

COMPARISON OF MACROINVERTEBRATE SAMPLING METHODS FOR NONWADEABLE STREAMS

KAREN A. BLOCKSOM* and JOSEPH E. FLOTEMERSCH

*National Exposure Research Laboratory, U.S. Environmental Protection Agency, Ecological
Exposure Research Division, 26 W. Martin Luther King Drive, Cincinnati, Ohio U.S.A.*

*(*author for correspondence, e-mail: blocksom.karen@epa.gov)*

(Received 1 July 2003; accepted 9 April 2004)

Abstract. Bioassessment of nonwadeable streams in the United States is increasing, but methods for these systems are not as well-developed as for wadeable streams. In this study, we compared six macroinvertebrate field sampling methods for nonwadeable streams adapted from those used by three major programs: the U.S. Environmental Protection Agency's Environmental Monitoring and Assessment Program-Surface Waters, the U.S. Geological Survey's National Water Quality Assessment Program, and the Ohio Environmental Protection Agency, Division of Surface Water Biocriteria Program. We performed all six methods at 60 sites across four rivers and measured water chemistry and physical habitat at each site to assess abiotic condition. Sites were divided into two groups: those influenced by navigational lock and dam structures (restricted flow, or RF) and those free-flowing or with lowhead dams (run-of-the-river, or ROR). Metrics based on passive Hester-Dendy artificial substrate samplers differed greatly from active sampling methods (i.e., using nets) but represented abiotic conditions well in both ROR and RF sites. Although metric values were similar across certain sampling methods, the metrics significantly correlated with abiotic variables varied among methods and between ROR and RF sites. These results emphasize that methods are not interchangeable, and the ability to detect certain stressors depends on sampling method.

Keywords: nonwadeable rivers, sampling methods, bioassessment, run-of-the-river, restricted flow, macroinvertebrates

1. Introduction

Following passage of the Clean Water Act of 1972, several bioassessment programs have been developed and are in use by state and local agencies throughout the United States. These programs initially focused on point source detection and monitoring, leading to the development of sampling methods oriented toward surveys of smaller, wadeable streams and rivers. However, the rise in awareness of the substantial problem of non-point source pollution has recently led to an increased interest in larger rivers where effects of several stressors are often integrated and sources are less clear.

Biological communities and the type and quality of habitats change with increasing stream size (Vannote *et al.*, 1980). Thus, there is a critical need for bioassessment

The U.S. Government's right to retain a non-exclusive, royalty free licence in and to any copyright is acknowledged.

methods for measuring aquatic community health that are specifically designed for larger rivers. Most current methods for nonwadeable systems have been developed from existing methods developed for wadeable streams (Flotemersch *et al.*, 2001). Often, wadeable stream methods are applied in shallow areas (e.g., near shore) of nonwadeable waters or from a boat with no further modifications (Flotemersch *et al.*, 2001). However, there has been no direct comparison among nonwadeable methods to understand potential differences and their relative effectiveness. For that reason, the primary focus of this study was to evaluate methods currently used to assess macroinvertebrate assemblages in nonwadeable flowing waters.

This study examined existing methods (with minor modifications) used by three bioassessment programs. These programs included the Environmental Monitoring and Assessment Program for Surface Waters (EMAP-SW) of the U.S. Environmental Protection Agency, the National Water Quality Assessment Program (NAWQA) of the U.S. Geological Survey and the Biological and Water Quality Monitoring and Assessment Program of the Ohio Environmental Protection Agency (Ohio EPA) Division of Surface Water. These programs were established to assess water bodies at a range of spatial scales (Flotemersch *et al.*, 2001). Thus, while the programs share similar goals, they vary in sampling design and in their specific measurements based upon the objectives of the program.

Our main objective was to evaluate differences among sampling methods, both in the macroinvertebrate assemblage collected and in the ability to detect various types of impairment. To make the results relevant to bioassessment programs, we chose to use metrics based on macroinvertebrate assemblage structure to assess these differences. For each sampling method, we directly compared metrics among methods and examined relationships between metrics and both univariate and multivariate measures of abiotic condition. In doing this, we also recognized that major hydrologic features (e.g., impoundment) could affect relationships between biota and abiotic variables. Thus, methods might be effective in one type of system and not another. In this paper, we examine, within each impoundment class, the influence of sampling method on biological metric values and on associations between biological metrics and abiotic measures of site condition, and we identify the methods most related to specific types of measures of abiotic condition within each impoundment type.

2. Methods

2.1. STUDY AREAS

We sampled during summer 1999 in four tributaries of the Ohio River, the Great Miami and Scioto rivers in Ohio and the Kentucky and Green rivers in Kentucky. These rivers were selected to reflect a range of watershed land uses, stressors, hydrogeological influences and habitat types (Table I). Twenty sites each were selected in the Great Miami and Scioto rivers, and ten sites each in the Kentucky

TABLE I

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

Basin descriptors	River basin			
	Great Miami (19 sites)	Scioto (17 sites)	Kentucky (9 sites)	Green (9 sites)
Drainage basin (km ²)	13,947 ^a	16,879 ^b	18,130 ^c	23,041 ^d
River length (km)	233.4 ^a	370.2 ^b	410.4 ^c	330.1 ^d
Average gradient (m/km)	0.74 ^a	0.44 ^b	0.13 ^c	e
Urban land use (%)	5.90 (2.64)	6.91 (1.87)	7.81 (0.52)	7.98 (0.32)
Agriculture land use (%)	82.55 (3.21)	76.39 (5.40)	80.18 (1.34)	80.56 (0.83)
Forested land use (%)	10.10 (1.44)	15.34 (5.74)	10.83 (1.60)	10.27 (0.69)

^aOhio EPA (1997a).

^bOhio EPA (1997b).

^cKentucky River Authority (1999).

^dKDEP (1976).

^eNot available.

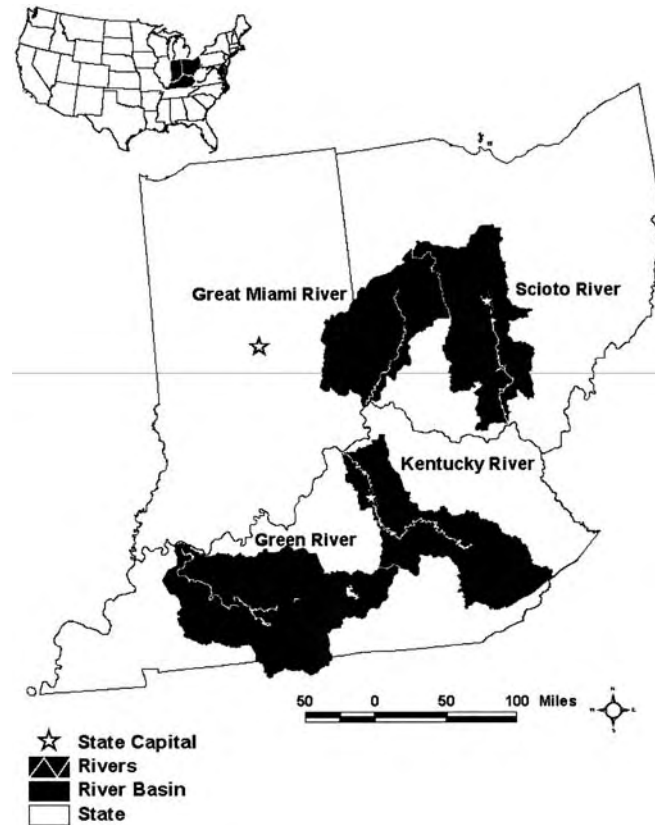


Figure 1. Great Miami, Scioto, Kentucky, and Green river basins.

and Green rivers (Figure 1). For site-specific reach placement, we avoided obvious stressors (e.g., major outfalls) and physical features (e.g., stream confluences or bridges) that might influence comparisons.

2.2. SAMPLE COLLECTION AND PROCESSING

The length of sampling reaches was equal to 40 times the mean wetted width, with a maximum length of 2000 m. Within each sampling reach, 11 evenly-spaced transects were set at intervals of four times the mean wetted width, with a maximum distance of 200 m between transects.

In each stream reach, macroinvertebrates were collected with six methods: two EMAP-based, two NAWQA-based, and two Ohio EPA-based methods (Table II). The EMAP-based methods were the drift net (DN) and the kick net (KN) methods using EMAP protocols (Lazorchak *et al.*, 2000). If current velocity was at least 0.05 m/s at the sampling point, two drift nets (30.48 cm × 45.72 cm opening, 595- μ m mesh) were deployed at the lower end of a site for 3–4 h during daytime hours, preferably with one net in shallower and one in deeper water. The KN

TABLE II
Macroinvertebrate sampling methods performed at each site, adapted for use from existing methods used by EMAP-SW, NAWQA, and Ohio EPA programs

Method (program)	Description	Distance (m)	Banks	Mesh size (μ m)
DN (EMAP-SW)	Two drift nets set for 3–4 h during daylight hours at lower end of reach, in water of adequate depth	Two locations	Either or mid-reach	595
KN (EMAP-SW)	Two 20-s kicks at each transect	11 transects over 2000 m	One	595
QMH500 (Ohio EPA)	All available habitats with D-frame net and hand-picking, minimum 30 min	500 m	Both	595
HD (Ohio EPA)	Five multi-plate samplers attached to a single block in lower 500 m of reach	Single location	One	595
QMH1000 (NAWQA)	All available habitats with D-frame net and hand picking, approximately 60 min	1000 m	Both	210
RTH (NAWQA)	Five–six sites of the richest targeted habitats (according to prioritized list)	2000 m	Both	425

method consisted of two 20-s kicks using a rectangular frame net (50 cm × 30 cm opening, 595- μ m mesh) at each of the 11 transects. All kicks were conducted on a single bank and composited into a single sample for the entire site. For both methods, samples were processed in the field with a 595- μ m sieve.

Two NAWQA-based methods were performed: the richest targeted habitat method (RTH) and the 1000-m qualitative multihabitat (QMH1000) (based on Cuffney *et al.*, 1993). For this study, the RTH method consisted of sampling five to six areas of the richest habitats (e.g., rocks, snags, macrophytes) along either bank of a 1000-m stretch of the reach. In each habitat, a 50 cm × 50 cm area was sampled. A variety of samplers were used, depending on the type of habitat sampled. All devices used to collect macroinvertebrates and process materials had a mesh size of 425 μ m. All samples were then composited into a single sample for each site. The QMH1000 method consisted of sampling macroinvertebrates from all habitat types along both banks of a 1000-m section of the reach using a D-frame dip net with a mesh size of 210 μ m. The material collected was processed in the field with a 210- μ m mesh sieve.

We executed two Ohio EPA-based methods: the 500-m qualitative multihabitat (QMH500) and Hester-Dendy multi-plate substrate samplers (HD) (based on Ohio EPA, 1987). The QMH500 method involved sampling all available habitat types on both banks over a 500-m distance using a D-frame net (595- μ m mesh). Habitats were sampled for a minimum of 30 min or until no new taxa were observed by gross examination. The HD samplers were deployed in a run area at each site for approximately 6 weeks. Total surface area of each sampler was approximately 0.092 m². Upon retrieval, samplers were disassembled, and organisms and debris were processed with a 595- μ m sieve.

All samples were preserved in the field with ethyl alcohol and transported to the laboratory, where they were processed following EMAP-SW laboratory protocols (Klemm *et al.*, 1990) so that field and laboratory methods were not confounded. Each sample was placed in a gridded ceramic pan with 40 squares. Squares were chosen randomly and were completely sorted until at least 270 (within 10% of 300 organisms) organisms were picked or until the entire sample was picked. All organisms sorted were identified to the lowest practical taxon depending on the condition of the specimen and availability of taxonomic keys.

Additional information was collected from each site to supplement the macroinvertebrate data. Crews collected physical habitat data following EMAP protocols for nonwadeable streams (Lazorchak *et al.*, 2000). An integrated water sample was collected from each sampling reach and evaluated for sulfate, nitrate, total Kjeldahl nitrogen, ammonia, total phosphorus, chloride, and total suspended solids (TSS) concentrations. Samples were evaluated using EMAP-SW laboratory protocols (Klemm *et al.*, 1990). Conductivity and water temperature were measured *in situ* using a YSI Model 85 meter in the center of the sampling reach.

Finally, 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. Then, proportions of forest, agriculture, and urban (including residential) land uses were calculated within the riparian corridor. This approach was used instead of calculating land use in the entire watershed because these sites represent very large watersheds, and using the watershed approach would mean that each site would incorporate all of the watersheds upstream of it. For this reason, the watershed approach resulted in very little overall variation in land use characteristics among sites. However, the riparian corridor approach provides an estimate of more immediate land use disturbances that are likely to impact biological condition.

2.3. DATA ANALYSIS

In light of the obvious hydrologic differences among sites, we divided them into two groups: (a) restricted flow (RF) sites, which are heavily influenced by navigational lock and dam structures built to support commercial traffic, and (b) run-of-the-river (ROR) sites, which are free-flowing or have only low head dams. The distributions of physical habitat and water chemistry characteristics for each group are provided in Table III.

As a way to visualize differences in species composition among methods and between impoundment types, we performed non-metric multidimensional scaling (NMS) in PC-ORD for Windows v. 3.20 (MjM Software, Gleneden Beach, Oregon) with generic level data for all sites and methods combined. This ordination technique positions sites relative to one another based on the ranks of dissimilarities in taxonomic composition and assumes only that the relationship between sample separation and dissimilarities is monotonic (Minchin, 1987; Tong, 1992). Only taxa present in at least 5% of samples were retained in the analysis. Abundances were transformed using $\log(x + 1)$ to control for skewed distributions, and the dissimilarity measure used was Euclidean distance. The stress function measures the fit of the relationship between the original dissimilarities, or distances, among samples and those in the ordination. Stress was calculated for solutions of 1, 2, 3, and 4 dimensions, and the solution beyond which the decrease in the stress function was relatively small was used. In plots of ordination axes, individual samples were identified by impoundment type and sampling method to look for patterns of taxonomic composition.

Using the data collected by each method, we calculated 42 macroinvertebrate assemblage metrics, both those traditionally considered in stream bioassessment and others more typically associated with lentic systems (Table IV). These metrics represented taxa richness, assemblage composition, pollution tolerance, and functional feeding group aspects of the macroinvertebrate assemblages. To narrow the set of metrics included in analyses, we evaluated metric distribution characteristics within each impoundment grouping and collection method. Metrics that provided little information, including richness metrics with a range of less than five taxa and percentage metrics with a range of less than 10%, were excluded. Finally, we

TABLE III

Range (median) of water chemistry and physical habitat variables as measures of abiotic condition in restricted flow (RF) and run-of-river (ROR) sites

Variable	ROR sites range (median) ^a	RF sites range (median)
Physical habitat^b		
Mean thalweg depth (m)	0.7–3.7 (1.6)	4.7–9.9 (7.1)
Mean wetted width (m)	28.0 – 149.1 (78.8)	64.3–104.4 (82.7)
Bankfull height (m)	1.0–4.4 (1.5)	1.1–6.0 (3.8)
Riparian human disturbance – non-agriculture	0–118.0 (30.7)	1.5–66.2 (17.0)
Mean areal fish cover – natural types	4.5–148.4 (56.1)	31.1–161.4 (107.7)
Canopy density at bank (%)	0–90.5 (59.1)	20.7–95.6 (64.6)
Substrate as gravel (%)	0–100 (66.3)	0–0 (0)
LWD quantity	1–95 (19)	1–155 (18)
Total suspended solids (TSS, mg/L)	8–48 (23)	1–22 (9.5)
Agriculture in riparian corridor (%)	8.2–87.9 (62.1)	9.7–79.0 (33.6)
Forest in riparian corridor (%)	6.8–57.5 (17.3)	9.4–85.8 (65.5)
Water chemistry		
Mean conductivity (uS/cm)	152 – 1176 (770)	337–786 (525)
SO ₄ (mg/L)	9.1–134.5 (89.6)	13.9–272.5 (83.0)
NO ₃ (mg/L)	0.1–3.6 (1.3)	0.10–0.15 (0.10)
Chloride (mg/L)	3.5–65.6 (41.3)	4.3–36.4 (11.7)
Ammonia (mg/L)	0.01–0.10 (0.03)	0.02–0.32 (0.06)
Total Khejldahl nitrogen (TKN) (mg/L)	0.35–1.57 (0.58)	0.18–1.84 (0.68)
Total phosphorus (ug/L)	0.045–0.706 (0.248)	0.016–0.164 (0.075)

^aVariables with the range not bolded were excluded from analyses due to limited range or redundancy with another variable.

^bRiparian human disturbance variables are based on an ordinal system of rating individual transects that evaluates the proximity of disturbance to the river. Areal fish cover variables are based on ordinal categories of areal percentage cover of several types of fish cover, resulting in possible totals greater than 100.

identified pairs of redundant metrics using Pearson correlations (>0.80) and visual examination of scatter plots and selected one metric from such pairs based on the relative range, ease of calculation, and correlations with other metrics. If samples for a particular method were missing for a particular site, we excluded the entire site from this analysis. For this and subsequent analyses, separate analyses were run for each impoundment class.

First, we directly compared metric values among sampling methods. For each metric, we used a nonparametric repeated measures ANOVA, the Friedman test, and its associated multiple comparison procedures (Hollander and Wolfe, 1999) at a Type I error rate of 0.05. This procedure tested for differences among methods within sites and the multiple comparisons identified which specific methods

TABLE IV
Metrics evaluated for each sampling method (X), excluding those with limited ranges or redundant with other metrics

Metric	ROR sites					RF sites				
	KN	QMH500	HD	QMH1000	RTH	KN	QMH500	HD	QMH1000	RTH
Richness and diversity										
Number of taxa	X	X	X	X	X	X	X	X	X	X
Number of Ephemeroptera taxa										
Number of Plecoptera taxa	X	X								
Number of Trichoptera taxa			X							
Number of Ephemeroptera, Plecoptera, Trichoptera (EPT) taxa			X	X						
Number of Ephemeroptera, Trichoptera, and Odonata (ETO) taxa	X	X	X	X	X	X	X	X	X	X
Number of Odonata taxa				X	X					
Number of Chironomidae taxa	X	X		X	X				X	
Number of Hemiptera taxa	X	X		X	X				X	
Number of Coleoptera taxa	X	X	X	X	X				X	X
Number of Mollusca + Crustacea taxa	X	X	X	X	X	X	X	X	X	X
Shannon diversity										
Composition and evenness										
No. individuals per taxon	X	X	X	X	X			X		X
Non-insects (%)	X	X	X	X	X			X		
Oligochaetes and leeches (%)	X	X	X	X	X					
Ephemeroptera individuals (%)	X	X	X	X	X				X	
Plecoptera individuals (%)										
Trichoptera individuals (%)								X		
Chironomidae individuals (%)	X	X	X	X	X	X	X	X	X	X

(Continued on next page)

differed. Within each impoundment class, we examined only metrics that were selected for analysis for all of the sampling methods.

Next, we evaluated relationships of metric values with the physical and chemical condition of each site for each method in two different ways. First, we ran a Principal Components Analysis (PCA) on the physical habitat and water chemistry variables (Table IV) to develop axes that represent multivariate abiotic condition gradients. For PCA, eliminating skewness in the contributing variables is typically all that is required (Legendre and Legendre, 1998). Thus, each abiotic variable was transformed by taking either the \log_{10} , square root, or square of the variable to reduce skewness in the data as much as possible. Skewness (S) and the standard error of skewness (SES) were calculated in SYSTAT (v. 10, SPSS, Chicago, Illinois), and a ratio of S to SES of greater than 2 was considered to indicate significant skewness. Variables were transformed and PCAs were run in SYSTAT for each impoundment group. We used the broken stick model to identify the meaningful principal component axes (Legendre and Legendre, 1998). We then performed a varimax rotation in SYSTAT for easier interpretation of the axes. For each sampling method and impoundment group, we calculated Spearman rank correlations between metric values and PCA axis scores. Because this was an exploratory data analysis and we were not interested in testing specific hypotheses, we considered correlation coefficients of 0.5 or higher and -0.5 and lower as "strong correlations."

We also examined Spearman rank correlations of metrics with individual physical habitat and water chemistry variables. We again examined relationships for each sampling method and impoundment type separately. We used the same 0.5 correlation strength as a cutoff.

3. Results

Five Hester-Dendy substrate samplers were not recovered, one each from the Great Miami, Kentucky, and Green rivers, and two from the Scioto River. The five sites missing HD data were excluded from all subsequent analyses. An additional RF site on the Scioto River was excluded from analyses as an anomaly because the reach sampled functioned more as a reservoir than as a river and was much wider (391 m wetted width) than other RF sites. Of the sites remaining, 39 were ROR and 15 were RF sites.

Only 39 samples (37 ROR, 2 RF) were collected using the DN method due to inadequate flow for deployment at many sites. Examination of data from the DN method showed that when drift nets could be deployed, this method usually collected very few organisms relative to all other methods (Figure 2). The mean number of organisms collected with the DN method was only 74 organisms. All other methods collected more than 250 organisms on average. Rather than drop a large number of sites due to missing DN samples, we dropped the DN method from further analyses, including the initial evaluation of metric characteristics, to ensure that more subtle differences among the remaining methods could be detected.

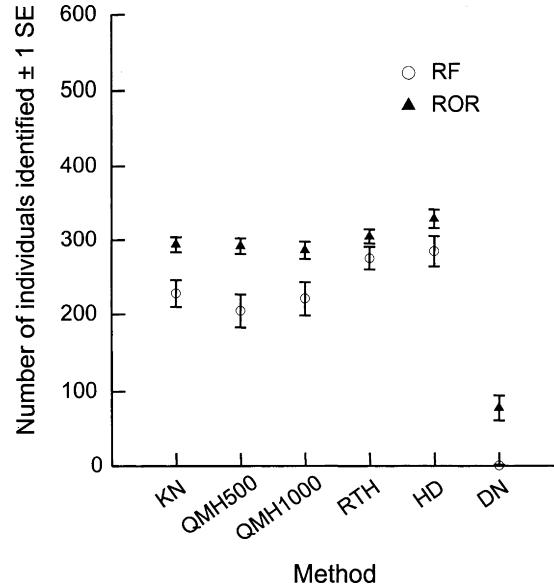


Figure 2. Mean number of individuals identified for each method and site class based on a fixed count of 300 organisms. Those values below 100 represent full counts of samples. Means for DN method represent 2 and 37 samples for restricted flow (RF) and run-of-the-river (ROR) sites, respectively.

At these remaining sites, the ROR and RF classes differed most strongly by mean thalweg depth, with no overlap of ranges between impoundment classes (Table III). In addition, there was little variation in substrate type among RF sites, with all sites dominated by sand or fines. Although the overall ranges of many abiotic variables were similar between ROR and RF sites, the medians were sometimes very different. For example, there was generally more forested riparian corridor among RF sites and more agriculture in the riparian among ROR sites. In addition, the distributions were somewhat complementary for total phosphorus and chloride concentrations between the two impoundment classes, in that ROR sites tended to have values at one end of the gradient and RF sites at the other end. Thus, the gradients of abiotic conditions showed some variation between impoundment classes.

The NMS ordination of genus-level composition data showed groupings of sites by impoundment status and sampling method (Figure 3). The three-axis solution had a final stress value of 14.90, and the three axes explained a total of 82.7% of the variation in the data. Both the second and third axes showed differences between ROR and RF sites. On the first and second axes, the HD and RTH samples were separated from other methods in ROR sites. These two sampling methods were also grouped away from other methods for RF sites on the second axis.

Within each impoundment type, the metrics evaluated varied among sampling methods. This difference in metrics evaluated is due to limited ranges of some metrics only for particular methods. In addition, certain metrics were redundant only for certain methods and not for others. Thus, the metrics eliminated from

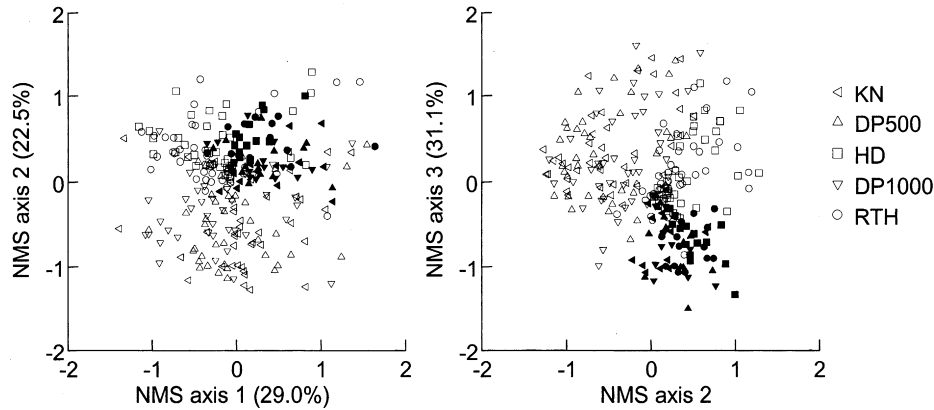


Figure 3. Nonmetric multidimensional scaling (NMS) ordination axes based on genus-level macroinvertebrate data from all sites and methods. The percentage of variance explained by each axis is provided in parentheses. Filled symbols represent restricted flow (RF) sites, and open symbols represent run-of-the-river (ROR) sites.

consideration were dependent on the method used to collect the data. Across the five methods, from 21 to 28 metrics for ROR sites and from 14 to 20 metrics for RF sites were further evaluated. There were 10 metrics retained across all methods for RF sites and 13 in ROR sites.

Among ROR sites, metric values often differed among sampling methods (Table V). Typically, the KN, QMH500, and QMH1000 methods resulted in similar metric values. The HD method metrics differed significantly from those of other methods most frequently, although the RTH method also differed from the KN, QMH500, and QMH1000 methods in some metrics tested. Typically, the HD method collected few taxa and higher percentages of chironomids, but the KN, QMH500, and QMH1000 methods collected more tolerant taxa than the HD and RTH methods.

The variables included in PCAs on water chemistry and physical habitat differed between ROR and RF sites, although the two types of sites shared some similar abiotic gradients. For ROR sites, sulfate was excluded because of a very high correlation with conductivity, reducing the number of variables in the PCA to 18, and the number of sites was limited to 37 because two ROR sites were missing TSS data. Four axes were meaningful for interpretation, with the total variance explained approximately 63% (Table VI). The first axis reflected higher canopy densities and more large woody debris at one end and more non-agricultural human disturbance and greater depths and widths at the other. The second axis reflected a nutrient gradient (nitrate and total phosphorus), and the third axis was most related to chloride and TKN at one end and more gravel substrate and ammonia at the other. The fourth axis was most correlated with bankfull height, natural fish cover, and TSS, reflecting steeper, more undercut banks.

TABLE V

Results of Friedman tests and multiple comparison group membership (values for A>B>C) for macroinvertebrate metrics in ROR ($N = 39$) and RF ($N = 15$) sites

Metric	Multiple comparison groups (overall $\alpha = 0.05$)						
	S'	p -value	KN	DP500	DP1000	RTH	HD
ROR sites							
Number of taxa	77.8	<0.0001	A	A	B	A	B
Number of Coleoptera taxa	52.5	<0.0001	A	AB	C	AB	B
Number of Mollusca + Crustacea taxa	40.9	<0.0001	A	AC	B	AC	BC
Non-insects (%)	26.3	<0.0001	A	BC	C	AB	BC
Chironomids (%)	30.6	<0.0001	B	B	A	AB	A
Number of tolerant taxa	99.2	<0.0001	A	A	B	A	B
Taxa as tolerant (%)	57.2	<0.0001	A	A	B	A	B
Tolerant individuals (%)	6.8	0.1474					
Collector-filterer individuals (%)	13.8	0.0079	B	B	AB	B	A
Collector-gatherers individuals (%)	12.8	0.0122	AB	AB	AB	A	B
Number of predator taxa	114.7	<0.0001	A	A	B	A	B
Number of scraper taxa	29.6	<0.0001	A	AB	A	A	B
Scraper individuals (%)	15.9	0.0032	A	AB	B	AB	A
RF sites							
Number of taxa	28.0	<0.0001	A	AB	C	AB	B
Number of ETO taxa	3.1	0.5370					
Taxa in EPT (%)	7.0	0.1341					
Taxa in Chironomidae (%)	13.1	0.0109	B	AB	A	AB	AB
Dominant taxon (%)	12.1	0.0170	B	B	AB	B	A
Taxa as tolerant (%)	3.3	0.5065					
Tolerant individuals (%)	4.1	0.3918					
Collector-gatherer individuals (%)	0.9	0.9312					
Predator individuals (%)	8.1	0.0896					
Scraper individuals (%)	3.4	0.4874					

Among ROR sites, there were few strong correlations ($|r| > 0.5$) with either multivariate PCA axes or individual abiotic variables, regardless of sampling method (Table VII). The KN and QMH methods were correlated with three of four PCA axes, but no method was correlated with the majority of water chemistry variables. The HD method produced metrics that were correlated with the largest number of physical habitat variables, followed by the KN and QMH500 methods. Overall, the HD and QMH500 methods were correlated with the largest numbers of abiotic variables. The third PCA axis was not correlated with any metrics, and there also were no correlations with the individual variables most correlated with this axis (percentage gravel substrate, chloride concentration, ammonia concentration,

TABLE VI
Correlations of water chemistry and physical habitat variables with PCA axes for ROR and RF sites

Variable	ROR sites (N = 37)				RF sites (N = 15)		
	Axis 1 ($\lambda = 3.6$, 20.1%) ^a	Axis 2 ($\lambda = 2.5$, 14.0%)	Axis 3 ($\lambda = 3.0$, 16.4%)	Axis 4 ($\lambda = 2.3$, 12.6%)	Axis 1 ($\lambda = 3.5$, 25.3%) ^a	Axis 2 ($\lambda = 3.0$, 21.6%)	Axis 3 ($\lambda = 3.3$, 23.9%)
Mean thalweg depth (m)	-0.680	-0.190	0.283	0.421	0.509	0.537	0.095
Mean wetted width (m)	-0.856	0.198	0.020	0.160	0.864	-0.188	0.130
Bankfull height (m)	-0.023	-0.096	-0.025	0.725	0.301	-0.769	0.170
Riparian human disturbance – non-agriculture	-0.730	0.047	0.047	-0.191	0.038	0.015	-0.799
Mean areal fish cover – natural types	0.238	-0.076	-0.182	0.682	-	-	-
Canopy density at bank (%)	0.759	-0.024	-0.125	0.332	-0.906	0.048	-0.196
Substrate as gravel (%)	0.049	0.002	-0.646	0.008	-	-	-
LWD quantity	0.674	-0.124	0.501	0.053	-0.745	0.493	-0.187
TSS	0.077	0.068	0.044	0.628	0.281	-0.040	0.797
Agriculture in riparian corridor (%)	0.591	0.372	-0.370	0.276	0.417	0.142	0.617
Forest in riparian corridor (%)	0.324	-0.302	0.551	0.123	-	-	-
Mean water temperature	-0.287	0.030	0.384	-0.253	0.808	0.288	0.101
Mean conductivity	-0.157	0.531	0.245	0.568	0.030	0.153	-0.857
NO ₃	-0.031	0.796	0.127	-0.092	-	-	-
Chloride	-0.254	0.552	0.702	0.155	-0.159	-0.764	0.085
Ammonia	0.405	0.229	-0.625	-0.025	0.215	-0.567	0.398
TKN	0.020	0.234	0.761	-0.106	0.302	-0.208	0.803
Total phosphorus	0.028	0.925	-0.177	-0.034	-0.013	-0.890	-0.084

^aEigenvalue, percentage of variance explained.

TABLE VII
Multivariate or individual abiotic variables with Spearman rank correlations of 0.5 or stronger with at least one metric for a given method for ROR and RF sites

Variable	ROR sites (N = 39 sites)					RF sites (N = 15 sites)				
	KN	QMH 500	HD	QMH 1000	RTH	KN	QMH 500	HD	QMH 1000	RTH
PCA axis 1 ^a	X	X	X	X	X	X	X	X	X	X
PCA axis 2 ^a		X				X	X		X	X
PCA axis 3 ^a								X	X	X
PCA axis 4 ^a	X	X		X		^b	-	-	-	-
Mean thalweg depth (m)	X	X	X	X		X	X	X	X	X
Mean wetted width (m)	X	X	X	X	X	X	X	X	X	X
Bankfull height (m)		X			X	X	X	X	X	X
Riparian human disturbance – non-agriculture			X		X	X	X	X	X	X
Mean areal fish cover – natural types	X		X			^b	-	-	-	-
Canopy density at bank (%)	X		X			X	X	X	X	X
Substrate as gravel (%)						^b	-	-	-	-
LWD quantity					X	X	X	X	X	X
TSS ^a		X	X			X	X	X	X	X
Agriculture in riparian corridor (%)						X	X	X	X	X
Forest in riparian corridor (%)						^b	-	-	-	-
Mean water temperature						X	X	X	X	X
Mean conductivity	X	X	X			X	X	X	X	X
NO ₃		X				^b	-	-	-	-
Chloride			X			X	X	X	X	X
Ammonia						X	X	X	X	X
TKN						X	X	X	X	X
Total phosphorus		X	X			X	X	X	X	X

^aFor ROR sites, N = 37 for correlations.

^bNot analyzed.

and TKN). If the metrics for a particular method were associated with a certain multivariate gradient (i.e., PCA axis), there were also generally strong associations between the metrics and most of the important individual component variables of that multivariate gradient.

Comparison of metric values for RF sites revealed few consistent differences among sampling methods (Table V). The KN method tended to collect the most taxa, and the HD method collected the fewest. The KN method tended to differ the most from the HD and RTH methods when methods differed at all in metric values.

The PCA for RF sites was based on 14 variables, excluding NO_3 due to limited range, percentage gravel substrate due to a lack of variation, and percentage forest in the riparian corridor, natural fish cover, and SO_4 because of strong correlations with other variables. Three axes were meaningful for interpretation based on the broken stick model, and the varimax rotation was applied to these three (Table VI). The first axis was strongly related to canopy density and large woody debris quantity at one end and wider, deeper, warmer conditions at the other end. The second axis reflected a gradient of total phosphorus, chloride, and ammonia concentrations, along with bankfull height. The third axis was most related to conductivity and non-agricultural human disturbance at one end and TKN, TSS, and agriculture at the other end of the gradient.

In RF sites, sampling methods varied in the metrics showing strong correlations ($|r| > 0.5$) with multivariate and individual abiotic variables (Table VII). Only the QMH1000 and RTH methods showed correlations with all three PCA axes. Most methods performed similarly with respect to water chemistry variables, related to four or five of the individual variables. The KN and HD methods were related to the most physical habitat variables, although only the HD method had metrics associated with TSS concentrations.

4. Discussion

Not surprisingly, the DN method was not effective using the current protocol of setting the nets for 4 h during full daylight. There are several factors that might have contributed to the low effectiveness of this method in this study. Studies have shown that drift is higher for many species at night (e.g., Brittain and Eikeland, 1988; Cellot, 1989; Sagar and Glova, 1992). Thus, the timing of drift net deployment probably contributed greatly to the low numbers or lack of individuals collected, and use of drift nets may work more effectively if deployed during nighttime hours at sites with sufficient flow. However, drift nets may be more problematic in larger systems in general because drift may not reflect local conditions well (Koetsier and Bryan, 1996).

Based on the NMS axes, there were clear differences in taxonomic composition among samples collected using certain sampling methods and from the two different impoundment classes. These results support the separation of the two types of sites for data analysis. Such differences are likely due largely to the large differences in

physical characteristics such as mean thalweg depth and substrate types, although other differences in abiotic gradients also may have played a role in these taxonomic differences. The differences among methods in the NMS plots were evident in examining metric data as well, reinforcing that the assemblages sampled differed among methods. In ROR sites, the RTH and HD methods differed most from other methods for both the NMS plots and the metric comparison. On the other hand, the RF samples were generally all grouped together in the NMS plots, and these same samples showed little consistent variation in metrics among sampling methods.

Although we made an effort to compare existing field sampling methods, with some adaptations, the complete method used by an agency was not always followed. In some cases, the laboratory method associated with a particular field method differed significantly from that used in processing samples for this study. Therefore, the use of the results from this study may not be as originally intended by the agency that developed a particular field sampling method. In fact, the methods on which the QMH500 and QMH1000 were based are used by the Ohio EPA (1987) and NAWQA (Moulton *et al.*, 2002) programs, respectively, only to measure taxa richness. Thus, the processing of these samples was performed with the intention of recording the presence of taxa in a sample and not necessarily relative abundances of taxa. Even so, we were interested in a comparison of field methodologies only, and by processing samples differently for different field methods, we would have been unable to distinguish differences due to the field method from those due to laboratory methods.

At ROR sites, no sampling method was very effective at detecting all three types of abiotic gradients (i.e., multivariate, physical habitat, and water chemistry). The HD method was correlated with the largest number of physical habitat gradients. Because the HD sampler was limited to the lower 500 m of the sampling reach, and the habitat variables were measured and summarized across the entire reach, the macroinvertebrate assemblage on the HD sampler may have represented the integration of upstream habitat conditions. However, water chemistry samples were collected from a single transect in the reach, making it less likely that the conditions experienced by the macroinvertebrates colonizing the HD sampler were similar. In fact, although the QMH500 method was correlated with the largest number of individual water chemistry variables, only three of six variables were represented. This might reflect the general difficulty in comparing macroinvertebrate samples collected from a variety of sampling points and water chemistry samples integrated across a single transect.

Surprisingly, all five methods performed similarly and collected more similar macroinvertebrate assemblages at RF sites. Perhaps this is primarily due to a greater mean depth at these sites, which limits the overall accessible sampling area in a reach. A limited sampling zone may result in greater habitat homogeneity and fewer available habitats that can be sampled for the QMH500, QMH1000, and RTH methods, resulting in more similar data. The placement of HD samplers in these reaches was also in a more limited zone, which may explain why this method

was not as dissimilar to other methods as it was for ROR sites. The tendency for sedimentation on these samplers may also account for the association between HD method metrics and TSS in slow-flowing RF sites, whereas other methods failed to detect this abiotic gradient.

Superficial similarities among certain methods do not indicate that data from these methods can be pooled into a single set of data. Metric values may have varied little among some methods at ROR sites and among all methods for RF sites, but relationships between macroinvertebrate data and abiotic conditions varied among methods. Even though the KN and QMH500 methods shared similar metric values among ROR sites, correlations of metrics with abiotic gradients were not consistent between the two sampling methods. These same issues were present for RF sites, though to a lesser degree. Even for those methods that showed correlations with the same abiotic variables, the metrics associated with those abiotic variables may vary widely among sampling methods. Thus, combining data indiscriminately from different sampling methods is likely inappropriate, as the data from different methods may provide us with different information about abiotic conditions in an area.

In this study, the HD method was relatively effective in representing abiotic conditions in both ROR and RF sites. In some types of systems, the HD method may provide one of the only viable sampling protocols, such as where substrates are too hard or the river is too deep for kick nets to be useful (see review by Rosenberg and Resh, 1982). Battagazzore *et al.* (1994) concluded that artificial substrates performed best in the larger rivers sampled in that study, and in our study, the HD method was among the methods performing consistently well across impoundment groups. However, a review of artificial substrates by Rosenberg and Resh (1982) identifies several disadvantages to using this sampling method. Artificial substrates have a higher likelihood of being lost due to high flow or vandalism, and a higher risk of drying if water levels drop. In addition, the requirement of at least two visits to a site for deployment and retrieval of the device may limit the appeal of this sampling method.

Considering the requirements associated with each sampling method and the results of this study, a combination of sampling methods may provide the most useful macroinvertebrate data, particularly in ROR sites. For example, the Ohio EPA uses a combination of a qualitative, active sampling method (similar to the QMH500 method) and a quantitative, passive method (substrate sampler) for calculation of the Invertebrate Community Index (ICI) (Ohio EPA, 1987). For reasons described, the use of artificial substrates (HD sampler) can be a drain on resources because multiple site visits are required. We suggest using a method that combines systematic, semi-quantitative characteristics of the KN method and the more qualitative QMH500 method. The KN method provides some insurance that macroinvertebrate habitats unrecognized by the human eye will be sampled, and the QMH500 method ensures that additional taxa obviously missed by the placement of the kick net are still collected.

Even with improved, more reliable sampling methods, bioassessment at RF sites is more problematic than at ROR sites. Relatively unaltered run-of-the-river sites can be identified using habitat, chemistry, and land cover data. However, defining a reference condition for sites with restricted flow is very difficult because comparing these sites with sites lacking lock-and-dam structures often sets an unrealistic expectation or goal. In addition, identifying otherwise unaltered RF sites is also challenging because sites often have restricted flow for the express purpose of promoting human activities in the watershed. Therefore, the identification of appropriate reference conditions for RF sites should be a focus for research related to bioassessment of nonwadeable rivers and streams.

Acknowledgements

We thank Donald Klemm (USEPA, NERL), Marc Smith (Ohio EPA), and Marty Gurtz (USGS) for their input on the proposed sampling protocols for macroinvertebrates. Thanks to Shawn Mills (SoBran, Inc.) for delineating study watersheds and providing watershed land cover data. The United States Environmental Protection Agency through its Office of Research and Development and the Regional Method Initiative funded the research described here. Support for field sampling and lab processing was provided by SoBran, Inc. under contract number 68-C6-0019. This paper has been subjected to Agency review and approved for publication. The views expressed in this paper are those of the authors and do not necessarily reflect the views and policies of the U.S. Environmental Protection Agency. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

References

- Battegazzore, M., Guzzini, A. and Pagnotta, R.: 1994, 'Compared use of macroinvertebrate samplers for the evaluation of water quality in rivers of different sizes', *Limnologica* **24**, 43–49.
- Brittain, J.E. and Eikeland, T.J.: 1988, 'Invertebrate drift – a review', *Hydrobiol.* **166**, 77–93.
- Cellot, B.: 1989, 'Macroinvertebrate movements in a large European river', *Freshw. Biol.* **22**, 45–55.
- Cuffney, T., Gurtz, M.E. and Meader, M.R.: 1993, *Methods for Collecting Benthic Invertebrate Samples as Part of the National Water Quality Assessment Program*, U.S. Geological Survey Open File Report 93-406, US Geological Survey, Raleigh, North Carolina.
- Flotemersch, J.E., Autrey, B.C. and Cormier, S.M.: 2001, *Comparisons of Boating and Wading Methods Used to Assess the Status of Flowing Waters*, EPA 600/R-00/108, U.S. Environmental Protection Agency, Cincinnati, Ohio.
- Hollander, M. and Wolfe, D.A.: 1999, *Nonparametric Statistical Methods*, 2nd ed., John Wiley and Sons, New York.
- KDEP: 1976, *The River Basin Water Quality Management Plan for Kentucky: Vol. 3, Green River Basin, 303(e) Report*, Kentucky Department of Environmental Protection, Commonwealth of Kentucky, Frankfort, Kentucky.
- 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.

- Klemm, D.J., Lewis, P.A., Fulk, F.A. and Lazorchak, J.M.: 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.
- Koetsier, P. and Bryan, C.F.: 1996, 'Is macroinvertebrate drift a density-dependent mechanism of the Benthos in the lower Mississippi River?', *J. Freshw. Ecol.* **11**, 1-10.
- Lazorchak, J.M., Hill, B.H., Averill, D.K., Peck, D.V. and Klemm, D.J.: 2000, *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, U.S. Environmental Protection Agency, Cincinnati, Ohio.
- Legendre, P. and Legendre, L.: 1998, *Numerical Ecology*, 2nd ed., Elsevier Science, The Netherlands.
- Minchin, P.R.: 1987, 'An evaluation of the relative robustness of techniques for ecological ordination', *Vegetatio* **69**, 89-107.
- Moulton II, S.R., Kennen, J.G., Goldstein, R.M. and Hambrook, J.A.: 2002, Revised Protocols for Sampling Algal, Invertebrate, and Fish Communities as Part of the National Water-Quality Assessment Program, *U.S. Geological Survey Open-File Report 02-150*, U.S. Geological Survey, Reston, Virginia.
- Ohio EPA: 1987, *Biological Criteria for the Protection of Aquatic Life: Vol. III. Standardized Biological Field Sampling and Laboratory Methods for Assessing Fish and Macroinvertebrate Communities*, Division of Water Quality Monitoring and Assessment, Ohio Environmental Protection Agency, Columbus, Ohio.
- Ohio EPA: 1997a, *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: 1997b, *Biological and Water Quality Study of the Upper Scioto River Basin*, Division of Surface Water, Ecological Assessment Unit, Ohio Environmental Protection Agency, Columbus, Ohio.
- Rosenberg, D.M. and Resh, V.H.: 1982, 'The Use of Artificial Substrates in the Study of Freshwater Benthic Macroinvertebrates', in: J. Cairns, Jr. (ed), *Artificial Substrates*, Ann Arbor Science Publishers, Ann Arbor, Michigan, pp. 175-235.
- Sagar, P.M. and Glova, G.J.: 1992, 'Invertebrate drift in a large, braided New Zealand River', *Freshw. Biol.* **27**, 405-416.
- Tong, S.T.Y.: 1992, 'The use of non-metric multidimensional scaling as an ordination technique in resource survey and evaluation: A case study from southeast Spain', *Appl. Geogr.* **12**, 243-260.
- Vannote, R.L., Minshall, G.W., Cummins, K.W., Sedell, J.R. and Cushing, C.E.: 1980. 'The river continuum concept', *Can. J. Fish. Aquat. Sci.* **37**, 130-137.
- Vogelmann, J.E., Sohl, T.L., Campbell, P.V. and Shaw, D.M.: 1998, 'Regional land cover characterization using landsat thematic mapper data and ancillary data sources', *Environ. Monit. Assess.* **51**, 415-428.