RAPID BIOASSESSMENT OF SEVEN STREAMS IN WEST SANDY CREEK WATERSHED, HENRY COUNTY, TENNESSEE UNDER DIFFERENT SAMPLING REGIMES

J. R. SCHILLER AND S. W. HAMILTON

Department of Biology and Center for Field Biology, Austin Peay State University, Clarksville, TN 37044

ABSTRACT—We compared seven streams in the West Sandy Creek Watershed based on their macroinvertebrate assemblages using the United States Environmental Protection Agency (EPA) Rapid Bioassessment Protocol III (RBP). Macroinvertebrates were collected with kick nets from submerged root mass habitat from all streams on six dates in spring and late summer, 1992–1995. Riffle, aquatic macrophytic plant beds, and snag/course particulate organic matter (S/CPOM) habitats also were sampled in those streams where these habitats occurred. We compared bioassessments obtained from four sampling regimes: (1) three semiquantitative root mass samples consisting of 100+ organism subsamples (100 picks), (2) the three 100 pick subsamples combined into a pooled sample (pooled samples), (3) the mean of the three 100 pick subsamples (sample means), and (4) the pooled samples combined with qualitative multiple habitat samples (multi-habitat samples). Metrics of community richness and structure were sensitive to sampling regimes that included more area and/or types of habitat. Metrics of community pollution tolerance and trophic organization were robust to sampling regime. Differences in bioassessment conclusions between sampling regimes for the same sample date and stream were compared. Bioassessments differed among 100 picks 53% of the time, between 100 picks and those from pooled samples 46% of the time, and between 100 picks and those for sample means 33% of the time. Bioassessments based on pooled samples differed with those from sample means 22% of the time, and differed with those from multi-habitat samples 31% of the time. Only bioassessments using 100 picks differed by more than one EPA RBP III Biological Condition Category (4.6% of the time). The differences among bioassessments using 100 picks are excessive and indicate that bioassessments of soft-bottomed streams using aquatic macroinvertebrates should utilize a minimum of three replicate samples, either combined or averaged, for more reliable bioassessments using the EPA RBP III. The difference in the bioassessments obtained using pooled samples compared to multi-habitat samples poses a dilemma since reasonable arguments can be made for the use of either method. Multi-habitat samples are not necessarily inferior for assessment, but differences in the distribution of habitat types among streams could reduce the reliability of RBP using this sample regime.

The United States Environmental Protection Agency Rapid Bioassessment Protocols (RBP) are variations of an approach to biological monitoring termed "multimetric techniques," and are derived from the Index of Biotic Integrity developed for assessing stream condition based on fish assemblages (Karr, 1981; Fausch et al., 1984; Karr et al., 1986). These techniques employ several ecologically relevant measures of community structure and function (metrics) that are sensitive to stream degradation (Plafkin et al., 1989). The metrics are selected to cover a wide range of structural and functional properties of aquatic biological communities. Proponents of RBPs maintain that when appropriate metrics are selected and data are collected and processed in standardized protocols, RBPs provide a cost-effective and accurate assessment of relative stream conditions (Karr, 1991; Gerritsen, 1995). Proponents also maintain that RBP methods are easier to apply, less expensive, and easier to communicate to managers and to the public when compared to standard multivariate statistical methods (Gerritsen, 1995). However, even proponents of RBPs often disagree about specific protocols used in their application. As a result, multimetric techniques have been the subject of much debate since their introduction (Gerritsen, 1995; Norris, 1995; Reynoldson et al., 1997). Much of this debate centers on issues of sample collection, processing, and analysis implemented as cost saving measures (Norris and Georges, 1993; Doberstein et al., 2000). Examples include the practices of not collecting replicate samples, subsampling and/or splitting of collected samples, and identifying collected specimens only to taxonomic family.

We applied a modified RBP to the macroinvertebrate assemblages we sampled from seven streams in the West Sandy Creek watershed of Henry County, Tennessee. The bioassessments obtained from these RBPs using different sampling regimes provide insight into two aspects of the ongoing debate about RBP methodology. One aspect was the effect of replication, and the other was the effect of pooling samples from multiple habitats versus stratifying sampling to reduce variability. We compared the bioassessments from four sampling regimes: (1) three semiquantitative root mass samples consisting of 100+ organism subsamples (100 picks), (2) the three 100 pick subsamples combined into a pooled sample (pooled samples), (3) the mean of the three 100 pick subsamples (sample means), and (4) the pooled samples combined with qualitative multiple habitat samples (multi-habitat samples). In addition, our data provide bioassessment results from habitats infrequently reported in the literature (submerged root masses, macrophytic plant beds, and snags with associated course particulate organic matter [S/CPOM]) because stream rif-

fles are recommended as the habitat of first choice in the RBP protocols and, consequently, are the most common habitat reported.

MATERIALS AND METHODS

Study Area—The dendritic West Sandy Creek watershed (WSCW) drains the southeastern two-thirds (approximately 52,000 acres) of Henry County, Tennessee. West Sandy Creek is a tributary of the Big Sandy River, flowing into a western embayment of the impounded Tennessee River. This watershed is in the Southeastern Plains and Hills Subregion of the Southeastern Plains Ecoregion (Griffith et al., 1998). The highly erodible soils are loose unconsolidated gravels, sands, and clays from the Cretaceous Period. Given the nature of these materials, the streams of this watershed have poorly developed riffles composed of either small gravels or hard clay. Most of the stream bottoms are composed of loose shifting sand. All of the streams, from the many first and second order tributaries draining fields to the fifth order main channel of West Sandy Creek, are affected to some degree by channelization.

Land utilization in the watershed includes agriculture (row crop, pasture), forestry, urban development (Paris, Tennessee) and light industry. Significant acreage in the lower watershed is in forested wetlands managed by the Tennessee Valley Authority (TVA). In 1990 the TVA established a targeted watershed monitoring program in the WSCW. In 1992 WSCW was selected as a Nonpoint Source Water Pollution Target Watershed National Demonstration Project. Funds from the United States EPA NPS 319(h) program were used to implement Best Management Programs (BMPs) and stream water quality monitoring. Various agricultural BMPs such as sediment retention basins, reduced tillage, nutrient and animal waste management were implemented. In-stream BMPs such as log weirs, log deflectors, and log-boulder deflectors were installed in some streams to mitigate the effects of channelization. Although many BMPs were implemented throughout the watershed, the urbanization that occurred during the study period may have mitigated against some of the BMP benefits.

Macroinvertebrate Sample Collection and Processing—Seven streams in the WSCW were selected for study. Each stream was sampled on 6–7 August 1992, 18–19 June 1993, 8–9 and 11–12 October 1993, 24 September and 1 October 1994, 24–25 May 1995, and 10–11 November 1995. Thus, each stream was sampled twice in the spring "wet" season, once before and once after implementation of agricultural and in-stream BMPs; and four times in the late summer/early fall "dry" season, twice before and twice after implementation of BMPs. Macroinvertebrates were collected with benthic kick nets. We collected two categories of sample: semiquantitative and qualitative. The number of organisms collected can be related to a standard unit of sampling effort for semiquantitative samples, however the number of organisms collected cannot be related to a standard unit of sampling effort for qualitative samples. Semiquantitative submerged root mass samples were standardized by collecting them using a timed effort approach. Each semiquantitative submerged root mass sample was collected by vigorously sweeping the net through submerged root masses for one minute. Three semiquantitative submerged root mass samples were collected from each stream on each sampling date. Qualitative samples were collected by picking macroinvertebrates from benthic kick net samples in proportion to their abundance for one hour or until no new taxa

were encountered. A single qualitative riffle, macrophytic plant bed, S/CPOM, and submerged root mass sample was collected from each stream in which they were present on each date.

Samples were preserved in the field using 80% ethanol and returned to the lab for further processing. A subsample of at least 100 macroinvertebrates (a 100 pick) was obtained from each semiquantitative submerged root mass sample by uniformly spreading the sample on a pan demarcated into twenty-eight 25 cm² cells and picking all of the macroinvertebrates out of randomly chosen cells in the grid. All macroinvertebrates were picked from each randomly selected cell until a sample of at least 100 organisms was obtained. In all samples the macroinvertebrates, including chironomids, were identified to genus when possible using standard taxonomic keys. The three 100 pick submerged root mass samples served as replicates in this study. Although three is a small number of replications, it is typical of many bioassessment studies.

Data Analysis—Bioassessments were calculated using a modified EPA RBP III (Plafkin et al., 1989). The following metrics and their derivation were included in these bioassessments.

Metric 1. Taxa richness (N): the total number of distinguishable taxa (i.e. genera) in the sample. Resh and Jackson (1993) consider taxa richness to be the most robust indicator of water quality. Taxa richness is hypothesized to decrease with decreasing water quality, with the possible exception of moderate organic enrichment of naturally nutrient-deficient streams.

Metric 2. Taxa richness of Ephemeroptera-Plecoptera-Trichoptera (EPT): the total number of distinguishable taxa (i.e. genera) of ephemeropterans (mayflies), plecopterans (stoneflies), and trichopterans (caddisflies) in the sample. These three taxa are generally less tolerant of pollution than other taxa and are reliable indicators of stream conditions in North Carolina (Lenat, 1993; Eaton and Lenat, 1991). Taxa richness and EPT are used in at least 95% of bioassessment programs in the United States (Tetra Tech, 1993).

Metric 3. Modified Hilsenhoff Biotic Index (HBI):

$$\text{HBI} = \sum x_i t_i / n$$

where: x_i = number of individuals within a taxon, t_i = tolerance value of a taxon, and n = total number of organisms in the sample. We used tolerance values from the North Carolina Department of Environmental Management (NCDEM) (Lenat, 1993) and from Hilsenhoff (1987, 1988) and Plafkin (1989) when not available from NCDEM. The HBI estimates a weighted average pollution tolerance of the macroinvertebrate assemblage of a stream. The rationale of this metric is that the average pollution tolerance of the macroinvertebrate assemblage in a stream will increase proportionately to the pollution load in the stream.

Metric 4. Ratio of the number of EPT individuals to the number of EPT + chironomid individuals (EPT/[EPT + chi]): This metric can be viewed as the proportional abundance of macroinvertebrates in the mostly pollution sensitive taxa to the mostly pollution tolerant taxa (Lenat, 1993). The original version of this metric suggested by the EPA (Plafkin et al., 1989) was the ratio of EPT to Chironomidae in a sample. The original metric was statistically unstable (Barbour et al., 1992; Resh and Jackson, 1993), and the metric used here was proposed (Hannaford and Resh, 1995) as a more stable alternate since it can only vary between 0–1. However, this modified metric, while improved, remains the most variable of the metrics in this analysis.

Metric 5. Percent abundance of the dominant taxon (Dom):

the most abundant single taxon in the sample as the percent of the total number of macroinvertebrates in the sample was used as a simple measure of dominance. Dominance is hypothesized to increase with pollution and/or habitat degradation because tolerant species become a larger proportion of the total.

Metric 6. Jaccard Community Similarity Index (Jccrd):

$$Jccrd = c/(a + b + c)$$

where: c = the number of taxa common to samples from both streams, a = the number of taxa found in samples from stream A but not from stream B, and b = the number of taxa found in samples from stream B but not from stream A. This index compares the taxonomic composition of two streams based on shared and unshared taxa using presence or absence of each taxon.

Metric 7. Ratio of shredders to total macroinvertebrate abundance (Shr): the ratio of macroinvertebrates considered to belong to the "shredder" functional feeding group (Merritt and Cummins, 1996) as a percent of the total number of macroinvertebrates in the sample. The River Continuum Concept (Vannote et al., 1980) postulates that the shredder functional feeding group is expected to be a large portion of the macroinvertebrate community in low order streams with healthy riparian zones. Reduction of the shredder portion of the macroinvertebrate community in such streams is hypothesized to result from riparian and hydrological degradation of streams.

All of the above metrics were used in bioassessments based on the 100 pick samples. The multi-habitat samples were combined with the pooled samples to assess the effect of sampling more habitats and a greater area of stream habitat. In these combined samples only the additional information for metrics 1, 2, and 6 was included because metrics 3–5 and 7 require at least semiquantitative sampling for validity. Thus values of metrics 3–5 and 7 were unchanged by combining qualitative multiple habitat samples with the pooled samples.

In the RBP analysis each stream receives a Biological Condition Score (BCS) based on the degree to which each of its metric scores differ from the same metric scores in the reference stream. Bomar Creek was chosen as the reference stream because it is not channelized, has healthy riparian zones, and a higher percentage of forested area compared to the other streams of this study. Four BCSs are possible: 0, 2, 4, or 6. The percent difference relative to the reference stream corresponding to any particular BCS varies for each metric. The BCSs are summed for all the metrics to yield a multimetric score. The percent similarity of the multimetric score of each stream compared to the multimetric score of the reference stream is used to assign each stream to a Biological Condition Category (BCC) in the final bioassessment. We assigned streams to BCCs based on similarities in multimetric scores of > 83% = not impaired, 54–83% = slightly impaired, 21–53% = moderately impaired and < 21% = severely impaired. The similarity criteria used here to assign BCCs are similar to the values proposed in the original EPA RBPs (Plafkin et al., 1989). However, these values are somewhat arbitrary, and other criteria, such as statistical quartiles and probability levels have been proposed and used to set BCC values (Resh et al., 1995).

RESULTS

The distribution of metric values among streams is compared for each sampling regime in Fig. 1. Each box and whisker chart in Fig. 1 displays the distribution of metric values obtained from

a sampling regime over the 6 sampling dates. Each box and whisker chart for 100 pick samples includes 18 observations, but those for pooled, means, and qualitative multi-habitat samples include 6 observations. The different sample size results from summing or averaging the three separate 100 pick samples into a single pooled or mean sample for the other sampling regimes.

The percent differences in bioassessment outcomes among the four different sampling regimes for the seven streams are summarized in Table 1. The bioassessments (BCC scores) resulting from the use of 100 picks, pooled samples, sample means, and qualitative multi-habitat samples are compared in Fig. 2. Bioassessments using individual 100 picks varied more than those for either pooled samples or the sample means. Bioassessments differed among 100 picks 53% of the time, and varied by two BCCs 4.6% of the time. Bioassessments based on 100 picks differed with those from pooled samples 46% of the time, but never by more than one BCC. Bioassessments based on sample means differed from the 100 picks 33% of the time, but never by more than one BCC. Bioassessments based on pooled samples differed from the sample means 22% of the time, but never by more than one BCC. Bioassessments based on the pooled samples differed from those obtained using multi-habitat samples 31% of the time, but never by more than one BCC.

DISCUSSION

Some patterns can be observed in the metric scores illustrated in Fig. 1. The range of metric scores is often greater for the 100 pick samples. The diversity metrics, N and EPT, seem to contradict this pattern, but this most likely results from the smaller number of species in the smaller 100 pick samples compared to pooled and qualitative multi-habitat samples. Smaller numbers cannot vary as much as larger numbers. One reason the 100 pick samples vary more than other sampling regimes may be that there are more samples that can vary (18 vs. 6 samples) and, therefore, more opportunity for atypical assemblages to be collected. More importantly, the smaller area of habitat sampled by an individual 100 pick would be expected to yield a less reliable estimate of the true macroinvertebrate assemblage of the stream. This would be true even if the spatial distribution of macroinvertebrates was random, but it is much more likely given that macroinvertebrate spatial distribution is usually aggregated.

The taxa richness metrics (N and EPT) and metrics of community structure (Jccrd and Dominance) are clearly sensitive to sampling regime. The number of taxa increased with the amount of habitat sampled which is the effect of combining 100 picks into pooled samples or combining pooled samples with qualitative multi-habitat samples. Community similarity (Jccrd) increased with sample area and habitats sampled. The increase in community similarity observed for pooled and multi-habitat samples suggests this metric may perform better if larger areas of habitat or more habitats are sampled. However, the relative similarities did not seem to differ among sampling regimes and probably did not affect the bioassessment outcomes. Dominance decreased with increased sample area indicating that this metric, like Jccrd, may require larger samples. Values for sample means for these four metrics were distributed similarly to those for 100 picks, but had less variance.

The HBI, EPT/(EPT + Chi), and Shr metrics were very robust to different sampling regimes and seem to be little affected by the area of habitat sampled. Recall that these metrics were not recalculated for qualitative multi-habitat samples. These
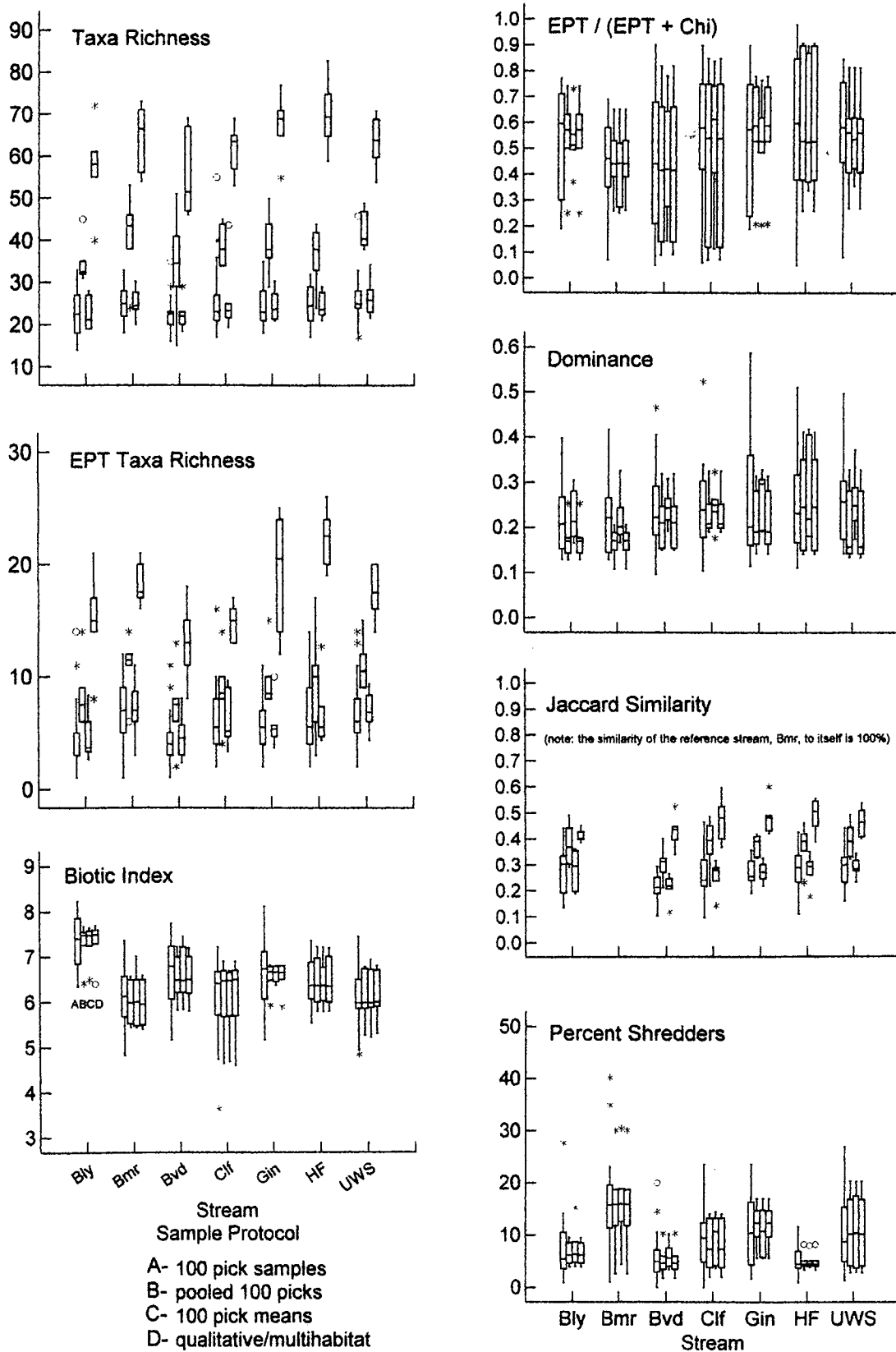
FIG. 1. The distribution of metric values among streams for each sampling regime. Each box and whisker displays the distribution of metric values obtained from a sampling regime over the 6 sampling dates. Each box and whisker chart for 100 pick samples includes 18 observations, but those for pooled, means, and qualitative multi-habitat samples include 6 observations. From left to right within each group of four box and whisker charts are: A) 100 pick samples, B) pooled 100 pick samples, C) 100 pick means, and D) combined qualitative multi-habitat samples. The horizontal line transects the box at the median of the data distribution,

TABLE 1. Differences in Biological Condition Categories (BCCs) for different sampling regimes.

Sample regime	Differ by one BCC	Differ by two BCC
Among 100 picks	53%	4.6%
Pooled samples versus 100 picks	46%	0.0%
Sample means versus 100 picks	33%	0.0%
Pooled samples versus sample means	22%	0.0%
Pooled versus multi-habitat samples	31%	0.0%

might be good metrics to use to compare streams that were sampled differently as often occurs when different agencies or research groups attempt to compare results.

Some differences among bioassessments are to be expected and accepted as statistically insignificant in a large number of comparisons. However, the differences among 100 picks for these data are excessive, and indicate that the streams in this study require more than a single 100-organism subsample for reliable bioassessment based on the EPA RBP III. Although Hilsenhoff (1987, 1988) found 100 pick subsamples were adequate for the Biotic Index he developed to classify hard-bottomed streams in Wisconsin, that does not seem to be the case for soft-bottomed streams in Tennessee (e.g. the West Sandy Creek watershed). Hilsenhoff's Biotic Index was developed to classify the extent of organic pollution, and although it has been widely used in bioassessments for other pollutants, we are not aware of any studies that measure its power to assess the effects of a variety of non-point source pollutants.

Both combining and averaging the three 100 pick subsamples, which effectively increased the area of stream habitat sampled, reduced the variability among bioassessments compared to the individual 100 picks. This is not surprising and would be predicted by basic statistical principles, i.e. lower standard error of the estimate for replicated samples. However, pooling samples reduced variability of some metrics differently than did averaging of samples. For example, as more samples are pooled, a larger area of stream habitat is included in the sample and the variability of species richness metrics tends to decline because the metric tends to asymptotically approach the maximum total species present in the stream (Preston, 1948; May, 1975; Slocomb and Dickson, 1978). The variation in diversity metrics is reduced also by averaging increasing numbers of samples, but in a different manner. By the Central Limit Theorem, larger numbers of samples are more likely to have a mean closer to the mean of the population from which they were drawn. The greater agreement between 100 pick and sample mean bioassessments compared to 100 pick and pooled sample bioassessments is also to be expected because the BCS of the sample means must always be within the range of the BCS for the three 100 picks, but this is not always true of the BCS of the pooled samples. Interestingly, Barbour and Gerritsen (1996) reported that analysis of variance

of taxa richness of benthic macroinvertebrate samples from Florida lakes discriminated differences among lakes better using fixed count subsamples (100 picks) than using either combined samples or sample means. The magnitude of some metrics from pooled samples is greater than that obtained from averaging samples. If the metric variability is similar for the two sample types, then the metric coefficient of variation from pooled samples will be smaller for the pooled samples than for sample means. Smaller coefficients of variation allow more powerful statistical analysis (Cohen, 1988; Sokal and Rohlf, 1995), a reduced chance of accepting a null hypothesis when it is false, and a more precise bioassessment (Pratt and Bowers, 1992; Resh and McElravey, 1993; Fore et al., 1994). The variability of other metrics, such as ratio metrics, may decline as more samples are averaged but not as more are combined, or vice versa. It is these metric by metric differences in response to pooling versus averaging of samples that explain the 22% differences in bioassessment results observed for pooled samples versus sample means.

The fact that the bioassessments obtained by adding qualitative sampling to the analysis differed from 31% of those obtained using strictly semiquantitative data is problematic, in part, because it is difficult to determine which of the two methods provided the most accurate assessment. The 31% level of differences between these two sampling regimes is quite high considering that the values of four of their seven metrics are identical. Only the values of taxa richness, EPT taxa richness, and Jaccard Similarity differed between the two sampling regimes. Data from the multi-habitat, qualitative samples was not included in the derivation of the remaining metrics because these metrics, being based on counts, require at least semiquantitative sampling (uniform sampling effort) for validity. Although sample stratification to reduce variation has been a basic principle of stream sampling since the work of Allen (1959), Lenat (1988) reported good results using assessment protocols based on combined multiple habitat samples. Several other state agencies have adopted multi-habitat sampling regimes (Massachusetts Department of Environmental Protection, 1995; Mid-Atlantic Coastal Streams Workgroup, 1996; Florida Department of Environmental Protection, 1996). In addition, since the metrics taxa richness and EPT taxa richness have proved to be among the most useful in a large number of studies, it seems logical a sampling regime that maximizes the accuracy and value of these metrics would be beneficial. Sampling a wide range of stream habitats includes a larger total proportion of the stream taxa in the sample. However, Kerans et al. (1992) obtained differing bioassessments from riffle habitats than pool habitats and they interpreted this to indicate that bioassessment accuracy would be diminished by combining samples from both habitats, i.e. multi-habitat sampling. Riffle habitat is the most commonly sampled when stratification is practiced (Resh and McElravey, 1993), although pool habitat was shown to be at least as accurate for bioassessment (Kerans et al., 1992).

The addition of qualitative multi-habitat samples to an analysis has the effect of increasing both the area and number of stream habitats sampled. However, it also adds variance resulting

←

the box edges mark the central 50% of the data distribution (1st and 3rd quartiles), and the whiskers show the range of observed values that fall within ± 1.5 * interquartile range (the range between the box edges). Asterisks mark observations that fall between 1.5 and 3 interquartile ranges, and open circles mark observations that fall outside of 3 interquartile ranges.
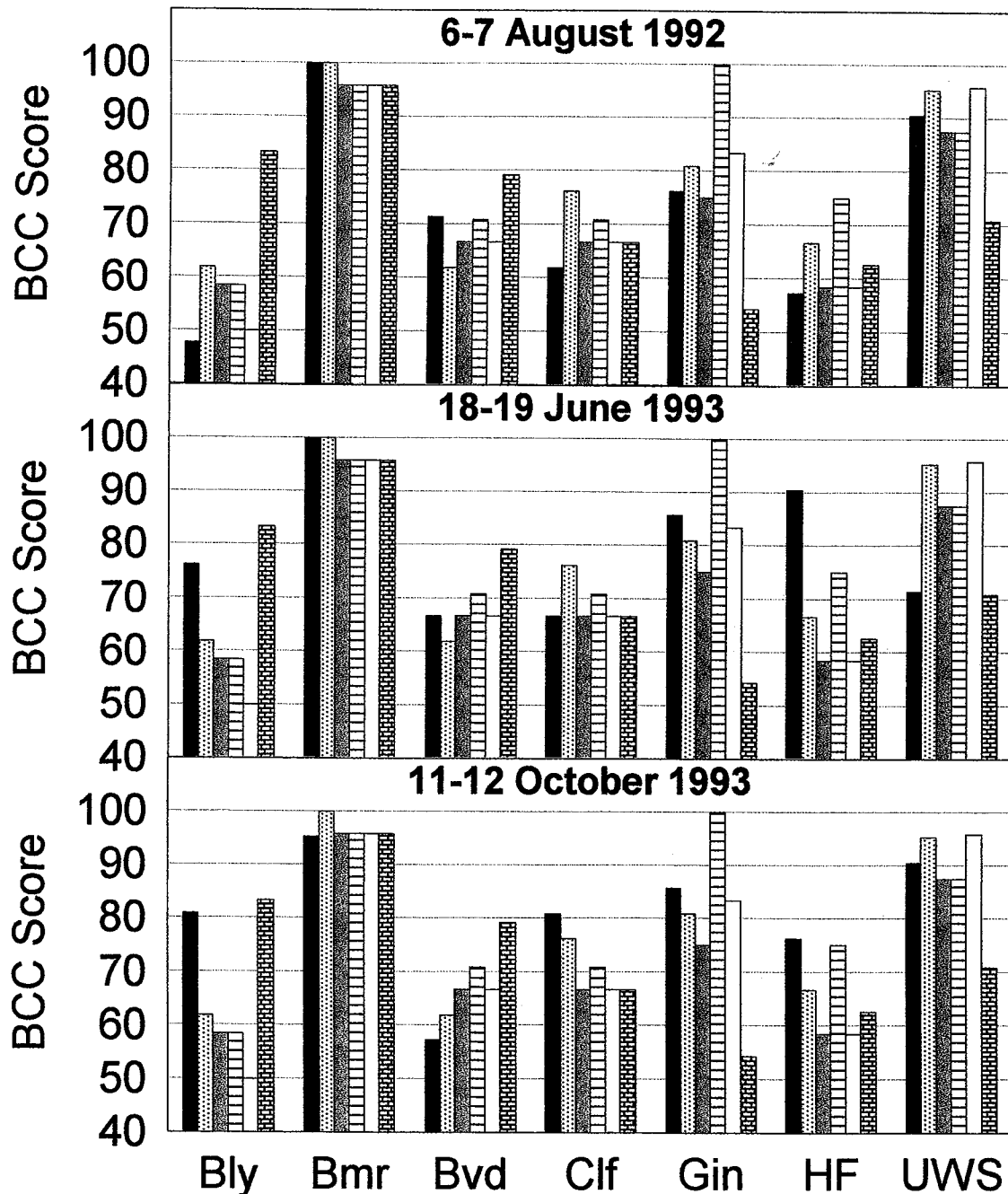
FIG. 2. Bioassessment results (Biological Condition Categories; BCC scores) for seven streams on six sample dates comparing use of pooled samples (pooled), sample means (mean), multi-habitat samples (multi), and individual 100 pick subsamples (P1–P3). Bly = Bailey Creek, Bmr = Bomar Creek (reference), Bvd = Beaverdam Creek, Clf = Clifty Creek, Gin = Gin Creek, HF = Holly Fork Creek, and UWS = Upper West Sandy Creek.

from differences in the relative amount of each habitat type in the streams. All habitats were not present in every stream. Few of the streams had riffles and macrophytes. Furthermore, the quality of riffles varied among the streams where they were present (substrate material and size distribution) and the plant species of the macrophyte habitats also differed. This additional variability implies that bioassessments based on stratified sampling, i.e. pooled root mass samples, would likely be more accurate than those based on multi-habitat samples. However, variability of the two species richness metrics used in the bioassessment may be reduced by sampling multiple habitats. Additionally, since sampling multiple habitats yields larger numbers of species, the coefficient of variation for these metrics may be reduced.

Resh and McElravy (1993) applied statistical power analysis to benthic macroinvertebrate data from a stream in California to assess the effect of replication on the minimal detectable differences of metrics used in bioassessments. They reported that the minimal detectable difference for four different metrics at a 5% significance level and at a statistical power of 95% was greatly reduced by increasing replicates from two to three, but little ad-
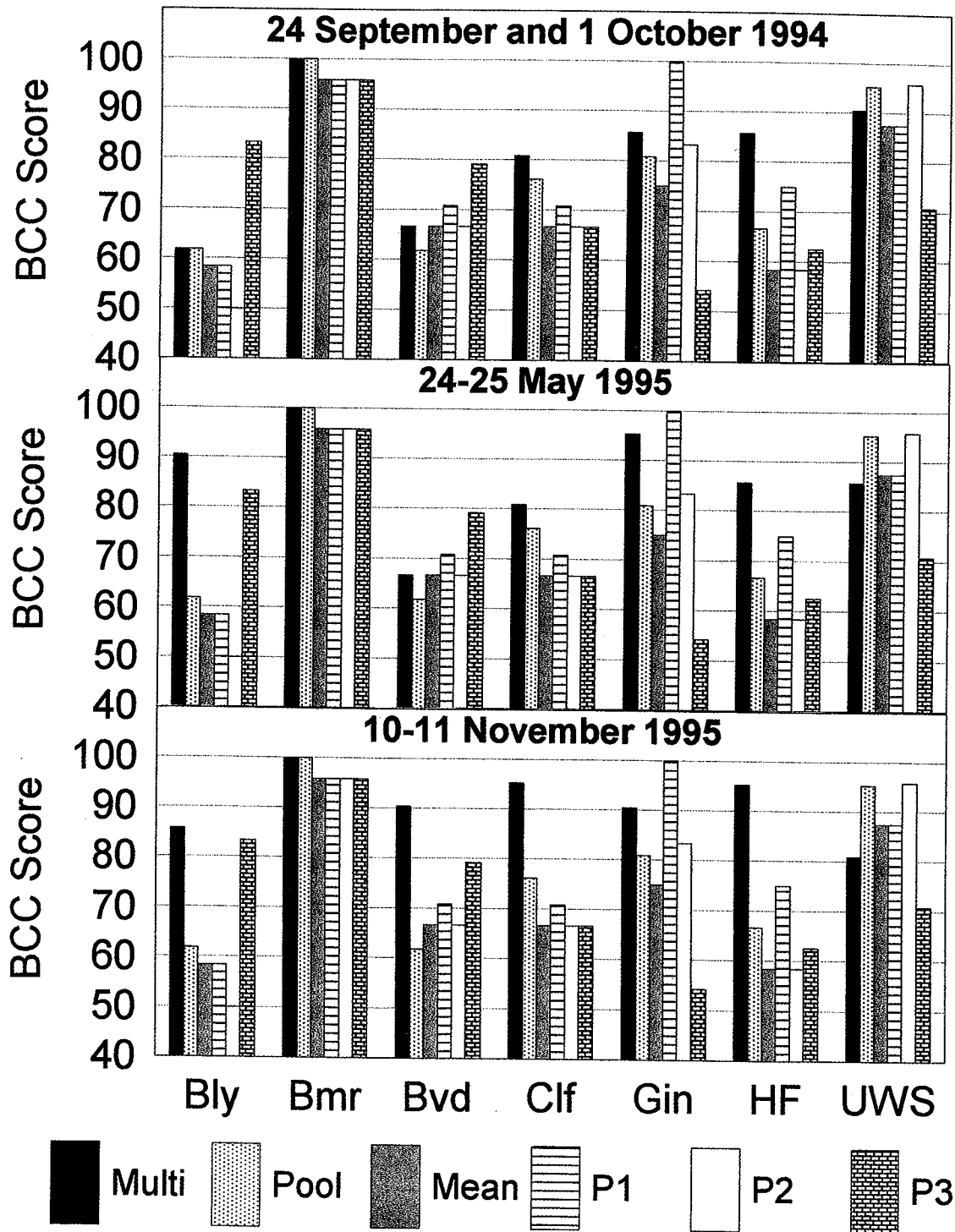
FIG. 2. Continued.

ditional reduction in minimal detectable difference was achieved by increasing the number of replicates above five. Pratt and Bowers (1992) reported similar conclusions for studies of laboratory microcosm studies of microbial communities. Smaller minimal detectable differences for the component metrics of the RBP should equate to more powerful and accurate bioassessments since the RBP is functionally an additive combination of its individual metrics. Fore et al. (1994) demonstrated that the benthic Index of Biotic Integrity (a multimetric technique they derived)

could accurately discriminate among 5 stream condition categories (BCCs). Application of these procedures to the WSCW data is beyond the scope of this study, but would be useful for understanding which metrics and sampling regimes provide the best assessment in terms of precision and power.

In summary, while each sampling regime revealed differences from the control stream, the bioassessments often differed among sampling regimes. Taxa richness and community composition metrics were sensitive to sampling regimes that in-

creased the area and diversity of stream habitats included. This result suggests that multimetric bioassessments including these metrics should contain at least three 100 pick samples. The modified Hilsenhoff Biotic Index, Dominance, and Percent Shredder metrics were robust to sampling regime. Bioassessments among 100 pick samples differed more than half the time. Given this result, the use of single 100 pick samples for performing multimetric bioassessments would seem invalid. Differences among the other sampling regimes, all of which include three replicates, also were significant. It is possible that a larger total number of replicates would achieve convergent results among these different sampling regimes, but it also may be that different sampling regimes will always yield somewhat different results. If the latter is true, then comparisons among bioassessments derived from different sampling regimes should not be attempted. Specific to the West Sandy Creek watershed, bioassessments using the pooled 100 pick samples appear the most reliable because they include larger sample area than individual 100 pick samples and avoid the additional variance that could result from sampling multiple habitats.

ACKNOWLEDGMENTS

LITERATURE CITED

ALLEN, K. R. 1959. The distribution of stream bottom fauna. Proc. New Zealand Ecol. Soc., 6:5–8.

BARBOUR, M. T., J. L. PLAFKIN, B. P. BRADLEY, C. G. GRAVES, AND R. W. WISSEMAN. 1992. Evaluation of EPA's rapid bioassessment benthic metrics: metric redundancy and variability among reference stream sites. Environ. Toxicol. Chem., 11:437–449.

BARBOUR, M. T., AND J. GERRITSEN. 1996. Subsampling of benthic samples: a defense of the fixed-count method. J. North Am. Benthol. Soc., 15:386–391.

COHEN, J. 1988. Statistical power analysis for the behavioral sciences. Lawrence Erlbaum Associates, Hillsdale, New Jersey.

DOBERSTEIN, C. P., J. R. KARR, AND L. L. CONQUEST. 2000. The effect of fixed-count subsampling on macroinvertebrate biomonitoring in small streams. Freshwater Biol., 44:355–371.

EATON, L. E., AND D. R. LENAT. 1991. Comparison of a rapid bioassessment method with North Carolina's qualitative macroinvertebrate collection method. J. North Am. Benthol. Soc., 10:335–338.

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.

FLORIDA DEPARTMENT OF ENVIRONMENTAL PROTECTION. 1996. Standard operating procedures for biological assessment. Florida Dept. Environ. Prot., Biol. Sect., July 1996.

FORE, L. S., J. R. KARR, AND L. L. CONQUEST. 1994. Statistical properties of an index of biological integrity used to evaluate water resources. Can. J. Fish. Aquat. Sci., 51:1077–1087.

GERRITSEN, J. 1995. Additive biological indices for resource management. J. North Am. Benthol. Soc., 14:451–457

GRIFFITH, G., J. OMERNIK, AND S. AZEVEDO. 1998. Ecoregions of Tennessee (map and narrative). US Geol. Surv., Denver, Colorado.

HANNAFORD, M. J., AND V. H. RESH. 1995. Variability in macroinvertebrate rapid-bioassessment surveys and habitat assessments in a northern California stream. J. Am. Benthol. Soc., 14:430–439.

HILSENHOFF, W. L. 1987. An improved biotic index of organic stream pollution. Great Lakes Entomol., 20:31–39.

———. 1988. Seasonal correction factor for the biotic index. Great Lakes Entomol., 21:9–13.

KARR, J. R. 1981. Assessment of biotic integrity using fish communities. Fisheries, 6:21–27.

KARR, J. R., K. D. FAUSCH, P. L. ANGERMEIER, P. R. YANT, AND I. J. SCHLOSSER. 1986. Assessment of biological integrity in running waters: a method and its rationale. Illinois Nat. Hist. Surv. Special Pub. 5.

KARR, J. R. 1991. Biological integrity: a long-neglected aspect of water resource management. Ecol. Applicat., 1:66–84.

KERANS, B. L., J. R. KARR, AND S. A. AHLSTEDT. 1992. Aquatic invertebrate assemblages: spatial and temporal differences among sampling protocols. J. North Am. Benthol. Soc., 11:377–390.

LENAT, D. R. 1988. Water quality assessment of streams using a qualitative collection method for benthic macroinvertebrates. J. North Am. Benthol. Soc., 7:222–233.

———. 1993. A biotic index for the southeastern United States: derivation and list of tolerance values, with criteria for assigning water quality ratings. J. North Am. Benthol. Soc., 12:279–290.

MASSACHUSETTS DEPARTMENT OF ENVIRONMENTAL PROTECTION. 1995. Massachusetts DEP preliminary biological monitoring and assessment protocols for wadeable rivers and streams. Massachusetts Dept. Environ. Prot., North Grafton, Massachusetts.

MAY, R. M. 1975. Patterns of species abundance and diversity. Pp. 81–120 in Ecology and evolution of communities (M. L. Cody and J. M. Diamond, eds.). Harvard Univ. Press, Cambridge, Massachusetts.

MERRITT, R. W., AND K. W. CUMMINS. 1996. An introduction to the aquatic insects of North America. Kendall/Hunt Publishing, Dubuque, Iowa.

MID-ATLANTIC COASTAL STREAMS WORKGROUP. 1996. Standard operating procedures and technical basis: Macroinvertebrate collection and habitat assessment for low-gradient nontidal streams. Delaware Dept. Nat. Resour. Environ. Conserv., Dover, Delaware.

NORRIS, R. H. 1995. Biological monitoring: the dilemma of data analysis. J. North Am. Benthol. Soc., 14:440–450.

NORRIS, R. H., AND A. GEORGES. 1993. Analysis and interpretation of benthic macroinvertebrate surveys. Pp. 234–286 in Freshwater biomonitoring and benthic macroinvertebrates(D. M. Rosenberg and V. H. Resh, eds.). Chapman & Hall, New York.

PLAFKIN, J. L., M. T. BARBOUR, K. D. PORTER, S. K. GROSS, AND R. M. HUGHES. 1989. Rapid bioassessment protocols

for use in streams and rivers: benthic macroinvertebrates and fish. EPA/440/4-89/001. US Environ. Protect. Agency.

PRATT, J. R., AND N. J. BOWERS. 1992. Variability of community metrics: detecting changes in structure and function. Environ. Toxicol. Chem., 11:451–457.

PRESTON, F. W. 1948. The commonness, and rarity, of species. Ecology, 29:254–83.

RESH, V. H., AND J. K. JACKSON. 1993. Rapid assessment approaches to biomonitoring using benthic macroinvertebrates. Pp. 195–233 *in* Freshwater biomonitoring and benthic macroinvertebrates (D. M. Rosenberg and V. H. Resh, eds.). Chapman & Hall, New York.

RESH, V. H., AND E. P. MCELRAVY. 1993. Contemporary quantitative approaches to biomonitoring using benthic macroinvertebrates. Pp. 159–194 *in* Freshwater biomonitoring and benthic macroinvertebrates (D. M. Rosenberg and V. H. Resh, eds.). Chapman & Hall, New York.

RESH, V. H., R. H. NORRIS, AND M. T. BARBOUR. 1995. Design and implementation of rapid assessment approaches for monitoring using benthic macroinvertebrates. Australian J. Ecol., 20:108–121.

REYNOLDSON, T. B., R. H. NORRIS, V. H. RESH, K. E. DAY, AND D. M. ROSENBERG. 1997. The reference condition: a comparison of multimetric and multivariate approaches to assess water-quality impairment using benthic macroinvertebrates. J. North Am. Benthol. Soc., 16:833–852.

SLOCOMB, J., AND K. L. DICKSON. 1978. Estimating the total number of species in a biological community. Pp. 38–52 *in* Biological data in water pollution assessment: quantitative and statistical analysis. (K. L. Dickson, J. Cairns, and R. J. Livingston, eds.). Am. Soc. Testing Materials STP 652.

SOKAL, R. R., AND F. J. ROHLF. 1995. Biometry, the principles and practice of statistics in biological research, 3rd ed. W. H. Freeman, New York.

TETRA TECH. 1993. Summary of state bioassessment programs. Prepared for US Environ. Protect. Agency, Monitoring Branch, Washington, DC.

VANNOTE, R. L., G. W. MINSHALL, K. W. CUMMINS, J. R. SEDELL, AND C. E. CUSHING. 1980. The river continuum concept. Can. J. Fish. Aquat. Sci., 37:130–137.