BIOLOGICAL MONITORING AND ASSESSMENT USING INVERTEBRATES

Ian C. Campbell

INTRODUCTION

This chapter will focus primarily on the use of benthic macroinvertebrates for monitoring freshwater ecosystems. Many of the techniques and principles discussed are equally applicable to marine systems or the use of zooplankton for water quality monitoring, but given the constraints of a single chapter, it is not possible to encompass these other systems in any detail.

Biological versus Chemical Monitoring Programs

The selection of biological versus chemical monitoring and the balance between the two will depend on the purpose of the monitoring program. Biological monitoring has the capability of detecting unexpected impacts on aquatic systems because the biota will respond to any impacts whether expected by the investigator or not. Biological data are more directly related to the ecological condition or "ecological health" of aquatic ecosystems than chemical data, although the concept of ecological health is itself somewhat problematic (Fairweather, 1999; Norris and Thoms, 1999; Wicklum and Davies, 1995). However, biological data are less specific than chemical data, and it may not be obvious whether an affected biological community is responding mainly to poor water quality or poor habitat quality, for example. The two are often linked.

Chemical water quality monitoring has the advantage of specificity. If a water body is being monitored for the pesticide lindane, for example, and high levels are found, it is clear that the water body has been contaminated with lindane. Continuing the monitoring will demonstrate clearly whether contamination is continuing and, when it ceases, the pattern of decline of lindane concentrations. This sort of information is essential to determine whether abatement or policing practices are effective and the safety of the water for human use. However, chemical monitoring requires the specific chemical contaminants to be known prior to the sampling program commencing. A monitoring program for lindane will not detect contamination by mercury, and if contamination by mercury is not expected and therefore not tested for, it will not be discovered.
In a number of examples, contamination of aquatic ecosystems remained undetected despite chemical monitoring programs because the nature of the contamination was unexpected. For example, McKaige (1986) conducted a 1-year study of the invertebrate assemblages of the Thredbo River in southeastern Australia. The stream, which is located in Kosciusko National Park, had been the subject of several previous investigations of chemical water quality because of particular concerns about the impact of sewage effluent from the Thredbo Ski Village on the river (Hogg, 1984). However, McKaige (1986) found that at a site downstream of the village in July (southern hemisphere winter), species richness was reduced to 50 per Surber sample compared with 60 and 63 at the two upstream sites. Total number of individuals was depressed to 330 per sample (0.05/m²) compared with 1500 and 1502 at the two upstream sites. These results are consistent with toxic pollution, presumably from leachate from the village waste dump, rather than sewage impacts, which would depress species numbers but increase total numbers of organisms. Toxic leachate had not been suspected previously as a problem, and thus no chemical monitoring for toxicants had been conducted previously.

Biological monitoring can detect the impacts of all stressors present. It is thus ideally suited to routine ambient monitoring programs and to broad-scale environmental condition surveys. Chemical monitoring is far more suitable for compliance monitoring and the monitoring of specific sites where problems have been clearly identified and the nature of the contamination is established.

Why Use Invertebrates?

The invertebrates are one of the components of aquatic ecosystems most widely used for biological monitoring for a number of reasons. These have been documented by numerous previous authors (e.g., Hynes, 1960; Hellawell, 1986; Rosenberg and Resh, 1993). The first and most obvious reason is the diversity of invertebrates, particularly insects, which make up about 54 percent of all described species of organisms (Wilson, 1988). Any biological monitoring program that does not include invertebrates is likely to be excluding much of the biodiversity present at a site. Intuitively, groups with many taxa present at a site are likely to contain more specialists, which are likely to be sensitive to environmental change.

Second, invertebrates are almost ubiquitous in aquatic systems—wherever there is water, there are likely to be invertebrates. Consequently, seldom can invertebrates not be used for monitoring because they are absent from the system to be monitored. Third, invertebrates have limited mobility, so they are not able to move out of an area if conditions deteriorate and then quickly return later. They thus reflect the history of the site, enabling intermittent contaminants to be detected. Finally, their life cycles are usually on the order of months to years long, which also limits their ability to recolonize sites rapidly.

Types of Biological Monitoring

Biological monitoring is carried out most frequently for large-scale assessment of ecosystem condition, routine ambient monitoring, or to assess local environmental impacts. Monitoring of the first type would be used for state-of-the-environment reporting, and the second type would be used to establish whether the waterways of a particular region are deteriorating, improving, or stable in terms of their overall ecological condition. Monitoring of the third type is at a smaller scale, attempting to answer questions about the impact of a particular discharge or other activity on a particular stream or other water body. Monitoring may use the entire biological community or, more commonly, a subset of it, such as invertebrates, algae, or fish. More restricted subsets also are used frequently, such as the EPT (Ephemeroptera,
Plecoptera, Trichoptera) orders (e.g., Lenat, 1988) or single species, as in fluctuating asymmetry in Chironomidae and other invertebrates (e.g., Clarke, 1993). Various physiological responses, such as changes in respiration rates in fish or invertebrates, also may be used for continuous effluent monitoring (e.g., Morgan, 1976). This chapter will emphasize invertebrate assemblage analysis because this is the most widely practiced form of invertebrate monitoring, but the use of fluctuating asymmetry also will be briefly discussed.

**DESIGN OF MONITORING PROGRAMS**

The design of biological monitoring programs has always been problematic (Rabeni et al., 1995; Resh et al., 1995). Monitoring at a single site over time is relatively straightforward. As long as the sampling and sorting methods and the taxonomic resolution remain uniform, change through time can be detected, although see Linke et al. (1999). For broadscale ambient monitoring, the detection and measurement of change may be sufficient to answer questions about the trends in environmental quality. However, there are two areas where difficulty arises.

The first difficulty occurs when we wish to identify causes of environmental impacts. Standard scientific procedure requires the use of controls, in this case a similar site or sites unaffected by the putative impact. The simplest experimental design involves selection of sites prior to impact commencing, including one site or group of sites that is subjected to the impact and one that is not—the so-called BACI design (before after control impact) (Green, 1979; Underwood, 1991). Difficulties arise in use of these designs in studies of streams and rivers because the control sites often are upstream of the impacted sites, which becomes a confounding factor in the design. There is also an assumption required for the statistical analysis of BACI designs that the control and impacted sites are independent of each other, even though much of stream ecological theory is based on the assumption that downstream sites are influenced by and not independent of upstream sites.

BACI designs are not always possible in biomonitoring. For example, there is often a need to assess the impact of an activity that has already commenced and may have been operating for a number of years. The difficulty lies in establishing what the biota of a site would have been in the absence of any impact. It is well established that spatial variations in invertebrate communities are considerable even in the absence of any impacts (e.g., see Hynes, 1970). Control sites, in the sense of sites apparently unaffected, may be available, e.g., upstream sites in rivers, but in many cases control sites on the same river may not be available. Where unaffected sites are available, it is important that several are sampled to provide data with which the data from the putatively affected site can be compared. This at least allows the investigator to see whether the data from the impact site lies within or outside the range of that from the control sites.

The second difficulty arises when there is a requirement to make a judgment about the environmental quality of a site. While detecting trends is simple, to judge the quality requires some assessment of what the assemblage at the site would have been in the absence of human disturbance. This problem is related to the preceding problem. Is there an objective method of assessing what the biological community at a site would have been prior to human impact? Three approaches have been employed to solve this problem. One is the selection of control sites on the same stream, as discussed previously. The second approach has been to find reference sites, i.e., sites with little or no impact that may be on the same stream or on other streams in the same region to which the fauna of the impact sites can be compared. The third approach has been to develop models to predict what the fauna at the impact site should have been in the absence of impact and compare the existing fauna with that. This third approach will be discussed in more detail below.
The difficulty with the reference-site approach is that of finding unaffected sites or even sites that are only slightly affected. For smaller streams, most regions have some sites that are in relatively good ecological condition, but the problem is acute for assemblages in large rivers. Each large river is virtually always the only large river in its region. With what could one compare invertebrate assemblages from the Mississippi, the Murray, or the Mekong? Yet large rivers are also the recipients of the problems of all their tributaries, so frequently the whole main channel is degraded to a significant extent.

For smaller streams, reference sites on other streams are more likely. One issue may be deciding when a site is too distant to act as an effective reference. Within the United States, ecoregions have been mapped that are based on soils, land use, land surface form, and potential natural vegetation (Bailey, 1976; Omernik, 1987; Gallant et al., 1989). This has been done to assist managers of aquatic resources to understand regional patterns of attainable quality. Aquatic biologists have suggested that reference sites should be located within the same ecoregion as the site to be evaluated. However, the ecoregions do not incorporate biogeographic information on aquatic invertebrates and a quick comparison of the patterns of ecoregions (Omernik, 1987) with the distribution patterns of North American mayflies (Allen, 1990) indicates little correspondence. Thus it is unlikely that the U.S. ecoregion approach will be a successful predictor of aquatic invertebrate assemblage composition. Marchant et al. (1999) found that terrestrial biogeographic regions were a poor predictor of stream invertebrate assemblage composition in southeastern Australia. The issue has been addressed recently by a full issue of the *Journal of the North American Benthological Society* (see Hawkins and Norris, 2000), with 15 papers all concluding that the terrestrial ecoregions approach was not successful in predicting invertebrate community distributions. Curiously, none of the papers took the relatively simple approach of comparing published biogeographic data with published ecoregions.

**SAMPLING METHODS**

**Standardization**

Biological monitoring and assessment are essentially comparative studies—comparing the biota of a site at a given time with that at the same site at another time or comparing unaffected with putatively affected sites. The key factor in selection of sampling methods for biological monitoring or assessment, therefore, is repeatability. The suite of techniques used must be equally applicable at all sampling locations and at all sampling times and must be adhered to rigidly. If this is not the case, there is no way of determining whether differences between samples are caused by environmental impacts or simply are artifacts of the sampling process or changes in the procedure. While these principles may appear obvious, it is astonishing how often they are not followed, often because someone decided that there was a better sampling procedure than that being followed. For most monitoring programs, it is far less important that the samples provide the best possible representation of the system being sampled than that the sampling procedure is replicated identically.

The particular sampling method selected will depend on the specific environment. For plankton samples, either nets that filter out the plankton or water samplers that collect a fixed volume could be used equally well at any site. Benthic environments are far more heterogeneous, and the Surber or Hess samplers that are suitable for use in stony cobble stream beds cannot be used to sample fine silt or wood habitats. For benthic studies, detailed descriptions of sampling equipment are provided by Welch (1948), Hauer and Resh (1996), Smock (1996), Hellawell (1978), Merritt and Cummins (1996), Southwood (1978), Weber...
Many rapid assessment methods now simply use standardized pond nets, which do not provide quantitative samples but can be used in a variety of habitats (e.g., see Wright, 2000).

**Sampling Natural Substrates**

Sampling of natural substrates generally is preferred in biologic monitoring of benthic assemblages primarily because the sample represents the natural system. The difficulty is that the spatial and temporal physical variability of the benthic environment adds a significant noise component to the data obtained. Benthic invertebrate abundance and distribution have long been known to be strongly influenced by the nature of the physical habitat (e.g., Cummins et al., 1966; Hynes, 1970). The variability in natural habitats, both spatial and temporal, is often a significant confounding problem. For example, if a benthic invertebrate monitoring program were required to sample the entire length of a river from the mountains to the sea, there would be great difficulty in using a single sampling procedure. The upland erosional section of the river probably would have a bed of rock, boulders, and stones, whereas the downstream depositional section would have a bed of sands and silts. No sampling apparatus appropriate for one type of substratum would be successful for the other. There are several possible solutions. One is to use a sampling technique that samples benthic invertebrates indirectly, such as drift sampling (e.g., McKaige, 1987). Drift sampling could be used wherever there was sufficient current, which may or may not be possible in the lower reaches of a river. A second possibility is to use an artificial substratum method, although an artificial substratum appropriate to the upper reaches may not be appropriate for the lower reaches. A third possibility is to use several different techniques and spatially overlap them. For example, Surber sampling could be used at erosional sites and grab sampling at depositional sites, and at several sites in the middle, where both sets of substrates occur, both erosional and depositional habitats could be sampled so that the two sampling methods could be calibrated against each other.

Temporal overlap also can be used where the sampling technique is to be changed for some reason. The new and old techniques can be operated in parallel for some period so that the new method is calibrated against the old. It will then be clear whether any changes in the data resulted from changes in the sampling method or from some other factor.

**Sampling Artificial Substrata**

Use of artificial substrata in invertebrate biomonitoring has been reviewed by Rosenberg and Resh (1982). Artificial substrata include structures such as Hester-Dendy plate samplers, bags of stones or wood, and crumpled nylon bags. Some, such as the bags of stone or wood, are intended to mimic natural substrata, but samplers such as the plate samplers or the ground-glass slides used for sampling stream algae are designed primarily to facilitate extraction or examination of the biota.

Artificial substrata have the advantage over natural substrata of greater uniformity. Variation between samples caused by differences in the nature or amount of the sampled substratum will be reduced. The disadvantage is that it is never quite certain what the biota that colonizes the substratum actually represents. If the biota in the artificial substratum is depauperate, does this indicate a depauperate biota in the river or merely that the site had few organisms capable of colonizing the substratum?
Sample Size and Number

Because sample processing (sorting, identifying, and counting of invertebrates) is the most time-consuming part of a biological monitoring program, there is considerable pressure to minimize the number of invertebrates processed to the fewest consistent with the sensitivity required for the study. Two approaches have been used to achieve this aim. The first is to subsample or reduce the area sampled but always to sample a fixed area. The other has been to identify a fixed number of invertebrates regardless of the number present in the sample (as long as this number is larger than the number required).

There has been considerable discussion in the literature about the influence of fixed-sampling-area and fixed-count sampling on the sensitivity of the sampling program. The discussion has focused particularly on the effect on species richness comparisons, a common suite of metrics in water quality assessment (see below). This is not surprising in view of the extensive discussions in the ecological literature about species area relationships (e.g., Arrhenius, 1921; MacArthur and Wilson, 1967; Douglas and Lake, 1994). Courtemanch (1996) argued that sampling a fixed number of organisms would produce unstable estimates of taxa richness and that the measurements thus produced would, as a consequence, not be usable for comparison with reference values. They suggested three alternative strategies: whole-sample processing but with the sampling area adjusted through experience so that median size of the samples was manageable within time and budgetary constraints; a two-phase approach as suggested by Vinson and Hawkins (1996), which used a search of the whole sample for large, rare species followed by a subsample; or a serial processing technique where a fixed number of organisms is counted followed by a search of the whole sample for additional species.

Barbour and Gerritson (1996) argue that fixed-count sampling provides good comparability among samples and provides more effective discrimination between sites than fixed area samples. Vinson and Hawkins (1996) reanalyzed an existing data set to simulate fixed-count sampling, which they compared with the taxon richness based on total taxa identified using area-based sampling. They suggested that where fixed-count sampling was based on less than 150 individuals there may be a loss of sensitivity and decreased ability to detect real differences between collections but that above this number the sensitivity of the technique was similar to that of total counts as long as only comparisons of taxon richness are being made. Where other metrics, such as community similarity or presence or absence of specific indicator groups, are to be used, then fixed-count enumeration may have low power to discriminate between invertebrate assemblages.

Larson and Herlihy (1998) conducted a field test by sampling 35 wadable streams in Oregon. They collected 16 to 50 Surber samples from each stream, calculated species richness at three different sampling areas, and used a rarefaction equation (Hurlbert, 1971) and data from the pooled set of samples at each site to estimate the numerical taxon richness for various fixed counts. They found a high degree of correlation between the two when counts and areas were both high (counts = 500; area = 0.45 m²) but poor correlations ($R^2 = 0.54$) where numbers and areas were low (count = 100; area = 0.09 m²).

One of the most thorough assessments of sample size and processing methods on rapid assessment outcomes was conducted by Growns et al. (1997). They concluded that a selective 100-animal subsample was the most cost-efficient technique to obtain effective discrimination between unimpacted and mildly polluted sites around Sydney, Australia. Somers et al. (1998) also found that subsamples of 100 animals were sufficient to distinguish littoral benthic communities of small inland lakes in southern Ontario. They concluded that increased sampling effort (double or triple) produced little additional benefit but that not all the indices they tested were equally effective.
Sorting and Identification

Sorting the invertebrates from the detritus and other debris present in the sample and identifying them are the most time-consuming components of invertebrate monitoring. A number of methods can be used to speed both processes. Samples may be collected and preserved in the field or sorted in the field. Several field sorting methods have been advocated in the literature (e.g., by Chessman, 1995; Plafkin et al., 1989). Several advantages may accrue from field sorting. The first is time saved through the absence of double handling of samples; the second is that field sorting picks out live organisms. Some organisms, such as many of the stick-dwelling caddises, are easier to locate when they are alive and moving. Finally, field sorting reduces the exposure of sampling staff to preservatives such as ethanol and formaldehyde, commonly used to preserve samples, and reduces the total amount of preservative used because only a small volume of invertebrates is preserved rather than a large sample of invertebrates, water, and debris. The disadvantage of field sampling is a possible bias toward selecting large, active or otherwise conspicuous organisms from the sample.

Where field sorting is not used, samples returned to the laboratory may be fully sorted, so that all invertebrates are removed, using several types of aids. Staining with stains such as rose-Bengal makes invertebrates in the sample more conspicuous, and flotation using calcium chloride, kerosene, or sugar solutions may speed removal of many species, particularly in samples that contain relatively little plant material. However, samples still must be sorted manually for gastropods, bivalves, and many of the cased caddisflies, which cannot be floated off.

Sample processing time can be reduced by subsampling. Subsampling may be carried out as a field exercise; for example, rapid assessment techniques often involve field sorting and the collection of a fixed number of invertebrates from the sample (Chessman, 1995). Alternatively, the sample may be returned to the laboratory and then subsampled, sometimes after the removal of large items of debris. Common splitting techniques include the Folsom plankton splitter (e.g., Waters, 1969; McKaige, 1986), which divides the sample into two equal subsamples on each iteration, or quadrant-based subsamplers such as that developed by Marchant (1989).

Invertebrate Identification

The level to which invertebrates should be identified in biological monitoring exercises has been the subject of considerable debate (e.g., Resh and Unzicker, 1975). There are three sets of issues. The first is the extent to which information is lost when invertebrates are not identified, as often as possible, to species. The second is the cost of identification in both time and money. The third is the practicality of identification.

Several authors have argued that the use of species-level identification in biological monitoring improves the sensitivity of the monitoring, particularly to subtle impacts (e.g., Furse et al., 1987). Resh and Unzicker (1975) pointed out that many genera of aquatic invertebrates include species with quite different tolerances to organic pollution, for example. Thus they suggested that the presence or absence of particular genera at a site is not likely to be a useful indicator of the water quality.

However, other authors have compared the outputs of water quality assessments or macroinvertebrate community analysis using different levels of identification of the invertebrates. For example, Marchant et al. (1995) compared the outputs of two different multivariate analyses on a suite of data sets from nine rivers that were based on species-level resolution. They then created data sets with taxonomic resolution reduced to the generic and family levels, a data set converted from abundance to presence-absence, and finally, a
data set based just on EPT (Ephemeroptera, Plecoptera, and Trichoptera). They concluded
that the same patterns were evident with both ordination methods and all data sets. Thus
species-level identification does not appear necessary at least where reasonably strong
environmental gradients are present.

DATA ANALYSIS AND INTERPRETATION

Multivariate Systems

Biological monitoring data are almost always multivariate, consisting of lists of taxa and
their abundances (Resh and McElravy, 1993). Multivariate analyses assess the similarities
or dissimilarities between samples (Gauch, 1982), and results normally are expressed as
cladograms or, more popularly, ordination plots. A wide variety of multivariate algorithms
is available, and the results may differ appreciably depending on the algorithm employed.
Since multivariate techniques increasingly are included in standard statistical packages and
specialized multivariate statistical packages are becoming relatively cheap and accessible,
data processing is relatively simple. It is usually advisable to analyze data using several
algorithms to test whether the patterns found are robust, i.e., are features of the data rather
than artifacts of the analysis. Specific computer packages available for multivariate statis-
tical analysis include MVSP, PC-ORD, PRIMER, and PATN. Most multivariate methods
are techniques for finding patterns in data, not for hypothesis testing or assessing environ-
mental quality. They can identify a sample or a group of samples with invertebrate assem-
blages that differ from others included in the analysis. Some techniques, such as ANOSIM,
can determine whether the differences are statistically significant (Clarke, 1993), but they
cannot determine whether patterns are attributable to water quality. When environmental
data are analyzed along with the invertebrate data using techniques such as MDS or DEC-
ORANA, one can identify which environmental parameters are correlated with the inver-
tebrate data. However, correlation does not equate to causality.

Indices and Metrics

The alternative to analyzing multivariate data with multivariate techniques is to collapse the
information to a single index or metric. The earliest techniques for biological assessment of
water pollution, proposed by Kolkwitz and Marsson (1902, 1906), used such metrics. The
simplest measures are those concerned with species richness or species diversity. Of the two,
species richness measures generally are to be preferred because diversity measures gener-
ally compound species richness and evenness, making comparisons more difficult to inter-
pret (Hurlbert, 1971). It should be noted that although diversity indices are intended to be
independent of sample size, they are usually not. Figure 5.1 presents diversity spectra
(Margalef, 1968) developed by calculating diversity indices for benthic invertebrate samples
collected using a Surber sampler from riffles in the Yarra River, Victoria, Australia
(Campbell et al., 1982). Five samples were collected per site per occasion. Index values were
calculated for each individual sample, then for each possible combination of two samples
pooled, and then each possible combination of three and so on up to five samples. The mean
diversity values for individual samples and the various combinations were then plotted.
Effectively, this demonstrates how diversity index values change as sample size, in this
case equivalent to sampling area, increases. Figure 5.1 presents data for the Shannon and
Weaver index (1949), but the same pattern was found using a species richness index proposed by Margalef (Hellawell, 1978) and Pielou’s (1969) modification of a diversity index proposed by Simpson (1949) (Campbell et al., 1982). At 11 of 15 sites, index values increased appreciably as sample size increased, but of greatest concern were the rates of change; these differed between sites so that the relative rankings of sites changed as sample area increased.

While various measures of taxa richness are now widely used (Resh and Jackson, 1993), there is still considerable debate about which are most appropriate.

Specific metrics for assessment of water pollution use the relative abundances of specific taxa selected for their known tolerance of or sensitivity to water quality or other environmental quality parameters. The most elaborate example is the saprobien system originally developed by Kolkwitz and Marsson (1902, 1906) but since elaborated by Sladecek (1979) and others. Saprobic-based indices are still used in Germany (Persoone and DePauw, 1979) and parts of the Netherlands (Tolkamp, 1985). In Great Britain, the Trent Biotic Index (Woodiwiss, 1964) was developed in the early 1960s but was replaced by the BMWP system (National Water Council, 1981) in the 1980s. In South Africa, an index initially proposed by Chutter (1972) and subsequently further developed and is used widely (Dallas and Day, 1993).

In the United States, a number of different metrics have been proposed. Beck (1955) and Beak (1964) both proposed biotic indices based on indicator groups, Goodnight and Whitley (1961) proposed an index based on the abundance of oligochaets, and Patrick (1949) suggested a method based on species abundances. Most recently, the Index of Biotic Integrity (IBI) developed by Karr and others (Karr, 1981, 1991; Karr et al., 1986) has been used in a number of North American studies (Karr, 1999). The index was based initially on an assessment of the fish assemblage, but modified versions have been based on invertebrates (Karr, 1999).

A problem with any of the metrics proposed for water quality assessment is that of validation. There is no reason to expect that any biological metric will correlate precisely with, for example, chemical or physical water quality parameters. Indeed, if it did, there would be little need for the metric. The difficulty, then, is how to tell whether the results calculated

![FIGURE 5.1](image_url) Changes in diversity index value at three sites on tributaries of the Yarra River, Victoria, Australia, as the effective sampling area was increased. Note that one site with the lowest diversity if only single sample size was increased had the highest diversity of the five pooled samples that were considered.
using the metric are expressing an ecological reality or are simply an artifact of the metric. The literature on biologic water pollution assessment is replete with examples of modified metrics. In some cases, the metrics have been adapted through biogeographic necessity, as was the case when the IBI was modified for Australia (Harris and Silviera, 1999). In other cases, metrics appear to have been modified because the modified metric gave a better result, which presumably is one that more closely matches the subjective impression of the investigator.

Chessman and coworkers (Chessman, 1995; Chessman et al., 1997) have developed a very successful index called SIGNAL for use in Australia. Many of the validation shortcomings of other indices have been avoided (Chessman et al., 1997), and the index does appear to be particularly sensitive (Chessman, 1999).

Predictive Models

Predictive models are one of the more recent developments in water quality assessment. As noted previously, one difficulty in assessing quality of a site using invertebrates is the lack of information on what the invertebrate assemblage at the site would have been before significant human impact. The modeling approach uses a model to predict the assemblage composition. The actual assemblage found can then be compared with the assemblage expected and the degree of correspondence scored.

The approach was first developed by the Freshwater Biological Association in Great Britain (Armitage et al., 1992; Moss et al., 1999; Wright et al., 1984; Wright et al., 1989; Wright, 2000) and named RIVPACS. As many unaffected sites as possible were sampled, the invertebrates collected, and a variety of physical and chemical parameters recorded. In particular, physicochemical parameters were selected that were unlikely to be altered by the most frequent types of human impact. Thus latitude and longitude, alkalinity, gradient, substrate type, and distance from the source were among those recorded, whereas dissolved oxygen concentrations and phosphorus and nitrogen levels were not measured. Multiple Discriminant Analysis (MDA) (Klecka, 1975) was then used to identify the suite of physicochemical parameters that gave the best fit to the invertebrate assemblage classification. This suite could then be used to predict the assemblage composition at a test site by assuming that the probability of occurrence of any given taxon at a test site was a function of its frequency of occurrence within the appropriate reference classification group and its frequency of capture. The test site could then be sampled and the appropriate physicochemical data collected to allow a prediction of the test-site invertebrate assemblage. An index of quality can be determined using a ratio of the number of taxa observed to number of taxa expected at some predetermined level of probability.

The technique has passed through a number of iterations in Great Britain (Wright et al., 1989; Wright, 2000) and has been based on data from 614 sample sites in Great Britain and a further 70 sites in Northern Ireland. It also has been applied with apparent success in Spain (Armitage et al., 1990). In Australia, the method has been developed further by combination with rapid-assessment protocols similar to those developed in the United States by Plafkin et al. (1989) and named AUSRIVAS (Davies, 2000).

Rapid-Assessment Methods

A problem with traditional biologic assessment techniques has been the length of time required to process samples and interpret the data. This slowness is problematic because of its impact on the speed of feedback, which made these methods unsuitable for short-term impact monitoring. If the samples take weeks to months to process, the results are obtained too late to
inform many of the decisions about effluent standards. Slow processing is particularly labor-intensive, which made large-scale programs using these methods extremely expensive.

Two solutions have been applied. The most widely applied solution has been rapid field assessment methods such as those developed in the United States by Plafkin et al. (1989) and discussed previously. These methods attempt to reduce the time taken to sort and identify invertebrates.

**Alternative Monitoring Strategies**

Alternatives to community monitoring use particular species or species groups. One alternative to ambient community monitoring, which has a growing number of adherents, is the evaluation of fluctuating asymmetries.

Fluctuating asymmetries are minor morphologic deviations from normal symmetries, which are detected as nondirectional differences between left and right members of paired bilateral characters in animals (van Valen, 1962). They have been suggested as potentially sensitive biologic indicators of environmental stress (Valentine et al., 1973; Leary and Allendorf, 1989; Parsons, 1990; Clarke, 1993). In aquatic systems, this approach has been applied most widely to the impacts of toxic contaminants such as trace metals and organic toxicants on fish (e.g., Utayopas, 1996) and chironomid mouthparts (Wiederholm, 1984; Diggins and Stewart, 1998). Several studies that considered the total chironomid fauna failed to find a significant relationship between the frequency of chironomid mouthpart deformities and levels of toxicants in sediments. However, investigations limited to particular suites of sensitive species such as the *Chironomus thummi* group (Diggins and Stewart, 1998) and the *Chironomus plumosus* group (van Urk et al., 1992) have been more successful, in both cases finding significant regressions between percentage of deformed individuals and principal components analysis (PCA) factor scores for toxic contaminants in the sediments.

Another alternative that can provide very rapid feedback in a monitoring program is the establishment of some form of continuous monitoring, either in-stream using enclosures or on effluent prior to discharge into a receiving water. In-stream enclosure methods have been applied most frequently with sedentary species such as mussels (e.g., see Walker, 1981) but also with plants and fish (Hellawell, 1986). Continuous biologic effluent monitoring has been practiced most widely with fish (e.g., see Morgan, 1976; Dickson et al., 1980), but similar systems also have been used for oyster larvae (Roberts, 1980).

**Statistical Power Considerations in Environmental Monitoring**

Environmental monitoring of any kind is carried out most frequently with the purpose of detecting whether or not a change has occurred. Standard statistical tests are designed primarily to avoid type I errors that occur when a test shows that a change has occurred when in fact no change has occurred. In environmental monitoring, it is type II errors that are normally of greater concern; these occur when the test shows no change when in fact there has been a change (Fairweather, 1991). In much of our environmental monitoring, particularly of relatively unaffected systems, we are hoping that the system has not changed since the previous set of measurements and that degradation has not occurred. In order to have confidence that results showing no statistical difference reflect the ecological reality, we need to know how large an effect must occur before our sampling program can detect it.

Power analysis can be used to determine the type II error rate, as well as the minimum effect size that could have been detected with a given sampling design. Alternatively, it can be used to determine the number of samples necessary to detect an effect of a given size (Cohen, 1988). Power analysis is as yet a greatly underused tool in water quality monitoring.
REFERENCES


BIOLOGICAL MONITORING AND ASSESSMENT USING INVERTEBRATES


