Environmental Monitoring and Assessment (2006) **119:** 83–96
DOI: 10.1007/s10661-005-9011-x

EVALUATING ECOLOGICAL INDICATORS: LAKES IN THE NORTHEASTERN UNITED STATES

PAUL A. MURTAUGH[1,*] and PENELOPE S. POOLER[2]
[1]*Department of Statistics, Oregon State University, Corvallis, Oregon 97331, U.S.A.;* [2]*Department of Statistics, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061-0439, U.S.A.*
*(*author for correspondence, e-mail: murtaugh@stat.oregonstate.edu)*

(Received 18 April 2005; accepted 22 August 2005)

Abstract. We use data from a survey of several hundred lakes in the northeastern United States by the U.S. Environmental Protection Agency to illustrate an approach to identifying promising indicators of lake condition. We construct a hypothetical gold standard of water quality from the first principal component of 16 chemical variables measured in the lakes, and examine its associations with 71 candidate indicators based on measurements of human activity, birds, fish and zooplankton in the lakes or their watersheds. Nonparametric summaries of these associations – based on rank correlations and receiver-operating-characteristic curves – suggest that variables summarizing the extent of human disturbance are generally the strongest indicators. To the extent that our water-quality variable is a useful proxy for ecological condition, our results suggest that easily-obtained measures of human activity are at least as predictive as many of the harder-to-measure biological indicators that have been proposed.

Keywords: indicator, monitoring, ROC curve, water quality, zooplankton

1. Introduction

Ecological indicators are commonly used to characterize features of communities and ecosystems. For example, biologists have developed indicators of ecosystem stress (Adams, 2002), biological "health" or integrity (Davis and Simon, 1995; Xu *et al*., 2001; Hughes *et al*., 2004), nutrient enrichment (Cottingham and Carpenter, 1998), and other forms of environmental pollution (Kottferova *et al*., 1996). The identification of suitable indicators is a key step in the development of environmental monitoring programs (Hunsaker and Carpenter, 1990).

Murtaugh (1996) considered indicators as surrogates for key response variables that are difficult to measure, drawing analogies with the evaluation of test results used as indicators of disease in clinical medicine. This approach assumes the existence of a "gold standard" – that is, an assessment of the "true" condition against which test results can be compared. The gold standard is typically too expensive or difficult to obtain routinely, and the goal is to identify surrogates that are more readily available but still capture the key features of the clinical condition. This view is consistent with the definition of an ecological indicator by Carignan and

Villard (2002) as "an element, process, or property of the ecosystem that for some reason . . . cannot be measured in a more direct way."

As implied by the above definition, there usually *is* no gold standard in environmental assessment. In some cases the indicators themselves seem to be the responses or endpoints of interest, and the identification of useful indicators then depends largely on expert judgement. For example, Jackson *et al*. (2000) define an ecological indicator as "a measure, an index of measures, or a model that characterizes an ecosystem or one of its critical components." In other cases, investigators seek variables that reflect ecosystem health or integrity (e.g., see Bryce *et al*., 2002; Hughes *et al*., 2004). Pristine or minimally disturbed systems are sometimes used as references against which to compare other systems (Hughes, 1995; Hughes *et al*., 1998), and correlation with the extent of human disturbance is often identified as a desirable feature of an indicator (Stemberger and Lazorchak, 1994; Allen *et al*., 1999; Carignan and Villard, 2002; Gergel *et al*., 2002).

Large pilot data sets, like that collected on lakes in the northeast United States by the U.S. Environmental Protection Agency, as described below, provide valuable opportunities for quantitative assessments of the usefulness of potential indicators. Such assessments require a clear statement of the ecological properties that are of key interest, and they require that some gold standard for those properties can be developed for use in comparing the predictive value of candidate indicators. This standard, which must obviously be based on information distinct from that contained directly in the indicators, is all-important in determining the ranking of the indicators.

We choose a standard based on water quality, as reflected by chemical measurements made on the lake water. Our water-quality summary is not a gold standard in the usual sense, because it is not the "true" condition that all ecologists want to assess when measuring lake indicators. Rather, we consider water quality to be a proxy for a variety of properties that vary with the pristineness of lakes, which we hope will provide a useful frame of reference for comparing indicators.

We evaluate individual indicators, as well as sets of indicators identified by multiple linear regression modeling, using several metrics for summarizing these indicators' associations with the water-quality variable.

2. Methods

2.1. The data

The data used here are from the U.S. Environmental Protection Agency's EMAP Surface Waters Northeast Lakes 1991–1994 program (Chaloud and Peck, 1994; Baker et al., 1997; http://www.epa.gov/emap/html/dataI/surfwatr/data/nelakes/index.html). This program demonstrated a probability-based survey design for assessment of the status of ecological resources over a large region. Data were

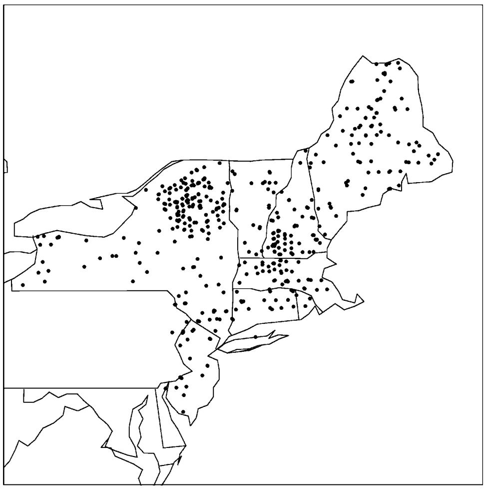

Figure 1. Locations of the 363 lakes that were not missing values of the water-quality variable.

available for up to 368 lakes that were visited one to several times during the summer; see Figure 1.

Included in the data set were (i) information on lake and watershed characteristics digitized from topographic maps; (ii) chemical measurements from water taken from the deepest part of each lake using a van Dorn sampler; (iii) information on the diversity and abundance of fishes based on sampling with multiple gears in pelagic and riparian zones of the lakes; (iv) information on breeding birds in the riparian zone, from visual and auditory identifications made at specified stops around the lake perimeter; (v) counts of zooplankton from vertical tows of coarse (202 μm) and fine (48 μm) mesh nets at the deepest part of the lake; and (vi) information on the diversity of diatoms in sediment cores.

We did not use every variable in the EMAP data set; we excluded some variables that took on only a few distinct values, that were missing for many lakes, or that were tightly correlated with other included variables. When there were multiple measurements of a variable in a lake, they were averaged. Acronyms and definitions of the variables are given in the Appendix.

2.2. Receiver-operating-characteristic curves

Among other metrics, we used receiver-operating characteristic (ROC) curves to evaluate sets of indicators (e.g., see Murtaugh, 1996). To understand this approach, suppose Y is a binary variable indicating whether or not the water-quality variable exceeds some threshold, and X is a marker – or a prediction of a regression model based on several markers – that we will use to guess whether or not $Y = 1$. If $X > c$, where c is some cutoff, we guess $Y = 1$ ("positive"); if $X \leq c$, we guess $Y = 0$

("negative"). The *sensitivity* of this procedure is defined as $P(X > c|Y = 1)$, which we estimate as the proportion of positive lakes for which $X > c$. The *specificity* is $P(X \leq c \mid Y = 0)$, estimated as the proportion of negative lakes for which $X \leq c$.

When we plot sensitivity vs. specificity calculated over a range of possible values of c, we get a receiver operating characteristic (ROC) curve. The more bowed this curve is toward the upper right corner, the better the marker is in predicting the value of Y. The area under the ROC curve is often used as a summary of the accuracy of the marker (e.g., see Hanley and McNeil, 1982). An area of 1.0 means that the marker perfectly discriminates "positive" $(Y = 1)$ and "negative" $(Y = 0)$ lakes, while an area of 0.5 indicates a marker that is of no use in predicting Y.

The value of this approach is that it is nonparametric, in the sense that it is invariant to tranformations of the indicator (X) variable, and it accommodates the frequent need of lake managers to decide whether individual lakes are in "acceptable" or "unacceptable" condition $(Y = 0$ or $1)$.

2.3. The gold standard

If one accepts that indicators are indirect reflections of key aspects of lake condition that are difficult or impossible to measure directly, then the comparison of candidate indicators requires some proxy for the unknown, true condition. We chose to use a standard based on water-quality measurements, as described below. We emphasize that water quality *per se* is not itself the true condition being assessed. If it were, direct measurements of water chemistry would provide the best available summary, and there would be no point in seeking out other kinds of indicators (such as birds, fish, etc.). Rather, we assume that water quality is closely correlated with aspects of lake condition that are of greatest interest to ecologists and lake managers.

As a proxy for "true" lake condition, we used the first principal component calculated for 16 water-chemistry variables measured in 365 lakes:

$$\begin{aligned}\text{CHEM} = {} & -0.160 \cdot \text{ALTD} + 0.329 \cdot \text{ANC} + 0.343 \cdot \text{CA} + 0.238 \cdot \text{CHLA} \\ & +0.361 \cdot \text{COND} + 0.058 \cdot \text{DOC} + 0.284 \cdot \text{K} + 0.358 \cdot \text{MG} \\ & +0.291 \cdot \text{NA} + 0.139 \cdot \text{NH}_4 + 0.020 \cdot \text{NO}_3 + 0.239 \cdot \text{NTL} \\ & +0.248 \cdot \text{PTL} + 0.112 \cdot \text{SIO}_2 + 0.257 \cdot \text{SO}_4 + 0.217 \cdot \text{TSS}, \qquad (1)\end{aligned}$$

where the variables were transformed as shown in the Appendix and then standardized by subtracting the mean and dividing by the standard error. This principal component explains 40% of the variation in the data.

Examining the coefficients of the principal component in (1), one can see that higher values of CHEM indicate more productive and/or impaired lakes. Most of these variables are known to be affected by anthropogenic disturbances (e.g., see Meybeck, 1998; Carignan and Steedman, 2000; Cude, 2001). The negative coefficient for dissolved aluminum was unexpected, and could be related to the higher solubility of aluminum in acidic lakes (Sullivan and Cosby, 1998) and the generally good water quality in those lakes, many of which are located in sparsely populated areas of the Northeast.

As explained earlier, ROC curves are based on a binary response. In the lake context, this might be a judgement by a manager about whether lake condition is acceptable or unacceptable. To obtain a reasonable cutoff value of CHEM, above which we might score lake condition as unacceptable, we compared the distribution of this variable in eutrophic vs. non-eutrophic lakes, where eutrophic lakes have total phosphorus exceeding 35 μg/L, chlorophyll *a* exceeding 8 μg/L, and Secchi depth less than 3 m (OECD, 1982). We chose as our cutoff the value 9.2, which is midway between the third quartile of CHEM in non-eutrophic lakes (8.7) and the first quartile of CHEM in eutrophic lakes (9.7).

Of course, since our gold standard is based on the 16 water-chemistry variables in Equation (1), we lose the ability to judge the value of those variables as indicators of lake condition.

3. Results

3.1. Univariate associations

The Appendix shows values of r_S, the Spearman rank correlation between indicator values and water-quality scores, and the area under ROC curves summarizing each indicator's ability to identify "impaired" lakes. Both of these measures are non-parametric, in the sense that they are invariant to monotonic transformations of the indicator values. Figures 2 and 3 present scatterplots and ROC curves, respectively, that illustrate the associations of four of the most informative indicators with water quality.

Most of the human variables have reasonably strong associations with the water-quality variable (Appendix). The "best" indicator, the percent of the watershed consisting of disturbed land, has an area of 0.847 under the ROC curve. That is, if two lakes are randomly selected from this population, there is an 85% chance that the lake in the more disturbed watershed will have poorer water quality (higher value of CHEM).

Among the bird variables, the percent of stops with tolerants present (FSTOTO) and the average number of individuals per stop (SIND) were the most strongly correlated with the water-quality variable. Both increased with decreasing water quality. Whether these trends hold in other systems is unclear. Studying bird assemblages along stream reaches in Oregon's Willamette Valley, Bryce *et al.* (2002) found no association between total number of individuals and an index representing human disturbance. In addition, Bryce *et al.* (2002) reported that native species richness tended to decrease with increasing disturbance. In the northeastern lakes, the average number of species per stop (SRICH) was positively correlated with CHEM

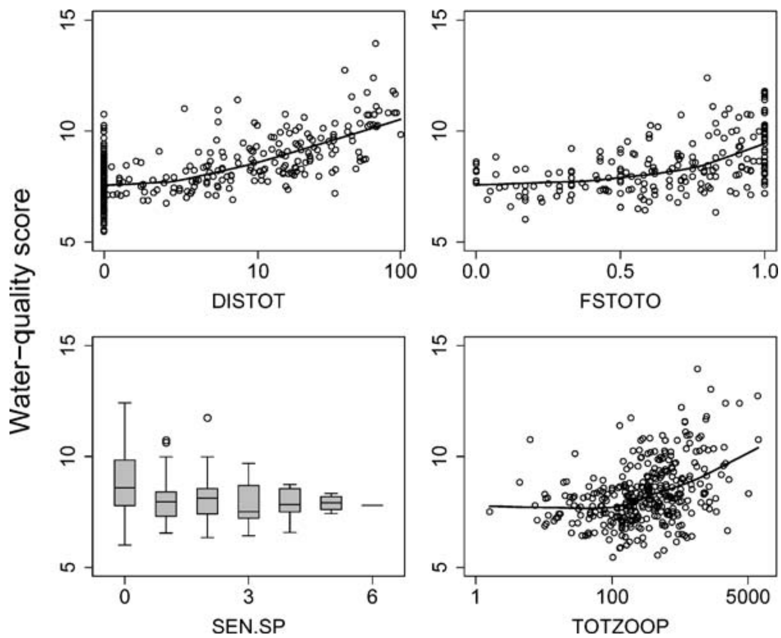

Figure 2. Scatterplots of the water-quality score (Equation (1)) vs. indicator values, for the human, bird, fish and zooplankton variables having the strongest correlations with CHEM (see Appendix). Larger values of CHEM indicate lower water quality. The lines are nonparametric smoothed fits to the scatterplots, from the "lowess" function in R (R Development Core Team, 2004). Note that the horizontal axes reflect the transformations given in the Appendix.

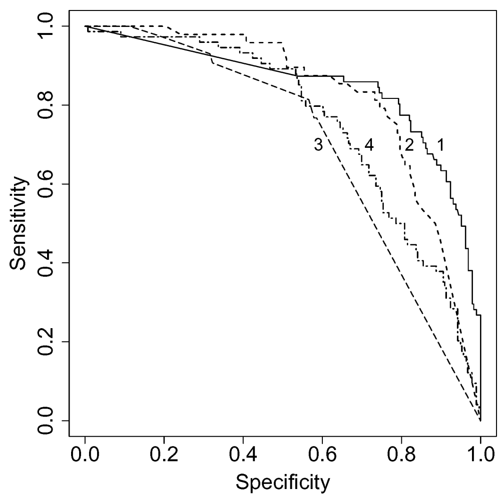

Figure 3. Receiver-operating-characteristic (ROC) curves for the four indicators shown in Figure 2 as predictors of values of CHEM exceeding 9.2: 1 = DISTOT, 2 = FSTOTO, 3 = SEN.SP, 4 = TOTZOOP.

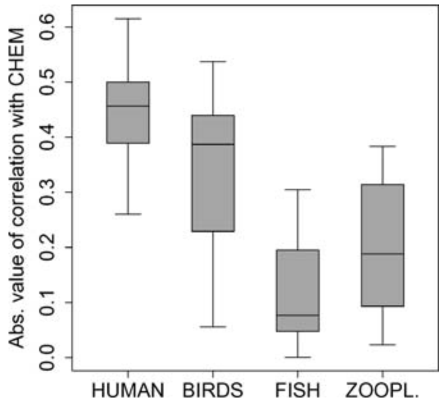

Figure 4. Boxplot of absolute values of r_S, the Spearman rank correlation of indicator values with water-quality scores, for four categories of indicators.

(negatively correlated with water quality), but native and non-native species were not discriminated in this data set.

The magnitudes of the correlations of the zooplankton indicators with the water-quality variable were generally smaller than those of human- and bird-based indicators. Numbers of individuals of several groups were positively correlated with CHEM, perhaps reflecting a productivity gradient. Species richness of larger grazers tended to be negatively correlated with CHEM (positively correlated with water quality).

Among the other indicators, two measures of diatom diversity were higher in lakes having better water quality, and there was a tendency for the diversity and abundance of sensitive fish species to decrease with deteriorating water quality. Interestingly, the only three indicators showing statistically significant evidence of non-monotonic relationships with the water-quality variable (INV.IND, INV.SP and OMNI.SP) were from the fish group; these variables would consequently be poorly suited as indicators of water quality.

Figure 4 summarizes the associations of different categories of indicators with the water-quality variable. The human-based indicators generally have the strongest correlations. Surprisingly, the bird indicators are the next best group, even though birds seem less directly connected to water quality than are fish and zooplankton, which showed relative weak associations. Since there are different numbers of indicators in the various categories, and since we do not know the rationale behind the EPA's choice of indicators to include in their survey, this comparison must be interepreted cautiously.

3.2. Multiple-variable models

Table I shows the results of regression modeling with two methods of variable selection: stepwise selection starting with no predictors, and an "all-subset" procedure

TABLE I

Regression models for the water-quality score as a function of predictors selected from a pool containing the "best" five indicators from each category (human, bird, fish and zooplankton variables)

Model	Terms	C_p	R^2	ROC area
1	4.89 + 0.43 DISTOT + 0.75 SIND + 0.30 CRH	25.8	0.60	0.90
2	4.35 + 0.47 DISTOT − 0.16 RD.DEN + 0.48 FSTOTO + 0.26 DIDTSE + 0.61 DIDTIN + 0.077 SEN.IND + 0.19 TOL.SP − 0.19 INV.IND + 0.45 INV.SP + 1.02 TOTZOOP − 1.09 MICRO.N + 0.24 ROTH + 0.22 CRH	12.4	0.67	0.93

Both models are fit to a set of 192 lakes not missing values of any of these 20 predictors. Model 1 was obtained by stepwise variable selection, starting with no variables in the model and using 0.05 for both the P-to-enter and P-to stay. Model 2 is from an "all-subset" algorithm (the "leaps" function in S-PLUS) designed to minimize C_p. All predictor variables are transformed as indicated in the Appendix. R^2 is the coefficient of determination, and ROC area is the area under a receiver-operating-characteristic curve summarizing a model's ability to identify "positive" lakes (arbitrarily defined as those having CHEM > 9.2).

designed to minimize the C_p statistic by exploring a large number of models of various sizes. The C_p statistic is based on a model's mean squared error, with a penalty related to the number of predictors (e.g., see Ramsey and Schafer, 2002). For least-squares regression, ranking of models by this statistic gives results very similar to rankings based on Akaike's Information Criterion (AIC), a popular metric these days in ecology (Burnham and Anderson, 2002).

Our purpose in presenting these models is not to advocate a particular variable-selection method or promote a single "best" model. Rather, we hope to illustrate the effect on our water-quality predictions of increasing the number of indicators contributing to those predictions.

The stepwise procedure gives a simple model based on the amount of disturbed land in the watershed, numbers of individual birds, and numbers of herbivorous crustacean zooplankters, while the all-subset approach yields a complicated model that includes two human variables, three bird variables, four fish variables, and four zooplankton variables (Table I). Because of the large number of collinear predictors in this model, the regression coefficients are essentially uninterpretable and are presented mainly for completeness. As shown in Table I, this model explains more variability in the data than does the three-variable model obtained by stepwise variable selection, but it requires considerably more measurement effort.

Figure 5 shows ROC curves for predictions based on these two models, as well as the curve corresponding to the best univariate indicator, DISTOT. By this metric, the single-variable model is only slightly less accurate than the two multiple-variable

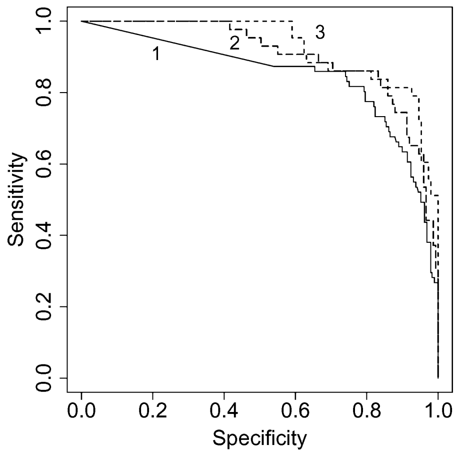

Figure 5. ROC curves for the best univariate indicator, DISTOT, and for predictions based on the multiple-variable models in Table I: 1 = DISTOT, 2 = three-variable model, 3 = thirteen-variable model.

models, and the accuracy of the three-variable model is quite similar to that of the thirteen-variable model.

4. Discussion

An important limitation of this study is our inability to explore temporal and within-lake variability of the indicators, since most of the EMAP lakes were visited just once. For indicators having roughly the same magnitude of association with the water-quality variable, considerations of the indicators' patterns of spatial and temporal variation could be very helpful in judging their usefulness (e.g., see Rusak *et al.*, 2002; Smith, 2002; Klemm *et al.*, 2003).

The human variables were better predictors of water quality than were any of the sets of biotic indicators (Figure 4), and the best human indicator, DISTOT, was only slightly less accurate than predictions from the two multiple-variable models (Figure 5). Given the relative ease of obtaining estimates of the extent of watershed disturbance (e.g., using remote sensing and GIS tools; Herlihy *et al.*, 1998), one must ask whether the discriminatory power added by monitoring communities of birds, fish and zooplankton is worth the extra time and effort. The answer of course depends on the specific goals of the monitoring. Biotic indicators have their own intrinsic value, beyond their ability to reflect water quality – e.g., the well-being of sports fisheries and the aesthetic value of diverse bird communities are important regardless of the background levels of water quality. Still, to the extent that water quality acts as a surrogate for key aspects of ecosystem health or integrity, our analyses suggest that relatively simple measures of human impacts can be strongly predictive of lake condition.

5. Conclusions

When the goal of indicator development is simply to choose among various possible representations of some ecological feature of interest (e.g., chlorophyll *a* concentration, carbon-14 measurements or Secchi depth as summaries of primary production in lakes), there is a suite of criteria that can be used to discriminate among candidate indicators (Hughes *et al*., 1998, 2004): the disparity between within-unit and between-unit variation, temporal variability, ease and cost of measurement, and so on. When the goal is to identify surrogates for less tangible or harder-to-measure properties, such as "integrity" of ecosystems or assemblages of organisms (e.g., Bryce *et al*., 2002), indicator development is facilitated when information from a few intensively-studied systems can be used to devise a "gold standard" for the feature(s) of interest. The tools discussed here can then be used for quantitative comparisons of the accuracy of candidate indicators in reflecting that standard. Without that information, one can never be sure whether an indicator – no matter how reliably, precisely and economically it is measured – is really assessing the ecological properties of greatest concern.

Acknowledgments

We thank Michael Schlax for helpful comments on an earlier version of the manuscript and an anonymous reviewer for many constructive criticisms and suggestions.

Appendix: Complete Results

Superscripts indicate the transformation that was done in regression modeling: 1 = $\log(Y + 1)$; 2 = $\log(\log(Y + 1))$; and 3 = $\log(Y + 100)$. An asterisk indicates evidence of non-monotonicity of a variable's relationship with the water-quality variable (based on a bootstrap test on the residuals from an isotonic regression fit to the original data; Murtaugh, 2003). n is the number of lakes having measurements of both the lake variable and the water-quality score; r_S is the Spearman correlation coefficient between the lake variable and the water-quality score; and ROC area is the area under an ROC curve relating the lake variable to a dichotomized version of the water-quality score, using a cutoff of 9.2 (see text). For variables having r_S less than zero, the ROC curves are constructed with lake variable values multiplied by -1. Within each category, variables are listed in order of their values of ROC area.

Name	Variable	n	r_S	ROC area
Human				
DISTOT[1]	% watershed human disturbed land	360	0.615	0.847
POPDENKM[1]	Population density (persons km^{-2})	360	0.536	0.797
HOUDENKM[1]	Housing unit density (housing km^{-2})	360	0.487	0.769
AG.TOT[1]	% watershed agricultural/range	360	0.500	0.748
RD.DEN[1]	Road density (m ha^{-1})	360	0.461	0.745
POPEST[1]	Est. human population in watershed	360	0.452	0.732
HOUSINGU[1]	Est. no. housing units in watershed	360	0.389	0.690
NONRES[1]	% watershed non-residential urban land	360	0.358	0.638
TOT.RD[1]	Meters of road in watershed	360	0.260	0.594
URB.TOT[1]	% watershed urban (including mines/quarries)	360	0.426	0.695
Birds				
FSTOTO	% stops present, toler. guild, tolerants	205	0.537	0.824
SIND[1]	Avg. no. of individuals per stop	205	0.499	0.771
DIDTOM[1]	Mean indiv./stop, diet guild, omnivore	205	0.463	0.755
DIDTSE[1]	Mean indiv./stop, diet guild, seeds	205	0.416	0.712
DIDTIN[1]	Mean indiv., insect diet guild	205	0.387	0.710
SRICH	Avg. no. of species per stop	205	0.384	0.706
FSFDHG	% stops pres., forag. guild, hover and glean	205	−0.343	0.683
FSNERE	% stops pres., neotr. guild, residents	205	0.405	0.663
RICH	total species on lake	205	−0.115	0.608
FSTOIN	% stops pres., toler. guild, intolerants	205	0.085	0.525
FSNENE	% stops pres., neotr. guild, migrants	205	0.056	0.524
Diatoms				
RICHNESS	Number of non-missing values	240	−0.311	0.773
DVRSTY	Diatom species diversity (H′)	240	−0.132	0.656
Fish				
SEN.SP	No. of sensitive (intol/hab-sen) species	195	−0.305	0.699
SEN.IND[1]	No. of sensitive (intol/hab-sen) indiv.	195	−0.300	0.696
TOL.SP[1]	No. of tolerant species	195	0.210	0.625
INV.IND[1] *	No. of insectivorous individuals	195	−0.211	0.619
INV.SP[1] *	No. of insectivorous species	195	−0.195	0.608
TOP.IND[1]	No. indiv. at top of food chain	195	0.120	0.600
EXOT.IND[1]	No. of non-native individuals	195	0.162	0.597
TOL.IND[1]	No. of tolerant individuals	195	0.065	0.569
NAT.IND[1]	No. of native individuals	195	−0.056	0.559
PISC.SP	No. of piscivorous species	195	0.069	0.558
OMNI.IND[1]	No. of omnivorous individuals	195	0.080	0.554

(*Continued on next page*)

(*Continued*)

Name	Variable	n	r_S	ROC area
EXOT.SP	No. of non-native species	195	0.111	0.549
NAT.SP[1]	No. of native species	195	−0.047	0.527
PISC.IND[1]	No. of piscivorous individuals	195	−0.023	0.515
TOT.SP	Total no. of species caught	195	−0.008	0.514
OMNI.SP*	No. of omnivorous species	195	0.025	0.507
TOT.IND[1]	Total no. of individuals caught	195	0.001	0.482
TOP.SP	No. of species at top of food chain	195	0.073	0.462
Zooplankton				
TOTZOOP[1]	Total number of zooplankton indiv.	350	0.383	0.756
MICRO.N[1]	No. indiv. collected with 48 μ net	349	0.381	0.753
ROTH[1]	No. of herbivorous rotifer individuals	350	0.344	0.735
CRH[1]	No. herbivorous crustacean individuals	350	0.342	0.726
MACRO.R	No. species collected with 202 μ net	350	−0.261	0.717
SCLADH[1]	No. small herbivorous cladoceran indiv.	350	0.314	0.716
LGCLAD.R	No. of large cladoceran species	350	−0.262	0.716
NAUPLIH[1]	No. nauplii individuals	350	0.346	0.711
CAL.R	No. of calanoid species	350	−0.306	0.709
CYCCOPH[1]	No. herbivorous cyclopoid copepodite indiv.	350	0.322	0.692
ACALH[1]	No. adult herbivorous calanoid indiv.	350	−0.340	0.688
CROM[1]	No. omnivorous crustacean indiv.	350	0.312	0.660
ACYCOM[1]	No. adult omnivorous cyclopoid indiv.	350	0.314	0.634
CR.R	No. of crustacean species	350	−0.096	0.630
ROTOM[1]	No. omnivorous rotifer individuals	350	0.176	0.613
CROM.R	No. omnivorous crustacean species	350	0.087	0.613
MACRO.N[1]	No. indiv. collected with 202 μ net	350	0.084	0.591
CRH.R	No. herbivorous crustacean species	350	−0.023	0.582
SMCLAD.R	No. small cladoceran species	350	0.170	0.579
ACLADOM	No. adult omnivorous cladoceran indiv.	350	−0.226	0.578
CALCOPH[1]	No. herbivorous calanoid copepodite indiv.	350	−0.209	0.558
MICRO.R	No. species collected with 48 μ net	350	0.188	0.539
CHAOB.R	No. of chaoborid larvae species	350	−0.092	0.531
LCLADH[1]	No. large herbivorous cladoceran indiv.	350	0.107	0.522
ROT.R	No. of rotifer species	350	0.151	0.517
CHAOBOM[1]	No. of chaoborid larvae individuals	350	−0.068	0.510
SMCYC.R	No. of small cyclopoid species	350	0.095	0.489
LGCYC.R	No. of large cyclopoid species	350	0.030	0.464
TOTZOOPR	Total zooplankton richness	350	0.071	0.462
ROTOM.R	No. omnivorous rotifer species	350	0.025	0.421

References

Adams, S. M. (ed): 2002, *Biological Indicators of Ecosystem Stress,* American Fisheries Society, Bethesda, Maryland.

Allen, A. P., Whittier, T. R., Kaufmann, P. R., Larsen, D. P., O'Connor, R. J., Hughes, R. M., Stemberger, R. S., Dixit, S. S., Brinkhurst, R. O., Herlihy, A. T. and Paulsen, S. G.: 1999, 'Concordance of taxonomic richness patterns across multiple assemblages in lakes of the northeastern United States', *Can. J. Fish. Aquat. Sci.* **56**, 739–747.

Baker, J. R., Peck, D. V. and Sutton, D. W. (eds): 1997, *Environmental Monitoring and Assessment Program Surface Waters: Field Operations Manual for Lakes,* EPA/620/R-97/001, U.S. Environmental Protection Agency, Washington D.C.

Bryce, S. A., Hughes, R. M. and Kaufmann, P. R.: 2002, 'Development of a bird integrity index: Using bird assemblages as indicators of riparian condition', *Environ. Manage.* **30**, 294–310.

Burnham, K. P. and Anderson, D. R.: 2002, *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach,* Springer, New York.

Carignan, R. and Steedman, R. J.: 2000, 'Impacts of watershed perturbations on aquatic ecosystems', *Can. J. Fish. Aquat. Sci.* **57**(Suppl. 2), 1–4.

Carignan, V. and Villard, M.-A.: 2002, 'Selecting indicator species to monitor ecological integrity: A review', *Environ. Monit. Assess.* **78**, 45–61.

Chaloud, D. J. and Peck, D. V. (eds): 1994, *Environmental Monitoring and Assessment Program: Integrated Quality Assurance Project Plan for the Surface Waters Resource Group, 1994 Activities,* EPA/600/X-91/080, U.S. Environmental Protection Agency, Office of Research and Development, EMSL, Las Vegas, Nevada.

Cottingham, K. L. and Carpenter, S. R.: 1998, 'Population, community, and ecosystem variates as ecological indicators: phytoplankton responses to whole-lake enrichment', *Eco. Appl.* **8**, 508–530.

Cude, C. G.: 2001, 'Oregon water quality index: A tool for evaluating water quality management effectiveness', *J. Am. Water Resour. Assoc.* **37**, 125–138.

Davis, W. S. and Simon, T. S. (eds): 1995, *Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making,* Lewis Publishers, Boca Raton, Florida.

Gergel, S. E., Turner, M. G., Miller, J. R., Melack, J. M. and Stanley, E. H.: 2002, 'Landscape indicators of human impacts to riverine systems', *Aquat. Sci.* **64**, 118–128.

Hanley, J. A. and McNeil, B. J.: 1982, 'The meaning and use of the area under a receiver operating characteristic (ROC) curve', *Radiology* **143**, 29–36.

Herlihy, A. T., Stoddard, J. L. and Johnson, C. B.: 1998, 'The relationship between stream chemistry and watershed land cover data in the mid-Atlantic region, U.S.', *Water Air Soil Pollut.* **105**, 377–386.

Hughes, R. M.: 1995, 'Defining acceptable biological status by comparing with reference conditions', in: W.S. Davis and T.P. Simon (eds), *Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making,* Lewis Publishers, Boca Raton, Florida, pp. 31–47.

Hughes, R. M., Howlin, S. and Kaufmann, P. R.: 2004, 'A biointegrity index (IBI) for coldwater streams of western Oregon and Washington', *Trans. Am. Fish. Soc.* **133**, 1497–1515.

Hughes, R. M., Kaufmann, P. R., Herlihy, A. T., Kincaid, T. M., Reynolds, L. and Larsen, D. P.: 1998, 'A process for developing and evaluating indices of fish assemblage integrity', *Can. J. Fish. Aquat. Sci.* **55**, 1618–1631.

Hunsaker, C. T. and Carpenter, D. E. (eds): 1990, *Environmental Monitoring and Assessment Program: Ecological Indicators,* Atmospheric Research and Exposure Assessment Laboratory, Office of Research and Development, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina.

Jackson, L. E., Kurtz, J. C. and Fisher, W. S. (eds): 2000. *Evaluation Guidelines for Ecological Indicators,* EPA/620/R-99/005, U.S. Environmental Protection Agency, Office of Research and Development, Research Triangle Park, North Carolina.

Klemm, D. J., Blocksom, K. A., Fulk, F. A., Herlihy, A. T., Hughes, R. M., Kaufmann, P. R., Peck, D. V., Stoddard, J. L., Thoeny, W. T. and Griffith, M. B.: 2003, 'Development and evaluation of a macroinvertebrate biotic integrity index (MBII) for regionally assessing mid-Atlantic Highlands streams', *Environ. Manage.* **31**, 656–669.

Kottferova, J., Korenekova, B., Breyl, I. and Nadaskay, R.: 1996, 'Free-living animals as indicators of environmental pollution by chlorinated hydrocarbons', *Toxicol. Environ. Chem.* **53**, 19–24.

Meybeck, M.: 1998, 'Man and river interface: Multiple impacts on water and particulates chemistry illustrated in the Seine river basin', *Hydrobiologia* **373–374**, 1–20.

Murtaugh, P. A.: 1996, 'The statistical evaluation of ecological indicators', *Ecol. Appl.* **6**, 132–139.

Murtaugh, P. A.: 2003, 'On detecting hump-shaped relationships in ecology: A bootstrap test for monotonicity', *Environmetrics* **14**, 1–6.

OECD: 1982, *Eutrophication of Waters. Monitoring, Assessment and Control.* Organisation for Economic Cooperation and Development, Paris.

R Development Core Team: 2004, R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-00-3, URL http://www.R-project.org.

Ramsey, F. L. and Schafer, D. W.: 2002, *The Statistical Sleuth: A Course in Methods of Data Analysis,* 2nd ed, Duxbury, Pacific Grove, CA.

Rusak, J. A., Yan, N. D., Sorners, K. M., Cottingham, K. L., Micheli, F., Carpenter, S. R., Frost, T. M., Paterson, M. J. and McQueen, D. J.: 2002, 'Temporal, spatial, and taxo-nomic patterns of crustacean zooplankton variability in unmanipulated north-temperate lakes', *Limnol. Oceanogr.* **47**, 613–625.

Smith, E. P.: 2002, 'Statistical considerations in the development, evaluation, and use of biomarkers in environmental studies', in: S. M. Adams (ed), *Biological Indicators of Ecosystem Stress,* American Fisheries Society, Bethesda, Maryland, pp. 565–590.

Stemberger, R. S. and Lazorchak, J. M.: 1994, 'Zooplankton assemblage responses to disturbance gradients', *Can. J. Fish. Aquat. Sci.* **51**, 2435–2447.

Sullivan, T. J. and Cosby, B. J.: 1998, 'Modeling the concentration of aluminum in surface waters', *Water Air Soil Pollut.* **105**, 643–659.

Xu, F.-L., Tao, S., Dawson, R. W., Li, P.-G. and Cao, J.: 2001, 'Lake ecosystem health assessment: Indicators and methods', *Water Res.* **35**, 3157–3167.