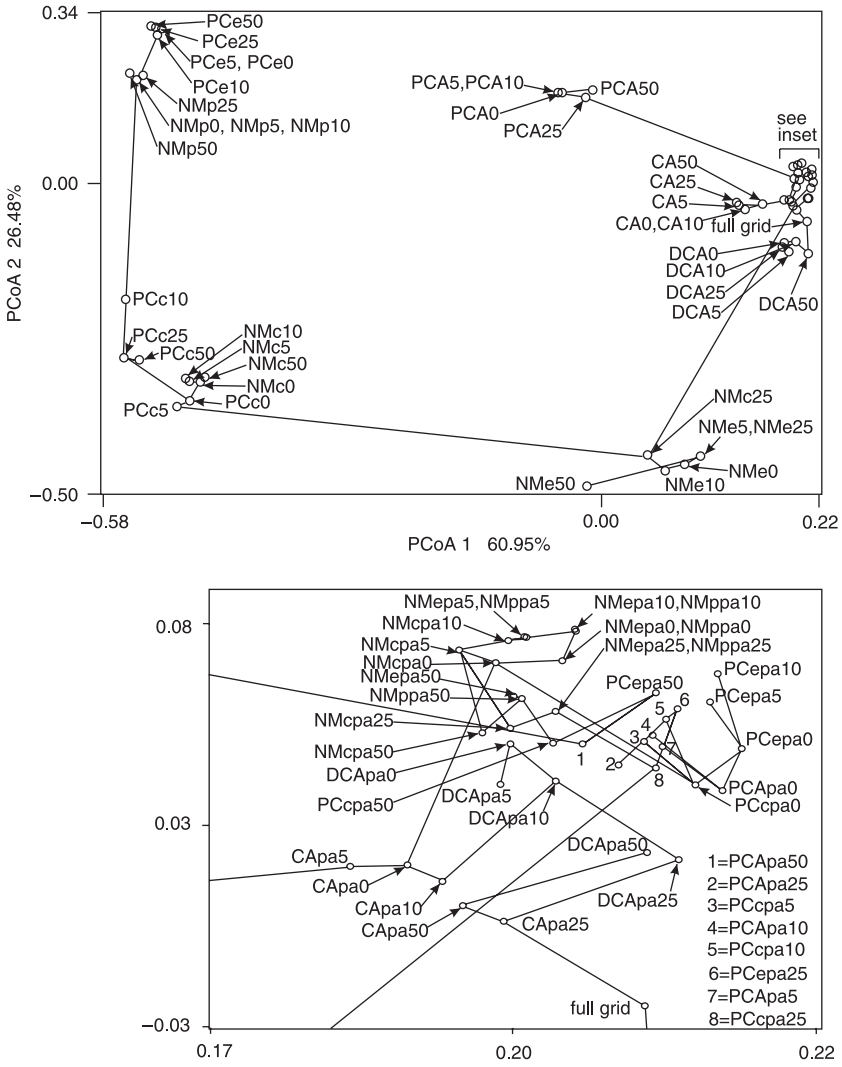

Figure 6 Principal coordinates analysis (PCoA) results from an analysis comprising all combinations of the ordinations and error levels with the original full grid. The inset, located below the main figure, magnifies a portion of the plot in order to more accurately portray certain relationships. The following abbreviations were used for the sake of space: NM represents NMDS (non-metric multidimensional scaling) and PC represents PCoA.

When abundance data were used, the behaviour of NMDS provided some particularly instructive results. An NMDS ordination that used a particular resemblance measure tended to fall out with a PCoA plot using the same measure, for the longer gradient only (Fig. 6). For the shorter gradient, $NMDS_P$, $NMDS_C$, and $NMDS_E$ ordinations clustered together and exhibited similar levels of distortion, regardless of the initial configuration or choice of resemblance measure (Fig. 7). This is not to imply that each variation of NMDS gave the same result; the ordinations indicated a range of data structures (Figs 2 & 3). These results show that the choice of resemblance measure, and initial configuration, for NMDS is of greater importance for long gradients, on which NMDS performed poorly, in any case. Other researchers have confirmed that the choice of coefficient is important for NMDS (Kenkel & Orlóci, 1986; Jackson, 1993b), but differed in their conclusion about the relative suitability of NMDS to accurately recover the data structure.

There are two additional, interesting results from NMDS using abundance data, one being the difference in distortion between gradients. The longer gradient produced more highly distorted plots than the shorter half gradient. The improved ability of the resemblance measures to accurately capture the relationships between observations with the shorter gradient is responsible for this situation. Although proponents of NMDS suggest that it is not affected by non-linearities between variables, it is still essential that the resemblance measure is able to adequately capture the relationships among the observations. Failure to do so results in distorted results and an inability to recover the true relationships. Other studies emphasize that NMDS is less sensitive to zero values and non-linearities (Fasham, 1977; Kenkel & Orlóci, 1986), but this sensitivity will vary depending on the choice of resemblance measure.

The second noteworthy result is that NMDS results demonstrate that low stress may in fact have higher levels of distortion as measured by m^2 values, although stress is used as a measure of goodness of fit (Legendre & Legendre, 1998). For $NMDS_C$ on the full gradient, distortion decreased markedly at 25% error (Table 1.2), and then increased at 50% error. This finding was unexpected and we returned to our $NMDS_C$ ordination results in order to examine whether this decrease in distortion was simply a chance effect. We selected a second result with higher stress than our presumably more optimal solution, and found that

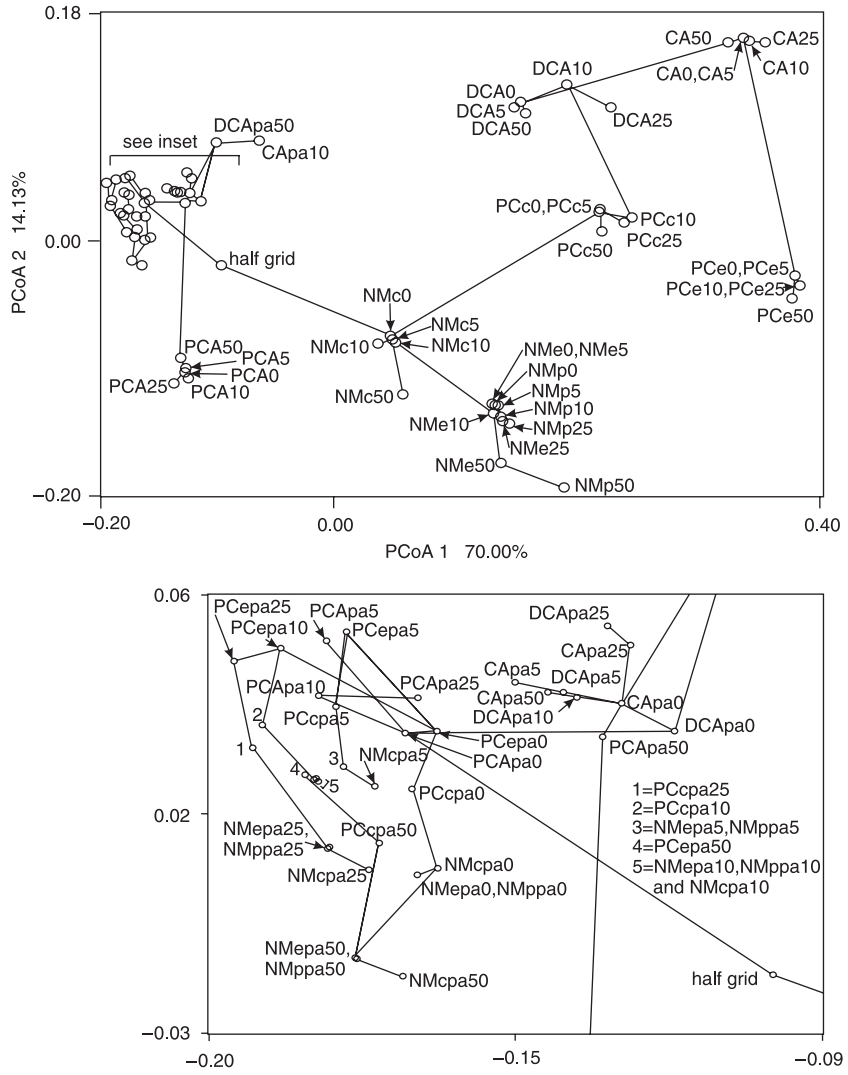

Figure 7 Principal coordinates analysis (PCoA) results from an analysis comprising all combinations of the ordinations and error levels with the original half grid. The inset, located below the main figure, magnifies a portion of the plot in order to more accurately portray certain relationships. The following abbreviations were used for the sake of space: NM represents NMDS (non-metric multidimensional scaling) and PC represents PCoA.

it was even less distorted, with an m^2 value of 0.39. This result indicates that although performance measures typically used in NMDS may indicate a good solution, these measures may not be adequately capturing the relationship to the original data. This may be due to the stress measures being dependent on monotonic or linear fits between resemblance measures, rather than a direct comparison of the ordination solution to the geometry of the original data pattern (similar relationships exist in the analysis of distance matrices with Mantel test vs. the direct analysis of the original variables, i.e. the analysis of distances is not identical to the analysis of the variables for those same observations). Unfortunately, we cannot know what the true pattern in the data is with field data and consequently we must use indirect measures such as stress, even though our results show that such measures may not provide the optimal solutions.

Kenkel & Orlóci (1986) found that NMDS (with abundance data) was a highly effective ordination strategy for a coenoplane model when chord distance was used as the measure of similarity. In contrast, we found that the usefulness of ordinations produced by NMDS depended heavily on gradient length. For short gradients, NMDS did indeed perform well, but not as well as PCA or analyses that used presence–absence data. Our study design differed from that of Kenkel & Orlóci (1986) given we had more observations, we used 40 species as opposed to 30, and we examined the effects of sampling error on coenoplane behaviour. Kenkel & Orlóci (1986) examined coenoplane behaviour at a range of species turnover rates (from 2.65 half changes to 7.50 half changes) rather than a range of sampling errors. While there is no simple or single explanation for the discrepancy between studies (e.g. we considered both random and metric starting configurations for our analyses), the difference between our results and those of Kenkel & Orlóci (1986), as well as the incongruous relationship between stress and distortion, indicates a need for further comparisons of NMDS results with those from other methods.

Various measures of similarity/distance have standardizations or transformations implicit in the measure, but in various instances the performance of several of the methods may be improved through the incorporation of additional transformations and standardizations (e.g. see Legendre & Gallagher, 2001; Peres-Neto *et al.*, 2006). While evaluations of the performance of such standardizations and transformations can be made when data are simulated communities, one cannot assess the relative

improvement of such modifications when analyses are based on field data and there is no 'known' or 'true' pattern of observations to which the ordinations can be compared.

It has been suggested that studies using simulated data do not adequately represent the range of community variability in natural studies. Clearly, every field-based study differs in the number of sites and species sampled, the degree of species turnover across the range of sites, and various other factors. Simulation-based studies may fail by not capturing some of these conditions or by creating very different conditions. Two relevant points of criticism relate to gradient length (i.e. is it too long or too short?) and the ratio of gradient lengths when using two-dimensional simulations. We (D.A. Jackson, R. Paavola, and T. Willis, unpubl. data) reviewed all papers published over a 2-year period in 12 ecological journals. From these studies encompassing terrestrial and aquatic communities of numerous types, we determined that the average length of the first gradient was 4.10 (range of 1.36–11.98 for field-based studies), but this mean is slightly overestimated as several studies simply reported their gradient length as < 2 or < 3 when justifying the use of particular ordination methods. This finding compares with that in our full-gradient simulation of 3.10. The average ratio of gradient length for the first and second axes reported in these various publications was 1.17 relative to the value of 1.14 used in our present study. Therefore our full-gradient simulation appears typical of results reported for field studies and our half-gradient simulation would have a shorter gradient than is typical, but within the range reported in the literature and both the full- and the half-gradient simulations would virtually match the typical ratio of gradient length for the first and second axes as reported in the literature.

As one might expect, lower levels of sampling error tend to produce more accurate, less distorted results overall. Consequently, the minimization of sampling error through consistent methodology among researchers is a concern that must be addressed in order for field data to produce ecologically meaningful ordinations. In order to confront this problem, one approach is to perform multiple ordinations, with varying resemblance measures, and look for consensus among the results. Further work could consider the use of replicated sampling to examine the effects of sampling error further. This would allow comparisons to be made between multivariate methods using a common set of replicates vs. comparisons between methods based on different replicates. Similarly pooling of replicates will reduced variance between replicates, but this still requires that all species be sampled equally well (i.e. without bias or at least a consistent bias) which will not be the case when species vary in life form, patchiness, habitat, or vulnerability to sampling regime. While many of these issues have been considered by plant ecologists (e.g. Whittaker, 1978), they remain a major problem in sampling many species of animals (e.g. fishes in Jackson & Harvey, 1997) with further field comparisons required in various cases. However, an alternative approach is to use a presence–absence data format that provides a very strong standardization for differences in abundance — all species are considered numerically equivalent — which as long as the species is detected during the sampling will provide correct assessments of species presence or absence. Our results indicate that the use of presence–absence data removes much of the noise induced by sampling error, and we found that presence–absence data in a correspondence analysis or even a principal component analysis provided the best means of recovering the true underlying pattern in the data regardless of the gradient length considered. The use of presence–absence data in community analyses is recommended given that: (1) presence–absence data are often easier to collect in the field (Green, 1979); (2) allows the inclusion of more observations with a given amount of effort than trying to adequately capture relative abundance data (Jackson & Harvey, 1989, 1997); and (3) provided a more accurate representation of the true community relationships as shown by the results, particularly from PCA_{pa} and CA_{pa}. In many cases, given the sampling error involved, we may be misleading ourselves with our use of relative abundance data in community multivariate analyses.

ACKNOWLEDGEMENTS

We would like to thank NC Kenkel, R Økland, R Paavola, PR Peres-Neto, BJ Shuter, and KM Somers for their comments on this manuscript. We thank PR Peres-Neto for his comments and providing a program for calculating all pairwise m^2 values. Funding was provided by an NSERC graduate scholarship to C.N.H. and an NSERC Discovery Grant to D.A.J.

REFERENCES

Bowman, M.F., Somers, K.M., Reid, R.A. & Scott. L.D. (2006) Temporal response of stream benthic macroinvertebrate communities to the synergistic effects of anthropogenic acidification and natural drought events. *Freshwater Biology*, **51**, 768–782.

ter Braak, C.J.F. (1986) *CANOCO: A FORTRAN program for canonical correspondence analysis*. IWIS-TNO, Wageningen, the Netherlands.

ter Braak, C.J.F. & Prentice, I.C. (1988) A theory of gradient analysis. *Advances in Ecological Research*, **18** (271), 317.

Bradley, D.C. & Ormerod, S.J. (2002) Evaluating the precision of kick-sampling in upland streams for assessments of long-term change: the effects of sampling effort, habitat and rarity. *Archiv Fur Hydrobiologie*, **155**, 199–221.

Brazner, J.C. & Beals, E.W. (1997) Patterns in fish assemblages from coastal wetlands and beach habitats in Green Bay, Lake Michigan: a multivariate analysis of abiotic and biotic forcing factors. *Canadian Journal of Fisheries and Aquatic Sciences*, **54**, 1743–1761.

Courtemanch, D.L. (1996) Commentary on the sub-sampling procedures used for rapid bioassessment. *Journal of the North American Benthological Society*, **15**, 381–385.

De'ath, G. (1999) Extended dissimilarity: a method of robust estimation of ecological distances from high beta diversity data. *Plant Ecology*, **144**, 191–199.

Digby, P.G.N. & Kempton, R.A. (1987) *Multivariate analysis of ecological communities*. Chapman & Hall, London.

Fasham, M.J.R. (1977) A comparison of nonmetric multidimensional scaling, principal compoenents and reciprocal averaging for the ordination of simulated coenoclines and coenoplanes. *Ecology*, **58**, 551–561.

Gauch, H.G. (1982) *Multivariate analysis in community ecology*. Cambridge University Press, Cambridge, England.

Gauch, H.G., Whittaker, R.H. & Wentworth, T.R. (1977) A comparative study of reciprocal averaging and other ordination techniques. *Journal of Ecology*, **65**, 157–174.

Green, R.H. (1979) *Sampling design and statistical methods for environmental biologists*. John Wiley & Sons, Inc., New York.

Green, R.H. & Young, R.C. (1993) Sampling to detect rare species. *Ecological Applications*, **3**, 351–356.

Jackson, D.A. (1993a) Stopping rules in principal components analysis: a comparison of heuristical and statistical approaches. *Ecology*, **74**, 2204–2214.

Jackson, D.A. (1993b) Multivariate analysis of benthic invertebrate communities: the implication of choosing particular data standardisations, measures of associations, and ordination methods. *Hydrobiologia*, **268**, 9–26.

Jackson, D.A. (1995) PROTEST: a PROcrustean randomisation TEST of community environment concordance. *Ecoscience*, **2**, 297–303.

Jackson, D.A. (1997) Compositional data in community ecology: the paradigm or peril of proportions? *Ecology*, **78**, 929–940.

Jackson, D.A. & Harvey, H.H. (1989) Biogeographic associations in fish assemblages: local versus regional processes. *Ecology*, **70**, 1472–1484.

Jackson, D.A. & Harvey, H.H. (1997) Qualitative and quantitative sampling of lake fish communities. *Canadian Journal of Fisheries and Aquatic Sciences*, **54**, 2807–2813.

Jackson, D.A. & Somers, K.M. (1991) Putting things in order: The ups and downs of detrended correspondence analysis. *The American Naturalist*, **137**, 704–712.

Jackson, D.A., Somers, K.M. & Harvey, H.H. (1989) Similarity coefficients: measures of co-occurrence and association or simply measures of occurrence? *The American Naturalist*, **133**, 436–453.

Jackson, D.A., Peres-Neto, P.R. & Olden, J.D. (2001) What controls who is where in freshwater fish communities: the roles of biotic, abiotic, and spatial factors. *Canadian Journal of Fisheries and Aquatic Sciences*, **58**, 157–170.

Jongman, R.H.G., ter Braak, C.J.F. & van Tongeren, O.F.R. (1987) *Data analysis in community and landscape ecology*. PUDOC, Wageningen, the Netherlands.

Karadzic, B. & Popovic, R. (1993) On the incompatibility of the chord distance with some classification and ordination algorithms used in studies of plant-communities. *Zhurnal Obshchei Biologii*, **54**, 430–437.

Kenkel, N.C. & Orlóci, L. (1986) Applying metric and nonmetric multidimensional scaling to ecological studies: some new results. *Ecology*, **67**, 919–928.

Kenkel, N.C., Derksen, D.A., Thomas, A.G. & Watson, P.R. (2002) Multivariate analysis in weed science research. *Weed Science*, **50**, 281–292.

Kerans, B.L., Karr, J.R. & Ahlstedt, S.A. (1992) Aquatic invertebrate assemblages: spatial and temporal differences among sampling protocols. *Journal of the North American Benthological Society*, **11**, 377–390.

Kessel, S.R. & Whittaker, R.H. (1976) Comparisons of three ordination techniques. *Vegetatio*, **34**, 191–197.

Kremen, C. (1992) Assessing the indicator properties of species assemblages for natural areas monitoring. *Ecological Applications*, **2**, 203–217.

Legendre, P. & Gallagher, N.E. (2001) Ecologically meaningful transformations for ordination of species data. *Oecologia*, **129**, 271–280.

Legendre, P. & Legendre, L. (1998) *Numerical ecology*. 2nd English edn. Elsevier Science B, V., Amsterdam, Netherlands.

Manly, B.J.F. (2005) *Multivariate statistical methods: a primer*, 3rd edn. Chapman & Hall, London.

McCoy, E.D. & Mushinsky, H.R. (1992) Rarity of organisms in the sand pine scrub habitat of Florida. *Conservation Biology*, **6**, 537–548.

McCune, B. (1994) Improving community analysis with the Beals smoothing function. *Ecoscience*, **1**, 82–86.

McCune, B. (1997) Influence of noisy environmental data on canonical correspondence analysis. *Ecology*, **78**, 2617–2623.

Minchin, P.R. (1987) Simulation of multidimensional community patterns: towards a comprehensive model. *Vegetatio*, **71**, 145–156.

Noy-Meir, I., Walker, D. & Williams, W.T. (1975) Data transformations in ecological ordination. II. On the meaning of data standardisation. *Journal of Ecology*, **63**, 779–800.

Oksanen, J. & Minchin, P.R. (1997) Instability of ordination results under changes in input data order: explanations and remedies. *Journal of Vegetation Science*, **8**, 447–454.

Oksanen, J. & Minchin, P.R. (2002) Continuum theory revisited: what shape are species responses along ecological gradients? *Ecological Modelling*, **157**, 119–129.

Ominski, P.D., Entz, M.H. & Kenkel, N.C. (1999) Weed suppression by *Medicago sativa* in subsequent cereal crops: a comparative survey. *Weed Science*, **47**, 282–290.

Ostermiller, J.D. & Hawkins, C.P. (2004) Effects of sampling error on bioassessment of stream ecosystems: application to RIVPACS-type models. *Journal of the North American Benthological Society*, **23**, 363–382.

Peres-Neto, P.R. & Jackson, D.A. (2001) How well do multivariate data sets match? The advantages of a Procrustean superimposition approach over the Mantel test. *Oecologia*, **129**, 169–178.

Peres-Neto, P.R., Jackson, D.A. & Somers, K.M. (2003) Giving meaningful interpretation to ordination axes: assessing loading significance in principal component analysis. *Ecology*, **84**, 2347–2363.

Peres-Neto, P.R., Legendre, P., Dray, S. & Borcard, D. (2006) Variation partitioning of species data matrices: estimation and comparison of fractions. *Ecology*, **87**, 2614–2625.

Resh, V.H., Beche, L.A. & McElravy, E.P. (2005) How common are rare taxa in long-term benthic macroinvertebrate surveys? *Journal of the North American Benthological Society*, **24**, 976–989.

Rohlf, F.J. (1997) *Ntsyspc. Numerical taxonomy and multivariate analysis system*. Exeter Publishing, Setauket, New York.

Tamas, J., Podani, J. & Csontos, P. (2001) An extension of presence/absence coefficients to abundance data: a new look at absence. *Journal of Vegetation Science*, **12**, 401–410.

Whittaker, R.H. (1978) *Ordination and classification of plant communities*. Junk, The Hague, the Netherlands.

Zimmer, K.D., Hanson, M.A. & Butler, M.G. (2000) Factors influencing invertebrate communities in prairie wetlands: a multivariate approach. *Canadian Journal of Fisheries and Aquatic Sciences*, **57**, 76–85.