SCIENCE AND SOCIETY SERIES PAPER NO. 2

A COMPARISON OF A PROFESSIONAL METHOD AND A VOLUNTEER METHOD FOR ASSESSING STREAM HEALTH, INCLUDING DISCUSSION OF AN IMPROVED VOLUNTEER METHOD

USEPA's Rapid Bioassessment Protocol II
vs.
West Virginia Save Our Streams

NICOLE NAVIS AND W. NEIL GILLIES, JUNE 2001

"SEEKING A MORE DEFINITIVE UNDERSTANDING OF WATER QUALITY ISSUES IN THE POTOMAC WATERSHED"

THE SCIENCE AND SOCIETY SERIES WAS DEVELOPED BY THE CACAPON INSTITUTE WITH FUNDING FROM THE MARPAT FOUNDATION

CACAPON INSTITUTE, RT. 1 BOX 326, HIGH VIEW, WV 26808
TEL: 304-856-1385 FAX:304-856-1386
Pcrel@mountain.net
WWW.CACAPONINSTITUTE.ORG

A COMPARISON OF A PROFESSIONAL METHOD AND A VOLUNTEER METHOD FOR ASSESSING STREAM HEALTH, INCLUDING DISCUSSION OF AN IMPROVED VOLUNTEER METHOD

USEPA's Rapid Bioassessment Protocol II
vs.
West Virginia Save Our Streams

Abstract

In recent years, the science of using animals to assess stream health has gone public. Volunteer monitoring programs, such as the Izaak Walton League's Save Our Streams (SOS) program, have sprouted up around the country. The SOS and other volunteer methods are similar in general design to the methods used by professional biologists, but tailored to the capabilities of non-professionals. Cacapon Institute compared results from WV's volunteer SOS monitoring and the more scientifically rigorous Rapid Bioassessment Protocol (RBPII) stream assessment methods used by WV's Division of Environmental Protection. We found that SOS Stream Scores as currently calculated don't provide stream assessment data that are consistently comparable to professional RBPII results because they lack abundance data and thereby lack critical information. The identification level used with the SOS method (mayfly, stonefly, clam, etc.) can provide a stream assessment comparable to professional methods if actual counts of the organisms collected are obtained in order to properly weight the importance of each organism. Based on our study, it is possible for non-professional, volunteer conducted, benthic stream assessments to obtain results that compare favorably to professional assessments. The proposed method utilizes the same level of identification skill currently required of SOS volunteers and the same collection technique. It differs by requiring samples to be preserved in the field for "picking" under slight magnification and good lighting at home, in counting the organisms obtained, and in noting different "kinds" within each of the current SOS identification categories.

Science and Society Series, Number 2

By Nicole Navis and W. Neil Gillies

June 2001

Cacapon Institute, High View, WV

A COMPARISON OF A PROFESSIONAL METHOD AND A VOLUNTEER METHOD FOR ASSESSING STREAM HEALTH, INCLUDING DISCUSSION OF AN IMPROVED VOLUNTEER METHOD

USEPA's Rapid Bioassessment Protocol II vs. West Virginia Save Our Streams

Cacapon Institute, High View, WV

Introduction

The science of using animals to assess the vitality of a river ecosystem has gone public. Since the Izaak Walton League began their pioneering Save Our Streams (SOS) program in 1969, volunteer stream monitoring programs have sprouted up throughout the country. These volunteer programs make the link between causes and effects of pollution using hands-on, in-stream activities, and play an essential role in making the societal benefits of clean water immediate and real to participants. The SOS and other volunteer methods are similar in general design to the methods used by professional biologists, but tailored to the capabilities of non-professionals.

Cacapon Institute (CI) received funding from the WV Division of Natural Resources Non-Game Wildlife & Natural Heritage Research/Cooperative Projects Program to compare results from WV's volunteer SOS monitoring and the more scientifically rigorous Rapid Bioassessment Protocol (RBPII) stream assessment methods used by WV's Division of Environmental Protection (WVDEP). Both methods assess stream health using benthic macroinvertebrates, the small animals without backbones (invertebrates) that live on the river bottom (benthos) and are visible without magnification (macro). They include the water-living larval stages of flying insects such as mayflies, dragonflies and gnats, small shellfish such as the small, white asian clams that litter the bottoms of many eastern rivers, and many other kinds of life.

The advantage of using aquatic organisms over chemical indicators of water quality – such as the amount of nitrogen in a water sample – is that animals are constantly "sampling" their environment and the communities found in benthic samples are indicative of water quality conditions over time. Chemical measures, in contrast, provide a momentary snapshot in a constantly changing environment.

The challenge in using aquatic organisms, however, comes in understanding what the animals are telling us. Historically, researchers looked for individual "indicator" species that, by their presence, absence, or abundance relative to other organisms, indicate specific environmental conditions. For example, large numbers of midge larvae can indicate severe organic pollution. In recent decades, researchers have moved away from using single indicator species and toward indicator groups of species — called "indices." A typical index, for instance, might look at the total number of different species or the relative abundance of different species. The professional RBPII and volunteer WVSOS methods both take the approach of looking at indices based groups of organisms.

Our study had two primary goals:

- The first was to confirm a study done by Virginia Polytechnic Institute in 1999/2000, which found that the SOS method tends to overestimate the health of a stream when compared to a professional assessment. Their data indicated that the difference was not a result of the physical method of collection (i.e. net size, etc.), but how the data was analyzed.
- The second goal was to "tweak" the information gained by SOS field methods to obtain a result similar to WV's professional RBPII method, without making the procedure too difficult for the SOS volunteer. For this reason, we tried to modify SOS protocols not in physical collection methods, but in tallying methods and calculations.

A central assumption of this project was that the biological assessment protocols utilized in the RBPII methods represent a high professional standard and therefore provide a suitable "yardstick" to measure the success of volunteer methods.

Methodology

The WV Division of Environmental Protection and Cacapon Institute collected benthic macroinvertebrates at 20 sampling sites in West Virginia's Cacapon River watershed during the summer of 2000. WVDEP used standard RBP II collection and identification procedures (see box this page for details), while CI used WV's SOS method (see box page 5).

The RBP II and SOS methods are similar in overall design; the differences are in the details:

- The RBP II method uses a smaller mesh net size than SOS to collect samples at eight riffle sites rather than three.
- RBPII preserves samples in the field for later "picking" of a subsample under slight magnification in the laboratory, while the SOS method involves fieldpicking live organisms.
- The RBPII counts organisms and trained taxonomists identify them to the family level while the SOS method does not utilize counts and volunteers identify organisms to general categories (mayfly, stonefly, etc.) in the field.
- The **RBPII WV Stream Condition Index (WVSCI**) rates a site using the average of a set of six standard Indices that "weight" the importance of different groups using counts and tolerance to pollution. In contrast, the **SOS Stream Score** rates a site with a single numeric indicator that considers the presence (not abundance) and tolerance to pollution of different groups.

Picking?

When you collect a river bottom sample using a net, the net fills not only with animals but with debris such as leaves, algae and rocks. To get the animals out of that debris, you have to pick through the sample with a pair of forceps, good light, and a lot of patience.

By adding counts to the SOS procedure, we were able to calculate a SOS Stream Condition Index (SOSSCI) based on a set of six indices directly comparable to those used in the RBPII method, but using SOS level identification of organisms.

WVDEP's Professional Bioassessment Method

The RBPII method requires collecting eight samples in a riffle at each site using a D-frame and/or Surber 500-micron mesh net. All eight samples are then combined into one sampling container and preserved. Subsamples of each sample are later "picked" under low power magnification in the laboratory until at least 200 organisms are recovered. Identifications are made to the family level and the organisms counted.

WVDEP analyzes their data using a number of standard indices, each of which assesses a different aspect of stream health:

- **Total taxa** is simply the count of the number of families found in the sample. It is a measure of diversity, or overall richness. This index increases with stream health.
- **Number of EPT Taxa** is the count of the number of EPT families found in the sample. EPTs are the Mayflies (Ephemeroptera), Stoneflies (Plecoptera), and Caddisflies (Trichoptera). Stream health increases as the diversity of these three, often pollution sensitive, insect orders increases.
- **% EPT**. Percent EPT abundance measures the contributions to total benthic macroinvertebrate abundance of these three generally pollution sensitive orders. Increases with stream health.
- **% Chironomidae**. The percentage of midge (gnat) larvae in the family Chironomidae tends to increase in streams subject to organic pollution.
- **%two dominant taxa**. Dominance of the top two taxa measures the contributions to overall abundance of a few species (indicating lower diversity and the possibility of pollution) and decreases with increasing stream health.
- **Hilsenhoff Biotic Index**. The HBI assigns a pollution tolerance value to each identified taxon, and weights their importance in the benthic community using relative abundance. This index increases as water quality decreases.

Each of these indices is designed to assess a different component of stream health. In combination, they provide a more powerful assessment of stream health than any one alone. **The WV Stream Condition Index (WVSCI) stream rating** is the average of the six indices above (each standardized to a 0-100 point scale). WVSCI scores are interpreted using the following scale: sites scoring >68 are non-impaired, sites scoring <68 and >60.6 have a degree of uncertainty and need further investigation, and sites scoring <60.6 have a 95% chance of being truly biologically impaired.

Volunteer SOS Method

In the WVSOS method, trained volunteers collect 3 samples in a riffle using a relatively coarse kick net. The organisms collected in each sample are then sorted live out of the net and debris and identified in the field using a visual key that groups organisms into general categories, such as mayfly, stonefly, aquatic worm, etc. The stream's biological health is then determined by calculating the SOS Stream Score, for each of the three samples separately, in which the presence/absence of an organism (1 or 0) is multiplied by the organism's pollution tolerance level, ranging from 3 (not tolerant of pollution) to 1 (very tolerant of pollution) and summing the results (see table below for example). Abundance of organisms does not factor into this rating. A score ranging from 0-10 indicates a Poor stream quality rating, 11-16 Fair, 17-22 Good, and 23+ Excellent and the best of the three scores is used.

Abbreviated example of a SOS Stream Score calculation. Actual form has 20 different organisms.

1. Organism	2. Sensitivity Rating	3. Present or absent	4. Sensitivity X Presence
Caddisfly	3	1	3
Mayfly	3	0	0
Stonefly	3	1	3
Clam	2	1	2
Crayfish	2	1	2
Dragonfly	2	0	0
Aquatic Worm	1	1	1
Blackfly larva	1	1	1
Midge	1	0	0
Sum of Column 4 = Stream Score =			12

CI modified the standard SOS protocol slightly. Each of the three samples per site was collected using a timed, two-minute effort, in order to obtain a catch per unit effort. After handpicking live in the field, organisms were preserved in alcohol and brought to the lab for identification to the standard SOS level. We also counted these organisms, and noted visually distinct "kinds" within each SOS taxonomic grouping. This allowed us to calculate additional indices, on combined (pooled) results from the three samples, similar to those used in the professional RBPII method:

- **Total Richness** — number of visually distinct "kinds"of organisms, similar to the RBP Total Taxa.
- **Number of EPT "kinds"** — similar to RBP # of EPT Taxa.
- **% EPT abundance** — identical to RBP.
- **%Midges** — similar to, and in many cases identical to, RBP % Chironomidae.
- **% dominance** — most numerically dominant SOS "taxa" or group
- **Very Modified Hilsenhoff Biotic Index — The VMHBI** is similar to the RBP HBI index, but uses SOS-level identification, SOS pollution tolerance levels (inverted) and counts.

The SOS Stream Condition Index (SOSSCI) stream rating is the average of the six indices above (each standardized to a 0-100 point scale), calculated in the same manner as the WVSCI.

Results and Discussion

We assumed that RBPII methods provide a high standard and a "professional yardstick" against which volunteer methods can be measured. For that reason, SOS results are presented in comparison to RBPII results.

The data collected by WVDEP and CI allowed us to ask many questions, the most important on three levels:

1. Do SOS and WVSCI stream scoring methods result in the same stream quality rating.?
2. Does adding counts to the SOS method to weight the contribution of each group to the total benthic community improve the accuracy of the method?
3. Does the SOS Stream Condition Index (SOSSCI), based on a set of six indices directly comparable to those used in the RBPII method but using SOS level identification of organisms and counts, improve the accuracy of the method?

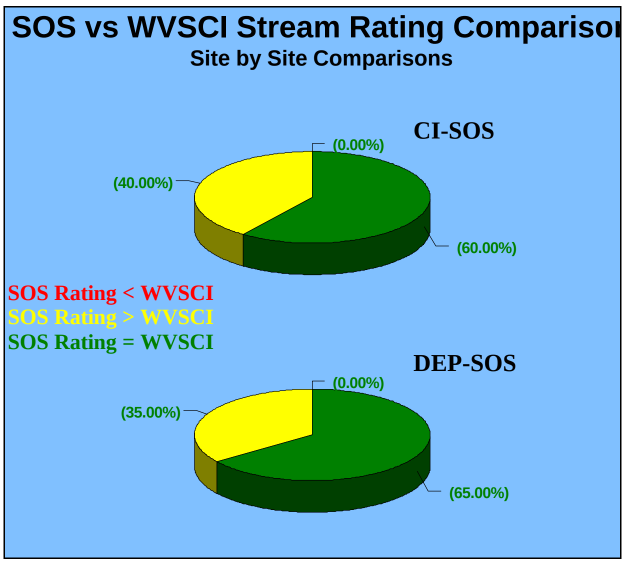
We were able to separate differences due to analytical methods and sampling method because WVDEP shared their raw data -- identified to the family level -- with us. We converted this data to SOS-level identifications and analyzed it according to SOS and the alternative SOS indices noted above. For this analysis, number of "kinds" was considered equivalent to number of families. This allowed us to compare SOS scores to the professional RBP data based on analytical methods alone, eliminating variation caused by differences in sampling and sample sorting methods.

In this report, we label the data collected, processed and analyzed by WVDEP under RBP protocols **DEP-RBP**. WVDEP data analyzed under SOS protocols is labeled **DEP-SOS**. Finally, data collected using SOS protocols is labeled **CI-SOS**.

Do SOS and WVSCI scoring methods result in the same stream quality rating? As noted above, WVSOS stream scores are used to determine a stream quality rating of Poor, Fair, Good, or Excellent. WVSCI scores result in a stream quality rating of Impaired, Uncertain Impairment, and Non-Impaired. For the purpose of comparing stream scores between the two analytical methods, SOS- Poor was considered equivalent to WVSCI- Impaired, SOS-Fair equivalent to WVSCI-Uncertain Impairment, and SOS-Good and Excellent equivalent to WVSCI-NonImpaired.

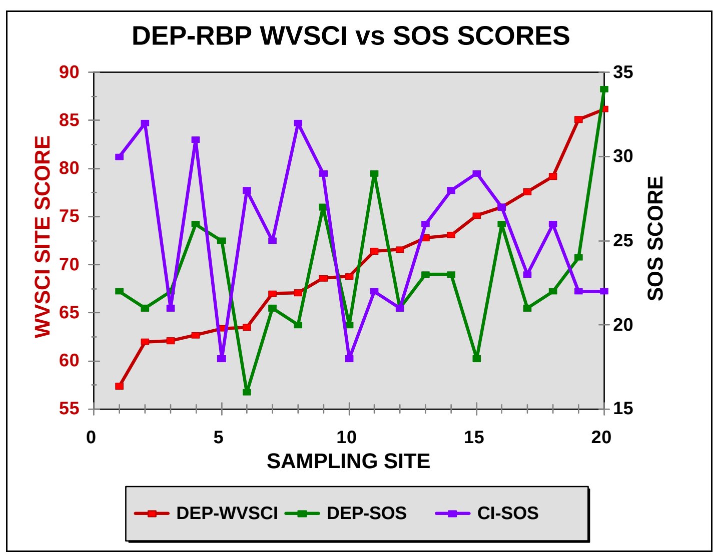
We compared overall stream ratings generated from the three analyses. The CI-SOS and DEP-SOS methods rated 100% and 95% of the sites to be unimpaired, respectively. Five percent of the DEP-SOS sites had uncertain impairment and none were scored impaired. The WVSCI rated 60% of sites unimpaired with 5% of the sites impaired and 35% of sites with uncertain impairment.

Analytical Method			
Site Rating	**DEP-WVSCI**	**DEP-SOS**	**CI-SOS**
Non-impaired	60%	95%	100%
Uncertain impairment	35%	5%	0%
Impaired	5%	0%	0%

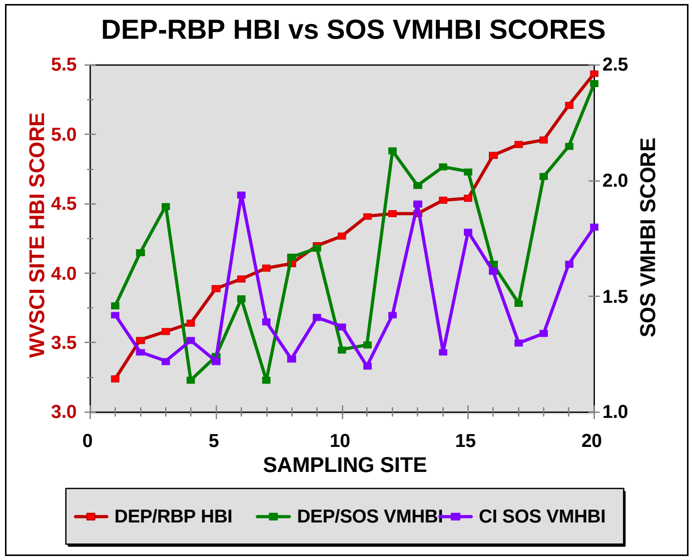


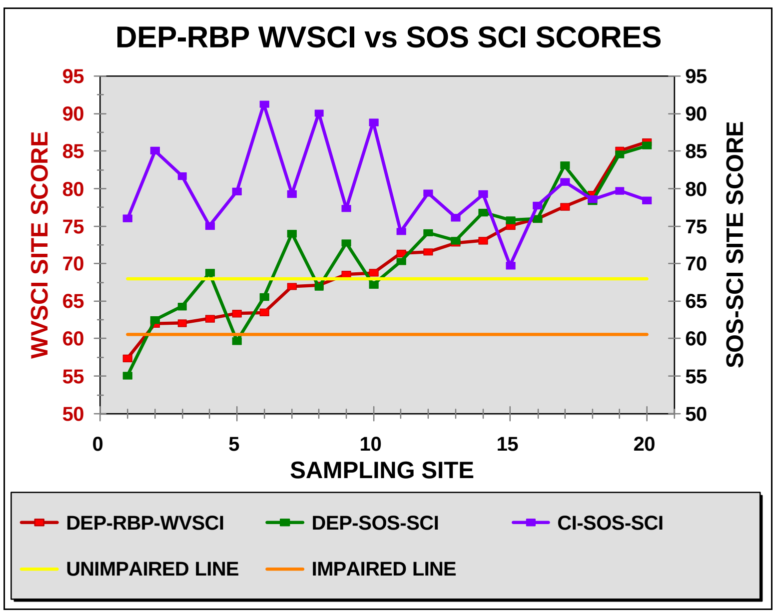
We also compared SOS and WVSCI stream ratings site by site (figure at left). The CI-SOS and DEP-SOS stream quality ratings were equivalent to the WVSCI stream quality ratings 60% and 65% of the time and exceeded the WVSCI rating 40% and 35% of the time, respectively. In other words, stream quality was overestimated by the standard SOS scoring method 35-40% of the time. SOS ratings never underestimated the health of the stream. The DEP-SOS scores overrated sites nearly as often as CI-SOS, indicating that the difference between RBP and SOS results lies in the analysis of the data and not the method of collection. However, six out of eight sites that were rated uncertain or impaired by WVSCI were rated excellent by the CI-SOS, compared to only 2 out of eight DEP-SOS sites - an indication that method of collection might also impact the result.

In the figure at right, the 20 sampling sites are arranged from lowest to highest WVSCI score, with DEP-SOS and CI-SOS scores plotted around that line. No relationship between any of the three datasets is apparent, and no significant correlation was found between the WVSCI and either SOS stream score, even though the DEP-SOS score was calculated using the same original data as the WVSCI.



The family-level Hilsenhoff Biotic Index (HBI) used in the DEP-RBP analysis is similar in concept to the SOS Stream Score; it uses the sensitivity of different benthic organisms to organic pollution (on a scale of 0 to 10, with 0 being most sensitive) and their relative abundance to see if the total benthic community has more pollution sensitive or more pollution tolerant organisms. The higher the score, the more the community is dominated by pollution tolerant organisms. The Very Modified HBI (VMHBI) described above for SOS level data is the standard SOS score weighted using counts of each "taxon", with tolerance values reversed so the most pollution sensitive organisms are assigned a score of 1 and most tolerant a score of 3 . In the figure at right, the 20 sampling sites are arranged from lowest to highest DEP-RBP-HBI score, with DEP-SOS-VMHBI and CI-SOS-VMHBI scores plotted around that line. A significant correlation (r=0.57) was detected between these DEP-SOS and DEP-RBP HBI indices, although considerable variability between the two is still apparent. Neither were correlated with the CI-SOS VMHBI index.





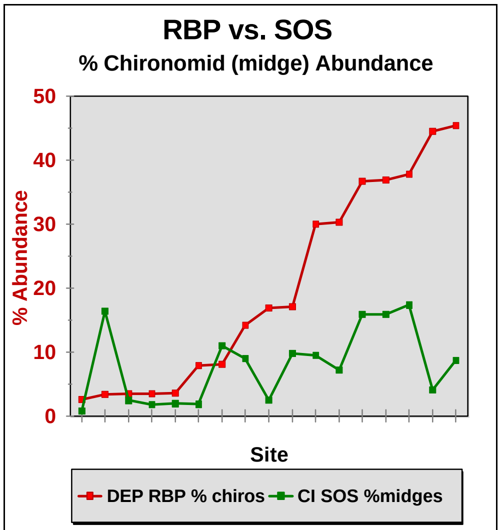
The SOS-SCI rating was developed to create a composite stream index using SOS level identifications that functions in the same manner as the WVSCI. The figure at left compares the WVSCI to SOS-SCI stream scores. The DEP-SOS-SCI has a strong, positive correlation (r=0.94) with RBP-WVSCI scores, and the graph shows a much tighter fit between these two indices than was seen in the HBI comparison. An 80% agreement between DEP-SOS-SCI and WVSCI in stream rating was achieved, with 2 out of 20 sites rated higher, and two out of twenty sites rated lower. The CI-SOS-SCI scores were not correlated with either of the above scores, and tended to rate those sites highest that were rated lowest by the other two indices.

Calculations based on DEP raw data at the SOS identification level had significant and positive correlations with RBP metrics and WVSCI scores. It therefore appears that SOS level identifications and the modified metrics that were developed by CI such as VMHBI and SOS-SCI have merit. Note, however, that four of the six component indices in the DEP-SOS-SCI were identical to the metrics in the WVSCI (total taxa, EPT taxa, %EPT and %chironomids/midges); we discuss this below.

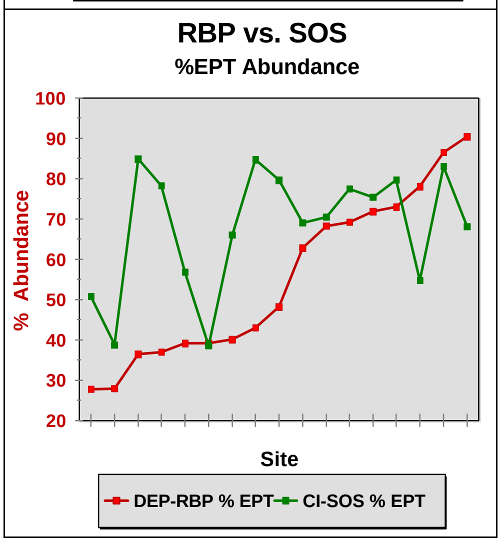
Why did data collected using SOS protocols compare poorly to data collected using RBPII protocols? One of the reasons may be that benthic macroinvertebrate distributions are quite variable within any given riffle (Hynes, 1970), and some of the difference could be caused by necessarily sampling different locations at each site. However, when comparing the proportions of different orders of organisms collected using SOS and RBP methods, the RBP samples often had higher proportions of smaller, cryptic organisms, especially chironomids/midges and blackfly larvae.

This indicates that differences in collection method may contribute to differences in indices and stream ratings, a result that differs with the results reported by Virginia Polytechnic Institute (see page 3). We expect that much of the difference is due to either the mesh size of the nets used by the different methods (SOS method uses a larger mesh size), or in the difference between field picking live samples (SOS) vs. laboratory picking of preserved samples (RBPII). Therefore, we conducted a preliminary comparison of the 2 different collection and picking methods. The goal was to compare the catches in the 2 different types of nets, and to compare field picking of live organisms with lab-picking of preserved organisms.

Preliminary results indicate that the main differences were between careful field picking of live organisms and lab-picking preserved samples, regardless of the net used. In particular, lab-picking of preserved samples under 5x magnification recovered a disproportionately large number of midges compared to other abundant taxa, despite the fact that most midges were brightly colored. The addition of lab picking resulted in large differences in important indices: %chironomids/midges (13% increased to 23%), % EPT abundance (83% dropped to 72%) and VMHBI (1.32 increased to 1.53). We will conduct further studies to test this effect more thoroughly. However, a review of this project's data (figures at right) supports the conclusion that SOS collection and "picking" methods may be disproportionately missing small organisms that have a big impact on overall stream rating when counts of organisms are used.



This is the case because caddisflies, mayflies, and midge larvae were typically the most abundant groups in both RBPII and SOS samples. The pollution sensitive caddisflies, mayflies and stoneflies make up % EPT abundance, which increases with better stream quality. Small, more cryptic chironomid/midge larvae are considered pollution tolerant and make up the % Chironomidae/midge index, which increases as stream quality decreases.



As the relative abundance of chironomids increased in RBP samples, the %chironomids/midges recovered from SOS samples did not and the difference between the two became increasingly large. The opposite was true of the EPTs; as their relative abundance decreased in RBP samples, the %EPT recovered from SOS samples did not and the difference between the two became increasingly large. Because the taxa used in these two induces were generally the most abundant in a sample, when %chironomids/midges is undervalued due to field picking, then %EPT will usually be overvalued, therefore inflating the estimate of stream health. This is the pattern we see.

If volunteer samples are collected and preserved in the field and later picked under magnification and counted, we expect that much of the difference noted above between the CI-SOS results and the other two data sets will disappear. The relative percentage of EPT fauna and of chironimids/midges should be nearly the same, as nearly as can be expected when sampling communities of benthic macroinvertebrates. Still to be determined is the relationship between the "number of kinds" measures of richness in our proposed modification to SOS methods and the "number of families" in RBP. However, summary

A comparison of summary statistics on measures of richness in CI-SOS ("number of kinds") and RBP ("number of families").

	RBP Total Taxa (Family)	CI-SOS Total Richness ("Kinds")	RBP EPT Taxa (Family)	CI-SOS EPT Richness ("kinds")
Mean	16.8	19.5	8.2	9.1
Std. Deviation	3.5	3.3	1.4	1.3
Median	16.0	20.0	8.0	9
Min	11	15	6	7
Max	24	26	11	12

statistics (see table at left) on total richness and EPT richness metrics for the two collection methods are quite similar and indicate that the "kinds" metric may prove valid.

Based on the above observations, we think that….

SOS Stream Scores as currently calculated don't provide stream assessment data that are consistently comparable to professional RBPII results because they lack relative abundance data and thereby lack critical information. As the Virginia Polytechnic Institute study found, SOS scores tend to overrate sites when compared to a professional stream assessment. However, the SOS method provides a conservative stream assessment tool – although it may miss some impaired streams, if the SOS score says it's bad, it is.

The identification level used with the SOS method can provide a stream assessment comparable to professional methods, based on our analysis. The main drawback, is that SOS simply categorizes the organisms using A = 1-9, B = 10-99 and C = 100 or more, and these "count" categories are then not used in the calculation of the stream score. Our study indicates that actual counts of the organisms collected are essential in order to properly weight the importance of each organism. These counts allow calculation of variations of the same biological indices used in professional stream assessments (RBP II), and also calculation of an overall stream index similar to DEP's West Virginia Stream Condition Index (WVSCI). The indices can be easily calculated using a computer spreadsheet, developed by Cacapon Institute, where volunteers enter counts and the number of different "kinds" and the computer does the rest.

The SOS method of fieldpicking live organisms appears to disproportionately miss some small organisms in specific groups, like midges and blackflies, when compared to picking preserved samples in the lab. This difference appears to be of paramount importance. We believe it is unlikely that field picked samples analyzed under either of the above methods (modified SOS or professional RBP) will produce results consistent with laboratory picked samples because certain critical groups are likely to be seriously under-represented in field picked samples.

Suggested modifications to WVSOS methodology to improve acceptance of volunteer data:

1. Count the organisms collected.

2. Samples should be preserved in alcohol in the field and picked in the lab (or on the kitchen counter) under slight magnification with good lighting. This will eliminate disproportionate "under-picking" of important groups such as midges and blackflies and will reduce the overestimation inherent in SOS methods. While this is a separate step for volunteers, it is actually easier (and more comfortable) than the current field picking - and more rewarding because the results will have greater validity and, hopefully, achieve greater acceptance by government agencies.

3. Note visually different "kinds" of organisms in SOS level identification categories, to obtain a measure of species richness, or diversity.

4. Supply an automated spreadsheet for volunteers to enter raw data and generate a composite index (based on a number of indices) similar to the WV Stream Condition Index.

Conclusion

Based on our study, it is possible for non-professional, volunteer conducted, benthic stream assessments to obtain results that compare favorably to professional assessments. The proposed method utilizes the same level of identification skill currently required of SOS volunteers and the same collection technique. It differs by requiring samples to be preserved in the field for "picking" under slight magnification and good lighting at home, in counting the organisms obtained, and in noting different "kinds" within each of the current SOS identification categories.

Acknowledgments

This project would not have been possible without a grant from the WV Division of Natural Resources Non-Game Wildlife & Natural Heritage Research/Cooperative Projects Program. It also would not have been possible without close cooperation from the WV Division of Environmental Protection and willingness to share data collected during their Watershed Assessment Program in the Cacapon River watershed in the summer of 2000.

Bibliography

Barbour, M.T., J. Gerritsen, B.D. Snyder, and J. B. Stribling. 1999. Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates and Fish, Second Edition. EPA 841-B-99-002. U.S. Environmental Protection Agency; Office of Water; Washington, D.C.

Craddock, Tim. 2000. West Virginia Save Our Streams: Citizens Volunteer Stream Monitoring Program. WVDEP; Watershed Assessment Program.

Engel, S. R. 2001. The effectiveness of using volunteers for biological monitoring of streams. M.S. thesis. Department of Entomology, Virginia Polytechnic Institute and State University, Blacksburg, Virginia. 41 pages.

Gerritsen, Jeroen, June Burton, and Michael T. Barbour. 2000. A Stream Condition Index for West Virginia Wadeable Streams.

Hynes, H.B.N. "The Ecology of Running Waters". 1970. Liverpool University Press.

Save Our Streams Program, Izaak Walton League of America, Inc.

For more information, contact
Cacapon Institute, Rt. 1 Box 326, High View, WV 26808
pcrel@mountain.net, (304) 856-1385