

## BRIDGES

*BRIDGES is a recurring feature of J-NABS intended to provide a forum for the interchange of ideas and information between basic and applied researchers in benthic science. Articles in this series will focus on topical research areas and linkages between basic and applied aspects of research, monitoring, policy, and education. Readers with ideas for topics should contact Associate Editors Nick Aumen and Ashley Moerke.*

Nick Aumen, nick\_aumen@nps.gov  
Ashley Moerke, amoerke@lssu.edu  
Co-editors

### Effect of sampling protocol and volunteer bias when sampling for macroinvertebrates

Julia Frost Nerbonne<sup>1,6</sup>, Brad Ward<sup>2,7</sup>, Ann Ollila<sup>3,8</sup>, Mary Williams<sup>4,9</sup>, AND  
Bruce Vondracek<sup>5,10</sup>

<sup>1</sup>Department of Fisheries, Wildlife, and Conservation Biology, University of Minnesota, 1980 Folwell Ave., St. Paul, Minnesota 55108 USA

<sup>2</sup>Forest Lake Area High School, 6101 Scandia Trail N., Forest Lake, Minnesota 55025 USA

<sup>3</sup>Natural Resources and Environmental Studies, College of Natural Resources, University of Minnesota, St. Paul, Minnesota 55108 USA

<sup>4</sup>Conservation Biology Program, University of Minnesota, 1980 Folwell Ave., St. Paul, Minnesota 55108 USA

<sup>5</sup>Department of Fisheries, Wildlife, and Conservation Biology, University of Minnesota, and <sup>11</sup>US Geological Survey, Minnesota Cooperative Fish and Wildlife Research Unit, University of Minnesota, 1980 Folwell Ave., St. Paul, Minnesota 55108 USA

**Abstract.** We evaluated the efficacy of different field sampling approaches for volunteers sampling macroinvertebrates in low-gradient streams. We used a series of analytical metrics to compare results using the Environmental Protection Agency (EPA) multihabitat, the Minnesota Pollution Control Agency multihabitat, and EPA single-habitat sampling protocols. We also investigated the effect of 2 scenarios in which volunteers fail to follow (and potentially bias) the widely used EPA multihabitat protocol by including either more snag and vegetated banks or more run and riffle habitat than prescribed by the protocol. We collected jab samples from cobble, snags, vegetated banks, submerged macrophytes, and sand in 4 contiguous 125-m reaches in an Anoka sand-plain stream in Minnesota. We identified up to 100 macroinvertebrates in each jab sample to family. We subjected a parent population of 40 jab samples/reach to a bootstrap analysis to sample and create metric or index scores 100 times without replacement for each of the 3 volunteer sampling methods and 2 biased scenarios. The EPA multihabitat protocol and the biased scenario in which woody debris and bank vegetation were oversampled yielded the highest diversity of organisms, whereas the biased scenario in which riffle and run habitats were oversampled yielded the lowest diversity. The EPA multihabitat protocol used correctly was more likely to indicate “good” water quality (on the basis of the EPA muddy-bottom narrative assessment tool designed for volunteers) than either biased sampling scenario. This result illustrates that poor field methods could result in underestimation of water quality.

**Key words:** macroinvertebrate monitoring, volunteer monitoring, citizen science, multihabitat sampling, bias.

<sup>6</sup> E-mail addresses: nerbonne@umn.edu

<sup>7</sup> bward@forestlake.k12.mn.us

<sup>8</sup> olli0011@umn.edu

<sup>9</sup> will1425@umn.edu

<sup>10</sup> vondr001@umn.edu

<sup>11</sup> The Minnesota Cooperative Fish and Wildlife Research Unit is jointly sponsored by the US Geological Survey, the University of Minnesota, the Minnesota Department of Natural Resources, the US Fish and Wildlife Service, and the Wildlife Management Institute.

An increasing number of volunteers around the country have been engaged in monitoring aquatic health (Ripley 1993, Lathrop and Markowitz 1995, Penrose and Call 1995, DiStefano 1999, Nerbonne and Nelson 2004). Volunteers are encouraged and contribute their data to local and state databases to track long-term trends in water quality (Heiskary et al. 1994, Mattson et al. 1994, USEPA 1996, Fore et al. 2001, Boylen et al. 2004). In particular, stream macroinvertebrate monitoring has become popular because biological monitoring offers the opportunity to understand the integrated effects of landuse changes on aquatic health across watersheds (Karr and Chu 1999) and can be an engaging activity that teaches local residents about effects of anthropogenic pollution on aquatic communities (USEPA 1997, Kudelka and Dates 2003, Nerbonne and Nelson 2004). For example, school groups in Hennepin County, Minnesota have been collecting stream macroinvertebrates since 1995 and have contributed to a countywide database that illustrates significant impacts of development on stream health (Fortin 1996). Interest in volunteer stream monitoring has grown since the first programs in Minnesota began during the mid-1990s, and in 2002, the Minnesota legislature passed a law that instructed the Minnesota Pollution Control Agency (MPCA) to support and encourage citizen monitoring (MPCA 2003).

A variety of protocols have been developed for conducting macroinvertebrate monitoring in the US. The simplest and most popular approach is to collect samples from a single habitat (usually cobble in riffle areas) with a kick or D-frame net (Hilsenhoff 1977, Plafkin et al. 1989). Although this approach works well in high-gradient streams, cobble often is not present in lower-gradient or muddy-bottom streams. When it is, sampling only cobble substrate might not result in complete representation of the macroinvertebrate community (Barbour et al. 1999, DiStefano 1999, MPCA 2004). Several different multihabitat methods have been proposed in response to the need for diversification of rapid bioassessment protocols. These new methods instruct practitioners to sample alternative habitats such as woody debris, leaf packs, and vegetated banks. They are more complex than single-habitat sampling and often demand that field technicians or volunteers be trained to identify different habitats and to collect samples from each in proportion to the availability of each habitat type. For example, the Environmental Protection Agency (EPA) multihabitat protocol instructs users to estimate the percentage of each habitat available and then to allocate sampling effort in rough proportion to the occurrence of each available habitat in the sampling reach

(Barbour et al. 1999). In contrast, the MPCA instructs users to identify the available habitats and sample equally from each (MPCA 2004).

Many factors can influence the reliability of data collected by volunteers, but little has been published to help practitioners assess potential sources of bias. For example, in one study, data collected by volunteer high school students were more variable than those collected by professional biologists, but data collected by the 2 groups indicated similar trends in water quality (Reynoldson et al. 1986). In another study, volunteers with more training were able to produce data that more closely resembled data from fisheries biologists than volunteers with less training (DiStefano 1999). Understanding the potential for bias using different methods is critical because volunteer data currently are being considered as a tool for informing local and regional decision makers about water-quality trends (MPCA 2003).

Volunteer organizers in Anoka County, Minnesota, began sampling Hardwood Creek in 1998. Riffles were present, so a single-habitat sampling method was considered despite the fact that the stream system was predominantly sandy and had a low gradient. Although other streams in the metropolitan area were being sampled using a single-habitat approach, leading organizers questioned whether Anoka County should follow this already-established volunteer protocol or use a multihabitat sampling approach. Many considerations arose during the decision-making process: How will the selection of sampling protocol alter the results? Will the multihabitat protocol be too challenging for volunteers? If so, how might volunteer sampling bias affect the results? Our study is a response to those conversations.

Our study was designed to model how protocol choice and volunteer bias might affect monitoring outcomes in low-gradient streams. We begin by comparing a suite of metrics applied to data obtained from 3 different sampling protocols (EPA single-habitat, EPA multihabitat, and MPCA multihabitat) used in a prototypical low-gradient stream in Anoka County. We then assume that the EPA multihabitat protocol is the standard and compare the results of 2 different biased sampling scenarios with results obtained when the EPA multihabitat sampling protocol is followed precisely. We ask: What would happen if volunteers sampled run or riffle substrate more intensely than woody debris and vegetated banks? Alternatively, what would happen if volunteers sampled woody debris and vegetated bank more intensely than run or riffle substrate? What is the variability of data produced by different methods?

Will volunteer sampling biases ultimately lead to significantly different assessments of stream quality?

### Methods

We conducted our study in Hardwood Creek, a 1<sup>st</sup>-order, low-gradient stream in the Anoka Sand Plain in the suburbs of the Twin Cities metropolitan area, which is undergoing rapid development. At present, the creek contains a diverse assemblage of macroinvertebrates ( $\geq 24$  families) including *Pteronarcys* stoneflies, which are found only in the less-affected stream systems in this region. The watershed upstream from our sampling reaches contains 22% agricultural land, 15% forest, 25% wetland, 20% grassland, 8% impervious surface, 5% mowed lawn, and 4% water. Students from an area high school have been sampling in Hardwood Creek since 1998, and a pilot adult-volunteer macroinvertebrate-monitoring program (the Stream Health Evaluation Program) was begun in 2006.

#### *Choosing methods and developing biased scenarios*

Volunteers in the Twin Cities metropolitan area began collecting macroinvertebrates during the mid-1990s (Fortin 1996). At first, all streams in the metropolitan area were sampled with the standard EPA riffle or rocky-bottom protocol. This protocol worked well in the high-gradient streams, but some metropolitan-area counties had streams with predominantly sandy substrate. Organizers of a volunteer program administered by the Anoka Conservation District decided to use the EPA volunteer multihabitat protocol in an effort to sample their low-gradient streams more effectively. The multihabitat sampling method promised to represent the stream systems in the Anoka Sand Plain better than did the single-habitat method, but organizers were concerned that the more-complicated protocol would introduce sources of volunteer bias.

To understand how volunteer bias might affect the outcome of sampling we: 1) conducted interviews with 5 leaders of volunteer monitoring programs in the Twin Cities metropolitan area, and 2) observed students from Forest Lake Area High School as they collected invertebrates in Hardwood Creek from 1999 to 2001. We developed several scenarios that represented realistic ways in which volunteers might assess stream quality with macroinvertebrates.

We used data obtained from field sampling in 2002 (see *Data collection* below) to test 5 sampling scenarios: 1) EPA single-habitat protocol (Barbour et al. 1999): riffles are present and volunteers precisely follow the EPA rocky-bottom protocol by collecting all organisms

from cobble habitat, 2) EPA multihabitat protocol (Barbour et al. 1999): volunteers precisely follow the EPA multihabitat protocol by collecting a total of 20 jabs in proportion to the habitat types present, 3) MPCA protocol (MPCA 2004): volunteers follow the Minnesota Pollution Control multihabitat protocol by collecting an equal (or nearly equal) number of jabs from each habitat type present, regardless of their relative proportion in the stream, 4) *oversample riffle and run* volunteer-biased scenario: volunteers (perhaps fearful of what might be lurking in the vegetation or woody debris) attempt to follow the EPA multihabitat protocol, but sample mostly in the riffle and run areas, and 5) *oversample snag and vegetation* volunteer-biased scenario: volunteers (perhaps excited about the potential for finding crayfish and other larger invertebrates) attempt to follow the EPA multihabitat protocol, but sample mostly in woody debris and vegetated banks and undersample riffle and run areas.

#### *Data collection*

*Field sampling and sample processing.*—We sampled a 500-m stretch of Hardwood Creek during September 2002. We divided the stretch into 4 contiguous 125-m reaches. We mapped instream habitat in each reach according to the EPA multihabitat protocol (Barbour et al. 1999). This protocol instructs practitioners to collect 1 composite sample/reach consisting of the contents of 20 jabs. A *jab* is defined as sampling 1 foot (30.48 cm) of substrate with a 1-foot-wide D-frame net. The number of jabs collected in each particular habitat (cobble, snags, vegetated banks, submerged macrophytes, and sand) is dependent on the relative proportion of that habitat in the stream. However, instead of collecting 1 composite sample/reach, we collected 10 jabs for each habitat represented in each reach (up to 40 jabs in any 1 reach) and processed each jab separately to provide a larger sample for a bootstrap procedure. We randomly sorted and identified up to 100 macroinvertebrates in each jab. To reduce error associated with sorting, the first 10 samples picked by an individual were cross-checked by another member of the research team. We identified macroinvertebrates to family because Fore et al. (2001) and O'Leary et al. (2004) found that overall water-quality ratings on the basis of family-level identification of macroinvertebrates were similar to ratings on the basis of genus-level identification. We also cross-checked identifications of macroinvertebrates.

*Bootstrap sampling.*—We created a parent population (data set) that consisted of the macroinvertebrate data from 10 individual jabs for each major habitat type and reach in our 2002 field samples, and simulated

TABLE 1. Mean (SD) number of taxa, number of Ephemeroptera, Plecoptera, and Trichoptera (EPT) taxa, % dominance, sensitive taxa index (STI), Environmental Protection Agency (EPA) multimetric score (biosurvey), and family-level biotic index (FBI) for each sampling scenario in Hardwood Creek, Anoka County, Minnesota. Values with the same superscript are not significantly different: a, b, and c refer to comparisons among sampling protocols (EPA single-habitat, EPA multimetric, Minnesota Pollution Control Agency [MPCA]); x, y, z refer to comparisons of sampling bias scenarios (EPA multimetric sampled correctly, EPA multimetric biased toward oversampling riffles and runs, EPA multimetric biased toward oversampling snags and bank vegetation). *p*-values for all comparisons < 0.001. —, metric not calculated.

Metric	EPA single habitat	MPCA multihabitat	EPA multihabitat	Oversampling snags and bank vegetation	Oversampling riffle and run habitat
Number of taxa	10.78 (1.92) <sup>a</sup>	19.46 (1.38) <sup>c</sup>	18.95 (1.64) <sup>b,z</sup>	18.12 (1.27) <sup>y</sup>	16.41 (2.23) <sup>x</sup>
Number of EPT taxa	3.74 (0.44) <sup>a</sup>	4.25 (0.43) <sup>b</sup>	4.25 (0.43) <sup>b,x</sup>	4.25 (0.43) <sup>x</sup>	4.16 (0.39) <sup>y</sup>
% dominance	42.7 (14.25) <sup>c</sup>	27.1 (4.00) <sup>b</sup>	24.2 (4.09) <sup>a,x</sup>	25.6 (6.06) <sup>y</sup>	27.5 (7.20) <sup>z</sup>
STI	4.69 (0.39) <sup>a</sup>	5.17 (0.20) <sup>b</sup>	5.21 (0.23) <sup>b,y</sup>	5.34 (0.27) <sup>z</sup>	5.07 (0.30) <sup>x</sup>
Biosurvey	—	16.34 (1.97) <sup>a</sup>	16.16 (1.55) <sup>a,z</sup>	14.96 (1.33) <sup>x</sup>	15.73 (1.76) <sup>y</sup>
FBI	4.69 (0.36) <sup>a</sup>	5.16 (0.16) <sup>b</sup>	5.29 (0.31) <sup>c,y</sup>	5.39 (0.27) <sup>z</sup>	5.16 (0.35) <sup>x</sup>

sampling with this data set to draw inferences about our scenarios. We wrote a bootstrap program in SAS (version 6; SAS Institute, Cary, North Carolina) to sample the data set 100 times without replacement for each scenario/reach combination. We then calculated metric scores for each simulated sample. For the EPA single-habitat protocol, we sampled only 3 times in the cobble habitat, whereas for the MPCA and EPA multihabitat protocols, we sampled equally among habitats or in proportion to the available habitat, respectively. For the oversample snag and bank vegetation scenario, we biased the sampling to include ~2× the number of jabs in woody debris and vegetated banks than found in the baseline EPA scenario, and for the oversample riffle and run scenario, we biased sampling to include ~2× the run and riffle habitat than woody debris and vegetated banks.

*Assessment metrics.*—We evaluated each scenario with the following metrics or composite indices proposed in the EPA *Volunteer Methods Manual* (USEPA 1997): 1) number of taxa (families) represented in a sample; 2) number of Ephemeroptera, Plecoptera, Trichoptera (EPT) taxa in a sample; 3) percentage composition dominance: [(number of individuals in the most abundant family in a sample/total number of individuals in the sample) × 100]; 4) sensitive taxa index (STI): calculated by multiplying the number of organisms in each family by the EPA pollution tolerance value (USEPA 1997) assigned to each taxon, adding these values for all taxa represented in the sample, and dividing by the total number of taxa in the sample (lower scores indicate better water quality); 5) EPA multimetric (biosurvey) score for muddy-bottom streams (USEPA 1997): an aggregate metric based on the number of taxa (families), EPT taxa, percentage dominance, and STI with scores that range from 0 to 24

(higher scores indicate better water quality). Biosurvey scores were assigned a rating of good (>18), fair (9–15), or poor (0–6) according to the muddy-bottom narrative quality-assessment parameters. We did not calculate a biosurvey score for the single-habitat sample because the rocky-bottom biosurvey score is calculated using a different set of criteria; and 6) family-level biotic index (FBI): calculated in a similar fashion to the STI, but uses Hilsenhoff's (1988) tolerance scores, rather than scores published by EPA (USEPA 1997) (lower scores represent better water quality).

*Data analysis.*—We calculated summary statistics for all metrics for each reach/scenario combination from the bootstrap program. We used analysis of variance (ANOVA) with stream reach as a blocking factor to compare values of metrics across scenarios. Significant (*p* < 0.05) ANOVAs were followed by Tukey's post hoc comparisons to identify differences between scenarios.

## Results

### Comparison of protocols

We collected 24 families of macroinvertebrates, of which 5 were EPT. The MPCA and EPA multihabitat protocols yielded higher numbers of taxa and numbers of EPT taxa than did the single-habitat protocol (Table 1). Percentage dominance was highest for the single-habitat protocol and lower for the MPCA and EPA multihabitat protocols. For the composite scores, the single-habitat protocol yielded significantly lower STI and FBI scores than did either multihabitat protocol (Table 1).

### Comparison of sampling bias

The EPA multimetric protocol used correctly and the oversample snags and vegetation scenarios yielded

significantly higher numbers of taxa and EPT taxa than did the oversample riffle and run scenario. The EPA multihabitat protocol used correctly and the oversample snags and bank vegetation scenarios also yielded slightly lower percentage dominance values than did the oversample riffle and run scenario. If only these metrics were used, a volunteer who sampled correctly or oversampled snags and bank vegetation might conclude that water quality was slightly better than would a volunteer who oversampled riffle and run habitat.

In contrast, the oversample riffle and run scenario yielded significantly lower STI and FBI scores than did the EPA multimetric protocol used correctly and the oversample snags and bank vegetation scenarios. If only these metrics were used, a volunteer who oversampled the riffle and run habitat might conclude that water quality was better than would a volunteer who sampled correctly or oversampled snags and bank vegetation. However, the oversample snag and bank vegetation scenario yielded significantly higher STI and FBI scores than did the other 2 scenarios. Thus, a volunteer who oversampled snag and bank vegetation might reach the conclusion that the water quality was lower than would a volunteer who sampled correctly or oversampled riffle and run habitat (Table 1). Last, the EPA multihabitat protocol used correctly was more likely to indicate “good” water quality according to the EPA narrative assessment than either the oversample riffle and run or oversample snag and bank vegetation scenarios (Fig. 1).

**Discussion**

Differentiating rocky-bottom and muddy-bottom streams seems straightforward. However, in the real world, many soft sediment systems also contain riffle areas, complicating the choice of sampling protocols for volunteers. Organizers of volunteer monitoring programs often strive to select a protocol that can be used to compare streams across broad regional areas. Throughout the 1990s, most volunteer stream programs in Minnesota used a single-habitat sampling approach for macroinvertebrates (e.g., Fortin 1996); today, many programs have adopted multihabitat methods. Our simulations illustrate that the choice of protocol and consideration of the expertise of the volunteers conducting the sampling are important when designing a volunteer monitoring program for a low-gradient stream system containing riffles.

Our results illustrate that a multihabitat sampling protocol, such as the EPA muddy-bottom protocol or the MPCA low-gradient protocol, will yield consistently different results than will sampling in riffles

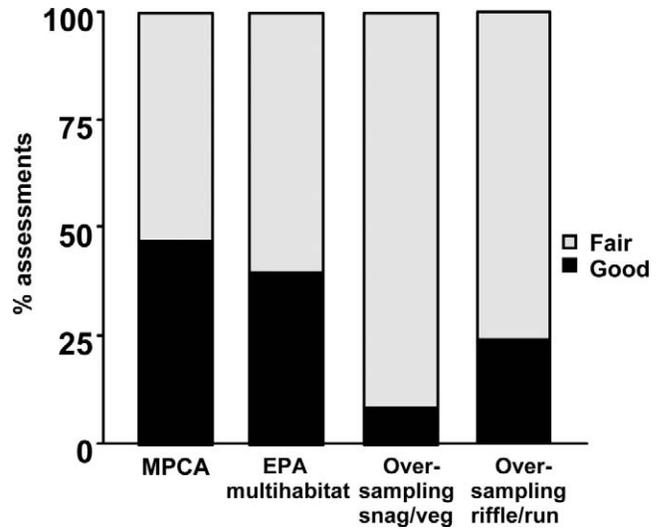


FIG. 1. Normalized percentage of narrative assessments of habitat quality in fair and good quality categories (US Environmental Protection Agency [EPA] *Volunteer Methods Manual*; USEPA 1997) for each multihabitat sampling scenario (Minnesota Pollution Control Agency [MPCA] protocol, EPA multimetric protocol sampled correctly, EPA multimetric protocol biased toward oversampling riffles and runs [oversampling riffle/run], EPA multimetric protocol biased toward oversampling snags and bank vegetation [oversampling snag/veg]).

only. Three of the 5 metrics we used (i.e., those based on richness or diversity) indicated poorer water quality when samples were collected with the single-habitat protocol than when collected with multihabitat protocols. Single-habitat sampling requires volunteers to collect fewer samples (jabs at 3 locations within a riffle) than multihabitat protocols (20 jabs from a variety of habitats). The smaller number of samples used in the single-habitat protocol might yield fewer taxa and higher percentage dominance than the larger number of samples used in the multihabitat protocols. We note with interest that evaluating a stream with either the STI or the FBI might lead to a different conclusion. The STI and FBI, which incorporate pollution tolerance of the taxa collected, indicated better water quality when samples were collected with the single-habitat protocol than with multihabitat protocols. Riffle habitat tends to harbor more sensitive taxa than do other habitats, and the single-habitat protocol calls for samples only from riffles. Thus, a higher proportion of macroinvertebrates with lower pollution tolerance might be collected with single-habitat than with multihabitat protocols.

Volunteer bias is also likely to affect stream assessment. Deviations from the EPA multihabitat protocol in either direction yielded lower numbers of

taxa and EPT taxa than when the EPA protocol was followed precisely. Our results are consistent with results obtained by Johnson et al. (2003), who found that volunteers were significantly more likely to judge that a stream was “fair” rather than “good” (based on narrative standards) when they deviated from the protocol. However, in our study, this trend was not consistent across all metrics. Biasing sampling toward riffle and run habitat yielded higher water-quality scores (on the basis of STI and FBI) than did biasing sampling toward snag and bank vegetation.

Our study illustrates that choice of field sampling protocols and choice of summary metrics might be critical when designing a volunteer monitoring program, especially in areas where riffles exist in low-gradient systems. Metric scores can be affected by the sampling approach. Single-habitat protocols can generate lower estimates of the quality of low-gradient streams on the basis of biodiversity indices (number of taxa or EPT taxa) and higher estimates of stream quality on the basis of metrics that assess average tolerance levels (STI and FBI) than do multihabitat protocols. This discrepancy might be indicative of the fact that sensitive-taxa indices were originally developed for use in high-gradient streams and are not appropriate for assessing low-gradient systems.

We also illustrated that field-sampling bias has the potential to affect the overall results of a monitoring project. Focusing more on riffles than on snags and vegetated banks might lead to lower tolerance ratings, but it can underestimate overall biodiversity. Thus, we recommend that volunteer training programs use a multihabitat approach in low-gradient streams, but we emphasize that programs should focus on field sampling as an important part of training and that multiple metrics should be used to assess overall stream quality.

### Acknowledgements

We thank Matt Kocian for calculating the proportions of land covers in the Hardwood Creek basin and Laurie Sovell and Joel Chirhart for reviewing our manuscript. Support for our project was provided by the Volunteer Stream Monitoring Partnership and the National Science Foundation Undergraduate Research Opportunities Program.

### Literature Cited

- 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. 2<sup>nd</sup> edition. EPA 841-B-99-002. Office of Water, US Environmental Protection Agency, Washington, DC.
- BOYLEN, C. W., E. A. HOWE, J. S. BARTKOWSKI, AND L. W. EICHLER. 2004. Augmentation of a long term monitoring program for Lake George, NY by citizen volunteers. *Lake and Reservoir Management* 20:121–129.
- DI STEFANO, C. M. 1999. An evaluation of a citizen monitoring method of assessing the benthic macroinvertebrate community in selected Missouri streams. MS Thesis, University of Missouri, Columbia, Missouri.
- FORE, L. S., K. PAULSEN, AND K. O'LAUGHLIN. 2001. Assessing the performance of volunteers in monitoring streams. *Freshwater Biology* 46:109–123.
- FORTIN, C. 1996. Macroinvertebrate education and monitoring. Pages 81–82 in E. Ely (editor). Proceedings of the 5<sup>th</sup> National Volunteer Monitoring Conference, University of Wisconsin, Madison, Wisconsin, 3–7 August 1996. EPA 841-R-97-007. Wisconsin Department of Natural Resources, Madison, Wisconsin. (Available from: <http://www.epa.gov/owow/volunteer/proceedings/toc.html>.)
- HEISKARY, S., J. LINDBLOOM, AND C. B. WILSON. 1994. Detecting water quality trends with citizen volunteer data. *Lake and Reservoir Management* 9:4–9.
- HILSENHOFF, W. L. 1977. Use of arthropods to evaluate water quality of streams. Technical Bulletin 100. Wisconsin Department of Natural Resources, Madison, Wisconsin. (Available from: Wisconsin Department of Natural Resources, P.O. Box 7921, Madison, Wisconsin 53707 USA.)
- HILSENHOFF, W. L. 1988. Rapid field assessment of organic pollution with a family-level biotic index. *Journal of the North American Benthological Society* 7:65–68.
- JOHNSON, L. B., D. H. BRENNEMAN, AND C. RICHARDS. 2003. Macroinvertebrate community structure and function associated with large wood in low gradient streams. *River Research and Applications* 19:199–218.
- KARR, J. R., AND E. W. CHU. 1999. Restoring life in running waters. Island Press, Washington, DC.
- KUDELKA, A. B., AND G. DATES. 2003. An evaluation of citizen volunteer water quality monitoring in Minnesota. Rivers Council of Minnesota, Walker, Minnesota. (Available from: Minnesota Waters, 720 West Saint Germain, Suite 143, St. Cloud, Minnesota 56301 USA.)
- LATHROP, J. E., AND A. MARKOWITZ. 1995. Monitoring water resource quality using volunteers. Pages 303–314 in W. S. Davis and T. P. Simon (editors). *Biological assessment and criteria: tools for water resource planning and decision making*. Lewis Publishers, Boca Raton, Florida.
- MATTSON, M. D., M.-F. WALK, P. A. KERR, A. M. SLEPSKI, O. T. ZAJICEK, AND P. J. GODFREY. 1994. Quality assurance testing for a large scale volunteer monitoring program: the Acid Rain Monitoring Project. *Lake and Reservoir Management* 9:10–13.
- MPCA (MINNESOTA POLLUTION CONTROL AGENCY). 2003. 2003 report to the legislature: encouraging citizen monitoring of water quality. Minnesota Pollution Control Agency, St. Paul, Minnesota. (Available from: Minnesota Pollution

- Control Agency, 520 Lafayette Road, St. Paul, Minnesota 55155 USA.)
- MPCA (MINNESOTA POLLUTION CONTROL AGENCY). 2004. Invertebrate sampling procedures. EMAPSOP4, Rev.0. Minnesota Pollution Control Agency, St. Paul, Minnesota. (Available from: Minnesota Pollution Control Agency, 520 Lafayette Road, St. Paul, Minnesota 55155 USA.)
- NERBONNE, J. F., AND K. C. NELSON. 2004. Volunteer macroinvertebrate monitoring in the United States: resource mobilization and comparative state structures. *Society and Natural Resources* 17:817–839.
- O'LEARY, N., A. T. VAWTER, L. P. WAGENET, AND M. PFEFFER. 2004. Assessing water quality using two taxonomic levels of benthic macroinvertebrate analysis: implications for volunteer monitors. *Journal of Freshwater Ecology* 19:581–586.
- PENROSE, D., AND S. M. CALL. 1995. Volunteer monitoring of benthic macroinvertebrates: regulatory biologists perspectives. *Journal of the North American Benthological Society* 14:203–209.
- PLAFKIN, J. L., M. T. BARBOUR, K. D. PORTER, S. K. GROSS, AND R. M. HUGHES. 1989. Rapid bioassessment protocols for use in streams and rivers: benthic macroinvertebrates and fish. EPA/440/4–89/001. US Environmental Protection Agency, Washington, DC.
- REYNOLDS, T., L. HAMPEL, AND J. MARTIN. 1986. Biomonitoring networks operated by school children. *Environmental Pollution (Series A)* 41:363–380.
- RIPLEY, P. 1993. Citizen volunteer monitoring programs of the upper Midwest. *LakeLine* 13:10–13.
- USEPA (US ENVIRONMENTAL PROTECTION AGENCY). 1996. The volunteer monitor's guide to quality assurance project plans. EPA 841-B-96–003. Office of Wetlands, Oceans and Watersheds, US Environmental Protection Agency, Washington, DC.
- USEPA (US ENVIRONMENTAL PROTECTION AGENCY). 1997. Volunteer stream monitoring: a methods manual. EPA 841-B-97–003. US Environmental Protection Agency, Washington, DC.

*Received: 12 August 2007*

*Accepted: 16 May 2008*