Comments on “A Field-Based Aquatic Life Benchmark for Conductivity in Central Appalachian Streams” EPA/600/R-10/023A

A WHITE PAPER

Submitted to:
National Mining Association
101 Constitution Ave. NM Suite 500 East
Washington, D.C. 20001

Date:
July 12, 2010

Norwest Corporation
950 S. Cherry Street, Suite 800
Denver, CO 80246
Tel: (303) 782-0164
Fax: (303) 782-2560

www.norwestcorp.com

Authors:
ART O’HAYRE AND FRAN AMENDOLA

TABLE OF CONTENTS

1 EXECUTIVE SUMMARY

In March, 2010, the Environmental Protection Agency (EPA) released a report that attempted to derive an aquatic life benchmark for conductivity that would apply to neutral to mildly alkaline waters in the Appalachian Region that are dominated by salts of sulfate and bicarbonate. The report, titled “Field-Based Aquatic Life Benchmark for Conductivity in Central Appalachian Streams” is being utilized by EPA as the basis for establishing a water quality benchmark for conductivity that would protect the aquatic life. The study and the calculated benchmark specifically targets coal mining activities in the region. At the request of the National Mining Association, Norwest has prepared this White Paper, which provides an analysis of the scientific basis for the statistical approach used by EPA to develop the proposed conductivity benchmark of approximately 300 µS/cm.

Conductivity is a measure of water’s capacity to conduct an electrical current and is correlated with the concentrations of the total dissolved ions or salts in solution (TDS). Appendix A of the Report provides background information and statistical methods used to address the causal assessment between conductivity and abundance of Ephemeroptera (mayfly) genera benthic macroinvertebrates. The causal assessment focused on mayflies because they are “among the most sensitive genera” according to EPA. While EPA asserts that there is a direct or causal relationship between conductivity and number of Ephemeroptera genera based upon field data, it is evident that the abundance of Ephemeroptera genera at sites included in the study is the result of a variety of physical and biological factors that are also weakly correlated with conductivity. Furthermore, laboratory studies do not support a causal relationship at the levels identified by the EPA benchmark. Several citations in the EPA Report identify much higher lowest observable effects concentrations (LOEC) for conductivity values (1,500 to 4,200 µS/cm) for mayflies and other aquatic genera.

An epidemiology approach is used in the causal assessment because the EPA believed that there was insufficient information available from controlled laboratory studies and they believed that they had sufficient field data. This mainly statistical approach is more common to studies of human populations, where laboratory experiments for causal relationships cannot be performed. It should be noted that a common issue in much of the epidemiological literature is that “correlation does not imply causation.” While the EPA study found a moderate correlation between conductivity and number of Ephemeroptera genera there is still a wide scatter in the relationship. This together with the relatively high conductivity LOEC values in laboratory studies casts doubt on conductivity as the dominant causal factor for the number of mayfly genera observed at sites included in the study.

In 2006 and 2009, both the Illinois EPA and the Iowa Department of Natural Resources have concluded that it is inappropriate to use TDS or conductivity as a water quality criteria to protect aquatic species because it is the concentration of individual ions rather than conductivity or TDS that is relevant to the toxicity to aquatic organisms. Consequently, the Illinois EPA proposed water quality criteria for specific ions such as sulfate and chloride based on laboratory toxicity studies on aquatic organisms. This approach was approved by USEPA.

In Appendix B of the EPA Conductivity Benchmark study, EPA attempts to support its conclusion that conductivity is the dominant causal factor by a cursory analysis of potential confounders given the *a priori* assumption that conductivity is a cause for the extirpation of Ephemeroptera species. The Causal Analysis/Diagnosis Decision Information System (CADDIS) is an online system developed by EPA to identify the causal factors that are contributing to biological impairment. While a few of these factors were considered by EPA in their statistical analysis of potential confounders, the approach was inadequate because many factors were ignored or were inadequately represented in the statistical approach used by EPA. For example, the effects of habitat on the number of Ephemeroptera genera was examined statistically using the Rapid Bioassessment Protocol (RBP) score derived by the West Virginia Department of Environmental Protection. The RBP is an index of habitat quality but it is not an adequate representation of all the habitat conditions that are favorable or unfavorable to various Ephemeroptera genera. Nevertheless, the EPA analysis still found that this index of habitat quality was correlated with conductivity and with the number of Ephemeropteran genera, but it was subsequently dismissed along with other potential confounders.

EPA’s analysis of confounding factors is insufficient and does not represent a rigorous analysis of the true factors affecting the number of Ephemeroptera genera at sites included in the study. Other conditions including channel alteration, the presence of upstream impoundments, dissolved oxygen, and temperature represent well-documented significant factors that need to be evaluated. The causal analysis of biological impairment needs to be performed at each site because multiple stressors can act together to cause impairment and these causal factors vary considerably among sites.

The conductivity benchmark is derived from a Species Sensitivity Distribution (SSD) that EPA generates from the field data based on the assumption that conductivity is the primary and only factor controlling the extirpation of aquatic genera from sampling locations in the study area. There is no scientific basis for this assumption. SSDs are supposed to be generated from **laboratory data** based on actual causal relationships and not from field biomonitoring studies based on moderate correlations.

The EPA attempt to develop a water quality criterion for conductivity from biomonitoring field data is not appropriate and should be replaced by the traditional methods for establishing water quality criterion. The traditional method for establish water quality criteria relies on information from controlled laboratory experiments to identify a direct relationship between organisms' response and exposure concentrations of a water quality constituent without the confounding influences. The biomonitoring approach is not suitable for establishing water quality criteria but is appropriate for detecting generalized and non-specific impairments to biological integrity. An analysis of the likely causes of biological impairment needs to be performed at each site following a structured approach such as that outlined in CADDIS.

The use of the conductivity benchmark as a water quality criterion would actually be detrimental because it would focus on the reduction of conductivity to protect benthic environments, when the cause of impairment at most of the sites is due to other factors, including a change in the flow regime, the presence of an impoundment, changes in the physical structure of the stream bed or changes in the quality and quantity of food resources preferred by various benthic insect genera. Thus, a mining company could invest in water treatment to meet conductivity limits for discharges and find that there is little or no change in the benthic insect genera downstream of the discharge following treatment. Finally, a water quality criterion for conductivity designed to protect biological conditions would result in a substantial increase in the number of streams classified as impaired even though the biomonitoring information does not indicate impairment. Examination of the West Virginia data base indicates that 990 of sites would be designated as impaired based on the EPA conductivity benchmark, while 409 of these sample sites would not be designated as impaired based on biomonitoring. Thus, the benchmark is found to have a 41% error rate as an additional criterion to protect water quality.

2 INTRODUCTION

Norwest Corporation (Norwest) is pleased to provide these technical comments in support of NMA’s comments on this precedent-setting approach to establishing water quality criteria proposed by EPA. The “Field-Based Aquatic Life Benchmark for Conductivity in Central Appalachian Streams” (herein referred to as the EPA Conductivity Benchmark study) is a significant departure from EPA’s regulatory approach for establishing water quality criteria. These comments address the methodologies employed in this document and in particular the causal assessment and related statistical methods and assumptions.

The EPA Conductivity Benchmark study develops a conductivity standard based on a species sensitivity distribution (SSD) that is determined empirically from a manipulated set of field data and not from controlled laboratory experiments. The EPA has calculated the species sensitivity distribution (SSD) based on the erroneous assumption that conductivity is the direct cause for the extirpation of genera from the Ephemeroptera order at levels far below the lowest observed effects concentrations results from controlled laboratory experiments. Upon review of the field data and the literature it is quite clear that the correlation between conductivity and number of Ephemeroptera genera were likely the result of a variety of physical and biological factors that are also weakly correlated with conductivity. The rationale behind using information from controlled laboratory experiments when developing water quality criteria is to identify a direct relationship between organisms’ response and exposure concentrations of a water quality constituent without the confounding influences of predation/competition from multiple species along with the influences of habitat, temperature, geology, flow and water chemistry. Furthermore, there is no scientific evidence for conductivity as a toxicity factor to benthic organisms at the low levels of conductivity proposed as the benchmark concentration of 300 µS/cm. Toxicity to aquatic organisms can occur at very high conductivity levels and varies depending on the specific aquatic organism and relative mix of ions in the water.

3 CAUSAL ASSESSMENT

Appendix A of the EPA Conductivity Benchmark study states that epidemiological arguments are applied in an attempt to show that mixtures of salts that elevate conductivity in streams in the Mountain and Plateau Regions of Central Appalachia are causing local extirpation of aquatic species. Epidemiology is the study of factors affecting the health and illness of populations. It generally applies to human populations, where laboratory experiments for causal relationships cannot be performed. This epidemiological approach used by EPA relies on statistical tools to establish an association between conductivity levels in a water body and different characteristics of the benthic organisms in the water body, primarily the presence of Ephemeroptera (mayfly) species. It should be noted that a common issue in much of the epidemiological literature is that "correlation does not imply causation."

Data from the West Virginia Department of Environmental Protection in-house Watershed Assessment Branch Data Base (WABbase) was used for both the derivation of the conductivity benchmark and for the causal assessment provided in Appendix A of the EPA Conductivity Benchmark study. The data set from the Kentucky Division of Water (KDOW) was not used in the causal assessment but was used for comparing the results of applying the same methodology to the Kentucky data. A copy of the West Virginia data set was obtained and evaluated by Norwest in preparation of these comments. A copy of the Kentucky data set was not provided.

The causal assessment references toxicity information from the literature involving laboratory studies to support the conductivity benchmark. The cited laboratory studies by Kennedy et al. (2003, 2004, 2005) and by Mount et al (1997) indicate that concentrations of major ions can adversely impact fish and benthic macroinvertebrates but at concentrations substantially higher than the conductivity benchmark proposed by EPA. As discussed in the EPA Conductivity Benchmark study report, the causal assessment for the conductivity benchmark is developed largely on the occurrence of Ephemeroptera (mayflies) genera "because they are among the most sensitive genera". The EPA Conductivity Benchmark study cites the study by Echols et al (2009), which reported a conductivity range from 1,508 to 4,101 µS/cm for the lowest observable effects concentration (LOEC) in laboratory tests of the mayfly *Isonychia bicolor* exposed to water samples that ranged from 13% to 39% coal-mine-processed effluent. This study also reported LOEC values for *Ceriodaphnia dubia* that ranged from 2,132 to 4,240 µS/cm.

Conductivity is a measure of the capacity of water can conduct an electrical current and is correlated with the concentrations of the total dissolved ions or salts in solution (TDS). However, it is the concentration of individual ions rather than conductivity or TDS that is relevant to the toxicity to aquatic organisms. The study by Mount et al. (1997) cited in the EPA Conductivity Benchmark study states that the toxicity of fresh waters with high TDS is dependent on the species ionic composition of the water. Mount et al. (1997) found that the relative toxicity of major ions in high TDS effluents was: K > HCO3= Mg > Cl > SO4. They also found that the presence of multiple cations reduces the toxicity of Cl, SO4 and K. Also, Chapman et al. (2000) studied toxicity of two mine effluents to early life stages of rainbow trout and Chironomid larvae and concluded that the toxicity is due to the specific combination and concentration of ions and is not predictable from TDS concentrations.

Furthermore, both the Illinois EPA[1] and the Iowa Department of Natural Resources[2] have concluded that it is inappropriate to use TDS or conductivity as water quality criteria based on toxicity to aquatic organisms. The Illinois EPA adopted water quality rules that replaced the TDS water quality standard with a numerical sulfate standard. This revision to the Illinois Water Quality Standards rules regarding sulfate and total dissolved solids were approved by the USEPA on March 19, 2009. In its technical justification, the Illinois EPA states that:

> *"While sulfate was being evaluated, it became increasingly obvious that TDS is a very inappropriate parameter for use in water quality standards. TDS is the sum of all dissolved substances in water and is dominated by the common ions of sulfate, chloride, sodium, calcium, carbonate and magnesium in various proportions. Our investigations into sulfate toxicity reinforced the notion that it makes little sense to have a standard that covers all these substances together when the toxicity of each constituent is really what is important. For example, a TDS concentration of 2,000 mg/L with chloride as the primary anion constituent is acutely toxic to aquatic life, but the same TDS concentration composed primarily of sulfate is nontoxic. With toxicity-based sulfate and chloride standards in force, there should be no need of a TDS standard that is incapable of predicting the threshold of adverse effects to aquatic life."*

On September 17, 2009 the Iowa Environmental Protection Commission adopted and filed amendments to the Commission's Water Quality Standards that removed the TDS site-specific approach and established numerical water quality criteria for chloride and sulfate. In its Responsiveness Summary to the Chloride, Sulfate and TDS Revisions, the Iowa Department of Natural Resources (August 24, 2009) cited "lack of scientific support" as the basis for eliminating

---

[1] Illinois Environmental Protection Agency. April 2006. "Preliminary Technical Justification for Changing Water Quality Standards for Sulfates, Total Dissolved Solids and Mixing Zones."

[2] Iowa Department of Natural Resources, February 9, 2009 "Water Quality Standards Review: Chloride, Sulfate and Total Dissolved Solids."

the site-specific TDS approach. Further discussion of the lack of scientific support for a TDS water quality standard for protection of aquatic life can be examined in the Iowa Department of Natural Resources, February 9, 2009 “Water Quality Standards Review: Chloride, Sulfate and Total Dissolved Solids.”

3.1 EXAMINATION OF THE EPA DATA SET

The data set used in the causal assessment and in developing the conductivity benchmark was manipulated by EPA to exclude samples from a number of locations. The following data exclusions occurred:

- Data outside Ecoregion 69 and 70,
- Samples where the aquatic organisms were not identified to the genus level,
- Samples from drainage areas greater than 155 km2,
- Conductivity was greater than 1,000 µS/cm, the SO4 less than 125 mg/L, and the Cl greater than 250 mg/L, and
- Sites with a pH of less than 6.

First, all samples that were not identified as occurring in either Ecoregion 69 or 70 were excluded. However, if conductivity as a measure of the concentration of dissolved salts in water is the primary factor causing the extirpation of aquatic species the effects should occur elsewhere and the geographic exclusion is not appropriate. This EPA Conductivity Benchmark report states that it is applicable to parts of Kentucky and West Virginia and is expected to be applicable to the same regions in Ohio, Pennsylvania, Tennessee, and Maryland. Next, all samples that did not have measurements of conductivity and aquatic organisms identified to the genus level were excluded as this information was needed to apply the methodology. All samples that were identified as being collected from locations with drainage areas greater than 155 km2 were excluded. The rationale for this exclusion was to eliminate samples from large rivers. Yet there was no explanation of the basis for the 155 km^2 threshold. Furthermore, if conductivity is a primary factor causing the extirpation of aquatic species in the Mountain and Plateau Regions of Central Appalachia, then larger streams should not be excluded from the study. Sample locations were also excluded if the conductivity was greater than 1,000 µS/cm, the SO4 less than 125 mg/L, and the Cl greater than 250 mg/L. This exclusion was applied to eliminate waters dominated by Cl rather than SO4 but there was no explanation for the selection of the exclusion thresholds. Furthermore, there were 149 samples with conductivity levels greater than 1,000 µS/cm that remained in the data set, where the exclusion criteria could not be applied because of lack of information on either SO4 or Cl.

An examination of the complete data set and the correlation with sample exclusions as provided in Figure A-1 of Appendix A of the EPA Conductivity Benchmark study indicates that the data set used in these correlations also excluded sites with a pH of less than 6. These sample exclusions resulted in the highest Spearman correlation between conductivity and the number of Ephemeroptera genera while other exclusions applied in the confounding analysis decreased the Spearman correlation between conductivity and the number of Ephemeroptera genera as shown in Figure A-2 of Appendix A of the EPA Conductivity Benchmark study. Figure A-2 shows that conductivity is a poor predictor for the number of Ephemeroptera genera present at sample sites. If conductivity were a causal factor for number of Ephemeroptera genera one would expect a much stronger relationship after confounding factors have been removed.

The correlations provided in both Figures A-1 and A-2 of Appendix A of the EPA Conductivity Benchmark study show substantial scatter in the relationship between the number of Ephemeropteran genera and conductivity, which casts doubt on the validity of the causal relationship.

There are a number of factors other than conductivity that may affect the number of Ephemeroptera genera present at a sample site. Examination of the data that were used to generate Figure A-1, found that 290 of the sample sites with conductivity levels below the EPA benchmark of 300 µS/cm were designated as biologically impaired based on the West Virginia Stream Condition Index (WVSCI) that is used to assess stream impairment. If conductivity were the primary cause of biological impairment, there would not be this large number (290) of biologically impaired sites with conductivity levels below the benchmark. Also, among the sample sites with conductivity levels below the EPA benchmark, the number of Ephemeroptera genera was 3 or less at 240 of the sample sites. These are sample locations where biological impairment occurs due to environmental conditions or stressors other than conductivity. The specific conditions that are causing biological impairment and limiting the number Ephemeroptera genera at low conductivity have not been assessed in the EPA study. This is due to an approach that focused on the correlation between the number of Ephemeroptera genera and conductivity, and not on identifying the causal factors influencing biological conditions and stream impairment at the specific locations included in the study.

It is logical to conclude that the range of factors that are the cause for biological impairment and limiting the number of Ephemeroptera genera at conductivity levels below the EPA benchmark are also relevant causal factors for impairment at sites with conductivity levels above the EPA benchmark. Furthermore, if the EPA conductivity benchmark is valid, then one would not expect multiple Ephemeroptera genera to occur at sites with conductivity levels higher than the EPA benchmark of 300 µS/cm. Yet it is apparent from an examination of Figure A-1 that multiple Ephemeroptera genera do occur at many sites with conductivity levels well above the EPA benchmark. As demonstrated in further analysis of the data, one reason that there is some correlation between the number of Ephemeroptera genera and conductivity is that there is some correlation between some of the actual causal factors and conductivity.

The use of the EPA conductivity benchmark as a water quality criterion would result in a substantial expansion of the number of streams designated as impaired. The WVSCI is used to identify sites that are impaired with respect to listing under Section 303(D) in the state of West Virginia. If the WVSCI is less than 60.6, the stream at the sample location is designated as biologically impaired. Examination of the EPA data base shows that, at sample locations with conductivity levels above the EPA benchmark, 581 of the sample sites would be designated as biologically impaired while 409 of the sample sites would not be designated as impaired using the WVSCI.

Table 1. Summary of Impairment Classifications that exceed the EPA conductivity Benchmark

Impairment Classification	**Number of samples exceeding EPA benchmark**	**Conductivity (µS/cm)**	
		Mean	**Maximum**
WVSCI Not Impaired	409	611	2,768
WVSCI Impaired	581	1,072	11,646

As shown in Table 1, the mean conductivity for the sites that were not designated as impaired index was 611 µS/cm while the maximum conductivity was 2,768 µS/cm. Thus use of the EPA conductivity benchmark could result in a 41% false designation of impairment with respect to the current designation of biological impairment using the WVSCI. Clearly, if conductivity levels above the benchmark were the cause of biological impairment, there would not be this large number (409) of unimpaired sites with conductivity levels above the benchmark.

The contingency table (Table A-4) that was supplied to support the causal assessment examined the ratio of presence or absence of mayflies at sample sites where conductivities were greater than 1,500 µS/cm. This is far above the EPA proposed benchmark of 300 µS/cm and is expected to include sites where particular ions such as bicarbonate, chloride or sulfate may be at concentrations that some Ephemeroptera genera may be affected. Still mayfly species are present at 21.7 % of the sites where conductivity is greater than 1,500 µS/cm. The presence of some Ephemeroptera species at these sites demonstrates that the mix of ions resulting in the high conductivity is not causing the extirpation of these Ephemeroptera species. Furthermore, it is not possible to conclude that at the 78.3 % of the sites where mayfly species were not observed when conductivity was greater than 1,500 µS/cm, that the mix of ions resulting in the high conductivity is the cause of the extirpation of the Ephemeroptera species at all of these sites.

3.2 POTENTIAL CAUSES OF BIOLOGICAL IMPAIRMENT

The Causal Analysis/Diagnosis Decision Information System (CADDIS) is an online system developed by EPA to help conduct causal evaluations for aquatic systems that are designated impaired (Ziegler, Suter, and Kefford, 2007). CADDIS lists the following conditions that could be a potential cause for biological impairment:

- Dissolved oxygen (DO) regime alteration,
- Hydrologic regime alteration (includes flow or depth conditions; timing, duration, frequency etc.),
- Nutrient regime alteration,
- Organic-matter regime alteration,
- pH regime alteration,
- Salinity regime alteration,
- Bed sediment load changes, including siltation,
- Suspended solids and/or turbidity alteration,
- Water temperature regime alteration,
- Habitat destruction,
- Habitat fragmentation (e.g., barriers to movement, exclusion from habitat),
- Physical crushing and trampling,
- Toxic substances,
    - Herbicides and fungicides,
    - Halogens and halides (e.g., chloride, trihalomethanes),
    - Fish-killing agents (e.g., rotenone),
    - Insecticides,
    - Lampricides,
    - Metals,
    - Molluscicides,
    - Organic solvents (e.g., benzene, phenol),
    - Other hydrocarbons (e.g., dioxins, PCBs),
    - Endocrine disrupting chemicals, and
    - Mixed, cumulative effect,
- Interspecies competition,
- Complications due to small populations (e.g., inbreeding, stochastic fluctuation, etc.),
- Genetic alteration (e.g., hybridization),
- Overharvesting or legal, intentional collecting or killing,
- Parasitism,
- Predation,
- Poaching, vandalism, harassment, or indiscriminate killing,

- Unintentional capture or killing (e.g., artillery explosions, roadway casualties),
- Vertebrate animal damage control (includes trapping, shooting, poisoning), and
- Radiation exposure increase (e.g., increased UV radiation).

A few of these factors were considered to some extent by EPA in their statistical analysis of potential confounders but were summarily eliminated from consideration because “effects appear to be minimal given the inevitable variability in sites to which the benchmark would be applied”. There are a number of flaws in the causal assessment performed in the EPA Conductivity Benchmark study but a primary flaw is that the analysis of potential causes of biological impairment needs to be performed at each site because of the variability among sites.

Appendix B of the of the EPA Conductivity Benchmark study attempts to support its conclusion that an increase in conductivity is the dominant causal factor in the local extirpation of aquatic genera through an analysis of potential confounders based on the *a priori* assumption that conductivity is a cause for the extirpation of Ephemeroptera species at sites included in the study. The assessment starts with the correlation between conductivity and the number of Ephemeroptera genera and then tries to determine whether any of the measured potential confounders interfere with this correlation.

For example, to assess whether stream habitat may be a potential confounder, they examine a qualitative index of habitat quality, the Rapid Bioassessment Protocol (RBP) score derived by the West Virginia Department of Environmental Protection. As shown in Table B-7 of the EPA Conductivity Benchmark study, the RBP score was found to be moderately correlated (based on EPA assessment criteria for correlation in the Table B-2) with conductivity (r = 0.29) and moderately correlated with the number of Ephemeropteran genera (r = -0.26). Also, as indicated in Table B-7 of the EPA Conductivity Benchmark study, the correlation between conductivity and the number of Ephemeropteran genera declined from an r = -0.63 to r = -0.50 when sites with an RBP score of less than 140 were removed from the analysis in an attempt to reduce the influence of poor habitat. Clearly this index of habitat quality is a confounder but it was dismissed in the development of the EPA benchmark along with other potential confounders, except for low pH. Furthermore, the rationale for selection RBP score threshold of less than 140 was not provided.

There is also the issue of whether the RBP score is the appropriate measure of habitat quality with respect to mayflies. As noted by GEI (2010), the nearly identical ratios of mayfly presence between poor quality and high quality habitat at low conductivity levels in the contingency table (Table B-8) suggests that that RBP score is not a good indicator of habitat quality for mayflies. The correlation between RBP and conductivity, and the correlation between RBP and Ephemeropteran genera, combined with the presence of mayfly in both poor and high quality habitat confirms that EPA's analysis of confounding factors is insufficient and does not represent a rigorous analysis of the true cause of Ephemeropteran or mayfly abundance.

In our analysis of the EPA data, the Channel Alteration score alone was found to have a better correlation with conductivity (r=-.31) and with number of Ephemeropteran genera (r = 0.27) than the RPB score. Furthermore, some of the habitat factors that have been found to be directly related to the presence of mayflies, such as the presence of an upstream impoundment, were not considered in the assessment of potential confounders by EPA. Bauernfeind and Moog (2000) indicate that channel alteration, irrigation, and impoundments generally lead to deficiencies in the mesohabitat structures which significantly affect mayfly diversity and abundance. Arnwine, Sparks and James (2006) found that impoundments on small first order to third order streams in Tennessee have adverse affects on the macroinvertebrate community downstream of these impoundments. The most frequent change in the benthic community structure was a loss of the orders Ephemeroptera, Plecoptera, and Trichoptera (EPT). Finally, the Technical Memorandum by GEI (2009) cites literature sources that assessed changes in trophic groups occurring downstream of lakes and impoundments. Filter-feeding insects such as hydropsychid tend to be found at higher abundance immediately below impoundments while mayflies typically decrease. These shifts apparently occur as a result of changes in trophic condition or food availability downstream of impoundments with the releases being rich in components fed on by hydropsychid and other filter-feeding organisms.

Furthermore, the EPA data base is not adequate to address the issue of confounding because it does not include all of the potential confounding parameters and it is missing data for some of the confounding parameters. For example for the final data set used by EPA in the conductivity benchmark, fecal coliform counts were missing in 149 of the samples, watershed area and land use designations were missing in 1,441 of the samples, and nutrient information from 1,032 of the samples. Temperature and dissolved oxygen are some of the most important factors influencing the occurrence and distribution of Ephemeroptera larvae (Bauernfeind and Moog, 2000). The EPA report completely failed to address dissolved oxygen as a potential confounding parameter. Perhaps the failure to address influences of dissolved oxygen are due to data limitations that report dissolved oxygen (and temperature) only at the point in time that the benthic sampling is performed. Examination of the EPA data indicates that only four samples had dissolved oxygen values less than 3.5 mg/L. All four samples were collected in August and the number of Ephemeropteran genera was found to be zero or one in these four samples. Dissolved oxygen could also be a factor limiting the number of Ephemeropteran genera but it is not possible to draw further conclusions based on the one-time samples. Likewise, the one-time measurement of temperature precludes the proper assessment of temperature as a confounding parameter. Nevertheless, the EPA report concluded that "elevated temperature does not appear to be associated with the loss of Ephemeroptera". In supporting this statement, the report includes the statement that “temperatures rarely exceed 20° C and are, therefore not likely to cause extirpation of genera.” Yet, only 780 of the 2,155 temperature samples were taken from the high temperature months of July and August. It should be noted that the median temperatures for July and August were 21.18 and 21.03° C, respectively.

Our analysis of the EPA data found that land use also had a statistically significant correlation with the number of Ephemeropteran genera observed. For example, percent woodland was moderately correlated (using the EPA criterion) with the number of Ephemeropteran genera ($r = 0.34$). We do not believe that there is a direct causal relationship between the percent woodland in a drainage basin and the number of Ephemeropteran genera observed. Rather, the correlation is spurious and is the result of a number of causal factors that influence the occurrence of Ephemeropteran genera, such as habitat and temperature, which are correlated with percent woodland. Conductivity, like percent woodland, appears to be correlated with several causal factors that may influence the number of Ephemeropteran genera present at a sample site. But it is not valid to calculate a species sensitivity distribution (SSD) based on conductivity, percent forest cover or any other correlated parameter obtained from uncontrolled field monitoring data.

The causal analysis of biological impairment needs to be performed at each site because multiple stressors can act together to cause impairment and these causal factors vary considerably among sites. There may also be an individual factor that is the primary cause of impairment at a limited number of locations that is not significant in a statistical analysis. Examples of this are low dissolved oxygen and toxicity to a specific water quality constituent (selenium, ammonia, etc.) In the analysis of potential confounders, the EPA Conductivity Benchmark study used fecal coliform counts as an indicator of organic enrichment and the presence of other toxicants. Of course a loss of information and greater uncertainty in the assessment of confounding is introduced when an indicator is used rather than direct information on the toxicants and organic enrichment. There are also data and conceptual problems in relying solely on the number of Ephemeropteran genera as the basis concluding that increasing conductivity levels is the direct cause for reductions in the number of Ephemeropteran genera. Bauernfeind and Moog (2000) indicate that the abundance of mayflies varies seasonally and that seasonal differences may also influence the identification of species that may be present at a site. They note that mayfly diversity at a given sampling station is “easily underestimated if certain (meso) habitats are unintentionally neglected” and that seasonal changes in habitat preference further complicate identification. Maggard (2009) work affirms this point as the number of Ephemeropteran genera observed in this study varied seasonally within the four steams included in his study of stream biota and water chemistry in the coal belt region of West Virginia.

Although the EPA Report mentions the on-line CADDIS system, they did not use the tools and step-by-step procedures to assess the conditions causing biological impairment at the specific impaired sites that were included in the data base. Earlier in these comments we summarized many of the potential causes for biological impairment that can act together to cause impairment. CADDIS also provides recommendations and cautions in the application of statistical tools. It states that statistical significance does not equate to biological significance. Statistical analysis does not tell us whether variability in observations is caused by the stressor being analyzed or whether variability in observations is biologically relevant. CADDIS cautions that "*Concluding that a candidate stressor is or is not the cause based on a correlation coefficient is inappropriate because:*

- *Stressors often covary with each other and with natural environmental attributes. A strong relationship between the biological response and candidate cause could reflect a covarying stressor or natural factor other than the candidate cause,*
- *Hypothesis testing was designed for interpreting controlled experiments with replicates and random assignment of treatments, and*
- *Field data from observational studies rarely include replicates and "treatments" are not randomly assigned, therefore even strong associations do not prove causation.*

These cautions in CADDIS have not been adequately addressed in the EPA Conductivity benchmark study.

4 CALCULATION OF SPECIES SENSITIVITY DISTRIBUTION

The EPA Conductivity Benchmark study calculates a Species Sensitivity Distributions (SSD) from the data set based on the assumption that conductivity is the primary and only factor responsible for the local extirpation of aquatic genera. On the other hand, a site specific analysis of potential causes of biological impairment would undoubtedly reveal that the causes of biological impairment are multifaceted and conductivity, as a measure of a particular mix of dissolved salts, may be a causal factor only at the sites with very high levels of conductance. The EPA study needs to utilize and reflect the relevant types of evidence as outlined in CADDIS in a specific causal analysis performed at each of the sites included in the study. One of relevant types of evidence that CADDIS recommends utilizing in the causal assessment is an SSD. As noted in CADDIS, "*SSDs are generated from **laboratory data** and can be used in at least two ways in causal analysis: to generate predictions that may be confirmed by site data and to quantify stressor-response models based on laboratory test data*." The EPA study has perverted the approach because it uses the field data to calculate the SSD based on the quickly arrived at and erroneous assumption that conductivity as the only factor that is causing the local extirpation of aquatic species observed in field studies.

Extirpation for a genus is defined in the EPA Conductivity Benchmark study as the conductivity value (referred to as the XC_{95}) above which very few, i.e., less than 5%, of the observations of a particular genus are likely to be found among the 2,145 sites in the West Virginia data set. A genus was excluded from the extirpation calculation if it was never observed at reference sites or if it was observed at less than 30 of the 2,145 sampling sites. EPA failed to provide any justification for the 30-site threshold.

For example, an XC_{95} conductivity value of 101µS/cm was calculated for the extirpation of the stonefly genus *Remenus*. This calculation assumes that conductivity of the water is the reason that less than 5% of the observations of the presence of the genus *Remenus* were found at sites with conductivity levels greater that 101µS/cm. Yet, *Remenus* species were found at only 35 of the sites in the West Virginia Data base. Kondratieff and Baumann (2000) show the occurrence of only one *Remenus* species (*Remenus bilobatus*) occurring in West Virginia and that its geographic distribution is limited to Webster and Mingo counties within the state. We are not aware of the factors that limit the geographic distribution of *Remenus* to a few isolated locations within the state but it is unlikely to be related solely to conductivity. Similar conclusions can also be reached concerning the stonefly genera *Alloperla* and *Utaperla*, although the estimated XC_{95} conductivity values for these genera were estimated at 228 µS/cm and 224 µS/cm, respectively. According to Kondratieff and Baumann (2000) the geographic distribution of *Alloperla* extends to 6 counties within the State while *Utaperla* was not shown to occur in West Virginia. One species *Utaperla gaspesiana* was found in two counties in neighboring Pennsylvania. *Utaperla* was observed at 47 of the sites in the West Virginia data set while Alloperla was observed at 96 of the sites.

The XC_{95} conductivity value for the 170 genera that were observed at more than 30 sampling sites in the West Virginia data set was used to develop an SSD based on the assumption that conductivity is the factor that is responsible for the extirpation of aquatic genera at all the sites in the data set. The conductivity benchmark was determined from the SSD as the conductivity level that is not exceeded by the XC_{95} conductivity value for 5% of the 170 genera. The assumption was made that if conductivity levels are below the benchmark, at least 95% of genera will be protected. Thus, the calculated benchmark is directly dependent upon the XC_{95} conductivity value estimated for the 8 genera in the data set with the lowest calculated XC_{95}. These are the genera *Alloperla*, *Cinygmula*, *Drunella*, *Lepidostoma*, *Leptophlebia*, *Pyenspsyche*, *Remenu*s, and *Utaperla*. Of these genera, only the mayfly genus *Drunella* was observed at more than 4.5% of the sample sites in the data set. *Drunella* was observed at approximately 8% of the sample sites based on year round data but at less than 1% of the sites based on summer samples.

It appears that these genera with the lowest calculated XC_{95} conductivity values are limited by more than conductivity as they are observed at very few of the sites. It is inappropriate to develop SSDs from field observations of specific genera where the presence and absence of the genera are related to factors other than conductivity. This is the reason why SSDs are calculated from acute or chronic toxicity results from controlled laboratory experiments.

The analysis of uncertainty presented in the EPA Benchmark study is an obfuscation of the actual sources of uncertainty concerning conductivity levels that are protective of aquatic genera in West Virginia. The uncertainty analysis involved a “Bootstrap” method for estimating the variation in the calculated XC95 values for each genus used in the derivation of the benchmark by sampling with replacement from the data set that was used to derive the benchmark. This involved generating 1000 data sets each with 2,145 observations that were obtained by random sampling from the data set with replacement. Thus, the generated data sets will exclude some observations and include duplicates or multiple observations from some of the sample sites. In effect this is a method of massaging the original data in an effort to examine the variation resulting from the observations obtained at each of the sampling sites included in the study. The benchmark conductivity value was then estimated for each of the 1000 “bootstrap samples” in order to generate an estimate of the mean and confidence interval in the benchmark value.

The EPA Conductivity report applied the same methodology to data from the State of Kentucky and obtained a conductivity benchmark estimate of 319 µS/cm with a lower confidence bound of 180 µS/cm and an upper bound of 439 µS/cm. The report implies that that the confidence interval generated from this Bootstrap analysis of these two data sets adequately represents the uncertainty in the conductivity benchmark. The report also states that the sampling variance may be the largest component of total uncertainty because the confidence intervals for the estimates of the benchmark conductivity from the two data sets overlap.

We strongly disagree with this conclusion. The largest source of uncertainty is in the application of the methodology and assumptions, particularly in the assumption that conductivity is the variable controlling the extirpation of aquatic species in the study area. An indication of the fallacy of this assumption is shown in the uncertainty analysis performed by EPA. Figure 7 of the EPA Conductivity Benchmark study shows the 95% confidence bounds for the XC95 values for the 35 genera with the lowest XC95 value estimated from the West Virginia data set. A copy of that figure is provided below. The bootstrap analysis of the West Virginia data set shows a considerable range in the XC95 values that occur just due to the variation in sampling. For example, the XC95 value calculated for the caddisfly genus *Lepitostoma* varies widely from less than 100 to over 700 µS/cm, due to the random variation in the sampling points as reflected in the bootstrap analysis. This high variability in XC95 values calculated for each genus in the data set clearly demonstrates that conductivity is not controlling the extirpation of aquatic species for most of the sampling points in the study area.

The uncertainty analysis does not address many of the other assumptions that were made. For instance, a genus was excluded from the extirpation calculation if it was observed at less than 30 of the 2,145 sampling sites. Perhaps a more appropriate threshold for exclusion is if it was not observed at more than 10% of the sites. This would eliminate the genera that are not commonly observed in the study area. Also, a more appropriate measure of the extirpation of a particular genus would be the XC100 or the conductivity above which observations of a particular genus are not found among the 2,145 sites. This criterion for extirpation should also have been examined in assessing the uncertainty in the calculated benchmark.

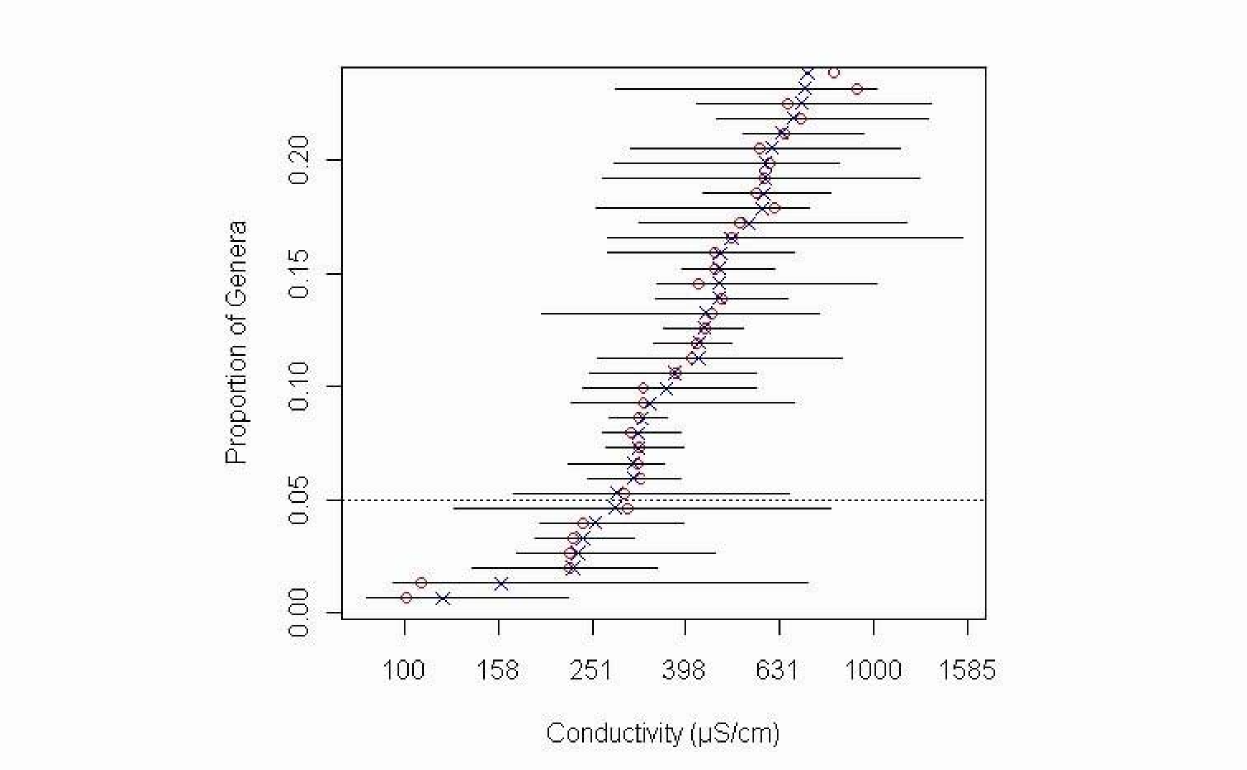

Figure 7 (from USEPA (2010). The cumulative distribution of XC95 values for the 35 most sensitive genera (red circles) and the bootstrap-derived means (blue x symbol) and two-tailed 95% confidence intervals (whiskers). (The 5th percentile is shown by the dashed line.)

5 CONCLUSIONS AND RECOMMENDATIONS

There is no scientific basis for the species sensitivity distribution (SSD) that was developed from biomonitoring field data in the EPA Conductivity Benchmark study. There are a multitude of causal factors that explain the extirpation of genera from sampling locations included in the EPA data set, including predation/competition from multiple species, influences of habitat, temperature, geology, flow, dissolved oxygen, water quality, and limited geographic distribution for some genera. The lack of scientific support is also demonstrated by the substantial scatter in the relationship between the number of Ephemeropteran genera and conductivity as shown Figures A-1 and A-2 of Appendix A of the EPA Conductivity Benchmark study and by the wide variability in XC95 values calculated for each genus in the bootstrap analysis of the data as shown in Figure 7 of the EPA Conductivity Benchmark study.

The EPA attempt to develop a water quality criterion for conductivity from biomonitoring field data is not appropriate and should be replaced by the traditional method for establishing water quality criteria. The traditional method for establish water quality criteria relies on information from controlled laboratory experiments to identify a direct relationship between organisms' response and exposure concentrations of a water quality constituent without the confounding influences. The development of the conductivity benchmark based on field data has little more validity than developing a percent woodland benchmark based on field data. It does not address the actual factors that account for the presence or absence of aquatic organisms at a particular site or the causes for biological impairment.

The biomonitoring approach is "best used for detecting generalized and non-specific impairments to biological integrity, and for assessing the severity of those impairments," as noted by Gibson et al. (2000). Chemical and toxicity tests and habitat assessments are needed to identify probable causes of impairment as the basis for corrective measures. An analysis of the likely causes of biological impairment needs to be performed at each site following a structured approach such as that outlined in CADDIS.

The use of the conductivity benchmark as a water quality criterion would actually be detrimental because it would focus on the reduction of conductivity to protect benthic environments, when the cause of impairment at most of the sites is due to other factors, including a change in the flow regime, the presence of an impoundment, changes in the physical structure of the stream bed or changes in the quality and quantity of food resources preferred by various benthic insect genera. The literature shows that the distribution and abundance of mayfly communities are strongly dependent on habitat composition structure and that often mayflies are absent from benthic environments downstream of impoundments. Thus, a mining company could invest in water treatment to meet conductivity limits for discharges and find that there is little or no change in the benthic insect genera downstream of the discharge following treatment.

Also, a water quality criterion for conductivity designed to protect biological conditions would result in a substantial increase in the number of impaired streams that are impaired with respect to listing under Section 303(D) in the state of West Virginia. As previously discussed, 409 of the sites in the West Virginia data base would be added when current information indicates that these streams are not biologically impaired.

6 REFERENCES

Arnwine, D. H., K. J. Sparks and R. R. James. (2006). Probabilistic Monitoring of Streams Below Small impoundments in Tennessee. Tennessee Department of Environment and Conservation, Division of Water Pollution Control. Nashville, Tennessee

Bauernfeind, E. and O. Moog. (2000). Mayflies (Insecta: Ephemeroptera) and the assessment of ecological integrity: a methodological approach. Hydrobiologia 422/423: 71–83, 2000. M. Jungwirth, S. Muhar & S. Schmutz (eds), Assessing the Ecological Integrity of Running Waters. Kluwer Academic Publishers. Printed in the Netherlands

Chapman, P.M., H. Bailey, and E. Canaria (2000). Toxicity of total dissolved solids associated with two mine effluents to chironomid larvae and early life stages of rainbow trout. Environ. Toxicol. Chem. 19: 210–214.

Echols, B.; R.J. Currie; and D.S. Cherry (2009) Preliminary results of laboratory toxicity tests with the mayfly, *Isonychia bicolor* (Ephemeroptera: Isonychiidae) for development as a standard test organism for evaluating streams in the Appalachian coalfields of Virginia and West Virginia. Environ Monitor Assess. published online 04 November 2009.

USEPA (2010). Field-Based Aquatic Life Benchmark for Conductivity in Central Appalachian Streams. EPA/600/R-10/023A. March 2010 External Review Draft

GEI Consultants, Inc. (GEI). (2009). Update on Issues for Further Investigation in Regard to the ―Pond et al. Study on Effects of Mountaintop Mining and Valley Fill on Benthic Invertebrate Communities. Technical Memorandum, GEI Consultants, Inc. Denver, CO.

GEI Consultants, Inc. (GEI). (2010). Preliminary Technical Review of “A Field-based Aquatic Life Benchmark for Conductivity in Central Appalachian Streams”. Technical Memorandum, GEI Consultants, Inc. Denver, CO.

Gibson, G.R., M.L. Bowman, J. Gerritsen, and B.D. Snyder. (2000). Estuarine and Coastal Waters: Bioassessment and Biocriteria Technical Guidance. EPA 822-B-00-024. U.S. Environmental Protection Agency, Office of Water, Washington, D.C.

Illinois Environmental Protection Agency. April 2006. “Preliminary Technical Justification for Changing Water Quality Standards for Sulfates, Total Dissolved Solids and Mixing Zones.”

Iowa Department of Natural Resources, February 9, 2009 “Water Quality Standards Review: Chloride, Sulfate and Total Dissolved Solids.”

Kennedy, A.J.; D.S. Cherry; R.J. Currie. (2003) Field and laboratory assessment of a coal processing effluent in the Leading Creek Watershed, Meigs County, Ohio. Arch Environ Contam Toxicol 44(3):324−331.

Kennedy, A.J.; D.S. Cherry; R.J. Currie. (2004) Evaluation of ecologically relevant bioassays for a lotic system impacted by a coal-mine effluent, using Isonychia. Environ Monit Assess 95:37−55.

Kennedy, A.J.; D.S. Cherry; and C.E. Zipper. (2005) Evaluation of ionic contribution to the toxicity of a coal-mine effluent using Ceriodaphnia dubia. Arch Environ Contam Toxicol 49:155−162.

Kondratieff, B. C. and R.W. Baumann. (2000). Stoneflies of the United States. Jamestown, ND: Northern Prairie Wildlife Research Center Online. http://www.npwrc.usgs.gov/resource/distr/insects/sfly/index.htm (Version 12DEC2003).

Maggard, R.R. (2009). The effects of specific conductance of water on the benthic macroinvertebrate community downstream of coal mining activities. Paper presented at the West Virginia Mine Drainage Task Force Symposium.

Mount, DR; D.D. Gulley; and R. Hockett. (1997). Statistical models to predict the toxicity of major ions to Ceriodaphnia dubia, Daphnia magna, and Pimephales promelas (fathead minnows). Environ Toxicol Chem 16(10):2009−2019.

Ziegler, CR; G.W. Suter; and B.J. Kefford. (2007) Candidate Cause: Ionic Strength. Available online at www.epalgov/caddis