

# Effects of taxonomic and numeric resolution on the ability to detect ecological patterns at a local scale using stream macroinvertebrates

Adriano S. Melo<sup>1</sup>

Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil

With 3 figures and 3 tables

**Abstract:** The increasing demand for methods of rapid stream bioassessment has stimulated the evaluation of data simplification. In particular, these studies have assessed how much power is lost when species/morphospecies identification is replaced by family identifications or use of EPT (Ephemeroptera, Plecoptera and Trichoptera) taxa only. A second simplifying factor commonly evaluated is the use of presence/absence data instead of density. These simplifications have provided valid results in most cases where differences among groups are large, particularly in studies comparing impacted vs. non-impacted stream sites and ecological studies involving large spatial scales. Here I evaluate whether data simplification, both in terms of taxonomic (families, morphospecies of EPT) and numeric (presence/absence) resolutions, is valid for ecological studies done at local scales, where differences among groups are subtle. Datasets used are derived from a five-year study of five stream sites situated in a catchment in south-east Brazil. Streams were sampled twice a year, in the rainy (summer) and dry (winter) seasons. I used Analysis of Similarity (ANOSIM) to evaluate if differences i) among stream sites and ii) between seasons within a stream site, revealed by using the full data set (morphospecies, quantitative data), were also detected when using the simplified datasets. The effect of taxonomic resolution was not significant; the two simplified levels of this factor (morphospecies of EPT, families) were able to recover the same groups revealed by the full dataset. However, the use of presence/absence data had a strong negative effect on the ability to distinguish groups, particularly when differences were small (between seasons within a stream site). The success in recovering groups using simplified taxonomic data agrees with previous evaluations done using datasets from applied fields and those from ecological studies involving large spatial scales. However, in contrast to results observed in applied and large-scale studies, use of simplified data quantification in local datasets resulted in significant loss of infor-

<sup>1</sup> **Author's address:** Dep. Ecologia, IB, Universidade Federal do Rio Grande do Sul. Caixa Postal 15007, 91501-970, Porto Alegre, RS, Brazil;  
E-mail: adrimelo@ecologia.ufrgs.br

mation. I suggest that the use of family identifications or morphospecies of EPT are reliable alternatives to the use of species/morphospecies in ecological studies at a local scale.

**Key words:** stream, aquatic insects, benthic invertebrates, morphospecies, data quantification, biological monitoring.

## Introduction

The ability of studies to detect patterns in community structure is strongly dependent on the quality of data available. For example, conclusions in stream benthic studies may differ when using different data quantification, taxonomic resolution of identifications, sample size, and inclusion or exclusion of rare species (CAO et al. 1998, LENAT & RESH 2001).

A number of recent studies dealing with stream benthos have evaluated the influence of data quality on the ability of protocols to distinguish impacted from reference/control sites (RESH & MCELRAVY 1993). Interest in this issue relates to the continuous search of ways to reduce costs, allowing in some cases inclusion of an increased number of sites (WARWICK 1993, LENAT & RESH 2001, BRADLEY & ORMEROD 2002). Despite current opposing views on specific topics (e. g. fixed-area vs. fixed-count of organisms; BARBOUR & GERITSEN 1996, COURTEMANCH 1996, VINSON & HAWKINS 1996), some convergent results have emerged. LENAT & RESH (2001) predict that family-level identifications might be useful in situations where differences between sites are large, but should fail in the distinction of sites that differ only slightly. In fact, studies assessing the distinction of impacted from reference assemblages have shown that identification to species-level does not add much information when compared to genus- or family-level identifications (BOWMAN & BAILEY 1997). In the field of marine biology, even identifications to phylum-level have produced similar patterns to those detected by species-level data (WARWICK 1993). Similarly, MARCHANT et al. (1995) used a dataset encompassing a wide geographic area to evaluate the effects of reducing the degree of detail in the original data. The results were similar to those observed in the applied studies cited before; patterns were still evident using the simplified datasets. A similar conclusion was attained by FURSE et al. (1984) in the classification of 268 sites in Great Britain. Compared to ecological studies done at local scales, where differences between groups are generally small, we could regard the two situations cited above as cases where differences between groups were large.

Few evaluations are available in the literature for cases in which differences between groups are small, as is often the case in ecological studies at a local scale. MARCHANT (1990) evaluated the effects of data simplification (family identifications and presence/absence data) in a study encompassing a

single river system. The effect of quantification was weaker than that observed when using family identifications. Field patterns recovered using the simplified datasets were generally similar to those revealed using the most complete dataset (species-level identification, quantitative data). However, weak differentiation observed between seasons was not revealed when using family level identifications. Despite the results of MARCHANT (1990) and the prediction of LENAT & RESH (2001) that the use of family-level identifications should not be sufficient to detect small differences between groups (e. g. differences between seasons in MARCHANT's study), rigorous tests on the validity of data simplification are still needed. Studies using family-level identifications are common in tropical areas, where complications due to poor taxonomic knowledge and high species richness occur (STOUT & VANDERMEER 1975). In South America, identification of Trichoptera, for example, even to genus level is not always possible, as some genera are only known from adult specimens (ANGRISANO 1995, FLINT et al. 1999) and discoveries of previously undescribed genera still appear in the literature (e. g. HARRIS & HOLZENTHAL 1997, HOLZENTHAL 1997, HOLZENTHAL & PES 2004). In extreme cases, even new family records for biogeographic provinces have been published (HOLZENTHAL 1997, HUAMANTINCO & NESSIMIAN 2003). As a consequence, many ecological studies carried out in these areas only use family-level identifications of the macroinvertebrate fauna (CLETO-FILHO & WALKER 2001, BUENO et al. 2003, JACOBSEN 2003), or families of EPT (Ephemeroptera, Plecoptera, Trichoptera) (BISPO & OLIVEIRA 1998, OLIVEIRA & BISPO 2001).

In this study, I used a large quantitative dataset sorted to morphospecies to evaluate whether simplifying data quantification (presence/absence) or taxonomic resolution (family-level or morphospecies of EPT) affects distinctions of two levels of grouping i) samples collected in different stream sites and ii) samples collected in different seasons within a stream site. All stream sites were in the same catchment and physical-chemical differences among sites and also between seasons within sites were small. As a consequence, differences in community structure among stream sites and between seasons within sites were assumed to be small as well. In contrast to studies available in the literature, mostly comparing distinct groups (e. g. impacted vs. reference sites), I evaluate whether data simplifications are useful in situations where differences among groups are small.

## Methods

### Data collection and processing

The data were derived from a five-year study of macroinvertebrates in five stream sites and two seasons (summer and winter), totalling 50 samples.

Stream sites are located in the Rio do Carmo catchment (24° 18' S, 48° 25' W), Parque Estadual Intervales, Iporanga, São Paulo state, Brazil. The vegetation of the catchment is Tropical Ombrophilous Submontane-montane Forest, known locally as Atlantic Rain Forest. The conservation status of the region is good and the catchment does not include any point source of pollution.

Streams differed from each other in size. Site 1 was a 1st order stream, 0.5–1 m in width and with a drainage basin area of 0.50 km<sup>2</sup>; the streambed was composed mainly of sand and stones (10–40 cm in size), and the reach studied had several pools and small waterfalls formed by fallen trees and accumulated twigs, leaves, and sand. Site 2 was 2nd order, 2.5–3.5 m wide and with a drainage basin area of 1.81 km<sup>2</sup> while site 3 was 3rd order, 3–4 m wide and with a drainage basin area of 3.62 km<sup>2</sup>; at both sites, the stream bottoms were predominantly composed of stones and boulders, with few pools and waterfalls. Sites 4 and 5 were fourth order streams, whose drainage basin areas were 25.19 km<sup>2</sup> and 25.69 km<sup>2</sup>, respectively, in both sites, the width ranges were 9–11 m and like sites 3 and 4, the streambeds were composed mainly of stones and boulders, with no pools and waterfalls in the studied reaches. Stream sites 1–5 represent sites 1, 4, 6, 8, and 9 respectively in a previous study, where additional information on physical characteristics of streams, invertebrate assemblages and a map are provided (MELO & FROELICH 2001).

The mean annual precipitation in the area is ~1700 mm, unevenly distributed across two seasons: one wet (130–270 mm/mo) and warm (15–30 °C) from September to March and another dry (60–95 mm/mo) and cold (0–25 °C) from April to August. Samples were taken twice a year, in February–March and July–August, during the period of 1997–2001. The timing of collection corresponds to the end of the (austral) summer and winter, respectively. Henceforth, samples accordingly will be called summer and winter samples.

Each sample was composed of 25 sampling units, each one consisting of all macroinvertebrates associated with a single stone (~18 cm). Sampling was done using a U-net sampler (SCRIMGEOUR et al. 1993) in riffles. Material collected in the net was transferred to a white tray and all visible invertebrates were removed and fixed in ethanol. Stones were examined for attached individuals.

Because of the poor taxonomic knowledge of aquatic macroinvertebrates in Brazil, particularly for immature insects, individuals were identified to family level and then separated into morphospecies, aided by comparisons to a reference collection, pictures and lists of diagnostic characteristics. As more individuals of morphospecies were observed, some initial morphospecies were split in two, and all previously assigned material was rechecked, while some others were pooled. First instar larvae of some insect groups were not included in the dataset, as they do not present diagnostic structures (e. g. Simuliidae, Perlidae). I believe that most morphospecies correspond to biological species, an assertion supported by previous work on terrestrial groups showing a good correspondence between morphospecies sorted by non-specialists and species identified by specialists (OLIVER & BEATTIE 1996). However, this certainly was not true in a few groups (Diptera: Simuliidae and Brachycera). Twenty-two individuals (0.05 % of total) belonged to muscoid Diptera (Brachycera), and these individuals were pooled and used in analyses as a single “family”. Because of difficulties in separating Chironomidae (Diptera) and Acari into morphospecies, these taxa were not included.

## Data analysis

I first explored the relationships among the 50 samples with a classification analysis, using the UPGMA linkage method applied to a Bray-Curtis (or Sørensen) similarity matrix. The similarity data were computed after transforming abundance to  $\text{Log}_{10}(x + 1)$ . The analysis revealed that samples were first grouped within stream sites and, in some cases, a second level of grouping within sites separated summer and winter samples. I used these two grouping-levels to evaluate the effects of data resolution.

I assessed the validity of two simplification factors: i) quantification of data, with two levels ( $\text{log}_{10}(x + 1)$ , presence/absence) and ii) taxonomic resolution, with three levels (morphospecies, morphospecies of EPT only, families). Factors were crossed and thus six datasets were produced. The most detailed dataset ( $\text{log}_{10}(x + 1)$  and morphospecies) was designated as the reference in the comparisons.

I used Analysis of Similarity (ANOSIM; CLARKE & GREEN 1988) to evaluate if the two levels of groupings (stream sites, seasons) were apparent after data simplification. ANOSIM relates distance ( $= 1 - \text{similarity}$ ) within groups to distances between groups. In the case of no difference between groups (the null hypothesis), distances within and between groups should be similar. In contrast, if a distinct group factor is present, distances within groups should be smaller than those between groups. A statistic called *R*, ranging from  $-1$  to  $+1$ , measures the distinctiveness of the grouping. *R*-values around 0 indicate no distinct groups, while those close to  $+1$  denote perfect groups. The significance of the observed *R* statistic is evaluated by comparison with a distribution of *R*-values generated by shuffling the samples from the different groups and randomly forming new groups. The procedure to produce the *R* statistic under the null hypothesis was repeated 1000 times. Calculation of *R* is done using rank of the distances instead of their absolute values. ANOSIM was computed using the Bray-Curtis distance index, which works both on quantitative and qualitative (presence/absence) data. Calculations were carried out using functions *vegdist* and *anosim* of the *vegan* package (OKSANEN 2004), run under the R software (The R Development Core Team 2003).

The two grouping levels revealed initially by the cluster analysis (stream, season) were evaluated using the ANOSIM analysis. The *R* statistic and its associated *p*-value evaluated whether groupings were retained using the different combinations of data simplifications. Also, I further explored the effects of the two simplification factors (quantification and taxonomic resolution) by subjecting the *R*-values obtained to evaluate "season groupings" to a two-way analysis of variance. The *R*-values obtained for all five streams were considered replicates in the analysis of variance.

## Results

### Summaries of the datasets and relationships among samples

The full dataset contained 40,490 individuals distributed among 185 morphospecies, 91 of them included in EPT orders (Table 1). A total of 50 families were identified. Samples contained in average 809.5 individuals and 53.5 mor-

**Table 1.** Abundance, number of morphospecies, number of morphospecies of EPT (Ephemeroptera, Plecoptera and Trichoptera) and number of families found in five sites and two seasons. Unbracketed numbers refer to means calculated from five samples collected between 1997–2001. Numbers in round brackets indicate the standard deviation. Numbers in square brackets indicate the proportion of EPT or families in relation to the mean number of morphospecies in the site.

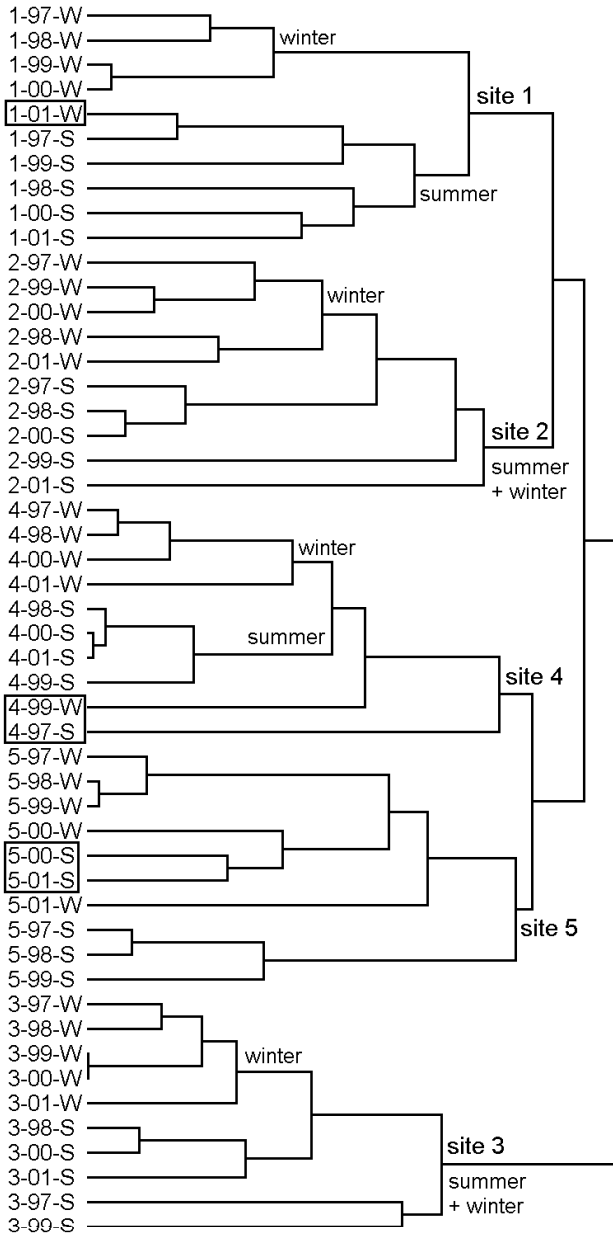
		Individuals	Morpho-species	Morphospecies of EPT	Families
Site 1	summer	712 (169)	54.0 (6.0)	31.8 (5.6) [0.59]	29.6 (2.3) [0.55]
	winter	626 (125)	51.6 (2.8)	31.2 (2.7) [0.60]	29.2 (2.2) [0.57]
Site 2	summer	1131 (318)	61.2 (6.1)	37.4 (3.4) [0.61]	30.2 (2.7) [0.49]
	winter	681 (142)	55.6 (5.0)	36.6 (4.6) [0.66]	28.2 (1.5) [0.51]
Site 3	summer	693 (269)	60.2 (5.7)	39.0 (5.2) [0.65]	28.4 (2.6) [0.47]
	winter	933 (202)	57.8 (7.8)	37.6 (5.6) [0.65]	28.0 (3.1) [0.48]
Site 4	summer	792 (254)	47.0 (4.7)	25.6 (2.3) [0.54]	26.6 (1.1) [0.57]
	winter	856 (242)	51.2 (5.6)	29.6 (3.0) [0.58]	28.6 (2.9) [0.56]
Site 5	summer	794 (349)	47.2 (2.9)	28.8 (1.8) [0.61]	24.0 (0.7) [0.51]
	winter	877 (239)	48.8 (5.9)	29.4 (5.0) [0.60]	25.6 (1.9) [0.52]
Mean		809.5	53.5	32.7 [0.61]	27.8 [0.52]
Total		40,490	185	91	50

phospecies. Nearly 60 % of the morphospecies belonged to EPT orders and each family contained on average two morphospecies at a given site and season (Table 1).

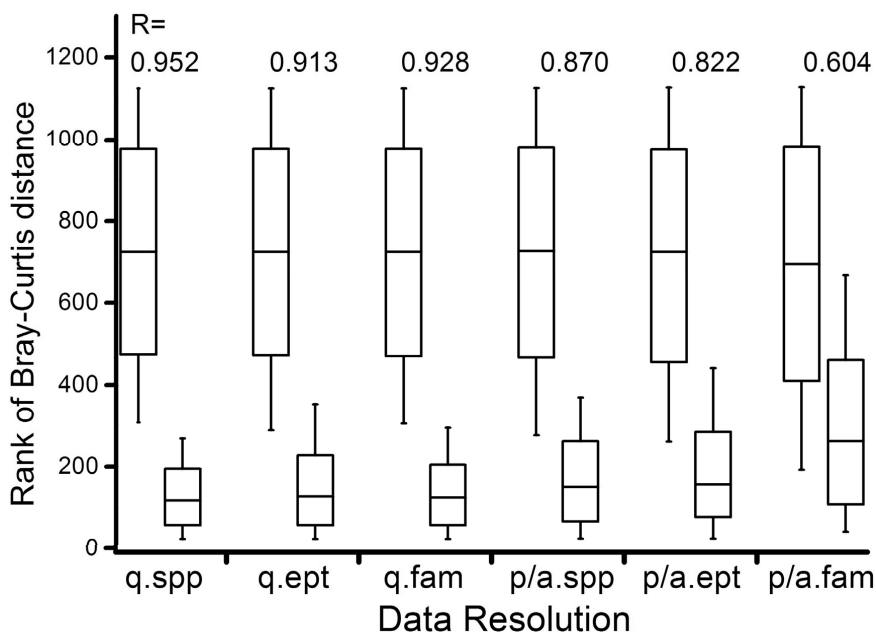
Similarity among the 50 samples used in the study is depicted by means of a dendrogram produced by UPGMA and using morphospecies and quantitative data (Fig. 1). Samples collected at the same stream site were always grouped. Within each stream group, there was a clear trend for samples to be arranged according to season. This was clearest at stream sites 2 and 3, and to a lesser degree in site 1. Winter samples tended to group at a higher similarity level than summer samples, in some cases forming a compact subgroup within a summer + winter group.

**Stream grouping**

Stream groups revealed using the UPGMA classification were distinctively recovered using ANOSIM on the dataset including quantitative data and morphospecies identification. The box-plot bars for distance ranks between- and within-groups shown in Fig. 2 are clearly separated. Distances within-groups were small when compared to distances between-groups. When using quantitative data, simplification in taxonomic resolution did not affect the distinctiveness of distance bars for between- and within-groups. However, taxonomic resolution was important when qualitative data were used. Distinctiveness of



**Fig. 1.** Classification of samples collected in five stream sites and two seasons during a five-year period (1997–2001) by unweighted pair-group method using arithmetic averages (UPGMA) and Bray-Curtis (or Sørensen) distance. 1, 2, 3, 4 and 5 = stream sites. 97, 98, 99, 00 and 01 = years. S = summer, W = winter. Terms indicate consistent groups, and boxes “misclassified” samples. Cophenetic correlation between dendrogram and similarity matrix is 0.8609.



**Fig. 2.** Box-plot of rank distances between- and within-groups used to evaluate consistency of stream-grouping by Analysis of Similarity (ANOSIM). The horizontal lines in the box denote the 25th, 50th, and 75th percentile values. The error bars denote the 5th and 95th percentile values. Six datasets were evaluated, differing in terms of taxonomic resolution and data quantification. spp = organisms identified to morphospecies. ept = organisms belonging to orders Ephemeroptera, Plecoptera and Trichoptera identified to morphospecies. fam = families. q = quantitative ( $\log_{10}(x + 1)$ ) data. p/a = qualitative (presence/absence) data. The left hand bar of each dataset corresponds to distances (rank) between-groups and the right hand bar to distances within-groups. The more separated the bars, the more distinct are the groups. The 50 samples composing each dataset produced a distance matrix containing 1225 values ( $= N(N-1)/2$ , where  $N = 50$  samples). The between-groups bar includes 1000 distance values and the within-groups bar 225 values ( $= 5 \text{ sites} * n(n-1)/2$ , where  $n = 10$  samples). Values on top of each pair of bar represent the R statistic of ANOSIM. In all six cases  $p < 0.001$ , indicating that the five stream-groups are consistent.

groups was reduced when qualitative datasets were identified to EPT and, more severely, when identification was to family. Despite the decrease in distinctiveness of groups using qualitative data, separations among stream-groups were still significant (in all six cases  $R > 0.6$  and  $p < 0.001$ ) (Fig. 2).

### Season grouping

The UPGMA classification indicated that, at least for some sites (streams 2, 3), groups corresponding to summer and winter samples were present (Fig. 1).



**Table 2.** R-values of the Analysis of Similarity (ANOSIM) used to test the consistency of season grouping using three different taxonomic resolutions and two data quantification. Analyses were done within each of the five stream sites. Values in brackets indicate the probability that the two assigned groups (summer, winter) did not differ statistically when compared to distribution of *p*-values generated by 1000 random permutations of samples between groups.

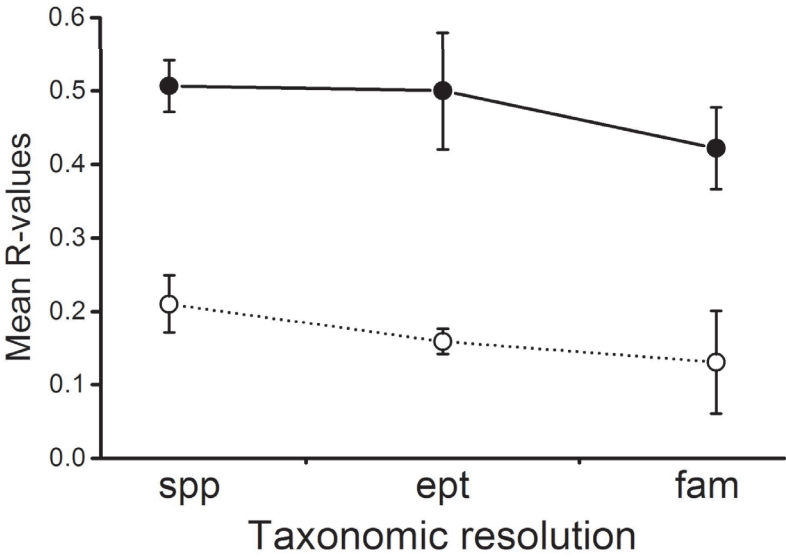
	Quantitative			Qualitative (presence/absence)		
	Morpho-species	Morpho-species of EPT	Families	Morpho-species of EPT	Morpho-species	Families
Site 1	0.552 (0.012)	0.564 (0.018)	0.276 (0.054)	0.228 (0.076)	0.172 (0.606)	-0.064 (0.606)
Site 2	0.564 (0.009)	0.676 (0.011)	0.448 (0.011)	0.200 (0.099)	0.212 (0.086)	0.122 (0.131)
Site 3	0.392 (0.012)	0.524 (0.007)	0.316 (0.036)	0.070 (0.306)	0.106 (0.231)	0.028 (0.376)
Site 4	0.456 (0.011)	0.200 (0.046)	0.492 (0.003)	0.248 (0.052)	0.146 (0.204)	0.314 (0.031)
Site 5	0.568 (0.007)	0.536 (0.014)	0.576 (0.011)	0.304 (0.011)	0.158 (0.152)	0.254 (0.069)

However, the distinctiveness of these groups was weak when compared to groups formed by streams. Accordingly, distance bars for between- and within-groups in ANOSIM overlapped to some degree in four out of the five sites when using quantitative data and morphospecies identification (results not shown). For the quantitative datasets, the effect of taxonomic resolution was not homogenous among the five stream sites. For stream 5, R statistics (which measure distinctiveness of groups) were similar among the three types of taxonomic resolution (Table 2). The lowest R-values were found for identifications to the family level in streams 1, 2 and 3. However, in stream 4 the identification to the family level produced the highest R-value. Except for one case, groups were statistically distinct ( $p < 0.05$ ) using the three taxonomic resolutions independent of stream site (Table 2).

Qualitative data sets including identifications to morphospecies produced the highest R in only two out of five stream sites. In two streams, best separations between distance bars were obtained using EPT identifications. In only two of the qualitative datasets (family for stream 4, morphospecies for stream 5) were groups formed by season statistically distinct ( $p < 0.05$ ) (Table 2).

### Effects of the two simplification factors

The effects of the two simplification factors (taxonomic resolution, data quantification) were further contrasted using the R statistics obtained in the ANO-



**Fig. 3.** Mean and 1 SE of R-values obtained from the season-grouping Analysis of Similarity. Closed circles correspond to quantitative datasets and open circles to qualitative datasets. spp = data identified to morphospecies. ept = morphospecies of EPT (Ephemeroptera, Plecoptera and Trichoptera). fam = families.

**Table 3.** Two-factor Analysis of Variance of the R statistic produced by Analysis of Similarity (ANOSIM) among season-grouping (values of Table 2). The taxonomic factor has three levels: morphospecies, morphospecies of EPT (Ephemeroptera, Plecoptera and Trichoptera) and families. The quantification factor has two levels: quantitative ( $\log_{10}(x + 1)$ ) and qualitative (presence/absence). Stream sites are used as replicates in the analysis.

	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>
Taxonomy (T)	2	0.0173	1.196	0.320
Quantification (Q)	1	0.7183	49.650	<0.001
Interaction (T * Q)	2	0.0019	0.132	0.877
Residual	24	0.1319		

SIM applied to season (Fig. 3). Reponses to the factors were independent of each other because the ANOVA interaction term was not significant (Table 3). Although there was a decrease in the R statistics with decreasing taxonomic resolution, the difference among levels was low and not significant (Table 3). In contrast, the effect of quantification was high and significant ( $p < 0.001$ ) (Table 3).

## Discussion

Changes in quantification had a stronger effect than changes in taxonomic resolution. This is indicated by the recovery of groups (streams or seasons) irrespective of taxonomic resolution when using quantitative data. Use of qualitative data was enough to recover groups when differences among them were large (stream grouping), but failed when differences were subtle (season grouping).

The dataset encompasses a small catchment (see map in MELO & FROEHLICH 2001) and the largest geographic distance between two streams (sites 1 and 5) is only 7.5 km. Streams are similar in terms of riparian vegetation and geologic composition, and are free of human disturbance. Size is the single, although only partially, distinguishing factor. Stream site 1 possesses the smallest discharge, stream sites 2 and 3 intermediate discharges, and sites 4 and 5 the largest discharges. If we were to classify them according to discharge, we would have three groups. Taking such similarities among stream sites into account, it is reasonable to assume that i) if they differ, this probably would occur among sites differing in size and ii) differences would be detected only when using a detailed sampling and processing procedure (i. e. quantitative data and morphospecies identification). It is quite surprising, therefore, to observe that i) even streams similar in size differed and that ii) simplified data were enough to reveal such differences.

Previous evaluations have shown that use of families is enough to clarify relationships among streams subjected to different degrees of anthropogenic impact (HILSENHOFF 1988, GROWNS et al. 1995, WAITE et al. 2004). This is conceivable if one considers that differences among reference and impacted streams are generally large. If a given pollutant is enough to cause local extinction of a species, it is likely that it will have the same effect on most of the other confamilial species (BOWMAN & BAILEY 1997). Accordingly, the same pattern will emerge no matter if one employs species or family identifications. The same line of reasoning may account for the success in use of family identifications in studies encompassing a large geographic region (CORKUM 1989, MARCHANT et al. 1995). Areas restrictive to a species are likely to be restrictive to other species of the family as well. In the present study nearly the same families are present in the five sites and the two seasons. Within families, most common morphospecies occur in all sites (and seasons), while only a minority are restricted to streams of a given size (MELO & FROEHLICH 2001). LENAT & RESH (2001) predict that patterns in data from sites with only slight differences would only be revealed when using species/morphospecies identifications. Contrary to their expectation, analysis using family data performed nearly equally well as did analyses using morphospecies data, indicating that responses of morphospecies within families were homogeneous. This may be

due to the fact that each family contained on average only two morphospecies (Table 1). Use of family was therefore equivalent to pooling pairs of morphospecies.

Most of the species belonging to orders EPT are thought to be more susceptible to anthropogenic impacts than species in other insect orders. As such, the number of EPT in a site or the proportion of species belonging to EPT is commonly employed as a metric in bioassessment programs (RESH & JACKSON 1993). In our study, where the distinguishing factors among samples were stream size and season (natural factors), we have no reason to suspect a drastic variation in numbers of EPT among levels of the two factors. In contrast to applied studies, therefore, we would not expect the EPT datasets to reveal groups effectively. However, datasets comprising morphospecies of EPT performed equally well as the full dataset containing morphospecies of all groups. Similarly to the observations made when using families, morphospecies of EPT acted as a subgroup including nearly 60 % of a redundant group, the full dataset (Table 1). This is an important result, because use of EPT or families in a study containing many samples would result in a saving of time and money (WARWICK 1993).

Simplification in numeric resolution had a stronger effect than simplification in taxonomic resolution in the present study. This contrasts with applied or large-scale studies, where the inclusion of quantitative data does not greatly improve the recovery of patterns in data (MARCHANT 1990). In cases where differences among groups are generally large, presence or absence of a species is enough to characterize the site, e. g. impaired or not-impaired. In the present study, samples were obtained from essentially the same regional species composition. This is particularly the case for samples collected from the same stream but different seasons, and for samples collected in streams similar in size (sites 1 and 2; sites 4 and 5) (MELO & FROELICH 2001). Given the high similarity in species composition among samples in the present study, differences in abundance among samples became the most important distinguishing factor. Conclusions of applied or large-scale studies recommending the use of presence/absence data should thus not be extended to ecological studies at local scale.

MARCHANT (1990) evaluated the effect of qualitative and family-level identification on the robustness of ordination and classification of sites included in the same catchment. Patterns that emerged from the best dataset (species-level identification, quantitative data) were also generally recovered with the two simplified datasets. However, small seasonal differences among samples were not clearly detected with simplified data. In these cases, use of family identification (and quantitative data) performed less well than qualitative data (and species-level identification). The recovery of contrasting groups, and blurring of only slightly different groups using simplified datasets in the

study of MARCHANT (1990), is similar to the results found in the present paper. However, the two studies differ in respect to the relative importance of the two factors evaluated. In Marchant's study family identifications caused a greater negative effect than qualitative data, while the opposite occurred in the present evaluation.

This study reinforces previous studies, dealing with large spatial scales and those from applied fields, which indicate the validity of the use of low taxonomic resolution (e. g. family), and extends the conclusion to situations where differences among groups are small. However, it differs from these studies in its demonstration that use of presence/absence data should not be used in situations where presumed differences among groups are small.

## Acknowledgements

The field guides and the administrative staff at P. E. Intervalos were of great value during the fieldwork. RODRIGO PEREIRA introduced me to S<sup>+</sup> and R and assisted me in the first steps. GERHARD OVERBECK reviewed the English. LUIS M. BINI, LÁZLÓ ORLÓCI, and RICHARD MARCHANT critically read the manuscript. MICHAEL DOBSON improved the final manuscript. I received a student fellowship from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and fellowships and grants from Fundação de Amparo à Pesquisa do Estado de São Paulo (Fapesp procs 98/00590-0, 02/12538-0, and 03/10936-1). This work is part of the Biota/Fapesp – The Biodiversity Virtual Institute Program ([www.biota.org.br](http://www.biota.org.br)).

## References

- ANGRISANO, E. B. (1995): Insecta Trichoptera. – In: LOPRETTO, E. C. & TELL, G. (eds): Ecosistemas de aguas continentales: metodologías para su estudio. – SUR, La Plata, pp. 1199–1237.
- BARBOUR, M. T. & GERRITSEN, J. (1996): Subsampling of benthic samples: a defense of the fixed-count method. – *J. North Amer. Benthol. Soc.* **15**: 386–391.
- BISPO, P. C. & OLIVEIRA, L. G. (1998): Distribuição espacial de insetos aquáticos (Ephemeroptera, Plecoptera e Trichoptera) em córregos de cerrado do Parque Ecológico de Goiânia, Estado de Goiás. – In: NESSIMIAN, J. L. & CARVALHO, A. L. (eds): Ecologia de Insetos Aquáticos. – PPGE-UFRJ, Rio de Janeiro, pp. 175–189.
- BOWMAN, M. F. & BAILEY, R. C. (1997): Does taxonomic resolution affect the multivariate description of the structure of freshwater benthic macroinvertebrate communities? – *Can. J. Fish. Aquat. Sci.* **54**: 1802–1807.
- BRADLEY, D. C. & ORMEROD, S. J. (2002): Evaluating the precision of kick-sampling in upland streams for assessment of long-term change: the effect of sampling effort, habitat and rarity. – *Arch. Hydrobiol.* **155**: 199–221.
- BUENO, A. A. P., BOND-BUCKUP, G. & FERREIRA, B. D. P. (2003): Estrutura da comunidade de invertebrados bentônicos em dois cursos d'água do Rio Grande do Sul, Brasil. – *Rev. Bras. Zool.* **20**: 115–125.

- CAO, Y., WILLIAMS, D. D. & WILLIAMS, N. E. (1998): How important are rare species in aquatic community ecology and bioassessment? – *Limnol. Oceanogr.* **43**: 1403–1409.
- CLARKE, K. R. & GREEN, R. H. (1988): Statistical design and analysis for a ‘biological effects’ study. – *Mar. Ecol. Prog. Ser.* **46**: 213–226.
- CLETO-FILHO, S. E. N. & WALKER, I. (2001): Efeitos da ocupação urbana sobre a macrofauna de invertebrados aquáticos de um igarapé da cidade de Manaus/AM – Amazônia Central. – *Acta Amazon.* **31**: 69–89.
- CORKUM, L. D. (1989): Patterns of benthic invertebrate assemblages in rivers of north-western North America. – *Freshwat. Biol.* **21**: 191–205.
- COURTEMANCH, D. L. (1996): Commentary on the subsampling procedures used for rapid bioassessments. – *J. North Amer. Benthol. Soc.* **15**: 381–385.
- FLINT, O. S. Jr., HOLZENTHAL, R. W. & HARRIS, S. C. (1999): Catalog of the Neotropical caddisflies (Insecta: Trichoptera). – Ohio Biological Survey, Columbus.
- FURSE, M. T., MOSS, D., WRIGHT, J. F. & ARMITAGE, P. D. (1984): The influence of seasonal and taxonomic factors on the ordination and classification of running-water sites in Great Britain and on the prediction of their macro-invertebrate communities. – *Freshwat. Biol.* **14**: 257–280.
- GROWNS, J. E., CHESSMAN, B. C., MCEVOY, P. K. & WRIGHT, I. A. (1995): Rapid assessment of rivers using macroinvertebrates: case studies in the Nepean River and Blue Mountains, NSW. – *Aust. J. Ecol.* **20**: 130–141.
- HARRIS, S. C. & HOLZENTHAL, R. W. (1997): *Mejicanotrichia*, a new genus of micro-caddisflies from Mexico and Guatemala (Trichoptera: Hydroptilidae). – In: HOLZENTHAL, R. W. & FLINT, O. S. Jr. (eds): Proceedings of the 8th International Symposium on Trichoptera. – Ohio Biological Survey, Columbus, pp. 129–137.
- HILSENHOFF, W. L. (1988): Rapid field assessment of organic pollution with a family-level biotic index. – *J. North Amer. Benthol. Soc.* **7**: 65–68.
- HOLZENTHAL, R. W. (1997): The caddisfly (Trichoptera) family Atriplectididae in the Neotropics. – In: HOLZENTHAL, R. W. & FLINT, O. S. Jr. (eds): Proceedings of the 8th International Symposium on Trichoptera. – Ohio Biological Survey, Columbus, pp. 157–165.
- HOLZENTHAL, R. W. & PES, A. M. O. (2004): A new genus of long-horned caddisfly from the Amazon basin (Trichoptera: Leptoceridae: Grumichellini). – *Zootaxa* **621**: 1–16.
- HUAMANTINCO, A. A. & NESSIMIAN, J. L. (2003): A new species of *Antactoercia* Ulmer, 1907 (Trichoptera: Limnephilidae) from Southeastern Brazil. – *Aquat. Insects.* **25**: 225–231.
- JACOBSEN, D. (2003): Altitudinal changes in diversity of macroinvertebrates from small streams in the Ecuadorian Andes. – *Arch. Hydrobiol.* **158**: 145–167.
- LENAT, D. R. & RESH, V. H. (2001): Taxonomy and stream ecology – The benefits of genus- and species-level identifications. – *J. North Amer. Benthol. Soc.* **20**: 287–298.
- MARCHANT, R. (1990): Robustness of classification and ordination techniques applied to macroinvertebrate communities from the La Trobe River, Victoria. – *Aust. J. Mar. Freshwat. Res.* **41**: 493–504.

- MARCHANT, R., BARMUTA, L. A. & CHESSMAN, B. C. (1995): Influence of sample quantification and taxonomic resolution on the ordination of macroinvertebrate communities from running waters in Victoria, Australia. – Mar. Freshwat. Res. **46**: 501–506.
- MELO, A. S. & FROELICH, C. G. (2001): Macroinvertebrates in neotropical streams: richness patterns along a catchment and assemblage structure between 2 seasons. – J. North Am. Benthol. Soc. **20**: 1–16.
- OKSANEN, J. (2004): Vegan – Community ecology package. [1.6–4]. <http://cran.r-project.org>.
- OLIVEIRA, L. G. & BISPO, P. C. (2001): Ecologia de comunidades das larvas de Trichoptera Kirby (Insecta) em dois córregos de primeira ordem da Serra dos Pirineus, Pirenópolis, Goiás, Brasil. – Rev. Bras. Zool. **18**: 1245–1252.
- OLIVER, I. & BEATTIE, A. J. (1996): Invertebrates morphospecies as surrogates for species: a case study. – Conserv. Biol. **10**: 99–109.
- RESH, V. H. & JACKSON, J. K. (1993): Rapid assessment approaches to biomonitoring using benthic macroinvertebrates. – In: ROSENBERG, D. M. & RESH, V. H. (eds): Freshwater biomonitoring and benthic macroinvertebrates. – Chapman and Hall, New York, pp. 195–233.
- RESH, V. H. & MCELRAVY, E. P. (1993): Contemporary quantitative approaches to biomonitoring using benthic macroinvertebrates. – In: ROSENBERG, D. M. & RESH, V. H. (eds): Freshwater biomonitoring and benthic macroinvertebrates. – Chapman and Hall, New York, pp. 159–194.
- SCRIMGEOUR, G. J., CULP, J. M. & GLOZIER, N. E. (1993): An improved technique for sampling lotic invertebrates. – Hydrobiologia. **254**: 65–71.
- STOUT, J. & VANDERMEER, J. (1975): Comparison of species richness for stream-inhabiting insects in tropical and mid-latitude streams. – Amer. Nat. **109**: 263–280.
- The R Development Core Team (2003): R – A Language and Environment. [1.7.0]. <http://cran.r-project.org>.
- VINSON, M. R. & HAWKINS, C. P. (1996): Effects of sampling area and subsampling procedure on comparisons of taxa richness among streams. – J. North Amer. Benthol. Soc. **15**: 392–399.
- WAITE, I. R., HERLIHY, A. T., LARSEN, D. P., URQUHART, N. S. & KLEMM, D. J. (2004): The effects of macroinvertebrate taxonomic resolution in large landscape bioassessments: an example from the Mid-Atlantic Highlands, U. S. A. – Freshwater Biol. **49**: 474–489.
- WARWICK, R. M. (1993): Environmental impact studies on marine communities: pragmatic considerations. – Aust. J. Ecol. **18**: 63–80.

Submitted: 14 December 2004; accepted: 20 July 2005.