This document was prepared in conjunction with work accomplished under Contract No. DE-AC09-96SR18500 with the U. S. Department of Energy.

WSRC-TR-2005-00524

Key Words:
Environment
Streams
Ecology
Biological Assessment
Invertebrates

Retention:
Permanent

APPLICATION OF THE SCDHEC MULTIPLE HABITAT SAMPLING PROTOCOL (MHSP) FOR MACROINVERTEBRATES IN SRS STREAMS

Michael H. Paller, F. Douglas Martin, Lynn D. Wike, and Winona L. Specht

Savannah River National Laboratory
Environmental Sciences and Technology Department
Environmental Analysis Section

October 2005

Westinghouse Savannah River Company
Savannah River Site
Aiken, SC 29808

Prepared for the U.S. Department of Energy Under
Contract Number DE-AC09-96SR18500

Table of Contents

List of Figures

List of Tables

EXECUTIVE SUMMARY

The MHSP is a SCDHEC bioassessment method that utilizes macroinvertebrate samples collected from natural substrates to evaluate the ecological health of South Carolina streams. It includes two metrics, the EPT (number of Ephemeroptera, Plecoptera, Trichoptera taxa) and the BI (a biotic index that reflects the relative pollution tolerances of individual taxa), each of which receives a score of 1 to 5 that is averaged to produce an overall bioclassification score ranging from 5=Excellent to 1=Poor. The MHSP score for many undisturbed, high quality streams on the SRS is only fair or poor. We suggest three possible reasons for the failure of high quality SRS streams to receive high MHSP scores: 1) the MHSP may not measure attributes of the macroinvertebrate community that accurately reflect the health of SRS streams, 2) the MHSP does not compensate for natural differences in the invertebrate fauna that are associated with differences in stream size, and 3) the scoring criteria used in the MHSP may not be appropriate for SRS streams.

Statistical analysis (simple correlations and multiple regression) showed that the BI and EPT were good predictors of disturbance in SRS streams, and that the addition of other metrics (total number of invertebrate taxa, percent clinger, and percent Ephemeroptera) did not improve the accuracy of the MHSP. These results demonstrate that the macroinvertebrate community attributes measured by the MHSP are good indicators of the health of SRS streams. However, linear regression models showed that both EPT and the BI were significantly affected by stream size, with the relationship being stronger for EPT. The influence of stream size can be eliminated by developing metric scores based on the residuals from the regression of the metrics on stream width (i.e., the deviation of the metrics from the expectation for a stream of given size) rather than the raw values of the metrics.

The potentially confounding effects of stream size on metric scoring did not fully account for the low MHSP scores received by many undisturbed SRS streams. A second problem with the MHSP is that the scoring criteria for EPT are set too high for SRS streams, possibly because the reference streams from which the criteria were developed support more EPT than do SRS streams. This type of problem can arise if there is substantial unrecognized heterogeneity among the sandhill/coastal plain streams from which the scoring criteria were developed. We corrected this problem by developing new scoring criteria from undisturbed SRS reference sites that support numbers of EPT taxa expected under natural conditions. This correction, together with the correction for stream size, resulted in substantial increases in the MHSP scores for many SRS streams and changed the average bioclassification for undisturbed streams from approximately “Good-Fair” to “Good”. A Receiver Operating Characteristic (ROC) curve was constructed to evaluate the ability of the adjusted MHSP to correctly identify disturbed and undisturbed sites. The area beneath the curve was about 91%, which is indicative of an excellent classification method.

1.0 INTRODUCTION

SCDHEC recently developed an environmental assessment methodology called the Multiple Habitat Sampling Protocol (MHSP) that is being used to evaluate the health of South Carolina streams (SCDHEC 1998). The protocol evaluates stream health based on the diversity and species composition of macroinvertebrates collected from natural substrates in the stream. Application of the MHSP to SRS streams has demonstrated that it produces unexpectedly low ratings for a number of streams that have been shown to be high quality in previous studies. An example is Upper Three Runs, which is considered an outstanding resource that supports an unusually high diversity of aquatic insects (based on studies by Dr. John Morse at Clemson University and others). No sites in Upper Three Runs received MHSP bioclassification ratings of Excellent and some received ratings of Fair, the latter of which is considered only partially supporting of aquatic life. Other unpolluted and largely or completely undisturbed SRS streams also received ratings of Fair or even Poor. These inaccuracies can have negative consequences for the SRS NPDES program and other environmental initiatives because they falsely indicate environmental degradation.

The MHSP is an example of a multimetric bioassessment index. Multimetric indices are composed of two or more community, population, or organism level variables (i.e., metrics) that are ecologically important and sensitive to environmental disturbances of various types. These variables are measured at assessment sites, compared with the same variables in a range of similar but undisturbed benchmark streams, and the results summarized in a single number that reflects the extent to which the assessment sites resemble the benchmarks. The MHSP includes two metrics: the EPT, which is the number of Ephemeroptera, Plecoptera, Trichoptera taxa and the BI, which is a biotic index that reflects the relative pollution tolerances of individual taxa (Lenat, 1993). Each of these metrics is assigned a score of 1 to 5 for each site (SCDHEC 1998, Table 1), and the two scores are averaged to produce an overall bioclassification score of 5=Excellent, 4=Good, 3=Good-Fair, 2= Fair, or 1=Poor. Excellent and Good indicate that biological assemblages are not significantly modified from a natural state, Good-Fair and Fair indicate moderate modification compared with reference conditions, and Poor indicates severe modification.

There are three possible reasons why the MHSP produces inaccurate scores in SRS streams:

1. The MHSP does not compensate for natural differences in the invertebrate fauna that are associated with differences in stream size. Such differences are predicted by the River Continuum Concept (RCC), an important conceptual framework for understanding changes in biotic communities with progression from headwater to higher order stream reaches (Vannote et al. 1980). The RCC predicts that the taxonomic richness of benthic communities increases with stream size, reaching a maximum in mid-order streams (Minshall et al. 1985). The consequence for the MHSP is that naturally occurring stream size related changes in EPT can be confounded with changes in EPT resulting from environmental degradation.
2. The MHSP may not measure attributes of the macroinvertebrate community that accurately reflect the health of SRS streams. Although the two metrics in the MHSP (EPT and BI) have proven bioassessment value (Plafkin et al. 1989), most multimetric indices use more than two metrics to accurately assess stream health. A recent study in the mid-Atlantic coastal plain states resulted in the development of a multimetric index with five metrics (total number of taxa, number of EPT, percentage

Table 1. Scores for EPT Taxa Richness and BI values for Sandhills/Coastal Plain Streams (from SCDHEC, 1998)

Score[a]	**Biotic Index Values**	**Number of EPT Taxa**
5	<5.42	>29
4.6	5.42 - 5.46	28
4.4	5.47 - 5.51	27
4	5.52 - 6.00	22 - 26
3.6	6.01 - 6.05	21
3.4	6.06 - 6.10	20
3	6.11 - 6.67	15 - 19
2.6	6.68 - 6.72	14
2.4	6.73 - 6.77	13
2	6.78 - 7.68	8 - 12
1.6	7.69 - 7.73	7
1.4	7.74 - 7.79	6
1	>7.79	0 - 5

[a] 5=excellent, 4=good, 3=good-fair, 2=fair, 1=poor = 1

of Ephemeroptera, Hilsenhoff Biotic Index, and percentage of taxa with a clinger mode of existence) that effectively identified impaired sites (Maxted et al. 2000). Inclusion of more or different metrics in the MHSP might result in more accurate bioclassification scores for SRS streams.

3. The scoring criteria used in the MHSP may not be accurate for SRS streams. Scoring criteria are developed by comparing impaired sites to undisturbed reference sites, with the magnitude of the score generally being inversely proportional to the deviation from the undisturbed condition. Thus, scoring is strongly dependent upon choosing appropriate reference sites that are comparable with potentially impaired sites in all respects except the occurrence of disturbance. This results in the common practice of using different scoring criteria for different ecological regions that support different types of macroinvertebrate communities. The MHSP incorporates different scoring criteria for Mountain, Piedment, and Sandhill/Coastal Plain ecoregions, the last of which includes SRS streams. However, it is possible that the reference site standards for the Sandhill/Coastal plain ecoregion do not adequately represent SRS streams.

The objective of this study was to determine why the MHSP does not produce accurate scores for SRS streams and to identify methods for improving its accuracy in SRS streams. Specific objectives were to determine the consequences of stream size related changes in macroinvertebrate community structure on the performance of the MHSP, assess whether the inclusion of additional metrics can improve the performance of the MHSP, and determine whether the MHSP scoring criteria for Sandhill/Coastal Plain streams are accurate for streams on the SRS.

2.0 MATERIALS AND METHODS

2.1 STUDY AREA AND FIELD METHODS

This study included 27 sites in 12 first through fourth order streams on the Savannah River Site (SRS), a 780 km^2 Department of Energy (DOE) reservation in the Sandhill ecoregion of South Carolina. Sand was the predominant substrate in most streams, and woody debris (e.g., snags, logs, twigs, and leaves) constituted most of the instream structure along with overhanging shoreline vegetation, undercut banks, root masses and aquatic plants (in larger streams). The streams were about 2-16 m wide, 0.6-2.5m/km in average gradient, and mildly acidic (pH 4.5-6.9) with relatively low conductivity (11-104 uS/cm). Some of the streams were largely undisturbed with little or no agriculture, urbanization, or industrialization in their watersheds. Others receive or received discharges from industrial facilities or seepage basins (where radioactive or other toxic materials were deposited) or were altered by changes in flow regime, construction activities, or prior discharge of heated nuclear reactor cooling water. Because of environmental remediation programs and the shutdown of the SRS nuclear reactors in the 1980's, conditions were improving in most disturbed SRS streams during this study.

Sixteen sites were sampled in 1997, 18 in 2000, and 22 in 2003; 12 were sampled in all three years. Sampling was conducted in the fall when most insect larvae were comparatively large prior to emergence. The SCDHEC (1998) MHSP was employed in 2003, in which all available natural habitats were qualitatively sampled with a D-frame dip net, kick net, hand sieve, white plastic pan and fine mesh sampler with the objective of collecting as many different macroinvertebrate taxa as possible during the allotted three man-hours of effort at each site. The last 1.5 hours were directed mainly towards collection of additional species rather than collecting more individuals of species already in the collections. The MHSP protocol is designed to ensure that all the habitats at a site are thoroughly sampled to obtain a good representation of the macroinvertebrate community. Generally similar methods were used in 1997 and 2002 except that sampling was conducted for two man-hours rather than three. Macroinvertebrates were placed in jars or vials containing 85% ethanol; labeled with the station number, collector, and collection date; and returned to the laboratory for microscopic identification to the lowest practical taxonomic level (generally genus or species).

Stream width at the water surface was measured at each sample site before or after macroinvertebrate sampling at seven to 12 evenly spaced transects across the stream perpendicular to the direction of water flow. Habitat quality was assessed using a methodology developed by SCDHEC (1998). Variables included epifaunal substrate, pool substrate, pool variability, sediment deposition, channel flow status, channel alteration, channel sinuosity, bank stability, vegetative protection, and riparian vegetation.

2.2 DATA ANALYSIS

To calculate the MHSP, the number of EPT (Ephemeroptera, Plecoptera, Trichoptera) taxa were tallied, and a biotic index (Lenat, 1993) was calculated for each sampling site by averaging the tolerances of the individual taxa. The EPT and BI metrics were assigned a score of 1 to 5 (Table 1), and the two scores were averaged to produce a combined score that determined the final bioclassification. When the final score fell between two bioclassification ratings, it was either rounded up or down based on whether the decimal fraction is larger or smaller than 0.5 following SCDHEC protocol (SCDHEC 1998). The MHSP was calculated only with 2003 data because the MHSP field protocol was followed in detail only in 2003. However, data from all three years were used in the other analyses described below.

Data analysis emphasized 1) the effects of stream size on bioassessment metrics, 2) the value of adding metrics to the MHSP to improve its ability to distinguish biological degradation, and 3) the adequacy of the MHSP scoring criteria. The value of adding metrics to the MHSP was evaluated with correlation coefficients and multiple regression analysis. The metrics evaluated were EPT and the BI, which are already included in the MHSP plus total number of taxa, percent clinger, and percent Ephemeroptera (three additional metrics found to be good discriminators of degraded conditions in mid-Atlantic coastal plain streams, Maxted et al. 2000). Spearman correlation coefficients were used to assess the relationships among metrics and between metrics and disturbance levels (described below). Multiple regression was used to determine if the addition of metrics could significantly increase the ability of the MHSP to predict disturbance level. In cases where “stepwise regression” was used, P for variables entering and exiting the model was set at 0.05. Stream size related effects on EPT and the BI were investigated using simple linear regression models with metric values at each site as dependent variables and stream widths at each site as the independent variable. Regression residuals were calculated using standard methods (Sokal and Rohlf 1995). The suitability of MHSP scoring criteria were evaluated by comparing existing and modified criteria with independently assessed levels of disturbance in each stream.

For the preceding analyses, each stream site was assigned an ordinal disturbance level ranging from 0 (undisturbed) to 4 (maximum disturbance) based on the history of the site, its nearness to industrial areas and outfalls, habitat evaluations, and the results of fish based bioassessments (IBI)(Paller and Dyer 2003). The use of ordinal dependent variables in multiple regression analysis is technically improper but is commonly done and generally yields acceptable results when the number of ordinal classes is five or more. (Berry 1993). The largest impact is on the accuracy of significance tests, which can also be compromised by the testing of numerous regression models in stepwise regression and by the assumption that sites sampled more than once are independent replicates. Therefore, P values generated by the regression analyses should be regarded as rough approximations only.

The ability of the MHSP to accurately classify sites as impaired or unimpaired was summarized by constructing a Receiver Operating Characteristic (ROC) curve. ROC curves are commonly used in the biomedical field to assess the discriminatory power of diagnostic tests (DeCarlo 1998, Motulsky 2005). They express the relationship between sensitivity (in this case the fraction of disturbed sample sites that are correctly identified as disturbed by a low MHSP score) and specificity (the fraction of undisturbed sample sites that are correctly identified as undisturbed by a high MHSP score). The ROC curve was constructed by plotting the true positive rate (sensitivity) against the false positive rate (1-specificity) for each value of the MHSP. The area under an ROC curve ranges from 0.5 to 1.0 and quantifies the effectiveness of a diagnostic test; 0.5 indicates a worthless test and 1.0 indicates a test that perfectly discriminates disturbed from undisturbed sites. The significance of the difference between the area under the curve and 0.5 was tested with a nonparametric method described in Motulsky (2005).

3.0 RESULTS AND DISCUSSION

3.1 EFFECTS OF STREAM SIZE ON BIOASSESSMENT METRICS

Previous work (Paller et al. in review) indicates that taxa richness metrics (e.g. total taxa richness and EPT) increase with stream width within the range of stream sizes (first through fourth order) found on the SRS, and that the slope of the regression line describing this relationship is similar for disturbed and undisturbed streams. As a result, taxa richness metrics should be adjusted for stream size to avoid confounding stream size related changes with changes resulting from environmental degradation. There are several ways to make such adjustments as illustrated with EPT data collected from SRS streams. An approach used in Specht and Paller (2004) is to draw a “maximum species richness line” (Fausch et al. 1984) through the highest points on a plot of number of EPT taxa versus stream size and divide the area below the line into intervals of equal size corresponding to different bioclassification scores (Figure 1). A second approach is to develop a regession model relating EPT to stream size (Figure 2) and compute the residuals for each site. The residuals correspond to the EPT values for each site with the effects of stream size removed. Positive values indicate more EPT than expected for a stream of given size and negative values indicate fewer. In the example shown in Figure 2, Lower Three Runs at Patterson Mill (Ltrp) had a negative residual (measured as the vertical distance from the regression line) indicating the presence of fewer species than expected. Uncorrected for stream size, this site possessed more than the average number of species and could be considered ecologically healthy. If the residual method is used, a remaining step in the development of scoring criteria is to arrange the residuals into scoring classes, as illustrated in Figure 3, which was constructed by separating the residuals into equally spaced intervals after excluding the upper 5% of the values (following Plafkin et al. 1989).

Analysis of the metrics included in this study indicated that EPT, taxa richness and %Clingers were directly related to stream width (R^2=0.46 [P<0.001], R^2=0.38 [P<0.001] and R^2=0.42 [P<0.001], respectively), and %Ephemeroptera was unrelated to stream width (R^2=0.06, P=0.076). The BI was inversely related to stream width, but the relationship was weaker than for EPT, total taxa, and %Clingers (R^2=0.18, P=0.001). The occurrence of an inverse relationship indicated that average organism tolerance increased as stream size decreased because the BI increases with organism tolerance (Lenat 1993). As a result, negative BI residuals receive higher bioclassification scores and positive BI residuals receive lower scores (Figure 4).

3.2 ADDITION OF METRICS TO THE MHSP

Metrics evaluated for addition to the MHSP included total number of taxa, % Ephemeroptera, and % Clingers; all of which were shown to effectively discriminate disturbed from reference sites in mid-Atlantic coastal plain states (Maxted et al. 2000). Spearman correlations between these variables and disturbance classes for SRS stream sites were substantial for number of taxa (-0.59) and % Clingers (-0.58), but relatively low (-0.34) for % Ephemeroptera (Table 2). The highest correlations with disturbance class were exhibited by the BI (0.67) followed by EPT (-0.60), both of which are included

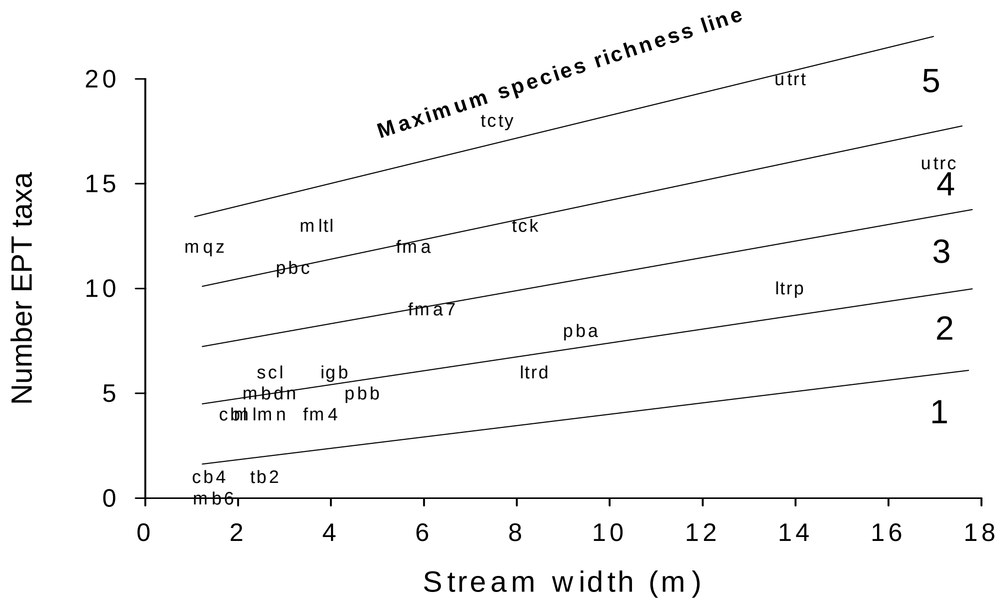

Figure 1. Relationship between number of EPT taxa and stream size. The area beneath the maximum species richness line has been subdivided to produce five bioclassification ratings.

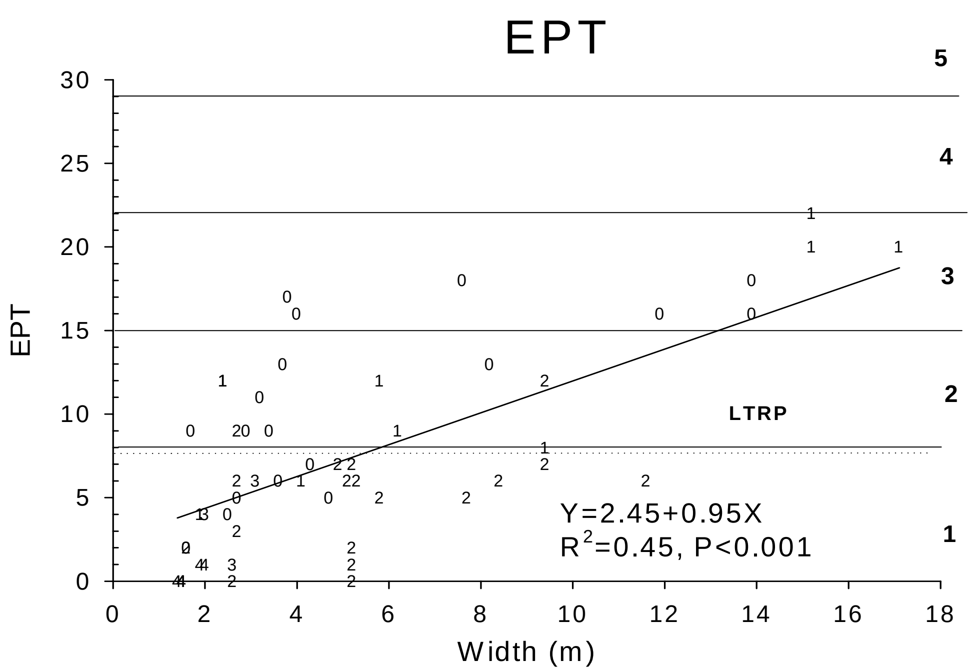

Figure 2. Regression of EPT on stream width. Sample site symbols indicate level of disturbance (0=undisturbed to 4=highly disturbed). Numbers 1-5 on the left indicate SCDHEC MHSP scoring intervals for EPT. Dotted line represents the average EPT for all sites. “LTRP” = EPT at Lower Three Runs, Patterson Mill.

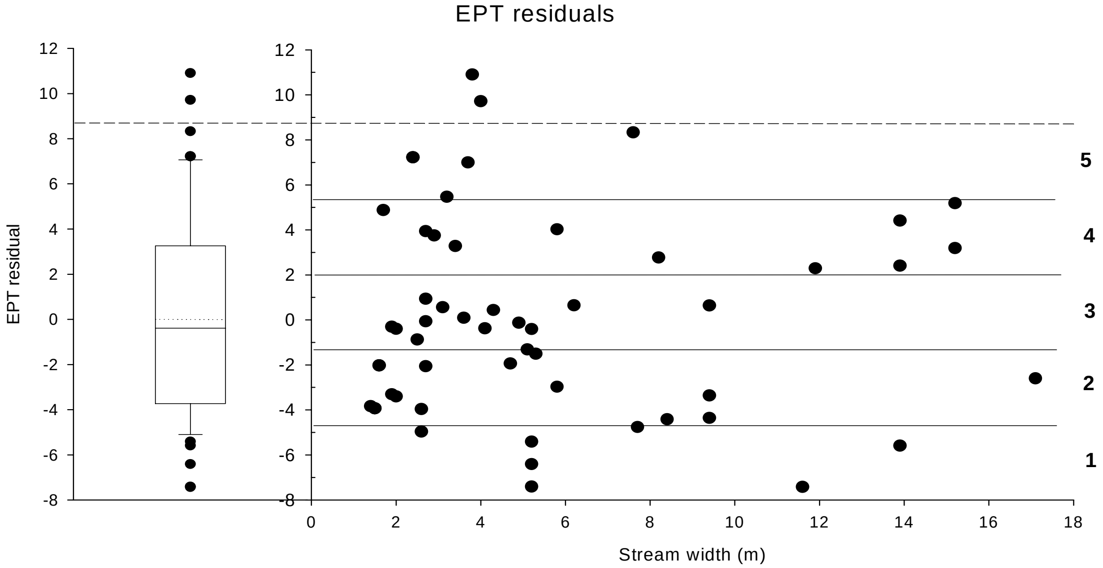

Figure 3. Distribution of EPT residuals divided into five scoring intervals after excluding the upper 5% (shown by dotted line). EPT residuals were derived from a regression of sample sites on stream width (shown in Figure 2). Higher residuals indicate greater than expected EPT and receive higher scores.

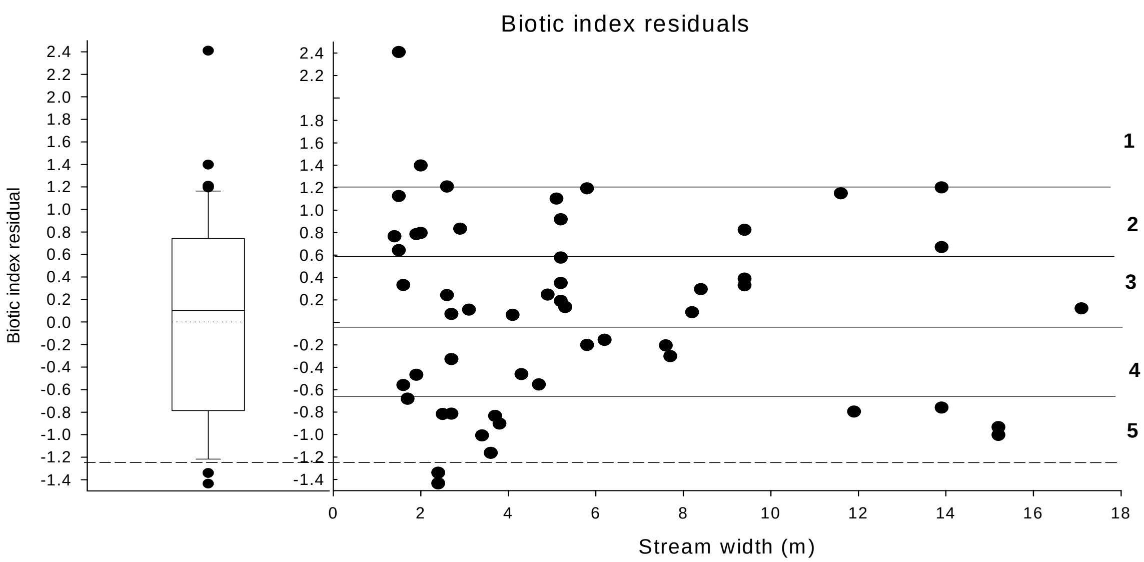

Figure 4. Distribution of BI residuals divided into five scoring intervals after excluding the lower 5% (shown by the dotted line). BI residuals were derived from a regression of sample sites on stream width. Lower BI residuals indicate a greater than expected BI and receive higher scores.

Table 2. Spearman correlation coefficients among bioassessment metrics, stream width, and stream disturbance level. Disturbance level was expressed on a scale from 0 (least disturbed) to 4 (most disturbed).

	BI	TAXA	EPT	%EPHEM	%CLING	WIDTH	DISTURB
BI	1.00						
TAXA	-0.63	1.00					
EPT	-0.74	0.85	1.00				
%EPHEM	-0.51	0.51	0.62	1.00			
%CLING	-0.71	0.87	0.86	0.52	1.00		
WIDTH	-0.37	0.58	0.62	0.42	0.60	1.00	
DISTURB	0.69	-0.59	-0.63	-0.45	-0.60	-0.27	1.00

in the MHSP. Correlations were also high among some of the metrics, such as 0.85 between EPT and total taxa and 0.86 between % Clingers and EPT.

Backwards stepwise regression beginning with all five bioassessment metrics as independent variables and disturbance class as the dependent variable showed that only the BI and total number of taxa met the P=0.05 criterion for retention in the model. Inclusion of total taxa richness in the model rather than EPT may seem surprising since the simple correlation was slightly higher between EPT and disturbance (-0.60) than between total taxa and disturbance (-0.59). However, EPT and the BI shared more variation (r =-0.74) than did total taxa and the BI (r=-0.63), indicating that total taxa contributed more unique information to a model that included the BI (which was the best single predictor). R^2 for the model that included the BI and total taxa was 0.49. Inclusion of EPT, % Clingers, and %Ephemeroptera in the model only increased the R^2 to 0.51 indicating they did little to increase the ability to predict disturbance level. When evaluating these R^2 values, it is important to consider that some of the unexplained variance likely reflects inaccuracy in the assignment of disturbance classes rather than inability of the metrics to accurately measure actual disturbance.

The simple correlation between EPT and total taxa was 0.85 (Table 2), indicating that these two variables contained similar information. As a result, a regression model with BI and EPT as independent variables explained almost as much variance in disturbance level (R^2=0.47) as a regression model with BI and total taxa as independent variables (R^2=0.49). In other words, EPT and total taxa can be used interchangeably with little change in predictive power. An additional consequence of the high correlation between EPT and total taxa is that inclusion of taxa richness in a regression model that already contains EPT and the BI will have only a small effect on the predictive power of the model (R^2 of 0.50 for BI, EPT, and total taxa compared with 0.47 for BI and EPT only). In summary, these results indicate that the ability of the MHSP (which includes BI and EPT) to predict disturbance is not significantly increased by the inclusion of the % Ephemeroptera or % Clinger. They also show that, although total taxa is a slightly better predictor than EPT, its inclusion in the MHSP will effect only a minor improvement in the accuracy of the index.

3.3 APPROPRIATENESS OF MHSP SCORING CRITERIA

In the MHSP, different values of EPT and the BI are assigned bioclassification scores of one through five that are averaged to produce a final bioclassification rating (shown in Table 1). Thus, the criteria for assigning the scores have a critical influence on the rating of the stream. Examination of the bioclassification scores for EPT indicates that no SRS streams attain the highest scores of 4 or 5 based on the existing criteria for assigning scores (Figure 2) despite the fact that many of them (Tinker Creek, Upper Three Runs, Meyers Branch, and Mill Creek) receive no effluents and have largely undisturbed watersheds. Some of these undisturbed or largely undisturbed sites receive EPT scores as low as 1, which indicates "Poor" quality and a "Not Supporting" Aquatic Life Use Support (ALUS) rating. "Not Supporting" is defined as "severe modification of the biological community compared to the reference condition" (SCDHEC 1998), which is clearly at variance with the undisturbed state of these sites.

One reason for the failure of the scoring criteria for EPT to adequately reflect the condition of SRS streams is that they do not adjust for the effects of stream size as previously mentioned. However, this does not entirely account for the problem because even larger undisturbed streams such as Upper Three Runs receive scores no higher than three. An alternative explanation is that the EPT criteria for sandhill/coastal plain streams are inappropriate for SRS streams because the reference sites from which they were developed support more EPT taxa than SRS streams. This type of problem could arise if there is substantial unrecognized heterogeneity among sandhill/coastal plain streams. A possible source of such heterogeneity is substrate type, which can have a strong influence on benthic species composition and diversity. SRS streams have sandy substrates, which typically support fewer EPT taxa than gravel, cobbles, or boulders (Jowett et al. 1991). Other possible sources of heterogeneity include stream size (as previously discussed) and amount and type of macrophyte growth.

Assuming that the EPT scoring criteria for SRS streams are inappropriate, it is possible to develop new criteria from reference sites that support numbers of EPT taxa expected under undisturbed conditions in SRS streams. Such criteria can be developed with data from the SRS because the SRS supports (in addition to disturbed streams) a substantial number of undisturbed streams and stream reaches (although ideally sites outside the SRS would also be included). Metric scoring is commonly based on the distribution of values in a population of comparable sites (i.e., ecologically similar except for the occurrence of disturbance) that includes both disturbed and reference streams (Plafkin et al. 1989). For metrics that decrease with disturbance, the upper 5% of the values are generally eliminated to exclude outliers and the remaining values are trisected, quadrisected, etc. to provide a range of scores (although other methods are also possible, Plafkin et al. 1989). Figure 2 demonstrates the use of this method with EPT residuals. Figure 2 has the combined effect of controlling for the effects of stream size as well as adjusting the scoring criteria for numbers of EPT taxa characteristic of undisturbed SRS streams.

Table 3 shows unadjusted MHSP scores and adjusted MHSP scores for data collected from SRS streams in 2003. Unadjusted scores were calculated using the scoring criteria used by SCDHEC (1998) (Table 1), and adjusted scores were calculated by correcting for stream size and using undisturbed SRS reference sites to establish scoring standards (Figures 3 and 4). Streams were divided into disturbed and relatively undisturbed categories for this comparison, with relatively undisturbed including disturbance categories 0 and 1 and disturbed including disturbance categories 2, 3, and 4. The adjusted EPT scores were substantially higher than the unadjusted EPT scores,

Table 3. Unadjusted and adjusted MHSP scores for SRS streams. Scores are adjusted for stream size and for SRS reference sites.

		MHSP			Adjusted MHSP		
Stream	Location	EPT	BI	Avg.	EPT	BI	Avg.
Undisturbed Streams							
Mcqueens Branch	near Z area	2	5	3.5	5	5	5.0
Mill Creek	Munroe Rd.	1	4	2.5	3	5	4.0
Meyers Branch	Dunbarton Rd.	1	2.4	1.7	3	3	3.0
Pen Branch	Rd. C	2	5	3.5	5	5	5.0
Mill Creek	Telephone cable Rd.	2.4	4	3.2	5	5	5.0
Indian Grave Branch	K cooling tower	1.4	3	2.2	3	3	3.0
Pen Branch	Rd. B	1	4	2.5	2	4	3.0
Fourmile Branch	Rd. A	2	3	2.5	4	4	4.0
Fourmile Branch	Rd. A7	2	3	2.5	3	4	3.5
Tinker Creek	Tyler bridge Rd.	3	4	3.5	5	4	4.5
Tinker Creek	Kennedy pond Rd.	2.4	3	2.7	4	3	3.5
Pen Branch	Rd. A	2	3	2.5	2	3	2.5
Upper Three Runs	Rd. C	3	5	4.0	3	4	3.5
Upper Three Runs	Tyler Bridge Rd.	3.4	3	3.2	3	2	2.5
Average		2.0	3.7	2.9	3.6	3.9	3.7
Disturbed Streams							
Meyers Branch	Rd. B6.2	1	1	1.0	2	2	2.0
Crouch Branch	Rd. 4	1	2	1.5	2	2	2.0
Lower Crouch Branch	lower reach	1	3	2.0	3	4	3.5
Fourmile Branch	Rd. 4	1	2	1.5	3	2	2.5
Tims Branch	Rd. 2	1	1	1.0	2	1	1.5
Steel Creek	Rd. C	1.4	4	2.7	1	5	3.0
Lower Three Runs	Donora Station	1.4	3	2.2	2	3	2.5
Lower Three Runs	Patterson Mill	2	2.4	2.2	1	2	1.5
Average		1.2	2.3	1.8	2.0	2.6	2.3

especially for the undisturbed streams. Prior to adjustment, the average EPT score for undisturbed streams was 2.0, which is equal to a bioclassification of “Fair”. Following adjustment, EPT increased to an average of 3.6, which falls between “Good-Fair” and “Good”. In contrast, adjustment resulted in relatively small changes in the BI (3.7 versus 3.9 for undisturbed streams), which had original scoring criteria less strongly affected by stream size and more suitable for SRS streams. The adjustments resulted in an average increase in the MHSP of 0.8 points for undisturbed streams (2.9 versus 3.7) and 0.5 for disturbed streams (1.8 versus 2.3), with most of the increase resulting from the change in EPT.

An ROC curve was constructed to summarize the sensitivity and specificity of the adjusted MHSP, with sensitivity being the proportion of impaired sites identified as impaired and specificity being the proportion of unimpaired sites identified as unimpaired. Sites in disturbance classes 0 and 1 were defined as impaired and sites in disturbance classes 2, 3, and 4 were defined as unimpaired. The area beneath the ROC was 91.1% which was significantly different ($P<0.0001$) from the 50% characteristic of a random classifier. The value, 91.1%, can be interpreted as the probability that a randomly selected impaired site will have a lower MHSP score than a randomly selected unimpaired site. Generally, a classification test is rated as excellent when the area under the curve exceeds 0.90 (Motulsky 2005).

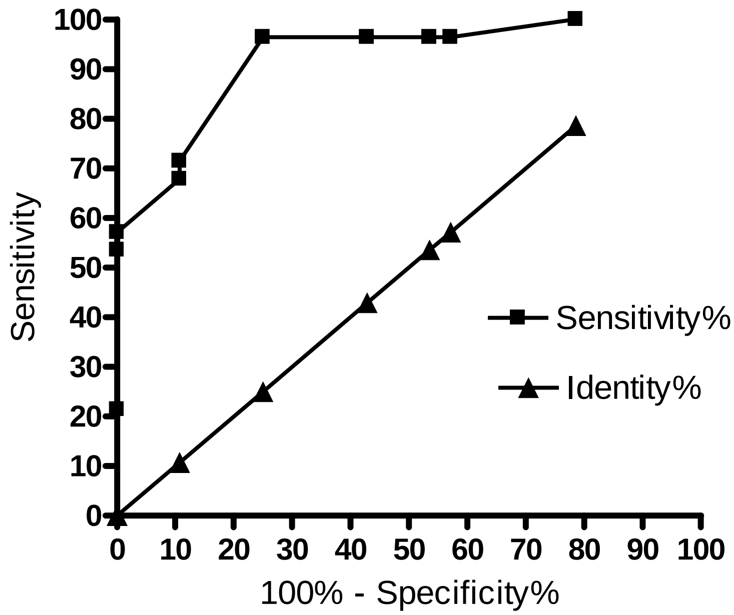

Figure 5. ROC curve for the MHSP with adjusted scoring criteria. The straight line represents a random classifier.

4.0 REFERENCES

Berry, W. 1993. Understanding regression assumptions. Series: Quantitative applications in the social sciences, no. 92. Sage Publications, Thousand Oaks, CA.

DeCarlo, L.T. 1998. Signal detection theory and generalized linear models. Psychological Methods 3:186-205.

Fausch, K. D., J. R. Karr, and P. R. Yant. 1984. Regional application of an index of biotic integrity based on stream fish communities. Trans. Am. Fish. Soc. 113: 39-55.

Jowett, I.G., J. Richardson, B.J.F. Biggs, C.W. Hickey and J.M. Quinn. 1991. Microhabitat preferences of benthic invertebrates and the development of generalized Deleatidium spp. habitat suitability curves, applied to four New Zealand rivers. New Zealand J. Mar. Freshw. Res. 25:187-199.

Lenat, D.R. 1993. A biotic index for the southeastern United States: Derivation and list of tolerance values, with criteria for assigning water-quality ratings. Journal of North American Benthological Society 12:279-290.

Maxted, J.R., M.T. Barbour, J. Gerritsen, V. Poretti, N. Primrose, A. Silvia, D. Penrose, and R. Renfrow. 2000. Assessment framework for mid-Atlantic coastal plain streams using benthic macroinvetertebrates. Journal of North American Benthological Society 19:128-144.

Minshall, G.W., R.C. Petersen, Jr. & C.F. Nimz, 1985. Species richness in streams of different size from the same drainage basin. The American Naturalist 125: 16-38.

Motulsky, H. 2005. GraphPad Prism, Version 4.0, Statistics Guide, Statistical analyses for laboratory and clinical researchers. GraphPad Software Inc., San Diego.

Paller, M.H., and S.A. Dyer. 2004. Biotic integrity of streams in the Savannah River Site integrator operable units, 1996 to 2003. WSRC-TR-2004-00562, Savannah River National Laboratory, Aiken, SC.

Paller, M.H., W.L. Specht, and S.A. Dyer. Effects of stream size on taxa richness and other commonly used benthic bioassessment metrics. Submitted to Hydrobiologia

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. U.S. Environmental Protection Agency, EPA/444/4-89-0001. Washington, D.C.

SCDHEC (South Carolina Department of Health and Environmental Control), 1998. Standard operating and quality control procedures for macroinvertebrate sampling. Technical Report No. 004-98. Bureau of Water; Division of water Monitoring, Assessment, and Protection; Aquatic Biology Section .

Specht, W.L. and M.H. Paller. 2004. Macroinvertebrate assessments of 22 locations in SRS streams, in support of the Integrator Operable Unit Program, July-August 2003. WSRC-TR-2004-00482.

Sokal, R.R., and F.J. Rohlf. 1995. Biometry, Third Edition. W.H. Freeman and Company, NY.

Vannote, R.L., G.W. Minshall, K.W. Cummins, J.R. Sedell & C.E. Cushing, 1980. The river continuum concept. Canadian Journal of Fisheries and Aquatic Sciences 37: 130-137.