CABIN
(Canadian Aquatic Biomonitoring Network)
Invertebrate Biomonitoring
Field and Laboratory Manual

Prepared by:

Trefor B. Reynoldson, Craig Logan, Tim Pascoe and Sherri .P. Thompson

National Water Research Institute

Environment Canada

Environment Canada Environnement Canada

1. Introduction

Over the past ten years Environment Canada's National Water Research Institute (NWRI) has conducted major programmes in both the Great Lakes and Fraser River as part of a programme to develop a national reference database on benthic invertebrates for Canada. From these studies has developed the concept of a Canadian Aquatic Biomonitoring Network (CABIN). The initial focus of the CABIN network is the use of benthic invertebrate communities in ecological assessment. A critical part of this programme is the establishment of a standard set of protocols and methods for the various phases of data collection and processing. This manual describes the methods recommended by NWRI for collection of both biological and habitat data, in running water habitats. In addition we have provided a set of tables and forms used by the programme. In laying out the manual we have attempted to provide a rationale for the measurement of many of the variables and the layout of the field sections of the manual follows that on the field sheets

The layout of the manual attempts to follow the logical progression and sequence of events in designing, collecting, analysing, and interpreting a project. There are 4 major sections:

- **Data Collection**: where there is some superficial discussion about study design, but the main focus is the site selection and on site procedures and protocols.
- **Laboratory Processing**: which describe the actual sequence of processes for handling samples, sorting invertebrate samples and data management and data entry.
- **Taxonomy**: which provides an identification key to the family level for the organisms.
- **Data Interpretation**: where some recommended summary descriptors are discussed.
- **Appendix**: which provides the various data sheets required for the project.

We hope that this will provide a useful resource for others involved in this type of work. In preparing this manula we have relied heavily on texts prepared by other authore and programmes. In particular we would like to acknowledge three sources:

Culp, J.C. and Halliwell, D.B. 1999. Volunteer based monitoring programme. Using benthic indicators to assess stream health. Instructors manual. Environment Canada, NWRI, Saskatoon.

D.M. Rosenberg, I.J. Davies, D.C. Cobb, and A.P. Weins. Freshwater Ecosystem Protocols: benthic macroinvertebrates. A web site developed by the Ecosystem Monitoring and Assessment Network:

http://www.eman-rese.ca/eman/ecotools/protocols/freshwater/benthics/intro.html

AUSRIVAS. The web site developed by the Australian National River Health Programme:

http://ausrivas.canberra.edu.au/

2. Pre-Field procedures

Study design

While this method does not attempt to address the many issues related to study design some mention should be made, as the success of any sampling programme is dependent on matching the questions being asked by the collection of data and how those data are collected. Detailed reference books have been prepared on study design and sampling two of the most useful are Elliott's (1977) publication on analysis methods for benthic invertebrates and Green's (1979) book on sampling design. In his book Green outlines five basic categories of study design:

Optimal;
Where impact is inferred form temporal change alone;
Baseline or monitoring studies;
Where impact is inferred from spatial pattern alone, and;
When there is no knowledge of where and when an impact occurred.

Optimal study designs

There are several different strategies in optimal design, and these are documented in the Environmental Effects Monitoring (EEM) Technical Guidance documents (http://www.ec.gc.ca/eem/english/default.cfm). The classic study design is that formalised by Stewart-Oaten *et al*. (1986) that uses before and after treatment at control and impact (BACI) sites to assess change, other study designs use a gradient approach, and more recently the Reference Condition Approach (Reynoldson *et al*. 1997) which selects multiple reference sites from a reference database as the "control" and individual test sites as the treatment (see Table 1). This latter RCA has been proposed as the basis of a national aquatic biomonitoring programme (CABIN) for Canada (Reynoldson *et al*. 1999), and is the basic study design for which this manual has been developed.

Table 1. Benthic sampling program designs

Design Type	Reference / "control" Areas	"Impact" Areas	Statistics
Before/After/Control/Impact (BACI)	A single reference area upstream of the discharge or future discharge	Near field/Far- field/Far-far field (optional) as defined by effluent concentration, or distance from exposure in receiving environment	ANOVA
Control/Impact (C/I)	A single reference area, upstream of discharge or stressor	Near field/Far- field/Far-far field (optional) as defined by effluent concentration in receiving environment, or distance from exposure in receiving environment	ANOVA
Multiple - Control/Impact (MCI)	Multiple reference areas in the same	Near field/Far- field/Far-far field (optional) as defined by effluent	MANOVA/ Multivariate

	or environmentally similar adjacent watersheds or bays	concentration in receiving environment	
Simple Gradient (SG)	a series of reference stations with no or low effluent levels, situated towards the end of a declining gradient of the stressor	single gradient through "near"-field, "far"- field and "far-far"-field (optional) levels of effluent in the receiving environment	Regression/ Multivariate
Radial or Multiple Gradient (RMG)	multiple series of reference stations with no or low effluent levels situated towards the end of several transects of declining levels of mine effluent	multiple gradients through "near"-field, "far"- field levels of effluent in the receiving environment	Regression/ Multivariate
Reference Condition Approach (RCA)	multiple series of reference stations with no or low effluent levels situated in similar drainage basins within the same ecoregion	series of stations within the zone of influence which are tested against the reference stations	Multivariate

Site selection/name

Once a study design has been developed (Table 1) then sites must be selected to address the question that has been formulated. For the purposes of this programme a site is determined as a river reach six times the bankfull width. This reach of stream contains a full pool/riffle sequence.

As a first step in determining the most appropriate study design for any monitoring program, it is necessary to locate suitable reference areas. Environmental and biological data obtained from reference areas when compared to exposure areas can detect impairment of aquatic life (Yoder 1991), diagnose stressors (Fletcher *et al.* 2001), provide data on temporal and spatial trends (Yoder 1989) and provide data for water resource summaries for government agencies (OEPA 1990). Identification of appropriate reference areas is essential for assessment and should initially be done in the pre-field phase.

A reference area should have no effluent and natural habitat features similar to those of exposure areas. However, reference areas are not to represent pristine (or pre-European settlement) conditions, but areas in which impacts are lowest or disturbance is minimal (Simon 1991, Omernik 1995). For example, Omernik (1995) identifies references sites in drainage basins that lack municipal and industrial influence even within the polluted Huron-Erie corridor of the Great Lakes. A second example can be found on the prairies where minimally impacted

reference areas may be located in drainage basins that lack point source effluents or logging activities but, unavoidably, have a high degree of agricultural activity. Therefore, the characteristics of these "least impacted" areas will, necessarily, differ across the country, however, the approach to defining and identifying reference areas should be consistent.

Because reference sites will vary among different landscapes, approaches have been developed to classify the land that rivers and lakes reside to predict aquatic biotic assemblages (Corkum 1989, 1992, Hughes 1995, Omernik 1995). A hierarchical classification system is a way of simplifying sampling procedures and management strategies by organizing a variable landscape by discriminating among features at several scales of resolution Two recently proposed classification schemes for characterizing aquatic habitats are compared and summarized by Bisson and Montgomery (1996). Both systems are based on hierarchies of topographic and fluvial characteristics and both employ descriptors that are measurable and ecologically relevant. There are several reasons to apply physical classification systems to ecological studies including; simple descriptions of physical changes in response to human or natural disturbances (Gordon et al. 1992), the grouping of sites into like physical units for comparisons, extrapolation to other areas with similar features, and, importantly, similar assemblages of macroinvertebrates are likely to occur at sites with similar landscape and biome features (Corkum 1992).

Careful consideration is required, not only for the consistent application of the selection criteria for reference areas, but also in determining how many reference areas or stations (RCA) will be examined for all study designs. Many authors (Reynoldson et al. 1997, Hawkins et al. 2000, Wright et al. 1984) suggest that reliance on only one field site as a control is not adequate to evaluate the effects of a potential impact. It is possible for natural communities at two areas to diverge or converge over time in the absence of any environmental impact (Underwood 1991) which would potentially obscure the effects from anthropogenic inputs. Some other common problems of study designs with a single reference area include; a confounded control area (Eberhardt 1978, Underwood 1994, Resh 1995, Environment Canada 1997), limited capacity for extrapolation to other sites, limited ability to calculate natural variability and an inability to address non-point source factors (Reynoldson et al. 1997).

*Table 2: Habitat and geographical features, which should be similar between, reference and exposure areas. Features in bold **must** be similar, while those in non-bold **should** be similar.*

Ecoregion	**Riparian vegetation**
Drainage basin	Stream order
Basin area if different basins	Channel gradient
Land use	**Depth**
Other inputs	**Velocity**
Dominant habitat types	**Bankfull width**
Substratum	Channel pattern

Table 2 contains a list of features that should be similar between the reference and the exposure areas for a benthic invertebrate survey. Most of these can be obtained from topographic maps. The following specific points should be considered during the selection of reference sites for a benthic invertebrate assessment study.

- As reference areas are to be selected to match the habitat features of the exposure area(s), identification of the exposure area and it's habitat features should precede the selection of reference areas.
- Selection of an appropriate classification scheme is important. As there are several different hierarchical classification schemes, which are described in detail in the original sources, they are referenced here as examples. (Conquest et al 1994, Meador et al 1993).
- The size of the drainage basin selected is based on stream order. If test site is located on a second order stream, the drainage basin should be delineated at a third order stream (i.e. at the junction of two second order streams).
- If upstream inputs are absent, reference area(s) should be within the drainage basin in which the test site is located.
- If non-point or other point source inputs occur upstream of the test sites, select reference area(s) in the nearest comparable drainage basin with minimal development that occurs within the same ecoregion.
- If physical disturbance of the river valley is associated with the test sites, effluent effects may be confounded by the disturbance. Accordingly, additional physically matched reference areas should be selected.

Finally, a site nomenclature should be established. As sites are going to be maintained in a database as unique entries then sites should be assigned a unique site code that identifies them in both space and time. The structure of this system depends in part on the spatial and temporal frequency of sampling. A typical nomenclature system uses:

3 letters for a basin/sub-basin code
2 numbers for a site number code
2 numbers for a year code,

e.g. GLD0702, for site 7 on the Gold River sampled in 2002. If there is a seasonal element, two extra numbers can be used with the year code.

Map scale attributes

A number of site attributes should be determined prior to going to the field these are all simply acquired, and are useful in characterising sites and assisting in site selection.

Location – latitude and longitude

Both latitude and longitude are readily measured and should be entered onto the field sheet, these will automatically be converted to decimal degrees when entered into the database.

Ecoregion

Rationale: Ecological land classification is a process of delineating and classifying ecologically distinctive areas. Each area can be viewed as a discrete system which has resulted from the mesh and interplay of the geologic, landform, soil, vegetative, climatic, wildlife, water and human factors which may be present. The dominance of any one or more of these factors varies with the given ecological land unit. This holistic approach to land classification can be

applied incrementally on a scale-related basis from site-specific ecosystems to very broad ecosystems. This is a hierarchical classification (i.e. ecozones, ecoprovinces, ecoregions, ecodistricts, etc.) where areas gain their identity through spatial differences in a combination of landscape characteristics. The factors that are important vary from one place to another at all scales. As an objective in site selection is to initially match reference and test sites, identification of ecoregions and ecodistrict is a useful attribute.

The ecoregion information can be obtained from the following website, and can be entered into the field sheet.

http://www.ec.gc.ca/soer-ree/English/Framework/Nardesc/default.cfm

Information at the ecodistrict level is available from:

http://sis.agr.gc.ca/cansis/nsdb/ecostrat/printed_maps.html

Altitude

Rationale: Altitude is important primarily as it locally effects the temperature regime at a site, while there is no direct correlation and other local climate factors are important, altitude is readily measured from topographic maps and using GPS.

Stream order

Rationale: An overarching development in riverine ecology was the *river continuum concept* (Vannote et al 1980). While this is not so widely regarded now it does contain important elements in how stream function along their length and the patterns of functional groups of organisms and their relationship to the different sources of energy from external through leaf litter (allocthonous) in headwater areas to internal, through algal and plant production (autochthonous) in lower reaches. Stream order is a simple quantitative method of assigning a number to stream segments, which indicates the relative importance of the segment within the drainage basin.

Method

Stream order is determined prior to arriving at a sampling location as follows:

a) The entire watershed for the river in question is mapped out, using 1:250,00 scale topographical map sheets.

b) First order streams are defined as those segments having no tributaries. First order streams are assigned a value of 1.

c) Subsequent stream orders are assigned values according to the method established by Strahler (1957, p914), Figure 1.

d) where stream segments of the same order come together, the resulting segment is assigned the next highest order. E.g.: where first order streams join, the resulting stream segment is elevated to the second order (2).

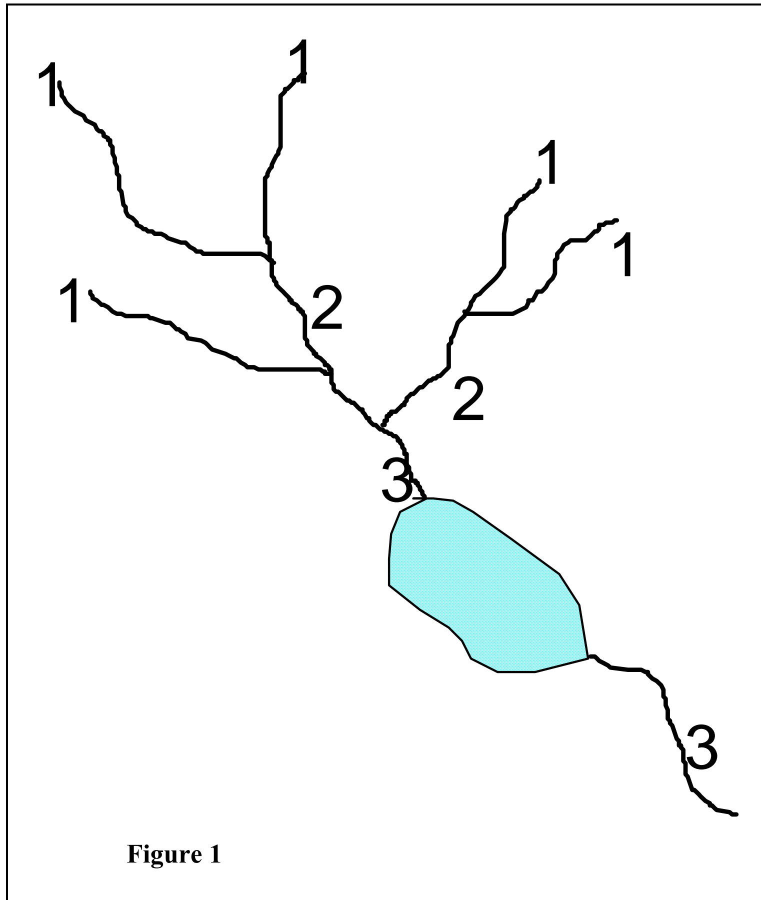


Figure 1

e) where segments of differing orders come together, the resulting segment retains the order of the highest contributing tributary.

f) where a stream enters a lake, the lake is treated as part of the stream. If more than one stream enters a lake, the outflow of the lake retains the order of the highest contributing tributary segment (Figure 1). If two streams of the same order run into a lake, the outflow increases to the next stream order.

Distance from Source

Rationale: Distance from source is also a measure of the position of the site along the stream and is an alternative indicator to stream order that conveys a different type of information related to the *River Continuum Concept.*

Method: the distance is easily calculated directly from topographic maps.

Catchment area

Rationale: Again the size of the catchment area and particularly land uses in the catchment area are important in determining the quality of the stream. Catchment area is readily calculated using planimetry on a topographic maps. Land use is not as readily obtained, but if available for a local drainage should be acquired as it will assist in determining the suitability of reference sites.

Method: Standard planimetric approaches with topographic maps or from GIS systems if available.

Before Departure

Ensure all equipment is present, use an equipment checklist to be sure (Table 3).

Table 3. Field equipment checklist for invertebrate biomonitoring.

Item	Number
Invertebrate sampling	
Kick net	1
Tweezers	2
Squeeze bottle	1
Sample jars	3 per site
Sample labels	1 per jar
Formalin (10% buffered)	2L
Ethanol	2L
Habitat measurement	
Tape measure (30m) stream width	1
Metre stick (stream depth)	1
Metal ruler	1
Orange tennis balls (flow velocity)	2
Stopwatch	1
Water chemistry bottles (2 per site)	2 per site
Miscellanous	
Clipboard	1
Binder (1”)	1
Fine markers	2
Fat markers	2
HB pencils	2
Field sheets	1 box
Topographic maps	As required
Manual (protocols, data analysis, key)	1
Tote box	1

Field Procedures

On arrival in the field there are two types of measurements/samples to be taken. A sample of the benthic invertebrate community, and measurements and samples describing the habitat. It is critical that all measurements be completed and before leaving the site the field sheet should be checked to ensure that all measurements and samples have been taken. One of the interpretation methods being used employs multivariate statistics, these methods are very sensitive to missing data and the only options available when data are missing are to either discard the site or discard the missing variable from all sites. In either case this is very wasteful. Multivariate methods are the most powerful that we can use in assessing the health of sites and therefore all care should be taken to measure all the variables.

In the description of methods below we have generally separated the habitat assessment from the invertebrate sample collection, it is desirable in the field to follow approximately the following order.

A. **General information** – this can largely be done before arriving at the site. Photographs should be taken on arrival.
B. **Reach characteristics** - can be assessed before entering the water.
C. **Water quality** - should be done first as disturbance of the sediment can affect some variables, particularly nutrients.
D. **Benthic invertebrates** - should be sampled before other stream measurements are taken (e.g. channel profile) which will disturb the stream bottom.
E. Substrate measures
F. Channel measures

Quality assurance

It is recommended that at approximately every tenth site the entire procedure be conducted in triplicate to provide an estimate of the replicability of the methods.

Habitat assessment

A. General description

Initial procedures

Upon arrival at the site, variables including latitude, longitude, and altitude are obtained (or verified) using a GPS or other similar equipment (if available), this may have already been completed from a topographic map. On site measurement may also be necessary if a site has to be moved in the field. The station number, stream name, and date should also be recorded on the field sheet, if this has not already been done. The actual sampling reach should now be determined, the sampling reach consists of a stream distance of six times the bankfull width, this typically represents a complete pool riffle sequence contained within the meander

wavelength. If desired marking tape sticks or some other markers can be used to define the reach (site). The site should be typical of the general area.

Site photographs

Due to the expense of field collection, it is often impossible to go back to a site. Photographs of the location can often help to solve any problems that may arise during laboratory analysis. First, a photograph of the field sheet, with the site number, is taken to identify the ensuing series of photographs. These include: an upstream, downstream and an across the stream photograph. In addition a photograph of the substrate in the area where the invertebrate sample will be collected should be taken at a still water location and on the bank in a dry area similar to the wetted sampling area: **Note**; *a metal ruler should be positioned in the photograph for scale*. These photographs provide a valuable record of conditions at the site. If a digital camera is available it should be used, as images can be stored in the electronic database.

B. Reach characteristics

Flow state

Rationale: This is a basic description of the type of habitat present at the site. This describes the stream and provides an indication of what type of organisms may be expected.

Method: This is a simple categorical description of whether the area sampled is predominantly riffle/rapids, a straight run, or a pool/back eddy. Simply circle the appropriate description for the sampling site.

Macrophyte coverage

Rationale: Many organisms are adapted to specifically living among emergent plants, which tend to be associated with slower flow conditions and higher nutrient levels. They also determine the form of material available to herbivores in the invertebrate community.

Method: This is an approximation of the amount of the stream bed that is covered by macrophyte vegetation. Simply circle the approximate degree of the wetted channel covered by aquatic plants.

Canopy coverage

Rationale: The degree of canopy coverage is important for two reasons. First, canopy cover provides shade and thus prevents the stream from overheating in the summer, and thus can affect the degree of both temperature and oxygen stress. Second, extensive canopy can determine the relative amount of external versus internal plant material in the stream and the types of organisms present.

Method: Again this is a simple approximation of the percentage of the stream covered by the tree canopy. Stand in the middle of the stream and estimate the percent shading provided

by overhanging vegetation. It is advisable to get all the individuals present to estimate the coverage and take the average value.

Riparian vegetation

Rationale: The forested land along rivers, streams, and lakes is known as the "riparian zone". Riparian zones are areas of transition between aquatic and upland ecosystems. Riparian (streamside) vegetation bordering the stream protects the water from disturbance and acts as a buffer between the stream and general activities in the watershed and protects the banks from erosion. The width of the riparian zone from the bank is coarsely correlated with the distance from the bank to the base of the tallest tree that could reach the channel

Method: This describes the vegetation found in the riparian zone along the stream sampling site, check if the vegetation type is present. There are four distinct categories:

- Grasses Score - 1
- Shrubs Score - 2
- Coniferous Trees Score - 3
- Deciduous Trees Score - 4

Riparian zone scoring: The presence of any of the four riparian vegetation categories results in a related score (above). Total score for a site is ranked with higher values representing greater stability and allocthonous food input, on a subjective basis:

grass → deciduous increasing bank stability
grass → deciduous - higher allocthonous input
deciduous > coniferous as higher quality food and higher stream temperatures

C. Water Quality

The dissolved materials in the water are determined by erosional processes in the catchment and the underlying bedrock and surficial geology. Measurement of several key variables that either directly or indirectly affect the invertebrates can provide a great amount of information about the types of pollutants and their impact on a stream. Specific activities produce specific pollutants, by identifying specific pollutants we can identify some activities that may be having an impact on the stream. For example nutrients are likely to come from animal feedlots, grazing land, and runoff from fertilized cropland, but could also come from changes in land use such as forestry harvesting or construction sites.

Field measurements

This includes temperature, pH, dissolved oxygen and conductivity. Water samples are also taken for total phosphorus, nutrients and alkalinity. Water is taken directly from the stream by submerging the sampling container to the middle of the water column. Water parameters (DO, pH, conductivity, temperature) are obtained using a Hydrolab or similar device, ***if such a device is available.***

Temperature

Rationale: Temperature is a key physical variable that directly affects many of the physical biological and chemical factors influencing aquatic organisms. If temperatures are outside the range of tolerance for organisms for extended periods of time they can become stressed and die, resulting in a change in the types of organisms inhabiting the stream. Temperature can be modified by among others, weather, removal of riparian vegetation, turbidity and dams.

Method: Sample away from the bank in the main current, place the thermometer or sensor at least 10 cm below the surface or halfway to the bottom in shallower streams. Allow the thermometer to stabilize for a minute before recording, if possible read the temperature with the bulb beneath the water. Take temperature readings in two other paces a few metres apart. Record the average temperature and the time of day

pH

Rationale: The relative acidity of water is ranked on a pH (percent hydrogen) of 0 – 14. A pH of 0 is strongly acidic while 14 is strongly basic (alkaline). Pure water has a pH of 7 (neutral). The pH scale is logarithmic thus every change in 1 unit there is a 10 fold change in acidity. A stream with a pH of 6 is 100 times more acidic than one with a pH of 8. Water with pH 6.5 to 8.5 is suitable for the greatest diversity of aquatic organisms. Young fish and aquatic insects are especially sensitive to extreme pH values outside the optimum range. Stream pH is usually determined by the surrounding geological makeup, but acid rain, wastewater discharges and drainage from coniferous forests (acidic) can decrease the pH of a stream.

Method: measurement of pH depends on the instrument used. For all types of metres the unit must be calibrated according to the manufacturer's instructions. A water sample from the main flow should be collected in a clean container and the pH is determined from this sample. Allow the reading to stabilize before recording the value.

Conductivity

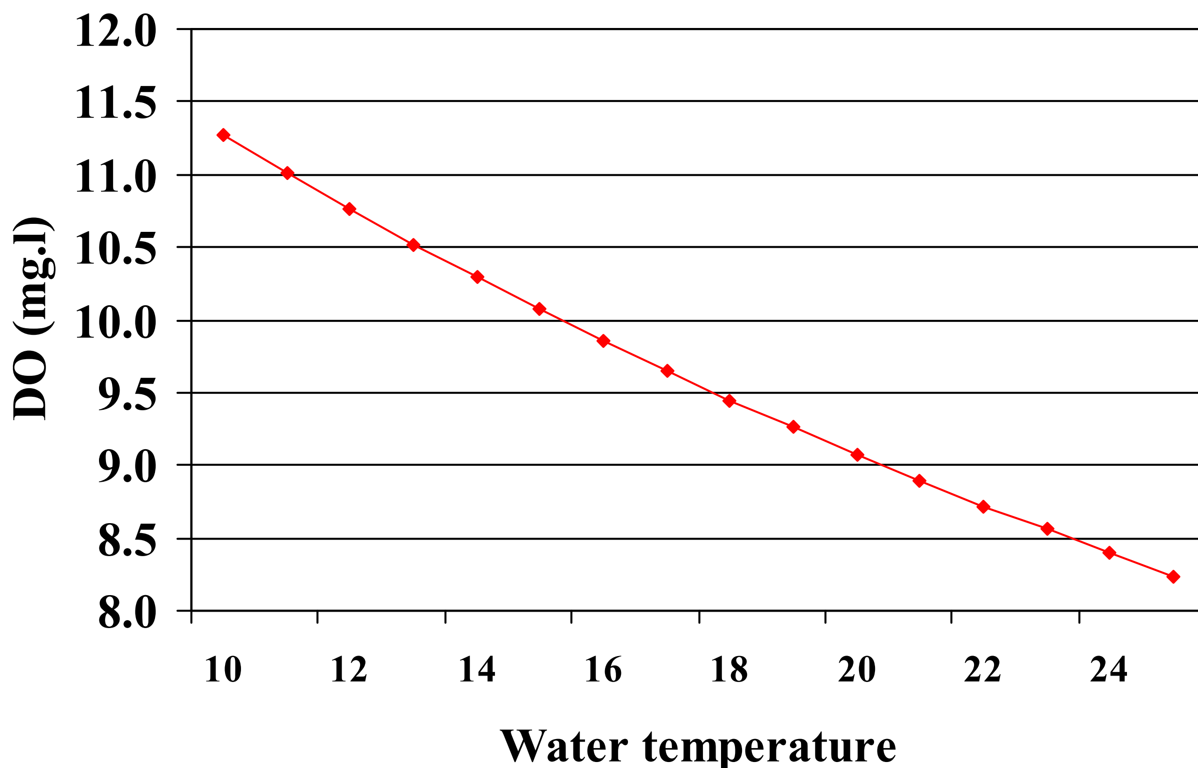
Rationale: conductivity is a measure of the dissolved salts present in the water and is determined by how well they conduct an electrical current. For example pure water has a conductivity of 0. Measuring conductivity is a good way of determining how much dissolved material is present in the water. Conductivity is a useful tracer of point source discharges and sudden increases along a stream can indicate a pollution source.

Method: conductivity is measured using a metre of some type. For all types of metres the unit must be calibrated according to the manufacturer's instructions. A water sample from the main flow should be collected in a clean container and the conductivity determined from this sample. Allow the reading to stabilize before recording the value.

Dissolved oxygen

Rationale: The amount of oxygen dissolved in water affects the kind of organisms found there. Water with higher concentrations of oxygen is generally considered to be of higher quality and supporting many types of animals. Although natural occurrences of low oxygen (hypoxia) can occur and many of the animals considered to be pollution tolerant have naturally evolved to live in these types of environments (e.g., bottom waters of productive lakes). Low oxygen levels can create unfavourable conditions for many organisms and can change population structure. Under conditions of extremely low dissolved oxygen organisms that require high oxygen levels (e.g. mayflies, stoneflies) will emigrate or die leaving other organisms that can tolerate low oxygen (chironomids, tubificids). Low dissolved oxygen in streams can be caused by several factors, temperature is a major influence as cold water can contain more oxygen (Fig 2). Streams with low flow in the summer when air temperature is high are subject to reduced oxygen levels. Slow moving water has less natural aeration. Organic wastes such as agricultural runoff and sewage discharges reduce oxygen because of bacterial decomposition of organic matter. Aquatic plants and algae replenish oxygen during daylight but consume oxygen during the dark.

Fig 2. Maximum dissolved oxygen (DO) concentration at specific water temperatures.



Methods: DO may vary significantly along the length of a river. In riffle areas where re-aeration occurs DO will be higher than in slow moving areas and pools. Therefore, it is important that DO readings be taken in a similar habitat at each site. Dissolved oxygen is

typically measured with a metre make sure the metre has been calibrated according to the manufacturers instructions, and allow time for the reading to stabilize. Where possible express results both as concentration (mg.l^{-1}) and percent saturation (Fig 2).

Water samples

Some measurements can only be measured from laboratory analysis. For these measures water samples will be taken in the field with the bottles provided. The samples taken include one for phosphorus, measured as total unfiltered phosphorus, one for nitrogen and one for alkalinity.

Rationale: Phosphorus and nitrogen are essential elements for both plant and animal life, and phosphorus is often the limiting nutrient in freshwater systems. Increased phosphorus input is frequently a result of human activity and results in increased plant growth. Nitrogen although an essential nutrient is not often limiting in freshwater, but it also responds to human activity. Alkalinity is a measure of the ability of a sample to neutralize strong acid. It is reported in equivalents per litre and consists of the sum of titratable carbonate and noncarbonate chemical species in a filtered water sample, it is largely determined by the geology of the streams watershed and is also a useful surrogate for nutrient levels.

Methods:

Well before sampling contact Dr T. Reynoldson at Acadia University to arrange for sample bottles to be shipped to you. Provide a phone number, contact person and mailing address, as the bottles will be shipped from the Environment Canada Laboratory in Moncton, and should be returned there.

Samples should be collected from the middle of the stream or in the vicinity of the benthic invertebrate sample. At each site a sample will be taken for alkalinity (large 500 ml plastic) and nutrients (small 250 ml plastic). The bottles are pre-rinsed but care should be taken not to touch the mouth of the bottles. Bottles should be filled leaving a slight air space at the top, and kept cool until submitted for analysis. The sample for nutrient analysis (small 250 ml) will have 20 drops of sulphuric acid added to it as a preservative. Each container must be marked with proper, legible labels, with the appropriate Site Code using a water- and solvent-proof marker. Samples should be kept cool once taken. Care should be taken not to spill the preservative. Samples should be shipped to the Environment Canada analytical laboratory in Moncton N.B.

D. Benthic invertebrate sampling:

Benthic macroinvertebrates are common inhabitants of lakes and streams where they are important in moving energy through food webs. The term "benthic" means "bottom-living", so these organisms usually inhabit bottom substrates for at least part of their life cycle; the prefix "macro" indicates that these organisms are retained by mesh sizes of ~200-500 mm (Rosenberg and Resh 1993).

The most diverse group of freshwater benthic macroinvertebrates is the aquatic insects, which account for ~70% of known species of major groups of aquatic macroinvertebrates in

North America. More than 4000 species of aquatic insects and water mites have been reported from Canada. Thus, as a highly diverse group, benthic macroinvertebrates are excellent candidates for studies of changes in biodiversity.

The use of benthic macroinvertebrates in assessment studies is also supported by the extensive background knowledge available for these organisms. This covers everything from study design to data analysis (Table 3). Other general sources of information include Rosenberg (1978), Elliott and Tullett (1978, 1983, 1993), and Murkin et al. (1994).

A number of technical developments enable the effective use of benthic macroinvertebrates in biodiversity studies (Rosenberg and Resh 1993):

1.qualitative sampling and sample analyses is possible using simple, inexpensive equipment;

2.the taxonomy of many groups is well known and identification keys are available; and

3.many well-developed methods of data analysis are available.

Methods: Kick-net sampling - The kick net (Fig. 4) is a triangular (or D-shaped) metal frame holding a mesh bag of 400-mm size (Fig. 4). One end of the metal frame is attached to a rake handle. The part of the bag that attaches to the frame is made of canvas or ripstop-plastic tarpaulin to withstand abrasion. A detachable cup can be added to the end of the bag to facilitate removal of the sample.. A 400-mm mesh net is recommended for general sampling.

The kick net is placed downstream of the collector, flat side of the triangle resting on the substrate of the stream. The collector walks backward, away from the net, kicking the substrate to disturb it to a depth of ~5 cm. For large boulders, the net is held downstream while the boulder is brushed by hand. The net is held near to the area being disturbed so the current will carry dislodged animals into it.

The collector zigzags over the stream bottom from bank to bank in an upstream direction for a timed period (e.g. 2-5 min). Standard time collections (e.g. 3 min) allow comparisons among sites. The zigzag coverage allows collection of invertebrates from a variety of stream habitats (pools, riffles, runs, etc.). It is important that sampling be extended directly adjacent to the stream bank because this region may have aquatic macrophytes that support a unique fauna.

When sampling is completed, the cup is removed and its contents are emptied into a plastic jar. Material remaining in the cup can

be washed into the jar by spraying the outside of the mesh (bottom of the cup) with water from a squeeze bottle. The net and cup should be checked for remaining invertebrates. A label (non-recycled photocopier paper or waterproof paper marked by soft pencil or alcohol-proof pen) accurately describing the sampling location (stream name, site number), date, replicate, and collector, is added to the inside of the jar. The outside and lid of the container should be similarly labelled using a waterproof felt pen.

The sample is preserved by 10% formalin (1:3 ratio formalin :sample) to the sample container. Samples should be transferred into 70% ethanol back in the laboratory after 72h. This process kills specimens quickly in the field with a minimum of preservative, provides tissue fixation without dissolving calcareous deposits in the exoskeletons of some taxa, preserves colour, replaces most of the water in organisms with alcohol, and makes sorting more comfortable by reducing the amount of formalin in the sample.

Summary of procedure:

- the area to kick should be defined and kicking should start at the downstream end.
- stepping into the stream the net should be placed firmly on the substrate surface and as close to the stream border as is possible.
- kicking begins by moving your feet through the substrate, rolling over rocks and stones, large rocks or those deeply embedded should be rubbed using your hand.

Note: the net should always be in contact with the substrate and should always be directly downstream of the operator.

- the area to be sampled should be traversed in a zigzag pattern going from bank to bank and always heading upstream. In large rivers or very fast flowing streams, bank to bank sampling may not be possible. In these cases the same procedure is followed, zigzagging through the area that has been defined for sampling.
- the sample is checked, with large rocks and sticks being removed after they have been washed and carefully checked for organisms.
- the sample is then washed into a container and the net checked carefully for organisms clinging to it.

Note: Seams and folds should be checked carefully for hidden organisms.

- 10% formalin is then added to the sample (approximately in a 1:3 ratio).
- containers are labelled outside and a label is also placed inside each sample.
- records of the person doing the kicking and the typical depth are noted on the field sheets.

E. Substrate characteristics

The composition of the stream bed material is important in identifying hydrological characteristics of the river and the type of habitat available to aquatic organisms. Insects need to attach themselves to the stream bottom or live within the bed materials. The more attachment or living spaces available the greater will be the variety and number of organisms found. Optimally

the stream bed will be dominated by cobbles, gravel and boulders. As the percentage of sand and silt increases the suitability and availability of living space for invertebrates decreases.

Dominant Substrate and Embededdness

Rationale: In streams, Nielsen et al. (1983) describe the use of a substrate score that is determined by the sizes of the two predominant substrates, the size of the material surrounding the predominant substrates, and the degree of embeddedness. In undisturbed streams, fine sediments (< 2mm) do not accumulate in large quantities on gravel and cobble in riffles. In areas modified by agriculture or other stream side activity increased erosion results in accumulation of fine material. Embedded riffle substrates provide less desirable habitat for invertebrates and reduce productivity.

Method:

Dominant substrate - Examine the area where the benthic sample was taken and refer to Table 1 in the field sheet to determine the two dominant substrate types and the material surrounding those. It is recommended that more than one individual does this independently and the median score is recorded.

Embeddedness - Wade into the middle of the area you sampled for invertebrates and pick up at least five pieces of large pebble or cobble, avoiding rocks disturbed by sampling. Estimate the percentage of rock buried in the fine material. A stain line on the rock may indicate the level of burial and aid in the estimation. Record the average degree of embeddedness using Table 1 on the field sheet.

Substrate dimensions

Rationale: a final estimate of the size of the substrate can de obtained by simply measuring a randomly selected sample of the substrate.

Method: walk through the area from which the invertebrate sample was taken, stopping at random and selecting a rock. To avoid bias select a rock closest to either the left or right toe. The maximum length width and height are recorded for 10 rocks and the average, maximum and minimum values determined.

F. Channel characteristics

Characteristics of the channel in a stream reach often determine the abundance and distribution of benthic macroinvertebrates, so it is important to describe these attributes. The dimensions and shape of the channel and the substrate paving the bottom of the channel - a factor of critical importance to the benthos - are a result of the geology of the area and peak flows. Peak flows occur, on average, in two out of three years, are called "bankfull discharge", and are related to spates caused by snow melt or summer rain storms. Erodible materials are carried through the stream reach, shape the dimensions of the channel (width, depth), and leave behind

substrate material that the stream does not have enough energy to transport. The substrate material is crucial to the development of benthic macroinvertebrate communities. It is possible to relate faunal distributions with a particular suite of hydrological variables by measuring channel characteristics (e.g. Cobb et al. 1992).

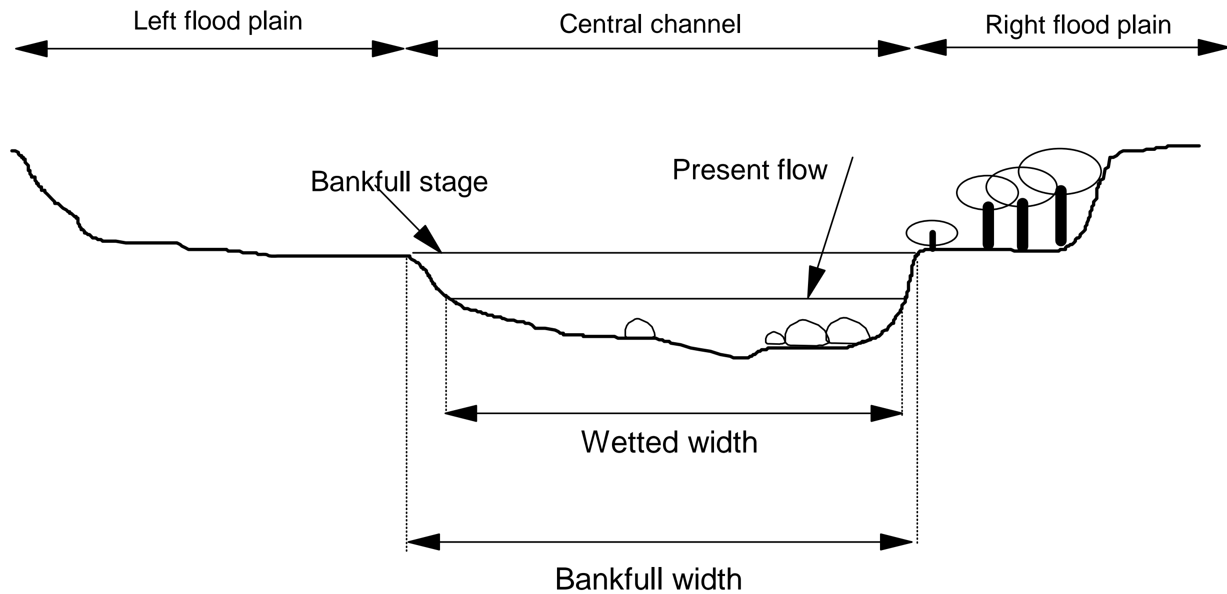


Figure 3. Various channel attributes (adapted form Newbury and Gaboury 1993).

Stream width

Rationale: As flow decreases, water will cover less of the stream bottom, which will limit the available habitat for aquatic organisms. Two measures of stream width are made, bankfull and wetted.

Method: Judgement is required in obtaining a representative cross section of the channel, the whole site should be examined before a section is measured. For example, areas with log jams, braided channels or beaver dams etc should be avoided. A transect is established at right angles to the flow; a tape measure is stretched across the stream and secured to each bank; and width and depth of the present (wetted) and bankfull flows are measured (Fig. 3). Bankfull flows are usually short in duration and seldom observed; however, they can be determined by locating points of vegetation change on the stream banks, where algae or marl have been scoured from boulders, or where sediment texture abruptly changes. Detailed determination of bankfull dimensions is described in Newbury and Gaboury (1993) and Harrelson et al. (1994).

Stream cross-section, flow and discharge:

Rationale: The shape of the channel and thus the habitat available for invertebrates is a consequence of the local geology and the total discharge of the stream. For many small streams there are no discharge records available and in these cases current discharge can be estimated form the cross sectional area of the stream and the flow of water through the stream. Ideally

flow can be calculated at several points using a flow meter if it is available. An alternative is to measure velocity using tennis balls travelling over a measured distance.

Stream cross section

Method: This requires measuring across the stream or section sampled using a standard metric tape measure at the site where bankfull and wetted width were determined as being representative. The cross sectional profile is determined by dividing the stream into segments. For a detailed estimate of discharge divide the wetted width by 10 and measure the depth at each segment. For example if the wetted width is 4.8 metres then the first point would be 48 cm from the bank, the second 96 cm and so on until the final segment at 4.32 m. If a flow measurement device (e.g., Pygmy meter) is available the stream flow should also be measured at each point following the procedures outlined in the operation of these meters.

A simpler alternative if a flow meter is unavailable is to estimate velocity using the tennis balls. In this case three cross sections will be done:

1. at the top of the sampling reach;
2. at the centre of the sampling reach;
3. at the bottom of the sampling reach.

In all cases the area chosen for taking the cross section should be unobstructed by logs and debris and braided channels (with islands) should be avoided.

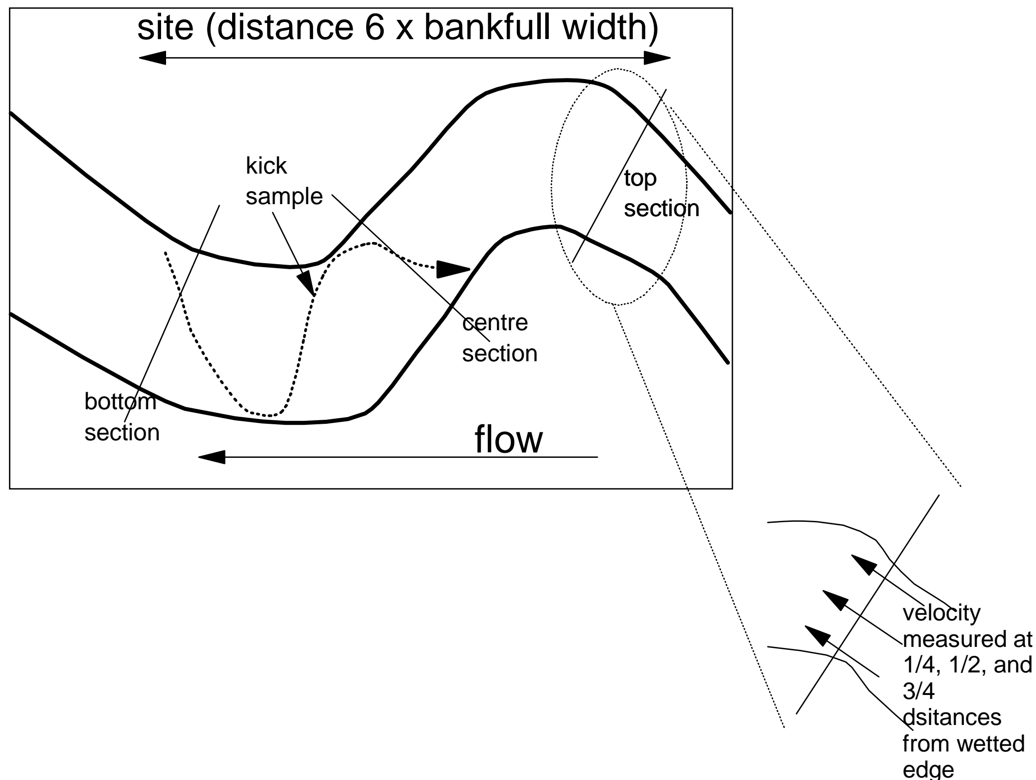


Figure 4. Cross section profiles.

At each cross section the depth and velocity should be measured at 3-5 (depending on width) locations across the stream. Velocity should be measured over a fixed distance, centred on the cross section. Distances should typically be in the 3-5m range, and the time of travel estimated (Fig 4). In smaller streams (5 m or less width) this should be done at ¼, ½ and ¾ of the way across the stream. In larger streams, (> 5m) we would recommend five equidistant measurement points across the stream.

Velocity

Rationale: In addition to determining discharge at selected points it is important to obtain a measurement of the velocity from the area from which the invertebrate sample was taken, as this represents the conditions to which the invertebrates are exposed.

Method: The area travelled during the kick sample should have been noted. The actual distance of the thalweg (the deepest channel, where the main flow resides) should be measured. The coloured tennis balls are released in the centre of the channel and the time required to travel a fixed distance is recorded. A stream reach is selected that covers approximately the distance sampled during the kick sample. Each end of the reach is marked and an individual releases the ball at the top end recording the time required to travel the complete distance. This is repeated 5 times and the average velocity calculated.

This completes the field measurements. As a final check verify that all the field measurements have been taken and are legibly recorded, ensure that all equipment is taken with you from the site.

F. Quality Assurance and Control

Quality Assurance and Control (QA/QC) is an ongoing process which has the goal of prevention, early detection and correction of field and analytical data collection errors (U.S. EPA 1995). Participants must ensure all data sheets are filled in correctly and completely. They must determine if the data are reasonable before they leave the field, and if not, the measurements should be repeated before leaving. It is also recommended that at every tenth site the entire process should be repeated in triplicate to provide some assurance of the variability associated with the measurements. QA/QC procedures for the laboratory are identified in that section of the manual.

Useful texts

Culp, J.M. and D.B. Halliwell. 1999. Volunteer-based monitoring program. Using benthic indicators to assess stream health. Instructors manual. Environment Canada, NWRI, Saskatoon.

Merritt, R.W. and K.W. Cummins (eds.) 1996. An introduction to the aquatic insects of North America. 3rd ed. Kendall-Hunt, Dubuque, Iowa. 862 p.

Plafkin, J.L., M.T. Barbour, K.D. Porter, S.K. Gross, and R.M. Hughes. 1989. Rapid bioassessment protocols for use in streams and rivers: benthic macroinvertebrates and fish. EPA/444/4-89-001. Assessment and Watershed Protection Division, Washington DC.

U.S. EPA 1995. Volunteer stream monitoring: a methods manual. EPA 841 D95-001 April. Office of Wetlands, Oceans and watersheds, 4503F Washington DC 20460.

References

Bisson, P.A. and D.R. Montgomery. 1996. Valley segments, stream reaches, and channel units. Pp. 23-52. *In* Methods in Stream Ecology. Edited by F.R. Hauer & G.A. Lamberti. Academic Press, San Diego 674p.

Cobb, D.G., T.D. Galloway, and J.F. Flannagan. 1992. Effects of discharge and substrate stability on density and species composition of stream insects. Canadian Journal of Fisheries and Aquatic Sciences 49: 1788-1795.

Conquest, L.L., S.C. Ralph, and R.J. Naiman. 1994. Implementation of large-scale stream monitoring efforts: sampling design and data analysis issues. Pp69-90. In Biological Monitoring of Aquatic Systems (Loeb S.L. and A. Spacie eds) Lewis Publ. Boca Raton FL. 381p.

Corkum, L.D. 1989. Discussion Paper for the Environmental Effects Monitoring Benthic Community Expert Working Group

Corkum, L.D. 1989. Patterns of benthic invertebrate assemblages in rivers of northwestern North America. Freshwater Biol. **21**: 195-205.

Corkum, L.D. 1992. Spatial distribution patterns of macroinvertebrates along rivers within and among biomes. Hydrobiologia. 239:101-114.

Elliott, J.M. 1977. Some methods for the statistical analysis of samples of benthic invertebrates. 2nd ed. Freshwater Biological Association Scientific Publication No. 25. 156 p.

Elliott, J.M. and P.A. Tullett. 1978. A bibliography of samplers for benthic invertebrates. Freshwater Biological Association Occasional Publication No. 4. 61 p.

Elliott, J.M. and P.A. Tullett. 1983. A supplement to a bibliography of samplers for benthic invertebrates. Freshwater Biological Association Occasional Publication No. 20. 27 p.

Elliott, J.M. and P.A. Tullett. 1993. A new bibliography of samplers for freshwater benthic invertebrates. Freshwater Biological Association Occasional Publication No. 30. 92 p. Environment Canada. 1997.

Fletcher, R., T.B. Reynoldson, and W.D. Taylor. 2001. The use of benthic mesocosms for the assessment of sediment contamination. Envir Poll. 115:173-182.

Gordon, N.D., T.A. McMahon, and B.L. Finlayson. 1992. *Stream hydrology: an introduction for ecologists*. John Wiley & Son, Inc. West Sussex, UK.

Green, R.H. 1979. *Sampling design and statistical methods for environmental problems*. Wiley, New York.

Harrelson, C.C., C.L. Rawlins, and J.P.Potyondy. 1994. Stream channel reference sites: an illustrated guide to field technique. General Technical Report RM-245. U.S. Department of Agriculture, Forestry Service, Ft. Collins, Colorado. 61 p.

Hawkins, C.P., R.H. Norris, J.N. Hogue, and J.W. Feminella. 2000. Development and evaluation of predictive models for measuring biological integrity of streams. Ecological Applications 10:1456-1477.

Hilsenhoff, W.L. 1988. Rapid field bioassessment of organic pollution with a family level biotic index. J. N. Am. Benthol. Soc. 7:65-68.

Hughes, R. M. 1995. Defining acceptable biological status by comparing with reference conditions. Pages 31-47 in W. S. Davis and T. P. Simon (editors). Biological assessment and

criteria. Tools for water resource planning and decision making. Lewis Publishers, Boca Raton, Florida.

Meador, M.R., C.R. Hupp, T.F. Cuffney, and M.E. Gurtz. 1993. Methods for characterizing stream habitat as part of the national water-quality assessment program. US. Geological Survey Open-File Report. Raleigh, North Carolina. USGS/OFR 93-408.

Murkin, H.R., D.A. Wrubleski, and F.A. Reid. 1994. Sampling invertebrates in aquatic and terrestrial habitats, p. 349-369. In: T.A. Bookhout (ed.) Research and management techniques for wildlife and habitats. The Wildlife Society, Bethesda, Maryland.

Newbury, R.W. and M.N. Gaboury. 1993. Stream analysis and fish habitat design. A field manual. Newbury Hydraulics, Gibsons, British Columbia. 256p.

Nielsen, L.A., D.L. Johnson, and S.S. Lampton. 1983. Fisheries techniques. American Fisheries Society, Bethesda, Maryland. 468 p.

OEPA (Ohio Environmental Protection Agency). 1990. Ohio Water Resource Inventory. Ohio Environmental Protection Agency, Columbus.

Omernik, J. M. 1995. Ecoregions: a spatial framework for environmental management. Pages 49-62 in W. S. Davis and T. P. Simon (editors). Biological assessment and criteria. Tools for water resource planning and decision making. Lewis Publishers, Boca Raton, Florida.

Resh, V.H. 1995. Freshwater benthic macroinvertebrates and rapid assessment for water quality monitoring in developing and newly industrialized countries. Pp167-177 in W. S. Davis and T. P. Simon (editors). Biological assessment and criteria. Tools for water resource planning and decision making. Lewis Publishers, Boca Raton, Florida.

Reynoldson, T.B., Rosenberg, D.R , Day, K.E., Norris, R.H., Resh, V.H.. 1997. The reference condition: a comparison of multimetric and multivariate approaches to assess water-quality impairement using benthic macroinvertebrates. J. N. Am. Benth. Soc. 16:833-852.

Reynoldson, T. B., M. Bombardier, D. Donald, H. J. O'Neill, D. M. Rosenberg, H. Shear, T. Tuominen and H. Vaughan. 1999. Strategy for a Canadian Aquatic Biomonitoring Network. NWRI Contribution No. 99-248. 26pp.

Rosenberg, D.M. 1978. Practical sampling of freshwater macrozoobenthos: a bibliography of useful texts, reviews, and recent papers. Canadian Fisheries and Marine Service Technical Report No. 790. 15 p..

Rosenberg, D.M. and V.H. Resh. 1993. Introduction to freshwater biomonitoring and benthic macroinvertebrates, p. 1-9. In: D.M. Rosenberg and V.H. Resh (eds.) Freshwater biomonitoring and benthic macroinvertebrates. Chapman and Hall, New York.

Simon, T.P. 1991. Development of biological integrity expectations for the ecoregions of Indiana. I. Central Corn Belt Plain. US. Environmental Protection Agency, Region V. Environmental Sciences Division, Monitoring and Quality Assurance Brabch, Chicago IL. EPA 905/9-91/025.

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

Strahler, A.N. 1957. Quantitative analysis of watershed geomorphology. Am. Geophys. Union Trans. 38:913-920.

Underwood, A.J. 1991. Beyond BACI: experimental designs for detecting human environmental impacts on temporal variations in natural populations. Australian J. Marine Freshwater Res. 42:569-587.

U.S. EPA 1995. Volunteer stream monitoring: a methods manual. EPA 841 D95-001 April. Office of Wetlands, Oceans and watersheds, 4503F Washington DC 20460.

Vannote, R.L, G.G. Minshall. et al. 1980. The River continuum concept. Can J. Fish Aq. Sci. 37:130-137.

Wright, J.F., Moss, D., Armitage, P.D., and Furse, M.T. 1984. A preliminary classification of running-water sites in Great Britain based on macroinvertebrate species and the prediction of community type using environmental data. Freshwater Biol. **14**: 221-256.Yoder. 1989.

Yoder, C.O. 1989. The integrated biosurvey as a tool for evaluation of aquatic life use attainment and impairment in Ohio surface waters. Pp 110-122 in Biological Criteria research and regulation, proceedings of a symposium (G. Gibson Ed). US Environmental Protection Agency, Office of water Washington DC EPA-440-5-91-005.

Yoder, C.O. 1991. Answering some concerns about biological criteria based on experiences in Ohio. Pp 95-104. In Water Quality Standards for the 21st century, US. Environmental Protection Agency, Office of Water, Washington DC 251p.

Laboratory Processing

Just as the sampling process for any study is critical to the success of the research, so is the management and analysis of those samples once they arrive at the laboratory from the field. A carefully structured framework for managing both the samples themselves, as well as data generated from their analysis is required. Without such a framework, samples can be lost or unanalysed and results can become misplaced or misconstrued. Guidelines have been established for sample and data management. These guidelines are broken down into sections, each identifying the proper procedure(s) at each stage.

Project Initialisation

Upon the initialisation of a project, a hard-copy Project Folder should be started by the project manager. The Project Folder acts as a data warehouse for raw field sheets and any other hard-copy data generated during a project, until its entry into the electronic data base (see Data Management). Individuals responsible for specific results must ensure that their data sheets are placed in the Project Folder once analysis has been completed. Project Folders should be stored in a single location.

Arrival of Samples

Delivery

When project samples collected in the field arrive at the laboratory, they should be inspected and distributed to the appropriate laboratories for analysis. The inspection of samples includes:

i. entering each sample into the Project Folder log.

ii. checking against the project field sheets to ensure samples from all sampling locations are accounted for.

iii. identifying any samples which have been damaged/lost during transport. Damaged/lost samples are noted, and a comment must be placed in the Comments section of the benthic data information system (see Laboratory Protocols for Data Management).

iv. ***Community Structure Samples:*** should be sent to the Benthic Ecology lab for storage and processing.

v. ***Water Chemistry Samples:*** stored at 4C until delivery to the Environment Canada laboratory for processing.

Invertebrate Sample Processing

Invertebrate samples should be checked against the Project Folder log to ensure that no samples have been lost during delivery to the laboratory. Once all of the samples have been verified, they should be prepared for storage until processing can be done.

Sample Storage

Preparation consists of exchanging the formalin (see Field Protocols) in each sample with 70% Ethanol. Stored samples should be checked every three months, to ensure the replacement of any ethanol lost due to evaporation.

Sample Processing

Sub-Sampling

Depending on the density of organisms presumed to be in a sample, the time required to process a single benthic sample can be extremely long. For this reason, a method of sub-sampling has been adopted, based on the Marchant (1989) sub-sampling device and procedure. The sub-sampler is a box (35 x 35 x 10 cm) divided into 100 equal cells.

- Only one person should process an individual sample, this ensures a consistent level of precision and accuracy, even if poor.
- the need to sub-sample a sample is based on expected yield of organisms.
- an unsorted sample is placed into the box, the box is covered and shaken to distribute the sample evenly among the 100 cells.
- We suggest removing a single cell and processing that to determine the approximate number of cells required to be sub-sampled. Then remove the anticipated number of cells and begin sorting and counting them sequentially
- cells are randomly sampled, using die or another random selection process, and then processed until at least 300 organisms are picked (see Sample Processing). This count does not include organisms from the following orders, Porifera, Nematoda, Copepoda, Cladocera, Platyhelminthes, and Ostracoda (see Sorting). Cells which have been started MUST be finished, after the 300 organism level has been achieved.
- the number of organisms in the entire sample is then estimated by extrapolation to the full 100 cells in the sub-sampler.

Sorting and Identification

Sample sorting consists of 5 steps, on a sample by sample basis:

- Samples are picked using a low power stereo microscope (16X with 10X eye piece).and organisms of the same type are placed into separate sample containers.
- Each type is identified, counted, and the counts are recorded on bench sheets (Appendix C)
- During identification, each organism is identified to family level using the taxonomic key and tallies are recorded on the bench sheets. Organisms from the same families

are kept in single vials in ethanol and labelled appropriately. Bench sheets are stored in the Project Folder until data entry takes place

- Sorted samples are stored for submission to Environment Canada (below) for verification and Genus/Species level identification

  Mr Craig Logan (Taxonomist)
  Benthic Ecology Laboratory
  National Water Research Institute
  Environment Canada
  867 Lakeshore Rd
  Burlington ON L7R 4A6

The taxa Porifera, Nematoda, Platyhelminthes, Ostracoda, Copepoda, and Cladocera are not included in the 300 organism sub sample count because they are not generally considered as part of the macroinvertebrate community. The organisms in these taxa are still counted however and their numbers are recorded in the database. Since these organisms can be found in extremely high numbers and the time to remove them from a sample can be substantial only the first fifty are removed and the rest counted but left in the sample. The taxa where this has occurred are then marked on the field sheet.

Tallies are transferred from bench sheets to the electronic database, and the bench sheets are placed in the Project Folder by the SCSL technician.

Quality Assurance/ Quality Control (QA/QC)

The picking and sorting of benthic organisms is a somewhat tedious process. There is a possibility that in some samples, organisms will be missed due to sheer numbers, or other confounding factors such as detritus in the sample. For this reason, quality assurance/quality control procedures (QA/QC) must be carried out on a regular basis, in order to establish a standard sorting efficiency for the sampling process.

Determination of sorting efficiency is established:

- the residue from every picked sample (enumerated cells only) is replaced in a sample container and stored.
- 20% of samples (minimum 3) are randomly selected from the total and re-picked; and the number of new organisms found is counted.
- the % OM (organisms missed) is calculated using the equation:
- $\frac{\#OrganismsMissed}{TotalOrganismsFound} * 100 = \%OM$
- The average %OM is calculated based on the 3 samples re-picked, and represents the standard sorting efficiency for that project.

Determination of taxonomic error

This will be determined by comparing identifications with sample verified by the Environment Canada Taxonomy laboratory. The following verification rate is proposed for those identifying invertebrates:

- First five samples — all are verified
- Next 10 samples — 2 are verified
- Every 30 samples — 2 are verified

Samples for verification should be sent to:

Mr Craig Logan (Taxonomist)
Benthic Ecology Laboratory
National Water Research Institute
Environment Canada
867 Lakeshore Rd
Burlington ON L7R 4A6

Water samples

On return to the laboratory water samples are stored according to Table 3. Once all samples have been collected they should be shipped to the Environment Canada analytical laboratory for analyses.

Environmental Quality Laboratory
Environment Canada
P.O.Box 23005
Moncton, NB E1A 6S8

Table 3. Preservation conditions for selected water chemistry analyses.

Parameter	Container	Preservation	Holding Time
General Water Quality (acidity, alkalinity, color, conductivity, hardness, pH, NO_3, turbidity, anions, cations)	500-mL plastic	refrigerate	7 days
NH_4, Total Phosphorus, Total Nitrogen and TOC	250-mL plastic	20 drops of H_2SO_4 (1:1) to pH<2 and refrigerate	7 days

Samples should be kept cold at 4°C. For additional info on supplies or preservation techniques, please contact the laboratory at (506) 851-2899.

References

Marchant, R. 1989. A subsampler for samples of benthic invertebrates. *Bull. Aust. Soc. Limnol*. Vol. 12. Pp. 49-52.

Data Management

The management of data from field collection and laboratory analysis is the final step in establishing a complete project data base. The success or failure of a project can be determined as much by the management of project data, as by the success or failure at any other stage of a project's duration. Like the collection and analysis of samples, resulting data must be handled within a well-established framework.

During the field sampling and laboratory analysis portion of a project, raw data sheets are warehoused in the Project Folder initiated by the project manager, which represents the first stage in data management. However, the Project Folder is not the final storage media for data collected during a project. Standard data storage methods enable scientists to manipulate and analyse data with much more accuracy and fewer problems. Environment Canada has established a computer-based information storage and retrieval system for use during projects involving Benthic Community Structure and Environmental Attributes in Aquatic Ecosystems. Initially this will be provided on CD ROM , however, we are currently working on a web based geographically referenced data entry system.

Electronic benthic data information system

Project Initialisation

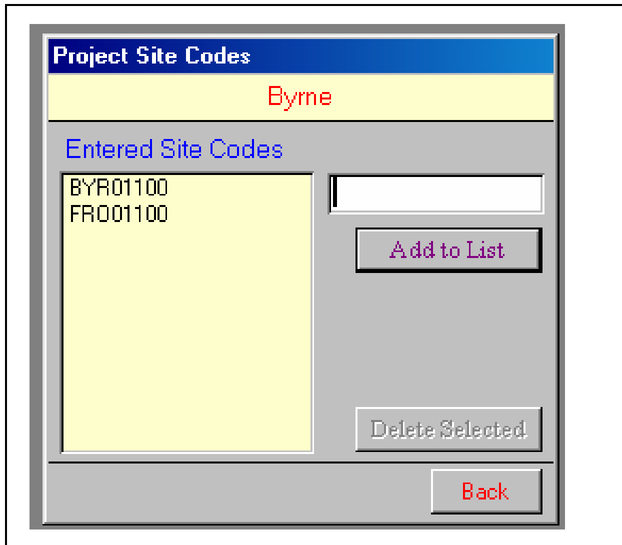
The Project Folder represents the hard-copy equivalent of the final electronic data base. Unlike the Project Folder, however, the Benthic Data Information System (BDIS) is designed to store data from multiple projects. For this reason, the first stage in transferring data from hard-copy to electronic format is establishing a unique Study Name for identification in BDIS. The initialisation of a new study within BDIS can be undertaken at any time during a project. However, the most efficient routine is to establish a new electronic project at the same time that field-work is begun. New projects are initiated by the Project Manager.

Study Name Protocol

Any alphanumeric name which uniquely identifies a project and associated data within the data base.

Site Code

Individual sample sites for a specific project are identified by unique Site Codes (see Field Methods page) within BDIS. These are established before field work is undertaken, and are used to identify samples as they are collected in the field. Each Site Code must be entered into the information system, and this should be the responsibility of the Project Manager. The entry of site codes is done only once during a project. Once these codes have been established, they are automatically maintained by BDIS as new data for each site is entered into the system from the Project Folder. This eliminates data entry errors, and the possibility of data being lost due to conflicting or contradicting site codes. No data entry can take place until the Project Manager has entered Site Codes for a particular project.

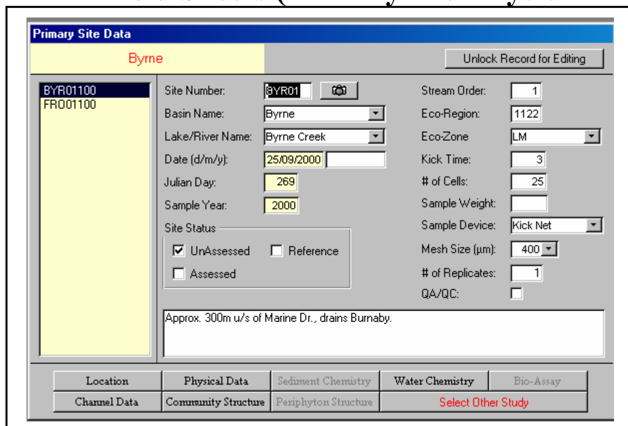


Data Entry

Hard-Copy Study Binder

As data for each stage is entered into the benthic data information system, the raw data sheets are returned to the Project Folder.

Field Sheets (Primary and Physical Data)



Field Sheets for benthic assessment include the site code, date sampled, location, and other physical parameters collected while in the field. Field sheets are placed in the Project Folder immediately after a return from the sampling trip, and are the responsibility of the Project Manager. The entry of data from the Field sheets is also undertaken by the Project Manager, and must be completed before any other data can be entered to BDIS. Once field data has been entered, field sheets are returned to the Project Folder.

Chemistry Data

Chemistry data are entered from the data sheets returned by the analysis labs listed in appropriate sections above. The entry of chemistry data is again the responsibility of the Project

Manager. Once data for a particular analysis (e.g. Water Chemistry) for a project has been entered, data sheets are returned to the Project Folder.

Community Structure

Community structure data are entered from the completed tally sheets compiled by the individuals responsible for sorting and identification. The entry of community structure data is the responsibility of the Project Manager. Data sheets are returned to the project folder once all data has been entered into the data base.

Automated Computations

Several forms in the data entry system automate the computation of values:

Decimal Degrees

Location data in the field is typically recorded using a Global Positioning System (GPS), and is measured in Degrees, Minutes, and Seconds of Latitude and Longitude. For statistical analysis, however, Latitude and Longitude must be converted to Decimal Degrees. BDIS automatically computes Decimal Degrees for the user, as Degrees, Minutes and Seconds are entered.

Mean Counts for Benthic Organisms

During sample collection in the field, up to 5 replicates at a single site can be taken. Statistical analysis typically requires a single value for each site, requiring a mean to be computed. BDIS automatically computes a mean count for each organism during data entry.

Note: If less than 5 replicates are taken, blanks must be used for each unsampled replicate (e.g. only three were taken, 4 and 5 are blank). If a zero is left in these boxes, BDIS assumes a replicate was taken, and bases the mean on the number on replicates filled in, causing an erroneous entry to the data base.

Data Interpretation

Studies of the effects of disturbance often require the determination of the difference in means over time or at different sites. This involves a different approach than estimating specific population or community measures with a given level of precision; the reader is directed to Resh and McElravy (1993), Norris and Georges (1993), and Merritt et al. (1996) for a discussion of the subject and an introduction to the literature. The detection of differences among sites or times has traditionally relied on quantitative approaches using inferential statistics (e.g. Green 1979; Stewart-Oaten et al. 1986; Underwood 1991). However, a new approach called the "reference condition" (Reynoldson et al. 1995, 1997; Wright 1995), which uses qualitative sampling and multivariate statistics, circumvents many of the problems inherent in quantitative, inferential approaches (Reynoldson et al. 1997). While the use of the Reference Condition Approach is the objective of the programme until sufficient data are collected to construct the various predictive models rapid assessment approaches are recommended (RAAs).

The high cost of quantitative approaches has led to the development of qualitative methods called RAAs (e.g. Plafkin et al. 1989). The original purpose of using RAAs was to identify water quality problems and to document long-term regional changes in water quality (Resh and Jackson 1993), but these methods can also be applied to measuring changes in biodiversity. Qualitative techniques have been used in Europe for decades (e.g. biotic indices and scoring systems; see Metcalfe 1989). Their chief advantage is the reduction of the intensity of study required at individual sites (relative to what is required by quantitative approaches: see above), which permits a greater number of sites to be examined (Resh and Jackson 1993). Key to the use of RAAs are the following considerations: (1) What population and/or community measures are relevant? (2) What are the baselines against which these measures are being compared? (3) How much deviation from a baseline indicates change? Resh and Jackson (1993) provide a comprehensive review of RAAs.

Simple Analyses – plots trends and coarse indicators

Some simple descriptors (metrics) of invertebrate community can be examined at the order level of taxonomy, however these metrics are as rigorous as family level identifications. These simple metrics involve counting the number of different taxonomic groups (taxonomic richness), number of insect groups, total number of invertebrates or the total number of EPTs (Ephemeroptera, Plecoptera, Trichoptera).

It is suggested that the following indices be considered to analyse the data and simple graphical results be presented:

- Taxonomic richness
- EPT richness
- Shannon-Wiener diversity
- Equitability
- Dominance
- Hilsenhoff family biotic index
- Bray-Curtis similarity

These indices are not intended to replace proper statistical analysis of data. These indices are not absolute measures of a community and are subject to sampling error and more importantly natural community variation. This latter concern will gradually diminish as more reference sites are accumulated and reference sites are shared amongst the participants in the network allowing finer and finer resolution and increased certainty in eliminating natural variation as a source of error. Initially replicate reference sites are required to establish natural variability before conclusions can be drawn from such indices. However, the indices are useful in condensing data down to single numbers and must be interpreted in combination withy other indices as well as other data collected at the sites.

Taxonomic richness

This measure reflects the health of the community by measuring the variety/ diversity of the taxa present. Richness generally increases with increasing water quality, habitat quality and availability or suitability. In some situations moderate organic enrichment may result in an increase in the number of taxa (including EPTs) form a less impacted reference site.

Taxa richness is determined by simply counting the total number of families present.

Total EPT

This measures the total number of Mayflies (Ephemeroptera), Stoneflies (Plecoptera), and Caddisflies (Trichoptera) which are typically most sensitive to habitat disturbance. Look for even distribution and high numbers of EPTs as indicators of good water quality. Special attention should be paid to the absence of any one of the three EPT groups at a site. As with taxonomic richness moderate organic enrichment may result in an increase in the number of EPTs present in samples.

Total EPTs is simply the total count of these organisms in the sample.

Shannon-Wiener diversity

Diversity measures tend to be more informative than simple abundance totals. There are a variety of diversity indices and most exploit aspects of the number of taxa and their relative abundance (see equitability and dominance below). While a difficulty of these indices is the absence of absolute values of what represents an effect they do provide a summary of the distribution of the taxa. Different diversity indices emphasize the species richness and equitability components of diversity to varying degrees. The most commonly used is the Shannon-Wiener diversity index:

$$H' = -\sum_i p_i \log(p_i)$$

Where:

p_i = proportion of the total count arising from the *i*th species. Caution is advised in using and comparing these indices as they are very susceptible to sampling effort. Also we recommend the use of the natural logarithm (to the base e).

Equitability

This expresses how evenly the individuals are distributed among the different species, and is often termed ***evenness***. Sites that are more even are considered to be more diverse and an effect of stress is often to reduce numbers of intermediate taxa and increase numbers of one or two taxa.

Equitability is expressed by Pielou's evenness index:

$$J' = H' / H'_{max}$$

Where:

H' = site Shannon-Wiener diversity index

H'_{max} = Shannon-Wiener index where all taxa are set at the average abundance per taxa

Dominance curves

Dominance is the inverse of equitability, and shows the pattern of dominance across the community without reducing to a single summary statistic such as a diversity index. Such techniques are intermediate between univariate summaries and multivariate analyses as they capture and present information on all the taxa present.

Ranked taxa dominance curves are based on ranking taxa in decreasing order of their abundance, expressed as a percentage of the total abundance of all species, and plotted against the relevant species rank. Curves become steeper as dominance increases.

Hilsenhoff family biotic index

The Hilsenhoff index was originally based on species level data but has been modified for use at the family level (Hilsenhoff 1988) and is widely used in North America. Although it must be noted that the Tolerance values (see Table 4) are largely based on the response to organic pollution, with sensitive species having low scores and tolerant species having high scores.

In streams with severe organic pollution the index may underestimate the degree of pollution because species within families can have a wide range of tolerance values, and this index may need adjustment for the Atlantic region.

The index is calculated:

$$\mathbf{HBI} = \sum x_i t_i / n$$

Where

x_i = number of individuals in each family

t_i = tolerance value for each family

n = number individuals in the sample

Table 4. Hilsenhoff family tolerance values.

Reference Number	Phylum	Class	Order	Family	Tolerance Value
162020002	Annelida	Hirudinoidea	Rhynchobdellida	Piscicolidae	NA
121020019	Annelida	Oligochaeta	Haplotaxida	Naididae	8
121040012	Annelida	Oligochaeta	Haplotaxida	Tubificidae	NA
122010001	Annelida	Oligochaeta	Lumbriculida	Lumbriculidae	7
31010001	Arthropoda	Arachnida	Hydroida	Hydridae	5
191080001	Arthropoda	Arachnida	Trombidiformes	Lebertiidae	4
191090001	Arthropoda	Arachnida	Trombidiformes	Limnesiidae	4
191050001	Arthropoda	Arachnida	Trombidiformes	Protziidae	4
191120001	Arthropoda	Arachnida	Trombidiformes	Sperchontidae	4
191160001	Arthropoda	Arachnida	Trombidiformes	Stygothrombidiidae	NA
191130001	Arthropoda	Arachnida	Trombidiformes	Torrenticolidae	4
321040002	Arthropoda	Crustacea	Amphipoda	Crangonyctidae	5
321010003	Arthropoda	Crustacea	Amphipoda	Gammaridae	4
321030001	Arthropoda	Crustacea	Amphipoda	Hyalellidae	8
291010003	Arthropoda	Insecta	Coleoptera	Elmidae	4
201020001	Arthropoda	Insecta	Collembola	Hypogastruridae	5
201010002	Arthropoda	Insecta	Collembola	Isotomidae	5
301030007	Arthropoda	Insecta	Diptera	Athericidae	6
301030128	Arthropoda	Insecta	Diptera	Blephariceridae	6
301080001	Arthropoda	Insecta	Diptera	Ceratopogonidae	2
301090001	Arthropoda	Insecta	Diptera	Chironomidae	0
301050002	Arthropoda	Insecta	Diptera	Deuterophlebiidae	6
301100001	Arthropoda	Insecta	Diptera	Dixidae	3
301110002	Arthropoda	Insecta	Diptera	Empididae	10
301060003	Arthropoda	Insecta	Diptera	Pelicorhynchidae	6
301180001	Arthropoda	Insecta	Diptera	Psychodidae	8
301120001	Arthropoda	Insecta	Diptera	Simuliidae	5
301150001	Arthropoda	Insecta	Diptera	Stratiomyidae	5
301070004	Arthropoda	Insecta	Diptera	Tanyderidae	3
301030103	Arthropoda	Insecta	Diptera	Thaumaleidae	6
301030005	Arthropoda	Insecta	Diptera	Tipulidae	6
211050001	Arthropoda	Insecta	Ephemeroptera	Ameletidae	0
211100001	Arthropoda	Insecta	Ephemeroptera	Ametropodidae	11
211040001	Arthropoda	Insecta	Ephemeroptera	Baetidae	4
211060002	Arthropoda	Insecta	Ephemeroptera	Ephemerellidae	1
211020003	Arthropoda	Insecta	Ephemeroptera	Ephemeridae	4
211070009	Arthropoda	Insecta	Ephemeroptera	Heptageniidae	4
211090001	Arthropoda	Insecta	Ephemeroptera	Leptohyphidae	4
211080001	Arthropoda	Insecta	Ephemeroptera	Leptophlebiidae	2
251010001	Arthropoda	Insecta	Megaloptera	Sialidae	4
231020001	Arthropoda	Insecta	Plecoptera	Capniidae	1
231070003	Arthropoda	Insecta	Plecoptera	Chloroperlidae	1
231040001	Arthropoda	Insecta	Plecoptera	Leuctridae	0
231050001	Arthropoda	Insecta	Plecoptera	Nemouridae	2
231090001	Arthropoda	Insecta	Plecoptera	Peltoperlidae	0
231010001	Arthropoda	Insecta	Plecoptera	Perlidae	1
231030001	Arthropoda	Insecta	Plecoptera	Perlodidae	2
231060002	Arthropoda	Insecta	Plecoptera	Pteronarcyidae	0
231080001	Arthropoda	Insecta	Plecoptera	Taeniopterygidae	2
241090001	Arthropoda	Insecta	Trichoptera	Apataniidae	4
241100002	Arthropoda	Insecta	Trichoptera	Brachycentridae	3
241060001	Arthropoda	Insecta	Trichoptera	Glossosomatidae	6
241170004	Arthropoda	Insecta	Trichoptera	Hydropsychidae	4
241110001	Arthropoda	Insecta	Trichoptera	Hydroptilidae	1
241120002	Arthropoda	Insecta	Trichoptera	Lepidostomatidae	1
241020016	Arthropoda	Insecta	Trichoptera	Leptoceridae	4
241030006	Arthropoda	Insecta	Trichoptera	Limnephilidae	4
241150001	Arthropoda	Insecta	Trichoptera	Philopotamidae	4
241130001	Arthropoda	Insecta	Trichoptera	Polycentropodidae	0
241010004	Arthropoda	Insecta	Trichoptera	Rhyacophilidae	4
241140002	Arthropoda	Insecta	Trichoptera	Uenoidae	0
104020001	Mollusca	Gastropoda	Basommatophora	Physidae	8
104030004	Mollusca	Gastropoda	Basommatophora	Planorbidae	5
101010006	Mollusca	Gastropoda	Heterostropha	Valvatidae	5
111010001	Mollusca	Lamellibranchia	Unionoida	Margaritiferidae	NA
112020006	Mollusca	Lamellibranchia	Veneroida	Pisidiidae	8

Generally the higher the index the more likely a pollution problem is present. Table 5 provides some guidance on water quality assessments and tolerance values. Perhaps more important than the absolute score is the similarity between the reference sites and the test sites. This can be expressed as percent similarity:

% Similarity = (reference HBI/test HBI) x 100

The US EPA (1989) suggest that the following degrees of similarity are indicative of impairment:

>85% similarity - unimpaired
84-70% similarity - slightly impaired
69-50% similarity - moderately impaired
< 50% similarity - severely impaired

Table 5. Evaluation of stream condition using the family biotic index (adapted from Hilsenhoff 1988)

Calculated Family Biotic Index	Stream condition
0.00 – 3.75	Excellent
3.76 – 4.25	Very good
4.26 – 5.00	Good
5.01 – 5.75	Fair
5.76 – 6.50	Fairly poor
6.51 – 7.25	Poor
7.26 – 10.00	Very poor

Bray-Curtis similarity

Most of the previous invertebrate community statistics are measures of total abundance and taxon richness but do not provide any information on what kind of organisms are present. A *similarity index* is also recommended as these summarize the overall difference in community structure between reference and exposed sites in a single number, they require no preconceived assumptions about the nature of the community and they only vary in one direction (Taylor and Bailey 1997). Of the various indices available many reviewers have indicated that the Bray-Curtis Index (Bray and Curtis 1957) is the most reliable (Pontasch et al 1989, Marchant et al 1984, Jackson 1993, Bloom 1981). Faith et al 1991 showed that for detecting effects of uranium and gold mines the Bray-Curtis (B-C) Index was superior to seven other indices and provided consistently high statistical power because of its low susceptibility to temporal variability at control sites and sensitive response to disturbance. It is also unaffected by the nature of communities being compared (Bloom 1981) and differences contribute the same to the Bray-Curtis (B-C) index regardless of the species being rare or abundant. Bloom (1981) showed that of four indices examined only the B-C index reflected accurately the true resemblance over its range.

The B-C Index is a distance co-efficient that reaches a maximum value of 1 for two sites that are entirely different and a minimum value of 0 for two sites that possess identical descriptors. These distance coefficients measure the amount of association between sites and the

B-C Index is a member of the class of distance coefficient known as a semi metric that some prefer to call dissimilarity coefficients. The B-C index measures the percentage of difference between sites (Legendre and Legendre 1983):

$$D = \frac{\sum_{i=1}^{n} |y_{i1} - y_{i2}|}{\sum_{i=1}^{n} (y_{i1} + y_{i2})}$$

where:
D = Bray-Curtis distance between sites 1 and 2;
Y_{i1} = count for species i at site 1;
Y_{i2} = count for species i at site 2;
N = total number of species present at the two sites.

To illustrate how the index should be used an example is presented where samples were taken from an exposed and reference area (Table 6) with a total of 5 species present. To calculate the Bray-Curtis similarity the absolute (remove sign) differences for each species are summed and simply divided by the total count for both samples. If multiple refernce sites are sampled then the average distance of the reference samples to the reference median is compared to the average distance of the exposed samples to the reference median

Table 6. example calculation of Bray-Curtis similarity for two samples.

	Sp 1	Sp 2	Sp 3	Sp 4	Sp 5	Sum diff.	Sum of Sp 1-5	BC dist
Reference	2	3	2	3	1		11	
Test	4	4	2	3	1		14	
Reference-test	2	1	0	0	0	3	25	0.2727

Where:

$$D_{ref-test} = \frac{2+1+0+0+0}{11+14} = \frac{3}{25} = 0.2727$$

Multivariate statistics

RCA model building procedures

Step 1. Data Preparation

Requirements, two data matrices:

Matrix 1(invertebrate data) consists of a square data matrix with sites as rows, and taxa (at selected taxonomic level as columns.

Matrix 2 (habitat data) consists of a square data matrix with sites as rows and habitat data as columns.

NOTE: the rows in the two matrices must match, i.e., each row should represent the same site in each matrix.

Decisions on appropriate taxonomic level, data transformation for up weighting and down weighting of taxa, and meeting requirements of normality in the habitat data matrix should be made at this time.

Step 2. Classification (for formation of biotic groups)

The predictive modeling approach used, multiple discriminant analysis, requires the formation of groups form the invertebrate data matrix. This is done using cluster analysis, the process using PATN is:

Data entry

- Preparation of three ASCII files for input to PATN. These files are created in PFE (Programme File Editor). The first file contains the data only (sites as rows, taxa as columns) and as give the file extension *.dta*. The second contains the row labels (site names), one per line extension *.row*. The third contains the column labels (taxa) file extension *.col*. These three files are saved in the PATN/DATA sub directory.
- Start PATN and run *Preparation, Input Output Menu, Parameter* – then describe the structure of the data, title, no rows and columns and file name.
- Run *Preparation, Input Output Menu, Data input* – this reads the data file into PATN, remember to use (*) as the data input format.
- Run *Preparation, Input Output Menu, labels* – to read in the two label files for rows and columns again (*) as the data input format.
- Run *Preparation, statistics, Presence absence* – to ensure the data are entered as expected.

Selection of an association measure.

- You need to create an association matrix as this is the basis of all the following analysis, the recommended association measure is Bray-Curtis option 1 in *Analysis, Association menu, Association measures.*

Classification

- You know need to generate a dendrogram of your sites so that you can divide the sites into groups representing different assemblages of taxa. This is done by using the

Analysis, Classification menu, Agglomerative clustering methods are recommended Option UPGMA (option 5).

- Use the default values
- To construct the dendrogram use the *Evaluation, Classification evaluation, Dendrogram Drawing menu.*
- It is easier to read the dendrogram file by opening it in Word (use a non-proportional font).
- Selection of groups is a subjective process, but ideally you want a minimum groups size of 10 sites. You can use ordination to help you select the groups.

Ordination

- Multi-dimensional scaling is the recommended method as it makes no assumptions about normality of the data. Use Analysis, Ordination, SSH. You will have to select the number of dimensions, try 2 then 3 and so on, until you get a suitable stress level, definitely < 0.2, preferably < 0.15.
- You can look at the relationship between the ordination and the original matrix (or the habitat matrix) by using PCC in *evaluation, ordination, fitting attributes*. This gives the position of the variable in the ordination matrix and the correlation.
- You can also examine the significance by using Monte Carlo simulation in *evaluation, ordination, Monte Carlo attributes.*

Groups

- **Once you have finalized which sites belong in which groups you have to insert a grouping variable in the two original matrices (invertebrate and habitat). Again this is a subjective and iterative analysis.**

Step 3. Model building (matching sites to habitat)

In this step you will try to relate the habitat attributes to the groups formed from the invertebrate matrix. The purpose for this is to develop a model using multiple discriminant analysis (MDA) that will assign a probability of a site belonging to each of the invertebrate groups. It is important to note that:

- Models are built using reference sites only, and are then applied to test sites – see Step 4 below.
- That variables used in the models should not, or only be minimally, affected by the suspected stressor. The best example would be not to use phosphorus as a predictor variable when investigating impacts from sewage discharges. While nutrients often show a relationship with biological structure in reference sites, use of phosphorus in the model would match the test site to nutrient rich reference sites which may not be appropriate.

There are three ways to select habitat variables for use in a model:

Principle Axis Correlation in PATN.

This relates the invertebrate ordination structure to that of the habitat matrix. The following steps are required:

- Generate three files for PATN as described above for the invertebrate data. Note that the habitat data need to be range standardized (*Preparation, Manipulation, Data standardizing*: option 4), as they are measured in different units. The association measure to use is Euclidean distance (see PRIMER for more discussion)
- Input the data as above for invertebrate data
- Run a multi dimensional scaling ordination (above)
- Run fitting attributes (PCC) as described above BUT select the invertebrate SSH file created above as the ordination file. This then generates correlations between the ordination scores for the invertebrate matrix with the habitat file, these can be assessed using a Monte Carlo simulation (MCAO)

ANOVA.

You can use the invertebrate groups as the categories and run ANOVAS on all the habitat data to determine which show significant differences among the groups and using Tukey's test can determine which variables discriminate which groups. This analysis can be done in SYSTAT or SIGMA STAT. Remember to test all data for normality as ANOVA assumes a normal distribution. This can be done by drawing probability plots.

Stepwise Discriminant Analysis (in SYSTAT)

To run this analysis you use the habitat matrix plus the groups from the invertebrate matrix. Run either a forward or backward stepwise to see which variables seem to contribute the most.

Model building

This is an iterative process until you get the best combination of groups and predictive capacity. I first run a complete MDA with all the variables and use the cross-validation (jackknifed) classification to assess the performance. This is the more rigorous test of the model. Compare the performance of various combinations of variables with the complete data set to get the set of variables with the lowest error rate. This is your model.

Step 4. Assessing test sites

All the preceeding steps are done once only, the objective being to develop a method for assessing test sites. Both methods require running MDA with reference sites and test sites :

- Generate a file in excel that contains, site number, Reference (1 for ref, 0 test) Gp (no value for test sites), and predictor variables, for the reference and test sites.
- Set data weight and select reference – this uses only refernce sites to generate the model but runs it against the test sites
- Open in systat and run Discriminat Analysis (Option – complete) and save distances and data into a file. You can then open that file and it contains the probability of each site belonging to one of the groups. These probabilities are then pasted back into the excel worksheet.

There are two methods for assessing test site with the RCA.

RIVPACS

This the method developed by Wright et al in the UK, and used by them and by the Australians (AUSRIVAS). This method predicts the probability of a taxon occurring at a site by adding the weighted probabilities of each taxa belonging to a group:

> Total P of taxa occurrence = (P of test site being in group 1 x % occurrence of taxa in group1)+ (P of test site being in group 2 x % occurrence of taxa in group 2) etc.

At a selected P level the number of taxa likely to occur is determined by adding the probabilities this gives the expected number of taxa (see Table 7).

Table 7. Calculation of O/E scores for invertebrate families.

Taxa	**P of occurrence (from model)**	**Actual occurrence**		
Baetidae	0.96	34		
Leuctridae	0.96	12		
Leptophlebiidae	0.82	0		
Hydropsychidae	0.79	2		
Caenidae	0.75	0		
Heptageniidae	0.74	3		
Ephemerellidae	0.71	0		
Leptoceridae	0.59			
Lepidostomatidae	0.42			
Glossosomatidae	0.34			
Goeridae	0.32			
Brachycentridae	0.04			

Expected taxa at P> 0.70 5.74

Observed taxa *5*

O:E ratio *5/5.74 = 0.87*

While there are 7 taxa with a probability of occurrence of more than 0.7. The expected number is less as this is determined by adding the probabilities. Sites are assessed using the O:E

ratio. Assessment Bands are determined from calculating the mean and SD for O:E ratios from the reference sites and setting bands based on the SD (see Wright 1995).

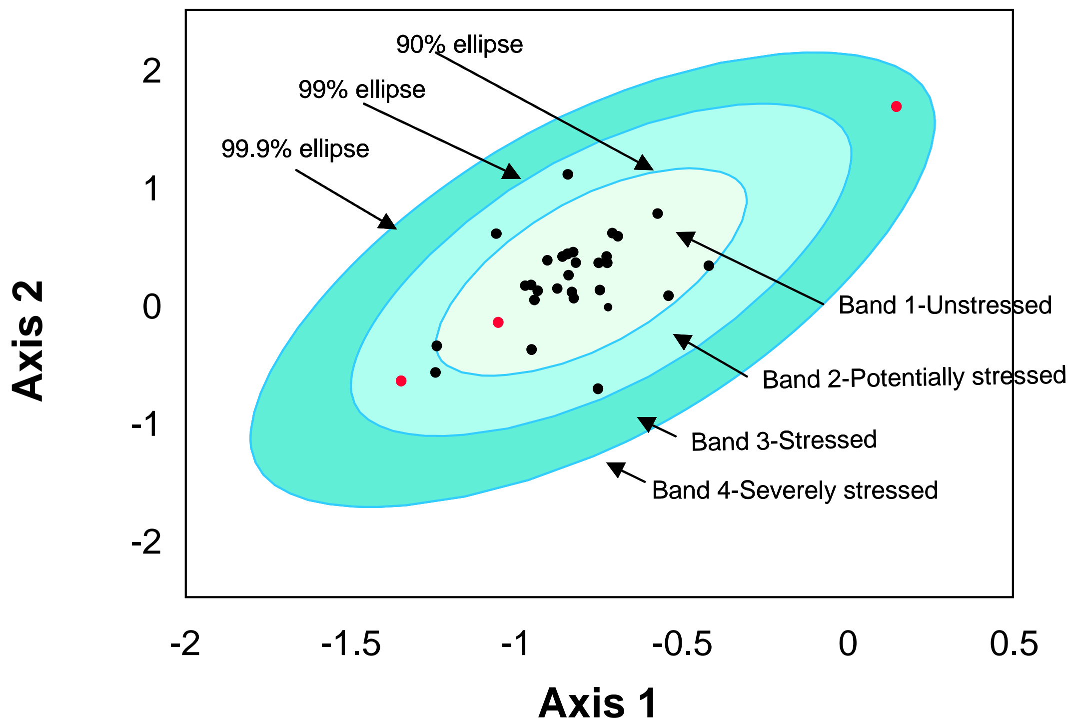
The advantages of this approach are that it uses all the reference sites in the assessment, and actually indicates which taxa are absent which helps in interpretation. Disadvantages are that it is based on presence absence only and relies solely on taxa richness.

BEAST

The second approach was developed by Reynoldson et al and uses ordination. Test sites are plotted in the same ordination space as matched reference sites, and the distance of the test site from the reference space is used as an indicator of impairment:

The probabilities of test site membership examined and the test site data is appended to the refernce site data of the group to which they have the greatest probability of belonging, e.g., a test site with the following probabilities: **Gp 1** 0.5103, **Gp 2** 0.1938, **Gp 3** 0.2959 would only be compared to the Gp 1 reference sites. The test site data would be appended to the Gp 1 reference sites. This new file would be ordinated in PATN as described above.

The new ordination file is then used in SYSTAT and the reference sites and test site plotted in SYSTAT. Probabilty ellipses (90, 99, 99.9%) are plotted around the reference sites only and the test site is assigned to the worst Band to which it resides.



NOTE: you should not use a ratio of less than 10 ref sites to 1 test site.

References

Bloom, S.A. 1981. Similarity indices in community studies: potential pitfalls. Mar. Ecol. Prog. Ser. 5:125-128.

Bray, J.R. and J.T. Curtis. 1957. Ordination of the upland forest communities of southern Wisconsin. Ecol. Monogr. 27:325-349.

Faith, D.P., C.L. Humphrey and P.L. Dostine. 1991. Statistical power and BACI designs in biological monitoring: Comparative evaluation of measures of community dissimilarity based on benthic macroinvertebrate communities in Rockhole Mine Creek, Northern Territory, Australia. Aust. J. Mar. Freshwat. Res. 42:589-602.

Green, R.H. 1979. *Sampling design and statistical methods for environmental problems*. Wiley, New York.

Hilsenhoff, W.L. 1988. Rapid field bioassessment of organic pollution with a family level biotic index. J. N. Am. Benthol. Soc. 7:65-68.

Jackson, D.A. 1993. Multivariate analysis of benthic invertebrate communities: the implication of choosing particular data standardizations, measures of association, and ordination methods. Hydrobiologia 268:9-26.

Legendre, L., and Legendre, P. 1983. Measures of ecological resemblance. In Numerical ecology. Elsevier Scientific Publishing Company, Amsterdam.

Marchant, R., P. Mitchell and R.H. Norris. 1984. Distribution of benthic invertebrates along a disturbed section of the LaTrobe River, Victoria: An analysis based on numerical classification. Aust. J. Mar. Freshwat. Res. 35:355-374.

Merritt, R.W. and K.W. Cummins (eds.) 1996. An introduction to the aquatic insects of North America. 3rd ed. Kendall-Hunt, Dubuque, Iowa. 862 p.

Metcalfe, J.L. 1989. Biological water quality assessment of running waters based on macroinvertebrate communities: history and present status in Europe. Environmental Pollution 60: 101-139..

Norris, R.H. and A. Georges. 1993. Analysis and interpretation of benthic macroinvertebrate surveys, p. 234-286. In: D.M. Rosenberg and V.H. Resh (eds.) Freshwater biomonitoring and benthic macroinvertebrates. Chapman and Hall, New York.

Plafkin, J.L., M.T. Barbour, K.D. Porter, S.K. Gross, and R.M. Hughes. 1989. Rapid bioassessment protocols for use in streams and rivers: benthic macroinvertebrates and fish. EPA/444/4-89-001. Assessment and Watershed Protection Division, Washington DC.

Pontasch, K.W., E.P. Smith and J. Cairns Jr. 1989. Diversity indices, community comparison indices and canonical discriminant analysis: Interpreting the results of multispecies toxicity tests. Water Res. 22:1229-1238.

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

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

Reynoldson, T.B., R.C. Bailey, K.E. Day, and R.H. Norris. 1995. Biological guidelines for freshwater sediment based on BEnthic Assessment of SedimenT (the BEAST) using a multivariate approach for predicting biological state. Australian Journal of Ecology 20: 198-219..

Reynoldson, T.B., Rosenberg, D.R , Day, K.E., Norris, R.H., Resh, V.H.. 1997. The reference condition: a comparison of multimetric and multivariate approaches to assess water-quality impairement using benthic macroinvertebrates. J. N. Am. Benth. Soc. 16:833-852.

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

Taylor, B.R. and R.C. Bailey. 1997. Technical evaluation on methods for benthic invertebrate data analysis and interpretation – final report. AETE project 2.1.3. Prepared for CANMET, Ottawa.

Underwood, A.J. 1991. Beyond BACI: experimental designs for detecting human environmental impacts on temporal variations in natural populations. Australian J. Marine Freshwater Res. 42:569-587.

Wright, J.F. 1995. Development and use of a system for predicting the macroinvertebrate fauna in flowing waters. Australian Journal of Ecology 20: 181-197..

Appendix:

List of Suppliers

Sample jars	
Kick nets	
Ethanol	
Formalin	
Waterproof paper	
Sub-sampling (Marchant) box	
Fine tweezers	

Data Sheets