CHAPTER II: EVALUATION OF A STANDARDIZED QUALITATIVE SAMPLING METHOD FOR RAPID BIOASSESSMENT APPROACHES WITH BENTHIC MACROINVERTEBRATES

Abstract

The objective of this study was to determine the accuracy of a standardized qualitative approach commonly used in rapid bioassessment for assessing the biological condition of lotic systems. First, the composition of benthic macroinvertebrate community estimated by a standardized qualitative approach was compared to that estimated by a quantitative approach. Benthic macroinvertebrate communities were surveyed at 19 stations along Peak Creek, Virginia. Peak Creek suffers from a wide assortment of impacts resulting in a general decline of water quality and biological condition along its length. The benthic macroinvertebrate community was also surveyed at an additional reference station on Wolf Creek, a relatively unperturbed stream within the same ecoregion as Peak Creek. At each station, one standardized qualitative sample, consisting of two composite kick screen samples with subsampling (KSS), and a quantitative sample of the macroinvertebrate community using a Portable Invertebrate Box Sampler (PIBS) were taken. The Index of Biotic Similarity and the Bray-Curtis Coefficient were used to compare the two methods at both of the reference sites. The Index of Biotic Similarity and the Bray-Curtis Coefficient indicated that the two methods estimate similar communities. The methods were also compared using a t-test on 10 benthic community metrics at both reference sites. The PIBS samples estimated significantly greater taxa richness, modified Hilsenhoff Biotic Index score (HBI) and % collector-gatherers abundance than the KSS samples at the Peak Creek Reference Station and significantly more taxa and Ephemeroptera, Plecoptera and Trichoptera (EPT) taxa at the Wolf Creek Reference Station. A correlation analysis comparing the methods metric by metric at Peak Creek Stations 1-18 also indicated the methods estimated similar communities. The slight differences found in the metrics and similarity indices

are most likely due to differences in abundance of organisms obtained by the two sampling methods. To determine if standardized qualitative sampling with subsampling leads to the same assessment of biological condition as quantitative sampling, each station along Peak Creek was compared with a 95% confidence interval based on the 5 Peak Creek Reference Station samples for each metric and method. Peak Creek Stations 1-12 and 18 were considered impacted and stations 13-17 relatively unimpaired. The two sampling methods made the same assessment an average of 73% of the time. Assessments of biological condition using MAIS, a family-level multimetric index, showed that the two sampling methods agreed 89% of the time. Assessment agreement between sampling methods was strongly affected by metric sensitivity. We found no pattern showing one method was more accurate in making assessments of biological condition than the other. Given the greater time and costs associated with quantitative sampling methods, such as the PIBS, we conclude that standardized qualitative methods, such as the KSS, are preferable for rapid bioassessment approaches to environmental assessment.

Introduction

Rapid bioassessment approaches using benthic macroinvertebrates have been widely heralded as fast, cost-effective techniques for assessing water quality (Reynoldson et al. 1986, Plafkin et al. 1989, Resh 1995, Resh et al. 1995). Independently developed by several state water quality agencies, rapid bioassessment was eventually integrated by the United States Environmental Protection Agency (USEPA) into 3 standardized protocols (Plafkin et al. 1989). These protocols were developed to assist and encourage states to add biological monitoring to their water quality programs.

Regulatory and natural resource agencies are often responsible for monitoring the water quality of hundreds of streams. Prior to the development of rapid bioassessment, agencies often relied on traditional quantitative bioassessments for detecting impairments in water quality. Though fairly accurate if enough replicates are collected, quantitative assessments quickly become expensive if many streams are involved. Processing, sorting and identifying benthic macroinvertebrates from multiple replicate samples is very time consuming and labor intensive. Securing results often takes months, thereby delaying possible management decisions. In the past, agencies were forced to either expend large amounts of time and money on a few streams and disregard the rest, or base their assessments on less accurate measures. (Lenat and Barbour 1994).

Rapid bioassessment approaches were developed to reach the same decisions obtained from quantitative assessments, but with less time and cost. Rapid bioassessment is considered "rapid" because replication is not emphasized and sample processing time is greatly reduced through subsampling. Lenat and Barbour (1994) estimate that three to five sites can be completely sampled, processed, and analyzed by a single worker in five working days. Rapid bioassessment enables natural resource agencies to assess more streams. However, the approach

was not designed to completely replace traditional quantitative sampling methods. Rapid bioassessment is primarily a screening tool. If an impairment is found, the site can then be more thoroughly investigated with quantitative methods (Resh and Jackson 1993).

Previous evaluations of rapid bioassessment approaches emphasized the accuracy of the method either by comparing single or multiple rapid bioassessments at various levels of impact (Plafkin et al. 1989, Hannaford and Resh 1995), or the variability of metrics in ascertaining impairment (Barbour et al. 1992, Barton and Metcalfe-Smith 1992, Resh and Jackson 1993, and Resh 1994). Little attention has been given to the effects of using different types of sampling gear that are uniquely designed either for standardized qualitative approaches or quantitative approaches. In standardized qualitative sampling, which is always employed in rapid bioassessment, samples are standardized by estimating the area of bottom habitat sampled or the amount of effort expended. Results can be expressed in terms of taxa richness and relative abundance of the taxa in the community, but results cannot be expressed in terms of absolute abundance. In quantitative sampling, the area of bottom being sampled is physically delimited, so results can be expressed in terms of absolute abundance.

In this paper, we present the results of a comparison of a typical standardized qualitative sampling method used in rapid bioassessment approaches with a typical quantitative sampling method in a stream known to be impaired. The study had two objectives: (1) to determine if standardized qualitative sampling with subsampling, such as is commonly done in rapid bioassessment approaches, estimates the same benthic macroinvertebrate community composition as quantitative sampling, and (2) to determine if standardized qualitative sampling with subsampling leads to the same assessment of biological condition as quantitative sampling.

Methods

Study area

This study was conducted in Peak Creek, a small mountain stream located in Pulaski County in southwest Virginia (37° 02’ 30”/80° 44’ 55”). Part of the New River drainage, Peak Creek lies completely within the Central Appalachian Ridges and Valleys ecoregion (Ecoregion 67, Omernik 1987). Near its headwaters, Peak Creek is impounded by Gatewood Reservoir. Below the reservoir, the stream quickly recovers from any impoundment effects and flows unimpaired for several kilometers through the George Washington and Jefferson National Forest and adjacent rural areas. As Peak Creek enters the town of Pulaski, Virginia, it receives an input of heavy metals (chromium, copper, zinc, iron, and manganese) from the permitted discharge of Magnox Incorporated. Magnox uses heavy metals in the manufacture of magnetic tape. Below Magnox, Tract Fork joins the stream and transforms Peak Creek into a fourth-order stream. The stream is partially channelized as it flows through the downtown area. Peak Creek receives another substantial input of heavy metals from the abandoned Allied Chemical Plant located downstream from Magnox. The Allied plant manufactured sulfuric acid and ferric sulfide. In 1976, Allied closed, leaving behind extensive waste piles (Willis 1989a). These waste piles were capped shortly before the beginning of this study. In addition to the two major inputs from Magnox and Allied, the town of Pulaski is also a source of heavy metal runoff for Peak Creek. Natural deposits of heavy metals surround the city. Iron and coal mines, now abandoned, supplied ore to three furnaces in Pulaski. Waste slag from these furnaces was used as fill for many of the construction sites in town.

Willis (1989a) analyzed sediments for heavy metals at nine stations along Peak Creek. Immediately above Magnox, he found 177 ppm of copper, 222 ppm of lead, 1070 ppm of zinc, 0 ppm of iron, 1 ppm of selenium, and 1 ppm of cadmium. Even at this control site, the concentration of copper, lead, and zinc were above the statewide 95 percentile. Percentiles represent the the probability of a stream having a lower concentration. Therefore, 95 percentile

means that there is a 5% probability of finding a stream within the state with a higher concentration of a given metal. Below Magnox, copper levels decreased to 63 ppm, lead increased to 346 ppm, zinc increased to 1150 ppm, iron skyrocketed to 34000 ppm, selenium and cadmium stayed at 1 ppm. Below the input of Allied Chemical, copper levels increased to 3120 ppm (exceeds the statewide 100 percentile), lead increased to 1200 ppm (exceeds the statewide 95 percentile), zinc increased to 1830 (exceeds the statewide 95 percentile), iron increased to 311000 ppm, selenium increased to 125 ppm (exceeds the statewide 100 percentile) and cadmium increased to 9 ppm (exceeds the statewide 95 percentile). Benthic macroinvertebrate surveys by Willis (1989b) found little or no impact upstream of Pulaski and moderate to severe impacts within Pulaski.

Study design

We selected a total of 19 stations along an 11-km reach of Peak Creek. We had five stations, approximately 0.5 km apart, starting immediately below the dam, plus an additional site 3.5 km below the dam, halfway to the town of Pulaski. We located one station immediately above Magnox and another directly below. We distributed the remaining stations approximately 0.5 km apart through Pulaski with five stations located above the Allied plant and six stations below it (Fig. 1).

Based on the results of a preliminary survey, we designated one of the original 19 sites on Peak Creek as a reference station. This reference station was located approximately 2 km downstream from the dam on land managed by the US Forest Service. We chose Wolf Creek as an off-site reference stream. Wolf Creek is a relatively unimpaired fourth-order stream situated near the town of Narrows in Giles County, Virginia (37° 18’ 30” / 80° 51’ 00”). We selected Wolf Creek as a reference site because like Peak Creek, it is a fourth-order stream, part of the New River drainage, and located within the Central Appalachian Ridges and Valleys ecoregion.

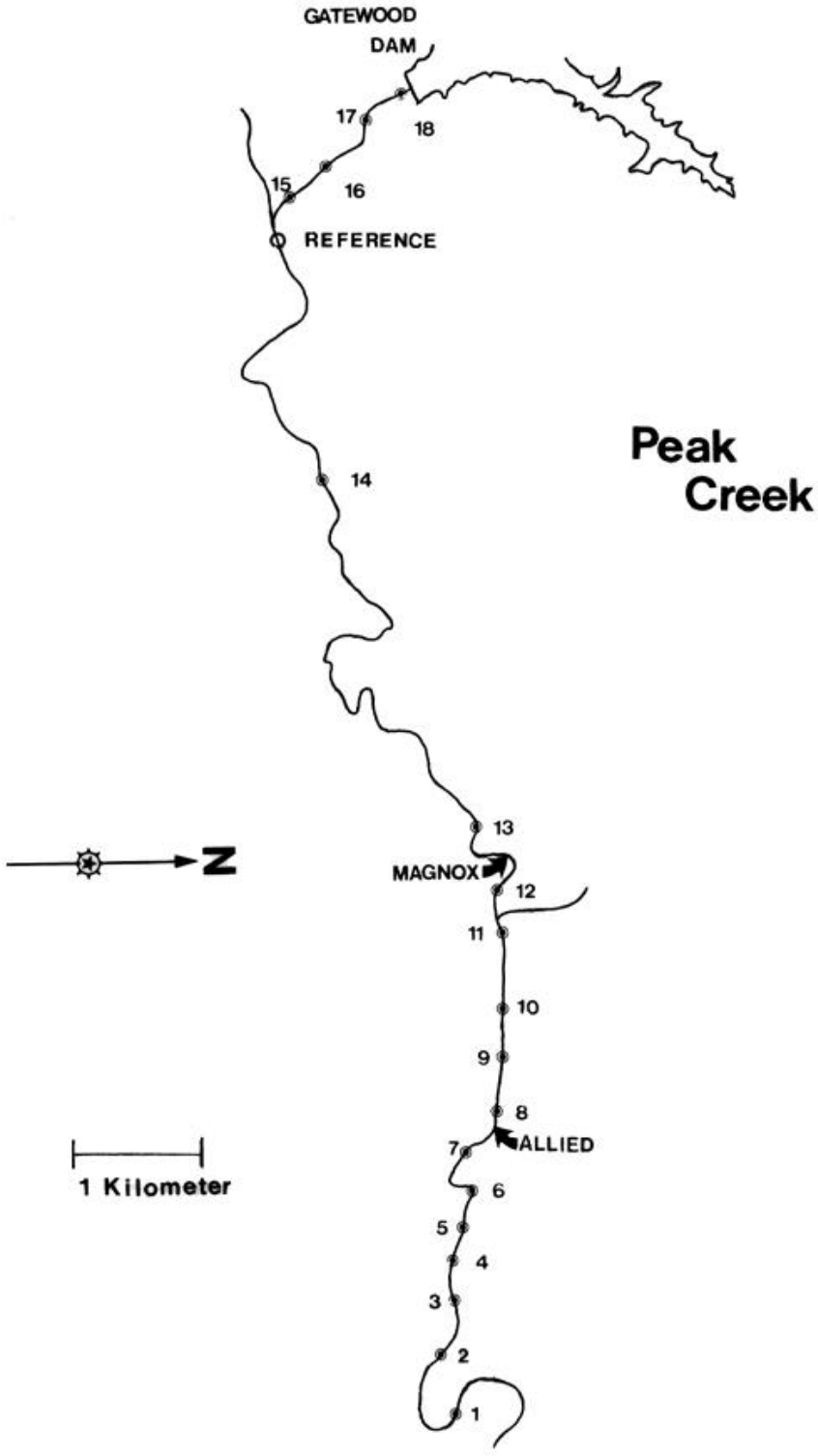

FIG. 1. Map of study area on Peak Creek.

Field methods

We used two sampling methods at each station. On August 5, 1993, we sampled the benthic macroinvertebrate communities at the 19 stations along Peak Creek. The following day, we sampled the remaining reference station on Wolf Creek. At each station, we took one quantitative sample using a Portable Invertebrate Box Sampler (PIBS) and one standardized qualitative sample using a 1-m^2 kick screen with subsampling (KSS). At the reference stations on Peak Creek and Wolf Creek, we also took five PIBS samples and five KSS samples. At each station, we sampled the best available riffle habitat (good flow, substantial substrate heterogeneity, and adequate depth).

The PIBS completely encloses a 0.1 m^2 area of stream bottom. The PIBS works especially well in riffle areas because its foam-lined bottom takes the shape of the often irregular stream bottom and creates a tight seal. Each rock in the sampling area was meticulously scrubbed, examined, and then removed. The remaining sediment was thoroughly raked to ensure the removal of all macroinvertebrates (Voshell, Layton, and Hiner 1989).

The USEPA protocol recommends using a 1-m^2 kick screen for taking a rapid bioassessment sample (Plafkin et al. 1989). Although a kick screen is not an enclosed sampling device like the PIBS, the sample is standardized by sampling an estimated area of stream bottom (1 m^2) (Hynes 1970). According to the protocol, two samples are taken, one in a riffle with a fast current and the other in a riffle with a slower current. The two samples are later combined into a single composite sample.

In addition to sampling the benthic macroinvertebrate communities along Peak Creek, we measured several physical and chemical parameters at each station. We measured temperature with a long stem thermometer and dissolved oxygen by the Winkler method. To determine alkalinity (mg/L $CaCO_3$), we collected 250 ml of stream water at each station and titrated the

sample with 0.02 N H_2SO_4, in accordance with the EPA's two-endpoint method [(USEPA 1983 (Method 310.1)]. We measured pH with an Orion SA250 pH meter, calibrated with pH 4.0 and 7.0 buffers, measured conductivity with a YSI Model 33 meter, and hardness (mg/L $CaCO_3$) with a Hach HA-71A kit. We also made an assessment of the habitat at each station (Barbour and Stribling 1991).

In the field, the benthic macroinvertebrate samples were preserved in a 95% ethanol solution and then transported to our laboratory. The PIBS samples were processed by sorting every macroinvertebrate from the sample. We subsampled the kick screen samples using a method similar to one devised by Caton (1991). Each kick screen sample was transferred to a rectangular gridded sieve. The sieve was divided into a grid of 30 5 x 5 cm^2 squares. After a sample was evenly distributed on the grid, we randomly selected a square from the grid and transferred the square's contents to a white sorting tray. The macroinvertebrates were separated from the debris and counted. Successive squares were selected until the total number of sorted organisms was within 10 % of 200. If we reached 200 organisms before a square was completely sorted, we finished the square.

We decided on using 200-organism subsamples based on the results of a preliminary study at another stream. We specifically examined the accuracy of assessments at different stopping points (100, 200, 300, etc.). We found that a 200 organism subsample was not only representative of the sample but also kept processing time to a minimum.

The benthic macroinvertebrates from both types of samples were preserved in a 70% ethanol solution and then identified to the lowest practical taxonomic level, usually genus, using keys by Wiggins (1977), Merritt and Cummins (1984) and Stewart and Stark (1988). Chironomidae were identified only to family. We used Pennak (1989) to identify the non-insect macroinvertebrates.

Data analysis

Community composition.-To determine if the two sampling methods make similar estimates of community structure, we made pairwise comparisons of the PIBS and KSS samples at each reference station with two similarity indices. The two similarity indices we used were: the Bray-Curtis Coefficient (Bray and Curtis 1957) and the Index of Biotic Similarity (Pinkham and Pearson 1974). We calculated the Bray-Curtis Coefficient percent similarity of community *a* and community *b* (PS_{ab}) as follows (Ludwig and Reynolds 1988):

$$PS_{ab} = \left(\frac{2W}{A + B}\right)(100)$$

where

$$W = \sum_{i}^{k} \left[\min\left(X_{ia}, X_{ib}\right)\right]$$

$$A = \sum_{i}^{k} X_{ia}$$

$$B = \sum_{i}^{k} X_{ib}$$

X_{ia} = abundance of the i^{th} species in community *a*

X_{ib} = abundance of the i^{th} species in community *b*

k = number of taxa compared

PS_{ab} ranges from 0 - 100, with 0 for communities with no taxa in common and 100 for identical communities. One shortcoming of the Bray-Curtis Coefficient is that the index is weighted toward dominant taxa and may neglect rare taxa. Because the abundance of rare taxa might be ecologically important in assessing differences in composition, we also calculated the Index of Biotic Similarity, B, (Pinkham and Pearson 1974). This index is more sensitive to rare taxa than the Bray-Curtis Coefficient because it weighs all taxa equally. The Index of Biotic Similarity ranges from 0 to 1, with 0 for entirely different communities and 1 for identical communities. It is calculated as follows:

$$B = \frac{1}{k} * \sum_{i}^{k} \frac{\min(X_{ia}, X_{ib})}{\max(X_{ia}, X_{ib})}$$

where

X_{ia} = abundance of the i^{th} species in community a

X_{ib} = abundance of the i^{th} species in community b

k = number of taxa compared

For an easier comparison with the Bray-Curtis Coefficient, we multiplied the results of the Index of Biotic Similarity by 100 so both similarity indices range from 0 to 100.

We also compared the different estimates of community composition at Peak Creek Stations 1-18 and the reference stations using 10 benthic macroinvertebrate community metrics. The 10 metrics were selected according to the following criteria: (1) categories that measure different components of benthic macroinvertebrate community structure (richness, composition, balance, tolerance, trophic status, habits), (2) respond in a predictable way to human-influenced disturbances, (3) low coefficients of variation (C.V.< 50%) and (4) means at unimpaired reference sites high enough to show a change caused by pollution or environmental stress (Karr et al. 1986, Karr 1991, Barbour et al. 1995, Smith and Voshell 1997). We used the reference stations on Peak Creek and Wolf Creek and the relatively unimpaired Peak Creek Stations 13-17 to calculate the means and C.V. for 35 candidate metrics. The 10 metrics that our analysis indicated would be the best to compare the two sampling methods are listed and explained in Table 1. The percentage and ratio metrics were transformed using the arcsine transformation (Krebs 1987). Statistical analyses included comparing the community composition estimated by the sampling methods with unpaired t-tests on individual metrics and correlation analysis at Peak Creek Stations 1-18.

TABLE 1. Definitions of benthic macroinvertebrate community metrics used in sampling method comparison.

Category	Metric	Definition
Richness measures	Taxa richness	Number of total taxa. Measures the collective variety of the community.
	EPT index	Number of taxa in the insect orders Ephemeroptera, Plecoptera, and Trichoptera. These orders are generally considered to be sensitive to a wide variety of impairments.
Composition measures	% EPT abundance	Percent abundance of the insect orders Ephemeroptera, Plecoptera, and Trichoptera. These orders are generally considered to be sensitive to a wide variety of impairments.
Balance measures	% 5 most dominant taxa	Measures the dominance of the 5 most abundant taxa. The greater the percentage, the greater the redundancy of taxa in the assemblage.
	Hydropsychidae/Trichoptera	Percentage of the moderately pollution-tolerant caddisfly family Hydropsychidae to total Trichoptera.
	Simpson's index of diversity	Integrates taxa richness and evenness into a measure of general diversity.
Tolerance measures*	modified HBI	Pollution tolerance values range from 0 to 10, with 0 representing complete intolerance of pollution. The index summarizes the overall tolerance of the community.
	% Intolerant	Pollution tolerance values range from 0 to 10, with 0 representing complete intolerance of pollution. Percent abundance of taxa with tolerance values equal to or less than 5.
Trophic Status measures	% Collector-gatherers	Percent abundance of collector gatherer functional feeding group.
Habit measures	% Haptobenthos	Percent abundance of taxa requiring clean coarse substrate.

* Pollution tolerance values were modifed from Hilsenhoff (1977, 1982, 1987a) and USEPA (1973) using the best professional judgment of the Aquatic Entomology Group at VPI&SU and biologists from the Virginia Department of Environmental Quality. These values were adjusted specifically for use in Virginia and surrounding states and to make them sensitive to a greater range of perturbations than just organic enrichment (J.R. Voshell, Jr. Virginia Polytechnic Institute and State University, personal communication).

Biological assessment.- To determine if the two sampling methods make the same assessment of biological condition, we compared each station along Peak Creek with a 95% confidence interval based on the five Peak Creek reference station samples obtained by each method. Relying on information gathered from biosurveys, chemical analyses, and visually-based habitat assessments by the Virginia Department of Environmental Quality, we classified Peak Creek Stations 1-12 and 18 as impacted and Stations 13-17 as relatively unimpaired. Taxa richness, EPT index, % EPT, Simpson's index of diversity, % Intolerant, % Haptobenthos, the Bray-Curtis Coefficient, and the Index of Biotic Similarity all decrease in response to increasing perturbations. Values for these metrics should fall below the reference interval at the impaired stations and within or above the reference interval at the relatively unimpaired stations. Modified HBI and % contribution of the five most dominant taxa increase in response to increasing perturbations. These metrics should rise above the reference confidence interval at the impaired stations and within or below the reference interval at the unimpaired stations. One metric, % collector-gatherers has a variable response to perturbed conditions. Therefore, we considered any value above or below the confidence interval to be impaired and within the confidence interval as unimpacted. We compared the assessments made by each method by counting the number of stations along Peak Creek correctly assessed as impacted or unimpacted. We also counted the total number of stations, both impacted and unimpacted, where each method made the same assessment.

We also compared assessments at each site using the Macroinvertebrate Aggregated Index for Streams (MAIS) (Smith and Voshell 1997). This family-level multimetric index was developed especially for streams in the mid-Atlantic highlands area. The MAIS combines nine benthic macroinvertebrate community metrics into a standardized scoring system with a maximum score of 18 (Tables 2-3). In an evaluation of the MAIS within the Central Appalachian Ridges andValleys ecoregion, Smith and Voshell (1997) found that the MAIS correctly classified reference sites 92% of the time and impaired sites 90% of the time.

TABLE 2. Metric category, metric and bioassessment scoring values of the Macroinvertebrate Aggregated Index for Streams (MAIS) (Smith and Voshell 1997).

	Bioassessment scoring		
Metric	2	1	0
Richness			
EPT index	≥ 8	3-7	≤ 2
# Ephemeroptera	≥ 4	1-3	0
Composition			
% Ephemeroptera	≥ 18	1-17	0
Balance			
% 5 dominant taxa	≤ 79	80-99	100
Simpson's index of diversity	≥ 0.83	0.67-0.82	≤ 0.66
Tolerance			
Modified HBI	≤ 4.21	4.22-5.55	≥ 5.56
# Intolerant taxa	≥ 10	2-9	≤ 1
Trophic status			
% Scrapers	≥ 11	1-10	0
Habit			
% Haptobenthos	≥ 84	52-83	≤ 51

TABLE 3. Macroinvertebrate Aggregated Index for Streams (MAIS) Bioassessment criteria for the Central Appalachian Ridges and Valleys ecoregion (Smith and Voshell 1997).

Biological Condition Category	Total Score
Acceptable	
Very Good	≥ 17
Good	13-16
Unacceptable	
Poor	7-12
Very Poor	≤ 6

Results

Physical and chemical data

Temperature and dissolved oxygen were lower at the downstream stations, reached their peak at the midstream stations, and then decreased again at the upstream stations (Fig. 2). We attribute these differences to natural changes in solar radiation during the day and our sampling schedule. We began sampling at Station 1 at 0600 hours and worked our way upstream throughout the day. We reached the middle stations around 1200 hours and sampled the last station around 2100 hours.

Alkalinity, pH, hardness, and conductivity were all higher at the downstream sites 1-12 (Figs. 3 & 4). These sites may represent a transition area between subregions of the ecoregion. Along its length, Peak Creek moves from the Sandstone Ridges subregion (67c) to the Limestone Valleys subregion (67a) within the Central Appalachian Ridges and Valleys ecoregion. Conductivity levels increased dramatically between Peak Creek Stations 12 and 13. Magnox, with its permitted discharge of heavy metals, lies between these two stations. A second peak occurred below Station 7 from the output of the Allied Chemical plant (Fig. 4). Habitat assessment scores were substantially lower at stations 5-12. These sites were located in downtown Pulaski and reflect the influence of urbanization on Peak Creek (Fig. 5). These sites had reduced substrate heterogeneity, reduced mean substrate size, increased sedimentation, some channelization and degraded riparian areas along the banks.

Community composition

We found that for the Index of Biotic Similarity and the Bray-Curtis Coefficient, the mean similarity of the KSS samples was greater than the mean similarity of the PIBS samples at either reference site (Table 4). The mean similarity between the PIBS samples and the KSS samples was considerably less, especially at Wolf Creek, than the mean similarities of the

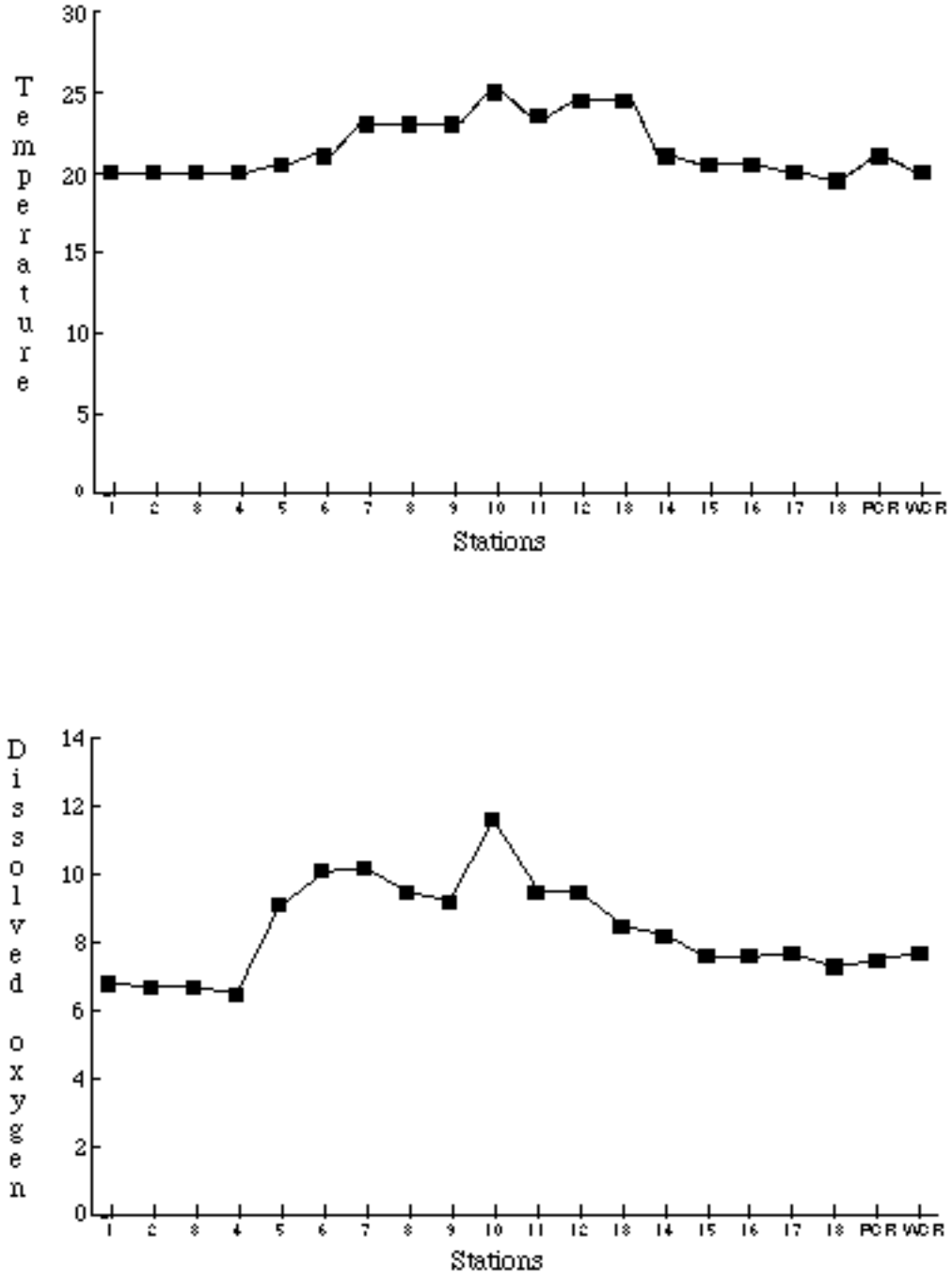

FIG. 2. Temperature (°C) and dissolved oxygen (mg/L) at Peak Creek Stations 1-18, Peak Creek Reference Station (PCR), and Wolf Creek Reference Station (WCR).

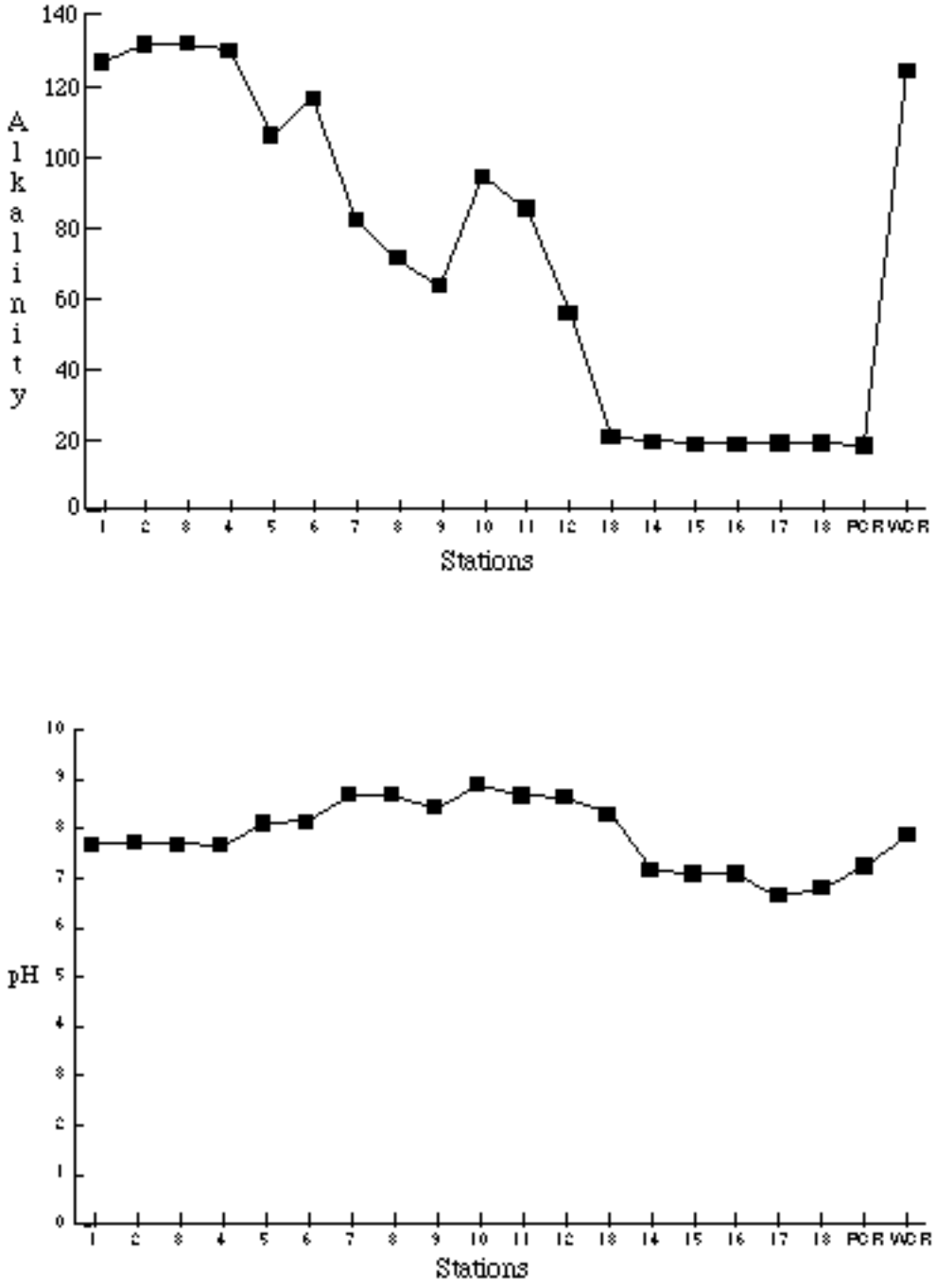

FIG. 3. Alkalinity (mg/L $CaCO_3$) and pH at Peak Creek Stations 1-18, Peak Creek Reference Station (PCR), and Wolf Creek Reference Station (WCR).

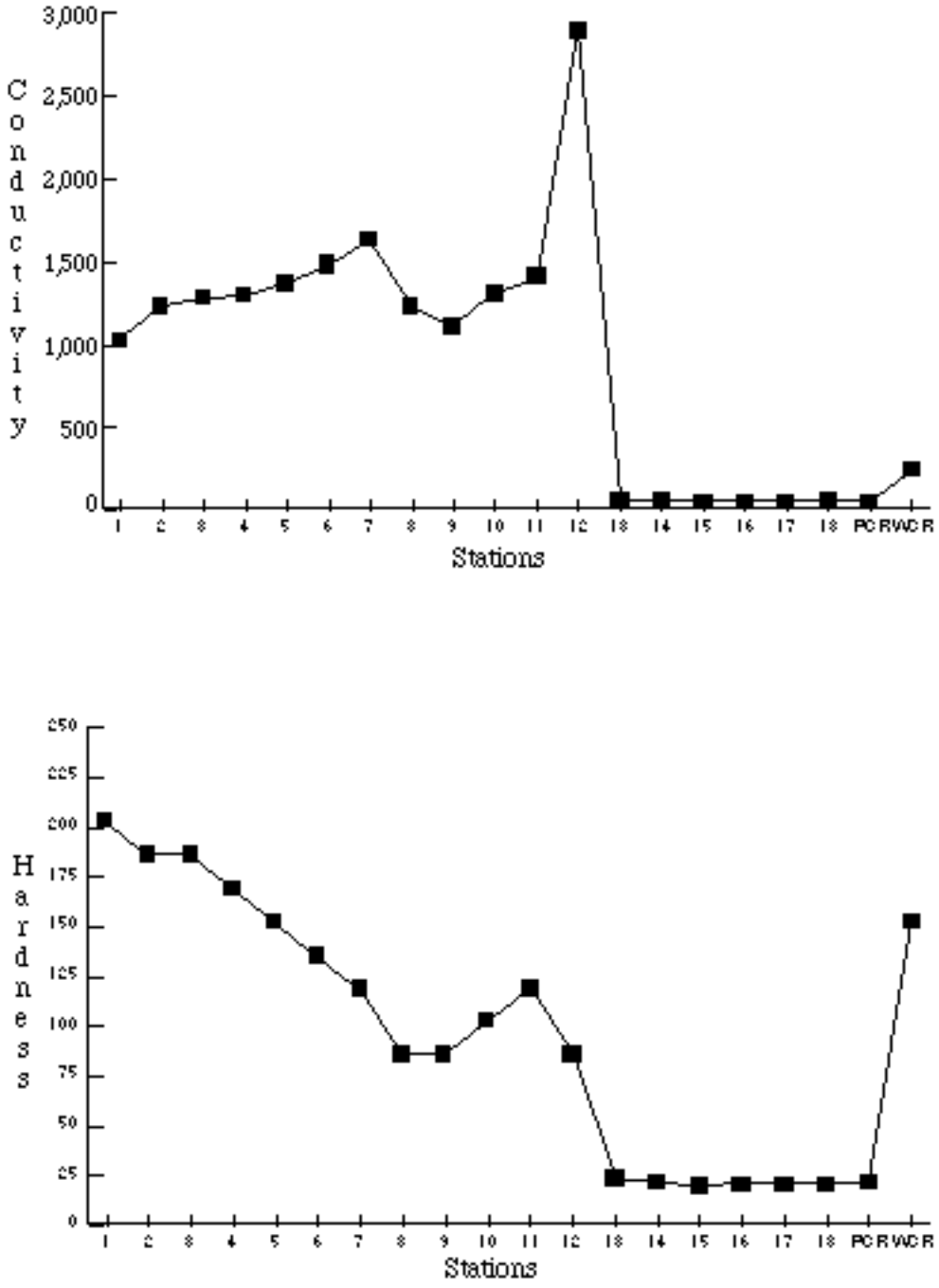

FIG. 4. Conductivity (μmhos) and hardness (mg/L $CaCO_3$) at Peak Creek Stations 1-18, Peak Creek Reference Station (PCR), and Wolf Creek Reference Station (WCR).

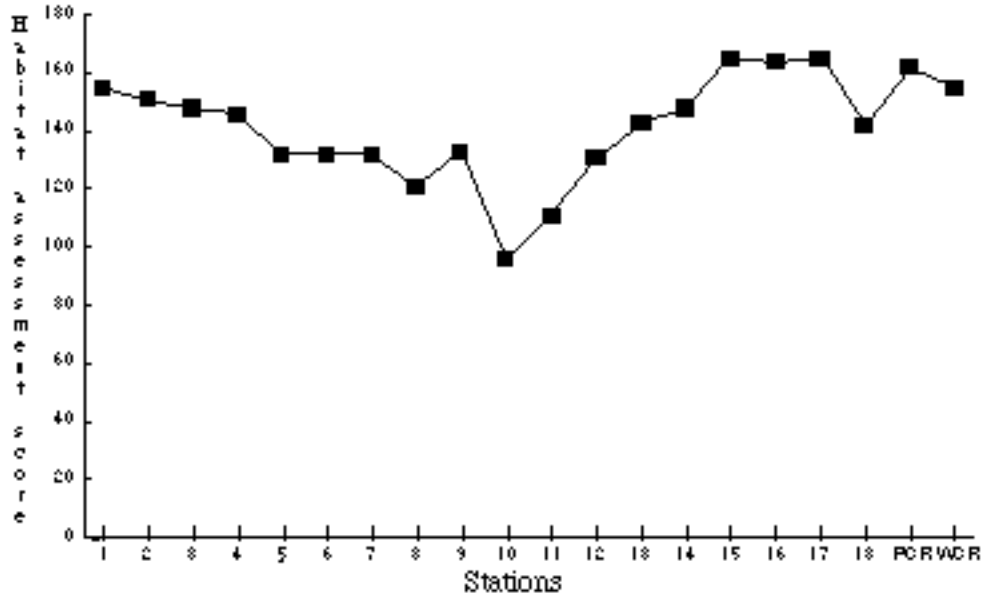

FIG. 5. Habitat Assessment Score at Peak Creek Stations 1-18, Peak Creek Reference Station (PCR), and Wolf Creek Reference Station (WCR).

TABLE 4. Mean similarities within and between PIBS and KSS using Index of Biotic Similarity and Bray-Curtis Coefficient at Peak Creek Reference and Wolf Creek Reference.

	PIBS	KSS	PIBS vs. KSS
Peak Creek Reference			
Index of Biotic Similarity	27.45	34.15	21.19
Bray-Curtis Coefficient	56.83	83.44	43.58
Wolf Creek Reference			
Index of Biotic Similarity	33.14	38.70	13.71
Bray-Curtis Coefficient	67.13	73.32	23.12

methods themselves at both reference sites. The Index of Biotic Similarity and the Bray-Curtis Coefficient are strongly affected by sample size. Wolda (1981) showed that for the Bray-Curtis Coefficient (Wolda did not examine the Index of Biotic Similarity) the expected maximum values decrease as differences in sample size increase. By interpolating the graphs in Wolda (1981), we estimated the expected maximum similarities for each method at each reference site. At Peak Creek Reference Station, the expected maximum similarity for the PIBS, KSS, and PIBS vs. KSS was 80.0, 78.0 and 40.0 respectively. The expected maximum similarity for the PIBS, KSS, and PIBS vs. KSS at Wolf Creek was 90.0, 78.0, and 20.0 respectively. With the possible exception of the PIBS at Wolf Creek, we found the mean similarities were very close to the expected maximum values.

Using unpaired t-tests on 10 metrics to compare PIBS and KSS samples at the Peak Creek Reference Station resulted in 3 metrics showing significant differences ($P \leq 0.05$) (Table 5). Likewise, two metrics had significant differences between the sampling methods at the Wolf Creek Reference Station. The PIBS samples had significantly more taxa at both the Peak Creek and Wolf Creek Reference Stations. At the Peak Creek Reference Station, there was no significant difference between the sampling methods for the EPT Index. The mean abundance of organisms per PIBS sample was twice as great in Wolf Creek than at the Peak Creek Reference Station. Wolf Creek was also more diverse. The greater abundance combined with greater diversity at Wolf Creek may explain why the PIBS samples from Wolf Creek had significantly more EPT taxa than the KSS samples.

At the Peak Creek Reference Station, the PIBS samples had significantly higher modified HBI values and % collector-gatherer abundance. We found substantially greater abundances of Baetidae and Chironomidae in the PIBS samples, which may explain the higher modified HBI values and % collector-gatherer abundance. Baetidae is considered fairly pollution tolerant and

TABLE 5. Mean (± 95% confidence interval) of 10 benthic macroinvertebrate community metrics at Peak Creek Reference and Wolf Creek Reference. Two-sample t-test on Peak Creek Reference Station PIBS samples (*n*=5) versus Peak Creek Reference Station KSS samples (*n*=5) and Wolf Creek Reference Station (*n*=5). Values with different letters indicate means that are significantly different ($P \leq 0.05$).

	Peak Creek Reference		Wolf Creek Reference	
Metric	PIBS	KSS	PIBS	KSS
Taxa richness	25.6 (2.02)a	21.2 (2.09)b	41.2 (2.93)a	27.2 (2.00)b
EPT index	13.4 (1.59)a	12.8 (1.14)a	24.0 (2.56)a	16.2 (2.27)b
% EPT	42.8 (7.08)a	50.3 (2.44)a	48.4 (4.35)a	52.8 (3.50)a
% 5 most dominant taxa	63.1 (2.23)a	64.6 (4.45)a	50.5 (1.80)a	53.2 (1.64)a
Hydropsychidae/ Trichoptera	41.4 (5.55)a	35.1 (3.43)a	42.0 (6.45)a	51.8 (6.57)a
Simpson's index of diversity	0.74 (0.12)a	0.80 (0.04)a	0.90 (0.02)a	0.90 (0.02)a
modified HBI	4.98 (0.24)a	4.58 (0.17)b	4.02 (0.26)a	4.06 (0.16)a
% Intolerant	37.5 (5.41)a	43.0 (2.62)a	63.7 (5.37)a	61.1 (3.06)a
% Collector-gatherers	49.1 (7.82)a	35.2 (2.43)b	31.0 (4.61)a	27.6 (3.88)a
% Haptobenthos	48.8 (8.97)a	56.9 (3.63)a	73.9 (3.65)a	79.5 (3.35)a

Chironomidae is highly pollution tolerant. These taxa are both collector-gatherers (Merritt and Cummins 1984).

A correlation analysis on the metrics was performed to compare the methods along Peak Creek Stations 1-18 (Table 6). A high correlation suggested a high level of similarity between sample communities. We designated correlation coefficients ≥7 as highly similar as was done by Barbour et al. (1992). Many of the metrics showed moderate (~0.50) correlation between the methods. However, the strong correlation between the sampling methods for EPT index (0.84), modified HBI (0.69), % intolerant (0.73), Index of Biotic Similarity (0.82) and Bray-Curtis Coefficient (0.80) indicated the sample communities were highly similar.

Biological assessment

Results of analyses to determine if the two sampling methods lead to the same assessment of biological condition are presented in Figs. 6-17 for individual metrics at each site by each sampling method. The assumption is that the two sampling methods indicate the same assessment of biological condition if the metric values for the PIBS and KSS are located in the same position relative to the 95% confidence limits of samples taken at the reference site (i.e., either both metric values within 95% confidence limits or both metric values outside of 95% confidence limits). Results are summarized in Table 7, in terms of number of sites correctly assessed.

The agreement of assessments made by each sampling method for each metric ranged from 9 (50%) to 17 (94%) of 18 sites and averaged 13 (72%) of 18 sites. The assessments made using the KSS correctly classified an average of 10 (77%) out of the 13 impacted sites and outperformed the assessments made with the PIBS (9 [69%] of 13 impacted sites). However, the assessments made using the PIBS (3.3 [66%] of 5 unimpacted sites) surpassed the assessments made with the KSS at the unimpacted stations (2.6 [53%] of 5 unimpacted sites). Some metrics

TABLE 6. Correlation between sampling methods (PIBS and KSS) based on 12 benthic macroinvertebrate community metrics from Peak Creek Stations (1-18).

Metric	*r*	*P*-value
Taxa richness	0.3365	0.1721
EPT index	0.8368	< 0.001
% EPT	0.4781	0.0448
% 5 most dominant taxa	0.5194	0.0272
Hydropsychidae/Trichoptera	0.4540	0.0584
Simpson's index of diversity	0.4870	0.0404
modified HBI	0.6905	0.0015
% Intolerant	0.7337	0.0005
% Collector-gatherers	0.5123	0.0297
% Haptobenthos	0.2625	0.2925
Index of Biotic Similarity	0.8177	< 0.001
Bray-Curtis Coefficient	0.8025	< 0.001

TABLE 7. Number of stations each method correctly assessed as impaired (out of 13) and unimpaired (out of 5) and total number of stations where each method made the same assessment.

Metric	Impaired Stations (1-12,18)		Unimpaired Stations (13-17)		Total no. of stations in agreement (out of 18)
	PIBS (out of 13)	KSS (out of 13)	PIBS (out of 5)	KSS (out of 5)	
Taxa richness	8	10	4	3	9
EPT Index	13	13	3	3	14
% EPT	1	7	5	4	9
% 5 most dominant taxa	12	8	4	4	12
Hydropsychidae/Trichoptera	12	13	1	1	15
Simpson's index of diversity	2	5	5	5	13
modified HBI	12	13	4	2	15
% Intolerant	12	13	4	2	15
% Collector-gatherers	10	10	2	2	9
% Haptobenthos	1	4	5	5	15
Index of Biotic Similarity	13	13	2	1	16
Bray-Curtis Coefficient	13	13	1	0	17

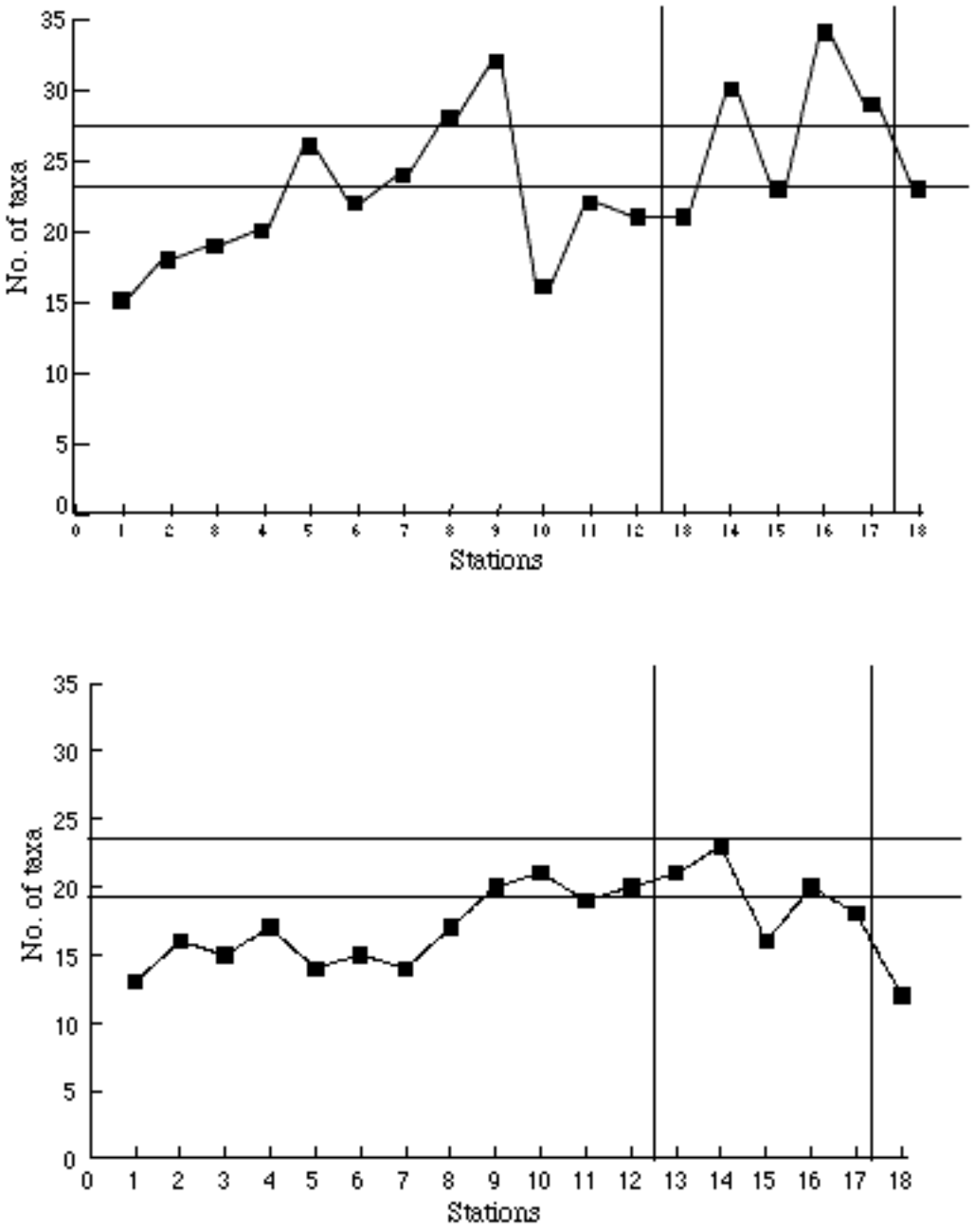

FIG. 6. Comparison of taxa richness along Peak Creek Stations 1-18 with the reference condition for each method (PIBS top panel, KSS bottom panel). Horizontal lines denote 95% confidence interval calculated from the 5 PIBS or KSS samples taken from the Peak Creek Reference Station. Vertical lines separate impaired stations (1-12,18) from relatively unimpaired stations (13-17).

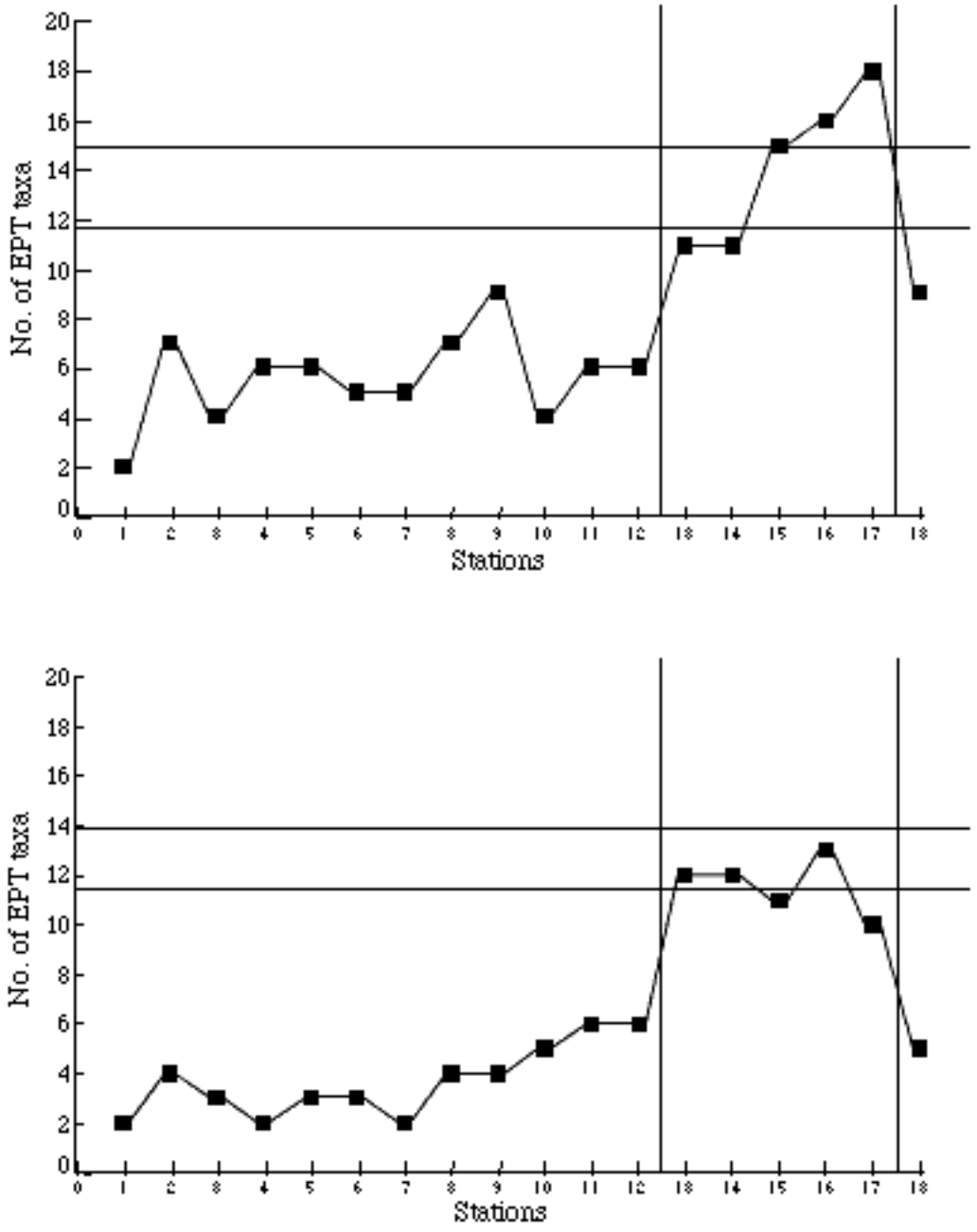

FIG. 7. Comparison of EPT index along Peak Creek Stations 1-18 with the reference condition for each method (PIBS top panel, KSS bottom panel). Horizontal lines denote 95% confidence interval calculated from the 5 PIBS or KSS samples taken from the Peak Creek Reference Station. Vertical lines separate impaired stations (1-12,18) from relatively unimpaired stations (13-17).

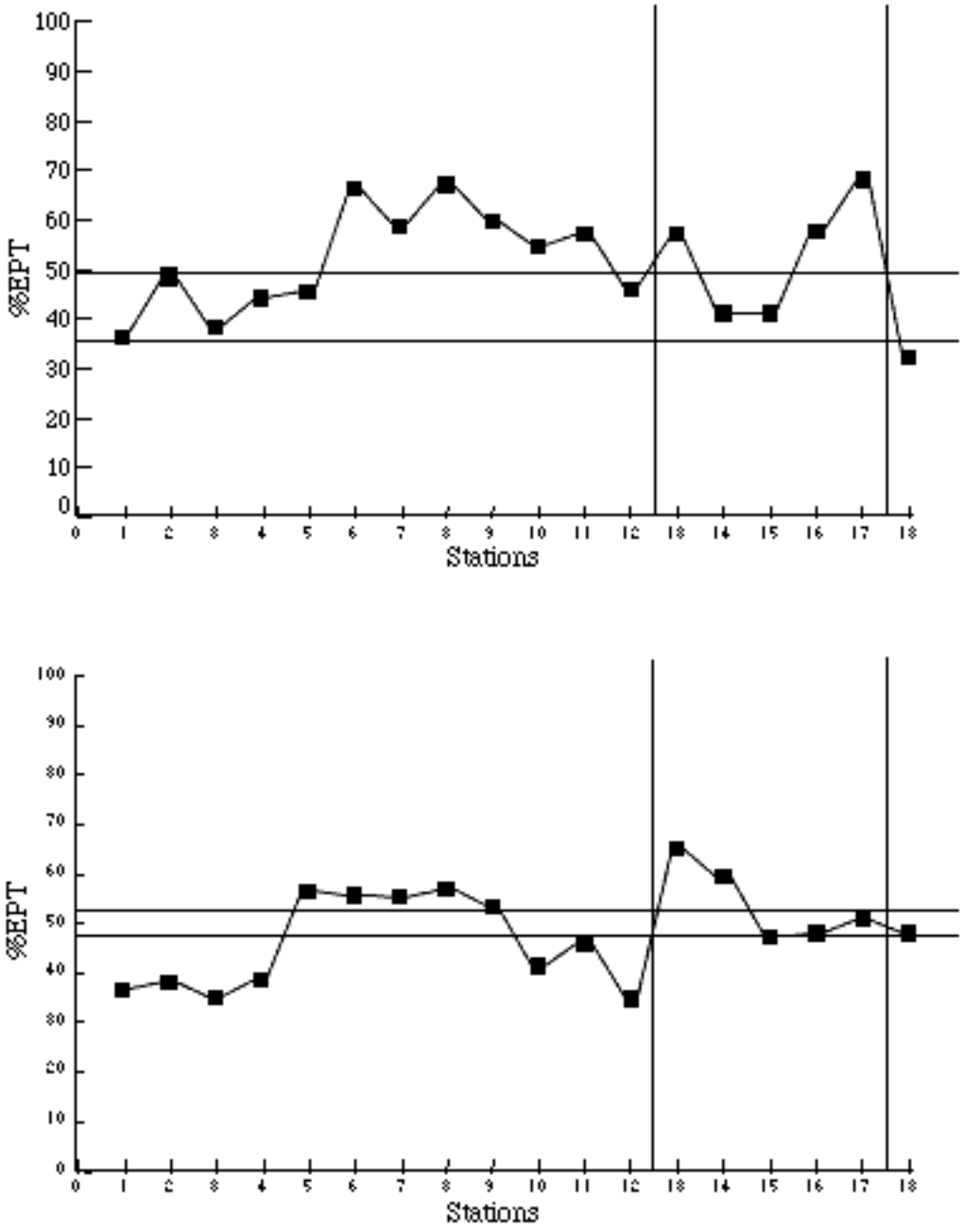

FIG. 8. Comparison of % EPT along Peak Creek Stations 1-18 with the reference condition for each method (PIBS top panel, KSS bottom panel). Horizontal lines denote 95% confidence interval calculated from the 5 PIBS or KSS samples taken from the Peak Creek Reference Station. Vertical lines separate impaired stations (1-12,18) from relatively unimpaired stations (13-17).

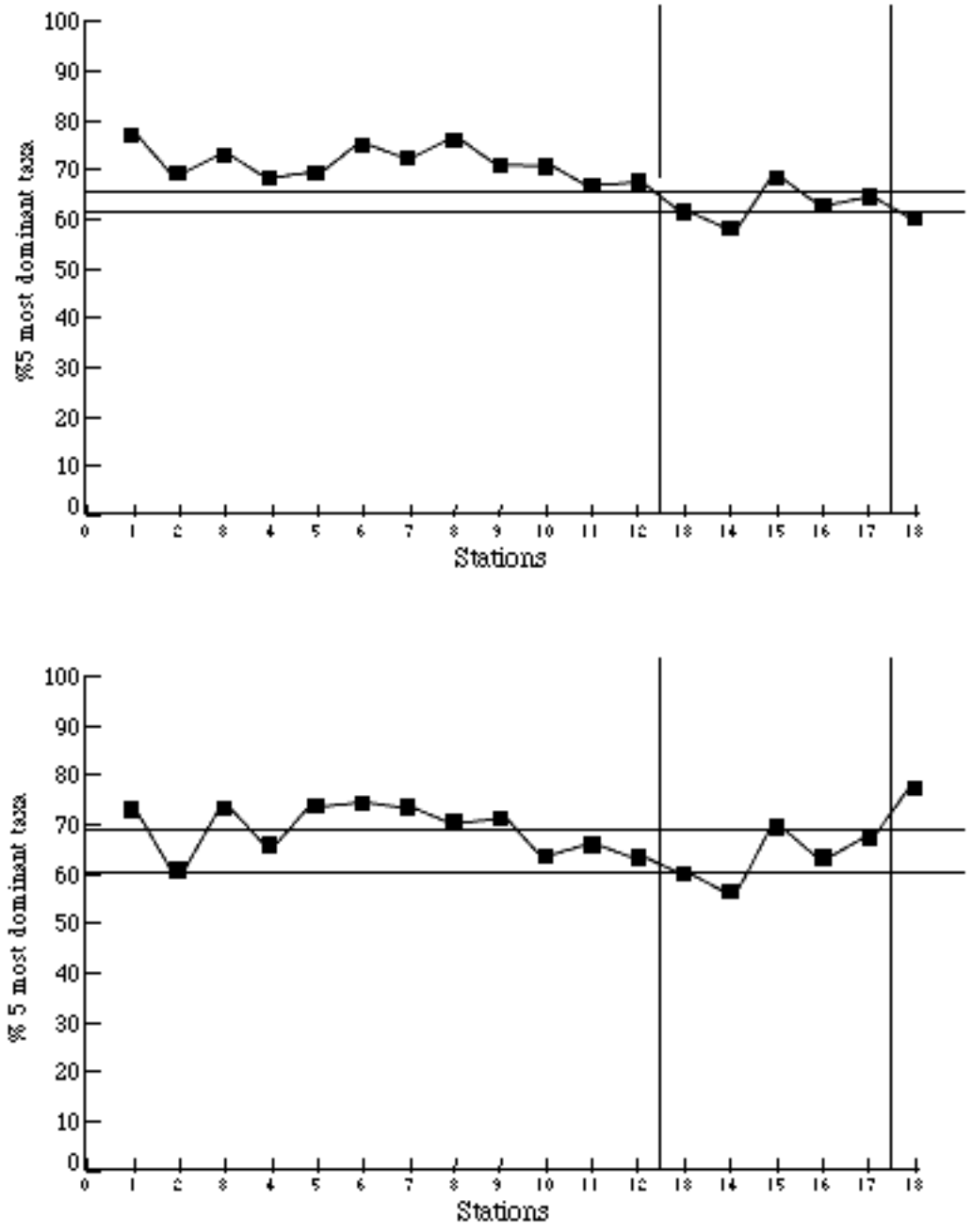

FIG. 9. Comparison of % 5 most dominant taxa along Peak Creek Stations 1-18 with the reference condition for each method (PIBS top panel, KSS bottom panel). Horizontal lines denote 95% confidence interval calculated from the 5 PIBS or KSS samples taken from the Peak Creek Reference Station. Vertical lines separate impaired stations (1-12,18) from relatively unimpaired stations (13-17).

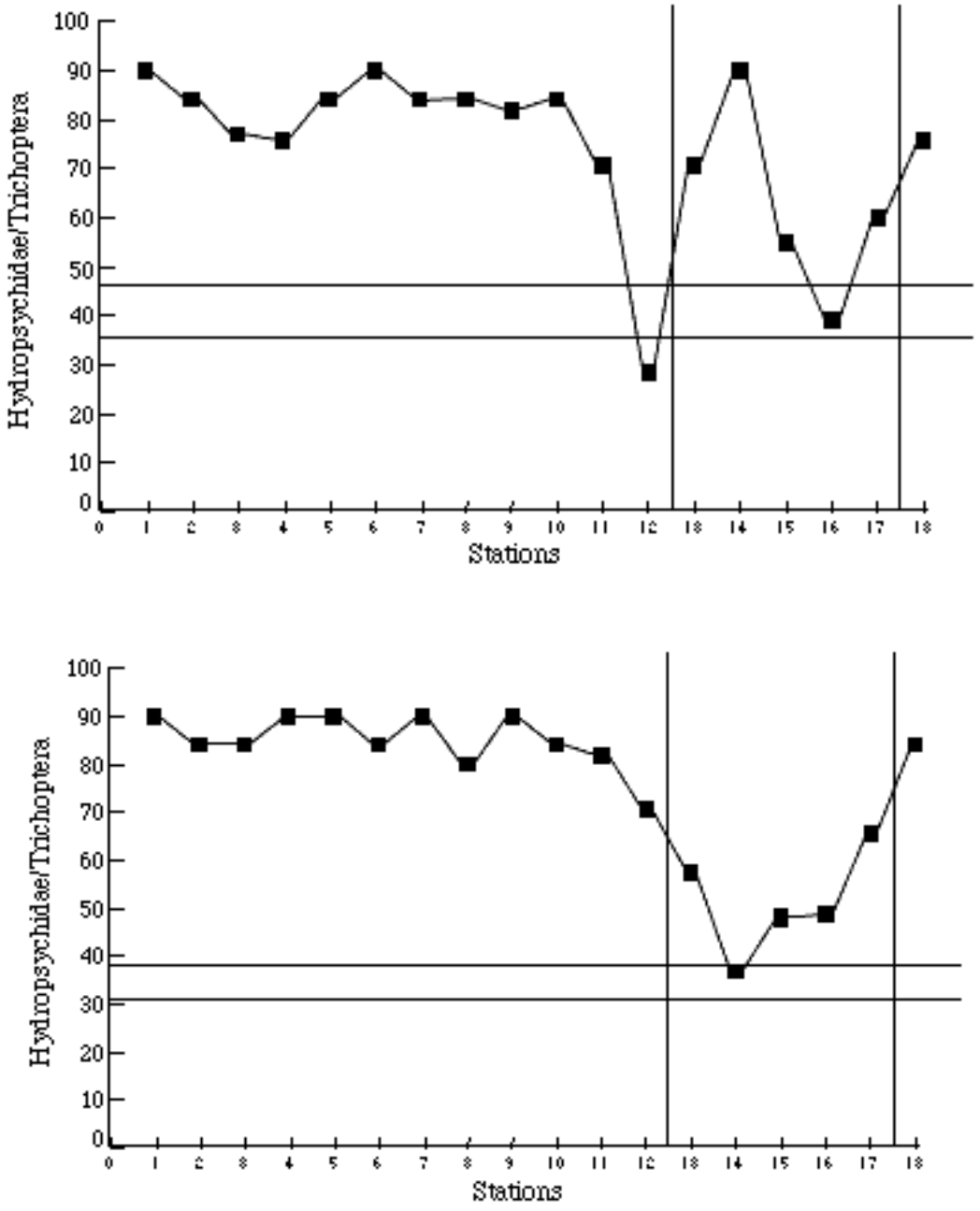

FIG. 10. Comparison of Hydropsychidae/Trichoptera along Peak Creek Stations 1-18 with the reference condition for each method (PIBS top panel, KSS bottom panel). Horizontal lines denote 95% confidence interval calculated from the 5 PIBS or KSS samples taken from the Peak Creek Reference Station. Vertical lines separate impaired stations (1-12,18) from relatively unimpaired stations (13-17).

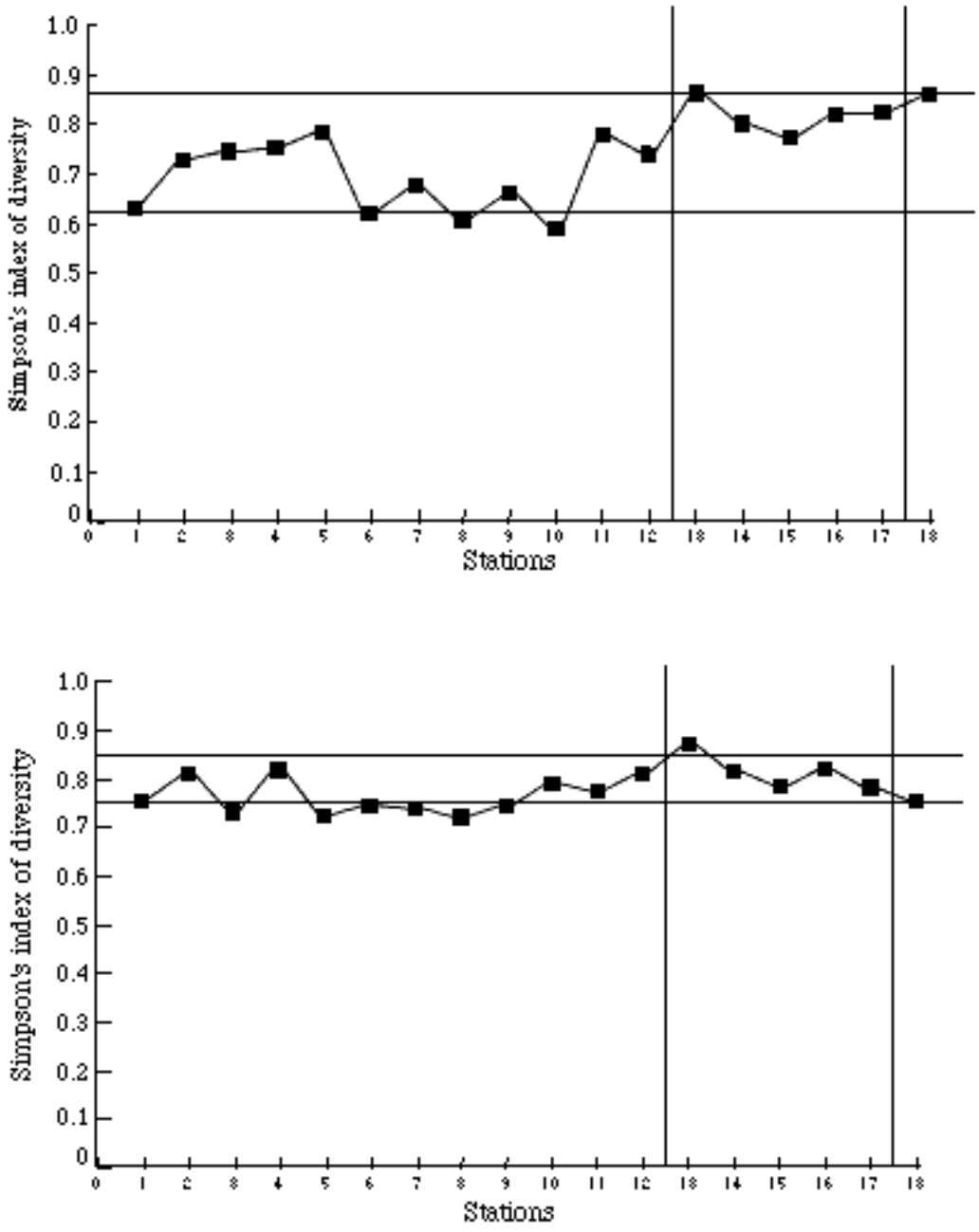

FIG. 11. Comparison of Simpson's index of diversity along Peak Creek Stations 1-18 with the reference condition for each method (PIBS top panel, KSS bottom panel). Horizontal lines denote 95% confidence interval calculated from the 5 PIBS or KSS samples taken from the Peak Creek Reference Station. Vertical lines separate impaired stations (1-12,18) from relatively unimpaired stations (13-17).

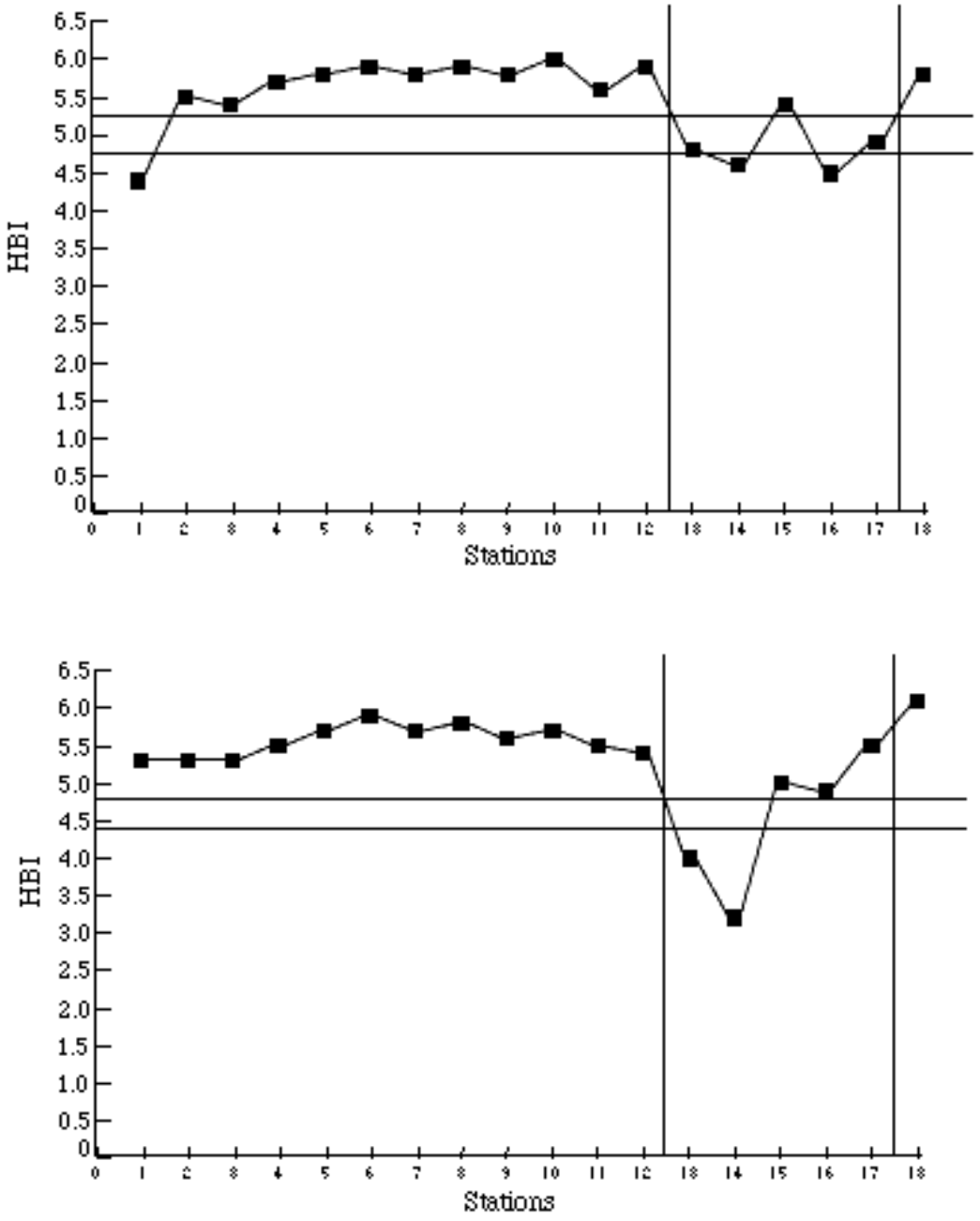

FIG. 12. Comparison of HBI along Peak Creek Stations 1-18 with the reference condition for each method (PIBS top panel, KSS bottom panel). Horizontal lines denote 95% confidence interval calculated from the 5 PIBS or KSS samples taken from the Peak Creek Reference Station. Vertical lines separate impaired stations (1-12,18) from relatively unimpaired stations (13-17).

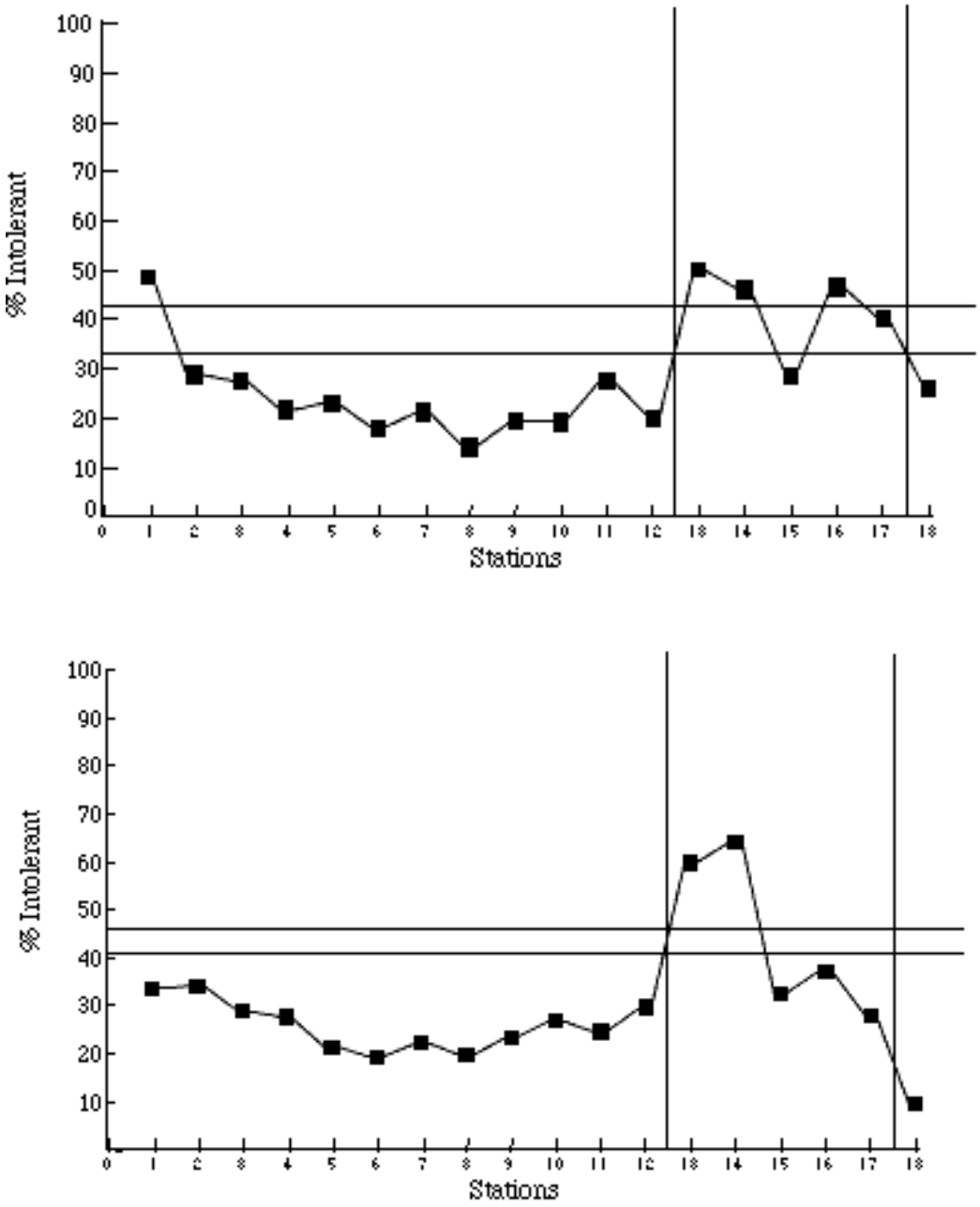

FIG. 13. Comparison of % Intolerant along Peak Creek Stations 1-18 with the reference condition for each method (PIBS top panel, KSS bottom panel). Horizontal lines denote 95% confidence interval calculated from the 5 PIBS or KSS samples taken from the Peak Creek Reference Station. Vertical lines separate impaired stations (1-12,18) from relatively unimpaired stations (13-17).

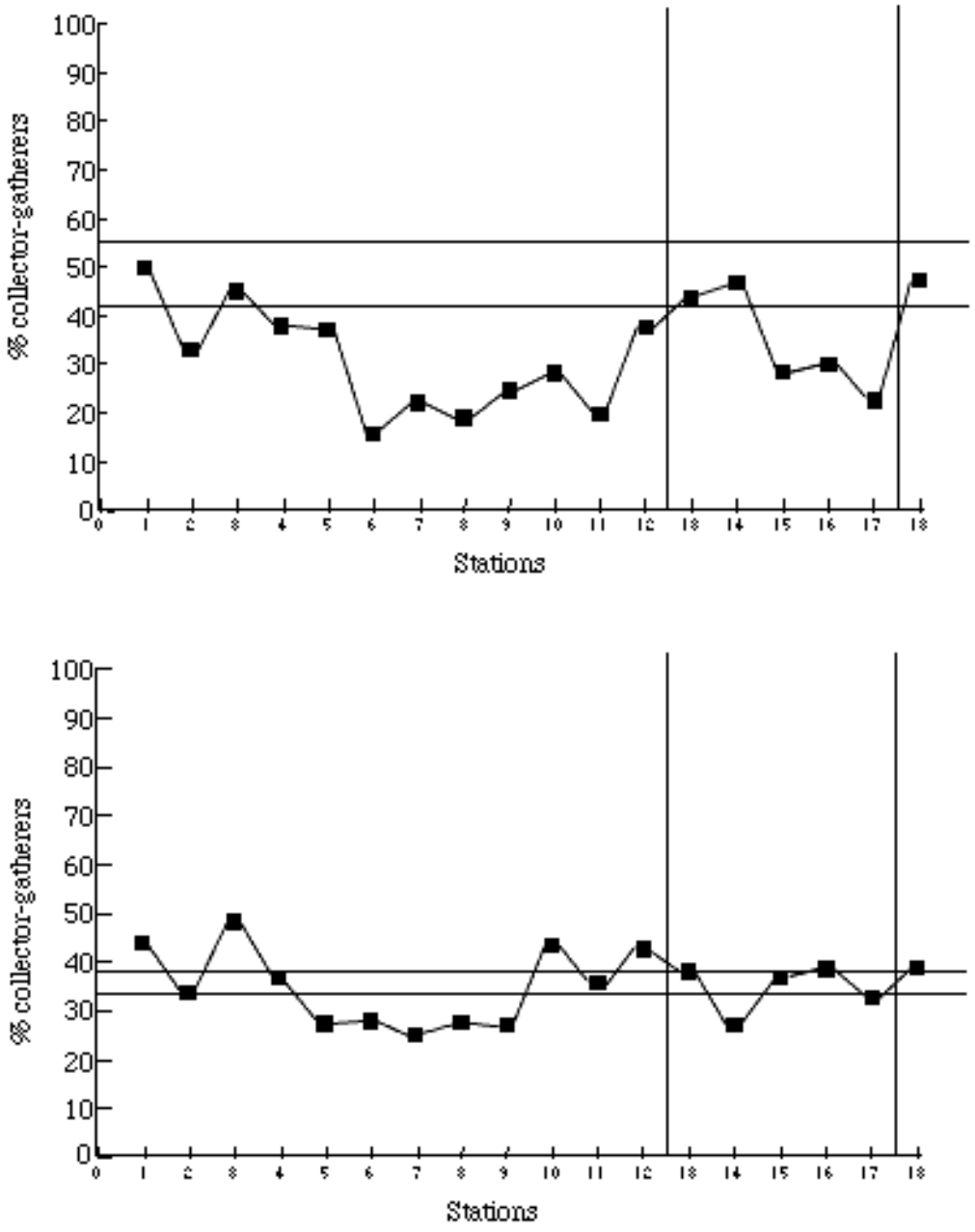

FIG. 14. Comparison of % collector-gatherers along Peak Creek Stations 1-18 with the reference condition for each method (PIBS top panel, KSS bottom panel). Horizontal lines denote 95% confidence interval calculated from the 5 PIBS or KSS samples taken from the Peak Creek Reference Station. Vertical lines separate impaired stations (1-12,18) from relatively unimpaired stations (13-17).

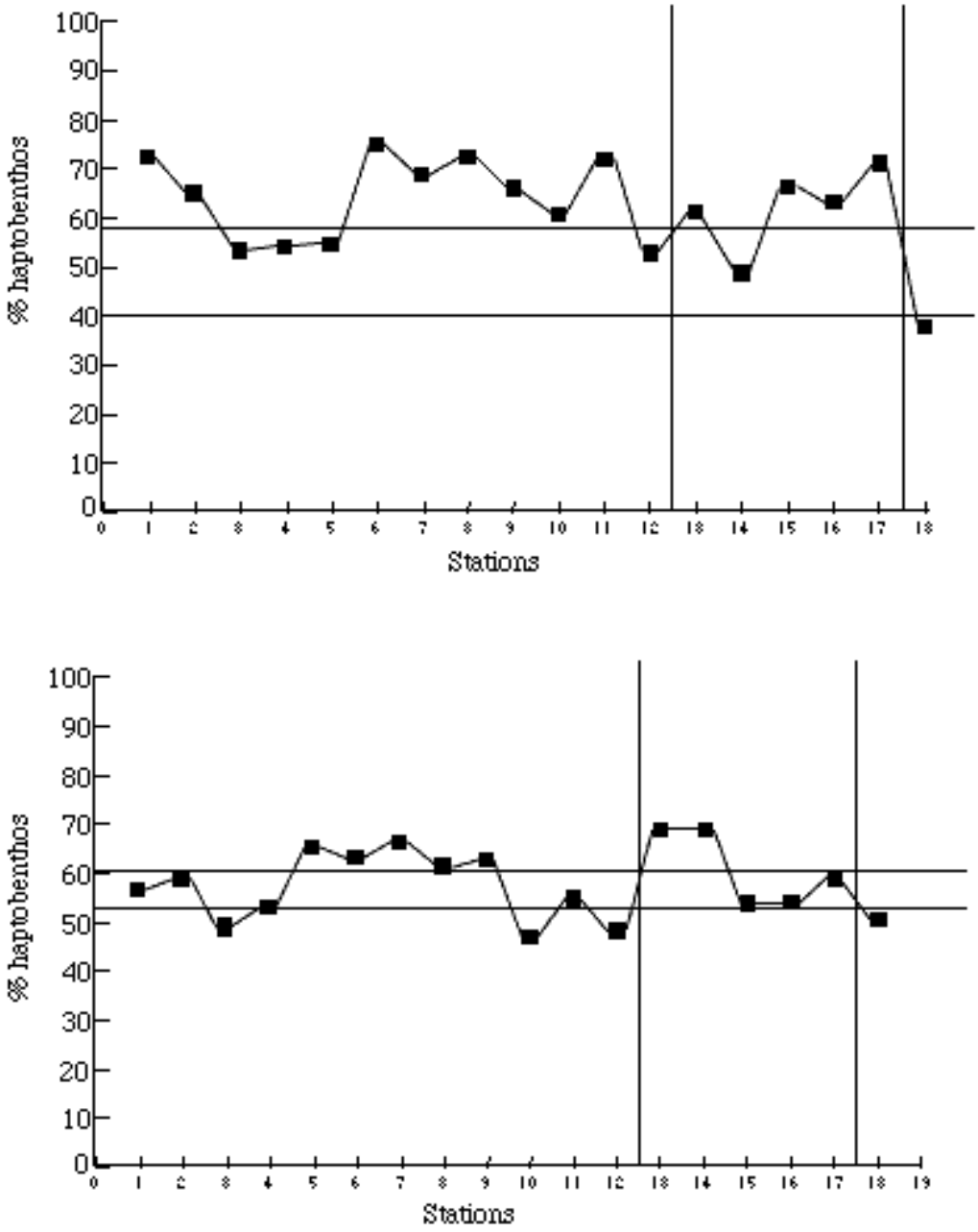

FIG. 15. Comparison of % haptobenthos along Peak Creek Stations 1-18 with the reference condition for each method (PIBS top panel, KSS bottom panel). Horizontal lines denote 95% confidence interval calculated from the 5 PIBS or KSS samples taken from the Peak Creek Reference Station. Vertical lines separate impaired stations (1-12,18) from relatively unimpaired stations (13-17).

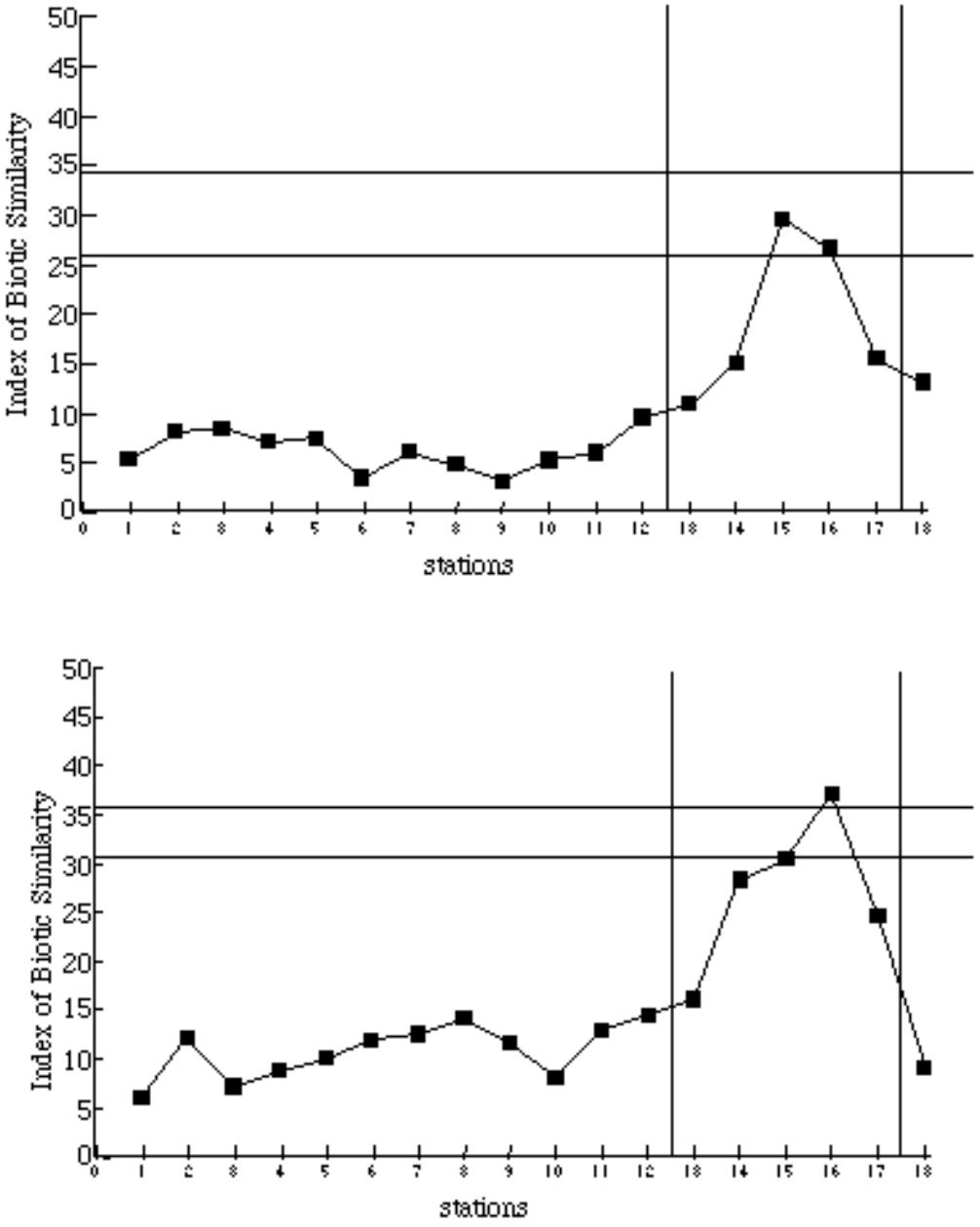

FIG. 16. Comparison of Index of Biotic Similarity along Peak Creek Stations 1-18 with the reference condition for each method (PIBS top panel, KSS bottom panel). Horizontal lines denote 95% confidence interval calculated from the 5 PIBS or KSS samples taken from the Peak Creek Reference Station. Vertical lines separate impaired stations (1-12,18) from relatively unimpaired stations (13-17).

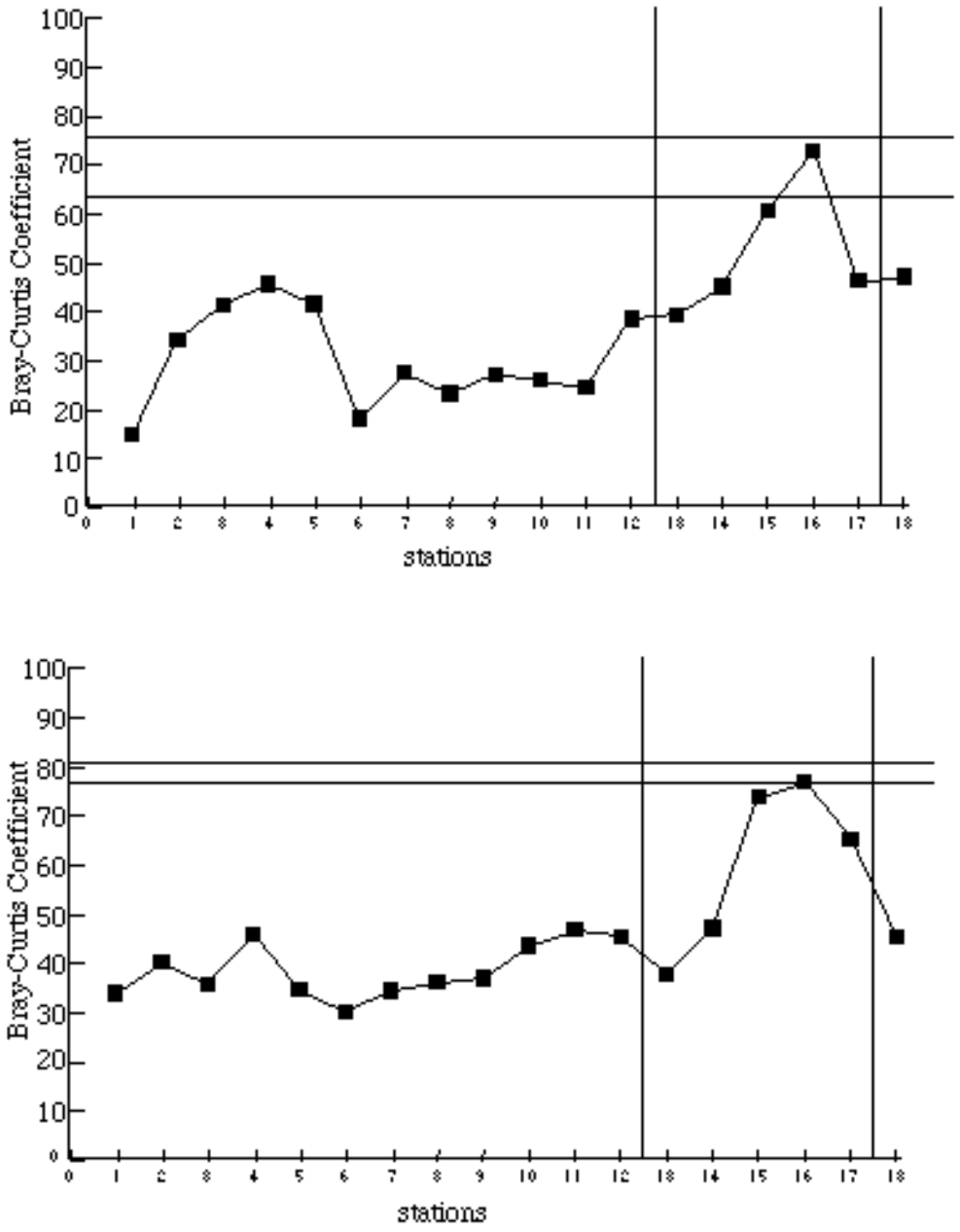

FIG. 17. Comparison of Bray-Curtis Coefficient along Peak Creek Stations 1-18 with the reference condition for each method (PIBS top panel, KSS bottom panel). Horizontal lines denote 95% confidence interval calculated from the 5 PIBS or KSS samples taken from the Peak Creek Reference Station. Vertical lines separate impaired stations (1-12,18) from relatively unimpaired stations (13-17).

were better than others in detecting impairment (Figs 6-17). EPT index, modified HBI, % Intolerant, Index of Biotic Similarity, and Bray-Curtis Coefficient had the best separation between the impacted and unimpacted stations. Taxa richness, % EPT and Simpson's index of diversity, changed very little between impacted and unimpacted conditions. For % EPT, the pollution-tolerant family Hydropsychidae was sometimes very abundant along the perturbed stations of Peak Creek, making this metric not particularly reliable in detecting differences between the communities. We found a strong relationship between the number of stations in agreement and the sensitivity of metrics to impairment. The most sensitive metrics, EPT index, HBI, % Intolerant, Index of Biotic Similarity and Bray-Curtis Coefficient, also had the greatest number of stations in agreement.

We also compared the MAIS scores at each station along Peak Creek (Table 8). According to Smith and Voshell (1997), MAIS scores ≥13 reflect acceptable biological conditions and scores ≤12 reflect unacceptable biological conditions. We found that the MAIS scores from each sampling method indicated unacceptable conditions at all 13 of the impacted sites. There was a disagreement between the methods at two of the five unimpacted sites. KSS assessed stations 15 and 17 as unacceptable. Nevertheless, the two sampling methods agreed regarding acceptable and unacceptable biological conditions 89% of the time.

Discussion

Community composition

The slight differences found in the metrics and similarity indices are most likely due to differences in abundance of organisms obtained by the two sampling methods. There was always a greater abundance of organisms found in the PIBS samples. In our study, the PIBS ranged from 246 to 2000 organisms, while the subsampling limited the KSS samples to approximately 200 organisms. Richness metrics, such as taxa richness and EPT Index, are

TABLE 8. Comparison of MAIS scores and Biological Condition Categories (A=acceptable, U=unacceptable) at Peak Creek Stations 1-18 and Reference Station.

Peak Creek Station	MAIS Score (Biological Condition Category)	
	PIBS	KSS
1	6 (U)	6 (U)
2	6 (U)	7 (U)
3	7 (U)	6 (U)
4	6 (U)	6 (U)
5	7 (U)	4 (U)
6	8 (U)	4 (U)
7	7 (U)	6 (U)
8	7 (U)	6 (U)
9	7 (U)	7 (U)
10	4 (U)	9 (U)
11	9 (U)	8 (U)
12	6 (U)	8 (U)
13	13 (A)	17 (A)
14	15 (A)	16 (A)
15	13 (A)	9 (U)
16	12 (U)	11 (U)
17	15 (A)	10 (U)
18	9 (U)	5 (U)
Reference	12 (U)	13 (A)

strongly affected by sample abundance. The greater the sample size, the greater the chance of finding additional taxa, especially at reference sites. Therefore, the PIBS samples usually had higher values for these metrics. In addition, the expected maximum values of similarity indices, like the Bray-Curtis Coefficient, decrease as differences in sample size increase. In this study, we found the mean similarity between the sampling methods was very close to the expected maximum values of the Bray-Curtis Coefficient adjusted for sample abundance. Although the richness metrics show slight differences between the methods in relation to sample abundance, we conclude that, overall, the PIBS and KSS give similar estimates of benthic macroinvertebrate community composition.

We did find, at at least one station, a difference in the estimations of community composition not directly related to sample size. Besides sampling area, PIBS and KSS also differ in sampling effort. Quantitative sampling methods usually involve more intensive efforts to dislodge clinging and burrowing organisms in the relatively small sampling area. In this study we brushed each rock and raked the underlying substratum in the PIBS. Rapid bioassessment approaches are intended to be rapid in the field as well as the laboratory so much less effort is expended for removing clinging and burrowing organisms in favor of sampling a larger area containing more microhabitats. In this study we overturned cobbles and pebbles in front of the kick screen and very quickly rubbed them with our hands, after which we quickly ran our fingers through some of the gravel and sand.

The metric Hydropsychidae/Trichoptera reflected these differences in sampling intensity (Fig. 10). At station 12, the caddisfly *Leucotrichia* was extremely abundant. This small caddisfly cements its silk case firmly to rock surfaces. *Leucotrichia* made up 26.7% of the PIBS but only 2.2% of the KSS. The intensive nature of taking a PIBS sample, scrubbing and examining each rock versus simply rubbing the stones in KSS, resulted in a difference in community composition between the sampling methods at least at one station.

Biological assessment

Standardized qualitative sampling methods are often used in studies intended to determine if biological condition has been impaired by human activities (biomonitoring). Rapid bioassessment approaches, which usually employ standardized qualitative sampling, have been developed to assess biological condition without using more rigorous quantitative sampling methods that tend to be time consuming and expensive. However, standardized qualitative sampling methods, especially when subsampling is employed, must lead to an accurate assessment of biological condition in order to be valid. Biological condition is assessed by determining if there are major differences in the composition of benthic macroinvertebrate communities between sites suspected to be impaired and reference sites representing best attainable conditions. Benthic macroinvertebrate communities exhibit natural variability so minor differences in composition usually do not indicate serious impairment of biological condition. Reliable assessment of biological condition includes analysis of a number of metrics that represent different ecological factors. Given the complexity of ecological interactions and responses to pollution, rarely do all metrics show the expected change in value. Decisions on impairment of biological condition are usually based on a weight-of-evidence approach.

In the present study, we found a strong relationship between assessment agreement and the sensitivity of an individual metric. The better the separation between acceptable and unacceptable biological conditions indicated by a metric (e.g., EPT index, modified HBI, % intolerant, Index of Biotic Similarity and Bray-Curtis Coefficient) the better the agreement between the sampling methods. In fact, using a rigorously-tested multimetric index like the MAIS substantially increased the overall agreement of assessments made using the two sampling methods (from 73% to 89%).

In this study, we found that the two sampling methods gave slightly different estimates of community composition, but only because a few metrics are related to sample abundance. We

suspect that the PIBS may do a better job of estimating the true community because more intensive sampling techniques lead to higher abundances. However, we found no pattern in our results that would indicate one sampling method was more accurate than the other for making assessments of biological condition. Reliable assessments of biological condition can be made with either sampling method. Given the greater time and costs associated with quantitative sampling methods, such as the PIBS, we conclude that standardized qualitative methods, such as the KSS, are preferable for rapid bioassessment approaches to environmental assessment.

Acknowledgments

We are grateful to Stephen Hiner, William Van Wart, Joel Herbein, Larry Willis, Ed Cumbow, Brett Marshall, Carolyn Bussi, and Tim Mack for their assistance with this study.

Literature Cited

Barbour, M.T., M.L. Bowman, and J.S. White. 1994. Evaluation of the biological condition of streams in the Middle Rockies Central Ecoregion, Wyoming. Wyoming Department of Environmental Quality.

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. Environmental Toxicology and Chemistry 11:437-449.

Barbour, M.T., and J.B. Stribling. 1991. Use of habitat assessment in evaluating the biological integrity of stream communities. Pages 25-38 *in* Biological criteria: research and regulation. EPA-440/5-91-005. US Environmental Protection Agency, Office of Water, Washington, DC.

Barton, D.R., and J.L. Metcalfe-Smith. 1992. A comparison of sampling techniques and summary indices for assessment of water quality in the Yamaska River, Quebec, based on benthic macroinvertebrates. Environmental Monitoring and Assessment 21:225-244.

Bray, J.R. and J.T. Curtis. 1957. An ordination of the upland forest communities of southern Wisconsin. Ecological Monographs. 27: 325-349.

Caton, L.W. 1991. Improved subsampling methods for the EPA "Rapid Bioassessment Benthic Protocols." Bulletin North American Benthological Society 8: 317-319

Hannaford, M.J. and V.H. Resh. 1995. Variability in macroinvertebrate rapid-bioassessment surveys and habitat assessments in a northern California stream. Journal of the North American Benthological Society 14:430-439.

Hilsenhoff, W.L. 1977. Use of arthropods to evaluate water quality of streams. Wisconsin Department of Natural Resources Bulletin 100:1-15.

Hilsenhoff, W.L. 1982. Using a biotic index to evaluate quality in streams. Wisconsin Department of Natural Resources Bulletin132:1-22.

Hilsenhoff, W.L. 1987a. An improved biotic index of organic stream pollution. Great Lakes Entomologist 20:31-39.

Hutchinson, G.E. 1993. A treatise on limnology: volume 4. John Wiley & Sons, Inc. New York, New York.

Hynes, H.B.N. 1970. The ecology of running waters. University of Toronto Press, Toronto, Canada.

Hruby, T. 1985. Using similarity measures in benthic impact assessments. Environmental Monitoring and Assessment 8:163-180.

Krebs, C.J. 1987. Ecological Methodology. Harper and Row, New York, New York.

Lenat, D.R., and M.T. Barbour. 1994. Using benthic macroinvertebrate community structure for rapid, cost effective, water quality monitoring: rapid bioassessment. Pages 187-215 *in* S.L. Loeb and A. Spacie (editors). Biological monitoring of aquatic systems. Lewis Publishers, Ann Arbor, Michigan.

Ludwig, J.A., and J.F. Reynolds. 1988. Statistical ecology: a primer on methods and computing. Wiley Interscience Publication, John Wiley and Sons. New York, New York.

Merritt, R.W. and K.W. Cummins (editors). 1984. An introduction to the aquatic insects of North America, 2nd edition. Kendall/Hunt Publishing Company, Dubuque, Iowa.

Norris, R.H., E.P. McElravy, and V.H. Resh. 1993. The sampling problem. Pages 282-306 *in* P. Callow and G.E. Petts (editors). The Rivers Handbook, Volume 1. Blackwell, Oxford, England.

Omernik, J.M. 1987. Ecoregions of the conterminous United States. Annals of the Association of American Geographers 77:118-125.

Pennak, R.W. 1989. Freshwater Invertebrates of the United States: Protozoa to Mollusca. 3rd edition. Wiley Interscience, John Wiley and Sons, New York, New York.

Pinkham, C.F.A. and J.G. Pearson. 1974. A new measure of biotic similarity between samples and its applications with a cluster analysis program. Edgewood Arsenal Technical Report. EB-TR-74062. Department of the Army. Headquarters, Edgewood Arsenal, Aberdeen Proving Ground, Maryland.

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/444/4-89/001. US Environmental Protection Agency, Office of Water. Washington, DC.

Resh, V.H. 1994. Variability, accuracy, and taxonomic costs of rapid assessment approaches in benthic biomonitoring. Bollettino di Zoologia 61:375-383.

Resh, V.H. 1995. The use of benthic macroinvertebrates and rapid assessment procedures for water quality monitoring in developing and newly industrialized countries. Pages 167-177 *in* W.S. Davis and T.P. Simon (editors). Biological assessment and criteria: tools for water resource planning and decision making. Lewis Publishers, Boca Raton, Florida.

Resh, V.H., and J.K. Jackson. 1993. Rapid assessment approaches to biomonitoring using benthic macroinvertebrates. Pages 195-233 *in* D.M. Rosenburg and V.H. Resh (editors). Freshwater Biomonitoring and Benthic Macroinvertebrates. Chapman and Hall, New York, New York.

Resh, V.H., R.H. Norris, and M.T. Barbour. 1995. Design and implementation of rapid assessment approaches using benthic macroinvertebrates for water resource assessment. Australian Journal of Ecology 20:108-121.

Reynoldson T., Hampel L. & Martin J. 1986. Biomonitoring networks operated by school children. Environmental Pollution 41:363-380.

SAS Institute Inc. 1985. SAS User's Guide:Statistics. 5th edition. SAS Institute Inc., Cary, North Carolina.

Smith, E.P., and J.R. Voshell, Jr. 1997. Studies of benthic macroinvertebrates and fish in stream within EPA Region 3 for development of biological indicators of ecological condition. Report submitted to US Environmental Protection Agency.

Sokal, R.R. and F.J. Rohlf. 1995. Biometry: The principles and practice of statistics in biological research. 3rd ed. W.H. Freeman and Company. New York, New York.

Stewart, K.W., and B.P. Stark. 1988. Nymphs of North American stonefly genera (Plecoptera). Thomas Say Foundation, The Entomological Society of America.

Suter, II, G.W. 1993. A critique of ecosystem health concepts and indexes. Environmental Toxicology and Chemistry 12:1533-1539.

USEPA. 1973. Biological field and laboratory methods for measuring the quality of surface waters and effluents. EPA-670/4-73-001 National Environmental Research Center, Office of Research and Development. U.S. Environmental Protection Agency, Cincinnati, Ohio.

USEPA. 1983. Method of chemical analysis of water and wastes. EPA-600/4-79-020 Environmental Monitoring and Support Laboratory, Office of Research and Development. US Environmental Protection Agency, Cincinnati, Ohio.

Voshell, J.R., Jr., R.J. Layton, and S.W. Hiner. 1989. Field techniques for determining the effects of toxic substances on benthic macroinvertebrates in rocky-bottomed streams. Pages 134-155 *in* U.M. Cowgill and L.R. Williams (editors). Aquatic Toxicology and Hazard Assessment: 12th Volume, ASTM STP 1027. American Society for Testing and Materials, Philadelphia, Pennsylvania.

Wiggins, G.B. 1977. Larvae of the North American Caddisfly Genera (Trichoptera). University of Toronto Press, Toronto, Ontario Canada.

Willis, L.D. 1989a. Benthic survey of Peak Creek. Virginia State Water Control Board-West Central Regional Office. Technical Report QL-89-008.

Willis, L.D. 1989b. Peak Creek sediment metals. Virginia State Water Control Board-West Central Regional Office.

Wolda, H. 1981. Similarity indices, sample size and diversity. Oecologia 50:296-302.

CONCLUSIONS

Currently, many regulatory and natural resource agencies across the US are adopting ecoregional multimetric biological criteria into their assessment programs. This framework assesses biological condition by comparing the benthic macroinvertebrate community of a test stream with the benthic macroinvertebrate communities found at unimpaired reference sites within the same ecological region. These reference sites are assumed to represent the best attainable condition of the water resource within the region. In the US, ecoregions are the principal classification scheme for organizing streams into homogeneous groups. The underlying assumption of using an ecoregional classification scheme in bioassessment is that streams within the same ecoregion should have similar benthic macroinvertebrate communities.

One reason why many state and federal agencies are adopting ecoregional multimetric biocriteria is due to the lower-cost sampling approach used in the assessment. The standardized qualitative approach, commonly used in rapid bioassessment, makes estimates of benthic macroinvertebrate community composition from a subsample of a large area of stream bottom (2-m^2). The underlying assumption in using a standardized qualitative approach is that the estimate of community composition is accurate enough to make reliable assessments. Therefore, the purpose of this study was to test these two critical assumptions essential to the success of ecoregional multimetric biocriteria.

LANDSCAPE AND BIOTIC CLASSIFICATION FRAMEWORKS

I compared the benthic macroinvertebrate communities of 3 ecoregions; Blue Ridge Mountains (66), Central Appalachian Ridges and Valleys (67), Central Appalachians (69), and their corresponding subregions. There was no difference between the Blue Ridge Mountains ecoregion (66) and the Central Appalachian Ridges and Valleys ecoregion (67) in individual

taxa, metrics or community similarity analysis. However, in one community similarity analysis, there was a significant difference between the Central Appalachians ecoregion and the Blue Ridge Mountains ecoregion and Central Appalachian Ridges and Valleys ecoregion. I also found significant differences between the subregions. The Forested Hills and Mountains subregion (69a) had a substantially different community than any other subregion. The valley/plateau subregions (66c, 67a and 67b) had substantially different communities from the mountain subregions (66a, 66b, 67c and 67d). The valley/plateau subregions also had significantly greater abundances of sediment and pollution-tolerant taxa as well as greater abundances of taxa that generally occur in warmer lowland streams. There was little or no difference in individual taxa, metrics, or communities between the montane subregions of the Blue Ridge Mountains ecoregion and the montane subregions of the Central Appalachian Ridges and Valleys ecoregion. Therefore, I found little or no correspondence between benthic macroinvertebrate distribution and ecoregions or subregions in western Virginia.

An alternative landscape classification to ecoregions is a rearrangement of ecoregions or subregions into more homogeneous regions for a given community. I found that it was more accurate to rearrange the subregions into three new regions called bioregions: forested hills and mountains (subregion 69a), valleys and plateaus (subregions 66c, 67a and 67b), and mountains (subregions 66a, 66b, 67c and 67d). Multiple discriminant analysis confirmed that bioregions (73%) were a more accurate stream classification scheme than ecoregions (60%) or subregions (55%) for western Virginia.

Classification schemes like the ecoregion approach are generally built on landscape variables first and then stream communities are assessed to see if they match. The biotic classification scheme works by first classifying reference sites into groups by their biota and then identifying the predominant environmental variables responsible for segregating the biotic groups. Test sites are then compared to reference sites with matching environmental factors. A

hierarchical cluster analysis classified the sites into 7 homogeneous biotic groups. Multiple discriminant analysis using 14 environmental variables correctly classified 69.8% of the reference sites. Stepwise multiple discriminant analysis and graphical analysis showed that sampling date, slope, pH, habitat assessment score and distance from source correctly classified 52.8% of the reference sites. These classification rates compare favorably with rates considered acceptable by Wright et al. (1984) [76.1%], Moss et al. (1987) [65.7%] and Norris (1996) [66%].

In western Virginia, ecoregions and subregions may not be the most accurate classification frameworks for setting biocriteria using benthic macroinvertebrates. However, information from this study showed that alternative classification systems, the bioregion approach and the biotic approach, correspond better than ecoregions with benthic macroinvertebrate distribution in western Virginia streams. Considering the high sampling cost and rigorous statistical assumptions required by the biotic approach, bioregions may be preferable.

STANDARDIZED QUALITATIVE APPROACH

I compared the benthic macroinvertebrate composition estimated by a standardized qualitative approach with the benthic macroinvertebrate composition estimated by a quantitative approach. Pairwise comparisons of the sampling methods showed had similar relative abundances of the more dominant taxa. Likewise, the estimates from the quantitative method had significantly greater taxa richness, modified HBI and % collector-gatherers than the standardized qualitative method at the Peak Creek Reference Station. At the Wolf Creek Reference Station, the estimates from the quantitative method also had significantly more taxa and EPT taxa. The slight differences found in the metrics and similarity indices are most likely due to differences in abundance of organisms obtained by the two sampling methods.

I also compared the assessments of these two approaches along a stream with various levels of impairment. Based on a comparison of 12 metrics individually at all 18 stations, the two methods made the same assessment an average of 73% of the time. Furthermore, assessments using the Macroinvertebrate Aggregated Index for Streams (MAIS) showed that the sampling methods made the same assessment of biological condition 89% of the time. Assessment agreement between sampling methods was strongly affected by metric sensitivity. I found no pattern showing one method was more accurate in making assessments of biological condition than the other. Given the greater time and costs associated with quantitative sampling methods, such as the KSS, are preferable for rapid bioassessment approaches to environmental assessment.

SUMMARY

I tested the two assumptions behind many multimetric biocriteria in place today: the ecoregion framework and standardized qualitative sampling. I found that an alternative landscape classification framework (bioregions) is better than ecoregions for establishing multimetric biological criteria in western Virginia. Since the Blue Ridge Mountains ecoregion, the Central Appalachian Ridges and Valleys ecoregion, and the Central Appalachians ecoregion extend across the Mid-Atlantic Highlands Area, this finding may positively influence the development of more accurate multimetric biocriteria within this region. In addition, I found no pattern showing that a quantitative approach was more accurate in making assessments of biological condition than a standardized qualitative approach. Therefore, multimetric biocriteria built on a standardized qualitative approach generate reliable assessments.

References Cited

Moss, D., M.T. Furse, J.F. Wright, and P.D. Armitage. 1987. The prediction of the macroinvertebrate fauna of unpolluted running-water sites in Great Britain using environmental data. Freshwater Biology 17:42-52.

Norris, R.H. 1996. Predicting water quality using reference conditions and associated communities. Pages 32-52 *in* R.C. Bailey, R.H. Norris and T.B. Reynoldson (editors). Study design and data analysis in benthic macroinvertebrate assessments of freshwater ecosystems using a reference site approach. Technical Information Workshop, North American Benthological Society, 44th Annual Meeting, Kalispell, Montana.

Wright, J.F., D. Moss, P.D. Armitage, and M.T. Furse. 1984. A preliminary classification of running-water in Great Britain based on macroinvertebrate species and the prediction of community type using environmental data. Freshwater Biology 14:221-256.

VITA

Shane Kent Evans

I was born in Westminster, California in 1963, but spent the first 6 years of my life living in Utah. My family moved to California in 1969. In 1989, I graduated from Brigham Young University with a bachelors degree in Biology Composite Teaching, with an emphasis in Zoology. Two years later in 1991, I received a masters degree in Zoology, with an emphasis in Biological Science Education. For my research project, I developed a set of curriculum materials to help secondary school students think about societal values, analyze environmental dilemmas, and explore current environmental issues. In August 1991, I started my Ph.D. in Entomology at Virginia Polytechnic Institute and State University. My dissertation research examined the correspondence of ecoregions with benthic macroinvertebrate distribution in the streams of western Virginia. In addition, I compared the effectiveness of a standardized qualitative sampling approach with a traditional quantitative approach for detecting impairments in water quality and biological condition. In addition to my own research, I have assisted in developing a biological monitoring program for the George Washington and Jefferson National Forest. I also assisted the USEPA in developing biocriteria for the Mid-Atlantic Highlands region (Maryland, Pennsylvania, Virginia and West Virginia). As a graduate student, I have served as a teaching assistant for introductory biology, general entomology, nature study methods, entomology for elementary school teachers, hands-on activities for elementary school teachers, aquatic entomology for elementary school teachers, medical-veterinary entomology, and aquatic entomology. From 1995-1996, I served as President of the W.B. Alwood Entomological Society. I also assisted with the Virginia Governor's School Field Biology Program during the summers of 1994-1996. I have been a member of the Entomological Society of America since 1991 and the North American Benthological Society since 1995. In 1994, I received the James M. Grayson Scholarship in Entomology.