Preprint 11-063

SHOULD THERE BE AN AQUATIC LIFE WATER QUALITY CRITERION FOR CONDUCTIVITY?

R. Gensemer, GEI Consultants, Denver, CO
S. Canton, GEI Consultants, Denver, CO
G. DeJong, GEI Consultants, Denver, CO
C. Wolf, GEI Consultants, Denver, CO
C. Claytor, GEI Consultants, Denver, CO

ABSTRACT

A number of state and federal agencies are calling for development of aquatic life criteria for conductivity. These aquatic life criteria or "benchmarks" are based on observed correlations between conductivity and benthic macroinvertebrate community composition, generally measured by composite metrics or indices. However, development of an aquatic life criterion for a composite variable like conductivity is made difficult by a number of factors. Based on a database from West Virginia, observed patterns of invertebrate community composition versus conductivity may instead be related to a combination of abiotic (e.g., ionic composition, habitat) and biotic (e.g., life history, colonization potential) factors. Many benthic invertebrate taxa also do not respond to increasing conductivity in a consistent relationship. Many states have determined that a composite variable, like conductivity or TDS, is not appropriate for criteria development, as many studies have shown that toxicity varies as a function of ion composition and can be mitigated by elevated hardness. We conclude that the relationships between conductivity and changes in benthic macroinvertebrate community composition are neither strong nor reliable enough to warrant derivation of a criterion based solely on conductivity.

INTRODUCTION

It has been recently proposed that coal mining and valley fill (CM/VF) activities in West Virginia lead to increases in the conductivity of surface waters located immediately downstream of activities, and that these increases in conductivity are related to adverse changes in the structure of benthic macroinvertebrate communities (Pond et al. 2008). In particular, reduced abundances of mayflies (represented by the aquatic insect order Ephemeroptera) were considered to be most closely related to elevated water conductivity. The relationships identified in Pond et al. (2008) were based purely on statistical correlations between water quality characteristics and benthic macroinvertebrate community structure and do not represent a formal or mechanistic test of the hypothesis that conductivity (or the chemical parameters detected by the composite measure of conductivity) is the primary cause of changes in the macroinvertebrate communities downstream of CM/VF activities.

The U.S. Environmental Protection Agency (EPA) is now proposing that the correlation between conductivity and benthic macroinvertebrate community structure is strong enough that an aquatic life "benchmark" can be derived. In a detailed guidance memorandum dated April 1, 2010, EPA provides guidance on the interpretation of narrative Water Quality Criteria for conductivity stating that "predicted conductivity levels below 300 µS/cm generally will not cause a water quality standard violation and that in-stream conductivity levels above 500 µS/cm are likely to be associated with . . . exceedances of the narrative state water quality standards" (EPA 2010). The specific benchmarks referenced therein are based on a draft report currently under review and thus are subject to change.

Given these potential regulatory implications of the April EPA memorandum, the biological plausibility of using conductivity as the basis for deriving an aquatic life benchmark must be carefully reviewed to determine whether it represents a scientifically reliable means of ensuring aquatic life protection. Correspondingly, as part of a larger evaluation of this issue (GEI 2010; http://www.nma.org/pdf/legal/092110_gei.pdf), we conducted an independent statistical evaluation of the ecological factors most likely associated with observed variation in benthic macroinvertebrate community structure in West Virginia headwater streams, including streams associated with CM/VF activities.

METHODS

Pond et al. (2008) and EPA (2010) appear to presuppose that conductivity is the best predictor that is both functionally and causally related to the response of macroinvertebrate communities in Central Appalachian streams, while disregarding many other factors that may influence community composition. The West Virginia Department of Environmental Protection (WVDEP) Watershed Assessment Branch Database (WABbase - http://oaspub.epa.gov/eims/eimscomm.getfile?p_download_id=496202) provides an opportunity to examine other possible factors that may contribute to shaping macroinvertebrate community structure. This dataset includes results for 3,286 sampling events representing 3,121 unique Station ID codes and contains a variety of variables that present site-specific information regarding regional landscape, water quality, and aquatic habitat conditions as well as macroinvertebrate community composition. Using this dataset, we conducted an independent analysis that considers all of the available information and strives to objectively identify key water quality and physical parameters that are most strongly associated with biotic responses.

This analysis was based on an integrated approach to identify environmental factors that best describe the observed variability between stream sampling sites that strongly correlate with each other, rather than trying to establish causal relationships. In the absence of a rigorous study design conducted under controlled experimental conditions, it is more important to identify data relationships rather than attempt to establish cause-effect relationships.

Our approach involved a series of statistical analyses that reduced the total number of parameters to a more ecologically meaningful subset of variables with respect to the available data. The original dataset was initially subdivided into independent stressor and dependent response variables. Independent stressor variables in a stream ecosystem include chemical and physical habitat variables, such as metal and ion concentrations in the water column and substrate particle-size composition (Paulson et al. 2001). Dependent response variables were selected to represent the biological components of the stream, with a focus on macroinvertebrate density or taxa richness. The independent stressor variables generally represent a mix of both quantitative (e.g., major ion or metal concentrations) and qualitative (e.g., embeddedness) variables, as well as composite habitat and water quality variables (e.g., Rapid Bioassessment Protocol [RBP] score, conductivity). Thus, understanding the general categories of each variable also helped reduce the overall list of variables.

The integrated analysis followed a series of statistical procedures (Paulson et al. 2001), as presented below, to identify key variables that could be used to characterize relationships between water quality, aquatic habitat, and macroinvertebrate communities. As shown in the step-by-step sequential analysis outlined below, this series of analyses is a strongly iterative process. In other words, as results of various tests are evaluated, decisions with respect to which variables are most statistically significant and or ecologically important may be adjusted, requiring the process to be repeated. Thus, this analytical process attempts to be as objective as possible, but some professional judgment is applied to ensure a meaningful outcome.

1. Apply basic statistics
   a. Generate descriptive summary statistics and data plots
   b. Normalize data as needed to meet statistical assumptions
   c. Compile correlation matrices
2. Identify key stressor and response variables using the following methods (described in detail below):
   a. Principle Components Analysis (PCA)
   b. All Possible Regressions (APR)
   c. Chi-square Automatic Interaction Detection (CHAID)
3. Rank variables according to relative influence
   a. Develop matrix of key independent stressor variables and relationships found in Step 2
   b. Repeat Steps 2 and 3 until the two most influential independent stressor variables are identified for each dependent response variable
4. Fit equation to describe interactions between stressor and response variables
   a. Use three-dimensional modeling program to identify non-linear relationships, and the extent to which these relationships are statistically significant

Using the dataset, basic statistical procedures (e.g., Spearman rank correlation, scatter and box plots) were used to summarize each of the independent stressor and dependent response variables, as well as evaluating general relationships between the two variable types. All variables were evaluated for approximation of a normal distribution using Shapiro-Wilkes normality tests and Q-Q probability plots. When appropriate, variables were transformed and re-evaluated for fit with an expected normal distribution. A logarithm base10 transformation (log) was used for water quality variables and macroinvertebrate density, while the arcsine-square root transformation was used for variables reported as percentages (e.g., percent fines and percent Ephemeroptera). The water quality variables—temperature and pH, as well as the physical habitat and macroinvertebrate variables such as embeddedness and genera-based metrics—did not require transformation. Two macroinvertebrate metrics (number of Trichoptera taxa and percent Trichoptera) were not included in the database, so were calculated based on subtraction of reported Ephemeroptera and Plecoptera metrics from summary Ephemeroptera, Plecoptera, Tricoptera (EPT) results that were provided in the dataset.

Using the basic summary statistics (e.g., frequency distributions, correlation analyses), as well as professional judgment, the entire list of variables was initially reduced to a smaller subset of variables that we believed to be the most ecologically relevant when evaluating factors that explained the variability observed between sites, in terms of macroinvertebrate communities in Central Appalachian streams (Table 1).

It is important to note that composite type variables are often not very useful when evaluating biological responses to environmental stressors. For example, the total RBP score for aquatic habitat evaluation may appear to strongly correlate with select biotic responses, yet this index provides little insight into the specific environmental characteristics that may be influencing biotic communities because it is comprised of many metrics. To the extent possible, we have excluded such composite independent stressor variables in our data analyses, including conductivity and hardness, because they provide little information above and beyond their component variables when trying to isolate water quality factors that may be most strongly associated with a biotic response. This is particularly important for conductivity, because biological responses are well known to result from exposure to individual ions, rather than composite descriptors of ionic strength such as conductivity or total dissolved solids concentrations (Mount et al. 1997).

Table 1. Refined list of independent stressor and dependent response variables used in the integrated analysis.

Independent Stressor Variables		Dependent Response Variables
Water Quality	**Physical Habitat**	**Macroinvertebrate**
Temperature	Bank stabilization	Clinger taxa, genera
Dissolved oxygen	Bank vegetation	Ephemeroptera, genera
Alkalinity	Undisturbed vegetation	EPT, genera
pH	Channel alteration	HBI, genera
Chloride	Channel flow	Intolerant taxa, genera
Sulfate	Riffle sinuosity	Plecoptera taxa, genera
Total aluminum	Embeddedness	Trichoptera taxa, genera*
Total calcium	Sediment deposition	Total taxa, genera
Total iron	Epifaunal substrate	Density
Total magnesium	Velocity of pool	Percent Chironomidae
Total manganese	Percent fines	Percent Ephemeroptera
Total suspended solids	Percent sand	Percent Ephemeroptera minus Baetidae
Total phosphorus	Percent silt	Percent EPT
Nitrate – Nitrite nitrogen		Percent EPT minus Cheumatopsyche
Fecal coliforms		Percent EPT minus Cheumatopsyche and Baetidae
		Percent Hydropsyche
		Percent Orthocladiinae
		Percent Plecoptera
		Percent Trichoptera*
		Percent Simuliidae
		Percent dominant 5 taxa, genera

* Calculated metric.

Principal Component Analysis

Principle Component Analysis (PCA) is a variable reduction procedure that helps identify redundancy among numerous variables, and is used to identify whether groups of observed variables tend to "move together" (i.e., were positively or negatively correlated with one another) or not (Johnson and Wichern, 1992). PCA also helps identify variables that best explain the variability observed between sites and how those variables relate to one another, as well as whether one variable could be used as a surrogate for other variables within each grouping (i.e., water quality, physical habitat, macroinvertebrate). When such variables are replaced with a surrogate which explains the same amount of variation, the power of the statistic to identify relationships is maximized (Paulson et al. 2001).

An iterative process was used for the PCA analyses, such that all variables from each grouping were loaded into separate PCA models. This initially created three distinct groupings, two for stressor variable groupings (i.e., water quality and physical habitat) and one for the response variable group (i.e., macroinvertebrates). The PCA extraction method was based on a correlation matrix with a varimax rotated solution, pairwise deletion of missing values, and extracted eigenvalues[1] greater than 1.0. The rotated component matrix[2] for each variable grouping was examined, with variables exhibiting coefficients greater

[1] An eigenvalue is a measure of the strength of a principal component axis, the amount of variation along the axis, and, ideally, the importance of an ecological gradient.

[2] A rotated component matrix is one showing the results of varimax orthogonal rotation that minimizes the number of heavily weighted variables on each principal component.

than 0.6 considered a significant part of the component. If the component contained multiple significant variables, the Spearman rank correlation values for those variables were also evaluated. If variables were highly correlated (i.e., > 0.6 or < -0.6) with each other, the variable with the largest component coefficient (i.e., most heavily weighted) was selected. Up to five components were examined with the most heavily weighted or unique variables (either positive or negative) being selected for inclusion in a subsequent PCA model.

All Possible Regressions

All Possible Regressions (APR) is another iterative method that combines one dependent response variable with many independent stressor variables, using all possible combinations of the stressor variables to maximize the variance explained in the response variable. This data exploration approach identifies the best single variable or subset of variables that explains the most variation observed in the biological response variable. The goal of APR analysis is to identify the smallest subset of variables that explains most of the variation, rather than to provide an actual predictive equation for the subset of variables.

For this analysis, the total taxa and percent EPT variables were selected as the biological response variables, as identified in the PCA analysis (see below). All of the independent stressor variables identified in Table 1 were initially included in each of the water quality and physical habitat APR models. Similar to the PCA approach, the water quality variables and physical habitat variables were first analyzed independently then combined in an overall APR model for each biological response variable. The R-squared (R^2) and root mean square error (RMSE) for each APR model were reviewed to identify a model with the largest R^2 and smallest RMSE, while minimizing the variable count.

Chi-square Automatic Interaction Detection

Chi-square Automatic Interaction Detection (CHAID) is a nonparametric exploratory model used to evaluate contingent relationships between a dependent variable and a series of independent stressor variables, including non-linear relationships (Paulson et al. 2001). CHAID selects a subset of stressor variables that best predicts the dependent variable, and presents these variables in a decision tree. The decision tree starts with the dependent variable and progressively splits into smaller branches (nodes) based on groupings of the stressor variables that best predict responses by the dependent variable. CHAID is a sequential fitting algorithm similar to a forward stepwise model, although the decision to split or combine independent variables is dependent or contingent upon earlier effects, rather than simultaneously as in regression analysis. Both the dependent and independent variables were raw untransformed values treated as interval scale variables, rather than nominal or ordinal variables.

Similar to the PCA and APR analyses, an iterative process was used to evaluate both water quality and physical habitat variables independently, and then select a subset of variables from each analysis to be combined in a final decision tree for each dependent variable. Individual CHAID models were developed for total taxa and percent EPT, which included all of the water quality or physical habitat parameters listed in Table 1. Thus, four separate CHAID decision trees were created: two for total taxa (water quality and physical habitat tree) and two for percent EPT (water quality and physical habitat tree). Each decision tree was evaluated and the most important independent stressor variables were selected from each analysis. The independent stressor variables listed for each dependent variable were included in a combined CHAID model to evaluate the relationships between both types of stressor variables and the biological response variable.

RESULTS

Principal Component Analysis – Water Quality

The goal of the PCA analysis was to understand whether the water quality variables "moved together" (i.e., were positively or negatively correlated with one another) and to select variables that may be a surrogate for other variables. For the first iteration of our PCA model, the first component included the log transformed variables for total magnesium, sulfate, and total calcium as weighted the most heavily. This weighting and movement (all positive) of the variables along the first component was to be expected, based on the chemical relationship between all of these variables and their Spearman Rank correlation values. In the second component, the log transformed variables for total iron, total aluminum, and manganese were weighted the most heavily, with all variables showing positive movement with each other. In the third component, fecal coliforms, pH, and alkalinity revealed the strongest weighting coefficients. Temperature and dissolved oxygen were key variables in the fourth component, and moved in opposite directions as is to be expected, while the nutrients total phosphorus and nitrate-nitrite were the most heavily weighted variables in the fifth component.

The selected variables within the first five components accounted for a total of 72% of the variation observed among sample sites with respect to the water quality variables contained within the WVDEP dataset. In contrast, parameters such as calcium, sulfate, and magnesium, along with parameters that characterize overall ionic strength explained approximately 38% of the variation among sample sites with respect to water quality.

The following variables were selected to be surrogates for other less heavily weighted variables in each component and were included in subsequent PCA analyses:

1. total magnesium
2. total iron
3. pH
4. fecal coliforms
5. dissolved oxygen
6. total phosphorus
7. total suspended solids (TSS)

TSS was selected even though it did not initially meet our original selection criteria. Based on its relatively moderate weighting in two of the five components, as well as its relationship to geological and hydrological underpinnings within the watersheds, we believed this to be an important variable that may influence macroinvertebrate communities.

The seven selected water quality variables were subsequently loaded into a second PCA model, with the same evaluative process being performed on the rotated component matrix. The rotated component matrix converged in the first two components, with the first component comprised of the log transformed variables for total magnesium (0.800), pH (0.692), and fecal coliform (0.638). In the second component, the log transformed variables for total iron (0.698) and dissolved oxygen (0.661) weighted the most heavily, while total suspended solids (0.780) and total phosphorus (0.743) were considered part of the third component.

The final water quality variables that were selected to be included in the overall PCA model evaluating relationships between water quality, habitat, and macroinvertebrate variables were:

1. total magnesium—also surrogate for Ca, SO_4, pH
2. fecal coliforms
3. total iron—also surrogate for Al and Mn
4. dissolved oxygen—also surrogate for temperature, and
5. total suspended solids—also surrogate for total phosphorus

Principal Component Analysis – Physical Habitat

The iterative PCA process described above was also performed using the independent physical habitat stressor variables. The initial PCA model using physical habitat characteristics extracted four components, with the first component being comprised of sediment deposition (0.832), embeddedness (0.735), riffle sinuosity (0.675), and epifaunal substrate (0.643), all of which exemplify substrate quality in these watersheds. The second component included undisturbed vegetation (0.855), bank vegetation (0.833), and channel alteration (0.755), which are characteristic of riparian habitat. The third component included the arcsine-square root transformation for percent fines (0.950), percent sand (0.844), and percent silt (0.678), which again characterize substrate composition. The fourth component only included channel flow, which had a weighting coefficient of 0.810. These four components accounted for a total of 66% of the variation observed among sample sites with respect to physical habitat conditions. The first component accounted for approximately 20% of the variation in physical habitat observed among sample sites.

From our initial analysis, we selected the following physical habitat variables to be included in a subsequent PCA analysis:

1. sediment deposition
2. undisturbed vegetation
3. percent fines, and
4. channel flow

The second physical habitat PCA extracted two components with sediment deposition (0.795) and percent fines (-0.769) weighted heavily and in opposite directions in the first component, even though they are not strongly correlated (Spearman, -0.376). Channel flow (0.909) weighted heavily in the second component. All three variables were selected to be included in the overall PCA model evaluating relationships between water quality, habitat, and macroinvertebrate variables.

Principal Component Analysis – Macroinvertebrates

The initial macroinvertebrate PCA model resulted in four components being extracted, with the first component comprised of the arcsine-square root transformations for the percent EPT variable and its derivatives, along with percent Ephemeroptera and its derivatives, percent Chironomidae, and the genera-based Hilsenhoff Biotic Index (HBI). Even though the genera-based HBI is not very informative from the standpoint of identifying key macroinvertebrate response variables, it is informative from a general community health perspective. The second component weighted the genera-based metrics for total taxa, clinger taxa, EPT taxa and its derivatives Ephemeroptera and Trichoptera taxa, as well as intolerant taxa and arcsine-square root transformed percent dominant 5 taxa (negative weighting). The third component was comprised of the arcsine-square root transformed percent Trichoptera, percent Hydropsyche, and the genera-based Trichoptera taxa, all of which characterize the caddisfly assemblage. The fourth component only included the log transformed macroinvertebrate density variable. All four components explained a total of 76% of the variation observed in sample sites with respect to the macroinvertebrate metrics contained in the WVDEP dataset. The first component, which was mainly comprised of EPT metrics and a Chironomidae metric, only accounted for approximately 31% of the variation among sample sites with respect to macroinvertebrates.

The percent EPT was strongly and negatively correlated with percent Chironomidae (Spearman, -0.686) and the EPT derivatives; therefore, the percent EPT variable was selected from the first component. Similarly, the genera-based total taxa (total taxa) was strongly correlated with the percent dominant 5 taxa (Spearman, -0.789), clinger taxa (Spearman, 0.763), Ephemeroptera taxa (Spearman, 0.625), EPT taxa (Spearman, 0.724), and intolerant taxa (Spearman, 0.662); therefore, the total taxa metric was selected from the second component. The third component was comprised of caddisfly metrics; thus, the most heavily weighted variable of percent Trichoptera was selected. Thus, from our initial macroinvertebrate PCA, we selected the following variables to be included in a subsequent PCA analysis:

1. percent EPT
2. genera-based total taxa
3. percent Trichoptera, and
4. density

The second macroinvertebrate PCA extracted two components with the arcsine-square root transformed variables of percent Trichoptera (0.801) and percent EPT (0.785) weighting in the first component and the total taxa (0.940) being heavily weighted in the second component. These two components explained approximately 64% of the variation observed among sample sites with respect to macroinvertebrate metrics. The percent EPT variable was selected from the first component due to its inclusion of both mayflies and stoneflies, and total taxa was also selected for inclusion in the overall PCA model evaluating relationships between water quality, habitat, and macroinvertebrate variables.

Principal Component Analysis – Overall

As a result of the individual PCAs described above, a total of 10 variables were selected for inclusion in the overall PCA to evaluate the relative importance of key water quality (5), physical habitat (3), and macroinvertebrate (2) variables in characterizing sample sites with respect to the available data. The overall PCA extracted four components, with the first component weighting the log transformed total magnesium with total taxa, and the second component weighting sediment deposition and arcsine-square root transformed percent fines. The log transformed total suspended solids and total iron were strongly weighted in the third component. Channel flow and log transformed dissolved oxygen were weighted heavily in the fourth component. These four components explained approximately 55% of the variation observed among sampling sites.

The first component in the overall PCA indicates that total macroinvertebrate taxa is moving in the opposite direction of (i.e., is negatively correlated with) major ions such as magnesium, indicating a strong relationship between the response of the macroinvertebrate community and ionic chemistry. In the initial water quality PCA, total magnesium was selected as a surrogate for sulfate, calcium, and pH, which may also be important factors to consider regarding biological response. The second component indicates that substrate characteristics also are an important factor when trying to explain the variation observed among sample sites in Central Appalachian streams. Lastly, total suspended solids, total iron, channel flow, and dissolved oxygen also appear to be important factors to consider when evaluating these stream site conditions. Notably, the percent EPT metric did not weight heavily in any of the components, although its coefficients for both the first and second component indicate this metric may be weakly related to ionic chemistry and substrate conditions.

The key variables identified in the PCA analyses were retained and placed into a matrix for further evaluation after the results from the APR and CHAID. This matrix was used to refine the key variables for inclusion in a possible three dimensional model to evaluate the non-linear relationships between water quality, physical habitat, and macroinvertebrate metrics.

All Possible Regressions

When the total taxa metric was regressed with the water quality variables, the best fit APR model was based on three variables that included log transformed alkalinity, sulfate, and total aluminum. However, these three variables only explained approximately 17% of the total variation observed in total taxa. The best fit physical habitat-based total taxa APR weighted four variables: bank stabilization, undisturbed vegetation, channel alteration, and embeddedness, although the maximized R^2 was even lower at 9%.

The six variables identified as contributing to the best fit APR models for macroinvertebrate total taxa were combined for an overall APR analysis. The best fit model using both water quality and physical habitat variables weighted three variables: undisturbed vegetation, channel alteration, and log transformed sulfate, and accounted for approximately 21% of the variation observed in total taxa.

The APR analysis of the transformed percent EPT with water quality variables resulted in a best fit model containing five variables: fecal coliform, total aluminum, total calcium, chloride, and total manganese, and accounted for approximately 24% of the variation observed in the percent EPT. The physical habitat APR resulted in a best fit model that included undisturbed vegetation, embeddedness, epifaunal substrate, and percent fines, which explained 16% of the variation in the percent EPT metric. When these water quality and habitat variables were combined in an overall APR analysis, the best fit model included five variables: epifaunal substrate, log transformed fecal coliforms, total aluminum, chloride, and total manganese. This model accounted for 28% of the variation observed in the percent EPT variable.

Chi-square Automatic Interaction Detection

When evaluating a CHAID decision tree, the first variable after the dependent response variable is considered the most important stressor in the tree (Figure 1). The CHAID decision tree presented in Figure 1 is the combined water quality – physical habitat CHAID model for percent EPT.

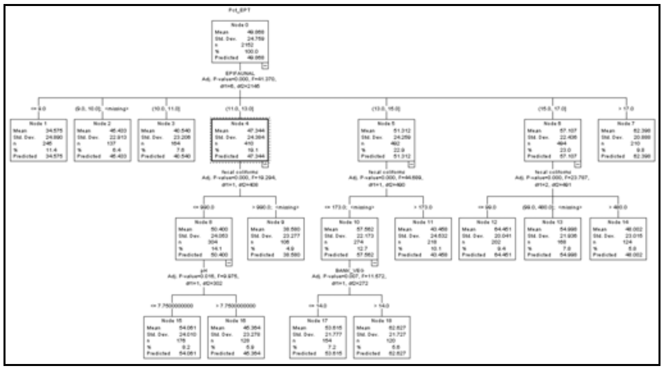

Figure 1. The combined water quality and physical habitat CHAID tree for percent EPT.

In this model, epifaunal substrate is the most important stressor variable (parent Node 0) for percent EPT. The box at each node shows the mean percent EPT value, the standard deviation for percent EPT, the number and percentage of sites with epifaunal substrate values in the listed range, and the predicted percent EPT at such sample sites. The nodes that branch from parent Node 0 (child nodes) list ranges of epifaunal substrate values (in brackets) such that Node 1 represents sample sites that scored less than or equal to 9.0 for epifaunal substrate. As epifaunal substrate scores increase (range from 0 to 20), the mean percent EPT value generally increases with each node. This response in percent EPT is to be expected, because as epifaunal substrate values increase, the quality of the habitat measure transitions from poor to optimal conditions. Sample sites that scored relatively high in this metric present a wide variety of natural structures in the stream, including fallen trees, large rocks, and cobble, all of which create a more complex habitat for aquatic life (Barbour et al. 1999).

Based on the information provided within Nodes 1-3, approximately 25% of the sample sites are categorized as having marginal to poor epifaunal substrate habitat (i.e., scored less than 11); thus, the habitat is less than desirable for benthic invertebrates, especially EPT taxa.

The second most important variable in the percent EPT CHAID analysis is fecal coliforms, which branch from three of the epifaunal substrate nodes. At sites that scored greater than 11 for epifaunal substrate (i.e., suboptimal to optimal), fecal coliform is an important secondary measure that influences the percent EPT metric. Sites that scored 11-13 for epifaunal substrate (Node 4) and exhibited fecal coliform levels less than or equal to 999 cfu/ml also exhibited a greater percent EPT value (50.4%) as compared to sites with fecal coliforms greater than 999 cfu/ml (38.6%). This relationship is consistent among all of the sample sites, such that greater levels of fecal coliforms result in a lower percent EPT value. This relationship suggests that other anthropogenic disturbances may be affecting the EPT taxa. Additional factors that influence percent EPT CHAID analysis were pH and bank vegetation, which branch out from two of the fecal coliform nodes. These factors appear to influence invertebrate communities in streams that scored 11-15 for epifaunal substrate (i.e., suboptimal range) and contained relatively low fecal coliform levels.

The combined water quality – physical habitat CHAID model for total taxa showed that sulfate concentration was the most important stressor variable (Figure 2). The model distinguished seven child nodes for sulfate concentrations, with the mean total taxa ranging from approximately 21 taxa for nodes that exhibited sulfate concentrations greater than 504 mg/L, to 31 taxa for nodes that exhibited concentrations less than 9.8 mg/L. However, these seven nodes essentially represent a breakpoint between sample sites that exhibit sulfate concentrations less than 61 mg/L or greater than 61 mg /L (i.e., between nodes 4 and 5).

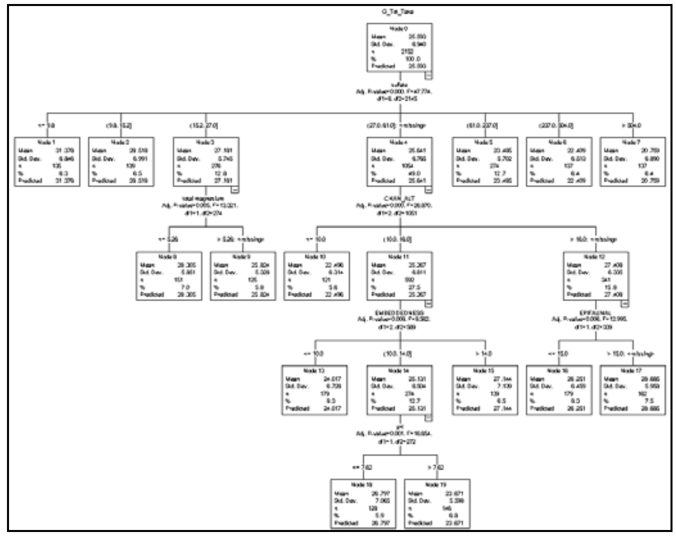

Figure 2. The combined water quality and physical habitat CHAID tree for total taxa.

In general, the mean total taxa ranged from 26 to 31 taxa for nodes that exhibited sulfate concentrations less than 61 mg /L. This range in sulfate concentrations is very similar to that observed for Level 1 Reference sites, which ranged from the detection limit to 65 mg/L. For the nodes representing sulfate concentrations greater than 61 mg/L, the mean total taxa ranged from 21 to 23 taxa. While mean total taxa varies by approximately 10 taxa across the full range of concentrations, the variability in mean total taxa for nodes representing concentrations greater than 61 mg/L is considerably less.

Secondary stressor variables for the combined total taxa model include total magnesium and channel alteration. These two variables are important variables to consider when sulfate concentrations are generally less than 61 mg/L. For sample sites characterized by Node 4, channel alteration is important to consider because this metric provides information regarding large-scale changes in the shape of the channel, such as channelization or bank stabilization using rip-rap (Barbour et al. 1999). Channel alteration values less than or equal to 10 (Figure 2, Node 10) represent poor to marginal conditions for this metric, whereas values greater than 16 represent optimal conditions for this metric. The total taxa metric responds predictably to channel alteration, such that poor to marginal conditions result in fewer total taxa when compared to optimal conditions. Other factors that influence total taxa are embeddedness and epifaunal substrate conditions. Both of these variables characterize the available substrate conditions, a critical consideration for benthic invertebrates.

Summary of PCA, APR, and CHAID Analyses

Our analyses indicate that a single composite parameter, like conductivity, cannot explain the variation observed among the Central Appalachian macroinvertebrate communities with respect to water quality and physical habitat. Rather, some combination of ionic composition, substrate, and channel features may be the most appropriate stressor variables to consider.

These analyses also indicate that total taxa and percent EPT abundance are the key response variables to consider when evaluating factors that shape the macroinvertebrate community, as opposed to a singular focus on Ephemeroptera. Furthermore, EPT abundance itself is a composite surrogate for taxa other than mayflies, and is a widely used indicator of impairment in benthic communities.

Additionally, total suspended solids, dissolved oxygen, and fecal coliforms appear to be key variables to consider when evaluating these stream sites, as they are strong indicators of other anthropogenic disturbances in the watersheds.

Despite the underlying assumption of Pond et al. (2008), and presumably EPA (2010), that conductivity is the primary driver in structuring macroinvertebrate community composition in the Central Appalachian streams, our analyses indicate that it is more appropriate to evaluate multiple possible stressors, including some of the specific ions that comprise the measure of specific conductance (Table 2). Furthermore, it is also important to consider substrate characteristics and habitat disturbance when evaluating macroinvertebrate responses.

This list of independent stressor variables represents the most important variables and their relative ranking of importance for each analysis. For example, the PCA model that considered percent EPT along with the key water quality and physical habitat variables revealed that percent fines and total magnesium weighted heavily in the first component and in the opposite direction of percent EPT. The total suspended solids weighted heavily in the second component. Thus, these three variables are considered important factors that influence the percent EPT metric. Similarly, for the total taxa APR model, three important factors, including undisturbed vegetation, channel alteration, and sulfate, were sequentially weighted into the APR model, indicating that of the three variables, undisturbed vegetation explained the most variation in the model.

Table 2. List of independent stressor variables considered important in the data reduction approach when evaluating stream sites and the two dependent response variables (genera-based total taxa and percent EPT).

Principal Component Analysis	All Possible Regressions	Chi-square Automatic Interaction Detection
Genera-based Total Taxa		
Total magnesium	Undisturbed vegetation	Sulfate
Percent fines	Channel alteration	Channel alteration
	Sulfate	Total magnesium
		Embeddedness
		Epifaunal substrate
Percent EPT		
Percent fines	Undisturbed vegetation	Epifaunal substrate
Total magnesium	Epifaunal substrate	Fecal coliforms
Total suspended solids	Fecal coliforms	Bank vegetation
	Chloride	pH
	Total manganese	

These independent stressor variables were further reviewed for their commonality among analyses, as well as their relative influence on each dependent response variable. The variables were then ranked to determine the most influential stressor variables for each biological response variable (Table 3). For example, based on our data reduction approach, channel alteration and sulfate concentration are the two most influential variables with respect to total taxa, while epifaunal substrate cover and fecal coliform concentrations are the two most influential variables with respect to percent EPT. The two primary stressor variables for each biological response variable are related to both physical habitat and water quality conditions, although ionic composition appears to be more influential on total taxa than percent EPT. The relative influence of fecal coliforms on percent EPT indicates that other anthropogenic disturbances are important factors to consider with respect to the benthic macroinvertebrate assemblages in West Virginia.

Based on the results of the PCA, APR, and CHAID analyses, the top two ranked stressor variables for each biological response variable were included in a 3-dimensional model (TableCurve 3D v4.0.01) to evaluate the non-linear relationships. Total taxa was modeled as a function of channel alteration and sulfate, while percent EPT was modeled as a function of epifaunal substrate cover and fecal coliforms. The best fit model for total taxa explained 21% of the variation observed in this metric, while the model for percent EPT explained only 14%. While the data reduction analyses provide insight into the key variables that influence total taxa and percent EPT, the outcome of the 3-dimensional modeling is not surprising. It is well known that multiple physicochemical and physical habitat characteristics elicit a variety of biological responses, thus a poorly fit model that explains little variation in a community composition metric is not unexpected.

Table 3. Matrix of sorted and ranked independent stressor variables for two dependent response variables (genera-based total taxa and percent EPT).

Principal Component Analysis	All Possible Regressions	Chi-square Automatic Interaction Detection
Genera-based Total Taxa		
1	Channel alteration	Channel alteration
2	Sulfate	Sulfate
3 Total magnesium		Total magnesium
4	Undisturbed vegetation	
5 Percent fines		
6		Embeddedness
7		Epifaunal substrate
Percent EPT		
1	Epifaunal substrate	Epifaunal substrate
2	Fecal coliforms	Fecal coliforms
3 Percent fines		
4	Undisturbed vegetation	
5 Total magnesium		
6		
7 Total suspended solids		
8		Bank vegetation
9	Chloride	
10	Total manganese	
11		pH

Despite the poor 3-dimensional modeling outcome, this data reduction approach indicates that physical habitat characteristics such as channel alteration, epifaunal substrate cover, and other sediment-based metrics are important factors to consider, in addition to ionic composition (e.g., sulfate and total magnesium), when evaluating macroinvertebrate responses. Additionally, the fecal coliforms variable indicates that other anthropogenic disturbances may play a key role in EPT composition of West Virginia streams.

DISCUSSION

Our analyses of the WABbase dataset indicate that conductivity alone cannot explain the variation observed among the Central Appalachian macroinvertebrate communities with respect to water quality and physical habitat. Rather, ionic composition, substrate composition, and channel features may be the most appropriate stressor variables to consider. Additionally, total suspended solids, dissolved oxygen, and fecal coliforms appear to be key variables to consider when evaluating these stream sites, as they are strong indicators of other anthropogenic disturbances in the watersheds. These analyses also indicate that total taxa and percent EPT abundance are the key response variables to consider when evaluating factors that shape the macroinvertebrate community, as opposed to a singular focus on Ephemeroptera. Therefore, we conclude that any regulatory benchmark based on conductivity and mayfly abundance is overly simplistic, and not likely to be an accurate measure of the biological condition or impairment of benthic macroinvertebrate communities.

It is noteworthy that three other states, Illinois, Indiana, and Iowa, have all rejected a composite variable like conductivity or TDS-based aquatic life standards in lieu of numeric standards for sulfate and chloride that also depend on water hardness. For Iowa, the current final rules (http://www.iowadnr.gov/water/standards/chloride.html) specifically state that the existing scientific data support the importance of individual ions over composite variables such as TDS because “chloride and sulfate are better indicators than integral parameters such as TDS, conductivity, and salinity for water quality protection” (IDNR 2009). Similarly, the Illinois EPA proposed a numeric sulfate standard, which was also ultimately approved by EPA, to replace TDS standards for the same technical reasons (Norwest Co. 2010). Indiana proposed essentially the same sulfate and chloride criteria equations, which were also approved by EPA because ". . . the TDS standard currently in place is inappropriate. By definition TDS is a measure of all dissolved solids, yet we know that the toxicity of TDS is exerted by its individual components" (EPA 2008). Therefore, the available scientific information does not support development of regulatory thresholds based on composite variables such as conductivity or TDS, but rather the development of individual numeric criteria for specific ions.

Illinois sulfate criteria

To illustrate the outcome of using the single-ion approach preferred by Illinois, the WABbase chemical data were used to derive aquatic life criteria for sulfate as modified by chloride and hardness. Using this example, the revised Illinois sulfate criteria are based on a range of total hardness and chloride concentrations (Table 4). Given site-specific conditions, sulfate criteria are either set at a constant 500 mg/L for samples exhibiting less than 100 mg/L total hardness, or a constant 2,000 mg/L for samples exhibiting total hardness greater than 500 mg/L and chloride concentrations greater than 5 mg/L. In addition, two equations are used to calculate site-specific sulfate criteria for samples exhibiting total hardness in the range of 100 to 500 mg/L and chloride in the range of 5 to 500 mg/L (Table 4).

The WABbase dataset contained 1,591 samples with paired hardness, chloride, and sulfate values, and represented a wide range of concentrations. Each sample was categorized based on total hardness and chloride concentrations and assigned a sulfate value based on the Illinois sulfate criteria rules (Table 4). The assigned sulfate value was then compared to the measured sulfate value to determine whether the sample achieved the Illinois sulfate criteria. Less than 1% (15 samples) of the WABbase samples exceeded the Illinois sulfate criteria, with the majority of exceedances occurring in the samples with hardness levels greater than 500 mg/L. There are a total of 54 samples exhibiting hardness values greater than 500 mg/L over a range of chloride concentrations, and 14 of these samples exceeded the sulfate criteria. In contrast, 26% of these WABbase samples exceed the proposed conductivity benchmark.

Table 4. Illinois sulfate criteria (mg/L, bold values) based on a range of hardness and chloride ion concentrations. The number of WABbase water samples within each range is identified by n.

Ion Ranges	Chloride <5 mg/L	Chloride 5 to <25 mg/L	Chloride 25 to <500 mg/L	Chloride ≥500 mg/L
Hardness <100 mg/L	**500** n = 696	**500** n = 350	**500** n = 23	**500** n = 0
Hardness 100 to <500 mg/L	**500** n = 113	**Eqn 1** n = 84 *1 of 84 exceeded criteria*	**Eqn 2** n = 270	**2,000** n = 1
Hardness ≥500 mg/L	**500** n = 10 *6 of 10 exceeded criteria*	**2,000** n = 26	**2,000** n = 15 *7 of 15 exceeded criteria*	**2,000** n = 3 *1 of 3 exceeded criteria*
Eqn 1: Sulfate = [-57.478 + 5.79(Hardness) + 54.163(Chloride)] x 0.65				
Eqn 2: Sulfate = [1,276.7 + 5.508(Hardness) – 1.457(Chloride)] x 0.65				

This analysis suggests that using a single ion criteria approach that incorporates the effects of hardness and chloride provides a significantly different indication of which and how many waters are likely to impair aquatic life. While elevated hardness and chloride concentrations are known to ameliorate sulfate toxicity (Soucek and Kennedy 2005, Soucek 2007), it is unknown whether the specific ionic composition of streams in West Virginia differs enough from Illinois streams in such a way that would make the single ion approach applicable. Notably, the State of Iowa is also considering adopting the same criteria that EPA and Illinois adopted in 2008. Given the empirical relationships between total hardness, chloride ions, and sulfate toxicity; the single ion approach warrants closer examination for use in Ecoregions 69 and 70 of West Virginia instead of a conductivity-based benchmark. However, even if this is an improvement over use of conductivity alone, it still does not incorporate any of the habitat or stream condition indicators that our statistical analysis identified as important predictors of macroinvertebrate community structure.

CONCLUSIONS

We conclude that the relationship between conductivity and changes in benthic macroinvertebrate community structure is neither strong nor reliable enough to warrant derivation of a regulatory benchmark at this time. For the most part, this is because Pond et al. (2008), and presumably EPA (2010) did not rigorously or independently test the primary hypothesis that elevated conductivity was the best predictor of changes in macroinvertebrate community structure in West Virginia streams associated with CM/VF activities. Rather, most of the analysis takes it as a given that conductivity is the best predictor. Furthermore, insufficient laboratory studies are available to confirm either the causal mechanisms or conductivity thresholds that would confirm the proposed benchmark under the specific ion composition of streams in this region. For similar reasons, Illinois, Indiana, and Iowa have rejected the use of TDS or conductivity-based criteria in lieu of criteria for individual ions such as sulfate or chloride.

Therefore, based on our statistical analysis, we conclude that it is inappropriate and inadvisable to adopt a conductivity benchmark until or unless such additional study is conducted. This is because many ecological factors other than water quality are strongly related to benthic macroinvertebrate community structure. To adopt this benchmark without the additional study runs a significant risk of expending significant financial resources to reduce conductivity from all possible ionic sources, with little confidence that this would provide any measureable environmental benefit.

REFERENCES

1. Pond G. J., M. E. Passmore, F. A. Borsuk, L. Reynolds, and C. J. Rose (2008), "Downstream effects of mountaintop coal mining: comparing biological conditions using family- and genus-level macroinvertebrate bioassessment tools", *Journal of the North American Benthological Society* 27:717-737.

2. U.S. Environmental Protection Agency (EPA) (2010), "Improving EPA Review of Appalachian Surface Coal Mining Operations Under the Clean Water Act, National Environmental Policy Act, and the Environmental Justice Executive Order (Detailed Guidance", *EPA-HQ-2010-0315*, U.S. Environmental Protection Agency, Office of Water, Washington D.C.

3. GEI Consultants, Inc. (GEI) (2010), "Final Technical Review: A Field-based Aquatic Life Benchmark for Conductivity", Available at http://www.nma.org/pdf/legal/092110_gei.pdf, GEI Consultants, Inc., Denver, CO.

4. Paulson, C. L., S. Reeves, S. P. Canton, and T. F. Moore (2001), "Distinguishing the relative influence of habitat and water quality on aquatic biota", *Water Environment Research Foundation,* 98-WSM-1.

5. Mount, D. R., D. D. Gulley, J. R. Hockett, T. D. Garrison, and J. M. Evans (1997), " Statistical models to predict toxicity of major ions to *C. dubia*, *D. magna* and *P. promelas*", *Environmental Toxicology and Chemistry* 10:2009-2019.

6. Johnson, R. A., and D. W. Wichern (1992), *Applied Multivariate Statistical Analysis, 3rd edition*, Prentice Hall, Englewood Cliffs, New Jersey.

7. Barbour, M. T., J. Gerritsen, B. D. Snyder, and J. B. Stribling (1999), "Rapid bioassessment protocols for use in streams and wadeable rivers: periphyton, benthic macroinvertebrates and fish, 2nd edition", *EPA 841-B-99-002*, U.S. Environmental Protection Agency, Office of Water, Washington, D.C.

8. Iowa Department of Natural Resources (IDNR) (2009), "Fact Sheet: Revising Criteria for Chloride, Sulfate, and Total Dissolved Solids" http://www.iowadnr.gov/water/standards/files/ws_fact.pdf, accessed July 7, 2010. Iowa Department of Natural Resources, Des Moines, IA.

9. Norwest Co. (2010), "Comments on A Field Based Aquatic Life Benchmark for Conductivity in Central Appalachian Streams *EPA/600/R-10/023A*", White paper prepared for National Mining Association, Washington D.C.

10. U.S. Environmental Protection Agency (EPA) (2008), "Supporting Rationale for Decision, Revisions to Indiana's Sulfate Water Quality Standards", *WQSTS IN2008-257*, U.S. Environmental Protection Agency, Region 5, Chicago, IL.

11. Soucek, D. J., and A. J. Kennedy (2005), "Effects of hardness, chloride, and acclimation on the acute toxicity of sulfate to freshwater invertebrates", *Environmental Toxicology and Chemistry* 24:1204-1210.

12. Soucek, D. J. (2007), "Comparison of hardness- and chloride-regulated acute effects of sodium sulfate on two freshwater crustaceans", *Environmental Toxicology and Chemistry* 26:773-779.