

# Virginia SOS Modifies Method, Improves Reliability

by Eleanor Ely

In the spring of 1998, Virginia Save Our Streams (SOS) coordinator Jay Gilliam enrolled in Reese Voshell's aquatic entomology class at Virginia Tech. At the time, Virginia SOS was engaged in a push to encourage the Virginia Department of Environmental Quality (DEQ) to use the program's macroinvertebrate data; but some DEQ biologists were skeptical of the data's quality. So one day,

issue. For a thorough discussion of the methods, findings, and analysis, readers are encouraged to read the published paper, which is available in a downloadable pdf format at [www.vasos.org/ValidationStudy.htm](http://www.vasos.org/ValidationStudy.htm).

## Virginia SOS's original method

Since the Virginia SOS program's inception in 1988, volunteers had followed

categories. Note that the actual number of organisms is not a factor in determining the final score. Only taxa richness—the number of different taxa present in each category—is reflected in the water quality rating. A separate score is calculated for each sample, and the highest of the three scores is considered the most accurate indication of ecological conditions.

STACEY BROWN



Jay Gilliam (left) and James Riverkeeper Lynn Ridley demonstrate the new SOS macroinvertebrate monitoring method.

Gilliam asked Voshell what he thought about the SOS data.

Voshell replied, "I'm a scientist—I'd have to study it scientifically in order to give you a scientific answer."

Out of this two-minute conversation grew an intensive two-and-a-half-year study that ultimately provided SOS with a modified method whose results correlate well with professional results. Even more significantly, according to Virginia SOS staff scientist Stacey Brown, the study "stands as a model that other volunteer monitoring groups can use to validate their work and find metrics appropriate to their area."

The study has been published in the journal *American Entomologist*, Fall 2002

the macroinvertebrate monitoring method developed by the Izaak Walton League of America (IWLA) Save Our Streams program (see Kellogg 1994). The IWLA method calls for collecting three individual kick-net samples. Macroinvertebrates are sorted and identified in the field (mainly to order level) using the IWLA picture key.

To rate the stream's water quality, the protocol assigns each group of organisms to a sensitivity category (sensitive, somewhat sensitive, or tolerant). For example, mayfly nymphs are rated sensitive; clams, somewhat sensitive; and leeches, tolerant. A water quality rating score is calculated based on the number of taxa present in each of the three sensitivity

## The study

For their study, Voshell and graduate student Sarah Engel chose 23 regular Virginia SOS monitoring sites that represented a range of ecological conditions. In 1998, Engel accompanied volunteer monitors on their visits to their sites and observed their technique. After the volunteers finished their sampling and identification, Engel preserved the samples for later verification in the laboratory. In addition, Engel preserved all debris left on the kick-net, to assess the completeness of the volunteers' picking.

Engel then collected her own sample (termed the "professional sample") at an undisturbed location at the same site, following the U.S. EPA's rapid bioassessment protocol (Barbour et al. 1999). The professional samples were preserved in ethanol, then sorted, identified to genus level, and counted in the laboratory. In addition, Engel sampled 122 "historical" sites that Virginia SOS volunteers had sampled during the past five years.

Two indexes were calculated for the professional samples: the Hilsenhoff biotic index (HBI) and a multimetric index called the macroinvertebrate aggregated index for streams (MAIS), which was specifically developed for streams in the mid-Atlantic highlands.

## Stream quality overrated

Engel and Voshell compared the SOS water quality rating score for each site with the MAIS score and HBI value

determined from the professional sample. They found that, for both concurrent and historical samples, volunteer scores did not correlate well with professional scores. R-values ranged from 0.36 to 0.58, well below the target value of 0.70 that Engel and Voshell had chosen as the criterion for acceptable correlation.

More importantly, the volunteer and professional methods also disagreed in categorizing stream water quality as acceptable or unacceptable. For the 23 concurrently sampled sites, conclusions about ecological conditions based on the SOS score differed from those based on the MAIS score in 8 cases, and each time the SOS method rated the stream acceptable while the MAIS score rated it unacceptable. Twelve sites were rated acceptable by both the volunteer and the professional methods, and 3 were rated unacceptable by both. The histori-

**THE PATTERN WAS CLEAR:  
THE SOS METHOD  
WAS CONSISTENTLY  
OVERESTIMATING  
STREAM QUALITY.  
NOW THE QUESTION  
WAS WHY.**

cal sites showed a similar pattern. In no case was a stream rated acceptable by MAIS but unacceptable by the SOS score.

The pattern was clear: the SOS method was consistently overestimating stream quality. Now the question was why. Were the discrepancies caused by the volunteers' field methodology (sampling, sorting, identifying), or did the problem reside in the scoring method itself?

After careful analysis of the data, Engel and Voshell concluded that the problem lay with the scoring method. As they wrote in their paper, "The SOS score is not rigorous enough to distinguish impaired ecological conditions because it is based solely on the presence or absence of taxa, without consideration of their abundance."

### Modifying the method

Engel and Voshell's next step was to see if they could modify the Virginia SOS

## Bias and Error in Bug Sorting and ID

Another academic who's been investigating volunteer macroinvertebrate monitoring methods is Julia Frost Nerbonne. For her master's thesis, Nerbonne examined some sources of bias and error in sorting and identifying macroinvertebrates, and for her Ph.D. she conducted an extensive survey of macroinvertebrate monitoring groups. She is still analyzing the results of this survey, which she promises to summarize in a future article in *The Volunteer Monitor*.

An article based on Nerbonne's master's thesis research has just been published in the *Journal of the North American Benthological Society*.<sup>\*</sup> The study identified errors made by previously untrained volunteers, mostly high school students, who received only a brief introduction to the methods. As Nerbonne notes in the article, it is critical that the findings "be evaluated in context and not be used to characterize all volunteer efforts." Briefly, the study revealed that when sorting macroinvertebrates from debris these novice volunteers were biased toward selecting larger rather than smaller organisms, and toward selecting organisms that moved slowly rather than those moving very fast or those not moving at all. These volunteers also encountered a number of difficulties with identification.

Clearly, Nerbonne's study underscores the need for thorough training, as well as the importance of using adequate lighting and magnification for both sorting and identification. Noting that "macroinvertebrate monitoring is difficult for volunteers," Nerbonne recommends that volunteer program leaders focus on attracting highly motivated volunteers who will commit sufficient time and effort to learning and performing the techniques. In addition, she suggests that volunteers work in teams so they can verify each other's sorting and identification efforts.

For more information contact Julia Frost Nerbonne at [jaf@fw.umn.edu](mailto:jaf@fw.umn.edu).

<sup>\*</sup>Nerbonne, Julia Frost and Bruce Vondracek. 2003. Volunteer macroinvertebrate monitoring: Assessing training needs through examining error and bias in untrained volunteers. *Journal of the North American Benthological Society* 22(1):152-163.

protocol to bring it into closer agreement with professional results. An important constraint was that any modification must still allow field identification of organisms. Gilliam felt that only a small percentage of SOS's 350-odd active volunteer monitors would be interested in identifying preserved specimens to lower taxonomic levels in the laboratory. Nor did SOS want to send the samples to a professional lab for identification. As Voshell explains, "We knew the volunteers wanted to be involved with the whole process."

Engel and Voshell focused on making both sampling and scoring more quantitative. In the modified method, volunteers sort, identify, and count the entire contents of the kick-net, which must total at least 200 organisms. If there are fewer than 200 organisms in the initial sample, the volunteers must collect additional samples.

In 1999, Engel repeated concurrent sampling with volunteers at 23 sites (not all the same sites as in 1998), but this

time the volunteers used the new sampling method. The professional samples were collected in the same manner as previously. Engel and Voshell used the results of the 1999 sampling to evaluate 24 potential metrics,<sup>\*</sup> all suitable for the level of identification that volunteers were performing. From these, they selected six: % EPT (percentage of sample that consists of mayflies plus stoneflies plus all caddisflies except net-spinners), % net-spinning caddisflies, % lunged snails (commonly called left-handed snails), % beetles, % tolerant,

*continued on next page*

<sup>\*</sup>Note: A metric is a measurable attribute of a biological system that provides a reliable indication of biological condition across a gradient of human influence. For example, % EPT—a measurement of the relative abundance of certain pollution-sensitive species in the macroinvertebrate population of a stream—is a useful metric because it tends to decrease as stream biological condition becomes more degraded. A *multimetric index* integrates several metrics into an overall score.

SOS, continued

and % non-insects. These six metrics are combined to yield the Virginia SOS multimetric index.

The R-value for the correlation of the SOS multimetric index score with the MAIS score for the 23 sites was 0.6923, only narrowly below the target value of 0.70. In classifying stream ecological condition as acceptable or unacceptable, the SOS multimetric index showed 95.7% agreement with the professional conclusion for the 23 sites.

No one claims that the new SOS method is the equivalent of a professional assessment, but Voshell stresses that it almost always reaches the same overall conclusion about acceptable versus impaired water quality. To reach more detailed conclusions—for example, about degree of impairment or cause and effect—would require more detailed professional methods, he notes.

Voshell says, “Critics may say, ‘They’re not identifying organisms to species’ or ‘The net is too coarse,’ and that’s true, but the critics aren’t looking at the end result. It’s the ultimate decision that we’re concerned about. We found a way to use a simpler method and still reach the same conclusion as the professional biologists in the state.”

“I think most of the volunteers feel excellent about the transition to the new method,” says Gilliam. “The only diffi-

cult aspect is that now we ask them to distinguish between the net-spinning caddisfly and the other caddisflies, because the netspinner is much more tolerant to pollution. But overall, the modified method requires fewer family-level identifications than the original method.”

What’s more, the new method only requires collecting one sample. Gilliam says, “You do need to carefully scrutinize your catch and count every single organism, but once you get used to the patience that’s needed, it’s a whole lot easier to do it one time thoroughly than to do it three times, the way we did before.”

Broader implications

What does all this mean for other volunteer programs that monitor macroinvertebrates, particularly those that use simplified methods such as the traditional IWLA method? Should they undertake a similar study? Not necessarily, says Gilliam. “Which method is most appropriate depends on your goals,” he says. “The traditional IWLA method is a tremendously valuable technique to introduce people to biological monitoring.”

For groups that want their data used by state agencies, a validation study is critical. As Voshell notes, “State agencies are in the heat of battle and have to be able to justify their decisions. If vol-

unteers want their data used for important decisions, they need to do validation and evaluation studies.” But Voshell cautions that other groups cannot simply adopt the Virginia SOS method, which was developed specifically for the western part of Virginia. In fact, a separate validation study is required even for the eastern (non-mountainous) part of Virginia. Such a study, following Engel and Voshell’s model, is currently being conducted at Randolph Macon College. Voshell says, “I think our modified method would work in rocky-bottom, steep-gradient, mountain streams in our ecoregion (West Virginia and parts of Kentucky, Maryland, North Carolina, Tennessee, and Pennsylvania), though I would recommend that people do their own validation study. But if you went out to California, there’s no reason to think the same metrics would work.” He adds that volunteer groups also need to make sure their methods produce results that agree with the results of their own state agency’s protocols.

Engel and Voshell’s study provides a model for other groups to follow. Those who are not able to partner with a nearby university may be able to get help from their state agency. Voshell suggests that the volunteers could monitor the same sites where the agency routinely monitors and see if they reach the same conclusions about ecological condition. They might also want to preserve their samples, and send them in to the agency for validation.

The payoff

The Virginia SOS study was possible because of the strong three-way partnership between Virginia Tech, Virginia SOS, and Virginia DEQ. The university provided not only the expertise of Voshell and Engel but also equipment and laboratory facilities. Virginia SOS staff enlisted volunteer monitors who were willing to participate in concurrent sampling, and helped coordinate the site visits. And the study was funded by Virginia DEQ—a testament both to the good relationship between Virginia SOS and DEQ and to DEQ’s recognition of the potential value of citizen-collected data. “The DEQ only monitors about 18

RESOURCES

Barbour, M., et al. 1999. *Rapid Bioassessment Protocols for Use in Wadeable Streams and Rivers: Periphyton, Benthic Macroinvertebrates, and Fish*. 2nd edition. U.S. EPA, publication number EPA 841-B-99-002. Order from National Service Center for Environmental Publications, 800-490-9198.

Engel, Sarah R. and Voshell, J. Reese, Jr. 2002. Volunteer biological monitoring: Can it accurately assess the ecological condition of streams? *American Entomologist* 48(3):164-177. Available in pdf format at [www.vasos.org/ValidationStudy.htm](http://www.vasos.org/ValidationStudy.htm).

Kellogg, Loren. 1994. *Monitor’s Guide to Aquatic Macroinvertebrates*, 2nd edition. Izaak Walton League of America. Complete instructions for the IWLA

macroinvertebrate monitoring method, including picture key. \$6. To order, call 1-800-BUG-IWLA or visit [www.iwla.org/iwlastore](http://www.iwla.org/iwlastore).

Voshell, J. Reese, Jr. 2002. *A Guide to Common Freshwater Invertebrates of North America*. Field guide used by Virginia SOS volunteers. Over 200 illustrations. \$29.95. Order from McDonald and Woodward Publishing Co., 800-233-8787; [www.mwpubco.com](http://www.mwpubco.com).



percent of stream miles in the state,” says Gilliam. “Once the Randolph Macon study is completed, SOS’s contribution should bring that figure up to 30 percent or more.”

Gilliam adds, “My main motivation for improving the reliability of our method was to be able to say to our volunteer monitors that the time and effort they were putting in was having significant usefulness to state and local governments in making decisions. And I think we’ve achieved that.” As evidence, he points to the 2002 Letter of Agreement between Virginia SOS, DEQ, and others, which specifies that DEQ will use the Virginia SOS data for baseline information, as a red flag for problems, and in the state 305(b) report. “This wouldn’t have happened without the study,” he says.

*For more information, contact Jay Gilliam, Coordinator, Virginia Save Our Streams, Raphine, VA; 540-377-6179; jay@vasos.org; or Reese Voshell, Department of Entomology, Virginia Tech, Blacksburg, VA 24061; rvoshell@vt.edu.*

## Secchi Dip-In

A university/volunteer monitoring partnership of unusually large scope is the Great North American Secchi Dip-In, which is directed by three scientists at Kent State University (Robert Carlson, David Waller, and Jay Lee), with sponsorship by the North American Lake Management Society and the U.S. Environmental Protection Agency.

Last July, more than 2,500 volunteers from volunteer monitoring programs in the United States and Canada measured transparency in their favorite lake, reservoir, river, or estuary and sent their data to the Dip-In website. The majority used the traditional black-and-white Secchi disk, but a few used an all-black disk and others—particularly those monitoring streams and rivers—used transparency tubes. (The Dip-In website includes some interesting articles about these different methods.)

Each year, the Dip-In data are displayed on maps which are posted at the website. Lakes in the northern part of the United States and Canada are typically the clearest, while those in agricultural regions of the Midwest tend to have some of the lowest transparencies.

The Dip-In also collects information on the volunteers’ perception of water quality, which varies considerably from region to region. For example, a person in Maine might think a lake with a transparency of 6 feet was degraded, while in another state a lake with a transparency of only 1 foot might be considered of high quality. Another interesting finding has been that in some states personal watercraft now equal or surpass algae and weeds as the chief perceived water quality problem.

To find out how you can participate in July 2003, visit the Dip-In website at <http://dipin.kent.edu>.

*ALLARM, continued from page 4*

community-driven partnerships—are valuable, but they have different strengths.

We feel very strongly that the bottom-up approach, in which we train community members to do their own work, leads to a much higher level of understanding and empowerment. Group members may initially be intimidated by their increased decision-making power and time commitments, but volunteers who have struggled with study design and data interpretation are better equipped to identify risks, assess mitigation options, and participate in policy decisions.

The bottom-up approach is also more effective at building a sense of community and an appreciation for the power of organized groups. It gives the volunteers a sense of ownership of the project, and encourages them to become environmental stewards and advocates committed to seeking community unity for action.

That said, the benefits of the acid rain

monitoring project should not be sold short. Some volunteers prefer to work on a “ready-made” project rather than investing the time and energy required for community-designed programs. Our acid rain monitors are empowered by the strength of their numbers across the state, as well as an increased knowledge derived from frequent streams-site visits throughout the year.

Furthermore, the greater degree of institutional control over study design, methods, and data analysis has made possible more robust data that are more acceptable to the scientific community. Faculty and students have presented ALLARM’s acid rain monitoring results—which indicate that the problem of acid deposition has been underestimated in Pennsylvania and that many more streams are impacted than was previously thought—at professional conferences. The data have also been used by state agencies to revise fish stocking practices, and by citizens to craft testimony

in support of state and federal acid-deposition legislation.

Whichever model is used, volunteer monitors collect independent data that can be used to either uphold or challenge public agency and industry data. This opens previously closed doors for meaningful participation by citizens in environmental decision-making. This kind of dialogue leads to the mutual trust that is essential to cooperative efforts to solve environmental problems—a trust that is sorely lacking in most communities.

*Candie Wilderman is a Professor of Environmental Science at Dickinson College and the founder and Science Director of ALLARM. Alissa Barron is ALLARM’s Assistant Director, and Lauren Imgrund is the program’s Director. For more information, contact ALLARM, Environmental Studies Department, Dickinson College, P.O. Box 1773, Carlisle, PA 17013; 717-245-1565; [allarm@dickinson.edu](mailto:allarm@dickinson.edu).*