Sub-sampling Techniques for Macroinvertebrates, Fish and Benthic Algae Sampled in Biological Monitoring of Streams and Rivers

Report No. 132 of the
Kansas Biological Survey

October 2005

By the
Central Plains Center for BioAssessment

Debra S. Baker
Donald G. Huggins

University of Kansas
Takeru Higuchi Building
2101 Constant Avenue, Room 35
Lawrence, KS 66047-3759
www.cpcb.ku.edu

Prepared in fulfillment of USEPA Award X7-99790401

INTRODUCTION

As the broad-scale use of biological monitoring and assessment increases in both the regulatory and research communities, the need for accurate, precise and cost efficient methods becomes more important. Subsampling of macroinvertebrates, fish, or benthic algae samples is one solution to address time and cost considerations. While researchers and regulators have employed subsampling over the last 50 or more years, more often entire samples are sorted and identified at great expense and time. However, recent trends in benthic macroinvertebrate sampling (e.g. Rapid Bioassessment Protocols, Barbour et al. 1999) have popularized the use of subsampling techniques. Davis (1996) indicated that 30 of the 44 states with bioassessment programs that use macroinvertebrates as the assessment community use some form of subsampling approach to economize and expedite sample processing and data production. Unfortunately, in many programs, subsamples are employed without enough thought (or statistical rigor) given to the consequences of examining only a portion of the sample. If time saving methods such as subsampling are applied to biological monitoring programs without prior analysis of the accuracy or precision of such methods, the information collected may be useless, resulting in a waste of resources or worse, the implementation of regulatory actions based on incorrect decisions. Conversely, the application of labor-intensive and time-consuming surveys are impractical for most bioassessment agencies, which may be responsible for the water quality monitoring of hundreds to thousands of streams (Lenat and Barbour 1994). In conservation or inventory studies, the collection of all or most species and the determination of their relative abundances may be required, but for biological monitoring studies only enough of the community needs to be collected and identified to determine differences among sites. One must

consider how much of the community must be sampled and whether this can be accomplished using only a portion of a sample.

The purpose of this paper is to describe existing subsampling techniques and their associated characteristics and to develop recommendations to aid the decisions of biological monitoring investigators. Macroinvertebrate, fish, and benthic algae subsampling is explored, however subsampling of macroinvertebrates constitutes the bulk of this paper as the literature and techniques of macroinvertebrate subsampling are extensive and diverse. Thus, many considerations of subsampling in general are covered in the macroinvertebrate section and one should read this section even if interested in fish or algae.

MACROINVERTEBRATES

The practice of subsampling has been prevalent for many years as a means to reduce the time and effort required to sample aquatic systems in order to increase the coverage of biological monitoring programs and to improve the feasibility of studies. Although subsampling is common practice, there is little consensus regarding which method is the most efficient and the least prone to bias and/or error. One of the major difficulties encountered by bioassessment workers is the wide variety of sampling techniques employed by multiple agencies. When designing a sampling protocol for lotic systems, investigators must make several important decisions that will impact their sampling strategy. Sampling strategy depends on purpose of the survey, method used by nearby monitoring agencies, expertise of collectors, use of quantitative or qualitative data and characteristics of the streams in the particular geographic area (Lenat and Barbour 1994). In order to select a subsampling method (or lack thereof), several factors must first be considered.

1. Sampling Decisions

A. Habitat Type

A decision faced by researchers in developing a monitoring and assessment plan is the adoption of a single-or multiple-habitat sampling design. Some programs recommend the use of single-habitat (usually riffles since the greatest diversity and abundance of algae and invertebrates are generally found there) sampling to decrease sampling costs, lessen variability among samples and to improve standardization. However, the use of single-habitat sampling is not appropriate for some ecoregions or ecosystems, particularly for lotic ecosystems that have no riffles or are not dominated by riffles. Invertebrate communities in non-riffle habitats can make important contributions to a stream reach and may better represent the structural and functional properties of these systems (Lenat 1988). In many cases, in some streams and regions, sampling only riffles is simply not feasible due to a lack of riffle habitat. Many investigators recommend the use of multiple-habitat sampling based on the importance of each type of habitat in a stream reach as it is applicable to all streams and it permits the sampling of a larger proportion of the taxa present at a site (Vinson and Hawkins 1996).

It is also worthwhile to point out here that something as simple as the mesh size of the sampler is important to consider as the mesh sizes influences the portion of the community collected by the sampler. For example, small-sized organisms and/or smaller early instars would not be retained by a sampler with a large mesh, therefore affecting the proportion of the sample required to discriminate among sites.

B. Processing Location

With the recent increase in the popularity of Rapid Bioassessment Protocols (RBPs), some agencies have decreased labor and time required to sort macroinvertebrate samples by sorting samples in the field. Field sorting may also have some utility for remote and rugged areas where the transport of samples is a limitation (Growns et al. 1997). However, traditionally, samples are collected and fixed in the field and sent to the laboratory for sorting, enumeration and identification. It has been argued that field sorting causes a large amount of error and limits modifications in sampling techniques. For instance, some agencies may wish to limit the count of abundant organisms in order to sample a larger proportion of the diversity, but this is impossible in the field without experienced workers, particularly for identifications below family. Although there are no studies utilizing statistical analyses that look at field versus laboratory sorting, the savings gained through field sorting is probably not an equal trade off for the loss of accuracy and precision and the introduction of bias. Since it is likely that the use of field sorting causes some degree of bias and error, its use should be limited until a proper statistical examination is performed.

C. Taxonomic Resolution

The choice of taxonomic resolution can also have a strong effect on subsampling decisions. Increasing or decreasing the taxonomic resolution (e.g. family vs. genus) of samples will result in a concurrent change in the percentage of taxa encountered in a sampling regime, which could permit the use of smaller samples (Bouchard et al. 2005).

2. Whole Samples

Given adequate time, money, and personnel, the sorting of an entire sample is the most straightforward method to avoid the error and bias associated with subsampling. There is little argument that given sufficiently large samples, the processing of whole samples can produce a more accurate determination of taxa richness at a site. Additionally, the use of whole samples is considered by some investigators to be the only suitable method to assess macroinvertebrate communities in lotic systems as whole counts produce the best return for the inherent high cost of sampling streams (DePauw and VanHooren 1983, Wright et al. 1993, Courtemanch 1996, Cao et al. 1998, Stroom and Richards 1999, Doberstein et al. 2000). Total abundance measures are possible with whole sample counts and if the sampling area or effort can be quantified, one can additionally quantify abundance or richness measures on a per unit of measure basis (Courtemanch 1996). Data collected by units of space or time permit standardization of data variables that allow comparisons between investigators and agencies when similar collecting procedures are employed. Subsampling techniques often result in biases due to differences among sorters or as a result of the subsampling devises, but through the examination of entire samples, much of this bias can be eliminated given competent sorters. Doberstein et al. (2000) argues that whole samples are the best way to avoid errors that potentially result in the incorrect assessment of a stream reach, thereby causing costly regulation or conservation mistakes. However, there are a number of problems inherent in the determination and use of whole samples.

The most obvious impediment to the use of whole samples is the large amount of time and effort often required to sort and identify organisms from an entire sample. The large amount

of effort is almost never proportional to the information gained for sorting the entire sample. Full count samples may also impair the ability to compare samples between investigators and agencies if there is no attempt to standardize sampling procedures. Courtemanch (1996) suggests that these difficulties can be minimized with the use of appropriately sized standard samples that estimate the sample size needed to collect a sufficient number of organisms. Although it is not possible to determine a sample size that will yield a constant number of organisms in each sample, it may be feasible to determine the size that will give relatively consistent abundances for specific regions (and possibly aquatic ecosystems types). This approach suggests that controlling the physical area collected will control the number of organisms collected and therefore the number that would have to be counted. This eliminates the need to subsample, but introduces possible variance due to changes in sample size (not the number of organisms, but the area measured). Once samples are collected there is no way to repeat or duplicate the sample so samples with too few organisms will be useless while others may have such high densities that time and effort required to sort samples will be prohibitive. Besides collecting a larger proportion of the community in areas where densities are patchy, another advantage of collecting large samples and then subsampling is that it allows the size of subsample to be corrected to account for a sufficient portion of the community.

3. Subsampling

A. Enumeration Type

RBPs have also increased the use of qualitative or semiqualitative enumerations to reduce time and effort. Semiqualitative or ordinal counts are a method for estimating relative abundances by counting macroinvertebrates as abundant, common or rare (Lenat 1988, Plafkin et

al. 1989). The method reduces the time necessary to pick and count a sample so that a larger portion can be analyzed thereby increasing the diversity sampled and improving the ability to detect rare species. Semiqualitative counts also permit some analyses using estimated abundances or densities (i.e. relative abundance of tolerant species) in water impairment detection (Lenat 1988). Qualitative enumerations are also commonly used to reduce time and effort and to allow investigators to focus on taxa richness as a measure of stream quality. Qualitative counting methods are considered acceptable because abundance measures are prone to a large amount of variation, even in undisturbed systems, thereby reducing their ability to detect perturbation.

B. Fixed Fraction

Fixed fraction subsampling is the traditional (i.e. most common, long history, well established) subsampling method in which a standard proportion of the sample (e.g. ¼, or ½) is chosen as the subsample of interest. In most cases a standard proportion based on total sample volume is removed from each sample, although a minimum amount of material or number of organisms is usually required. An alternative to division of total volume or area is to subsample based on weight allocations. This approach has been proposed for samples with a large amount of algae or other material (Sebastien et al., 1988). Sampling devices for fixed fraction subsampling are numerous and diverse. For methods on subsampling devices that methods and have been used to subsample very small (e.g. zooplankton) to rather large organisms, refer to Waters 1969, Hynes 1970, Hickley 1975, and Wrona 1982. One advantage of the use of a fixed fraction method is that it yields an estimation of areal taxa richness or richness per unit area (Hurlbert 1971, Barbour and Gerritsen 1996). This areal taxa richness is considered important as

it provides accountability for taxa richness and allows taxa richness density to be quantified per unit area or volume, permitting comparisons (Courtemanch 1996). However, abundance metrics are considered highly variable even in the absence anthropogenic impact and thus has inherently high noise to signal ratio (Hynes 1970, Lenat 1988, Resh and Rosenberg 1989, Lenat 1990, Lenat and Barbour 1994, Barbour and Gerritsen 1996).

C. Fixed Count

The fixed count subsampling method consists of a random picking of a fixed number of organisms (i.e. 100, 200, 300, etc.). Most fixed count subsamplers are performed using a gridded pan, which inexpensively and simply allows a random selection of a standard number of organisms to be picked from a sample (Hilsenhoff, 1987). A sample is added to the gridded pan and dispersed evenly on the bottom. Several random grids (four or more to ensure proper representation) are selected and picked of all organisms until the standard count is passed or approached (Barbour and Gerritsen 1996). Barbour and Gerritsen (1996) recommend that the number of organisms picked during subsampling needs to remain within 20% of the targeted number.

Fixed count subsamples have become widespread and are an important part of RBP protocols. Initially designed for use with plankton, the method was adapted by Hilsenhoff (1987, 1988) for use with freshwater benthic samples and later modified by Plafkin et al. (1989) for use in RBPs (Barbour and Gerritsen 1996). The use of the fixed count method reduces time and effort required to process samples, particularly for samples with high abundances (Growns et al. 1997; Larsen and Herlihy 1998). Some researchers have indicated that the use of subsampling does not compromise data quality (Barbour and Gerritsen 1996; Somers et al. 1998). However, most researchers suggest that the reduction in time and effort is an acceptable trade off from the

complete enumeration of samples as the sorting of whole samples does not contribute proportionally to the cost of whole sampling processing. However, due to this error, many recommend the use of fixed count subsamples when they are employed cautiously and enough analysis is employed to minimize error (Resh and Jackson 1993; Vinson and Hawkins 1996; Walsh 1997). Besides a reduction in time and effort required to process samples, it has also been argued that fixed count subsampling also improves standardization, allowing comparisons between agencies (given similar sampling methodologies) as fixed counts are achievable in a variety of stream types and are feasible for most agencies. The improvement of standardization is related to the ease of obtaining multiple investigator and agency agreements and the ability to sample those habitats that provide contributions to the stream community, which are difficult to express in area (e.g. snags). It has also been suggested that the use limited taxa counts produces greater differences between polluted and unpolluted sites (Growns et al. 1997).

In order to further reduce sampling time and effort as well as the size of samples returned to the laboratory, in-field fixed count subsampling is sometimes employed (Chessman 1995; Growns et al. 1997). In-field (Chessman and Robinson 1987; Chessman 1995; Growns et al. 1997) or laboratory (Wright et al. 1997) methods may also involve the use of timed count procedures. However, timed counts are not utilized often as timed count procedures introduce uncontrollable bias due to worker variations and large organism bias (Growns et al. 1997). For example, a more experienced sampler can pick a greater number of species than a less experienced sampler in the same period of time (Growns et al. 1997). There are problems with bias against cryptic or relatively immobile taxa for both timed and fixed count in-field methods. The use of limited counts for abundant taxa reduces some of the error associated with in-field

counts (Growns 1997). However, this method will be limited to high level identifications (e.g. family), which may be dependant on the expertise of the sampler.

The most important argument for the use of fixed count subsamples is the idea that if subsamples can reach the asymptote of the collector's curve for most samples in a bioassessment program an accurate description of macroinvertebrate community can be made. Most collections can be described using the collector's curve, where a regression of taxa richness over number of organisms enumerated (or sorting effort) produces an asymptotic curve that is characterized by a rapid initial increase in taxa richness followed by a leveling out where few new taxa are enumerated. The idea behind the successful use of a fixed count method is that the fixed number of organisms is sufficient to reach the asymptote where an increase in the subsample size does not greatly increase taxa richness. Essentially, a laboratory will want to sort enough organisms to reach the asymptotic portion of the curve in order to insure that most of the abundant and common taxa are enumerated in for most samples. If the subsample count is in the steep portion of the curve, samples can be highly variable. Unfortunately, the count necessary to reach the collector's curve asymptote is variable as a result of total abundance differences and sampling variations. The shape of the curve is also affected by the distribution of the relative abundance of the macroinvertebrate communities (i.e. a few numerically abundant taxa and many rare taxa or a relatively even distribution of abundance for all species) (Larsen and Herlihy 1998). Basically there is need to determine the count number that will reach the asymptote that encompasses all or most of the stream communities from a variety of stream types affected by gradients of various forms of perturbation in a given region.

A drawback of fixed count subsampling is that the use of total abundance measures is not permitted. Total abundance metrics can be useful when considered with other metrics (e.g.

richness) as some types of pollution increase total macroinvertebrate abundance, but there is generally an associated decrease in taxa richness. However, the loss of the ability to employ total abundance metrics (considered a highly variable metric anyway) may be a justifiable compromise for increased standardization. The fixed count method also results in the loss of a sizable proportion of the rare species, but this characteristic should have little affect on the interpretation of the community as rare taxa generally contribute little to studies that detect differences in stream communities (Marchant 1989, Barbour and Gerritsen 1996).

A more troubling problem of fixed counts is the lack of account of richness per unit area for a given area or volume, which Hurlbert (1971) and Courtemanch (1996) argue is the most meaningful method to describe macroinvertebrate communities. A measure of richness should be accompanied by a measure of unit area or sampling effort. In order to account for richness per unit, adjustments can be used (e.g. rarefaction) or the standard, quantitative sample must remain constant (Williams and Gaston 1994). For instance in enriched systems, densities will be higher compared to an unperturbed system, which means that a larger number of individuals will need to be picked to reach the asymptote in the enriched system (Courtemanch 1996). Consequently, given that both systems have an equal number of species, more taxa will generally be picked from the undisturbed system thereby overestimating richness. The opposite would be true of a system experiencing increased stress as the result of a toxic substance where total abundances are decreased compared to an undisturbed system. In this case the taxa richness would be overestimated in the perturbed system. Essentially, by using the fixed count method, a fixed number of organisms collected at one site represents a different area unit compared to another site with differing densities (Courtemanch 1996). The degree of variation in macroinvertebrate densities may have an effect on the success of a sampling program employing

subsampling procedures. However, if subsamples are sufficient to count enough taxa to reach to extend beyond the collector's curve asymptote for most samples, the hope is that variance resulting from density differences will be minimal. The shape of a collector's curve determines how much error is present after the asymptote is reached. If the relative abundances are close to being even among taxa, the asymptote will not be reached until all or most taxa have been collected, a large amount of error will be introduced. Conversely, if the community is dominated by a few abundant species and a lot of rare species, the asymptote will be reached quickly and error will be minimal. It is assumed that the majority of lotic systems are more similar to the latter abundance pattern, but there is still variability in this pattern among streams that will contribute various levels of error. These relationships need to be examined along pollution gradients and for various stream types in each region in order to predict the number of individuals needed to reach the asymptote.

Modifications of the fixed count subsampling method have come about to attempt to eliminate some of the error associated with this form of subsampling and to allow taxa richness to be expressed per unit area. A two-phase subsampling method is suggested by Cuffney et al. (1993) and Vinson and Hawkins (1996), which consists of an initial search for large or rare organisms followed by a standard subsampling procedure. This compromise selectively allows the use of accurate taxa richness metrics as standardized sample size are often considered important at a minimum of cost (Courtemanch 1996).

4. Recommendations for Subsampling

The sorting and identification of macroinvertebrates from entire large samples is probably not a feasible solution for most agencies. Many papers indicate that subsampling can

be employed cautiously or even without a drop in data quality. Unfortunately, as pointed out by Doberstein et al. (2000), many studies on the effect of subsampling are insufficient to determine if it is possible to use subsamples in benthic macroinvertebrate monitoring studies. Bioassessment protocols only need to determine the most efficient method that accurately and precisely describes stream reaches that fits the purposes of the study. The development of standard protocols for ecoregions that are tailored to the suite of streams (from reference to enriched to toxic systems and stream size) encountered in the region will promote the sharing of information and thus advance the science of bioassessment.

FISH

1. Subsampling Decisions

As with macroinvertebrates, a sampling design for fish studies depends on the study objectives. Yoder and Smith (1999) provide an excellent review of the various sampling methods for fish, and how study objectives and stream conditions (size, geography, etc.) determine the most appropriate methods to use. Also, Standard Methods (1998) provides general guidelines as to study design, gear type, and data analysis. Rather than repeating the information, only the generalities are present here. They recommend that only a single sampling gear type be used, and that where electrofishing can be used, it is much more effective than seining.

2. Taxonomic Resolution and Enumeration

Unlike macroinvertebrates and benthic algae, fish samples are of such a nature that the entire sample can relatively easily be enumerated and identified to species in the field (Yoder

and Smith 1999). Thus, subsampling is not an issue. In addition, other measurements typically taken in the field are weight, length, and notes of external anomalies (Yoder and Smith 1999).

ALGAE

1. Sampling Decisions

Subsampling algae presents many of the same considerations as with macroinvertebrates and fish: single- vs. multi-habitat sampling (Rosen 1995), taxonomic resolution, etc. (see Aloi 1990 for a review of field methods and Standard Methods (1998) for sample collection and laboratory analytical methods.). Substrate type will determine the type of sampler used. The RBP manual (Barbour et al. 1999) presents three methods of sampling periphyton, based on substrate type. The Central Plains Center for BioAssessment developed a periphyton sampler that is adaptable to various substrates (Anderson and Bouchard 2000). In all these cases the sample equipment samples a known area, and samples can be standardized by unit area.

2. Taxonomic Resolution

Once in the lab, the researcher must decide what taxonomy to identify and enumerate. Identifying a limited taxonomic assemblage will save time and money, but at the cost of loosing ecological information. The RBP manual (Barbour et al. 1999) recommends only identifying diatoms of which the species show diverse ecological preferences. Additionally, taxonomic assemblage and level identified may come down to the expertise of the taxonomist. (Barbour et al. 1999).

3. Enumeration

By necessity, enumeration of periphyton requires subsampling. Unlike macroinvertebrates and fish, subsampling involves examining a specific volume of water, thus results can be standardized and samples can be more easily compared than macroinvertebrates and fish samples (Barbour et al. 1999). One must, however, pre-determine how many specimens will be enumerated. DARES 2004 sets a protocol on taxonomic abundance. If one taxonomic group represents the bulk of the subsamples, then count additional specimens until a specific count is reached. Specific procedures for subsampling and enumeration are presented in Rosen 1995, Barbour et al. 1999, Bahls 1993, and DARES 2004. Rationale of subsampling for time and cost efficiency versus adequate counts for statistical and taxonomic analyses is presented in the section on macroinvertebrates in this paper.

RECOMMENDATIONS

The ideal biomonitoring macroinvertebrate, fish, or benthic algae sampling regime would involve a large number of whole samples from a known area, sampled from all of the important habitats within a reach with taxa identifications taken to species. However, a protocol such as this is unachievable for biological monitoring purposes, so compromises must be made to permit the sampling of a large number of sites in a relatively short period of time in order to obtain accurate and relevant information at a minimum of cost. A large amount of money and effort would be required to collect and identify all (or even most) species in a stream reach and to determine their relative abundances. This sampling regime would be impossible or would at least limit agencies to such as small number of sites that little could be accomplished in terms of biological monitoring. Conversely, a very rapid, simplified sampling protocol may not have

enough power to detect the difference between subtle levels of pollution or even gross pollution, making it useless and a waste of time and resources. A better approach is the establishment of a middle ground or better yet, the determination of the lowest effort that will provide relevant information.

Problems with Subsampling Studies

A difficulty of the subsampling dilemma is the lack of appropriate studies addressing the subject. Doberstein et al. (2000) points out that some of the proponents of 100-count subsampling (e.g. Barbour and Gerritsen 1996) use examples from lentic systems that may not be comparable to lotic systems for two reasons: (1) the richness and relative abundances may not be equivalent in both systems and; (2) the communities in both systems may not respond in similar manners. Several papers arguing for fixed counts also do not use multiple subsamples from each site to assess the variability among replicates. Other studies (e.g. Barbour and Gerritsen 1996; Vinson and Hawkins 1996) only examine one or two metrics (taxon richness and/or total abundance) as a basis for their conclusions (Doberstein et al. 2000). In order to fully understand the effect of subsampling, all possible metrics potentially used in biological monitoring studies, must be assessed (Doberstein et al. 2000). Somers et al. (1998) indicated that 100 organism counts are sufficient to detect differences and do not compromise results, but since the results were not compared to whole count samples, these conclusions are inconclusive (Doberstein et al., 2000). Some studies do not look at the effect of pollution gradients on subsampling results (Doberstein et al., 2000). Doberstein et al. (2000) also points out that the use of ANOVA to evaluate the effect of subsamples is problematic because subsamples violate the variance of

homogeneity assumption required for most parametric tests and a large number of replicates are needed (>100) to estimate the true population variance and error resulting from subsampling.

How much difference is enough?

In many biological monitoring studies, the differences among sites are determined or sites are compared to a baseline or reference condition to assess the level of impact. However, the question should be asked regarding what constitutes an important difference among sites and how can that be used to determine actual anthropomorphic impact? When comparing a reference site to a study site, looking at the difference between the two sites without any knowledge of how much of the observed difference is the result of sampling error and natural factors makes monitoring difficult because unless there is severe impact at the study site, decisions are based on conjecture. For instance, if there is a 20% difference between two sites, how much is a result of actual perturbations and how much is the result of sampling errors and natural or stochastic factors? Furthermore, how can recommendations regarding regulatory decisions be made without that knowledge? The RPB (Plafkin et al. 1989) uses 83% similarity (17% difference) to indicate difference between reference and study sites. However, it has been indicated that 83% is possibly too high and does not account enough for sampling or natural differences (Resh and Jackson 1993). Hannaford and Resh (1995) uses both 83% and 65% similarity to detect impact in the effectiveness of the RBPs. These numbers are rather arbitrary and do not represent any real threshold. Once there is some account for the importance of the difference or similarity between sites, the development of sampling methodologies would become clearer since a percent difference could be identified as meaningful. It is unlikely that any specific percent difference could ever be considered more than arbitrary (much like the choice of the alpha value 0.05) but

through analysis of relevant percent similarities, these values could be tailored to regions and stream types. This reasoning could help to reduce error associated with misrepresenting differences between site samples.

Power-Cost Efficiency

The decision of subsampling methods need be evaluated from a cost efficient point of view in order to maximize the number of streams or sites that can be monitored while maintaining a high degree of power to detect relevant differences between sites. Traditionally, rarefaction was used to determine the sufficient sample size required to reach the collector's curve asymptote. Ferraro and Cole (1989) use Power-Cost Efficiency (PCE) as a method to determine the least costly sampling regime that provides the higher power for detecting desired differences between sites. Since the PCE analysis uses a t-test, it may not be a good idea to apply PCE to subsampling. PCE may be good to determine the best subsampling methods, but may not stand up statistically. Barbour and Gerritsen (1996) apply PCE to subsampling and give an explanation.

CONCLUSION

Although in bioassessment, especially in terms of regulation, investigators would like to determine the sampling scheme that performs best for all regions, habitats, and study objectives, this ideal is an impossible objective (Ferraro et al. 1989). As information and studies continue to improve our current knowledge of aquatic ecosystems and the effectiveness of methods used to sample the communities and identify perturbations, it becomes necessary to attempt to develop some standardization. Without standardization, the growth of bioassessment will be limited as

the sharing of information and ideas between practitioners and agencies is becoming more important to advance our current knowledge of complex lotic systems and their macroinvertebrate communities.

Investigators must realize that it is impossible to completely describe the biotic community in a reach of stream. Rather, it is important to characterize the stream in a way that allows it to be separated from other streams or sites in a relevant manner for use in bioassessment. For example, Vinson and Hawkins (1996) determined that although 100-200 counts greatly underestimate the true richness, they are robust enough to detect differences between sites. The question needs to be asked, how much difference between sites is important for bioassessment.

There will be no standard method that will fit all regions (this ideal can never be achieved), but within regions protocols can tested and designed to reduce variation associated with sampling. The resulting standardization will ease the pooling of information from multiple agencies thereby improving biological monitoring programs. The attempt to thoroughly describe a stream community in a given reach is impractical and faulty in its reasoning. It is not necessary, since in terms of regulation and management of aquatic systems, it is only important to collect information required to separate sites or streams based on perturbations in a meaningful manner. It is possible that this may be accomplished when studies rigorously examine the effects of subsampling on sampling benthic macroinvertebrates. However, the use of subsampling needs to be examined on a regional basis to determine the most cost efficient method to sample and monitor lotic systems.

LITERATURE (both cited and other sources)

Aloi, J. E. 1990. A critical review of recent freshwater periphyton field methods. Canadian Journal of Fisheries and Aquatic Sciences. 47: 656-670.

APHA, AWWA and WPCF. 1998. Standard Methods for the Examination of Water and Wastewater. 20th ed. Amer. Publ. Health Assoc., Washington, D.C. 1134 pp.

A.R.C. 1998. Analysis of benthic invertebrate samples from the region 1 pilot study for the determination of biocriteria reference conditions. Aquatic Resources Center, Franklin, TN.

Bahls, L. L. 1993. Periphyton bioassessment methods for Montana streams. Montana Water Quality Bureau, Department of Health and Environmental Science, Water Quality Bureau, Helena, Montana.

Barbour, M. T., J. Gerritsen, G. E. Griffith, R. Frydenborg, E. McCarron, J. S. White, M. L. Bastian. 1996. A framework for biological criteria for Florida streams using benthic macroinvertebrates. Journal of the North American Benthological Society 15:185-211.

Barbour, M. T., J. Gerritsen, B. D. Snyder and J. B. Stribling. 1999. Rapid bioassesment protocols for use in streams and wadeable rivers: periphyton, benthic, macroinvertebrates and fish, second edition. EPA 841-B99-002. U.S. Environmental Protection Agency; Office of Water, Washington D.C.

Barbour, M. T., J. Gerritsen. 1996. Subsampling of benthic samples: a defense of the fixed-count method. Journal of the North American Benthological Society 15:386-391.

Bouchard, R. W. Jr., D. G. Huggins, and J. Kriz. 2005.A Review of the Issues Related to Taxonomic Resolution in Biological Monitoring of Aquatic Ecosystems with an Emphasis on Macroinvertebrates. USEPA white paper in progress.

Cao, Y., D. D. Williams and N. E. Williams. 1998. How important are rare species in aquatic community ecology and bioassessment? Limnology and Oceanography 43:1403-1409.

Chessman, B. C. 1995. Rapid assessment of rivers using macroinvertebrates: a procedure based on habitat-specific sampling, family level identification and a biotic index. Australian journal of ecology 20:122-129.

Ciborowski, J. J. H. 1991. Estimating processing time of stream benthic samples. Hydrobiologia 222:101-107.

Courtemanch, D. L. 1996. Commentary on the subsampling procedures used for rapid bioassessments. Journal of the North American Benthological Society 15:381-385.

DARES. 2005. Enumeration of Diatom Samples. http://craticula.ncl.ac.uk/DARES/methods.htm

Davis, W. S., B. D. Snyder, J. B. Stribling and C. Stoughton. 1996. Summary of state biological assessment programs for streams and wadeable rivers. US Environmental Protection Agency; Office of Policy Planning and Evaluation, Washington D.C.

DePauw, N. G. V. 1983. Method for biological quality assessment of watercourses in Belgium. Hydrobiologia 100:153-168.

Diamond, J. M., M. T. Barbour and J. B. Stribling. 1996. Characterizing and comparing bioassessment methods and their results: a perspective. Journal of the North American Benthological Society 15:713-727.

Doberstein, C. P., J.R. Karr and L. L. Conquest. 2000. The effect of fixed-count subsampling on macroinvertebrate biomonitoring in small streams. Freshwater Biology 44:355-371.

Ferraro, S. P., F. A. Cole, W. A. DeBen and R. C. Swartz. 1989. Power-cost efficiency of eight macrobenthic sampling schemes in Puget Sound, Washington, USA. Canadian Journal of Fisheries and Aquatic Sciences 46:2157-2165.

Ferraro, S. P. and F. A. Cole. 1990. Taxonomic level and sample size sufficient for assessing pollution impacts on the Southern California Bight macrobenthos. Marine Ecology Progress Series 67:251-262.

Furse, M. T., J. F. Wright, P. D. Armitage and D. Moss. 1981. An appraisal of pond-net samples for biological monitoring of lotic macro-invertebrates. Water Research 15:679-689.

Growns, J. E., B. C. Chessman, J. E. Jackson and D. G. Ross. 1997. Rapid assessment of Australian rivers using macroinvertebrates: cost and efficiency of 6 methods of sample processing. Journal of the North American Benthological Society 16:682-693.

Hannaford, M. J. and V. H. Resh. 1995. Variability in macroinvertebrate rapid-bioassessment surveys and habitat assessments in a northern California stream. Journal of the North American Benthological Society 14:430-439.

Hartley, J. P. 1982. Methods for monitoring offshore macrobenthos. Marine Pollution Bulletin 13:150-154.

Heck, K. L., G. Van Belle and D. Simberloff. 1975. Explicit calculation of the rarefaction diversity measurement and the determination of sufficient sample size. Ecology 56:1459-1461.

Hellawell, J. M. 1977. Change in natural and managed ecosystems: detection, measurement and assessment. Proc. B. Soc. Lond. B. 197:31-57.

Hewlett, R. 2000. Implications of taxonomic resolution and sample habitat for stream classification at a broad geographic scale. Journal of the North American Benthological Society 19:352-361.

Hickley, P. 1975. An apparatus for subdividing benthos samples. Oikos 26:92-96.

Hilsenhoff, W. L. 1977. Use of arthropods to evaluate water quality of streams. Department of Natural Resources, Madison, WI. pp. 15.

Hilsenhoff, W. L. 1982. Using a biotic index to evaluate water quality in streams. Department of Natural Resources, Madison, WI. pp. 22.

Hilsenhoff, W. L. 1987. An improved biotic index of organic stream pollution. The Great Lakes Entomologist 20:31-39.

Hilsenhoff, W. L. 1988. Rapid field assessment of organic pollution with a family-level biotic index. Journal of the North American Benthological Society 7:65-68.

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 12:359-364.

Hurlbert, S. H. 1971. The nonconcept of species diversity: a critique and alternative parameters. Ecology 52:577-586.

James, R. J., S. M. P. Lincoln and P.G. Fairweather. 1995. Sieve mesh-size and taxonomic resolution needed to describe natural spatial variation of marine macrofauna. Marine Ecology Progress Series 118:187-198.

Kingston, P. F. and M. J. Riddle. 1989. Cost effectiveness of benthic faunal monitoring. Marine Pollution Bulletin 20:490-496.

Larsen, D. P. and A. T. Herlihy. 1998. The dilemma of sampling streams for macroinvertebrate richness. Journal of the North American Benthological Society 17:359-366.

Lenat, D. R. 1988. Water quality assessment of streams using a qualitative collection method for benthic macroinvertebrates. Journal of the North American Benthological Society 7:222-233.

Lenat, D. R. and M. T. Barbour. 1994. Using benthic macroinvertebrate community structure for rapid, cost-effective, water quality monitoring: rapid bioassessment, p. 187-215. *In:* Biological monitoring of aquatic ecosystems. S. L. L. A. Spacie (ed.). Lewis Publishers, Boca Raton.

Marchant, R. 1989. A subsampler for samples of benthic invertebrates. Bulletin of the Australian Society of Limnology 12:49-52.

Marchant, R., L. A. Barmuta and B. C. Chessman. 1995. Influence of sample quantification and taxonomic resolution on the ordination of macroinvertebrate communities from running waters in Victoria, Australia. Marine and Freshwater Research 46:501-506.

May, R. M. 1975. Pattern of species abundance and diversity, p. 81-120. *In:* Ecology and evolution of communities. M. L. C. J. M. Diamond (ed.). Harvard University Press, Cambridge.

McIntyre, A. D., J. M. Elliott and D. V. Ellis. 1984. Chapter 1. Introduction: Design of sampling programs, p. 1-26. *In:* Methods for the study of marine benthos. Vol. 16. N. A. H. A. D. McIntyre (ed.). Blackwell Scientific Publishers, Oxford, UK.

Merritt, R. W. and K. W. Cummins. 1996. An introduction to the aquatic insects of North America. Kendall/Hunt, Dubuque, Iowa. pp. 862.

Morin, A. 1985. Variability of density estimates and the optimization of sampling programs for stream benthos. Canadian Journal of Fisheries and Aquatic Science 42:1530-1534.

Plafkin, J. L., M. T. Barbour, K. .D. Porter, S. K. Gross and R. M. Hughes. 1989. Rapid bioassessment protocols for use in streams and rivers: benthic macroinvertebrates and fish. U.S. Environmental Protection Agency, Washington, D.C. pp. 88.

Pringle, J. D. 1984. Efficiency estimates for various quadrat sizes used in benthic sampling. Canadian Journal of Fisheries and Aquatic Sciences 41:1485-1489.

Resh, V. H. 1994. Variability, accuracy and taxonomic costs of rapid assessment approaches in benthic macroinvertebrate biomonitoring. Bollettino di Zoologia 61:375-383.

Resh, V. H. 1995. Freshwater benthic macroinvertebrates and rapid bioassessment procedures for water quality monitoring in developing and newly industrialized countries, p. 167-177. *In:* Biological assessment and criteria: tools for water resource planning and decision making. W. S. D. T. P. Simon (ed.). CRC Press, Boca Raton.

Resh, V. H., D. M. Rosenberg, and J. W. Feminella. 1985. The processing of benthic samples: responses to the 1983 NABS questionnaire. Bulletin of the North American Benthological Society 2:5-11.

Resh, V. H. and E. P. McElravy. 1993. Contemporary quantitative approaches to biomonitoring using benthic macroinvertebrates, p. 159-194. *In:* Freshwater biomonitoring and benthic macroinvertebrates. D. M. R. V. H. Resh (ed.). Chapman and Hall, New York, NY.

Resh, V. H. and J. K. Jackson. 1993. Rapid assessment approaches to biomonitoring using benthic macroinvertebrates, p. 195-233. *In:* Freshwater biomonitoring and benthic macroinvertebrates. R. M. R. V. H. Resh (ed.). Chapman and Hall, New York, NY.

Rosen, B. H. 1995. Use of periphyton in the development of biocriteria. *In*: W.S. Davis and T.P. Simon (eds.) Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making. Lewis Publishers, Boca Raton, FL. pp. 415.

Schwenneker, B. W. and R. A. Hellenthal. 1984. Sampling considerations in using stream insects for monitoring water quality. Environmental Entomology 13:741-750.

Sebastien, R. J., D. M. Rosenberg and A. P. Wiens. 1988. A method for subsampling unsorted benthic macroinvertebrates by weight. Hydrobiologia 157:69-75.

Sell, D. W. and M. S. Evans. 1982. A statistical analysis of subsampling and an evaluation of the Folsom plankton splitter. Hydrobiologia 94:223-230.

Sheldon, A. L. 1984. Cost and precision in a stream sampling program. Hydrobiologia 111:147-152.

Skalski, J. R. and D. H. McKenzie. 1982. A design for aquatic monitoring programs. Journal of Environmental Management 14:237-251.

Sodergren, S. 1974. A simple subsampler for streams-bottom-fauna samples. Archiv fur Hydrobiologie 73:549-551.

Somers, K. M., R. A. Reid and S. M. David. 1998. Rapid biological assessments: how many animals are enough? Journal of the North American Benthological Society 17:348-358.

Stroom, K. T. and C. Richards. 1999. Development of a biocriteria for streams on Minnesota's Lake Superior watershed. Bulletin of the North American Biological Society 16:202.

Underwood, A. J. 1996. Detection, interpretation, prediction and management of environmental disturbances: some roles for experimental marine ecology. Journal of Experimental Marine Biology and Ecology 200:1-27.

Vanderklift, M. A., T. J. Ward and J. C. Phillips. 1998. Use of assemblages derived from different taxonomic levels to select areas for conserving marine biodiversity. Biological Conservation 86:307-315.

Vezina, A. F. 1988. Sampling variance and the design of quantitative surveys of the marine benthos. Marine Biology 97:151-155.

Vinson, M. R. and C. P. Hawkins. 1996. Effects of sampling area and subsampling procedure on comparisons of taxa richness among streams. Journal of the North American Benthological Society 15:392-399.

Walsh, C. J. 1997. A multivariate method for determining optimal subsample size in the analysis of macroinvertebrate samples. Marine Freshwater Research 48:241-248.

Waterhouse, J. C. and M. P. Farrell. 1985. Identifying pollution related changes in chironomid communities as a function of taxonomic rank. Canadian Journal of Fisheries and Aquatic Sciences 42:406-413.

Waters, T. F. 1969. Need to get this. Limnology and Oceanography 14:813-815.

Wilhm, J. L. and T. C. Dorris. 1968. Biological parameters for water quality criteria. BioScience 18:477-481.

Williams, P. H. and K. J. Gattson. 1994. Measuring more of biodiversity: can higher-taxon richness predict wholesale species richness. Biological Conservation 67:211-217.

Wright, I. A., B. C. Chessman, P. G. Fairweather and L. J. Benson. 1995. Measuring the impact of sewage effluent on the macroinvertebrate community of an upland stream: The effect of different levels of taxonomic resolution and quantification. Australian Journal of Ecology 20:142-149.

Wright, J. F., D. Moss, P. D. Armitage and M. T. Furse. 1984. A preliminary classification of running-water sites in Great Britain on macro-invertebrate species and the prediction of community type using environmental data. Freshwater Biology 14:221-256.

Wrona, F. J., J. M. Culp and R.W. Davies. 1982. Macroinvertebrate subsampling: a simplified apparatus and approach. Canadian Journal of Fisheries and Aquatic Science 39:1051-1054.

Yoder, C. O. and M. A. Smith. 1999. Chapter 2: Using fish assemblages in a state biological assessment and criteria program: essential concepts and considerations. *In*: Simon, T.P. (ed) Assessing the Sustainability, and Biological Integrity of Water Resources Using Fish Communities. CRC Press, Boca Raton, FL. pp 671.