Evaluation of an alternate method for sampling benthic macroinvertebrates in low-gradient streams sampled as part of the National Rivers and Streams Assessment

Joseph E. Flotemersch · Sheila North ·
Karen A. Blocksom

Received: 10 April 2013 / Accepted: 6 September 2013 / Published online: 1 October 2013
© The Author(s) 2013. This article is published with open access at Springerlink.com

Abstract Benthic macroinvertebrates are sampled in streams and rivers as one of the assessment elements of the US Environmental Protection Agency's National Rivers and Streams Assessment. In a 2006 report, the recommendation was made that different yet comparable methods be evaluated for different types of streams (e.g., low gradient vs. high gradient). Consequently, a research element was added to the 2008–2009 National Rivers and Streams Assessment to conduct a side-by-side comparison of the standard macroinvertebrate sampling method with an alternate method specifically designed for low-gradient wadeable streams and rivers that focused more on stream edge habitat. Samples were collected using each method at 525 sites in five of nine aggregate ecoregions located in the conterminous USA. Methods were compared using the benthic macroinvertebrate multimetric index developed for the 2006 Wadeable Streams Assessment. Statistical analysis did not reveal

any trends that would suggest the overall assessment of low-gradient streams on a regional or national scale would change if the alternate method was used rather than the standard sampling method, regardless of the gradient cutoff used to define low-gradient streams. Based on these results, the National Rivers and Streams Survey should continue to use the standard field method for sampling all streams.

Keywords Benthic macroinvertebrates · Wadeable Streams Assessment · National Rivers and Streams Assessment · Biological condition

Introduction

States, tribes, federal agencies, and other organizations in the USA collect water quality data to respond to many Clean Water Act (CWA) requirements (USGPO 1989) for protection and monitoring of inland waters. Section 305(b) of the CWA also requires that the US Environmental Protection Agency (USEPA) submit biennial national reports to congress that summarize water quality by states, territories, tribes, and jurisdictions of the USA. While many of these requirements are met using field data, summarizing condition at a national scale has been problematic because reporting entities use many differing monitoring designs, indicators, and methods. Consequently, these data cannot be merged and used to effectively answer questions about the quality of the nation's waters or to track changes over time at a regional or national scale (USEPA 2011).

J. E. Flotemersch (✉)
Office of Research and Development, US Environmental
Protection Agency,
26 W. Martin Luther King Dr., Mail Stop 591, Cincinnati,
OH 45268, USA
e-mail: flotemersch.joseph@epa.gov

S. North
Office of Research and Development, Dynamac
Corporation, c/o U.S. Environmental Protection Agency,
26 W. Martin Luther King Dr., Cincinnati, OH 45268, USA

K. A. Blocksom
Office of Research and Development, US Environmental
Protection Agency,
200 S.W. 35th Street, Corvallis, OR 97333, USA

In response to this problem, the USEPA Office of Water initiated the National Aquatic Resource Surveys (NARS; USEPA 2012a). These probability-based surveys were designed to provide nationally consistent and scientifically defensible assessments of our nation's waters and to track condition over time. These surveys use standardized field and lab methods to yield unbiased estimates of condition for the national water resource type in question (i.e., rivers and streams, lakes, wetlands, or coastal waters; USEPA 2012a).

The first of the NARS surveys was the Wadeable Streams Assessment (WSA; USEPA 2006). This survey examined the biological condition of wadeable streams throughout the USA (USEPA 2006). Between 2000 and 2004, USEPA, states, and tribes collected chemical, physical, and biological data at 1,392 wadeable, perennial stream locations throughout the USA using the same standardized methods (USEPA 2004). As part of the survey, benthic macroinvertebrates were collected as an indicator of biological condition. However, in the study's 2006 report (USEPA 2006), it was recommended that future surveys evaluate the use of different methods for collecting benthic macroinvertebrates in different stream types. This recommendation was based on observations made by USEPA regional scientists who were concerned that in some locations of the nation, the method used for the survey underrepresented the existing benthic macroinvertebrate fauna in some stream types. Low-gradient streams were highlighted as an example of one stream type where alternative methods may be necessary. Beyond under representing the fauna, there was also a concern that the existing method did not collect adequate numbers of organisms in low-gradient streams. This is highly relevant as insufficient numbers can impact assessment endpoint scores or even prohibit their calculation (e.g., Mazor et al. 2010).

Previous studies have variously defined low-gradient streams. Mazor et al. (2010) defined low-gradient as having a stream gradient or slope $\leq 1\%$. This same definition has been used by Virginia Department of Environmental Quality (L. Willis, personal communication). One USEPA study defined low-gradient as having typical velocities of <0.5 fps (0.1524 mps) and lacking riffle habitats (USEPA 1997). These and other studies have further described low-gradient streams as generally having smaller-sized substrates (e.g., sand and silt) and differences in bed morphology and microhabitat distribution (Montgomery and Buffington 1997) when

compared to their higher-gradient counterparts (Smock and Gilinsky 1992; Rinella and Feminella 2005). With little doubt, such differences in habitat would give rise to differences in the structure and function of the biotic communities (Minshall 1984) and, consequently, those available for collection as part of a bioassessment effort. For USEPA surveys sampling streams, the central question then became whether these changes necessitate use of a different field collection method. This question was considered carefully, as developing and maintaining unique assessment tools for multiple habitat types may be prohibitively expensive and make comparisons of results from different stream types or regions problematic (Mazor et al. 2010), a reality that contributed to the creation of the NARS program in the beginning.

In 2008–2009, the USEPA conducted a follow-up survey to the WSA. The design was expanded to include nonwadeable rivers and streams and thus was renamed the National Rivers and Streams Assessment (NRSA; USEPA 2012b). This survey incorporated a research element to evaluate the performance of an alternate method that was specific for sampling benthic macroinvertebrates in low-gradient wadeable streams and rivers. An expert panel was consulted for development of the alternate low-gradient sampling method (LG). Summing up, the goals of this manuscript are to determine the following: (1) is there a difference between the LG and historically used methods for number of organisms collected and, if so, is there a definable gradient (e.g., $\leq 1\%$) that differentiates when to use which method; and (2) is there a difference between IBI scores by method and, if so, what is the slope cutoff to determine which method to use? These questions were considered for the entire study area and for each of the five aggregate ecoregions (defined below) it contained.

Methods

Study site As part of the NRSA survey, benthic macroinvertebrate samples were collected from wadeable streams and rivers across the conterminous USA between 2008 and 2009. The total NRSA study area was divided into nine aggregated Omernik ecoregions (AOEs) to account for regional differences due to geographical phenomena associated with differences in the quality, health, and integrity of ecosystems (Omernik 2004; Olsen 2007; Fig. 1). The nine AOEs are aggregations of smaller level III ecoregions with similar landform and climate characteristics (USEPA 2006). Based

Fig. 1 Sampling locations included in this study were located in five of the nine NRSA aggregate Omernik ecoregions (AOEs)

on review of the 2000–2004 WSA data, five of the nine AOEs were determined to contain high numbers of low-gradient sites and were selected for inclusion in this study. These were the Coastal Plains (level III ecoregions 33–35, 63, 65, 73–76, 84), Northern Plains (NPL; level III ecoregions 42, 43), Southern Plains (SPL; level III ecoregions 25–27, 29–32, 44), Temperate Plains (TPL; level III ecoregions 28, 40, 46–48, 53–55, 57, 72), and the Xeric (XER; level III ecoregions 6, 7, 10, 12–14, 18, 20, 22, 27, 79, 80, 81, 85). A description of each level III ecoregion included in the study is available at http://www.epa.gov/wed/pages/ecoregions/level_iii_iv.htm (accessed 2 Nov 2012).

The NRSA dataset included a total of 1,924 wadeable and nonwadeable sites for potential analysis. Sites within each of the AOEs were selected using a random sampling design to provide regional and national estimates of the condition of rivers and streams (Olsen

2007). Of the wadeable sites in the five AOEs targeted for this study, complete data for use in analyses (total number of organisms collected, benthic multimetric index (MMI) metrics and score, and stream gradient) were available for 78 CPL, 138 NPL, 129 SPL, 103 TPL, and 77 XER sites for a total of 525 sites across all five AOEs (Fig. 1). When a site was visited more than once (i.e., repeat visit), only samples from the first visit were retained for all statistical procedures.

Field and laboratory methods Benthic macroinvertebrate data were collected at each study site over a length of 40 times the channel width, with a minimum reach length of 150 m and a maximum reach length of 4 km (USEPA 2007). Data were collected using two different sampling methods, with the first being the standard reachwide method (RW) used to collect data supporting the WSA (USEPA 2004), and the second

being the alternate LG method (USEPA 2007). For each method, a 0.09-m² quadrat sample was collected at each of 11 transects equally distributed along the same sample reach (D-frame net: 500 μ m mesh). Samples were alternately collected at either a left, center, or right point along each transect with the initial location on the first transect being randomly selected (USEPA 2007). For the RW method, collection points were located at 25, 50, or 75 % of the stream width; common habitats sampled included bottom substrate, woody debris, macrophytes, and leaf packs. For the LG method, collection points were located at 0, 50, or 100 % of the stream width; this method included stream edge habitats (e.g., undercut banks and root wads) that the RW method frequently did not. The LG method's inclusion of edge habitats also likely increased the frequency of sampling snags and macrophyte beds. The initial location for the RW sample was randomly selected. Then, the LG sample was shifted to the next location in a left, center, or right configuration. All subsequent transects were shifted one position or location to the "right" until all 11 transects were sampled.

For each method, collected samples were composited and field preserved with ~95 % ethanol. In the laboratory, with the goal of having 300 organisms available for identification, a randomized 500-organism subsample was

sorted using a gridded screen and preserved separately from the rest of the sample (USEPA 2008). If a sample contained <300 organisms, the entire sample was sorted. Sorted organisms were then identified to the taxonomic level specified for the study (usually genus; USEPA 2008). Field methods for all parameters are described in detail in the NRSA field manual (USEPA 2007).

Analysis A stream gradient threshold to identify low-gradient streams for assessment purposes has not officially been defined by the USEPA. Therefore, all sites sampled in the five targeted AOE's with complete data were retained for analysis, and stream gradient was evaluated as a continuous variable in conjunction with biological data collected at each site (described below). To evaluate whether the two methods collected similar numbers of organisms, total organism estimates were generated by extrapolating from the laboratory subsample to the full sample volume, accounting for those samples where the targeted number of organisms (i.e., 500; USEPA 2008) was sorted from a fraction of the sample. Nonparametric Wilcoxon signed-rank tests (with a continuity correction) were used to compare the estimated total organisms between paired LG and RW samples, both overall and by AOE. Two-way contingency tables were computed and McNemar's test of symmetry was used to test for

Table 1 Wadeable streams assessment multimetric index metrics, by AOE (X indicates metric part of MMI for that AOE)

Metric	Metric Code	CPL	NPL	SPL	TPL	XER
No. EPT taxa	EPT_NTAX	X		X		X
% EPT taxa	EPT_PTAX		X*			
% EPT individuals	EPT_PIND			X	X	
No. Ephemeroptera taxa	EPHENTAX		X		X	
% super tolerant taxa (PTV $\geq$ 8)	STOLPTAX				X	
% Non-insect indiv.	NOINPIND	X				X
No. shredder taxa	SHRDNTAX	X				
No. scraper taxa	SCRPNNTAX		X*	X*	X	X
No. clinger taxa	CLNGNTAX				X*	
% burrower taxa	BURRPNTAX		X	X		
% clinger taxa	CLNGPTAX	X				X
Shannon diversity	HPRIME	X	X	X	X	
% indiv. in dominant 5 taxa	DOM5PIND					X
% indiv. with PTV<6	NTOLNTAX		X			
% tolerant (PTV $\geq$ 6) taxa	TOLRPNTAX	X				X
No. of intolerant (PTV $\leq$ 3) taxa	INTLNTAX			X ^a		

* $p=0.05$, significant paired difference between methods

statistical differences between methods in the proportion of sites meeting the NRSA target of having at least 300 identifiable organisms for metric calculation and model development. To investigate patterns in discrepancy between methods along a continuum of stream gradient, we ran a generalized linear model (GLM) with paired differences in the total numbers of organisms (RW-LG) as the response, and stream gradient, AOE, and a stream gradient–AOE interaction as predictors. We then re-ran the Wilcoxon test and the GLM analysis on the subset of sites with a stream gradient $\leq 1\%$ ($N=493$).

To determine whether the two methods resulted in different assessment endpoints, we calculated both metric values and MMI scores for all samples. The metrics and MMI used were those developed to assess condition of

individual AOE of the WSA (USEPA 2006; Stoddard et al. 2008; Table 1). Metric values were calculated based on a random subsample of 300 organisms (or the full sample if it contained <300 organisms), and the metrics included in the MMI varied by AOE. We calculated paired differences by subtracting the LG index value from the RW index value for each visit. We ran a Wilcoxon Signed-rank test with continuity correction on these differences for each metric, as well as for the MMI, both overall and by AOE. We again used GLM to model differences on AOE, stream gradient, and the interaction of these two factors. Also as described above, we re-ran both sets of analyses with only sites having a stream gradient of $\leq 1\%$. This stream gradient threshold was recommended by Mazor et al. (2010) and by the Virginia Department of Environmental Quality

Fig. 2 Box plots showing the distributions of Total Number of Organisms Collected and MMI scores by method and AOE. Boxes capture interquartile range (IQR) with notches marking

the approximate 95 % confidence interval around the median. Whiskers show 1.5 times the IQR, and outliers are marked with empty circles

Fig. 3 Pairwise differences in total number of individuals collected (RW-LG), plotted across the full range of observed stream gradients on a logarithmic scale

(L. Willis, personal communication). While it was used to aid in the interpretation of results, it should not be considered an official criterion proposed by the USEPA.

We used R (version 2.15.2 Patched; R Core Team 2013) to perform all statistical analyses and create plots. We used the base statistics and graphics for all analyses and some plots, and the lattice package (Sarkar 2008) for panel plots.

Results

The total number of organisms collected showed similar distributions between methods for most AOE (Figs. 2 and 3), and the only significant difference (p value, <0.05) was for the NPL region (Table 2), with the RW method collecting more organisms per sample. The significant difference (p value, <0.05) detected for ALL

Table 2 Results of Wilcoxon signed-rank test on total number of individuals collected and multimetric index (MMI) scores, and average percentage slope, by AOE

AOE	Total number of individuals		MMI score		% Slope		
	Median difference (p value)		Median difference (p value)		Mean	Min	Max
CPL	45.5 (0.30)		1.97 (0.28)		0.198	0	2.570
NPL	42.5 (0.04)		-0.38 (0.32)		0.123	<0.001	1.797
SPL	-8.0 (0.48)		-1.25 (0.10)		0.265	0	8.128
TPL	54.9 (0.08)		1.08 (0.39)		0.186	0	3.956
XER	-1.0 (0.38)		-0.33 (0.56)		0.908	0.010	4.916
ALL	22.0 (0.03)		-0.08 (0.44)		0.297	0	8.128

Table 3 McNemar's test of symmetry results and contingency table counts of samples for each method with fewer than 300 and at least 300 organisms (NRSA threshold)

AOE	χ^2	<i>p</i> value	<i>N</i>	RW<300, LG<300	RW≥300, LG<300	RW<300, LG≥300	RW≥300, LG≥300
CPL	2.58	0.11	78	33	13	6	26
NPL	3.57	0.06	138	36	19	9	74
SPL	1.88	0.17	129	48	13	21	47
TPL	0.00	1.00	103	33	14	14	42
XER	0.20	0.65	77	12	2	3	60
All	0.56	0.45	525	162	61	53	249

AOEs was therefore driven by the differences in the NPL region. The GLM of differences in total organisms between methods showed no significant effect (all *p* values, >0.05) of either slope (log-transformed), AOE, or the interaction. In addition, when examined by both AOE and overall, one method did not exhibit a greater tendency to meet the 300-organism target than the other as measured by McNemar's test of symmetry (Table 3). Considering only sites with a slope of ≤1 % changed the results very little, primarily eliminating any significant differences in total number of organisms between method (Table 4).

MMI scores did not differ significantly between methods either overall or by AOE (Table 2), and differences between methods did not appear to be related to the condition of the site (as judged by the MMI score), regardless of AOE (Fig. 4). Only five differences in metric values between methods were significant (Table 1). For the NPL region, the number of scraper taxa differed significantly between methods, although the median difference was 0. This same situation occurred in the SPL region for numbers of intolerant taxa and scraper taxa. This occurs when there are large numbers of zero values because differences of zero are ignored in the Wilcoxon signed-rank test. When only non-zero differences were considered, the median difference was −1 (RW-LG) for all three metrics. In addition, percent EPT taxa was slightly higher in the

NPL for the RW method (difference=1.075), and there were slightly more clinger taxa in the TPL region (difference=1). However, there were no other differences between methods at the metric level. Modeling differences in MMI scores between methods showed no effect of slope (log-transformed), AOE, or the interaction (all *p* values, >0.10; Fig. 5).

By looking only at sites with slope of ≤1 %, we obtained slightly different results. There were still no significant differences in MMI scores, but one additional metric exhibited a significant difference between methods and one metric no longer exhibited a difference, both in the SPL region (Table 4). Median differences in metric values were again quite small. For the number of scraper taxa (in NPL and SPL), the issue of relatively large numbers of zero values for differences again created a situation where the median difference was zero, but if all zero values were ignored, the median differences were −1 and 1, respectively.

Discussion

Previous research has demonstrated that stream habitat differs between low- and high-gradient streams (e.g., Smock and Gilinsky 1992; Montgomery and Buffington

Table 4 Wilcoxon signed-rank test results (median *p* value) comparing total number of individuals, MMI scores, and significantly different metrics between methods (RW-LG) by AOE for sites with stream gradient ≤1 %

AOE	<i>N</i>	Total individuals	MMI score	Significantly different metrics (<i>p</i> <0.05)
CPL	76	36.1 (0.37)	1.87 (0.33)	
NPL	137	40.9 (0.06)	−0.59 (0.31)	EPT_PTAX, SCRPNTAX
SPL	125	−8.0 (0.63)	−1.15 (0.17)	SCRPNNTAX, EPT_PIND
TPL	100	53.5 (0.08)	1.06 (0.46)	CLNGNTAX
XER	55	−64.6 (0.78)	−0.65 (0.69)	
ALL	493	15 (0.08)	−0.08 (0.50)

Fig. 4 Bivariate plots of pairwise MMI score distributions plotted for all sites and by AOE

1997; Rinella and Feminella 2005), and that benthic community structure and composition change in response to such habitat changes (Minshall 1984). These changes, however, do not mandate that different field collection methods and assessment endpoints be used to assess condition. A single method may indeed suffice as long as the field method used adequately samples the assemblage in a manner that permits condition assessment (e.g., numbers of organisms), and the assessment and measurement endpoints selected for establishing condition respond appropriately across the range of site types. For example, the individual metric values of a MMI may change as samples are collected from high- to low-gradient streams, but if the metrics of the index are selected, or even adjusted, to account for this, the index can be effective over the full range of stream gradients. In the case of multivariate-predictive models (i.e., RIVPACS-type O/E model; Wright et al. 1993), predictors can be selected to incorporate this type of variation and would only be included in the model if they are relevant.

In this study, two primary questions were addressed. The first was whether the two sampling methods (i.e., LG and RW) differed in the total number of organisms collected at sites, overall, or at a definable percent slope (e.g., $\leq 1\%$). As stated before, this is a highly relevant question as insufficient numbers of organisms can impact assessment endpoint scores or even prohibit their calculation (e.g., Mazor et al. 2010). No significant differences were detected that support the use of the LG method over the RW method at sites of any slope. However, it is important to note that neither method performed particularly well in collecting the targeted number of organisms for MMI calculation (i.e., 300), especially in the four AOE that included plains. This is likely the result of both the LG and RW field methods sampling inadequate substrate area in the field, a problem easily remedied by increasing the area of subsamples or the number of subsamples prescribed by the field method. It could also be the result of, or a problem exacerbated by, how samples were processed in the laboratory. For example, of those

Fig. 5 MMI score differences (RW-LG), plotted against stream gradient on a logarithmic scale for all sites and by AOE

organisms sorted from debris, a considerable number may have not met the criteria outlined in the laboratory manual as to what was to be retained for identification (USEPA 2008). Regardless of cause, the target number of organisms was not collected at the majority of sites sampled in the five AOE's included in this study. We recommend that this element of the study design be carefully examined to discover how existing study protocols might be modified to increase the number of sites at which the target number of organisms is successfully collected. If a change is made to the field protocols, we recommend that the change apply to all AOE's of the NRSA, not just those included in this study, to support the NRSA goal of method consistency across the entire study area.

Our finding of no significant differences in the number of organisms collected by the LG and RW methods can be compared to those of Mazor et al. (2010) who compared three methods in California low-gradient streams, two of which were very similar to the two tested in the present

study (i.e., margin–center–margin and reach-wide benthos methods, respectively). Mazor et al. (2010) had the same field target of collecting a minimum of 500 organisms in each sample. They found that the RW-analogous method did not collect adequate numbers of organisms in nearly half of all the samples ($n=21$). However, the LG-analogous method collected adequate numbers at the majority of sample sites in the Mazor et al. (2010) study. Based on this, Mazor et al. (2010) made the recommendation that the LG-analogous method be used in future studies sampling California low-gradient streams. No obvious explanation for this differing result from our study is apparent. The total area sampled in the field and net dimensions are comparable. Possible explanations could be differences in mesh size (not listed by Mazor et al. (2010)), differences in how the samples were collected in the field, or differences in how the samples were processed in the laboratory.

The second and more noteworthy question addressed by our study was whether use of an alternative sampling

method resulted in different assessment endpoints (community condition scores). In agreement with Mazor et al. (2010), our study did not find significantly different index scores between the two methods, although significant differences were found among some metric values in some regions. These AOE-specific metric differences should be noted by scientists working in the identified AOE, especially if the scoring of the affected metrics might influence interpretation of study results.

In summary, all statistical tests conducted in this study support the finding that the “alternate” (LG) method and the “standard” (RW) collection methods result in a similar assessment of wadeable streams of the five AOE included in this study. These results were true across and within all AOE when all stream gradients were considered and when only those with a slope $\leq 1\%$ were examined. In short, use of the LG method would not ultimately impact USEPA assessments of national stream health in the studied AOE even for those streams with a slope $\leq 1\%$. This is not to say that differences do not exist between high- and low-gradient streams. Our results demonstrate that even though differences do exist (Smock and Gilinsky 1992; Montgomery and Buffington 1997; Rinella and Feminella 2005), especially with regard to physical habitat, a single well-designed benthic macroinvertebrate sampling method can be used to sample and assess site conditions across a mixture of stream habitat types. This finding is of high value to the NRSA program, as well as other bioassessment programs, as the development and maintenance of unique assessment tools for differing habitat types can be prohibitively expensive and potentially impede comparisons of results from different regions and stream types. However, it is important to note that while neither method performed better than the other, neither collected sufficient numbers of organisms, a conclusion that should be considered in future surveys.

Acknowledgments USEPA funded and collaborated in the research described herein, with analysis under contract number EP-D-06-096 to Dynamac Corporation. Although this work was reviewed by USEPA and approved for publication, it may not necessarily reflect official Agency policy. Dave Peck, Greg Toth, Lou Reynolds, Phil Kaufmann, Ellen Tarquinio, and Richard Mitchell of USEPA and Marlys Cappaert of SRA International provided assistance in data acquisition, while Erica Grimmer and William Thoeny of Dynamac Corporation and Greg Pond of USEPA provided taxonomic expertise. GIS support was provided by Ellen D’Amico of Dynamac Corporation. B. Johnson, C. Lane, M. Bagley and several anonymous reviewers are acknowledged for astute and constructive comments on the previous drafts of this manuscript.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

- Mazor, R. D., Schiff, K., Ritter, K., Rehn, A., & Ode, P. (2010). Bioassessment tools in novel habitats: an evaluation of indices and sampling methods in low-gradient streams in California. *Environmental Monitoring and Assessment*, 167, 97–104.
- Minshall, G. W. (1984). Aquatic insect–substratum relationships. In V. H. Resh & D. M. Rosenberg (Eds.), *The Ecology of Aquatic Insects* (pp. 358–400). New York: Praeger.
- Montgomery, D., & Buffington, J. (1997). Channel-reach morphology in mountain drainage basins. *Geological Society of America Bulletin*, 109, 596–611. doi:10.1130/0016-7606(1997)109<0596:CRMIMD>2.3.CO;2.
- Olsen, T. (2007). National Flowing Waters Assessment Survey Design: 2008–2009. http://water.epa.gov/type/oceb/upload/2008_02_06_riverssurvey_flowingsurvey0809.pdf. Accessed 30 Aug 2012.
- Omernik, J. M. (2004). Perspectives on the nature and definition of ecological regions. p. 34—Supplement 1, pp.27–38.
- R Core Team. (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. <http://www.R-project.org/>.
- Rinella, D. J., & Feminella, J. W. (2005). Comparison of benthic macroinvertebrates colonizing sand, wood, and artificial substrates in a low-gradient stream. *Journal of Freshwater Ecology*, 20(2), 209–220.
- Sarkar, D. (2008). *Lattice: multivariate data visualization with R*. New York: Springer. ISBN 978-0-387-75968-5.
- Smock, L. A., & Gilinsky, E. (1992). Blackwater streams of the Atlantic Coastal Plain. In C. T. Hackney, S. M. Adams, & W. A. Martin (Eds.), *Biodiversity of the Southeastern United States* (pp. 215–249). New York: Wiley.
- Stoddard, J. L., Herlihy, A. T., Peck, D. V., Hughes, R. M., Whittier, T. R., & Tarquinio, E. (2008). A process for creating multimetric indices for large-scale aquatic surveys. *Journal of the North American Benthological Society*, 27, 878–891.
- US Environmental Protection Agency (USEPA). (1997). *Field and laboratory methods for macroinvertebrate and habitat assessment of low gradient nontidal streams; Mid-Atlantic Coastal Streams Workgroup, Environmental Services Division*. Wheeling: S. Environmental Protection Agency, Region 3.
- US Environmental Protection Agency (USEPA). (2004). Wadeable streams assessment: field operations manual. Washington, D.C., USA. Environmental Protection Agency, Office of Research and Development and Office of Water and Office of Research and Development, EPA841-B-04-004.
- US Environmental Protection Agency (USEPA). (2006). Wadeable streams assessment: a collaborative assessment of the Nation’s streams. Washington, D.C., USA. Environmental Protection Agency, Office of Research and Development and Office of Water, EPA841-B-06-002.

- US Environmental Protection Agency (USEPA). (2007). National rivers and streams assessment: Field Operations Manual. Washington, D.C., USA, Office of Water, EPA-841-B-07-009.
- US Environmental Protection Agency (USEPA). (2008). National rivers and streams assessment: laboratory methods manual. Washington, D.C., USA, Office of Water and Office of Research and Development, EPA-841-B-07-010.
- US Environmental Protection Agency (USEPA). (2011) National aquatic resource surveys: an update. <http://water.epa.gov/type/watersheds/monitoring/upload/nars-progress.pdf>. Accessed: 2 Nov 2012.
- US Environmental Protection Agency (USEPA). (2012a) National aquatic resource surveys. http://water.epa.gov/type/watersheds/monitoring/aquaticsurvey_index.cfm. Accessed: 2 Nov 2012.
- US Environmental Protection Agency (USEPA). (2012b) National rivers and streams assessment. <http://water.epa.gov/type/rsl/monitoring/riverssurvey/index.cfm>. Accessed 2 Nov 2012.
- US Government Printing Office (USGPO). (1989). Federal Water Pollution Control Act (33 U.S.C. 1251 et seq.) as amended by a P.L. 92–500. In: *Compilation of selected water resources and water pollution control laws*, Washington, D.C. USA, Printed for use of the Committee on Public Works and Transportation: US Government Printing Office.
- Wright, J. F., Furse, M. T., & Armitage, P. D. (1993). RIVPACS: a technique for evaluating the biological water quality of rivers in the UK. *European Water Pollution Control*, 3, 15–25.