Ordinating tropical moth ensembles from an elevational gradient: a comparison of common methods

Gunnar Brehm and Konrad Fiedler

Department of Animal Ecology I, University of Bayreuth, 95440 Bayreuth, Germany

(Accepted 25 March 2003)

Abstract: The analysis of beta diversity (inter-habitat diversity) of very species-rich and incompletely sampled tropical arthropod communities requires the choice of appropriate statistical tools. The performance of the three commonly employed ordination methods, correspondence analysis (CA), detrended correspondence analysis (DCA), and non-metric multidimensional scaling (NMDS), was compared on a large empirical data set of geometrid moths sampled along an altitudinal gradient in an Andean montane rain forest. Despite the high species richness and incompleteness of the ensembles, all methods depicted the same, readily interpretable patterns. Both CA and NMDS showed an arch-like structure, which hints at an underlying coenocline, whereas this arch was computationally eliminated in DCA. For this particular data set, CA and NMDS both provided convincing results while the detrending algorithm of DCA did not improve the interpretability of the data. Of the large number of similarity indices available to be used in combination with NMDS, the binary Sørensen and the abundance-based Normalized Expected Species Shared (NESS) index were tested. Performance of the indices was measured by comparing stress, a measure of poorness-of-fit in NMDS. NMDS ordinations with lowest values of stress were achieved by the NESS index with the parameter \( m \) set to its maximum \( (m_{\text{max}}) \). In contrast, ordinations based on NESS values with the parameter \( m \) set to 1 (identical with Morisita’s index), had consistently higher stress values and performed worse than ordinations using Sørensen’s index. Hence, if high values of \( m \) can be achieved in similar data sets, the NESS index with \( m_{\text{max}} \) is recommended for ordination purposes and Morisita’s index should be avoided.

Key Words: arches, beta diversity, correspondence analysis, detrended correspondence analysis, Geometridae, incomplete sampling, NESS index, non-metric multidimensional scaling, Sørensen index, tropical arthropods

INTRODUCTION

Changes in communities of animals or plants along habitat gradients have attracted attention since Alexander von Humboldt’s investigations of altitudinal patterns in plant diversity in the Andes (von Humboldt & Bonpland 1807). Whittaker (1972) coined the term beta diversity (inter-habitat diversity) for the change of organismic diversity along habitat gradients. Such gradients offer systems in which complex and species-rich communities and their responses to environmental factors can be studied (Körner 2000). Beta diversity of a variety of organisms in tropical rain forests is thought to be particularly high on small spatial scales (Gentry 1988, Hammond 1994, Stevens & Willig 2002, Wolda 1996). Tropical arthropod communities pose specific methodological challenges. Two major problems are (1) that their coverage by sampling is incomplete because communities are extremely species rich (e.g. Price et al. 1995), and (2) that many species are represented by only one or very few individuals (Floren & Linsenmair 1998, Wagner 2000). In particular this latter property of the samples is frequently regarded as a nuisance in the statistical analysis and interpretation of the data (Novotný & Basset 2000). Studies on tropical arthropods which have applied ordination methods are still scarce (but see Beck et al. 2002, Kitching et al. 2000, Schulze & Fiedler 2003, Wagner 2000, Willott 1999). Here we focus on the performance of selected ordination methods and similarity indices on a large empirical data set in order to give recommendations on which of these techniques are most appropriate to use. The data set resulted from light-trapping of moths of the family Geometridae (Lepidoptera) in a montane rain forest in southern Ecuador. Samples are characterized by a very high species richness and

1 Corresponding author. Email: Gunnar_Brehm@yahoo.com
alpha diversity as well as by incomplete coverage and a high ratio of rare species (Brehm et al. 2003a). A detailed interpretation of the beta diversity patterns and its possible determinants was presented by Brehm et al. (2003b).

It is often difficult to choose the most appropriate method from the great number of available multivariate techniques and similarity indices because no definitive guideline exists. Ordination techniques aim to find a parsimonious representation of objects such as sites or species in a space of low dimensionality (Kenkel & Orlóci 1986). Two-dimensional representations are easiest to interpret and most readily compared visually (Jackson & Somers 1991, Shepard 1974). Unfortunately, no objective method exists to assess the efficiency of different ordination methods (Kenkel & Orlóci 1986). We focused our analyses on three commonly employed methods: the metric techniques correspondence analysis (CA) and detrended correspondence analysis (DCA), as well as non-metric multidimensional scaling (NMDS). CA and DCA can be calculated directly from the original species × site matrices, whereas NMDS requires the calculation of a similarity (or dissimilarity) matrix from the original matrix. The choice of an appropriate similarity index is important, since different indices might lead to divergent results. The Morisita index is almost uninfluenced by the sample sizes unless samples are small (Morisita 1959) and was recommended by Wolda (1981). All other indices tested by Wolda depended on sample size and/or species diversity to a varying extent. Later, Wolda (1983) identified the NESS index (Normalized Expected Species Shared; Grassle & Smith 1976) as ‘far superior’ to any other. This index is a generalization of Morisita’s index and considers that two random samples even when drawn from the same community are usually not identical but differ due to stochastic effects. The NESS index can be adjusted by choice of the sample size index, \( m \) (Grassle & Smith 1976). Therefore, NESS is not just one similarity index, but rather a family of indices. A popular and widely used binary index of similarity is the Sorensen index (Sørensen 1948) that is identical to the Czekanowski index and closely related to other indices such as the Bray–Curtis and Jaccard index. It is not based on species abundances, but on presence–absence data. In many biodiversity inventories, data often do not exceed this presence–absence quality. Therefore, we aimed at assessing as to how strongly ordination results for a highly complex fauna would differ when relying on this coarser source of information.

**METHODS**

Within three field periods in the years 1999 and 2000, a total of 13 938 specimens of geometrid moth representing 1010 species were sampled by using portable weak light-traps (2 × 15 W) in a tropical montane rain forest in southern Ecuador. The alpha diversity of the investigated ensembles (terminology according to Fauth et al. 1996) of the moths was very high with values of Fisher’s alpha ranging from 69 to 131 and estimated species numbers reaching up to 400 and more per site (Brehm et al. 2003a). At single sites, 52 ± 6% (\( n = 22 \)) of all species were singletons, i.e. species represented by only one specimen. The estimated species number per site, based on the Chao 1 estimator, exceeded the actual species number greatly by 76 ± 27% (\( n = 22 \)) (Brehm et al. 2003a). Geometrid moths are taxonomically relatively well known (Pitkin 1996, 2002; Scoble 1999) and are a suitable model group for tropical biodiversity studies (Beck et al. 2002, Intachat et al. 1997). The study area is situated within the East Cordillera of the Andes in Zamora–Chinchipe province (Reserva Biológica San Francisco, 3°58'S, 79°5'W). A description of the area, including exact geographical and elevational positions of all sites, and sampling methods was provided by Brehm (2002) and Brehm & Fiedler (2003). Specimens were sampled at 11 elevations (two sites at each elevation). The lowest sites (elevational level number 1) were situated at 1040 m above sea level, while the highest sites (elevational level number 11) were situated at 2670 and 2677 m. The three lowest sampling levels were spaced at approximately 400 m of elevation, whereas all other levels had an altitudinal distance of c. 100 m between each other. Moths were identified using large reference collections (Zoologische Staatssammlung, Munich; The Natural History Museum, London).

We ordinated the samples using CA, DCA as well as NMDS (based on four different indices of similarity: Sørensen index, NESS index with \( m = 1 \), \( m = 50 \) and \( m_{\text{max}} = 192 \)). The ordinations obtained were checked for their representation of the expected altitudinal pattern of species turnover in geometrid ensembles by correlating the scores of sites with regard to the first extracted dimension with elevation of the sampling sites (Pearson correlation coefficient). In order to more precisely assess differences between non-metric ordinations derived from different similarity indices by an objective measure, the behaviour of ‘stress’ in NMDS was investigated. Stress is a measure that reflects the degree of deviation of NMDS distances from original matrix distances. These distances can be plotted against each other (Shepard 1974). A line denotes the best-fit regression while the scatter about this line is defined as stress (Clarke 1993). Analyses were carried out at the level of the whole family and for the seven largest subordinated taxa (Table 2), and with three of the four similarity measures used for ordination (Sørensen and the NESS indices \( m = 1 \) and \( m = 192 \)).

All standard statistical methods were performed using the software package Statistica 5.5 (StatSoft 1999).
Ordinating tropical moth ensembles from an elevational gradient

Figure 1. Ordinations of samples of geometrid moths from a montane rain forest in southern Ecuador. (a–c) Non-metric two-dimensional scaling based on the NESS index, (d) non-metric ordination based on the Sørensen index, (e) correspondence analysis (CA): inertia of first axis 18.5%, inertia of second axis 13.2%, (f) detrended correspondence analysis (DCA): inertia of first axis 18.5%, inertia of second axis 6.3%. Each dot represents the geometrid ensemble from one site. Numbers from 1 to 11 correspond to altitudinal levels from low (1040 m) to high elevations (2677 m). All ordinations arrange the ensembles according to their elevational position (see Table 1 for statistics).

The Sørensen and NESS indices were calculated with a program provided by S. Meßner (Meßner 1996). A modified version of the program DECORANA in the software PC-Ord (26 segments; McCune & Mefford 1999) was used to perform DCA.

RESULTS

Performance of ordination methods

Figure 1 shows two-dimensional ordinations of NMDS, CA and DCA of all geometrid moths. All results are readily interpretable since a gradual structure from site levels 1 to 11 (i.e. revealing the altitudinal sequence) is depicted along the first axis. In contrast, interpretation of the second axis or even higher (CA- and DCA-) axes is questionable for this particular data set (Brehm 2002). Visually, the CA ordination shows the ‘clearest’ picture. Among the NMDS ordinations (Sørensen index: NESS \( m = 1, m = 50, m_{\text{max}} = 192 \)) the differences are slight. The extracted scores of the first dimension of all ordinations strongly correlate with altitude (all \( P < 0.001 \), Table 1). This is not only the case in geometrid moths as a whole, but also for all seven subordinate taxa. Correlations in the five smaller taxa investigated are weaker than in Geometridae, Ennominae and Larentiinae. However, none of the ordination methods performs better or worse than the others with regard to the reflection of the elevational sequence. All NMDS ordinations as well as CA reveal an arch-like structure which was eliminated in DCA by the detrending algorithm.

Performance of similarity indices

Figure 2 shows two Shepard diagrams for ordinations of all Geometridae \( (m = 1 \) and \( m_{\text{max}} = 192 \)) after two-dimensional scaling. Both results are readily interpretable, but the value of stress with \( m_{\text{max}} \) is only half of that with \( m = 1 \). Moreover, the Shepard plot of the \( m_{\text{max}} \) ordination shows a smooth, almost linear relationship \( (y = -0.35 + 2.84x, R^2 = 0.93, P < 0.0001) \). In contrast, distances between sampling sites in the
two-dimensional ordination based on \( m = 1 \) abruptly increase at NESS distances > 0.9. In Figure 3, stress values were compared at a range of intermediate \( m \) values, and for the Sørensen index, for one-, two- and three-dimensional (1D, 2D, 3D) scaling. The results show two major trends: first, stress decreases with increasing \( m \) in asymptotic curves. Major changes occur in the range from \( m = 1 \) to \( m = 20 \), whereas changes at higher levels of \( m \) are less pronounced and almost level off at values of \( m \geq 100 \). Stress in the ordinations using the Sørensen index has intermediate values (corresponding to approximately NESS \( m = 20 \) in 3D scaling, \( m = 50 \) in 2D scaling and \( m = 70 \) in 1D scaling). Second, stress is lower in 3D representations of data as compared with 2D or 1D ordinations. The difference between 1D and 2D ordinations is about twice as large compared with that from 2D to 3D ordinations. Hence, little would be gained from using a three-dimensional ordination compared with the visually more comprehensible two-dimensional plots.

Table 2 provides stress values in two-dimensional scaling for eight taxa. Only in the three largest groups, high NESS \( m_{\text{max}} \) values (> 30) are available. In these cases, stress in non-metric ordinations using NESS \( m_{\text{max}} \) is considerably lower than with NESS \( m = 1 \) and the Sørensen index. In the remaining taxa represented by fewer species, \( m_{\text{max}} \) reaches only values between 3 and 10, and no consistent stress pattern can be seen in relation to the similarity measures applied. As expected, absolute stress values decrease as numbers of species and

![Shepard diagram](image)

**Figure 2.** Shepard diagrams of non-metric two-dimensional ordinations of two different NESS indices of Geometridae \((m = 1 \text{ and } m_{\text{max}} = 192 \text{ respectively})\). The diagrams correspond to the ordinations (a) and (c) shown in Figure 1. Stress is defined as scatter around the best-fit regression line.

![Stress plot](image)

**Figure 3.** Stress in non-metric scaling as a function of the size of the parameter \( m \) of the NESS index and of the Sørensen index. Data are from 13 938 specimens of 1010 species of geometrid moths from 22 sampling sites in southern Ecuador. Stress is systematically higher in one-dimensional ordinations (1D) than in two-dimensional (2D) and three-dimensional (3D) ordinations. Stress decreases in all cases with increasing \( m \), except in three-dimensional ordinations where \( m > 150 \).
specimens increase. For the taxa represented by fewer species, values are below 0.2, whereas in the three largest taxa, values are below 0.11. The smallest stress value is achieved for the total geometrid sample with 0.04 for $m_{\text{max}}$ (see also Figure 3).

**DISCUSSION**

**Which ordination method is preferable?**

All ordination methods examined provide essentially the same, readily interpretable results on the first axis. With regard to Geometridae, sites are arranged according to their elevation, reflecting a successive replacement of species along the gradient. Very similar patterns emerged when the two large geometrid subfamilies Ennominae and Larentiinae were ordinated separately (figured in Brehm et al. 2003b). Even small taxa represented by average specimen numbers of only 44–97 per site are ordinated accurately. The correlations between the first ordination axis and altitude are generally less strong than in the three largest taxa, because stochastic effects play a greater role in small samples than in samples with many species and specimens. However, the correlations are still remarkably high. No consistent differences between the different ordination methods tested were detected with regard to their reflection of the altitudinal sequence. Hence, none of the methods tested failed to ordinate species-rich and incomplete data sets as presented in this study.

Using simulated data, Faith et al. (1987) found that the use of chi-square distances, as inherent in CA, is often inappropriate. Legendre & Legendre (1998) stated that rare species are relatively over-emphasized in CA, and subsequently Legendre & Gallagher (2001) suggested a transformation of data before the application of CA and related ordination methods. The emphasis of rare species in CA could turn out as an advantage in the analysis of tropical communities that are characterized by the occurrence of a great number of singletons. In our data set, 20% of all species were singletons, i.e. species that were collected only once in the entire study, and $52 \pm 6\%$ ($n=22$) were singletons at single sites. The high proportion of singletons did not result in any distortions in CA as compared with non-metric ordinations.

Hill & Gauch (1980) criticized CA because it often shows an ‘arch’ form (or ‘horseshoe’) in the first two dimensions if a coenocline is investigated. In addition, distances at the lower and upper end of the parameter space covered by the samples to be ordinated are compressed relative to the middle. The nature of this arch is a controversial issue. It has been interpreted as an artefact (Hill & Gauch 1980), but in strong contrast it was also described as an important and inherent property of successive replacement data (Wartenberg et al. 1987). Such successive replacement is to be expected along extended altitudinal gradients with a pronounced species turnover. Arches appear prominently in ordinations of the present empirical data set (Figure 1a–e). The arches probably indicate that only one dominant gradient (viz. altitude) governs our data set, i.e. a coenocline situation. In the second axis, sites from the extremes of the gradient have similar values despite their very low faunal similarity (Sørensen values between 0.05 and 0.11, NESS ($m_{\text{max}}$) values between 0.06 and 0.10). This phenomenon can be interpreted as involution, i.e. the apparent closeness of dissimilar extremes of an environmental gradient (Wartenberg et al. 1987). Brehm et al. (2003b) showed that arches also appear in simulated coenocline model data sets.

In DCA, arches are eliminated by the detrending algorithm. There has been a controversial debate about
this method. Kenkel & Orlóci (1986) emphasized that detrending can distort underlying data structure. Moreover, Wartenberg et al. (1987) fundamentally criticized the concept as having no empirical or theoretical justification. To the contrary, Peet et al. (1988) defended detrending and rescaling as necessary steps in order to make higher dimensions interpretable. However, Jackson & Somers (1991) showed that the choice of axis segmentation in DCA could substantially affect the interpretation of the results. The selection of the number of segments is arbitrary, since no guidelines exist for different numbers of samples. Jackson & Somers (1991) concluded that multidimensional configurations obtained by using DCA might be unstable and potentially misleading. Brehm (2002) showed that DCA did not improve the interpretability of simulated model data as compared with CA. For the data analysed here, DCA did not represent the altitudinal gradient more clearly than the other methods, and did not reveal any additional information.

In NMDS, unlike metric ordinations, the total number of dimensions can a priori be fixed to a low number (e.g. two or three dimensions). To the contrary, both metric methods have a high number of further axes that explain a decreasing proportion of the variability of the data and that cannot be eliminated in advance. The first two dimensions in CA and DCA only explain 32% and 25% of the variability (‘inertia’) of the data, respectively (Figure 1). The a priori fixing of dimensions can be regarded as an advantage of NMDS. Furthermore, the underlying rank statistics of NMDS minimizes the influence of outliers and leads to very robust results (Clarke 1993, Minchin 1987). Transformation of data as suggested by Legendre & Gallagher (2001) for metric techniques is not required. Until the 1980s, NMDS was criticized as being computationally more demanding than CA (Hill & Gauch 1980, Kenkel & Orlóci 1986), but this problem has vanished following progress in computer technology. This historical background might be one of the reasons why NMDS is still less commonly used than CA.

Overall, metric and non-metric ordination techniques resulted in remarkably similar representations of tropical moth ensembles in two-dimensional space. All performed equally well in depicting the elevational gradient in the first dimension, despite the extreme diversity and the occurrence of many rare species. Our results show that all methods applied appear to be appropriate for the data analysed. Whether these findings are valid for similar data sets of tropical arthropods, needs to be tested in the future. Overall, we recommend the use of both NMDS and CA. DCA appears to be less appropriate than the other methods for our data set since the detrending algorithm did not add additional information.

Which similarity index is most appropriate for non-metric ordination?

The NESS index is still rarely used, but has become more widely applied in the past decade, e.g. by Beck et al. (2002), Schulze & Fiedler (2003), Willott (1999) and Wolda (1992, 1996). Some authors have arbitrarily set the parameter \( m \) to 20 without explaining their choice (Willott 1999, Wolda 1983), while others have alternatively presented \( m = 1 \) and \( m_{\text{max}} \) (Schulze & Fiedler 2003). We have shown that for ordination purposes a maximum \( m \) performs better in terms of minimizing stress than a low \( m \), and that intermediate \( m \) values do not differ strongly in their performance from very high \( m \). Therefore, in future studies, \( m_{\text{max}} \) or a high \( m \) should be preferred over a low to intermediate \( m \), since the results are the easiest to interpret. Trueblood et al. (1994) stated that benthic animal communities with a very high proportion of singletons may not be properly analysed with the NESS index. In such communities, NESS can overestimate similarities and may reach values \( > 1.0 \). However, our study shows that the analysis of geometrid moth communities characterized by many singletons poses no serious problem. Among 231 pairwise comparisons calculated with NESS \( m_{\text{max}} \), there is only one similarity with a value \( > 1.0 \) (1.04 between the two sites at elevational level 3; Brehm et al. 2003b). We calculated the Chord-NESS (CNESS) index introduced by Trueblood et al. (1994) and got very similar ordinations and a high correlation between similarities of NESS and CNESS (not shown).

The Sørensen index also proved to be a suitable measure of similarity in this particular data set. Stress was low \( (< 0.10) \) for larger taxa (Geometridae, Ennominae, Larentiinae), and was still acceptable in the others \( (< 0.18) \) according to Clarke’s (1993) criteria. Hence, our results show that at least under certain conditions simple presence–absence measures such as Sørensen’s index can be useful. Although data on species abundances add valuable information, insects can change in abundance from night to night or during the year. This might explain the somewhat ‘uneven’ picture of the NESS \( m = 1 \) ordination in which the most abundant species are highly emphasized. Sørensen’s index proved to be useful in the data set investigated because sample sizes and coverage by actual species numbers differed only moderately between sites (Brehm 2002). However, in other data sets that are more heterogeneous with regard to sample size, Sørensen’s index will be less appropriate, as was shown e.g. for pyraloid moths in the same study area (Süßenbach et al. 2001). The Sørensen index calculates the ratio of species that two samples have in common. If one sample is very small, there are only very few shared species. In such cases, similarity values are low, even if the samples are drawn from identical communities.
In conclusion, the NESS index appears to be well suited for arthropod communities that are incompletely sampled and contain many rare species. This index or the related CNESS index deserve to attract much wider usage than they currently have. So far, the concept of NESS might have remained widely unaccepted because standard computer programs that are commonly used in biodiversity research do not compute this index.

Only very few studies have applied the combination of NMDS ordination and NESS index (Beck et al. 2002, Schulze & Fiedler 2003). In future studies, several alternative techniques should be explored in order to check the robustness of resulting patterns since no single method will perform perfectly under any given set of circumstances.

ACKNOWLEDGEMENTS

We would like to thank two anonymous reviewers, and Johannes Refisch for valuable comments on earlier drafts of the manuscript. Dirk Süssenbach was of great help in the field work. David J. Carter, Malcolm J. Scoble and Linda M. Pitkin (The Natural History Museum, London) and Axel Hausmann (Zoologische Staatssammlung, Munich), kindly allowed access to reference collections under their care. Giovanni Onore and Christoph L. Häuser supported the project logistically. Adrienne Hogg helped to database moths and improved the text linguistically. The Ministerio de Medio Ambiente del Ecuador granted research permits, Nature and Culture International (Loja) allowed access to the study sites, and the Deutsche Forschungsgemeinschaft financed our work (Fi 547/5-1, FOR 402/1-1 TP 15).

LITERATURE CITED


