United States Environmental Protection Agency

Association among invertebrates and habitat indicators for large rivers in the Midwest

How sampling distance, point-sampling of habitat, and subsample size effect measures of large river macroinvertebrate assemblages

RESEARCH AND DEVELOPMENT

EPA 600/R-04/177
October 2004
www.epa.gov

Association among invertebrates and habitat indicators for large rivers in the Midwest

How sampling distance, point-sampling of habitat, and subsample size effect measures of large river macroinvertebrate assemblages

Joseph E. Flotemersch*, Karen Blocksom,
John J. Hutchens, Jr. [1], Bradley C. Autrey

National Exposure Research Laboratory, U.S. Environmental Protection Agency, 26 W. Martin Luther King Dr., Cincinnati, OH 45268 USA, [1]Current address: Department of Biology, Coastal Carolina University, Conway, SC 29528 USA

U.S. Environmental Protection Agency
Office of Research and Development
Washington, DC 20460

Executive Summary

Six reaches in each of two large rivers (one each in Kentucky and Ohio) were sampled using a prototype benthic macroinvertebrate sampling technique. The intent was to better understand the relationship between large river macroinvertebrate assemblages and habitat features. This information was to determine an acceptable sampling design to support development of a large river bioassessment protocol (LR-BP). Specific objectives included determining the appropriate number of habitat point-samples to be collected, examining how varying reach length affects assemblage characteristics, and determining an appropriate laboratory subsample size to accompany the resulting field sampling method.

At each site, both banks of 12 transects separated by increasingly larger distances were sampled. Analyses were conducted using Monte Carlo methods. Interpretation of results relied on the metric values of total taxa richness, mayfly taxa richness, caddisfly taxa richness, Diptera richness, % mayflies, % caddisflies, % Tanytarsini, % non-Tanytarsini dipterans and non-insects, and % tolerant individuals.

This research indicates that, using the sampling technique discussed herein, a representative sample of the benthic macroinvertebrate fauna in the study reaches was collected by sampling both banks of 6 transects spaced at 100 m intervals over a 500-m distance. It is hypothesized that these results were achieved because the sampling method and design effectively sampled the benthic macroinvertebrate fauna of the dominant habitat types within the reach.

It is recommended that the field method be coupled with a fixed laboratory subsample size of 300 organisms for bioassessment purposes, with the recognition that a subsample size of 500 organisms may be needed to meet the objectives of some studies. This recommendation is based on the response of the tested metrics, and the observation that the ability to separate sites of different macroinvertebrate composition generally did not increase with larger subsample sizes.

It is likely this approach will over-sample sites of uniform composition, but the goal was to develop a standardized LR-BP that would perform well across sites of differing habitat composition. It should be noted that the LR-BP for macroinvertebrates has only been tested in main-channel habitats. It may work equally well on off-channel habitats, but this remains to be tested.

While the method has been designed to perform well in a variety of habitats, resulting data should be interpreted with appreciation for coarse habitat characteristics. This information would ideally be derived from habitat data collected concomitantly with the faunal data. Sites can then be categorized into river types (e.g., impounded vs. free-flowing or lowland vs. upland rivers) or even habitat types within a specific river (e.g., sandy- vs. cobble-bottom) in a more controlled environment (i.e., in the laboratory), thus increasing the overall integrity of any and all site assessments. Possible modifications to the method to streamline its future application in the field are provided.

Key Words: benthic macroinvertebrates, bioassessment, sampling method, LR-BP, non-wadeable streams.

Introduction

Wadeable streams and smaller rivers are abundant and relatively easy to sample compared to large rivers. As a result, efforts to develop appropriate sampling protocols for the bioassessment of lotic ecosystems have been focused largely on smaller systems (e.g., Barbour et al. 1999). As these methods become increasingly refined and accepted, a growing number of government agencies are starting to better understand (Humphries et al. 1998) and develop sampling protocols for non-wadeable large rivers.

Large rivers differ from wadeable systems in some important ways. In rivers, stressor sources are generally more numerous (Sweeting 1994) and almost certainly more rapidly diffused as discharge rates are generally higher in higher-order streams (Allen 2000). Consequently, individual stressor effects are masked by the presence of other stressors and their impacts less conspicuous. Additionally, biological communities change with stream size, as do habitat type and quality (Vannote et al. 1980). Assemblages adapted to deeper, wider streams with limited canopy cover are more likely to occur in downstream higher order reaches. Thus, expectations for communities in large rivers may be very different from those in smaller systems. In addition, the thalweg of a large river often may not be accessible for sampling as it is in wadeable streams, precluding the use of some wadeable stream sampling protocols. Hence, resource managers need clear and consistent protocols available for measuring ecological integrity that are designed specifically for large river systems (Loucks 2003).

To effectively support a bioassessment program, protocols for sampling fauna of large rivers should be clear, consistent, and reproducible. In order to be applicable to a wide audience, they should perform well across numerous non-wadeable habitats and river types, represent site conditions accurately, and ideally, identify the presence of stressors. Protocols should also be cost effective, logistically feasible, and meet or be adaptable to multi-purpose sampling needs of researchers and managers (e.g., trend analysis, point-source and non-point source programs) if they are to be accepted by regulatory organizations.

Benthic macroinvertebrates are one of the most common faunal assemblages used in the bioassessment of aquatic ecosystems (Rosenberg and Resh 1993, Metcalfe-Smith 1994, Barbour et al. 1999, Karr and Chu 1999). Many macroinvertebrate collection methods currently used in non-wadeable systems are derived from wadeable methods (Ohio EPA 1989, Barbour et al. 1999, Klemm et al. 2000, Flotemersch et al. 2001, Moulton et al. 2002,). These methods often involve wading in shallow areas (e.g., near the shore) of larger rivers or sampling from a boat in deep areas without additional modification. The exception is the use of artificial substrates, which were developed largely for non-wadeable invertebrate sampling applications (Cairns 1982).

Blocksom and Flotemersch (in press) compared six sampling techniques used by three government agencies to sample benthic macroinvertebrate assemblages of large rivers. They found that these methods resulted in different metric values. Additionally, metric response (i.e., positive vs. negative) to certain

stressors varied among sampling methods, and metrics detecting a specific stressor were not consistent across methods (Blocksom and Flotemersch, in press). Differences among the methods and the relatively poor performance of some methods were hypothesized to be due partly to the inadequacy of using a single sampling technique (e.g., kick-net, dip-net, artificial substrates) when sampling large rivers. For example, a method that produced a representative sample in a large river with abundant epifaunal substrate and low embeddedness might not reflect a highly embedded reach well. The research of Bartsch et al. (1998) and Poulton et al. (2003) corroborate this hypothesis, with both concluding that an approach employing multiple sampling techniques was needed to effectively sample all components of a macroinvertebrate assemblage in riverine ecosystems.

To support the development of a more consistent Large River Bioassessment Protocol (LR-BP) for benthic macroinvertebrates, three fundamental issues must be addressed. First, a collection technique is needed that secures a representative sample of benthic macroinvertebrates across the broad range of habitats that occur within and across rivers. Second, an appropriate sampling design will be needed for application of the developed sampling technique. And third, an appropriate laboratory method (e.g., laboratory subsample size) must be determined. Through this entire process, it is important to keep in mind the logistical realities faced by regulatory agencies.

To address this issue of collection technique, data from the different sampling techniques compared by Blocksom and Flotemersch (in press) were

analyzed to identify those techniques that when integrated, produced a potentially more representative sample across a broad range of large river habitat types. Consideration was given to how the critical elements of each method could be combined and applied in a standardized manner to support a Large River Bioassessment Protocol (LR-BP). The result was a sampling technique that hypothetically should overcome the limitations of previous approaches and permit standardized sampling across all non-wadeable habitats and river types of varying impoundment status. The approach consisted of features of the Environmental Monitoring and Assessment Program – Surface Waters kick net sampling method (Klemm et al. 2000), and the multiple habitat dip net methods of the Ohio Environmental Protection Agency (Ohio EPA 1989) and the United States Geological Survey (Moulton et al. 2002) that sample all available habitats.

Critical elements of the development of a scientifically sound sampling design include the spatial scale over which sub-samples should be collected (i.e., reach length), the number of sub-samples needed, and the manner in which sub-samples should be distributed within the sample reach. The use of “reach” in this study follows that of Frissell et al. (1986) who defined it as a length of stream between breaks in channel slope, local side-slopes, valley floor width, riparian vegetation, and bank material.

For bioassessment purposes, determination of appropriate sample reach lengths are typically linked to measures of geomorphology (e.g., channel widths, meander wavelengths, riffle pool sequences) (Barbour et al. 1999, Herlihy and

Lazorchak 2000, and Moulton et al. 2002) or evaluation of species accumulation curves. Several studies have focused on appropriate reach lengths for macroinvertebrates in wadeable streams (e.g., Li et al. 2001), and for fish in both wadeable and non-wadeable streams (e.g., Lyons 1992, Hughes et al. 2002). However, an appropriate assessment reach for macroinvertebrates in non-wadeable streams has not been estimated. One difficulty is that benthic macroinvertebrates are usually sampled at specific points, whereas sampling for fish (e.g., fish by means of electrofishing techniques) is continuous over the whole reach. Hence, the approach of determining an appropriate sampling reach length for macroinvertebrates using species accumulation curves as a direct function of distance is logistically impractical (due to the large number of contiguous samples that would be required).

Similar challenges are encountered using measures of geomorphology for reach determination on large rivers. The majority of streams in the U.S. have been anthropogenically altered (especially through dam construction) to the extent that <2% are of a quality worthy of federal protection status (Benke 1990). This reality limits the utility of geomorphology (e.g., riffle-pool sequences, multiples of the natural channel width) as a determinant of reach length because it would only apply to a small subset of sites. Even if either of these traditional approaches did work, the question still remains as to the appropriate number and distribution of sub-samples within the designated reach to effectively represent the reach for bioassessment purposes.

As for the development of an appropriate laboratory method, the procedures for sample processing are typically assumed to be readily transferable from wadeable streams to large rivers, but this has not been evaluated. Hence, this study also investigated the efficacy of sample processing in the laboratory. The methods used for laboratory processing of invertebrates can greatly influence sample results and ultimately determine the value of a method for bioassessment purposes. A full count of all invertebrates may provide a more accurate assessment (Doberstein et al. 2000), but is usually not feasible when large numbers of organisms are collected (Barbour and Gerritsen 1996). As a result, samples are often subsampled in the laboratory, typically using either a fixed-organism count or a fixed sample proportion (Barbour et al. 1999, Carter and Resh 2001).

The primary objectives of the study were to: 1) determine the appropriate number of sampling points needed using a new LR-BP for macroinvertebrates in nonwadeable rivers, 2) determine an appropriate laboratory subsample size to accompany this sampling method, and 3) examine how varying reach lengths affect assemblage characteristics. Examination of these features relied on evaluation of the quantitative metrics of the Ohio EPA Invertebrate Community Index (ICI).

Methods

Study Sites

We collected data during late July through August 2001 from the Kentucky (n = 6 sites) and Great Miami (n = 6 sites) rivers, both of which are major tributaries of the Ohio River in the east-central United States (Figure 1). The 12

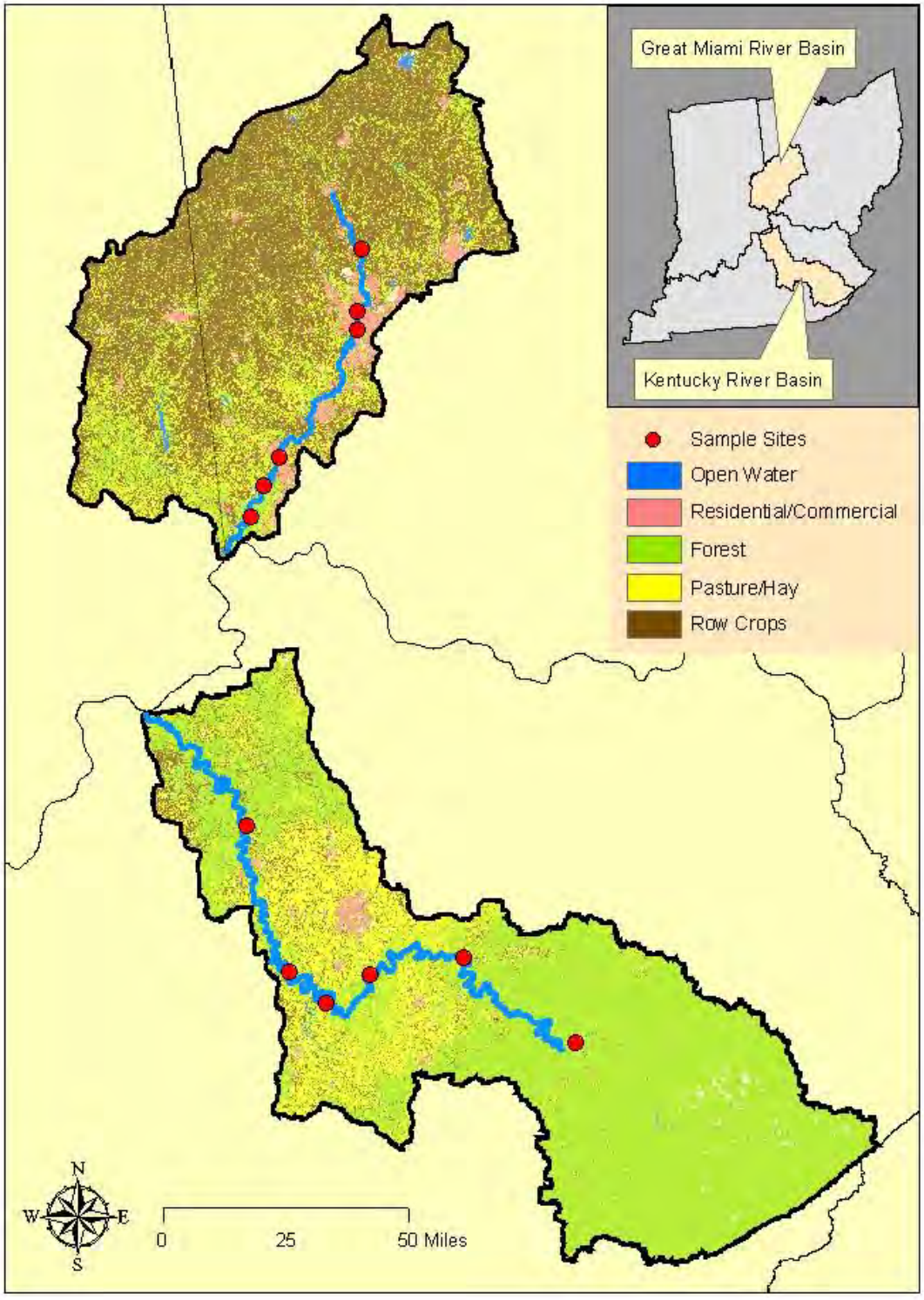

Figure 1. Sample sites on the Great Miami and Kentucky rivers.

sites were selected from 30 sites over 2 midwestern large rivers that were sampled in a previous study to compare existing large river sampling methods (Blocksom and Flotemersch, in press)(Table 1). Sites were selected to partition the sampling effort evenly between impounded and relatively free-flowing sites, and across a gradient of habitat conditions within each river. Gradients were based on existing instream and riparian physical habitat data collected using EMAP protocols (Kaufmann 2000), land use data, and best professional judgment.

The Great Miami River flows through several urban and industrial corridors in Ohio (e.g., Dayton, Springfield, Hamilton, and Middletown) before reaching the Ohio River. However, the dominant land use in the basin is agriculture (80.3%) (www.miamiconservancy.org). The river has sections with exposed riffles and rapids and sections with restricted flow associated with low-head dams that store, rather than regulate, waters.

Table 1. Physical characteristics and mean percent (standard deviation) of land use types in the study basins, with means and standard deviations based on sites used in analyses.

Parameter	River Basin	
	Great Miami (n=6)	Kentucky (n=6)
Drainage Basin (km^2)	13,947[a]	18,130[b]
River Length (km)	233.4[a]	410.4[b]
Average Gradient (m/km)	0.74[a]	0.13[b]
% Urban land use[c]	5.90 (2.64)	7.81 (0.52)
% Agriculture land use[c]	82.55 (3.21)	80.18 (1.34)
% Forested land use[c]	10.10 (1.44)	10.83 (1.60)

[a] Ohio EPA, 1997
[b] Kentucky River Authority, 1999
[c] Multi-Resolution Land Characteristics Consortium (Vogelmann *et al.*, 1998)

The Kentucky River has a series of 14 lock-and-dam structures that span the length of the mainstem, which historically supported commercial traffic. The watershed has some large forested sections and some small areas with mining, agricultural and urban influences (e.g., Lexington). As a result of impoundment, all Kentucky River sites sampled in this study were much deeper than those of the Great Miami River (Table 2).

Final site selection resulted in sites well-distributed longitudinally along the mainstem of each river and included a mixture of habitat types. Study reaches were positioned so that stream confluences, bridges and obvious stressor

Table 2. Ranges and medians of chemical and physical habitat variables at study sites.

	Great Miami River (N=6)		Kentucky River (N=6)	
	Range	Median	Range	Median
Physical habitat				
Mean thalwag depth (m)	1.2–2.3	2.03	5.2–9.7	7.0
Mean wetted width (m)	45.8-154.3	94.4	69.9-97.3	80.5
Mean bankfull height (m)	0.7-2.9	1.3	1.7-2.2	2.1
LWD quantity	11-70	30.0	20-43.0	30.5
% Canopy density at bank	0-92	37	71-92	78
% Substrate as large gravel and larger at bank	0-89.5	25.5	0-0.92	0.78
% Urban in riparian	2.1-83.1	42.4	0.4-7.7	1.1
% Agriculture in riparian	9.0-77.1	27.7	9.7-33.6	14.0
% Forest in riparian	6.8-57.5	12.2	65.4-85.8	83.6
Water Chemistry				
Mean conductivity (uS/cm)	521.2-857.2	664.6	270.6-435.2	334.2
SO_4 (mg/L)	33.0–64.4	45.3	33.9–104.6	79.8
NO_3 (mg/L)	1.37–5.69	1.95	0.37–0.80	0.56
Chloride (mg/L)	25.06–71.64	44.9	3.30–6.61	5.74
Ammonia (mg/L)	0.07–0.23	0.09	0.02–0.07	0.04
Total Khejldahl Nitrogen (TKN) (mg/L)	0.53–0.95	0.61	0.15–0.29	0.23
Total Phosphorus (ug/L)	0.05–0.28	0.17	0.01–0.04	0.02

sources, such as major outfalls, did not occur within the reach as this might complicate and confuse the analysis of data within sites.

Sampling Design

An appropriate reach length for macroinvertebrates in non-wadeable streams has not been estimated. However, benthic macroinvertebrate and fish assemblage structure are often correlated (e.g., Kilgour and Barton 1999). Therefore, the available literature on appropriate assessment units for fish in large rivers was used for setting a maximum size for the study reach.

Measures of fish species richness is a function of the number of channel units sampled (Gormann and Karr, 1978; Angermeier and Schlosser 1989, Lyons 1992), and the size and spacing of these units are functions of stream size (Leopold et al. 1964). The assessment unit length required can also vary by study objectives (Cao et al. 2001, Hughes et al. 2002). Lyons (1992) concluded that for assessments of environmental quality or community-level ecological analyses, a distance of 35 times the mean stream width, or a length equal to three complete riffle-pool sequences, was sufficient. Pilot studies for the Environmental Monitoring and Assessment Program suggested that in eastern non-wadeable streams and rivers, a length of 40 channel widths was necessary to characterize the fish community of a site (Herlihy and Lazorchak 2000). Using this information, a reach length of 40 times the estimated mean wetted width of the channel at each river site was selected. In hydrologically formed channels,

this reach length would include approximately four meander wavelengths (Leopold et al. 1964).

The downstream end of the study reach at each site was set at a randomly determined point on one bank and marked with flagging. A systematic sampling design was applied to establish 12 transects within the reach. This design has many desirable features for field studies, and as long as the first point is selected at random, remaining points based on that point can be considered random as well (Cochran 1977). The simplicity of the design makes it easy to execute without mistakes and results in significant time saving in the field. It also results in the drawn sample being spread more evenly over the population (Cochran 1977, Manly 2001).

Proceeding upstream from the initial point, 11 transects spanning the width of the river were identified and flagged. The first 4 were spaced at a distance equal to the mean wetted width of the channel, followed by 2 spaced at 2 times the wetted width, 2 at 4 times the wetted width, and 3 at 8 times the wetted width (Figure 2). This identified 24 stations in the reach (e.g., 2 per transect, 12 on each bank) where macroinvertebrate samples would be collected.

The size of the sampling zone at each sampling station was proportional to the mean wetted width of the river. At each sampling station, a shoreline sampling zone was defined as 0.1 times the estimated wetted width in shoreline length and extended from the shore to the non-wadeable point of the river. Therefore, if the river was 70 m wide, the sampling zone for each station would be 7 m. This served to keep the sampling zone in proportion to the increasing

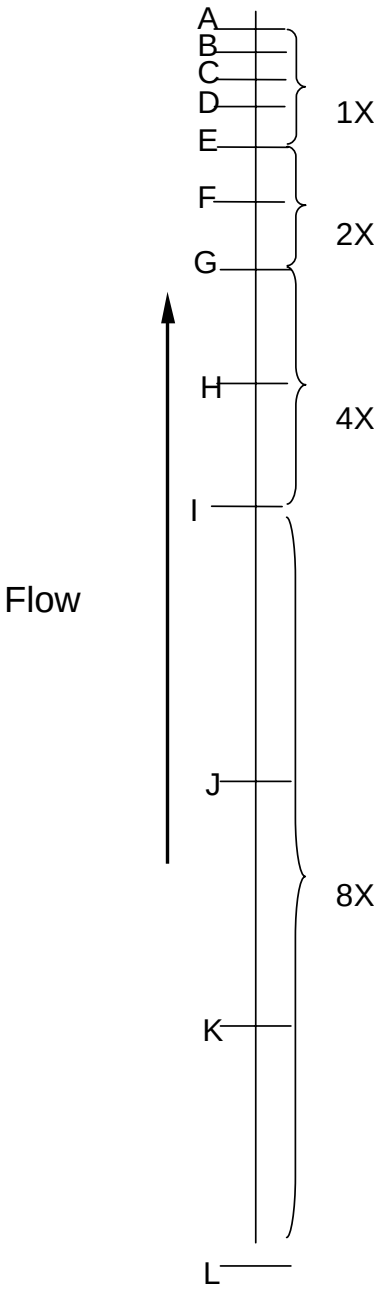

Figure 2. Sampling scheme used to examine the effect of distance on metric values.

size of habitat features as the size of a river increased. Zone placement was centered on the station transect.

Invertebrate Sampling

The technique used to collect benthic macroinvertebrate subsamples in each station zone was a hybrid of existing techniques and consisted of two distinct sample-collection procedures. The use of two distinct collection approaches provides a representative benthic macroinvertebrate sample from the different types of habitats encountered within and across rivers.

At each sampling station, two samples were collected with a modified kick-net (50-cm wide x 30-cm tall x 60-cm bag-depth; 595-µm mesh). The area in front of the net equal to the width and length of the net frame (0.5 m; total area = 0.25m^2) was then vigorously kicked for 20 seconds (see Klemm et al. 2000).

Next, a D-frame net (25.4-cm wide x 30.5-cm high x 25.4-cm bag-depth; 595 µm mesh) was used to sample other available habitats in the sampling zone (e.g., root wads, undercut banks, steep banks, vegetation). These habitats can be quite difficult to sample with a kick-net procedure, and therefore, are often underrepresented. For example, a river constrained by valley walls can be nearly impossible to sample safely with a kick net. At each station the sampling effort was standardized to 3 minutes per 5 m of sample zone width. While a standardized sampling time may be sufficiently quantitative (Hynes 1970) and used for quantifying effort, the main purpose of the timed effort was to control for

the amount of time field personnel spent at any single station, thus assuring ample time to cover the entire reach or multiple reaches in a day.

At some stations, only one collection procedure was feasible and suitable for collection of a representative sample of the fauna of the prevailing habitat. However, both were performed at every sampling zone if logistically safe and practical.

Samples from the kick and D-frame nets were composited into a single transect sample (n=12 per reach) for use in determining the appropriate number of point-samples that needed to be collected and examining the effects of reach length on sample results.

Each sample was processed in the field with a 595-µm sieve, preserved with 100% ethanol, and diluted to a final concentration approximating 70% ethanol (following Klemm et al. 2000). In the laboratory, individual samples were completely sorted. All organisms were identified to the lowest practical taxonomic level, usually species when specimen condition was adequate and taxonomic keys were available.

Additional information was collected from each sample station (n=24) to supplement the macroinvertebrate data, characterize each station and the study reach, and document the gradient of conditions over which samples were collected. Crews collected physical habitat data following EMAP protocols for nonwadeable streams (Kaufmann 2000). A single depth-integrated water sample was collected from each site and analyzed for sulfate, nitrate, total Kjeldahl nitrogen, ammonia, total phosphorus, and chloride concentrations. Chemical

analyses were conducted using EMAP-SW laboratory protocols (Klemm et al. 1990). Conductivity and water temperature were measured *in situ* using a YSI Model 85 meter at the center of the sampling reach. Land cover data developed by the Multi-Resolution Land Characteristics Consortium (Vogelmann et al. 1998) were overlaid on a riparian corridor 500 m in width on each side of the river for a distance of 4 km upstream of the center of the sampling reach. Proportions of forest, agriculture, and urban (including residential) land uses were then calculated within the riparian corridor.

Data Analyses

The differences in assemblage characteristics between the two banks at a given transect were sometimes quite large. To encompass the spatial variability present at each transect, samples were combined from the two banks at each transect. Thus, all analyses described in this paper use samples from both banks composited at each transect.

Subsample size.—Prior to analyses for estimating the minimum number of transects required per site, an appropriate laboratory subsample size was determined. The entire combined sample for each site (all transects combined) was used to simulate fixed-count subsamples of 100 to 1000 organisms in steps of 100. Simulations were run in C++ (Borland C++ Builder 4.0, Inprise Corporation, Scotts Valley, California). It was assumed that organisms were distributed randomly within each sample. Random sampling without replacement

was used to simulate each subsample, and 100 subsamples were generated for each site at each fixed count size to estimate laboratory sampling variability.

The effect of subsample size was measured on the quantitative metrics in the Ohio EPA Invertebrate Community Index (ICI) because these metrics are used to assess the macroinvertebrate assemblage in larger streams and rivers in Ohio (Ohio EPA 1998). These metrics included total taxa richness, mayfly taxa richness, caddisfly taxa richness, Diptera richness, % mayflies, % caddisflies, % Tanytarsini, % non-Tanytarsini dipterans and non-insects, and % tolerant individuals.

Since taxa richness metrics did not tend to level off with increasing subsample size, the difference in a metric value between sites was used as a way to measure the effect of sample size. The change in this absolute value of the difference in the metric from one sample size (X_i) to the next higher sample size (X_{i+1}) is defined as the "return", and the percent of the return relative to the maximum value achieved for that metric ($|X_i-X_{i+1}|/\max(X_{i+1})$) as the "relative return". In this calculation, the maximum value was set as the maximum for the next higher sample size.

The subsample size at which the average relative return leveled off for most metrics was selected for subsequent analyses on the effect of the number of transects on metrics.

Number of transects.--After the subsample size was determined, the number of transects needed per site was evaluated. Following the concept of a species

area curve, metric values were plotted as a function of the number of transects or samples. This required randomizing the order of the transects to ensure that the results were not affected by sequence of samples. However, the nature of the sampling design meant that transects were not equidistant from one another. If there was strong spatial autocorrelation among samples, randomizing the order of transects would not be appropriate. Thus, spatial autocorrelation in assemblage composition was tested by calculating the Coefficient of Community (CC) similarity index (Sorensen 1948) for each pair of transects within each site. The CC for each pair of transects was plotted against the distance between them (Figure 3). There was no strong trend apparent between the CC and distance and it was concluded that spatial autocorrelation was not prevalent.

Next, 100 randomizations in C++ were used to determine the number of transects required before metric values leveled off. For each randomization, transects were randomly ordered within each site. Next, transect data for successively larger numbers of transects within each site was combined, beginning with the first transect in the sequence. At each step in each randomization, a simulated subsample was generated based on the *Subsample size* results. For each metric and site, the average metric value across the 100 randomizations was plotted against the number of transects. The point at which each metric leveled off was identified by visual inspection. Finally, a similar set of simulations was run for smaller subsample size(s) to examine the influence of subsample size on these plots.

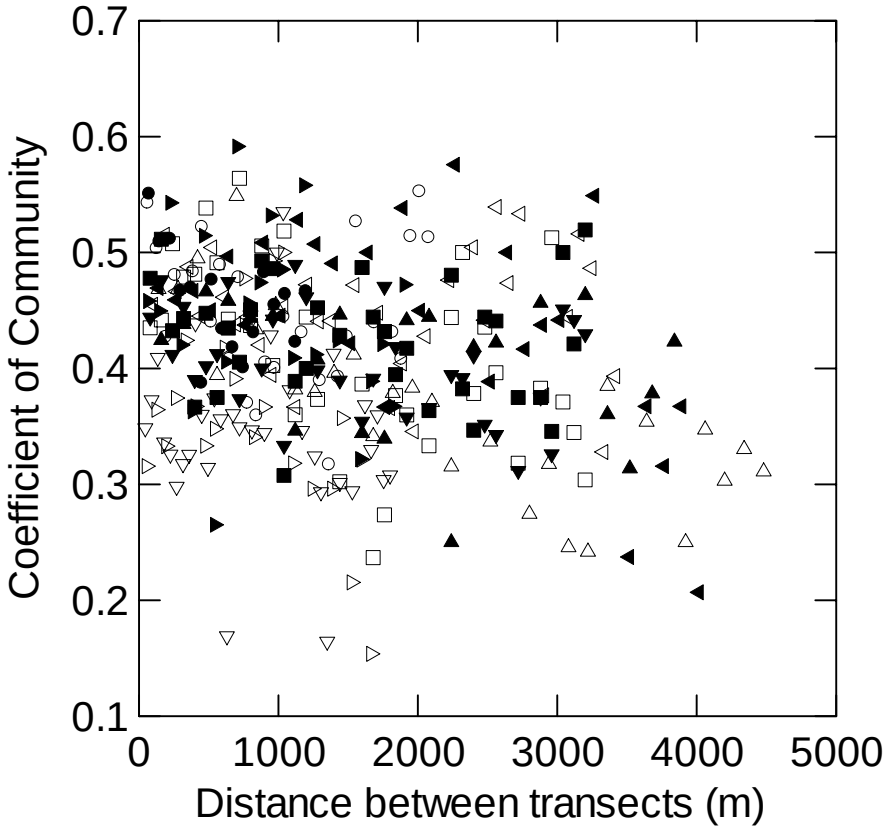

Figure 3: Average Coefficient of Community Index value for each possible pair of transects within each site as a function of the distance between transects. Each type of symbol (shape and fill) represents a different site.

Distance between transects.--After determining the number of transects required per site, the affect of distance between transects on metric values was examined. In each site, pairs of transects were grouped by the distance between them. This created up to four groups with five pairs of transects each with inter-transect distances of 1, 2, 4, and 8 times the mean wetted width. Not all sites had four groups because of shortened reaches as described earlier. For each site, data for all samples within a group were combined and 100 simulations of a 500-organism subsample on each group were performed. The within-site differences among groups were assessed qualitatively from plots of the mean metric value (± 1 SE).

Results

Although all 12 selected sites were sampled, the distance between dams prevented sampling of all transects at three sites, severe weather conditions at two sites, and loss of daylight at one site. The impact of these logistical limitations on data analysis was negligible. Total reach length sampled at individual sites ranged from 1200 to 4480 m in the Great Miami River and from 1680 to 4000 m in the Kentucky River. The range and median of water chemistry and physical habitat variables at study sites are presented in Table 2. The number of organisms per transect sample ranged from 63 to 2369 with a mean of 477.

Subsample size.--The metric values for simulated samples quickly leveled off for percentage metrics, but not for richness metrics (Figure 4). However, the

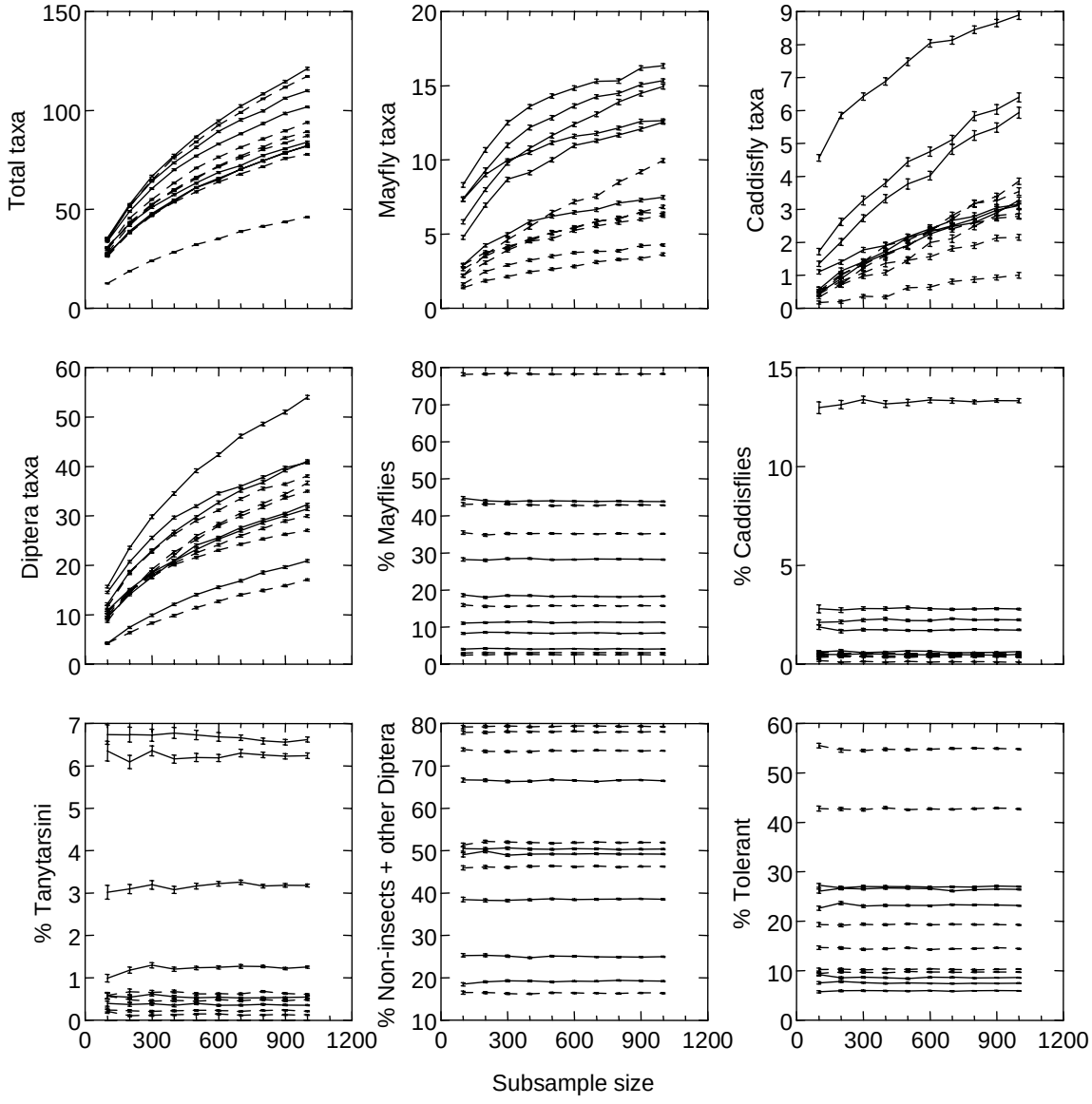

Figure 4: Results of subsample size simulations for each site and metric. Solid lines represent sites in the Great Miami River and dashed lines represent sites in the Kentucky River. Error bars represent 1 standard error of the mean.

difference in richness metric values between any two sites did not change as rapidly after approximately 500 organisms. In fact, the relative return dropped below about 2% beyond 500 organisms (Figure 5), indicating that additional sorting would not provide sufficient additional information in separating sites from one another. There was also a significant drop between 200 and 300 organisms, resulting in relative returns below 5% for 300 or more organisms. This information is useful to note because most state programs subsample 300 or fewer organisms for bioassessment samples in streams. Nonetheless, a subsample of 500 organisms was used for further simulation analyses.

Number of transects.--There was a strong leveling off of richness metric values at approximately six transects (Figure 6). For percentage metrics, the asymptote typically was reached in fewer transects. When this analysis was rerun using a subsample size of only 300 organisms, similar results were achieved (Figure 7).

Distance between transects.--Across all sites, there was no consistent pattern in metrics based on five transects one wetted width apart to five transects eight wetted widths apart (Figure 8). However, within individual sites, there were sometimes very strong differences among the four groups, particularly between the group of transects separated by a distance of eight times the wetted width and the other three groups of transects separated by smaller distances (Figure 8). Retrospective analysis of the physical habitat data suggests that at some

sites, as the distance between transects increased, the likelihood of encountering large

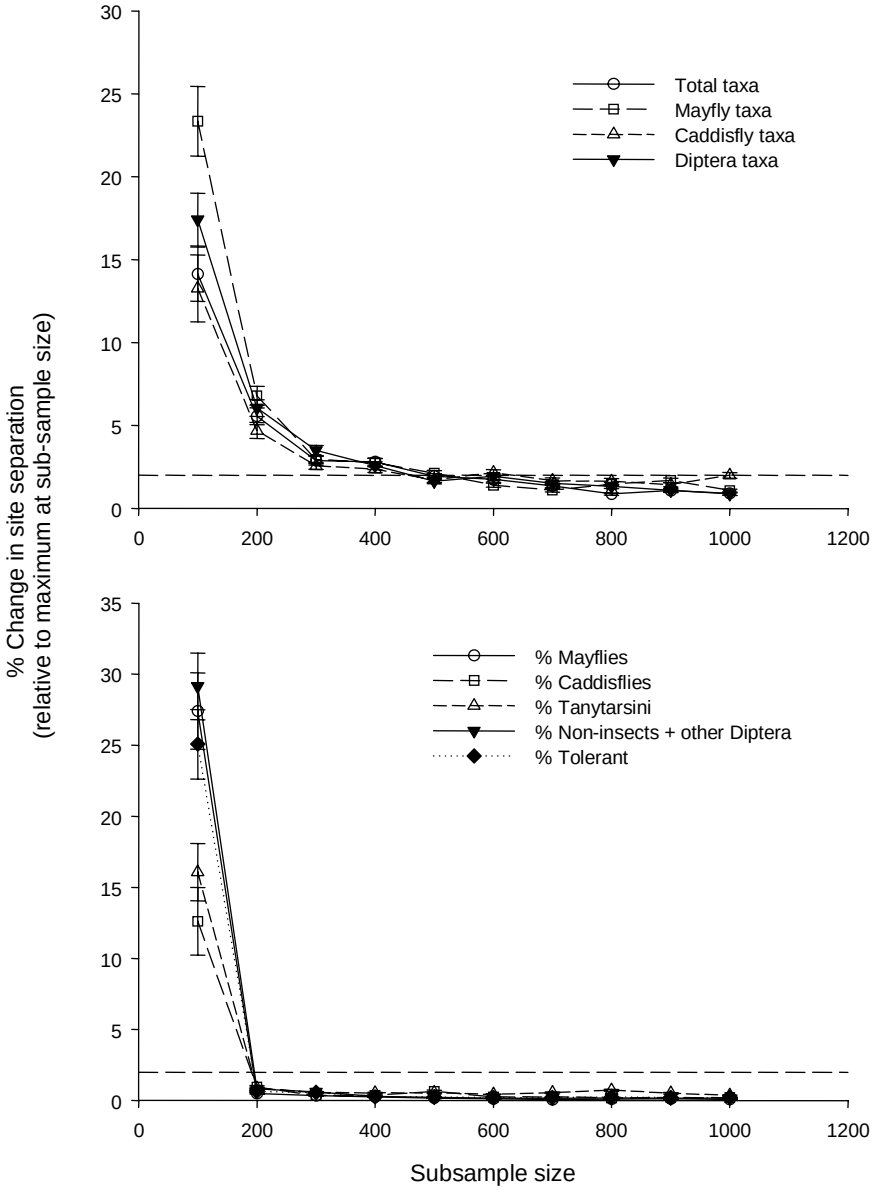

Figure 5: Relative return as a function of subsample size for each richness (top) and percentage (bottom) metric and site. Error bars represent 1 standard error of the mean value.

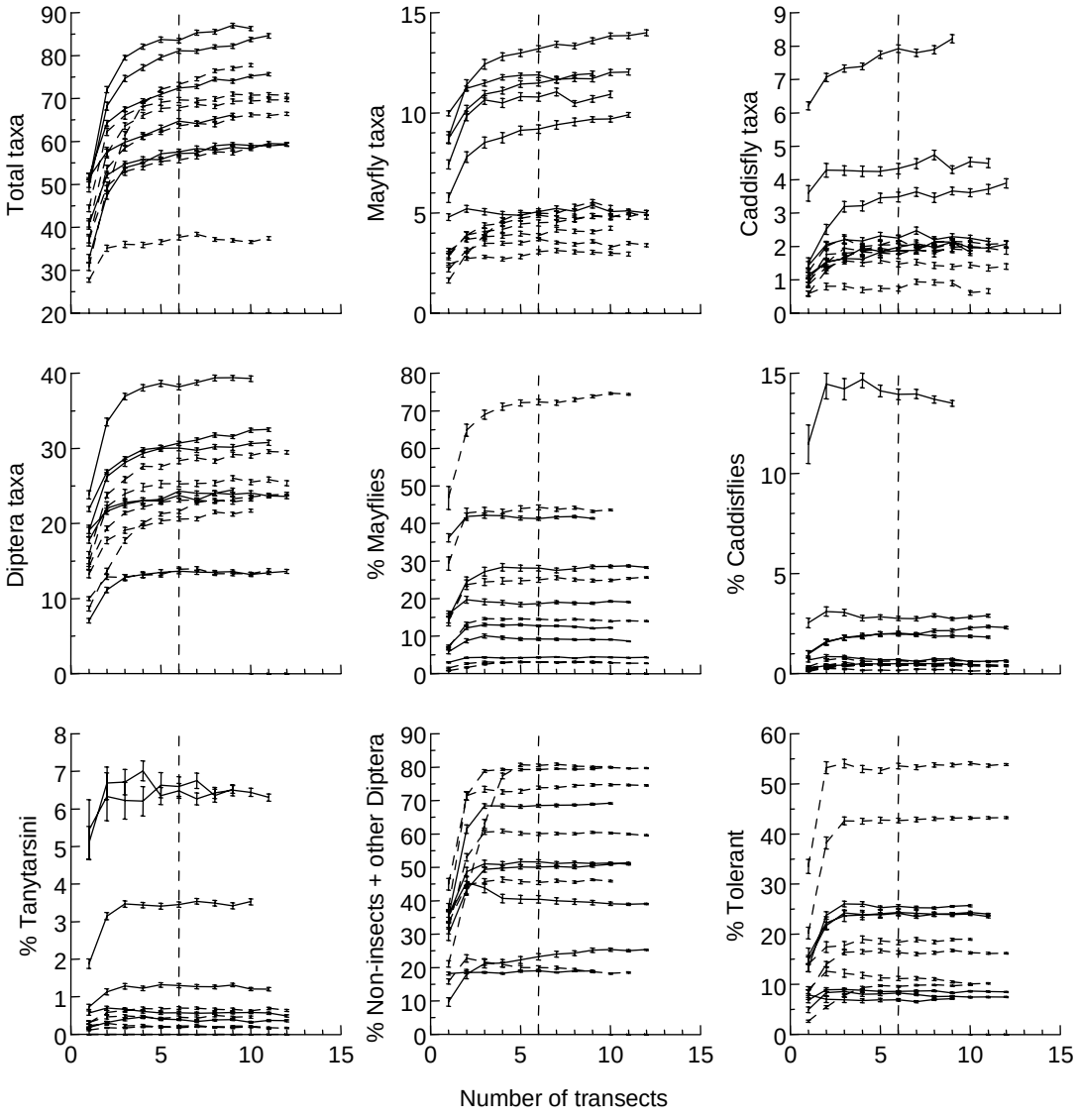

Figure 6: Metric values with increasing number of transects based on 500 organism simulated counts. Solid lines represent Great Miami River and dashed represent Kentucky River sites. Error bars represent 1 standard error. Vertical dashed line represents estimated point at which leveling-off occurs.

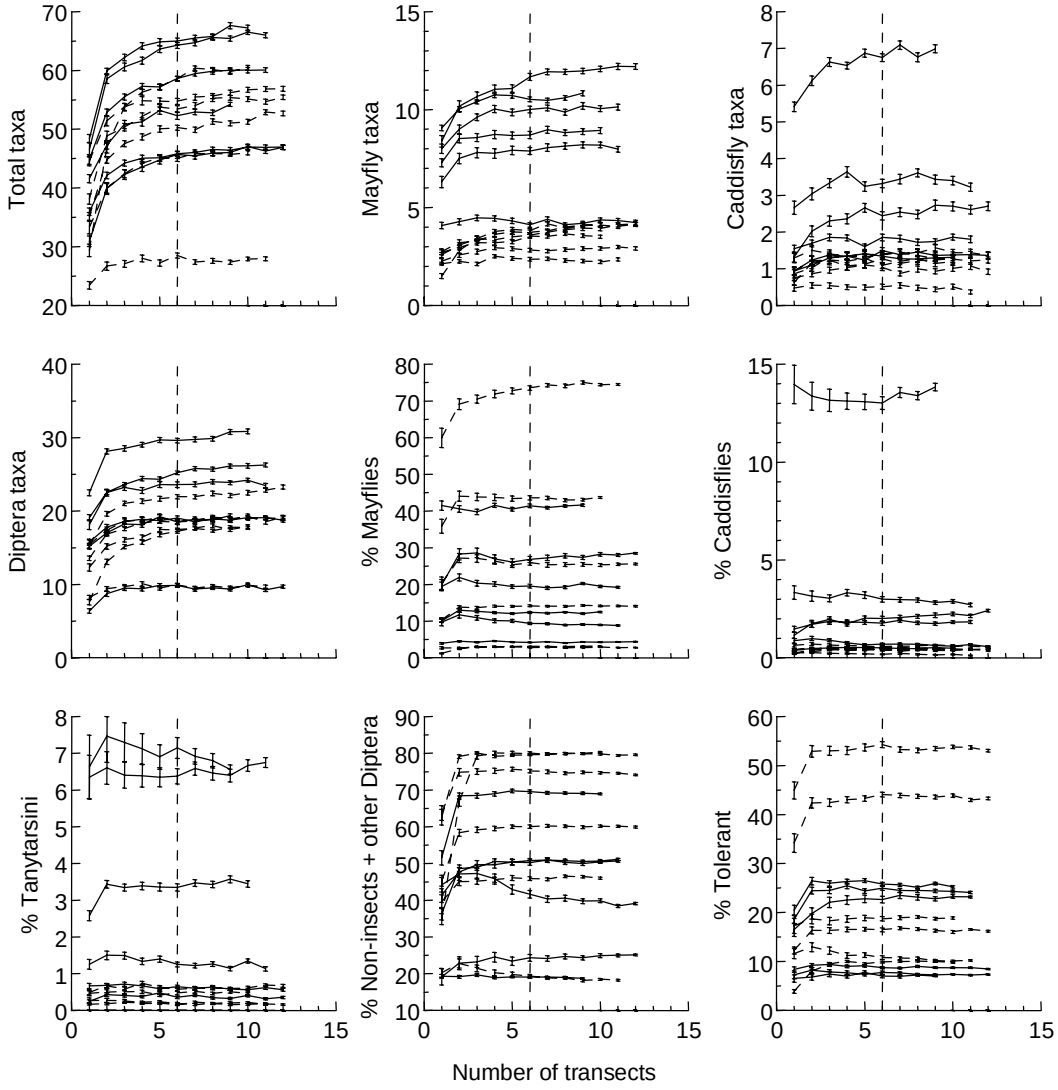

Figure 7: Metric values with increasing number of transects based on 300 organism simulated counts. Solid lines represent Great Miami River and dashed represent Kentucky River sites. Error bars represent 1 standard error. Vertical dashed line represents estimated point at which leveling-off occurs.

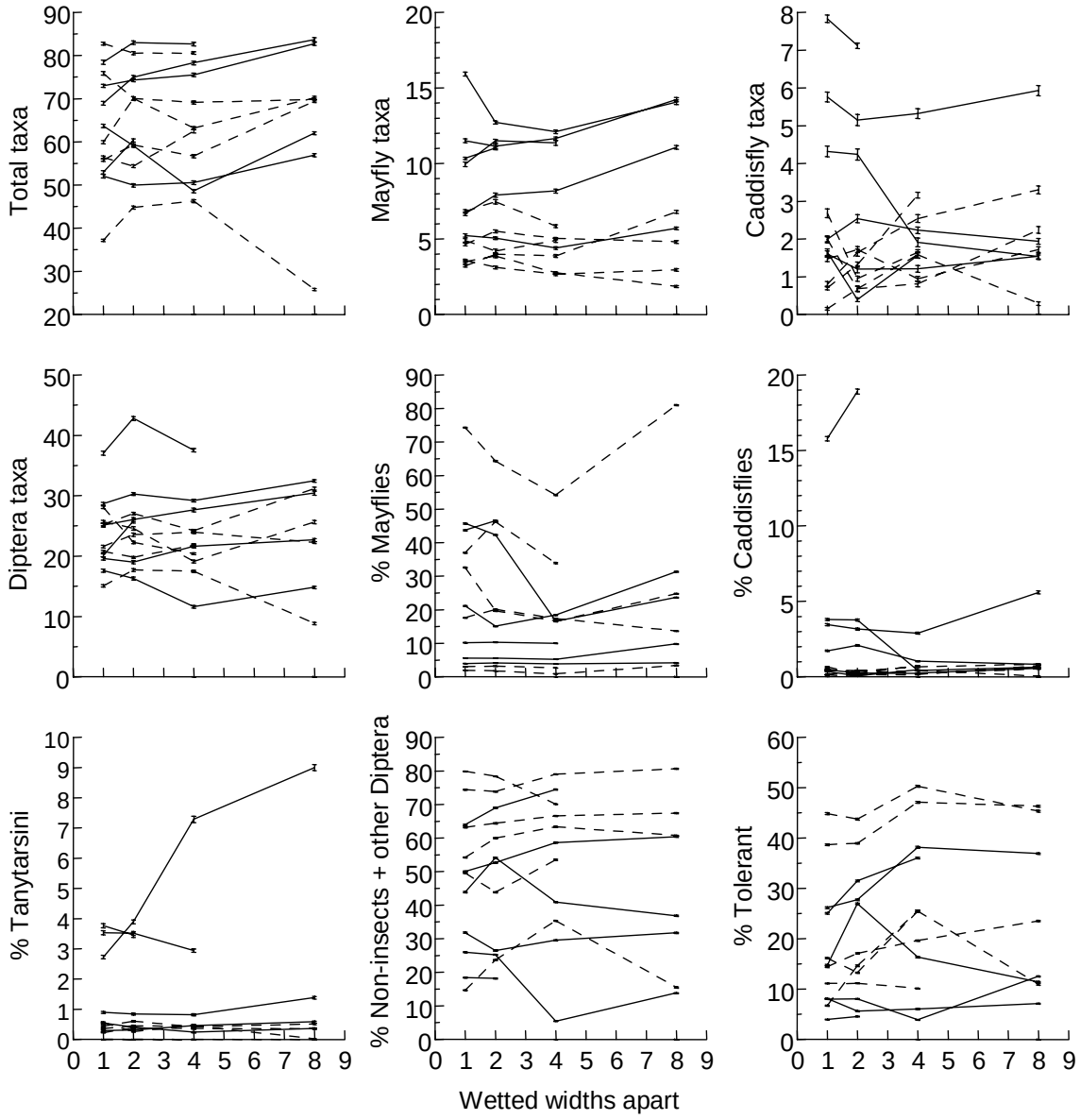

Figure 8: Metric values across groups (define) equidistant by varying numbers of channel widths for each site. Solid lines represent Great Miami River sites and dashed lines represent Kentucky River sites.

variability in one or more gross physical habitat features (e.g., thalweg depth, substrate composition) increased, but the habitat feature(s) causing the variability was not the same across sites.

Given the study results and the reality that the vast majority of non-wadeable streams and rivers have been modified in ways that undermine the rationale of using geomorphology as a determinant of reach length, a fixed reach length rather than one based on multiples of the wetted widths is appropriate. From the study data, the range of distances covered by six transects was 270 m to 960 m, with a median of 480 m and a mean of 542.5. Based on these results, and taking into consideration the logistics of field sampling, the reach length was set at 500 m. This results in six transects separated by a fixed distance of 100 m each for the LR-BP.

Discussion

A new sampling protocol was developed for sampling large river macroinvertebrates that is specifically designed to perform well across all shore-line habitats and river types, integrate different habitats, and thus represent site conditions accurately. The appropriate fixed count for laboratory subsampling size to use with this method was also determined, based on the ability to separate sites of differing macroinvertebrate composition.

This research indicates that, using the sampling technique discussed herein, a representative sample of the benthic macroinvertebrate fauna of the study reaches was collected by sampling both banks of six transects. These

results were achieved because the sampling method and design effectively sampled the benthic macroinvertebrate fauna of the dominant habitat types within the reach. It is further recommended that transects be evenly-spaced over a 500-m distance, a distance equivalent to that used for many years by several state agencies for sampling riverine benthic macroinvertebrate and fish assemblages (e.g., Ohio EPA, 1989, Royer et al. 2001).

The recommended design results in a composite sample consisting of 24 20-second kick-net samples and 12 timed samples collected with a D-frame net in habitats complementing those sampled by the kick-net. Therefore, the final composite sample from the 500-m reach consists of 36 subsamples collected by two complementary sampling techniques. It is likely this approach will over-sample sites of uniform composition, but the goal was to develop a standardized LR-BP that would perform well across sites of differing habitat composition. These conclusions agree with Bartsch et al. (1998), who stated that 18 to 40 subsamples may be required to adequately sample large flood-plain rivers, and that no single sampling technique would efficiently and adequately sample all components of a riverine macroinvertebate community. It should be noted that this LR-BP for macroinvertebrates has only been tested in main-channel habitats. It may work equally well on off-channel habitats, but this remains to be tested.

It is recommended that the field method be coupled with a fixed laboratory subsample size of 300 or 500 organisms to maximize effectiveness of the LR-BP for bioassessment purposes. The fixed laboratory subsample size of 500 organisms does offer lower variability for percentage metrics, but variability for

richness metrics was higher. Nonetheless, 300 is in all probability sufficient for most study needs. This recommendation is based on the response of the tested metrics and the observation that the ability to separate sites of different macroinvertebrate composition generally did not increase with larger subsample sizes. Studies on other systems have recommended a broad range of subsample sizes as sufficient (Barbour and Gerritsen, 1996, Vinson and Hawkins, 1996, Growns et al. 1997, Somers et al. 1998). However, a one-size-fits-all subsample size should not be expected, since the quality of information needed by researchers and managers can vary depending on individual studies (Doberstein et al. 2000). The best strategy for determining an appropriate subsample size is to first determine the data quality requirements to meet study objectives and then determine the appropriate subsample size from collected data. This seems especially appropriate when developing new or modifying existing field methods.

An advantage of the proposed protocol is that field crews targeting non-wadeable streams and rivers can be sent into the field with a single method that works well across a variety of site types. While the method has been designed to perform well in a variety of habitats, resulting data should be interpreted with appreciation for coarse habitat characteristics. This information would ideally be derived from habitat data collected concomitantly with the faunal data. Sites can then be categorized into river types (e.g., impounded vs. free-flowing or lowland vs. upland rivers) or even habitat types within a specific river (e.g., sandy- vs. cobble-bottom) in a more controlled environment (i.e., in the laboratory).

The structured nature of the LR-BP provides a standardized sampling protocol that produces a representative sample from the varying habitats and changing impoundment conditions (through time and space) encountered within and across large rivers. Sites in this study varied from free-flowing to those with hydrologic modifications associated with lock-and-dam systems, habitat modifications due to channelization, and the presence of low-head dams. In habitats where both sampling methods could be performed, one method did not supersede the other, and both were performed. At others, for example, the banks of a sampling station may have been too steep, rendering the collection of a sample via the kick-net method logistically impossible. However, using the D-frame net from the boat, the benthos would be sampled. Hence, a habitat that would have gone unrepresented using a sampling approach that relied purely on kick-net sampling, was still represented in the composite sample of the site.

As a result of the additional equipment required to work in non-wadeable streams and rivers (i.e., boats and associated equipment), the effort required to secure a representative sample for bioassessment generally exceeds that required in wadeable streams. Given these realizations, the proposed sampling method is cost effective, logistically feasible, and collects a representative sample for bioassessment purposes. Critical elements of the LR-BP include the complementary sampling techniques, the distance of the sample reach, the number of transects at which both banks are sampled, and the subsample size in the laboratory.

Future Research.-- With development of this initial design, additional field sampling has been conducted to allow performance-based testing (Diamond et al. 1996) of the field and laboratory components of the LR-BP. Additional research may also be needed to determine applicability of the LR-BP for use in riverine ecosystems functioning differently than those described in this study (e.g., floodplain-river ecosystems, riverine-influenced reservoirs, fast-flowing rivers). Possible modifications to the method to streamline its application in the field include using the D-frame net configuration for both the kick- and dip-net sampling, setting a depth criterion for kick-net sampling (e.g., 1 m), and using a fixed distance for sample zones (e.g., 10 m). Experimenting with an area quantification of the dip-net sampling may also be considered for use in studies requiring full quantification of sampling effort.

Literature Cited

Allen, J. D. 2000. Stream Ecology: Structure and function of running waters. Kluwer Academic Publishers. Dordrecht, Netherlands.

Angermeier, P. L., and I. J. Schlosser. 1989. Species-area relationships for stream fishes. Ecology 70:1450-1462.

Barbour, M. T., and J. Gerritsen. 1996. Subsampling of benthic samples: a defense of the fixed-count method. Journal of the North American Benthological Society 15:386-391.

Barbour, M. T., J. Gerritsen, B. D. Snyder, and J. B. Stribling. 1999. Rapid bioassessment protocols for use in streams and wadeable rivers: periphyton, benthic macroinvertebrates, and fish, second edition. EPA 841-B-99-002. U.S. Environmental Protection Agency; Office of Water; Washington, D.C.

Bartsch, L. A., W. B. Richardson, and T. J. Naimo. 1998. Sampling benthic macroinvertebrates in a large floodplain river: considerations of study design, sample size, and cost. Environmental Monitoring and Assessment 52:425-439.

Benke, A. C. 1990. A perspective on American's vanishing streams. Journal of the North American Benthological Society 9(1):77-88.

Blocksom, K. A., and J. E. Flotemersch. 2004. Comparison of Macroinvertebrate Sampling Methods for Nonwadeable Streams. Environmental Monitoring and Assessment (In Press).

Cairns, J., Jr. 1982. Artificial Substrates. Ann Arbor Science Publications, Inc., Ann Arbor, Michigan.

Cao, Y., D. P. Larsen, and R. M. Hughes. 2001. Evaluating sampling sufficiency in fish assemblage surveys: a similarity-based approach. Canadian Journal of Fisheries and Aquatic Sciences 58:1782-1793.

Carter, J. L., and V. H. Resh. 2001. After site selection and before data analysis: sampling, sorting, and laboratory procedures used in stream benthic macroinvertebrate monitoring programs by USA state agencies. Journal of the North American Benthological Society 20:658-682.

Cochran, W. G., 1977. Sampling Techniques. John Wiley and Sons, New York.

Diamond, J. M., M. T. Barbour, and J. B. Stribling. 1996. Characterizing and comparing bioassessment methods and their results: a perspective. Journal of the North American Benthological Society 15:713-727.

Doberstein, C. P., J. R. Karr, and L. L. Conquest. 2000. The effect of fixed-count subsampling on macroinvertebrate biomonitoring in small streams. Freshwater Biology 44:355-371.

Flotemersch, J. E., B. C. Autrey, and S. M. Cormier. 2001. Comparisons of boating and wading methods used to assess the status of flowing waters. U.S. Environmental Protection Agency, Cincinnati, Ohio, EPA/600/R-00/108.

Frissell, C. A., W. J. Liss, C. E. Warren, and M. D. Hurley. 1986. A hierarchical framework for stream habitat classification: viewing streams in a watershed context. Environmental Management 10(2):199-214.

Gormann, O. T., and J. R. Karr. 1978. Habitat structure and stream fish communities. Ecology 59:507-515.

Growns, J. E., B. C. Chessman, J. E. Jackson, and D. G. Ross. 1997. Rapid assessment of Australian rivers using macroinvertebrates: cost and efficiency of 6 methods of sample processing. Journal of the North American Benthological Society 16:682-693.

Herlihy, A. T., and J. Lazorchak J. 2000. Initial Site Procedures. Pages 4.1-4.8 *in* J. M. Lazorchak, B. H. Hill, D. K. Averill, D. B. Peck, and D. J. Klemm (editors) Environmental monitoring and assessment program-surface waters: field operations and methods for measuring the ecological condition of non-wadeable rivers and streams. EPA 620/R-00/007. Office of Research and Development, US Environmental Protection Agency, Cincinnati, Ohio.

Humphries, P., J. E. Growns, L. G. Serafina, J. H. Hawking, A. J. Chick, and P. S. Lake. 1998. Macroinvertebrate sampling methods for lowland Australian rivers. Hydrobiologia 364(2):209-218.

Hughes, R. B., P. R. Kaufmann, A. T. Herlihy, S. S. Intelmann, S. C. Corbett, M. C. Arbogast, and R. C. Hjort. 2002. Electrofishing distance needed to estimate fish species richness in raftable Oregon rivers. North American Journal of Fisheries Management 22:1229-1240.

Hynes, H. B. N. 1970. The ecology of running water. Liverpool University Press, Liverpool.

Karr, J. R., and E. W. Chu. 1999. Restoring Life in Running Waters: Better Biological Monitoring. Island Press, Washington, D.C.

Kaufmann PR. 2000. Physical Habitat Characterization – non-wadeable rivers. Pages 6.1-6.29 *in* J. M. Lazorchak, B. H. Hill, D. K. Averill, D. B. Peck, and D. J. Klemm (editors) Environmental monitoring and assessment program-surface waters: field operations and methods for measuring the ecological condition of non-wadeable rivers and streams. EPA 620/R-00/007. Office of Research and Development, US Environmental Protection Agency, Cincinnati, Ohio.

Kentucky River Authority. 1999. Kentucky River Basin Status Report. http://water.nr.state.ky.us/dow/pubs/kybasin97/status1.htm. Kentucky River Authority, Frankfort, Kentucky.

Kilgour, B. W., and D. R. Barton. 1999. Associations between stream fish and benthos across environmental gradients in southern Ontario, Canada. Freshwater Biology 41:553-566.

Klemm, D. J., P. A. Lewis, F. A. Fulk, and J. M. Lazorchak. 1990. Macroinvertebrate field and laboratory methods for evaluating the biological integrity of surface waters, EPA 600/4-90/030, U.S. Environmental Protection Agency, Cincinnati, Ohio.

Klemm, D. J., J. M. Lazorchak, and D. V. Peck. 2000. Benthic macroinvertebrates. Pages 9.1-9.14 *in* J.M. Lazorchak, B. H. Hill, D. K. Averill, D.B. Peck, and D. J. Klemm (editors) Environmental monitoring and assessment program-surface waters: field operations and methods for

measuring the ecological condition of non-wadeable rivers and streams. EPA 620/R-00/007. Office of Research and Development, US Environmental Protection Agency, Cincinnati, Ohio.

Leopold, L. B., M. G. Wolman, and J. P. Miller. 1964. Fluvial processes in geomorphology. W.H. Freeman: San Francisco.

Li, J., A. Herlihy, W. Gerth, P. Kaufmann, S. Gregory, S. Urquhart, and D. Larsen. 2001. Variability in stream macroinvertebrates at multiple spatial scales. Freshwater Biology 46:87-97.

Loucks, D. P. 2003. Managing America's rivers: who is doing it? International Journal of River Basin Management 1:23-31.

Lyons, J. 1992. The length of stream to sample with a towed electrofishing unit when fish species richness is estimated. North American Journal of Fisheries Management 12:198-203.

Manly, B. F. J. 2001. Statistics for environmental science and management. Chapman and Hall/CRC. Boca Raton, FL.

Metcalfe-Smith, J. L. 1994. Biological water-quality assessment of rivers: use of macroinvertebrate communities. Pages 144-170 *in* P. Colow and G. E. Petts (editors). The rivers handbook: Hydrological and ecological principles. Blackwell Scientific. Oxford, United Kingdom.

Moulton, S. R., II, J. G. Kennen, R. M. Goldstein, and J. A. Hambrook. 2002. Revised protocols for sampling algal, invertebrate, and fish communities as part of the National Water Quality Assessment Program. U.S. Geologic Survey Open File Report 02-150.

Ohio EPA. 1997. Biological and Water Quality Study of the Middle and Lower Great Miami River and Selected Tributaries, 1995. Division of Surface Water, Monitoring and Assessment Section, Ohio Environmental Protection Agency, Columbus, Ohio.

Ohio EPA. 1988. Biological criteria for the protection of aquatic life: Volume II: users manual for biological field assessment of Ohio surface waters. Ohio EPA, Division of Water Quality Monitoring and Assessment, Surface Water Section, Columbus, Ohio.

Ohio EPA. 1989. Biological criteria for the protection of aquatic life: Volume III: Standardized biological field sampling and laboratory methods for assessing fish and macroinvertebrate communities. Ohio EPA, Division of Water Quality Monitoring and Assessment, Surface Water Section, Columbus, Ohio.

Poulton, B. C., M. L. Wildharber, C. S. Charbonneau, J. F. Fairchild, B. G. Mueller, and C. J. Schmitt. 2003. A longitudinal assessment of the aquatic macroinvertebrate community in the channelized lower Missouri River. Environmental Monitoring and Assessment 85:23-53.

Rosenberg, D. M., and V. H. Resh. 1993. Freshwater Biomonitoring and Benthic Macroinvertebrates, Chapman and Hall, London.

Royer, T. V., C. T. Robinson, and G. W. Minshall. 2001. Development of macroinvertebrate-based index for bioassessment of Idaho rivers. Environmental Management 27:627-636.

Somers, K. M., R. A. Reid, and S. M. David. 1998. Rapid biological assessments: how many animals are enough? Journal of the North American Benthological Society 17:348-358.

Sorensen, T. 1948. A method of establishing groups of equal amplitude in plant sociology based on similarity of species content. Det Kongelige Danske Videnskabernes Selskab, Biologiske Skrifter, (Copenhagen) 5:1-34.

Sweeting, R. A. 1994. River Pollution. Pages 23-32 *in* P. Colow and G.E. Petts (editors) The rivers handbook: Hydrological and ecological principles. Blackwell Scientific. Oxford, United Kingdom.

Vannote, R. L., G. W. Minshall, K. W. Cummins, J. R. Sedell, and C. E. Cushing. 1980. The river continuum concept. Canadian Journal of Fisheries and Aquatic Science 37:130-137.

Vinson, M. R., and C. P. Hawkins. 1996. Effects of sampling area and subsampling procedure on comparisons of taxa richness among streams. Journal of the North American Benthological Society 15:392-399.

Vogelmann, J. E., T. L. Sohl, P. V. Campbell, and D. M. Shaw. 1998 Regional land cover characterization using landsat thematic mapper data and ancillary data sources. Environmental Monitoring and Assessment 51:415-428.

www.miamiconservancy.org. 4/22/04. Great Miami River Watershed.

United States
Environmental Protection
Agency

Office of Research
and Development (8101R)
Washington, DC 20460

Official Business
Penalty for Private Use
$300

EPA 600/R-04/177
October 2004
www.epa.gov

Please make all necessary changes on the below label, detach or copy, and return to the address in the upper left-hand corner.

If you do not wish to receive these reports CHECK HERE

☐; detach, or copy this cover, and return to the address in the upper left-hand corner.

PRESORTED STANDARD
POSTAGE & FEES PAID
EPA
PERMIT No. G-35