species. Scuds are laterally compressed and often swim on their sides, hence the common name sideswimmers. **Ephemeroptera (Mayflies):** Presently, 115 species found within 52 genera and 19 families have been identified from Wisconsin. Larvae of all Ephemeroptera species inhabit a variety of permanent or temporary lentic habitats. Mayfly larvae are aquatic insects by having lateral or ventrolateral gills on most of the basal abdominal segments. Mayfly larvae mostly crawl about on aquatic substrates, and some burrow for feeding. Almost all Ephemeroptera larvae are herbivores or detritivores; only a few are known to prey on other invertebrates. Ephemeroptera species may live for as little as a few hours, but other species may live for as long as a few weeks. **Odonata (Damselflies and Dragonflies):** Presently, 154 species in the order odonata have been found in Wisconsin, represented by 3 families, 19 genera, and 45 species in the sub-order Zygoptera (damselflies), and 6 families, 38 genera, and 109 species in the suborder Anisoptera (dragonflies). The larvae of all species are aquatic with about two-thirds inhabiting lentic environments and one-third inhabiting lotic environments. Most larvae inhabit all types of permanent streams but some are also found in gravel and rock riffles, debris along streambanks, and soft sediments and sand; occasionally they are found along the wind-swept shores of lakes. Lentic larvae inhabit permanent and temporary ponds, wetlands, and littoral zones and shoreline areas of lakes. Lifecycles are relatively long, varying from one to four years. Larvae found in Wisconsin can be identified to species. **Plecoptera (Stoneflies):** Presently, 58 species of stoneflies in 8 families and 25 genera have been identified from Wisconsin. All larvae are aquatic, and almost all inhabit streams; larvae of a few species may

**MACROINVERTEBRATE Data Interpretation Guidance Manual**

Richard A. Lillie, Stanley W. Szczytko, and Michael A. Miller

Wisconsin Department of Natural Resources
Madison, WI
2003
Contents

Overview ............................................................. 2
Introduction ....................................................... 3
Metrics Generated by the BUGPROGRAM ......................... 5
Interpretation of Metrics ........................................ 5

Descriptive Metrics .................................................. 6
  Biotic Indices .................................................. 6
  Richness Measures ........................................... 9
  Diversity Measures .......................................... 14
  Functional Feeding Classes .................................... 16
  Ratio Measures ............................................... 18
  Dominance Measures ......................................... 19

Comparison Metrics ............................................... 19
  Paired Comparison Metrics .................................... 20
  Ratio Metric Comparisons .................................... 20
  Presence/Absence and Percentage Comparison Metrics ....... 20

Recommendations .................................................. 23
  Objectives ..................................................... 23
  Microhabitat Considerations .................................. 25
  Choosing the Appropriate Metrics ............................. 25

Future Directions .................................................. 28

References .......................................................... 31

Appendix A. Selected Literature Review .......................... 47

Appendix B. A History of the Development of
  Macroinvertebrate Biological Indices in Wisconsin ............ 57
Overview

This guidance manual is designed to assist Departmental staff interpret macroinvertebrate data reported to the Department through contractual arrangements with the University of Wisconsin Stevens Point (UWSP). Data are analyzed, and the computer software program BUGPROGRAM 1 version 6.01 reports results.

The electronic database as of November 2003 had 8,940 individual samples (11,970 samples when including replicate samples) and includes data for all macroinvertebrate samples processed by the UWSP from 1983 to the present. Approximately 90 percent of the samples in the database have associated latitude and longitude data. The database includes macroinvertebrate samples collected by: Wisconsin Department of Natural Resources (DNR), United States Geological Survey (USGS), and United States Forest Service (USFS), University, and private research projects.

By querying the electronic database, the BUGPROGRAM produces reports consisting of a series of over 30 biological community attributes for each sample. These results include measures of richness, diversity, dominance, functional-feeding classes, and biotic indices. This guidance manual is intended to: 1) provide a detailed description of each biological attribute (commonly referred to as a metric), 2) provide a simplified explanation of the biological significance of each metric, and 3) instructions as to how and when each metric should be applied. A comprehensive list of references on biological indices is provided at the end of this document for those who wish to explore these issues in greater depth.

Please note, this document does not tell you how, when, or where to sample (please see DNR Field Procedures Manual at http://intranet.dnr.state.wi.us/int/es/science/ls/fpm/table.htm). It does not tell you how to conduct statistical procedures or suggest which procedure is best to use under given circumstances, nor does it specifically address other macroinvertebrate biomonitoring issues such as lake, wetland, or contaminated sediment assessments. Order descriptions of insects, presented in this document as side-bars, are summarized from Hilsenhoff (1995).

1 The BUGPROGRAM, program manual, installation instructions, and historical databases are currently available using the following URL: http://www.uwsp.edu/water/biomonitoring/BUGPRO.HTM. Program structure and capabilities of the BUGPROGRAM 6.01 are described in further detail in the program manual online. Biomonitoring data are available at http://www.uwsp.edu/water/biomonitoring/index3.htm.
Ten orders of insects found in Wisconsin have species with aquatic lifestages. Five (Ephemeroptera, Odonata, Plecoptera, Megaloptera, and Trichoptera) are aquatic orders in which almost all species have aquatic larvae. The remaining orders (Heteroptera, Coleoptera, Diptera, Lepidoptera, and Neuroptera) are partially aquatic orders in which most species are terrestrial, but in which certain families, genera, or one or more species have a lifestage adapted to the aquatic environment. The insect orders Collembola, Orthoptera, and Hymenoptera are primarily terrestrial orders that have some semi-aquatic species. However, these are not used in biomonitoring studies.
### Table 1. Individual sample metrics generated by the BUGPROGRAM. n.a. = not applicable

<table>
<thead>
<tr>
<th>Metric</th>
<th>Name or Description</th>
<th>Taxonomic Level</th>
<th>Range</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Biotic Indices</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HBI</td>
<td>Hilsenhoff Biotic Index</td>
<td>Mixed (genus and species)</td>
<td>0-10</td>
<td>Hilsenhoff 1987</td>
</tr>
<tr>
<td>FBI</td>
<td>Family-Level Biotic Index</td>
<td>Families</td>
<td>0-10</td>
<td>Hilsenhoff 1988a</td>
</tr>
<tr>
<td>MTV</td>
<td>Mean Tolerance Value</td>
<td>Mixed (genus and species)</td>
<td>0-10</td>
<td>Lillie and Schlesser 1994</td>
</tr>
<tr>
<td>Max x</td>
<td>Max-x HBI</td>
<td>Mixed (genus and species)</td>
<td>0-10</td>
<td>Hilsenhoff 1998</td>
</tr>
<tr>
<td><strong>Taxa Richness</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SR</td>
<td>Total Species Richness</td>
<td>Mixed (genus and species)</td>
<td>0-∞</td>
<td>Plafkin et al. 1989</td>
</tr>
<tr>
<td>GR</td>
<td>Generic Richness</td>
<td>Genus</td>
<td>0-∞</td>
<td>Plafkin et al. 1989</td>
</tr>
<tr>
<td>EPTG</td>
<td>Ephemeroptera-Plecoptera-Trichoptera Generic Richness</td>
<td>Genus</td>
<td>0-∞</td>
<td>Plafkin et al. 1989</td>
</tr>
<tr>
<td><strong>Diversity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d</td>
<td>Margalef’s Diversity Index</td>
<td>Genus</td>
<td>0-?</td>
<td>Margalef 1958</td>
</tr>
<tr>
<td>H’</td>
<td>Shannon’s Index of Diversity</td>
<td>Genus</td>
<td>0-?</td>
<td>Magurran 1988</td>
</tr>
<tr>
<td><strong>Trophic Function</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percent Scrapers</td>
<td>Percent of total represented by Scrapers</td>
<td>Individuals and genus</td>
<td>0-100</td>
<td>Cummins and Merritt 1996</td>
</tr>
<tr>
<td>Percent Filterers</td>
<td>Percent of total represented by Filterers</td>
<td>Individuals and genus</td>
<td>0-100</td>
<td>Cummins and Merritt 1996</td>
</tr>
<tr>
<td>Percent Shredders</td>
<td>Percent of total represented by Shredders</td>
<td>Individuals and genus</td>
<td>0-100</td>
<td>Cummins and Merritt 1996</td>
</tr>
<tr>
<td>Percent Gatherers</td>
<td>Percent of total represented by Gatherers</td>
<td>Individuals and genus</td>
<td>0-100</td>
<td>Cummins and Merritt 1996</td>
</tr>
<tr>
<td>Percent Collectors</td>
<td>Percent of total represented by Collectors</td>
<td>Individuals and genus</td>
<td>0-100</td>
<td>Cummins and Merritt 1996</td>
</tr>
<tr>
<td>Percent Scrapers/</td>
<td>Ratio percent Scrapers to percent Filterers</td>
<td>Individuals and genus</td>
<td>0-100</td>
<td>Cummins and Merritt 1996</td>
</tr>
<tr>
<td>percent Filterers</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percent Scrapers/</td>
<td>Ratio percent Scrapers to percent Gatherers</td>
<td>Individuals and genus</td>
<td>0-100</td>
<td>Cummins and Merritt 1996</td>
</tr>
<tr>
<td>percent Gatherers</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percent Scrapers/</td>
<td>Ratio percent Scrapers to percent Collectors</td>
<td>Individuals and genus</td>
<td>0-100</td>
<td>Cummins and Merritt 1996</td>
</tr>
<tr>
<td>percent Collectors</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Dominance</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dominant families</td>
<td>Percent of total count represented by top five families</td>
<td>Family</td>
<td>n.a.</td>
<td>Plafkin et al. 1989</td>
</tr>
<tr>
<td>Dominant genera</td>
<td>Percent of total count represented by top five genera</td>
<td>Genus</td>
<td>n.a.</td>
<td>Plafkin et al. 1989</td>
</tr>
<tr>
<td>Dominant species</td>
<td>Percent of total count represented by top five species</td>
<td>Species</td>
<td>n.a.</td>
<td>Plafkin et al. 1989</td>
</tr>
</tbody>
</table>

a Note: Occasionally higher taxonomic levels used.
b This metric is also reported individually for each order (Ephemeroptera generic richness, Plecoptera generic richness, and Trichoptera generic richness) and also reports the ratio of numbers of individuals of EPT to total number of individuals of all taxa in a sample.
c Although not reported in BUGPROGRAM 6.01, Margalef’s Diversity Index was reported in the first version of the BUGPROGRAM and it was thought that future comparative studies might evolve.
d This metric is reported in version 6.01 of BUGPROGRAM.
e This taxonomic level is computed separately on the basis of total individuals in sample and total genera in sample.
f Note: BUGPROGRAM may list as ‘unidentified’ because some taxonomic identification is only to genus.

### Table 2. Paired community comparison metrics available through the BUGPROGRAM.

<table>
<thead>
<tr>
<th>Metric</th>
<th>Name</th>
<th>Taxonomic Level</th>
<th>Range</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCL</td>
<td>Coefficient of Community Loss</td>
<td>Lowest taxa</td>
<td>0-∞</td>
<td>Courtemanch and Davies 1987</td>
</tr>
<tr>
<td>CS</td>
<td>Coefficient of Similarity</td>
<td>Genus</td>
<td>0-1</td>
<td>Pinkham and Pearson 1976</td>
</tr>
<tr>
<td>SIMI</td>
<td>Stander’s Similarity Index</td>
<td>Lowest taxa</td>
<td>0-1</td>
<td>Stander 1970</td>
</tr>
<tr>
<td>PS</td>
<td>Percentage Similarity</td>
<td>Genus</td>
<td>0-1</td>
<td>Whittaker 1952, Rabeni and Gibbs 1980</td>
</tr>
<tr>
<td>B</td>
<td>Coefficient of Similarity</td>
<td>Genus</td>
<td>0-1</td>
<td>Pinkham and Pearson 1976</td>
</tr>
<tr>
<td>EDIS</td>
<td>Ecological Distance</td>
<td>Genus and species</td>
<td>0-1</td>
<td>Clark 1952, Rhodes et al. 1969</td>
</tr>
</tbody>
</table>
Metrics Generated by the BUGPROGRAM

Currently for each macroinvertebrate sample, version 6.01 of the BUGPROGRAM generates several measures of taxonomic richness, one diversity index, four biotic index values, eight trophic function ratio metrics at two taxonomic levels, and three sets of dominant taxa lists (Table 1). In addition, the BUGPROGRAM provides five metrics to compare macroinvertebrate communities between pairs of selected samples (Table 2).

In addition to the previously named metrics (Tables 1 and 2), the BUGPROGRAM reports counts of individuals, percentages of individuals, or percent specific taxa present in the sample or used in computing the biotic indices. The various counts are used in computing some individual sample metrics and counts are useful in evaluating the validity of a particular sample. UWSP standard procedure is to subsample until 125 organisms are found (Hilsenhoff 1977, Hilsenhoff 1982, Hilsenhoff 1987, Hilsenhoff 1998) that have assigned HBI values. However, in some cases this goal is not achieved. This situation can result from a number of circumstances. First, not all arthropods present in a sample are currently assigned pollution tolerance values and not all specimens in a sample can be identified to a taxonomic level at which a tolerance value has been assigned (i.e., underdeveloped or damaged specimens lacking key structures necessary for identification to lower taxonomic level). Second, the following organisms are not included in the HBI count according to the UWSP and DNR protocols: hemipterans, non-dryopid coleopterans, insect pupae, insect adults, mollusks, annelids, decapods, nemata and nematophans, hydrocarinians, and tubellarians. The inclusion of underdeveloped or damaged specimens in the 125 count occasionally may result in reporting of HBI values based on less than the 100 specimen minimum recommended by Hilsenhoff (1987). For example, there may have been 165 organisms in a particular sample (total sample sort) but only 90 specimens may have had assigned HBI values.

Interpretation of Metrics

In this section we explain how each macroinvertebrate metric is derived and how to interpret values. In many cases there are no strict rules available to assign qualitative designations to a sample metric. That is, it is difficult to say one value is “bad” while another is “good”. Judgments under such circumstances remain subjective and open to debate. Nevertheless, the metric may be valuable in making relative comparisons of water resource quality among streams (or among stations within streams) or in identifying possible pollution sources. Some metric values may represent a continuum or gradient, while for other metric values a threshold may be established that links a specific value with a specific dose of an environmental stressor. For a more comprehensive review and coverage of the following topics we recommend the following three books: 1) Biological Monitoring of Aquatic Systems (Loeb and Spacie 1993) 2) Biological Assessment and Criteria (Davis and Simon 1994) and 3) Freshwater Biomonitoring and Benthic Macroinvertebrates (Rosenberg and Resh 1993).
Descriptive Metrics

Biotic Indices

HBI and FBI. The Hilsenhoff Biotic Index (HBI) and Family Level Biotic Index (FBI) represent the average weighted pollution tolerance value of all arthropods present in a sample (excluding organisms either too immature or damaged to allow for correct identification and organisms which have not been assigned a pollution tolerance value). The HBI is a well-tested metric that has been incorporated into national protocols for rapid bioassessment (Plafkin et al. 1989). The North Carolina Biotic Index (Lenat 1993) is a modified HBI index applicable for use in the ecoregions of the southeastern United States.

For HBI determinations, identification is carried to the lowest possible taxonomic level necessary to assign a pollution tolerance value (Hilsenhoff 1987). In many cases this means that an identification at the genus level is sufficient to assign a corresponding HBI value to that organism. All identifications for the FBI are made at the family level (Hilsenhoff 1988a). Anyone using the HBI or the FBI should be familiar with the field and laboratory procedures as described by Hilsenhoff (1987, 1988a). Field and laboratory procedures used for DNR can be found at the following intranet site: http://intranet.dnr.state.wi.us/int/es/science/ls/fpm/table.htm.

Sampling in sites not meeting established criteria (i.e., inadequate flow velocities, snags or pools rather than riffles or runs) seriously limit the use of the resulting data. In recent studies, comparing samples taken in close proximity within the same stream and that use a minimum of five replicates, snag data are more variable and generally give higher biotic index (BI) values than riffle samples (S. Szczytko, UW Stevens Point, pers. comm 2000). It is extremely important to emphasize that the HBI and FBI are indices of organic pollution and are based on a community’s response to the combination of high organic loading and decreased dissolved oxygen levels. The HBI or the FBI was not intended for use outside the purpose of detecting or monitoring organic pollution. The relation of HBI values to water quality is presented in Table 3. Narf et al. (1984) provide a way to test for the detectable differences (DDs) in HBIs between samples. However, the subsequent range expansion of the HBI scale from 0-5 to 0-10 and assignment of new tolerance values makes the work of Narf et al. (1984) obsolete. To correct for the expansion of the HBI scale it is not statistically valid to multiply the DDs or HBIs by a factor of two (R. Narf and G. Lange, Wisconsin DNR, pers. comm 1997). Instead, historical BI data will have to be recalculated using the new BUGPROGRAM after changing
macroinvertebrate identification to correspond with changes in nomenclature and updating tolerance values. The current version of the BUGPROGRAM is able to update older databases including organism identification numbers, name changes, and tolerance values.

The HBI is seasonally dependent (Hilsenhoff 1988b). Higher HBI values occur in summer because organisms present during the summer months generally tend to be more tolerant than taxa inhabiting the same stream during the spring. To insure comparable HBI values, Hilsenhoff (1988b) recommends sampling warm water streams in the spring before 440 degree days accumulated and cold water streams until 1050 degree days accumulated. Fall sampling can be resumed 60 days after the 440 degree day mark in warm-water streams, and 45 days after the 1050 degree day mark in cold-water streams. Please note, sampling after November 1st is not recommended. Sampling within these time windows should avoid exceeding the 95 percent confidence limits for the HBI. If sampling is conducted outside these time windows, Hilsenhoff (1988b) recommends subtracting 0.50 units from the reported HBI value to compensate for the bias.

The FBI was designed as a rapid field assessment tool. As a consequence, the FBI can be less precise than the HBI. Generally the FBI underestimates the severity of pollution in highly polluted streams and overestimates the degree of impact in clean streams (Hilsenhoff 1988a, Szczytko 1988). To compensate for this bias Hilsenhoff (1988a) proposed a modified water quality rating system for the FBI (Table 4). Reporting FBI values may seem redundant when the more accurate HBI values are also available (i.e., lab-processed samples), in view of the fact that large disparities commonly occur between the two indices (Szczytko 1988). However, the HBI values generated by laboratory processed samples are potentially valuable for making direct comparisons with field monitoring assessments using the FBI technique. Irrespective of its limitations the FBI remains a valuable field tool for rapid bioassessments and screening of potential sources of impact.

Table 3. Water quality ratings for HBI values (from Hilsenhoff 1987).

<table>
<thead>
<tr>
<th>HBI Value</th>
<th>Water Quality Rating</th>
<th>Degree of Organic Pollution</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤ 3.50</td>
<td>Excellent</td>
<td>None Apparent</td>
</tr>
<tr>
<td>3.51-4.50</td>
<td>Very Good</td>
<td>Possible Slight</td>
</tr>
<tr>
<td>4.51-5.50</td>
<td>Good</td>
<td>Some</td>
</tr>
<tr>
<td>5.51-6.50</td>
<td>Fair</td>
<td>Fairly Significant</td>
</tr>
<tr>
<td>6.51-7.50</td>
<td>Fairly Poor</td>
<td>Significant</td>
</tr>
<tr>
<td>7.51-8.50</td>
<td>Poor</td>
<td>Very Significant</td>
</tr>
<tr>
<td>8.51-10.00</td>
<td>Very Poor</td>
<td>Severe</td>
</tr>
</tbody>
</table>

2 Defined as streams experiencing summer maxima above 20º C.
3 Hilsenhoff (1988b) used a base of 4.5º C to calculate degree-days. Use mean daily air temperatures to compute degree-days.
4 Defined as streams experiencing summer maxima below 20º C.
The True HBI (THBI) was calculated by Hilsenhoff (1998) as the mean annual HBI excluding the three consecutive late spring or summer samples that had the highest HBI values and November samples. These dates are the times the HBI is not recommended by Hilsenhoff (1988b). Hilsenhoff (1998) viewed the THBI as the most accurate bioassessment of water quality for streams. Unfortunately this assumption is not supported by hard data and is open to debate. (i.e., the summer HBI values or MTV actually may be more closely correlated with trends or rankings in true water resource quality). Hilsenhoff (1998) found the metrics closest to the THBI were the annual mean HBI and the 25-Max HBI in clean and polluted streams respectively. In addition, Hilsenhoff (1998) also found that the annual mean HBI was always higher than the THBI in the streams in his study because the three summer sampling dates with the highest HBI were not included in the calculation of the THBI. Hilsenhoff (1998) recommends that diversity and species richness metrics not be used to supplement the THBI because factors other than pollution effects have a great impact on these metrics.

MTV. The Mean Tolerance Value (MTV) represents the average tolerance value of all taxa (as opposed to all individuals in the HBI) present in a sample. The MTV gives equal weight to rare and dominant taxa regardless of the abundance of the taxa (Lillie and Schlesser 1994) while the HBI places more weight on the dominant taxa. The MTV may be less susceptible to temporal changes and sample sizes than the HBI. The MTV is similar to other metrics that de-emphasize relative numerical abundance such as: the Macroinvertebrate Community Index (MCI) in New Zealand (Stark 1993), the Stream Invertebrate Grade Number Average Level biotic index (SIGNAL) in Australia (Chessman 1995), and the Average Score Per Taxon (ASPT) in Great Britain (Armitage et al. 1983, Wright et al. 1988). The ASPT is less sensitive to sampling effort and seasonal effects than its counterpart quantitative metric, the

---

**Table 4. Water quality index for the family-level HBI (from Hilsenhoff 1988a).**

<table>
<thead>
<tr>
<th>FBI Value</th>
<th>Water Quality Rating</th>
<th>Degree of Organic Pollution</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤ 3.75</td>
<td>Excellent</td>
<td>Unlikely</td>
</tr>
<tr>
<td>3.76-4.25</td>
<td>Very Good</td>
<td>Possible Slight</td>
</tr>
<tr>
<td>4.26-5.00</td>
<td>Good</td>
<td>Some Probable</td>
</tr>
<tr>
<td>5.01-5.75</td>
<td>Fair</td>
<td>Fairly Substantial</td>
</tr>
<tr>
<td>5.76-6.50</td>
<td>Fairly Poor</td>
<td>Substantial Likely</td>
</tr>
<tr>
<td>6.51-7.25</td>
<td>Poor</td>
<td>Very Substantial</td>
</tr>
<tr>
<td>7.26-10.00</td>
<td>Very Poor</td>
<td>Severe</td>
</tr>
</tbody>
</table>

---

The MTV is intended to serve as a companion metric to the HBI and should not be used independently from HBI evaluations! In most cases the MTVs should exhibit a pattern similar to that presented by the HBI. Large discrepancies between the two metrics should indicate that further examination of the data is needed. For example, if large abundances of a relatively cosmopolitan species with high tolerance values are found in a sample otherwise dominated by less tolerant species, this may result in an inaccurate HBI value that is not reflective of the ‘true’ water quality of a particular stream. Likewise, the introduction of a few relatively intolerant taxa from a clean upstream tributary into a poor stream may substantially lower the MTV and bias the interpretation of the data. In general MTVs are similar to FBIs in that they have the same apparent bias relative to HBI values. The MTVs tend to be lower than the corresponding HBIs in highly polluted situations and higher than HBI scores in clean streams (Lenz and Miller 1996, Hilsenhoff 1998). This suggests that the HBI water quality rating system is not appropriate for use with MTV scores and that a separate rating system needs to be developed for MTVs.

After evaluating the effects of a range of maximum numbers (5, 10, and 25) on HBI values Hilsenhoff (1998) proposed a new index, the 10-Max BI. Hilsenhoff (1998) suggests limiting the maximum number of individuals of each taxon in an HBI sample to 10 as a means to reduce the effects of seasonal variability on the HBI. The 10-Max BI limits the effect of dominant organisms and does not elevate the importance of very rare taxa to the same level as the MTV. The use of the 10-Max BI may be superior to the HBI in evaluating polluted streams (Hilsenhoff 1998). The BUGPROGRAM allows the user flexibility in selecting a maximum number of individuals (i.e., 5, 10, 25, or any other number) to calculate a new HBI referred to as the Max-\(x\) HBI value. Setting the maximum value to 1 will produce the equivalent to the MTV. Setting the maximum value to 10 will produce the 10-Max BI recommended by Hilsenhoff (1998).

**Richness Measures**

Richness measures represent the number of distinctly different taxa found in a sample. A richness value does not represent the total number of taxa at a site, but rather it is a relative measure or index. Often it is only necessary to process a small fraction of a sample to compute an HBI value. In the case of a BUGPROGRAM report a richness metric is reported on the basis of the numbers of taxa present in a HBI subsample (i.e., modified fixed count subsample). The remainder of the sample is not included in the calculations and any information regarding additional taxa present at the site is lost. This is not intended to be a criticism of the HBI but is reflective of the established laboratory procedures and the need to keep processing costs down. The loss of information is an unfortunate by-product of the established fixed count laboratory procedure. This has significant ramifications

**Odonata (Damselflies and Dragonflies):**

Presently, 154 species in the order Odonata have been found in Wisconsin, represented by 3 families, 19 genera, and 45 species in the sub-order Zygoptera (damselflies), and 6 families, 38 genera, and 109 species in the suborder Anisoptera (Dragonflies). The larvae of all species are aquatic with about two-thirds being lentic and one-third inhabiting lotic environments. Lotic-dwelling larvae occur in all types of permanent stream habitats, including gravel and rock riffles, debris along streambanks, bank vegetation, soft sediments and sand; occasionally they are found along the wind-swept shores of lakes. Lentic larvae inhabit permanent and temporary ponds, wetlands, and littoral zones and shoreline areas of lakes. Life-cycles are relatively long and range from one to four years. Most Odonate larvae found in Wisconsin can be identified to species.
with respect to calculations and the use of other metrics derived from the sample. Consequently, the data derived from the HBI subsamples represent relative measures per total number of specimens examined.

Although high taxa richness is generally associated with good water quality, low taxa richness does not necessarily indicate poor water quality, nor does high richness always indicate good water quality. Some habitats (i.e., small cold headwater streams, or oligotrophic streams, or mineral poor waters) may naturally have low numbers of taxa density per unit area. Crunkilton and Duchrow (1991) demonstrated a peak in total taxa richness in middle order streams in Missouri. In some cases, intermediate or low levels of disturbance can cause an increase in taxa richness (Townsend et al. 1997). This could be obtained either through physical mechanisms (i.e., increased heterogeneity of substrate niches), biological interactions (i.e., addition of invasive or pioneering species without loss of native species), or increased nutrients and energy. While some water quality rating scales have been established (Szczytko 1988, Barbour et al. 1996) for various richness metrics in other regions, no scale has been established for Wisconsin streams. However, see discussion of La Liberte (unpub. data) in the Future Directions section of this document. Until a Wisconsin stream water quality scale is established relative comparisons of richness values with reference stations of a similar stream class and ecoregion is recommended.

There are three taxa richness metrics reported by BUGPROGRAM at two taxonomic levels: Species Richness (SR), Generic Richness (GR), and EPT Generic Richness (EPTG).

**Species Richness (SR).** SR is based on a count of the number of species identified in a HBI sample (Note: If an unidentified specimen is keyed to the genus level and not the species level the BUGPROGRAM will count the specimen as a species within the genus if no other specimens in that genus are identified. Therefore it is possible that unidentified specimens will count as only one species instead of multiple unidentified species within a genus). This count is not a true species count due to: 1) the lack of adequate keys for many aquatic larvae among some orders, 2) the condition and life stage of particular specimens, and 3) excellent taxonomic keys are available to species but identification beyond genus is not completed or required in calculating HBIs because all members of the genus have been assigned the same pollution tolerance value (i.e., elmid riffle beetles). As a result, SR may be based on a mixture of taxonomic levels and may result in an underestimation of true total taxa richness. If the influence discussed above is consistent across samples SR will continue to have merit in evaluating wide disparities among samples.
Generic Richness (GR). GR refers to the number of different genera represented in a biotic index subsample. As noted in the case of SR above, a single unidentified specimen that is identified to family but not to genus will be counted as a genus within that family if no other specimens in that family are identified. The points discussed above for SR also hold true for GR with the exception that most specimens are identified to the genus level.

EPT Generic Richness (EPTG). The third richness metric reported by the BUGPROGRAM is EPTG. This metric represents the number of distinct genera found only among the orders Ephemeroptera, Plecoptera, and Trichoptera (EPT) in a biotic index subsample. These three orders are separated from other aquatic taxa because they generally represent the more organic pollution intolerant organisms present in rivers and streams. Consequently the EPTG is believed to be a more sensitive metric to organic pollution than SR or GR. The number of EPT taxa generally decreases with increasing perturbation (Barbour et al. 1996, Wallace et al. 1996, Growns et al. 1997). Use of the EPTG metric is somewhat redundant to the HBI (similar response characteristics) and has a higher degree of variability based on the Coefficients of Variation (CVs). This may be due to differences in the scales used in the two metrics (see highlight box below). Hilsenhoff (1998) reports that EPTG index values are seasonally variable and dependent on stream-order size. The EPTG metric has been applied successfully in numerous studies and has been incorporated into various state biomonitoring programs.

Metrics are often judged on the basis of their relative coefficients of variation (CVs) and recent studies have addressed the statistical properties of benthic macroinvertebrate metrics (Szcztko 1988, Resh 1994, Rheaume et al. 1996). One aspect of the metrics that has not been fully explored is that of scale. Many metrics are open-ended (i.e., scales range from zero to infinity) while others are restricted. The influence that restricted scales may have on CVs and other mathematical or statistical properties of the various metrics may be significant (author’s conclusions). Direct comparisons between open ended richness metric CVs and other restricted metrics CVs may not be a fair method of estimating the significance of richness metrics (Diamond et al. 1996). Biological relevance is most important when determining value of richness metrics. For example, one metric may have higher CV than another but that same metric may display a wider range in values, thus having greater discriminatory power (Barbour et al. 1996). Likewise, while some metrics may be less significantly correlated with resource conditions than other metrics, those same metrics may be more effective in distinguishing between disturbed and undisturbed conditions (Fore et al. 1996). Highly variable metrics may exhibit a threshold response to human influence at a point in their range (Fore et al. 1996) which may make them useful in assessments. Consequently, direct comparisons of CVs among different metrics should not be the sole means of judging the merit of metrics. Rather an evaluation using a combination of metric CV and the discriminatory power of the metric to detect differences between test and reference sites should be applied (Diamond et al. 1996).

Caution should be taken when interpreting various richness metrics. Richness measurements derived from biotic index subsamples may be misleading. It is well documented that there are pros and cons associated with fixed count subsampling and problems related to interpretations of

Plecoptera (Stoneflies):

Presently, 58 species of stoneflies in 8 families and 25 genera have been identified from Wisconsin. All larvae are aquatic, and almost all inhabit streams; larvae of a few species may live in cold oligotrophic lakes. Stonefly larvae differ from most aquatic insects by having two long filamentous caudal cerci (tail filaments) and an elongate or flattened appearance. Larvae obtain respiratory-oxygen from water through their cuticle or primitive gill structures, and as a result are relegated to cold, fast-flowing, highly oxygenated streams. Feeding habits vary among families with most being herbivore-detritivores, although several families have predatory species. Lifecycles are relatively long with some taxa living 3 years or more in the nymphal stage. Some species of stoneflies have some of the earliest emergences in the year, hatching as early as January and February, and the adults are often seen crawling on the snow along streambanks.
richness metrics derived from such counts (Courtemanch 1996, Barbour and Gerritsen 1996, Vinson and Hawkins 1996, Somers et al. 1998, Grown et al. 1997, Larsen and Herlithy 1998, Cao et al. 1998). The standard protocol used in sorting HBI samples outlines counting a minimum of 100 organisms (Hilsenhoff 1987; note the UWSP Benthic Macroinvertebrate Laboratory counts a minimum of 125 organisms). Thus, in most samples only a subsample of the organisms present are counted while noting what percent of the total sample is examined.

For example, in one sample only ten percent of the total number of organisms present is counted and included in the taxa richness determination. In a different sample the entire sample is searched meticulously to reach the minimum number of organisms required to include in a taxa richness evaluation. If the SR is 15 in the first sample (ten percent examined), and the SR is 25 in the second sample (100 percent examined), one should note that these values represent the number of distinctly different taxa present in the portion of the sample that was enumerated. Because the percent of the sample that was examined to reach a ‘fixed-count’ is different between the two samples, it is quite possible that both samples contained a true total of 25 taxa. Barbour and Gerritsen (1996) propose to define this value as taxa density (as opposed to taxa richness). They suggest that while taxa richness may be a desirable metric in some circumstances (assuming equal areas sampled) relative taxa density is more economical and has greater discriminatory power than taxa richness. Others argue that comparing relative taxa richness (or any other relative richness metric) between any two samples may be questionable because one sample may have an inordinate abundance of a particular taxon that may skew the count and result in missing the rarer taxa.

Richness metrics are strongly dependent upon a correlation between the area sampled and the number of individuals examined (Vinson and Hawkins 1996). It is important that field sampling efforts in collecting biotic index samples be standardized to the greatest extent possible (i.e., choose equal sample areas of representative substrates5). As outlined earlier, fixed count subsampling may underestimate the true or absolute taxa richness of a sample (or stream). However, Vinson and Hawkins (1996) found that estimated taxa richness (using a rarefaction technique presented by Hurlbert 1971) that "normalizes" taxa counts to a standard

---

5 It is the opinion of the authors that the DNR Field Manual should be amended to reflect the fact that restricting kick sampling to too small of an area within a riffle may seriously bias a sample due to the tendency for aquatic organisms to be highly aggregated. It is preferable to collect a composite sample representing a wider area of a riffle to gain a better representation of the entire macroinvertebrate community present.
The number of organisms examined in samples containing as few as 200 individuals were equally sensitive in defining differences in richness among streams from several ecoregions (Omernik 1987). Vinson and Hawkins (1996) also suggest that fixed-count subsampling is a mechanical form of rarefaction and that data derived from counts ranging between 100-200 individuals may be sufficient to detect differences among samples. Barbour and Gerritsen (1996) suggest that fixed count subsampling provides a repeatable estimate of taxa richness (i.e., relative number of taxa per standardized number of organisms).

If the objective of estimating taxa richness is to compare differences in relative taxa richness among samples, then data from fixed count subsamples may be adequate. However, it should be emphasized that this metric does not equate to true total taxa richness but rather to area density (taxa density per unit area). Larsen and Herlithy (1998) suggest that a fixed count of 300-500 individuals will produce a richness value that will approximate area density. Somers et al. (1998) found that counts of 200 or 300 produced very little gain in discriminatory power over counts of 100 among lakes using a variety of indices except for richness metrics where larger counts were deemed better! Doberstein et al. (2000) show that subsampling severely limits the discriminatory power of most metrics and the ability to accurately assess biological condition. Courtemanch (1996) explains that expression of richness metrics based on fixed counts lacks ecological interpretative value. All in all, interpretation of BUGPROGRAM taxa richness values should be made with care.

Richness measurements derived from quantitative samples (i.e., Hess stovepipe samples, cores, and some artificial substrate samplers) are less controversial when the entire sample is examined (see precautions listed by Palomaki and Paasivirta 1993). However, there are other alternatives to processing the entire sample. Vinson and Hawkins (1996) suggest that implementation of a two phase subsampling technique (Cuffney et al. 1993) with more extensive habitat sampling (and pooling of samples) would represent a compromise to expand the taxa lists for particular sites. The two phase subsampling includes searching the entire sample for large or rare taxa present. Courtemanch (1996) suggests a method which consists of serial processing of additional subsample cells until the recovery of new taxa levels off, or reaches an asymptote. Both of these alternatives have a common weakness associated with the visual recognition of new taxa. While a larger rare taxon may be immediately recognized as representing a new taxon, scanning additional cells may overlook numerous small, nondescript species, which normally would only be identified as a new species after examination at a higher magnification. Barbour and Gerritsen (1996) clearly point out that subsampling is intended to provide an unbiased estimate of a larger sample. Interpretations of data and comparisons based on relative compositions of randomly selected subsamples are indeed valid if standardized and rationalized appropriately. Much of the controversy regarding richness metrics appears to be clouded by semantics and lack of clear definitions. The issue is certainly not resolved and further thought and discussion are warranted.

Example of a case-building caddisfly case.

Trichoptera (Caddisflies):

Presently, 245 species of caddisflies in 19 families and 72 genera have been identified from Wisconsin. Caddisfly larvae are a very important faunal component of most streams, but half of the families also have species occurring in lentic environments as well, some even in temporary ponds. Caddis larvae use silk to spin nets for food collection or to construct tubular shelters. The structures vary by size, shape, and use of streambed organic and inorganic materials for construction. For some taxa the cases can be used to identify the animal to genus and sometimes species. Cryptic cases protect some species from predation, while certain case-forms allow other taxa to inhabit fast currents, or provide protection from abrasive scour. All taxa pupate in closed silken cocoons. The adults tend to be drab-colored and moth-like in their shape and erratic flight behavior.
Diversity Measures

Diversity indices represent a measure of the distribution of individuals among different taxa present in a sample. In theory, a macroinvertebrate community consisting of many taxa of even distribution (i.e., relative abundance) is considered more natural than a simple community dominated by one or few taxa (see review in Winget and Mangum 1979). However, it should be noted that a simple community can result from natural conditions (Winget and Mangum 1979).

Diversity Index \((d)\). Former versions of the BUGPROGRAM reported values for Diversity Index \((d)\) which was calculated at the generic level based on the logarithmic relationship between the number of individuals in each genus and the total number of organisms in the sample (see equation below).

\[
d = \left( N \log_2 N - \sum n_i \log_2 n_i \right) / N
\]

Where \(N\) = total of all arthropods in sample, and \(n_i\) = number of individuals in each genus.

The formula used to compute the Diversity index was actually a variant of the Margalef index as modified for stream arthropod analysis by Hilsenhoff 1977 (see Washington 1984 for further explanation). Consequently, when the BUGPROGRAM was developed, the Diversity index was referred to as the “Margalef Diversity index” despite the fact that technically it was not the same.

The Margalef index is seldom used today (Washington 1984). Boyle et al. (1990) showed that the response of Margalef’s diversity value to theoretical perturbations was erratic. Resh and McElravy (1993) conclude that less than five percent of macroinvertebrate studies surveyed used the Margalef’s Diversity index. Despite the low usage, Resh and Jackson (1993) report that Margalef’s Index proved superior over other diversity indices in detecting impacts of acid and toxic metals inputs. Magurran (1988), in her review of ecological diversity measures, considers the Margalef’s index advantageous to use due to its simplicity. She admits that Margalef’s index is applied infrequently but concludes that the index could be an important tool because it is easy to interpret and statistically and ecologically sound (Magurran 1988).

Shannon’s Index Of Diversity \((H')\). A recent modification to the BUGPROGRAM included replacing Margalef’s Index (as modified by Hilsenhoff 1977) with the Shannon Diversity index \((H')\). Resh and McElravy (1993) found that \(H'\) is the most commonly applied index in a survey of 90 lentic and lotic benthos field studies. The Shannon Diversity
index, which often is called mistakenly the Shannon-Wiener index (see Magurran 1988 for further discussion), is represented by the formula:

\[ H' = -\sum p_i \ln p_i \]

Where \( p_i \) = the proportion of individuals represented by each taxon.

\( H' \) values range between 1.5 and 3.5 with higher values representing higher diversity.

The Shannon Diversity index is based on two assumptions that may be violated in BI sampling (Magurran 1988). These assumptions are: 1) random sampling of an infinite population occurs, and 2) that all species present are sampled. It becomes important to define adequately and sample the particular areas that we wish to compare. Are we interested in comparing diversity among microhabitats, riffles, reaches, or rivers? It is important to randomly sample an area adequate enough to provide a representative picture of what taxa are present and the relative proportions that are represented by these taxa. Samples not collected randomly (i.e., artificial substrate samplers or light trapping) promote the use of an alternative diversity index such as the Brillouin index (see Magurran 1988).

DNR HBI samples may not be conducive to interpretation with the Shannon Diversity index. During HBI lab processing some individuals are identified to species level, some to genus level, and some to even coarser taxonomic levels. During field collections riffle kick sampling may include taxa from fine, depositional sediment areas, while other riffle kick samples may be restricted to coarser sediments. It is not clear how these differences influence \( H' \) values.

Diversity values, irrespective of the index chosen, may vary directly with water quality and low diversity may indicate an unstable community. However, cold, clean, headwater streams may have low diversity and still represent excellent water quality. In some cases the diversity index value does not respond as expected due to the replacement of intolerant taxa by an equal number of tolerant taxa. Hilsenhoff (1977) found that the best stream (ranked by the lowest mean HBI value) in his study was only ranked 32 out of a total of 53 by Margalef’s diversity index (using Hilsenhoff’s formula) and the second best stream ranked only 45 out of 53.

Diversity indices have undergone severe criticism (Hurlbert 1971, Norris and Georges 1993). Many diversity indices are dependent upon sample size, which may not be controlled in BI sample analysis (i.e., while a minimum of 125 organisms are counted, the actual number counted and included in the diversity index computation may exceed this number by more than 50 percent). In addition, no consensus has been reached as to what levels of change in diversity index values when shown to be statistically significant are biologically significant. As a result, little value can be placed on Margalef’s Diversity index or Shannon Diversity index values as computed by the BUG-PROGRAM. Perkins (1983) reviewed the performance of eleven diversity indices and concluded that Shannon’s index should not be used alone, however, Florida currently uses the Shannon-Wiener Index as a biocriterion (Barbour et al. 1996). Although Norris and Georges 1993 suggest that diversity indices may be of some value if used with caution, interpretation of results remains open to debate and is largely inconclusive.

Diptera (Aquatic Flies and Midges):
While primarily a terrestrial order Diptera is the dominant order of aquatic insects in Wisconsin, with 660 species estimated to occur in the state distributed among 19 families and at least 185 genera. Larvae and pupae of many species are aquatic, and account for more than one-third of all aquatic insect larvae found in Wisconsin. More than one-third of aquatic Diptera are in the family Chironomidae. Larvae inhabit all types of aquatic environments, and often dominate the invertebrate fauna in lentic environments. Diptera larvae are easily recognized by their lack of segmented thoracic legs. Taxonomic identification to the species level is not possible for the larvae of most families.
Functional Feeding Classes


The trophic structure of the community reveals much about the character of a stream. Food sources, substrates, and contaminants play important roles in shaping (i.e., establishing a templet for) the invertebrate community. The effects of riparian land use and stream side vegetation potentially have strong influence on trophic metrics (Sweeney 1993, Weigel et al. 2000). The classic river continuum concept states that trophic functional feeding classes reflect a combination of external (allochthonous) and internal (autochthonous) energy inputs to the stream. As a result, location in the stream relative to stream order or size (i.e., headwaters versus large, high order rivers), dams, lake outlets, tributary streams, riparian or general land use in the watershed, and ecoregion or climatic influences each impact upon the character of the invertebrate community inhabiting any reach of stream (Vannote et al. 1980).

Biotic indices are designed to detect organic impacts whereas functional feeding group metrics are potentially useful for detecting the impact of both organics and toxicants (Plafkin et al. 1989). Fore et al. (1996) found that feeding ecology metrics were not able to distinguish between most and least disturbed sites in evaluating the effects of logging operations in Oregon streams. Similarly, James Karr (University of Washington, pers. comm. 1988) concluded that trophic measures are not especially useful in examining the effects of human actions on invertebrate assemblages.

The BUGPROGRAM reports relative composition represented by five functional feeding classes among total individuals and total genera in a sample. The five functional feeding classes reported are: scrapers, filterers, shredders, gatherers, and collectors (note: filterers and gatherers are subcategories of collectors). The ratios of scrapers to filterers, gatherers, and collectors are also reported. Predators, parasites, and piercers are not reported in the BUGPROGRAM report. Assignment of Wisconsin taxa to these classes is based on a current literature review and classification by Merritt and Cummins (1984).
These metrics are useful in characterizing the food base of a community. However, effects of toxicants adsorbed onto particulate food items often confound interpretations. Stream size and the concepts of river continuum theory are also important considerations in interpreting the metrics derived from functional feeding classes. It should be added that many organisms shift from one feeding class to another as they advance through their respective life stages from early instar to adult.

**Scrapers.** The class scrapers include various herbivores and detritivores that graze periphyton (in particular diatoms) and attached microflora and fauna on mineral or organic surfaces. Examples include many mayflies (especially Heptageniidae) and caddisflies (including Glossosomatidae). A high proportion of scrapers is indicative of an abundant supply of periphyton. Filamentous algae and mosses can interfere with feeding by scrapers. Kerans et al. (1992) and Barbour et al. (1996) report that the proportion of grazers to scrapers decreases with increased human impact. However, Fore et al. (1996) reports a variable response to human impact.

In a California study, Hannaford and Resh (1995) found that scraper metrics were not useful in discriminating among restored, unrestored, and reference sites. Conversely, in another California study, Resh (1994) found that percent scrapers was the only functional measure to indicate a significant change following an oil spill.

**Filterers.** Filterers are a subcategory of collectors. The class filterers include various suspension feeders consisting of a combination of detritivores, herbivores, and carnivores. Examples include the simuliiids and net spinning caddisflies. Because filter feeders feed on suspended fine particulate organic material (FPOM) and often use filamentous algae for attachment sites, a high preponderance of filterers may suggest increased organic enrichment. Filterers are sensitive to toxicants often associated with fine particles, and as a result, large inputs of toxicant laden particles may cause a reduction in filterers under certain circumstances. The presence of rich deposits of FPOM and filamentous algae in combination with low counts of filterers may be a good indication of such a condition.

In one selected study (Kerans et al. 1992) percent filterers was one of only six metrics that responded in a similar direction in both pools and riffles. Kerans and Karr (1994) include percent filterers as one of 13 community attributes in their benthic index of biotic integrity. According to Barbour et al. (1996) percent filterers decreases with increasing perturbation. However, Fore et al. (1996) indicate a variable response to human disturbance.

**Shredders.** The class shredders includes a large group of detritivores and herbivores that feed on both live and dead matter. Examples include many stoneflies especially Pteronarcyidae, Peltoperlidae, Leuctridae, Taeniopterygidae, Nemoridae, and Capniidae. Among the Diptera, Tipulidae are the representative shredders. Shredders feed primarily on the coarse particulate organic matter (CPOM) and subsequently Resh and Jackson (1993) suggest that percent shredders should be assessed using leaf pack samples.

This class is particularly sensitive to riparian zone influences and land use. Barbour et al. (1996) found that percent shredders decreases with
increasing perturbations. The lack of allochthonous inputs of leaves and associated microbial colonizers due to the lack of wooded areas adjacent to and upstream from a stream site will be reflected in a decreased importance of shredders. Hannaford and Resh (1995) reported that percent shredders discriminated among restored, non restored, and reference streams despite a relatively low level of precision. The accumulation of terrestrial toxicants (i.e., pesticides and herbicides) on leaf surfaces can contribute to a substantial loss of shredders (Plafkin et al. 1989), but Fore et al. (1996) showed a variable response in this metric to human disturbance. Kerans and Karr (1994) excluded shredders from their benthic index of biotic integrity because the metric misclassified sites in comparison with the fish IBI assessments.

Gatherers. The class gatherers include detritivores and herbivores that are deposit feeders (taxa living in and feeding upon items found in or on the surface of surficial deposits). Representatives include many dipterans (particularly chironomids), some trichopterans, and mayflies.

Kerans et al. (1992) suggest that the proportion of gatherers increases with increases in human impacts, while Barbour et al. (1996) and Fore et al. (1996) indicate a variable response for this metric. Barbour et al. (1996) point out that gatherers may dominate under conditions of organic enrichment, but they can be severely reduced if the stressor is a form of toxicant. Kerans and Karr (1994) tested but did not use gatherers in their benthic index of biotic integrity because the metric misclassified sites as compared with fish IBI assessments.

Collectors. This class is the sum of gatherers and filterers. Because the response of gatherers may be variable (Barbour et al. 1996, Fore et al. 1996), the response of collectors may also be indeterminant. Barbour et al. (1996) suggest that because collectors and filterers are generalists, they may be more tolerant to certain forms of pollution. Consequently, the relative contribution of collectors in a sample may increase with more severe perturbations due to the loss of more intolerant taxa.

Ratio Measures
Metrics derived from ratios generally exhibit wider fluctuations than non-ratio metrics (Plafkin et al. 1989, Barbour et al. 1992), are difficult to interpret biologically (Kerans et al. 1992) and may have serious statistical shortcomings (Jackson et al. 1990, Berges 1997).

There are three ratio metrics computed by the BUG-PROGRAM. The ratio of percent scrapers to percent filterers has the clearest meaning. A high percentage of scrapers
to filterers generally reflects an unbalanced community and points to the relative abundance ratio of CPOM to FPOM. Generally stream orders four through six (Cummins 1974) contain organisms that are dependent upon light reaching the stream bottom to support attached algae (diatoms and filamentous forms). Increased algae growth will in turn support grazers and scrapers. However in deep high order streams (orders seven through twelve) shading and increased concentrations of suspended organic and inorganic particles reduce light penetration and reduces attached algae growth. A reduction of algae contributes to an increase in filterers relative to scrapers.

**Dominance Measures**

The contribution a particular taxon (family, genus, or species) to the total number of individuals represented in a sample provides an indication of relative dominance, and is a simple estimator of evenness (Plafkin et al. 1989). A healthy and stable community should contain a diverse group of organisms with few dominant taxa. The percentage of a dominant organism (irrespective of the identity) increases with increasing perturbation (Barbour et al. 1996).

The BUGPROGRAM reports the five most dominant families, genera, and species in a sample. Communities dominated by a few taxa with high combined percentages generally reflects a disturbed situation, whereas communities dominated by a good mixture of taxa with lower combined percentages indicates a more balanced and healthy condition. It is difficult to provide guidelines for interpreting community structure based on dominants because a myriad of natural communities exist.

Aside from dominance by functional feeding groups (see section on functional feeding classes), certain taxa may indicate that particular conditions exist. For example, percent Ephemeroptera, percent Plecoptera, percent Trichoptera, and percent Amphipoda all decrease with increasing perturbation; however, percent Odonata, percent Diptera, and percent Isopoda all increase with increasing perturbation (Barbour et al. 1996). The percent contribution of dominant taxon metric has been shown to be sensitive to sub-sampling and collector bias (Hannaford and Resh 1995). Additionally, Plafkin et al.(1989) found that dominance metrics are not particularly sensitive to moderate amounts of organic or toxic loadings. Impairment (or nonimpairment) response is most evident at the extremes.

**Comparison Metrics**

There are a number of ways by which macroinvertebrate communities may be compared. Various non-parametric and parametric tests may be used to compare single test metrics between pairs of stations or among multiple stations. These comparisons may represent upstream-downstream studies (i.e., above and below discharge point), before-after (i.e., pre-post treatment or spill), or control-impact (i.e., treated versus untreated). Bioassessment of biological condition is best conducted by comparison with a reference site of a similar position or class (Plafkin et al. 1989). The current emphasis is on comparison with reference or least-disturbed sites by ecoregion and stream type. Staff may best be advised to consult with a biostatistician prior to beginning any detailed investigation or analyzing existing data.
Paired Comparison Metrics

In addition to descriptive characteristics, six paired community comparison metrics are available through the DNR-BUG program. These include both similarity and dissimilarity indices (Table 2). Despite the high degree of variability associated with similarity indices (Szczytko 1988) their application to bioassessments is not entirely without merit. In fact, community similarity indices are listed among recommended rapid bioassessment protocols (Plafkin et al. 1989) and continue to be applied successfully in many studies (Pearson and Pinkham 1992, Lillie and Isenring 1996). Interpretation of similarity index values with respect to condition or biological impairment is difficult. Determinations of similarity or dissimilarity are often obscure and arbitrary with discriminatory threshold values differing among indices. However, in rapid bioassessments, Plafkin (1989) and Hannaford and Resh (1995) recommend percent similarity thresholds of 83 percent and 65 percent respectively. Comparisons with a reference stream (or historical reference condition on the same stream) and other known impacted sites are recommended. As is the case with most metrics replication in terms of either samples, closely spaced stations, or different dates is recommended if time and funding permits.

Ratio Metric Comparisons

The Coefficient of Community Loss (CCL) may be used to compare changes (before-after) or differences (control-impact) in communities based on simple taxa presence or absence. The CCL represents a ratio metric that corresponds to loss of taxa at an impacted site relative to a control site. The CCL differs from other similarity metrics in that the values range from 0 to infinity. Low values suggest little change or difference between samples while high values indicate greater impacts. Courtemanch and Davies (1987) discuss the range in CCL values relative to different levels of impacts. Barbour et al. (1992) reported finding the CCL to be unacceptable in discriminating differences among macroinvertebrate data in eight ecoregions.

6 Judging the relative value of individual metrics purely on the basis of CV comparisons is not recommended. Metrics with open-ended scales may have higher CVs than metrics with restricted or limited scales purely as an artifact of their mathematical properties. Likewise, caution should be taken in evaluating the utility of similarity indices based on CVs. By the very nature of the paired comparisons, variation is magnified. The use of replicates actually introduces greater opportunity for increased variation because the weight of outliers is twice that of its representation.
Presence/Absence and Percentage Comparison Metrics

The BUGPROGRAM computes five other similarity-dissimilarity metrics with values ranging from 0 to 1. These metrics include: coefficient of similarity ($CS$ and $B$), Stander’s similarity index (SIMI), percentage similarity ($PS$), and ecological distance (EDIS). With the exception of the EDIS, which is a dissimilarity index, low values represent dissimilar communities while high values represent more similar communities. The calculation of each of these metrics differs slightly (see Szczytko 1988 for formulas). The Coefficient of Similarity ($CS$) is based on simple presence or absence of taxa which allows application to both quantitative and semi-quantitative data. The remaining similarity metrics are based on relative taxa abundance in paired samples.

Each metric has its own set of strengths and weaknesses which derive from their mathematical calculation. Some are subjective in their interpretation (i.e., SIMI and EDIS) while other indices are more sensitive to dominant (i.e., SIMI) or rare taxa (Coefficient $B$; Brock 1977). Currently the CCL, $B$ similarity, and the Jaccard Coefficient of Community are recommended in rapid bioassessment protocols (Plafkin et al. 1989). The Bray-Curtis Index is favored by some investigators because it varies linearly relative to changes in species numbers and abundance (Norris and Georges 1993) and is said to reflect ‘true’ similarity (Bloom 1981). While similarity-dissimilarity indices have great potential in comparing invertebrate communities (Plafkin et al. 1989), no single index can be recommended over another. It is recommended that more than one metric be applied whenever possible.

Coefficient of Similarity ($CS$) And Pinkham Pearson $B$ Similarity Index ($B$). Brock (1977) argues that the Pinkham Pearson $B$ similarity index (Pinkham and Pearson 1976) is too sensitive with rare species, not sensitive enough with dominant forms, and may be subject to bias or sampling error. Other studies consider the sensitivity of the $B$ similarity index to be an asset. Boyle et al. (1990) reported the $B$ similarity index was preferred over the Jaccard Coefficient for detecting impairment. Barbour et al. (1992) recommends the $B$ similarity index as a community similarity metric. Lillie and Isenring (1996) found the $B$ similarity index useful when comparing EPTG community structure among streams in the region of Baraboo Hills, WI. Note, $CS$ and $B$ similarity (Pinkham and Pearson 1976, Pearson and Pinkham 1992) can be calculated using the software BIOSIM1 (Gonzales et al. 1993).

Stander’s Similarity Index (SIMI). This index is based on the relative proportion of individuals in a community. Values range from 0 (no taxa in common) to 1 (the relative abundance of all taxa are equal at both sites) (Stander 1970). The interpretation of SIMI values, and assigning categorical ranks to the degree of similarity observed (i.e., low, medium, high) is difficult because the SIMI is a correlation measurement (Johnson and Millie 1982). Based on this fact SIMI is best used in making general comparisons between two or more sites with a common reference site.

Various ranking scales have been proposed for the SIMI (Rohr 1977, Tuchman and Blinn 1979). Johnson and Millie (1982) recommend the Lepidoptera (Aquatic Caterpillars):

Nine species and six genera in the family Pyralidae are the taxa of moths found in Wisconsin that have aquatic larvae. The larval stages inhabit rooted aquatic macrophytes and feed by mining the plant stems or by feeding on the attached algae. Adults are nondescript moths that hold their wings “tent-like” over their bodies.
most appropriate application of SIMI is based on calculating confidence intervals using replication and computer simulation. However, Smith et al. (1986) suggest that the jackknife or bootstrap methods are better for estimating confidence intervals.

Czarnecki (1979) believed that a weakness of the SIMI is in that it is more sensitive to changes in dominant taxa than in the rare or uncommon taxa. But Johnson and Millie (1982) suggest that the occurrence of rare taxa in low abundance would not have a great effect on the overall structure of the community. For additional insight regarding the ecological significance and importance of intolerant but rare taxa please see Fausch (1990).

The SIMI has been found to be relatively insensitive to low and moderate changes in community structure (Boyle et al. 1990). Despite the fact that the SIMI has probably been used more frequently (particularly in algae studies) than any other single similarity index, there is considerable confusion and lack of accord in the literature with respect to the interpretation of SIMI values. Most investigators agree that more than one similarity index should be used when making community comparisons.

**Percentage Similarity (PS).** The PS is based on the comparison of percentages of organisms at the genus level in two communities (Whittaker 1952). Values for this index can range from 0 (no taxa in common) to 1 (all taxa in common and in same percentages). The Vermont Department of Environmental Conservation has developed guidelines for interpreting PS values: \( PS > 0.75 \) represents no significant biological alterations (typical for comparisons between replicates), \( PS 0.25 – 0.75 \) indicates a probable change has occurred, and \( PS < 0.25 \) represents a significant change has occurred (S. L. Fisk, Ohio EPA, pers. comm. 1987).

There has been much debate on the value of PS as similarity index. Whittaker (1952) reports that the PS was inaccurate when the relative proportions of the taxa were equal but the overall abundance was great. However, Brock (1977) points out that this inaccuracy may not be important in pollution studies where severe effects usually change dominance; if dominance shifts, and the taxa are in balance between the two communities, then they may be functioning the same. Boyle et al. (1990) found that the PS was less sensitive to changes in community structure than Jaccard’s Index. Additionally, Brock (1977) found that for separating structural and functional differences between communities the PS might be better than the B similarity index. Cao et al. (1997a) demonstrated that PS (along with several other similarity measures) are sensitive to sample size. He argues that because PS is more sensitive to sample size, it should therefore be
more sensitive to community changes and thus a good similarity measure. In a companion paper Cao et al. (1997b) berates PS for overweighing abundant species and ignoring ‘sample-specific’ species. Apparently some people make a living at this sort of thing!

**Ecological Distance (EDIS).** EDIS measurements have largely been used by terrestrial ecologists to measure dissimilarity (Clark 1952, Rhodes et al. 1969). The EDIS is based on the number of different genera (or species) in a set of paired samples and the differences in abundance of each taxa between the samples. Values can range from 0 (least dissimilar = identical) to 1 (greatest dissimilarity = no genera or species in common). The similarity between sites is estimated by subtracting the EDIS value from 1. The assignment of discrete values or ranges of values representing different degrees of dissimilarity has been arbitrary and there is no general consensus in the literature of what these values or ranges should be. Values also may vary widely depending on the design of studies and the specific organisms used in the analysis of community structure.

**Recommendations**

Macroinvertebrate data can be used in a variety of ways for making bioassessments. The method of analysis differs based on the project objectives (i.e., are the data to be used for baseline monitoring, trends analysis, or detecting unknown sources of pollution?). Whatever the case, one first must define a question. This determines a path to follow, allows determination of data collection issues (i.e., sampling or monitoring plan or design), and helps focus on the metrics that should be examined. Evaluations may be inconclusive using existing data; it is ok to design a new sampling strategy tailored to answer a specific question.

---

For additional recommendations, there is material available at the US EPA Rapid Bioassessment Protocols website. [http://www.epa.gov/owow/monitoring/rbp/](http://www.epa.gov/owow/monitoring/rbp/)

Macroinvertebrate data, as with all biological data, exhibit a considerable amount of natural variability both temporally (i.e., seasonal) and spatially (i.e., erosional versus depositional areas). However, understanding what an increase or decrease in a metric means and which direction indicates a condition of improvement or deterioration is vital. It is important to recognize an impact when one occurs and one should have an idea of how much change must occur in order to represent a significant impact (keeping in mind that some changes are natural and may not be related to human impacts). If the sampling design is carefully planned to answer a question, the results will be able to separate and discriminate between human induced impacts and background variation or noise.

**Objectives**

The first step in making a bioassessment is to define the objectives. It is important to clearly define what do you want to know or examine. For example, you are interested in measuring the biological condition at a particular...
site on a stream and have reason to believe (or are certain) that some form of impact has occurred at or above the sampling site. You would like to determine whether that impact has had a significant effect on the macroinvertebrate community at that site.

If you have access to historical data recorded at the sites of interest, it is possible to attempt a Before-After, Control-Impact (BACI) sampling design. You could then collect a new set of samples (preferably three or more replicates) from the site and compare the results with the historical data using any of the various statistical procedures. However, this simple method of before-after comparison may allow you to detect a difference or change but the comparison may not necessarily indicate a direct cause-effect relationship between the change and the suspected factor. Without sampling other sites in the watershed you cannot be sure that some other factor (i.e., drought or flooding) may have caused the significant change in the macroinvertebrate community at your sampling sites. Additionally, the differences seen between the new set of data and the historical data may be the result of slight differences in emergence patterns between the two data sets (i.e., climatic noise). If you are in the position of gathering new data it may be beneficial to sample comparable sites in adjoining basins or upstream of the suspected point of impact. These sites also should have historical data available along the same time scale as at the ‘impacted’ sites.

In other circumstances, historical data do not exist and you may be sampling a site for the first time. In this case the best approach is to compare the new data (a single sample may be adequate, but always take replicates if budget and time allow) with existing data representing least-disturbed reference sites of a similar stream order, ecoregion, watershed (Omernik et al. 2000), or Relatively Homogeneous Units (RHU) (Rheaume et al. 1996, Robertson and Saad 1996). An analysis of the FBI value may be a good start. For comparisons, it is important to use data collected during the same season, within the same habitat, and using the same sampling and laboratory protocols. While this type of analysis can not determine if or when a change has occurred, it can quantify the relative difference between the macroinvertebrate community at the site in question and the RHU. The reference site could be a ‘least-disturbed’ site or even represent an upstream (possibly unimpaired?) or downstream (recovery?) site. Careful examination of the differences in community structure among sites may help form a hypothesis as to what caused the differences. Keep in mind that some differences among sampling sites in community structure may result from subtle differences in substrate (particle size, texture, composition, algae, macrophytes), flow (velocity, depth), temperature (springs, shading), or channel position (head versus bottom of riffle or snag).
Microhabitat Considerations
Most lotic macroinvertebrate populations exhibit a contagious or clumped distribution pattern in a stream matrix. Because of this, most macroinvertebrate data exhibit a non-normal distribution pattern and require the use of non-parametric statistics. Macroinvertebrate population data can vary depending on where and when samples are taken. Similar microhabitats must be sampled both temporally and spatially to insure comparable data. In some cases streams do not have suitable riffle substrates to sample (i.e. many Wisconsin streams in the central sands area). In these situations it is acceptable to sample snag habitat in place of riffle habitat.

In a recent study of central Wisconsin trout streams that included comparisons of 82 sets of 5 replicate samples (total of 410 samples), snag samples exhibited greater variability than riffle samples using the HBI, 5-25 Max HBI, FBI, and MTV (S. Szczytko, UWSP, unpublished data). The yearly mean HBI value from snag samples was always higher and at least one water quality classification poorer than the THBI determined from riffle samples. Annual mean values of 5-25 Max HBI, FBI and MTV were also higher using snag samples than riffle samples (S. Szczytko, UW Stevens Point, unpublished data). These results indicate that if you are relegated to sample snag habitats instead of the recommended riffles, the resulting metric values are likely to be higher than the THBI.

Choosing the Appropriate Metrics
There are a number of choices available to analyze the metric data. Depending on the quantity, quality, distribution characteristics of the data, and the objectives any number of statistical tools can be applied. These tools may range from simple graphic analysis techniques (i.e., univariate plots or bivariate regression plots) to more complex procedures (i.e., ANOVA, t-tests, chi-square, etc.) and a wide range of other parametric and non-parametric analyses.

The various similarity (or dissimilarity) indices are particularly useful for making relative comparisons among streams or when searching for trends in water quality at a particular site. Each metric comes with its own set of strengths and weaknesses and each may be more sensitive to different forms of impacts. Table 5 provides a general overview of which metrics are useful for selected impacts. To select what statistical analysis to use you should be aware of the data assumptions that affect the correct application (i.e., data normally distributed, equal variance, etc.) and be prepared to make the necessary data transformations if required. When transformations fail to ‘normalize’ the data you may need to use non-parametric statistics. If you are in doubt, please consult the computer software help advisor or speak to a Department statistician.

\footnote{When using historical data it is best to: determine who collected the data, what type of data they collected, how they collected the data, where exactly in the stream reach the samples were collected, and why the data were collected in the first place.}

\footnote{See also: NABS (1993) workshop on Use of Biostatistics in Benthic Ecological Studies; Fore et al. (1994). Statistical properties of the IBI are discussed including inherent bias related to the way metric values are scored.}
<table>
<thead>
<tr>
<th>METRIC</th>
<th>NAME or DESCRIPTION</th>
<th>STRESSOR</th>
<th>STRESSOR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Organics, Nutrients, Low DO</td>
<td>Toxics, Contaminants, Heavy Metals</td>
</tr>
<tr>
<td>Biotic Indices</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HBI</td>
<td>Hilsenhoff Biotic Index</td>
<td>Increases</td>
<td>n.r.</td>
</tr>
<tr>
<td>FBI</td>
<td>Family Biotic Index</td>
<td>Increases</td>
<td>n.r.</td>
</tr>
<tr>
<td>MTV</td>
<td>Mean Tolerance Value</td>
<td>Increases</td>
<td>n.r.</td>
</tr>
<tr>
<td>MAX-10</td>
<td>Max-10 HBI</td>
<td>Increases</td>
<td>n.r.</td>
</tr>
<tr>
<td>Taxa Richness</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SR</td>
<td>Total Species Richness</td>
<td>Decreases</td>
<td>Decreases</td>
</tr>
<tr>
<td>GR</td>
<td>Generic Richness</td>
<td>Decreases</td>
<td>Decreases</td>
</tr>
<tr>
<td>EPTG</td>
<td>Ephemeroptera-Plecoptera-Trichoptera Generic Richness</td>
<td>Decreases</td>
<td>Decreases</td>
</tr>
<tr>
<td>Diversity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d</td>
<td>Margalef’s Diversity Index</td>
<td>Increase or Decrease</td>
<td>Decreases</td>
</tr>
<tr>
<td>Trophic Function</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percent Scrapers</td>
<td>Percent of total represented by Scrapers</td>
<td>Decreases</td>
<td>Decreases</td>
</tr>
<tr>
<td>Percent Filterers</td>
<td>Percent of total represented by Filterers</td>
<td>?</td>
<td>Decreases?</td>
</tr>
<tr>
<td>Percent Shredders</td>
<td>Percent of total represented by Shredders</td>
<td>Decreases</td>
<td>Decreases</td>
</tr>
<tr>
<td>Percent Collectors</td>
<td>Percent of total represented by Collectors</td>
<td>Increases</td>
<td>Increases</td>
</tr>
<tr>
<td>Percent Scrapers/percent Filterers</td>
<td>Ratio percent Scrapers to percent Filterers</td>
<td>n.r.</td>
<td>n.r.</td>
</tr>
<tr>
<td>Percent Scrapers/percent Gatherers</td>
<td>Ratio percent Scrapers to percent Gatherers</td>
<td>n.r.</td>
<td>n.r.</td>
</tr>
<tr>
<td>Percent Scrapers/percent Collectors</td>
<td>Ratio percent Scrapers to percent Collectors</td>
<td>n.r.</td>
<td>n.r.</td>
</tr>
<tr>
<td>Dominance</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dominant Families</td>
<td>Percent of total count represented by top five families</td>
<td>Response at extremes</td>
<td>Some chironomids tolerant to metals</td>
</tr>
<tr>
<td>Dominant Genera</td>
<td>Percent of total count represented by top five genera</td>
<td>Response at extremes</td>
<td></td>
</tr>
<tr>
<td>Dominant Species</td>
<td>Percent of total count represented by top five species</td>
<td>Response at extremes</td>
<td></td>
</tr>
<tr>
<td>Comparison Metrics</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCL</td>
<td>Coefficient of Community Loss</td>
<td>Increases</td>
<td>Increases</td>
</tr>
<tr>
<td>CS</td>
<td>Coefficient of Similarity</td>
<td>Decreases</td>
<td>Decreases</td>
</tr>
<tr>
<td>SIMI</td>
<td>Stander’s Similarity Index</td>
<td>Decreases</td>
<td>Decreases</td>
</tr>
<tr>
<td>PS</td>
<td>Percentage Similarity</td>
<td>Decreases</td>
<td>Decreases</td>
</tr>
<tr>
<td>B</td>
<td>B Similarity Index</td>
<td>Decreases</td>
<td>Decreases</td>
</tr>
<tr>
<td>EDIS</td>
<td>Ecological Distance</td>
<td>Increases</td>
<td>Increases</td>
</tr>
<tr>
<td>STRESSOR</td>
<td>STRESSOR</td>
<td>STRESSOR</td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td></td>
</tr>
<tr>
<td>Inorganics, Sediments</td>
<td>Flow Disruption</td>
<td>Thermal</td>
<td></td>
</tr>
<tr>
<td>n.r.</td>
<td>n.r.</td>
<td>n.r.</td>
<td></td>
</tr>
<tr>
<td>n.r.</td>
<td>n.r.</td>
<td>n.r.</td>
<td></td>
</tr>
<tr>
<td>n.r.</td>
<td>n.r.</td>
<td>n.r.</td>
<td></td>
</tr>
<tr>
<td>n.r.</td>
<td>n.r.</td>
<td>n.r.</td>
<td></td>
</tr>
<tr>
<td>?</td>
<td>Decreases</td>
<td>Decreases</td>
<td></td>
</tr>
<tr>
<td>?</td>
<td>Decreases</td>
<td>Decreases</td>
<td></td>
</tr>
<tr>
<td>?</td>
<td>Decreases</td>
<td>Decreases</td>
<td></td>
</tr>
<tr>
<td>?</td>
<td>Decreases</td>
<td>Decreases</td>
<td></td>
</tr>
<tr>
<td>Increase or Decrease</td>
<td>Increase or Decrease</td>
<td>Decreases</td>
<td></td>
</tr>
<tr>
<td>Decreases</td>
<td>?</td>
<td>Increases if periphyton production enhanced?</td>
<td></td>
</tr>
<tr>
<td>Increases?</td>
<td>?</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>Decreases</td>
<td>?</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>Increases</td>
<td>?</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>n.r.</td>
<td>n.r.</td>
<td>n.r.</td>
<td></td>
</tr>
<tr>
<td>n.r.</td>
<td>n.r.</td>
<td>n.r.</td>
<td></td>
</tr>
<tr>
<td>n.r.</td>
<td>n.r.</td>
<td>n.r.</td>
<td></td>
</tr>
<tr>
<td>Loss of many filterers, scrapers, shredders with increased imbeddedness.</td>
<td>Favors mobile forms over sessile forms above line of low water levels.</td>
<td>Interference of life cycles difficult to predict, but some insects may be eliminated while others fill vacated niches (try before/after examinations).</td>
<td></td>
</tr>
<tr>
<td>Increases</td>
<td>Increases</td>
<td>Increases</td>
<td></td>
</tr>
<tr>
<td>Decreases</td>
<td>Decreases</td>
<td>Decreases</td>
<td></td>
</tr>
<tr>
<td>Decreases</td>
<td>Decreases</td>
<td>Decreases</td>
<td></td>
</tr>
<tr>
<td>Decreases</td>
<td>Decreases</td>
<td>Decreases</td>
<td></td>
</tr>
<tr>
<td>Decreases</td>
<td>Decreases</td>
<td>Decreases</td>
<td></td>
</tr>
<tr>
<td>Increases</td>
<td>Increases</td>
<td>Increases</td>
<td></td>
</tr>
</tbody>
</table>
We can compare the percent change (before-after) or differences (control-impact) using a variety of methods. These include upstream-downstream site analysis, treatment versus non-treatment areas; or, using comparisons with ‘reference’ streams that represent undisturbed, least-disturbed, or ‘typical’ conditions. Statistical significance of percent change may be accomplished through repeated measurements (i.e., replication among sites or among similar stations) and the subsequent determination of change associated with varying levels of disturbance. Qualitative ratings (i.e., no impact, moderate impact, severe impact) can be assigned only after extensive sampling and analysis. If categorical ratings have been assigned to the metric we can assign a water quality rating score based on the data (i.e., good, fair, poor, etc.) and examine changes or differences in ratings between pairs of data or multiple observations. Comparisons to reference streams (Lenat and Barbour 1994, Hughes 1995, Bailey 1995, Reynoldson et al. 1997) may take into account a variety of factors that are known to influence community composition. These factors may include: ecoregions, habitat types, stream order, longitudinal position, coldwater, warmwater, and many others. Regardless of the methods chosen, you should always identify and account for the expected seasonal influences on the involved metrics.

Future Directions

Do not overlook new developments in the field of macroinvertebrate community analysis. The current literature is exploding with articles incorporating new and innovative approaches to interpret macroinvertebrate data. The following recommendations refer only to typical kick-net sampling assessments. Other macroinvertebrate assessment techniques come with their own set of metrics and recommendations. For example, in the assessment of sediment contaminated sites, oligochaetes and chironomids become more important and separate indices have been formulated for sediment sample analysis (Krieger and Ross 1993, Reynoldson et al. 1995, Hoke et al. 1995, Canfield et al. 1996). In addition, sampling methods, laboratory analysis, and data interpretation of sediment samples warrant a different set of recommendations (Aartila 1996, North American Benthological Society 1995). The following summaries present only a few of the various approaches that are being used.

Faunal Typology. Castella and Amoros (1983) used multivariate analysis to illustrate the faunal affinities of 26 macroinvertebrate communities in the Rhone River. Cluster analysis using different sets of aquatic orders, aquatic vegetation, sediment, and water quality represent an innovative approach to search for faunal descriptors and group associations.

Bioequivalence. McDonald and Erickson (1994) present a novel approach to determine whether two sites are ‘bioequivalent’. The authors discuss an alternative procedure to apply in cases where use of the classical null and alternative hypotheses is inappropriate. Although too lengthy to describe here, the method is founded in the medical field (i.e., drug-testing) and may be used to assign confidence limits for regulatory purposes.

Percent Modal Affinity. Novak and Bode (1992) developed the Percent Model Affinity (PMA) index to compare the percent similarity (Whittaker and Fairbanks 1958) of macroinvertebrate communities in selected streams to typical communities representative of unpolluted sites (reference sites). The degree of impact was judged by the ranges of similarity to the ‘model’ community. Barton (1996) found the PMA useful in detecting land use impacts on 213 Ontario streams. The sensitivity of the PMA index increases with the level of taxonomic resolution used (Barton 1996).

Other Dominance Indices. Beisel et al. (1996) compares the performance of four dominance indices: the Simpson Index, McIntosh dominance index, Berger-Parker index, and the Camargo dominance measure. They favor the McIntosh dominance index for its ability to detect small differences in the abundant taxa in communities and its relative insensitivity to rare taxa. However, Camargo (1997) refutes Beisel’s conclusions and provides evidence to suggest that the Camargo index is the best dominance and diversity metric available.

Other Biological Indices. Cairns et al. (1968) present the Sequential Comparison Index (SCI) (Resh and Price 1984). Death (1996) discusses species abundance patterns as indicative of habitat stability. A number of new indices have been proposed based on the relationship between morphological deformities in
multimetric indices for use in evaluating the biological integrity of freshwater ecosystems is reviewed extensively by Simon and Lyons (1995). Multimetric indices offer an advantage over individual metrics in that they integrate effects of multiple stressors (Barbour et al. 1995). Karr (1981) presented the first Index of Biotic Integrity (IBI) based on fish community attributes in Illinois and Indiana streams. Since its introduction, the IBI and its many modifications have become very popular and incorporated into many state and agency management protocols. Currently multimetric indices have been developed or proposed for: Ohio (DeShon 1995), Florida (Barbour et al. 1996), the Tennessee Valley (Kerans and Karr 1994), Oregon (Fore et al. 1996), the Chesapeake Bay (Weisberg et al. 1997), Washington (Kleidal 1995), Arkansas (Shackleford 1988), Michigan (Grant and Thorpe 1993), Texas (Twidwell and Davis 1989), and several developing countries (Thorne and Williams 1997). La Liberte (1996) developed a macroinvertebrate multimetric for use in Wisconsin using data from 14 West Central Region counties with restricted application in the West Central Region. The macroinvertebrate multimetric is composed of six metrics, some adjusted to compensate for effects of stream width (La Liberte 1996).

Other Multivariate Methods. An increasing number of studies are using multivariate analyses (ordinations, canonical correspondence analysis) to examine macroinvertebrate communities (Culp and Davies 1980, Cortes 1992, Jackson 1993). Reynoldson et al. (1997) compared the performance characteristics of selected multivariate methods with selected multimetric methods. They concluded that multivariate approaches had higher precision and accuracy than the multimetric approaches. Fore et al. (1996) found that a multimetric index performed better than a principal components analysis in detecting impacts of disturbance on Oregon streams. Other studies use a combination of both methods to make water quality assessments (Zamora-Munoz and Alba-Tercedor 1996). It is evident that more studies are applying multivariate procedures to understand and explain macroinvertebrate community structure.

aquatic larvae and various contaminants (Warwick and Tisdale 1988, Vermeulen 1995). A number of deformities in Odonate larvae have been discovered in Wisconsin streams (B. Smith, Wisconsin DNR, pers. comm. 1998). Done and Reichelt (1998) suggest an innovative approach to evaluate the effectiveness of management decisions using four new indices based on biomass and trophic structure, endangered species, representativeness or uniqueness of species composition, and successional stage. They introduce the concept of ‘desired’ and ‘degraded’ states and discuss disturbance regimes in reference to management goals. Although the paper is directed at fisheries management it is clearly applicable to any resource.

Multimetric Indices. The development of multimetric indices for use in evaluating the biological integrity of freshwater ecosystems is reviewed extensively by Simon and Lyons (1995). Multimetric indices offer an advantage over individual metrics in that they integrate effects of multiple stressors (Barbour et al. 1995). Karr (1981) presented the first Index of Biotic Integrity (IBI) based on fish community attributes in Illinois and Indiana streams. Since its introduction, the IBI and its many modifications have become very popular and incorporated into many state and agency management protocols. Currently multimetric indices have been developed or proposed for: Ohio (DeShon 1995), Florida (Barbour et al. 1996), the Tennessee Valley (Kerans and Karr 1994), Oregon (Fore et al. 1996), the Chesapeake Bay (Weisberg et al. 1997), Washington (Kleidal 1995), Arkansas (Shackleford 1988), Michigan (Grant and Thorpe 1993), Texas (Twidwell and Davis 1989), and several developing countries (Thorne and Williams 1997). La Liberte (1996) developed a macroinvertebrate multimetric for use in Wisconsin using data from 14 West Central Region counties with restricted application in the West Central Region. The macroinvertebrate multimetric is composed of six metrics, some adjusted to compensate for effects of stream width (La Liberte 1996).

Other Multivariate Methods. An increasing number of studies are using multivariate analyses (ordinations, canonical correspondence analysis) to examine macroinvertebrate communities (Culp and Davies 1980, Cortes 1992, Jackson 1993). Reynoldson et al. (1997) compared the performance characteristics of selected multivariate methods with selected multimetric methods. They concluded that multivariate approaches had higher precision and accuracy than the multimetric approaches. Fore et al. (1996) found that a multimetric index performed better than a principal components analysis in detecting impacts of disturbance on Oregon streams. Other studies use a combination of both methods to make water quality assessments (Zamora-Munoz and Alba-Tercedor 1996). It is evident that more studies are applying multivariate procedures to understand and explain macroinvertebrate community structure.

aquatic larvae and various contaminants (Warwick and Tisdale 1988, Vermeulen 1995). A number of deformities in Odonate larvae have been discovered in Wisconsin streams (B. Smith, Wisconsin DNR, pers. comm. 1998). Done and Reichelt (1998) suggest an innovative approach to evaluate the effectiveness of management decisions using four new indices based on biomass and trophic structure, endangered species, representativeness or uniqueness of species composition, and successional stage. They introduce the concept of ‘desired’ and ‘degraded’ states and discuss disturbance regimes in reference to management goals. Although the paper is directed at fisheries management it is clearly applicable to any resource.

Multimetric Indices. The development of multimetric indices for use in evaluating the biological integrity of freshwater ecosystems is reviewed extensively by Simon and Lyons (1995). Multimetric indices offer an advantage over individual metrics in that they integrate effects of multiple stressors (Barbour et al. 1995). Karr (1981) presented the first Index of Biotic Integrity (IBI) based on fish community attributes in Illinois and Indiana streams. Since its introduction, the IBI and its many modifications have become very popular and incorporated into many state and agency management protocols. Currently multimetric indices have been developed or proposed for: Ohio (DeShon 1995), Florida (Barbour et al. 1996), the Tennessee Valley (Kerans and Karr 1994), Oregon (Fore et al. 1996), the Chesapeake Bay (Weisberg et al. 1997), Washington (Kleidal 1995), Arkansas (Shackleford 1988), Michigan (Grant and Thorpe 1993), Texas (Twidwell and Davis 1989), and several developing countries (Thorne and Williams 1997). La Liberte (1996) developed a macroinvertebrate multimetric for use in Wisconsin using data from 14 West Central Region counties with restricted application in the West Central Region. The macroinvertebrate multimetric is composed of six metrics, some adjusted to compensate for effects of stream width (La Liberte 1996).

Other Multivariate Methods. An increasing number of studies are using multivariate analyses (ordinations, canonical correspondence analysis) to examine macroinvertebrate communities (Culp and Davies 1980, Cortes 1992, Jackson 1993). Reynoldson et al. (1997) compared the performance characteristics of selected multivariate methods with selected multimetric methods. They concluded that multivariate approaches had higher precision and accuracy than the multimetric approaches. Fore et al. (1996) found that a multimetric index performed better than a principal components analysis in detecting impacts of disturbance on Oregon streams. Other studies use a combination of both methods to make water quality assessments (Zamora-Munoz and Alba-Tercedor 1996). It is evident that more studies are applying multivariate procedures to understand and explain macroinvertebrate community structure.

Isopoda (Aquatic Sowbugs):
Members of the class Crustacea (including isopods, amphipods, mysids, and decapods) are very important components of the invertebrate community found in both lentic and lotic habitats in Wisconsin. Aquatic sowbugs are a significant group of crustaceans occurring throughout North America. Of the four families, Asellidae is the most important in terms of taxa numbers. Asellus (= caecidotea; Jass and Klausmeier 1997) commonly are omnivore-detritivores found living in leaf detritus. Some taxa are relatively tolerant of organic pollution, whereas others are relegated to springheads.

Amphipoda (Aquatic Amphipods):
Approximately 90 species found in 3 families and 8 genera occur in North America. The families Gamaridae and Talitridae are common in Wisconsin. Scuds are benthic, primarily occur in shallow waters, feed upon detritus, and in-turn are important prey of a number of fish species. Scuds are laterally compressed and often swim on their sides, hence the common name sideswimmers.

Amphipod or scud.
**Predictive Models.** An alternative approach to using multimetric indices is predictive modeling. Developed primarily in England and Australia, predictive modeling compares measured and predicted communities based on selected environmental attributes (Winget and Magnum 1979, Wright et al. 1984, Armitage et al. 1987, Ormerod and Edwards 1987, Wright 1995, Norris 1995). The degree of impact can be measured by comparing predicted assemblages in the absence of impact with observed communities. Investigators have used different taxonomic levels from species to the family level with generally good success. Predictive models in England helped create the River Invertebrate Prediction and Classification System (RIVPACS). This system, which has been tested successfully in a number of European countries, is based on the application of TWINSPLAN computer technology. For information on the future of this application in the United States see the EPA web site at [http://www.epa.gov/OWOW/monitoring/techmon.html](http://www.epa.gov/OWOW/monitoring/techmon.html).

**Species Trait Modeling.** Many studies include elaborate examinations of species traits characteristics which appear to show exciting possibilities for the future (Richoux 1994, Cellot et al. 1994, Tachet et al. 1994, Townsend and Hildrew 1994, Usseglio-Polatera 1994, Townsend et al. 1997, Poff 1997, and Richards et al. 1997). This approach is based on the premise that habitat, in the form of both abiotic and biotic features, forms a templet from which aquatic insect species traits (i.e., age, life history characteristics, body form, etc.) evolve. As a result, this templet may predict what type of species may occur in any particular habitat. This approach is often used in combination with ‘fuzzy coding’ procedures (Chevenet et al. 1994, Roberts 1986) to graphically present relationships between species traits and environmental variables.

**Landscape Classification In Macroinvertebrate Bioassessments.** The ability to predict and classify macroinvertebrate community composition is of interest to water resources managers. The ability to detect human impacts often depends on our knowledge of what the biota was like prior to human influence. To assist in this effort, various ecological frameworks have been developed to characterize the biological properties of various geographic regions (i.e., ecoregions or provinces) (Omernik 1987, Bailey 1995). A recent volume of the *Journal of the North American Benthological Society* (Vol. 19, No. 3, September 2000) is devoted to this topic and should be reviewed in its entirety.
References


Hilsenhoff, W.L. 1977. Use of arthropods to evaluate water quality of streams. Technical Bulletin (100)1-15. Wisconsin Department of Natural Resources, Madison, WI.

Hilsenhoff, W.L. 1982. Using a biotic index to evaluate water quality in streams. Technical Bulletin (132)1-22. Wisconsin Department of Natural Resources, Madison, WI.


Murphy, P.M. 1978. The temporal variability in biotic indices. Environmental Pollution 17:227-236.


Ohio Environmental Protection Agency. 1990. The use of biocriteria in the Ohio EPA surface water monitoring and assessment program. Ohio Environmental Protection Agency, Ecological Assessment Section, Division of Water Quality Planning and Assessment, Columbus. 52pp.


Appendix A. Selected Literature Review

General Reviews
Cairns 1977. Biological integrity; diversity indices; functional integrity.
Debinski and Humphrey 1997. Biological diversity assessments; design considerations.
Farnsworth et al. 1979. Sediment and nutrients; biota.
Fausch et al. 1990. Concerns fish but applies to inverts as well.
Guhl 1987. Biotic, diversity, and similarity indices in aquatic ecosystems.
Leopler, Merry. No date. DNR review of published literature on biological indicators.
Noss 1990. Biodiversity; definitions, indicators, and scales.
Rosenberg and Resh 1993. Biomonitoring; statistical considerations (pp. 252-253).
Somers et al. 1998. Lakes; ordination, metrics, sample sizes.
Wallace and Webster 196. Review of macroinvertebrate functional groups.
Winget and Mangum 1979. Biotic Condition Index (USFS); integrates aspects of habitat, water quality, and macroinvertebrate community.

Evaluations of Rapid Bioassessment Protocols
Courtemanch 1996. Subsampling; effect on richness metrics.
Diamond et al. 1996. Comparisons CVs, discriminating power, ecoregions.
Growns et al. 1997. Effect of subsampling, family level, SIGNAL biotic index, EPT, taxa richness, discriminatory power.
Hannaarford and Resh 1995. California; habitat, metric evaluations.
Rabeni et al. 1999. Metric sensitivity to sample size; similarity comparisons.

9 Copies of all references listed in the selected literature review are available upon request from the DNR library. It is anticipated that this section of the guidance document will be updated periodically in an attempt to incorporate new and relevant literature on specific issues important to the Department. Any assistance with this effort will be greatly appreciated. We hope to make this reference collection searchable on the web and to create a more thorough index of the content of the included articles. Please note that this list of references is by no way complete, but is intended to serve as a starting point to find specific references that may support you in your work.
Resh 1995. Good review article on design issues.
Vinson and Hawkins 1996. Effect of sampling area and subsampling on taxa richness metrics.

**Biotic Indices**
Cao et al. 1997a-c. Modified Chandler; ASPT Index, Great Britain.
Chandler 1970. Precursor to HBI in Britain.
Chessman and McEvoy 1998. Three disturbance sensitivity indices; family level.
Chessman et al. 1997. Australia; ‘SIGNAL’ biotic index, family level, average score per family index.
DeShon 1995 (also Ohio EPA 1987 and 1988). Invertebrate Community Index, Ohio.
Grotheer et al. 1994. Colorado system of "biologically determined values" based on environmental preferences of taxa present.
Hilsenhoff 1977. Original HBI publication.
Hilsenhoff 1987. Revised Biotic Index.
Hilsenhoff 1988a. Family level biotic index.
Hilsenhoff 1988b. Seasonal corrections to biotic index.
Hilsenhoff 1998. Further modifications to HBI to account for seasonality; MAX-10 HBI.
Lang and Reymond 1995. RIVAUD 95 Index, Switzerland.
Lillie and Schlesser 1994 Mean tolerance value biotic index.
Murphy 1978. Temporal variability in six BIs and diversity indices, Wales.
Narf et al. 1984. Statistical procedures; HBIs in Wisconsin (out of date).
Shackleford 1988. Arkansas; Rapid bioassessments; Indicator Assemblage Index.
Stark 1993. MCI; mean tolerance value, New Zealand.
Wright 1995. United Kingdom RIVPACS. Review of the development of prediction and classification of invertebrates based on selected environmental attributes.
Wright et al. 1988. Family level; environmental data, RIVPACS predictive modeling.

**New Diversity Indices**
Camargo 1997. Dominance Index, importance of rare taxa.
Docampo and de Bikuna 1994. Complex paper dealing with development and testing of yet another diversity index.

**Taxa/Species Richness**
Cao et al. 1998. Importance of rare species in bioassessments, species richness.
Clements 1991. Species richness better than abundance (except heavy metals).
Death and Winterbourn 1995. Species richness and evenness influenced by disturbance; habitat stability.
Douglas and Lake 1994. Stone sizes and areas; species richness, indicators.
Fausch et al. 1990. Fish communities, but discussion applies to inverts.
Kerans et al. 1992. Influence of sampling protocol taxa richness; other metrics; pool versus riffles.
Larsen and Herlihy 1998. Richness; area versus fixed count.
Lenat 1988. Multihabitat sampling; EPT, taxa richness by ecoregions.
Lenz and Miller 1996. Wisconsin; Invertebrate comparisons using different collection methods.
Palomaki and Paasivirta 1993. Species richness influenced by area and sampling sizes in Finland.
Rae 1990. Species richness; depth and discharge.
Siegel and Gegrman 1982. Rarefaction, taxonomic diversity and richness.
Vinson and Hawkins 1996. Effect of sample area and subsampling on taxa richness.

Indicator Species
Fausch et al. 1990. Fish, but good discussion of indicator taxa.
Johnson et al. 1993. Review article
Patrick and Palavage 1994. General review species as indicators.

Specific Effects: Sediment/Suspended Sediment
Angradi 1999. Forestry management and sediments, macroinvertebrate metric response.
Armitage 1989. Substrate; community associations, predictive RIVPACS system used in Great Britain.
Farnsworth et al. 1979. EPA review article on sediments and nutrients.
Krieger and Ross 1993. Great Lakes sediments; changes in oligochaetes and chironomids
Luedtke and Brusven 1976. Sand; stream insects, drift and colonization issues in Idaho.
Newcombe and MacDonald 1991. Effects of sediments on fish and invertebrates.
Quinn and Hickey 1990b. Substrate index, invertebrates, floods, land use.
Reynoldson et al. 1995. BEAST index, multivariate tests, invertebrates, sediments.
Roback 1978. Chironomids and Mayflies; sediments and thermal effects.
Schrank 1982. Wisconsin review article on the effects of erosional sediment on streams.

Specific Effects: Physical Disturbance Including Flows, Floods, and Drought
Anderson 1992. Species richness and diversity by age of forested watershed.
Chessman and Robinson 1987. Effect of drought on macroinvertebrate richness; wastewater.
Death 1996. Habitat stability, shape of distribution curve.
Doeg et al. 1989. Australian study; invertebrate richness and densities; recovery after disturbance.
Duchrow 1983. Study done in the Ozarks; stonefly and mayfly response to disturbance.
Quinn and Hickey 1990a, 1990b. Flooding and substrate size; impact on species richness in New Zealand.
Rader 1997. Species traits for aquatic insects; drift, availability to salmonids.
Rae 1990. Effect of discharge on species richness.
Rosillon 1989. Belgium streams; yearly changes in invertebrates related to floods.

Specific Effects: Land Use
Collins and Pess 1997a and 1997b. Forestry practices, riparian buffers, large woody debris.
Not an invertebrate paper, but useful for management issues.
Corkum 1990. Invertebrate distributions and assemblages as influenced by land use.
Johnson and Gage 1997. Landscape approaches using GIS technology; review of multivariate techniques used in watershed analysis.
Lenat and Crawford 1994. Land use; effects on taxa richness and biotic index in North Carolina.
Richards and Host 1994. GIS study of relation between land use and macroinvertebrates.
Richards et al. 1993. Michigan studies; multivariate analysis CCA ordination; EPT.
Trimble 1997. Classic Coon Creek study in Wisconsin, riparian land use, grazing; physical effects (no invertebrates).
Wang et al. 1997. Wisconsin land use, habitat quality, and fish IBI.

Specific Effects: Herbivores/Grazing
Gilliam et al. 1989. Fish interactions, invertebrate distributions and community structure.
Hart and Robinson 1990. Periphyton, phosphorus enrichment and grazing, caddisflies.
Peterson and Tuchman 1994. Substrates, associations with microbiota; a review article.
Stewart 1987. Fish grazing and algae, fertilization.

Specific Effects: Gasoline and Road Runoff

Specific Effects: Climate
Hann et al. 1992. Climate effects on aquatic insects; debate article.
Sweeney et al. 1991. Climate change effects on aquatic insect distribution patterns.
Walker et al. 1992. Further discussion on climate effects; debate article.
Specific Effects: Chlorine/Chlorides
Williams et al. 2000. Ontario springs; Chloride Contamination Index.

Specific Effects: Trace or Heavy Metals/Contaminants
Canfield et al. 1996. Great Lakes, sediments, TRIAD approach, chemistry, toxicity, benthos.
Diamond et al. 1993. Acid mine drainage; metal toxicity; EPT, Bray-Curtis ordination.
Duchrow 1983. Ozark study; sensitivities; stoneflies and mayflies, lead tailings.
Faith et al. 1995. Mining aspects in Australia; BACI, family level, Bray-Curtis analysis.
Garie and McIntosh 1986. Urban runoff, heavy metals in New Jersey.
LaPoint et al. 1984. Number of taxa; community structure in 15 United States streams.
Perkins 1983. Effects of copper on aquatic insects; community comparison index.
Roline 1988. Diversity; Baetis tolerant, heavy metals in the Arkansas River.

Specific Effects: Acid Rain
France 1992. Lakes in Canada; amphipods as biotic index to acidification.
Stephenson et al. 1994. Littoral benthic macroinvertebrates; lake acidification, NMDS ordination, taxonomic importance index.

Specific Effects: Phosphorus and Other Nutrients
Clarke and Ainsworth 1993. Study dealing with Diatoms, but approach is applicable to other organisms; CANOCO, diatoms and environment.
Monda et al. 1995. Euclidean similarity index to detect effects of ammonia.

Specific Effects: Dams and Hydropower
Barwick et al. 1985. Differences in prey and fish feeding patterns below dams.
Bournaud et al. 1996. Longitudinal pattern in Rhone River; PCA analysis.
Cereghino and LaVandier 1998a. French study; Plecoptera, hydropeaking.
Cereghino and LaVandier 1998b. French study; hydropeaking, mayflies.
Camargo and de Jalon 1995. Spanish study; taxonomic richness, scrapers, BA-design, impoundments.
Converse et al. 1998. Colorado River flow regulation, shoreline habitat loss. Although orientated to fish, results are significant to insects.


Fisher and LaVoy 1972. Hydroelectric dams; effects on macroinvertebrate community.


Hooper 1993. Wisconsin study; seasonality, habitat, impoundment effects, snag sampling.

Inverarity et al. 1983. Welch study; use of Jaccard Similarity coefficients; clustering.

Knight and White 1982. Unpublished review article on environmental effects of dams on fish and insects. (available from DNR library)


Olmsted and Bolin 1996. Industrial perspective; electric utility; dams.

Pardo et al. 1998. Dams, mayflies in Australian streams.


**Biological Community Analyses**


Anderson and Day 1986. Invertebrate habitat association on Mississippi River.


Bargos et al. 1990. Use of correspondence analysis ordination; biotic index in Spain.


Barton 1996. Use of Percent Model Affinity index in Ontario streams.


Bournaud et al. 1996. France study; large river; family level ordination analysis, taxonomic resolution issues.


Bruns and Minshall 1985. Functional feeding groups; ordination river continuum.

Bunn 1995. Australian study; overview, community persistence using ordination.


Chessman 1995. Family level BI; multihabitats in Australia.

Clements 1991. Discharge impacts taxa richness and CV, percent dominance, sampling methods.


Corkum and Ciborowski 1988. Community associations with environmental data using CANOCO.


Culp and Davies 1980. Saskatchewan study; substrate, invertebrate densities, reciprocal averaging, polar ordination.

Elstad 1986. General community composition; Mississippi River.


Fleituch 1992. Uses Jaccard Similarity Index; biotic indices; environmental variables.


Growns et al. 1992. Australian wetlands; ordination MDS.
Growns et al. 1995. Australian rivers; ordination in rapid assessments.
Jackson 1993. Use of multivariate analysis ordination invert communities; good review of methods.
Johnson and Gage 1997. Aquatic ecosystems, landscape scale, GIS.
Koetsier et al. 1996. Alkalinity; community associations, landscape scale.
Lyons 1996. Wisconsin fish communities; ecoregions; methods.
Mackay 1992. Invertebrate colonization, good review article.
Moss et al. 1987. Prediction of total community in Great Britain; Twinspan.
Pearson 1986. Species co-occurrence emphasis.
Poff 1997. Landscape habitat filters, species traits, prediction of community structure.
Quinn and Hickey 1990b. New Zealand rivers; taxa richness, invert community index.
Rabeni and Gibbs 1980. Ordination and Bray-Curtis; PS similarities, Maine rivers.
Rae 1985. PCA chironomid community comparisons in Ohio; sediments.
Richards and Host 1994. Land uses; communities; PCA analysis, richness, filterers.
Richoux 1994. Habitat templetts; beetles; species traits, Upper Rhone River.
Rossaro and Pietrangelo 1993. Correspondence analysis; BI’s; richness, etc.
Sheaffer and Nickum 1986. Invertebrates, habitats; backwaters of the Mississippi River.
Sheldon and Haick 1981. PCA and canonical correlation in Montana.
Smock et al. 1989. LWD (large woody debris); percent shredders.
Sollak 1985. Chironomids in Canada shifting sand bottomed rivers.
Tachet et al. 1994. Habitat templetts; caddisflies; species traits, Upper Rhone River.
Usseglio-Polatera 1994. Habitat templet, aquatic insect richness, species traits, Upper Rhone River.
Walley and Fontama 1998. Use of ‘neural networks’ to predict invertebrate index scores in Britain; ASPT, family level.
Wohl et al. 1995. Euclidean similarity/Cluster analysis, southern Appalachian streams.
Wright 1995. RIVPACS; biological assessment Britain.
Wright et al. 1993. British study; Biological Monitoring Working Party System and RIVPACS.
Wright et al. 1995. NMDS ordination; Bray-Curtis, taxa richness; tax. resolution.
Zamora-Munoz and Alba-Tercedor 1996. Spanish study; BMWP score, twinspan and CCA ordination.

**Paired Comparisons: Similarity/Dissimilarity**

Barton and Metcalfe-Smith 1992. Comparison of eight indices or metrics in Canada.
Beisel et al. 1996. Comparison of four dominance indices.
Cao et al. 1997a. Effect of sample size on eleven similarity measures.
Faith et al. 1995. BACI and paired comparisons; Bray-Curtis, Kulczynski; family level.
James and Rathbun 1981. On avian communities, but valuable discussion on rarefaction and diversity indices.
Novak and Bode 1992. Presentation of Percentage Similarity Index.
Pedersen and Perkins 1986. Comparisons of six similarity coefficients; Canberra Metric Index.
Perkins 1983. Diversity and similarity comparisons; Bray-Curtis review.
Pinkham and Pearson 1976. Presentation of ‘B’ Similarity Index and others.
Wolda 1981. Comparison of similarity indices; Morisita Index preferred.

**Special Statistics**

Bunn 1995. Biological monitoring study design considerations.
Humphrey et al. 1995. Study design considerations; BACI.
McDonald and Erickson 1994. Bioequivalence; hypothesis testing.
Nemec and Brinkhurst 1988a and 1988b. Ordination techniques; multivariate analysis.
Stewart-Oaken et al. 1986. More on pseudoreplication.
Other Indices: Morphological Deformities
Lemly 1982. Wax and bacteria buildup on filamentous gills.

Other Methods
Cairns et al. 1968. Sequential Sampling Index.
Castella and Amoros 1988. Graphic presentations and comparisons; faunal typology.
Peters et al. 1989. Utilization and habitat preference metrics for macroinvertebrates and fish in Nebraska.
Resh and Price 1984. Test of Sequential Sampling Index.
Southwood 1977. Classic work on habitat templet theory.
Townsend and Hildrew 1994. Habitat templet theory, species traits.

Composite Indices/ Multimetric
Fore et al. 1996. Multimetric and multivariate approaches evaluated.
Gaunt and Thorpe 1993. Michigan biological condition scoring criteria; EPT, dominance, richness, CLI, etc.
Ohio Environmental Protection Agency 1987 and 1990. Invertebrate community Index; based on 10 metrics.
Resh et al. 1995. Rapid bioassessment metrics; richness, diversity, similarity, BIs, etc.
Shackleford 1988. Bioassessment in Arkansas; dominants, taxa richness, indicators, etc.

Applications
Crunkilton and Duchrow 1991. Missouri; density, diversity, richness, modified Margalef; variation with stream order.
Day et al. 1992. Horn's Similarity Index; cluster analysis, drift, Upper Mississippi River.
Dilley 1992. Test evaluation of Ohio’s ICI.
Heckman et al. 1990. German study; evaluation of several multivariate methods and indices.
Hooper 1993. Use of 25 metrics Wisconsin; riffle/snag and dam effects; HBI, FBI, richness, trophic classes, and Margalef’s diversity.
Hughes et al. 1990. Incorporating ecoregions into bioassessments.
Johnson 1998. Detecting impacts in Swedish lakes; benthic quality index, taxa richness, ASPT.
Lewis and Smith 1992. HBI, Shannon-Weaver diversity, CLI, similarity indices including Jaccard’s.
Lillie and Schlesser 1994. Southwest Wisconsin streams; Bray-Curtis, functional classes; BIs.
Maret 1988. Use of Shannon Diversity; modified HBI in Nebraska streams.
Modde and Drewes 1990. Artificial vs. natural substrates; HBIs, percent similarities; effect of flows on biotic index scores.
Murphy 1978. Temporal variability in BIs and diversity indices; Margalef’s.
Oda et al. 1991. Japanese study; diversity and BIs; PCA analysis.
Pedersen and Perkins 1986. Compares 6 similarity coefficients; Canberra-Metric best; discusses weakness of Bray-Curtis.
Rabeni et al. 1985. Response of inverts to pollution abatement; BIs and PS similarity and ordination.
Rheume et al. 1996. Invertebrate in eastern Wisconsin; PCA and other metrics, geography.
Rosillon 1989. Richness, diversity (Shannon-Weaver), density, annual changes.
Rossaro and Pietrangelo 1993. Test of CA ordination and biotic indices in Italy.
Twidwell and Davis 1989. Texas study; least disturbed streams, richness, diversity, dominance, etc.
Zamora-Munoz and Alba-Tercedor 1996. CANOCO; twinspan evaluation of BMWP biotic index in Spanish rivers.

**Biological Integrity**

Angermeier and Karr 1994. Definitions and policy applications; Biotic integrity versus diversity.
Barbour et al. 1996. Florida Stream Condition Index tested using reference sites; metric sensitivity discussed.
Fore et al. 1994. Statistical properties of the IBI; important to benthic BIs also.
Harig and Bain 1998. Biological integrity of wilderness lakes in the Adirondacks using fish, phytoplankton, and zooplankton.
Jennings et al. 1999. IBI for inland lakes in Wisconsin using fish, but good for concepts and principles.
Kerans and Karr 1994. TVA rivers (B-IBI) index, 13 invert attributes; pools/riffles.
Polls 1994. Use of reference sites, etc; industry perspective.
Appendix B. A History of the Development of Macroinvertebrate Biological Indices in Wisconsin

The use of macroinvertebrate biotic indices for water quality assessment in Wisconsin began with the development of the Hilsenhoff Biotic Index (HBI) at the University of Wisconsin Madison in cooperation with the DNR (Hilsenhoff 1977). The original index was based on collections from a set of 53 Wisconsin streams (29 undisturbed and 24 polluted or disturbed). Tolerance values were assigned to arthropod taxa on the basis of each taxon’s affiliation with an assigned (subjective) stream disturbance class from 0 (undisturbed) to 5 (severely polluted). The original index provided for five water quality classes. Hilsenhoff (1982) added improvements to the index and methodology after sampling over 1000 streams, changing tolerance values for some taxa, assigning tolerance values for additional taxa, and expanding the water quality rating scale to include six water quality classes. Hilsenhoff modified and refined the HBI in a series of publications. These modifications include: incorporating additional tolerance values, expanding the scale from 0-5 to 0-10, adding a seventh water quality class, deriving seasonal correction values, and modifying the index to minimize undue effects arising from dominant taxa (Hilsenhoff 1987, Hilsenhoff 1988b, Hilsenhoff 1998).

Hilsenhoff (1988a) developed a family level version of the HBI with seven water quality classes and Lillie and Schlessier (1994) proposed a complementary index measure based on a non weighted, mean tolerance value of all taxa represented in an HBI sample. In 1989, the standard HBI was incorporated into the national “Rapid Bioassessment Protocols for Use in Streams and Rivers” (Plafkin et al.1989). Lillie (2000) has developed a multi-metric macroinvertebrate index for use in wetlands.

Early assessments of water quality in Wisconsin using the HBI were done by Dr. Hilsenhoff at UW-Madison. Beginning in 1983, a contractual agreement was established with the UWSP to process HBI macroinvertebrate samples. At that time there was a basic program that calculated the HBI from sample data. This program was used by Bob Young (DNR) and further developed at the North Central District headquarters by a computer programmer. The program was developed in the dBase® III environment and has since been modified to suit DNR research needs and interests. In 1987 the program was modified to include the various metrics and utilities that are currently used. In 1992 the program was moved to the dBase® IV environment to increase speed and allow functionality as a stand alone program outside of the dBase® IV environment. At this time additional utilities were added to make the program easier to use.

A ‘Taxamaster’ file, which includes unique organism ID numbers and their associated tolerance values, has been maintained since 1983. Because the taxonomic literature is continually monitored by UWSP staff, information derived from this effort is used to update the ‘Taxamaster’ file on a routine basis. The latest version of the program (BUG-PROGRAM 6.0) includes computation of the HBI (MAX x) (Hilsenhoff 1998) and replaces Margalef’s Diversity Index with the Shannon Index.

When examining historical macroinvertebrate data, it is important to know why the samples were collected and where the collections were made. Some historical macroinvertebrate data may represent atypical samples (i.e., snags, pools, impounded areas) or anomalous conditions (i.e., fertilizer spill, during or after severe flooding or drought). In other cases, documentation and field records associated with collections may be incomplete or simply missing. Such data should be used with caution to avoid misapplication.

A high degree of taxonomic training and experience is required to correctly identify aquatic insects. Under the direction of highly qualified taxonomists, macroinvertebrate work (including most HBI calculations) conducted in Wisconsin has been done at UW-Madison, UWSP, and the UW Superior Lake Superior Research Institute. The Lake Superior Research Institute has been contracted by the DNR to do taxonomic and field sampling for in-place pollutant studies, endangered resource inventories, and dam removal assessments. In addition, many independent research studies have been conducted outside of DNR by various universities and colleges and the resultant data are available in unpublished theses or published articles. Other studies have been conducted by the US Fish and Wildlife Service, US Geological Survey, Natural Resources Conservation Service, The Nature Conservancy, and several private consultants. Within the DNR, the Bureau of Integrated Science Services has conducted several research studies involving macroinvertebrate and HBI data. The Bureau of Endangered Resources has conducted survey work to maintain the Natural Heritage Inventory Database that tracks the occurrences and distributions of species listed as endangered, threatened, or of special concern. Some regional offices have also conducted their own family level biotic assessments and several department staff members maintain private collections. More recently, numerous studies (in some cases only collections) have been conducted by schools and non-profit organizations using simplified rapid bioassessment protocols.
About the Authors

Richard A. Lillie, Wisconsin Department of Natural Resources
Bureau of Integrated Science Services
1350 Femrite Road, Monona, WI 53716
Retired

Stanley W. Szczytko, College of Natural Resources
University of Wisconsin – Stevens Point
Stevens Point, WI 54481
E-mail: sszczytk@uwsp.edu

Michael A. Miller, Wisconsin Department of Natural Resources
Bureau of Fisheries Management and Habitat Protection
101 S. Webster St., PO Box 7921
Madison, WI 53707
E-mail: Michael.A.Miller@dnr.state.wi.us

Production Credits
Managing Editor: Martin P.A. Griffin
Layout and Production: Michelle Voss
Illustrations: Courtesy of Aquatic Insects of Wisconsin: Keys to Wisconsin Genera and Notes on Biology, Habitat, Distribution and Species (Natural History Museum Council, University of Wisconsin Madison 1995), Wildlife, People and the Land (Wisconsin Department of Natural Resources), Water Resources ClipArt Collections I & II (University of Wisconsin-Extension Environmental Resources Center), Jessie Hewitt, Jim McEvoy, and Jens von Sivers.
Photographs: Courtesy of the North American Benthological Society (NABS).

The Wisconsin Department of Natural Resources provides equal opportunity in its employment, programs, services, and functions under an Affirmative Action Plan. If you have any questions, please write to Equal Opportunity Office, Department of Interior, Washington, D.C. 20240.

This publication is available in alternative format (large print, Braille, audio tape, etc.) upon request. Please call (608) 266-0531 for more information.