

Assessing Metal Mine Effects using Benthic Invertebrates for Canada's Environmental Effects Program

NANCY E. GLOZIER,^{1*} JOSEPH M. CULP,¹ TREFOR B. REYNOLDSON,²
ROBERT C. BAILEY,³ RICHARD B. LOWELL^{1,4} AND LISE TRUDEL⁴

¹National Water Research Institute, Environment Canada, 11 Innovation Boulevard,
Saskatoon, Saskatchewan S7N 3H5

²Centre for Estuarine Research, Box 115 Acadia University, 23 Westwood Avenue,
Wolfville, Nova Scotia B0P 1X0

³Department of Zoology, University of Western Ontario, London, Ontario N6A 5B7

⁴National EEM Office, Environment Canada, 351 St. Joseph Boulevard, Hull, Quebec
K1A 0H3

In Canada, the Metal Mining Effluent Regulation is a mechanism developed from the *Fisheries Act* (R.S., c. F-14, s.1) under which the effects of mine effluent on fish and fish habitat (i.e., benthic invertebrate communities) is determined by Environmental Effects Monitoring (EEM) studies. The Metal Mining EEM (MM EEM) program proceeds in a tiered manner, commencing with determining whether an effect is present and continuing with determining extent, magnitude and cause of the effect. The benthic invertebrate monitoring component of the MM EEM program includes consideration of study design elements such as confounding factors, monitoring frequency, statistical study design, appropriate community endpoints and standardized approaches to site descriptions, field and laboratory methods and data interpretation. We present the approaches and rationale recently adopted for the benthic component of Canada's Metal Mining EEM program. A primary objective of this program was to outline a consistent national program that was scientifically defensible and that would produce the necessary information to evaluate the effectiveness of current pollution regulations.

Key words: aquatic, bioassessment, benthic invertebrates, Environmental Effects Monitoring, metal mining

Introduction

Although environmental monitoring has traditionally been based on chemical and physical assessments, there has been a recent movement towards greater incorporation of biological information into these programs. This new assessment approach recognizes the complex nature of interactions within the environment and attempts to relate biological patterns to impacts arising from anthropogenic activities. For example, biologically based environmental monitoring programs have been imple-

*Corresponding author; nancy.glozier@ec.gc.ca

mented at various spatial scales (i.e., site specific, regional and national) in the United Kingdom (Wright et al. 2000), Australia (Faith et al. 1995; Davies 2000) the United States (Messer et al. 1991; Stevens 1994; Powell 1995) and in Canada (Environment Canada 1997, 1998, 2002). These programs range in scope from voluntary efforts of citizens' groups designed to assess ecosystem health (Taccogna and Munro 1995; Environmental Protection Agency 1995; Culp and Halliwell 1999), to detailed assessments implemented by regulators in an attempt to identify management priorities (Nusser and Goeble 1997; Roux et al. 1999). These monitoring programs have generated large data sets for effects evaluation and have led to the development of new protocols (e.g., Rapid Bioassessment Protocols, Barbour et al. 1999), indices (e.g., Stream Condition Index, Barbour et al. 1996), study designs (e.g., reference condition approach, Reynoldson et al. 1997) and alternative methods (e.g., mesocosm methodologies for bioassessment, Culp et al. 2000; Dubé et al. 2001, 2002). In this paper, we discuss considerations for the design and implementation of bioassessment programs intended to measure metal mining effects on benthic invertebrates. These protocols are based on monitoring frameworks developed for Canada's Metal Mining (MM) Environmental Effects Monitoring (EEM) program (Environment Canada 2002). A unique feature of this EEM approach is that the standardized sampling methodology and study design frameworks facilitates the application of the approach across regional and national scales.

Development of Benthic Invertebrate EEM Programs in Canada

In Canada, waste disposal for the pulp and paper sector is regulated by the Pulp and Paper Effluent Regulations of the *Fisheries Act* (R.S., c. F-14, s.1). The effectiveness of these regulations in protecting the receiving environment is assessed through Canada's EEM Program (Walker et al. 2002). It consists of a series of assessment cycles with the first cycle completed in April 1996. Preliminary examination of the initial results indicated that many (69%) pulp mills reported differences between reference and exposure sites for benthic invertebrate endpoints of species richness and abundance (Environment Canada 1997). However, data interpretation was often confounded by factors such as multiple effluent discharges, inappropriate reference site selection, lack of standardized methods (e.g., sampler type and mesh size) and inappropriate data analyses. A Benthic Community Expert Working Group (BCEWG) was formed in November 1996 in order to provide recommendations for future EEM cycles. Their major recommendations addressed three issues: (1) site selection protocols and study design, (2) sampling methodologies, and (3) data interpretation (Environment Canada 1997). Specific recommendations in the first category included development of protocols for *a priori* reference site selection, adoption of standard nomenclature for site descriptions and the development of criteria for the determination of sampling effort (i.e., number of replicates). In addition, they recommended that standard study designs include options for more than one reference area.

Recommendations for improving EEM sampling methods focused on the standardization of habitat type, sampling device, mesh size and laboratory sorting and sub-sampling protocols. Many recommendations to improve data interpretation capabilities were forwarded, ranging from standardized supporting variables and community descriptors, simple data handling, screening and transformation protocols. The BCEWG noted that a key deficiency was the absence of thresholds to define effects of ecological importance (i.e., critical effect size thresholds, Lowell 1997). This detailed evaluation of the first cycle results was instrumental in advancing EEM programs in Canada because BCEWG recommendations were quickly implemented into the Technical Guidance Document that was provided to the pulp and paper sector for use in the second monitoring cycle (Environment Canada 1998).

Preliminary analyses of Cycle 2 EEM results indicate that the second cycle was successful in addressing the concerns raised by the BCEWG (R. Parker, R. Lowell, S. Blenkinsopp, pers. comm.), and an evaluation currently underway will provide a more comprehensive analysis (Lowell et al. In prep.). With the pulp and paper EEM program established, a program was initiated for the metal mining industry (Aquamin 1996). In addition, an extensive government-industry research program, the Aquatic Effects Technology Evaluation Program (AETE) reviewed and evaluated environmental monitoring technologies as assessment tools for the determination of mining-related impacts on the aquatic environment. AETE research projects included field evaluations at mine sites (Beak 1996), method development (Taylor 1997), data analysis and interpretation (Taylor and Bailey 1997) and quality assurance programs (Beak 1999). Along with the Pulp and Paper EEM program, Aquamin and the AETE provided the basis from which frameworks and protocols for metal mining monitoring were developed.

Canada's Metal Mining Effluent Regulation (MMER) is a regulatory mechanism developed from the Canadian *Fisheries Act*. Under the MMER the effects of mine effluent on the aquatic receiving environment (i.e., effects on fish, fish habitat, and the use of fish by man) are evaluated. Because of their prominence in the aquatic food web, the benthic invertebrate community is considered the key component of fish habitat. Thus, to evaluate the effect of mine effluent on fish habitat, mines are required to determine if the mine effluent has modified the benthic invertebrate community (Fig. 1; Dumaresq et al. 2002). The EEM program for metal mining was designed to address a sequence of questions in a tiered manner, with the particular sequence dependent on results. This approach differs from the repeated cycles of the pulp and paper EEM program in that the scope of the questions addressed in each tier became more progressively focused. For example, the first step in the metal mining program (i.e., Site Characterization) involves summarizing existing information for the mine site, including the identification of potential reference and exposure areas, habitat mapping and a review of previously collected benthic invertebrate data. Subsequent phases involve determining if an effect on the benthic invertebrate community exists within the immediate vicinity

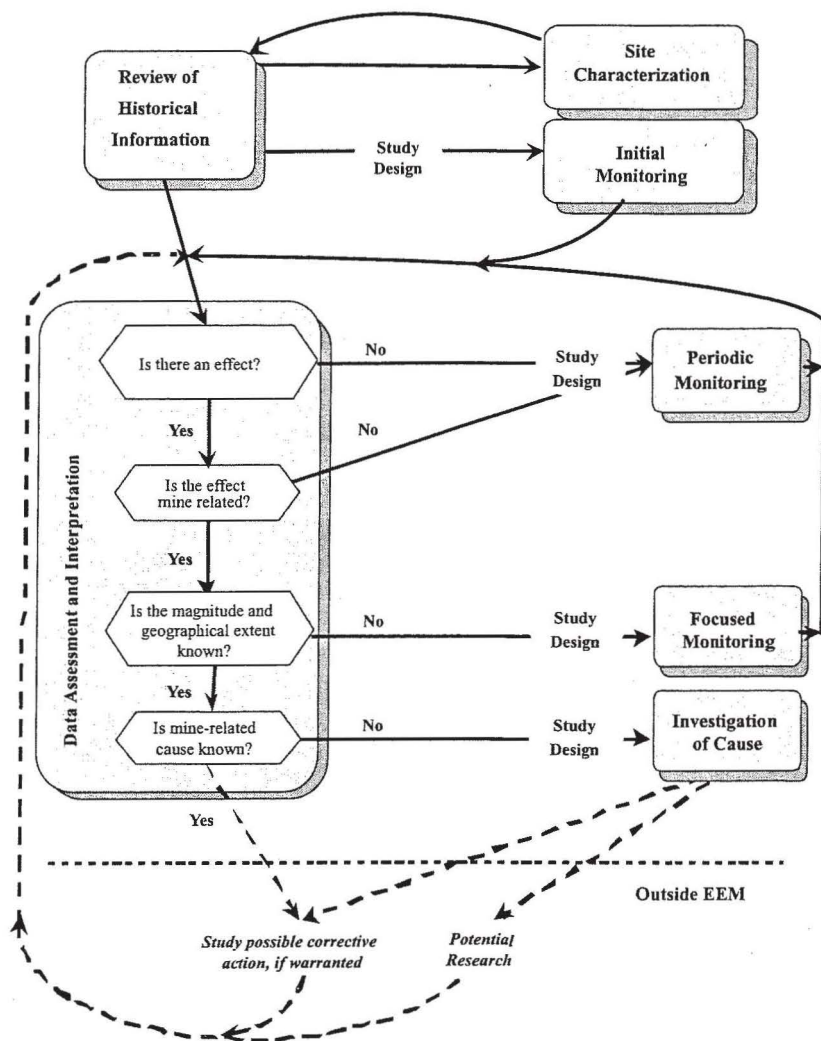


Fig. 1. Overall decision-making framework for Canada's Metal Mining EEM program, illustrating the decision path through four different monitoring phases, namely, Site Characterization, Initial Monitoring, Focused Monitoring and Investigation of Cause.

of the effluent discharge (Initial Monitoring), determining the extent and magnitude of these effects (Focused Monitoring) and determining the mine-related cause of the effect (Investigation of Cause). If an effect is not detected, the program also includes a phase where the aquatic receiving environment is monitored on an ongoing basis to detect changes from the current conditions, and to allow for the generation of longer term trend data (Periodic Monitoring). The tiered nature of the metal mining program was important in being able to couple the MMER mandate of a nationally consistent program with the site-specific flexibility required for

a monitoring program designed to address the diverse nature of Canada's mining industry. The key elements in designing a tiered monitoring program for benthic invertebrates included: consideration of historic and confounding factors, standardized site descriptions, appropriate study designs and their relevant statistical approaches, ecologically relevant seasons, monitoring frequency and habitats, appropriate supporting environmental variables and benthic invertebrate community descriptors, and the standardization of field and laboratory methods. In addition, it was important to apply consistent data assessment and interpretation procedures and to define an effect on the benthic invertebrate community.

Study Design Considerations for Metal Mining EEM Field Monitoring Programs

Consideration of Historic Information and Confounding Factors

The first step in designing a field monitoring program is the collation and review of physical habitat information and historic benthic invertebrate data and sampling protocols for representative watersheds within the ecoregion. This information can be used to aid in aspects of study design such as determining appropriate field sampling seasons, habitats and procedures, including an estimate of the number of samples required (Environment Canada 2001). A decision tree is used for reviewing and evaluating existing data during examination of historic information and during each study design phase (Fig. 2). This decision tree poses a series of questions regarding site-specific conditions that may affect potential study design options. It is intended for use at the onset of all study design exercises for any of the phases of the monitoring program (i.e., this series of questions is applicable for the Initial, Periodic or Focused Monitoring phases of the Metal Mining EEM program). A key element in the decision tree is the consideration of potential confounding factors. For example, environmental variables may confound the interpretation of the data if it is not possible to separate the effect of the mine effluent from the effects of covarying changes in natural habitat variables. While modifications to study and sampling design elements can often resolve confounding factors, consideration of alternatives such as mesocosms (Culp et al. 2000; Dubé et al. 2001, 2002) may be necessary to assess effects.

Monitoring Frequency

The frequency of benthic invertebrate surveys for the MM EEM program is dependent on the phase of the program and the results of previous monitoring (Dumaresq et al. 2002). For the Initial Monitoring phase, a benthic survey is to be completed within three years of the commencement of the MM EEM program. For Periodic Monitoring there are three potential sampling options: the more frequent *Confirmation Monitoring* that requires a survey in the next appropriate season, to *Surveillance Monitoring* which requires a survey within three years and finally to *Minimal Monitoring*, requiring a survey within six years. For Focused

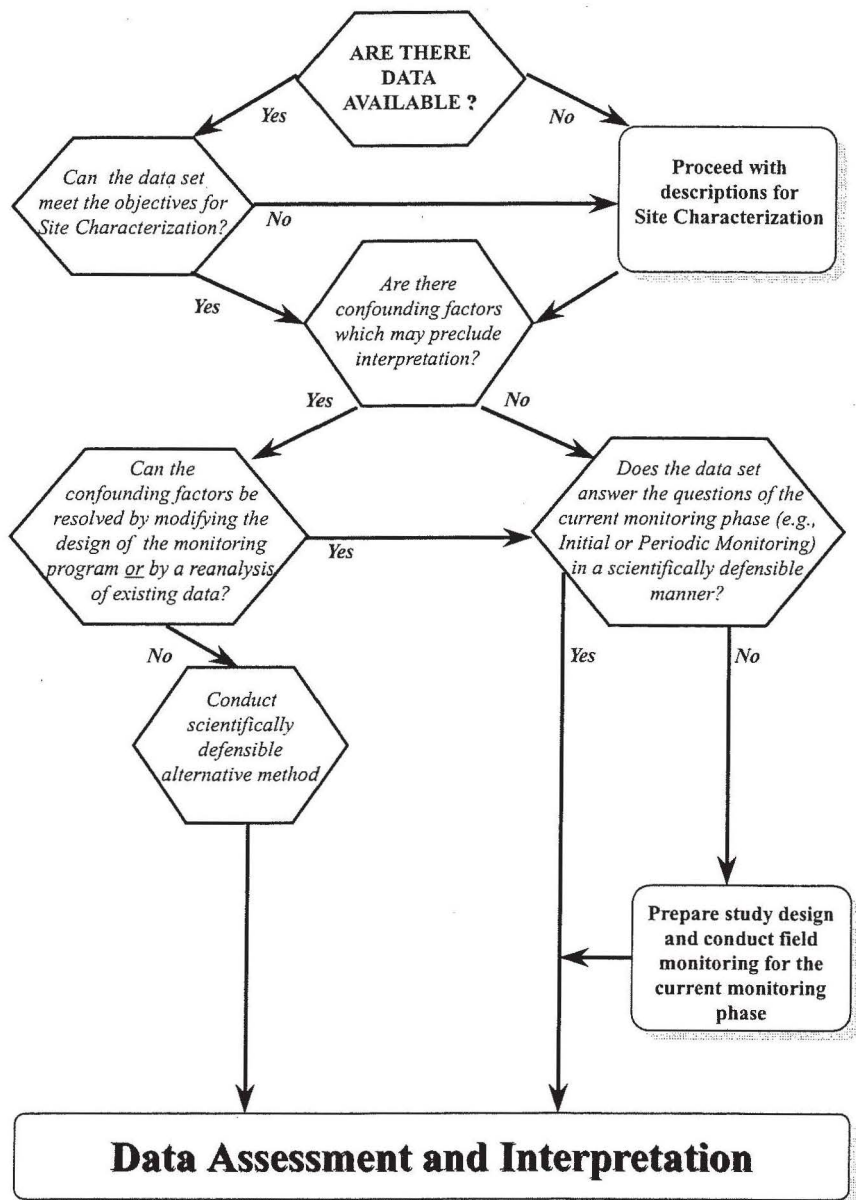


Fig. 2. Decision tree for assessing whether existing data is adequate for data assessment and interpretation. An important aspect of the tree is the determination of whether confounding factors preclude the interpretation of monitoring data.

Monitoring and Investigation of Cause, the investigation is to be conducted within the next appropriate season. The rationale for these frequencies is largely based on the requirement of coupling realistic field logistics for performing scientifically valid surveys with a need to address

the demands of the regulators. See Dumaresq et al. (2002) for a full explanation of sampling frequency rationale.

Standardized Site Descriptions

Definitions of sampling locations or sites in an aquatic ecosystem can often depend on the scale of the project or the background of the technical crew. After a review of data from cycle 1, it was recommended that a standard nomenclature be developed and used for site descriptions. Standardized definitions for sampling locations are essential to allow the delivery of consistency between monitoring studies conducted at different mines, as well as studies at the same mine conducted over time. The standardized terminology for Canada's EEM programs was developed in the context of an ANOVA study design (i.e., Control/Impact or Multiple Control/Impact) and included the following terms: Reference Area, Exposure Area, Replicate Station and Field Sub-Sample (Fig. 3 and Table 1, see also Environment Canada 2001). As the objective of ANOVA study designs is to detect differences between or among areas, each reference or exposure area consists of a number of replicate stations (i.e., the *error variation* in ANOVA) and each replicate station consists of a number of pooled field sub-samples. With the adoption of this standardized terminology, improvement has been achieved in the second cycle of monitoring for the Pulp and Paper EEM program (R. Lowell, pers. comm.).

For the tiered MM EEM program, the geographical extent of an area will differ between the program phases because they each have unique objectives. For Initial and Periodic Monitoring, the exposure area to be studied is defined as the aquatic receiving environment that is closest to the effluent discharge. However, for Focused Monitoring the additional objective of determining the spatial extent of an effect requires the definition of the exposure area to be expanded, including areas with lower levels of effluent to a point where a return to reference condition is reached.

Sampling Program Designs and Statistical Approaches

Difficulties in comparing national or regional impacts in receiving environments were encountered during the review of the first cycle of the Pulp and Paper EEM because little effort had been made to ensure large-scale consistency (Environment Canada 1997). The BCEWG's review of this first cycle recommended six basic sampling designs for the benthic invertebrate survey for the MM EEM program including: (1) Control/Impact design (C/I); (2) Multiple Control/Impact design (MC/I); (3) Simple Gradient (SG); (4) Radial Gradient (RG); (5) Multiple Gradient design (MG); and (6) Reference Condition Approach (RCA) (Environment Canada 1997, 2002). These designs fall into three basic categories with different philosophical approaches and a summary of their attributes, applicability and limitations is provided below (see also Table 2). If these field study design options will not address site-specific issues (Fig. 2), a recommended alternative method is available and also described below.

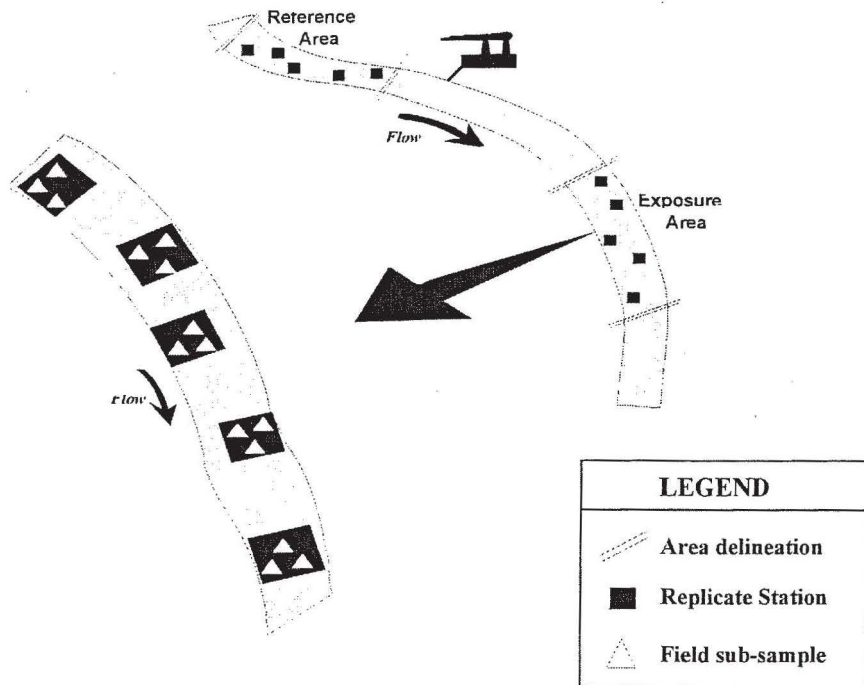


Fig. 3. Schematic diagram with examples of Area, Replicate Station and Field Sub-sample spatial scales for a basic Control/Impact Design used in Canada's EEM program.

Control/impact and multiple control/impact designs

The simplest study design for use in environmental effects monitoring is the Control/Impact (or reference exposure) design. For Initial and Periodic Monitoring, this would include at least one exposure area. Levels of exposure to mine effluent differ between impacted and reference areas, but must be similar between stations within an area. Thus, with the exception of effluent exposure level, these areas should be as similar as possible in habitat characteristics (e.g., substrate, current velocity, water chemistry properties, environmental gradients, land use, etc.). For Focused Monitoring, the C/I design can be used to ascertain the geographic extent of an effect by targeting additional exposure areas in localities where the effect is suspected to dissipate. Careful consideration of existing data and potential confounding factors during Site Characterization and study planning should minimize the chance of confounded study results. However, in most cases additional reference areas can alleviate these potential difficulties. These Multiple Control/Impact study designs include additional reference areas located in adjacent watersheds or bays with comparable habitat and environmental characteristics found within the exposure area. Note that the C/I design is only recommended for simple, homogeneous rivers and streams or lakes without confounding factors.

Simple, radial or multiple gradient designs

Gradient designs (i.e., simple, radial or multiple) for EEM are intended to examine changes in community structure along a physical and/or effluent gradient. The spacing of stations in these designs requires a continuum of sampling stations along the gradient and, except when there are discontinuities along the gradient, stations are evenly distributed along the effluent gradient. Simple or radial gradient designs are suitable for situations where rapid changes in effluent dilution within the receiving environment preclude the selection of an exposure area with relatively homogeneous effluent concentration. These designs are also useful when suitable reference areas are not available. In cases where the exposure gradient is unavoidably correlated with a co-occurring environmental gradient (e.g., depth gradients in lakes), a multiple gradient design may be appropriate. In these cases, it is useful to compare exposure versus reference gradients since it would be possible to make statistical comparisons of the exposure area gradient with a similar environmental gradient in a refer-

Table 1. Standardized definitions of sample site description terminology (see Fig. 3 for schematic diagrams)

Term	Definition
Field sub-sample	Field sub-samples consist of individual area or time-limited collections of benthic invertebrates (e.g., one grab or cylinder sample).
Replicate station	A replicate station is a fixed sampling location that can be re-sampled and defined quantitatively (e.g., latitude and longitude). Sufficient physical separation must exist between the replicate stations to allow statistical independence of stations. Stations located within the exposure area should be positioned to ensure exposure to the effluent plume.
Area	An area is a portion of the receiving environment having a predefined level of effluent exposure and that contains appropriate habitat to accommodate the necessary number of replicate stations.
Reference area	A reference area should have no exposure to effluent from the mine under study, and biotic and abiotic characteristics similar to the exposure area.
Exposure area	An exposure area is that part of the aquatic environment close to the effluent discharge and continuing to a point where the benthic invertebrate community characteristics return to those found in reference areas and may extend through several types of receiving environments (e.g., streams, lakes or wetlands).

Table 2. Summary of benthic sampling program designs, including brief descriptions of appropriate reference or exposure areas and statistical analyses

Design type	Receiving environment	Reference areas	Exposure areas	Statistics
Control/Impact (C/I)	Homogeneous freshwater rivers or lakes	A single reference area, upstream of mine discharge	High Exposure Area for Initial Monitoring (additional exposure areas for Focused Monitoring)	ANOVA/ ANCOVA
Multiple Control/Impact (MC/I)	Freshwater rivers or lakes with geographically homogenous lake shores	Multiple reference areas in the same or environmentally similar adjacent watersheds or bays	High Exposure Area for Initial Monitoring (additional exposure areas for Focused Monitoring)	ANOVA/ ANCOVA
Simple Gradient (SG)	Freshwater rivers or geographically restricted lakes	A series of reference stations with no or low effluent levels, situated towards the end of a declining gradient of mine effluent	Single gradient through declining levels of effluent in the receiving environment	Regression/ ANOVA/ ANCOVA
Radial Gradient (RG)	Freshwater lakes, non-homogeneous open marine bays and coastal areas	As above, but situated toward the end of several radially oriented transects	As above, but repeated in a radially oriented design	As above

(continued)

Table 2. (concluded)

Design type	Receiving environment	Reference areas	Exposure areas	Statistics
Multiple Gradient (MG)	Freshwater lakes or rivers	A series of reference stations with no effluent situated on a transect along the same kind of environmental gradient observed in the exposure area	Gradient through declining levels of effluent and a co-occurring environmental gradient in the receiving environment	ANCOVA, with reference and exposure transects considered as treatment groups
Reference Condition Approach (RCA)	Freshwater rivers or lakes, particularly for co-operative investigations or where there is an existing reference database	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 exposure area which are tested individually against the reference station distribution	Multivariate/ANOVA/ANCOVA (if possible)
Mesocosm Bioassays	Freshwater lakes or rivers with confounding factors that preclude interpretation of field survey results	A single reference area, upstream of mine discharge or in an adjacent watershed from which seed invertebrate communities and substratum can be obtained	Appropriate exposure levels (e.g., 5% effluent) for the bioassays are determined from ecologically relevant concentrations in the receiving water	ANOVA/Regression

ence area(s). In gradient designs, reference information is obtained from the stations furthest away from the effluent source. For radial gradients, a comparison of regression patterns for each gradient (e.g., regressions of faunal abundance versus distance from the outfall) may help to illuminate the direction and extent of effects. Alternatively, all data from the gradients can be included in one regression if the comparison is between biotic and physical factors unrelated to geographic or natural habitat factors.

The reference condition approach

The fundamental concept of the Reference Condition Approach is to establish a database of sites that represents unimpaired conditions (reference stations) at which biological and environmental attributes are measured. This database is used to develop predictive models that match a set of environmental variables to biological conditions. The predictive models then allow a set of environmental measurements to be made at a new station and used in the model to predict the station's invertebrate community composition. A comparison of the community present at the new station with the predicted community allows an assessment of the condition of the new station to be made. The RCA can reduce the need to find appropriate reference areas within the same watershed which can be problematic in some traditional approaches. Rather than identifying and sampling reference areas upstream or in nearby bays of a lake, the RCA uses a set of biologically equivalent reference sites selected from an existing database to evaluate an exposure site. Due to the intensive initial sampling effort required to establish the reference database, it would be practical and cost effective for multiple EEM participants in a given region to collaborate in the development of the reference database. Furthermore, provided that it is kept up-to-date (i.e., less intensive sampling in subsequent years at a representative group of reference stations), the reference condition database can be used over a number of MM EEM phases.

Alternative monitoring methods

If confounding factors cannot be resolved through modification of the design of field surveys, other methods are available for determining an effect on the benthic invertebrate community. The recommended alternative method for the benthic invertebrate component in the MM EEM program is the application of mesocosms to conduct on-site community bioassays (Fig. 2). Mesocosms provide a means for the manipulation of aquatic populations and communities under controlled effluent exposure conditions and allow sufficient replication, and precise control over environmentally relevant variables under relevant field conditions (i.e., temperature, water quality, photoperiod). Thus, properly designed mesocosm studies will allow for the impact of the mine effluent to be examined independently from confounding factors present at a particular mine site. Mesocosms range widely in design (Graney et al. 1994; Pontasch 1995), and recent examples of the use of mesocosms for the assessment of effects in freshwater and marine receiving environments within Canada include Culp et al. (2000) and Dubé et al. (2002).

Selection of Appropriate Reference Areas

The procedure for choosing a reference area is based upon finding habitats that have not been exposed to the mine effluent under investigation and have natural habitat features similar to those of exposure areas (Reynoldson et al. 1997). Reference areas do not necessarily represent pristine (e.g., pre-European settlement) conditions, rather they are areas in which anthropogenic impacts are similar to exposure areas. Appropriate selection of reference areas is aided by classification schemes which classify river and lake ecosystems. An objective of these classification schemes is to predict the aquatic biotic assemblages that are likely to occur in a particular environment (Corkum 1992; Hughes 1995). Normally, the reference area is selected within the same drainage basin and upstream of the mine. However, when mine effluent discharges occur near the stream headwaters they may be selected in nearby drainage basins which have comparable habitat features.

Selecting Ecologically Relevant Habitats and Seasons

The choice of relevant habitat types for benthic invertebrate sampling is determined by an assessment of site-specific characteristics. Although, more than one habitat may be ecologically relevant, it is recommended (at least for Initial Monitoring) that only one habitat be sampled. This recommendation is similar to other biomonitoring studies, where the sampling of a single habitat type is founded on the fact that, given similar sampling efforts, variance is higher when several natural substrates are sampled compared to focusing the sampling effort on one habitat type (Plafkin et al. 1989; Environment Canada 1998). For the MM EEM program, determination of the most ecologically relevant habitat requires consideration of the following questions:

1. Which habitat type is present in the highest proportion in the exposure area?
2. In the absence of human influences, which habitat supports the richest assemblages of invertebrates within the study area? In addition, which habitat is likely to detect deviations from the reference condition?
3. In which habitat are the invertebrates most likely to be exposed to sediment or waterborne contaminants for extended periods of time?
4. Is historical information available for a particular habitat?

Note that in determining ecological relevance, the sampling of highly sensitive and diverse habitats can override the importance of sampling the spatially dominant habitat. Stream riffles can support a diverse assemblage of species that display a wide range of sensitivities to water-quality changes. Therefore, the community in this habitat has the potential for greater change than less species-rich communities. In contrast, the fauna of depositional areas, which are generally of lower taxonomic richness, are of interest during biomonitoring exercises because they may be directly exposed to concentrations of sediment-borne con-

taminants for long periods. Finally, if there is extensive historic data available for a particular habitat, consideration should be given to maintaining and extending this data set for the watershed as a valuable part of the EEM study.

In northern aquatic ecosystems, biological diversity often varies with season and the preferred sampling season for an effects-based monitoring program is at the time of maximum biological diversity. In general, diversity corresponds with the seasonal recruitment cycles of benthic organisms that are related to climate and food supply. An understanding of the seasonal patterns and life cycles of the taxa, along with changes in the hydrologic regime (e.g., high flows, ice scour) found within the specific system, is helpful to determine the appropriate timing for the survey. Sampling during periods when no effluent is discharged should be avoided. For many lotic habitat types, sampling is often conducted during the fall (September/October) when flow regimes allow access for sampling and the majority of taxa are present and/or large enough to be collected by the sampling equipment. In large lakes where the benthic community is often dominated by annelids, crustaceans, and molluscs, insect emergence periods and hydrologic regimes are of less importance for determining the sampling period. The basic EEM objectives are addressed by sampling in a single season, but additional seasons may be warranted in order to determine the magnitude of an effect.

Selecting Supporting Environmental Variables

Supporting environmental variables are useful in providing explanatory evidence for data interpretation. In addition to the general information obtained from maps and historic data during Site Characterization, appropriate environmental supporting variables are measured during the benthic survey and, if possible, a subset of variables measured more frequently (e.g., continuous measurement of water temperature with data loggers). These supporting environmental variables are divided into exposure variables that provide an indication of the extent of effluent exposure, and explanatory variables which assist in the interpretation of the monitoring results and allow investigators to determine if natural environmental variables influence, or are related to, the observed benthic invertebrate community structure. For the MM EEM program, exposure variables are selected from the list of water-quality parameters recommended in Parker and Dumaresq (2002). These exposure variables should be appropriate for site-specific conditions and measured within each sampling area at the time of the benthic survey.

Explanatory variables should include appropriate, site-specific variables from the following categories: physical-chemical characteristics, morphometric and flow characteristics, substratum composition, and riparian zone condition. Sources of information on selecting habitat descriptors in each of these categories include Corkum (1992) and Environment Canada (2001). These variables should be measured quantitatively at each replicate station during the benthic survey in order to

relate environmental change with invertebrate community structure. Although this level of sampling effort for explanatory variables is important for all MM EEM phases and study design options, it is particularly important for the gradient and reference condition approaches. Additional supporting environmental variables could include biotic indicators of effects other than benthic invertebrates (e.g., periphyton biomass). The ecological relevance of sampling for other indicators such as periphyton, phytoplankton, macrophyte or zooplankton biomass and taxonomic composition is determined on a site-specific basis.

Standardizing Methods

The Benthic Community Expert Working Group also recommended that future EEM programs standardize sampling methodologies (Environment Canada 1997). These recommendations were in response to the diverse methods used in the EEM studies from the first cycle in Pulp and Paper. For example, a range of sampling devices (e.g., four types and various sizes of grabs) were used, sometimes within the same habitat and study, mesh sizes of capture nets ranged from 180 to 1000 mm.

Field Sampling

Two major considerations for standardizing methods for benthic invertebrate surveys are mesh size and quantitative sampling equipment. In fresh water, macroinvertebrates are defined as those retained by mesh sizes of 200 to 500 μm , although immature life stages of some taxa may be smaller and some adult life stages may be larger. The recommendation for sieve and/or mesh size for all freshwater MM EEM programs is 500 mm. However, in site-specific circumstances it may be desirable for the field samples to be screened for smaller organisms by using a smaller sieve size for comparison with historic benthic invertebrate data. Quantitative sampling of benthic communities is carried out using devices that sample a known area or volume of habitat, such as grab samplers or stream net samplers. Each sampling device should be non-selective and suitable for a particular substrate condition. Standardization of techniques does not only apply to the actual samplers but to the level of expertise required in using the sampler. A wide range of quantitative sampling techniques is available for the collection of benthic invertebrates in lotic and lentic habitats (Elliott and Tullett 1983; Scrimgeour et al. 1993). The use of artificial substrates for benthic invertebrate collections was not recommended as a standard sampling protocol for the MM EEM program (Environment Canada 2002) but their use may be considered when there is a long history of using artificial substrates, or when extreme habitat conditions (e.g., very deep, fast water systems) prohibit the application of standard protocols.

Laboratory Methods

Standardized laboratory protocols are essential to the success of any monitoring program and study-specific sorting and sub-sampling proce-

dures should be determined *a priori*. Schematic diagrams that list the sequence of processing steps are often useful (Environment Canada 2002). For example, to facilitate the sorting process, benthic samples can be divided into appropriate size fractions (i.e., coarse >1.00 mm and fine 500 μm –1.00 mm). Verification of sorting efficiency is performed on a spot-check basis if the debris from a sample is retained. For the MM EEM program at least 10% of all samples must be re-sorted, and the criteria for an acceptable sort of animals is that $\leq 10\%$ of the total number of organisms are missed. If >10% of the total number is found during verification sorting, the re-sort, then all the samples within that group of samples requires re-sorting (Environment Canada 2002). Furthermore, large samples or samples with excessive numbers of invertebrates may require quantitative sub-sampling in the laboratory. Various methods of laboratory sub-sampling are available and are outlined in Wrona et al. (1982) and Marchant (1989). The method used should be documented thoroughly and the error associated with the sub-sampling technique estimated by sorting all sub-samples (i.e., the entire sample) for at least 10% of all samples.

Taxonomic Level of Identification

The necessary level of taxonomic resolution depends on the purpose of a study. Accurate species-level identification is necessary if the physiology, toxicological response, population dynamics or secondary production for a particular species or group of species is of interest. However, there are a variety of methods used to evaluate a test site in a bioassessment, and therefore a variety of approaches are needed to answer the question regarding level of taxonomic resolution. The relevant questions to be considered in determining the necessary taxonomic resolution for a particular bioassessment study include:

1. Does genus/species identification add significant information to our description of variation among benthic macroinvertebrate communities?
2. Is a test site's deviation from reference condition easier to detect with genus/species identification?
3. Are there sufficient resources available to ensure that sample maintenance and data quality are sufficient?

In some cases, species sensitivities may make them good indicator candidates (Lenat and Resh 2001). At the community level, where the estimated tolerances of genera or species to some specific or general stressor is often integrated into an index (e.g., Hilsenhoff 1987), the value of genus/species identification is less important. Bailey et al. (2001) have shown that when using tolerance values and calculating indices of pollution effects on ecosystems, genus/species identification does not appear to add much information to the analysis. Repeated analysis of benthic macroinvertebrate community data has shown little change in a multivariate description of community variation at taxonomic levels

from genus to order (Furse et al. 1984; Marchant et al. 1995; Bowman and Bailey 1998). Thus, in most cases the multivariate characterization of community composition is not sensitive to taxonomic resolution when invertebrates are identified to the level of family (Hawkins and Norris 2000; Hewlett 2000).

While the recommended level of taxonomic identification for freshwater invertebrates is family for Initial and Periodic Monitoring phases of the MM EEM program, there may be site-specific conditions that warrant a lower taxonomic level for some or all familial groups. For example, if historical benthic invertebrate information has been identified to lower taxonomic levels, it may be desirable to identify subsequent surveys to a similar level for comparative purposes. Focused Monitoring may require different levels of taxonomic resolution in order to determine the geographical extent or the magnitude of the effect. For example, a detailed determination of the magnitude of the effect on sensitive taxa could be investigated in Focused Monitoring (i.e., how many genera or species within a sensitive family are affected?). Finally, for comparative purposes and quality control of taxonomic identification, the MM EEM program recommends the maintenance of reference collections of organisms and the verification of identification for each taxon in the collection by a professional taxonomist. Reference collections are beneficial for confirming identifications, ensuring consistent taxonomy between surveys, and in the training of personnel. In addition, consideration is also being given to national voucher collections, confirmation of identification by experts, and a national certification program for taxonomists.

Statistical Considerations

Statistical considerations applicable for each phase of the MM EEM program include the following three sequential steps: 1) selection of the effect size for the required power analysis, 2) selection of α and β levels for the power analysis (the same levels are also to be used during the statistical analyses on the compiled monitoring data), and 3) a power analysis to determine the appropriate level of sampling effort at the station scale (i.e., number of stations per area).

Setting Effect Size

A critical aspect of MM EEM for mines is the decision on the magnitude of an environmental impact that is deemed to be ecologically important (Mapstone 1995). An estimate of this effect size (ES) must be decided upon before a monitoring program is designed (i.e., *a priori* determination) in order to determine the sample size to have sufficiently high statistical power to detect an impact equal to or greater than the ES. Quantitative guidelines providing ES values have not been widely developed for ecological communities, such as benthic invertebrate assemblages in marine or freshwater systems (Environmental Protection Agency 1990). In large part, the lack of quantitative guide-

lines has resulted from the complex nature of aquatic communities and the difficulty in making quantitative predictions of the expected magnitude of change following an impact, such as might result from the discharge of mining effluents. Lowell (1997) discusses how ES values for EEM programs can be determined. The three ES options discussed here include:

1. ES Option 1: *a priori* setting of ES as a fixed percent change of the assessment endpoint based on existing ecological information on the magnitude of an effect which is known to have long-lasting deleterious effects on the aquatic ecosystem;
2. ES Option 2: *a priori* determination of a standardized, habitat-specific ES calculated from previous data obtained from appropriate reference areas within the same ecoregion; and
3. ES Option 3: *a priori* setting of ES at ± 2 SD calculated across stations from reference area data.

Option 1 would be the preferred choice when extensive background data is available. Extensive ecological data is needed to determine the ecological meaning of a change in community composition. Unfortunately, this ecological approach to setting ES is likely not suitable unless extensive, high-quality historical data are available. Alternatively, if sufficient ecoregion-specific data are available, Option 2 is recommended. This option involves calculating the normal range of between station variability (i.e., ± 2 SD) for each reference data set available within the ecoregion. The median of all the calculated values is then used as a habitat-specific ES for the ecoregion (see Lowell 1997). The median is used instead of the mean to minimize the influence of outlier (non-typical) reference areas on the calculation. This procedure yields a fixed value for ES (e.g., $\pm 30\%$ change in the number of taxa from reference), which may or may not be equal to ± 2 SD for any particular reference area used in the calculation. In the absence of any background information, it is recommended that ES initially be set as any change in a measured benthic descriptor that is ± 2 SD from reference (Option 3). This suggested approach defines ES in terms of the normal range of variability observed in the reference community at the mine site. Assuming the data are normally distributed, a commonly accepted criterion for the normal reference condition is ± 2 standard deviations (± 2 SD) of the reference station values about the reference area mean.

Selecting α and β

Type I and Type II errors must be considered in statistical evaluations of EEM data. **Type I error** occurs (at probability α) if the null hypothesis that there is no effect is rejected when in fact it is true (i.e., a site is declared as being different to reference when it is not). **Type II error** occurs (at probability β) if the null hypothesis is accepted when it is false (i.e., the site is declared as being the same as reference when it is actually

impaired). The **power** of a statistical test is defined as $1-\beta$, the probability associated with correctly rejecting the null hypothesis when it is false. Therefore, α is the risk to industry, and β is the risk to the environment.

To determine the sampling effort necessary to detect the recommended ES, decisions must be made about the magnitude of Type I and Type II errors that are acceptable. In testing whether an exposed community falls outside the pre-determined guidelines for a reference community, a low probability of a Type I error (α) is usually allowed so that a reference community will not be mistaken for an impacted one. Traditionally, α has been set at 0.05 but there is little literature or consensus on what is an appropriate level for β (Lowell 1997). However, the monitoring program should be designed to provide a reasonably high probability of detecting a predetermined ES if it has occurred. In other words, the power of the test should be high. Recently, because of the potentially high cost (both ecological and monetary) of failing to detect environmental impacts, many benthic impact researchers argue that α should be set at least to the same level as β (Underwood 1993; Mapstone 1995). In designing the MM EEM program, we considered it important that α and β be set to the same level, recommended that α and β be set equal and at 0.10 or less. Although these levels of α or β are discussed in relation to the power analysis exercise, they would equally apply to the statistical analysis of the results from the monitoring survey for determination of effects.

Power Analysis to Determine Sampling Effort for Stations

The number of stations and their allocation between discreet reference and exposure areas, or along a gradient, must be determined such that the sampling design will have the power to satisfy the MM EEM monitoring requirements and provide adequate data for decision-making for each mine site. The allocation and distribution of these replicates is dependent on the site-specific conditions and the sampling design used. Power analysis is used to determine the sample size necessary to achieve a test adequate to detect an impact equal to or greater than ES, where ES has been determined prior to sampling. For ES options 1 and 2, a scientifically defensible sampling strategy can be devised using ES, α and β , together with an estimate of reference SD and assumptions about the distribution of the data. For the basic C/I ANOVA design, the estimated sample size to detect a change of ± 2 SD at a given level of power can be conveniently calculated by arranging the standard power analysis equation as follows (Green 1989):

$$n = 2(t_{\alpha} + t_{\beta})^2 (SD/ES)^2 \quad (1)$$

where n is the sample size (i.e., replicate stations); t_{α} is t value for a significance level (Type I error; two-tailed); t_{β} is t value for β significance level (Type II error; one-tailed); SD is standard deviation; and ES is effect size.

For ES Option 3, both SD and ES cancel out of the power analysis equation (see Lowell 1997) because $ES = 2$ SD. The power analysis equa-

tion can then be modified by expressing ES in terms of d , the number of standard deviations separating the reference and impact means ($d = ES/SD$), and then substituting to yield (assuming equal reference and exposure variances):

$$n = 2(t_{\alpha} + t_{\beta})^2 / d^2 \quad (2)$$

Using these power analysis equations, the various sample sizes required to detect effects of ± 2 SD for α and β at different levels can be determined. Pre-calculated tables of n are available for a variety of values of α , β and SD (Allredge 1987; Environment Canada 1998). Thus, with the recommended case scenario of $\alpha = \beta = 0.1$ and $ES = 2$ SD, the sampling effort will require a sample size of five which is within the range used in many benthic surveys (Resh and McElvray 1993). Similar considerations and procedures apply to power analysis and determining sample size for regression and ANCOVA monitoring designs (simple, radial, and multiple gradient designs) (Lowell et al. 2002; Environment Canada 2002).

The issue of replication is somewhat different when using the RCA. Replication is at the station scale and it has been shown that variation within a station is much less than among stations (Reynoldson et al. 1997). Thus, single composite samples are taken at stations and variation among stations is used to describe the reference condition. The number of reference replicates is determined by the number of stations in the group to which the exposed station is predicted to belong using the RCA. This is determined when forming the groups of reference stations in the initial classification, but has been set to a minimum of 10 stations. The variation among the reference stations forming the reference group determines the Type I error which has been set at 0.1 by using a 90% probability ellipse. Because this approach compares single exposure stations to multiple reference stations (minimum of 10), it is not possible to set Type II error which requires an estimate of the variance associated with a single station. A surrogate can be applied by taking more than one sample at the exposed station, but this is estimating within-station error rather than the appropriate variation at the among-station level. Clearly Type II station error cannot be determined when there is only one member of the population of exposed stations. Therefore, the power analyses referred to above would not be applicable for the RCA study design. Consequently, RCA studies should be designed in a way that provides an accurate and precise determination of reference conditions so as to maximize the likelihood of detecting departures from reference conditions at exposure stations, when they exist.

Data Assessment and Interpretation

Benthic Invertebrate Community Endpoints

The objective of the benthic invertebrate component of the MM EEM program is to determine if the mine effluent has modified the benthic

invertebrate community (i.e., is there an effect). An effect on the benthic invertebrate community is defined as a statistical difference between reference and exposure areas for any of the following key descriptors:

1. Total invertebrate density;
2. Taxon (i.e., Family) Richness;
3. Simpson's Diversity Index; and
4. Bray-Curtis Index.

Evenness, taxon density, proportion or presence/absence are recommended to allow for the interpretation of effects (Environment Canada 2002). These descriptors are summary metrics selected to encompass the range of responses which may result from mine effluent, including changes in productivity, species composition and biodiversity. The Bray-Curtis *similarity index* is recommended as it summarizes the overall difference in community structure between reference and exposed sites in a single number and requires no preconceived assumptions about the nature of the community (Taylor and Bailey 1997). Many other benthic invertebrate descriptive metrics are available in the literature (for a review see Resh and McElravy 1993) and may be used, if applicable, on a site-specific basis to aid in the interpretation of effects determined with the key descriptors listed above.

Data Assessment Decision Tree:

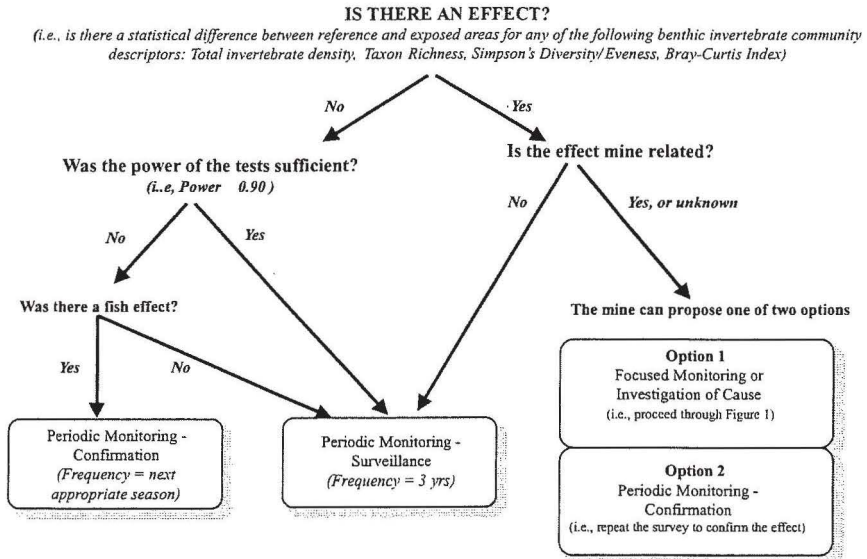
What is the Next Phase in the MM EMM?

Once the monitoring data has been analyzed, decisions regarding the next step in the MM EEM program are required. For this tiered monitoring program a decision tree (Fig. 4) was developed to aid in standardizing this decision process. The key factors in the decision tree include not only the statistical outcome from benthic surveys, but also results from fish surveys and changes in environmental or mine operating conditions. The multi-level approach encompassed by this decision-making framework allows for an evaluation of the effects of mining effluent on the benthic invertebrate community in concert with other key aspects of the aquatic ecosystem. The decision tree is divided into two sections, one which is applicable for use with monitoring data from the Initial Monitoring phase or appropriate historic data (Fig. 4A), and a second tree which is applicable if more than one MM EEM survey has been conducted (Fig. 4B). The four key factors which contribute to decision-making within this framework for benthic invertebrate community assessments are listed and discussed in detail below.

1. The statistical outcome of the benthic invertebrate survey:

The objective of the benthic invertebrate component of the MM EEM program is to determine if the mine effluent has modified the benthic invertebrate community (i.e., is there an effect). An effect on the benthic invertebrate community has been defined as a statistical difference

First Benthic Invertebrate Survey Results
(or acceptable historical data)



Second or Subsequent Benthic Invertebrate Survey Results

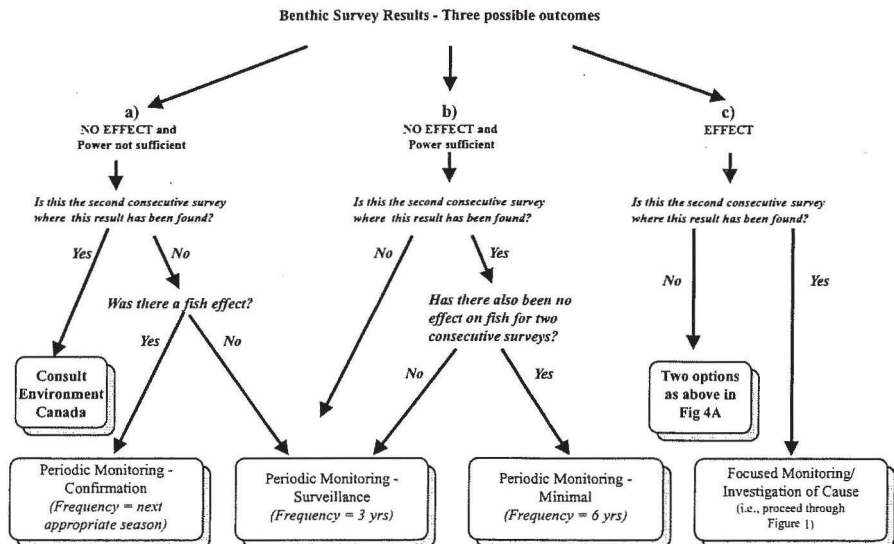


Fig. 4. Decision trees for data assessment. Two decision trees are provided for use with data from: A) The First Benthic Invertebrate Survey (or historic data), or B) The Second or Subsequent Surveys.

between reference and exposure areas (Environment Canada 2002) for any of the key benthic invertebrate community descriptors. If *no effect* is observed, a further caveat for this case includes a demonstration of sufficient power. For the statistical analyses, and determination of sufficient power, the recommendations for the setting of α and β discussed previously are applicable (i.e., setting α and β equal, at 0.10 or less). Therefore, there are three possible outcomes from the benthic invertebrate survey: (1) no effect is seen but power is not sufficient (i.e., power < 0.90), (2) no effect is seen and power is sufficient (i.e., power \geq 0.90), or (3) an effect is seen. After determination of the statistical outcome from the benthic survey, the additional factors listed below are included in the decision making framework.

2. Results from MM EEM fish surveys:

An effect in the MM EEM fish survey is defined in a similar manner to the benthic invertebrate component; a statistical difference for defined endpoints between reference and exposure areas (for details on the MM EEM fish component see Environment Canada 2001 and Ribey et al. 2002). The three possible outcomes from the MM EEM fish survey which need to be considered during the determination of the next monitoring phase for the benthic invertebrate component include a fish effect is seen, no fish effect is seen in one survey and no fish effect is seen in *two or more* consecutive surveys. These outcomes from the fish survey are used to aid in making decisions regarding the benthic monitoring program when (1) benthic invertebrate statistical results are inconclusive (i.e., insufficient power), or (2) when decisions are required regarding whether the MM EEM program should proceed to a Minimal Monitoring frequency (i.e., 6-year frequency) (Fig. 4).

3. Changes in condition of the environment or mining operation:

Effluent characterization and water-quality monitoring continues throughout the MM EEM program (see Parker and Dumaresq 2002 for details). This information can be utilized to detect whether any changes are occurring in either the mine operating conditions (i.e., effluent characteristics) or the natural environmental conditions (in either the reference or exposure areas). These changes are relevant if the mine is in the Minimal Monitoring frequency. The mine is required to move from a Minimal to Confirmation Monitoring frequency in the next appropriate sampling season if either the environmental conditions or the mine operating conditions change in a manner which may cause an increased risk of detrimental impacts to the benthic community.

4. MM EEM program options after an effect on the benthic invertebrate community has been established:

As indicated in the Data Assessment decision tree (Fig. 4), if a statistical difference in the measurement endpoints is detected between refer-

ence and exposure areas, then the MM EEM program proceeds (1) through the overall MM EEM Decision tree (Fig. 1; Environment Canada 2002; Dumaresq et al. 2002) to either Focused Monitoring or Investigation of Cause, or (2) to the Periodic Monitoring—Confirmation phase. If, after an effect is observed, the mine proceeds to this phase, the results of the new survey are evaluated with the Data Assessment decision tree. If an effect on the benthic community is confirmed, then the mine *must* proceed to the next phase determined through Fig. 1. The series of questions are laid out in the overall decision tree, more detailed questions are included below as a guide to the intent of the program with regards to the benthic invertebrate community assessments.

Is the effect mine related?

An assessment of whether the effect is mine-related could include the following questions. Is the cause of the effect known or suspected? Can the effect be related to a natural change in the aquatic receiving environment? Can the effect be reasonably correlated to an anthropogenic cause, other than the mine effluent? Is there a weight-of-evidence approach which can indicate a causal link? If the cause of the effect is unknown, and there are no confounding factors which could potentially cause the effect, then the mine should proceed to the next step in Data Assessment and Interpretation, determining the magnitude and geographic extent of the effect.

Is the magnitude and geographical extent known?

An assessment of the magnitude of the effect would include the following questions. How many taxonomic groups are affected? What is the magnitude (e.g., the amount of change in abundance) of the effect on the taxonomic group(s) affected? Is there an effect in other benthic invertebrate community members, such as periphyton and macrophytes, which are present in the reference area and would be expected to be present in the exposure area? An assessment of the geographical extent of the effect would include consideration of questions such as; what was the geographical area affected and were the benthic invertebrate communities at the sampling stations furthest from the effluent discharge similar to reference conditions?

Is the mine-related cause known?

An assessment of the cause of the mine-related effect would include the following questions. What are the possible mine-related causal agents? Is there a correlation between the type of effect observed on the benthic invertebrate community and a mine-related causal agent? Is there a weight-of-evidence approach which can indicate a causal link?

Conclusions and Summary

Canada's Metal Mining EEM program for assessments of benthic invertebrate communities has been developed over the previous three

years with participation of an expert working group and has drawn from existing programs within Canada including the AETE, Aquamin and the Pulp and Paper EEM. The primary objective of the program is to evaluate the effectiveness of current pollution regulations with a nationally consistent, scientifically defensible program, while considering the diverse nature of the mining industry in Canada. The program proceeds in a tiered manner, commencing with determining whether an effect is present and continuing with determining extent, magnitude and cause of the effect. This program is unique in that the decision frameworks are based on interpretation of effects on biological endpoints, and that results from several trophic levels and physical parameters are considered in a weight-of-evidence type of approach. Program implementation is proceeding as outlined in Dumaresq et al. (2002) and future modifications to monitoring programs within Canada will likely follow the guidelines and frameworks developed within the MM EEM program. In fact, much of the information presented here and elsewhere (Environment Canada 2001) is widely applicable to the development of other benthic invertebrate monitoring programs.

Acknowledgments

Financial support for this project was provided by the National EEM Office and the National Water Research Institute of Environment Canada. We thank the additional members of the MM EEM Benthic Subgroup and their affiliations for their commitment to the development of this national program, including: Mark Wiseman (industry co-chair, Falconbridge Limited), Glen Watson (Inco Limited), Yves Couillard and Bernadette Pinel-Alloul (Université de Montréal), Bill Keller (Ontario Ministry of Northern Development and Mines), Hélène Dupuis (Fisheries and Oceans Canada) and Karen Mailhot, Nardia Ali, Uwe Borgmann and John Lawrence (Environment Canada).

References

- Aldredge JR. 1987. Sample size for monitoring of toxic chemical sites. *Environ. Monit. Assess.* 9:143–154.
- Aquamin. 1996. Assessment of the aquatic effects of mining in Canada: AQUAMIN final report. April 1996. Aquamin Working Groups 7 & 8. Environment Canada, Ottawa, Ontario.
- Bailey RC, Norris RH, Reynoldson TB. 2001. Taxonomic resolution of benthic macroinvertebrate communities in bioassessments. *J. N. Amer. Benthol. Soc.* 20:280–286.
- Barbour MT, Gerritsen J, Griffith GE, Frydenborg R, McCarron E, White JS, Bastian ML. 1996. A framework for biological criteria for Florida streams using benthic macroinvertebrates. *J. N. Amer. Benthol. Soc.* 15:185–211.
- Barbour MT, Gerritsen J, Snyder BD, Stribling JB. 1999. Rapid bioassessment protocols for use in streams and Wadeable rivers: periphyton, benthic macroinvertebrates and fish, second edition. EPA 841-B-99-002. U.S. Environmental Protection Agency, Office of Water, Washington, D.C.

- Beak.** 1996. 1995 field evaluation of aquatic effects monitoring methods, pilot study. AETE project 4.1.1 Natural Resources Canada, CANMET, Ottawa, Ont.
- Beak.** 1999. Quality assurance program for assessing mine-related effects using benthic invertebrate communities. AETE project 2.1.4 Natural Resources Canada, CANMET, Ottawa, Ont.
- Bowman MF, Bailey RC.** 1998. Does taxonomic resolution affect the multivariate description of the structure of freshwater benthic macroinvertebrate communities? *Can. J. Fish. Aquat. Sci.* **54**:1802–1807.
- Corkum LD.** 1992. Spatial distributional patterns of macroinvertebrates along rivers within and among biomes. *Hydrobiologia* **239**:101–114.
- Culp JM, Halliwell DB.** 1999. Volunteer-based monitoring program using benthic indicator to assess stream health. NWRI, Environment Canada, Saskatoon, Saskatchewan.
- Culp JM, Lowell RB, Cash KJ.** 2000. Integrating mesocosm experiments with field and laboratory studies to generate weight-of-evidence risk assessments for large rivers. *Environ. Toxicol. Chem.* **19**:1167–1173.
- Davies PE.** 2000. Development of a national river bioassessment system (AUSRI-VAS) in Australia, p. 113–124. *In* Wright JF, Sutcliffe DW, Furse MT (ed.), *Assessing the biological quality of fresh waters: RIVPACS and other techniques*. Freshwater Biological Association, Ambleside, U.K.
- Dubé M, MacLatchy D, Culp J, Gillis G, Parker R, Courtenay S, Gilman C.** 2001. Utility of mobile, field-based artificial streams for assessing effects of pulp mill effluents on fish in the Canadian environmental effects monitoring program. *J. Aquat. Ecosyst. Stress Recovery*, In press.
- Dubé MG, Culp JM, Cash KJ, Glozier NE, MacLatchy DL, Podemski CL, Lowell RB.** 2002. Artificial streams for environmental effects monitoring (EEM): development and application in Canada over the past decade. *Water Qual. Res. J. Canada* **37**(1):155–180.
- Dumaresq C, Hedley K, Michelutti R.** 2002. Overview of the metal mining environmental effects monitoring program. *Water Qual. Res. J. Canada* **37**(1): 213–218.
- Elliott JM, Tullett PA.** 1983. A supplement to a bibliography of samplers for benthic macroinvertebrates. *Freshwat. Biol. Assoc. Occas. Publ.*, no. 20.
- Environment Canada.** 1997. Benthic invertebrate community expert working group final report: recommendations from cycle 1 review. EEM/1997/7. Environment Canada, National EEM Office, Science Policy and Environmental Quality Branch, Ottawa, Ont.
- Environment Canada.** 1998. Pulp and paper technical guidance for aquatic environmental effects monitoring. EEM/1998/1. Environment Canada, National EEM Office, Science Policy and Environmental Quality Branch, Ottawa, Ont.
- Environment Canada.** 2002. Metal mining guidance document for aquatic environmental effects monitoring. www.ec.gc.ca/eem. Environment Canada, National EEM Office, Science Policy and Environmental Quality Branch, Ottawa, Ont.
- Environmental Protection Agency.** 1990. Biological criteria: national program guidance for surface waters. EPA-440/5-90-004. US Environmental Protection Agency, Washington, D.C.
- Environmental Protection Agency.** 1995. Volunteer stream monitoring: a methods manual. EPA 841 D95-001 April. Office of Wetland, Oceans and Watersheds, 4503F. Washington, D.C.
- Faith DP, Dostine PL, Humphrey CL.** 1995. Detection of mining impacts on aquatic macroinvertebrate communities: results of a disturbance experiment and the design of a multivariate BACIP monitoring programme at Coronation Hill, Northern Territory. *Aust. J. Ecol.* **20**:167–180.

- Furse MT, Moss D, Wright JE, Armitage PD. 1984. The influence of seasonal and taxonomic factors on the ordination and classification of running-water sites in Great Britain and on the prediction of their macro-invertebrate communities. *Freshwat. Biol.* 14:257–280.
- Graney FL, Kennedy JJ, Rodgers JH (ed.). 1994. Aquatic mesocosm studies in ecological risk assessment. Special publication of SETAC Press, Pensacola, Florida, U.S.A.
- Green RH. 1989. Power analysis and practical strategies for environmental monitoring. *Environ. Res.* 50:195–205.
- Hawkins CP, Norris RH. 2000. Effects of taxonomic resolution and use of subsets of the fauna on the performance of RIVPACS-type models, p. 217–228. *In* Wright JE, Sutcliffe DW, Furse MT (ed.), *Assessing the biological quality of fresh waters: RIVPACS and other techniques*. Freshwater Biological Association, Ambleside, U.K.
- Hewlett R. 2000. Implications of taxonomic resolution and sample habitat for stream classification at a broad geographic scale. *J. N. Amer. Benthol. Soc.* 19:352–361.
- Hilsenhoff WL. 1987. An improved index of organic stream pollution. *Great Lakes Entomol.* 20:31–39.
- Hughes RM. 1995. Defining acceptable biological status by comparing with reference conditions, p. 31–47. *In* Davis WS, Simon TP (ed.), *Biological assessment and criteria: tools for water resource planning and decision making*. Lewis Publisher, Boca Raton.
- Lenat DR, Resh VH. 2001. Taxonomy and stream ecology—the benefits of genus- and species-level identifications. *J. N. Amer. Benthol. Soc.* 20:287–298.
- Lowell RB. 1997. Discussion paper on critical effect size guidelines for EEM using benthic invertebrate communities. EEM/1997/9. Environment Canada, National EEM Office, Science Policy and Environmental Quality Branch, Ottawa, Ont.
- Lowell RB, Hedley K, Porter E. 2002. Data interpretation issues for Canada's environmental effects monitoring program. *Water Qual. Res. J. Canada* 37(1): 101–117.
- Mapstone BD. 1995. Scalable decision criteria for environmental impact assessment: effect size, Type I and Type II errors. *Ecol. Appl.* 5:401–410.
- Marchant RA. 1989. A subsampler for samples of benthic invertebrates. *Bull. Aust. Soc. Limnol.* 12:49–52.
- Marchant R, Barmuta LA, Chessman BC. 1995. Influence of sample quantification and taxonomic resolution on the ordination of macroinvertebrate communities from running waters in Victoria, Australia. *Mar. Freshwat. Res.* 46:501–506.
- Messer JJ, Linthurst RA, Overton WS. 1991. An EPA program for monitoring ecological status and trends. *Environ. Monit. Assess.* 17:67–78.
- Nusser SM, Goebel JJ. 1997. The national resources inventory: a long-term multi-resource monitoring programme. *Environ. Ecol. Stat.* 4:181–204.
- Parker R, Dumaresq C. 2002. Effluent characterization, water quality monitoring and sediment monitoring in the metal mining EEM program. *Water Qual. Res. J. Canada* 37(1):219–228.
- Plafkin JL, Barbour MT, Porter KD. 1989. Rapid bioassessment protocols for use in streams and rivers: benthic macroinvertebrates and fish. EPA/444/4-89-001.
- Pontasch KW. 1995. The use of stream microcosms in multispecies testing, p. 169–191. *In* Cairns J, Niederlehner BR (ed.), *Ecological toxicity testing: scale, complexity and relevance*. Lewis Publishers, Boca Raton, Florida, U.S.A.
- Powell M. 1995. Building a national water quality monitoring program. *Environ. Sci. Tech.* 29:458A–463A.

- Resh VH, McElravy EP. 1993. Contemporary quantitative approaches to biomonitoring using benthic macroinvertebrates, p. 159–194. *In* Rosenberg DM, Resh VH (ed.), *Freshwater biomonitoring and benthic macroinvertebrates*. Chapman & Hall, New York, U.S.A.
- Reynoldson TB, Norris RH, Resh VH, Day KE, Rosenberg DM. 1997. The reference condition: comparison of multimetric and multivariate approaches to assess water-quality impairment using benthic invertebrates. *J. N. Amer. Benthol. Soc.* 16:833–852.
- Ribey SC, Munkittrick KR, McMaster ME, Courtenay S, Langlois C, Munger S, Rosaasen A, Whitley G. 2002. Development of a monitoring design for examining effects in wild fish associated with discharges from metal mines. *Water Qual. Res. J. Canada* 37(1):
- Roux DJ, Kleynhans CJ, Thirion C. 1999. Biological monitoring and assessment of rivers as a basis for identifying and prioritising river management options. *Water Sci. Technol.* 39:207–210.
- Scrimgeour GJ, Culp JM, Glozier NE. 1993. An improved technique for sampling lotic invertebrates. *Hydrobiologia* 254:65–71.
- Stevens DL. 1994. Implementation of a national monitoring program. *J. Environ. Manage.* 42:1–29.
- Taccogna G, Munro K (ed.). 1995. *The stream keepers handbook: a practical guide to stream and wetland care*. Salmonid Enhancement Program, Department of Fisheries and Oceans, Vancouver, B.C.
- Taylor BR. 1997. Optimization of field and laboratory methods for benthic invertebrate monitoring, final report. AETE project 2.1.2 Natural Resources Canada, CANMET, Ottawa, Ont.
- Taylor BR, Bailey RC. 1997. Technical evaluation on methods for benthic invertebrate data analysis and interpretation, final report. AETE project 2.1.3. Natural Resources Canada, CANMET, Ottawa, Ont.
- Underwood AJ. 1993. The mechanics of spatially replicated sampling programs to detect environmental impacts in a variable world. *Aust. J. Ecol.* 18:99–116.
- Walker SL, Hedley K, Porter E. 2002. Pulp and paper environmental effects monitoring in Canada: an overview. *Water Qual. Res. J. Canada* 37(1):7–19.
- Wright JF, Sutcliffe DW, Furse MT. 2000. Assessing the biological quality of fresh waters: RIVPACS and other techniques. *Freshwater Biological Association*. Ambleside, Cumbria, U.K.
- Wrona FJ, Culp JM, Davies RW. 1982. Macroinvertebrate subsampling: a simplified apparatus and approach. *Can. J. Fish. Aquat. Sci.* 39:1051–1054.