7

Analysis and Interpretation of Benthic Macroinvertebrate Surveys

Richard H. Norris and Arthur Georges

7.1. Introduction

Biologists often find themselves in the position of having to use patterns of distribution and abundance of organisms to detect environmental change and to infer the cause of the change by associating changes in biological variables with corresponding changes in physicochemical variables (Norris and Georges, 1986; Barmuta, 1987). Inevitably, biological and environmental correlates of water quality are compared across sites and times, or against set standards, to assess impacts of disturbances or management initiatives, to develop models useful for prediction, or to establish cause and effect. Study designs should facilitate the making of these comparisons through the collection of relevant data, elimination of confounding effects, and selection of appropriate analyses.

Biologists working with benthic macroinvertebrates have long been aware of the problems of variability in what they measure (e.g., Needham and Usinger, 1956) and have emphasized the need to account for variability of benthic macroinvertebrate data in their sampling designs (Downing, 1979; Resh, 1979; Allan, 1984; Morin, 1985; Norris and Georges, 1986; Canton and Chadwick, 1988; Resh and McElravy, Chapter 5). These sampling designs usually include some level of replication to enable subsequent analyses to be performed (see Resh and McElravy, Chapter 5).

Several approaches have been developed to cope with the need for sometimes large numbers of replicate collections of macroinvertebrates (Needham and Usinger, 1956; Chutter and Noble, 1966; Downing, 1979; Resh, 1979; Allan, 1984) and the need for rapid return of results (Cairns and Van Der Schalie, 1980). Many indices have been created and used in benthic monitoring studies (see reviews by Cairns et al., 1972; Fager, 1972; Poole, 1974; Washington, 1984; Hellawell, 1986), but many workers have avoided the problem altogether by relying on direct measurement or by presenting data on number of individuals or taxa with little or no analysis (lotic environ-

7.2. Elucidating Mechanisms

A problem often encountered in assessing water quality concerns the effects of a range of environmental variables on species that comprise a biological community. As our knowledge and understanding of aquatic systems improve, an increased need also exists for assessment of the effects of interrelationships of factors having minor or subtle outcomes on such communities. Studies assessing water quality often provide large data sets with many variables. The relationships of cause and effect between the variables may be complex and difficult to analyze, often needing multivariate methods (Green, 1979). Approaches used for the collection of physicochemical data have been different from those used for the collection of biological data, and this has resulted in difficulties in associating the two types of data during analysis. These broad divisions in data types need not be considered separately (Norris and Georges, 1986).

When interpreting the relationships between biological and physicochemical data, it is important to base the interpretations on real biological properties that relate to the environment rather than just on statistical interpretations (Taylor, 1980; Anderson et al., 1982; Fryer, 1987). Strictly speaking, causal relationships only can be determined through direct experimental work in which the features of concern are tested under controlled conditions (see Cooper and Barmuta, Chapter 11).

7.3. Data Quality Assurance and Exploratory Analysis

7.3.1. Precision and Analysis

Assurance of the quality of data from sampling involves sound measurement procedures and an understanding of sampling theory. For data to be useful, knowledge of the variability of repeated measurements is required. If single samples of benthic macroinvertebrates, or a single biological indicator such as a diversity index, are highly variable and differ at two stations, the researcher is faced with an ambiguity: do they differ because values of the indicator actually differ at the two stations, or do they differ solely because any two samples would be expected to differ as much, even if collected at the same time and from the same station? Environmental variability is a fundamental problem facing those interested in assessing changes in water quality through space and time. High environmental variability and logistical constraints on sample collection and analysis often may result in data that are too variable to demonstrate the impact of a disturbance, or management initiatives, on water quality (see also Resh and Jackson, Chapter 6).

Conclusions that water quality at particular sites or times actually differs

ranges (e.g., a species with the incorrect number allocated to it during sorting or data entry). Further checks will be necessary.

2. Calculation of means and standard deviations for subsets (possibly by date or site) of the data is useful because aberrantly high or low means, or aberrantly high standard deviations, indicate errors.

3. Univariate checks for outliers can be performed by testing for values more than three standard deviations from overall means. Since counts rarely are distributed normally, often being highly skewed, this procedure is bound to produce a high number of suspicious observations. For skewed data, calculating percentiles and scrutinizing values that fall above the ninety-fifth or ninety-ninth percentiles may be preferable.

4. The frequency of occurrence of particular species in a set of replicates from particular sites can be cross-checked independently. These occurrences should fall within specified acceptable limits based on previous work. The calculation of cumulative numbers for the different species collected can be compared with the total numbers recorded for each replicate.

5. Calculation of ratios of the numbers of selected species, which will be site- or subcatchment-specific, will help to locate which counts of particular species, or identifications, are in error.

6. Bivariate checks for outliers, using regressions of the variable under scrutiny with other related variables in the data set, may enable identification of erroneous data. For example, numbers of animals collected may be related linearly to distance from a point source of organic pollution. If sources are relatively constant, then departures of three or four standard deviations from the regression will highlight procedural or real environmental problems.

7. Plots of variables against each other or against time or distance from a point of impact are useful for visual checks on whether the data follow logical sequences.

8. A check on total counts is useful for verifying that counts of individual species have been entered correctly. Data sheets may consist of a list of macroinvertebrate species, and the data may be counts of individuals in each species. The total number of animals counted, although redundant because it can be calculated easily by statistical computer packages, enables a valuable check against mispunching of counts of individual species that are otherwise very difficult to verify.

man, 1972), a so-called biotic index (Washington, 1984). All biological measures are designed to capture some aspect of the community at a particular site or time.

Nearly 40% of all lake and river studies surveyed in Chapter 5 (Table 5.6) used some type of index. The most common one was the Shannon (1948) index, followed by percent similarity (Whittaker, 1952) in lentic studies and the Simpson (1949) diversity index in lotic studies. Reviews of indices relative to aquatic systems can be found in Cairns et al. (1972), Washington (1984), Hellawell (1986, p. 430), Abel (1989), and Johnson et al. (Chapter 4). According to Green (1979), Auclair and Goff's (1971) field data on upland forests provide an excellent example of spatial and temporal patterns of variation, and covariation, of all commonly used diversity indices.

Indices may be used for the following reasons: (1) they are seen as a useful way to condense complex data and thus aid interpretation (Wilhm, 1972; Hellawell, 1986); (2) people with little biological expertise can understand them easily (Cairns et al., 1968; Wilhm and Dorris, 1968) and can gather the data to create some of them (e.g., Cairns et al., 1968); (3) they are of more general value than physical and chemical measures (Hellawell, 1986); (4) they allow comparisons of sites or times where collections have been made using different sample sizes, methods, or habitats; and (5) their data needs are seen as relatively less expensive than other more traditional statistically based approaches. Clearly, the use of indices is popular in studies of benthic macroinvertebrates. Therefore, it is appropriate to consider data needs of these indices and the assumptions on which they are based, because these will affect data analysis and interpretation.

In his extensive review of indices, Washington (1984) identified three groups: diversity or community structure indices, biotic indices (which includes most of the rapid assessment indices discussed in Resh and Jackson, Chapter 6), and similarity indices.

7.4.1.1. Diversity Indices

Diversity indices usually require a count of the total number of individuals and a total count for each of the taxa. The taxa need to be separated but not necessarily identified. Separation is often at the species level, but it is sometimes at the generic or family level (Hughes, 1978).

The combination of abundance and richness in a diversity index supposedly indicates the state of the community. It seems to be generally accepted that values of most indices decrease with decreasing water quality. Also, low diversity supposedly indicates a stressed community that tends to be unstable (Goodman, 1975).

Washington (1984) divided the diversity indices he reviewed into eight

quency distributions of species abundance, including the "characteristic" of Fisher et al. (1943) and Yule (1944) and the log-normal curve of Preston (1948) (Washington, 1984). The log-normal distribution is probably the most widely used of these approaches. Some doubt exists as to the biological meaning of frequency distributions, and no consideration seems to have been given to how environmental stress (including pollution) will affect the relationship. The indices developed by this procedure only should be used when the curves are a good fit to the data (Pielou, 1975), which may be difficult to satisfy because the relationship is likely to change with environmental stress. Krebs (1985) and Goodman (1975) have discussed the log-normal distribution particularly in relation to estimating the total number of taxa, or the number of rare taxa, at a site. This suggests the existence of a true biological relationship (i.e., niche subdivision; Sugihara, 1980) represented by the log-normal distribution of species frequency curves.

The most widely used diversity indices are those derived from information theory (e.g., Shannon, 1948). Washington (1984) provided a full discussion of this type of index and pointed out the rather tenuous biological links that have been attributed to them. Indices from information theory purport to measure "uncertainty" in the data, which may be considered to be the same as "information content," and consequently "diversity" (Washington, 1984). However, a direct link of biological relevance between these factors is doubtful (Goodman, 1975). The Shannon Index (H') reaches its maximum value when all species are distributed evenly. Biologically, this is assumed to be the most desirable situation, although it contradicts the evidence provided by the log-normal distribution for many different communities (Goodman, 1975; Krebs, 1985). Hurlbert (1971) showed that many indices derived from information theory are correlated because they use the same variables in their calculation, and Krebs (1985, p. 523) concluded that "in practice it seems to matter very little which of these different measures of diversity we use." Goodman (1975) pointed out that these indices may be affected by the degree of clumping, problems of different body size in the organisms collected, and different habitat needs. Hughes (1978) listed six factors, other than pollution, that affected these diversity indices, including sampling method, sample size, depth of sampling, duration of sampling, time of year, and taxonomic level used. Thus, diversity indices based on information theory should be interpreted and compared with caution because their values will depend on study design.

Diversity indices often are based on ecological theories, such as the diversity/stability hypothesis (Goodman, 1975) or competitive interaction (Hurlbert, 1971). As such, authors promoting their use argue that real ecological properties are being measured. For example, the meaning of diversity and how it might be measured has been the subject of considerable debate (Hurlbert, 1971; Goodman, 1975), which remains inconclusive (Washing-

and is yet to be repeated in other areas of the world. The approach has considerable merit and may greatly strengthen the basis of some biotic indices.

Ephemeroptera, Trichoptera, and Plecoptera are sensitive to most types of pollution, so the numbers of individuals in these orders will decrease with a decrease in water quality. The numbers of some Diptera and tubificid worms may increase in response to organic pollution. These responses have been used as indices (e.g., Balloch et al., 1976), as the ratios between tubificids and other organisms (e.g., King and Ball, 1964), or just as counts of the number of taxa belonging to the sensitive groups (e.g., Plafkin et al., 1989) (see Johnson et al., Chapter 4). Virtually all of the indices or other measurements using these assumptions have been developed in relation to organic pollution of rivers. However, some species of Trichoptera and Ephemeroptera are highly tolerant of trace metal pollution (Norris et al., 1982; Norris, 1986), so caution is advised in the general application of indices based on the assumptions just discussed. Other difficulties include: the large amount of initial work that may be needed to define pollution tolerances and "clean" freshwater communities, and the limited number of taxonomic keys to many species (Resh and Jackson, Chapter 6).

Some biotic indices are based on the assumption that the ratios of organisms with different feeding strategies will change with pollution (e.g., collectors will be more abundant than shredders under polluted conditions) or that trophic generalists will be more pollution-tolerant than trophic specialists. Some doubt exists as to whether these general rules hold true and even whether it is possible to assign taxa to different feeding strategies (see Chapter 6).

7.4.1.3. Statistical Needs of Indices

Indices sometimes are applied on the assumption that their calculation in some way replaces the need for hypothesis testing (Resh and Jackson, Chapter 6) or statistical calculations (Beck, 1955). Clear questions must be asked before studies are designed and before the data from them can be analyzed in any sensible way (Green, 1979). Often, temporal trends in a diversity or biotic index are presented, or values of a diversity index are compared with little or no statistical analysis (e.g., Gupta and Pant, 1983; Chadwick and Canton, 1984), as if they were absolute measures characteristic of the community in question and not subject to sampling error. Clearly, this assumption is no truer than for any other finite set of measurements made, and replicated determinations of a diversity or biotic index can be expected to vary by chance alone. Many variance formulae for diversity indices are for the sample variance and not the variance of the sampling distribution of the diversity index. This latter value is needed for statistical inference. Without

7.4.2. Analysis of Variance

Analysis of variance (ANOVA) is a well-established parametric technique for comparing means of a single variable (Sokal and Rohlf, 1981). Its major uses in the analysis of data from benthic surveys include: assessing the statistical significance of differences in a biological measure among various water bodies, among different locations in the same water body, or among samples taken from one location at different times; and partitioning total variability in values of a biological measure into components attributable to variation among sites or times and components attributable to each of one or more levels of replication. This latter analysis is usually a prelude to optimizing sampling design.

ANOVA arose from the study of replicated samples and is founded in particular on a linear relationship between the variance of means of replicated samples and the average variance of the values that make up each sample. The observed variance among the means for different sites or times can be compared with that expected to arise from this relationship; the ratio of the two variances can be compared using an F-test.

An example of this kind of analysis is provided by Tiller (1988) who studied the effects of human disturbance on the benthic macroinvertebrate fauna of the Thredbo River in Kosciusko National Park, New South Wales, Australia. The river passes by a ski resort, Thredbo Village, which discharges treated sewage 1.5 km downstream. To assess the effects of this potential source of pollution on the fauna, Tiller chose sampling stations above and below the village itself (sites I, II), below a refuse dump and above the sewage outflow (site III), and at various distances downstream of the outflow (sites IV-VIII) (Fig. 7.1). Ten replicate collections of benthic macroinvertebrates were taken at each sampling station and total numbers were used as an appropriate biological measure likely to be affected by the sewage outflow. A separate analysis also was performed on species richness. Both of these variables generally are accepted as having skewed distributions (Elliott, 1977), so they were log-transformed before analysis. The results of the ANOVA for logged abundance are shown in Table 7.1. The significance of the F-value in the ANOVA table indicates that variation of macroinvertebrate abundance among sites was unlikely to have occurred by chance alone.

The next step in the analysis was to determine which sites were significantly different from the others, and Tiller chose from a range of possible procedures (reviewed by Keppel, 1973; Sokal and Rohlf, 1981; Day and Quinn, 1989) to do pair-wise multiple comparisons using the Student-Newman-Kuels procedure. Site II, below the village, and site IV, immediately below the sewage outflow, each had significantly greater numbers of benthic macroinvertebrates than any of the other sites, demonstrating an impact on

Table 7.1. Analysis of variance table of results from the effects of sewage effluent on abundance ($\log_{10}$ X + 1) of macroinvertebrates in the Thredbo River, Australia.

Source	DF	Sum of Squares	Mean Square	F	Probability
Site	7	6.7255	0.9608	28.03	0.0001
Error	72	2.4679	0.0343		
Total	79	9.1934			

the stream by both the village itself and the effluent outflow. The site upstream of the village and sites some distance downstream of the outflow were not significantly different, so a persistent effect on the stream fauna could not be demonstrated.

The advantage of ANOVA over more qualitative approaches lies in its ability to distinguish between true trends that occur in the river (significant results) and those trends likely to have arisen in the sample through chance alone (nonsignificant results), because of sampling error. However, the technique has a number of limitations, which are described next.

7.4.2.1. The Need to Replicate

Replication of the biological measure used as a summary of community structure or community conditions is a necessary prerequisite to ANOVA. Only by replication can the magnitude of differences between sites or times be compared against the magnitude of differences that would be expected to occur by chance.

The number of replicates needed in a study are decided at the design stage (see Resh and McElravy, Chapter 5), and are related to the interpretations to be made with the data, the magnitude of differences to be detected, and the type of analyses to be performed (see Section 7.3.1, "Precision and Analysis," above). Replication occurs at two levels: within a site and time and among sites and times.

The design of many benthic macroinvertebrate studies in rivers, and to a lesser extent in lakes, is difficult because of problems measuring changes before and after the beginning of an impact and because sites cannot be replicated easily. Consider the case where sites are located upstream and downstream of a harmful discharge and replicated collections are made on each sampling occasion before and after the discharge is released. This would appear to be a simple two-way ANOVA with two factors, each with two levels: area (control and impact) by time (before and after) (Barmuta, 1987). However, treatments within the design are not properly replicated. Hurlbert (1984) referred to this design as "pseudoreplicated" because it is possible

Figure 7.2. Relationship between standard deviations and means for counts of benthic macroinvertebrates from the South Esk River, northeastern Tasmania, Australia. n = 80, r = 0.96, P < 0.001.

extreme cases, no amount of transforming will render such data normal, and the researcher may need to resort to ANOVA models based on alternatives to the normal distribution (e.g., the GLIM package; Numerical Algorithms Group, 1986) or to nonparametric alternatives to ANOVA (Siegel, 1956; Conover, 1980). This situation might occur in water bodies that are highly polluted by some toxic waste, but fortunately the interpretation of biological data such as these usually is a trivial matter hardly needing the use of statistics (Norris and Georges, 1986).

More sophisticated approaches to transformation include those of Taylor (1961, 1980) and Box and Cox (1964). Box and Cox recommended the following transformations, which are attuned to characteristics of the data:

$$Y' = (Y^k - 1)/k \qquad \text{if } k \neq 0$$
$$Y' = \log_e(Y) \qquad \text{if } k = 0$$

where k is the maximum value of the expression:

$$(-v/2)\log_e S^2 + (k - 1)(v/n)\log_e Y$$

When $b = 0$, the distribution is considered normal and no transformation is necessary.

Downing (1979) examined the relationship between mean and variance for benthic macroinvertebrates collected with several types of samplers and from various substrata. He found that b was surprisingly constant at 1.5, suggesting that a fourth-root transformation was appropriate for all samples examined. Subsequent studies have failed to confirm Downing's suggested universality of the fourth-root transformation. Allan (1982) collected 12 replicate samples at each of several sites on 31 occasions over two years. Although these data showed a tendency for b to fall near 1.5 for seven out of 21 taxa, estimates of b were significantly different from 1.5 for the other taxa. Taylor (1980) found that, in some instances, square-root and log transformations were superior to the fourth-root transformation. Morin (1985) showed that, because the relationship between the logarithm of the variance and the logarithm of the mean is quadratic, the best transformation will depend upon the range of the means being compared.

Caution also is required when using Taylor's procedure because a log-log transformation may yield a grossly biased estimate of b (Zar, 1968; Sprugel, 1983), and the degree of bias will depend, in part, on the amount of scatter about Taylor's power relationship. A more appropriate solution for b can be obtained by nonlinear least squares regression (NLIN procedure; SAS Institute, 1987) applied directly to the means and variances suspected of following a power curve. Caution also should be exercised with the procedures of Box and Cox (1964) and Taylor (1961, 1980) where the transformation is optimized for the sample at hand, when what really is required is a transformation that will correct the population from which the sample was drawn. Sampling error may lead to selection of an inappropriate transformation, so the traditional log and square-root transformations should not be rejected too readily.

Finally, transformations alter the statistical properties of the data, and they also may alter the way in which underlying biological-environmental relationships are expressed. If a habitat had all but one of the characteristics that were necessary to support a species, this habitat would not support the species. A logarithmic transformation of habitat characteristics would render them multiplicative, rather than additive; a zero value would correctly define the site as uninhabitable, whereas untransformed data would remain additive and possibly obscure the true situation (Meffe and Sheldon, 1988). The consequences of transformation on the interpretations also may need to be considered.

7.4.2.3. Suitability of Diversity and Biotic Indices for ANOVA

Most of what has been said above about transformations applies to simple biological measures, such as total abundance or species richness. The be-

each environmental variable and the evaluation of their rank order of importance are meaningless when the "independent" variables are correlated (Green, 1979). Application of Principal Components Analysis to environmental variables before multiple regression will overcome violations of assuming independence among variables but will frustrate the major objective of providing a minimum subset of variables for prediction.

Contrary to common perception, multiple regression cannot provide insight into causal relationships between environmental variables and the biological measure of interest. Causal relationships only can be established by experimentation, where environmental factors can be manipulated (see Cooper and Barmuta, Chapter 11). At the very best, multiple regression may provide a subset of environmental variables that can be considered to be the most parsimonious explanation for variation in the biological measure, but then parsimony serves only as a foundation for future experimentation and testing. On its own, a multiple regression seldom provides great insight into the true state of affairs. However, where true controls are unavailable and true replication of treatments (sites or times) often is impossible (as in most field situations), meeting the requirements for manipulative experiments may not be achievable. Under these circumstances, many take the view that, as an exploratory tool for gaining insight into causal relationships, only multiple regression and related procedures are available. Although multiple regression is unable to demonstrate conclusively a causal link between variables, it is superior to just staring at a table of data or at a neat representation of an ordination.

The strength of multiple regression lies in its ability to yield predictive models. For example, Downing (1986) successfully used regression techniques to establish predictive relationships between the number of organisms for each of several epiphytic invertebrate taxa and the biomass of each macrophyte species. In this example, the relative significance of environmental variables in the model, or whether or not they were causally related to the dependent variable, were not of concern. An empirical tool useful for prediction was of interest, based on some minimum set of environmental variables.

If used in stepwise fashion, multiple regression will yield a subset of available environmental variables that is best able to predict the value of the biological index of interest. Only these variables need be measured to obtain the prediction, often with considerable savings in time and cost. The resulting model may be used to predict the value of the biological index at sites where only measurements of environmental variables are available or to predict changes in the biological index on the basis of postulated changes in environmental variables. The model may be used to predict the value of a biological index in an impacted area, which then can be compared with the value observed, to assess the impact.

cated together in a table; (2) nonhierarchical classification procedures that group the most homogeneous samples or taxa (the investigator usually predetermines the number of groups that are formed); (3) hierarchical classification, which forms groups but arranges them using a hierarchy (dendrogram); and (4) ordination.

Groups generally are formed by convenient management of information. The groups so formed usually are informationally somewhat homogeneous, a feature in which grouping offers a simplification of the complexities of the natural world (Gauch, 1982). The relationships among several group-forming procedures are illustrated in Table 7.2.

Table arrangement, which is used for classifying a "taxa by sample" matrix, was suggested first by Braun-Blanquet (1932) for work with plant communities. The approach subsequently has been used widely in Europe by plant ecologists (Gauch, 1982). The data originally were compiled manually for display in a table of compositionally similar samples and distributionally similar species. The table indicates relationships among groups of sites and/or species along a continuum, rather than in distinct groups as is the case with classification. This is the basis for the Saprobien system (Kolkwitz and Marsson, 1909; see Cairns and Pratt, Chapter 2, and Johnson et al., Chapter 4). The method has been used by Norris et al. (1982) in their "total numbers classification," which grouped taxa based on abundance and distribution relative to effluent from an abandoned metal mine (Table 7.3). Here, sites 1–3 are upstream controls, whereas sites 5–8 indicate a gradient of recovery downstream. Group 1 taxa were tolerant of trace metals, Group 2 were sensitive, and Group 3 were taxa that reached maximum numbers downstream of the effluent inflow (Norris et al., 1982). Allocation of taxa to groups is subjective (e.g., helminthid beetles in Group 1), so table arrangements by others may differ, usually at group boundaries (Gauch, 1982). Nevertheless, the method is useful for showing gradients in abundance of taxa, in groups or community types, and in sites.

Numerical techniques generally begin with the calculation of some measure of association between pairs of samples. The indices then are grouped using a mathematical technique. Many different indices are available (see reviews by Goodman and Kruskal, 1954, 1959; Sokal and Sneath, 1963; Southwood, 1978; Washington, 1984; Hruby, 1987; Johannsson and Minns, 1987). Many of the indices presented have different theoretical justifications (Grassle and Smith, 1976; Southwood, 1978; Hruby, 1987; Johannsson and Minns, 1987), so it is important for anyone approaching the field to decide which index is best suited to their particular needs. The methods used to group the indices are basically of two types (Table 7.2): classification procedures and ordination (e.g., Principal Components Analysis, Principal Coordinates Analysis, and Multidimensional Scaling).

Several features of classification procedures need to be considered when

Table 7.3. An example of a table arrangement: the "total numbers classification" of Norris et al. (1982). Mine waste pollution enters between sites 3 and 4. Numbers are totals collected in 30 replicates on 10 occasions over two years.

	Site							
	1	2	3	4	5	6	7	8
Group 1								
Ephemeroptera								
Baetidae *Baetis baddamsae*	252	616	681	366	28	299	416	63
Coleoptera Helminthidae								
Austrolimnius sp. (adult)	27	70	40	15	8	3	2	1
Austrolimnius sp. (larvae)	314	285	148	20	2	1	6	0
Trichoptera								
Rhyacophilidae *Taschorema ferulum*	65	100	56	65	14	35	13	5
Leptoceridae *Oecetis* sp. 1	75	2,302	1,404	1,030	1,681	672	203	2
Hydropsychidae *Asmicridea* sp.	504	23	24	231	1	127	1	33
Group 2								
Basomatophora								
Hydrobiidae *Rivisessor gunnii*	597	13,511	4,491	0	0	0	0	0
Heterodonta								
Sphaeriidae *Sphaerium tasmanicum*	7	643	107	0	0	1	0	1
Ephemeroptera								
Leptophlebiidae								
Atalophlebia australis	0	127	65	0	0	0	0	0
Atalophlebioides sp. 1	773	342	205	28	2	0	0	0
Atalophlebioides sp. 2	75	724	480	25	1	0	12	0
Atalonella sp.	356	485	337	27	2	1	2	0
Caenidae *Tasmanocaenis* sp.	4	27	99	0	0	0	0	1
Diptera								
Chironomidae Orthocladiinae sp.	104	3	3	8	4	0	12	1

choosing an approach (Gauch, 1982). Nonhierarchical methods will produce groups of sites or samples that are as similar as possible, but they do not provide information on the relationships among groups. Hierarchical methods, the most commonly used procedures, seek to find the most efficient step at each stage either in the progressive synthesis (agglomerative) of the population or in its subdivision to individuals (divisive), but the route may sacrifice homogeneity of the groups through which it passes. It is uncertain whether any method simultaneously can maximize hierarchical efficiency and cluster homogeneity (Lance and Williams, 1966). Classification methods can use either qualitative (presence or absence data: e.g., Crossman et al., 1974; Wright et al., 1984; Ormerod, 1987; Ormerod and Edwards, 1987) or quantitative data (abundance: e.g., Norris et al., 1982; Barnes, 1983; Osborne and Davies, 1987).

Divisive classification strategies use mathematical techniques that begin with all entities together and divide them into successively smaller groups until each one contains a single member or until a limit is reached that is determined by the researcher. Predetermined limits are useful because they save computing time and because individual entities are difficult to interpret (otherwise a classification procedure would not have been used in the first place). Agglomerative techniques begin with individual entities and form successive groups until all are included.

Monothetic approaches divide the sets of entities (usually sites or times) according to presence or absence of a single species. Such an approach would prove useful for determining individual species that are most indicative of particular pollution or habitat conditions (e.g., Murphy and Edwards, 1982). Polythetic methods use the entire taxonomic composition of samples when deriving clusters. When single indicator taxa are used to split groups, such monothetic methods may be only divisive, whereas polythetic techniques may be divisive or agglomerative (Gower, 1967). Polythetic, agglomerative approaches are the most commonly used methods (Gauch, 1982).

An important initial decision is whether or not the *relationships* among groups are needed (hierarchical vs. nonhierarchical clustering). Hierarchical methods will yield groups that are in some order. A hierarchy is the most efficient pathway for obtaining a number of groups, but not necessarily the most efficient means of obtaining final subdivisions. For example, if a group of taxa does not appear near a closely related group early in the hierarchy, it will be more dissimilar from succeeding groups, and its final representation in the hierarchy may be most dissimilar from groups to which it is, in reality, closely related. Nonhierarchical classification methods such as the REMUL program (Lance and Williams, 1975) may form groups, the members of which are as similar to each other as possible.

For example, several clustering methods were used by Norris et al. (1982) to create groups of taxa, the distribution and abundance of which responded
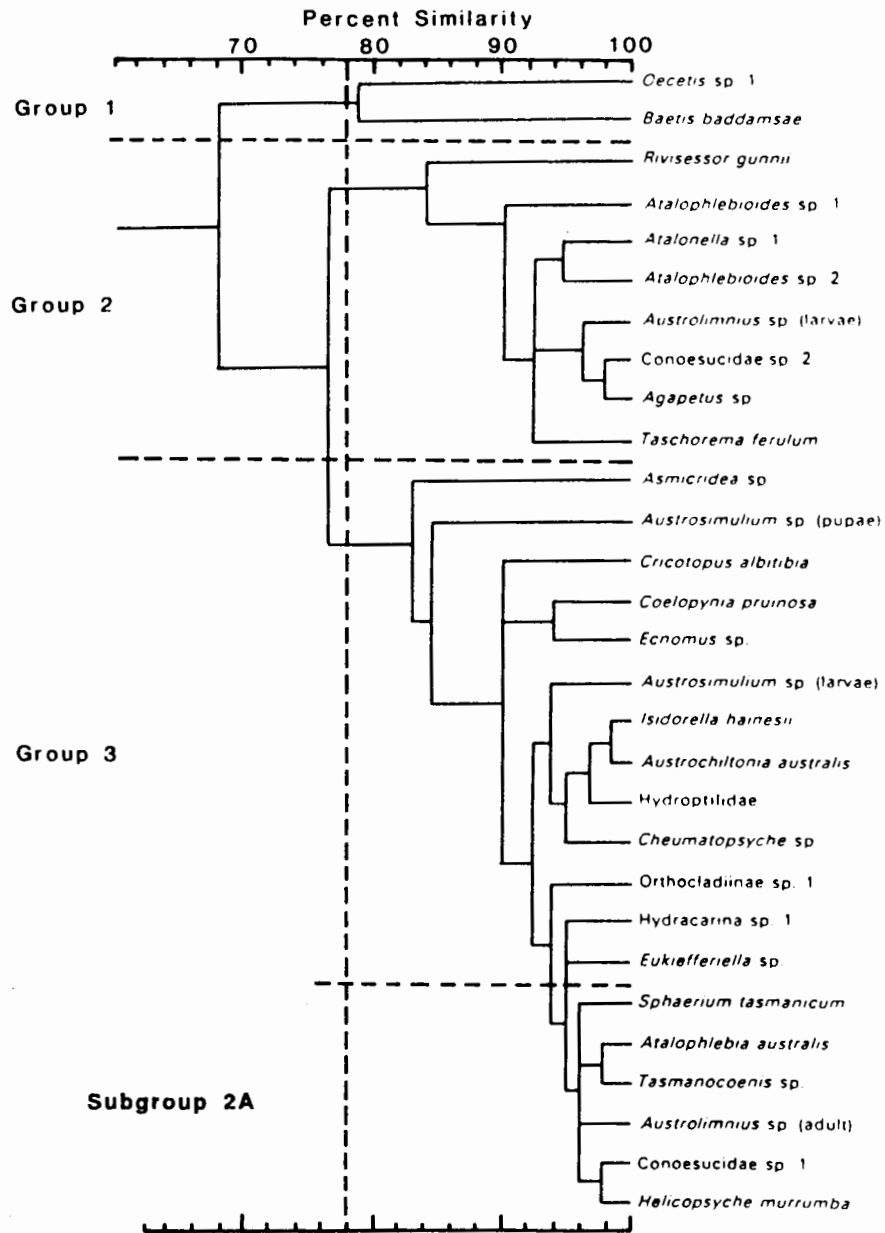
Figure 7.4. Dendrogram resulting from the hierarchical classification using an agglomerative polythetic clustering method with average linkage sorting. Subgroup 2A is composed of taxa whose distribution and abundance were close to Group 2, but which joined at the end of the hierarchy through a chaining effect. From Norris et al. (1982), reprinted by permission of the CSIRO Editorial and Publishing Unit.

data should be transformed (Noy-Meir, 1973; Gauch, 1982; Hruby, 1987), or an index should be chosen that accounts for these differences (e.g., the Canberra Metric Index, Clifford and Stephenson, 1975; or the Morisita Index, Hruby, 1987). The need for transformation was discussed by Gauch (1982) who suggested that ranges of 0 to 10 are best, but that ranges of 0 up to 300 can be tolerated. When community samples are relatively homogeneous, and variation is manifested in small differences in abundance, a transformation may compress the values, thereby destroying important information (Gauch, 1982). Additionally, rare taxa, which usually are defined by some arbitrary limit (e.g., Gauch, 1982; Norris et al., 1982; Marchant et al., 1984), often are deleted from the data matrix before analysis. Such taxa should not be disregarded totally because they may provide useful data particularly in relation to conservation (e.g., reanalysis by Faith and Norris, 1989, of Metzeling et al., 1984, from the La Trobe River, Victoria).

Indices of association that are based on the total number of taxa found in all samples and that treat the abundance of a taxon from samples being compared as a point of similarity, should be disregarded (Hruby, 1987). Co-absence is considered to be of little ecological significance (Clifford and Stephenson, 1975; Boesch, 1977), especially in data sets that may have many zero entries (Hruby, 1987). Additionally, Morisita's Index has abundance in the denominator, so its use can be rejected for any data sets that include sample records with no animals, a situation that may be common in polluted areas.

Euclidean distance gives more weight to abundant taxa (Clifford and Stephenson, 1975; Washington, 1984; Hruby, 1987). Pinkham and Pearson's Coefficient is sensitive to changes in rare taxa, which also may make it sensitive to normal sampling error.

The Bray-Curtis Index is favored by some (e.g., Boesch, 1977) because it varies linearly to changes in species numbers and abundance; the Canberra Metric Index does not (Bloom, 1981). Percent similarity also is useful (Johannsson and Minns, 1987) and also has been shown to respond linearly to community overlap (Bloom, 1981; Gauch, 1982).

Whittaker (1952) noted that percent similarity failed when relative proportions of taxa remained similar but overall abundance varied, because it only can be sensitive to changes in relative abundance.

Indices should be chosen relative to the type of data being analyzed, which may be transformed accordingly. Investigators should use several indices, rather than relying on a single, all-purpose one (Brock, 1977).
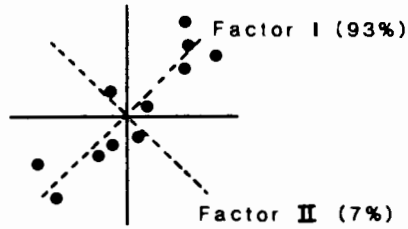
7.4.7. Clustering Strategies

A procedure must be selected to build up the hierarchy; such procedures are described by Sneath and Sokal (1973), Gauch (1982), and Belbin (1987).
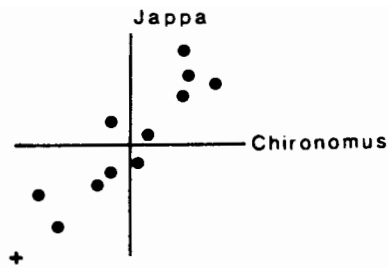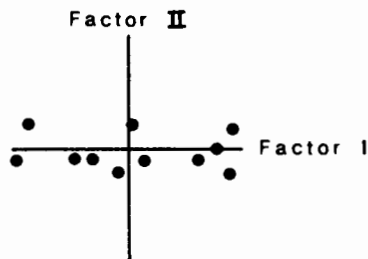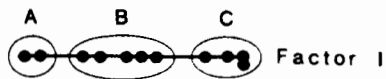
SITES PLOTTED IN A SPECIES SPACE

SELECTION OF PRINCIPAL COMPONENTS (FACTORS)

SITES PLOTTED IN A FACTOR SPACE

ORDINATION WITH 7% OF VARIATION DISCARDED

CLASSIFICATION

Figure 7.5. A visual rationale for ordination by the method of Principal Components. The first step is to plot sites or samples in a space defined by taking each species as an axis and letting abundance represent the position of the sample along each axis. The coordinate axes then are centered on the multivariate mean (+ is the origin before centering) and new orthogonal factors are defined progressively, each in the direction of maximal remaining variation. Sites or samples then are visualized in the new factor space, and least informative dimensions are discarded, with some loss of information. Patterns in the reduced space then are identified and trends or clusters defined.

ity between each pair of samples is measured, using some chosen measure of dissimilarity, and the points then are plotted to make the squared distance between every pair of points correspond to the squared dissimilarity. An initial "sample by species" matrix is not required by the analysis. Any metric distance matrix summarizing the relationships between samples can be used so that the researcher is not limited to similarities based on the correlation coefficient.

Should the dissimilarity measure be nonmetric, the procedure will produce satisfactory results provided that care is taken not to include axes in the final solution for which the cumulative percentage variation explained equals or exceeds 1.0, or for which the associated eigenvalues are negative. The often-quoted assumption of normality of the measurements that form the data set also may be disregarded, provided no statistical tests or inferences are to be made and that the technique is to be used only in the search for pattern (Marriott, 1974).

One difficulty with PCA and PCoA is that the solution is optimized in terms of squared distances, so that larger distances between samples are given disproportionate weight. This may be an advantage if one is seeking clusters, because the first few axes will emphasize distances between natural groupings at the expense of distances between samples within the groupings. However, less information will be summarized in the same number of axes compared with a technique such as Multidimensional Scaling (MDS) (Kruskal, 1964a,b). MDS endeavors to find, using an iterative procedure, the best fit between the input dissimilarities and the distances between samples in the resulting ordinated space. MDS now is chosen by many over PCA and PCoA when dealing with ecological data. MDS also is less likely to produce the distorted representations of underlying gradients that affect PCA (Noy-Meir and Austin, 1970; Austin and Noy-Meir, 1971). PCA and PCoA assume a linear response in the abundance of species along environmental gradients, which is a poor reflection of reality. Nonmetric Multidimensional Scaling (NMDS), so-called because only the rank order of the dissimilarities between samples is preserved in the geometric representation, can accommodate a much wider range of response functions (Minchin, 1987) but still assumes monotonicity in the response of species abundances to environmental gradients. A disadvantage of MDS in comparison with PCA and PCoA is that the analyst must provide the dimension of the solution in advance, and the most appropriate dimension may not be very evident. One must repeat the computationally expensive analysis for each of several dimensions and look for a significant decline in "stress" (Manly, 1986) with increasing dimension. Stress is a measure of the fit of the dissimilarity between samples and the distance between them in the ordinated space. In PCA and PCoA, one needs only to peruse the list of eigenvalues to decide on the most appropriate dimension for the solution.
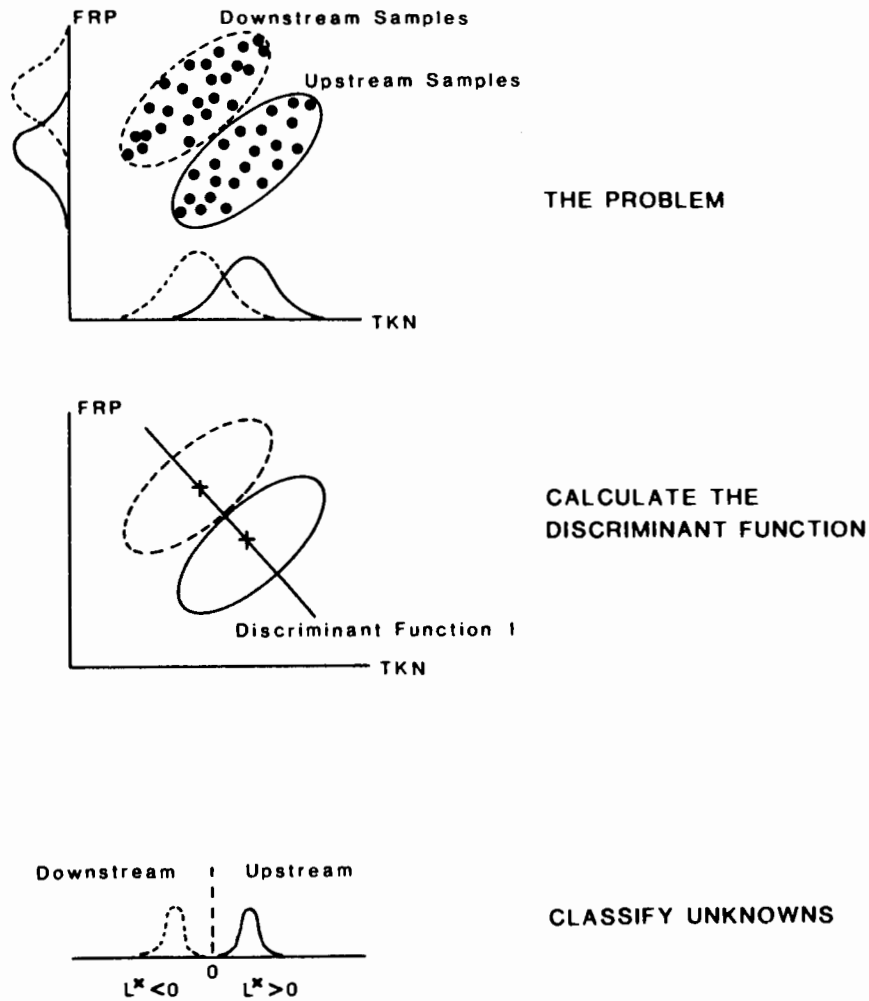
Figure 7.6. A visual rationale for Discriminant Function Analysis. In this hypothetical example, samples are classified before the analysis as being either upstream or downstream of a point impact. When the environmental variable, filterable reactive phosphorus (FRP), is considered alone, discrimination is poor between upstream and downstream sites. Another variable, total Kjeldahl nitrogen (TKN), is equally poor for discrimination. When taken together and used to define a linear discriminant function, these variables are able to discriminate absolutely between upstream and downstream samples.

discriminant functions to determine the proportion of sites that are misclassified (e.g. Marchant et al., 1984). This approach is circular because the sites used to validate the predictive model are those used to formulate it initially, but it provides a rough indication of the power of discrimination. A more rigorous approach is to validate on the basis of a set of sites of known group membership, but which were not used in developing the discriminant functions, or to use bootstrapping (where DFA is calculated from a subset of the data and validated using the remainder, and the process is repeated for an estimate of the reliability of discrimination).

7.4.10. Time-Series and Trend Analyses

7.4.10.1. Time Series

Many study designs for the collection of benthic macroinvertebrates involve sampling through time. The numbers of animals collected at successive samplings in time and space may be correlated highly with the numbers collected in previous samples. Such autocorrelation (Hurlbert, 1984; Stewart-Oaten et al., 1986; Barmuta, 1987) may invalidate the use of many parametric statistical tests because the assumption of independence is violated. A family of statistical methods called "time-series analysis" may be appropriate for analyzing these types of data, but the methods seem to have been overlooked in benthic macroinvertebrate studies. The approach has been reviewed by Green (1979) and Millard et al. (1985), and Van Latesteijn and Lambeck (1986) have applied time-series analysis to the effects of estuary closure on oyster catchers in Holland. An easily read introductory text on the subject is provided by Chatfield (1984). Examples of the application of time-series analysis to benthic biomonitoring studies are lacking, but because it is a potentially useful approach with which few seem to be familiar, a description of the procedure follows.

Time-series analysis involves the calculation of autocorrelations and cross-correlations. The autocorrelation of a series of observations taken over time is a measure of the extent to which a variable is dependent on its own past values. Calculation of the autocorrelation function (ACF) of a time series involves determining the correlation of each value with values 1, 2, 3, . . . k intervals (lags) preceding it. "k" of these autocorrelations will exist for a series, together making up the autocorrelation function (Box and Jenkins, 1976). The ACF of a time series with a seasonal pattern will exhibit a peak at a lag that corresponds to the length of the cycle. A time series with no serial dependence will produce an ACF with scatter around zero. The ACF of a random series can be shown to have a mean of zero, a variance of $1/N$ and a normal distribution, so values that lie outside $\pm 1.96 / \sqrt{N}$ are significant at the 5% probability level (Chatfield, 1984). Calculation of the

a closer fit to the data than a global polynomial because the fit of the function for each region is determined by local, not global, conditions (Wold, 1974).

Time-series analysis techniques require that variables be distributed normally and that no relationship exist between the mean and variance (Chatfield, 1984). Many limnological variables are distributed approximately lognormally, so a logarithmic transformation is appropriate (Platt et al., 1970; Rutherford, 1984). This has a further benefit: if variables have a seasonal component, it will convert any multiplicative relationship between the mean, the seasonal term, and the error term to an additive one, which allows the series to be fitted to a linear equation (Chatfield, 1984).

Time-series analysis further requires that any long-term trend in variables be removed before analysis (Chatfield, 1984). This can be done by "differencing" the data, that is, producing a new series made up of the differences between successive pairs of observations in the original series. A linear trend will be removed by differencing once, a quadratic trend by differencing the already differenced series, and so on. Seasonal variation also may be removed by seasonal differencing, that is, subtracting pairs of observations that are a season apart. This allows resolution of features that may have been concealed by the magnitude of seasonal variation.

The first steps in analysis are interpolation of missed observations and, where necessary, transformation of variables (pH excepted). ACFs of each variable then may be calculated and, using these functions, ARIMA models of the series can be fitted. The residual series from these models, the filtered series, are used to calculate the CCFs between variables.

7.4.10.2. Trends

Trends in water quality data may be detected by a variety of methods, including nonparametric, parametric, spectral, and time-series analysis. Nonparametric methods have been suggested for water quality data by Hirsch et al. (1982) and Van Belle and Hughes (1984) because of the problems of nonnormality, nonlinearity, nonindependence, missing values, censored data, and periodic cycles. The seasonal Kendall test can be recommended (Hirsch et al., 1982) because, under realistic stochastic processes (exhibiting seasonality, skewness, and serial correlation), it is robust in comparison to parametric alternatives, although it is not an exact test in the presence of serial correlation (autocorrelation). A second test recommended by Hirsch et al. (1982) is the Kendall slope estimator, which estimates trend magnitude and is an unbiased estimator of the slope of a linear trend. It has considerably higher precision than a regression estimator in which data are highly skewed, but it has lower precision when the data are distributed normally. The third procedure recommended provides a means of testing for change over time in the relationship between constituent concentrations and flow, thus avoid-

of occurrences, ratios, regressions, plotting, and subcounts vs. total counts. Missing data also need to be identified and a strategy to deal with them adopted (either by deletion of records or by some means of estimating the missing value).

4. Diversity and biotic indices are used commonly to condense presence/absence and abundance data for further analysis and interpretation. Many of the arguments in favor of using indices are not theoretically valid, many shortcomings are evident in their application, and the biological or ecological meaning of what they purport to measure is poorly understood. Most biotic indices are specific to the conditions and sites of their development, and much background information on responses of taxa to pollution usually is needed before they can be implemented. They are as subject to error as any other community measures would be, and the level of uncertainty in their use needs to be established before they can be interpreted properly. It is concluded that indices should be used with caution.

5. ANOVA is a major method of analysis used to make comparisons and to partition total variability into components of the study. It relies on replicated sampling and requires that certain assumptions be met. Normally distributed data is one assumption that may be attended to after data collection; this is done by transformation. Caution is advised in the use of diversity or biotic indices with ANOVA, unless the statistical distribution of the data is known and appropriate transformations (if needed) can be performed.

6. Multiple regression may be a useful technique for developing predictive models and for hypothesis generation. When a subset of environmental variables is required that is best able to predict the biological measure of interest, a stepwise application of the method should be used.

7. Many benthic studies will be multiple-variable problems and, therefore, are amenable to the use of multivariate statistical techniques. Multivariate procedures consider each species to be a variable and the presence/absence or abundance of each species to be an attribute of a site or time. Subtle changes in the species composition or in the abundance of particular species across sites are not inherently masked by the need to summarize the combined characteristics of the site as a single value. Multivariate techniques, therefore, show greater promise than univariate comparisons for detecting and understanding spatial and temporal trends in the benthic macroinvertebrate fauna.

8. Several methods are available for the classification or grouping

series analysis relies on data that are collected regularly at intervals more frequent than the period of variation among the variables of interest. These requirements, although quite stringent, are fulfilled by most monitoring programs.

Acknowledgments

References

Abel, P.D. 1989. *Water Pollution Biology*. Ellis Horwood, Chichester, England.

Allan, J.D. 1982. The effects of reduction in trout density on the invertebrate community of a mountain stream. *Ecology* 63:1444–55.

Allan, J.D. 1984. Hypothesis testing in ecological studies of aquatic insects. In *The Ecology of Aquatic Insects*, eds. V.H. Resh and D.M. Rosenberg, pp. 484–507. Praeger Pubs., New York.

Anderson, R.M., D.M. Gordon, M.J. Crawley, and M.P. Hassell. 1982. Variability in the abundance of animal and plant species. *Nature* (London) 296:245–8.

Anonymous. 1977. *General Statistical Programme*. Rothamsted Experimental Station, Harpendon, Hertfordshire, England.

Armitage, P.D., R.J.M. Gunn, M.T. Furse, J.F. Wright, and D. Moss. 1987. The use of prediction to assess macroinvertebrate response to river regulation. *Hydrobiologia* 144:25–32.

Auclair, A.N. and F.G. Goff. 1971. Diversity relations of upland forests in the western Great Lakes area. *American Naturalist* 105:499–528.

Austin, M.P. and I. Noy-Meir. 1971. The problem of non-linearity in ordination: experiments with two-gradient models. *Journal of Ecology* 59:763–73.

Balloch, D., C.E. Davies, and F.H. Jones. 1976. Biological assessment of water quality in three British Rivers: the North Esk (Scotland), the Ivel (England) and the Taf (Wales). *Water Pollution Control* 76:92–110.

Barmuta, L.A. 1987. Polemics, aquatic insects and biomonitoring: an appraisal. In *The Role of Invertebrates in Conservation and Biological Survey*, ed. J.D. Majer, pp. 65–72. Western Australian Department of Land Management Report, Western Australian Department of Land Management, Perth, Western Australia.

Barnes, L.E. 1983. The colonization of ball-clay ponds by macroinvertebrates and macrophytes. *Freshwater Biology* 13:561–78.

Beak, T.W. 1965. A biotic index of polluted streams and its relationship to fisheries. In *Advances in Water Pollution Research. Proceedings of the Second International Conference, Tokyo, August, 1964, Vol. 10*, ed. O. Jaag, pp. 191–210. Pergamon Press, Oxford, England.

Bechtel, T.J. and B.J. Copeland. 1970. Fish species diversity indices as indicators

Chandler, J.R. 1970. A biological approach to water quality management. *Water Pollution Control* 69:791–2.

Chatfield, C. 1984. *The Analysis of Time Series: Theory and Practice*, 3rd ed. Chapman and Hall, London.

Cherry, D.S., R.K. Guthrie, E.M. Davis, and R.S. Harvey. 1984. Coal ash basin effects (particulates, metals, acidic pH) upon aquatic biota: an eight-year evaluation. *Water Resources Bulletin* 20:535–44.

Chutter, F.M. 1972. An empirical biotic index of the quality of water in South African streams and rivers. *Water Research* 6:19–30.

Chutter, F.M. and R.G. Noble. 1966. The reliability of a method of sampling stream invertebrates. *Archiv für Hydrobiologie* 62:95–103.

Clarke, R.McV. 1977. The use of multivariate techniques in analysing effects of industrial effluents on benthic communities in central Canada. In *Biological Monitoring of Inland Fisheries*, ed. J.S. Alabaster, pp. 133–41. Applied Science Pubs., London.

Clifford, H.T. and W. Stephenson. 1975. *An Introduction to Numerical Classification*. Academic Press, New York.

Cochran, W.G. 1947. Some consequences when the assumptions for the analysis of variance are not satisfied. *Biometrics* 3:22–38.

Conover, W.J. 1980. *Practical Nonparametric Statistics*, 2nd ed. John Wiley, New York.

Crossman, J.S., R.L. Kaesler, and J. Cairns, Jr. 1974. The use of cluster analysis in the assessment of spills of hazardous materials. *American Midland Naturalist* 92:94–114.

Culp, J.M. and R.W. Davies. 1980. Reciprocal averaging and polar ordination as techniques for analysing lotic macroinvertebrate communities. *Canadian Journal of Fisheries and Aquatic Sciences* 37:1358–64.

Cushman, R.M. and J.C. Goyert. 1984. Effects of a synthetic crude oil on pond benthic insects. *Environmental Pollution* (Series A) 33:163–86.

Dale, M., D. Hain, G. Lance, P. Milne, D. Ross, M. Thomas, and W.T. Williams. 1980. *Taxon Users Manual*, 2nd ed. Division of Computing Research, Commonwealth Scientific and Industrial Research Organization, St. Lucia, Qld., Australia.

Day, R.W. and G.P. Quinn. 1989. Comparisons of treatments after an analysis of variance in ecology. *Ecological Monographs* 59:433–63.

de March, B.G.E. 1976. Spatial and temporal patterns in macrobenthic stream diversity. *Journal of the Fisheries Research Board of Canada* 33:1261–70.

Descy, J.-P. 1973. La végétation algale benthique de la Meuse Belge et ses relations avec la pollution des eaux. *Lejeunia* (nouveau Series) 66:1–62.

Downing, J.A. 1979. Aggregation, transformation, and the design of benthos sampling programs. *Journal of the Fisheries Research Board of Canada* 36:1454–63.

Downing, J.A. 1986. A regression technique for the estimation of epiphytic invertebrate populations. *Freshwater Biology* 16:161–73.

Dyer, D.P. 1978. An analysis of species dissimilarity using multiple environmental variables. *Ecology* 59:117–25.

Efron, B. 1979. Bootstrap methods: another look at the jackknife. *Annals of Statistics* 7:1–26.

Hall, A.V. 1969. Avoiding informational distortion in automatic grouping programs. *Systematic Zoology* 18:318–29.

Harman, W.N. 1972. Benthic substrates: their effect on fresh-water Mollusca. *Ecology* 53:271–7.

Hellawell, J.M. 1986. *Biological Indicators of Freshwater Pollution and Environmental Management.* Elsevier, London.

Hill, M.O. 1979a. *DECORANA—A FORTRAN Program for Detrended Correspondence Analysis and Reciprocal Averaging.* Cornell Univ., Ithaca, NY.

Hill, M.O. 1979b. *TWINSPAN—A FORTRAN Program for Arranging Multivariate Data in an Ordered Two-Way Table by Classification of the Individuals and Attributes.* Cornell Univ., Ithaca, NY.

Hirsch, R.M., J.R. Slack, and R.A. Smith. 1982. Techniques of trend analysis for monthly water quality data. *Water Resources Research* 18:107–21.

Hope, A.C.A. 1968. A simplified Monte Carlo significance test procedure. *Journal of the Royal Statistical Society* (Series B) 30:582–98.

Hottelling, H. 1933. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology* 24:417–41.

Hruby, T. 1987. Using similarity measures in benthic impact assessments. *Environmental Monitoring and Assessment* 8:163–80.

Hughes, B.D. 1978. The influence of factors other than pollution on the value of Shannon's diversity index for benthic macro-invertebrates in streams. *Water Research* 12:359–64.

Hurlbert, S.H. 1971. The nonconcept of species diversity: a critique and alternative parameters. *Ecology* 52:577–86.

Hurlbert, S.H. 1984. Pseudoreplication and the design of ecological field experiments. *Ecological Monographs* 54:187–211.

Hynes, H.B.N. 1960. *The Biology of Polluted Waters.* Liverpool Univ. Press, Liverpool, England.

Hynes, H.B.N. and U.R. Yadav. 1985. Three decades of post-impoundment data on the littoral fauna of Llyn Tegid, North Wales. *Archiv für Hydrobiologie* 104:39–48.

Jenkins, G.M. and D.G. Watts. 1968. *Spectral Analysis and Its Applications.* Holden-Day, London.

Johannsson, O.E. and C.K. Minns. 1987. Examination of association indices and formulation of a composite seasonal dissimilarity index. *Hydrobiologia* 150:109–21.

Keppel, G. 1973. *An Introduction to the Design and Analysis of Experiments.* Prentice-Hall, Englewood Cliffs, NJ.

King, D.L. and R.C. Ball. 1964. A quantitative biological measure of stream pollution. *Journal of the Water Pollution Control Federation* 36:650–3.

Kleinbaum, D.G. and L.L. Kupper. 1978. *Applied Regression Analysis and Other Multivariable Methods.* Duxbury Press, North Scituate, MA.

Kolkwitz, R. and M. Marsson. 1909. Ökologie der tierischen Saprobien. Beiträge zur Lehre von der biologischen Gewässerbeurteilung. *Internationale Revue der gesamten Hydrobiologie und Hydrographie* 2:126–52.

Krebs, C.J. 1985. *Ecology. The Experimental Analysis of Distribution and Abundance,* 3rd ed. Harper and Row, New York.

Narf, R.P., E.L. Lange, and R.C. Wildman. 1984. Statistical procedures for applying Hilsenhoff's biotic index. *Journal of Freshwater Ecology* 2:441–8.

Needham, P.R. and R.L. Usinger. 1956. Variability in the macrofauna of a single riffle in Prosser Creek, California, as indicated by the Surber sampler. *Hilgardia* 24:383–409.

Nemec, A.F.L. and R.O. Brinkhurst. 1988a. Using the bootstrap to assess statistical significance in the cluster analysis of species abundance data. *Canadian Journal of Fisheries and Aquatic Sciences* 45:965–70.

Nemec, A.F.L. and R.O. Brinkhurst. 1988b. The Fowlkes-Mallows statistic and the comparison of two independently determined dendrograms. *Canadian Journal of Fisheries and Aquatic Sciences* 45:971–5.

Norris, R.H. 1986. Mine waste pollution of the Molonglo River, New South Wales and the Australian Capital Territory: effectiveness of remedial works at Captains Flat mining area. *Australian Journal of Marine and Freshwater Research* 37:147–57.

Norris, R.H. and A. Georges. 1986. Design and analysis for assessment of water quality. In *Limnology in Australia*, eds. P. De Deckker and W.D. Williams, pp. 555–72. Commonwealth Scientific and Industrial Research Organization, Melbourne, Australia, and Junk Publs., Dordrecht, The Netherlands.

Norris, R.H., P.S. Lake, and R. Swain. 1982. Ecological effects of mine effluents on the South Esk River, north-eastern Tasmania. III. Benthic macroinvertebrates. *Australian Journal of Marine and Freshwater Research* 33:789–809.

Noy-Meir, I. 1973. Data transformations in ecological ordination. I. Some advantages of non-centering. *Journal of Ecology* 61:329–41.

Noy-Meir, I. and M.P. Austin. 1970. Principal component ordination and simulated vegetational data. *Ecology* 51:551–2.

Numerical Algorithms Group. 1986. *The Generalized Linear Interactive Modelling System. Release 3.77. Reference guide*. Numerical Algorithms Group Central Office, Oxford, England.

Odum, H.T., J.E. Cantlon, and L.S. Kornicker. 1960. An organizational hierarchy postulate for the interpretation of species-individual distributions, species entropy, ecosystem evolution, and the meaning of a species-variety index. *Ecology* 41:395–9.

Ormerod, S.J. 1987. The influences of habitat and seasonal sampling regimes on the ordination and classification of macroinvertebrate assemblages in the catchment of the River Wye, Wales. *Hydrobiologia* 150:143–51.

Ormerod, S.J. and R.W. Edwards. 1987. The ordination and classification of macroinvertebrate assemblages in the catchment of the River Wye in relation to environmental factors. *Freshwater Biology* 17:533–46.

Osborne, L.L. and R.W. Davies. 1987. The effects of a chlorinated discharge and a thermal outfall on the structure and composition of the aquatic macroinvertebrate communities in the Sheep River, Alberta, Canada. *Water Research* 21:913–21.

Patrick, R. 1950. Biological measure of stream conditions. *Sewage and Industrial Wastes* 22:926–38.

Pielou, E.C. 1975. *Ecological Diversity*. John Wiley, New York.

Pielou, E.C. 1984. *The Interpretation of Ecological Data*. John Wiley, New York.

Southwood, T.R.E. 1978. *Ecological Methods with Particular Reference to the Study of Insect Populations,* 2nd ed. Chapman and Hall, London.

Sprugel, D.G. 1983. Correcting for bias in log-transformed allometric equations. *Ecology* 64:209–10.

Sprules, W.G. 1977. Crustacean zooplankton communities as indicators of limnological conditions: an approach using principal component analysis. *Journal of the Fisheries Research Board of Canada* 34:962–75.

Stewart-Oaten, A., W.W. Murdoch, and K.R. Parker. 1986. Environmental impact assessment: "pseudoreplication" in time? *Ecology* 67:929–40.

Sugihara, G. 1980. Minimal community structure: an explanation of species abundance patterns. *American Naturalist* 116:770–87.

Taylor, L.R. 1961. Aggregation, variance and the mean. *Nature* (London) 189:732–5.

Taylor, L.R. 1980. New light on the variance/mean view of aggregation and transformation: comment. *Canadian Journal of Fisheries and Aquatic Sciences* 37:1330–2.

Tiller, D. 1988. The impact of sewage effluent on the benthic macroinvertebrate community of the upper Thredbo River. M.Sc. dissertation, Canberra College of Advanced Education, Canberra, Australia.

Van Belle, G. and J.P. Hughes. 1984. Nonparametric tests for trend in water quality. *Water Resources Research* 20:127–36.

Van Latesteijn, H.C. and R.H.D. Lambeck. 1986. The analysis of monitoring data with the aid of time-series analysis. *Environmental Monitoring and Assessment* 7:287–97.

Voshell, J.R., Jr. and G.M. Simmons, Jr. 1984. Colonization and succession of benthic macroinvertebrates in a new reservoir. *Hydrobiologia* 112:27–39.

Washington, H.G. 1984. Diversity, biotic and similarity indices. A review with special relevance to aquatic ecosystems. *Water Research* 18:653-94.

Whiting, E.R. and H.F. Clifford. 1983. Invertebrates and urban runoff in a small northern stream. Edmonton, Alberta, Canada. *Hydrobiologia* 102:73–80.

Whittaker, R.H. 1952. A study of summer foliage insect communities in the Great Smoky Mountains. *Ecological Monographs* 22:1–44.

Wilhm, J.L. 1972. Graphic and mathematical analyses of biotic communities in polluted streams. *Annual Review of Entomology* 17:223–52.

Wilhm, J.L. and T.C. Dorris. 1968. Biological parameters for water quality criteria. *BioScience* 18:477–81.

Wold, S. 1974. Spline functions in data analysis. *Technometrics* 16:1–11.

Woodiwiss, F.S. 1964. The biological system of stream classification used by the Trent River Board. *Chemistry and Industry* 11:443–7.

Wright, J.F., D. Moss, P.D. Armitage, and M.T. Furse. 1984. A preliminary classification of running-water sites in Great Britain based on macro-invertebrate species and the prediction of community type using environmental data. *Freshwater Biology* 14:221–56.

Yule, G.U. 1944. *The Statistical Study of Literary Vocabulary.* Cambridge Univ. Press, London.

Zar, J.H. 1968. Calculation and miscalculation of the allometric equation as a model in biological data. *BioScience* 18:1118–20.