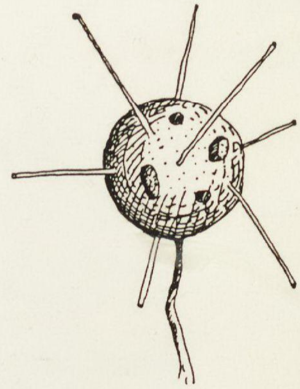


EPA REGION VIII

Bioassessment Workshop:

Analysis of Biological Data

The College Inn Conference Center
Boulder, CO

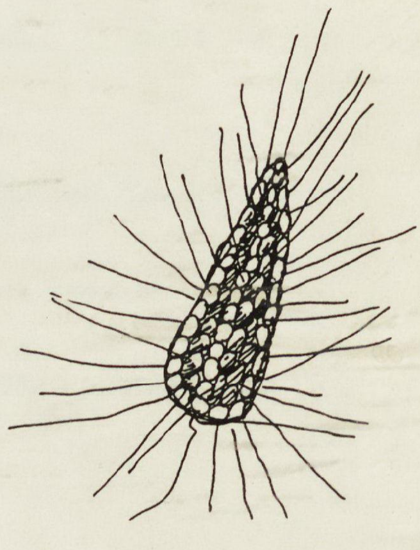
September 20, 21, and 22, 1995

Steering Committee:

Loren Bahls, Montana Department of Health & Environmental Sciences
Mike Barbour, US Environmental Protection Agency
Chris Faulkner, US Environmental Protection Agency
Susan Foster, Thorne Ecological Institute
Jeroen Gerritsen, Tetra Tech, Inc.
Susan Jackson, US Environmental Protection Agency
Phil Johnson, US Environmental Protection Agency
Bob McConnell, Colorado Department of Public Health & Environment
Toney Ott, US Environmental Protection Agency
Bill Wuerthele, US Environmental Protection Agency

Workshop Organizer:

Thorne Ecological Institute
5398 Manhattan Circle, Suite 120
Boulder, CO 80303
(303) 499-3647



257

R8
0035
c.1

EPA REGION VIII BIOASSESSMENT WORKSHOP : ANALYSIS OF BIOLOGICAL DATA

September 20, 21 and 22, 1995
The College Inn Conference Center, University of Colorado, Boulder, CO

Wednesday, September 20

1:00 p.m. **WELCOME,** *Betty Lou Carpenter*, Thorne Ecological Institute, Boulder, CO and *Phil Johnson*, U.S. EPA Region VIII, Denver, CO

SESSION I : AN OVERVIEW OF DATA ANALYSIS AND RELATED ISSUES

1:15 p.m. **Multimetric Approaches to Bioassessment,** *Jeroen Gerritsen*, Tetra Tech, Inc., Owings Mills, MD

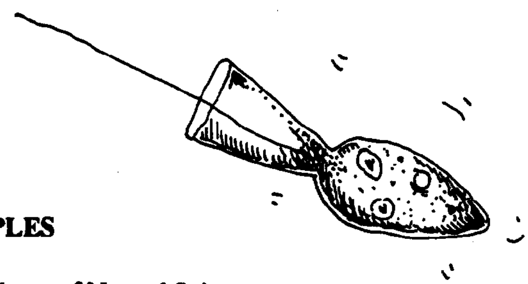
2:15 p.m. **Bioassessment Using Predictive Multivariate Models: Clarity, Not Smoke and Mirrors,** *Robert Bailey*, Dept. of Zoology, University of Western Ontario, London, Ontario

3:15 p.m. **BREAK**

3:30 p.m. **Data Management Issues and GIS Approaches for Biological Assessment,** *Tony Selle*, U.S. EPA Region VIII, Denver, CO

4:30 p.m. **DISCUSSION**

5:00 p.m. **RECEPTION**



Thursday, September 21

SESSION II: CASE STUDIES AND EXAMPLES

8:30 a.m. **Use of Periphyton as a Bioassessment Tool,** *Donald Charles*, Academy of Natural Sciences, Philadelphia, PA

9:00 a.m. **Trend Detection in Biological Monitoring: Evaluating Patterns of Community Similarity,** *Donald Charles*, Academy of Natural Sciences, Philadelphia, PA

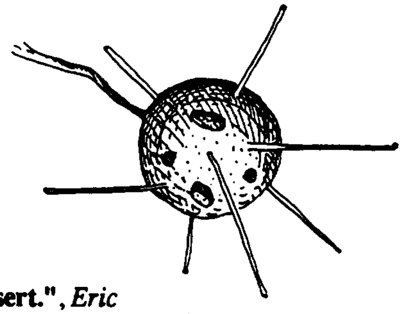
9:30 a.m. **Invertebrate Assemblages and Regional Classification: The Interior Columbia Basin as a Case Study in Scaling Up,** *Judith Li*, Oregon State University, Corvallis, OR

10:00 a.m. **BREAK**

10:15 a.m. **Bioassessment Using Macroinvertebrates in Low Gradient Reaches of the South Platte River,** *Jill Minter*, Dept. of Earth Resources, Colorado StateUniversity, Ft. Collins, CO

10:45 a.m. **Developing Biological Criteria for Pacific Northwest Streams,** *Leska Fore*, Institute for Environmental Studies, University of Washington, Seattle, WA

11:15 a.m. **Developing Bioassessment Protocols for Montana Wetlands,** *Randy Apfelbeck*, Montana Dept. of Environmental Quality, Helena, MT



11:45 a.m. **DISCUSSION**

12:15 p.m. **LUNCH**

1:15 p.m. **An Index of Biological Integrity for the Red River Ecoregion: A Biological "Desert."**, *Eric Pearson*, North Dakota Dept. of Health, Bismarck, ND

1:45 p.m. **Update on the Use of Macroinvertebrates in Montana Stream Bioassessments,** *Bob Bukantis*, Montana Dept. of Environmental Quality, Helena, MT

2:15 p.m. **Use of Lake Benthos as a Bioassessment Tool,** *Malcolm Butler*, Dept. of Zoology, North Dakota State University, Fargo, ND

2:45 p.m. **BREAK**

3:00 p.m. **Assessing Effects of Metals on Benthic Macroinvertebrate Communities in Rocky Mountain Streams,** *Peter Kiffney*, Dept. of Fishery and Wildlife Biology, Colorado State University, Ft. Collins, CO

3:30 p.m. **The Ordination of Benthic Invertebrate Communities in the South Platte River Basin in Relation to Environmental Factors,** *Cathy Tate*, U.S. Geological Survey, Denver, CO

4:00 p.m. **DISCUSSION AND SUMMARY**

5:00 p.m. **ADJOURN**

Friday, September 22

REGION VIII WORKING GROUP MEETING

(A facilitated discussion among EPA, States and Tribes examining challenges to program implementation)

8:00 a.m. **Introduction**

8:15 a.m. **Sample Collection and Analysis Needs for Region VIII States and Tribes**

9:00 a.m. **Database Needs within Region VIII and Possible Solutions**

9:45 a.m. **BREAK**

10:00 a.m. **Data Analysis Needs within Region VIII and Possible Solutions**

10:45 a.m. **Maximizing Program Effectiveness through Interagency Coordination: Defining Roles for EPA Region VIII, the States and the Tribes**

11:30 a.m. **Next Steps and Action Items**

12:30 p.m. **ADJOURN**

Organized by Thorne Ecological Institute

Thorne Ecological Institute is a non-profit environmental education organization founded in 1955. TEI promotes positive change by encouraging individual behaviors based on ecological principles that achieve environmental, economic and social harmony. For more information call (303) 499-3647.

MULTIMETRIC APPROACHES TO BIOASSESSMENT

Jeroen Gerritsen
Tetra Tech, Inc.
10045 Red Run Blvd., Suite 110
Owings Mills, MD 21117

NOTES:

The last decade has seen the development and increasing application of multimetric biological assessment methods, with the purpose of expressing and interpreting whether a water body is similar to an undisturbed reference state. Multimetric indices are additive; i.e., the sum of several measurements or calculated variables, termed metrics, obtained from sampling an aquatic biological assemblage. A metric is an ecological attribute of the assemblage, and should be responsive to perturbation or pollution. The multimetric approach was designed to be simple to implement for routine, operational monitoring, although establishment of a program requires substantial sampling and analysis effort.

Assessments are based on comparison of biological metrics at a site to their expected value under regional reference conditions. Reference conditions are not *ad hoc* special cases, such as upstream or paired reference sites, but should reflect regional conditions, and regional variability, under minimal human disturbance. Reference conditions are established by identifying sites that meet selected criteria for minimal disturbance, and characterizing their biological condition. A critical element in characterization of reference conditions is appropriate classification of site types, to ensure that apples are kept separate from oranges. A common approach to classification has been geographic categories such as ecoregion or biogeographic province, and continuous covariates such as stream size or salinity.

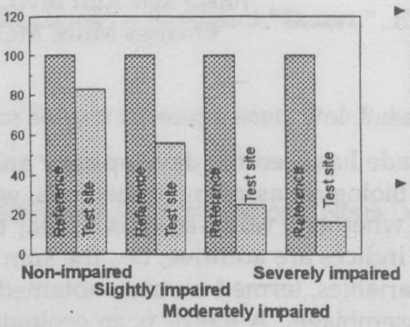
Metrics are selected based on low variability in the reference population, and responsiveness to disturbance or pollution. Metrics are scored on a common scale and the index is the sum of the selected metrics. Metrics are commonly scored on an ordinal scale of 1, 3, 5, or as a percent of the reference value. Experience with the multimetric approach has shown it to be reliable for detecting impairment of water bodies. Successful application requires a step-by-step progression to establish regional reference conditions, to select appropriate metrics, and to test the component metrics on a regional basis. Significant unresolved issues of biological assessment include spatial and temporal variability, and optimal sampling to minimize the effects of variability.

Jeroen Gerritsen
Tetra Tech, Inc.

Multimetric Approaches to Biological Assessment

1.1

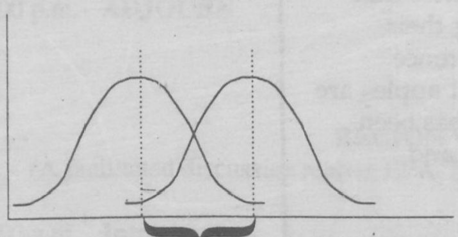
Biological Criteria and Bioassessments



- Biological criteria, determined from the reference condition, are used as the measuring stick for biological integrity
- Biological assessment scores are compared to these criteria to determine if a site is impaired

1.2

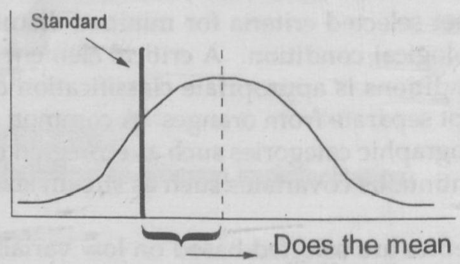
Traditional Statistical Tests



Are these means significantly different?

1.3

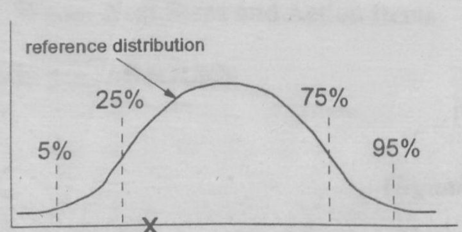
Standards and Criteria



Does the mean significantly exceed the standard?

1.4

Bioassessment



Is this observation (or mean of several observations) different from the reference population?

1.5

Assessment Alternatives

- Single indicator
- Multimetric index (this talk)
- Multivariate models (Bob Bailey)
- Indicator species models (Don Charles)
- Discriminant function models (in use in Maine)

1.6

Multimetric Approach

- Development of an index that is the sum of several metrics, or indicators, of biological condition
- The value of each component metric and the resultant index is based on reference condition

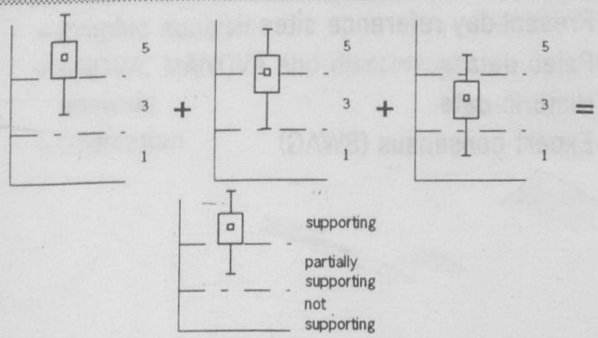
1.7

Definition

- A *metric* is a characteristic of the biota that changes in some predictable way with increased human influence.

1.8

Simple and Stupid Index; The index is the sum of several metrics:



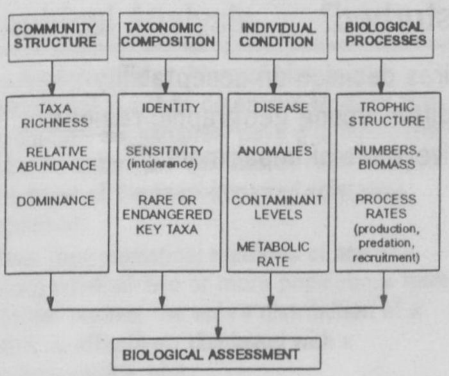
1.9

Metrics

- Measurements to determine biological integrity should:
  - be relevant to societal concerns
  - be responsive to environmental stresses
  - have low uncertainty
  - be cost-effective
  - be environmentally benign to measure

1.10

POTENTIAL ATTRIBUTES THAT MAY SERVE AS METRICS



1.11

Multimetric Approach

- The strength of using multiple metrics (attributes) is the ability to integrate information from individual, population, community, zoogeographic, and ecosystem level of a single assemblage into a single ecologically based index of the quality of a water resource.

1.12

Multimetric Approach

- Multimetric bioassessment is not a ready-made, off-the-shelf instrument. It is an approach that must be modified to specific regional conditions before it can be applied.

1.13

Topics

- Reference characterization
- Metric selection
- Index development
- Performance and unresolved issues

1.14

Reference Conditions for Biological Assessment

- Reference Conditions: Expectations on the state of biological communities in the absence of anthropogenic disturbance and pollution.
- Reference Sites: Real sites that form a database for characterizing reference conditions. These sites should be minimally impaired by human pollution and disturbance.

2.1

Characterization of Reference Conditions

- Present-day reference sites
- Paleo data
- Historic data
- Expert consensus (SWAG)

2.2

Reference Site Criteria

- All within category
- Least impacted within the context of the ecoregion or class
- Least impacted must be determined by direct evidence of human activity, <u>not</u> by biological responses
  - population
  - roads
  - discharges
  - structures (canals, dikes, armor)
  - hydrological modification

2.3

How Undisturbed is Undisturbed?

- Requires decision on acceptability
- May differ among geographic regions
- Relative scale of impairment

2.4

Regionalization and Preliminary Classification

- The intent of classification is to identify groups of sites that under ideal condition would have comparable biological communities.
- Classification should rely on those characteristics of sites that are intrinsic, or natural, and not the result human activities.

2.5

Classification Approaches

- Two fundamental approaches exist for classification: a priori and a posteriori.
  - *a priori* consists of developing logical rules for classification based on observed patterns in the characteristics of the objects (e.g., classifying lakes on ecoregion, surface area, and maximum depth)
  - a posteriori develops groups from a database of observations from the sites and is restricted to those sites in the database

2.6

Testing the *a priori* classification

- graphic analysis
- ANOVA, MANOVA and discriminant analysis
- Ordination

2.7

Florida Ecoregions

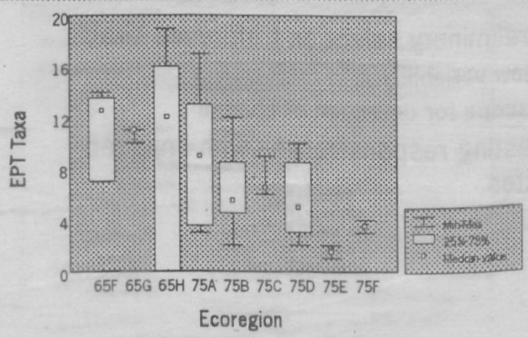
2.8

Graphical Analysis

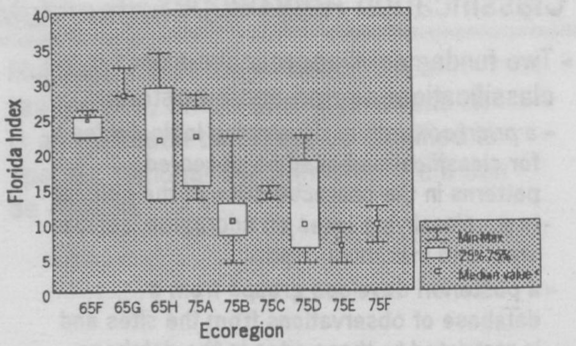
- The fundamental problem of biological assessment is not to determine whether two populations (or samples) have a different mean, rather to determine whether an individual site is a member of the least-impaired reference population.
- Rather than statistical methods of analysis testing whether two or more populations have different means, the entire distribution of a metric is effectively displayed with a box-and-whisker plot.

2.9

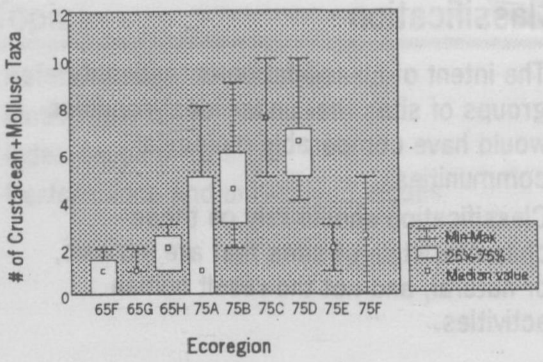
Distribution of EPT Taxa Metric Among Florida Subregional Reference Sites



2.10

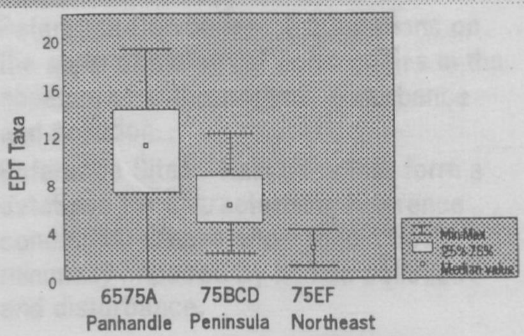


2.11



2.12

Distribution of EPT Taxa in Aggregated Ecoregions



2.13

MDS

2.14

Metric Selection

- Preliminary selection (reference sites)
  - low variability
  - scope for detection of change
- Testing responsiveness with impaired sites

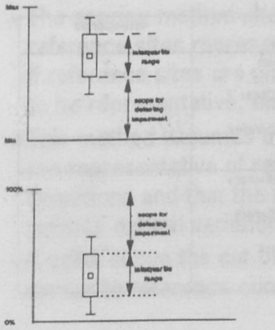
3.1

Metric Variability

- Metrics that are too variable within the reference sites are unlikely to be effective for assessment.
- A measure of metric variability is the ratio of the interquartile range to the distance between the lower quartile and the minimum possible value of the metric.

3.2

Assessing Candidate Metrics



- Metrics that have high values under reference (unimpaired) conditions.
- Metrics that have low values under reference conditions.

3.3

Candidate Macroinvertebrate Metrics to be Used for Site Classification and Discrimination

Richness Measures	Composition Measures	Tolerance Measures	Trophic Measures
# of Total Taxa	Shannon-Wiener Index	Florida Index	% Collector-Gatherers
EPT Index	% Dominant Taxon	% Class 1 and Class 2	% Collector-Filterers
# of Chironomidae Taxa	% Diptera	Hilsenhoff Biotic Index	% Shredders
# of Crustacean + Mollusc Taxa	% Crustacean + Mollusc		

3.4

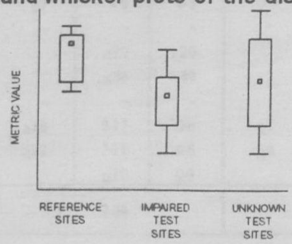
Metric Evaluation and Index Development

- Selection of metrics and development of a multimetric index requires a test data set composed of reference sites and nonreference (test) sites that might be impaired or that simply do not meet the criteria for reference sites.
- Ideally, the test sites should include at least some sites that are severely impaired by different stressors.
- Reference condition characterization used only the reference site data; metric evaluation and index development use both reference and test site data.

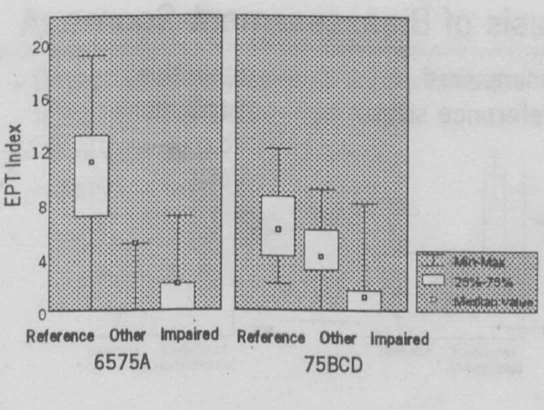
3.5

Metric Response

- Response of metrics to stresses is evaluated by comparison of reference sites to test sites using box-and-whisker plots of the distribution.

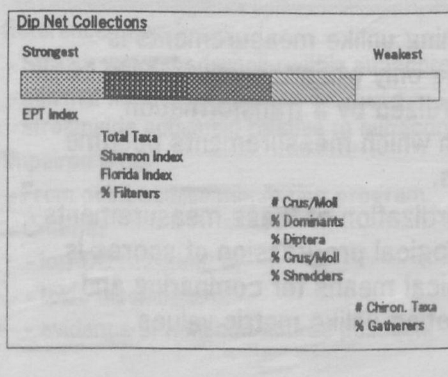


3.6



3.7

Discriminatory Power of the Macroinvertebrate Metrics



3.8

Metric Response to Increasing Biological Condition

Category	Metric	Response
Richness Measures	# of Total Taxa	+
	EPT Index	+
	# of Chironomidae Taxa	+
	# of Crustacean + Mollusc Taxa	+
Composition Measures	Shannon-Wiener Index	+
	% Dominant Taxon	-
	% Diptera	-
	% Crustacean + Mollusc	-

3.9

Metric Response to Increasing Biological Condition

Category	Metric	Response
Tolerance Measures	Florida Index	+
	% Class 1 and Class 2	+
	Hilsenhoff Biotic Index	-
Trophic Measures	% Collector-Gatherers	-
	% Collector-Filterers	-
	% Shredders	+

3.10

Multimetric Index Development

- Index scoring alternatives
- Examples

4.1

Index Development

- Following classification and characterization of the reference condition, metrics are evaluated for suitability in a multimetric index.
- Suitable metrics are those that respond in a predictable way to stressors on the system and that have low noise or variability.

4.2

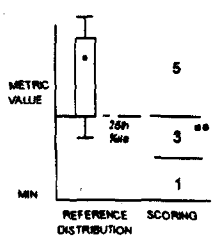
Scoring and Index Development

- Combining unlike measurements is possible only when the values have been standardized by a transformation through which measurements become unitless.
- Standardization of these measurements into a logical progression of scores is the typical means for comparing and interpreting unlike metric values.

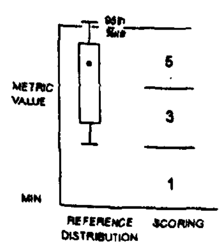
4.3

Basis of Bioassessment Scores

- unimpaired reference sites



- population distribution



4.4

Scoring and Index Development

- The scoring method should reflect how well the reference sites represent unimpaired conditions. If reference sites are unimpaired and considered to be representative, bisection is recommended.
- This method assumes that the reference sites are representative of relatively unimpaired conditions and that the metric distribution reflects natural variation of the metric.
- A value above the cut off is then assumed to be similar to reference conditions.

4.5

Scoring and Index Development

- The trisection method is best for scoring in regions where impacts might be so pervasive that nearly all reference sites are thought to be impacted or for assessment of reservoirs where reference sites cannot be defined.
- In trisection, it is assumed that at least some of the lakes attain an excellent value for the metric, but that many reference lakes are impaired and hence the lower limit of the reference distribution is not known.
- The 95th percentile is thus taken as the 'best' value and the range is trisected below it.

4.6

Scoring and Index Development

- Choice of scoring method should be based on confidence in the reference sites, rather than on the method that will produce the most conservative or most liberal scoring.
- If confidence is high that reference sites are representative of relatively unimpaired conditions, then the lower percentile cutoff and bisection are preferred.
- If confidence is low, then trisection below the 95th percentile is preferred.

4.7

Scoring Criteria for the Core Metrics as Determined by the 25th Percentile of the Metric Values for the Two Aggregated Subecoregions

Metric	Panhandle			Peninsula		
	5	3	1	5	3	1
# of Taxa	≥31	16-30	0-15	≥27	14-26	0-13
EPT Taxa	≥7	4-6	0-3		≥4	0-3
Crust + Moll Taxa	-	-	-		≥4	0-3
% Dom Taxon		≤20	>20		≤37	>37
% Diptera		≤38	>38		≤32	>32
% Crus + Moll	-	-	-		≥16	0-15
Florida Index	≥18	9-17	0-8	≥7	4-6	0-3
% Filterers	≥12	7-11	0-6	≥8	4-7	0-3
% Shredders		≥10	0-9		≥13	0-12
Range of Aggregated Score	7-29			9-33		

4.8

Assessment

- Detection of impairment, as deviation from reference condition
  - Real sites
  - Impaired scenarios

4.9

Site Criteria

- Reference Sites
  - Stream watershed wholly within subecoregion.
  - minimal human disturbance (land use, streamside activities) relative to subecoregion.
- Impaired Sites
  - From point source monitoring program.
  - Criteria:
    - low DO (<2ppm); or
    - toxic discharge; or
    - evidence of nonpoint source pollution.

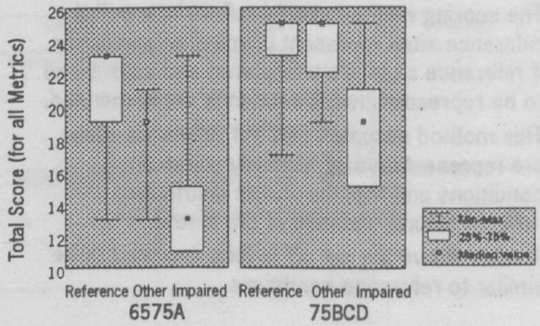
4.10

Impairment Scenarios

- Scenario 1: Toxic substance similar to pesticide. Extirpation of 50% of species of hemimetabolous insects and crustaceans, sel;ected randomly (emphemeroptera, plecoptera, odonata, hemiptera, crustacea).
- Scenario 2: Toxic substance similar to metals. Extirpation of 50% of species that are suspension feeders or deposit feeders. Includes representatives of most insect orders and non-insect phyla.

4.11

Discrimination of the Invertebrate Stream Index for Florida



4.12

Bioassessment Decision Criteria

Example Criteria (a)	Biological Condition Category	Attributes
>83%	Non-impaired	Optimum community structure (composition & dominance) for stream size and habitat quality in ecoregion.
54-79%	Slightly impaired	Composition (species richness) lower than expected due to loss of some intolerant forms. Percent contribution of tolerant forms increases.
21-50%	Moderately impaired	Fewer species due to loss of important forms. Reduction in EPT index.
<17%	Severely impaired	Few species present. If high densities of organisms, then dominated by one or two taxa.

(a) percent of reference

4.13

Research Issues

- Index performance
- Spatial and seasonal variability
- Optimal field methods
- Annual variability
- Optimal taxonomic level

5.1

Spatial Variability

- within habitat
  - (meters)
- among habitats
  - (meters to tens of meters)
- among sampling sites within stream
  - (100s to 1000s of meters)
- among streams within class
  - (kilometer and up)

5.2

Temporal Variability

- Seasonal: requires selection of index period when target organism populations are least variable; i.e., before or after major recruitment or spawning events.
- Annual: requires annual monitoring of subset sites to characterized annual variability.

5.3

Optimal Sampling

- Replication is expensive! Allocation of sampling effort requires careful definition of the questions being asked of the monitoring program, and of the sampling unit.
- A sample observation seeks to characterize its sampling unit, and should be done in a way to minimize the variability of that characterization. Typically, a composite sample form multiple habitats or multiple net hauls is the most cost-effective way to characterize the unit.

5.4

Repeated Observations

- It is usually more cost-effective to sample more units (sites) than to repeat samples at a site, but there are important exceptions:
- Exceptions
  - Compliance or attainment studies of index sites
  - Quality control replication to estimate measurement error (typically 10%)

5.5

BIOASSESSMENT USING PREDICTIVE MULTIVARIATE MODELS: CLARITY, NOT SMOKE AND MIRRORS

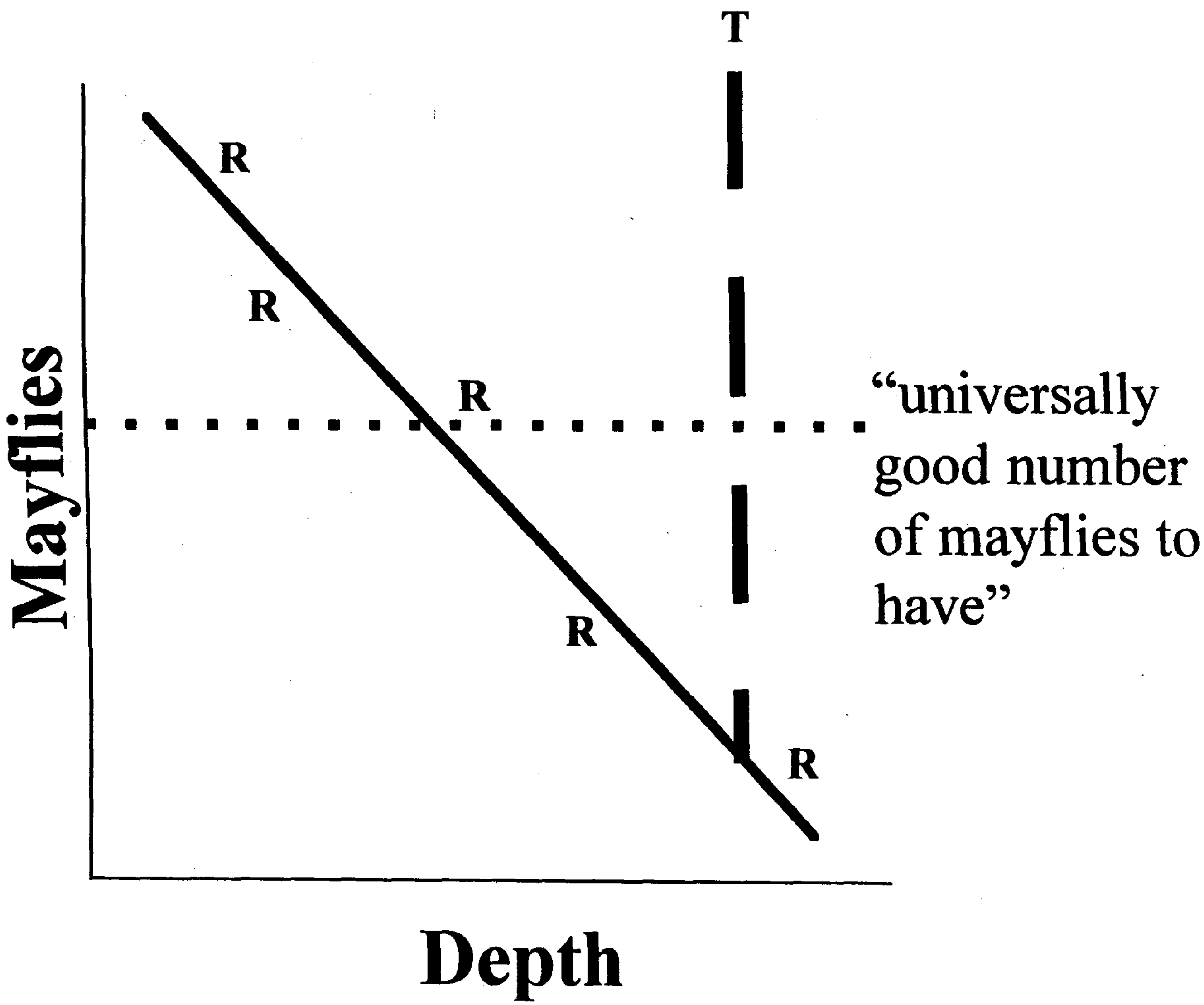
Robert Bailey
Department of Zoology
The University of Western Ontario
London, Ontario, Canada N6A 5B7

NOTES:

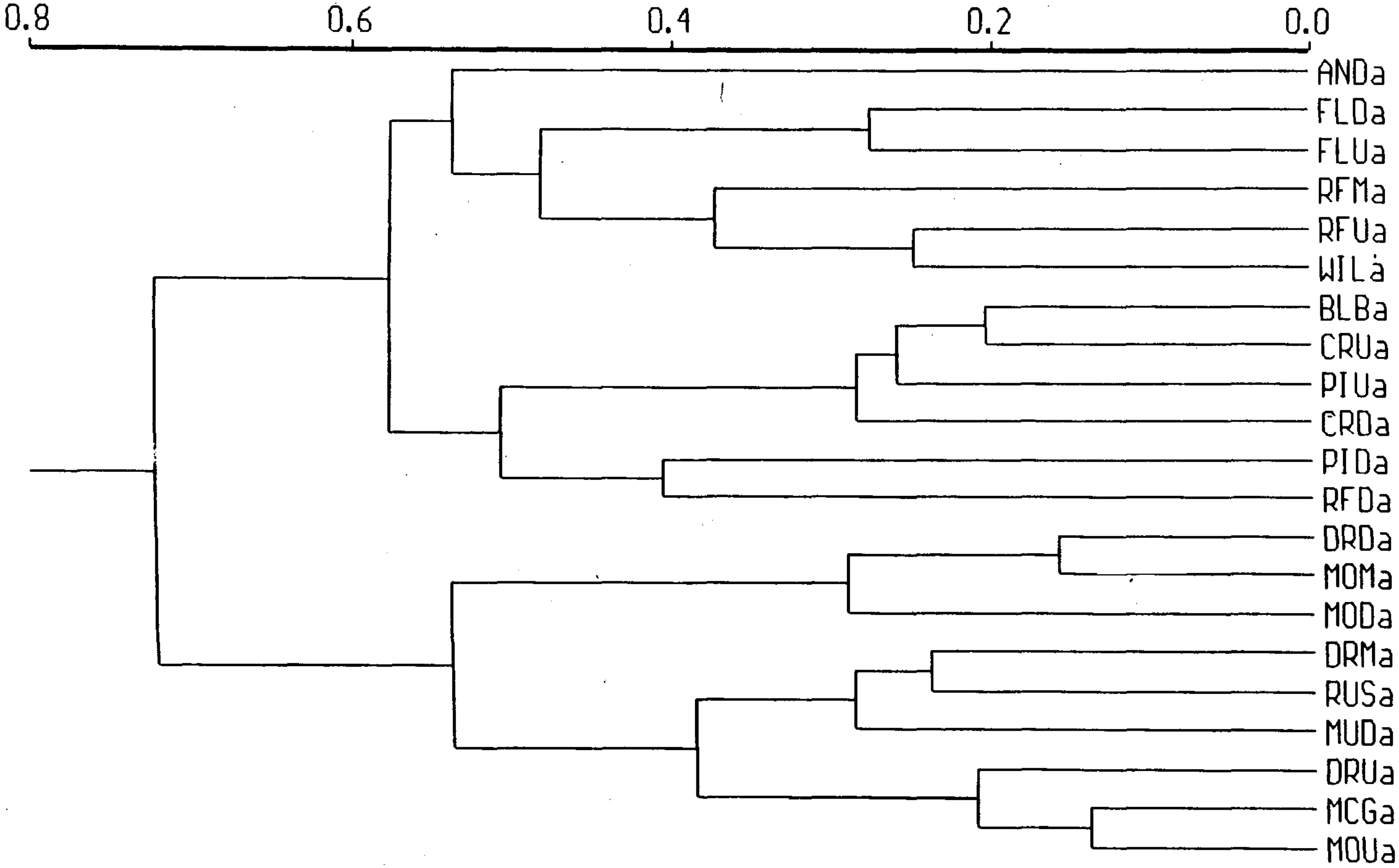
The primary consideration in bioassessment should be that "good is variable." Regardless of the temporal and spatial scale of the assessment, we cannot arbitrarily decree that the organisms we find at one particular site are representative of a healthy community. It is also unrealistic to test real communities against a hypothetical community that only exists in a benthologist's fantasy streamwalk. Thus, the first step in our approach is to define criteria for Reference sites, and then sample them and describe the variation among their biological communities. We also describe the habitat of these same Reference sites, and quantify correlations between the habitat and the community. Ultimately, we need to assess one or more "Test" sites. We can use correlations between the habitat and community of the Reference sites to predict the structure of a Test site's community from its habitat. This answers the question posed in most bioassessment studies... "Is the community at the Test site close to what we would expect if it was one of the Reference sites?" Comparing the predicted to the actual community allows us to assess human impacts. If we see a community similar to that predicted, the site "passes." If the community is quite different from the prediction, the site "fails."

None of this approach necessarily requires multivariate statistics, but it takes more than one variable to adequately describe either a community or its habitat. Once we decide which variables to use for such a description (e.g., family-level abundances or biomasses for communities; particle-size distribution, flow for habitat), we let their variation and covariation describe the structure. It's risky to count on arbitrary functions of the variables (indices) to do the job, but they may be useful in helping to better describe the structure after it has been revealed using multivariate analysis of the original variables. So far, researchers using the predictive multivariate approach have used some form of cluster analysis to describe the structure of Reference site communities. They have then contrasted the habitats of groups of sites with similar communities using Discriminant Functions Analysis (DFA). Finally, they've taken the predictive equations from the DFA, along with habitat data from a Test site, to predict which group the Test site's community would be in if it was "healthy."

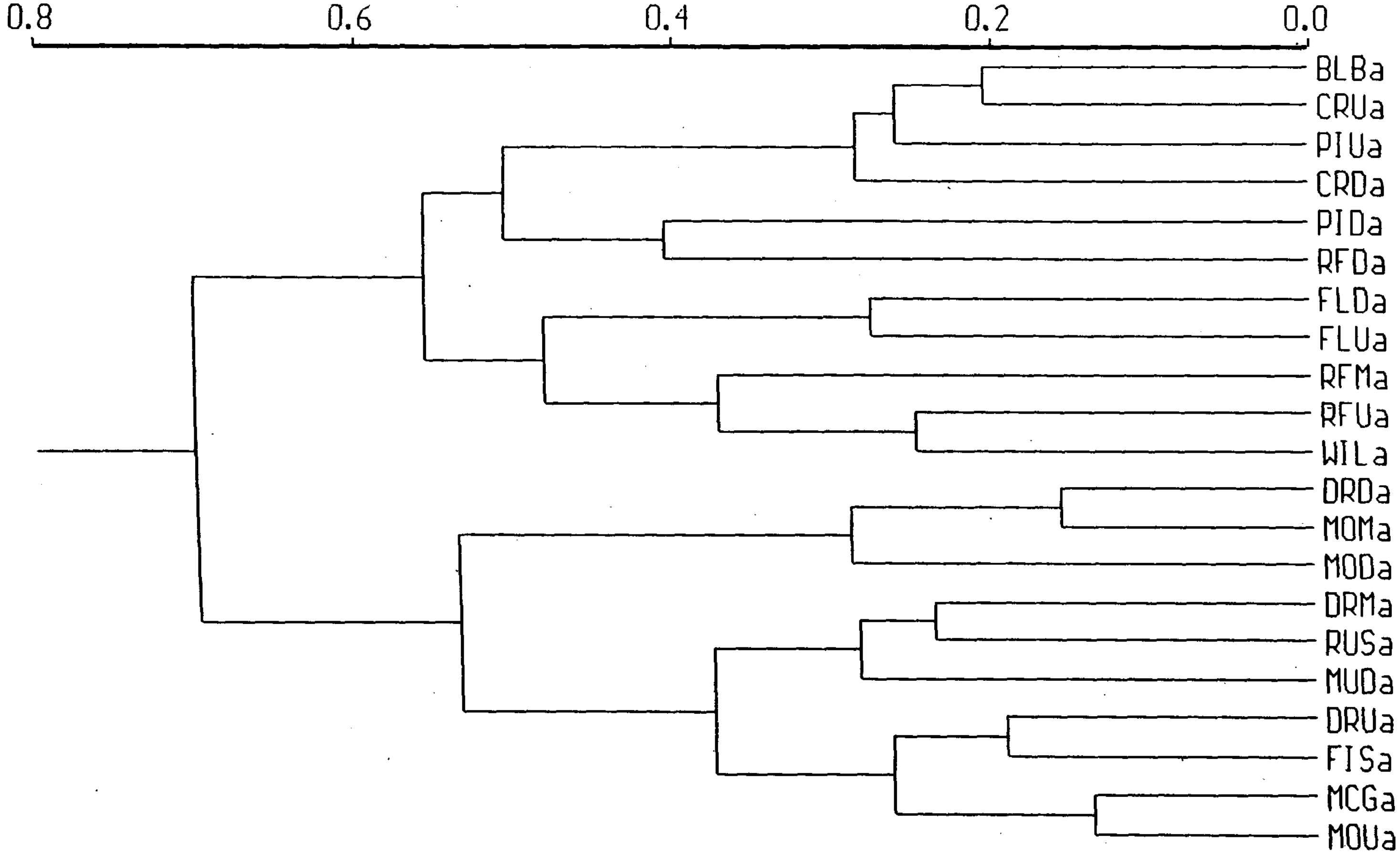
The approach has been used in streams across the United Kingdom (RIVPACS), the North American Great Lakes (BEAST), and Yukon streams (no acronym yet!). I'll look at some results from each of these studies to illustrate and provoke discussion about the techniques.


T
R
R
R
R
R
Mayflies
Depth
“universally good number of mayflies to have”

Yukon Streams: ANDERSON CK. (INDEX=1.34)



Yukon Streams: FISHER CK (INDEX=0.42)



Australian Journal of Ecology (1995) **20**, 198–219

Biological guidelines for freshwater sediment based on BEnthic Assessment of SedimenT (the BEAST) using a multivariate approach for predicting biological state

TREFOR B. REYNOLDSON,[1] R. C. BAILEY,[2] K. E. DAY[1] AND R. H. NORRIS[3]

[1]*National Water Research Institute, Environment Canada CCIW, Burlington, Ontario L7R 4A6, Canada,* [2]*Department of Zoology, University of Western Ontario, London, Ontario, Canada and* [3]*University of Canberra Water Research Centre, Canberra, Australia*

Abstract This paper describes the first results for an alternative approach to the development of sediment quality criteria in the nearshore areas of the Laurentian Great Lakes. The approach is derived from methods developed in the United Kingdom for establishing predictive relationships between macroinvertebrate fauna and the physico-chemistry of riverine environments. The technique involves a multivariate statistical approach using (i) data on the structure of benthic invertebrate communities, (ii) functional responses (survival, growth and reproduction) in four sediment toxicity tests (bioassays) with benthic invertebrates; and (iii) selected environmental variables at 96 reference ('clean') sites in the nearshore areas of all five Great Lakes (Lakes Superior, Huron, Erie, Ontario and Michigan). Two pattern recognition techniques (using the computer software package PATN) are employed in the analysis: cluster analysis and ordination. The ordination vector scores from the original axes of the pattern analysis are correlated (using CORR in SAS) with environmental variables which are anticipated to be least affected by anthropogenic activities (e.g. alkalinity, depth, silt, sodium etc.). Multiple discriminant analysis (MDA) is used to relate the site groupings from the pattern analysis to the environmental variables and to generate a model that can be used to predict community assemblages and functional responses at new sites with unknown but potential contamination. The predicted community assemblages and functional responses are then compared with the actual benthic communities and responses at a site, and the need for remedial action is determined.

The predictive capability of the discriminant model was confirmed by performing several validation runs on subsets of the data. An example of the use of the model for sediment in Collingwood Bay (an area of concern designated by the IJC in Georgian Bay, Lake Huron) is presented and the technique is shown to be more precise in determining the need for remediation than the currently used provincial sediment quality criteria based on Screening Level Concentration (SLC) and laboratory toxicity tests. The ultimate goal of the study is the development of a method to determine the need for, and the success of, remedial action and to predict what benthic communities should look like at a site if it were clean and what responses of organisms in sediment toxicity tests constitute an acceptable end-point.

Key words: benthic, communities, criteria, invertebrates, multivariate analysis, sediment, water quality.

INTRODUCTION

Environmental managers and regulatory decision makers have traditionally set water and sediment quality guidelines based on chemical concentrations. The primary advantage of a chemical approach is the apparent ease of simple numerical comparison of concentrations of chemicals found in environmental matrices with levels of these same compounds known 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 & van der Schalie 1980; Long &

Accepted for publication October 1994.

Chapman 1985; Chapman 1986; Chapman 1990) or because it is so excessively rigorous that it has limited value (Painter 1992; Zarull & Reynoldson 1993).

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, methods developed in the United Kingdom (Wright *et al.* 1984; Moss *et al.* 1987; Armitage *et al.* 1987; Ormerod & Edwards 1987) and elsewhere (Corkum & Currie 1987; Johnson & Wiederholm 1989) 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 approach 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 & Zarull 1993), the USA (Hunsaker & Carpenter 1990) and the United Kingdom (the RIVPACS method; Wright *et al.* 1984) and recent initiatives in Australia (R. H. Norris, pers. comm.).

This paper describes the development of biological objectives for sediments in nearshore habitats in the North American Great Lakes using a modification of the technique developed in the UK (Wright *et al.* 1984; Furse *et al.* 1984; Armitage *et al.* 1987). A large data base is being assembled from reference sites in Lakes Ontario, Erie, Michigan, Superior and Huron and includes information on: (i) the structure of the benthic invertebrate communities; (ii) measured environmental variables and (iii) the responses of four species of benthic invertebrates (*Hyalella azteca*, *Chironomus riparius*, *Hexagenia* spp. and *Tubifex tubifex*) exposed in the laboratory to sediment collected from the same sites. Benthic invertebrates were selected as the most appropriate biological indicators because they are most directly associated with contaminants in sediments through their feeding and behavioral activities. Laboratory sediment testing was included in addition to estimates of benthic invertebrate community structure, to identify the sediment rather than the water column or other physical disturbances as the cause of the observed effect at any given site. These data are being used to develop numeric biological sediment objectives for the Great Lakes and the results presented in this paper incorporate data collected from 96 out of a potential 250 reference sites sampled during 1991–93.

METHODS

Reference sites

The study area encompasses the entire basin of the Great Lakes. To ensure the range of habitat characteristics were adequately represented, a preliminary list of 250 sites were identified and stratified among 17 ecoregions described by Wickware and Rubic (1989) for the Great Lakes. These ecoregions are defined from characteristics such as climate, vegetation, bedrock geology, flora and fauna et cetera. The reference sites were selected to represent 'unpolluted' conditions within an ecodistrict and the inclusion of each site required it to meet the following criteria: the site be located well away (>10 km) from known discharges as described in the Ontario Intake and Outfall Atlas (Ontario Ministry of Environment 1990); the site be located within 2 km of the shore and at a depth of less than 30 m (with the exception of Lake Michigan) and; the site be known or suspected to have a fine-grained substrate. The sites were sampled over a 3 year period (1991–93). This paper presents preliminary results from the first 2 years of the study (i.e. 50 sites from 1991 and 46 sites from 1992; Fig. 1).

General

The location of each site was established in the field using either Loran C or a hand-held Geographical Positioning System (GPS). At each reference site, samples were taken of sediment, water and pore-water for chemical and physical analysis; in addition, samples were collected for the determination of the community structure of benthic macroinvertebrates and for laboratory sediment bioassays with selected species of benthic invertebrates. Each site was sampled once in late summer or early fall over a 3 year period. In addition, a sub-set of sites (10%) were sampled in each of the three field years, and four sites have been sampled monthly over 2 years. This will allow a subsequent determination of the effects of both annual and seasonal variation on the outcome of the predictions of community structure and toxicity. This will be the subject of later publications.

Environmental variables

A list of the variables measured at a site is presented in Table 1. Samples for water chemistry were taken using a Van Dorn sampler from 0.5 m above the sediment–water interface. A 1 L sample was stored at 4°C prior to analysis of total phosphorus, Kjeldahl nitrogen, nitrate–nitrite and alkalinity at the National Water Research Laboratory in Burlington, Ontario, Canada. Measurements of pH, dissolved oxygen and temperature were made in the field.

Sediment and sediment pore-water were characterized from samples taken from a mini-box core. The mini-box core takes a 40×40 cm section of sediment to a depth of 25–30 cm. Samples for geochemical analysis were taken from the surface 2 cm of the box core. After sampling the sediment was homogenized in a glass dish with a nalgene spoon. The sample was divided as follows.

(1) An aliquot of sediment for organic contaminants

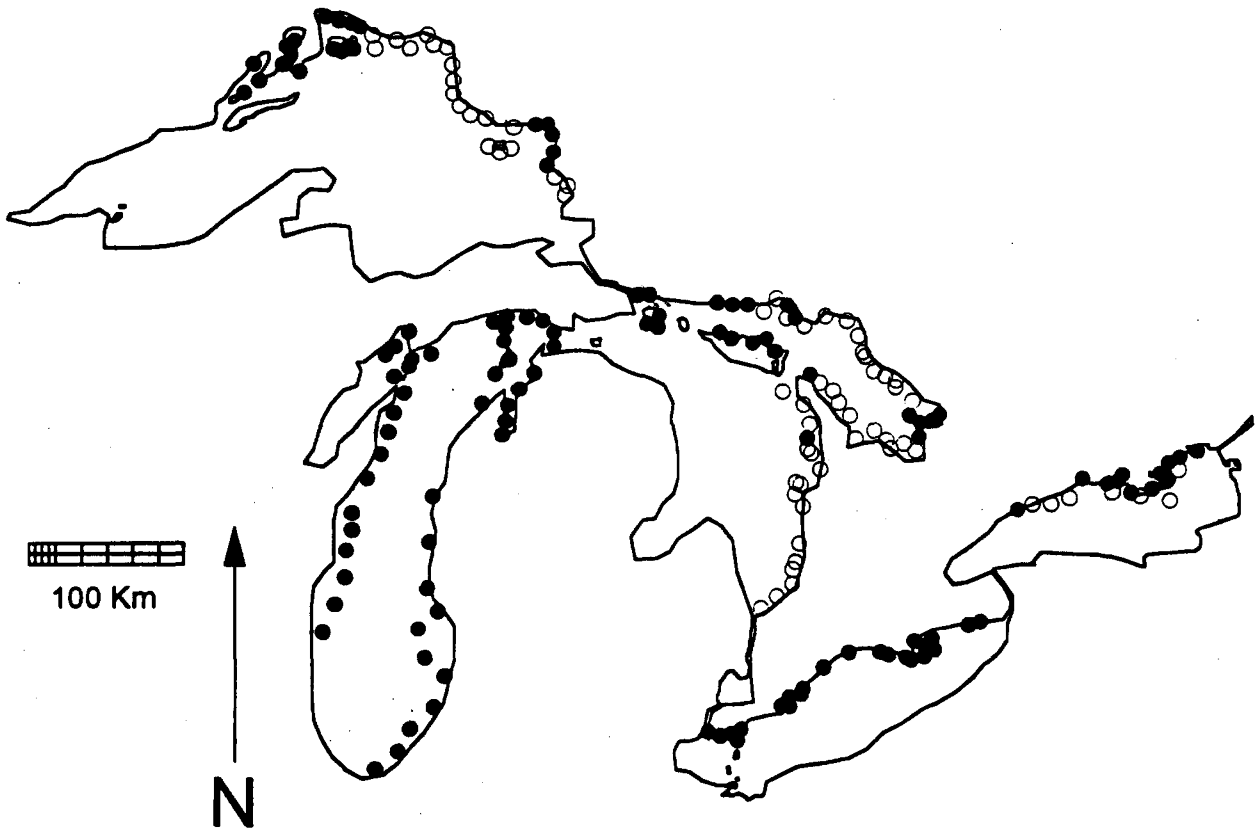


Fig. 1. Location of reference sites in the Great Lakes. (●) Sites included in analysis; (○) sites sampled since analysis.

Table 1. Summary of measured environmental variables and abbreviations used

Field (6 variables)	Water (mg/L) (5 variables)	Sediment (ug/g dry wt) (32 variables)	
Latitude (LAT)	**Alkalinity (AKW)**	**Silica (SI)**	Cobalt
Longitude (LON)	**T. phosphorus (TPW)**	Titanium	Nickel
Water depth (m) (DP)	**Kjeldahl Nitrogen (TKN)**	**Aluminium (Al)**	Copper
Oxygen (mg/L) (OXW)	**Nitrate–nitrite (NOW)**	**Iron (Fe)**	Zinc
Bottom temperature (°C) (TMW)	Ammonia	**Manganese (Mn)**	Arsenic
pH (PHW)		**Magnesium (Mg)**	Strontium
		Calcium (Ca)	Yttrium
		Sodium (Na)	Molybdenum
		Potassium (K)	Silver
		T. Nitrogen (TN)	Cadmium
		T. Phosphorus (TP)	Tin
		T. Org. Carbon (TOC)	Lead
		Loss on ignition (LOI)	**% Gravel (GR)**
		Selenium	**% Sand (SN)**
		Vanadium (V)	**% Silt (SL)**
		Chromium	**% Clay (CL)**

Variables in bold were considered for use as predictors in MDA.

was placed into a hexane-prewashed glass bottle with a hexane rinsed aluminium foil liner. Samples were sealed and stored frozen (or at 4°C in the field) for subsequent freeze-drying and storage. These samples were not normally analysed but were archived in the event of a site being suspected as contaminated.

(2) Samples for the determination of particle size distribution were placed into a plastic pill jar and stored at ambient temperature in the field. Upon return to the laboratory, samples were freeze-dried and analysed following the method of Duncan and LaHaie (1979).

(3) The remaining sediment in the glass dish was stored in a 500 mL plastic container at 4°C in the field and shipped to the National Water Research Laboratory for freeze drying and analyses for metals, major ions and nutrients.

Invertebrate community structure

Samples for the identification and enumeration of benthic invertebrates were taken by inserting five 10 cm plexiglass tubes (internal diameter 6.6 cm) into the sediment in the box core. Each core tube was considered to be a replicate sample unit and was removed and the contents placed into a plastic bag and kept cool until sieved. The contents of each bag were sieved through a 250 μm mesh in the field as quickly as possible after sampling. If sieving could not be done in the field, 4% formalin was added to the bag and the replicate samples were stored at 4°C and sieved as soon as possible thereafter. After sieving the samples were placed in plastic vials (50 mL) and preserved with 4% formalin. Replicates with large amounts of organic material were placed in larger containers and again preserved with 4% formalin. After 24 h the formalin was replaced by ethanol.

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. The confirmed voucher specimens are being maintained as a reference collection.

Sediment toxicity

A mini-ponar sampler was used to obtain five replicate field samples of sediment for laboratory bioassays with four species of invertebrates. Each replicate sample was placed in a plastic bag and held at 4°C until tests could be conducted.

Tests were conducted, in sets of six to seven, over a period of approximately 6 months. A clean control sediment from the Canadian Wildlife Bird Sanctuary, Long Point, Lake Erie was also tested with each set of samples to provide biological quality assurance. Complete details of the culture of organisms and conditions for each toxicity test with *C. riparius* and *T. tubifex* are described elsewhere (Reynoldson *et al.* 1991; Day *et al.* 1994; Reynoldson *et al.* 1994). Culture of *H. azteca* was conducted according to the procedure described in Borgmann *et al.* (1989). Eggs of the mayfly, *Hexagenia* spp. (both *H. limbata* and *H. rigida*), were collected during late June and July in 1991 according to the method of Hanes and Ciborowski (1992) and organisms were cultured using the procedure of Bedard *et al.* (1992). Tests with *H. azteca*, *C. riparius* and *T. tubifex* were conducted in 250 mL glass beakers containing 60–100 mL of sieved (500 μm mesh), homogenized sediment with approximately 100–140 mL of overlying carbon-filtered, dechlorinated and aerated Lake Ontario water (pH 7.8–8.3, conductivity 439–578 μohms/cm, hardness 119 to 137 mg/L). Tests with the mayfly, *Hexagenia*, were conducted in 1 L glass jars with 150 mL of test sediment and 850 mL overlying water. The sediment was allowed to settle for 24 h prior to addition of the test organisms. Tests were initiated with the random addition of 15 organisms per beaker for *H. azteca* and *C. riparius*, 10 organisms per jar for *Hexagenia* spp. and four organisms per beaker for *T. tubifex*. Juveniles of *H. azteca* were 3 to 7 days old at test initiation; *C. riparius* larvae were first instars and were approximately 3 days post-oviposition; *Hexagenia* nymphs were 1.5 to 2 months old (approximately 5 to 10 mg wet weight) and *T. tubifex* adults were 8 to 9 weeks old. Tests were conducted at 23 ± 1°C with a 16L:8D photoperiod (*T. tubifex* 24 h dark). Tests were static with the periodic addition of distilled water to replace water lost due to evaporation. Each beaker was covered with a plastic petri dish with a central hole for aeration using a Pasteur pipette and air line. Dissolved oxygen concentrations and pH were measured at the beginning, middle and end of each exposure period. Tests were terminated after 10 days for *C. riparius*, 21 days for *Hexagenia* and 28 days for *H. azteca* and *T. tubifex* by passing the sediment samples through a 500 μm mesh sieve. Sediment from the *T. tubifex* test was passed through an additional 250 μm mesh sieve at test completion. End-points measured in the tests were survival and growth for *C. riparius*, *Hexagenia* spp. and *H. azteca* and for *T. tubifex* survival and production of cocoons and young. Mean dry weights of *H. azteca*, *C. riparius* and *Hexagenia* spp. were determined after drying the surviving animals from each treatment replicate as a group to a constant weight in a drying oven (60°C).

Data analysis

The analytical strategy used is similar to that proposed by Wright *et al.* (1984). Pattern analysis was used to describe the biological 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.

Classification of biological data

The biological structure of the data was examined using two pattern recognition techniques, cluster analysis and ordination. The mean values from the five replicates for the species counts were used as descriptors of the benthic invertebrate community. These community data were not transformed and the raw scores were used as we considered numeric differences to be important community descriptors. The Bray and Curtis association measure was used because it performs 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 multidimensional scaling (MDS) method of ordination was used (i.e. Semi-Strong-Hybrid multidimensional scaling; Belbin 1991). Multi-dimensional scaling methods use metric and non-metric rank order rather than metric information and thus provide a robust relationship with ecological distance and do not assume a linear relationship, which is 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).

Correlation of biological data with environmental characteristics

Of the 43 environmental variables measured in this study, 25 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, particle size and organic material as a potential indicator of nutritive quality. These together with physical attributes such as water depth and general water chemistry were considered to be the most appropriate general habitat descriptors that are not as subject to modification from human activity.

The relationship with the biological data was examined by correlation analysis of environmental characteristics with ordination axis vector scores using the procedure CORR in SAS.

Prediction of biological groupings

Based on the results from correlation analysis, a suite of 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.

Validation of discriminant model

To test the predictive capability of the discriminant model, five validation runs were performed. In each validation test, ten sites were randomly removed from the reference data set. The discriminant model was calculated for the remaining sites and tested on the ten removed sites. The predicted grouping and its probability were compared with the actual group identified from the initial classification analysis.

Testing of model: Collingwood Harbour case study

As a demonstration of the practical application of this approach, it was used to determine the need for remedial dredging in Collingwood Harbour, Georgian Bay, Lake Huron. This harbour has been identified as an 'Area of Concern' (International Joint Commission 1987) in part because of sediments defined as contaminated using chemical guidelines (Persaud *et al.* 1992). Twenty-five sites were sampled and compared with the reference sites.

RESULTS

Classification of sites

At 93 reference sites, 103 species and 44 genera were identified. Certain groups were not identified below higher taxonomic levels, such as the Porifera, Platyhelminthes and Empididae. The most diverse taxonomic groups were the Chironomidae (43 genera), the Oligochaeta (37 species) and the Mollusca (36 species). Because of the large number of taxa (150) that were available for classification analysis we reduced the data set by including only those taxa with an abundance equal or greater to 0.05% of the total number of organisms (excluding the Porifera). This is because large numbers of rare species tend to add noise to the output from classification analysis.

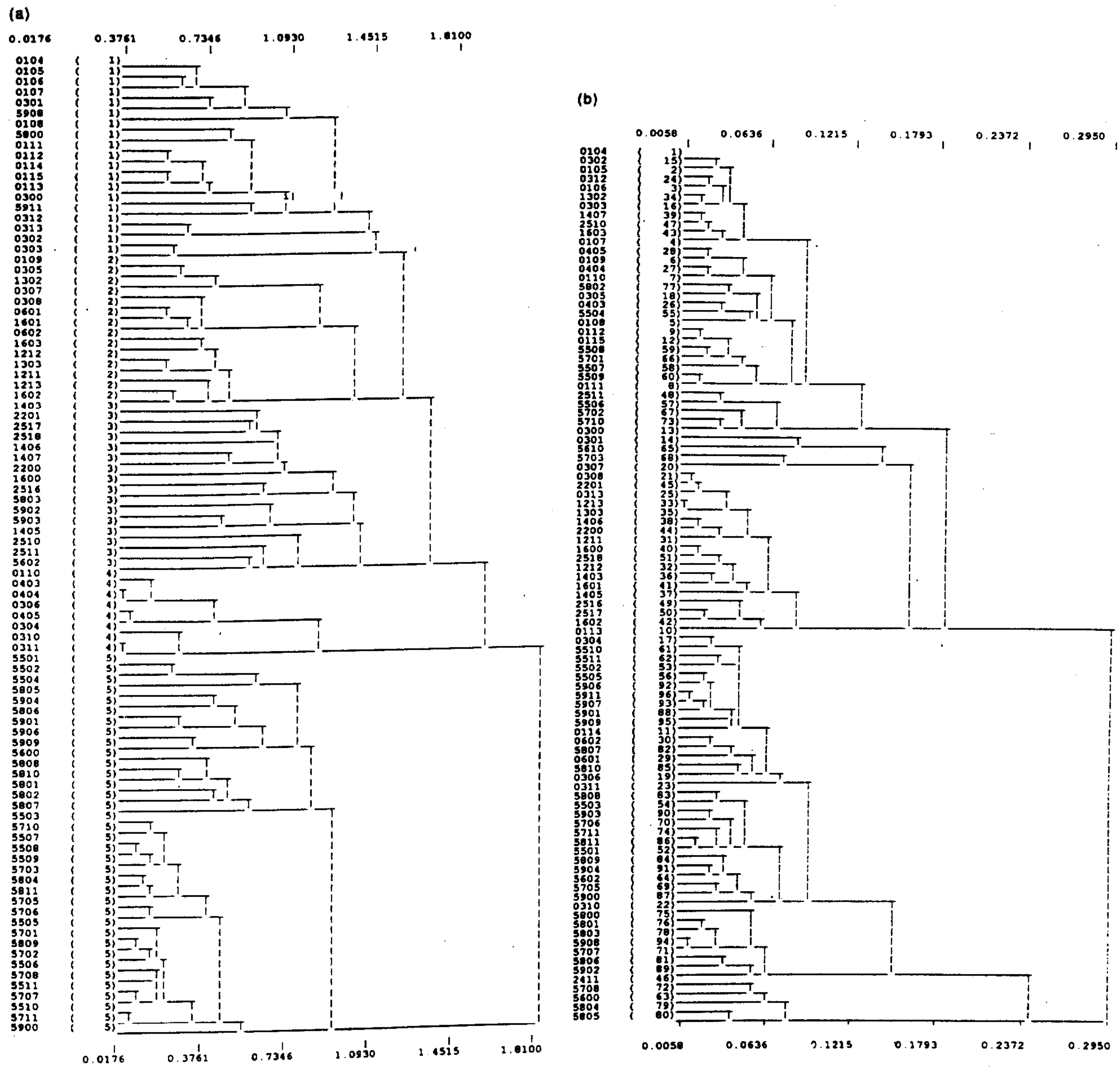


Fig. 2. Dendrogram of reference sites based on (a) invertebrate community structure and (b) toxicity.

Table 2. Geographic distribution of sites in five groups from classification of benthic invertebrate community structure

	Gp 1 (19 sites)	Gp 2 (14 sites)	Gp 3 (16 sites)	Gp 4 (8 sites)	Gp 5 (36 sites)
Lake Ontario		2		3	
Lake Erie	16	4		5	
Lake Huron		8	1		
Georgian Bay			6		
North Channel			5		
Lake Superior			4		36
Lake Michigan	3				

This reduced the number of taxa used in the analysis to 55 (Appendix I).

The results of the cluster analysis for the 93 sites are shown in Fig. 2a. The first group of sites to be distinguished are 36 sites, identified as Gp 5, from Lake Michigan. The remaining sites fall into four groups, beyond which the structure breaks down and small groups of sites begin to form. The geographical distribution (Table 2) of the sites in these five groups suggests that there is a strong spatial signal in the observed grouping. The sites forming Gp 1 are predominantly from mesotrophic Lake Erie together with three shallow sites from Lake Michigan. Eight of the nine Lake Erie sites not included in Gp 1, are from Long

Table 3. List of species that occur in at least 50% of sites in a group, ordered by declining frequency of occurrence

Gp 1	Gp 2	Gp 3	Gp 4	Gp 5
Tubificidae (co h)	**Porifera**	**Tubificidae (co h)**	**Porifera**	**_Diporeia hoyi_**
Tubificidae (c h)	**Tubifidae (co h)**	Tubificidae (c h)	**_Procladius_ spp.**	**_Stylodrilus heringianus_***
Procladius spp.	**_Procladius_ spp.**	_Procladius_ spp.	**_Chironomus_ spp.**	**_Pisidium_ spp.**
Pisidium spp.	**_Cryptochironomus_ spp.**	_Pisidium casertanum_	**_Dicrotendipes_ spp.***	**_Vejdovskella intermedia_***
Pisidium casertanum	**_Chironomus_ spp.**	_Diporeia hoyi_	_Microtendipes_ spp.*	Platyhelminthes
Porifera	**_Pisidium casertanum_**	_Micropsectra_ spp.*	_Cryptochironomus_ spp.	_Heterotrissocladius_ spp.*
_Spirosperma ferox_*	**Tubificidae (c har)**		Tubificidae (co h)	_Pisidium casertanum_
_Dreissenia polymorpha_*	**_Tanytarsus_ spp.***		_Physella_ spp.*	Tubificidae (co h)
Platyhelminthes	_Valvata tricarinata_*		_Polypedium_ spp.	Tubificidae (c h)
_Limnodrilus hoffmeisteri_*	_Aulodrilus pigueti_		_Pisidium casertanum_	
Chironomus spp.	Platyhelminthes		_Pisidium nitidum_*	
Aulodrilus pigueti	_Pisidium_ spp.		_Endochironomus_ spp.*	
	Polypedium spp.		_Pseudochironomus_ spp.*	
			Pisidium spp.*	
			Aulodrilus pigueti	
			_Amnicola limosa_*	

Species in bold have >70% occurrence. Taxa marked with * are unique to that group of sites.

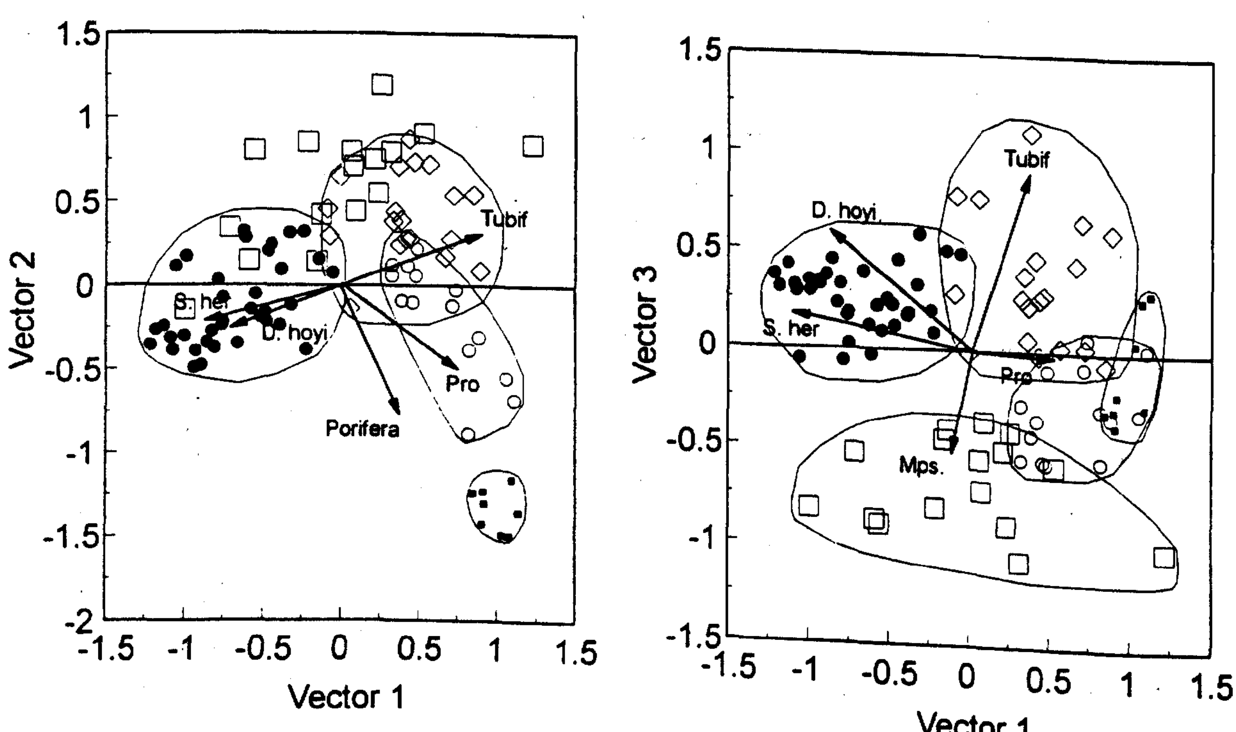


Fig. 3. Ordination of reference sites based on invertebrate community structure. (◇) Gp 1; (○) Gp 2; (□) Gp 3; (■) Gp 4; (●) Gp 5.

Point Bay and are shallow inner bay sites classified as Gp 4. Group 4 also includes three shallow sites from Presque'Ile Bay in Lake Ontario. The Gp 3 sites are mostly from the oligotrophic North Channel of Lake Huron and Lake Superior. Group 2 includes southern Georgian Bay sites together with a few Lake Erie and Lake Ontario sites.

The more common species present in each of the five groups, those found at 70% and 50% of the sites in a group, are shown in Table 3. In Gps 1 and 3, the Tubificidae are the most frequently occurring taxa and in Gp 1, which represents the more mesotrophic sites, the oligochaete species *Spirosperma ferox*, *Limnodrilus hoffmeisteri* and *Aulodrilus pigueti* are commonly found, as are *Chironomus* spp. These species are all typically associated with greater amounts of organic material. In contrast, Gp 3 and, particularly, Gp 5 have species that are characteristic of oligotrophic conditions; for example, the amphipod *Diporeia hoyi* and the lumbriculid worm *Stylodrilus heringianus* are important components of the benthic assemblage at the sites which make up these two groups. The shallow sites (Gp 4) are characterized by a community dominated by the Porifera (sponges) and a large number of chironomid species. Some taxa are common to all the groups, notably the Tubificidae and *Pisidium casertanum*, and the presence of these organisms in a group is not a useful indicator; however, their abundance does vary between groups. Other taxa are restricted to single groups: *Spirosperma ferox*, *Limnodrilus hoffmeisteri* and *Dreissenia polymorpha* are common in Gp 1 only; *Tanytarsus* spp. and *Valvata tricarinata* in Gp 2; *Microspectra* spp. in Gp 3; several chironomid species and molluscs in Gp 4; and *Stylodrilus heringianus* and *Heterotrissocladius* spp. in Gp 5.

The location of the reference sites in ordination space is shown in Fig. 3. With the exception of the sites forming Gp 3, there is good discrimination on the first two ordination vectors. The Gp 3 sites are separated on the third vector. The relative contribution of the species to the ordination vectors has been determined by principal axis correlation (Table 4). Those taxa contributing most to the pattern of site distribution are indicated by arrows in Fig. 3; the direction of the arrow indicates the direction of the loading and the length of the arrow the importance of the taxa. Sites with low scores on vector 1 and located on the left of the plot (Gp 5) are dominated by the presence of *D. hoyi* and *S. heringianus*. Those sites scoring low on vector two have communities dominated

Table 4. Correlation of taxa with vectors from SSH ordination

Taxa	r	Taxa	r
S. heringianus	0.7465	*Heterotrissocladius* spp.	0.3203
Tubificidae A	0.7447	*N. variabilis*	0.3076
D. hoyi	0.7446	*P. moldaviensis*	0.2956
Porifera	0.6718	*Chironomus* spp.	0.2941
Cryptochironomus spp.	0.5082	*A. pluriseta*	0.2838
Micropsectra spp.	0.4633	*P. vejodoskyi*	0.2747
Procladius spp.	0.4626	*V. piscinalis*	0.2747
Tubificidae B	0.4378	*L. hoffmeisteri*	0.2727
Gammarus pseudolimnaeus	0.4266	*Glypotendipes* spp.	0.2711
S. lacustris	0.4229	*A. lomondi*	0.2666
Dicrotendipes spp.	0.4157	*Q. multisetosus*	0.2663
Dreissenia polymorpha	0.4146	*Pisidium* spp.	0.2661
V. intermedia	0.4126	*Cryptotendipes* spp.	0.2604
A. pigueti	0.4071	*P. casertanum*	0.2592
Microtendipes spp.	0.4058	*S. josinae*	0.2446
M. speciosa	0.4009	*Cladopelma* spp.	0.2401
Psectrocladius spp.	0.3912	*Chaoborus* spp.	0.2309
Physella sp.	0.3840	*P. henslowanum*	0.2193
Polypedium spp.	0.3821	*Stictochironomus* spp.	0.2147
Endochironomus spp.	0.3816	*T. tubifex*	0.2022
S. ferox	0.3784	*Tanypus* spp.	0.1913
Platyhelminthes	0.3680	*Demicryptochironomus* spp.	0.1837
Coelotanyopus spp.	0.3485	*H. americana*	0.1572
A. limosa	0.3359	*M. securis*	0.1347
P. nitidum	0.3348	*P. compressum*	0.1144
Enchytraediae	0.3312	*Caecidotea* spp.	0.0537
C. intermedius	0.3279		
V. tricarinata	0.3271		

by sponges and *Procladius* spp. While tubificid worms are found at many sites they are most abundant at those sites scoring high on vectors 1 and 3 (Gps 1 and 2).

The same approach was used to define the pattern in the physiological responses of the test organisms from the chronic sediment bioassays. The similarity scores in the dendrogram (Fig. 2b) show that there is less structure in these data compared to that found in the community structure. Examination of the physiological data suggests that a three-cluster solution is most appropriate. Beyond this, it was not possible to relate the biological groupings to the environmental data. In fact, even at this level, there is little difference between the groups for several of the test end-points (Table 5); for example, survival of *C. riparius* and *Hexagenia* spp., growth of *C. riparius* and cocoon production of *T. tubifex*. The end-points which appear sensitive to site differences are survival of *H. azteca*, which was very low in the five sites comprising Gp 3, and growth of both *H. azteca* and *Hexagenia* spp., which is reduced at the Gp 3 sites. Conversely, production of young by *T. tubifex* is highest at these same sites and reduced in the Gp 1 sites.

Table 5. Toxicity test end-points in the three reference site groups defined by classification analysis

End-point	Overall		Gp 1 (n = 53)		Gp 2 (n = 38)		Gp 3 (n = 5)	
C. riparius								
% Survival	82.9	(8.6)	80.2	(7.8)	86.6	(8.6)	82.9	(6.2)
Growth (mg dry wt)	0.34	(0.08)	0.33	(0.08)	0.37	(0.08)	0.30	(0.04)
H. azteca								
% Survival	83.7	(19.0)	88.7	(11.8)	84.7	(12.9)	24.5	(22.4)
Growth (mg dry wt)	0.50	(0.14)	0.51	(0.11)	0.50	(0.17)	0.29	(0.11)
Hexagenia								
% Survival	96.9	(4.1)	97.4	(3.4)	96.4	(4.8)	95.5	(4.6)
Growth (mg dry wt)	3.5	(3.6)	3.07	(2.56)	4.43	(4.64)	1.49	(0.31)
T. tubifex								
Cocoons	34.9	(5.8)	32.8	(6.1)	37.6	(4.2)	35.8	(2.8)
Young	87.7	(40.0)	58.5	(23.5)	124.0	(23.0)	122.2	(24.9)

Values are means with SD in parentheses.

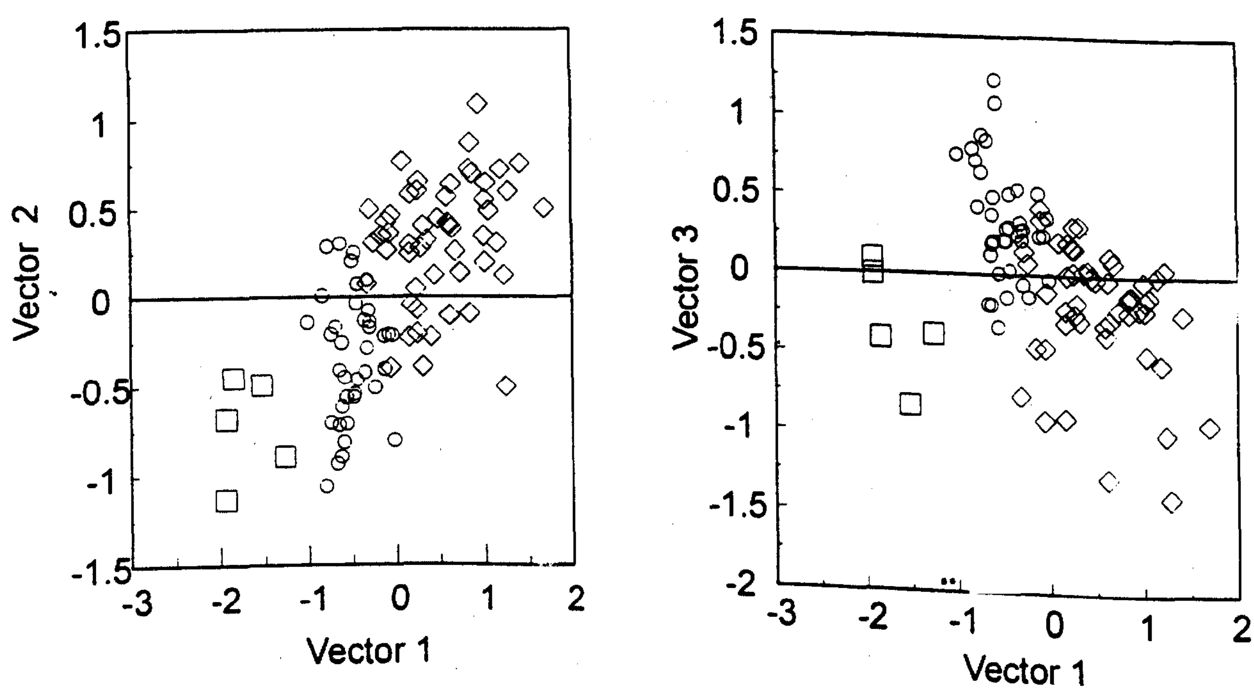


Fig. 4. Ordination of reference sites based on sediment toxicity. (◇) Gp 1; (○) Gp 2; (□) Gp 3.

Table 6. Correlation coefficients between selected environmental variables and ordination vectors from community structure

Vector 1		Vector 2		Vector 3	
TMW	**0.6456**	**SI**	**0.3122**	**AKW**	**0.6990**
TKN	**0.5760**	**V**	**0.2905**	**DP**	**0.4669**
SL	**0.5541**	**K**	**0.2757**	**PHW**	**0.4656**
NA	**0.5035**	NOW	0.2427	**NOW**	**0.4188**
AL	**0.4903**	AL	0.2322	MG	0.2597
TOC	**0.3642**	NA	0.1807	K	0.1984
TN	**0.3617**	TMW	0.1761	OXW	0.1624
CA	**0.3557**	FE	0.1527	CA	0.1237
LOI	**0.3265**	GR	0.1222	SN	0.0964
FE	0.1747	TPW	0.1145	MN	0.0729
V	0.1528	CY	0.0964	LOI	0.0683
CY	0.1358	SN	0.021	SI	0.0086
TPW	0.0763	MN	0.0143	CY	0.0014
K	0.0736	OXW	0.0128	GR	−0.0134
GR	0.0649	MG	0.0017	TP	−0.0208
PHW	0.062	TPO	−0.0002	SL	−0.1376
MN	−0.055	DP	−0.1094	FE	−0.1622
TP	−0.103	SL	−0.1114	TPW	−0.2192
MG	−0.16	TKN	−0.1881	**TKN**	**−0.2781**
OXW	**−0.377**	AKW	−0.2331	**TOC**	**−0.2959**
SI	**−0.464**	**PHW**	**−0.2922**	**TN**	**−0.3149**
SN	**−0.465**	**TOC**	**−0.3088**	**AL**	**−0.3271**
AKW	**−0.476**	**TN**	**−0.3169**	**V**	**−0.4156**
NOW	**−0.711**	**LOI**	**−0.4446**	**NA**	**−0.5347**
DP	**−0.8073**	**CA**	**−0.4815**	**TMW**	**−0.6761**

Each vector ranked from high positive to high negative; variables in bold $P<0.001$.

Table 7. Variables correlated ($P<0.01$) with three ordination vectors from toxicity end-points in rank order

Vector 1	Vector 2	Vector 3
Alkalinity (w)	Alkalinity (w)	Vanadium
Sodium	Aluminium	
Aluminium	Sodium	
Depth	Temperature (w)	
Temperature (w)	Depth	
Vanadium	Sand	
pH (w)	Oxygen (w)	
Iron	Vanadium	
	Nitrate (w)	
	Kjeldahl Nitrogen (w)	
	Silt	
	Iron	
	Silica	
	Clay	

(w) Indicates those variables measured in water.

The ordination of the toxicity data matrix shows a strong axis on the first vector on which the three groups of sites can be separated (Fig. 4a). Sites in the lower left hand corner on vectors 1 and 2 have low *Hyalella* survivorship and high *Tubifex* reproduction and at the top right of the same plot *Tubifex* reproduction is lower.

Correlation with environmental characteristics

To establish the relationship between community structure, toxicity and the environmental characteristics of the sites, correlation coefficients and probabilities were calculated for the 25 variables considered to be the most useful in developing a predictive model (Table 1) and the vector scores from ordination.

The results of the correlation with the first community structure vector (Fig. 3) showed a negative correlation with depth ($r = -0.80733$; Table 6) and nitrate–nitrite, and a positive correlation with temperature, Kjeldahl nitrogen and silt. None of the measured variables was well correlated with the second community vector (Table 6). On the third community vector, which discriminates the Gp 5 sites, water temperature (and depth), sodium and alkalinity were important. In the ordination plot (Fig. 3a), sites toward the left on vector 1 tend to represent deeper, more oligotrophic conditions and sites toward the bottom and right of the plot represent shallower, warmer and more mesotrophic conditions.

The variables best correlated with the toxicity ordination vectors were the major ions. Surprisingly, particle size and the organic carbon content were not correlated with the toxicity ordination vectors (Table 7). Water chemistry descriptors, particularly alkalinity, were well correlated with the bioassay end-points, although the test water is dechlorinated Lake Ontario water. The five sites forming Gp 3 were distinguished by being deeper (mean depth 67 m cf. 20.1 and 38.0 m) with a very low silt content (mean 8.7%) compared with the other two site groups, which have mean silt contents of 24.2% and 33.2%.

Prediction and validation of biological groups

Of the 25 variables examined, only those strongly related ($P<0.0001$) to the ordination vectors from the community structure were considered for use in discriminant analysis for predicting group membership derived from the species data. Water temperature (TMW) was also discarded as it is dependent on the time of sampling and was highly correlated with depth (DP).

Table 8. Environmental variables used to predict community and toxicity group membership

Community predictors	Bioassay predictors
Depth	Silica
Nitrate (w)	Aluminium
Silt	Iron
Aluminium	Manganese
Calcium	Calcium
LOI	Sodium
Alkalinity (w)	Potassium
Sodium	Phosphate
pH (w)	LOI
	Total Nitrogen
	Total Organic Carbon
	Sand
	Silt
	Clay
	Vanadium
	Depth
	Alkalinity (w)
	Feeding regime

(w) Indicates those variables measured in water.

The use of nine variables (Table 8) in the discriminant model produced the best prediction of group membership from community structure data (Table 9). Results showed that 79 of 91 (86.8%) sites were correctly predicted by the nine variables. The discriminant model had the most difficulty in classifying the Gp 2 sites (64.3% correct, 9 of 14) with four of the five incorrectly classified sites being predicted as Gp 1. Group 1 is adjacent to Gp 2 in ordination space and there is considerable overlap in the two groups (Fig. 3). Examination of the discriminant functions (Fig. 5) shows a clear separation of the groups based on community structure. The first discriminant function explained 73.6% of the variance with the greatest contributions being made by nitrate, depth and calcium, respectively. On the second function, which explains a further 14.8% of the variance, loss on ignition (%LOI) and alkalinity were major contributors.

The mean values for the nine environmental variables, for each of the groups, are shown in Table 10. The sites forming Gp 4 are shallow and have the greatest silt content and per cent loss on ignition (LOI). These sites are dominated by the presence of sponges (Porifera) which are found at densities of over 1900 per 100 cm^2. These shallow sites also have the greatest number of frequently occurring organisms (Table 4), particularly chironomids. The deep Gp 5 sites are characterized by the presence of the amphipod, *Diporeia hoyi*, and the lumbriculid oligochaete, *Stylodrilus heringianus*, both of which are typical of more oligotrophic conditions in the Great Lakes.

The same approach was used in classifying the site groups based on toxicity from the environmental data. The measured environmental variables (Table 9) were less able to discriminate between the groups derived from the toxicity test end-points. Only 70.2% of the sites were correctly identified using 18 variables compared

Table 9. The number and percentage of 91 sites predicted to the correct biological group using discriminant analysis with selected environmental variables

Group member	Predicted Gp 1	Predicted Gp 2	Predicted Gp 3	Predicted Gp 4	Predicted Gp 5
Community structure					
1	19 (100%)	0	0	0	0
2	4	9 (64.3%)	0	1	0
3	3	1	11 (73.3%)	0	0
4	0	2	0	6 (75%)	0
5	1	0	0	0	34 (97.1%)
Toxicity					
1	43 (82.7%)	5 (9.6%)	4 (7.7%)		
2	6 (16.2%)	25 (67.6%)	6 (16.2%)		
3	0	0	3 (100%)		

Table 10. Mean (SD) for nine environmental variables from five groups of sites grouped by benthic invertebrate community structure

Variable	Gp 1	Gp 2	Gp 3	Gp 4	Gp 5
Depth m	15.9 (9.3)	8.4 (3.9)	7.2 (4.8)	1.8 (0.6)	57.9 (22.0)
Nitrate (w) mg L^{-1}	0.22 (0.08)	0.10 (0.08)	0.20 (0.07)	0.01 (0.00)	0.33 (0.07)
Silt %	40.3 (19.8)	42.9 (27.0)	27.7 (24.5)	50.7 (24.6)	13.2 (19.4)
Aluminium $\mu g\ g^{-1}$	9.5 (2.2)	10.9 (2.1)	9.2 (3.4)	8.0 (2.1)	6.1 (2.3)
Calcium $\mu g\ g^{-1}$	8.1 (3.3)	6.8 (5.6)	3.6 (3.0)	15.8 (9.4)	5.5 (3.5)
Loss on ignition %	12.4 (2.8)	13.5 (6.8)	7.3 (4.3)	21.9 (11.4)	11.1 (6.4)
Alkalinity mg L^{-1}	92.5 (7.3)	82.2 (11.2)	64.3 (25.5)	78.6 (10.4)	111.2 (4.3)
Sodium $\mu g\ g^{-1}$	1.26 (0.39)	1.8 (0.6)	1.8 (0.8)	1.5 (0.8)	0.8 (0.2)
pH (w)	8.1 (0.2)	7.6 (0.5)	7.3 (0.5)	8.2 (0.1)	7.9 (0.6)

(w) Indicates those variables measured in water.

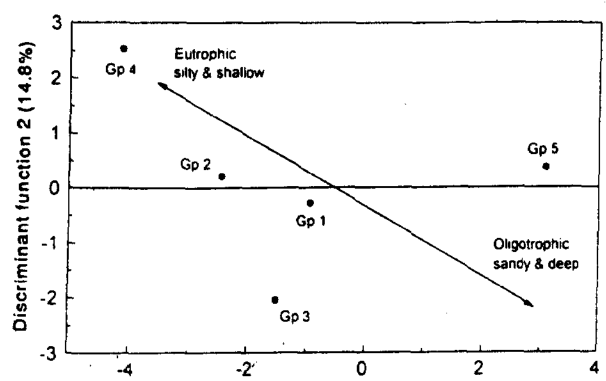


Fig. 5. *Community structure site group means in discriminant space.*

Table 11. Summary of five validation runs of discriminant analysis site predictions

Run	Correct (%)	Group	Correct	
Community structure				
1	80	1	15 of 15	(100%)
2	90	2	4 of 5	(80%)
3	80	3	5 of 6	(83%)
4	100	4	6 of 7	(86%)
5	100	5	15 of 17	(88%)
Overall	90			
Toxicity				
1	80	1	22 of 29	(76%)
2	80	2	11 of 20	(55%)
3	70	3	1 of 1	(100%)
4	80			
5	30			
Overall	68			

with 86.8% for the groups based on community structure using nine variables (Table 8). Again, variables such as water column alkalinity and depth, which correlated well with the ordination vectors, were important contributors to the discriminant functions. This poorer discrimination between the toxicity based groups is primarily due to the greater degree of similarity between them.

While these predictions provide an estimate of the ability of a suite of environmental variables to predict biological structure, a more realistic indicator of predictive capability is provided by using a separate set of test sites. Ten sites were randomly removed from the reference data base. A discriminant model based on the remaining reference sites was developed. The environmental variables for the 10 test sites were then substituted and a biological grouping predicted which was compared with the actual grouping from the original classification. This procedure was repeated separately five times for both the community structure and toxicity data matrices.

For the community validation, between 80 and 100% of the sites were correctly predicted from run to run with an overall average of 90% correct (Table 11). In each of the five runs, the Gp 1 sites were always correctly predicted. In each of Gps 2, 3 and 4, one site was incorrectly predicted overall. The Gp 5 sites were correctly identified 88% of the time. These data suggest high confidence in an assemblage of organisms being correctly predicted at a new site.

The results of the validation runs on the basis of the bioassay responses (Table 11) confirm their reduced predictive ability as only 68% of the test sites were correctly assigned. The discriminant model had the greatest difficulty in predicting the Gp 2 sites (55% correct). To determine the implications of this, we have compared the observed result at the incorrectly predicted sites with the average for the correct group and the average for the predicted group (Table 12). Chironomid survival showed little difference between the three groups; for example, the Gp 1 sites all fell within the actual Gp 1 range and within the range of the predicted group. The Gp 2 sites were also within the Gp 2 range and, with the exception of site 5808, were within the range of the predicted group. Survivorship in site 5808 was slightly higher than for the predicted group. Chironomid growth

Table 12. Values of bioassay endpoints for those sites incorrectly predicted and expected values for each test endpoint

	Expected	CRSU	CRGW	HLSU	HLGW	HASU	HAGW	TTCC	TTYG
	Gp 1	65–96	0.17–0.49	91–100	0.51–5.63	65–100	0.29–0.73	27–39	35–82
	Gp 2	70–100	0.21–0.53	87–100	0.00–9.07	59–100	0.16–0.84	33–42	99–147
	Gp 3	70–95	0.22–0.38	86–100	1.18–1.80	0–69	0.07–0.51	33–39	97–147
Gp 1 sites	Predicted to								
0106	2	80	0.28	96.7	3.83	86.7	0.50	37.4	70.0
0107	2	77.8	0.20	93.3	5.17	98.2	0.62	39.7	93.0
0108	2	88.9	0.21	100	5.28	97.8	0.47	37.7	79.3
5506	3	75.0	0.25	98.0	1.70	60.0	0.49	34.0	82.0
5507	2	88.0	0.31	90.0	2.09	96.7	0.48	31.0	90.0
5710	3	92.0	0.31	100	1.25	69.3	0.36	38.0	83.0
Gp 2 sites	Expected								
0310	1	91.1	0.52	100	15.43	95.6	0.68	47.0	153.7
0311	1	95.6	0.36	96.7	13.99	91.1	0.61	40.3	118.3
5502	3	90.7	0.33	94.0	3.15	92.0	0.46	38.0	113.0
5705	3	92.0	0.33	84.0	1.66	78.7	0.12	38.0	136.0
5807	3	85.3	0.44	100	1.48	83.1	0.54	41.0	100.0
5808	3	97.3	0.40	96.0	2.10	83.1	0.79	40.0	120.0

Expected value ranges are 2 standard deviations about the mean, and 1 standard deviation about the mean for TTCC and TTYG.

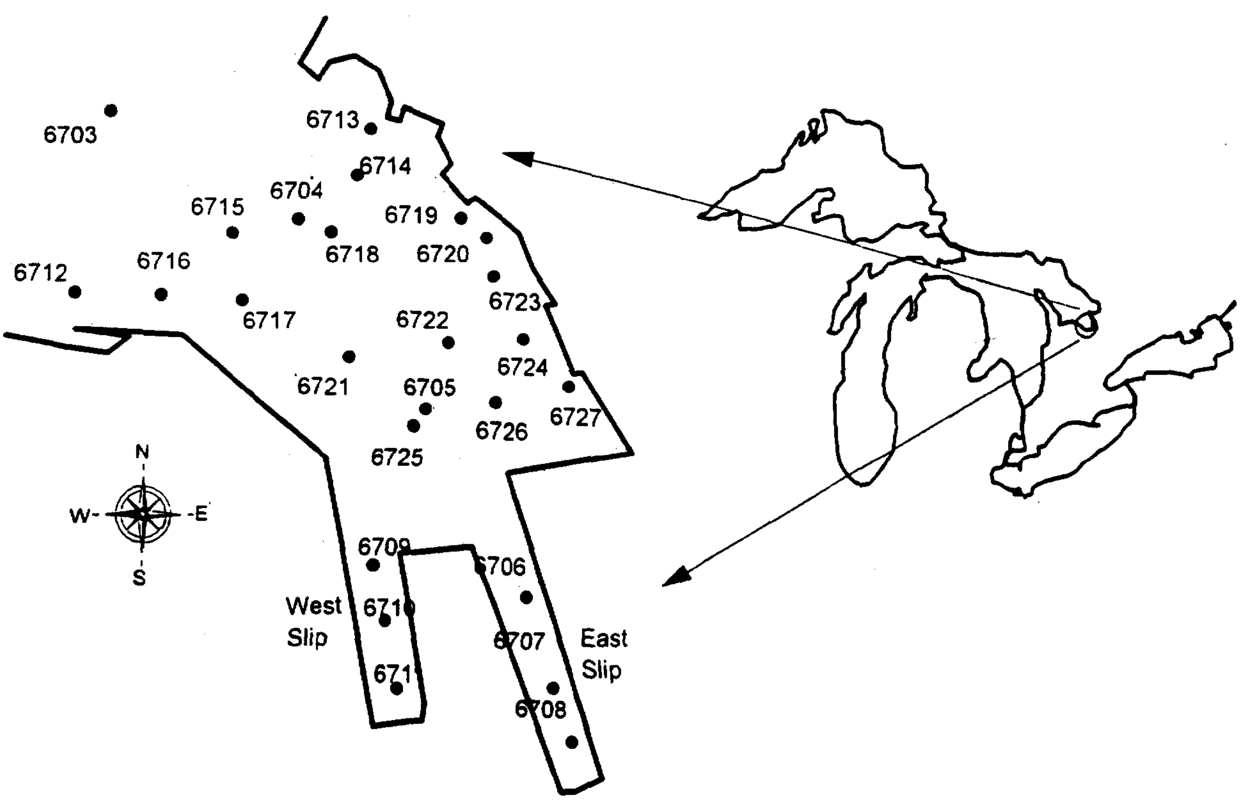


Fig. 6. Location of sampling sites in Collingwood Harbour.

was within the expected range or higher (0310, 5807, 5808), with the exception of site 0107 in which growth was lower than expected for the predicted group (Gp 2) but not for the actual group (Gp 1). This is a case where a site could be deemed toxic when in fact it is not. For *Hexagenia* survival, only site 5705 is lower than predicted.

Hexagenia growth is the most variable of the end-points. None of the Gp 1 sites is affected by the incorrect predictions. Of the Gp 2 sites, 0310 and 0311 show much greater growth than expected but this would not necessarily produce a toxic designation. Survival of *H. azteca* is an end-point that could provide a false toxic designation if not correctly predicted to Gp 3. The one Gp 3 site in the validation tests was correctly identified and site 5506 would not be considered toxic as it was predicted to Gp 3, although the survivorship of *Hexagenia* is lower than expected for Gp 1.

Application in Collingwood Harbour

Collingwood Harbour is located in Collingwood Bay, Georgian Bay, Lake Huron (Fig. 6) and has been identified as an area of concern by the International Joint Commission (International Joint Commission 1987) partly because of sediment contamination by various metals and partly based on eutrophication. As part of a remedial programme for the harbour, sediment removal is being considered. To define the extent of the sediment contamination, it was decided to compare the more generally used chemical criteria with the biological approach developed in this paper.

Based on sediment chemistry and the Province of Ontario's sediment quality criteria, the harbour was heavily contaminated by metals (Table 13), with three sites in the east boat slip (6706, 6070 and 6708) and one site in the west boat slip (6709) exceeding the Ontario Ministry of Energy and Environment's (Persaud *et al.* 1992) severe effects criteria for copper, zinc, lead, arsenic and iron. Furthermore, all the sites in both the boat slips and the outer harbour exceeded low effect concentrations for at least eight sediment variables and several sites for all 12 variables for which criteria have been established. This prompted the Canadian federal government (Environment Canada) and the Province of Ontario to consider removal of the contaminated material. However, the large area of removal and the anticipated cost prompted examination of the biological significance of the contamination and the biological objectives developed in this project were used to help make a managerial decision.

Table 13. Sediment chemistry for contaminants for which guidelines are available

	Total N	Total P	TOC	Total Fe (%)	Mn	Cr	Ni	Cu	Zn	As	Cd	Pb
Low	550	600	1	2	460	26	16	16	120	6	0.6	31
Severe	4800	2000	10	4	1100	110	75	110	820	33	10	230
Site												
6703	975	1095	1.8	1.6	522	25	21	37	129	2.5	0.6	76
6704	2205	1290	2.1	1.5	480	24	21	35	123	2.5	0.5	63
6705	2042	1335	3.5	2.2	570	29	24	101	411	2.5	1.2	329
6706	1524	1230	2.3	11.2	594	65	33	2835	10 780	105	3.9	7.9
6707	1244	945	1.8	14.4	642	85	36	4170	13 943	137	4.9	974
6708	1364	1080	2.1	8.9	646	58	31	2121	7 527	70	4.2	724
6709	1825	1200	2.4	5.0	1114	47	27	893	3 154	25	1.8	430
6710	2441	1515	2.7	3.0	647	40	29	201	669	2.5	1.5	200
6711	2211	1560	2.6	2.2	583	32	25	109	380	2.5	0.4	149
6712	1661	607	2.1	1.5	355	18	12	49	216	7	0.4	105
6713	2499	866	2.6	1.7	491	23	19	36	145	12	<0.2	71
6714	2439	821	2.4	1.7	495	22	20	35	141	9	0.3	71
6715	2371	801	2.5	1.6	460	20	19	32	125	<5	0.6	57
6716	2299	1032	3.1	1.6	405	25	17	65	247	6	0.4	88
6717	2431	850	2.9	1.5	444	20	17	35	151	10	0.3	55
6718	2534	988	2.6	1.6	453	21	20	34	138	8	0.5	59
6719	2128	902	2.3	1.8	487	23	22	50	161	13	0.3	94
6720	2082	1124	2.8	2.0	480	25	23	54	204	11	0.3	115
6721	2239	779	2.1	1.7	472	22	20	36	145	11	0.8	68
6722	2315	949	2.2	1.8	489	23	21	39	155	18	0.54	92
6723	2194	1029	2.4	1.9	525	25	23	41	165	12	<0.2	91
6724	2230	868	2.3	1.8	507	25	23	41	169	<5	0.4	92
6725	2293	998	2.4	2.2	508	26	22	96	467	12	0.5	111
6726	2368	630	2.7	1.7	464	20	19	54	198	10	<0.2	195
6727	2305	852	2.3	2.0	526	26	26	45	182	15	0.9	105

Values in ppm dry wt.

Table 14. Probability of Collingwood Harbour sites being a member of one of five community groups, using MDA with nine environmental variables

SITE	Gp 1	Gp 2	Gp 3	Gp 4	Gp 5	Predicted Gp
6703	0.791	0.006	0.202	0.000	0.000	1
6704	0.826	0.006	0.168	0.000	0.000	1
6705	0.906	0.000	0.087	0.000	0.006	1
6706	0.885	0.000	0.113	0.000	0.001	1
6707	0.748	0.000	0.242	0.000	0.010	1
6708	0.924	0.000	0.071	0.000	0.004	1
6709	0.893	0.009	0.098	0.000	0.000	1
6710	0.839	0.000	0.158	0.000	0.002	1
6711	0.002	0.904	0.001	0.093	0.000	2
6712	0.993	0.003	0.005	0.000	0.000	1
6713	0.987	0.008	0.004	0.000	0.000	1
6714	0.987	0.010	0.002	0.000	0.000	1
6715	0.974	0.016	0.009	0.001	0.000	1
6716	0.769	0.095	0.095	0.042	0.000	1
6717	0.972	0.017	0.010	0.001	0.000	1
6718	0.933	0.032	0.032	0.003	0.000	1
6719	0.989	0.009	0.001	0.000	0.000	1
6720	0.991	0.007	0.002	0.000	0.000	1
6721	0.986	0.009	0.004	0.000	0.000	1
6722	0.993	0.006	0.001	0.000	0.000	1
6723	0.994	0.005	0.001	0.000	0.000	1
6724	0.992	0.006	0.002	0.000	0.000	1
6725	0.990	0.007	0.003	0.000	0.000	1
6726	0.965	0.027	0.001	0.007	0.000	1
6727	0.992	0.007	0.001	0.000	0.000	1

Using the nine predictor variables selected previously (Table 8) and the reference site data matrix to develop discriminant equations, we predicted community assemblages for each of the 25 sites in Collingwood Harbour (Table 14) for which chemical data were available. With the exception of site 6711, all the sites were predicted as having a Gp 1 community assemblage (see Table 4). This type of community is represented by the Tubificidae and the chironomid, *Procladius* spp., as the most common organisms, but several other species are also frequently found (Table 4).

To determine whether the observed community was, in fact, similar to the predicted community, we repeated the ordination and examined the location of the Collingwood Harbour sites relative to the Gp 1 reference sites in ordination space (Fig. 7). The results show that of the 24 sites predicted as having a Gp 1 community most fall within the range of variation found in the reference sites that comprise Gp 1. Exceptions were sites 6708 and 6717 on the second vector (Fig. 7a), and sites 6706, 6707, 6708 and 6709 on the third vector. One site (6711) was predicted as having a community represented by Gp 2. This site was well outside the range observed in the reference sites (Fig. 7c) on vector 2.

Using the environmental data, we also predicted site groupings (Table 15) for the expected responses in sediment bioassays and sites were classified as being members of either Gp 1 or Gp 2. Survival and growth of

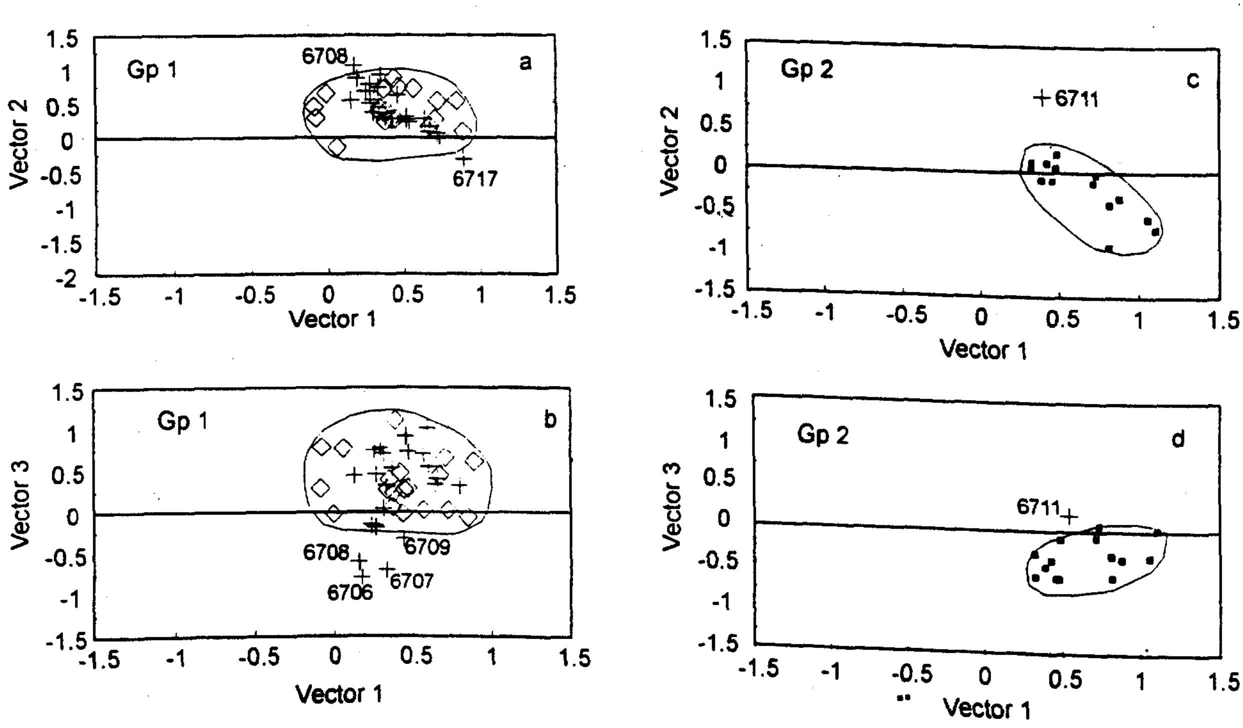


Fig. 7. Comparison of Collingwood Harbour sites with reference sites in community structure ordination space. (+) Collingwood sites; (◇) reference sites Gp 1; (■) reference sites Gp 2.

Table 15. Collingwood Harbour sites, predicted group and mean values for test end-points

Site	To Gp	CRSU	CRGW	HXSU	HXGW	HASU	HAGW	TTCC	TTYG
6703	1	88.0	0.38	98.0	6.87	89.3	0.72	*22.4*	*22.0*
6704	1	81.3	0.38	98.0	8.11	94.7	0.75	*24.4*	*22.8*
6705	1	80.0	0.43	100.0	5.78	90.0	0.66	*21.8*	*14.6*
6706	1	72.0	0.39	98.0	5.62	93.3	0.50	27.2	*22.4*
6707	1	86.6	0.33	100.0	6.17	90.7	0.42	*25.8*	*28.8*
6708	1	82.6	0.36	94.0	5.35	94.7	0.53	*24.4*	*22.3*
6709	1	68.0	0.46	100.0	3.86	94.7	0.53	36.4	42.4
6710	1	78.6	0.40	100.0	5.04	88.0	0.60	31.8	42.2
6711	1	85.3	0.40	100.0	4.56	84.0	0.50	29.0	55.8
6712	2	89.3	0.52	80.0	10.12	94.7	0.70	44.7	171.7
6713	2	76.0	0.51	100.0	10.93	88.0	0.66	43.5	125.5
6714	2	86.6	0.61	100.0	10.59	84.0	0.74	45.6	135.4
6715	2	85.3	0.49	100.0	10.94	82.7	0.76	43.0	126.0
6716	2	90.6	0.65	100.0	10.58	89.3	0.79	NA	NA
6717	2	89.3	0.64	94.0	10.40	86.7	0.87	43.4	153.6
6718	2	86.6	0.66	98.0	11.16	93.3	0.82	46.6	146.8
6719	2	88.0	0.45	100.0	8.56	92.0	0.62	40.0	*74.2*
6720	2	90.6	0.56	100.0	6.70	89.3	0.74	43.8	*36.6*
6721	2	84.0	0.58	98.0	9.97	89.3	0.74	45.0	117.2
6722	2	88.0	0.55	98.0	9.11	77.3	0.71	42.8	*74.2*
6723	2	96.0	0.35	100.0	7.98	90.7	0.71	39.2	*42.8*
6724	2	76.0	0.51	98.0	8.53	90.7	0.66	42.4	*62.4*
6725	2	90.6	0.37	84.0	8.34	94.7	0.61	38.4	*77.8*
6726	2	92.0	0.34	96.0	7.98	72.0	0.75	39.4	*48.8*
6727	2	92.0	0.38	98.0	6.28	85.3	0.56	37.8	*45.8*

Values in italic are below the range expected for the Gp from Table 10.

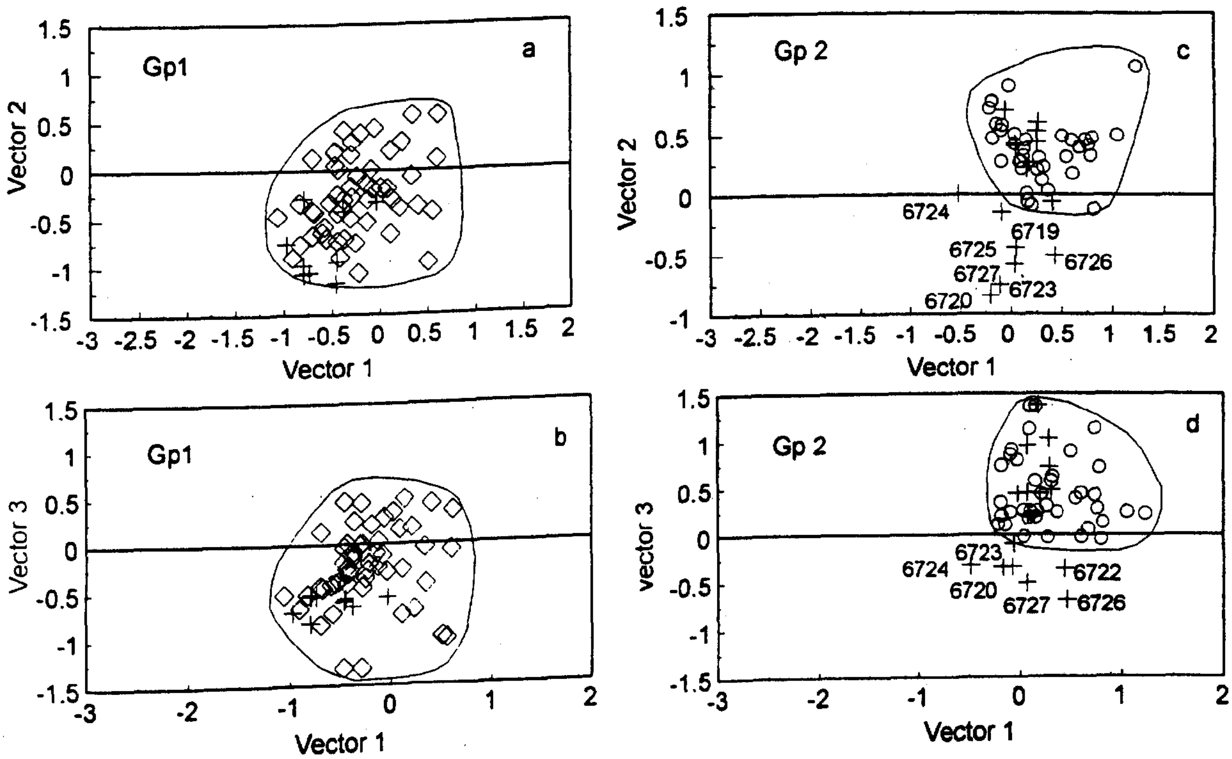


Fig. 8. Comparison of Collingwood Harbour sites with reference sites in toxicity ordination space. (◇) Gp 1; (○) Gp 2; (+) Collingwood Harbour.

C. riparius and *H. azteca* were not significantly reduced below levels in all the reference sites and thus, these end-points did not show evidence of toxicity at any of the sites. In the case of the amphipod, *H. azteca*, two sites showed slightly more growth than expected (6704 and 6717). Similarly nine sites had greater levels of growth of *Hexagenia* spp. than expected for the range of sediments in the bay. Two sites had slightly reduced survival (6712 and 6725). The effects of sediment contamination were most notable in the data for reproduction of *T. tubifex*. All the sites in the east slip and 1–2 sites in the outer harbour showed both reduced production of cocoons and total young. Several sites in the southeast corner of the outer harbour also showed a reduction in the number of young tubificids produced.

When plotted in ordination space (Fig. 8), the sites predicted as being Gp 1 all fell within the boundary of the reference sites. However, a number of sites predicted as Gp 2 (Fig. 8c,d) fell outside the reference site boundary, viz. 6719, 6720, 6723, 6724, 6725, 6726, 6727 on vector 2, and 6720, 6722, 6723, 6724, 6726 and 6727 on vector 3. These are the same sites that had reduced production of young tubificids (Table 15). While the reduction in production of young could suggest a low level of chronic toxicity at these outer harbour sites, the fact that reproduction does occur and that tubificid oligochaetes are abundant at these same sites in the harbour suggests another mechanism may be responsible. The chemical data do not indicate higher levels of contaminants at these specific Gp 2 sites (Table 13). However, these sites did have a slightly higher clay content (mean 25%) compared with the other harbour sites (19% clay). Tubificid worms are known to be sensitive to higher clay contents (Reynoldson *et al.* 1991).

Sediment remediation may not be warranted at all harbour sites in Collingwood Bay based on the data from this analysis. The spatial pattern (Fig. 9) from the analysis of community structure shows good correspondence with the chemical data. For example, the *in situ* benthic community structure was only outside the range of what should be present if the sediment were clean at sites where the levels of metals exceed the severe effects levels based on sediment quality criteria. Only one site in the outer harbour was identified as being outside the range of our reference data base (i.e. 6717) but the levels of contaminants at this site were all below the severe effects limits and this outlier can be attributed to the high numbers of Porifera at this site rather than toxicity.

The results from the laboratory tests do not indicate a sediment toxicity problem in the harbour with the exception of the results for reproduction of the oligochaete worm, *T. tubifex*. Total production of young is reduced in the east boat slip, at three sites in the outer harbour and in the southeast corner of the harbour. The reduction

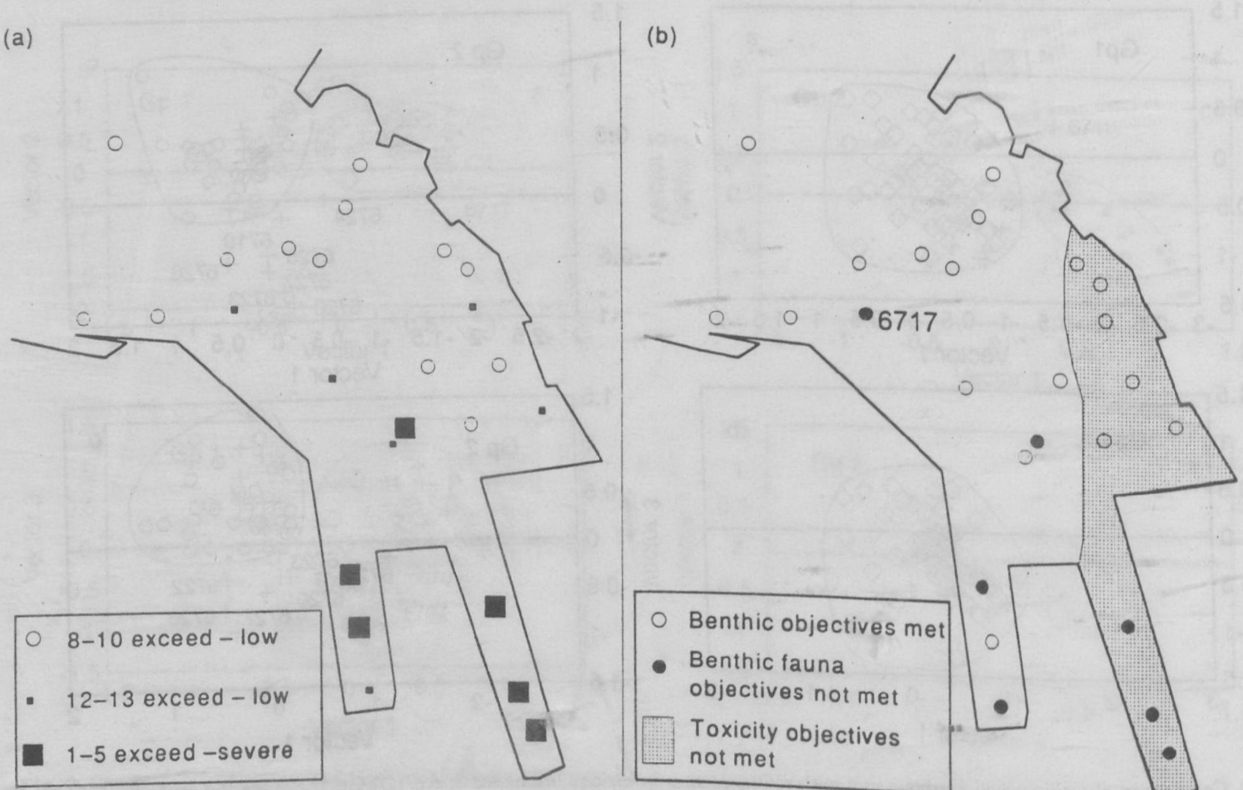


Fig. 9. Comparison of (a) chemical guidelines and (b) biological sediment guidelines in Collingwood Harbour.

in reproduction in the east slip may be attributed to the high levels of metals at these sites. Oligochaetes are known to be particularly sensitive to metal contamination (Hynes 1960; Brinkhurst & Jamieson 1971; Aston 1973; Chapman *et al.* 1980; Wachs 1980). In addition, the number of tubificids in the field-collected benthic invertebrate samples at these same sites were very low.

Based on the above results, removal of contaminated sediment is warranted in the boat slips. The remainder of Collingwood Harbour cannot be considered as having a degraded benthic community despite the fact that several of the provincial chemical sediment criteria are exceeded for metals. There is very little weight of evidence in the laboratory toxicity test data for an extensive remedial action programme in the outer harbour. The reduction in total production of young in tubificids occurs at only a few sites and this effect may not be in response to toxicity but rather to physical characteristics of the sediment. This case study emphasizes the value of a combined laboratory and field approach using both chemical and biological data in interpreting the effects.

DISCUSSION

This study was undertaken to provide an alternative to environmental guidelines and criteria for sediments in the Laurentian Great Lakes based solely on comparisons of bulk chemical concentrations of contaminants in sediments to levels of these same compounds known to cause a toxic response in biota (International Joint Commission 1988; Persaud *et al.* 1992). In our view, the results provide a more relevant and realistic method for determining environmental impact. We believe that ecosystem integrity is primarily a biological concern. The past trend of developing decision-making criteria based on chemical concentrations has primarily been in response to the inability of biologists to provide environmental managers with the information they require to make decisions. However, the development of standardized methodologies and multivariate statistical analysis, which allows prediction of biological responses based on simple environmental variables, have been revolutionary in the application of biological data to the environmental decision-making process. We view this paper as a demonstration of one approach to the use of biological data in decision making.

Comparison with other multivariate studies

A number of other studies have demonstrated the ability to predict the structure of benthic invertebrate communities from a set of environmental variables (Furse *et al.* 1984; Wright *et al.* 1984; Armitage *et al.* 1987; Corkum & Currie 1987; Moss *et al.* 1987; Ormerod & Edwards 1987; Johnson & Wiederholm 1989). Most of the studies predicting community assemblages have been conducted in lotic systems and their prediction accuracy is in the range of 68.9 to 79.6% (Reynoldson & Metcalfe-Smith 1992). The only other example of this method being used in lake systems is the work of Johnson and Wiederholm (1989) who showed that they could correctly predict 90% of the benthic invertebrate assemblages in a set of Swedish lakes using variables such as depth, silica, bicarbonate and phytoplankton volume. From our data, we were similarly able to correctly predict the community structure more than 86% of the time. Furthermore, as far as we know, this is the first time that an attempt has been made to examine and predict both structural (communities) and functional (survival, growth and reproduction) biological attributes.

Variation in functional responses

While exposure of benthic invertebrates to whole sediments is frequently used in the assessment of contamination, there has been little examination of the natural range of responses in clean sediments with a variety of geochemical characteristics. The data from this study show that there is considerable variation in the measured end-points to sediment attributes, and this is dependent on the species of invertebrate used in the bioassay and the response being measured. For example, *C. riparius* is very robust in its response to sediment type and neither growth nor survival are notably affected by sediment quality or nutrition. Similarly, *H. azteca* shows low variation in growth but survival is more variable. *Chiconomus riparius* and *H. azteca* were both fed over the period of the bioassay with fish food flakes (Nutrafin™) and therefore are less likely to respond to the nutritive quality of the various sediments. Ankley *et al.* (1994) have also shown that addition of exogenous food to test sediments with *H. azteca* and *C. riparius* significantly reduced variability in the test end-points. It is surprising, however, that there was little correlation in response to particle size distribution, mineralogical composition or organic carbon content.

Hexagenia spp. showed the least variation in survival but growth was highly variable. This variation is likely due to the fact that *Hexagenia* were not fed during the laboratory tests and the nutritive quality of the sediment may have a greater influence on the growth of the organisms than expected. We suspect that a lack of exogenous food also explains the greater variability in the reproductive end-points measured in the oligochaete test with *T. tubifex*, and total numbers of young oligochaetes has been shown to be sensitive to the amount of available food in the sediment, as measured by organic content (Reynoldson *et al.* 1991).

Comparison with other chemical and biological approaches

The use of this approach in Collingwood Harbour illustrates its advantages over the more traditional assessment methods that rely on chemical guidelines. Comparison of the concentrations of metals measured in sediments collected from the harbour with provincial sediment quality criteria (Persaud *et al.* 1992) demonstrates the inability of the criteria to determine non-impact. In addition, such criteria were only capable of providing a gradient in terms of high and low concentration(s) for an array of contaminants.

The biological approach on the other hand was able to specify where remediation was required and where contamination was not eliciting a biological response. In fact, there was good concordance between the chemical and biological data for the most severely contaminated sites (i.e. both the *in situ* data on community structure and the data from the laboratory toxicity tests indicated toxicity). The most contaminated sites (6706, 6707, 6708, 6709; Table 13) had the most depauperate communities with few organisms present, particularly the Oligochaeta. These same sites also had lower than expected *T. tubifex* reproduction in the chronic assays (Table 15).

Other sites in the outer harbour with lower contaminant concentrations required the biological data to define whether an impact was occurring. Two sites exceeded the severe effects level based on chemical criteria yet biological effects were not demonstrated at these sites. Site 6710, which had a high copper concentration (210 ppm) and site 6705 with high lead concentration (329 ppm) had reduced numbers of organisms but were within the range of variation found in reference sites. No evidence of toxicity was demonstrated in laboratory tests at these sites. None of the sites which exceeded any of the low effect criteria demonstrated any divergence from an expected community as defined by the reference sites.

The toxicity data demonstrate the importance of using a range of species rather than relying on a single species of invertebrate. In the Collingwood study, the only bioassay that showed any negative response was reproduction in the oligochaete, *T. tubifex*. The other three species showed no divergence from results found in clean sediment. This contradicts the commonly held belief that oligochaete worms are insensitive to contamination. The only other unexpected divergence was the enhanced growth of *Hexagenia* at several sites in the outer harbour. This may be attributed to a slight effect of eutrophication (G. Krantzberg, pers. comm. 1994).

We view these data as providing excellent validation of a multivariate approach for establishing reference or nominal conditions for biological variables which can be used in a practical manner to assist in management decisions. In the case of Collingwood Harbour we would fully support the remediation of the east and west boat slips. However, the outer harbour appears to be unimpacted and may require no remedial activity as both the benthic community structure and the results of chronic bioassays in four invertebrate species do not distinguish these sites from reference conditions. This is clearly something that chemical criteria were unable to do.

These data from 96 sites are the first portion of a data set that will finally consist of more than 250 sites and thus have even greater power for impact resolution. While this approach has numerous merits, particularly its ability to incorporate normal variability in predicting biological state and to define ecological targets, there are some disadvantages. The approach is initially labour intensive, requiring the acquisition of a large data base. This data base is necessary to ensure that the range of potential natural variation is captured by the reference data set and that the appropriate statistical degrees of freedom are available for performing the multivariate algorithms. Further, the model can only be applied to test sites within the range encompassed by the reference sites. Lastly, it is not a method that is intuitively easy to comprehend and thus methods will be required to transfer the technology to potential users.

Finally, there are a number of issues that have yet to be addressed. The grouping methodology is somewhat arbitrary. The degree to which the selected number of groups represent true community assemblages is unknown. The need to identify groups is a requirement of the analytical method selected; discriminant analysis requires a grouping variable in order to predict group membership from a set of variables. In this case we used two criteria to define group membership. First, the structure of the data, selecting as many groups as possible (five) before the structure began to be lost and many small groups of few sites formed. Second, the distribution of the selected groups; the strong geographic correlation between the groups and the lakes suggests that these are meaningful groupings. However, further investigation of an appropriate a priori method for group definition is desirable. A possible approach could be to use replicate samples from sites and set the number of groups by the point at which replicate samples are retained within a group. The majority of the reference sites have only been visited on one occasion. The robustness of the reference state must be established and the temporal trajectory of a site in ordination space must be determined. This will require resolution of both seasonal and annual trajectories. At present the majority of sites have been sampled in the fall and this may require sampling of test sites to be restricted to that season. The appropriate taxonomic level for incorporation into the model requires further determination. For practical considerations, it would be advantageous if a reduced taxonomic effort provided sufficient predictive capability for the model to be effective. This would ultimately make this method more attractive. Identification to the taxonomic level of species or genus is a major effort and if family or higher taxonomic levels of identifications are shown to be adequate in their diagnostic ability and predictability, then considerable time, effort and expense will be saved. Similarly, it is unreasonable to expect all investigators to use the same equipment as used in this study. Therefore, the effects of sampler type on the prediction of sites will also need to be established. These issues are currently under investigation.

CONCLUSIONS

We are confident that the preliminary results of this study demonstrate the ability to develop a reference data base that can provide meaningful information on community assemblages of benthic invertebrates at a 'clean' site and can also predict the responses of selected invertebrates to natural sediments. These biological attributes can be predicted from a relatively small set of environmental characteristics. The ultimate objective is to develop a set of numerical guidelines based on biological attributes that can be used in making appropriate management decisions.

ACKNOWLEDGEMENTS

The authors wish to thank Mr Griffin Sherbin, Great Lakes CleanUp Fund, and Dr Steven Lozano of the US EPA, Duluth, MN, provided partial financial support to this study. In particular, we would like to thank Susan Humphrey of Environment Canada who has been a stalwart financial and moral supporter. We also acknowledge Dan Faith and Dr Norris Sr for their comments on this paper. Finally, the assistance of Craig Logan, Danielle Milani, Cheryl Clarke, Leeanne Gris and Scott Kirby is gratefully acknowledged.

REFERENCES

Ankley G. T., Benoit D. A., Balogh J. C., Reynoldson T. B., Day K. E. & Hoke R. A. (1994) Evaluation of potential confounding factors in sediment toxicity tests with three freshwater benthic invertebrates. *Envir. Toxicol. Chem.* **13**, 627–35.

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

Aston R. J. (1973) Tubificids and water quality: A review. *Envir. Pollut.* **5**, 1–10.

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

Belbin L. (1991) Semi-strong hybrid scaling, a new ordination algorithm. *J. Veg. Sci.* **2**, 491–6.

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

Borgmann U., Ralph K. M. & Norwood W. P. (1989) Toxicity test procedures for *Hyalella azteca* and chronic toxicity of cadmium and pentachlorophenol to *H. azteca*, *Gammarus fasciatus* and *Daphnia magna*. *Arch. Envir. Contam. Toxicol.* **18**, 756–64.

Brinkhurst R. O. & Jamieson B. G. M. (1971) Aquatic Oligochaeta of the World. University of Toronto Press, Toronto, 860pp.

Cairns J. J. & van der Schalie W. H. (1980) Biological monitoring Part 1: Early warning systems. *Water Res.* **14**, 1179–96.

Chapman P. M. (1986) Sediment quality criteria from the sediment quality triad: An example. *Envir. Toxicol. Chem.* **5**, 957–64.

Chapman P. M. (1990) The sediment quality triad approach to determining pollution-induced degradation. *Sci. Total Envir.* **97/98**, 815–25.

Chapman P. M., Churchland L. M., Thomson P. A. & Michnowsky E. (1980) Heavy metal studies with oligochaetes. In *Aquatic Oligochaete Biology* (eds R. O. Brinkhurst & D. G. Cook). Plenum Press, New York.

Corkum L. D. & Currie D. C. (1987) Distributional patterns of immature Simuliidae (Diptera) in northwestern North America. *Freshwat. Biol.* **17**, 201–21.

Day K. E., Kirby R. S. & Reynoldson T. B. (1994) Sexual dimorphism in *Chironomus riparius* (Meigen): Impact on interpretation of growth in whole sediment toxicity tests. *Envir. Toxicol. Chem.* **13**, 35–9.

Duncan G. A. & LaHaie G. G. (1979) Size analysis procedures used in the sedimentology laboratory. NWRI, CCIW, Hydraulics Division Manual. September, 1979.

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

Furse M. T., Moss D., Wright J. F. & Armitage P. D. (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**, 257–80.

Hanes E. C. & Ciborowski J. H. (1992) Effects of density and food limitation on size variation and mortality of larval *Hexagenia rigida* (Ephemeroptera: Ephemeridae). *Can. J. Zool.* **70**, 1824–32.

Hunsaker C. T. & Carpenter D. E. (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, North Carolina.

Hynes H. B. N. (1960) *The Biology of Polluted Waters* Liverpool University Press, UK 202 pp.

International Joint Commission (1987) Report on Great Lakes water quality. Report to the International Joint Commission, Great Lakes Water Quality Board. Windsor, Ontario. 236 pp.

International Joint Commission (1988) Procedures for the assessment of contaminated sediment problems in the Great Lakes. Report to the Great Lakes Water Quality Board. Windsor, Ontario. 140 pp.

Johnson R. K. & Wiederholm T. (1989) Classification and ordination of profundal macroinvertebrate communities in nutrient poor, oligo-mesohumic lakes in relation to environmental data. *Freshwat. Biol.* **21**, 375–86.

Long E. R. & Chapman P. M. (1985) A sediment quality triad: Measures of sediment contamination, toxicity and infaunal community composition in Puget Sound. *Mar. Poll. Bull.* **16**, 405–15.

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

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

Ormerod, S. J. & Edwards R. W. (1987) The ordination and classification of macroinvertebrate assemblages in the catchment of the River Wye in relation to environmental factors. *Freshwat. Biol.* **17**, 533–46.

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

Persaud D., Jaagumagi R. & Hayton A. (1992) Guidelines for the protection and management of aquatic sediment quality in Ontario. Water Resources Branch, Ontario Ministry of Environment. Toronto, Ontario, 23pp.

Reynoldson T. B., Thompson S. P. & Bamsey J. L. (1991) A sediment bioassay using the tubificid oligochaete worm *Tubifex tubifex*. *Envir. Toxicol. Chem.* **10**, 1061–72.

Reynoldson T. B. & Metcalfe-Smith J. (1992) An overview of the assessment of aquatic ecosystem health using benthic invertebrates. *J. Aquat. Ecosystem Health* **1**: 295–308.

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

Reynoldson T. B., Day K. E., Clarke C. & Milani D. (1994) Effect of indigenous animals on chronic endpoints in freshwater sediment toxicity. *Envir. Toxicol. Chem.* **13**, 973–7.

Wachs B. (1980) Bioakkumulation und biogene retention von metallen in Fliessgewasseren. Munchener Beitrage zur Abwasser-, Fischerei- und Flussbiologie 32, 387–414.

Wickware G. M. & Rubic C. D. A. (1989) Ecoregions of Ontario. Environment Canada, Ecological Land Classification Series No. 26. 37pp.

Wright J. F., Moss D., Armitage P. D. & Furse M. T. (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. *Freshwat. Biol.* **14**, 221–56.

Zarull M. A. & Reynoldson T. B. (1992) A management strategy for contaminated sediments: Assessment and remediation. *Water Poll. Res. J. Canada* **27**, 871–82.

Appendix I. Great Lakes Benthic Community Structure Study: Species list

Gastropoda
- Bithyniidae
  - *Bithynia tentaculata*
- Hydrobiidae
  - ***Amnicola limosa*** **(Amn lim)**
  - *A. walkeri*
  - *Marstonia decepta*
  - *Probythinella lacustris*
  - *Hydrobiidae immatures*
- Lymnaeidae
  - *Fossaria obrussa*
- Physidae
  - *Physella integra*
  - ***P. species*** **(Phy spp)**
- Planorbidae
  - *Armiger crista*
  - *Gyraulus circumstriatus*
  - *G. deflectus*
  - *Helisoma anceps*
  - *Promenetus exacuous*
- Valvatidae
  - *Valvata lewisi*
  - ***V. piscinalis*** **(Val pis)**
  - *V. sincera*
  - ***V. tricarinata*** **(Val tri)**
- Viviparidae
  - *Campeloma decisum*
- Unknown spp (damaged)

Pelycepoda
- Sphaeridae
  - ***Pisidium casertanum*** **(Pis cas)**
  - ***P. compressum*** **(Pis com)**
  - *P. ferrugineum*
  - ***P. henslowanum*** **(Pis hen)**
  - ***P. nitidum*** **(Pis nit)**
  - *P. ventricosum*
  - ***P. unknown*** **(Pis unk)**
  - *Sphaerum nitidum*
  - *S. simile*
  - *S. striatum*
  - *S.* unknown
  - ***Musculim partinium*** **(Mus par)**
  - *M. securis*
  - *M. transversum*
- Unionidae
  - *Elliptio camplanata*
  - *Lampris radiata*
- Dreissenidae
  - ***Dreissena polymorpha*** **(Dre pol)**

Diptera
- Chironomidae
  - ***Chironomus*** **(Chi spp)**
  - ***Cladopelma*** **(Cpe spp)**
  - *Clanotanytarsus*
  - ***Cryptochironomus*** **(Cch spp)**
  - ***Cryptotendipes*** **(Cte spp)**
  - ***Dicrotendipes*** **(Dic spp)**
  - ***Demicryptochironomus*** **(Dem spp)**
  - ***Endochironomus*** **(End spp)**
  - ***Glypotendipes*** **(Gly spp)**
  - *Harnischia*
  - ***Micropsectra*** **(Mps spp)**
  - *Microtendipes*
  - *Nilothauma*
  - *Pagastiella*
  - *Parachironomus*
  - *Paracladoplema*
  - *Paralauterborniella*
  - *Paratendipes*
  - ***Polypedium*** **(Pol spp)**
  - ***Pseudochironomus*** **(Pse spp)**
  - ***Stictochironomus*** **(Sti spp)**
  - ***Tanytarsus*** **(Tan spp)**
  - *Tribelos*

Appendix I. continued

Stempellina
Zavreliella
Unknown chironominae
Potthastia
Protanypus
Corynoneura
Cricotopus
Epoicricotopus
Heterotrissocladius **(Het spp)**
Nanocladius
Parakiefferiella
Psectrocladius
Unknown orthocladinae
Ablabesmyia
Clinotanypus
Coelotanypus **(Coe spp)**
Larsia
Procladius **(Pro spp)**
Tanypus
Monodiamesia
Ceratopogonidae
Bezzia sp.
Mallochohelea sp.
Probezzia sp.
Chaoboridae
Chaoborus **sp. (Cha spp)**
Empididae
Ephmeroptera
Ephemeridae
Hexagenia limbata
Caenidae
Caenis sp.
Colembola
Trichoptera
Polycentropodidae
Polycentropus sp.
Phylocentropus sp.
Helicopsychidae
Helicopsyche
Leptoceridae
Leptocerus americanus
Mystacides sp.
Nectopsyche sp.
Oecetis sp.
Molannidae
Molanna sp.
Hydroptilidae
Agraylea sp.
Polychaeta
Sabellidae
Manayunkia speciosa **(Man spp)**
Oligochaeta
Lumbriclidae
Eclipidrilus lacustris
Lumbriculus variegatus
Stylodrilus heringianus **(Sty her)**
Enchytreidae
Naididae
Arcteonais lomondi **(Arc lom)**
Amphichaeta leydigi
Chaetogaster diaphanus
Nais barbata
N. elinguis
N. pseudobtusa
N. simplex
Dero digitata **(Der dig)**
Pristina leidyi
Pristinella acuminata
Specaria josinae **(Spe jos)**
Stylaria lacustris **(Sty lac)**
Uncinais uncinata
Ophidonais serpentina
Piguetiella michiganensis
Vejdovskyella intermedia **(Vej int)**
Tubificidae
Immatures with hair chaetae (A) (Imm chr)
Immatures without hair chaetae (B) (Imm coh)
Aulodrilus americana
A. limnobius
A. pigueti **(Aul pig)**
A. pluriseta **(Aul plu)**
Branchiura sowerbyi
Ilyodrilus templetoni
Limnodrilus claparedianus
L. cervix
L. hoffmeisteri **(Lim hof)**
L. profundicola
Potamothrix bedoti
P. moldaviensis **(Pot mol)**
P. vejdovskyi **(Pot vej)**
Quastadrilus multisetosus **(Qui mul)**
Spirosperma ferox **(Spi fer)**
Tasserkidrilus kessleri
Tubifex tubifex **(Tub tub)**
Hirudinea
Glossiphoniidae
Alboglassophonia heteroclita
Gloiobdella elongata
Helobdella stagnalis
Piscicolidae
Myzobdella lugubris
Platyhelminthes (Platy)
Isopoda
Asellidae
Caecidotea racovitzai
C. intermedius **(Cae int)**
C. **sp. (Cae spp)**
Amphipoda
Gammaridae
Gammarus pseudolimnaeus **(Gam pse)**
Haustoriidae
Diaporeia hoyi **(Dia hoy)**
Taliridae
Hyllela azteca
Coelenterata
Hydridae
Hydra americana **(Hyd ame)**
Porifera (Porif)
Tardigrada
Macrobiotidae
Dactylobiotus

Note those taxa used in classification analysis are shown in bold.

DATA MANAGEMENT ISSUES AND GIS APPROACHES FOR BIOLOGICAL ASSESSMENT

Tony Selle
U.S. Environmental Protection Agency, Region VIII
999 18th Street, Suite 500
Denver, CO 80202

NOTES:

Within the federal government, environmental samples are traditionally collected for a single purpose, i.e. to answer important questions about a specific area of concern. Unfortunately, once serving that purpose the data immediately becomes devalued because no plan or mechanism is put in place to extend the data use beyond the original purpose.

This presentation will focus on how we can make the biological data collected retain its value and increase its utility over time by applying sound data management principles. We will discuss developing and implementing a data management plan as part of the monitoring or sampling plan, integrating field and laboratory data with ancillary data, and managing data and information for the widest possible array of applications and audiences.

We will then take a more detailed look at one application, the use of Geographic Information Systems (GIS) for biological assessment. We will discuss the various ways GIS can be employed, including logistical support, data management, data development, analysis, and data/information display and distribution.

Why Do We Collect Biological Data?

- To determine what lives (or doesn't live) at the sample site
- To create biota-relevant environmental health metrics
- To measure changes to the biota over time

What Do We Do With Biologic Data?

- **Problem specific analysis for species**
- **Site/area assessment for environmental health**
- **Trends/change detection for environmental and species health**

EPA R8 Bioassessment Workshop, T. Selle 8/28/95

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY

What Don't We Do (Usually)?

- Store data in a common electronic format
- Integrate sampled data
- Make it available to everyone else
- Use the data for assessment other than original purpose

EPA R8 Bioassessment Workshop, T. Selle 8/28/95

How Do We Change?

Integrated Data Management!

EPA R8 Bioassessment Workshop, T. Selle 8/28/95

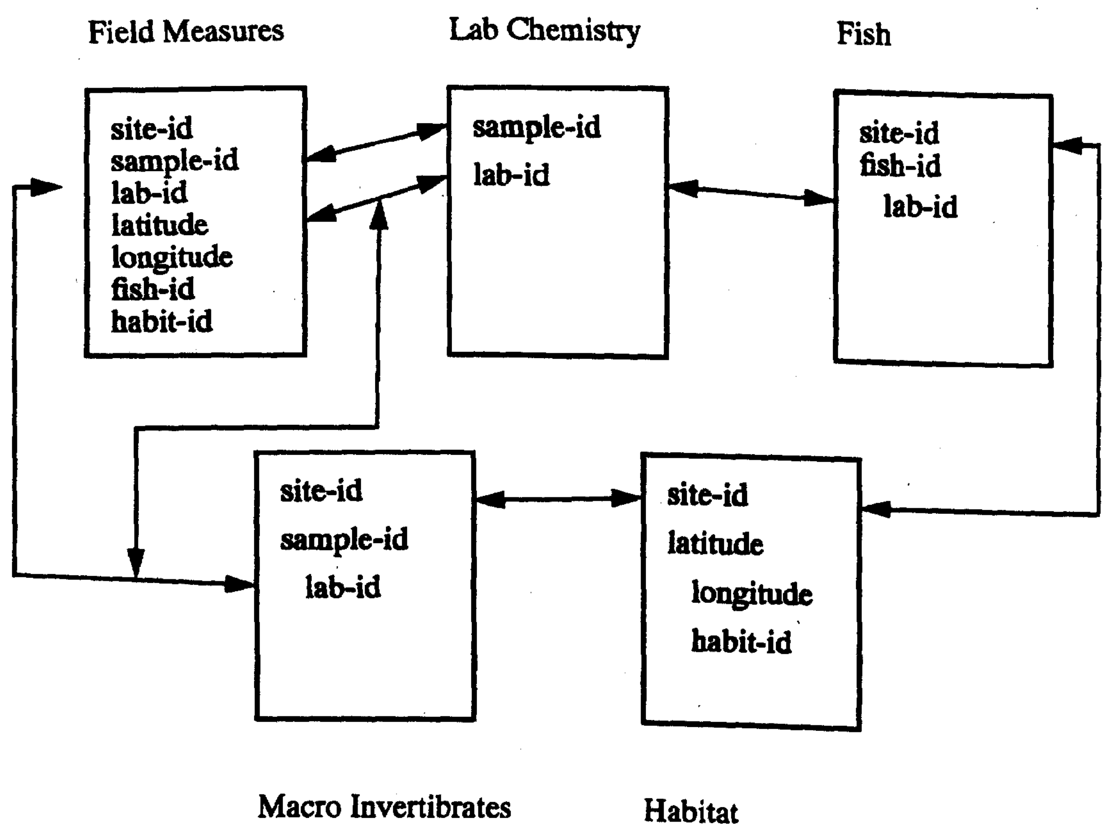
Data Management Plan Is Part of Sampling Plan

- Details sample integration
- Details data standardization (IDs, Locations)
- Details data distribution
- Describes data dictionaries and other metadata

EPA R8 Bioassessment Workshop, T. Selle 8/28/95

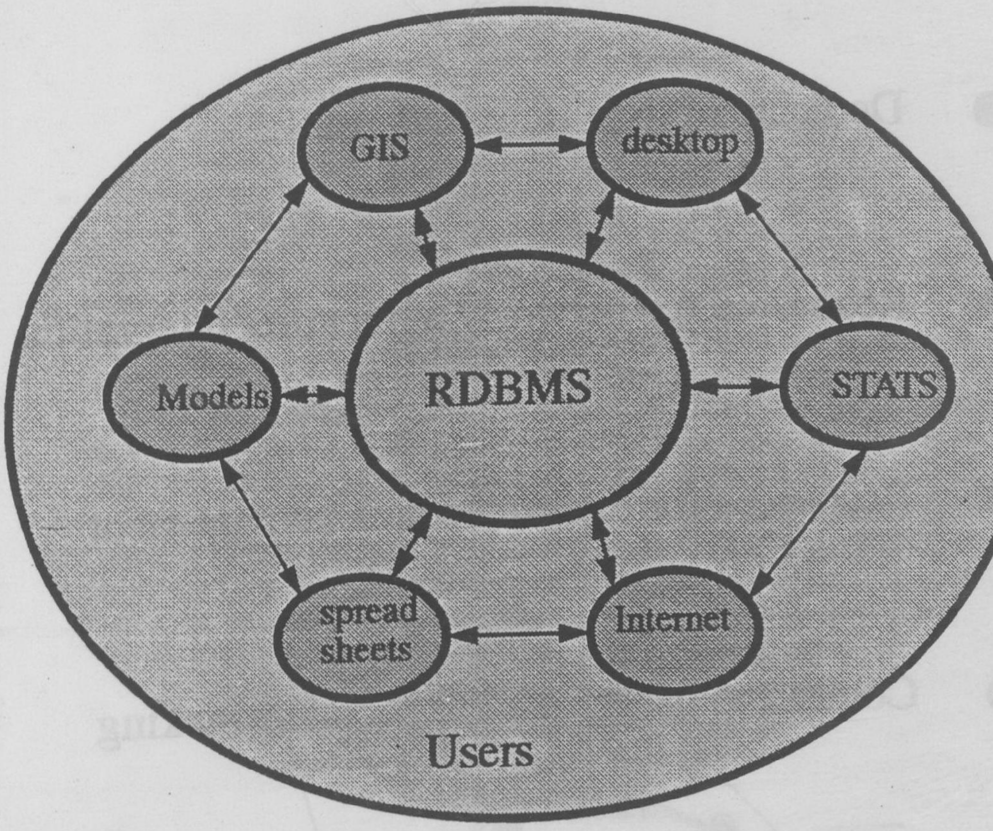
Data Management Plan

Sample Integration



EPA R8 Bioassessment Workshop, T. Selle 8/28/95

Integrating Computer Systems And Applications



EPA R8 Bioassessment Workshop, T. Selle 8/28/95

Institutional Changes Needed

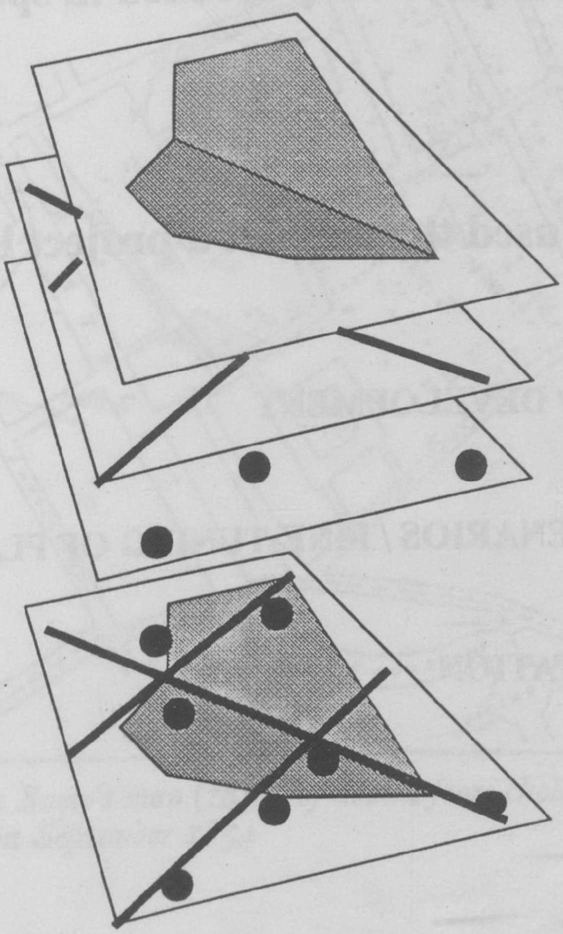
- Dedicate to long term data use
- Resources for network/application integration
- Resources for integrated data management
- Commit to using data for decision-making

EPA R8 Bioassessment Workshop, T. Selle 8/28/95

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY

What Is A GIS?

How Does It Work?



EPA R8 Bioassessment Workshop, T. Selle 8/28/95

WHAT MAKES GIS UNIQUE?

1. All data is physically indexed in space.

2. GIS can be used throughout a project lifecycle

A. PLANNING / DEVELOPMENT

B. DESKTOP SENARIOS / FINE-TUNING OF PLANS

C. IMPLIMENTATION

D. ANALYSIS

E. REPORTING

F. DECISION SUPPORT

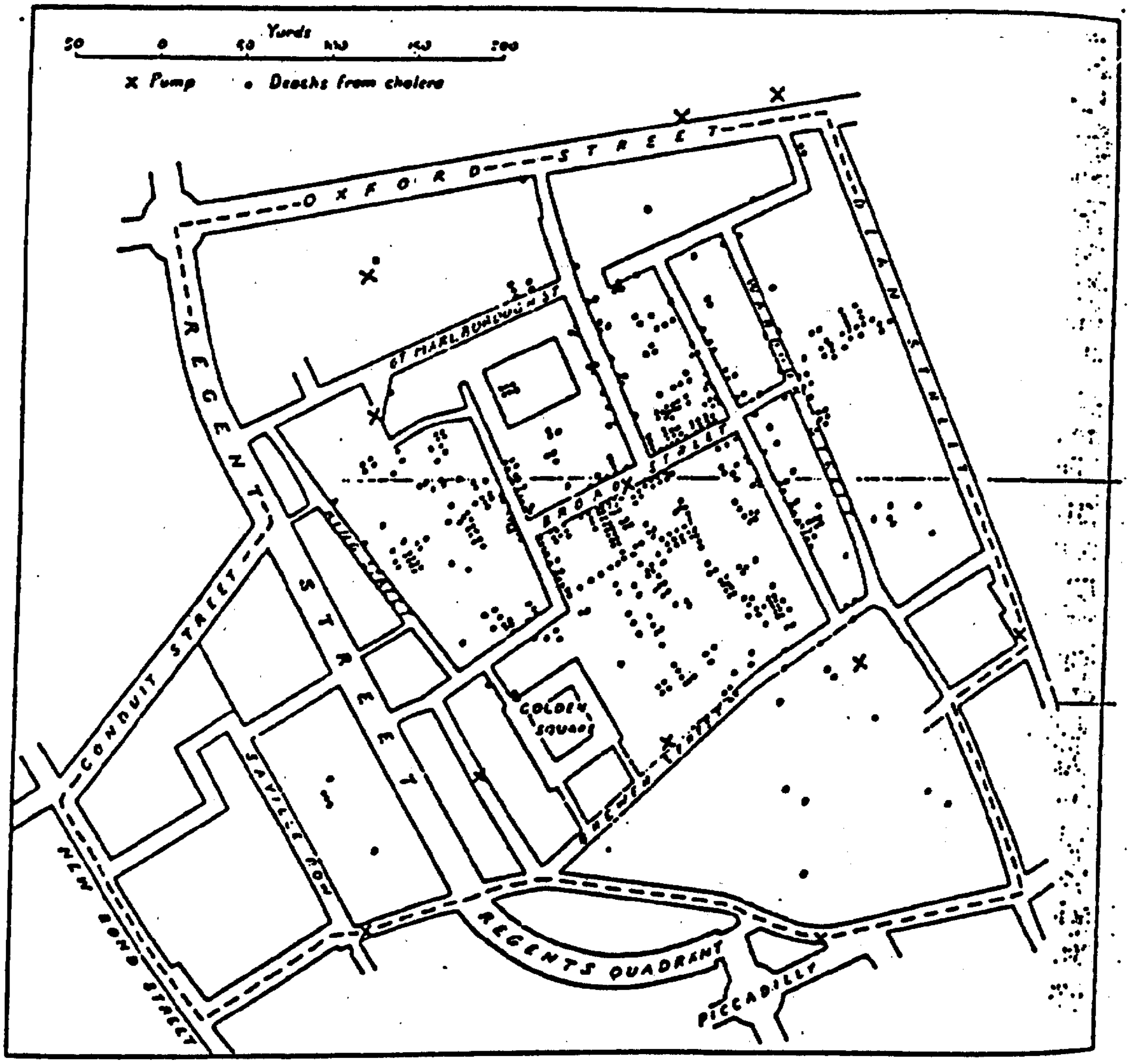
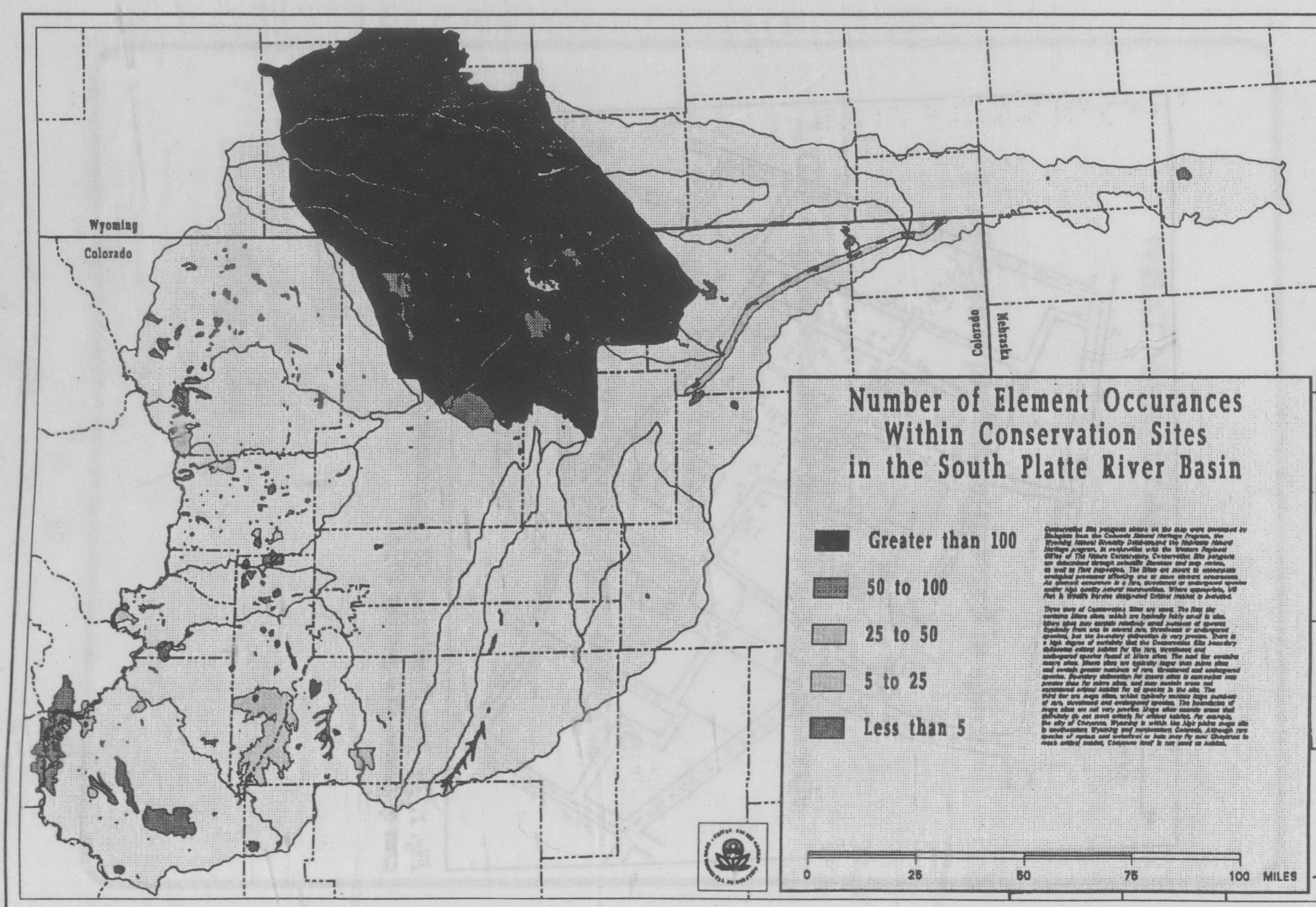
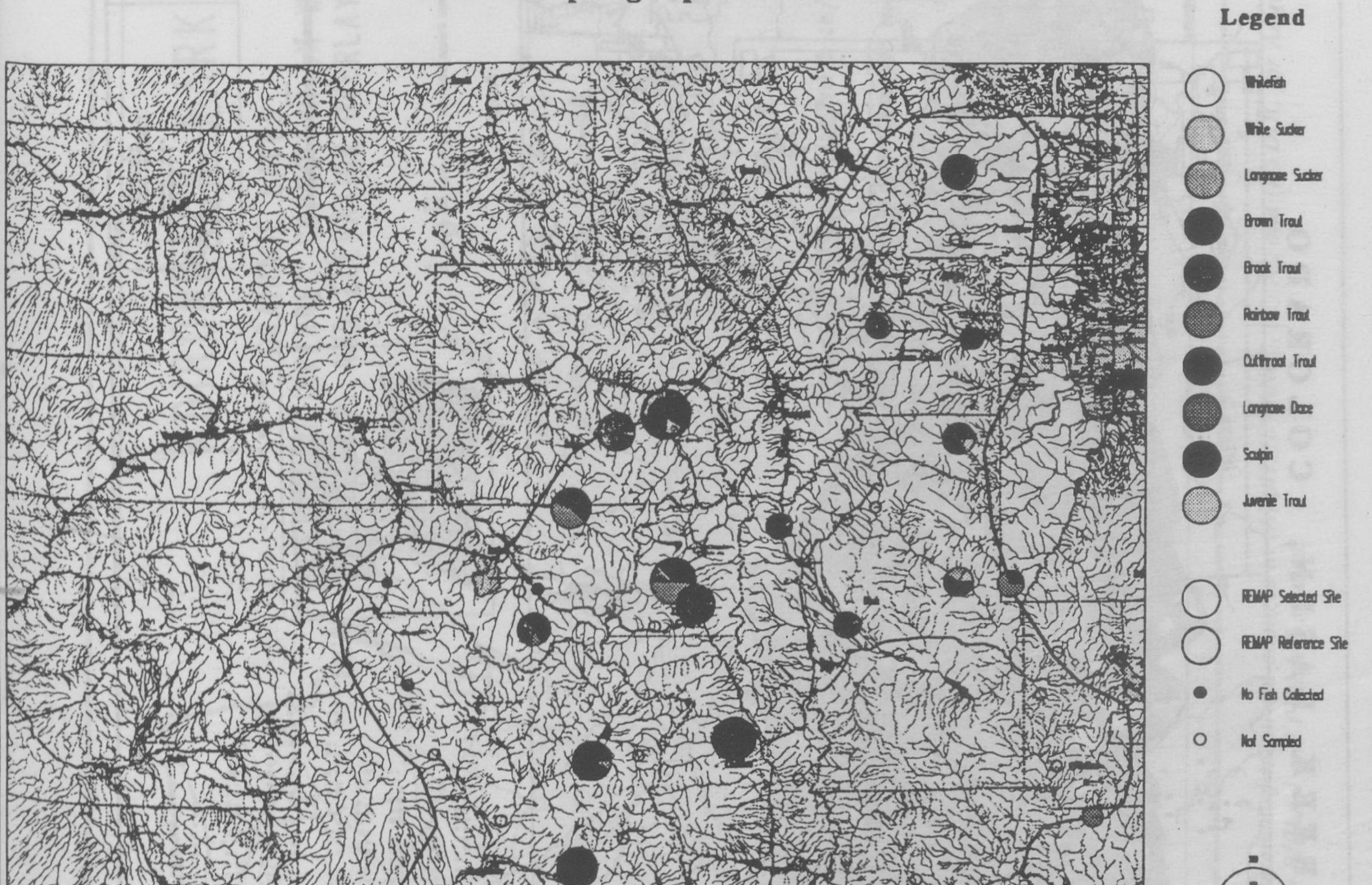


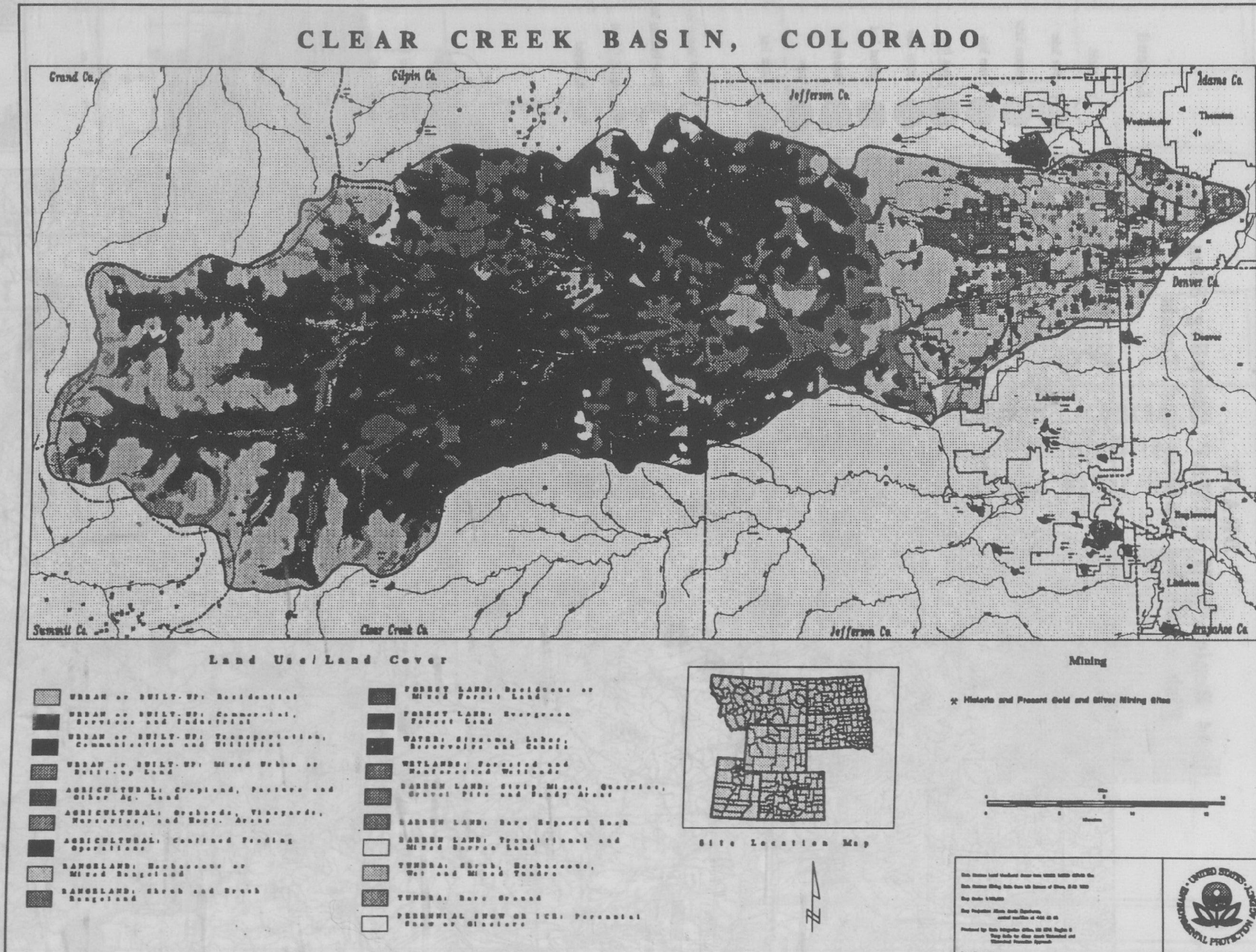
Fig. 1. Dr. John Snow's map (1855) of deaths from cholera in the Broad Street area of London in September 1854



Number of Element Occurances
Within Conservation Sites
in the South Platte River Basin
Greater than 100
50 to 100
25 to 50
5 to 25
Less than 5
Wyoming
Colorado
Nebraska
0
25
50
75
100 MILES

DRAFT
1994 Region VIII REMAP Sampling Sites
Fish Sampling Species and Totals




CLEAR CREEK BASIN, COLORADO
Grand Co.
Gilpin Co.
Jefferson Co.
Adams Co.
Westminster
Thornton
Denver Co.
Denver
Lakewood
Englewood
Littleton
Summit Co.
Clear Creek Co.
Jefferson Co.
Arapahoe Co.
Land Use/Land Cover
URBAN or BUILT-UP: Residential
URBAN or BUILT-UP: Commercial, Services and Industrial
URBAN or BUILT-UP: Transportation, Communication and Utilities
URBAN or BUILT-UP: Mixed Urban or Built-up Land
AGRICULTURAL: Cropland, Pasture and Other Ag. Land
AGRICULTURAL: Orchards, Vineyards, Nurseries, and Hort. Areas
AGRICULTURAL: Confined Feeding Operations
RANGELAND: Herbaceous or Mixed Rangeland
RANGELAND: Shrub and Brush Rangeland
FOREST LAND: Deciduous or Mixed Forest Land
FOREST LAND: Evergreen Forest Land
WATER: Streams, Lakes, Reservoirs and Canals
WETLANDS: Forested and Nonforested Wetlands
BARREN LAND: Strip Mines, Quarries, Gravel Pits and Sandy Areas
BARREN LAND: Bare Exposed Rock
BARREN LAND: Transitional and Mixed Barren Land
TUNDRA: Shrub, Herbaceous, Wet and Mixed Tundra
TUNDRA: Bare Ground
PERENNIAL SNOW OR ICE: Perennial Snow or Glaciers
Site Location Map
Mining
Historic and Present Gold and Silver Mining Sites
UNITED STATES ENVIRONMENTAL PROTECTION AGENCY

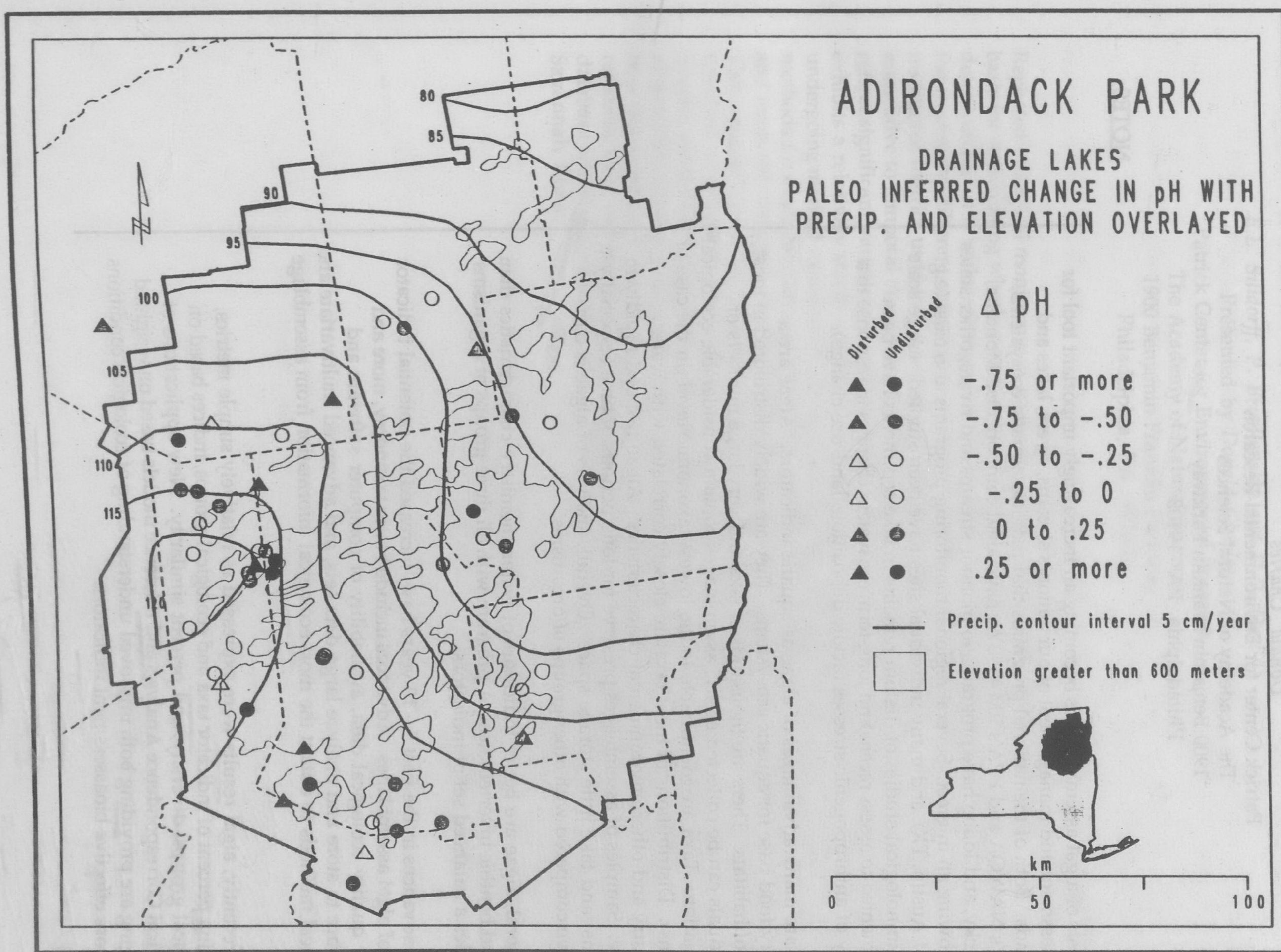


Figure 11-43a. Maps of paleolimnological inference of historical changes in (A) pH, (B) ANC_G, (C) Al_p and (D) DOC for Adirondack drainage lakes included in the PIRLA-I and PIRLA-II studies.

USE OF PERIPHYTON AS A BIOASSESSMENT TOOL

Donald Charles
Patrick Center for Environmental Research
The Academy of Natural Sciences
1900 Benjamin Franklin Parkway
Philadelphia, PA 19103

NOTES:

Analysis of algal assemblages is becoming an increasingly important tool for bioassessment and management of our nation's streams, rivers, lakes and wetlands. Both of the national programs that monitor aquatic ecosystems, USGS's NAWQA and EPA's EMAP-SW, have algal components. Montana, Kentucky, and Idaho have programs to monitor streams and rivers; other states are showing an interest. Several regional monitoring programs also exist (e.g. City of Austin, TX), and many individual sites have been studied. Large scale paleolimnological studies of diatom remains in lake sediments have been implemented to assess recent and long-term ecosystem changes in response to a variety of anthropogenic stresses (acidic deposition, land-use change).

There are many advantages of algae as aquatic indicators. Algae are a primary food base for aquatic ecosystems. They are widely distributed in most types of habitats. There are thousands of taxa and very large numbers of individuals can be collected easily; assemblages contain considerable ecological information. Taxa are identifiable to the lowest taxonomic level (in the case of diatoms). Distributions of most taxa are closely correlated with water chemistry and other environmental characteristics. Algae respond rapidly to change. Samples of assemblages preserve for long periods of time (especially diatoms) and take little storage space. Overall, analysis of algae is cost effective compared with other groups of organisms.

In general, algae are better indicators of water chemistry characteristics than fish and benthic invertebrates. A program with all three groups of organisms provides a balanced set of indicators.

Many advances in the past 5 - 10 years have increased the potential indicator value of algal assemblages. Advances include better taxonomy, more and higher quality ecological data, availability of computer software and hardware to store and analyze large data sets, and advanced multivariate and statistical methods to obtain the most ecological information from assemblage data.

Until recently, algal results were expressed as relatively simple metrics, including percent of indicator taxa and ecological groups, indices based on ecological groups, diversity, and percent similarity. New applications of Canonical Correspondence Analysis and inference models based on weighted averaging are providing both improved understanding of ecological conditions and more effective bioassessment indicators.

TREND DETECTION IN BIOLOGICAL MONITORING: EVALUATING PATTERNS OF COMMUNITY SIMILARITY

E.L. Silldorff, P. Brusilovskly and D.D. Hart
Presented by *Donald Charles*
Patrick Center for Environmental Research
The Academy of Natural Sciences
1900 Benjamin Franklin Parkway
Philadelphia, PA 19103

NOTES:

Repeated biological monitoring of particular water bodies provides a powerful basis for determining whether ecological conditions are improving or deteriorating. It is widely recognized that such biological assessments should focus on many different components of the aquatic community. Some of the traditional metrics used in bioassessment, however, can be relatively insensitive to temporal changes in community characteristics that may in turn reflect significant variations in environmental quality. Even if a community exhibits a relatively stable level of species diversity, for example, it may be undergoing marked variations in species composition. To supplement current methods for evaluating trends, we describe an approach for quantifying spatial and temporal variations in community compositions based on similar analyses. This approach uses Autosimilarity Analysis (a technique that evaluates temporal variations in similarity at a single site) in conjunction with Synchrosimilarity Analysis (a technique that evaluates temporal variations in similarity between two sites) to evaluate community trends. We demonstrate how this approach can quantify trends in community composition relative to reference sites, and we illustrate the approach using a long-term ecological database which describes variations in aquatic macroinvertebrates of the Savannah River.

Some Recent References on Use of Periphyton for Monitoring Rivers and Streams

Source: D. Charles, Academy of Natural Sciences of Philadelphia, 9/95

Whitton, B.A., E. Rott, and G. Friedrich. 1991. Use of algae for monitoring rivers. E. Rott, Innsbruck, Austria. 193 pages.

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

Bahls, L.L. 1993. Periphyton bioassessment methods for Montana streams. Water Quality Bureau, Dept. of Health and Environmental Sciences, Helena, MT.

Lynch, R.A. 199?. Development of rapid bioassessment protocols for Oklahoma utilizing characteristics of the diatom community. Oklahoma Conservation Commission, Oklahoma City, OK. 106 pages.

McCormick, P.V. and J. Cairns. 1994. Algae as indicators of environmental change. Journal of Applied Phycology, 6:509-526.

Mills, M.R., G. Beck, J. Brumley, S.M. Call, J. Grubbs, R. Houp, L. Metzmeier, and K. Smathers. 1993. Methods for assessing biological integrity of surface waters. Kentucky Department for Environmental Protection, Division of Water, Water Quality Branch, Ecological Support Section, Frankfort, KY. 139 pages (includes chapter on algae).

Porter, S.D., T.F. Cuffney, M.E. Gurtz, and M.R. Meador. 1993. Methods for collecting algal samples as part of the national water quality assessment program. Open-File Report 93-409. U.S. Geological Survey, Raleigh, NC. 39 pages.

Reid, M.A., J.C. Tibby, D. Penny, and P.A. Gell. 1995. The use of diatoms to assess past and present water quality. Australian Journal of Ecology, 20:57-64.

Rosen, B.H. 1994. Use of periphyton in the development of biocriteria. Pages 209-215. *In* W.S. Davis and T.P. Simon, editors. Biological Assessment and Criteria. Tools for Water Resource Planning and Decision Making. Lewis Publishers, Boca Raton.

Round, F.E. 1991. Diatoms in river water-monitoring studies. Journal of Applied Phycology, 3:129-145.

Stevenson, R.J. and R.L. Lowe. 1986. Sampling and interpretation of algal patterns for water quality assessments. Pages 118-149. *In* B.G. Isom, editor. Rationale for Sampling and Interpretation of Ecological Data in the Assessment of Freshwater Ecosystems, ASTM STP 894. American Society for Testing and Materials, Philadelphia.

INVERTEBRATE ASSEMBLAGES AND REGIONAL CLASSIFICATION: THE INTERIOR COLUMBIA BASIN AS A CASE STUDY IN SCALING UP

Judith Li
Department of Fisheries and Wildlife
Oregon State University
Corvallis, OR 97331

NOTES:

We chose five assemblage measures to capture essential qualities of macroinvertebrate populations of the Inner Columbia River Basin. Taxa richness, EPT taxa, EPT/Chironomids, total abundance, and percentage represented by the most dominant taxa, were used for our multivariate analyses. Despite the variations in methodology and other limitations of survey data made available to us, the database was extensive and suitable for comparisons of these metrics. Principal components analysis proved to be a useful tool in making distinctions between stream assemblages in the Blue Mountain, Eastern Cascade and Columbia Plateau ecoregions. Comparisons of taxa among streams and ecoregions were made by combining summary taxa tables and DECORANA analyses of the ten most dominant taxa for each stream. Surprisingly few taxa were ubiquitous, and a limited number of taxa were identified as characteristic of particular ecoregions. Whereas few distinctions between streams in the High Desert ecoregion of southern Oregon and Idaho were apparent using assemblage metrics, this was the region where most unique taxa were found. These combined techniques provided a preliminary template of assemblage distributions scaled to an expansive inter-regional landbase, readily modified by new surveys and more detailed information.

DRAFT, INCOMPLETE

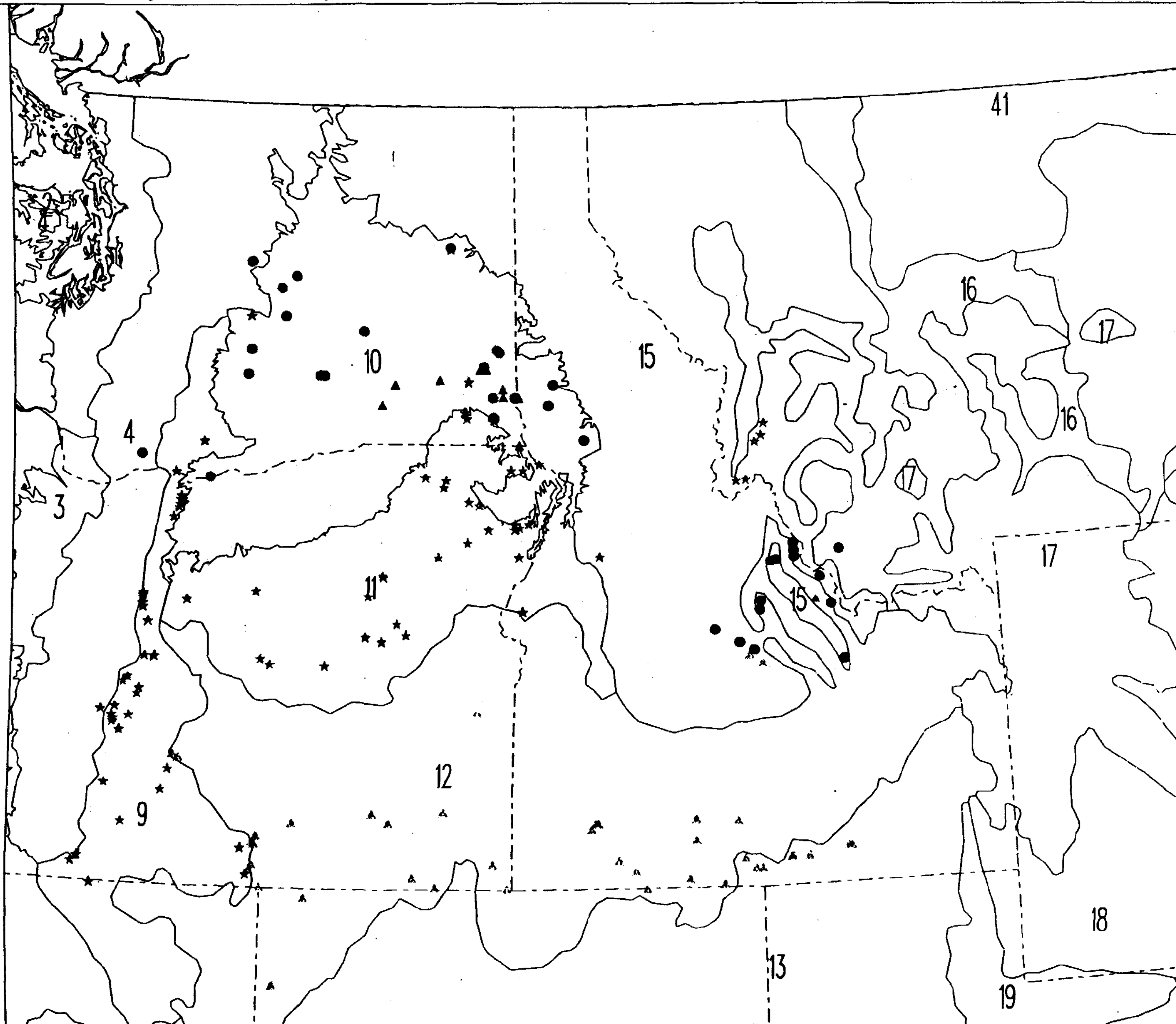
Benthic Invertebrates

Source: "A Survey of Eastside Ecosystem Benthic Invertebrates" J. Li, K. Wright, and J. Furnish 1995.

LEGEND

- State Boundaries
- Omernik Ecoregions
- Quality N/A
- Riverine
- Low Diversity, Flashy
- Low Diversity, High Desert
- Moderate Diversity
- High Diversity, Cool Habitat

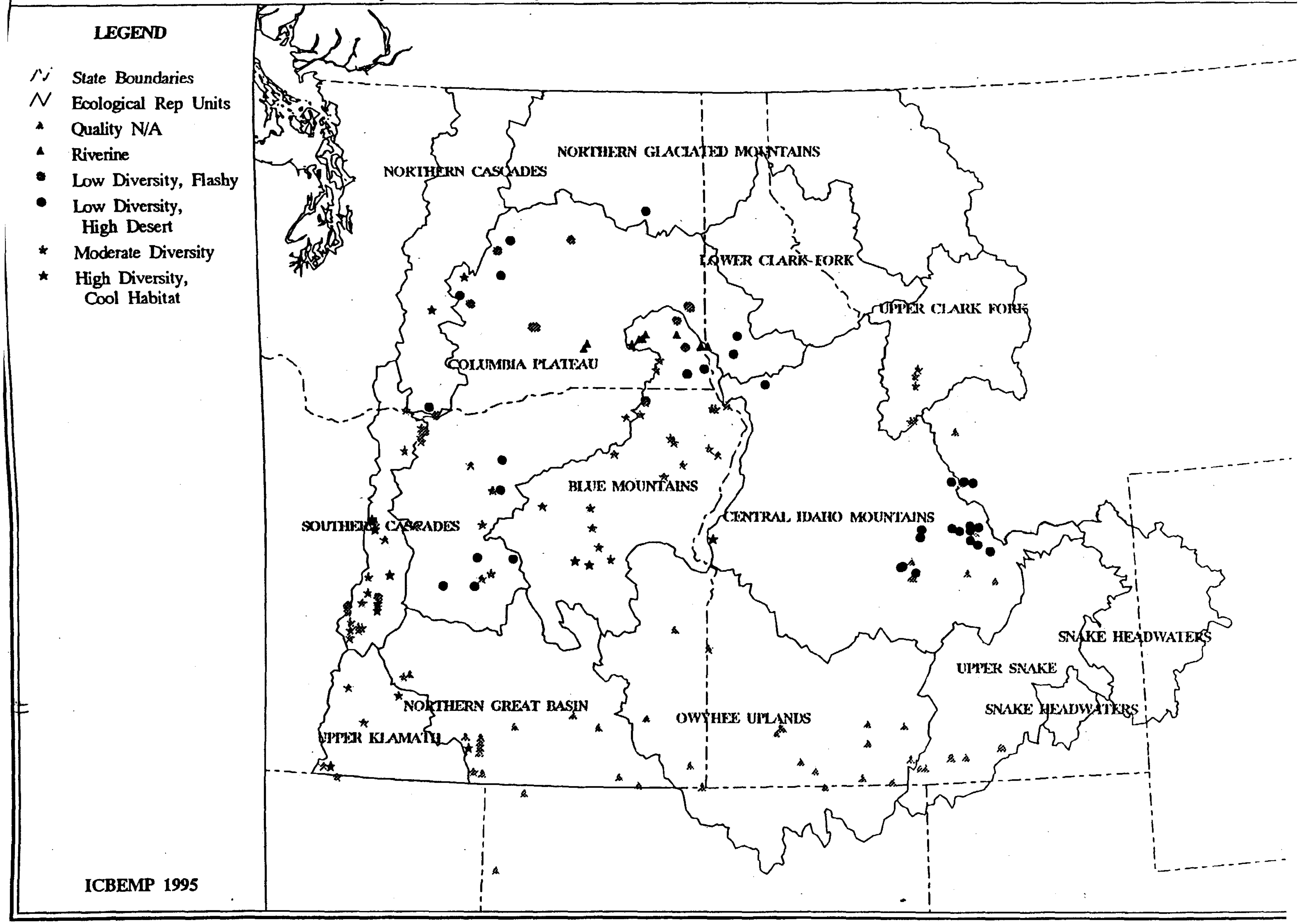
ICBEMP 1995



DRAFT

Benthic Invertebrates

Source: "A Survey of Eastside Ecosystem Benthic Invertebrates" J. Li, K. Wright, and J. Furnish 1995.

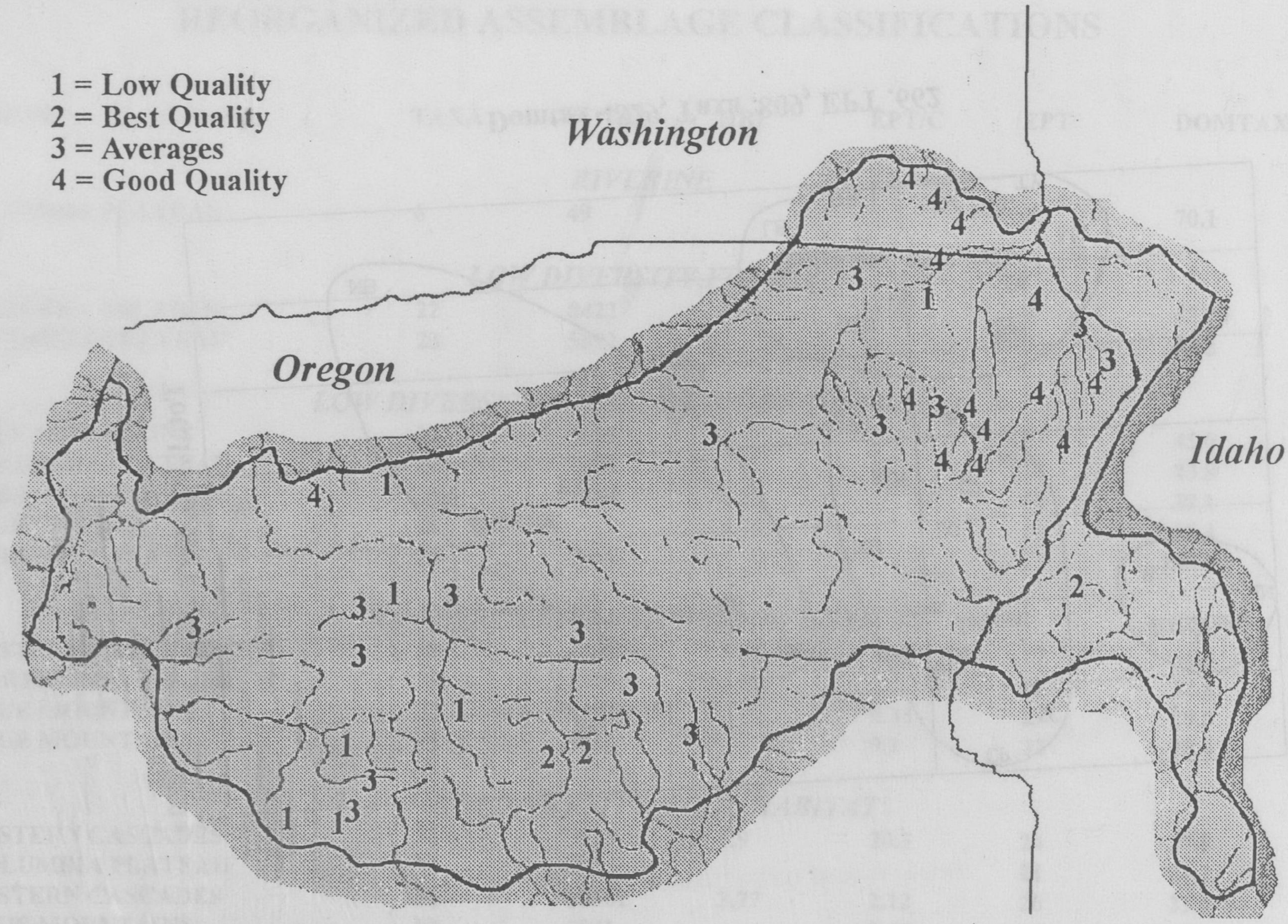


AVERAGE ASSEMBLAGE MEASURES
Based on Principal Components Analysis

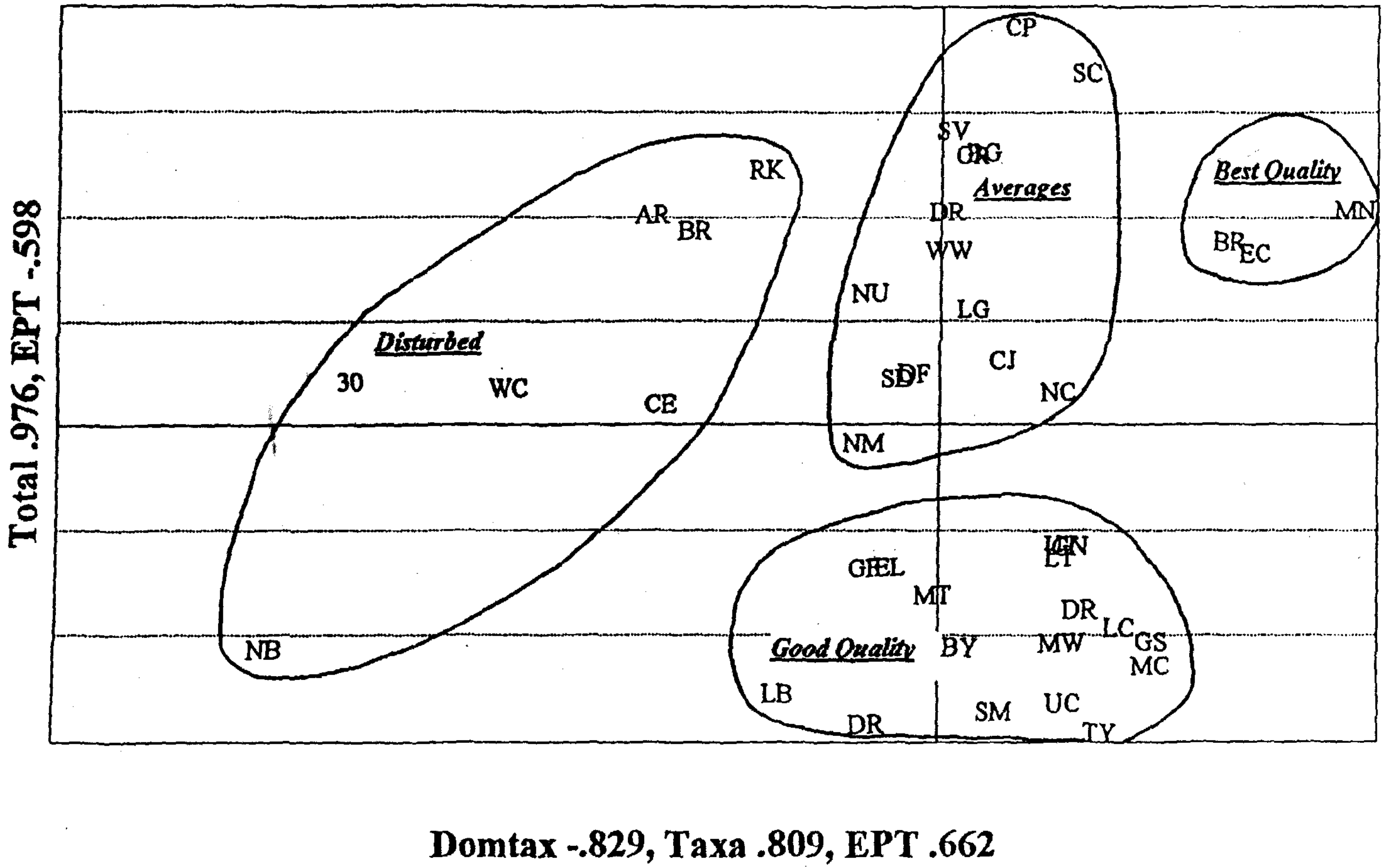
	Total Number/m^2	Total Taxa per stream	EPT	ETP/C	% Dominant Taxa	HBI
Eastern Cascades	6707	37	18	8.5	39.0	3.6
Columbia Plateau	2881	23	11	5.5	43.7	2.7
Northern Rockies	5635	25	13		36.0	4.3
Blue Mountains	4236	42	23	5.4	28.4	3.9
High Desert	5936	26	10	5.5	43.4	5.0

EPT:Ephemeroptera,Plecoptera & Trichoptera; C:Chironomidae
HBI:Hilsenhoff Biotic Index

Map of stream clusters for the Blue Mountains ecoregion.



Clusters of streams plotted from PCA analysis within the Blue Mountains ecoregion.



REORGANIZED ASSEMBLAGE CLASSIFICATIONS

REGION	TAXA	TOTAL	HBI	EPT/C	EPT	DOMTAX
		RIVERINE				
COLUMBIA PLATEAU	6	49			2	70.1
		LOW DIVERSITY,FLASHY				
EASTERN CASCADES	22	8422	4.88	2.33	8	67.79
COLUMBIA PLATEAU	23	5892			8	53.2
		LOW DIVERSITY,HIGH DESERT OR DISTURBED				
BLUE MOUNTAINS	23	4688		2.4	8	45.0
COLUMBIA PLATEAU	22	1879			11	43.8
NORTHERN ROCKIES	20	1404	4.13		11	39.1
HIGH DESERT	25		4.76		10	37.4
COLUMBIA PLATEAU	31	7968			12	33.7
		MODERATE DIVERSITY, GOOD HABITAT				
EASTERN CASCADES	38	4433	3.85	3.92	17	36.3
NORTHERN ROCKIES	38	17790	4.57		21	32.8
BLUE MOUNTAINS	42	4314		5.35	23	29.0
BLUE MOUNTAINS	41	5850		9.1	22	28.4
		HIGH DIVERSITY, COOL HABITAT				
EASTERN CASCADES	38	3915	1.9	20.3	24	27.8
COLUMBIA PLATEAU	32	652			21	24.2
EASTERN CASCADES	54	10561	3.77	2.12	26	23.5
BLUE MOUNTAINS	59	5941		3.44	30	18.2

Unique Taxa in Each Ecoregion

Based on ten most common taxa per stream

	Eastern Rockies	Columbia Plateau	Northern Rockies	Blue Mountains	High Desert
Number Taxa per Ecoregion	64	47	60	60	117
Number Taxa Unique to Each Ecoregion	14	4	11	4	35
Percent of Basinwide Unique Taxa	23	10	18	7	30

WIDESPREAD and COMMON TAXA

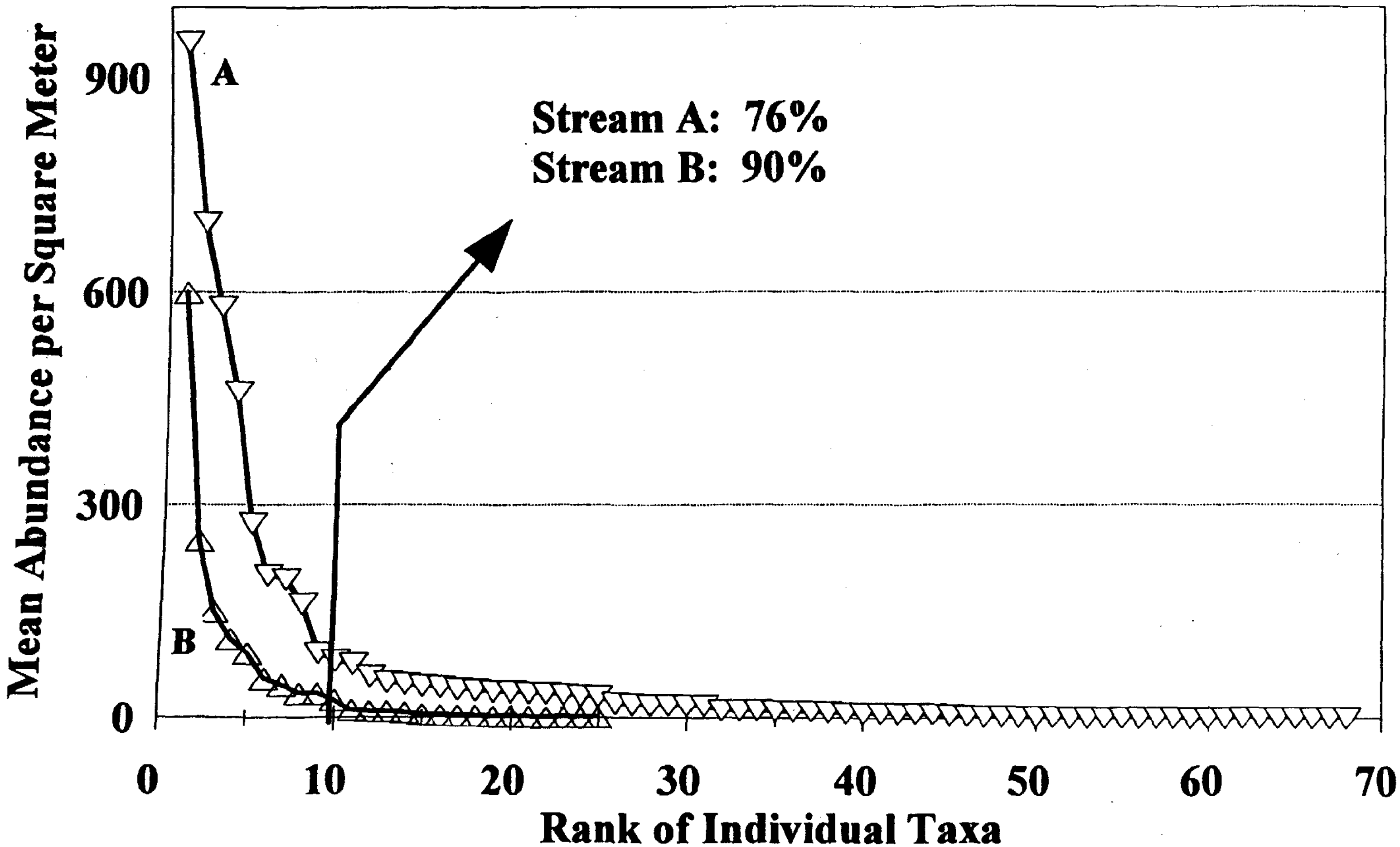
TAXA	Entire Basin	Blue Mountains	Northern Rockies	Columbia Plateau
Baetis	*			
Chironomidae	*			
Cinygmula		*	*	
Rhyacophila		*	*	
Hydropsyche			*	*
Optioservus			*	*

Widespread: among top 10 taxa in more than 1/2 streams within a region

CHARACTERISTIC TAXA

Taxa	Watershed
Yoraperla *Cinygmula*	Eastern Cascades: Crater Lake
Paraleptophlebia	Columbia Plateau: Disturbed streams
Drunella *Seratella* *Epeorus*	Blue Mountains: Grande Ronde
Rithrogena	Blue Mountains: High elevation streams Eastern Cascades: High elevation streams

Figure 2. Taxa abundance curves for two streams in western Oregon, illustrating log series abundance. Each symbol denotes abundance of a particular taxon. Arrow is drawn to show percent of total abundance represented by 10 taxa.



BIOASSESSMENT USING MACROINVERTEBRATES IN LOW-GRADIENT REACHES OF THE SOUTH PLATTE RIVER

Jill Minter
Department of Earth Resources
Colorado State University
Fort Collins, CO 80523

NOTES:

Biomonitoring in low-gradient rivers with shifting-sand substrates has not been fully addressed. This research examined the composition and variability of benthic communities from substrates representing three major habitats in low-gradient reaches of the South Platte River in Colorado and Nebraska. Benthic macroinvertebrates were collected during the summer season, 1993, from shifting-sand, submerged woody debris, and rock or rock-like material at six stations representing sixth and seventh order streams. This research also examined the effects of nutrients and habitat quality on those benthic communities and whether a particular substrate or combination of substrates is more appropriate for distinguishing between stations of differing water quality.

Fifty-seven taxa were collected from all substrates. Thirty-three taxa occurred in the sand substrate and 47 taxa occurred on wood and rock substrates. The proportion of major macroinvertebrate groups varied among substrate and station. Chiromonids predominated sand substrates except for one station impacted by wastewater discharge where Oligochaeta was the predominant group. Wood substrates contained a larger proportion of other Diptera, mayflies (Ephemeroptera) and caddisflies (Trichoptera) and contained very few Oligochaeta. The major groups found on rock substrates were similar to wood substrates, but there was a larger proportion of mayflies and fewer chironomids at reference stations. The number of taxa per replicate was less variable than macroinvertebrate abundance for all substrates and stations. In general, rock substrates were less variable than sand or wood, but clear patterns were not observed. Combining all six stations, 8-40 sand samples would be necessary to estimate number of taxa $\pm$ 10% of the mean compared to 1-16 wood samples and 1-14 rock samples.

Twelve taxa accounted for 91.0% to 99.5% of all organisms collected from any single replicate. Taxa examined individually from wood and rock substrates showed similar changes in patterns of abundance along a longitudinal gradient, and were dissimilar to the sand substrate. Canonical discriminant analysis, based on the abundance of the 12 dominant taxa, was used to examine the overlap and separation of stations for each substrate. Replicates from each station grouped together and were clearly separated from other stations. The dominant taxa responsible for this separation varied by substrate. There was a consistent pattern for all three substrates of the three less impacted stations being grouped together and separated from the three more impacted stations. Correlations were conducted between macroinvertebrate communities, represented by canonical axes 1 and 2, nutrient levels, and measures of habitat quality. High correlations (0.77 to 0.96) were found between macroinvertebrate communities and levels of ammonia, nitrite and nitrate, and phosphorous as well as habitat quality (e.g. riparian and bank structure). High correlations varied by substrate and may indicate those abiotic factors responsible for

shaping macroinvertebrate communities. The results of this research suggest that the choice of substrate or combination of substrates for biomonitoring in low-gradient streams may be dependent upon which abiotic factors are of most interest.

DEVELOPING BIOLOGICAL CRITERIA FOR PACIFIC NORTHWEST STREAMS

Leska Fore and James Karr
Institute for Environmental Studies
Box 352200
University of Washington
Seattle, WA 98195

NOTES:

The goal of biological monitoring is to evaluate the effect of human activities on biological resources. In this study, we linked human activities across landscapes to specific changes in assemblages of benthic macroinvertebrates in streams that drain those landscapes. We used data from second to fourth-order streams in southwestern Oregon to test approximately 30 hypotheses about how macroinvertebrates respond to several common human actions, especially logging and associated road construction. We found 11 attributes of macroinvertebrate assemblages (or metrics) to be reliable indicators of degradation. We outline methods for selecting reliable attributes of the assemblage to monitor changes in water-resource quality.

We constructed a multimetric index (benthic index of biological integrity, or B-IBI) from metrics that distinguished degraded stream sites from minimally degraded sites. Using an independent data set, B-IBI scores were significantly lower for streams whose watersheds were more degraded by human activities. We also tested rapid bioassessment protocol (RBP) III as modified by Oregon Department of Environmental Quality. RBP III failed to detect differences among sites to which B-IBI responded. In a comparison of multivariate and multimetric approaches to explore pattern in multidimensional data, we found that principal components analysis responded primarily to the absence of species--that is, zeros in the data matrix.

The collection of data for biological monitoring is not a goal in itself; data should be collected to answer specific questions about the management of environmental resources. Simple statistical tests are appropriate in some situations, such as when the goal is to regulate a polluter by assessing the sites upstream and downstream of a point source. But when a biologist or statistician reports a significant difference, an alert audience will ask, "How different?" and, "In what?" In a monitoring context, we are less interested in testing for statistical significance than we are in comparisons, but they also provide a yardstick that can be used to rank sites, for example, to determine which sites are the best candidates for restoration.

Draft manuscript submitted to
Journal of the North American Benthological Society

July 20, 1995

A Benthic Index of Biotic Integrity for Streams in the Pacific Northwest

Leska S. Fore and James R. Karr
Institute for Environmental Studies, Box 352200
University of Washington
Seattle, WA 98195

Robert W. Wisseman
Aquatic Biology Associates
3490 NW Deer Run Rd.
Corvallis, OR 97330

Introduction

The Pacific Northwest is on the brink of losing its salmon (Frissell 1993, Nehlsen et al. 1991). In 1994, the salmon fishery was closed along the West Coast from Mexico to Washington because of low returns of spawning adults. The decline in salmon and closure of sport and commercial fisheries (Wilderness Society 1993) represents an annual economic loss in excess of $1 billion and 60,000 jobs (Pacific Rivers Council 1995). This trend is simply the most recent chapter in a history of decline and loss of regional aquatic resources that spans the continent (Hughes and Noss 1992, Allan and Flecker 1993) --Atlantic salmon (Salmo salar) of New England rivers, sport and commercial species in the Midwest (Karr et al. 1985), and native fishes of California (Moyle and Williams 1990). Salmon and other aquatic organisms are disappearing because dams, timber harvest, agriculture, urbanization, and other human actions alter watersheds and degrade rivers that support populations of aquatic species.

Widespread recognition of this degradation has stimulated numerous efforts to improve our ability to track the condition of aquatic ecosystems (Davis and Simon 1995). Comprehensive, multimetric indexes (Barbour et al. 1995) were first developed in the Midwest for use with fishes (Karr 1981, Fausch et al. 1984, Karr et al. 1986) and modified for use with invertebrates (Ohio EPA 1988, Plafkin et al. 1989, Kerans and Karr 1994, DeShon 1995). Although the original index of biotic integrity (IBI) has been adapted for fish in the Willamette River (a large river in Oregon -- Hughes and Gammon 1987), broad application of a fish IBI to the Pacific Northwest is restricted by three problems: 1) streams in the Pacific Northwest have few species of fish; 2) many fish spend a significant portion of their life cycle in marine waters; and 3) salmonids, the dominant taxa in most streams, are affected by harvesting and stocking.

In contrast, streams west of the Cascades host several hundred easily identifiable taxa of macroinvertebrates that complete their life cycle in freshwater. The variety of taxa and knowledge of their natural history makes it possible to distinguish different types and levels of human-induced degradation. Thus, we evaluate the biotic integrity of streams based on analysis of the benthic invertebrate assemblage.

In this paper, we test and modify the benthic IBI (B-IBI) developed for the Tennessee Valley (Kerans and Karr 1994). We test hypotheses about invertebrate responses to degradation and construct an index with reliable statistical properties. Further, we test the resultant B-IBI on other data sets and compare B-IBI with a similar index, Rapid Bioassessment Protocol III (Plafkin et al. 1989) as modified by Oregon Department of Environmental Quality (DEQ) (Mulvey et al. 1992). We also compare the multimetric B-IBI with multivariate approaches to biotic assessment.

Methods

Study location

As is typical of many monitoring studies, our analysis of existing data sets was retrospective. Of the data available from Alaska, Idaho, Oregon, and Washington (EPA's Region 10), we chose streams from southwestern Oregon for initial studies (Table 1) because a large number of sites had been sampled over three years and land-use data were some of the best available. Macroinvertebrates were collected from mid-order streams within the drainages of the Umpqua and Rogue rivers in the western foothills of the Cascade Mountains south of Eugene, Oregon. Watershed area ranged from 15 to 400 km^2 with a median area of 54 km^2. Original vegetation was primarily Douglas Fir (Pseudotsuga menziesii).

Our study area was located in the Umpqua National Forest and in the Roseburg and Medford Districts of the Bureau of Land Management, the latter in a checkerboard pattern with private owners. Within the National Forest, clearcuts have increased the frequency and intensity of runoff which has simplified channel morphology and increased debris torrents (USFS 1992). Logging and road-building have increased sediment and significantly widened stream channels since 1937 (Dose and Roper 1994). Watersheds under BLM and private ownership are modified by timber harvest as well as grazing, agriculture, and urbanization.

Data sets

An ideal data set for testing metrics would include sites that were physically identical and differed only in their degree of human influence. Data available for the Pacific Northwest were not ideal, and so we had to first identify a subset of the data that included a large enough number of physically similar sites. We controlled for the effect of large-scale geographic features by selecting sites within southwest Oregon. We minimized seasonal effects with fall sampling and for year effects by testing within a single year. We selected second and third-order streams and excluded sites from the Umpqua River because the mainstem is a much larger river system with increased taxa richness, different species composition, and markedly different biological dynamics. We divided the data into subsets for each phase of index development according to year and location:

Metric testing - 1990 Umpqua - We tested hypotheses about benthic invertebrate responses to timber harvest and road building with data from 45 sites collected in 1990 from the Umpqua National Forest.

Metric validation - 1992 Umpqua - We validated patterns observed in 1990 metric scores with data from 24 Umpqua sites collected in 1992.

Index testing - 1991 Umpqua and Roseburg - We compared B-IBI scores of 20 Umpqua sites with 30 site scores from the Roseburg area. The different levels of human activity in these two areas made them ideal for a comparison of B-IBI and RBP.

Data collection

Invertebrates. Macroinvertebrates were collected with a 500 micron kicknet. All individuals collected were identified and counted, unless the sample was greater than 1000 individuals, in which case a subsample was processed. Insects were dominant in the benthic assemblages, and were usually identified to genus or species. Level of taxonomic identifications for non-insects varied from species to phyla. Invertebrates were collected in the fall because the largest number of species are typically present at this time of year and are, in general, as large as they get after feeding all summer.

A single location in each watershed yielded one collection of invertebrates for each stream. We analyzed riffle data for this study for three reasons: 1) riffles are easier to define operationally than pools or margins; (2) riffles are more uniform than other stream microenvironments; and (3) riffles have high current velocity and shallow depth, facilitating sampling with kicknets or Surber samplers. Pools provide information that riffles do not and there is some evidence that they are degraded by sediment before riffles (Kerans et al. 1992); however, our priority was to develop a sampling protocol that could be easily applied by others.

Land use. Elevation, location of roads and urban areas, location of private and federal land, and stream size were obtained from USGS topographic maps. Estimates of total watershed area, watershed area logged, and miles of road within each watershed were provided by the U. S. Forest Service (USFS) for Umpqua sites. Unfortunately, information about the type, location and date of timber harvest was not available. For instream condition of sites within the Umpqua NF, we asked USFS hydrologist Mikeal Jones to rank streams based on observations of the riparian corridor, stream bed, bank stability, and the influence of road building and logging on the stream channel. No water chemistry data were available. For BLM sites west of the National Forest, specific information was available for only a few watersheds.

Data analysis

Our goal was to extract relevant biological pattern from taxa lists and counts of macroinvertebrates and relate those patterns to human-induced changes in the watersheds. We evaluated data sets with a multimetric approach (Karr et al. 1986, Kerans and Karr 1994, Barbour et al. 1995) and a multivariate approach (principal components analysis) (Norris and George 1993) to identify patterns in macroinvertebrate assemblages.

Index development. For the multimetric analysis, metrics started out as hypotheses about how invertebrates respond to disturbance. The hypotheses we tested came from published literature (Cummins et al. 1989, Karr and Kerans 1991), existing protocols and multimetric indexes (Plafkin et al. 1989, Hayslip 1993, Kerans and Karr 1994, DeShon 1995), and our field experience. We grouped sites based on instream condition, level of logging, and other important human activities into three groups -- least, moderately, and most degraded. Instead of statistical correlation we used graphical analysis to judge candidate metrics. Metrics that showed little or no overlap between most degraded and least degraded sites and had moderately degraded sites arrayed mostly in between (Figure 1) were selected and tested again with data from a subsequent year.

Metrics from three general classes: taxonomic richness and composition, tolerance and intolerance, and population attributes were selected and combined into an index by transforming the metric values to a score of 5 (similar to or deviated slightly from undisturbed condition), 3 (moderate deviation), or 1 (strong deviation from undisturbed condition) (Karr 1981, 1991). The sum of the metric scores gave an index score for each site. We tested B-IBI by comparing index scores for sites from two areas (Umpqua NF and Roseburg BLM) experiencing different levels of human influence.

We tested the Oregon DEQ version of the RBP because it resolves some of the problems of the original RBP III. Ratio metrics in the original RBP, such as scrapers/filterers and EPT/chironomids, are replaced with metrics such as percentage scrapers and percentage filterers. The original RBP calculated scoring criteria as a percentage of a reference site which we could not do because reference sites were not sampled. In general, setting scoring criteria as a percentage of a reference site is a poor approach because it fails to recognize that reference condition is more meaningful when defined as a range of values rather than as a single value. We used the scoring criteria developed by Oregon DEQ (Mulvey et al. 1992).

Statistical considerations. We could have used a statistical test (e.g., Mann-Whitney test for two independent samples) to decide if the most and least degraded sites were significantly different. Instead, we based our decisions on graphical analysis because it provided more insight into the biology than a simple p-value could. For example, we were interested in whether scores for best and worst sites clumped together tightly or were spread more evenly. Graphs allowed us to examine each metric's range and evaluate where along the continuum of degradation the metric was most sensitive.

An alternative approach might base metric selection on statistical correlations with water chemistry, land use, or similar variables. The main problem with statistical correlation is that no single variable can summarize all the human activities that degrade streams and provide a linear ranking of sites. Watersheds may be degraded by mining, in which case water chemistry data are appropriate; or by channel modification by dams or roads, in which case physical data are relevant; or by hatcheries and stocking, in which cases fish density may be an appropriate correlate for human-induced degradation. In our study area, although logging was the primary land use, the percentage area of the watershed logged did not differentiate between sites that had been clearcut to the river bank and those sites that had left riparian areas intact. Another limitation was that the year of the cut and the frequency of cutting were also unavailable.

We doubt that any single land-use or chemical variable can reliably rank sites, except in unusual circumstances; indeed, that is a primary reason for evaluating river condition according to the resident biology. We relied on our knowledge of freshwater systems to judge which human actions were most important and destructive. We are confident that sites that appear to a human observer to be heavily used and visibly damaged really are more degraded than sites that appear pristine, except, perhaps, in the case of chemical contamination. Along a continuum of degradation, we used the extreme sites to test metrics. Reliable metrics were then combined into an index which can be used to sort out the middle, or moderately degraded sites.

Principal components analysis. For the multivariate analysis, land-use data were not sufficient to perform a canonical correspondence analysis; therefore, we chose principal components analysis of the macroinvertebrate data and identified important site characteristics afterwards on the principal component plots. Principal components analysis calculates the line ("the component") that extracts the maximum amount of statistical variance from a cloud of points (Tabachnick and Fidell 1989). Each point in the cloud represents a stream site. The number of dimensions through which the line passes is equal to the number of taxa collected. The typical application of this procedure to monitoring data uses species lists and abundances to interpret differences between stream sites (Norris and Georges 1993). PCA was based on 46 sites from the 1990 Umpqua National Forest data, the same data set used to test metrics.

Multivariate analyses are founded on the assumption that data are distributed normally; abundance data (as we used for this PCA) typically are not. Two problems in particular are, first, the presence of numerous zeros and, second, frequent high abundances, or right-skewed data. Abundance data are typically transformed by normalizing (i.e., subtracting the mean and dividing by the standard deviation) or by taking the logarithm. Neither approach was very satisfactory because neither transformation resulted in normally distributed variables. We performed PCA for three versions of the same data: normalized, log transformation, and no transformation.

Results

Our first task was to select a subset of physically similar sites to test metrics. Next, we divided sites into three groups according to the intensity of land use in their watersheds. Careful attention to outliers in the land-use data led to data corrections and new hypotheses. We tested metrics by evaluating their response along a gradient of the dominant human influence, with special emphasis on their ability to clearly distinguish the most and least degraded sites. Next, we constructed B-IBI from component metrics and tested B-IBI with a different data set. Finally, we tested the RBP index and analyzed the invertebrate data with PCA.

B-IBI development

Classifying sites and evaluating human influence. Our goal in classifying sites was to obtain a subset of sites that were physically similar and did not have human influence confounded with other physical attributes. For example, high elevation sites were logged less (Figure 2) and had fewer roads in their watershed: thus, high elevation sites were biased toward more pristine conditions. We excluded high elevation sites from the analysis and chose sites below 2500 ft because they formed a cohesive group and were spread more evenly across the range of human influence. High elevation sites also had much higher taxa richness; but we could not determine if high taxa richness was due to pristine conditions or elevation.

Percent area logged measured human influence at the watershed scale; instream condition measured human influence at the riparian scale. We expected measurements at two different scales to roughly agree for these sites because the primary human activity (logging) affected both the watershed and the instream channel (Figure 3). Given the many sources of variability and error in land-use variables, plotting land-use variables against each other provided a check for data consistency. Unexpected outliers pointed us toward data errors and new hypotheses to be tested. For example, two sites with low harvested area but very poor instream condition made us aware of several watersheds for which timber harvest was underestimated. Harvested area was underestimated because estimates were only for public land and did not include logging on private land within the watershed. We corrected these points by estimating that the proportion of area harvested on private land was the same as on forest land, probably an underestimate. Sites with the best instream condition had been logged, but had been left to recover (upper forks of Cow Creek) or had been logged from ridgeline roads rather than roads along the valley floor (Quartz Creek).

For 1990, we selected four sites (Calf Creek, South and East Forks of Cow Creek, and Quartz Creek) as the least degraded and four sites (Brownie and Tom Creek, lower Cow Creek, upper and lower Elk) as the most degraded.

The 1992 data set was smaller and three best sites (upper Cow Creek, N. Boulder Creek, and Quartz Creek) and two worst sites (Elk Creek, lower Jackson Creek) were selected to determine if results were consistent with the 1990 metric analysis. Least degraded sites were not pristine: all watersheds have been logged; rather we chose sites that were the best available. Best sites did not have roads paralleling the mainstem channel, had been logged least or longest ago, and were not easily accessible to humans. The most degraded sites in this data set were affected by roads along the bank of the stream, heavy logging, grazing and urbanization.

Decline in taxa richness is one of the most reliable indicators of degradation (Ford 1989, Barbour et al. 1995) for diverse taxa: periphyton (Bahls 1993); phytoplankton (Schelske 1984); riverine fish (Karr 1981, Miller et al. 1988, Ohio EPA 1988, Lyons 1992, Lyons et al. 1995); estuarine fish (Deegan et al. 1993); and invertebrates (Ohio EPA 1988, Kerans and Karr 1994, DeShon 1995). Therefore, we used taxa richness to test whether road density, presence/absence of roads in the channel, or percent area logged most accurately reflected the impact of human activities on the biota. Taxa richness declined significantly (Spearman's rho, $p < 0.05$) as area logged increased (Figure 4), but did not show a clear response to road density or presence. We were initially concerned that most of the sites were similarly degraded and, therefore, might not provide a broad enough range of degradation to test metrics. Significant correlation between area logged and taxa richness in spite of the inaccuracy of the land-use data gave us confidence that the sites were sampled from a gradient of human-induced change, and that the best and worst sites were truly different. Because our analysis was limited to existing data sets, our sites probably do not include the true extremes.

Testing metrics. Our goal was to construct an index with metrics from each of four broad classes: taxa richness and composition, tolerant and intolerant taxa, feeding ecology, and population attributes. We selected metrics that best distinguished most and least degraded sites for 1990 and 1992 (Table 2). Only metrics that worked for both 1990 and 1992 were included in B-IBI (Table 3). Because 1992 was a drought year, we had greater confidence that metrics selected were minimally influenced by this source of natural variability.

Taxa richness and composition. Total taxa richness and richnesses of Ephemeroptera, Plecoptera, and Trichoptera easily separated the best sites from the poor sites, and, in general, declined monotonically across the range of degradation (See figure 5 for examples). Plecoptera taxa richness declined first in response to human-induced changes, Ephemeroptera next, and Trichoptera last. The presence of Pteronarcys indicated better stream condition. Pteronarcys, a large, long-lived plecopteran, seems susceptible to scour during high flows (R. Wisseman, pers. observ). Neither dipteran nor chironomid taxa richnesses responded reliably to human-induced changes.

Tolerant and intolerant metrics. For tolerant and intolerant metrics, we identified the most tolerant (about 35 out of 200 total taxa) and extremely intolerant taxa (about 35 taxa). We also identified taxa that were specifically tolerant or intolerant to fine sediment which included some of the taxa in the general tolerant/intolerant metrics but others as well. Our goal is to eventually replace the general metrics with more specific metrics that identify what the taxa are tolerant to or intolerant of, for example, toxic contamination, increased temperatures, or scour.

Typically, tolerant metrics are calculated as a percentage of the total number of individuals in a sample;

intolerant metrics are calculated as taxa richness. Tolerant organisms occur at all sites but tolerant taxa tend to dominate as conditions are degraded. Intolerant organisms, when present, are typically a small percentage of the total. Accurate estimates of relative abundance for intolerant taxa are virtually impossible to obtain with a reasonable sampling effort; yet, the presence of these taxa sends a strong signal about local environmental quality. We predicted, therefore, that sediment tolerant organisms scored as a percentage would perform better than as a taxa richness metric and, furthermore, that the opposite would be true for sediment intolerant organisms. Empirical results confirmed theoretical expectations based on previous studies: tolerants performed better calculated as percentages of total abundance and intolerants performed better as taxa richness.

Feeding ecology. We were surprised to find that feeding ecology metrics (Merritt and Cummins 1984) calculated as either percentage of individuals or taxa richness, failed to distinguish the most and least degraded sites (See Figure 5 for examples). No doubt human disturbance affects the trophic composition of the resident biotic assemblage; we suspect the inability to detect a change was caused by the plasticity of many organisms as they develop from nymphs to adults, thus making it difficult to assign taxa to a single group. Percentage of scrapers in the sample distinguished good from poor sites in both 1990 and 1992, but the relationship was reversed: in 1990 the best sites had a smaller percentage of scrapers, but in 1992 the best sites had a higher percentage of scrapers. The year 1992 was unusually dry suggesting that temporal variation in environmental conditions may affect scrapers and possibly the relative abundances of other trophic groups as well.

Population attributes. High abundance relative to other species was tested as percentage of the most abundant taxa in the total sample for one through five most numerically abundant taxa. The best form for this metric calculated dominance as the relative abundance (percentage) of the three most abundant taxa. Total abundance did not satisfy our criteria for selection; nonetheless, we included it because we expect very low abundances only in the most extreme situations, such as when high levels of urban or industrial contaminants are present (unpublished data from urban streams, Karr 1981). Very high abundances may also indicate degradation, particularly in agricultural areas where nutrient enrichment affects invertebrates. In general, abundance (or population size or density) is one of the single most variable aspects of an assemblage; thus, all other metrics were calculated as a percentage of total abundance (relative abundance) or as taxa richness.

Constructing B-IBI. Metrics differed in their range of values; therefore, we transformed them to a similar scale before combining them into a summary index, B-IBI. Following the method developed for the fish IBI (Karr 1981, Karr et al. 1986), we divided the range of each metric into three parts and assigned a new score of 5 (indicates similar to expected, or reference, condition), 3, or 1 (indicates strong deviation from expected condition) (Table 3, Figure 6). Metrics were added to get a final B-IBI score (maximum = 55; minimum = 11). We determined the potential range of metrics based on 82 sites: 34 from the Umpqua NF (1990), 30 from the Roseburg BLM (1991), and 18 from the Medford BLM (1991).

Simple, general rules for setting metric scoring criteria are difficult to define because they depend on the original sampling design that generated the data. If sites from a region are sampled such that one-third are located in pristine areas, a third in moderately degraded areas, and a third in very degraded areas, then one can start by dividing the range of each metric at the 33rd and 67th percentiles. In our case, we suspected that our data sets included neither very degraded sites, e.g., sites from urban areas, nor pristine sites with no human influence. Therefore, we tried to err on the conservative side by broadening the middle category (3) to include more sites, thus making it more difficult for a site to score a 5 or 1. Natural breakpoints in the distribution probably reflect relevant biology. Where these occurred, we adjusted the scoring criteria to fall at these points. Setting scoring criteria is an iterative process and should be revisited as more regional data become available across a broader range of human influences.

Part of the data (30 Roseburg sites) used to set scoring criteria were used to test the B-IBI. We do not expect their inclusion to bias our test of the index because they were only used to assess the range of metrics. For most metrics, very low and very high scores were found in the Umpqua NF, Roseburg, and Medford BLM areas. It was not the case that Roseburg sites scored 1 for most metrics and Umpqua scored 5; rather, average metric scores for the Umpqua were higher than for Roseburg, which provided another check of the metrics.

Testing B-IBI

We tested the hypothesis that B-IBI scores for Umpqua stream sites would be higher, reflecting higher biotic integrity, than for Roseburg sites. Although the range of B-IBI scores was similar for the two areas, the poorest sites in Roseburg were more degraded than the poorest sites in the Umpqua. The Roseburg area includes BLM and private land in a checkerboard pattern. Private land, in general, has been logged more heavily and has been degraded further by urbanization, roads, mining and agriculture. A significant difference between index scores for the two areas ($p < 0.05$; Figure 7) indicated that B-IBI reflects the differing effects of human-induced degradation on instream biota in the two study regions.

Testing RBP

The RBP index as modified by Oregon DEQ (Table 4) failed to detect a difference between sites in Roseburg BLM and Umpqua NF ($p > 0.05$; Figure 7). RBP metrics represent a set of working hypotheses that must be tested (like all metrics) with data from the region to be monitored. RBP metrics have not been tested in the Northwest (and only minimally tested elsewhere). We tested the component RBP metrics with the same criteria described above and found that only four of the metrics could distinguish between most and least degraded sites (Table 4). Two of those metrics were included in B-IBI.

Multivariate Analysis

PCA results were similar for the three versions of the data (normalized, log transformed, and no transformation). Sites identified as most and least degraded, large river, and high elevation showed some tendency to group together for all three data versions of PCA when plotted against the first and second principal components; however, the groups defined by the sites were not exclusive and showed considerable overlap (Figure 8). Sites that were outliers from their groups in one analysis were also outliers in the other two analyses. The most degraded sites grouped together better than the least degraded sites for all three analyses.

We suspect that zeros in the data matrix had a strong effect on the solution of the PCA (Tabachnick and Fidell 1989) for two reasons. First, data transformation had little effect on the PCA outcome, probably because zero abundances were not changed by transformation. Second, degraded sites grouped together best, probably because, with the fewest taxa, they contained the largest number of zeros in the data matrix.

We looked for patterns in the loadings of the 198 taxa on the principal components. We found a slight tendency for Ephemeroptera, Plecoptera, and Trichoptera taxa to load more frequently on the component which separated most and least degraded sites (PCI). We did not detect patterns in the loadings for intolerant and tolerant taxa. Other taxa grouped together on principal component axes provided little additional insight into the structure of the assemblage. Rather than trying to work through the statistical analysis and determine how and why species were grouped together, we found it easier to compare sites and search for patterns in the raw data by simply noticing which taxa were present or absent in similar sites. By focusing on biological patterns associated with specific, testable hypotheses, we were better able to describe and detect the environmental condition of sites than we could from multivariate plots.

Discussion

In this paper, we emphasize the process of testing metrics and developing an index. Even with a rather narrow range of disturbance, we were able to test metrics and develop an index that responded to human-induced degradation. We selected metrics according to their response to levels of human disturbance in the watershed, primarily logging and road building. One may ask, if the biological condition of sites can be determined so easily, why bother with biological monitoring? Although sites near the ends of the spectrum are easy to judge, moderately degraded sites are not. Thus we started with the extremes, tested our hypotheses, and applied the results to sort the middle. Judging the condition of streams based on the instream biota allows us to look for smaller differences. Large scale differences that a human observer notices, such as large woody debris, may not be as important to the resident biology as other aspects of the river that humans can not see or measure.

Too often biological patterns in complex data sets are reported in terms of complicated, high level statistics, such as principal components analysis. In contrast, we demonstrated how graphical analysis and simple statistics can be used to interpret complex data and understand biological patterns. To many field biologists, "statistics" means "multivariate statistics." Monitoring reports have been using the same multivariate techniques since the 1960's in spite of a plethora of statistical techniques available (Potvin and Travis 1993). Multivariate statistical analyses make decisions based on statistical properties of the data, e.g., the structure of the covariance matrix, rather than the biological judgement of the investigator. Because most multivariate analyses are based on taxa lists and abundances, they fail to incorporate our biological knowledge of the animals' natural history, and especially their responses to human activities. Multivariate analyses are most appropriate for exploratory analysis, that is, developing hypotheses about systems for which little is known. We know enough about freshwater systems to test hypotheses rather than continue generating them; we must move forward with testing hypotheses directly rather than simply reporting complexity.

The object of a good index is not to include and measure every aspect of stream biology that responds to human influence; biological systems are too complicated to capture completely. The effects of human-induced degradation on rivers and streams that biologists are being asked to measure in a monitoring context are not subtle. Multimetric indexes formalize what any good biologist, familiar with local biota, knows about the biological condition of a stream. Narrative and numeric indexes are essential tools for communicating biological condition to policy makers and concerned citizens so that they can make informed decisions about the management of aquatic resources.

Metric development

Metrics were calculated as taxa richness or percentages depending on the attribute they were measuring. A good biomonitoring index should not be sensitive to slight differences in rainfall or location; therefore, we are less interested in which particular species of stonefly is found at a site because the identity of the species may depend on subtle microhabitat differences or broader biogeographic patterns. Similarly, species abundance responds dramatically to small changes in the environment. On the other hand, taxa richness of stoneflies, for example, is robust to natural variability. We measure how many stonefly taxa are present, because most stoneflies require large cobble, and fast, clean, cool water.

Trophic composition metrics (e.g., shredders, scrapers, and predators), tolerance metrics, and dominance (relatively most abundant taxa) metrics are scored as percentages rather than taxa richness because they evaluate processes at the level of the assemblage. Taxa in these groups may be present at all sites, but their relative proportion changes with degradation. We also tested trophic composition metrics as taxa richness metrics when they failed to discriminate sites as percentages; neither approach worked. We suspect that trophic composition is affected by human disturbance and that these metrics did not respond because the feeding ecology of many taxa change as they develop into adults.

Although ratio metrics have been included in other monitoring protocols, e.g., ratio of scrapers to filterers, we recommend against them for two reasons. First, the numerator and denominator vary together; therefore, a pair of very large values for scrapers and filterers may yield the same ratio value as a pair of very small values. We expect that large values for these two groups may have a different biological meaning than small values. In short, if two attributes of an assemblage are potentially important, they should be evaluated individually. Second, when two variables are combined as a ratio, the ratio tends to have higher variance than either variable alone (Sokal and Rohlf 1981). Barbour et al. (1992) found ratio metrics for invertebrates to be more variable than other metrics. On the other hand, percentages, or relative abundances, are based on the binomial distribution and are much more reliable.

Our goal is to replace the general tolerant and intolerant metrics with more specific metrics that describe what the organisms are tolerant to; for example, the sediment tolerant and intolerant metrics developed for these data.

As appropriate data sets become available, we suspect another more inclusive metric will include Pteronarcys along with other taxa, such as freshwater clams, that are sensitive to scour because of their large adult size and time needed to complete their life cycle.

"Indexes"

To most biologists, the word "index" carries with it semantic baggage that is the legacy of diversity indexes. Multimetric indexes are often condemned along with diversity indexes (Suter 1993, Norris and Georges 1993) although they differ fundamentally in what they measure and how they are constructed (Fausch et al. 1990). Biomonitoring indexes can be divided into three types: (1) diversity and similarity indexes; (2) pollution tolerant indexes, or "biotic indexes"; and (3) multimetric indexes.

Diversity indexes and most similarity indexes combine information about abundance (or density) and species richness into a measure of "evenness." The concept of evenness was developed in the context of information theory, but has not been meaningfully interpreted in terms of ecological theory (Hurlburt 1971). Diversity indexes sometimes respond to degradation because they are strongly correlated with taxa richness (Camargo 1992), an attribute of the assemblage that, on its own, is easy to interpret in terms of biology. The response of these indexes to systematic changes in the assemblage are often erratic, inconsistent, dependent on initial conditions, and can give misleading interpretations of biological data (Wolda 1981, Boyle et al. 1990).

Pollution tolerance indexes, or biotic indexes, assign a pollution tolerance value to every species and calculate an index score for a site as a function of the number of individuals of each tolerance class (Chutter 1972, Winget and Mangum 1979, Hilsenhoff 1982, Lenat 1993). These indexes estimate an average pollution tolerance for the assemblage. In western streams, chemical and organic pollution resulting from agriculture and urban sewage represent only part of the problem. Grazing, timber harvest, and road building affect streams and their resident biota differently than organic pollution. A potential problem with this index format is that by calculating the tolerance of every organism, even those that are not particularly sensitive, a strong response by a few taxa can be missed if the assemblage is dominated by a large number of insensitive taxa; that is, the signal may be lost in the noise. We avoided this problem with tolerant and intolerant metrics by concentrating on the responses of only the most and least tolerant organisms.

Each component metric of a multimetric index measures an attribute of the assemblage that is the product of evolutionary and biogeographic processes at a site (Karr 1995). Metrics are interpreted (and assigned a score of 5, 3, or 1) based on metric values observed in pristine, or best available, sites. The n metrics represent n measures of the biotic integrity of a stream. Variance is reduced and precision increased by taking an average (or sum) of several measurements because variance decreases as a function of sample size. The functional form of multimetric indexes allows us to take advantage of properties of the mean. The Central Limit Theorem (Cassella and Berger 1990) states that averages tend to be distributed normally; consequently, multimetric index scores can be tested with familiar statistics, such as ANOVA or regression (Fore et al. 1994).

Multimetric indexes are structured more like familiar economic indexes such as the consumer price index or the index of leading indicators (Mitchell and Burns 1938). These indexes summarize the current condition or state of the economy and reduce the inherent natural variability by including the prices of several items or the current levels of several indicators. As for multimetric indexes, economic indicators are combined into an overall index by comparing them to a standard year; after conversion, index components are summed. These composite indexes have survived 60 years of discussion and criticism and are still widely used to interpret economic trends by economists, policy makers, and citizens (Auerbach 1982).

Making decisions

The collection of data for biological monitoring is not a goal in itself; data should be collected to answer specific questions about the management of environmental resources (Yoder 1994). Simple statistical tests are appropriate in some situations, such as, when the goal is to regulate a polluter by assessing sites upstream and downstream of a point source. But when a biologist or statistician reports a significant difference, an alert audience will ask, How different? and, In what way? Detecting significance is dependent on sample size: given a large enough sample size a difference can always be detected (Peterman 1990). In a monitoring context, we are less interested in testing for statistical significance than we are in evaluating differences (Stewart-Oaten et al. 1992), and especially their biological relevance. Multimetric indexes can be used for statistical comparisons, but they also provide a yardstick that can be used to rank sites, for example, to determine which sites are the best candidates for restoration. Index scores for similar sites within a region provide a context for interpreting scores and identifying trends.

The complexity captured by multivariate analyses make them poor communication tools for an audience composed of non-scientists. Management decisions can (and have been) based on multivariate statistical analysis of biological data (Davies et al. 1993, Reynoldson and Zarull 1993, Wright et al. 1993); but the decision process is opaque to anyone less than an expert in these types of analyses. Complicated statistical decision rules are not easily adapted to a simple task like ranking streams from most to least degraded. Multimetric indexes condense and summarize information and can be used to compare sites over a large geographic area. On a smaller scale, the component metrics are available to make more site-specific assessments, such as, to pinpoint sources of degradation (Yoder and Rankin 1995b) or identify what aspects of the assemblage a restoration affects. Although a single number, the index, is used to evaluate the relative condition of sites within a region, detailed information about individual sites is not lost, but is recorded in the individual metrics (Simon and Lyons 1995).

The temptation (and the pressure) exists to establish biological criteria as soon as possible; but if an index is not composed of biologically meaningful and reliable metrics, the biological criteria will be useless. The RBP index failed to detect differences in southwestern Oregon because the component metrics were never adequately tested.

Fine-tuning B-IBI

Some of the steps involved in developing a biomonitoring index may need to be revisited as data accumulate, but we expect many of the conclusions from this study to hold true across the Pacific Northwest region.

In particular, the metrics that responded reliably to the effects of logging will likely respond to other types of degradation such as mining or agriculture. Whether the metric increases or decreases as human influence increases will probably also be consistent. New metrics may be tested and included in the index, but we expect the majority of the metrics to be retained. The ranges for the scoring criteria will probably shift as we evaluate data from more severely degraded streams (e.g., Puget Sound Basin). Our goal is for B-IBI to be composed of metrics that are generally applicable for large regions.

A good multimetric index includes a balance of metrics that respond across the range of degradation. Because our data sets were from moderately degraded sites, we were unable to specifically test the sensitivity of metrics under extremely degraded conditions. Abundance and percent sediment tolerants were sensitive at the most degraded end of the spectrum while taxa richness of Plecoptera and sediment intolerants were more sensitive for less degraded sites. Total taxa richness has a more monotonic response across the range of disturbance. As new data are evaluated, we intend to assess the redundancy of metrics and the overall balance of B-IBI (Kerans and Karr 1994).

By carefully selecting component metrics we can minimize the effects of natural geographic variability and weather so that changes in the assemblage that result from human actions are revealed in sharp relief. Complete separation of the effects of natural variability from human-induced variability is not realistic because disturbance exacerbates the effects of natural events (Schlosser 1990); for example, flood events are more extreme in damaged watersheds. Multimetric index scores reflect this reality and consequently are more variable for heavily degraded sites (Karr et al. 1987, Steedman 1988, Fore et al. 1994, Yoder and Rankin 1995a).

Statistical variability of the index determines how precise a measurement tool it is for comparing sites and making management decisions. State and federal agencies have not yet agreed on the number of individuals to be identified from an invertebrate sample or the number of replicates needed from a stream site. We intend to evaluate the statistical power of B-IBI for the different sampling protocols in order to determine how many classes of integrity B-IBI can distinguish, as was done for the fish IBI (Fore et al. 1994). If we can reduce the amount of effort involved in identifying samples without sacrificing reasonable precision, the money saved for each site assessment translates into more streams visited each year.

Acknowledgments

We thank M. Jones and J. Dose, Umpqua Ranger District, L. Lindell, Medford BLM, and E. Rumble and S. Hofford, Roseburg BLM for providing information on land use. M. Lagerloef (EPA, Region 10), G. Hayslip (EPA, Region 10), R. Hafele (Oregon Department of Environmental Quality), R. Plotnikoff (Washington Department of Ecology), and D. Zaroban (Idaho Department of Environmental Quality) participated in many discussions that guided the development of this manuscript. P. Larsen, R. O'Connor, and R. Hughes reviewed an early version of this manuscript. The research described in this article has been funded wholly or in part by the U.S. Environmental Protection Agency under Cooperative Agreement #3B1053NAEX to JRK.

Literature Cited

Allan, J. D., and A. S. Flecker. 1993. Biodiversity conservation in running waters. BioScience 43:32-43.

Auerbach, A. J. 1982. The index of leading indicators: "Measurement without theory", thirty-five years later. Review of Economics and Statistics 64(4):589-595.

Bahls, L. L. 1993. Periphyton Bioassessment Methods for Montana Streams. Water Quality Bureau, Department of Health and Environmental Sciences, Helena, MN.

Barbour, M. T., J. L. Plafkin, B. P. Bradley, C. G. Graves, R. W. Wisseman. 1992. Evaluation of EPA's rapid bioassessment benthic metrics: metric redundancy and variability among reference stream sites. Environmental Toxicology and Chemistry 11:437-449.

Barbour, M. T., J. B. Stribling, and J. R. Karr. 1995. Multimetric approach for establishing biocriteria and measuring biological condition. Pages 63-77 in W. S. Davis and T. P. Simon (editors). Biological assessment and criteria: tools for water resource planning and decision making. Lewis Publishing, Boca Raton.

Boyle, T. P., G. M. Smillie, J. C. Anderson, D. R. Beeson. 1990. A sensitivity analysis of nine diversity and seven similarity indices. Research Journal of the Water Pollution Control Federation 62:749-762.

Camargo, J. A. 1992. New diversity index for assessing structural alterations in aquatic communities. Bulletin of Environmental and Contaminant Toxicology 48:428-434.

Casella, G., and R. L. Berger. 1990. Statistical inference. Wadsworth, Inc., Belmont, CA.

Chutter, F. M. 1972. An empirical biotic index of the quality of water in South African streams and rivers. Water Research 6:19-30.

Cummins, K. W., M. A. Wilzbach, D. M. Gates, J. B. Perry, and W. B. Taliaferro. 1989. Shredders and riparian vegetation. BioScience 39:24-30.

Davies, S. P., L. Tsomides, D. L. Courtemanch, and F. Drummond. 1993. Maine Biological Monitoring and Biocriteria Development Program. Maine Department of Environmental Protection, Bureau of Water Quality Control, Division of Environmental Evaluation and Lake Studies, Augusta, Maine.

Davis, W. S., and T. P. Simon (editors). 1995. Biological assessment and criteria: tools for water resource planning and decision making. Lewis Publishing, Boca Raton.

DeShon, J. E. 1995. Development and application of the invertebrate community index (ICI). Pages 217-244 in W. S. Davis and T. P. Simon (editors). Biological assessment and criteria: tools for water resource planning and decision making. Lewis Publishing, Boca Raton.

Deegan, L. A., J. T. Finn, S. G. Ayvasian, and C. Ryder. 1993. Feasibility and Application of the Index of Biotic Integrity to Massachusetts estuaries (EBI). Final Project Report to Massachusetts Executive Office of Environmental Affairs, Department of Environmental Protection, North Grafton, MA.

Dose, J. J., and B. B. Roper. 1994. Long-term changes in low-flow channel widths within the South Umpqua Watershed, Oregon. Water Resources Bulletin 30:993-1000.

Fausch, K. D., J. R. Karr, and P. R. Yant. 1984. Regional application of an index of biotic integrity based on stream fish communities. Transactions of the American Fisheries Society 113:39-55.

Fausch, K. D., J. Lyons, J. R. Karr, and P. L. Angermeier. 1990. Fish communities as indicators of environmental degradation. American Fisheries Society Symposium 8:123-144.

Ford, J. 1989. The effects of chemical stress on aquatic species composition and community structure. Pages 99-144 in S. A. Levin, M. A. Harwell, J. R. Kelly, K. D. Kimball (editors) Ecotoxicology: problems and approaches. Springer Verlag, New York, NY.

Fore, L. S., J. R. Karr and L. L. Conquest. 1994. Statistical properties of an index of biotic integrity used to evaluate water resources. Canadian Journal of Fisheries and Aquatic Sciences 51:1077-1087.

Frissell, C. A. 1993. Topology of extinction and endangerment of native fishes in the Pacific Northwest and California (USA). Conservation Biology 7(2):342-354.

Hayslip, G. A. 1993. EPA Region 10 in-stream biological handbook (for wadable streams in the Pacific Northwest). EPA-910-9-92-013. U. S. Environmental Protection Agency-Region 10, Environmental Services Division, Seattle, Washington.

Hilsenhoff, W. L. 1982. Using a biotic index to evaluate water quality in streams. Department of Natural Resources, Madison, Wisconsin. Technical Bulletin Number 132.

Hughes, R. M., and J. R. Gammon. 1987. Longitudinal changes in fish assemblages and water quality in the Willamette River, Oregon. Transactions of the American Fisheries Society 116:196-209.

Hughes, R. M., and R. F. Noss. 1992. Biological diversity and biological integrity: current concerns for lakes and streams. Fisheries 17(3):11-19

Hurlburt, S. H. 1971. The nonconcept of species diversity: a critique and alternative parameters. Ecology 52(4):577-586.

Karr, J. R. 1981. Assessment of biotic integrity using fish communities. Fisheries 6(6):21-27.

Karr, J. R. 1991. Biological integrity: a long-neglected aspect of water resource management. Ecological Applications 1:66-84.

Karr. J. R. 1995. Ecological integrity and ecological health are not the same. In P. Schulze (editor). Engineering within ecological constraints. National Academy of Engineering. National Academy Press, Washington D. C.

Karr, J. R., K. D. Fausch, P. L. Angermeier, P. R. Yant, and I. J. Schlosser. 1986. Assessment of biological integrity in running water: a method and its rationale. Illinois Natural History Survey Special Publication, Number 5, Champaign, Illinois.

Karr, J. R., and B. L. Kerans. 1991. Components of biological integrity: their definition and use in development of an invertebrate IBI. Pages 1-16 in W. S. Davis and T. P. Simon (editors). Proceedings of the 1991 Midwest pollution control biologists meeting. EPA 905-R-92-003. U. S. Environmental Protection Agency, Environmental Sciences Division, Chicago, Illinois.

Karr, J. R., L. A. Toth, and D. R. Dudley. 1985. Fish communities of Midwestern rivers: a history of degradation. BioScience 35:90-95.

Karr. J. R., P. R. Yant, and K. D. Fausch. 1987. Spatial and temporal variability of the index of biotic integrity in three midwestern streams. Transactions of the American Fisheries Society. 116: 1-11.

Kerans, B. L., and J. R. Karr. 1994. A benthic index of biotic integrity (B-IBI) for rivers of the Tennessee Valley. Ecological Applications 4(4):768-785.

Kerans, B. L., J. R. Karr, and S. A. Ahlstedt. 1992. Aquatic invertebrate assemblages: spatial and temporal differences among sampling protocols, Journal of the North American Benthological Society 11:377-390.

Lenat, D. R. 1993. A biotic index for the southeastern U. S.: derivation and list of tolerance values, with criteria for assigning water-quality ratings. Journal of the North American Benthological Society 12(3):279-290.

Lyons, J. 1992. Using the index of biotic integrity (IBI) to measure environmental quality in warmwater streams of Wisconsin. US Forest Service, United States Dept. of Agriculture, St. Paul, MN. General Technical Report NC-149.

Lyons, J. S. Navarro-Perez, P. A. Cochran, E. Santana C., and M. Guzman-Arroyo. 1995. Index of biotic integrity based on fish assemblages for the conservation of streams and rivers in west-central Mexico. Conservation Biology 9:569-584.

Merritt, R. W., and K. W. Cummins (editors). 1984. An Introduction to the Aquatic Insects of North America. 2nd edition. Kendall/Hunt Publishing Company, Dubuque, Iowa.

Miller, K. L., and 13 other authors. 1988. Regional applications of an index of biotic integrity for use in water resource management. Fisheries 13(5):12-20.

Mitchell, W. C., and A. F. Burns. 1938. Statistical Indicators of Cyclical Revivals. National Bureau of Economic Research, New York, NY.

Moyle, P. B., and J. E. Williams. 1990. Biodiversity loss in the temperate zone: decline of the native fish fauna of California. Conservation Biology 4:275-284.

Mulvey, M., L. Caton, and R. Hafele. 1992. Oregon nonpoint-source monitoring protocols and stream bioassessment field manual for macroinvertebrates and habitat assessment. Oregon Department of Environmental Quality Laboratories, Portland, Oregon.

Nehlsen, W., J. E. Williams, and J. A. Lichatowich. 1991. Pacific salmon at the crossroads: stocks at risk from California, Oregon, Idaho, and Washington. Fisheries 16(2):2-21.

Norris, R. H., and A. Georges. 1993. Analysis and interpretation of benthic macroinvertebrate surveys. Pages 234-285 in D. M. Rosenberg and V. H. Resh (editors). Freshwater biomonitoring and benthic macroinvertebrates. Chapman and Hall, New York.

Ohio EPA. 1988. Biological criteria for the protection of aquatic life, volumes 1-3. Ecological Assessment Section, Division of Water Quality Monitoring and Assessment, Ohio EPA, Columbus, Ohio.

Pacific Rivers Council. 1995. A call for a comprehensive watershed and wild fish conservation program in eastern Oregon and Washington. 2nd edition. Pacific Rivers Council, Eugene, Oregon.

Peterman, R. M. 1990. Statistical power analysis can improve fisheries research and management. Canadian Journal of Fisheries and Aquatic Sciences. 47:2-15.

Plafkin, J. L., M. T. Barbour, K. D. Porter, S. K. Gross, and R. M. Hughes. 1989. Rapid bioassessment protocols for use in streams and rivers: benthic macroinvertebrates and fish. Assessment and Water Protection Division, U.S. Environmental Protection Agency, Report EPA/440/4-89-001. Washington: EPA.

Potvin, C., and J. Travis. 1993. Special feature. Statistical methods: an upgrade for biologists. Ecology 74: 1614-1676.

Reynoldson, T. B., and M. A. Zarull. 1993. An approach to the development of biological sediment guidelines. Pages 177-200 in S. Woodley, J. Kay, and G. Francis (editors). Ecological integrity and the management of ecosystems. St. Lucie Press, Delray Beach, Florida.

Schelske, C. L. 1984. In situ and natural phytoplankton assemblage bioassays. Pages 15-47 in L. E. Shubert (editor). Algae as ecological indicators. Academic Press, London, England.

Schlosser, I. J. 1990. Environmental variation, life history attributes, and community structure in stream fishes: implications for environmental management and assessment. Environmental Management. 14:621-628.

Simon, T. P., and J. Lyons. 1995. Application of the index of biotic integrity to evaluate water resource integrity in freshwater ecosystems. Pages 245-262 in W. S. Davis and T. P. Simon (editors). Biological assessment and criteria: tools for water resource planning and decision making. Lewis Publishing, Boca Raton.

Sokal, R. R., and F. J. Rohlf. 1981. Biometry. 2nd edition. Freeman and Company, New York.

Steedman, R. J. 1988. Modification and assessment of an index of biotic integrity to quantify stream quality in southern Ontario. Canadian Journal of Fisheries and Aquatic Sciences 45: 492-501.

Stewart-Oaten, A., J. R. Bence, and C. W. Osenberg. 1992. Assessing effects of unreplicated perturbations: no simple solutions. Ecology 73:1396-1404.

Suter, G. W., II. 1993. A critique of ecosystem health concepts and indexes. Environmental Toxicology and Chemistry 12:1533-1539.

Tabachnick, B. G., and L. S. Fidell. 1989. Using multivariate statistics. 2nd edition. HarperCollins Publishers, Inc., New York.

US Forest Service (USFS). 1992. Wild and scenic river eligibility and determination. Umpqua National Forest, Roseburg, Oregon.

The Wilderness Society. 1993. The living landscape. Volume 1. Wild salmon as natural capital; volume 2. Pacific salmon and federal lands. The Wilderness Society, Washington, D.C.

Winget, R. N., and F. A. Mangum. 1979. Biotic condition index: integrated biological, physical and chemical stream parameters for management. US Forest Service, Intermountain Region, Provo, Utah.

Wolda, H. 1981. Similarity indices, sample size, and diversity. Oecologia 50:296-302,

Wright, J. F., M. T. Furse, and P. D. Armitage. 1993. RIVPACS -- a technique for evaluating the biological quality of rivers in the U. K. European Water Pollution Control 3:15-25.

Yoder, C. O. 1994. Integrating basic and applied benthic science. Journal of the North American Benthological Society 13(3):391-398.

Yoder, C. O., and E. T. Rankin. 1995a. Biological criteria program development and implementation in Ohio. Pages 109-144 in W. S. Davis and T. P. Simon (editors). Biological assessment and criteria: tools for water resource planning and decision making. Lewis Publishing, Boca Raton.

Yoder, C. O., and E. T. Rankin. 1995b. Biological response signatures and the area of degradation value: new tools for interpreting multimetric data. Pages 263-286 in W. S. Davis and T. P. Simon (editors). Biological assessment and criteria: tools for water resource planning and decision making. Lewis Publishing, Boca Raton.

TABLE 1. For each data set, number of sites available, number that were appropriate for B-IBI development (in parentheses), and analysis for which data were used (footnotes). High elevation sites and large river sites were excluded for index development but were used in PCA. After exclusion of these sites, some data sets had too few sites remaining for analysis.

Year	Medford	Roseburg	Umpqua
1990			46[a] (34[b,c])
1991	22 (18[c])	30 (30[c,d])	30 (20[d])
1992	21 (0)	12 (0)	30 (24[b])

[a] PCA

[b] Metric testing

[c] Determine ranges of metrics

[d] Test B-IBI and RBP

TABLE 2. Metrics tested; their predicted response to degradation; and their ability to clearly distinguish least degraded from most degraded sites.

Metric	Predicted response	Metrics that could distinguish sites
Taxa richness & composition		
Total number of taxa [a]	Decrease	X
Ephemeroptera taxa [a]	Decrease	X
Plecoptera taxa [a]	Decrease	X
Pteronarcys taxa [a]	Decrease	X
Trichoptera taxa [a]	Decrease	X
Dipteran taxa	Decrease	
Chironomid taxa	Increase	
Tolerant / intolerant		
Intolerant taxa [a]	Decrease	X
Sediment intolerant taxa [a]	Decrease	X
Sediment tolerant taxa	Increase	X
% Tolerant [a]	Increase	X
% Sediment tolerant [a]	Increase	X
% Sediment intolerant	Decrease	
% Oligochaetes	Increase	
% Chironomids	Increase	
Feeding ecology		
% Scraper		
% Predator	Decrease	
% Gatherer		
% Shredder		
% Filterer		
% Omnivore	Increase	
Scraper taxa		
Shredder taxa		
Omnivore taxa	Increase	
Population attributes		
Abundance [a]	Decrease	
% Dominance (3 taxa) [a]	Increase	X

[a] Component metric included in benthic index of biotic integrity (B-IBI).

TABLE 3. Metrics included in B-IBI for Oregon streams, their response to human influence, and scoring criteria. A score of 5 indicates a metric score that deviates little or none from expected condition, 3 indicates moderate deviation, and 1 strong deviation from expectation.

Metric	Response to degradation	1	3	5
Total number of taxa	Decrease	< 40	40 - 54	> 54
Ephemeroptera taxa	Decrease	< 8	8 - 11	> 11
Plecoptera taxa	Decrease	< 6	6 - 9	> 9
Trichoptera taxa	Decrease	< 6	6 - 9	> 9
Pteronarcys taxa	Decrease	0	NA	> 0
Intolerant taxa	Decrease	< 2	2-5	> 5
% Tolerant	Increase	> 0.4	0.2 - 0.4	< 0.2
Sediment intolerant taxa	Decrease	0	1	> 1
% Sediment tolerant	Increase	> 0.15	0.05 - 0.15	< 0.05
% Dominance (3 taxa)	Increase	> 0.55	0.40 - 0.55	< 0.40
Abundance	Decrease	< 500	500 - 1500	> 1500

TABLE 4. Metrics included in Oregon DEQ's RBP, their predicted response to human influence, and whether they could distinguish least degraded from most degraded sites.

Metric	Predicted response	Could it distinguish?
Total number of taxa	Decrease	yes
EPT taxa	Decrease	yes
% EPT taxa	Decrease	yes
% Chironomids	Increase	no
% Dominance (1 taxon)	Increase	no; 2 or 3 taxa works better
% Shredders	Increase	no
% Filterers	Increase	no
% Scrapers	Decrease	no
Hilsenhoff biotic index	Increase	no
Diversity index, H'	Decrease	yes; but strongly correlated with taxa richness

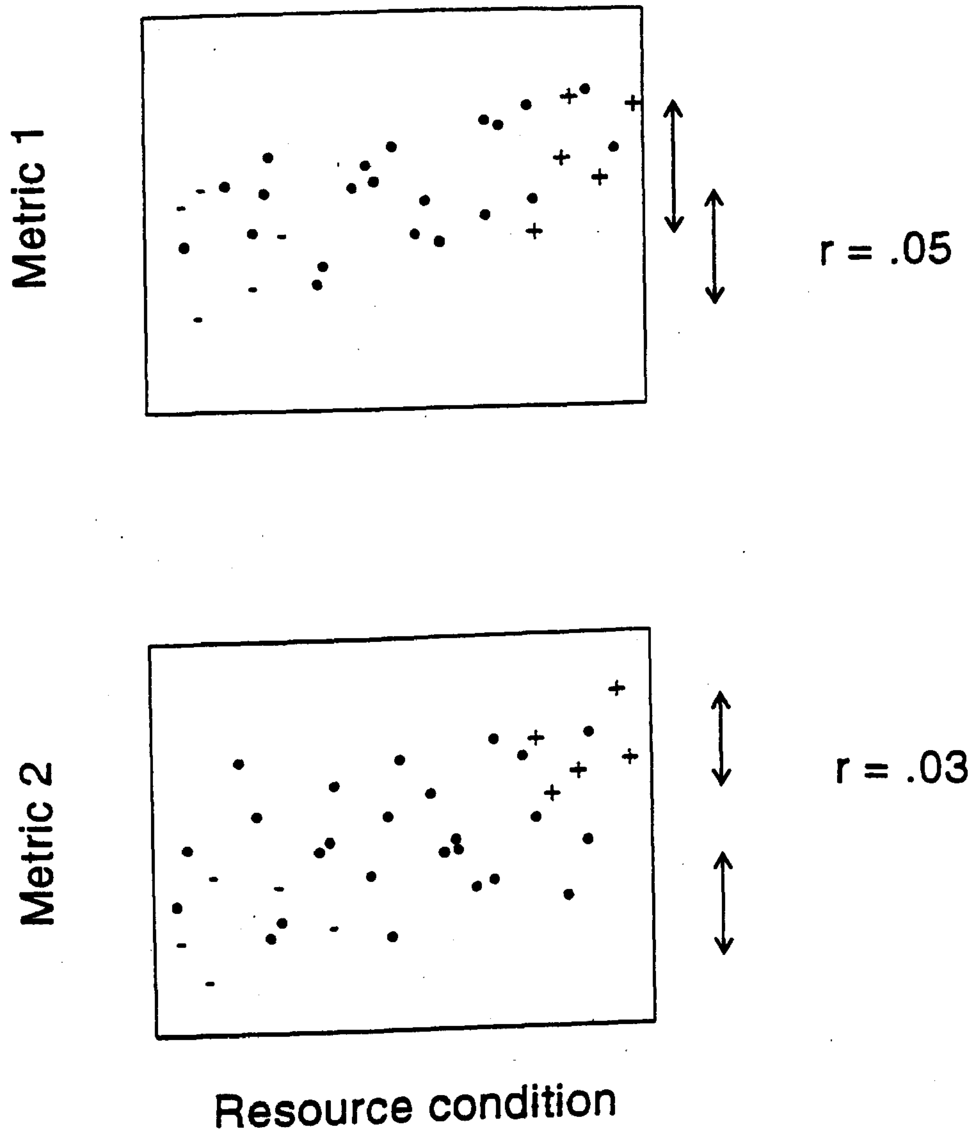


FIGURE 1. Hypothetical relationships between resource condition and candidate biological metrics. Metric 1 is more strongly correlated with resource condition (or r^2 is higher if using regression) than metric 2, suggesting it is a better metric. A more reliable test compares the metric's ability to distinguish (noted by arrows) between pristine ("+") and degraded ("-") sites. Metric 2 is the more effective metric in spite of its poor statistical correlation.

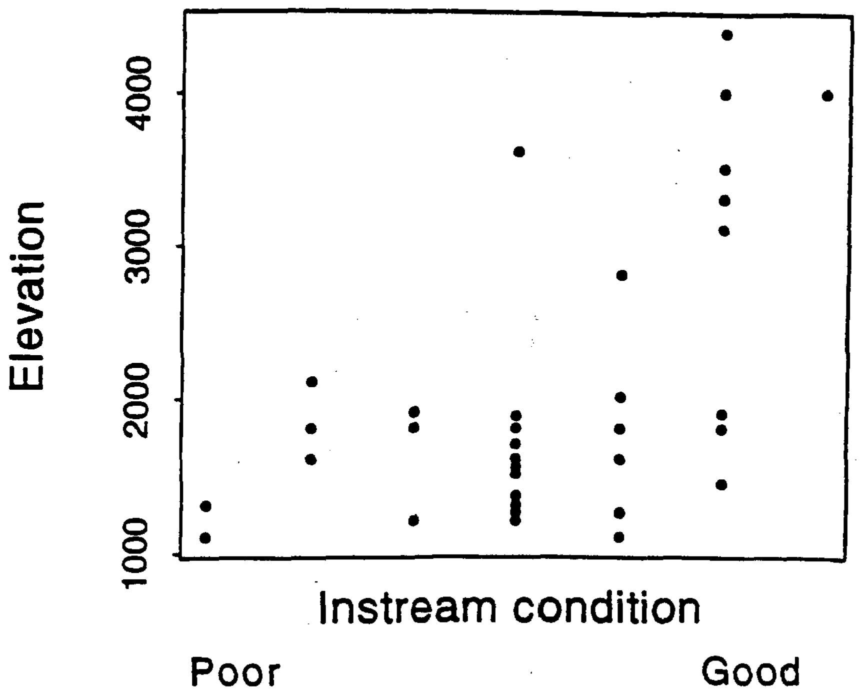


FIGURE 2. Relationship between elevation and instream condition (based on channel morphology, bank erosion and stability, and riparian zone features) in the Umpqua NF (n=41). Sites above 2500 ft. were excluded from analysis because there were no degraded sites at high elevation; thus, elevation was confounded with degradation.

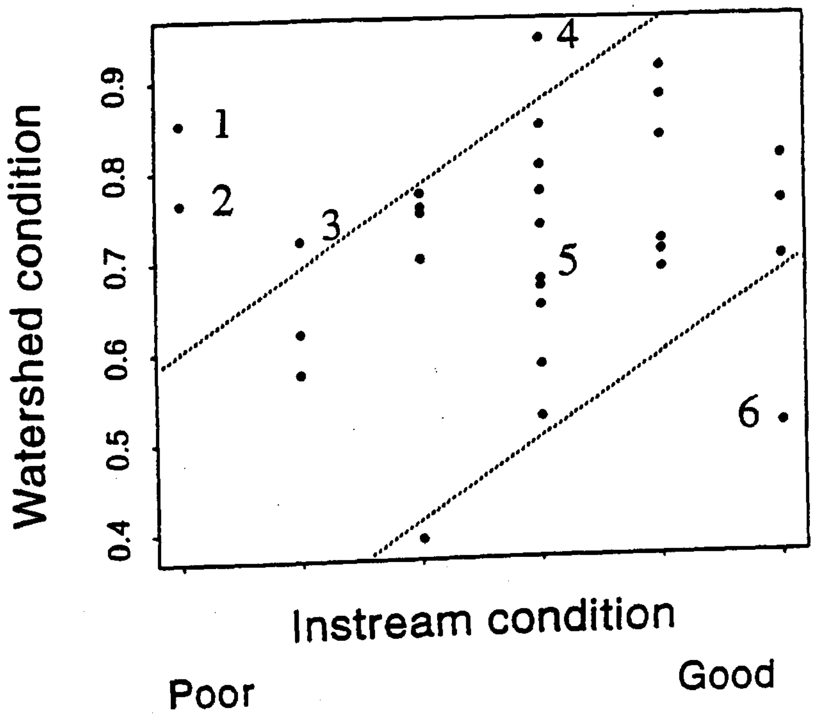


FIGURE 3. Relationship between watershed condition at a site (percent area *not* logged) and instream condition (riparian cover, bank stability, and other physical features). Dashed lines denote direction of expected correlation. When outliers were re-evaluated we found that area logged was underestimated (1, 2, 4); data were tabulated incorrectly (3, 5); and some logging had occurred long ago (6).

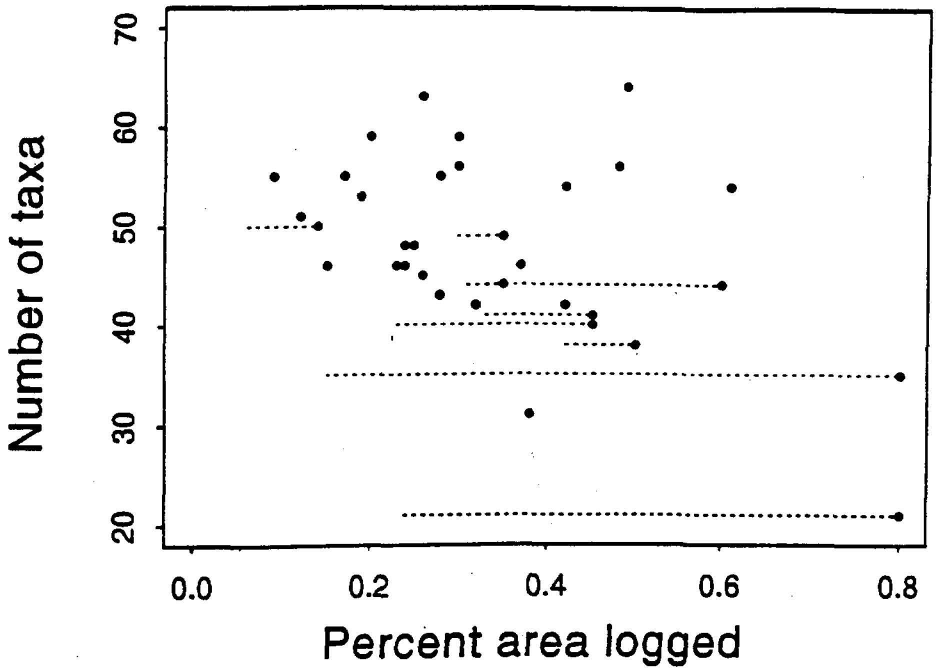


FIGURE 4. Number of taxa decreases as logged area of the watershed increases in Umpqua NF. For some sites, logged area was increased (dotted lines) to reflect private-land logging within the watershed (see text for details). Before data correction, correlation (Spearman's rho) was not significant; after correction, it was significant ($p < 0.05$).

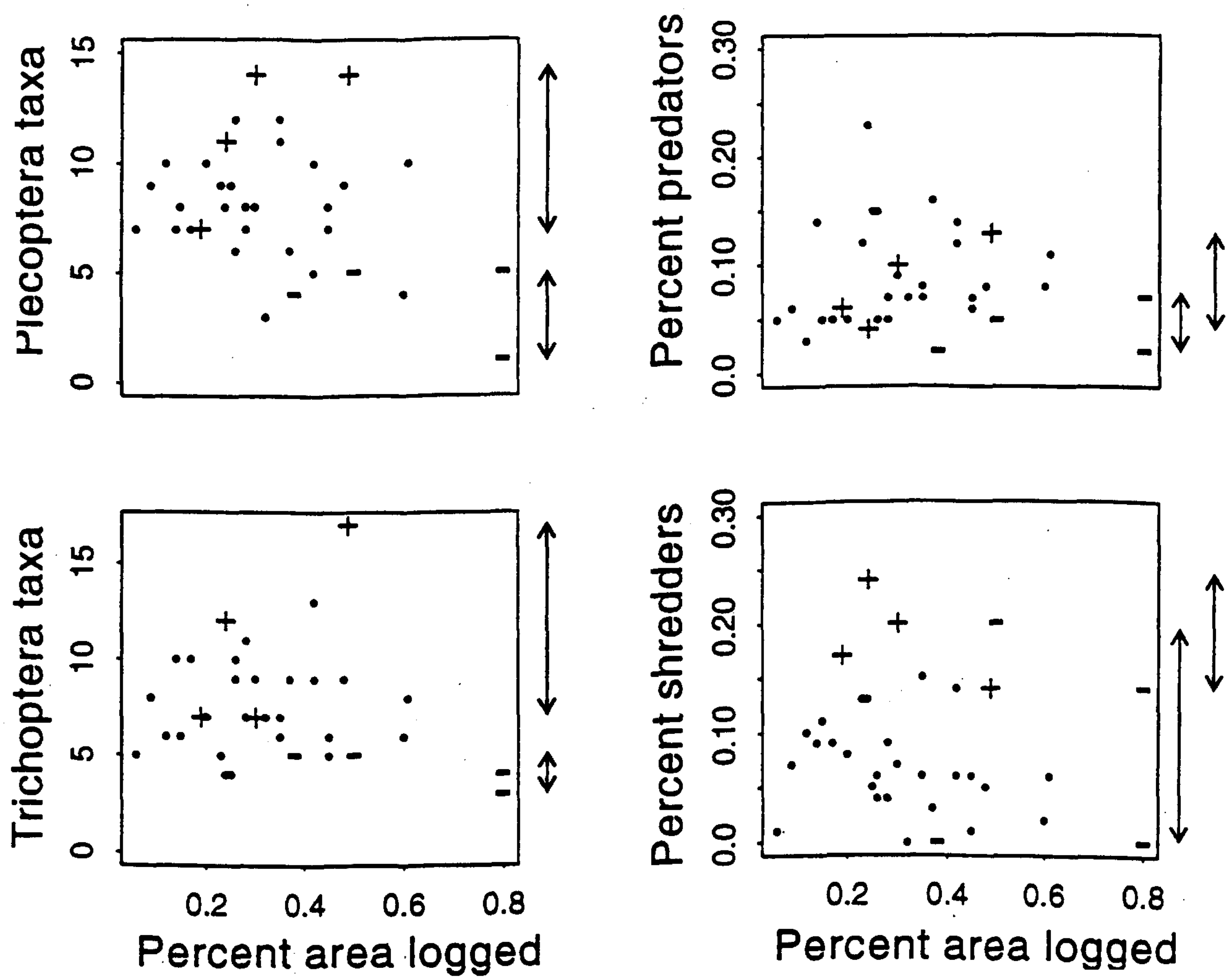


FIGURE 5. Metrics that could clearly distinguish least degraded ("+") from most degraded stream sites ("-") were used to construct B-IBI. Two taxa richness metrics (on left) distinguished the two groups of sites. Trophic structure metrics, such as percent of predators and shredders (on right), failed to distinguish the groups of sites. Arrows to the right of each plot indicate the range of metrics scores for good and poor sites.

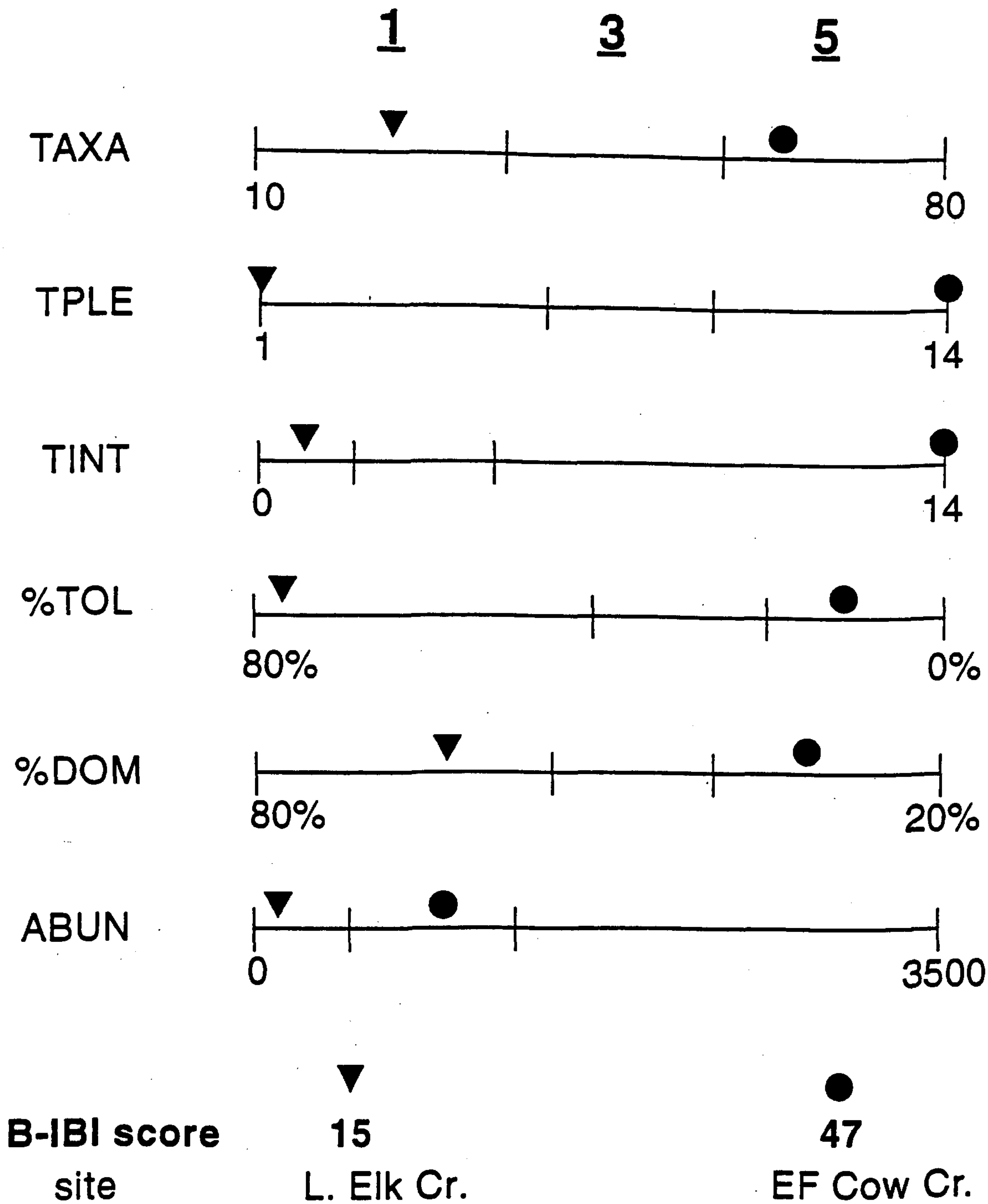


FIGURE 6. Because metrics have different ranges they are scored as 5 (indicating little or no deviation from expected, or reference, condition), 3, or 1 (indicating strong deviation from expected condition) in order to put them on the same scale. B-IBI score equals the sum of eleven metric scores. Human-induced degradation in and around L. Elk Creek was reflected in a low B-IBI score. TAXA=total taxa richness; TPLE=Plecoptera taxa richness; TINT=intolerant taxa richness; %TOL=Percent tolerants; %DOM=Percent dominance; ABUN=total abundance.

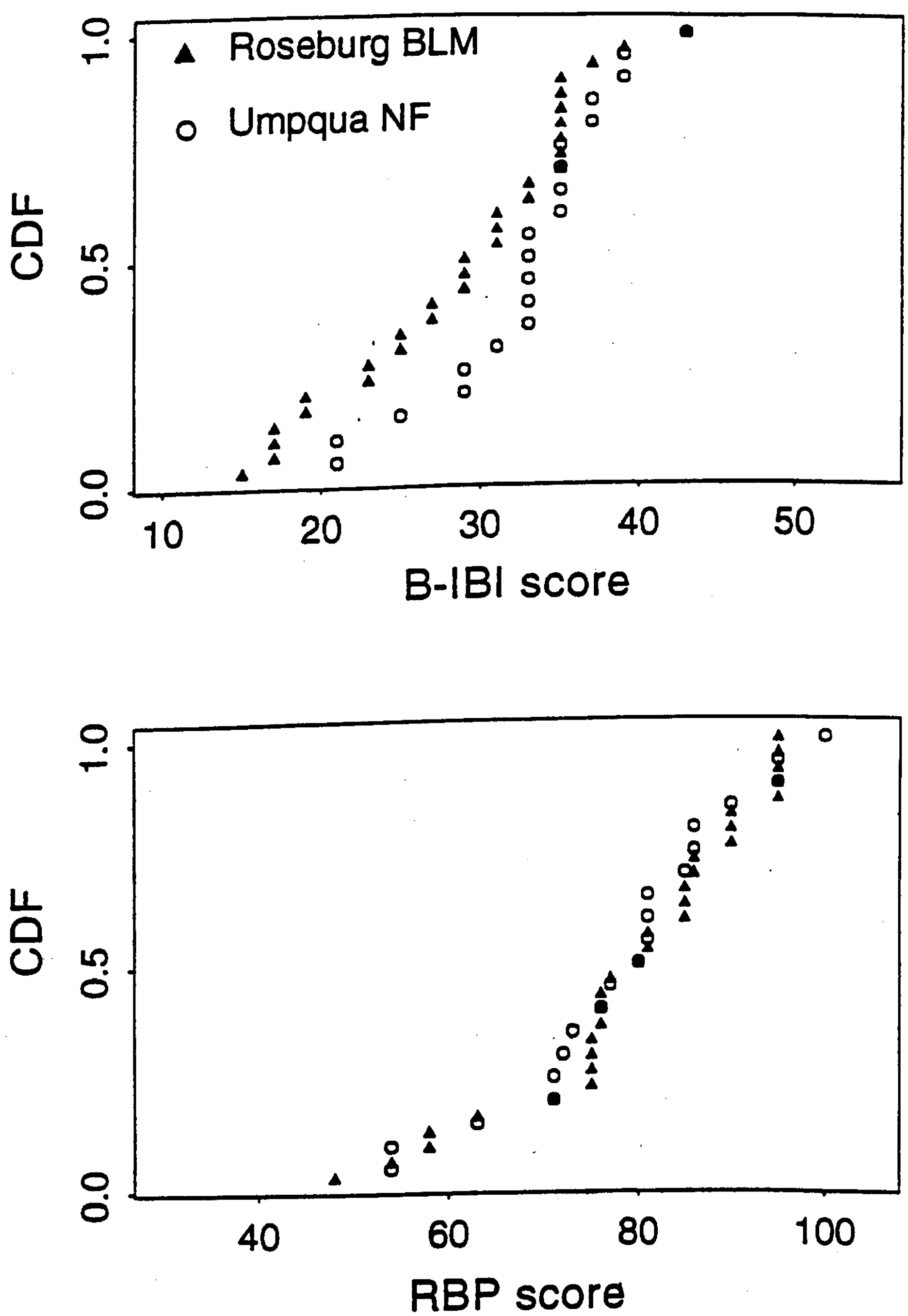


FIGURE 7. The cumulative distribution function (CDF) of B-IBI scores and RBP scores for sites in the Roseburg district (n = 30) and the Umpqua district (n = 20). The CDF is the percentage of sites with B-IBI scores lower than the B-IBI score on the x-axis. Distance between the two CDFs indicates the degree of difference between scores for the two areas. Lower B-IBI scores (one-sided t-test; $p < 0.05$) in the Roseburg reflected higher human use. The RBP index failed to detect a difference between sites in the Roseburg and Umpqua.

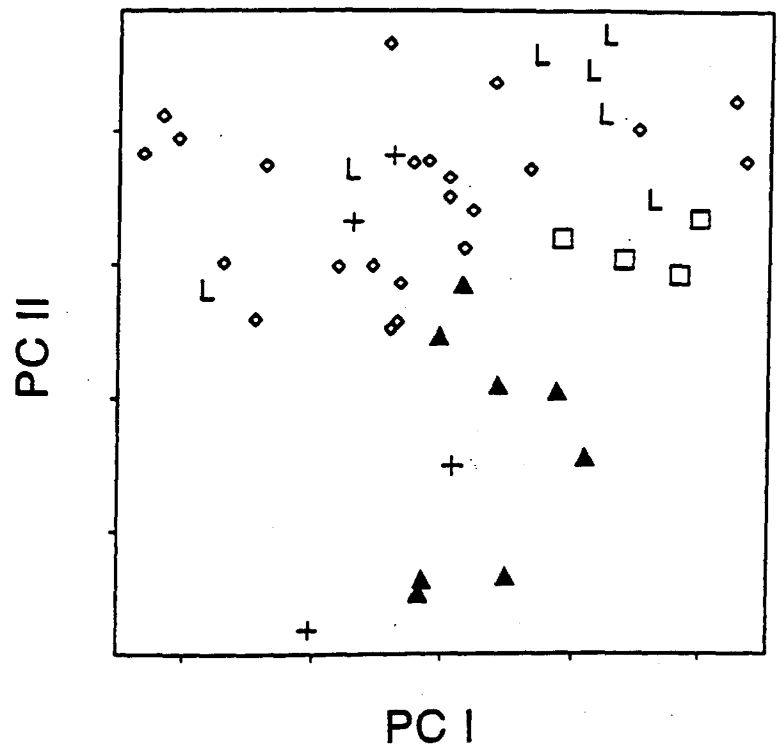


FIGURE 8. Sites from Umpqua NF sampled in 1990 plotted according to a principal components analysis of the "normalized" data. Most degraded (□), high elevation (▲), and large river (L) sites showed a tendency to group together. Least-degraded sites (+) and other remaining sites (◇) did not clump together for these or other component plots. PCA of log transformed and untransformed data gave similar plots.

DEVELOPING BIOASSESSMENT PROTOCOLS FOR MONTANA WETLANDS

R. Apfelbeck [1]; *Bahls, L.* [1]; *Shapley, M.* [2]; *Gerritsen, J.* [3]; *Barbour, M.* [3]; *Stribling, J.* [3]; *Charles, D.* [4]; *Acker, F.* [4]
Presented by *Randy Apfelbeck*
Montana Department of Environmental Quality
Cogswell Building, 1400 Broadway
Helena, MT 59601

NOTES:

Eighty wetlands were sampled throughout Montana from April through September during 1993 and 1994 for developing wetland bioassessment protocols using macroinvertebrate and periphyton. Wetlands were sampled for macroinvertebrate, periphyton, water column chemistry and sediment trace metal chemistry. Hydrologic, geologic and climatic data were collected from maps and existing databases. Additional hydrologic field measurements were taken to assist in developing a better understanding of the water quality for sites that were highly evaporative or predominantly groundwater supported.

A wetland classification system is currently being developed that will group reference wetlands with similar habitats and water quality in order to improve detection of anthropogenic impairments. The classification system incorporates ecoregions and geomorphic processes. Presently 10 wetland classes have been delineated.

Macroinvertebrate metrics identified as being useful in developing an index to assess wetland water quality include: 1) # taxa; 2) % dominant taxa (1, 2, 5); 3) # of Plecoptera, Odonata, Ephemeroptera and Trichoptera taxa; 4) # of individuals; 5) % and # of Chironomidae and % Orthocladiinae/total Chironomidae; 6) % and # of Crustacea and Mollusca; 7) # of Hirudinea, Spongillidae and Sphaeriidae taxa; and 8) Hilsenhoff Biotic Index. Preliminary results indicate detection of impairment that appear to be caused by heavy metals, nutrients, salinity, sediment and fluctuating water levels. Scuds (*Hyallela azteca*) were identified as being a potential indicator of ephemeral wetlands (absence or very low numbers), saline wetlands (absence) and alkali wetlands (very abundant). Periphyton bioassessment protocols are currently being developed by The Academy of Natural Sciences of Philadelphia.

[1]Montana Department of Environmental Quality
[2]Montana Natural Heritage Program
[3]Tetra Tech, Inc., Owings Mills, MD
[4]Academy of Natural Sciences, Philadelphia, PA

PROJECT OBJECTIVES

1. Develop Rapid Bioassessment Protocols
2. Determine Natural Spatial Variability
3. Develop Classification System
4. Detect Wetland Water Quality Impairment
5. Simple And Cost Effective Approach
6. Collect Quality Baseline Data

SAMPLING DESIGN OBJECTIVES

To Sample:

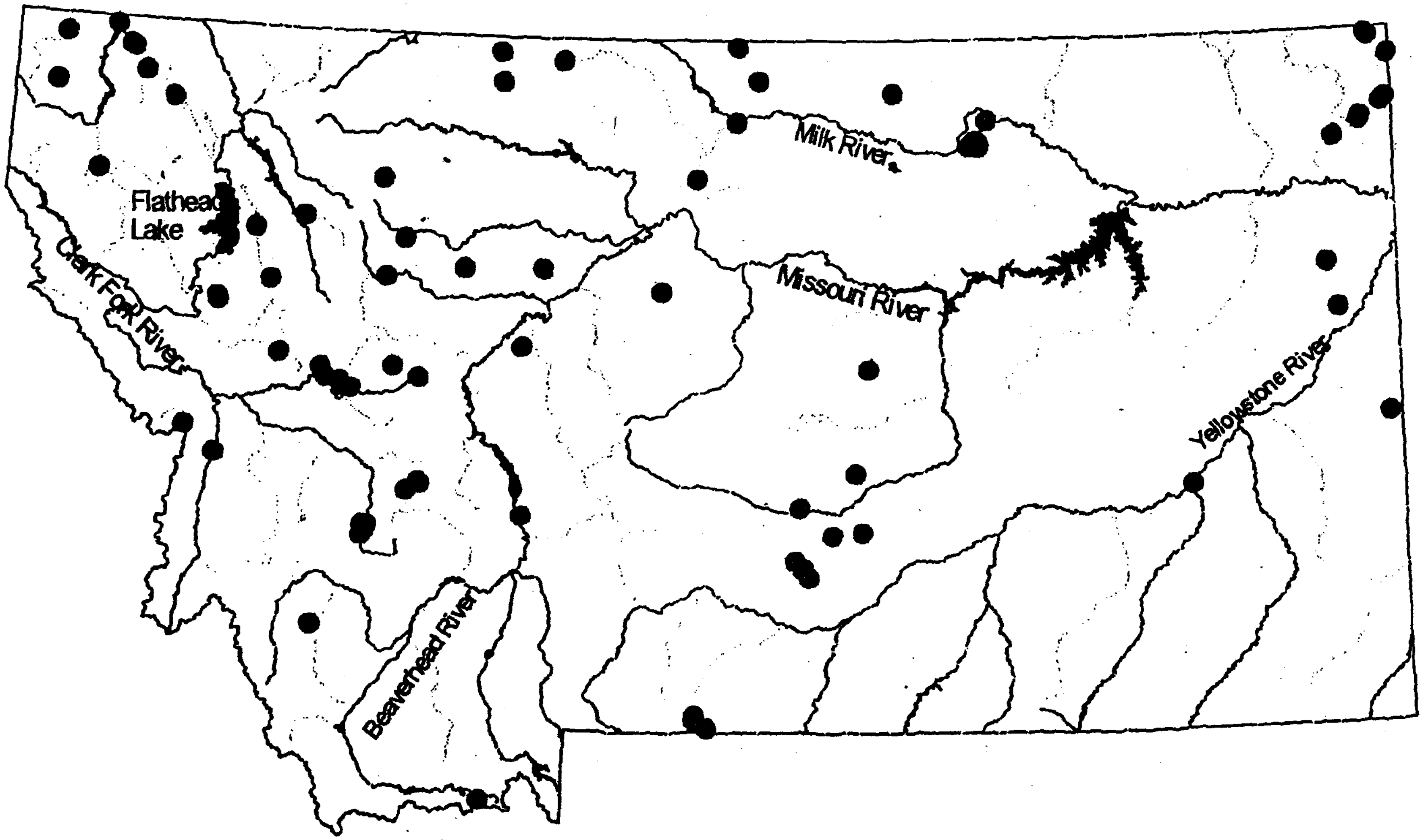
1. 80 Wetlands Throughout Montana
2. Wetlands From Each Major Ecoregion
3. Least Impaired (Reference) Wetlands
4. Impaired Wetlands
5. The Full Spectrum Of Wetland Types

SELECTION CRITERIA

1. Historical Data
2. Special Interest
3. Least-Impaired Wetlands
4. Wetlands With Known Impairments
5. Cooperation By Landowners
6. Access
7. Paired Wetlands

OWNERSHIP OF WETLANDS EVALUATED

1. U. S. Fish & Wildlife Service
2. U. S. Forest Service
3. The Nature Conservacy
4. State Lands
5. Private Ownership
6. Department Of Transportation



Montana DEQ Wetland Characterization Sites

SAMPLING METHODS

1. Location
2. Sampling Period
3. Macroinvertebrates
4. Periphyton
5. Water Column
6. Field Analysis
7. Sediment
8. Photo Documentation
9. Field Observations

WETLAND CLASSIFICATION

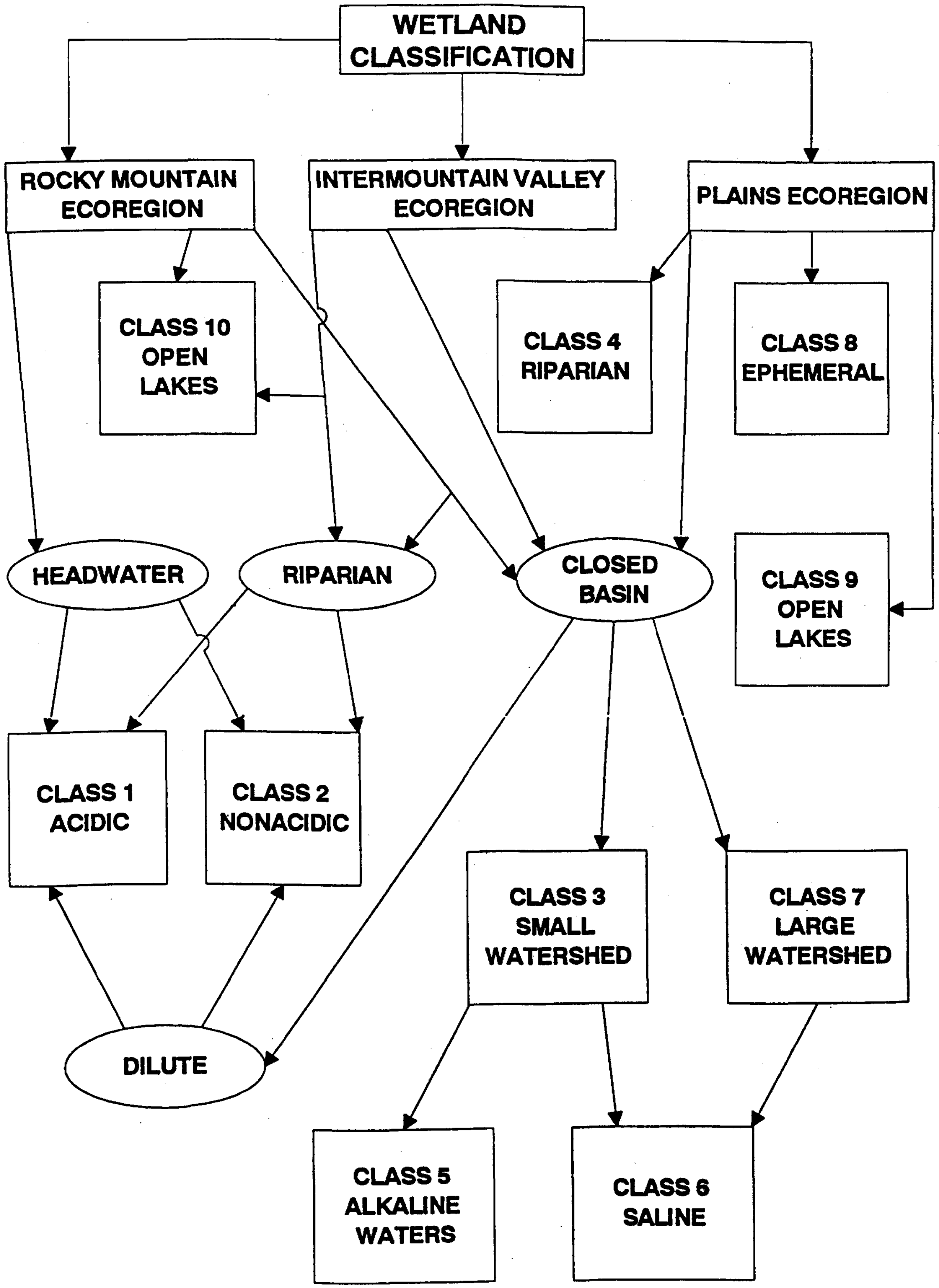
1. Ecoregion
   - A) Rocky Mountain Ecoregion
   - B) Intermountain Valley Ecoregion
   - C) Plains Ecoregion
2. Geomorphic
   - A) Open Lakes
   - B) Closed Basin
   - C) Riparian
   - D) Ephemeral
3. Chemical Delineation
   - A) Conductivity
   - B) pH
   - C) Alkalinity

CLASSIFICATION INDICATORS

1. Vegetation
2. Alkalinity
3. Scuds
4. Wildlife
5. Evaporative Salts
6. Watershed / Wetland Area
7. Topographic Position

MACROINVERTEBRATE INDEX DEVELOPMENT

1. Identify And Count Taxa
2. Develop And Test Potential Metrics
3. Select Core Metrics
4. Metric Scoring
5. Aggregation Of Metric Scores
6. Index Scores
7. Assessment Of Sites


WETLAND CLASSIFICATION
ROCKY MOUNTAIN ECOREGION
INTERMOUNTAIN VALLEY ECOREGION
PLAINS ECOREGION
CLASS 10 OPEN LAKES
CLASS 4 RIPARIAN
CLASS 8 EPHEMERAL
HEADWATER
RIPARIAN
CLOSED BASIN
CLASS 9 OPEN LAKES
CLASS 1 ACIDIC
CLASS 2 NONACIDIC
CLASS 3 SMALL WATERSHED
CLASS 7 LARGE WATERSHED
DILUTE
CLASS 5 ALKALINE WATERS
CLASS 6 SALINE

METRICS SELECTED		STRESSOR
#	Taxa	Decrease
%	Dominance Taxa (1,2,5)	Increase
#	Plec / Odon / Eph / Tric Taxa	Decrease
#	Individuals	Decrease
%	Chironomidae	Increase
#	Chironomidae	Decrease
%	Orthocladiinae / Total Chir	Decrease
#	Crustacea & Mollusca Taxa	Decrease
%	Crustacea / Mollusca Taxa	Increase
#	Leech, Sponge and Clam Taxa	Decrease
	Hilsenhoff Biotic Index	Increase

INDEX SCORE

METRICS	RANGE
Hilsenhoff Biotic Index	1-14
Leech / Sponge / Clam	0-18
Crustacea / Mollusca	0-22
Chironomidae (midge)	0-22
# Individuals	0-16
Plec / Odon / Eph / Tric	0-24
Inverse % Dominance	0-25
# Taxa	1-28
Total Index Score	7-119

ASSESSMENT

For Each Wetland Class:

1. Evaluate Reference Sites
2. Evaluate Sites With Know Impairment
3. Use Physicochemical Data To Assist In Evaluation
4. Use Field Notes And Historical Information
5. Evaluate Variability
6. Evaluate Outlier Sites
7. Determine Condition

FUTURE OBJECTIVES

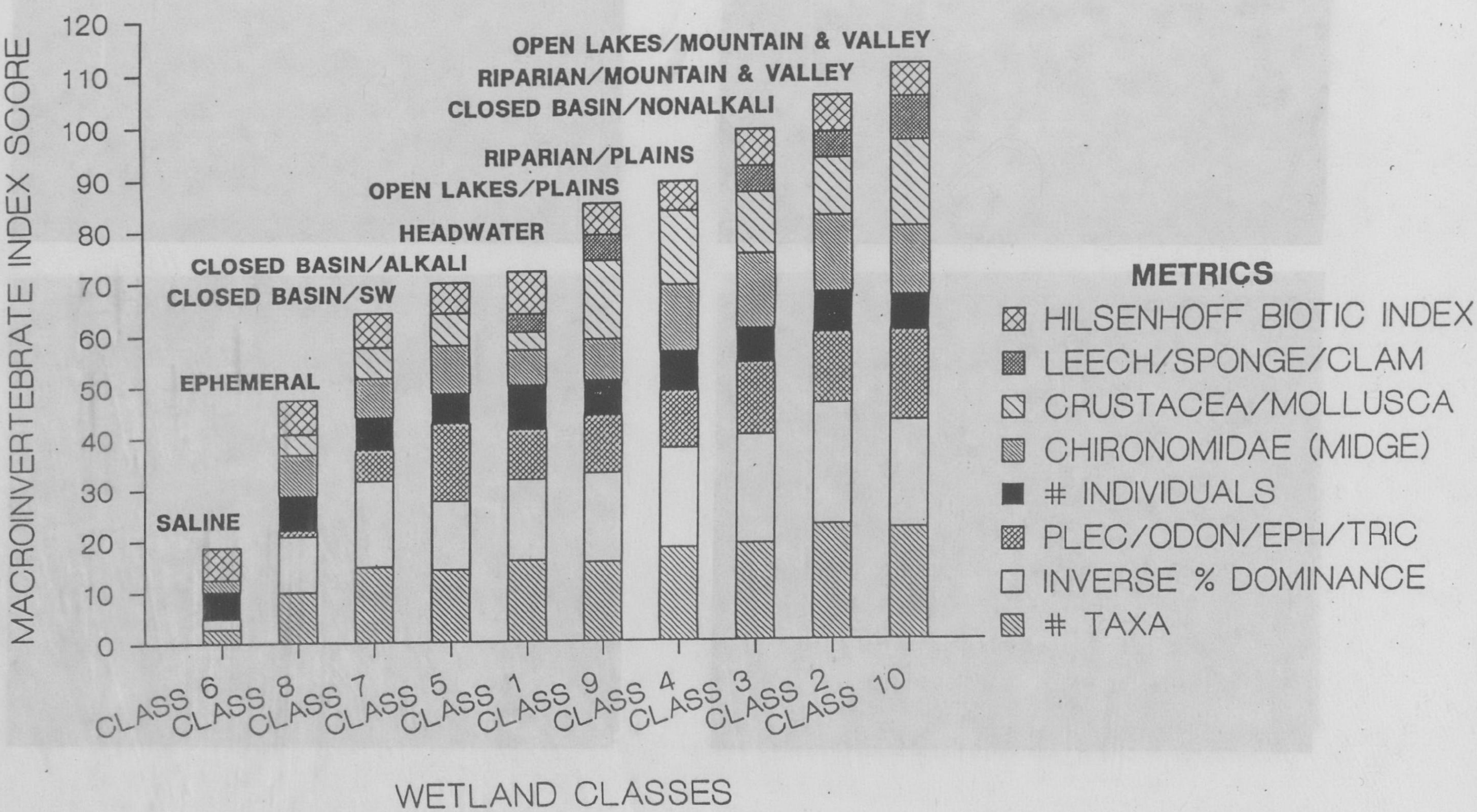
1. Refine The Classification System
2. Determine Temporal Variability Through Replication
3. Evaluate Seasonality And Determine Index Period
4. Refine The Habitat Assessment Approach To Reduce Subjectivity
5. Evaluate Using The Wetland Bioassessment Approach For Low Gradient Prairie Streams

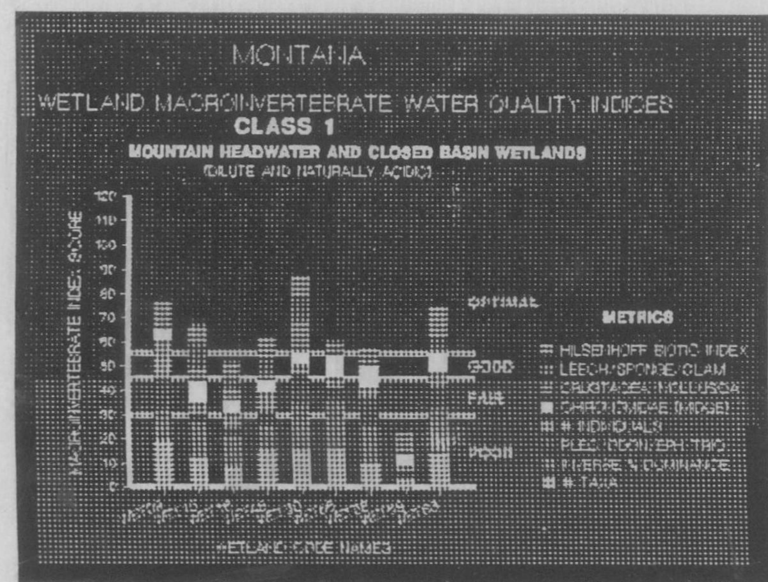
MONTANA

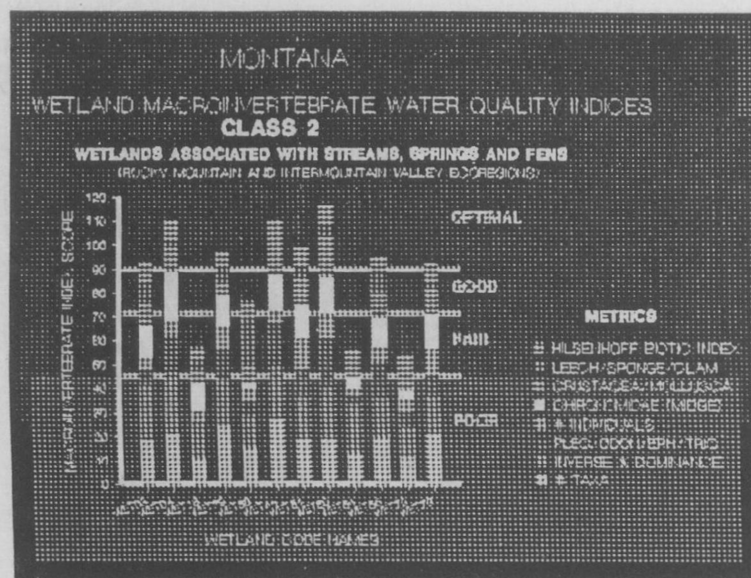
WETLAND MACROINVERTEBRATE WATER QUALITY INDICES

REFERENCE WETLANDS

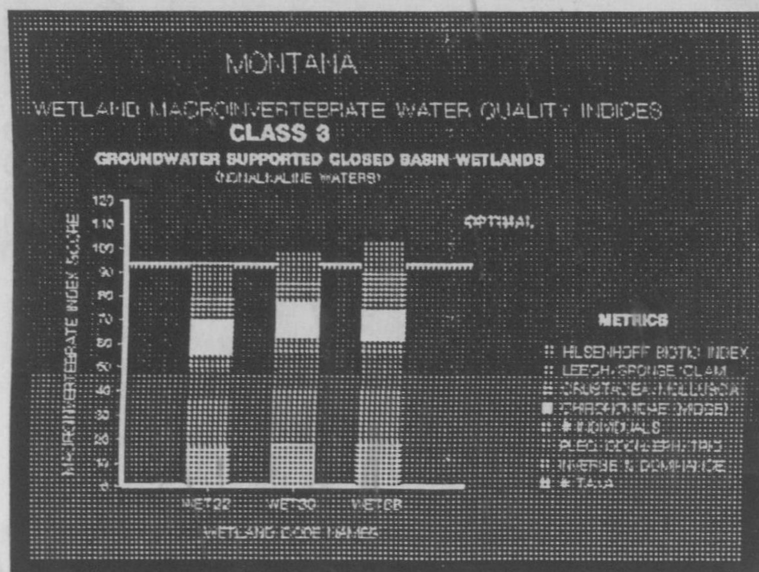
AVERAGE INDEX SCORES

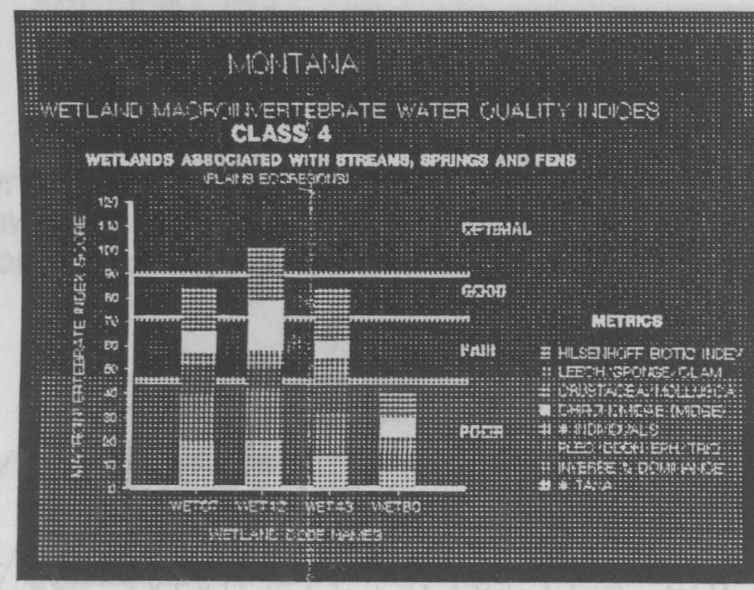


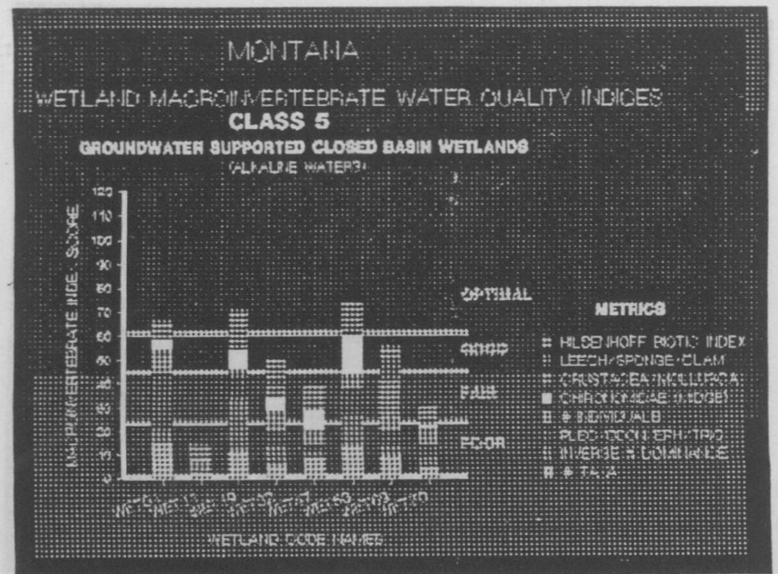

MONTANA
WETLAND MACROINVERTEBRATE WATER QUALITY INDICES
CLASS 1
MOUNTAIN HEADWATER AND CLOSED BASIN WETLANDS
(DILUTE AND NATURALLY ACIDIC)
MACROINVERTEBRATE INDEX SCORE
OPTIMAL
GOOD
FAIR
POOR
METRICS
HILSENHOFF BIOTIC INDEX
LEECH/SPONGE/CLAM
CRUSTACEA/MOLLUSCA
CHIRONOMIDAE (MIDGE)
INDIVIDUALS
PLEC/ODON/EPH/TRIC
INVERSE % DOMINANCE
TAXA
WETLAND CODE NAMES

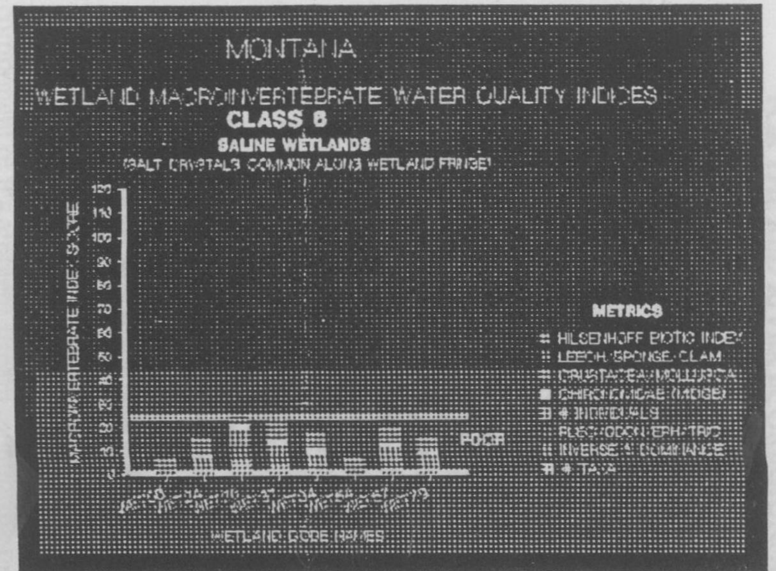

MONTANA
WETLAND MACROINVERTEBRATE WATER QUALITY INDICES
CLASS 2
WETLANDS ASSOCIATED WITH STREAMS, SPRINGS AND FENS
(ROCKY MOUNTAIN AND INTERMOUNTAIN VALLEY ECOREGIONS)
MACROINVERTEBRATE INDEX SCORE
OPTIMAL
GOOD
FAIR
POOR
METRICS
HILSENHOFF BIOTIC INDEX
LEECH/SPONGE/CLAM
CRUSTACEA/MOLLUSCA
CHIRONOMIDAE (MIDGE)
INDIVIDUALS
PLEC/ODON/EPH/TRIC
INVERSE % DOMINANCE
TAXA
WETLAND CODE NAMES

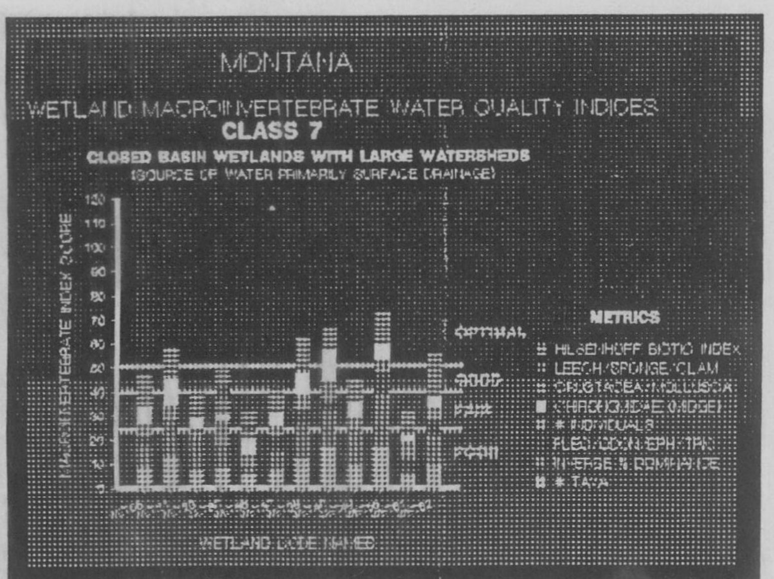
78

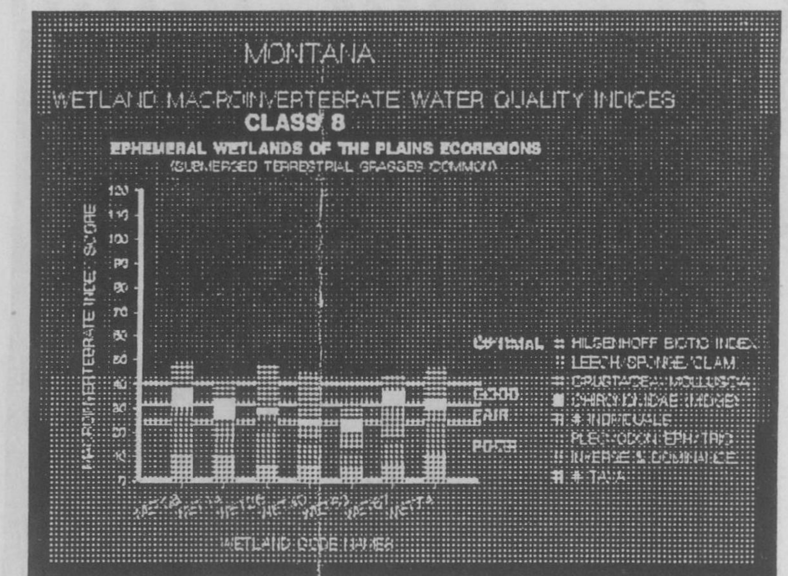

MONTANA
WETLAND MACROINVERTEBRATE WATER QUALITY INDICES
CLASS 3
GROUNDWATER SUPPORTED CLOSED BASIN WETLANDS
(NONALKALINE WATERS)
MACROINVERTEBRATE INDEX SCORE
OPTIMAL
METRICS
HILSENHOFF BIOTIC INDEX
LEECH/SPONGE/CLAM
CRUSTACEA/MOLLUSCA
CHIRONOMIDAE (MIDGE)
INDIVIDUALS
PLEC/ODON/EPH/TRIC
INVERSE % DOMINANCE
TAXA
WET22
WET30
WET68
WETLAND CODE NAMES

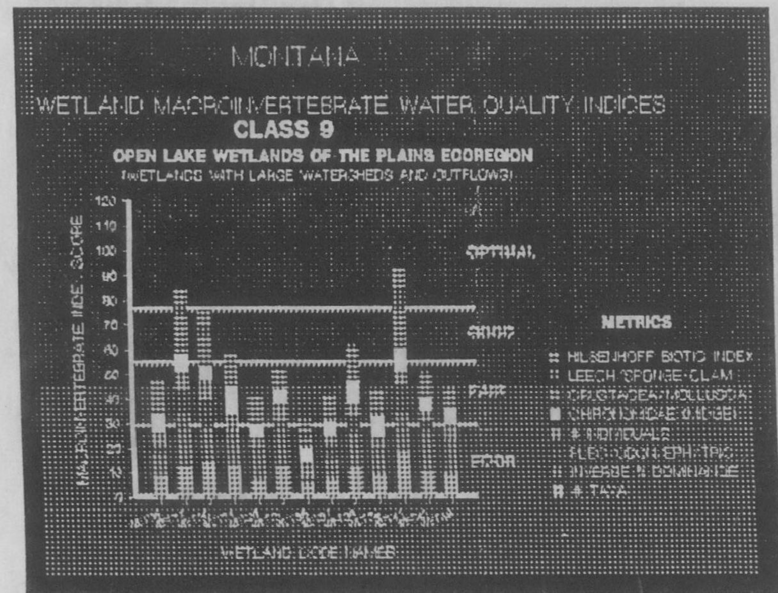

MONTANA
WETLAND MACROINVERTEBRATE WATER QUALITY INDICES
CLASS 4
WETLANDS ASSOCIATED WITH STREAMS, SPRINGS AND FENS
(PLAINS ECOREGIONS)
MACROINVERTEBRATE INDEX SCORE
OPTIMAL
GOOD
FAIR
POOR
METRICS
HILSENHOFF BIOTIC INDEX
LEECH/SPONGE/CLAM
CRUSTACEA/MOLLUSCA
CHIRONOMIDAE (MIDGE)
INDIVIDUALS
PLEC/ODON/EPH/TRIC
INVERSE % DOMINANCE
TAXA
WET42
WET43
WET80
WETLAND CODE NAMES

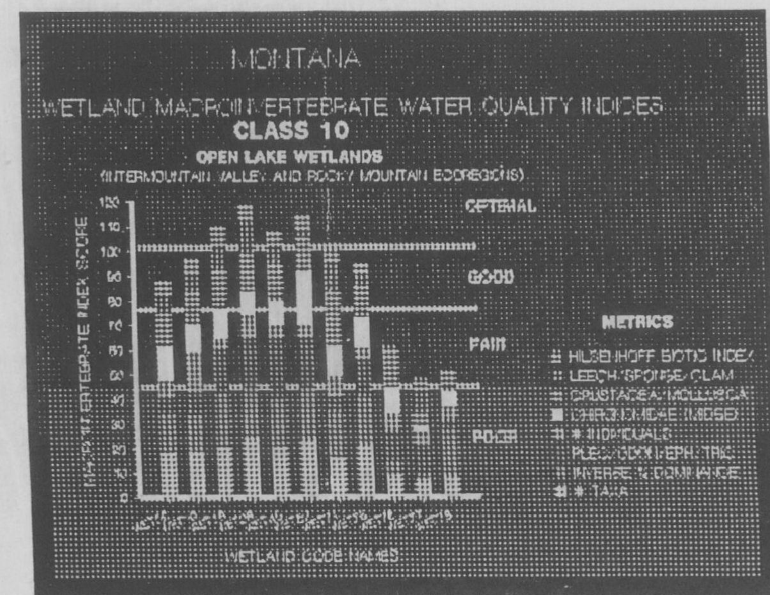

MONTANA
WETLAND MACROINVERTEBRATE WATER QUALITY INDICES
CLASS 5
GROUNDWATER SUPPORTED CLOSED BASIN WETLANDS
(ALKALINE WATERS)
MACROINVERTEBRATE INDEX SCORE
OPTIMAL
GOOD
FAIR
POOR
METRICS
HILSENHOFF BIOTIC INDEX
LEECH/SPONGE/CLAM
CRUSTACEA/MOLLUSCA
CHIRONOMIDAE (MIDGE)
INDIVIDUALS
PLEC/ODON/EPH/TRIC
INVERSE % DOMINANCE
TAXA
WETLAND CODE NAMES

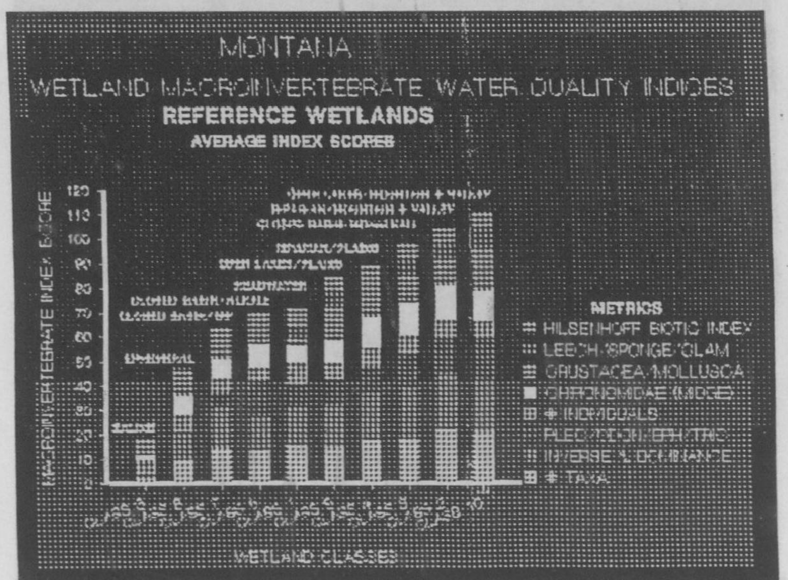

MONTANA
WETLAND MACROINVERTEBRATE WATER QUALITY INDICES
CLASS 6
SALINE WETLANDS
(SALT CRYSTALS COMMON ALONG WETLAND FRINGE)
MACROINVERTEBRATE INDEX SCORE
POOR
METRICS
HILSENHOFF BIOTIC INDEX
LEECH/SPONGE/CLAM
CRUSTACEA/MOLLUSCA
CHIRONOMIDAE (MIDGE)
INDIVIDUALS
PLEC/ODON/EPH/TRIC
INVERSE % DOMINANCE
TAXA
WETLAND CODE NAMES


MONTANA
WETLAND MACROINVERTEBRATE WATER QUALITY INDICES
CLASS 7
CLOSED BASIN WETLANDS WITH LARGE WATERSHEDS
(SOURCE OF WATER PRIMARILY SURFACE DRAINAGE)
MACROINVERTEBRATE INDEX SCORE
OPTIMAL
GOOD
FAIR
POOR
METRICS
HILSENHOFF BIOTIC INDEX
LEECH/SPONGE/CLAM
CRUSTACEA/MOLLUSCA
CHIRONOMIDAE (MIDGE)
INDIVIDUALS
PLEC/ODON/EPH/TRIC
INVERSE % DOMINANCE
TAXA
WETLAND CODE NAMES


MONTANA
WETLAND MACROINVERTEBRATE WATER QUALITY INDICES
CLASS 8
EPHEMERAL WETLANDS OF THE PLAINS ECOREGIONS
(SUBMERGED TERRESTRIAL GRASSES COMMON)
MACROINVERTEBRATE INDEX SCORE
OPTIMAL
GOOD
FAIR
POOR
HILSENHOFF BIOTIC INDEX
LEECH/SPONGE/CLAM
CRUSTACEA/MOLLUSCA
CHIRONOMIDAE (MIDGE)
INDIVIDUALS
PLEC/ODON/EPH/TRIC
INVERSE % DOMINANCE
TAXA
WETLAND CODE NAMES


MONTANA
WETLAND MACROINVERTEBRATE WATER QUALITY INDICES
CLASS 9
OPEN LAKE WETLANDS OF THE PLAINS ECOREGION
(WETLANDS WITH LARGE WATERSHEDS AND OUTFLOWS)
MACROINVERTEBRATE INDEX SCORE
OPTIMAL
GOOD
FAIR
POOR
METRICS
HILSENHOFF BIOTIC INDEX
LEECH/SPONGE/CLAM
CRUSTACEA/MOLLUSCA
CHIRONOMIDAE (MIDGE)
INDIVIDUALS
PLECO/ODON/EPH/TRIC
INVERSE % DOMINANCE
TAXA
WETLAND CODE NAMES


MONTANA
WETLAND MACROINVERTEBRATE WATER QUALITY INDICES
CLASS 10
OPEN LAKE WETLANDS
(INTERMOUNTAIN VALLEY AND ROCKY MOUNTAIN ECOREGIONS)
MACROINVERTEBRATE INDEX SCORE
OPTIMAL
GOOD
FAIR
POOR
METRICS
HILSENHOFF BIOTIC INDEX
LEECH/SPONGE/CLAM
CRUSTACEA/MOLLUSCA
CHIRONOMIDAE (MIDGE)
INDIVIDUALS
PLECO/ODON/EPH/TRIC
INVERSE % DOMINANCE
TAXA
WETLAND CODE NAMES


MONTANA
WETLAND MACROINVERTEBRATE WATER QUALITY INDICES
REFERENCE WETLANDS
AVERAGE INDEX SCORES
MACROINVERTEBRATE INDEX SCORE
OPEN LAKES/PLAINS
HEADWATER
METRICS
HILSENHOFF BIOTIC INDEX
LEECH/SPONGE/CLAM
CRUSTACEA/MOLLUSCA
CHIRONOMIDAE (MIDGE)
INDIVIDUALS
PLECO/ODON/EPH/TRIC
INVERSE % DOMINANCE
TAXA
WETLAND CLASSES

AN INDEX OF BIOLOGICAL INTEGRITY FOR THE RED RIVER ECOREGION: A BIOLOGICAL "DESERT"

Eric Pearson
North Dakota Department of Health
1200 Missouri Avenue
Bismarck, ND 58502

NOTES:

The Index of Biotic Integrity (IBI) is a numerical expression of the fish community structure. This study, initiated with a grant from the U.S. Environmental Protection Agency, involves a number of agencies including the State Health Department, Division of Water Quality, the North Dakota Game and Fish Department, Minnesota Pollution Control Agency, Minnesota Department of Natural Resources, the U.S. Geological Survey, NAWQA program and Regions 5 and 8 of the U.S. EPA.

We currently are proposing 17 metrics which are measures of species richness and composition, trophic composition, and fish abundance and condition. These IBI scores will allow the Health Department to assess the aquatic life impacts along the Red River and at various sites throughout the basin. These impacts may be anything from habitat degradation (e.g., isolation or snag removal) to toxic impacts such as sewage outfalls or chemical spills. The Health Department hopes to expand this program to other basins as an ongoing component of it's ambient stream monitoring program.

UPDATE ON THE USE OF MACROINVERTEBRATES IN MONTANA STREAM BIOASSESSMENTS

Bob Bukantis
Montana Department of Environmental Quality
Water Quality Division
A206 Cogswell Building
Helena, MT 59620-0901

NOTES:

We have been using EPA's Rapid Bioassessment Protocol III modified as follows: sampling is done with a D-net (1 mm mesh) employing a 1 minute traveling kick method (2 minutes in mountain streams). Samples are taken from riffles during the summer and subsampled to 300 +/- 30 organisms in the lab. Level of taxonomic determination is specified, usually to genus or species.

Analysis of data so far suggests that Montana can be lumped into Mountain, Foothill, Spring Creek, and Plains Bioregions. Work done on methods development/testing suggests that the sampling method produces less variable data than Surber samplers, and shows good repeatability among nearby riffles sampled. There is good repeatability between years in Mountain streams, but not for Plains streams.

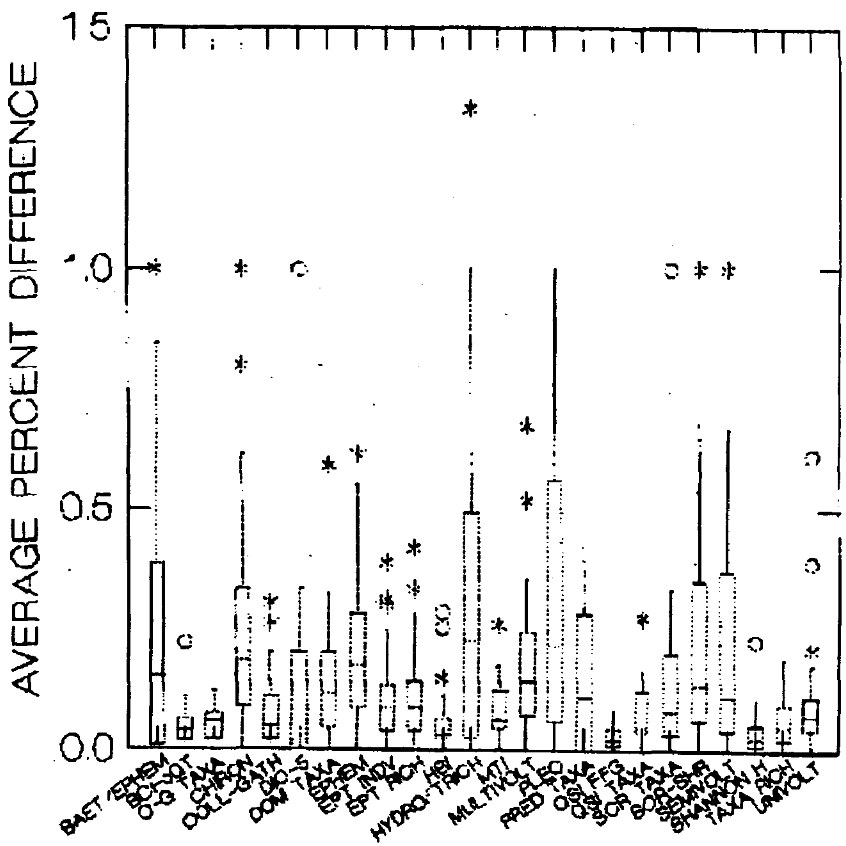
Investigator bias was not evident with 2 people taking samples on a foothills stream. However, there appeared to be a significant effect of investigator bias on a large Spring Creek.

Provisional, regional reference communities are currently being used in evaluation of macroinvertebrate samples taken for stream assessments. These reference communities have been modeled by describing community characteristics for non-impaired, slight, moderate, and severe impairment categories with expected ranges of metric values. Data from reference streams were used to set limits of impairment; a subjective evaluation of existing data was used to set levels of impairment.

Specific reference communities are developed on long-term projects by compositing the best attributes (metric values) available from the pool of samples taken from the project system. Use of these so-called "Internal Reference Communities" increases the ability to discriminate impairment between sites.

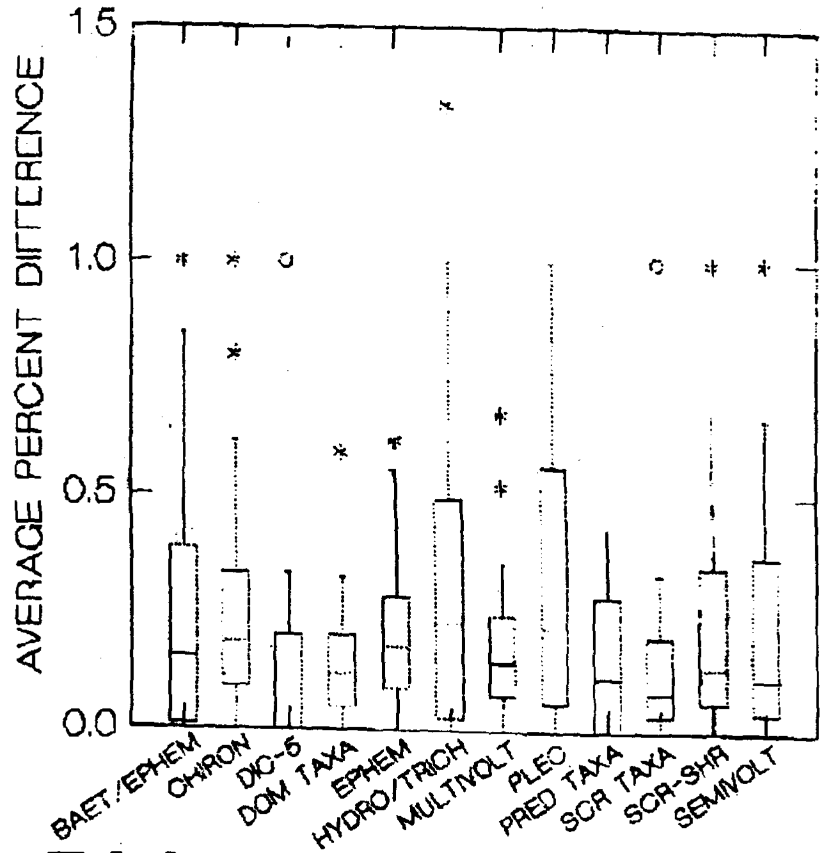
Analysis of sampling variation shows promise for use in selection of appropriate batteries of metrics for each Bioregion, and for refinement of setting impairment limits.

SAMPLE VARIATION IN MACROINVERTEBRATES



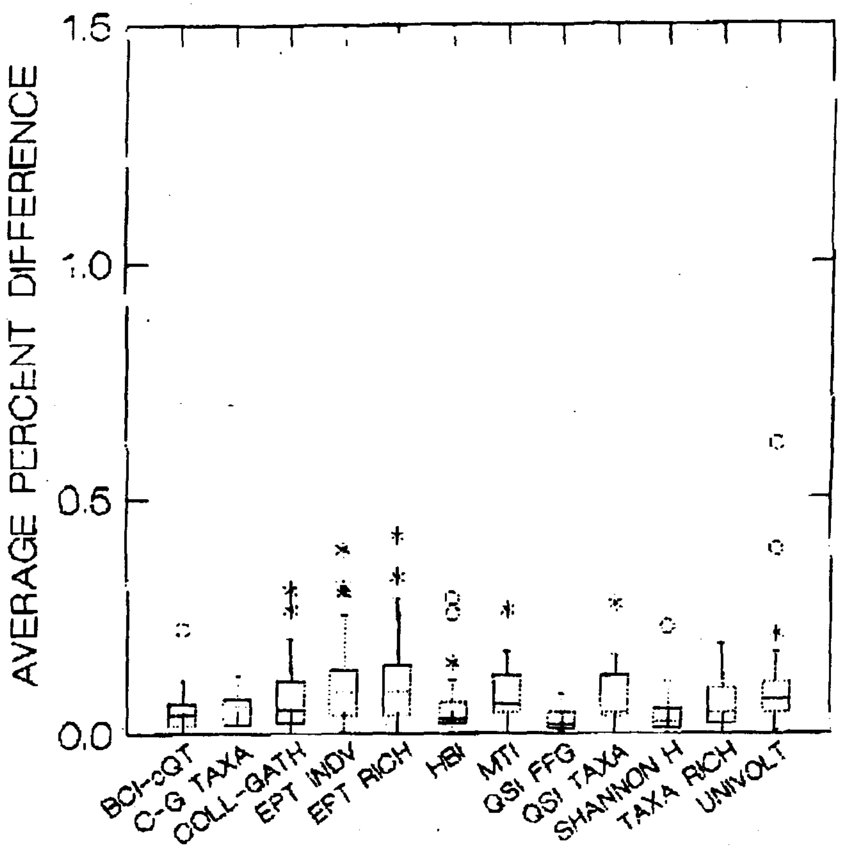
FOOTHILLS ECOREGION

SAMPLE VARIATION IN MACROINVERTEBRATES



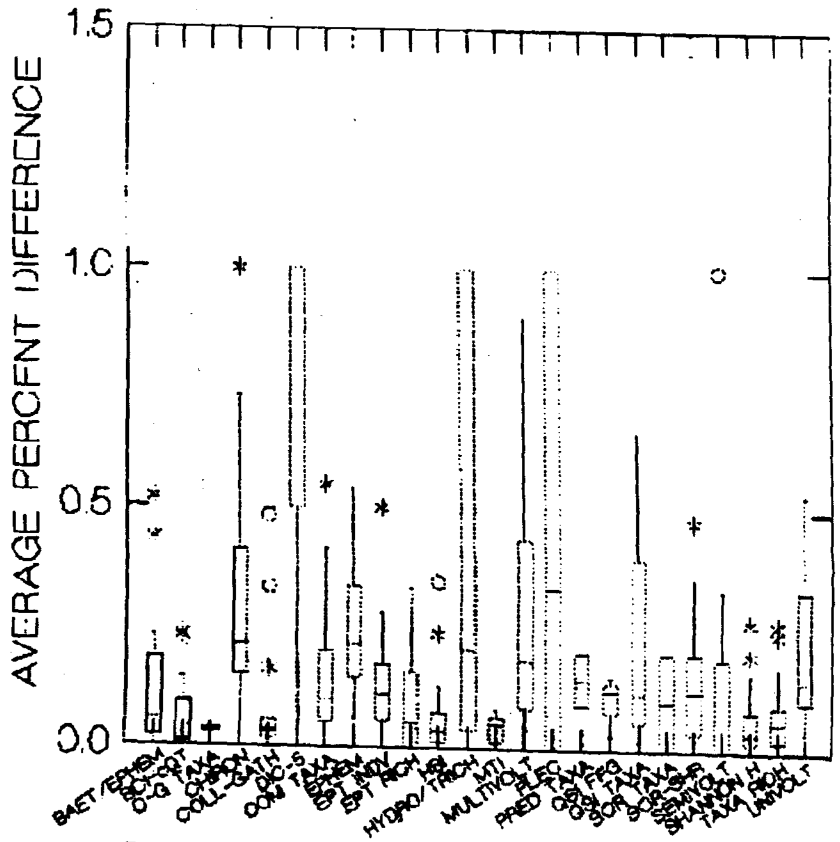
FOOTHILLS ECOREGION

SAMPLE VARIATION IN MACROINVERTEBRATES



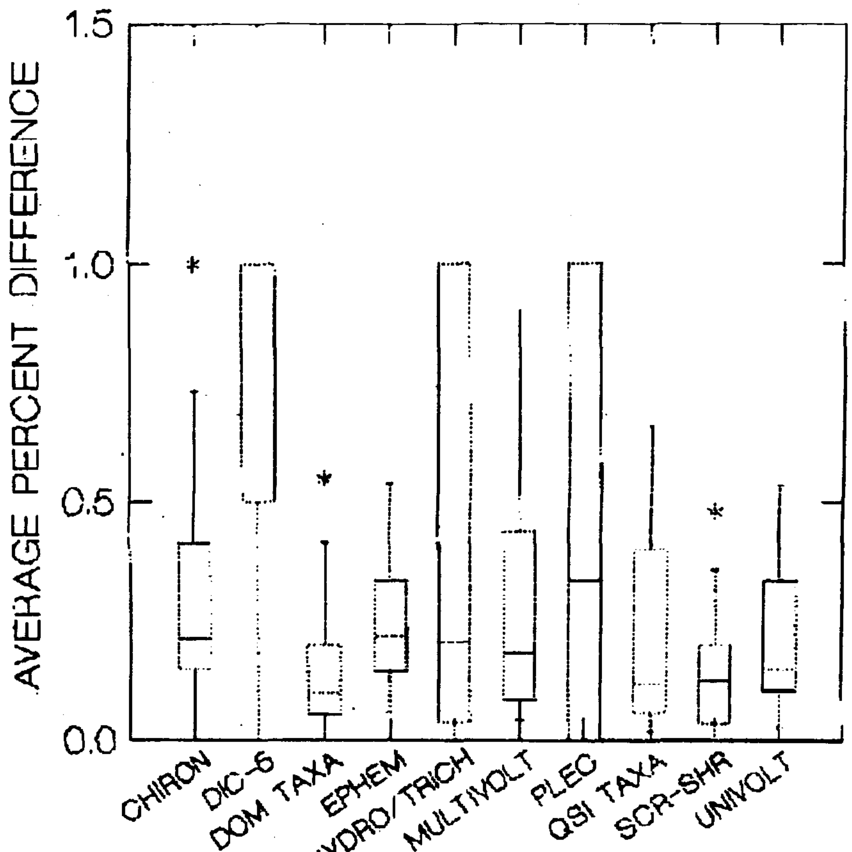
FOOTHILLS ECOREGION

SAMPLE VARIATION IN MACROINVERTEBRATES



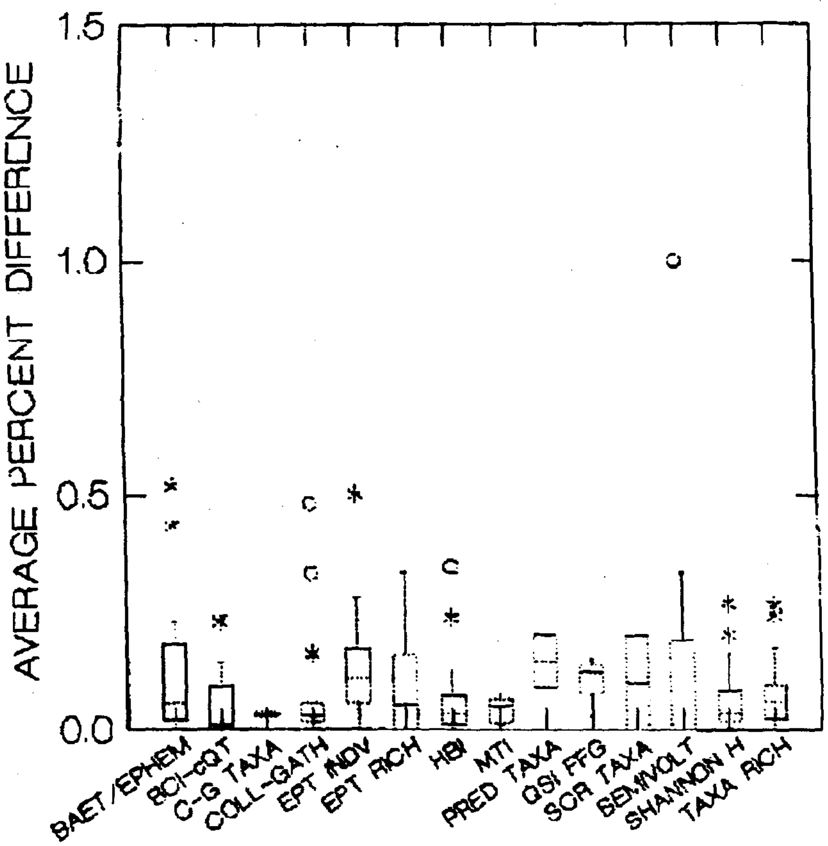
SPRING ECOREGION

SAMPLE VARIATION IN MACROINVERTEBRATES



SPRING ECOREGION

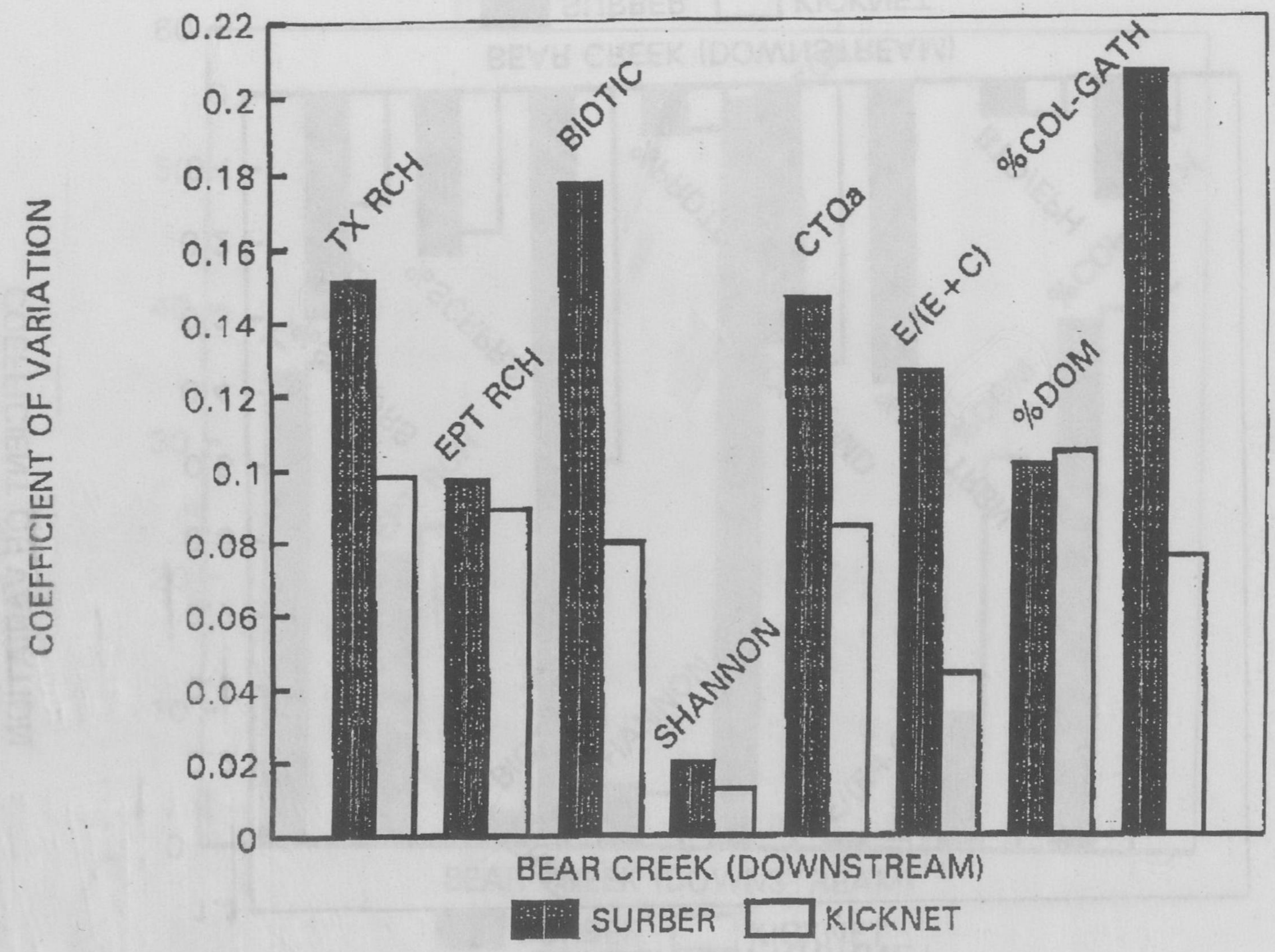
SAMPLE VARIATION IN MACRCINVERTEBRATES



SPRING ECOREGION

SAMPLING VARIATION

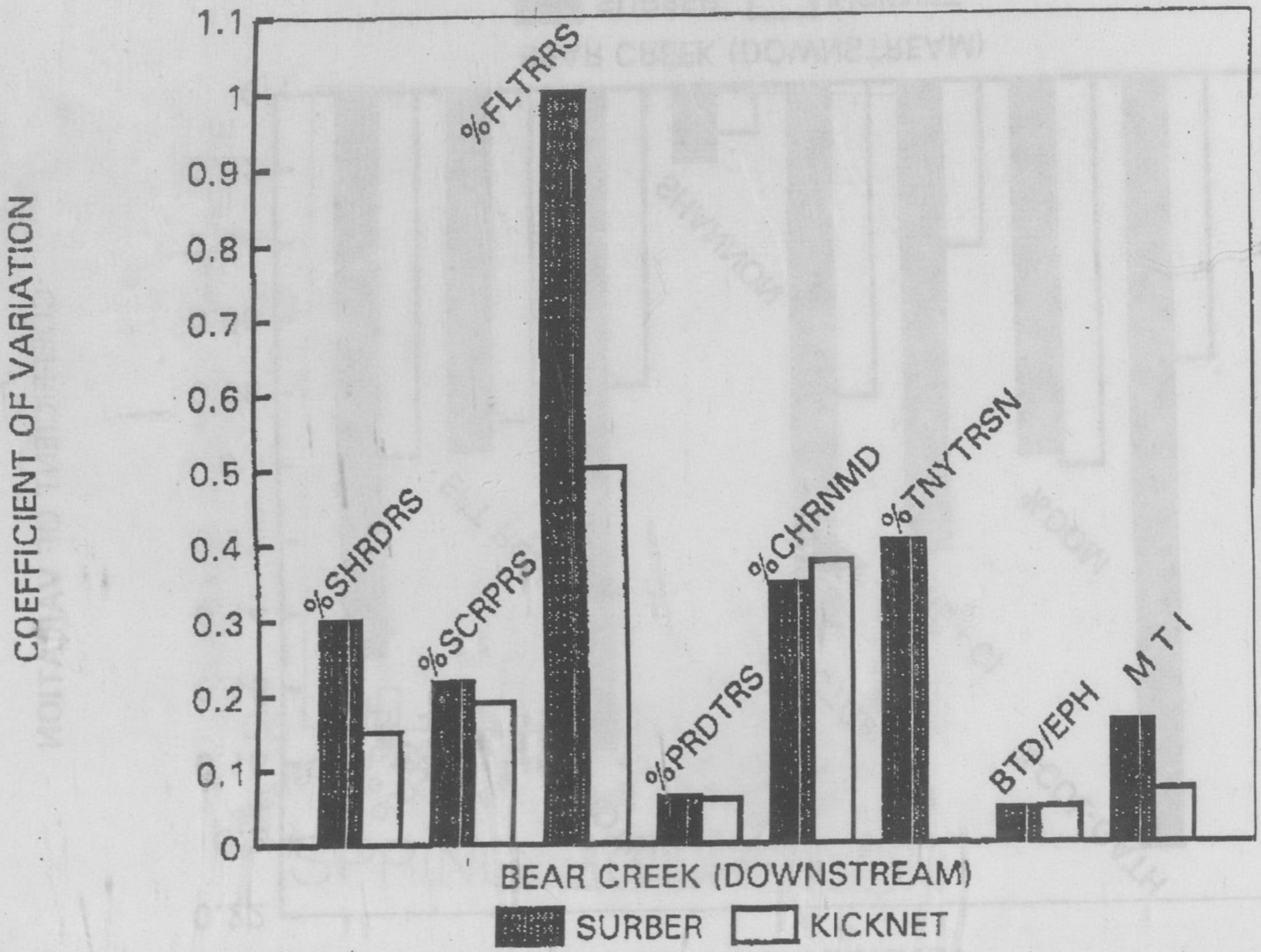
SURBER VS. KICKNET



DATA BY: INTER-FLUVE, INC., 1994

SAMPLING VARIATION

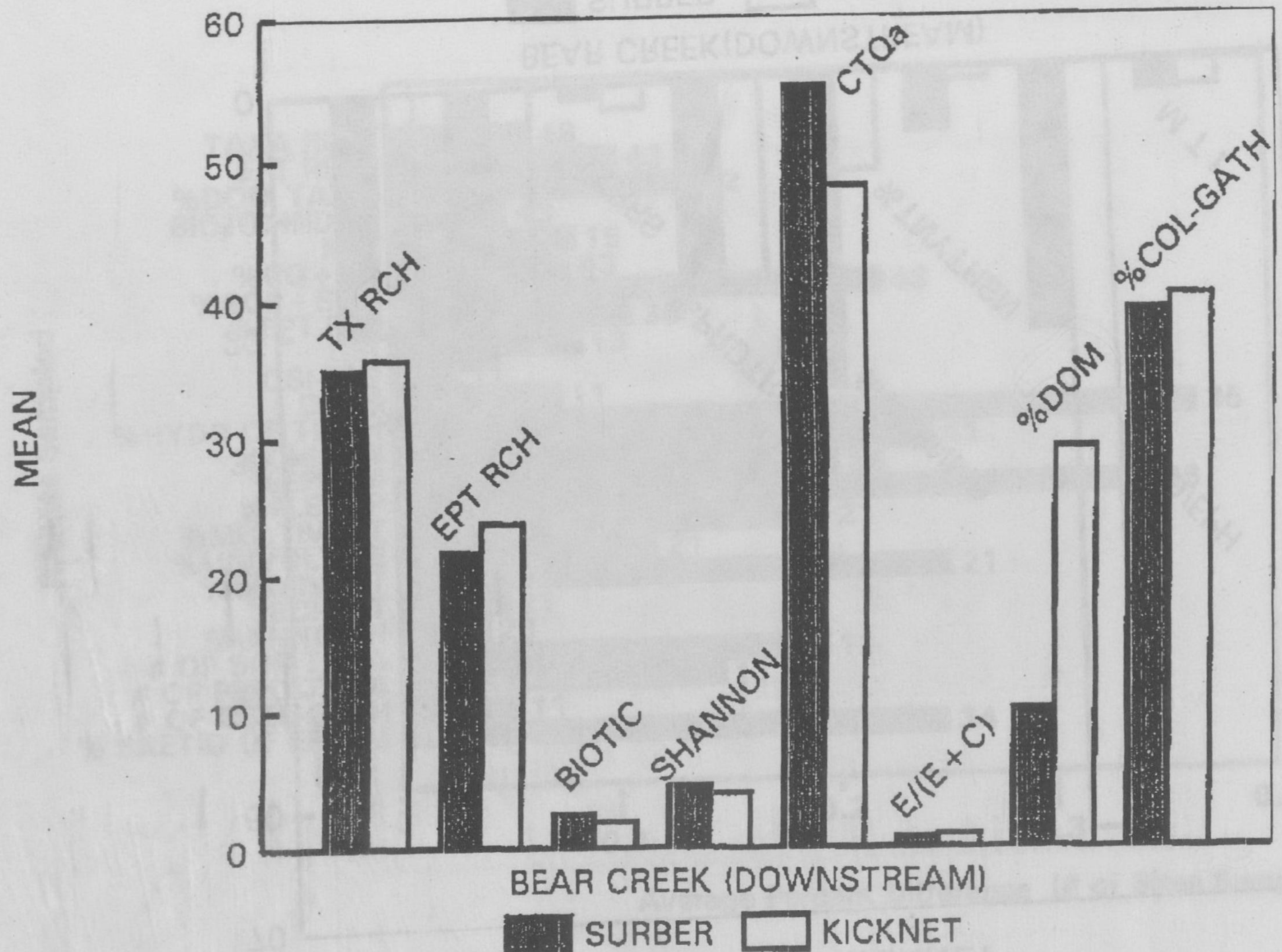
SURBER VS. KICKNET



DATA BY: INTER-FLUVE,INC., 1994

SAMPLING VARIATION

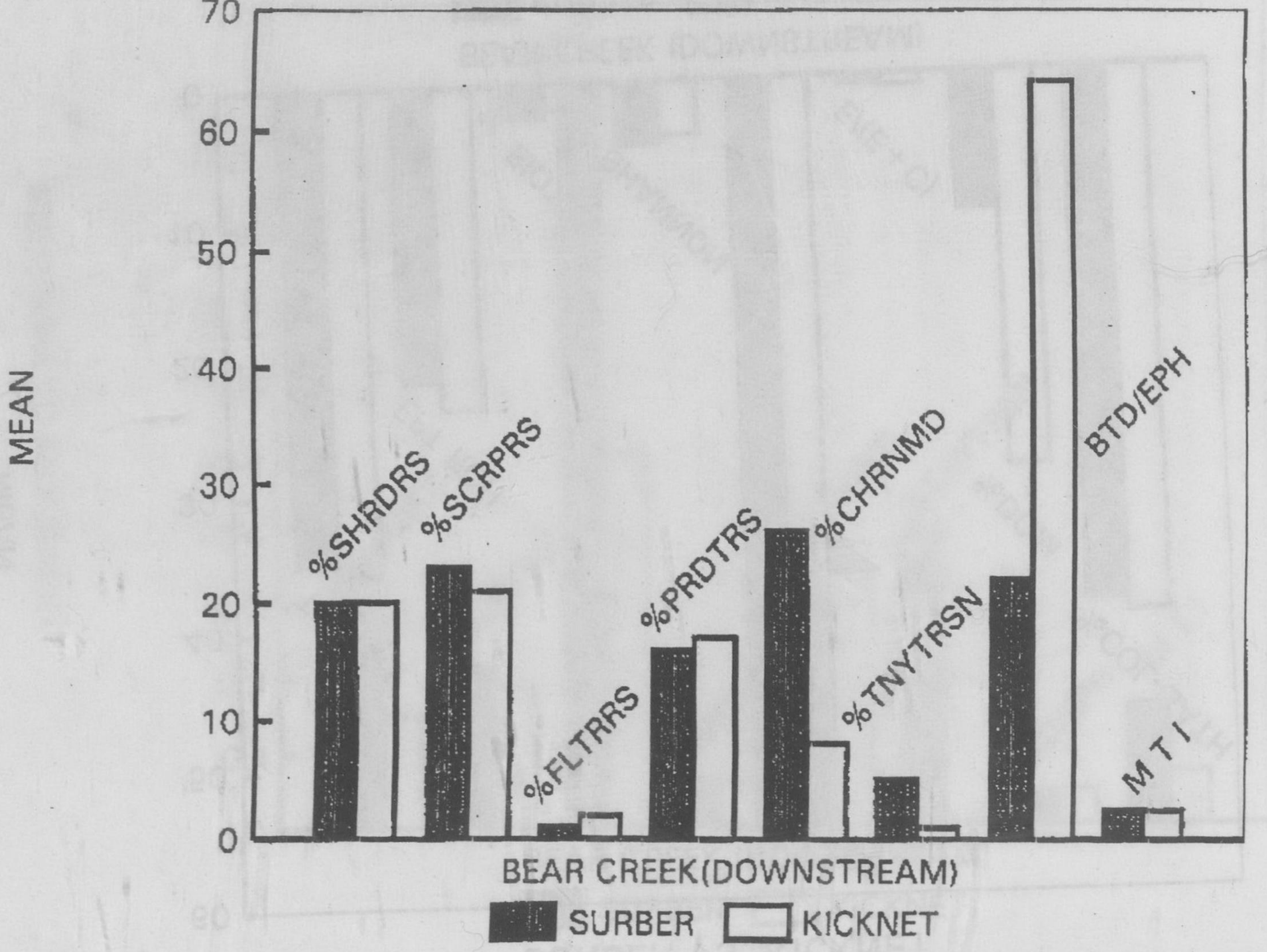
SURBER VS. KICKNET



DATA BY: INTER-FLUVE, INC., 1994

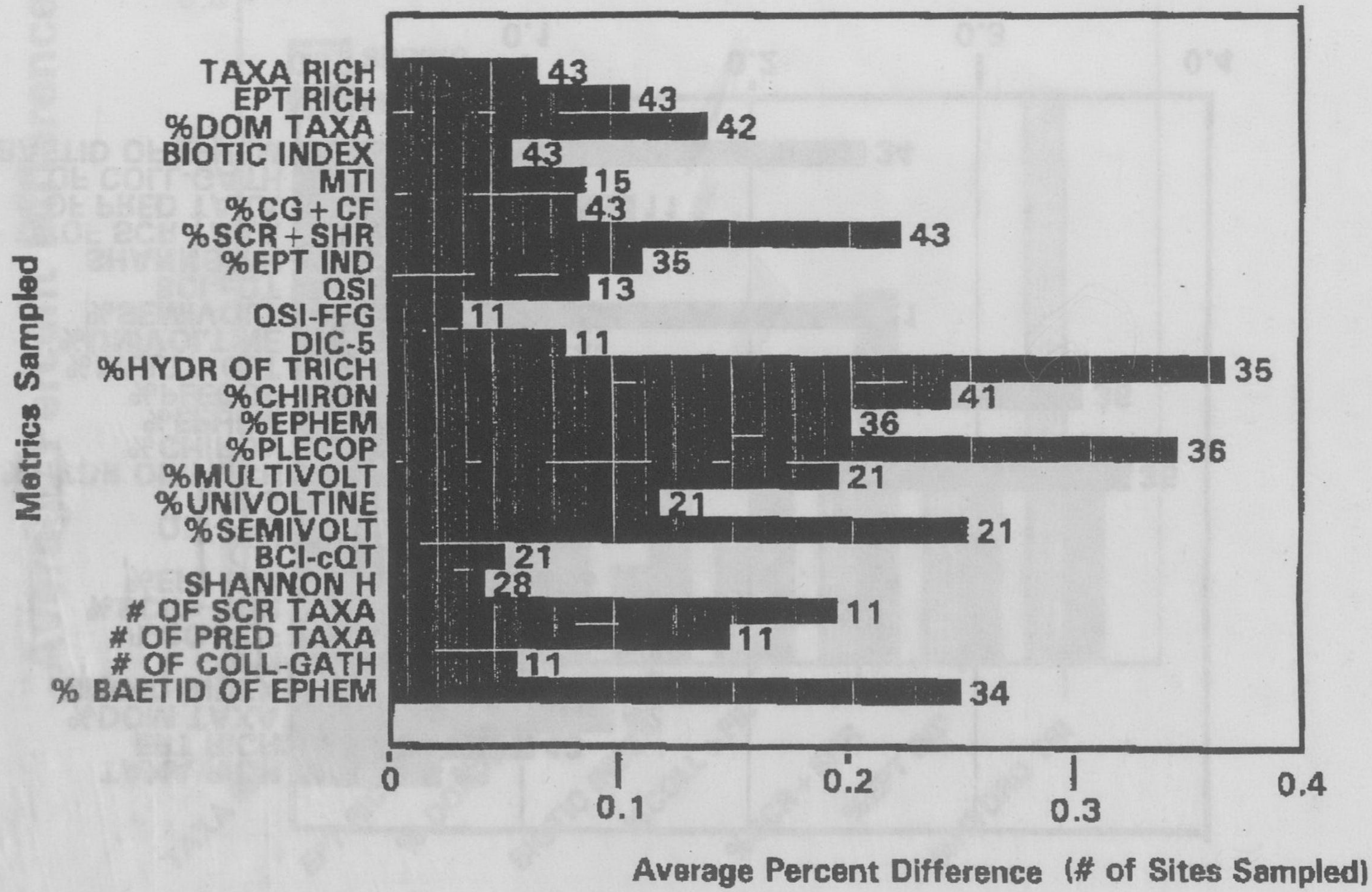
SAMPLING VARIATION

SURBER VS. KICKNET

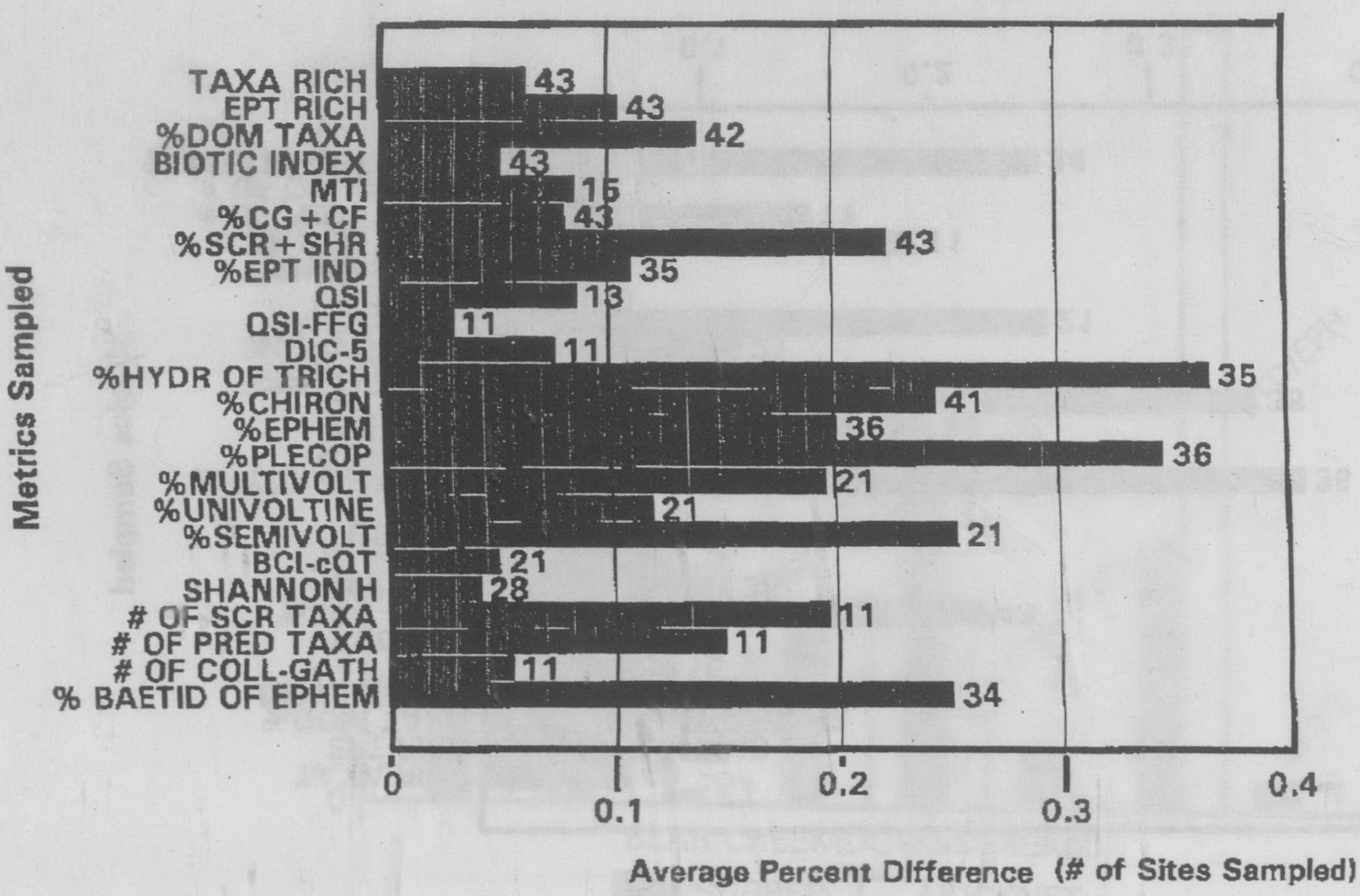


DATA BY: INTER-FLUVE, INC., 1994

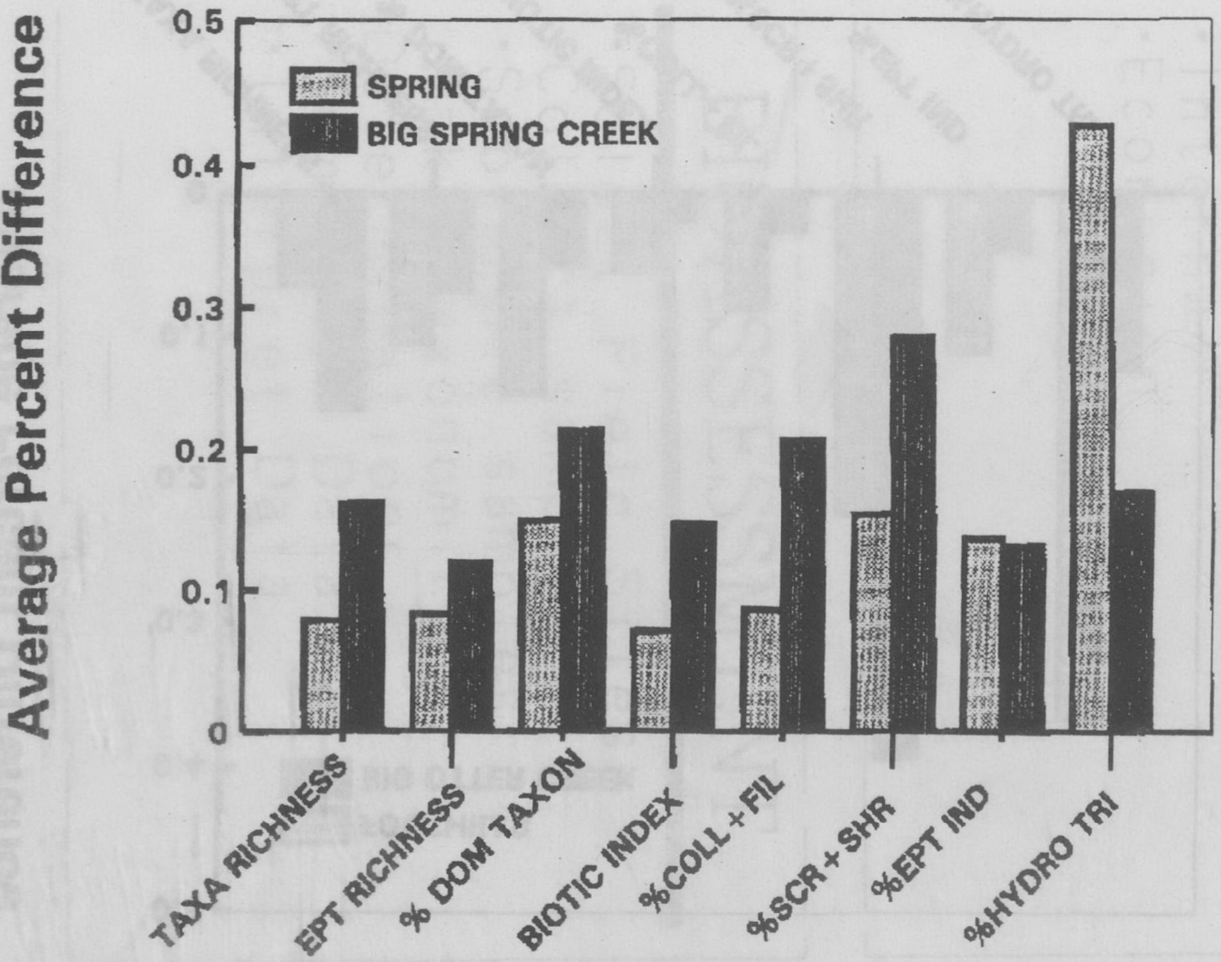
FOOTHILLS STREAMS



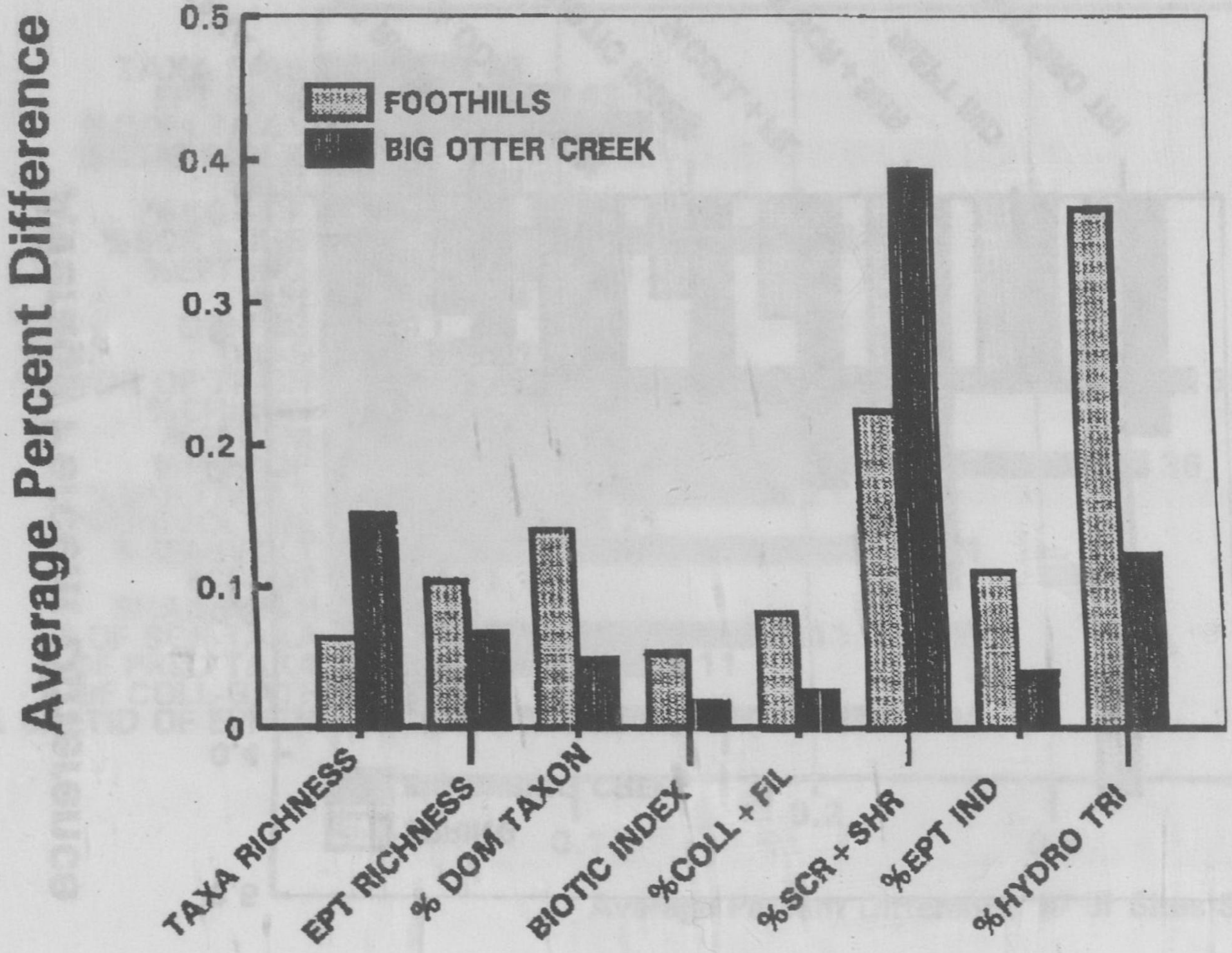
FOOTHILLS STREAMS



INVESTIGATOR BIAS COMPARISON



INVESTIGATOR BIAS COMPARISON



96

REFERENCE COMMUNITIES

- Upstream-Downstream
- Paired Watershed
- Internal
- Ecoregion

BIOASSESSMENT

- Select Field Sites
- Collect samples
- Subsample samples
  Make Taxonomic
Determinations
- Summarize Data
- Interpret Data

DATA SUMMARY AND INTERPRETATION

- COMPUTE metrics
- COMPARE metrics to REFERENCE
  - assign metric scores
- SUM metric scores
- CONVERT to PERCENT

MOUNTAIN REFERENCE STREAMS

	SCORING CRITERIA			
	non	slight	moderate	severe
metric	6	4	2	0
Taxa rich	>28	28-24	24-19	<19
EPT rich	>19	19-18	17-16	<16
Biotic	<3	3-4	4-5	>5
%DOM taxon	<25	25-35	35-45	>45
%Coll(G+F)	<60	60-70	70-80	>80
%Scr+Shr	>55	55-40	40-25	<25
%Chiron	<25	25-35	35-45	>45

(McGuire 1994)

PLAINS REFERENCE STREAMS

metric	SCORING CRITERIA non 6	slight 4	moderate 2	severe 0
Taxa richness	>24	24-18	18-12	<12
EPT richness	>8	8-6	5-3	<3
Biotic index	<5	5-6	6-7	>7
%DOM taxon	<30	30-45	45-60	>60
%Coll(G+F)	<60	60-80	80-95	>95
% EPT	>50	50-30	30-10	<10
%Scr+Shr	>30	30-15	15-3	<3
# Predator taxa	>5	4-5	3-4	<3
S. Diversity	>3.0	3.0-2.4	2.4-1.8	<1.0
%Multivolt	<40	40-60	60-80	>80

(McGuire 1995)

FOOTHILL AND VALLEY REFERENCE STREAMS

metric	SCORING CRITERIA non 6	slight 4	moderate 2	severe 0
Taxa richness	>28	28-21	21-14	<14
EPT richness	>14	14-13	12-11	<11
Biotic index	<4	4-5	5-6	>6
%DOM taxon	<30	30-40	40-50	>50
%Coll(G+F)	<60	60-75	75-90	>90
%Scr+Shr	>30	30-20	20-10	<10
Hydropsych.	<75	75-85	85-95	>95
EPT	<60	60-45	45-30	>30

(McGuire 1995)

THREEMILE CREEK

	Illinois Bench Rd		
	8/90	8/93	8/93
Taxa richness	32	31	31
EPT richness	15	13	16
Biotic index	4.7	3.6	3.0
% Dom taxon	22	23	30
% Coll-FFG	80	60	52
% Scr + Shr	19	29	38
% EPT	28	41	50
% Multivoltine	68	31	31

(McGuire 1995)

THREEMILE CREEK

	Illinois Bench Rd		
	8/90	8/93	8/93
Taxa richness	4	4	4
EPT richness	0	0	0
Biotic index	2	2	2
% Dom taxon	6	6	6
% Coll-FFG	2	6	6
% Scr+Shr	2	4	6
% EPT	2	4	4
% Multivoltine	0	4	4
Total score	16	30	32
% of Reference	33%	63%	67%
Classification	MOD	SLI	SLI

(McGuire 1995)

THREEMILE CREEK REFERENCE VALUES

metric	non 6	slight 4	moderate 2	severe 0
	SCORING CRITERIA			
Taxa richness	>33.6	33.6-25.2	25.2-16.8	<16.8
EPT richness	>22.5	22.5-20	20-17.5	<17.5
Biotic index	>1.96	1.96-1.64	1.64-1.15	<1.15
%DOM taxon	>8.4	8.4-6.3	6.3-4.2	<4.2
%Coll(G+F)	>40.8	40.8-35.7	35.7-30.6	<30.6
%Scr+Shr	>31.2	31.2-23.4	23.4-15.6	<15.6
%EPT	>51.75	51.75-34.5	34.5-17.25	<17.25
%Multivolt.	>4.8	4.8-3.6	3.6-2.4	<2.4

(McGuire 1992)

Relative Percent Difference Between 1990-94 In Reference Streams

	Plains	Mountains
Taxa Richness	16	6
EPT Richness	19	7
% Dominant	30	17
% Collectors (g+f)	11	19
% Scrapers	82	16
% Coll-gath	49	21
% Scr+Shr	54	10
% Chiron	56	54
% EPT	48	7
Biotic index	11	23
S. Diversity	18	6
# Predator taxa	26	13

(McGuire 1995)

USE OF LAKE BENTHOS AS A BIOASSESSMENT TOOL

Malcolm Butler
Department of Zoology
North Dakota State University
Fargo, ND 58105

NOTES:

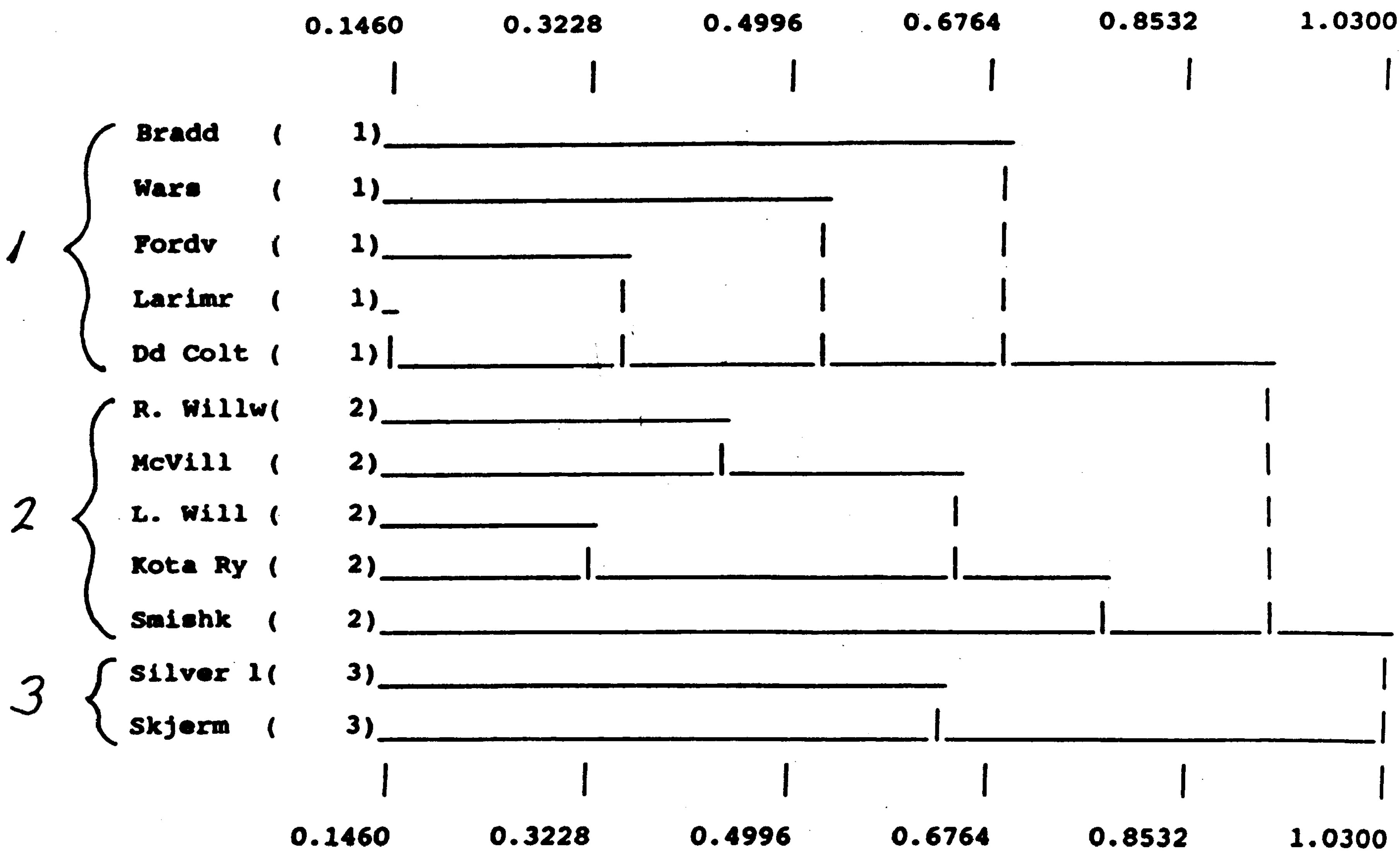
The profundal benthos has been used to classify lakes into types for over 70 years. Numerous classifications and indices have been generated based on (a) observed patterns in the distribution of taxa, and (b) knowledge of the ecological requirements or tolerances of specific benthic animals. Such indicator-based systems have served well to monitor lakes for impacts whose effects on the fauna are known. Multivariate approaches permit a broader focus in the search for patterns within faunistic and environmental data.

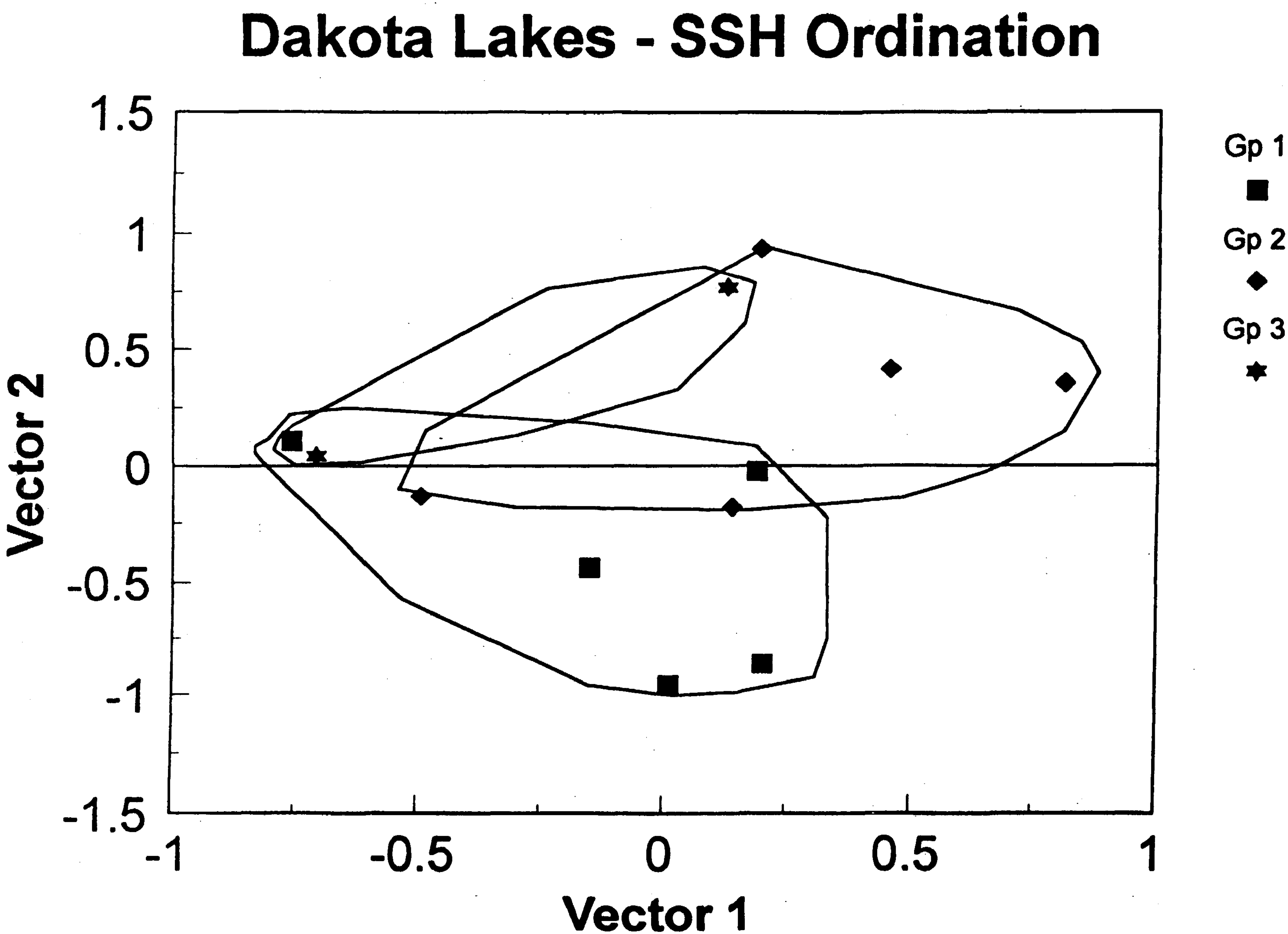
We are surveying the profundal benthos across a wide diversity of lakes in Region VIII states, for which basic limnological data are available. We hope to identify characteristic benthic communities and indicator taxa that reflect environmental differences such as lake depth, size, region, and trophic state. I will illustrate the general approach with an initial data set on 12 North Dakota lakes that was analyzed by Dr. Trefor B. Reynoldson of the Canadian National Water Research Institute. Three groups of lakes resulted from cluster analysis of the benthic macroinvertebrate communities, and non-metric multidimensional scaling identified potential indicator taxa. Discriminant function analysis showed that lake depth, size, trophic state index, and dissolved oxygen at Z_{max} did a fair job of distinguishing the three lake groups independently defined by benthic macroinvertebrate communities.

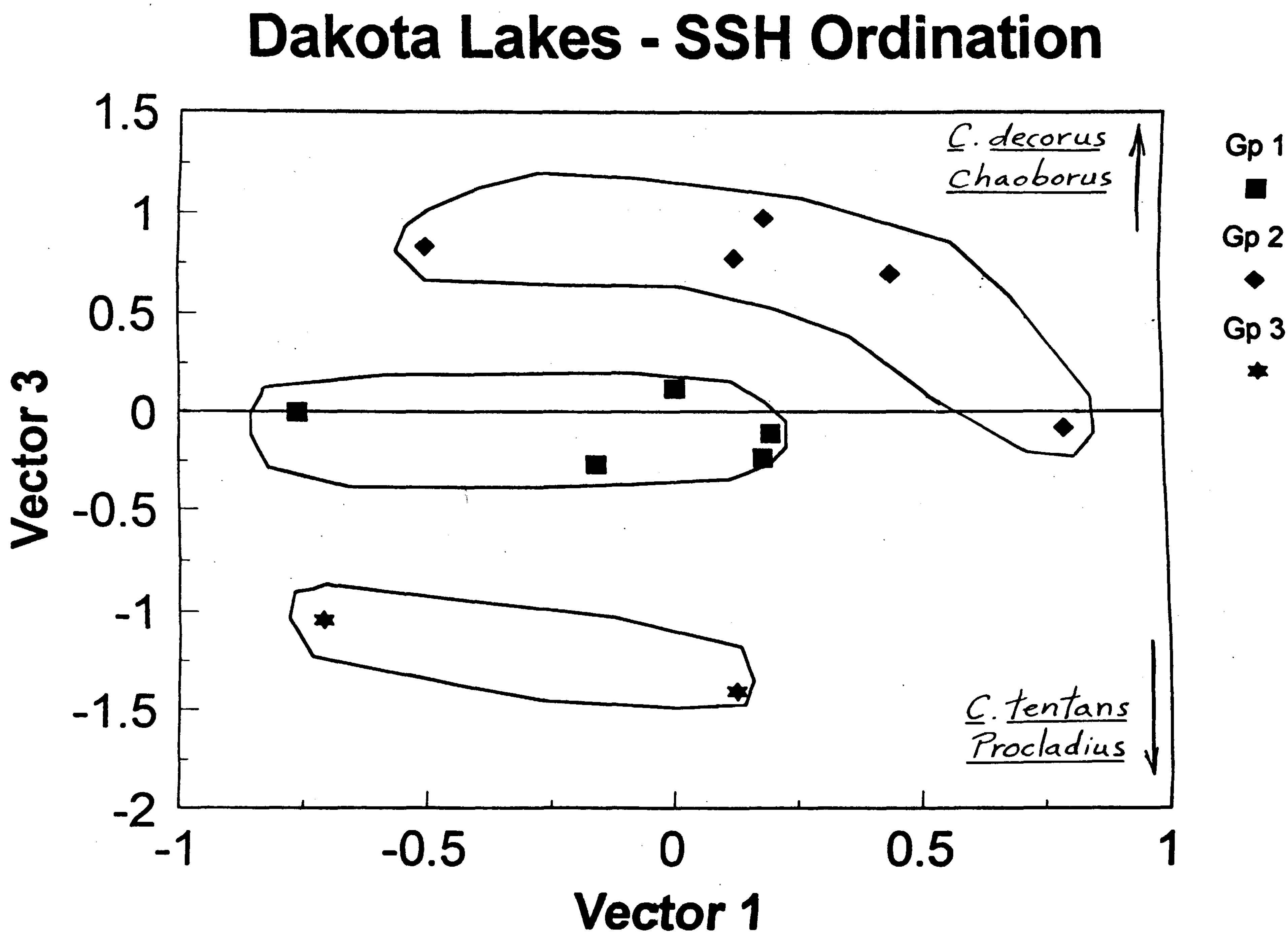
To further illustrate the generality and potential of this approach, I present a similar analysis of 26 Swedish lakes published by Dr. Richard K. Johnson (Department of Environmental Assessment, Swedish University of Agricultural Science). Johnson uses alternative multivariate methods to classify lakes according to their biota (TWINSPAN) and to seek environmental correlates with species abundance patterns (Canonical Correspondence Analysis) prior to producing a predictive model of species occurrence with discriminant function analysis.

The most obvious results from these analyses often reiterate the limnological wisdom of seat-of-the-pants biologists from decades past. However, multivariate approaches have the potential to identify subtle, confounded, or unexpected patterns in extensive data sets.

03/08/94 16:21:33.14 DEND Dakota Lakes Invertebrate Communities



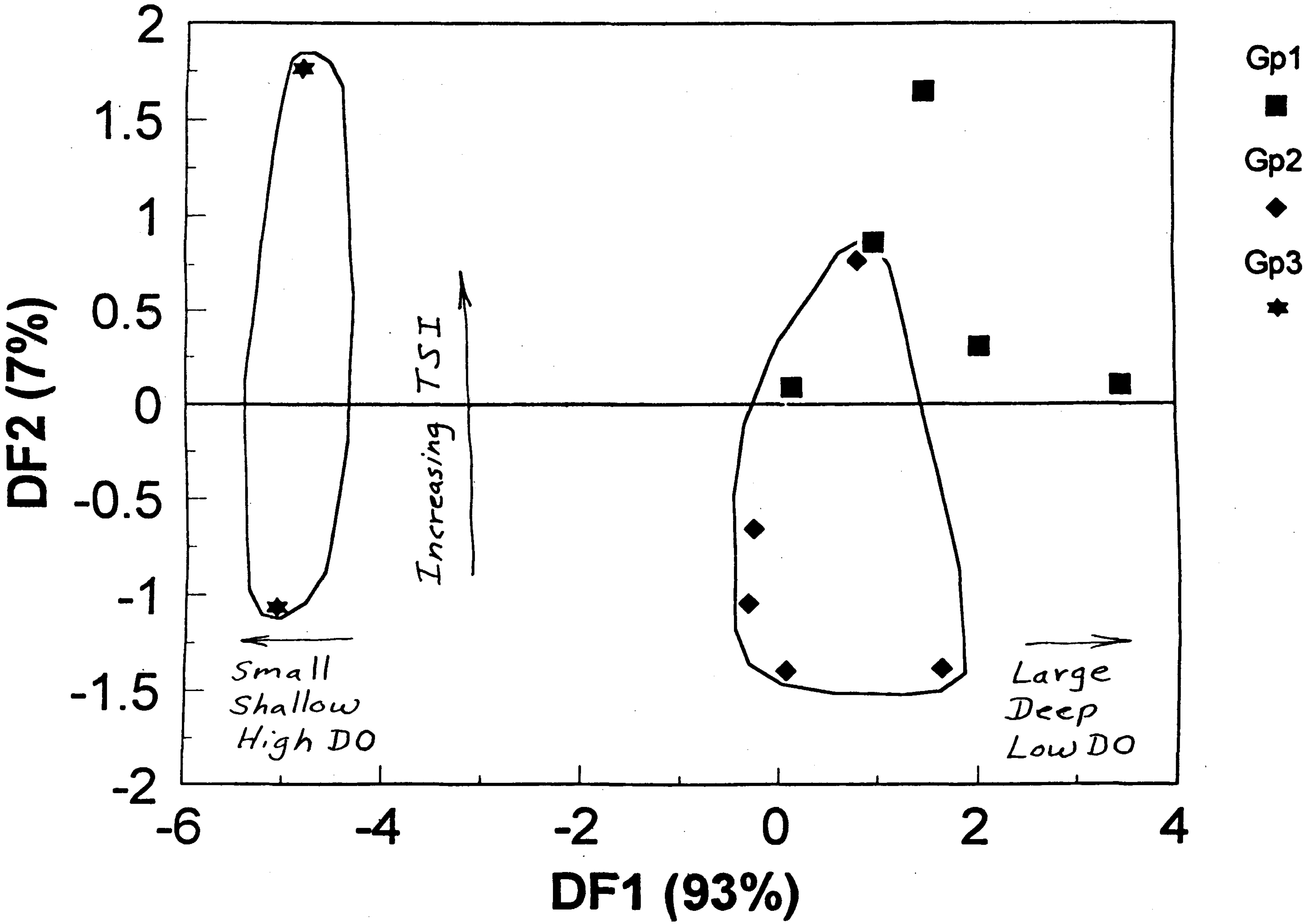

Dakota Lakes - SSH Ordination
1.5
1
0.5
0
-0.5
-1
-1.5
Vector 2
-1
-0.5
0
0.5
1
Vector 1
Gp 1
Gp 2
Gp 3


Dakota Lakes - SSH Ordination
Vector 3
Vector 1
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-1
-0.5
0
0.5
1
C. decorus
Chaoborus
C. tentans
Procladius
Gp 1
Gp 2
Gp 3

SAS

OBS	SITE	ZMAX	ZMN	AREA	DO	TSI	GP
1	Bradd	4	1.8	28.1	5.0	70	1
2	Wars	6	3.4	21.6	1.1	61	1
3	Fordv	6	3.2	75.7	2.3	69	1
4	RWillw	6	3.1	12.1	2.5	55	2
5	LWillw	7	2.9	62.6	5.9	51	2
6	Silverl	3	2.6	36.8	4.1	70	3
7	Skjerm	5	3.0	16.2	2.9	47	3
8	Smishk	7	2.9	75.9	2.9	54	2
9	Larimr	9	4.5	24.1	0.6	59	1
10	McVill	6	3.6	13.1	0.6	66	2
11	DdColt	10	5.6	45.7	0.3	63	1
12	KotaRy	7	3.5	11.9	1.1	49	2

Discriminant Analysis - Dakota Lakes



Prediction of Dakota lakes invertebrate assemblages

Discriminant Analysis

Cross-validation Summary using Linear Discriminant Function

Number of Observations and Percent Classified into Group:

From Group	1	2	3	Total
1	2 40%	3 60%	0	5 100%
2	2 40%	2 40%	1 20%	5 100%
3	0	0	2 100%	2 100%
Total	4 33%	5 42%	3 25%	12 100%

Error Count Estimates for Group:

	1	2	3	Total
Rate	60%	60%	0	40%

Prediction of Dakota lakes invertebrate assemblages

Discriminant Analysis

Resubstitution Summary using Linear Discriminant Function

Number of Observations and Percent Classified into Group:

From Group	1	2	3	Total
1	4 80%	1 20%	0	5 100%
2	1 20%	4 80%	0	5 100%
3	0	0	2 100%	2 100%
Total	5 42%	5 42%	2 17%	12 100%

Error Count Estimates for Group:

	1	2	3	Total
Rate	20%	20%	0	13%

Dr. R.K. Johnson: Dept. of Environmental Assessment, Biodiversity Section, Swedish Univ. of Agricultural Science, Box 7050, S-750 07 Uppsala, Sweden

Swedish reference lakes
(n = 26)

(n = 22)
C. flavicans (Meigen)
Ch. anthracinus Zett.
Z. zalutschicola Lipina
S. coracina (Zett.)

(n = 5)
P. triquetra Tshern.
Dicrotendipes spp.
P. prasinatus (Staeg.)
Microtendipes spp.
A. aquaticus L.
M. angustata Curtis

G4
BD7
F4
G9
BD15

Group 2

(n = 17)
S. coracina
S. rosenschoeldi (Zett.)
C. flavicans

(n = 9)
C. flavicans
Ch. plumosus L.
Ch. tenuistylus Brundin

E6 O4
F10 O5
H8 P6
K2 R4
N4

Group 3

(n = 8)
Monodiamesa spp.
S. rosenschoeldi
Pisidium spp.

A4 W9
E5 X3
N6 X11
S18 Y1

Group 4

(n = 4)
Micropsectra spp.
H. subpilosus (Kieff.)
Paracladoplema spp.
S. ferox (Eisen)

H1
AC16
BD20
U3

Group 1

Figure 5 from: R.K. Johnson (1995), The indicator concept in freshwater biomonitoring. In P. Cranston (ed.) Chironomids: From genes to ecosystems. CSIRO Publications, East Melbourne, Australia

Figure 6 from: R.K. Johnson (1995)

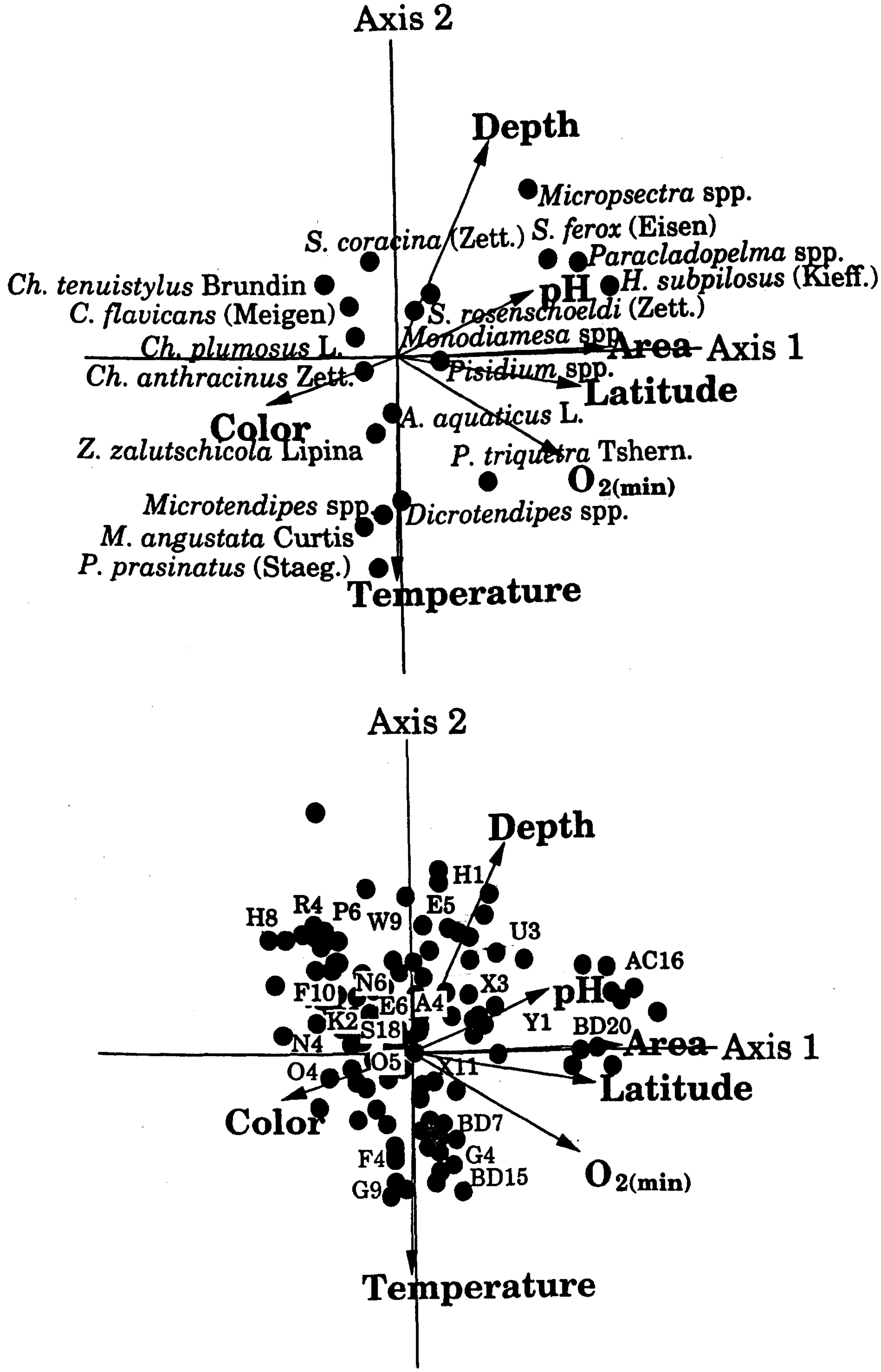


Table 3 from: R.K. Johnson (1995)

Table 3. The predicted probability of occurrence of selected taxa in limed reference lakes compared with observed values (frequency of lakes where the taxon occurred). Discriminant function analysis was performed with seven environmental variables and TWINSPAN group classification of 26 Swedish reference lakes (see text). MP = median probability of taxon occurrence and % O = percent of lakes where taxon was found occurring.

TWINSPAN group 1			TWINSPAN group 4		
	MP	% O		MP	% O
Oligochaeta	99	100	*Procladius* spp.	75	100
Procladius spp.	75	100	*S. coracina*	75	100
Tanytarsus spp.	75	75	*Tanytarsus* spp.	75	100
Pisidium spp.	75	67	Oligochaeta	75	100
S. ferox	75	58	*Pisidium* spp.	75	0
H. apicalis	75	58	*C. flavicans*	66	100
Nematoda	75	50	*Monodiamesa* spp.	66	0
Protanypus spp.	75	42	Hydracarina	57	100
Micropsectra spp.	75	33	*Z. zalutschicola*	56	100
Turbellaria	75	25	*S. rosenschoeldi*	56	50
H. subpilosus	75	25	*Ch. anthracinus*	47	50
Monodiamesa spp.	75	17	*H. apicalis*	47	0
Paracladopelma spp.	75	8	*Polypedilum* spp.	38	0
S. coracina	50	92	Nematoda	38	0
S. rosenschoeldi	50	67	*Micropsectra* spp.	28	0
Hydracarina	50	58	Turbellaria	28	0
Psectrocladius spp.	50	25	*T. tubifex*	28	0
H. marcidus	50	25	*S. ferox*	19	0
Polypedilum spp.	50	17	*H. marcidus*	19	0
Gammarus spp.	50	0	*Psectrocladius* spp.	19	0
T. tubifex	50	0	*Protanypus* spp.	9	0
C. flavicans	25	58	*Gammarus* spp.	0	0
Z. zalutschicola	0	33	*H. subpilosus*	0	0
Ch. anthracinus	0	17	*Paracladopelma* spp.	0	0

ASSESSING EFFECTS OF METALS ON BENTHIC MACROINVERTEBRATE COMMUNITIES IN ROCKY MOUNTAIN STREAMS

Peter Kiffney
Department of Fishery and Wildlife Biology
Colorado State University
Fort Collins, CO 80523

NOTES:

Biological assessment to determine ecological integrity of aquatic systems has become increasingly important with the realization that chemical monitoring is not always protective of aquatic life. Our research group has used a number of approaches in assessing the biological integrity of Rocky Mountain streams, such as whole effluent toxicity tests, field biomonitoring, and toxicity tests using indigenous organisms in experimental streams. I will discuss the merits of these methods, and provide an example of each approach in evaluating the effects of metals on stream benthic organisms. Because of the inherent natural temporal and spatial variability of Rocky Mountain streams, I suggest that an integrative approach that incorporates both experiments and field observations be used when assessing the effects of anthropogenic inputs on stream organisms.

The traditional approach to assessing the biological effects of contaminants on aquatic systems has been the single-species toxicity test. These tests can be reproducible, easy to perform, comparable across laboratories, and defensible in court. However, single-species tests are also ecologically unrealistic for a number of reasons. For example, test organisms are typically not found in mountain streams (i.e., *Ceriodaphnia dubia*) and, thus, responses to an effluent observed in a single species test may not be applicable to responses of natural populations in the receiving stream.

An alternative method to single-species toxicity testing is the use of indigenous organisms in experimental streams. In a series of experiments, I used artificial substrates to collect invertebrates and transfer them to experimental streams where they were exposed to metals. At the end of 10d, the effects of metal treatment on population-, and community-level indices were examined using ANOVA or regression techniques. With this experimental system, I examined the influence of stream altitude on the response of benthic macroinvertebrate communities to metals. Results from these experiments showed that the response of invertebrates to metals was affected by stream altitude, as insects from high-altitude streams were more sensitive to metals than similar groups from low-altitude streams. Additionally, my research showed that the most sensitive group to metals was mayflies, especially heptageniid mayflies. Thus, this family may be a reliable and sensitive indicator of metal-pollution in western streams.

Field biomonitoring using stream invertebrates has become increasingly important in biological assessment of aquatic systems. Biomonitoring studies have limitations, however, because of a number of statistical problems. Nevertheless, the use of field biomonitoring can provide insights into processes occurring in metal-polluted streams. In the Arkansas River, our group has collected stream invertebrates to measure metal levels in tissue and for community structure. Because this is a single system study, inference from these results are limited; therefore, we have also collected similar data from five other metal-polluted streams in the central mountains of Colorado. Results from these surveys have allowed us to develop metrics for metal-contaminated streams. Some metrics, such as abundance of caddisflies, were confounded by other abiotic factors (altitude), whereas abundance of mayflies or mayfly species richness were consistently lower at metal-polluted sites than at reference sites. Because this study included surveys of six streams, results have a broader implication for a larger population of metal-polluted streams in Colorado.

Since there are advantages and disadvantages associated with all methods of biological assessment, I suggest that biologists responsible for evaluating the ecological integrity of aquatic systems use an integrative approach. For example, in our research group we have combined stream biomonitoring with whole effluent toxicity tests, and/or toxicity experiments with indigenous organisms to evaluate the ecological effects of metals on natural populations of stream invertebrates in Rocky Mountain streams. This approach has provided metrics that have been found to be consistently sensitive to metal pollution in Colorado Rocky Mountain streams, and may be applicable to other regions of the intermountain west.