98-130

# Environment Canada
## Water Science and Technology Directorate

## Direction générale des sciences et de la technologie, eau
# Environnement Canada

The reference condition: a comparison of multimetric and multivariate approaches to assess water-quality impairment using benthic invertebrates

By:
T.B. Reynoldson, K.E. Day, T. Pascoe

98-130

# MANAGEMENT PERSPECTIVE

Title: The reference condition: a comparison of multimetric and multivariate approaches to assess water-quality impairment using benthic invertebrates

Authors: T.B. Reynoldson, K.E. Day and T. Pascoe

NWRI Publication No: 98-130

Citation: Proceedings of the RIVPACS workshop, published by Freshwater Biological Assos, U.K.

EC Priority/Issue: The purpose of environmental assessment and management is the maintenance of biological integrity; thus, setting of water and sediment quality objectives should consider the use of biological criteria as well as chemical surrogates. The objective of this study was to develop biological guidelines for sediment assessment in the Great Lakes based on invertebrate community structure and laboratory responses in survival, growth and reproduction of benthic invertebrates. These guidelines can be used for assessment of sediment in harbours and embayments for possible remediation or for assessment of sediment removed in navigational or other dredging projects.

Current status: A total of 345 site visits to 245 different locations were made over the study period (1991-93). A total of 162 taxa (genus or species level) were identified. Using cluster analysis and ordination six groups of sites supporting different communities were identified. Using 11 habitat variables a discriminant model was constructed that predicted 88% of a set of test sites to the correct community group. This predictive model allowed the six community types to be used as the community structure guidelines. Test sites being compared to reference sites in multivariate ordination space. From the reference site data three toxicological categories were established for each of 10 test endpoints, in four species. The categories were - nontoxic, warning of potential toxicity and toxicity. The delineations for the three categories were developed from the standard statistical parameters of population mean and standard deviation of each endpoint measured in all reference sediments.

Next steps: The NWRI is finishing the development of management software that can be provided to regional managewrs to allow them to apply the guidelines. This software will allow input of biological and environmental data and provide site specific assessment of the status of the invertebrate community and the toxicity of the sediment.

# The development of the BEAST: a predictive approach for assessing sediment quality in the Great Lakes

TREFOR B. REYNOLDSON, KRISTIN E. DAY, AND T. PASCOE

*National Water Research Institute, Environment Canada, CCIW,*

*867 Lakeshore Rd., Burlington, Ontario, Canada, L7R 4A6*

## Summary

The purpose of environmental assessment and management is ultimately the maintenance of biological integrity; thus, we suggest that the setting of water and sediment quality objectives should involve the use of biological criteria rather than chemical surrogates. The objective of this study was to develop biological guidelines using community structure together with laboratory responses in survival, growth and reproduction of benthic invertebrates. These guidelines can be used for assessment of sediment in harbours and embayments for possible remediation or for assessment of sediment removed in navigational or other dredging projects. A total of 345 site visits were made over the study period (1991-93). A total of 245 different locations were visited. A total of 162 taxa (genus or species level) were identified. Using cluster analysis and ordination six groups of sites were identified that support different communities. Using 11 habitat variables a discriminant model was constructed that predicted 88% of the test sites to the correct community group. This predictive model allowed the six

community types to be used as the community structure guidelines. Test sites being compared to reference sites in multivariate ordination space. From the reference site data three toxicological categories were established for each of 10 test endpoints, in four species. The categories were were - nontoxic, warning of potential toxicity and toxicity. The delineations for the three categories were developed from the standard statistical parameters of population mean and standard deviation ($x \pm$ S.D.) of each endpoint measured in all reference sediments.

## Introduction

Environmental managers and regulatory decision-makers have traditionally set water and sediment quality guidelines based on chemical concentration. The primary advantage of a chemical approach is the apparent ease of simple numerical comparison. Concentrations of chemicals found in environmental matrices are compared with levels of these same compounds suggested to cause a toxic response in biota. However, the chemical approach has been criticized in recent years because it frequently fails to achieve its objectives (Cairns and van der Schalie 1980, Long and Chapman 1985, Chapman 1986, Chapman 1990) or because it is so excessively rigorous that it has limited value (Painter 1992, Zarull & Reynoldson 1992).

The purpose of environmental assessment and management is ultimately the maintenance of biological integrity; thus, we suggest that the setting of water and sediment quality objectives should involve the use of biological criteria rather than chemical surrogates. Until recently, the development of numeric biological objectives was considered too difficult due to the temporal and spatial variability inherent in biological systems. However, over the past 10 years, multivariate methods developed in the United Kingdom (Wright *et al.* 1984, Moss *et al.* 1987, Armitage *et al.* 1987, Ormerod and Edwards 1987) and elsewhere (Corkum and Currie 1987, Johnson and Wiederholm 1989, Parsons and Norris 1996, Reynoldson *et al.* 1995) have demonstrated the ability to predict the community structure of benthic invertebrates in clean (or 'uncontaminated') sites using simple habitat and water quality descriptors. This has been described as the *reference condition approach* (Reynoldson *et al.* 1997) and allows appropriate site-specific biological objectives to be set for ecosystems from measured habitat characteristics and also provides an appropriate reference for determining when degradation at a site due to anthropogenic contamination is occurring. The acceptance by regulatory agencies of biological water and sediment quality objectives has been slow but is now being given serious consideration as shown by current work in Canada (Reynoldson and Zarull 1993, Reynoldson *et al.* 1995), the U.S.A. (Hunsaker and Carpenter 1990, Canfield *et al.* 1996, Besser *et al.* 1996) and the United Kingdom (Wright 1995) and recent initiatives in Australia (Parsons and Norris 1996).

Criticisms of the complexity of the multivariate methods used in the reference condition approach have been made (Gerritsen 1995). There is validity to these criticisms in that application of the methods for routine site assessment can be difficult. Consequently, in three of the areas where large reference databases have been developed (Australia, Canada and the UK), complementary software has also been developed to make the complex analyses routinely available to environmental managers. In the United Kingdom the RIVPACS software is in a third iteration (Wright 1995). In Australia the

AusRivAS software has just recently become available (http://enterprise.canberra.edu.au/AusRivAS.nsf). This paper describes the development of the BEAST software (http://www.cciw.ca//nwri-e/aerb/sediment-remediation/BEAST/) based on a large data base that has been assembled from reference sites in the Laurentian Great Lakes.

## Methods

The fundamental concept behind the *reference condition approach* is to establish a data base of sites that represents unimpaired conditions (reference sites) at which biological and environmental attributes are measured. The data-base is then used to:

1. Examine and classify the normal variability in biological response, in the case of the Great Lakes invertebrate community structure and sediment toxicity.

2. Develop methods to relate the variability in biological response to environmental attributes at reference sites.

3. Develop methods to compare the observed biological conditions at a test site with the expected conditions based on the reference site data base.

### *Selecting Reference Sites*

The objective of this study was to develop biological guidelines to be used for assessment of sediment in harbours and embayments for possible remediation or for assessment of sediment removed in navigational or other dredging projects. Therefore, the selection criteria for the Great Lakes reference sites were that they should: capture the entire range of biological variability expressed by the benthic invertebrate communities of the Great Lakes; represent fine grained sediment environments; be located in shallow nearshore areas, and; the sites should be unimpacted or clean. Seventeen ecodistricts were used to stratify sample sites. Within each ecodistrict hydrographic charts were used to identify areas having fine grained sediment material and boundaries were drawn around those areas indicated on the hydrographic charts as having either a silt or mud substratum. These were identified as regions where sites could be potentially located. Initially sites were restricted to a maximum 30m water depth limit and within 2 km of shore. However, the absence of fine grained material up to this depth in some geographic areas did require the inclusion of some deeper sites. Areas within 10km of a point source (Ontario Environment 1990) were excluded as potential reference site locations. Avoidance of areas likely to be affected by non-point source was done using topographic maps to locate areas that had minimal agricultural or urban shoreline development.

*Environmental Variables*

Three categories of environmental variables were measured; 1) geographic descriptors, which were large spatial scale descriptors of where the site was located; 2) sediment descriptors, which were physical or chemical measurements that would relate to small scale interactions between the organisms and their surrounding environment, and; 3) limnological descriptors, which relate the sediment to overlying lake scale processes. A total of 44 variables were measured (Table 1), details on sample handling, processing and analysis are available in Reynoldson *et al.* (1995).

The location of each site was established in the field using either Loran C or a hand-held Geographical Positioning System (GPS). Samples of water for chemical analyses were taken using a Van Dorn sampler from 0.5 m above the sediment-water interface. Sediments were characterized from samples taken from either a large box core

(50 x 50 cm) or mini box core (40 x 40 cm). Samples for geochemical analysis were taken from the surface 2 cm of the box core.

### *Invertebrate Community Structure*

Samples for the identification and enumeration of benthic invertebrates were taken using either a large box corer (1991) or a mini box core (1992-93). Comparison of matched samples using the two box cores showed no differences in estimates of invertebrate community structure. The intact box core was treated as a section of sediment that has simply been translocated to the surface. This was sampled by inserting five 10 cm plexiglass tubes (i.d. 6.5 cm, enclosed area 34.2cm$^2$) into the sediment in the box core. Each core tube was considered a replicate sample unit. The contents of each core tube were removed and placed into a plastic bag and kept cool until sieved. The contents of each bag were sieved through a 250µm mesh in the field immediately after sampling or preserved in 4% formalin. Samples were sorted with a low power stereo microscope and identified to species or genus level where possible. As required (Chironomidae and Oligochaeta) slide mounts were made for high power microscopic identification. Appropriate identification guides were used and voucher specimens of all identified specimens were submitted to experts for confirmation.

### *Whole Sediment Toxicity Tests with Reference Sediments*

A mini-ponar sampler was used at each site to obtain five separate samples of sediment for use in benthic invertebrate laboratory bioassays. Care was taken at each site during the sampling process to obtain sediment that was not disrupted by a previous sample collection method. The contents of each mini-ponar were placed in a food grade quality plastic bag and the bag was tightly tied with a plastic tie. All samples of sediment were placed in a cooler on ice until they were returned to the laboratory. In the laboratory, the bags of sediment were placed in plastic pails with lids and refrigerated at 4°C in the dark until bioassays could be conducted with the sediment.

Sediment toxicity tests were conducted using functional responses as endpoints. Four benthic invertebrate species were used as test organisms. In three species, the midge *Chironomus riparius*, the amphipod *Hyalella azteca* and the burrowing mayfly *Hexagenia limbata*, survival and growth were the endpoints. In a fourth species, the oligochaete worm *Tubifex tubifex*, survival and reproduction (cocoon/adult, hatch rate, young/adult) were the test endpoints. Thus, a total of four acute and six chronic endpoints were measured.

The culture and testing of the chironomid, *Chironomus riparius*, *Hexagenia limbata* and the oligochaete worm, *Tubifex tubifex*, are described in Reynoldson *et al.* (1991), Bedard *et al.* (1992), Hanes and Ciborowski (1990), Day *et al.* (1994) and Reynoldson *et al.* (1995). The culture of *H. azteca* was conducted according to the procedure described in Borgmann *et al.* (1989).

*Data Analysis*

Frequency distributions of the toxicity test data on survival, growth and reproduction for the four test species were plotted as histograms to present a graphical picture of the responses of each organism to a variety of reference sediments collected throughout the Great Lakes. In addition, the descriptive statistics of mean, median, standard error, standard deviation, maximum and minimum values and range were determined for each endpoint. The data were tested for normality and homogeneity of variance using SigmaplotR V.1.02 (Jandel Scientific). For purposes of analysis, the data pertaining to percent survival were transformed using the arcsine square root transformation.

The community structure data were examined using pattern analysis to investigate the structure of the data at the reference sites and correlation and multiple discriminant analysis (MDA) to relate the observed biological structure to the environmental characteristics. Structure in the data was examined using two pattern recognition techniques, cluster analysis and ordination. The mean values from the five replicates for both the species counts and test endpoints were used as descriptors of biological condition. The Bray and Curtis association measure was used as the association because it has performed consistently well in a variety of tests and simulations on different types of data (Faith *et al.* 1987). Clustering of the reference sites was done using an agglomerative hierarchical fusion method with unweighted pair group mean averages (UPGMA). The appropriate number of groups was selected by examining the group structure and, particularly, the spatial location of the groups in ordination space. Ordination was used to reduce the variables required to identify the structure of the data. A non-metric hybrid multi-dimensional scaling (MDS) method of ordination was used, *i.e.*, Semi- Strong- Hybrid multidimensional scaling (Belbin 1991). Hybrid multi-dimensional scaling methods use metric and non-metric rank order rather than metric information and thus provide a robust relationship with ecological distance. They do not assume a linear relationship an inherent assumption in some dissimilarity measures used by other ordination techniques (Faith *et al.* 1987). This is of particular value when relating ordination scores to environmental characteristics. All clustering and ordination was done using PATN, a pattern analysis software package developed by CSIRO in Australia (Belbin 1993).

Of the 44 environmental variables measured in this study, 27 were examined for their relationship with the biological structure of the data (Table 1). We excluded those variables most likely to be influenced by anthropogenic activity, particularly those associated with sediment contamination. Thus, all the metals were excluded from consideration as potential predictor variables. The variables used were general descriptors of sediment type such as the major elements and particle size as well as organic material, a potential indicator of nutritive quality. These together with physical attributes such as water depth and general water chemistry were considered as the most appropriate general habitat descriptors that would not be as subject to modification from human activity. The relationship with the biological data was examined two ways:

Using principal axis correlation in PATN; this is a multiple-linear regression method that describes how well a set of attributes (environmental data) can be fitted to ordination axes (species matrix). The method takes each environmental attribute and determines the location of the best fitted vector in ordination space. These can be represented as axes on an ordination plot and a correlation of the axis with the ordination is provided. A Monte Carlo simulation was then performed to establish the statistical significance of the correlations.

Stepwise discriminant analysis (Procedure STEPDIS in SAS) was used to establish which variables best described the biological groupings of the environmental data set.

Based on the results from these analyses, environmental variables were selected for use in multiple discriminant analysis (MDA) to relate the biological site groupings to the environmental characteristics of the sites. The SAS version of MDA was used with raw environmental data to generate discriminant scores, and to predict the probability of group membership. The more rigorous cross-validation method was used to verify the accuracy of the predictions from the discriminant model. Using this method each of the sites is in turn removed from the data set, a new model is generated without that site and then the removed site predicted to a group. The predicted groupings and actual groupings can then be compared to provide a group and total error rate.

Selection of the optimal predictor variable data set was done by iteration. Various combinations of predictor variables were selected from the stepwise discriminant analyses and principle axis correlation. The optimal set was defined as that with the lowest error rate from cross-validation in discriminant analysis.

## Results

A total of 345 site visits were made over the study period (1991-93). This included 41 seasonal visits to four sites visited monthly in 1992-93; together with repeat visits to 14 sites in each of the three years and to 17 sites for two of the three years. A total of 245 different locations were visited (Figure 4). Sampling was normally done in the autumn and 304 sites were included in the primary analysis. The 41 seasonal site visits have been examined separately to test the seasonal application of the reference data set.

Initial screening of the data set based on both toxicity and community structure data resulted in 252 sites being included as potential reference sites (Figure 1). Fifty two sites were excluded because either, two or more toxicity endpoints were below the acceptability criteria ( < lower 5th percentile of the distribution), the site had less than 50% survival for any test species or a site had no invertebrates present. This does not preclude a site from being re-instated as a reference site if it is equivalent to reference in future testing.

*Invertebrate communities*

A total of 162 taxa (genus or species level) were identified from the 252 reference sites in the Great Lakes. Using cluster analysis and ordination six groups of sites were identified that support different communities, these are summarised in Table 2.

Community 1 is characterised by *Chironomus* and *Dreissena*, another common chironomid, *Procladius* is also abundant in this community as is the sphaerid clam *Pisidium casertanum*. Numbers of *Chironomus* in community 1 are significantly greater than in the other 5 communities and the leech *Helobdella stagnalis* is also characteristic of this community. This community group contains 29 sites, the majority located in western and central Lake Erie.

Community 2 is characterised by the fingernail clam *Pisidium casernatum* and the amphipod *Diporeia hoyi* which is indicative of a more oligotrophic community. This was also suggested by the location of these communities in ordination space in which the first dimension represents a trophic and geographic. Community 2 is also more diverse with more taxa (Table 2) than the other community groups, but is also the least spatially defined. While the majority of sites in this group are located in Georgian Bay, it also includes sites from Lakes Erie, Ontario, Huron, Michigan and the North Channel.

Community 3 is characterised by the predatory midge *Procladius spp* and the fingernail clam, *Pisidium casernatum*, however total abundance is generally low at these sites. Half the sites in this community are from Georgian Bay, together with sites from Lake Erie (eastern basin) and the North Channel.

Community 4 consists of only 9 sites and these are dominated by very high numbers of the exotic species *Dreissena polymorpha* and *D. bugensis*, however this community is similar to community 1 with regard to the other taxa present. Both the location of the sites in Lake Erie and the sites position in ordination space suggest that these two communities are typical of the more mesotrophic Lake Erie (Table 2).

The last two communities, 5 and 6, both represent a *Diaporeia hoyi/Stylodrilus heringianus* assemblage. The major differences being the abundance of the two species and the presence of the oligotrophic chironomid *Heterotrissocladius spp*.

Community 5 has higher abundance of both *Diporeia* and *Stylodrilus* and *Heterotrissocladius* is less numerically important. This community is primarily found in Lake Michigan. Community 6 is the largest assemblage of sites (77) and includes more than 90% of the Lake Superior sites together with a large number of Georgian Bay and the North Channel sites (Table 2).

*Relationship with habitat structure*

From principle axis correlation 18 of the 27 variables considered as suitable predictor variables were significantly correlated ($P < 0.01$) with the species ordination matrix. Eleven variables were selected using stepwise discriminant analysis (Table 3) as being best related to the group structure. Two of these variables, pH (water) and MnO (sediment) were not significantly correlated with the species ordination matrix from principle axis correlation. From the relationship between site habitat attributes and species composition a model can be constructed that allows a prediction to be made of the type of community assemblage that should occur at any site in the Great Lakes basin. The predictive model is based on multiple discriminant analysis (MDA). This is a statistical method which enables one to statistically distinguish between two or more "groups" using a set of discriminating variables. In this case the "groups" have been established based on species composition and are defined as *community types* and the *discriminating variables* are a set of habitat attributes which are minimally affected by human activity. The results (Table 4) for three models show little difference in total error rates between the stepwise (32.4%) and optimal models (30.1%). The third model based on variables selected from principle axis correlation was not as accurate with a total error rate of 35.8%. The STEPWISE model uses 11 variables that are easily measured and describe geographic location (latitude and longitude), simple sediment attributes (total nitrogen and oxides of potassium, calcium, magnesium, manganese and silica) and general limnological conditions (water depth, alkalinity and pH of the water 0.5 m above the sediment water interface). In an independent test using 20 sites this model predicted 88% of the test sites to the correct community group (Table 3).

*Assessing impact*

While there are a number of possible approaches for comparing reference and test site(s), our approach is again a multivariate method in which the entire community assemblage can be used. In this method the reference and matching test site(s) are ordinated and plotted in ordination space. The distribution of the reference sites provides the range of variation in unimpaired communities. By constructing probability ellipses built around the reference sites the community at the test site can be compared to the range of conditions found amongst the reference sites with a given probability of belonging to the reference group. The greater the departure from reference state as measured in ordination space then the greater the difference from the reference condition. However, determining the actual degree of impact and what departure from the reference state defines an unacceptable impact is ultimately a subjective decision.

A large river water quality survey conducted in the UK in 1990 provided the impetus for the development of methods to circumscribe the continuum of responses into a series of bands that represented grades of biological quality (Wright *et al.* 1991). This is a simplification of what is a continuum of responses in sites ranging from good to poor biological quality. Despite the simplification, it was seen as an appropriate mechanism for obtaining a simple statement of biological quality allowing broad comparisons in either space or time that would be useful for management purposes.

We have adopted a similar approach for defining degrees of impact using a multivariate approach, using three probability ellipses (Figure 2). Sites inside the smallest ellipse (90% probability) would be considered *unstressed*; sites between the smallest and next ellipse (99% probability) would be considered *mildly stressed*; sites between the 99% probability and the largest ellizse (99.9% probability) would be considered *stressed*, and; sites located outside the 99.9% ellipse would be designated as *severely stressed*.

*Toxicity Endpoints and Target Values*

Several statistical analyses were conducted to try to correlate the responses for each end point and each species with sediment characteristics such as particle size distribution, TOC, LOI, MgO, SiO2,TP, TN, etc., for the reference sites. Both univariate (regression analysis with single variables) and multivariate statistics were used to determine if the range in any given response for a particular species in clean sediments could be correlated with specific characteristics of sediments. Although some trends were noted, especially with regard to growth and % silt or total organic matter in sediment, statistical significance with a single parameter could not be demonstrated. It was therefore concluded that the range in each endpoint noted for the reference sediment dataset represents the natural range in the responses of each organisms in laboratory bioassays. Based on these results, a decision was made to treat each response for a species as a continuum of data points with a range rather than to separate the responses into groups using multivariate analyses.

Mean percent survival of *C. riparius* in 212 reference sediments was 85.9 with a range of 53.5 to 100% and a CV of 10.2%. Only 4.7% of the reference sediments collected over the three-year period from all five of the Great Lakes caused mortality of *C. riparius* to be greater than 30%. ASTM (1994) have set a minimum acceptable criterion of 70% for survival of *Chironomus* spp. in uncontaminated sediments used in toxicity tests. Our results show that this criterion is achievable in the majority of sediments collected from reference areas in the Great Lakes. The few reference sediments for which % survival was <70% was well within the 1 in 20 results which would fall outside the 95% confidence limits for any given test.

Growth of larval chironomids in a variety of reference sediments with a range of physico-chemical characteristics was variable with dry weight of individual 4th instar larvae at test termination (10-d) ranging from 0.19 to 0.60 mg with a mean of 0.35 and a CV of 20.6%. All attempts to correlate this variability in growth to sediment characteristics were negative although some parameters such as TOC, % sand, % clay, total nitrogen, total phophorus and concentrations of lead, zinc, and copper in the reference sediments were implicated in both single parameter regressions and multivariate analyses.

As with midge larvae, survival of juvenile *H. azteca* in 212 reference sediments was good with a range of 52.0 to 100%, a mean of 87.5% and a CV of 10.6%. However, in 18.4% of the sediments tested, survival was below the minimum acceptable criterion of 80 % which has been set for *H. azteca* in control sediments in a 10-d lethality test by ASTM (1994).

The growth of 3 to 9 day-old *H. azteca* in reference sediments with a variety of physico-chemical characteristics over a 28-d exposure to sediments was more variable than growth in the midge bioassay and ranged from 0.10 to 0.80 mg dry wt. per juvenile with a CV of 26.5%. A negative correlation with % clay in the sediments was noted.

Percent survival of the mayfly nymph *Hexagenia* spp. was excellent in all types of sediment (167 sites) and ranged from 66.0 to 100% with a mean of 98.0% and a CV of 5.5%.

Growth of mayfly nymphs during the 21-d test was more variable than survival and ranged from 0.5 to 6.4 mg dry weight per individual with a CV of 34.0%. Strong positive correlations with LOI, TOC, TN, TP and SiO2 as well as negative correlations with % sand and % silt were noted in regressions.

Percent survival of adult *T. tubifex* was usually 100% in all bioassays with reference sediments (167 sites); only 3.6% of sediments tested recorded mortality between 10 and 20%. Based on these results, the acceptability criterion for % survival of adult worms in nontoxic sediments can be set quite high, i.e., >90%. Percent hatch of cocoons was also fairly high and constant with a mean of 58 ± 10% and a CV of 17.3%. The acceptability criterion for % hatch of cocoons is thus set at >35%. The number of cocoons produced per adult worm was consistent with a range of 4.8 to 14.5, a mean of 9.9 and a CV of 13.8% being recorded.

As the purpose of a toxicity test with whole sediment(s) is to determine if the biological response(s) of a cohort of organisms exposed to potentially contaminated sediment differ from the response(s) of a similar cohort of organisms to a negative control or reference sediment, the data from the reference sites was used to establish three categories of responses to test sediments. The three categories were - nontoxic, warning of potential toxicity and toxicity. The delineations for the three categories were developed from the standard statistical parameters of population mean and standard deviation ($x \pm$ S.D.) of an endpoint measured in all reference sediments (Table 5). For each endpoint, the nontoxic category was set at two standard deviations (S.D.) below the mean for the reference dataset; this represents the 95% confidence limit for that response (Table 5). At the 95% confidence level, 1 in 20 results (5%) would be expected to fall outside of the limits by chance alone. The toxic category was set at three S.D. below the mean of an endpoint which represents the 99.7% confidence limit (Table 5). At this confidence level, the probability of data falling outside of the limits by chance alone is only 0.3% (one out of every 333 tests). The range of responses between two and three times the S.D. represent the warning of potential toxicity category and would indicate sediment(s) which have some detrimental effects.

## Applying The Guidelines - The B.E.A.S.T. Software

Employing the reference condition approach for the benthic assessment of sediment has the potential to provide an alternative to current environmental guidelines and criteria. It has been suggested that multivariate methods such as those developed in this report are too complex, require specialized practitioners, and are difficult to convey to managers and the public (Gerritsen, 1995). Limitations associated with multivariate methods, however, can be attributed to the lack of a comprehensive tool for application. To date, someone wishing to employ multivariate methods for sediment analysis has required several expensive, cumbersome software packages to achieve their goals.

The need for a simple, inexpensive software tool which encapsulates the requirements for multivariate analysis has led to the development of the B.E.A.S.T. Designed exclusively for the **BE**nthic **A**ssessment of **S**edimen**T**, the software automates the methodology outlined in this report. Employing the RAISON Mapping and Analysis package from Environment Canada as a foundation, the B.E.A.S.T. combines new methods with a simple, straight-forward software user interface. The result is a powerful new tool for sediment analysis.

### *Conceptual Software Design*

The conceptual design for the BEAST calls for seven modules surrounding a central core of data. A method of automating the entry of data to be compared to the reference data base is the first module. Data in the BEAST was to be stored in an easily accessible, standard format to limit the problems normally associated with complex data sets. Once data to be analyzed by the BEAST has been entered, data handling and statistical modules are required. One module would predict the reference group membership of each test site using established predictor variables. The next is responsible for combining each test site with the appropriate group of reference data. The analysis of a site's community structure is the fourth module, with the final three modules responsible for graphic analysis and comparison of the BEAST results.

*Data Entry and Storage*

Microsoft™ Access file formats have been adopted for information storage and retrieval. A widely available Relational Data Base, Access is designed to accommodate the kind of large, complex data sets common to benthic analysis. All of the data related to BEAST operation is stored in this format.

Entry of Test data to the BEAST is achieved through the Benthic Data Information System (BDIS). Developed using Microsoft™ Access, BDIS is an automated data entry/management tool which provides a simple graphic user interface (GUI). Data entry errors are reduced significantly by providing validation routines to ensure that data falls within acceptable ranges, and conforms to previously established formats and standards.

Unlike may other software packages, the manual generation of complex input files for analysis is also eliminated by BDIS. Test data can be selected by a user and a data base file containing all of the information, in the proper format for successful BEAST analysis, automatically generated and placed in the appropriate location.

The B.E.A.S.T is also designed to maintain any number of reference data sets, without the need for continual updating of the software itself. When a new reference data base is developed, the resulting Access file can simply be placed in the same directory as other reference data base files within the B.E.A.S.T file structure. Once there, it is automatically available for analysis in the B.E.A.S.T.

*The BEAST Software*

The BEAST maintains information for various analysis projects in a hierarchical format. The first step in analysis is the creation of a *Project* with a unique name. Projects in the BEAST act as a container, establishing which reference and test data bases are to used each time the project is opened, and storing the results of any analysis undertaken. Any number of projects can be maintained within the B.E.A.S.T at any one time, and can be deleted when they are no longer needed.

A project in the BEAST also has a sub-set of containers within it called *scenarios*. Scenarios represent variations on the analysis of test data contained within a single project. Although the BEAST supplies an optimal set of predictor variables for Multiple Discriminant Analysis, some cases may occur where these variables are not available to the user. In these cases, the user must employ alternate variables, and the results for each of these discriminant models is stored as a scenario. This permits the user to compare the error rates of various discriminant functions, and select the most precise for use in the BEAST.

Results from BEAST analysis can be viewed several different ways. Error Rates and Probabilities of Prediction are generated in a table format. Using the RAISON mapping engine, thematic maps of group membership and toxicology for each site can be produced. Bar graphs comparing key species and environmental variables of a test site to the average of the related reference group can be generated. Finally, bivariate probability ellipse plots showing a test site's location in ordination space with relation to associated reference sites can be displayed.

## Summary

The process of determining whether an invertebrate community is impaired involves:

1. Sampling the community and measuring the predictor variables at the site of interest;

2. Running the discriminant model with the reference data base and test site(s) to predict the expected community at the test site(s);

3. Comparing test site(s) to the reference community sites, from the group to which the test site(s) were predicted.

Once a site is predicted to have a specific community structure, the actual structure is compared to the communities at the equivalent reference sites. A data file is constructed that includes the species counts for the appropriate reference sites. This is ordinated so that a matrix is calculated for both reference and test sites. The sites can then be plotted in ordination space showing the ordination dimensions that synthesise the biological attributes of the sites. Probability ellipses are calculated for the reference sites ONLY. The location of the tests sites relative to the reference sites can then be determined. The overall assessment of stress is based on site locations all the ordination vectors, and the overall assessment based on its worst position. The scale of response and the actual species that have been lost determine the degree of stress.

The assessment of the contribution of sediment toxicity to changes in community structure using multiple endpoints is simplified by criteria that rank sediment as either non-toxic, potentially toxic or toxic based on data from multiple reference sites.

## Acknowledgements

We wish to thank the following individuals without whom this study could never have been completed, Sue Humphrey of Ontario Region, Environment Canada whose efforts provided the resources by which this study was conducted. Danielle Milani and Cheryl Clark who performed most of the sediment toxicity tests. Craig Logan and Scott Hughson who sorted and identified the invertebrate samples. Finally, Dave Lam and Isaac Wong who converted our concepts into software that can actually be used.

## References

ASTM. (1994). *Standard guide for conducting sediment toxicity tests with freshwater invertebrates. American Society for Testing and Materials.* E 1383 -94a. Philadelphia, Pa. 33 pages.

Armitage, P.D., R.J.M. Gunn, M.T. Furse, J.F. Wright and D. Moss. (1987). The use of prediction to assess macroinvertebrate response to river regulation.. *Hydrobiologia* 144:25-32.

Bedard, D., A.Hayton and D. Persaud. (1992). *Ontario Ministry of Environment laboratory sediment biological testing protocol. Water Resources Branch, Ontario Ministry of Environment.* Toronto, Ontario. 26p.

Belbin, L. (1991). Semi-strong hybrid scaling, a new ordination algorithm. *J. Vegetation Science*, 2:491-496.

Belbin, L. (1993). *PATN, pattern analysis package*. Div. Wildlife and Ecology, CSIRO, Australia.

Besser, J.M., J.P. Giesy, J.A. Kubitz, D.A. *et al.* (1996). Assessment of sediment quality in dredged and undredged areas of the Trenton Channel of the Detroit River, Michigan, USA, using the sediment quality triad. *J. Great Lakes Res.* 22:683-696.

Borgmann, U., K.M. Ralph, and W.P. Norwood. (1989). Toxicity test procedures for *Hyalella azteca* and chronic toxicity of cadmium and pentachlorophenol to *H. azteca, Gammarus fasciatus* and *Daphnia magna*. *Arch. Environ. Contam. Toxicol.* 18:

Cairns, J.J. and W.H. van der Schalie. (1980). Biological monitoring Part 1 - Early warning systems. *Water Research* 14:1179-1196.

Canfield, T.J., F.J. Dwyer, J.F. Fairchild, P.S. Haverland, *et al.* (1996). Assessing contamination in Great Lakes sediments using benthic invertebrate communities and the sediment quality triad approach. *J. Great Lakes Res.* 22:565 583.

Chapman, P.M. (1986). Sediment quality criteria from the sediment quality triad: An example. *Environ.Toxicol.Chem.* 5:957-964.

Chapman, P.M. (1990). The sediment quality triad approach to determining pollution-induced degradation. *Sci. Total Environ.* 97/98: 815-825.

Corkum, L.D. and D.C. Currie. (1987). Distributional patterns of immature Simuliidae (Diptera) in northwestern North America. *Freshwater Biology* 17:201-221.

Day, K.E., R.S. Kirby and T.B. Reynoldson. (1994). Sexual dimorphism in *Chironomus riparius* (Meigen): impact on interpretation of growth in whole sediment toxicity tests. *Environ. Toxicol. Chem.* 13:35-39.

Faith, D.P., P.R. Minchin and L. Belbin. (1987). Compositional dissimilarity as a robust measure of ecological distance: A theoretical model and computer simulations. *Vegetatio* 69:57-68.

Furse, M.T., D. Moss, J.F. Wright and P.D. Armitage. (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. *Freshwater Biology* 14:257-280.

Gerritsen, J. (1995). Additive biological indices for resource management. *J. North Amer. Benth. Soc.* 14:451-457.

Hanes, E.C., Ciborowski, J.J.H. and Corkum, L.D. (1990). Standardized rearing materials and procedures for *Hexagenia*, a benthic bioassay organism: comparison of sediment types. In: *Proceedings 1990 Ontario Ministry of the Environment Technology Transfer Conference on Environmental Research*, edited by Anonymous. Toronto: 1990, p. 374-382.

Hunsaker, C.T. and D.E. Carpenter, eds. (1990). *Ecological Indicators for the Environmental monitoring and Assessment Program. EPA 600/3-90/060. U.S. Environmental Protection Agency, Office of Research and Development, Research Triangle Park*, N.C.

Johnson, R.K. and T. Wiederholm. (1989). Classification and ordination of profundal macroinvertebrate communities in nutrient poor, oligo-mesohumic lakes in relation to environmental data. *Freshwater Biology* 21:375-386.

Long, E.R. and P.M. Chapman. (1985). A sediment quality triad: Measures of sediment contamination, toxicity and infaunal community composition in Puget Sound.

*Mar. Poll. Bull.* 16:405-415.

Moss, D., M.T. Furse, J.F. Wright and P.D. Armitage. (1987). The prediction of the macroinvertebrate fauna of unpolluted running-water sites in Great Britain using environmental data. *Freshwater Biology* 17: 41-52.

Ontario Ministry of Environment. (1990). *The Canadian Great Lakes basin intake outfall atlas (8 vol.) Ed. M. Griffiths. Water Resources Branch*, Toronto, Ontario.

Ormerod, S.J. and R.W. Edwards. (1987). The ordination and classification of macroinvertebrate assemblages in the catchment of the River Wye in relation to environmental factors. *Freshwater Biology* 17: 533-546.

Painter, Scott. (1992). *Regional variability in sediment background metal conentrations and the Ontario sediment Quality Guidelines, NWRI Report No. 92-85. Environment Canada*, Burlington, Ontario.

Parsons, M. and R.H. Norris. (1996). The effect of habitat-specific sampling on biological assessment of water quality using a predictive model. *Freshwater Biology*. 36:419-434.

Reynoldson, T.B., S.P. Thompson and J.L. Bamsey. (1991). A sediment bioassay using the tubificid oligochaete worm *Tubifex tubifex*. *Environ. Toxicol. Chem.* 10:1061-1072.

Reynoldson, T.B. and M.A. Zarull. (1993). An approach to the development of biological sediment guidelines. In *Ecological Integrity and the Management of Ecosystems* (eds. G. Francis, J. Kay and S. Woodley). St Lucie Press, Florida.

Reynoldson, T. B., R. C. Bailey, K. E. Day, and R. H. Norris. (1995). Biological guidelines for freshwater sediment based on **BE**nthic **A**ssessment of **S**edimen**T** (the

BEAST) using a multivariate approach for predicting biological state. *Australian Journal of Ecology* 20:198-219.

Reynoldson, T.B., R. H. Norris, V. H. Resh, K. E. Day, and D. M. Rosenberg. (1997). The reference condition: a comparison of multimetric and multivariate approaches to assess water quality impairment using benthic macroinvertebrates. *J. N. Amer. Benthol. Soc.* 16:(in press)

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.

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

Wright, J.F., Furse, M.T., Clarke, R.T., and Moss, D. (1991). *Testing and further development of RIVPACS*. IFE Interim report to the National Rivers Authority, December 1991. pp. 1-141..

Zarull, M.A. and T.B. Reynoldson. (1992). A management strategy for contaminated sediments: assessment and remediation. *Water Poll. Res. J. Canada.* 27:871-882.

## List of Tables

Table 1.Measured Environmental Variables at Reference Sites.

| Measured Variable | Rationale | Use as Potential predictor |
|---|---|---|
| **Geographic** (5 variables) | geographic descriptors provide a synthesis of the effects of spatial processes on animal distribution | |
| latitude | | yes |
| longitude | | yes |
| lake basin | | no - non-continuous |
| ecodistrict | | no - non quantitative |
| date | | no - temporal effects examined separately |
| **Limnological** (8 variables) | | |
| water depth | integrates effects of temperature and oxygen on organisms | yes |
| dissolved oxygen | critical for most aerobic organisms | no - modified by seasonal processes |
| pH | modifies chemical interactions | yes |
| temperature | effects growth and reproductive processes | no - requires temporal integration |
| alkalinity | summarizes dissolved materials | yes |
| total phosphorus | effects nutrient status and primary producers | no - modified by anthropogenic inputs |
| kjeldahl nitrogen | effects primary producers | no - modified by anthropogenic inputs |
| nitrate-nitrite nitrogen | effects primary producers | no - modified by anthropogenic inputs |
| **Sediment** (31 variables) | | |
| Particle Size - 7 variables (% gravel, sand, silt clay, mean, $75^{th}$, $25^{th}$ %le) | Effects burrowing organisms, modifies bioavailability of materials. | yes |
| Major elements - 11 variables (oxides of Si, Ti, Al, Fe, Mn, Mg, Ca, Na, K, P) | Provide a good descriptor of overall sediment conditions, provides a regional signal. | yes |

| | | |
|---|---|---|
| Nutrients - 4 variables (TP, TN, loss on ignition, TOC) | Provide an indicator of food availability | yes |
| Metals - 9 variables (Totals for V, Cr, Co, Ni, Cu, Zn, As, Cd, Pb) | Provide a descriptor of anthropogenic inputs and general contaminant levels, allow verification of reference status | no - modified by anthropogenic inputs |

Table 2. Geographic distribution and taxonomic composition (major taxa, mean number per 34.2 $cm^2$ with SD) of sites in six communities formed from cluster analysis

| Lake | Gp 1 | Gp 2 | Gp 3 | Gp 4 | Gp 5 | Gp 6 |
|---|---|---|---|---|---|---|
| Erie | 17 | 3 | 11 | 9 | 1 | 0 |
| Ontario | 7 | 6 | 9 | 0 | 2 | 4 |
| St Clair | 1 | 0 | 0 | 0 | 0 | 0 |
| Huron | 2 | 4 | 1 | 0 | 6 | 4 |
| Georgian Bay | 0 | 11 | 32 | 0 | 0 | 18 |
| North Channel | 2 | 8 | 11 | 0 | 1 | 14 |
| Michigan | 0 | 7 | 0 | 0 | 22 | 8 |
| Superior | 0 | 0 | 0 | 0 | 2 | 29 |

| Taxa | Gp 1 | Gp 2 | Gp 3 | Gp 4 | Gp 5 | Gp 6 |
|---|---|---|---|---|---|---|
| *Chironomus spp* | 5.7 (5.8) | 0.9 (3.1) | 0.8 (1.3) | 1.3 (1.8) | 0.0 (0.0) | 0.1 (0.4) |
| *Tanytarsus spp* | 0.4 (1.3) | 4.7 (7.8) | 0.8 (1.5) | 0.1 (0.1) | 0.1 (0.3) | 0.5 (1.5) |
| *Heterotrissocladius spp* | 0.2 (1.1) | 0.8 (2.5) | 0.0 (0.0) | 0.4 (0.7) | 1.2 (1.7) | 1.6 (1.8) |
| *Procladius spp* | 1.5 (1.9) | 1.9 (2.3) | 3.2 (2.7) | 2.0 (1.4) | 0.1 (0.3) | 0.2 (0.6) |
| *Diaporeia hoyi* | 0.0 (0.0) | 2.8 (6.2) | 0.3 (0.9) | 0.0 (0.0) | 65.1 (41.8) | 10.4 (5.1) |
| *Valvata tricarinata* | 0.2 (0.4) | 0.7 (1.9) | 0.1 (0.2) | 1.7 (2.0) | 0.0 (0.0) | 0.0 (0.0) |
| *Dreissena polymorpha* | 5.1 (7.7) | 0.2 (1.0) | 0.2 (0.6) | 101.4 (78.1) | 0.0 (0.1) | 0.1 (0.7) |
| *Dreissena bugensis* | 1.8 (7.2) | 0.0 (0.1) | 0.0 (0.0) | 122.8 (181.2) | 0.0 (0.0) | 0.0 (0.0) |
| *Pisidium casertanum* | 2.5 (2.8) | 4.4 (8.7) | 1.0 (1.8) | 0.8 (0.8) | 5.0 (8.4) | 0.6 (1.1) |
| *Stylodrilus heringianus* | 0.0 (0.0) | 0.8 (1.8) | 0.0 (0.0) | 0.0 (0.0) | 10.2 (8.9) | 2.0 (3.8) |
| *Vejdovskia intermedia* | 0.2 (0.5) | 2.3 (6.5) | 0.1 (0.2) | 1.5 (3.9) | 1.1 (2.0) | 0.3 (1.0) |
| *Potamothrix vejdovskii* | 0.3 (1.2) | 4.8 (14.4) | 0.0 (0.2) | 6.2 (18.2) | 1.1 (3.9) | 0.1 (0.4) |
| *Spirosperma ferox* | 1.5 (5.2) | 4.0 (9.6) | 0.2 (0.5) | 0.4 (0.5) | 0.2 (0.7) | 0.5 (0.7) |

Table 3. Relationship between environmental variables and species based on principle axis correlation (PAC) and stepwise discriminant function analysis (stepwise).

| Variable | Partial $r^2$ (stepwise) | Prob. (stepwise) | r (PAC) |
|---|---|---|---|
| Depth | 0.519 | 0.0001 | 0.7458 |
| Latitude | 0.405 | 0.0001 | 0.6496 |
| Total Nitrogen (sediment) | 0.216 | 0.0001 | 0.4131 |
| Alkalinity (water) | 0.196 | 0.0001 | 0.5183 |
| $K_2O$ (sediment) | 0.098 | 0.0001 | 0.3138 |
| Longitude | 0.085 | 0.0007 | 0.5639 |
| CaO (sediment) | 0.064 | 0.0066 | 0.4650 |
| MgO (sediment) | 0.089 | 0.0004 | 0.2436 |
| pH (water) | 0.065 | 0.0063 | 0.1815 (n.s. $P > 0.01$) |
| MnO (sediment) | 0.054 | 0.0216 | 0.1362 (n.s. $P > 0.01$) |
| $SiO_2$ (sediment) | 0.056 | 0.0170 | 0.3449 |

Table 4. Error rate estimates for species level discriminant models constructed using three sets of variables.

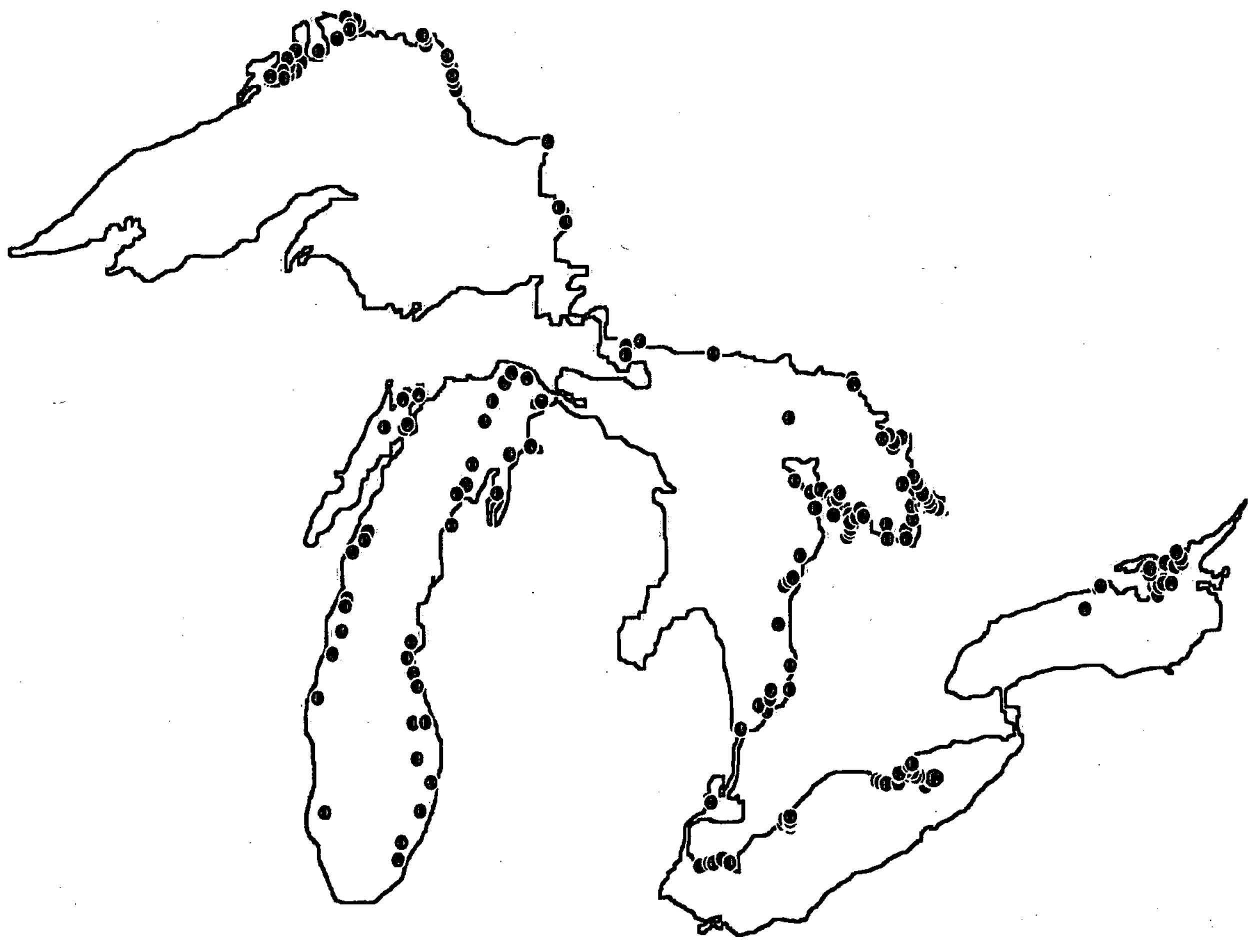
| Variable selection method | Stepwise discriminant analysis | Principle axis correlation | Optimal Iteration |
|---|---|---|---|
| Variables used | 11 (depth, lat, lon, alkw, pH, TN, $K_2O$ CaO, MgO, MnO, $SiO_2$) | 18 (depth, lat, lon, alkw, TN, TP, TOC, LOI, CaO, Al2O3, SiO2, K2O, MgO, Na2O, Sil, San, Cly, Psz75) | 12 (depth, lat, lon, alkw, pH, TN, TOC, $K_2O$ CaO, MgO, MnO, $SiO_2$) |
| Total Error rate for 252 sites | 32.4 % | 35.8% | 30.1% |
| Error rates for 20 site subset | | | |
| Gp 1 (n=2) | 0.00 | 0.00 | 0.00 |
| Gp 2 (n=3) | 0.00 | 0.33 | 0.00 |
| Gp 3 (n=5) | 0.40 | 0.60 | 0.40 |
| Gp 4 (n=1) | 0.00 | 0.00 | 0.00 |
| Gp 5 (n=3) | 0.00 | 0.00 | 0.00 |
| Gp 6 (n=6) | 0.33 | 0.33 | 0.33 |
| Total error rate | 0.12 | 0.20 | 0.12 |

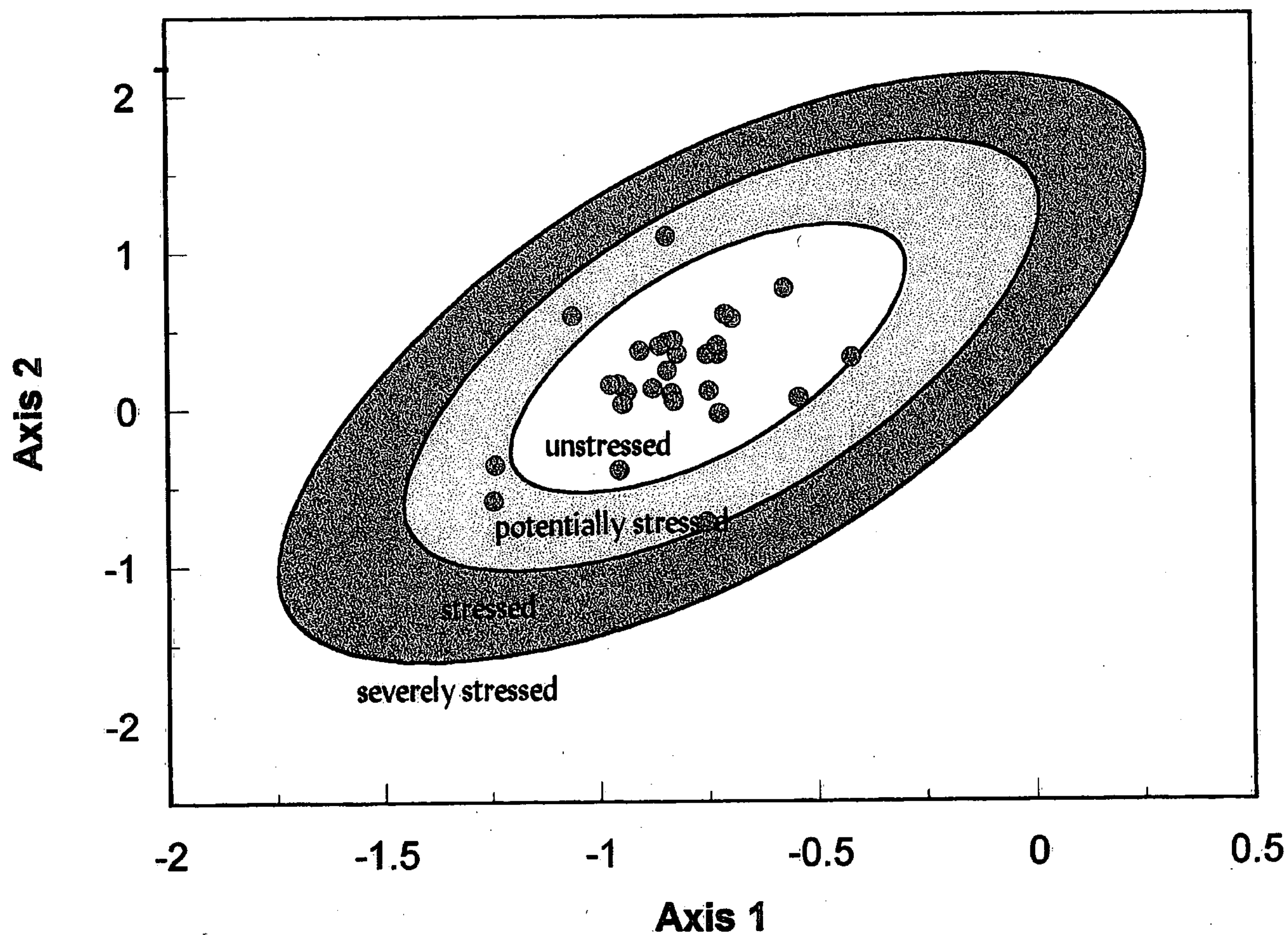
Table 5. Criteria for determination of toxicity for nearshore sediments of the Great Lakes.

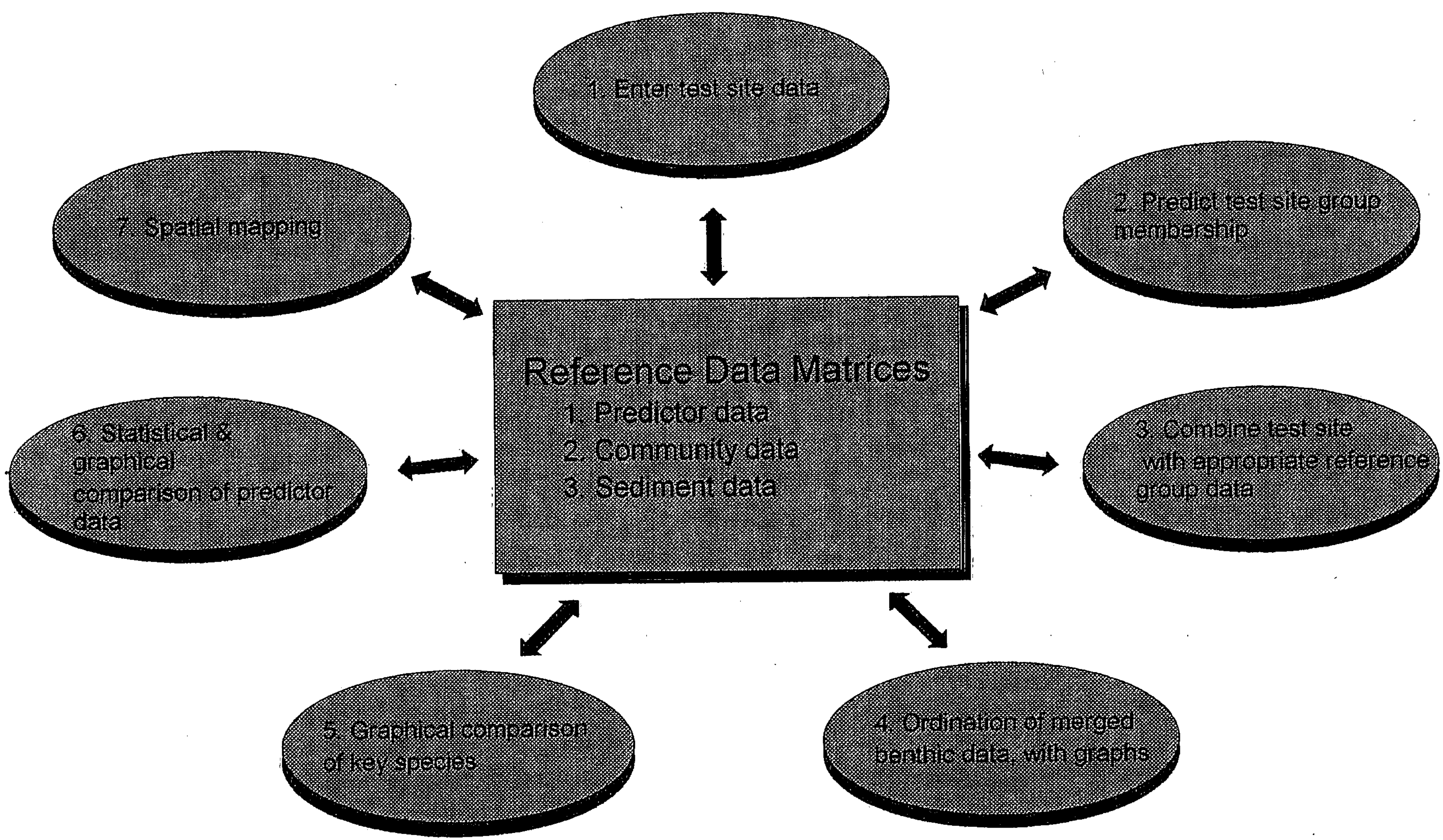
| | 1<br>Non-toxic* | 2<br>Warning of Potential Toxicity minus twice S.D. | 3<br>Toxicity minus three times S.D. |
|---|---|---|---|
| Chironomus riparius | | | |
| % Survival | ≥ 69.0 | 60.0 - 68.9 | < 60.0 |
| Growth | 0.21 - 0.49 | 0.14 - 0.20 | < 0.14 |
| Hyalella azteca | | | |
| % Survival | ≥ 69.0 | 60.0 -68.9 | < 60.0 |
| Growth | 0.24 - 0.76 | 0.11 - 0.23 | <0.11 |
| Hexagenia spp. | | | |
| % Survival | ≥ 85.0 | 80.0 - 84.9 | < 80.0 |
| Growth | 1.00 - 5.00 | 0 - 0.99 | - |
| Tubifex tubifex | | | |
| % Survival | ≥ 88.0 | 81.0 - 89.9 | < 81.0 |
| % Hatch | 37.8 - 78.2 | 27.7 - 37.7 | <27.7 |
| No. Coc/Worm | 7.1 - 12.7 | 5.7 - 7.0 | < 5.7 |
| No. Young/Worm | 11 - 47 | 2.0 - 10.9 | < 2.0 |

* upper limit for nontoxic category is set using twice the S.D. of the mean and indicates excessive growth or reproduction

## List of Figures




2
1
0
-1
-2
Axis 2
unstressed
potentially stressed
stressed
severely stressed
-2
-1.5
-1
-0.5
0
0.5
Axis 1


1. Enter test site data
2. Predict test site group membership
3. Combine test site with appropriate reference group data
4. Ordination of merged benthic data, with graphs
5. Graphical comparison of key species
6. Statistical & graphical comparison of predictor data
7. Spatial mapping
Reference Data Matrices
1. Predictor data
2. Community data
3. Sediment data

Environment Canada Environnement Canada

Canada


**Canada Centre for Inland Waters**
P.O. Box 5050
867 Lakeshore Road
Burlington, Ontario
L7R 4A6 Canada

**National Hydrology Research Centre**
11 Innovation Boulevard
Saskatoon, Saskatchewan
S7N 3H5 Canada

**St. Lawrence Centre**
105 McGill Street
Montreal, Quebec
H2Y 2E7 Canada

**Place Vincent Massey**
351 St. Joseph Boulevard
Gatineau, Quebec
K1A 0H3 Canada

**Centre canadien des eaux intérieures**
Case postale 5050
867, chemin Lakeshore
Burlington (Ontario)
L7R 4A6 Canada

**Centre national de recherche en hydrologie**
11, boul. Innovation
Saskatoon (Saskatchewan)
S7N 3H5 Canada

**Centre Saint-Laurent**
105, rue McGill
Montréal (Québec)
H2Y 2E7 Canada

**Place Vincent-Massey**
351 boul. St-Joseph
Gatineau (Québec)
K1A 0H3 Canada