RIVER RESEARCH AND APPLICATIONS

River. Res. Applic. **23**: 683–697 (2007)

Published online 20 April 2007 in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/rra.1011

FIRST STEPS IN DEVELOPING A MULTIMETRIC MACROINVERTEBRATE INDEX FOR THE OHIO RIVER

JEROMY M. APPLEGATE,[a*] PAUL C. BAUMANN,[b] ERICH B. EMERY[c] and MATTHEW S. WOOTEN[d]

[a] *United States Fish and Wildlife Service, Reynoldsburg, OH 43068, USA*
[b] *United States Geological Survey, 2021 Coffey Road, Columbus, OH 43210, USA*
[c] *Ohio River Valley Water Sanitation Commission, 5735 Kellogg Avenue, Cincinnati, OH 45228, USA*
[d] *Northern Kentucky Sanitation District No. 1, 1045 Eaton Drive, Fort Wright, KY 41017, USA*

ABSTRACT

The causes of degradation of aquatic systems are often complex and stem from a variety of human influences. Comprehensive, multimetric biological indices have been developed to quantify this degradation and its effect on aquatic communities, and measure subsequent recovery from anthropogenic stressors. Traditionally, such indices have concentrated on small-to medium-sized streams. Recently, however, the Ohio River Fish Index (ORFIn) was created to assess biotic integrity in the Ohio River. The goal of the present project was to begin developing a companion Ohio River multimetric index using benthic macroinvertebrates. Hester–Dendy multiplate samplers were used to evaluate benthic macroinvertebrate assemblages in relation to a gradient of water quality disturbance, represented by varying distances downstream of industrial and municipal wastewater outfalls in the Ohio River. In August 1999 and 2000, samplers were set every 100 m downstream of outfalls (12 outfalls in 1999, 22 in 2000) for 300–1000 m, as well as at upstream reference sites. Candidate metrics ($n = 55$) were examined to determine which have potential to detect changes in water quality downstream of outfalls. These individual measures of community structure were plotted against distance downstream of each outfall to determine their response to water quality disturbance. Values at reference and outfall sites were also compared. Metrics that are ecologically relevant and showed a response to outfall disturbance were identified as potentially valuable in a multimetric index. Multiple box plots of index scores indicated greater response to outfall disturbance during periods of low-flow, and longitudinal river-wide trends. Evaluation of other types of anthropogenic disturbance, as well as continued analysis of the effects of chemical water quality on macroinvertebrate communities in future years will facilitate further development of a multimetric benthic macroinvertebrate index to evaluate biotic integrity in the Ohio River. Copyright © 2007 John Wiley & Sons, Ltd.

KEY WORDS: benthic macroinvertebrates; biological monitoring; biotic integrity; large river; multimetric index; Ohio River

Received 10 January 2007; Accepted 25 January 2007

INTRODUCTION

The index of biotic integrity (IBI) (Karr, 1981; Karr *et al*., 1986) was the first widely accepted multimetric index to monitor ecological integrity of aquatic systems. The IBI uses fish assemblages to evaluate biotic integrity of small to medium-sized streams. Each metric in the index measures a different component of the community, and together they give a total measure of ecological health at a site. This method is an improvement over traditional physicochemical monitoring of aquatic systems because it integrates many types of stream disturbance (Karr, 1991). The IBI was originally developed for application in wadeable warmwater streams of the Midwestern United States, but has been adapted for use in other aquatic systems including cool-water (Leonard and Orth, 1986) and cold-water streams (Mundahl and Simon, 1999), wetlands (Mack *et al*., 2000), marine estuaries (Deegan *et al*., 1993; Engle and Summers, 1999) and Great Lakes near-shore waters (OEPA [Ohio Environmental Protection Agency], 1997; Thoma, 1999). Others have modified it for use on several continents (Ganasan and Hughes, 1998;

---

*Correspondence to: Jeromy M. Applegate, 6950 Americana Parkway, Suite H, Reynoldsburg, OH 43068, USA.
E-mail: jeromy_applegate@fws.gov

Copyright © 2007 John Wiley & Sons, Ltd.

Toham and Teugels, 1999; Belpaire *et al*., 2000), and variations have been developed using different taxonomic groups, including aquatic vegetation (McCormick and Stevenson, 1998; Hill *et al*., 2000; Mack *et al*., 2000) and benthic macroinvertebrates (Kerans and Karr, 1994; Deshon, 1995; Barbour *et al*., 1996; Fore *et al*., 1996).

Steps have recently been taken to develop multimetric indices to assess biotic integrity in large river systems (Lyons *et al*., 2001). Emery *et al*. (2003) and Simon and Emery (1995) developed the Ohio River Fish Index (ORFIn) to measure biotic integrity in the Ohio River. With the completion of the ORFIn, the need has arisen for a second index, using a different organism group, to complement the existing fish index. Seegert (2000) noted that index development has been considerably slower in large rivers, and overall, macroinvertebrate research in these systems has been scarce (Payne and Miller, 1989; Thorp, 1992). Lack of macroinvertebrate index development in large rivers has been attributed to the physical difficulties involved in sampling (Beckett and Keyes, 1983; Johnson *et al*., 1995), as well as data variability and lack of reference sites (Seegert, 2000).

The objective of this research was to begin the process of developing a macroinvertebrate index by (1) identifying metrics that respond to changes in chemical water quality in the Ohio River, (2) compiling these metrics into a preliminary index and (3) evaluating how this preliminary index responds to changes in chemical water quality downstream of outfalls.

STUDY AREA

The Ohio River begins at the confluence of the Monongahela and Allegheny Rivers in Pittsburgh, Pennsylvania and flows in a southwest direction for 1578 km to Cairo, Illinois, where it empties into the Mississippi River. The river crosses four Level III ecoregions: the Western Allegheny Plateau, the Interior Plateau, the Interior River Lowland and the Mississippi Alluvial Plain (Omernik, 1987). It is impounded by 20 navigational dams that, in combination with regular dredging, provide a 2.75-m minimum water depth, facilitating transport of more than 200 million tons of cargo each year (ORSANCO [Ohio River Valley Water Sanitation Commission], 2001). The Ohio River has a 528 000 km^2 drainage area that encompasses parts of 14 states. More than 25 million people, almost 10% of the US population, live in the Ohio River Basin. Over 600 permitted dischargers are located on the Ohio River, originating from various industries, including metal and chemical manufacturers, power-generating facilities and municipal wastewater treatment plants (ORSANCO, 1994).

SAMPLE COLLECTION

Macroinvertebrate sample data used in this project were compiled from samples collected by the ORSANCO over a 10-year period from 1991–2000, as part of its bioassessment programme for the Ohio River. Macroinvertebrates were sampled using modified Hester–Dendy multiplate samplers (Hester and Dendy, 1962), which consist of a series of 3-mm-thick masonite squares (60 cm^2 each), held together with an eyebolt and spaced with smaller masonite squares. For this study, one sampling unit included five individual Hester–Dendy samplers, with a total substrate area of 0.465 m^2, attached to a cement foundation block. Sampling coincided with the low-flow period of the year (mid August to early October). Sampling units were submerged in near-shore locations in approximately 0.75–1.0 m of water and secured to the river bottom with steel reinforcement bar. They were left to colonize for 6 weeks, and were retrieved by cutting them from the concrete blocks and carefully placing them in buckets to minimize loss of organisms. Samplers were disassembled in the buckets and organisms were scraped from the plates into the bucket. The resulting slurry was poured through a number 30 standard sieve (0.595-mm openings) and organisms were placed in plastic jars and preserved in 10% formalin. Macroinvertebrates were identified to the lowest possible taxonomic level as outlined by ORSANCO (1999).

From 1991–1999, macroinvertebrate sampling locations varied longitudinally throughout the river. They were located without regard to wastewater discharges, and produced a river-wide dataset of near-shore macroinvertebrate colonization of Hester–Dendy samplers (Table I). In 1999 and 2000, samplers were placed at varying distances below wastewater outfalls, as well as at upstream reference sites for each outfall (Table I). Outfalls originated from various industries, including power-generating facilities, chemical manufacturers, metals manufacturers and municipal wastewater treatment plants.

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/rra

Table I. Summary of Ohio River macroinvertebrate sampling design, 1991–2000

Year	Sampling design
1991–1996	Sampling concentrated in one or two navigational pools each year Approximately 15–25 sampling units retrieved each year Samplers located without regard to outfalls
1997–1998	Samplers placed every 8 km along entire length of river Locations offset from 1997 to 1998, resulting in river-wide dataset at 4 km intervals
1999–2000	Wastewater outfall sites sampled river-wide 1999: 12 outfalls 2000: 22 outfalls
	One sampling unit placed at each distance downstream of an outfall 1999: 0, 50, 100, 200, 300, 400, 500, 600, 700, 800, 900 and 1000 m downstream 2000: 0, 50, 100, 200 and 300 m downstream
	Reference samplers placed at an upstream site for each outfall 1999: same scheme as those at the outfall (0, 50, 100 m, etc.) 2000: 3 sampling units placed in triplicate at a single location

METRIC EVALUATION

Examination of potential metrics was conducted by observing each metric's response to a water quality gradient represented by varying distances downstream of wastewater outfalls, and comparison with upstream reference sites for each outfall. Potential metrics ($n = 55$) were gathered from existing literature and from a list previously proposed by the ORSANCO Macroinvertebrate Advisory Panel, a group of agency, industry and academic macroinvertebrate experts assembled to assist in macroinvertebrate biocriteria development. Other prospective metrics were proposed based on macroinvertebrate assemblages observed from Hester–Dendy samplers in this study. All proportional metrics (e.g. per cent Diptera, per cent Hydroptilidae) were calculated omitting zebra mussels (*Dreissena polymorpha*), which irregularly dominated some samples.

Initial screening of the metrics was conducted using outfall data collected in 1999. Because of low sampler retrieval rates in 1999, analysis was limited to five outfalls where a sufficient number of sampling units were collected. Metric values were plotted at reference sites and at each distance downstream of each individual outfall to determine whether the metric exhibited the hypothesized response to disturbance. A metric was deemed to have responded in a hypothesized manner to disturbance if it was depressed (or enhanced in negative metrics) in relation to reference samples, but never recovered to reference values at the greatest distance sampled below the outfall (Figure 1A), was depressed and then recovered to reference values a short distance downstream (Figure 1B), or was depressed and recovered gradually downstream (Figure 1C). If a metric did not respond at an outfall in one of these three ways, we determined that it had not responded in the hypothesized manner to that outfall (Figure 1D). At some outfalls, some metrics responded in an opposite manner from what was hypothesized (Figure 1E).

Each metric was evaluated at each of the five outfalls in 1999, following the criteria mentioned above, and unresponsive metrics were eliminated from further consideration if they fell into one of the following four categories (Table IIA):

1. Low numbers: metric values were prohibitively low in all of the samples collected from at least four of the five sites including reference samples. A prohibitively low metric value was less than 1% in proportional metrics (e.g. per cent *Corbicula*), and less than 3% in structural metrics (e.g. number of Coleoptera taxa).
2. Low response: metric exhibited the hypothesized response at only one or fewer of the five outfalls.
3. Variable response: metric exhibited the hypothesized response at two of the five outfalls, but exhibited the opposite of the hypothesized result at at least two other outfalls.
4. Redundant: metric was redundant (Pearson correlation coefficient ≥ 0.99, $p < 0.001$) with other, more commonly accepted metrics.

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/rra

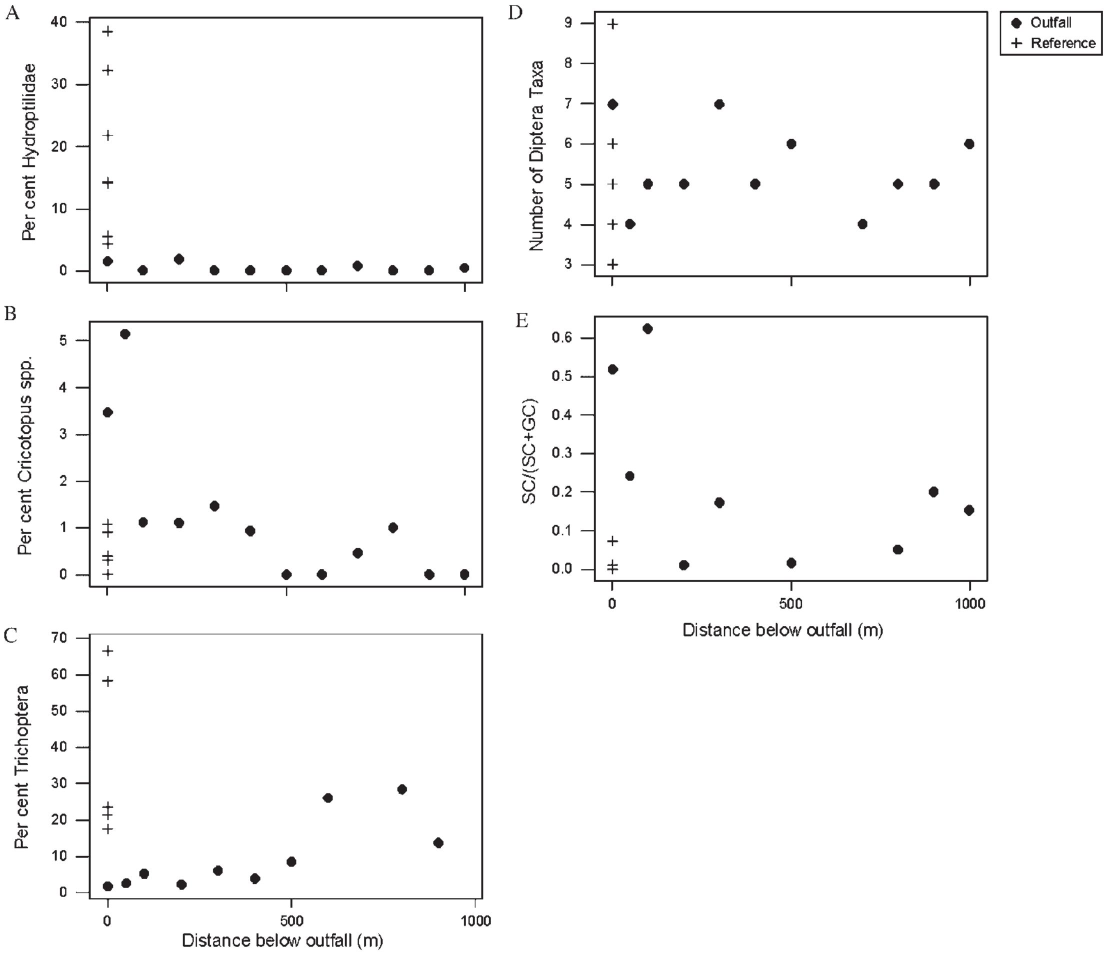

Figure 1. Examples of metric response at individual reference and outfall sites in 1999: (A) per cent Hydroptilidae at a metals manufacturer, illustrating a metric that was depressed in relation to upstream reference samples and never recovered to reference values at 1000 m downstream of the outfall; (B) per cent *Cricotopus* at a metals manufacturer, illustrating a metric that was enhanced in relation to reference samples, and then recovered to reference values a short distance downstream of the discharge; (C) per cent Trichoptera at a municipal wastewater treatment plant, illustrating a metric that was depressed, and recovered to reference levels gradually; (D) number of Diptera Taxa at a thermal outfall, illustrating a metric that did not respond to disturbance in a hypothesized manner; (E) the Ratio of Scrapers to Gathering Collectors at a municipal wastewater treatment plant, illustrating a metric that responded in an opposite manner from what was hypothesized

Screening the metrics using 1999 data resulted in elimination of approximately half of the least-responsive metrics, leaving a much more manageable group of variables to analyze with the larger dataset from 2000. Using the three criteria mentioned above to determine if a metric responded in a hypothesized manner to an individual outfall, the metrics retained for further analyses were evaluated with sample data collected at 21 outfalls in 2000 (one outfall was eliminated because of low sampler retrieval). In addition to these metrics, individual taxa were evaluated in the same manner, for potential use in a per cent tolerant individuals metric and a per cent intolerant individuals metric. Multimetric biotic indices have historically included metrics that measure the per cent of a sample comprised of tolerant or intolerant taxa (OEPA, 1988; Kerans and Karr, 1994). The taxa evaluated for potential use in either of these two metrics were those considered by the Ohio Environmental Protection Agency to be intolerant, moderately intolerant, moderately tolerant, tolerant or very tolerant, based on tolerances to a variety of anthropogenic disturbances. Many of these taxa were sampled very infrequently during our research in the Ohio River, and were, therefore, removed from consideration. Because samples at all sites contained very low numbers of individuals from intolerant taxa, no intolerant individuals metric could be created.

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/rra

Table IIA. Metrics eliminated from consideration after initial screening with 1999 outfall (n = 5) data

Metric	Hypothesized effect of disturbance	Reason for rejection
Per cent leeches	Increase	Low numbers
Number of zebra mussels	Decrease	Variable response
Per cent pulmonate snails	Increase	Variable response
Per cent prosobranch snails	Decrease	Variable response
Per cent *Corbicula* spp.	Decrease	Low numbers
Per cent Ephemeroptera	Decrease	Variable response
Number of Ephemeroptera taxa	Decrease	Low numbers
Per cent Odonata	Increase	Low numbers
Number of Odonata taxa	Decrease	Low numbers
Number of Trichoptera taxa	Decrease	Low numbers
Per cent Coleoptera	Decrease	Low numbers
Number of Coleoptera taxa	Decrease	Low numbers
Per cent Chironomidae	Increase	Redundant with % Diptera
Number of Chironomidae taxa	Decrease	Redundant with # Diptera taxa
Per cent Tanytarsini	Decrease	Low numbers
Per cent *Ablabesmyia* spp.	Increase	Low response
Per cent *Dicrotendipes* spp.	Increase	Variable response
Per cent *Phaenopsectra* spp.	Increase	Low numbers
Per cent *Polypedilum* spp.	Increase	Low response
Per cent dominant taxon	Increase	Low response
Per cent 2 dominant taxa	Increase	Low response
Per cent 3 dominant taxa	Increase	Low response
Per cent 4 dominant taxa	Increase	Low response
Per cent 5 dominant taxa	Increase	Low response
Per cent predators	Decrease/variable	Low response
Per cent filtering collectors (FC) and gathering collectors (GC)	Increase	Variable response
Per cent predators (except for chironomids and flatworms)	Decrease	Variable response
Per cent scrapters (SC)	Decrease	Low numbers
(SC/(SC+GC))	Decrease	Variable response

After determining whether each of the remaining metrics showed a predictable response at each outfall, individual metric responsiveness was quantified by determining the proportion of outfalls to which they responded in the hypothesized manner. It was noted in the analysis that some outfalls (two from 1999 and six from 2000) predictably influenced only a few potential metrics. These outfalls were therefore removed in order to improve resolution in detecting those metrics that were most responsive to wastewater disturbance. Outfalls that showed less than 40% metric response, and less than 20% response to the tolerant/intolerant taxa list were removed. We then ranked the metrics according to the proportion of outfalls to which they responded in the hypothesized manner (Tables IIB and IIC).

Using this ranking as a guide, we selected metrics to include in preliminary indices. For comparison purposes, two different groups of metrics were chosen to be included in two separate indices. One group, assembled into the ‘Panel Index’, was a compilation of metrics selected during conference with the ORSANCO Macroinvertebrate Advisory Panel. This group of 12 metrics consisted of those that were responsive to outfall disturbance and are considered to be ecologically significant (Table IIB). A second group of metrics was assembled into the ‘Percentage Index’, and included only those metrics that exhibited the hypothesized response at more than 50% of outfalls (Table IIB).

Although functional feeding group metrics were evaluated, and some appeared to exhibit the hypothesized response (e.g. shredder metrics), they were not chosen for inclusion in either of these two indices because feeding guild can vary significantly in a single species with life-history stage, season and other ecological features (Thorp and Covich, 1991).

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/rra

Table IIB. Metrics included in preliminary multimetric macroinvertebrate indices

Metric	Hypothesized effect of disturbance	Per cent of outfalls (n = 18) showing predicted response	Method of data division
Per cent Diptera[a,b]	Increase	78	Quadrisection
Per cent Hydroptilidae[a,b]	Decrease	67	Double bisection
Per cent tolerant individuals[a,b]	Increase	61	Triple bisection
Menetus dilatatus[c]		60	
Chironomus spp.[c]		40	
Polypedilum illinoense[c]		27	
Physella spp.[c]		27	
Tricladida[c]		60	
Dicrotendipes lucifer[c]		33	
Glyptotendipes spp.[c]		27	
Cricotopus bicinctus[c]		20	
Ablabesmyia rhamphe[d]		13	
Hydrobiidae[d]		13	
Nandocladius distinctus[d]		13	
Number of Ephemeroptera (E) individuals[b]	Decrease	56	Double bisection
E and Trichoptera (T)/Chironomidae[a,b]	Decrease	56	Triple bisection
Per cent Amphipoda[a,b]	Decrease	56	Double bisection
Number of E, Plecoptera (P), and T taxa[a]	Decrease	50	Quadrisection
Per cent Tanypodinae[a]	Increase	50	Double bisection
Total number of individuals[a]	Decrease	44	Triple bisection
Per cent EPT individuals[a]	Decrease	44	Quadrisection
Total number of taxa[a]	Decrease	39	Quadrisection
Per cent Oligochaeta[a]	Increase	28	Double bisection
Number of Diptera taxa[a]	Decrease	22	Quadrisection

[a] Metric included in Panel Index.
[b] Metric included in Percentage Index.
[c] Taxon included in per cent tolerant individuals metric.
[d] Taxon not included in per cent tolerant individuals metric due to low response.

RATIONAL FOR METRIC INCLUSION

Total number of taxa

This metric is a measure of the overall taxonomic diversity of a site. It is a primary component of ecological integrity (OEPA, 1988) and is based on the principle that stable, healthy communities in warm-water midwestern streams have high species richness. Increasing diversity correlates with increasing health of the assemblage and suggests that niche space, habitat and food sources are adequate to support survival and reproduction of many species (Barbour *et al*., 1996). Beckett and Keyes (1983) found that heavily polluted areas of the Ohio River contained reduced species richness in comparison to less polluted sections. Taxa richness has also been found to decrease in the presence of acid mine drainage (Dills and Rogers, 1974) and insecticide contamination (Schulz and Liess 1999).

Total number of individuals

This metric can be a measure of disturbance at a site, because total density often decreases in response to toxic pollutants such as heavy metals (Lenat *et al*., 1980; Beckett and Keyes, 1983), acid mine drainage (Dills and Rogers, 1974) and insecticides (Schulz and Liess, 1999). This metric may be variable, however, in that density of some species may increase in response to organic enrichment.

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/rra

Table IIC. Metrics retained for evaluation following initial screening using 1999 data, but not included in either preliminary index

Metric	Hypothesized effect of disturbance	Per cent of outfalls (n = 18) showing predicted response	Reason for rejection
Per cent non-insects	Increase	22	Low response
Per cent Ephemeroptera and Trichoptera taxa	Decrease	44	Replaced with # of EPT taxa
Per cent Trichoptera	Decrease	44	Chose per cent Hydroptilidae (greater response)
Per cent Polycentropodidae	Decrease	39	Chose per cent Hydroptilidae (greater response)
Per cent Chironominae	Increase	56	Chose per cent Diptera (greater response)
Per cent Orthocladiinae	Increase	44	Chose per cent Diptera (greater response)
Per cent *Cricotopus* spp.	Increase	44	Chose per cent Diptera (greater response)
Per cent filtering collectors (FC)	Increase/variable	39	Functional feeding group
Per cent gathering collectors (GC)	Increase/variable	39	Variable response
Per cent shredders (SH)	Decrease	56	Functional feeding group
Per cent shredders and scrapers (SC)	Decrease	50	Functional feeding group
(SC/(SC+FC))	Decrease	39	Functional feeding group/low response
(SH/(SH+FC))	Decrease	56	Functional feeding group
(SH/(SH+GC))	Decrease	44	Functional feeding group/low response

Number of Diptera taxa

The order Diptera (true flies) represents a major component of Ohio River macroinvertebrate samples. In large rivers, as habitat conditions become more homogenous, there tends to be a greater number of dipteran individuals, but a reduction in species diversity within the order (OEPA, 1988). Losos (1984) found that the diversity of Chironomidae taxa decreased as pollution and eutrophy increased in mountain streams. However Lenat (1983) suggested that Chironomidae taxa richness may increase under moderate pollution, responding in part to reduced competition with less tolerant groups such as Trichoptera and Plecoptera.

Per cent Diptera

Dipterans display a wide range of tolerances but generally are thought to be a tolerant order and increase in proportion under disturbed conditions (Barbour *et al*., 1996). Losos (1984) found that the density of chironomids increased as the level of pollution and eutrophy increased in mountain streams, and Dills and Rogers (1974) found dipterans to be proportionally greater at sites polluted with acid mine drainage.

Per cent Tanypodinae

The chironomid subfamily Tanypodinae is a group of active predators (Hilsenhoff, 1991) that are thought to be tolerant. An increase in individuals from this group has been shown to coincide with an increase in disturbance (Britt *et al*., 1973).

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/rra

Number of EPT taxa, per cent EPT individuals, number of Ephemeroptera individuals

Number of EPT taxa is comprised of the number of taxa from the insect orders Ephemeroptera, Plecoptera and Trichoptera, and per cent EPT individuals is the per cent of the sample comprised of individuals from these three orders. These metric values are thought to decrease as disturbance levels increase.

Plecoptera larvae (stoneflies) can be found in every type of unpolluted lotic habitat (Pennak, 1978). They generally lack extensive gills and are intolerant of low dissolved oxygen concentrations (Hilsenhoff, 1991), and taxa richness in this order typically decreases as a result of pollution (Lenat, 1983). Ephemeroptera larvae (mayflies) exhibit a variety of feeding functions (Pennak, 1978) and most species are highly pollution-sensitive. They are highly vulnerable to acidification (Fiance, 1978; Hall *et al*., 1985; Hermann and Anderson, 1986) and are usually one of the first groups to disappear in the presence of disturbance (OEPA, 1988). Trichoptera (caddisflies), like mayflies, exhibit a variety of feeding functions. Caddisflies are generally viewed as a moderately tolerant group (OEPA, 1988); however, few species can tolerate heavy pollutional stress (Kerans and Karr, 1994) and, as a result, their abundance is expected to decrease with disturbance (Barbour *et al*., 1996).

Per cent Hydroptilidae

This caddisfly family feeds primarily by scraping periphytic diatoms and piercing and sucking the cellular content of filamentous algae (Pennak, 1978; OEPA, 1988; Hilsenhoff, 1991). They are primarily a large-river family, relying on open canopies and slower-moving water for optimal plant growth (OEPA, 1988). Because of their narrow feeding niche, they comprise one of the more sensitive families of caddisflies, and should decline with disturbance. Because they feed on plant material, however, hydroptilids may be fairly tolerant to nutrient enrichment.

Ratio of Ephemeroptera and Trichoptera to Chironomidae

This metric is a tolerance metric that is expressed as the ratio of the relatively intolerant Ephemeroptera and Trichoptera orders to the relatively tolerant dipteran family Chironomidae.

Per cent Amphipoda

Amphipods are a group of crustaceans typically restricted to permanent bodies of relatively cool, clean and well-oxygenated water (Pennak, 1978; Covich and Thorp, 1991). Juvenile amphipods eat microbial foods such as bacteria and some algae, while adults are omnivorous scavengers (Pennak, 1978). Amphipods are sensitive to a number of toxic heavy metals (Abel and Barlocher, 1988; Borgmann and Munawar, 1989) and are thought to decrease as the level of disturbance increases (Barbour *et al*., 1996).

Per cent tolerant individuals

This metric is composed of a group of eight taxa that showed an increase in relative abundance at a variety of outfall sites, and are viewed by the OEPA to be tolerant to anthropogenic disturbance.

Per cent Oligochaeta

Aquatic oligochaetes include segmented worms of which most species directly consume substrate and digest the organic components (Pennak, 1978). Oligochaetes have been collected most abundantly in streams and rivers polluted with sewage (Pennak, 1978) and have been found to increase in abundance with an increase in pollution (Barbour *et al*., 1996). They can dominate the macroinvertebrate fauna of heavily polluted rivers (Eyres *et al*., 1978).

Copyright © 2007 John Wiley & Sons, Ltd.

SCORING METHODS

The scoring procedures used in the two preliminary indices were based on established methods used in previous indices (OEPA, 1988; Mack *et al*., 2000; Emery *et al*., 2003). After selecting metrics to include in our two test indices, we assigned index scores to each of these metrics, based on distribution of metric values recorded at sample sites ($n = 322$) throughout the entire river. In an effort to minimize the effects of point source discharges, all sites less than 1 km from a permitted discharger were excluded from analysis. For each metric, values were plotted against river km (which served as a surrogate for drainage area) (Figure 2), and examined to detect any natural river-wide trends in metric values. Although some metrics exhibited depressed values in the upper portion of the Ohio River, as in per cent EPT (Figure 2A), we felt this anomaly was most likely a function of the historically degraded conditions caused by heavy industry in the upper river (Beckett and Keyes, 1983), as opposed to a natural upstream to downstream trend for which metric scoring lines should be adjusted.

To divide the range of metric values into scoring categories, for 9 of the 12 metrics a straight line was first drawn on the graph at the 95th percentile of the data (Figure 2A–C). In structural metrics (e.g. total number of taxa), a line drawn at the maximum value (MV) was used (Figure 2D) because structural metrics lacked the significant outliers often seen in proportional metrics. In addition, we felt that the present macroinvertebrate species richness in the Ohio River is well below historical levels, and using the maximum value line would reduce the potential for artificially inflating metric scores.

The area beneath the 95% or MV line was then divided into four sections, with each section receiving a score of 0, 2, 4 or 6. A score of 0 was given to the area of the graph representing degraded conditions, while a score of 6 was assigned to the area of the graph representing the best-observed conditions. Scoring lines were drawn using one of the three methods, to achieve relatively equal distribution of sites into each scoring category: quadrisection, double bisection or triple bisection. The quadrisection technique was used when data points were relatively evenly distributed below the 95th percentile or MV (Figure 2A). The double bisection technique was used when there was moderate clumping of data at the bottom of the graph, including many zero values, as in the metric per cent Tanypodinae (Figure 2B). Double bisection consisted of bisecting the area below the 95th percentile, then bisecting the lower half again and assigning a separate scoring category to all sites with a value of zero (score of 0 in positive metrics, 6 in negative metrics) (Figure 2B). The triple bisection technique was used when the data were clumped near the bottom of the graph, but relatively few samples had zero values. This method was similar to double bisection, except that three separate bisections occurred at progressively lower levels, and the zero value was included in the lowest section of non-zero metric values (Figure 2C). Scoring procedures for each selected metric are listed in Table IIA.

INDEX RESPONSE TO OUTFALL DISTURBANCE

The 12-metric Panel Index had a minimum possible score of 0 and a maximum possible score of 72. Boxplots of scores for the Panel Index in 2000 showed a slight difference between reference and outfall sites (median difference between reference and 0-m outfall = 6), with subsequent recovery downstream, but differences were not significant (Mood's median test, $p = 0.376$, number of outfalls = 15) (Figure 3A). However, in 1999 (Figure 3B) (number of outfalls = 3) the difference between median values at the reference and 0-m outfall sites was 30 and the differences among plots were significant ($p = 0.009$). Median values also never returned to the level of the reference values up to 300 m downstream of the outfalls. Because only 3 outfalls were used in the 1999 index analysis, as opposed to 15 in 2000, one could conclude that the differences seen between 1999 and 2000 were because these three outfalls happened to be exceptionally responsive. An examination of the same three outfalls in 2000, however, reveals that these three sites showed trends similar to all the 2000 outfalls as a whole (Figure 3C). When combining the 2 years of data, the difference in medians between reference and 0-m outfall sites remains significant at 7 ($p = 0.09$) (Figure 3D).

The Percentage Index included 6 metrics and had a minimum possible score of 0 and a maximum of 36. Differences in index scores among reference and outfall sites in 2000 (number of outfalls = 15) were significant ($p = 0.016$), even though the difference in medians between reference and 0-m outfall was 5 (Figure 4A). Boxplots from 1999 (number of outfalls = 3) (Figure 4B), however, had a significant median difference of 20 (Mood's

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/rra

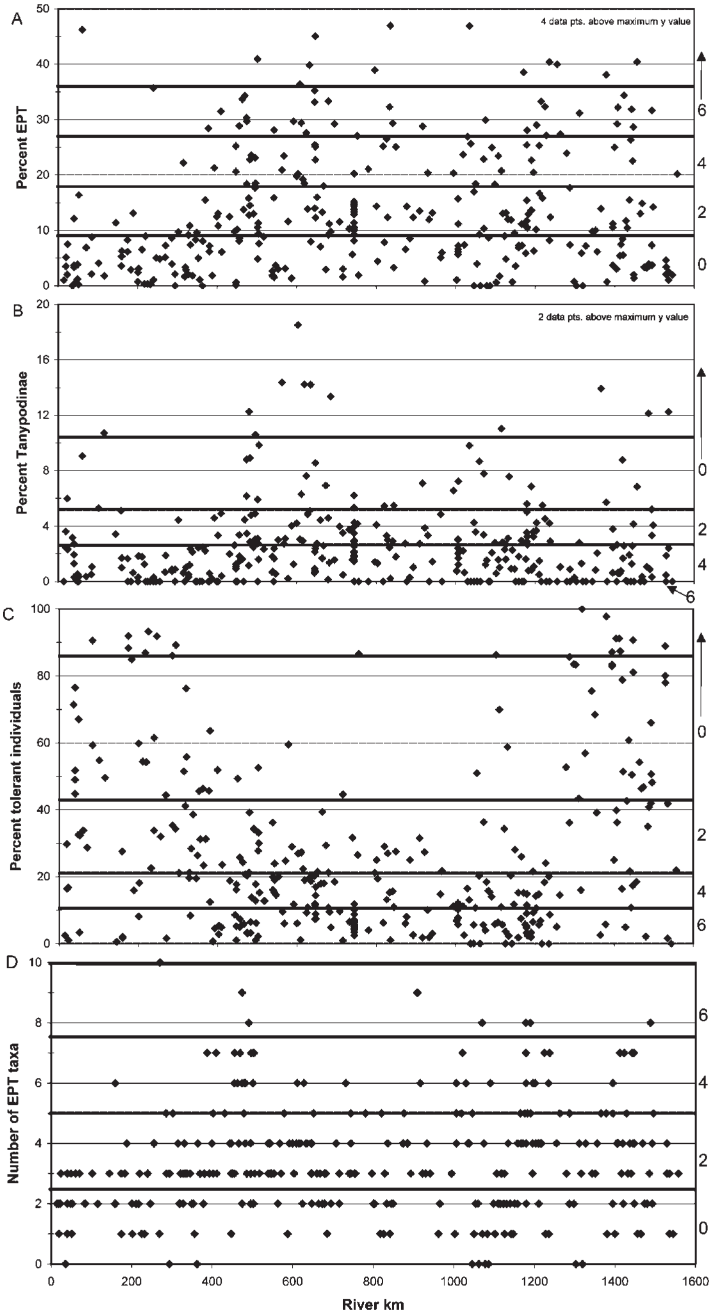

Figure 2. Metrics plotted against river km at non-outfall sites, illustrating examples of metric scoring procedures (numbers on right y-axis are assigned metric scores): (A) per cent EPT, scored by quadrisection, and showing depressed metric values in the upper river (upper bold line is 95% line); (B) per cent Tanypodinae, scored by double bisection (upper bold line is 95% line); (C) per cent tolerant individuals, scored by triple bisection (upper bold line is 95% line); and (D) number of EPT taxa (upper bold line is maximum value (MV)), scored by quadrisection

median $p = 0.039$). Also, the median score downstream of the outfalls did not return to reference condition values up to a distance of 300 m. Combining the 2 years' data (Figure 4C) resulted in significant differences ($p = 0.003$), with a median difference between reference and 0-m outfall sites of 6.

The difference in index response to outfalls in 1999 and 2000 is striking, and can probably be attributed to differences in flow between 1999 and 2000. Hester–Dendy samplers were set in late August of 1999 and 2000, and

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/rra

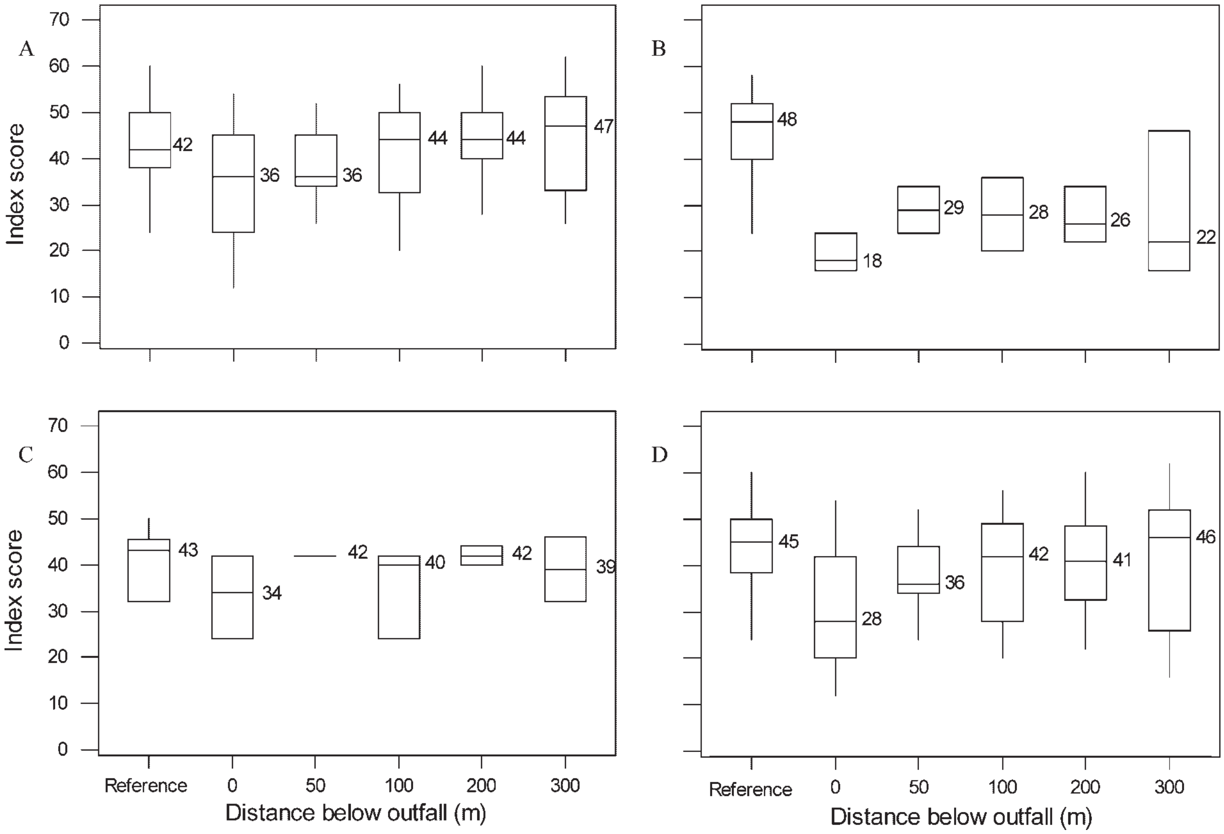

Figure 3. Box plots of Panel Index scores using (A) 2000 test outfalls ($n = 15$, median differences are not significant: Mood's median test, $p = 0.376$), (B) 1999 test outfalls ($n = 3$, median differences are significant: Mood's median test, $p = 0.009$), (C) 3 outfall sites from 2000 (same outfalls used in 1999 analysis, median differences are not significant: Mood's median test, $p = 0.361$) and (D) 1999 and 2000 test outfalls ($n = 18$, median differences are significant: Mood's median test, $p = 0.09$)

were retrieved in early October of these years. In 1999, water discharge levels for the Ohio River were one-third to one-half of the average discharge for these months, and approached 7-day, 10 year low-flow rates (pers. comm., Dennis L. McClain, United States Geological Survey, Louisville, KY). In 2000, flow levels were close to the long-term average. The greater flow in 2000 would probably cause increased dilution of contaminants at outfalls, and quicker flushing of pollutants downstream. Both versions of the macroinvertebrate index appear to effectively detect differences in water quality downstream of outfalls during low-flow periods, when contaminants are more heavily concentrated. When river discharge is greater and contaminants are quickly diluted, however, water quality conditions downstream of outfalls do not appear to be degraded to a point that results in consistent changes in the two draft indices.

The two draft indices also responded to outfall disturbance differently from one another. Mood's median test showed significance in only the Percentage Index during 2000, which also had lower *p*-values than the Panel Index when all of the 1999 and 2000 outfall sites were grouped together. However, Mood's median test was highly significant for the Panel Index in 1999, and, when grouping the 1999 and 2000 data together, the Panel Index showed a greater median difference between reference and 0-m outfall sites.

The Panel index may also be useful in discriminating between moderately and highly impacted sites. For instance, per cent Oligochaeta performed better in 1999 than in 2000, and showed a predictable response at all of the test outfall sites in 1999, indicating that this metric only responds under the most degraded conditions (e.g. downstream of outfalls during periods of very low river flow). The Panel Index also included metrics that may help detect water quality changes over time, such as number of EPT taxa and number of Diptera taxa, both of

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/rra

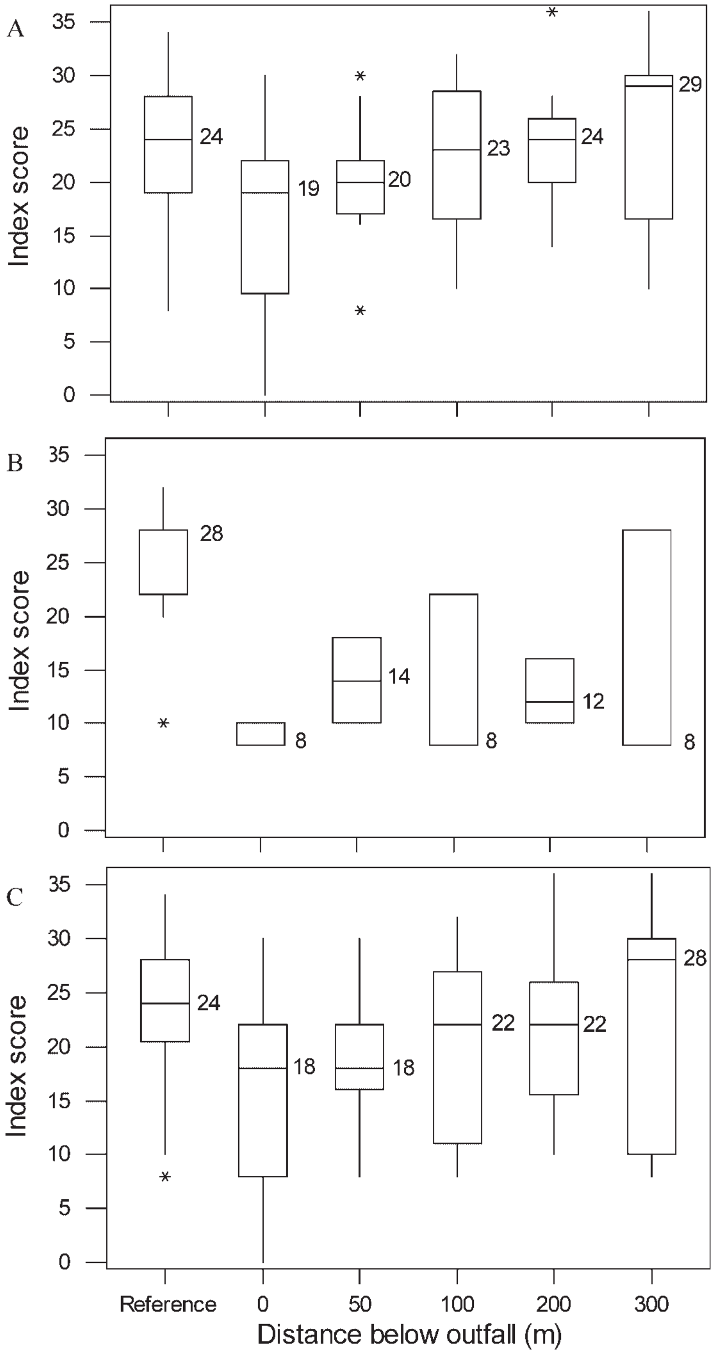

Figure 4. Box plots of Percentage Index scores using (A) 2000 test outfalls ($n = 15$, median differences are significant: Mood's median test, $p = 0.016$), (B) 1999 test outfalls ($n = 3$, median differences are significant: Mood's median test, $p = 0.039$) and (C) 1999 and 2000 test outfalls ($n = 18$, median differences are significant: Mood's median test, $p = 0.003$). Asterisks represent outliers

which have been shown to increase with declining anthropogenic disturbance (Losos, 1984; OEPA, 1988; Kerans and Karr, 1994; Barbour *et al*., 1996). In addition, the Percentage Index consisted of only six metrics, limiting its ability to fully describe the macroinvertebrate community. The Panel Index appears to exhibit several advantages over the Percentage Index, and therefore, the remaining discussion will be simplified by focusing only on the Panel Index.

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/rra

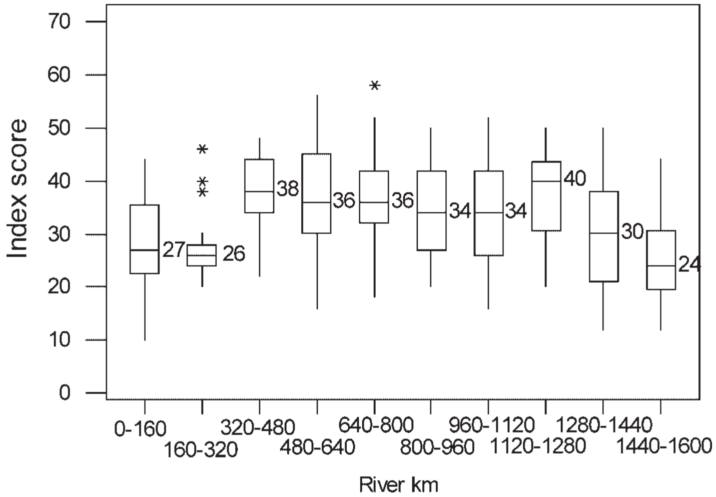

Figure 5. Boxplots of Panel Index scores at non-outfall ($n = 322$) sites in 160 km sections. Asterisks represent outliers

Analysis of index scores at non-outfall sites ($n = 322$) grouped along 10 equal sections of the river (160 km each) revealed that scores were relatively stable throughout the middle section of the river but were lower in the first 320 km of the river as well as the last 320 km (Figure 5). Lower values in the first 320 km of the river tend to correspond to the high concentration of industry in the upper river, which may be compounded by the effect of lower total river discharge available in the upper river to dilute the industrial effluent (Beckett and Keyes, 1983). Lower index scores in the lower 320 km of the river may be a result of several variables, including reduced water quality, increased habitat homogeneity, or increased flow. The three Ohio River tributaries with the largest total drainage areas flow into the Ohio in the last 300 km: the Wabash River (85 700 km^2 drainage area), the Cumberland River (17 920 km^2 drainage area) and the Tennessee River (105 900 km^2 drainage area). The average discharge of the Ohio River in August, September and October more than doubles from the gauge at river km 1274 to the gauge at river km 1532 (ORSANCO, 1994). If this increase in river volume or habitat homogeneity in the lower river have resulted in ecosystem change, and thus a natural alteration of the macroinvertebrate community structure, adjustment of scoring procedures may be necessary. Lower index scores may, however, also be a reflection of the addition of low-quality water from these tributaries, particularly the Wabash River, which flows through an intensively cultivated agricultural region.

CONCLUSION

We evaluated benthic macroinvertebrate metric response to wastewater discharges on the Ohio River and identified macroinvertebrate community metrics that are sensitive to water quality changes in the river. We assembled these metrics into two separate indices, and evaluated each for its ability to detect water quality differences. We determined that incorporating best professional judgement into metric selection resulted in an index (i.e. Panel Index) that was more likely to respond to water quality changes than one composed of metrics selected by a strictly objective approach (i.e. Percentage Index). We also found indications that flow differences in the Ohio River can affect macroinvertebrate community structure, both below outfalls and on a river-wide scale.

Although the research described herein identified the effects of chemical water quality on macroinvertebrate community structure, for an index to measure biotic integrity, it must also be sensitive to other variables impacting aquatic communities, including flow regime, habitat structure and energy flow. Future research into developing a benthic macroinvertebrate IBI for the Ohio River will require evaluation of these variables, and may include additional sampling designs, identification of least disturbed areas on the Ohio River, and analysis of historical data.

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/rra

In addition, the metrics and indices presented in this paper should be further refined using an independent dataset, with sufficient replication to permit statistical analyses of the results.

ACKNOWLEDGEMENTS

This work was supported by ORSANCO. We thank the ORSANCO field crews, especially Jeff Thomas, Robert Row, Dusty Jones and Bill Cassidy, for valuable assistance in macroinvertebrate collection. Thanks to the ORSANCO Macroinvertebrate Advisory Panel for guidance throughout the research process, and David Johnson and G. Thomas Watters for comments on earlier drafts of this paper. Macroinvertebrate sample identification was conducted by Pennington and Associates, Cookeville, Tennessee.

REFERENCES

Abel T, Barlocher F. 1988. Uptake of cadmium by *Gammarus fossarum* (Amphipoda) from food and water. *Journal of Applied Ecology* **25**: 223–231.

Barbour MT, Gerritsen J, Griffith GE, Frydenburg R, McCarron E, White JS, Bastian ML. 1996. A framework for biological criteria for Florida streams using benthic macroinvertebrates. *Journal of the North American Benthological Society* **15**: 185–211.

Beckett DC, Keyes JL. 1983. A biological appraisal of water quality in the Ohio River using macroinvertebrate communities with special emphasis on Chironomidae. *Memoirs of the American Entomological Society* **34**: 15–33.

Belpaire C, Smolders R, Auweele I, Ercken P, Breine J, Van Thuyne G, Ollevier F. 2000. An index of biotic integrity characterizing fish populations and the ecological quality of Flandrian water bodies. *Hydrobiologia* **434**: 17–33.

Borgmann U, Munawar M. 1989. A new standardized sediment bioassay protocol using the amphipod *Hyalella azteca* (Saussure). *Hydrobiologia* **188/189**: 425–531.

Britt NW, Addis JT, Engle R. 1973. Limnological studies of the island area of Western Lake Erie. *Bulletin of the Ohio Biological Survey* **4**: 1–88.

Covich AP, Thorp JH. 1991. Crustacea: introduction and Peracarida. In *Ecology and Classification of North American Freshwater Invertebrates*, Thorp JH, Covich AP (eds). Academic Press, Inc.: San Diego, California, USA; 665–689.

Deegan LA, Finn JT, Ayvazian SG, Ryder C. 1993. Feasibility and application of the index of biotic integrity to Massachusetts estuaries (EBI). *Report to the Massachusetts Executive Office of Environmental Affairs*, Massachusetts, USA.

Deshon JE. 1995. Development and application of the Invertebrate Community Index (ICI). In *Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making*, Davis WS, Simon T (eds). Lewis Publishers: Boca Raton, Florida, USA; 217–243.

Dills GD, Rogers DT. 1974. Macroinvertebrate community structure as an indicator of acid mine drainage. *Environmental Pollution* **6**: 239–262.

Emery EB, Simon TP, McCormick FH, Angermeier PL, Deshon JE, Yoder CO, Sanders RE, Pearson WD, Hickman GD, Reash RJ, Thomas JA. 2003. Development of a multimetric index for assessing the biological condition of the Ohio River. *Transactions of the American Fisheries Society* **132**: 791–808.

Engle VD, Summers JK. 1999. Refinement, validation, and application of a benthic condition index for northern Gulf of Mexico estuaries. *Estuaries* **22**: 624–635.

Eyres JP, Williams NV, Pugh-Thomas M. 1978. Ecological studies on Oligochaeta inhabiting substrata in the Irwell, a polluted English river. *Freshwater Biology* **8**: 25–32.

Fiance SB. 1978. Effects of pH on the biology and distribution of *Ephemerella funeralis* (Ephemeroptera). *Oikos* **31**: 332–339.

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.

Ganasan V, Hughes RM. 1998. Application of an index of biological integrity (IBI) to fish assemblages of the rivers Khan and Kshipra (Madhya Pradesh), India. *Freshwater Biology* **40**: 367–383.

Hall RJ, Driscoll CT, Likens GE, Tratt JM. 1985. Physical, chemical, and biological consequences of episodic aluminum additions to a stream. *Limnological Oceanography* **30**: 212–220.

Hermann J, Anderson KG. 1986. Aluminum impact on respiration of lotic mayflies at low pH. *Water, Air, and Soil Pollution* **30**: 703–709.

Hester FE, Dendy JS. 1962. A multiplate sampler for aquatic macroinvertebrates. *Transactions of the American Fisheries Society* **91**: 420–421.

Hill BH, Herlihy AT, Kaufmann PR, Stevenson RJ, McCormick FH, Burch-Johnson C. 2000. Use of periphyton assemblage data as an index of biotic integrity. *Journal of the North American Benthological Society* **19**: 50–67.

Hilsenhoff WL. 1991. Diversity and classification of insects and Collembola. In *Ecology and Classification of North American Freshwater Invertebrates*, Thorp JH, Covich AP (eds). Academic Press, Inc.: San Diego, California, USA; 593–663.

Johnson BJ, Richardson WB, Naimo TJ. 1995. Past, present, and future concepts in large river ecology. *BioScience* **45**: 134–141.

Karr JR. 1981. Assessment of biotic integrity using fish communities. *Fisheries* **6**: 21–27.

Karr JR. 1991. Biological integrity: a long-neglected aspect of water resource management. *Ecological Applicatons* **1**: 66–84.

Karr JR, Fausch KD, Angermeier PL, Yant PR, Scholosser IJ. 1986. Assessing biological integrity in running waters: a method and its rationale. *Illinois Natural History Survey Special Publication No. 5*. Illinois, USA.

Kerans BL, Karr JR. 1994. A benthic index of biotic integrity (B-IBI) for rivers of the Tennessee Valley. *Ecological Applications* **4**: 768–785.

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/rra

Lenat DR. 1983. Chironomid taxa richness: natural variation and use in pollution assessment. *Freshwater Invertebrate Biology* **2**: 192–198.

Lenat DR, Smock LA, Penrose DL. 1980. Use of benthic macroinvertebrates as indicators of environmental quality. In *Biological Monitoring for Environmental Effects*, Worf DL (ed.). D.C. Heath and Co.: Lexington, Massachusetts, USA; 97–112.

Leonard PM, Orth DJ. 1986. Application and testing of an index of biotic integrity in small, coolwater streams. *Transactions of the American Fisheries Society* **115**: 401–414.

Losos B. 1984. The influence of pollution on the density and production of Chironomidae (Diptera) in running waters. *Limnologica* **15**: 7–19.

Lyons J, Piette RR, Niermeyer KW. 2001. Development, validation, and application of a fish- based index of biotic integrity for Wisconsin's large warmwater rivers. *Transactions of the American Fisheries Society* **130**: 1077–1094.

Mack JJ, Micacchion M, Augusta LD, Sablak GR. 2000. Vegetation indices of biotic integrity (VIBI) for wetlands, and calibration of the Ohio rapid assessment method for wetlands v 5.0. *Report of The Ohio Environmental Protection agency to The United States Environmental Protection Agency*, Washington DC, USA.

McCormick PV, Stevenson RJ. 1998. Periphyton as a tool for ecological assessment and management in the Florida Everglades. *Journal of Phycology* **34**: 726–733.

Mundahl ND, Simon TP. 1999. Development and application of an index of biotic integrity for coldwater streams of the upper Midwestern United States. In *Assessing theSustainability and Biological Integrity of Water Resources Using Fish Communities*, Simon TP (ed.). CRC Press: Boca Raton, Florida, USA; 383–415.

OEPA (Ohio Environmental Protection Agency). 1988. *Biological Criteria for the Protection of Aquatic Life, Volume 2: Users Manual for Field Assessment of Ohio Surface Waters*. OEPA, Division of Water Quality Monitoring and Assessment, Surface Water Section: Columbus, Ohio, USA.

OEPA (Ohio Environmental Protection Agency). 1997. *Development of Biological Indices Using Macroinvertebrates in Ohio Nearshore Waters, Harbors, and Lacustuaries of Lake Erie in Order to Evaluate Water Quality.* Draft report in fulfillment of Lake Erie Protection Fund Grant LEPF-06-94. Division of Surface Water: Columbus, Ohio, USA.

Omernik JM. 1987. Ecoregions of the conterminous United States. *Annals of the Association of American Geographers* **77**: 179–190.

ORSANCO (Ohio River Valley Water Sanitation Commission). 1994. *Ohio River Fact Book*. ORSANCO: Cincinnati, Ohio, USA.

ORSANCO (Ohio River Valley Water Sanitation Commission). 1999. *Quality Assurance Document for Macroinvertebrates*. ORSANCO: Cincinnati, Ohio, USA.

ORSANCO (Ohio River Valley Water Sanitation Commission). 2001. *Ohio River Fact Sheet*. ORSANCO: Cincinnati, Ohio, USA.

Payne BS, Miller AC. 1989. Growth and survival of recent recruits to a population of *Fusconaia ebena* (Bivalvia: Unionidae) in the lower Ohio River. *American Midland Naturalist* **121**: 99–104.

Pennak RW. 1978. *Fresh-Water Invertebrates of the United States* (2nd edn). John Wiley & Sons: New York, New York, USA.

Schulz R, Liess M. 1999. A field study of the effects of agriculturally derived insecticide input on stream macroinvertebrate dynamics. *Aquatic Toxicology* **46**: 155–176.

Seegert G. 2000. Considerations regarding development of index of biotic integrity metrics for large rivers. *Environmental Science & Policy* **3**: S99–S106.

Simon TP, Emery EB. 1995. Modification and assessment of an index of biotic integrity to quantify water resource quality in great rivers. *Regulated Rivers: Research & Management* **11**: 283–298.

Thoma RF. 1999. Biological monitoring and an index of biotic integrity for Lake Erie's nearshore waters. In *Assessing the Sustainability and Biological Integrity of Water Resources Using Fish Communities*, Simon TP (ed.). CRC Press: Boca Raton, Florida, USA; 417–461.

Thorp JH. 1992. Linkage between islands and benthos in the Ohio River, with implications for riverine management. *Canadian Journal of Fisheries and Aquatic Science* **49**: 1873–1882.

Thorp JH, Covich AP. 1991. An overview of freshwater habitats. In *Ecology and Classification of North American Freshwater Invertebrates*, Thorp JH, Covich AP (eds). Academic Press, Inc.: San Diego, California, USA; 17–36.

Toham AK, Teugels GG. 1999. First data on an index of biotic integrity based on fish assemblages for the assessment of the impact of deforestation in a tropical West African river system. *Hydrobiologia* **397**: 29–38.

.

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/rra