Terrestrial and Aquatic Environmental Toxicology ©2007 Global Science Books

The Impact of Golf Course Construction on Benthic Macroinvertebrate Communities: An Evaluation of Bioassessment Techniques

J. C. Schulenburg[1,4*] • P. J. Dillon[1] • K. M. Somers[1,2] • J. G. Winter[1,3]

[1] Trent University, Environmental and Resources Studies, 1600 West Bank Drive, Peterborough, Ontario K9J 7B8 Canada
[2] Ontario Ministry of the Environment, Dorset Environmental Science Centre, Dorset, Ontario P0A 1E0 Canada
[3] Ontario Ministry of Environment, 125 Resources Road, Toronto, Ontario M9P 3V6 Canada
[4] Current Address: Quinte Conservation, 2061 Old Highway 2, R.R. 2, Belleville, Ontario K8N 4Z2 Canada

Corresponding author: *** jschulenburg@quinteconservation.ca**

ABSTRACT

There are numerous methods for assessing Benthic Macroinvertebrate (BMI) community heath using the Reference Condition Approach (RCA); however, no standard method of BMI analysis has been established. This study attempted to evaluate four distance measures and determined the most robust method for producing benthic indices that can be used in a multivariate approach. In Muskoka, BMI stream communities were collected from test sites located on golf courses that were at various stages of construction to being fully operational. BMI communities were also collected at references sites from streams that were considered to be relatively pristine. With this study design each site was classified according to one of four land-use categories. The classified BMI data were then used to calculate the four distance measures. These were Jaccard (presence and absence), Chi-square, Bray-Curtis based on BMI taxonomic group abundances, and Bray-Curtis on benthic indices. Criteria used to evaluate the distance measures were strength of correlations with NMDS axes; average Euclidean distance among land-use group means; temporal variability; number of Mantel tests that were significant between BMI and land-use matrices; and number of Mantel tests that were not significant between BMI and physical-geographic matrices. Each measure had different strengths, but Bray-Curtis based on benthic indices scored the best according to the criteria. In conclusion the first three axes from this distance measure are recommended to be used as BMI indices in bioassessment. Testing and compensating for temporal and spatial variation were found to be necessary when using any distance measure to calculate BMI indices.

Keywords: Benthic indices, Bray-Curtis, distance measure, multivariate approach, Reference Condition Approach, streams
Abbreviations: AusRivAs, Australia River Assessment Scheme; **BACI**, Before-After/Control-Impact; **BEAST**, BEnthic Assessment of SedimenT; **BMI**, benthic macroinvertebrates; **EPT**, Ephemeroptera, Plecoptera, and Trichoptera; **GPS**, Global Positioning System; **HBI**, Hilsenhoff's Biotic Index; **MANOVA**, non-parametric multivariate analysis of variance; **NMDS**, Non-metric Multidimensional Scaling; **OBM**, Ontario Base Maps; **RCA**, Reference Condition Approach; **RIVPACS**, River InVertebrate Prediction And Classification System

INTRODUCTION

Biomonitoring or biological assessment is the use of biological variables to survey the environment and assess human impact (Bonada *et al*. 2006). Rapid bioassessment is designed to reduce the cost and effort of sampling by reducing the number of habitats and/or replicate sampling units per site (David *et al*. 1998; Barbour *et al*. 1999) allowing more time to sample different sites. In addition, there may be only a set number of organisms sub-sampled (e.g. 100) from each replicate taking less time to process each site.

In aquatic bioassessments, the biotic community at a test site is compared to a desired level defined by the pre-treatment condition or a reference condition. Sampling designs traditionally compare treatment to control. For example in the Before-After/Control-Impact (BACI) design comparisons are made through either time or space (i.e. downstream vs. upstream locations). Alternatively, the Reference Condition Approach (RCA) is used when there are no appropriate before treatment or upstream sample sites available (Barbour *et al*. 1999). The RCA involves the comparison of a test site to a group of reference sites that are similar except for the variables that may be altered by the treatment, e.g. fertilizer runoff. The reference condition represents minimally impaired conditions that are defined by physical, chemical and/or biological characteristics (Reynoldson *et al*. 1997).

The RCA has been used to assess non-point source pollution in streams associated with urbanization, agriculture, mining, and golf course activities (Bailey *et al*. 1998; Linke *et al*. 1999; Davidson 2002; Winter *et al*. 2002). Although various methods in selecting reference sites have been discussed and compared, no standard method has yet been established (Moss *et al*. 1987; Gerritsen 1995; Hughes 1995; Reynoldson *et al*. 1997; Gerritsen *et al*. 2000). Generally, there are two accepted approaches; multimetric (used typically in the U.S.A) and multivariate (used more often in Canada, the U.K, and Australia; Bonada *et al.* 2006). The multimetric approach classifies reference sites based on predetermined physical or geographic criteria. A number of metrics are measured at reference and test sites to establish biocriteria and subsequently calculate a multimetric index score for each site. A multimetric index is the sum of ordinal-scaled standardized scores that are based on a variety of metrics, e.g. Hilsenhoff's Biotic Index (HBI), Shannon-Wiener diversity, percent abundance of functional feeding groups, percent abundance of intolerant groups (Gerritsen 1995). The multimetric index score for each test site is then compared to a threshold defined from regional reference sites (Gerritsen 1995; Norris 1995; Barbour *et al*. 1999).

The strengths of the multimetric method are that information is summarized and can be used to compare sites

Received: 25 February, 2007. Accepted: 2 October, 2007.

Original Research Paper

over a large geographic range, and component matrices (also referred to as indices) are available to make site-specific assessments in order to test a hypothesis (Fore *et al.* 1996). The method is also easy to understand and interpret because of the use of a single value to describe a condition. The disadvantage in using the multimetric approach is the manner in which reference sites are classified. Reference classification is accomplished using ecoregions defined by geographical boundaries. Evidence has shown that communities, such as benthic macroinvertebrate (BMI), are controlled more by local conditions, i.e., habitat rather than by regional conditions (climate and location) (Richards *et al.* 1993), and that multimetric approaches do not have large-scale applicability across ecoregions (Bonada *et al.* 2006).

Traditionally, multivariate approaches use relative abundances of aquatic communities, e.g. BMI, for statistical analyses (e.g. cluster analysis) to classify reference sites. Physical or geographic criteria are determined by either selecting a reference group and matching it with a test site using probabilities calculated in a discriminant model (BEAST-BEnthic Assessment of SedimenT model, Reynoldson *et al.* 1995) or by using all reference sites that are weighted by the probability of group membership (Aus RivAs -Australia River Assessment Scheme (Parsons and Norris 1996) and RIVPACS-River InVertebrate Prediction And Classification System models (Wright *et al.* 1998; Clarke *et al.* 2003). The presence and absence of BMI in ordination (BEAST) or the probabilities of expected presence of BMI to observed BMI (AusRivAs and RIVPACS) are then used for comparing BMI at test sites to reference sites in a test-site assessment.

The strength in using the multivariate method is that reference sites are classified using local habitat or faunal conditions, which is more objective than the multimetric approach that uses ecoregions (Norris 1995). The assumption is that regional conditions are relatively homogeneous compared to differences among regions, therefore the multivariate approach to reference site classification is sufficiently accurate at predicting expected biological conditions at test sites (Hawkins and Norris 2000). The disadvantages of using the multivariate approach are that it is not easily understood nor applied by non-specialists, and there is a multitude of statistical approaches from which to choose (Gerritsen 1995). Furthermore, the approach lacks the use of significance testing, it assumes that all relevant environmental variables are being measured and, it is intolerant of missing data because variables and sites with missing data are simply eliminated (Norris 1995).

The overall goal of this study is to identify a single distance measure that can be used in bioassessments of BMI. A distance measure represents the differences between samples in a matrix. A mathematical matrix is (usually) a symmetrical square where the rows and the columns represent (usually) samples. The diagonal elements (the difference between a sample and itself) in a matrix are usually zeros. A matrix can represent either differences or similarities among samples. The measure could be based on Euclidean distance, Manhatten (City Block) distance, Bray-Curtis dissimilarity, the Jaccard coefficient to name a few.

A distance measure is designed to obtain multivariate indices from a distance matrix that is calculated from a set of variables. The multivariate index is a way to summarize a BMI community. The distance matrix and its accompanying multivariate indices can be used in a RCA for bioassessment. The objective of the study is to evaluate four distance measures that summarize BMI community structure and to determine which approach is best for bioassessment based on five criteria.

METHODS

Study area

The study area is located in the District of Muskoka, approximately 200 km north of Toronto, ON, Canada. Muskoka is situated

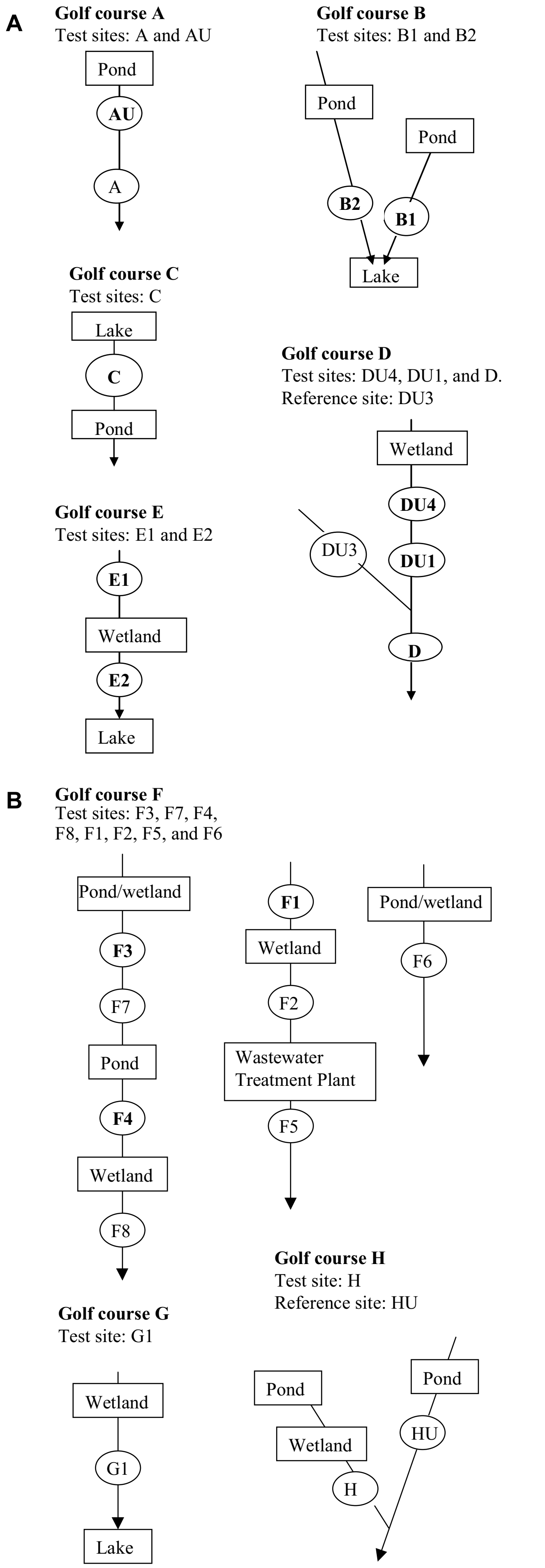

Fig. 1 (**A**) Schematic diagrams for stream sampling sites at each golf course that were in operation during the study in 1999 to 2001. Each diagram represents separate streams and arrows represent direction of flow. The site codes in bold represent sites located on the golf course. (**B**) Schematic diagrams for stream sampling sites at each golf course that were under construction for part of the study. The site codes in bold represent sites located on the course.

near the southern boundary of the Canadian Precambrian Shield, an area characterized by very thin soil/till on silicate bedrock, composed of granitic gneisses and migmatites with some marble, quartzite, amphibolites and various other igneous rocks (Jeffries and Snyder 1983). The numerous lakes and streams in the study area are sensitive to impacts from anthropogenic activities long-range atmospheric transport of acid precursors, such as sulphur and nitrogen; the thin soils result in hydrologic cycles that have shallow flow paths that may not neutralize strong acids or export nutrients, such as phosphorus (Dillon and Molot 1996). The surficial geology consists of till plains (continuous moraine deposits > 1 m thick) and thin till deposits (< 1 m thick), interrupted by rock ridges (Dillon and Molot 1997). The area is generally lightly developed with some recreational cottages and a few small urban areas.

The 19 forested reference sites were located in streams of undeveloped, forested watersheds with no golf courses upstream. Alternatively, the 19 test sites were located on or downstream from golf courses that were under construction or in operation. The test sites were classified into three categories based on the development stage of the golf course i.e. 1) golf course clear-cut, 2) turf establishment, and 3) golf course operation (**Fig. 1A, 1B**). Typically the first year of golf course construction involves clearing trees and other vegetation and prescribed burning. Reshaping the basic topography follows in the first or second year of construction. Turf establishment takes place during fall in year one or in year two. It involves installing drainage systems, spreading sand, sod, and seed, and the application of fertilizers and pesticides. By year two or three and thereafter, the course opens to the public and routine turf management involves irrigation, mowing, and fertilizer and pesticide application.

Golf courses A, B, C, D, and E were in full operation throughout the study. Test site D upstream 4 (DU4) was an exception, as site DU4 was located on the part of the property with plans of expansion. Clearing and burning of trees and vegetation was done in 1999, however the expansion was never completed during the study. Golf course F was cleared of trees, and topography was reshaped in 1999, turf was established in the fall of 2000, and by the summer of 2001 the course was opened to the public. Golf course G underwent clearing of trees and topography was reshaped in the spring and summer of 1999. Turf was established in the fall of 1999, and by spring of 2000 the golf course was opened to the public. The property of golf course H was partially cleared of trees and topography was reshaped during the study period, but the development of the golf course was never completed.

Study design

The RCA involves sampling a set of reference sites that are as similar to the test sites as possible, with the exception of the effects due to the treatment in question. In this study the effect of various stages of golf course construction on stream BMI communities was tested. Streams were sampled three times a year, in spring (April to June 20), summer (June 21 to September 22) and fall (September 23 to November 13) for three years, 1999 to 2001. Samples were not collected in winter as the organisms were considered to be immature at this time of year and therefore very small making it difficult to catch and identify a representative sample. Samples for BMI and water chemistry at the test sites were taken usually within a few days of the reference sites, although there was an interval of up to several weeks during the summers when some streams had low or no flow and could not be sampled. During the study period, water samples for chemical analysis were collected weekly at sites that were downstream from golf courses under construction and monthly at reference sites and established golf course sites. Averages for each season (i.e. spring, summer, and fall) were calculated for each water chemistry parameter to match the sampling period of BMI data.

Physical and geographic attributes

The latitude and longitude for each site was recorded in the field using a hand held Global Positioning System (GPS) accurate within 5 m. The dry bank width (i.e. width of the channel perpendicular to the two banks) was estimated from the dry line observed in the field. Stream order, site elevation, and catchment area were calculated using Ontario Base Maps (OBM) at a 1:10,000 scale. Stream order was based on the stream order nomenclature proposed by Strahler (1957) and described in Horne and Goldman (1994). Site elevation and catchment area for the sampling site were estimated with the OBMs. Beaver dams and the sewage treatment plant were noted when observed directly upstream in the field.

Stream width for test and reference sites ranged from 0.5 m to 4 m, representing first to third order streams, elevations from 228 m to 377 m asl, and catchment areas from 8 ha to 805 ha (**Table 1**). Two test sites were downstream from a sewage treatment plant. Several reference and test sites had a beaver dam directly above the site.

In the summer of 2001, a survey was conducted at each reference and test site to collect qualitative data on a 30 by 30 m riparian buffer strip and the stream channel. The quality of habitat represented by a set of variables included in the survey assisted in understanding the changes in community dynamics that was observed in BMI bioassessment. Information on slope of the stream channel, presence of bank erosion, grasses, and woody debris, canopy closure (cover), and general forest type (see **Table 2**) was recorded using methods modified from Moore *et al*. (1995). Canopy cover was measured with a densiometer taking 4 readings, i.e. facing north, east, south, and west (Lemmon 1957). Readings were averaged for the replicate site, converted to percent coverage, and coded in the range of 1 to 5 where 1 is very poor canopy cover (less than 20%) and 5 is excellent canopy cover (greater than 80%). The three replicate sites were averaged to get one class value that described canopy cover at each site.

The sites are described as having a slope of the stream channel ranging from 1% to 40%. Bank erosion was observed at only 3 test sites and none of the reference sites. Five test sites and none of the reference sites had grasses growing in the stream channel. Three test sites and one reference site had no woody debris or log dams in the stream. Test site values averaged 40% to 59% (class 3) canopy cover/closure, while the reference group averaged over 80% (class 5). In the first 30 meters parallel to the stream, forest type for all sites was generally deciduous or mixed.

Benthic macroinvertebrates

The BMI community was sampled using protocols for rapid bioassessments of streams developed by David *et al*. (1998) at the Ontario Ministry of the Environment, Canada (David *et al*. 1998). This technique differs from the more traditional biological assessments in that only 100 animals are sub-sampled randomly from the samples and these are identified to the taxonomic level of order, and in some cases to sub-order, class, family or only phylum, depending on the invertebrate (Somers *et al*. 1998).

In the field, the collection of water samples for chemical analysis was done prior to invertebrate sampling to minimize stream and sample disturbances. Replicate samples were taken from three shallow riffles within the stream reach, no more then 1-meter deep using the kick-and-sweep method and a 250-µm mesh D-net (David *et al*. 1998). Each replicate was sampled from a 1-m² quadrat area for 1 minute or until at least 100 organisms were caught. For each replicate, the material and organisms were placed in a labeled container that was partially filled with stream water and transported back to the laboratory for identification and enumeration.

In the laboratory, the samples were refrigerated and analyzed within 48 hours of collection. The contents of each bottle were first rinsed through a 1000-µm mesh sieve. A teaspoon method was used to randomly collect some of the material and placed into a sorting tray where all organisms were removed, enumerated, and identified to a coarse taxonomic level (David *et al*. 1998). Subsequent subsamples were taken for analysis until at least 100 specimens were counted. For statistical purposes, counts for each taxonomic group were expressed as relative abundance (proportional data) and then the 3 replicates were averaged.

A variety of summary indices were calculated from the BMI taxonomic abundances (**Table 3**); e.g. David *et al*. (1998) and Somers *et al*. (1998). The benthic indices were: i) Amphipoda, Chironomidae, Diptera, non-Diptera (abundance of all taxa groups

Table 1 The physical and geographic attributes of the test and reference sites. Construction land use category represents both clear-cut and turf establishment phases, unless otherwise stated.

Site Code	Land Use Activity	Latitude *	Longitude *	Width (m)	Order	Site Elevation (m)	Catchment Area (Ha)	Upstream Activity **
Test site, n=19								
DU1	established	45.1067	-79.2761	3	2	280	255	–
D	established	45.1097	-79.2797	1	3	277	805	–
B1	established	45.3425	-79.1531	0.5	1	305	34	–
B2	established	45.3608	-79.1553	0.5	1	300	115	–
AU	established	45.2175	-79.4603	1	1	300	8	Dam
A	established	45.2094	-79.4539	0.5	1	293	16	–
H	clear-cut	45.3467	-79.0339	0.75	1	334	71	–
F5	construction	45.3444	-79.1703	3	3	293	368	STP
F1	construction	45.3497	-79.1606	1	1	313	39	–
F6	construction	45.3500	-79.1906	1	2	305	92	STP
C	construction	45.3489	-79.1558	0.5	1	285	25	–
F2	construction	45.3533	-79.1683	2.5	3	301	306	–
F7	construction	45.3586	-79.1686	1	2	330	167	Dam
F3	construction	45.3578	-79.1711	2.5	2	323	159	–
F4	construction	45.3561	-79.1711	1	2	323	229	Dam
F8	construction	45.3489	-79.1778	2.5	2	313	266	–
E1	established	45.1322	-79.6064	0.8	1	234	9	–
E2	established	45.1325	-79.6103	0.5	1	228	18	–
G1	turf establishment	45.2233	-79.7797	1.5	2	242	120	–
Reference Site, n=19								
DU4	reference	45.1056	-79.2667	1	2	282	200	–
DU3	reference	45.1050	-79.2764	0.8	1	282	38	–
S5	reference	45.2661	-79.0856	3	2	317	222	Dam
S6	reference	45.2117	-78.9931	1	2	374	60	–
S9	reference	45.1506	-79.1050	2	2	359	30	–
S6	reference	45.1461	-79.0869	2	2	360	22	–
HU	reference	45.3703	-79.0283	1.5	1	333	582	–
S15	reference	45.2542	-79.0869	4	3	317	742	Dam
S25	reference	45.3744	-79.1406	2	1	328	26	–
S26	reference	45.3803	-79.1539	2	2	330	20	Dam
S27	reference	45.3842	-79.1514	2	2	330	119	Dam
S28	reference	45.3861	-79.1322	2	1	377	10	–
S31	reference	45.3114	-79.0928	1.5	1	320	135	–
S32	reference	45.2572	-79.0711	2	1	327	38	–
S33	reference	45.1983	-78.9628	2	1	349	70	–
S34	reference	45.1992	-78.9597	2	1	351	45	Dam
S36	reference	45.2792	-79.4661	1	1	295	73	–
S37	reference	45.2286	-78.9275	4	3	320	418	Dam
S38	reference	45.2342	-78.9528	2	2	318	502	Dam

* decimal degrees, ** no upstream features (-), beaver dam (Dam), and sewage treatment plant (STP)

with the exception of Diptera), Gastropoda, Odonata, Oligochaeta, and Plecoptera expressed as proportions by dividing by total abundance; ii) Oligochaeta and Nematoda combined as a proportion of total abundance (Worms) (David *et al*. 1998); iii) total abundance of Ephemeroptera, Plecoptera, and Trichoptera (EPT) (Plafkin *et al*. 1989); iv) the ratio of EPT to Chironomidae abundance (EPT/Chironomidae) (Plafkin *et al*. 1989); v) total number of different taxonomic groups in a sample (Richness); vi) abundance of most abundant taxon as a proportion of total abundance (Dominants) (Plafkin *et al*. 1989). Rescaling EPT, EPT/Chironomidae, and Richness by dividing each index value by their corresponding maximum value across all samples, gave equal weight to each index by expressing all values from 0 to 1.0 in the indices.

The benthic indices in **Table 3** were used to evaluate the BMI communities in various ways (Barbour *et al*. 1999). The first 11 indices, with the exception of Dominants, measure community composition because they measure the contribution by some tolerant or intolerant group to the sample population (proportion or percent). The ratio of EPT/Chironomidae measures community evenness (balance) by comparing the ratio of intolerant taxa (i.e. Ephemeroptera, Plecoptera. and Trichoptera) to a more tolerant taxonomic group (i.e. Chironomidae). Richness is a measure of community diversity and is one of the most commonly used indices (Lenat and Barbour 1994).

DATA ANALYSIS

Physical and geographic attributes

Using the RCA it is assumed that the test and reference stream sites are similar with respect to stream order or permanency, except for the effects of the specific treatment or event being assessed, i.e. golf course construction. However, this is not always true; BMI communities may be influenced by differences in physical and geographic attributes of the randomly selected sites (Charvet *et al*. 2000). To identify which physical and geographic stream characteristics differed between the reference group sites and the test group sites a two-sample *t*-test was run using STATISTICA (StatSoft 2000). Data assumptions for normality and homogeneity of variance were investigated prior to the *t*-test using Shapiro-Wilk's W test of normality and Levene's test of homogeneity of variances, respectively in STATISTICA (Manly 1994; Zar 1999). If the variables were not normally distributed, the data were transformed with the appropriate transformation to approximate normality (Zar 1999). The physical-geographic variables that had significantly different means among the reference and test site groups were used to create a distance matrix for Mantel tests and partial Mantel tests.

BMI community measure evaluation

Numerous distance and similarity measures are available to choose that summarize BMI communities (Legendre and Legendre 2000).

Table 2 Site description of buffer strip attributes from the riparian bank survey that was conducted in summer of 2001. DF = deciduous forest, Mix = mix of deciduous and coniferous in the buffer strip.

Site Code	Slope of Channel (%)*	Bank Erosion	Grass	Debris/ Log dam	Canopy Closure Class**	Forest Type
Test site, n=19						
DU1	1	No	Yes	Yes	1	Mix
D	1	No	Yes	No	1	DF
B1	5	Yes	No	Yes	4	DF
B2	1	No	Yes	No	3	Mix
AU	1	No	No	Yes	5	DF
A	1	No	No	Yes	4	Mix
H	1	No	No	Yes	4	Mix
F5	1	No	No	Yes	1	DF
F1	1	No	No	Yes	2	DF
F6	1	No	Yes	Yes	3	Mix
C	1	No	Yes	Yes	3	DF
F2	5	No	No	Yes	3	DF
F7	5	Yes	No	Yes	5	Mix
F3	1	Yes	No	Yes	1	none
F4	1	No	No	Yes	2	Mix
F8	5	No	No	Yes	5	Mix
E1	1	No	Yes	No	3	Mix
E2	1	No	No	Yes	5	Mix
G1	40	No	No	Yes	5	Mix
Reference Site, n=19						
DU4	1	No	No	No	1	DF
DU3	1	No	No	Yes	5	Mix
S5	10	No	No	Yes	5	Mix
S6	1	No	No	Yes	5	Mix
S9	1	No	No	Yes	4	Mix
S6	5	No	No	Yes	5	Mix
HU	1	No	No	Yes	4	Mix
S15	5	No	No	Yes	5	Mix
S25	28	No	No	Yes	5	DF
S26	10	No	No	Yes	5	Mix
S27	5	No	No	Yes	5	Mix
S28	30	No	No	Yes	5	DF
S31	1	No	No	Yes	5	Mix
S32	35	No	No	Yes	5	DF
S33	40	No	No	Yes	5	Mix
S34	40	No	No	Yes	5	DF
S36	1	No	No	Yes	5	DF
S37	1	No	No	Yes	5	Mix
S38	1	No	No	Yes	5	Mix

* average %slope of the 3 transect zones along the stream bank, ** average values from the 3 replicate stations at each site where 1 is less than 20% canopy cover and 5 is greater than 80% canopy cover.

Bray-Curtis is a popular measure that summarizes the percent difference between pairs of samples (Bray and Curtis 1957; Faith *et al*. 1987). Two Bray-Curtis distance measures were calculated: the first was the traditional Bray-Curtis distance among sites using the abundances for BMI taxa. The second Bray-Curtis distance was based on a collection of summary benthic indices. The Jaccard similarity also was used to summarize similarity among sites based on the presence and absence of various BMI taxa (Manly 1994; van Tongeren 1995; Legendre and Legendre 2000). However, the Jaccard similarity was converted into a distance measure, d, using the following formula (Rohlf 1993):

$$d = \sqrt{1 - s}$$

where d is distance, and s is similarity.

The fourth distance matrix determined from the BMI taxa abundance data was chosen because it is similar to the matrix used in Correspondence Analysis, a popular ordination method used in ecological community analysis (ter Braak 1995).

Mantel tests were used to relate each of these four BMI community distance matrices to environmental factors associated with land use categories and physical-geographic attributes (**Table 1**). The Mantel test is a type of non-parametric multivariate analysis of variance (MANOVA) that contrasts two distance or similarity matrices (Manly 1994; Legendre and Legendre 2000). One matrix is based on a series of response variables (a type of X matrix) and the other is based on a series of predictor variables (Y matrix). Sometimes a third matrix based on covariables (predictor variables) is used in a partial Mantel test to hold constant the differences in the response variables that were associated with the covariables. A hypothesis matrix is the predictor matrix that is correlated with the response matrix. In this study, three types of distance matrices were correlated with each other. These distance matrices were based on a BMI response matrix, the land use predictor matrix, and the physical-geographic predictor matrix (**Fig. 2**).

Since the Mantel test association is being measured between two distance matrices only positive Mantel types of correlation results were considered important and negative correlations were ignored (i.e. one-tailed test) (Legendre and Legendre 2000). The null hypothesis for the Mantel test states that the distances among sites (in this case samples from sites) based on BMI communities are not correlated with the corresponding inter-site distances based on (1) land use distance matrix or the (2) physical-geographic distance matrix. A correlation coefficient (r) value is calculated for each combination of distance matrices. A positive coefficient states that when the distances increase among sites in the X-matrix then the distances also increase among sites in the Y-matrix. The greater the increase, the more BMI differences are correlated with differences in land use or physical-geographic variables.

The significance of the correlation was evaluated using the randomization (permutation) test to evaluate whether the observed correlation is greater than what would be expected due to chance alone. If the observed correlation coefficient is greater than 95% of the permuted correlations then the null hypothesis is rejected. In the statistical package PASSAGE (Rosenberg 2001), the one-tailed Student *t*-test was used to test for the significance of the Mantel correlation.

After calculating the BMI distance matrices (and before conducting the Mantel tests), Non-metric Multidimensional Scaling (NMDS) in NTSYS-pc was run to ordinate each of the BMI distance matrices and reduce the number of variables to three sets of ordination scores (Rohlf 1993). NMDS, a non-parametric ordination analytical method, is a distance-based ordination that maxi-

Table 3 A list of benthic indices and definitions.

Benthic Indices	Description
Amphipoda	Amphipoda abundance / total sample abundance
Chironomidae	Chironomidae abundance / total sample abundance
Diptera	Diptera abundance / total sample abundance
Gastropoda	Gastropoda abundance / total sample abundance
non-Diptera	non-Diptera abundance / total sample abundance
Odonata	Odonata abundance / total sample abundance
Oligochaeta	Oligochaeta abundance / total sample abundance
Plecoptera	Plecoptera abundance / total sample abundance
Worms	Oligochaeta and Nematoda abundances / total sample abundance
Dominants	Abundance of most abundant taxon / total sample abundance
EPT	Abundance Ephemeroptera, Plecoptera, and Trichoptera
EPT/Chironomidae	EPT abundance / Chironomidae abundance (Plafkin *et al*. 1989)
Richness	total number of different taxa groups in sample

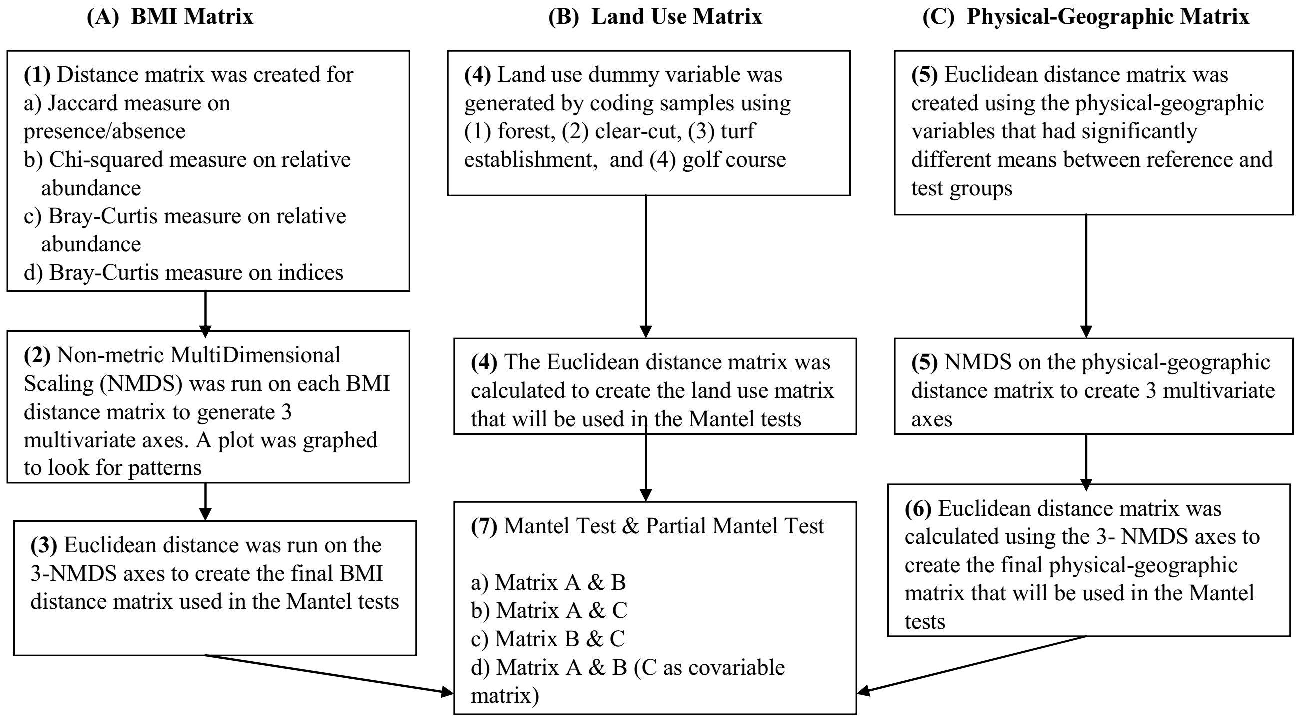

Fig. 2 Framework for the Mantel and partial Mantel tests which were used to select the benthic macroinvertebrate (BMI) distance measure that best detected BMI community differences between land use categories. The numbers in brackets on the top left corner of the boxes refer to the steps.

mizes rank-order correlation between original distance matrix and distance in ordination space by minimizing stress (Manly 1994; Legendre and Legendre 2000). Stress is a measure that describes the lack of fit between the original (e.g. community data) and final configurations in ordination space. Scores from the three NMDS axes for each of the BMI distance matrices were used to calculate a Euclidean distance matrix that was used in the Mantel tests.

Before running Mantel tests two predictor matrices were created. A 'land use' dummy variable was formed by coding the samples using: (1) reference, (2) clear-cut phase of golf course construction, (3) turf establishment, or (4) golf course operation categories. NMDS was not used to ordinate the land use variable, as the variable was already made-up of ranked datum. Using Microsoft Excel (PopTools add-in), a Euclidean distance measure was used to calculate the land use predictor distance matrix for use in the Mantel tests (Hood 2000).

The second predictor matrix was then calculated using the Euclidean distance measure on the physical and geographic variables that were found to have significantly different means between the reference sites and test sites (see section C in **Fig. 2**). To be consistent with the BMI distance matrices, NMDS was used to ordinate the physical-geographic Euclidean distance matrix to create 3 sets of ranked scores in NTSYS-pc. A final physical-geographic matrix was calculated using the Euclidean distance measure for the second time, but based on the NMDS axes scores. The Mantel test contrasted a BMI Euclidean distance matrix with the land use Euclidean distance matrix, and then a partial Mantel test held constant the differences in BMI distances associated with the physical-geographic Euclidean distance matrix. Where more than one test was performed on the same variable the p level of significance was corrected for pairwise tests using the Bonferroni procedure, because more than one partial Mantel test was calculated using the same variables (Zar 1999). The Bonferroni correction method (i.e. p-level of 0.05 divided by 6 pairwise tests = 0.008) reduces the possibility of falsely rejecting at least one $H_{0,}$ when it is true (Type I error).

The best BMI distance measure was selected using five criteria. Each criterion was considered equally important. The BMI distance measures were ranked in terms of their performance. If ties occurred, BMI measures carried the same rank. For all criteria, the ranks were then summed to retrieve a score. The BMI distance measure with the greatest score was declared the best for use in bioassessments.

NMDS criteria

The first criterion was designed to identify the distance matrix with the greatest percentage of BMI taxa (abundance or presence/absence) or benthic indices that were correlated with the first 3 axes of the corresponding NMDS ordination. This criterion is based on the amount of information retained in the NMDS scores prior to calculating the final Euclidean distance matrix for the Mantel tests. Pearson correlations were calculated between the scores for each NMDS axis and BMI abundances, presence/absence and benthic indices to identify the important taxa or indices represented by the first 3 axes (r=±0.40). Assumptions of normality and homogeneity of variance were tested prior to calculating the correlations using Shapiro-Wilk's W test of normality and Levene's test of homogeneity of variances where data were not normally distributed the appropriate transformations were applied.

The final NMDS solution is heavily dependent upon the number of axes specified before running the ordination. Any one NMDS axis is not necessarily more important than another. A score of 4 was given to the BMI matrix for the criterion with the highest percentage of BMI variables that strongly correlated with the three NMDS axes. The rest of the BMI measures were ranked accordingly.

The second criterion identified the distance measure that produced the greatest separation between land use groups based on group mean scores of the three NMDS axes. First, the mean NMDS score for each land use group of each NMDS axes was calculated. A Euclidean distance matrix among the four groups of mean scores for the three NMDS axes was generated. The average distance of the pairwise Euclidean distances was calculated indicating strength of group separation, with a range from 0 (no group separation) to 1 (perfect group separation). The BMI distance measure that had the greatest average Euclidean distance was given a score of 4 for this criterion.

Temporal variability criterion

Temporal effects have been observed in many studies that use abundance and benthic indices in bioassessments and so temporal variability was considered as the third criterion because BMI were collected in three seasons of three years (Lenat 1987; Hilsenhoff 1988; McElravy *et al.* 1989). To test whether the effects of seasonal and annual variability were significant on BMI distance matrices, a two-factor MANOVA was performed on the 3 NMDS axes for each of the BMI distance matrices using season (spring, summer, and fall) and year (1999, 2000, 2001) as the two factors with three levels within each factor. The dependent variables are the first three NMDS axes (see step 2 in **Fig. 2**) for the BMI distance matrix that contained the ranked data from the ordination method. Data assumptions no longer apply when ranked data are used therefore the MANOVA with NMDS axes becomes a non-parametric test (Conover and Iman 1981). The BMI distance matrix that had no seasonal and annual effect was given a score of 4 for this criterion.

Mantel test criteria

The fourth criterion identifies the BMI distance measure that produced the greatest number of significant Mantel tests and partial Mantel tests ($p<0.05$) between the BMI matrix and the land use matrix. For each BMI distance, the number of significant tests (positive r) was counted for all tests. The BMI measure with the greatest number (maximum = 14) was assigned a score of 1 for this criterion.

The fifth criterion identifies the BMI distance measure that was considered minimally affected by differences in physical-geographic stream attributes. It was estimated using the results of the Mantel tests that correlated between a BMI matrix and the physical-geographic matrix. The number of Mantel tests with no significant correlation between the BMI distance matrix and the physical-geographic matrix was recorded for each BMI measure. A score of 1 was given to the BMI measure with the highest frequency (maximum = 7) and a zero were given to all others.

RESULTS

Physical and geographic attributes

Three of the six physical-geographic variables were found to have significantly different means between the reference group sites and test group sites (**Table 4**). These variables were longitude, stream width, and site elevation. These three physical-geographic variables were used to create one of the distance matrices for the Mantel tests.

Macroinvertebrate community measure evaluation

According to the five criteria, Bray-Curtis based on benthic indices was the best overall BMI distance measure (**Table 5**). The Bray-Curtis measure had the highest percentage of benthic indices that were strongly correlated with the first 3 NMDS axes ($r>\pm0.40$), it was not affected by temporal variability, and it had the greatest number of Mantel tests where correlations between BMI and physical-geographic distances were not significant. However, Bray-Curtis based on indices ranked second last in criterion 2 as the average distance of Euclidean distances among land use group means was the second shortest (criterion 2). It also ranked second highest in having the greatest number of Mantel and partial Mantel tests with significant correlations between BMI and land use distances (criterion 4).

Jaccard distance measure based on presence and absence BMI data had different strengths than Bray-Curtis indices measure. Jaccard distance measure had the greatest Euclidean distance average among land use group means (criterion 2), and it had the greatest number of Mantel and partial Mantel tests with significant correlations between BMI and land use distances (criterion 4). Jaccard, Bray-Curtis abundance, and Bray-Curtis based on benthic indices were the best BMI summaries to distinguish BMI communities between the four land use categories (criterion 4 in **Table 5**). Jaccard had the greatest average Euclidean distance between land use group means based on NMDS scores. Similar to Bray-Curtis indices measure, chi-square distance based on abundance also had no significant seasonal, annual or interaction effects. Bray-Curtis distance based on BMI abundance did not rank first in any of the 5 criteria.

NMDS criteria

The first criterion involved evaluating the four measures on their ability to retain the BMI information in the 3 NMDS axes scores. For the Jaccard measure, 12 of 25 presence/absence taxonomic groups were correlated with the first three NMDS axes (**Table 6**). Chi-square distance had 7 of 25 taxonomic groups of abundance correlated with the three NMDS axes (**Table 7**). For the Bray-Curtis distance based on abundance data, 10 of 25 BMI abundance taxonomic groups were correlated with the first three NMDS axes (**Table 8**). Bray-Curtis distance based on BMI indices had 9 of 13 indices correlated with the three NMDS axes (**Table 9**).

For the second criteria, the quality of land use group separation for each BMI distance measure was estimated using the average distance of Euclidean distances among land use group means based on NMDS axes scores. In the

Table 4 Basic statistics and *t*-tests comparing the means of physical-geographic variables between the reference group and test group (n=38 df=36). P values in bold indicate a 95% confidence level ($p<0.05$). n = number of samples, SE = standard error.

Variable	Reference group			Test group			*t*-test	
	n	mean	SE	n	mean	SE	t value	p value
Latitude *	69	45.2630	0.01	127	45.2762	0.01	-0.905	0.37
Longitude *	69	-79.0870	0.01	127	-79.2727	0.02	7.23	<0.001
Stream width	69	2.01	0.11	127	1.26	0.07	5.88	<0.001
Stream order	69	1.59	0.08	127	1.70	0.06	-1.01	0.32
Elevation	69	333.1	3.06	127	293.7	2.48	9.72	<0.001
Catchment area	69	205.5	28.34	127	167.7	16.53	1.24	0.22

* decimal degrees

Table 5 Summary of results for the criteria used to evaluate the best Benthic Macroinvertebrate (BMI) community measure. The best score for each criterion is in bold. NMDS = Non-metric Multidimensional Scaling.

	Criteria		Temporal variability	Mantel tests		
	NMDS axes					
Distance Measure	% of BMI taxa or indices that were strongly correlated with the three NMDS axes	Average distance of Euclidean distances among land use group means based on NMDS axes scores	What type of temporal effect?	# of Mantel tests and partial Mantel tests that were significant between BMI and land use matrices	# of Mantel tests that were not significant between BMI and physical-geographic matrices	Score *
Jaccard	48%	0.75	interaction	9	1	2
Chi-square	28%	0.49	none	6	3	2
Bray-Curtis abundance	40%	0.58	seasonal	9	1	1
Bray-Curtis indices	69%	0.51	none	9	3	4

* sum score of ranks in descending order

Table 6 Correlation coefficients (r) between the presence/absence of macroinvertebrate taxa and the first 3 axes of the Non-metric Multidimensional Scaling (NMDS) on the Jaccard distance matrix, n=196. Variables with a bold r are highly correlated (r>+/-0.40) with the axis.

Taxa	NMDS1	NMDS2	NMDS3
Amphipoda (Order)	0.06	0.49	0.47
Anisoptera (Sub Order)	0.28	-0.01	-0.19
Bivalvia (Class)	0.35	0.38	0.07
Ceratopogonidae (Family)	-0.23	-0.13	-0.51
Chironomidae (Family)	-0.32	-0.27	-0.24
Coleoptera (Order)	-0.32	0.25	-0.11
Culicidae (Family)	-0.21	-0.11	-0.15
Decapoda (Sub Order)	0.14	0.01	-0.37
Ephemeroptera (Order)	0.36	-0.26	-0.27
Gastropoda (Class)	0.22	0.46	0.29
Hemiptera (Order)	0.23	0.28	0.13
Hirudinea (Class)	0.32	0.48	0.16
Isopoda (Order)	0.10	0.48	0.41
Lepidoptera (Order)	-0.05	0.17	-0.05
Megaloptera (Order)	0.36	-0.11	-0.08
Nematoda (Phylum)	-0.43	0.05	-0.09
Oligochaeta (Class)	0.06	-0.04	0.16
Plecoptera (Order)	-0.45	-0.48	-0.43
Simuliidae (Family)	-0.43	-0.05	-0.18
Tabanidae (Family)	-0.31	-0.18	-0.16
Tipulidae (Family)	-0.14	-0.48	-0.08
Trichoptera (Order)	-0.20	-0.39	-0.17
Trombidiformes (Sub Order)	-0.06	0.42	-0.57
Turbellaria (Class)	0.22	0.29	-0.05
Zygoptera (Sub Order)	0.54	0.42	-0.16

Table 7 Correlation coefficients between macroinvertebrate taxa (abundance) and axes from Non-metric Multidimensional Scaling (NMDS) on chi-square distance (as in Correspondence Analysis) matrix for macroinvertebrate abundance, (n=196). Variables with r in bold are highly correlated (r>+/-0.40) with the axis.

Taxa	NMDS1	NMDS2	NMDS3
Amphipoda (Order)	0.51	-0.02	0.17
Anisoptera (Sub Order)	-0.02	-0.02	0.01
Bivalvia (Class)	-0.02	0.21	0.29
Ceratopogonidae (Family)	-0.28	-0.18	0.07
Chironomidae (Family)	-0.03	-0.05	-0.16
Coleoptera (Order)	0.15	-0.07	0.16
Culicidae (Family)	-0.06	0.00	-0.29
Decapoda (Sub Order)	-0.10	0.06	0.05
Ephemeroptera (Order)	-0.43	-0.15	-0.08
Gastropoda (Class)	0.28	0.15	0.31
Hemiptera (Order)	0.14	0.10	0.23
Hirudinea (Class)	0.16	0.41	0.15
Isopoda (Order)	0.55	0.47	0.17
Lepidoptera (Order)	-0.04	0.13	0.02
Megaloptera (Order)	-0.06	-0.18	0.08
Nematoda (Phylum)	-0.08	-0.08	-0.20
Oligochaeta (Class)	0.18	-0.05	0.24
Plecoptera (Order)	-0.41	-0.34	-0.23
Simuliidae (Family)	-0.14	0.01	-0.58
Tabanidae (Family)	-0.11	-0.20	-0.06
Tipulidae (Family)	-0.32	-0.25	0.00
Trichoptera (Order)	-0.58	-0.08	-0.01
Trombidiformes (Sub Order)	-0.02	0.02	0.13
Turbellaria (Class)	0.05	0.22	0.23
Zygoptera (Sub Order)	0.04	0.26	0.28

Table 8 Correlation coefficients between the taxa abundances and axes for Non-metric Multidimensional Scaling (NMDS) that was based on Bray-Curtis distance on macroinvertebrate taxa abundance, n=196. Variables with r in bold are highly correlated (r>+/-0.40) with the axis.

Taxa	NMDS1	NMDS2	NMDS3
Amphipoda (Order)	0.40	0.47	0.26
Anisoptera (Sub Order)	-0.07	0.17	-0.03
Bivalvia (Class)	-0.11	0.02	0.51
Ceratopogonidae (Family)	-0.20	-0.18	0.03
Chironomidae (Family)	0.46	-0.52	0.12
Coleoptera (Order)	0.09	0.03	0.18
Culicidae (Family)	0.13	-0.21	-0.12
Decapoda (Sub Order)	-0.12	0.01	-0.10
Ephemeroptera (Order)	-0.50	0.29	-0.46
Gastropoda (Class)	0.05	0.37	0.25
Hemiptera (Order)	0.11	0.28	0.16
Hirudinea (Class)	0.19	0.23	0.27
Isopoda (Order)	0.48	0.39	0.25
Lepidoptera (Order)	-0.02	0.13	-0.02
Megaloptera (Order)	-0.19	0.22	-0.08
Nematoda (Phylum)	0.13	-0.23	-0.08
Oligochaeta (Class)	0.22	0.11	0.42
Plecoptera (Order)	-0.32	-0.58	-0.39
Simuliidae (Family)	0.25	-0.37	-0.45
Tabanidae (Family)	-0.12	-0.25	0.04
Tipulidae (Family)	-0.37	-0.34	0.01
Trichoptera (Order)	-0.68	-0.11	-0.16
Trombidiformes (Sub Order)	-0.06	0.07	0.04
Turbellaria (Class)	-0.05	0.21	0.04
Zygoptera (Sub Order)	-0.08	0.40	0.13

Table 9 Correlation coefficients between the benthic indices and Non-metric Multidimensional Scaling (NMDS) axes that was based on Bray-Curtis distance matrix on benthic indices, n=196. Variables with bold r are highly correlated (r>+/-0.40) with the axis.

Benthic Index	NMDS1	NMDS2	NMDS3
Chironomidae	0.71	-0.64	0.43
Diptera	0.65	-0.88	0.45
non-Diptera	-0.65	0.88	-0.45
Amphipoda	0.28	0.34	0.23
Gastropoda	-0.10	0.35	-0.08
Odonata	-0.04	0.15	-0.24
Oligochaeta	0.53	0.29	-0.02
Plecoptera	-0.26	-0.30	-0.40
Dominants	-0.32	0.03	0.61
Worms	0.55	0.25	-0.01
Richness	0.25	0.05	-0.33
EPT/Chironomidae	-0.64	0.13	-0.69
EPT	-0.64	-0.07	-0.69

Table 10 Euclidean distance matrix of the Non-metric Multidimensional Scaling (NMDS) group means for Jaccard distance measure based on presence and absence BMI data.

	Forest	Clear-cut	Turf establishment	Golf course
Forest	0	0.57	1.17	0.70
Clear-cut	0.57	0	0.97	0.27
Turf establishment	1.17	0.97	0	0.80
Golf course	0.70	0.27	0.80	0
Mean	0.75			

NMDS group mean Euclidian matrices, the lower triangle of the Euclidiean distance matrix was averaged (**Tables 10, 11, 12, 13**). The order, from greatest average distance to least, was Jaccard, Bray-Curtis based on abundance, Bray-Curtis based on indices, and chi-square (see **Table 5**).

Temporal variability

The influence of temporal variability, i.e. seasons and years on NMDS axes 1, 2, and 3 scores from the 4 BMI distance matrices, was tested using a two-factor MANOVA. The results show that there was no significant difference in means of the 3 NMDS axes among seasons (spring, summer, and fall), and among years (1999, 2000, 2001) for Jaccard, chi-square, and Bray-Curtis indices (n=196, $p<0.05$). Bray-Curtis abundance was the only measure with significant difference in means of the 3 NMDS axes among seasons. Jaccard was the only measure with a significant interaction effect between seasons and years.

Table 11 Euclidean distance matrix of the Non-metric Multidimensional Scaling (NMDS) group means for chi-square distance measure based on BMI abundance.

	Forest	Clear-cut	Turf establishment	Golf course
Forest	0	0.37	0.57	0.66
Clear-cut	0.37	0	0.39	0.38
Turf establishment	0.57	0.39	0	0.59
Golf course	0.66	0.38	0.59	0
Mean	0.49			

Table 12 Euclidean distance matrix of the Non-metric Multidimensional Scaling (NMDS) group means for Bray-Curtis distance measure based on BMI abundance.

	Forest	Clear-cut	Turf establishment	Golf course
Forest	0	0.54	0.60	0.71
Clear-cut	0.54	0	0.64	0.43
Turf establishment	0.60	0.64	0	0.54
Golf course	0.71	0.43	0.54	0
Mean	0.58			

Table 13 Euclidean distance matrix of the Non-metric Multidimensional Scaling (NMDS) group means for Bray-Curtis distance measure based on benthic indices.

	Forest	Clear-cut	Turf establishment	Golf course
Forest	0	0.57	0.32	0.58
Clear-cut	0.57	0	0.63	0.65
Turf establishment	0.32	0.63	0	0.29
Golf course	0.58	0.65	0.29	0
Mean	0.51			

Mantel tests

Mantel test results indicated that Jaccard distance (**Table 14**), chi-square distance on abundance (**Table 15**), Bray-Curtis distance on abundances (**Table 16**), and Bray-Curtis distance on indices (**Table 17**) were significantly correlated with the four land use categories ($p<0.05$). When partitioning out differences associated with physical-geographic attributes, chi-square based on BMI abundance was the only BMI measure with a non-significant partial Mantel. Test sites among the chi-square BMI matrix were not significantly correlated with corresponding distances between the four land use categories when differences in longitude and elevation were controlled. For pairwise tests, all four BMI distance measures had significant positive correlations (Bonferroni corrected $p<0.008$) with land use distances between land uses 1 and 2 and between 1 and 4 (1-forest, 2-clear-cut, 4- golf course in operation), with and without controlling for differences in physical-geographic variables.

The pairwise correlations between each BMI matrix and the physical-geographic matrix indicated that distances in BMI between samples were significantly correlated with the corresponding distances in physical-geographic variables for two of the BMI distance measures. These pairwise Mantel tests were also not significant for every land use scenario. Both Jaccard presence/absence distances and Bray-Curtis abundance distances among sites were significantly correlated with corresponding distances in physical-geographic variables in the pairwise Mantel tests when testing between forest and golf course operational land use categories. In addition, Bray-Curtis abundance distances also were significantly correlated with corresponding distances in physical-geographic variables when testing between forest and turf-establishment land use categories. There was no significant correlation between chi-square abundance and Bray-Curtis indices distances and corresponding distances of the physical-geographic variables in the pairwise tests.

DISCUSSION AND CONCLUSION

The rationale for bioassessment includes the capacity for biological communities to respond to and thus reveal up-stream sources of pollutants post discharge that may not be measured during routine chemical analysis (Bonada *et al.* 2006). Biota integrate environmental conditions, including short-term changes in water quality that may be missed by intermittent or routine chemical sampling. Lewis *et al.* (2001) concluded that the impact of golf course runoff in particular on sediment quality can be subtle and may require the use of biological assessment methods. While observed changes in water chemistry can provide direct evidence of water quality impacts, biological assessments are often used to indicate the relevance of water chemistry changes to the aquatic community. Since the 1970s, ecosystem health *per se* has increasingly been seen as a valuable water quality objective (*sensu* Bonada *et al.* 2006). Benthic macroinvertebrates are the most widely used organisms in freshwater biomonitoring, and a large number of different methods have been developed and are used in various countries or regions of the world (e.g. Walsh *et al*. 2001; Clarke *et al*. 2003; Hering *et al*. 2003; Bonada *et al*. 2006; Carlisle *et al*. 2007).

Norris (1995) stated that including a variety of indices in multivariate approaches might yield powerful predictive tools when studying ecosystem function and structure. In this study the predictive power of four sets of common summary measures using NMDS and Mantel tests were compared. The four BMI distances matrices, calculated using Jaccard, chi-square, Bray-Curtis based on abundance, and Bray-Curtis based on benthic indices measures, generally, discriminated among the four land use categories. Distances of BMI measured by Jaccard, Bray-Curtis abundance and Bray-Curtis benthic indices were significantly correlated with the corresponding distances of the four land use groups. Similar to the other three distance measures, chi-square on abundances was significantly (linearly) correlated with the land use distance matrix with 4 land use groups. However, the correlation was not significant when physical-geographic (longitude and elevation) differences were controlled, as they were in the partial Mantel test. This means that with respect to being sensitive to differences in longitude and elevation between sites, the chi-square distance measure was not very powerful.

Evaluation of the best BMI measure revealed that Bray-Curtis distance between benthic indices scored the highest according to five criteria: 1) the percentage of BMI taxonomic groups or benthic indices that were highly correlated with the three NMDS axes (Pearson correlations); 2) the average distance of Euclidean distance among land use group means based on the NMDS axes scores; 3) the effect if any of temporal variability on the NMDS axes (two-factor MANOVA); 4) the number of Mantel tests and partial Mantel tests between BMI matrix and the land use matrix with significant correlations; and 5) the number of Mantel tests correlating BMI distances with corresponding distances of physical-geographic variables that were not significant. Bray-Curtis distance calculated using abundance data did not rank best with any of the criterion, chi-square was ranked best in the third criterion, Jaccard distance was ranked best in the second and the fourth criteria, and Bray-Curtis indices ranked best in the first, third, and fifth criteria (see **Table 5**). Each BMI distance measure had different strengths, but Bray-Curtis based on benthic indices was best overall.

Temporal variability can influence abundance data (Milner *et al*. 2006) and so to avoid seasonal changes in BMI abundance the sampling for littoral macroinvertebrates should be limited to three weeks (Reid *et al*. 1995). From a study on the Fraser River in British Columbia, it was recommended that sampling of river test sites using the RCA be conducted in autumn only or over multiple sampling dates to reduce seasonal shifts or stochastic events (Reece *et al*. 2001). In this study three seasons were sampled over

Table 14 Summary results of Mantel test correlations using the Jaccard distance matrix based on BMI presence/absence. The p values that are bold were positively significant at p<0.05 (or Bonferroni corrected at p<0.008 in the pairwise tests). Land use scenario codes refer to (1) forest, (2) clear-cut/reshaping, (3) turf establishment, and (4) golf course operation, n=196. Matrix codes refer to the two main matrices used in the Mantel Test (m1 and m2), and the third as the covariable matrix (m3) in the Partial Mantel Test.

Test Type	Land Use Scenario	Jaccard P/A	Land use	Physical-geographic	Matrix r	p value
Mantel Test	1, 2, 3, 4	m1	m2		0.07	0.001
Mantel Test	1, 2, 3, 4	m1		m2	0.15	0.001
Mantel Test	1, 2, 3, 4		m1	m2	0.18	0.001
Partial Mantel Test	1, 2, 3, 4	m1	m2	m3	0.04	0.001
Pairwise Tests						
Mantel Test	1 & 2	m1	m2		0.23	0.001
Mantel Test	1 & 2	m1		m2	0.17	0.001
Mantel Test	1 & 2		m1	m2	0.16	0.001
Partial Mantel Test	1 & 2	m1	m2	m3	0.20	0.001
Mantel Test	1 & 3	m1	m2		0.38	0.001
Mantel Test	1 & 3	m1		m2	0.14	0.001
Mantel Test	1 & 3		m1	m2	0.13	0.003
Partial Mantel Test	1 & 3	m1	m2	m3	0.36	0.001
Mantel Test	1 & 4	m1	m2		0.07	0.001
Mantel Test	1 & 4	m1		m2	0.12	0.001
Mantel Test	1 & 4		m1	m2	0.25	0.001
Partial Mantel Test	1 & 4	m1	m2	m3	0.04	0.016
Mantel Test	2 & 3	m1	m2		0.35	0.001
Mantel Test	2 & 3	m1		m2	0.01	0.483
Mantel Test	2 & 3		m1	m2	0.02	0.403
Partial Mantel Test	2 & 3	m1	m2	m3	0.35	0.001
Mantel Test	2 & 4	m1	m2		-0.03	0.197
Mantel Test	2 & 4	m1		m2	0.09	0.001
Mantel Test	2 & 4		m1	m2	-0.02	0.309
Partial Mantel Test	2 & 4	m1	m2	m3	-0.03	0.205
Mantel Test	3 & 4	m1	m2		0.01	0.292
Mantel Test	3 & 4	m1		m2	0.11	0.001
Mantel Test	3 & 4		m2	m1	-0.01	0.479
Partial Mantel Test	3 & 4	m1	m2	m3	0.01	0.304

Table 15 Summary results of Mantel test correlations using the chi-square (χ^2) distance matrix based on macroinvertebrate abundance. The p values that are bold were positively significant at p<0.05 (or Bonferroni corrected at p<0.008 in the pairwise tests). Land use scenario codes refer to (1) forest, (2) clear-cut/reshaping, (3) turf establishment, and (4) golf course operation, n=196. Matrix codes refer to the two main matrices used in the Mantel Test (m1 and m2), and the third as the covariable matrix (m3) in the Partial Mantel Test.

Test Type	Land Use Scenario	χ^2 Abundance	Land use	Physical-geographic	Matrix r	p value
Mantel Test	1, 2, 3, 4	m1	m2		-0.05	0.007
Mantel Test	1, 2, 3, 4	m1		m2	0.09	0.001
Mantel Test	1, 2, 3, 4		m1	m2	0.18	0.001
Partial Mantel Test	1, 2, 3, 4	m1	m2	m3	-0.07	0.001
Pairwise Tests						
Mantel Test	1 & 2	m1	m2		0.18	0.001
Mantel Test	1 & 2	m1		m2	0.14	0.001
Mantel Test	1 & 2		m1	m2	0.16	0.001
Partial Mantel Test	1 & 2	m1	m2	m3	0.16	0.001
Mantel Test	1 & 3	m1	m2		0.27	0.001
Mantel Test	1 & 3	m1		m2	0.19	0.001
Mantel Test	1 & 3		m1	m2	0.13	0.002
Partial Mantel Test	1 & 3	m1	m2	m3	0.26	0.001
Mantel Test	1 & 4	m1	m2		-0.09	0.002
Mantel Test	1 & 4	m1		m2	0.07	0.008
Mantel Test	1 & 4		m1	m2	0.25	0.001
Partial Mantel Test	1 & 4	m1	m2	m3	-0.11	0.001
Mantel Test	2 & 3	m1	m2		0.15	0.018
Mantel Test	2 & 3	m1		m2	-0.02	0.417
Mantel Test	2 & 3		m1	m2	0.02	0.431
Partial Mantel Test	2 & 3	m1	m2	m3	0.15	0.018
Mantel Test	2 & 4	m1	m2		-0.09	0.003
Mantel Test	2 & 4	m1		m2	0.08	0.001
Mantel Test	2 & 4		m1	m2	-0.02	0.323
Partial Mantel Test	2 & 4	m1	m2	m3	-0.09	0.003
Mantel Test	3 & 4	m1	m2		-0.08	0.001
Mantel Test	3 & 4	m1		m2	-0.02	0.348
Mantel Test	3 & 4		m2	m1	-0.01	0.475
Partial Mantel Test	3 & 4	m1	m2	m3	-0.08	0.001

three years and the BMI distance measures were evaluated to test whether temporal variability influenced the NMDS axes of each BMI distance matrix. The results showed that Jaccard had an a significant interaction effect with a combination of seasonal and annual effects, Bray-Curtis abundance had a significant seasonal effect, while chi-square

Table 16 Summary results of Mantel test correlations using Bray-Curtis (BC) distance matrix based on macroinvertebrate abundance. The p values that are bold were positively significant at p<0.05 (or Bonferroni corrected at p<0.008 in the pairwise tests). Land use scenario codes refer to (1) forest, (2) clear-cut/reshaping, (3) turf establishment, and (4) golf course operation, n=196. Matrix codes refer to the two main matrices used in the Mantel Test (m1 and m2), and the third as the covariable matrix (m3) in the Partial Mantel Test.

Test Type	Land Use Scenario	BC Abundance	Land use	Physical-geographic	Matrix r	p value
Mantel Test	1, 2, 3, 4	m1	m2		0.05	0.001
Mantel Test	1, 2, 3, 4	m1		m2	0.12	0.001
Mantel Test	1, 2, 3, 4		m1	m2	0.18	0.001
Partial Mantel Test	1, 2, 3, 4	m1	m2	m3	0.03	0.032
Pairwise Tests						
Mantel Test	1 & 2	m1	m2		0.19	0.001
Mantel Test	1 & 2	m1		m2	0.10	0.001
Mantel Test	1 & 2		m1	m2	0.16	0.001
Partial Mantel Test	1 & 2	m1	m2	m3	0.18	0.001
Mantel Test	1 & 3	m1	m2		0.15	0.001
Mantel Test	1 & 3	m1		m2	0.18	0.001
Mantel Test	1 & 3		m1	m2	0.13	0.002
Partial Mantel Test	1 & 3	m1	m2	m3	0.13	0.001
Mantel Test	1 & 4	m1	m2		0.06	0.003
Mantel Test	1 & 4	m1		m2	0.15	0.001
Mantel Test	1 & 4		m1	m2	0.25	0.001
Partial Mantel Test	1 & 4	m1	m2	m3	0.02	0.136
Mantel Test	2 & 3	m1	m2		0.16	0.028
Mantel Test	2 & 3	m1		m2	-0.01	0.388
Mantel Test	2 & 3		m1	m2	0.02	0.415
Partial Mantel Test	2 & 3	m1	m2	m3	0.16	0.019
Mantel Test	2 & 4	m1	m2		-0.07	0.004
Mantel Test	2 & 4	m1		m2	0.14	0.001
Mantel Test	2 & 4		m1	m2	-0.02	0.306
Partial Mantel Test	2 & 4	m1	m2	m3	-0.07	0.008
Mantel Test	3 & 4	m1	m2		-0.04	0.075
Mantel Test	3 & 4	m1		m2	0.16	0.001
Mantel Test	3 & 4		m2	m1	-0.01	0.492
Partial Mantel Test	3 & 4	m1	m2	m3	-0.04	0.078

Table 17 Summary results of Mantel test correlations using the Bray-Curtis (BC) distance matrix based on benthic indices. The p values that are bold were positively significant at p<0.05 (or Bonferroni corrected at p<0.008 in the pairwise tests). Land use scenario codes refer to (1) forest, (2) clear-cut/reshaping, (3) turf establishment, and (4) golf course operation, n=196. Matrix codes refer to the two main matrices used in the Mantel Test (m1 and m2), and the third as the covariable matrix (m3) in the Partial Mantel Test.

Test Type	Land Use Scenario	BC Indices	Land use	Physical-geographic	Matrix r	p value
Mantel Test	1, 2, 3, 4	m1	m2		0.06	0.001
Mantel Test	1, 2, 3, 4	m1		m2	0.04	0.017
Mantel Test	1, 2, 3, 4		m1	m2	0.18	0.001
Partial Mantel Test	1, 2, 3, 4	m1	m2	m3	0.05	0.001
Pairwise Tests						
Mantel Test	1 & 2	m1	m2		0.18	0.001
Mantel Test	1 & 2	m1		m2	0.08	0.001
Mantel Test	1 & 2		m1	m2	0.16	0.001
Partial Mantel Test	1 & 2	m1	m2	m3	0.17	0.001
Mantel Test	1 & 3	m1	m2		0.17	0.001
Mantel Test	1 & 3	m1		m2	0.14	0.001
Mantel Test	1 & 3		m1	m2	0.13	0.003
Partial Mantel Test	1 & 3	m1	m2	m3	0.16	0.001
Mantel Test	1 & 4	m1	m2		0.06	0.003
Mantel Test	1 & 4	m1		m2	0.06	0.003
Mantel Test	1 & 4		m1	m2	0.25	0.001
Partial Mantel Test	1 & 4	m1	m2	m3	0.04	0.015
Mantel Test	2 & 3	m1	m2		0.20	0.004
Mantel Test	2 & 3	m1		m2	0.09	0.164
Mantel Test	2 & 3		m1	m2	0.02	0.437
Partial Mantel Test	2 & 3	m1	m2	m3	0.20	0.007
Mantel Test	2 & 4	m1	m2		0.03	0.105
Mantel Test	2 & 4	m1		m2	-0.01	0.261
Mantel Test	2 & 4		m1	m2	-0.02	0.341
Partial Mantel Test	2 & 4	m1	m2	m3	0.03	0.107
Mantel Test	3 & 4	m1	m2		-0.01	0.268
Mantel Test	3 & 4	m1		m2	0.02	0.170
Mantel Test	3 & 4		m2	m1	-0.01	0.506
Partial Mantel Test	3 & 4	m1	m2	m3	-0.01	0.286

and Bray-Curtis indices were not influenced by seasonal and/or annual effects. When using either chi-square abundance measure or the Bray-Curtis benthic indices measure, seasonal and annual effects were eliminated.

In the District of Muskoka, 1999 and 2000 were relatively normal rainfall years. In 2001, there was a summer drought as a result of the El Niño effect according to observed conditions and seen in the southern oscillation index (http://www.bom.gov.au/climate/current/soi2.shtml). McElravy *et al*. (1989) found that BMI density decreased and abundance of Chironomidae increased in a wet year in northern California. In this study, annual and seasonal effects on Jaccard and Bray-Curtis measures were most likely associated with the lack of precipitation Muskoka received in 2001. These two distance measures detected changes in BMI communities that reflected most likely on the differences in flow regimes in streams from year to year and season to season.

Charvet *et al*. (2000) found that abundance data should not be used over large geographic regions because both altitude and geological characteristics create differences in taxonomic composition at reference sites. However, the present geographic area was small even though there were significant differences in longitude, and elevation between the reference group sites and test group sites. In addition, when the four BMI distance measures were evaluated, all four approaches (distance measures) had some Mantel tests that showed that BMI distances among samples were significantly correlated with corresponding distances in physical-geographic variables. This suggests that spatial variability in BMI distances exists among samples with the particular distance measure. The Bray-Curtis measure that was based on benthic indices was affected the least by influences in spatial variability.

Axes from NMDS are summary indices where Jaccard reflected BMI richness among samples, chi-square is the goodness-of-fit, and the Bray-Curtis measures are percent difference among samples. Generally, richness measures have higher statistical power in detecting changes caused by perturbation and lower spatial, temporal, and sample variability than abundances (Sandin and Johnson 2000). In this study, Jaccard distance measure of richness was influenced by both temporal and physical-geographic variability, whereas chi-square of goodness of fit (linearly) was not influenced by either temporal or physical-geographic variability. This shows that temporal and spatial variability should be tested when using any distance measure.

Rapid bioassessment was used as an approach to collect and analyze BMI. This approach revealed that sub-samples of 100 organisms collected from 3 replicates at a site and identified only to the taxonomic level of family and order was sensitive enough to detect differences among land use categories associated with forested reference sites, and golf course construction and operational sites. Rapid bioassessment should be used as a screening tool for impact on Precambrian Shield streams. A more detailed survey including water chemistry analysis would be useful for assessing streams deemed biologically impaired during the screeing process.

ACKNOWLEDGEMENTS

The authors wish to thank the NSERC and Ontario Ministry of Natural Resources for funding.

REFERENCES

Bailey RC, Kennedy MG, Dervish MZ, Taylor RM (1998) Biological assessment of freshwater ecosystems using a reference condition approach: comparing predicted and actual benthic invertebrate communities in Yukon streams. *Freshwater Biology* **39**, 765-774

Barbour MT, Gerritsen J, Snyder BD, Stribling JB (1999) Rapid bioassessment protocols for use in streams and wadeable rivers: periphyton, benthic macroinvertebrates, and fish (2nd Edn), U.S. Environmental Protection Agency, Office of Water, Washington, D.C. USA, 339 pp

Bonada N, Prat N, Resh VH, Statzner B (2006) Developments in aquatic insect biomonitoring: A comparative analysis of recent approaches. *Annual Review of Entomology* **51**, 495-523

Bray JR, Curtis JT (1957) An ordination of the upland forest communities of southern Wisconsin. *Ecological Monographs* **27**, 325-349

Carlisle DM, Meador MR, Moulton II SR, Ruhl PM (2007) Estimation and application of indicator values for common macroinvertebrate genera and families of the United States. *Ecological Indicators* **7**, 22-33

Charvet S, Statzner B, Usseglio-Polatera P, Dumont B (2000) Traits of benthic macroinvertebrates in semi-natural French streams: An initial application to biomonitoring in Europe. *Freshwater Biology* **43**, 277-296

Clarke RT, Wright JF, Furse MT (2003) RIVPACS models for predicting the expected macroinvertebrate fauna and assessing the ecological quality of rivers. *Ecological Modelling* **160**, 219-233

Conover WJ, Iman RL (1981) Rank transformations as a bridge between parametric and nonparametric statistics. *The American Statistician* **x**, 124-129

David SM, Somers KM, Reid RA, Hall RJ, Girard RE (1998) Sampling protocols for the rapid bioassessment of streams and lakes using benthic macroinvertebrates (2nd Edn), Dorset Environmental Science Centre, Ontario Ministry of the Environment, Dorset, 29 pp

Davidson J (2002) Applying the reference condition approach to monitor invertebrates in streams of the Sudbury mining area. In: *Cooperative Freshwater Ecology Unit and Department of Biology*, Laurentian University, Sudbury, 115 pp

Dillon PJ, Molot LA (1996) Long-term phosphorus budgets and an examination of a steady-state mass balance model for central Ontario Lakes. *Water Research* **30**, 2273-2280

Dillon PJ, Molot LA (1997) Effect of landscape form on export of dissolved organic carbon, iron, and phosphate from forested stream catchments. *Water Resources Research* **33**, 2591-2600

Faith DP, Minchin PR, Belbin L (1987) Compositional dissimilarity as a robust measure of ecological distance. *Vegetatio* **69**, 57-68

Fore LS, Karr JR, Wisseman RW (1996) Assessing invertebrate responses to human activities: Evaluating alternative approaches. *Journal of the North American Benthological Society* **15**, 212-231

Gerritsen J (1995) Additive biological indices for resource management. *Journal of the North American Benthological Society* **14**, 451-457

Gerritsen J, Barbour MT, King K (2000) Apples, oranges, and ecoregions: on determining pattern in aquatic assemblages. *Journal of the North American Benthological Society* **19**, 487-496

Hawkins CP, Norris RH (2000) Performance of different landscape classification for aquatic bioassessments: introduction to the series. *Journal of the North American Benthological Society* **19**, 367-369

Hering D, Buffagni A, Moog O, Sandin L, Sommerhäuser M, Stubauer I, Feld C, Johnson R, Pinto P, Skoulikidis N, Verdonschot P, Zahrádková S (2003) The development of a system to assess the ecological quality of streams based on macroinvertebrates - design of the sampling programme within the AQEM project. *International Review of Hydrobiology* **88**, 245-361

Hilsenhoff WL (1988) Seasonal correction factors for the biotic index. *Great Lakes Entomologist* **21**, 9-14

Hood G (2000) PopTools. Canberra, Australia: Pest Animal Control Co-operative Research Centre, based at Wildlife and Ecology, CSIRO

Hughes RM (1995) Defining acceptable biological status by comparing with reference conditions. In: Davis WS, Simon TP (Eds) *Biological Assessment and Criteria: Tools for Water Resouce Planning and Decision Making*, Lewis Publishers & CRC Press Inc., Boca Raton, FL, USA, pp 31-47

Jeffries DS, Snyder WR (1983) Geology and geochemistry of the Muskoka-Haliburton study area. Ontario Ministry of the Environment, Dorset, Ontario, Canada, 101 pp

Legendre P, Legendre L (2000) *Numerical Ecology: Developments in Environmental Modelling*, Elsevier Science, The Netherlands, 870 pp

Lemmon PE (1957) A new instrument for measuring forest overstory density. *Journal of Forestry* **55**, 667-668

Lenat DR (1987) The macroinvertebrate fauna of the Little River, North Carolina: Taxa list and seasonal trends. *Archiv fur Hydrobiologie* **110**, 19-43

Lenat DR, Barbour MT (1994) Using benthic macroinvertebrate community structure for rapid, cost-effective, water quality monitoring: Rapid bioassessment. In: Loeb SL, Spacie A (Eds) *Biological Monitoring of Aquatic Systems*, CRC Press, Inc., Boca Raton, FL, pp 187-215

Lewis MA, Foss SS, Harris PS, Stanley RS, Moore JC (2001) Sediment chemical contamination and toxicity associated with a coastal golf course complex. *Environmental Toxicology and Chemistry* **20**, 1390-1398

Linke S, Bailey RC, Schwindt J (1999) Temporal variability of stream bioassessments using benthic macroinvertebrates. *Freshwater Biology* **42**, 575-584

Manly BFJ (1994) *Multivariate Statistical Methods: A Primer*, Chapman and Hall, New York, 232 pp

McElravy EP, Lamberti GA, Resh VH (1989) Year-to-year variation in the aquatic macroinvertebrate fauna of a northern California stream. *Journal of the North American Benthological Society* **8**, 51-63

Milner AM, Conn SC, Brown LE (2006) Persistence and stability of macroinvertebrate communities in streams of the Denali National Park, Alaska: Implications for biological monitoring. *Freshwater Biology* **51**, 373-387

Moss D, Furse MT, Wright JF, Armitage PD (1987) The prediction of the macro-invertebrate fauna of unpolluted running-water sites in Great Britain using environmental data. *Freshwater Biology* **17**, 41-52

Norris RH (1995) Biological monitoring: The dilemma of data analysis. *Journal of the North American Benthological Society* **14**, 440-450

Parsons M, Norris RH (1996) The effect of habitat-specific sampling on biological assessment of water quality using a predictive model. *Freshwater Bio-*

logy **36**, 419-434

Plafkin JL, Barbour MT, Porter KD, Gross SK, Hughes RM (1989) Rapid bioassessment protocols for use in streams and rivers: Benthic macroinvertebrates and fish. U.S. EPA, Office of Water, 125 pp

Reece PF, Reynoldson TB, Richardson JS, Rosenberg DM (2001) Implications of seasonal variation for biomonitoring with predictive models in the Fraser River catchment, British Columbia. *Canadian Journal of Fisheries and Aquatic Sciences* **58**, 1411-1418

Reid RA, Somers KM, David SM (1995) Spatial and temporal variation in littoral-zone benthic invertebrates from three south-central Ontario lakes. *Canadian Journal of Fisheries and Aquatic Sciences* **52**, 1406-1420

Reynoldson TB, Bailey RC, Day KE, Norris RH (1995) Biological guidelines for fresh-water sediment based on benthic assessment of sediment (the Beast) using a multivariate approach for predicting biological state. *Australian Journal of Ecology* **20**, 198-219

Reynoldson TB, Norris RH, Resh VH, Day KE, Rosenberg DM (1997) The reference condition: A comparison of multimetric and multivariate approaches to assess water-quality impairment using benthic macroinvertebrates. *Journal of the North American Benthological Society* **16**, 833-852

Richards C, Host G, Arthur J (1993) Identification of predominant environmental factors structuring stream macroinvertebrate communities within a large agricultural catchment. *Freshwater Biology* **29**, 285-294

Rohlf FJ (1993) NTSYS-pc: Numerical taxonomy and multivariate analysis system. Exeter Software, Setauket, NY

Rosenberg MS (2001) PASSAGE. Pattern analysis, spatial statistics and geographic exegesis. Department of Biology, Arizona State University, Tempe, AZ, 12 pp

Sandin L, Johnson RK (2000) The statistical power of selected indicator metrics using macroinvertebrates for assessing acidification and eutrophication of running waters. *Hydrobiologia* **422**, 233-243

Somers KM, Reid RA, David SM (1998) Rapid biological assessments: how many animals are enough? *Journal of the North American Benthological Society* **17**, 348-358

StatSoft (2000) STATISTICA for Windows [Computer program manual]. Tulsa, OK: StatSoft, Inc

ter Braak CJF (1995) Ordination. In: Jongman RHG, ter Braak CJF, Van Tongeren OFR (Eds) *Data Analysis in Community and Landscape Ecology*, Cambridge University Press, New York, pp 91-173

Van Tongeren OFR (1995) Cluster analysis. In: Jongman RHG, ter Braak CJF, Van Tongeren OFR (Eds) *Data Analysis in Community and Landscape Ecology*, Cambridge University Press, New York, 174-212

Walsh CJ, Sharpe AK, Breen PF, Sonneman JA (2001) Effects of urbanization on streams of the Melbourne region, Victoria, Australia. I. Benthic macroinvertebrate communities. *Freshwater Biology* **46**, 535-551

Winter JG, Somers KM, Dillon PJ, Paterson C, Reid RA (2002) Impacts of golf courses on macroinvertebrate community structure in Precambrian Shield streams. *Journal of Environmental Quality* **31**, 2015-2025

Wright JF, Furse MT, Moss D (1998) River classification using invertebrates: RIVPACS applications. *Aquatic Conservation: Marine and Freshwater Ecosystems* **8**, 617-631

Zar JH (1999) *Biostatistical Analysis*, Prentice-Hall, Inc., Upper Saddle River, New Jersey, 929 pp