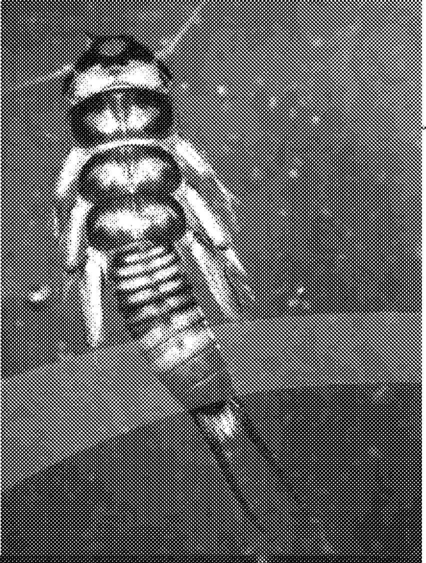


Consulting
Engineers and
Scientists

Review of EPA's Draft Field-Based Methods for Developing Aquatic Life Criteria for Specific Conductivity

EPA-822-R-07-010

PREPARED ON BEHALF OF THE NATIONAL MINING ASSOCIATION

Submitted to Docket ID No.: EPA-HQ-OW-2016-0353

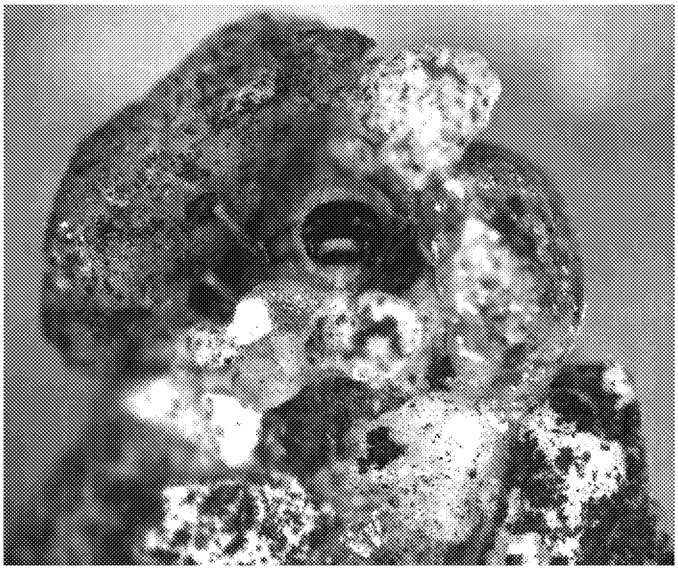
April 2017

Review of EPA's Draft Field-Based Methods for Developing Aquatic Life Criteria for Specific Conductivity

GEI Consultants
Consulting
Engineers and
Scientists

PREPARED ON BEHALF OF THE NATIONAL MINING ASSOCIATION

Submitted to Docket ID No.: EPA-HQ-OW-2016-0353



Submitted by:
GEI Consultants, Inc.
4601 DTC Boulevard, Suite 900
Denver, CO 80237

April 2017
Project 127260

Suzanne Pargee

Suzanne Pargee, Project Manager

Robert W. Gensemer

Robert W. Gensemer, Ph.D., Reviewer

Table of Contents

Figures

Tables

Appendices

List of Abbreviations and Acronyms

CCC	criterion continuous concentration
CDF	cumulative distribution function
CMC	criterion maximum concentration
CMEC	criterion maximum exposure concentration
ECDF	empirical cumulative distribution function
EC_{20}	20% effect concentration
EPA	U.S. Environmental Protection Agency
FFG	functional feeding group
GEI	GEI Consultants, Inc.
HC_{05}	hazard concentration 5th percentile
LC_{50}	median lethal concentrations
NMA	National Mining Association
PaFBC	Pennsylvania Fish and Boat Commission
SC	specific conductivity
SSD	species sensitivity distribution
WABbase	water analysis database
WVDEP	West Virginia Department of Environmental Protection
WVSCI	West Virginia Stream Condition Index
XC_{95}	extirpation coefficient 95th percentile
XCD	extirpation coefficient distribution
XC	extirpation coefficient

Executive Summary

On behalf of the National Mining Association (NMA), GEI Consultants, Inc. (GEI) has previously conducted detailed technical reviews of the draft and final EPA Benchmark Reports. Overall, as noted in those prior reviews, the methodology recommended for development of field-based conductivity criteria is inherently flawed and does not provide a reliable means of truly ensuring protection for 95% of the aquatic community as set forth in U.S. Environmental Protection Agency (EPA) guidance for development of aquatic life criteria.

EPA uses a risk assessment framework to develop the draft conductivity criteria. While we support this approach in general, there are substantial flaws in the Problem Formulation that cannot be adequately addressed in the Analysis Plan to achieve appropriate goals and endpoints. In particular, the XCD method does not develop criteria that would be protective of 95% of aquatic species, but instead effectively sets conductivity criteria concentrations very close to natural background concentrations, which is not consistent with the broad ecological integrity goals of the Clean Water Act. EPA appears to only be concerned with protecting against presumed extirpation at the genus level, but gives no consideration to diversity or abundance or other straightforward methods of evaluating overall health of aquatic communities, even though diversity is a well-established means of evaluating benthic invertebrate community health and structure.

While the EPA states that this Draft Conductivity Criteria methodology is modeled after the 1985 Guidelines, the method described in the EPA Draft Conductivity Criteria differs significantly in that the points in the species sensitivity distribution (SSDs) consist of extirpation coefficients (XCs) rather than median lethal concentrations (LC_{50}s) or chronic response values (e.g., EC_{20}s) from exposure to a single chemical in controlled laboratory studies. The resulting criterion continuous concentration (CCC) derived using this method does not represent a specific conductivity (SC) concentration that would protect 95% of the species with respect to SC. Rather, the CCC identifies the lowest possible SC concentration higher than undisturbed "background" SC that is associated with subtle changes in the presence/absence of select macroinvertebrate genera. This is a fundamentally inappropriate interpretation of the 1985 Guidelines and its goal to derive aquatic life protection criteria that are protective of all but 5% of the most sensitive species to a pollutant or stressor.

One of the major conceptual issues with the XCD-based CCC derivation method is the assumption that the absence of any benthic invertebrate genera is solely due to a conductivity level that exceeds the physiological limits of that genus, thereby leading to extirpation. However, their own data show substantial differences in XC_{95} values within a genus across ecoregions, thereby highlighting our concern that genera characterized as being sensitive to conductivity in one ecoregion may in fact not be sensitive to conductivity in another ecoregion. Such large variability in the purported physiological limits of "sensitive" genera raises considerable uncertainty regarding the applicability of conductivity as a determinant of the frequency of occurrence (or absence/extirpation) of a given taxa. This is an extremely important issue that puts into question the entire approach used in EPA's Benchmark Report and needs to be fully investigated by EPA before any conductivity criterion is adopted on a more national basis.

Traditionally, EPA methods for derivation of protective aquatic life criteria are based on the existence of consistent (i.e., unimodally decreasing, or negative) stressor-response relationships. In the case of the SC criteria, if conductivity were indeed the primary response driver, a consistent stressor-response should be apparent, where each of the taxa would respond negatively (e.g., decreased probability of occurrence) to increased conductivity. However, when evaluating taxa response profiles from Ecoregions 69 and 70, only about 40% of taxa show this negative response. In contrast, 20% actually show a positive response to increased SC, 20% show a negative response at both low and high SC, and 15% show little to no response. There is simply no way to reconcile these widely conflicting stressor-responses into a single SC criteria concentration that would actually be protective of 95% of all macroinvertebrate taxa expected to be present at a site.

Conductivity tolerance (or perhaps it's actually "preference"), as a surrogate summary metric for total ionic content of water, is evidently more akin to temperature tolerance/preference than to that of traditional toxic chemicals such as metals or pesticides. Just as there are taxa (*e.g.*, specific species of fishes) that prefer warm waters and those that prefer cold waters, there are invertebrate taxa that appear to prefer low conductivity and those that appear to prefer higher conductivity. We do not agree that conductivity is necessarily the direct causative agent for the presence and absence of taxa in any particular location in the datasets used by EPA. Ionic tolerance may play one role in establishing which species inhabit a site, but other chemical, physical, and biological factors discussed in our review likely have equally important roles.

EPA believes the various genera do not need to exhibit the same responses to conductivity for the calculation of the HC_{05}, and this is evident when looking at the top ten "most sensitive" taxa (i.e., those with the lowest XC_{95} values) across selected ecoregions. Using data from seven ecoregions selected from the list of 24 ecoregions used in the EPA's development of the Background to Criteria Regression approach, the top ten most sensitive taxa from each ecoregion were identified and capture probability/response-curves developed. There was little overlap in the "stressor response profiles" for these ten genera across ecoregions, which further highlights contradictions between assumed physiological sensitivities to conductivity across ecoregions. The resulting differences in the genus' extirpation coefficients (XC95) across multiple ecoregions provide further evidence that conductivity is likely not the causative factor in the presence/absence of these genera.

For the EPA Draft Conductivity Criteria, EPA correctly points out that the data do not generally exist to directly evaluate what SC level might protect aquatic life from acutely toxic (i.e., short-term) exposures. Instead, EPA suggests derivation of a Criterion Maximum Exposure Concentration (CMEC) which is the 90th percentile of SC observations at site with water chemistry conditions such that they meet the CCC. While this 90th percentile-based CMEC may indeed represent a less stringent SC criterion than the CCC, this is a largely arbitrary percentile selection that has little direct relationship to a SC concentration that would protect from acute exposures to nine taxa at relevant times of year. Therefore, we suggest that until such data can be made available, EPA should not propose a method for derivation of CMECs for SC.

We agree with EPA that an evaluation of causality is critically important in any scientific endeavor, particularly when proposing the use of a field-based method using macroinvertebrate community structure data that is subject to a potentially high degree of confounding. For the EPA Draft Conductivity Criteria, however, no new causality analysis has been conducted. Rather, EPA only choses to suggest that "it is good practice" to further evaluate the performance of the XCD model. We agree, but are very concerned that EPA has not chosen to conduct such an analysis here.

We also reviewed Appendix G, which discusses the extent to which EPA feels the ecoregional criteria for benthic macroinvertebrates are protective of fish. While EPA notes that the observed, and presumed, tolerance to conductivity may be due to the probability of capturing and enumerating fish, additional analysis is needed. This describes one of the main issues with using this method on fish. The problem formulation section earlier in the EPA Draft Conductivity Criteria suggests that indirect effects to fish, such as prey (benthic invertebrate) abundance, are possible. Yet, Appendix G does not provide any further examination on this issue, making this a critical but untested statement by EPA.

In summary, there are a number of significant weaknesses in the approach used by EPA to develop conductivity criteria on a national scale that preclude its adoption. Setting criteria with this approach would not be protective of the overall aquatic invertebrate community.

1. Introduction

GEI Consultants, Inc. (GEI) toxicologists, water quality specialists, and regulatory strategists are recognized experts in water quality effects on aquatic life. We frequently provide expert testimony and support for regulatory water quality hearings, environmental assessments, use-attainability analyses, and ambient water quality standards development. Our personnel have served as invited experts for tiered aquatic life use evaluations, provided peer review and independent development of multiple aquatic life criteria using U.S. Environmental Protection Agency (EPA) methods, and provided technical review of mountaintop mining and other land use issues—including the original EPA Conductivity Benchmark for the National Mining Association, and participated in development of a structured framework for stressor analyses for the Water Environment Research Foundation. We have been involved with water quality and aquatic life issues in waters throughout the U.S., including the collection of considerable water quality and biological data from a wide variety of waterbodies and development of site-specific standards for many metals and metalloids. As such, we can provide a unique perspective on EPA's draft criteria document and respectfully submit the following comments.

This report summarizes GEI's technical review of EPA's public review draft of Field-Based Methods for Developing Aquatic Life Criteria for Specific Conductivity (EPA 2016; hereafter referred to as "EPA Draft Conductivity Criteria"). This review is being conducted on behalf of the National Mining Association (NMA) in support of their comments being submitted as part of Docket EPA-HQ-OW-2016-0353.

The EPA Draft Conductivity Criteria document describes field-based methods for derivation of aquatic life protection criteria in flowing waters for dissolved inorganic ions, measured collectively as specific conductivity (SC). This method relies heavily on documents such as the EPA's Guidelines for Deriving Numerical National Water Quality Criteria for the Protection of Aquatic Organisms and Their Uses (Stephan et al. 1985, hereafter referred to as the "1985 Guidelines") and A Field-Based Aquatic Life Benchmark for Conductivity in Central Appalachian Streams (EPA 2011, hereafter referred to as the "EPA Benchmark Report"). EPA's Benchmark Report used field data from stream benthic macroinvertebrate surveys to derive an aquatic life benchmark for SC that, according to the EPA, may be applied to waters in the central Appalachian Region that are dominated by salts of calcium (Ca^{2+}), magnesium (Mg^{2+}), sulfate (SO_4^{2-}) and bicarbonate (HCO_3^-) at circumneutral to mildly alkaline pH. While the EPA states that this conductivity benchmark was derived using a method modeled after the 1985 Guidelines, the use of field benthic macroinvertebrate community data as opposed to individual species laboratory toxicity data represents a significant technical departure from this guidance.

The EPA Draft Conductivity Criteria builds upon methods described in the EPA Benchmark Report to estimate a protective maximum exposure concentration, duration, and frequency for SC on a more national basis. As such, the field-based approach includes methods for assessing the application of field-based criteria developed in one geographic region to another. Such methods are being proposed to develop SC criteria on the scale of Level III ecoregions (Omernik 1987 and 1995); however, EPA suggests that in some cases it may be appropriate to develop criteria on a different scale due to significant variation in background SC levels across an ecoregion. The method is intended to protect 95% of resident macroinvertebrate genera present based on field data from within the ecoregion.

On behalf of NMA, GEI previously conducted detailed technical reviews of both the draft and final EPA Benchmark Reports (GEI 2010 and 2012). It is important to reinforce the issues and concerns raised in our prior reviews, as they are directly relevant to the scientific basis for this draft conductivity criteria document. Based on our reviews of both the draft and final aquatic life benchmark for conductivity, the data and assumptions used by the EPA to develop the benchmark and the subsequently proposed methods for deriving a field-based aquatic life criteria are inherently flawed. Simply put, EPAs proposed derivation of field-based SC criteria does not represent a plausible and reliable means of truly ensuring protection for 95% of the full aquatic community.

This report presents GEI's technical evaluation of the EPA Draft Conductivity Criteria we provide comments beginning with the problem formulation (Section 2), including a discussion of the nature of and measures of effects proposed by EPA. We then discuss in Section 3 EPAs analysis plan which represents the core quantitative elements of EPAs proposed SC criteria. In this section, we include a discussion of how these criteria compare to EPA's 1985 Guidelines for development of ambient aquatic life criteria, a discussion of our concerns with the extirpation coefficient distribution (XCD) methodology used to derive the Criterion Continuous Concentration (CCC), concerns with development of EPA's background-criteria estimation model, and the lack of discussion of causality. We also provide a brief review of Appendix G which proposes fish as an alternative assessment endpoint.

2. Comments on Problem Formulation

As is becoming standard practice for new or revised aquatic life criteria documents, the EPA Draft Conductivity Criteria is built upon a risk assessment-based framework (EPA 1992), the core elements of which are a Problem Formulation and Analysis Plan. Perhaps the most critical phase of a properly conducted ecological risk assessment is the Problem Formulation which sets the stage for what problem is being addressed by the assessment, identification of the stressor in question, and how to measure the adverse effects in an ecologically meaningful way. From there, the Analysis Plan describes the numeric methods needed to achieve the goals and endpoints set forth in the Problem Formulation. This is a solid foundation upon which to build aquatic life criteria with which we support.

The Problem Formulation phase of any ecological risk evaluation is the critical first planning step that "establishes the goals, breadth, and focus of the assessment" (EPA 1992). As a result, no matter how thorough or rigorous the Analysis Plan may appear to be, any flaws in the Problem Formulation can significantly call into question the entire assessment. In the case of EPA's Draft Conductivity Criteria, it is our opinion that the assessment and measurement endpoints in the XCD method contain fundamental flaws in reasoning that do not support use of the species sensitivity distribution-based (SSD-based) analytical framework set forth in the 1985 Guidelines, as discussed in Section 3.1 (Stephan et al. 1985). Specifically, the hazard concentration 5th percentile (HC_{05}) values for SC derived from distributions of extirpation coefficient 95th percentile (XC_{95}) values do not represent protection of all but 5% of even macroinvertebrate genera from extirpation. Instead, the XCD method calculates a lower percentile of genera (ranked by their XC_{95} distributions) that is closest to the background SC of minimally affected or undeveloped macroinvertebrate assemblages.

A central theme for our concerns with EPAs XCD method is that we are not convinced that the XC_{95} values represent the true sensitivities of macroinvertebrate genera to SC. In large part, this is because of the disparate kinds of SC "response" curves observed between genera (GEI 2010 and 2012); and indeed wildly different XC_{95} values even within a genus (see Section 3.2.2.2 below). Even if we were to accept that the XC_{95} for each genus represents a reliable representation of the physiological sensitivity of organisms within that genus to SC, many genera with higher XC_{95} values represent likely preferences for higher SC values, not simply tolerance. As such, these organisms would naturally not be expected to be observed at lower SC values near the criterion concentration (Section 3.2.2.3). As discussed later in this review, the variable response curves should be interpreted as representing the variable "preferences" of organisms for different ranges of conductivity – not differential sensitivity or physiological "tolerance" assumed by EPA. In broader terms, these different "response" curves may not at all represent physiological tolerance to SC as a discrete stressor, but rather the culmination of the many ecological factors shaping which taxa will or will not be observed in a particular sample.

Therefore, it is essentially impossible for the XCD method to develop criteria concentrations that are protective of 95% of species as is the overall protection goal stated in the 1985 Guidelines (Stephan et al. 1985). Given that this represents a significant departure from how the 1985 Guidelines are constructed (see Section 3.1), the EPA Draft Conductivity Criteria does not adequately defend how EPA's 95% protection goals can be achieved. This represents a fundamental flaw in the Problem Formulation portion of the proposed SC criteria methods presented therein.

Other comments regarding elements of the Problem Formulation are presented in the following subsections.

2.1 Nature of the Effect

In this section, EPA develops the argument that "the background SC of an ecoregion is strongly associated with a predictable extirpation of 5% of species or genera." While EPA correctly points out that it is inappropriate to set SC criteria below the natural background of an ecoregion, the XCD method effectively sets criteria concentrations very close to (i.e., just above) that natural SC background of the ecoregion. EPA correctly points out that "species do not occur where the SC is lower or higher than their SC tolerance" (pg. 2-14), and goes on in this section to use ecological niche theory to explain how physiological tolerances to SC can help determine where an organism is expected (or most likely) to be observed. However, EPA appears to only be interested in organisms with realized niches in which their presumed "optimum" condition is near/equal to its lower tolerance limit (e.g., as exhibited by the mayfly genus *Ephemerella*). Little to no explanation is given to what this means for organisms for which SC concentrations near EPA's proposed criteria would be significantly lower than their SC "optimum."

This, in effect, means that the nature of the ecological effect to be addressed by these SC criteria is the smallest detectable difference in presence/absence patterns of genera with XC_{95} values close to the HC_{05}. Again, this does not represent a SC concentration that would be associated with protection of 95% of the macroinvertebrate (or even all animal) genera; indeed, no such single SC concentration can do so (Section 3.2.2.3). Setting an ecological effect to be so similar to natural background is not, in our opinion, consistent with the broad ecological integrity protection goals of the Clean Water Act (EPA 1991). At the very least, EPA must better explain how the nature of the effect is consistent EPA's well-established protection goals for the entire aquatic community.

2.2 Assessment Endpoints and Measures of Effect

The EPA Draft Conductivity Criteria are based on an assessment endpoint that is largely based on macroinvertebrate assemblages (the "entity" to be protected) that provide multiple ecosystem services. EPA discussed the importance of this entity at some length, but does not provide much, if any, discussion regarding the "attribute" to be protected by the SC criteria. If one assumes the attribute EPA wishes to protect is extirpation at the genus level, they do not justify why this attribute is superior over other possible attributes such as macroinvertebrate diversity and abundance (which are cited as examples of ecosystem services on pg. 2-21 of the EPA Draft Conductivity Criteria). Indeed, macroinvertebrate community diversity is an extremely well-established means of evaluating community structure that has been widely used in the development of numeric multimetric indices for environmental protection purposes (Barbour et al. 1999; Bukantis 1998; Jessup 2010; Jessup and Gerritsen 2002; Royer and Mebane 2002).

Therefore, EPA should at a minimum explain why such a common metric was not selected in lieu of what was selected for the EPA Draft Conductivity Criteria. We are concerned over this obvious omission because, as illustrated in Section 2.2.1 below, relationships between SC and macroinvertebrate diversity can be extremely different than the distribution of XC_{95} values might suggest. Another equally plausible ecological "attribute" EPA could have considered would the functional feeding groups represented by different macroinvertebrate genera; this is explored in Section 2.2.2 below.

2.2.1 Macroinvertebrate Diversity

The West Virginia Department of Environmental Protection (WVDEP) uses a multi-metric diversity index called the West Virginia Stream Condition Index (WVSCI; EPA 2000) which was developed specifically for use in interpreting compliance with their state narrative water quality standards which include "No significant adverse impact to the chemical, physical, hydrologic, or biological components of aquatic ecosystems shall be allowed" (WVDEP 2010; Appendix B). The threshold WVSCI score for inclusion of a stream on the 303(d) impairment list is 60.6, and a WVSCII score of 68 represents acceptable ecological integrity based on the 5th percentile of available reference site scores.

The protection goals inherent to this narrative standard would appear to be equally plausible as assessment endpoints that could be used to support these SC criteria. However, WVDEP's guidance for interpretation of this narrative water quality standard clearly shows that the relationship of WVSCI score vs. conductivity is highly variable at conductivity values near the CCCs derived for ecoregions 69 and 70 of 310 µS/cm and 340 µS/cm SC, respectively. As illustrated in the figure (Figure 2-1) below, a very high proportion of streams with less than a WVSCI score of 60.6 or 68 would occur at SC concentrations at or below the CCCs EPA would propose for these regions.

This points another fundamental flaw in EPA's method. The data being used across the ecoregions were from sampling efforts designed to evaluate overall macroinvertebrate community health. As such, they are generally comprised of single samples across multiple sites in ecoregions, with laboratory processing protocols to minimize effort (i.e., often fixed-count subsampling), and development of metrics to discriminate differences in community structure between background and perturbed sites (including multiple stressors, like water quality, habitat loss, flow modifications, etc.). The sampling was not designed, nor ever intended, to capture ALL invertebrate taxa at a location. As such, it is not appropriate to use such data to attempt to prove "presence" or "absence" at any particular site – meaning any estimation of "extirpation" is meaningless.

When the data are used correctly, in the context of overall invertebrate community health, it's clear that the relationship to conductivity is less clear than the EPA approach would indicate. This is discussed further in other sections below.

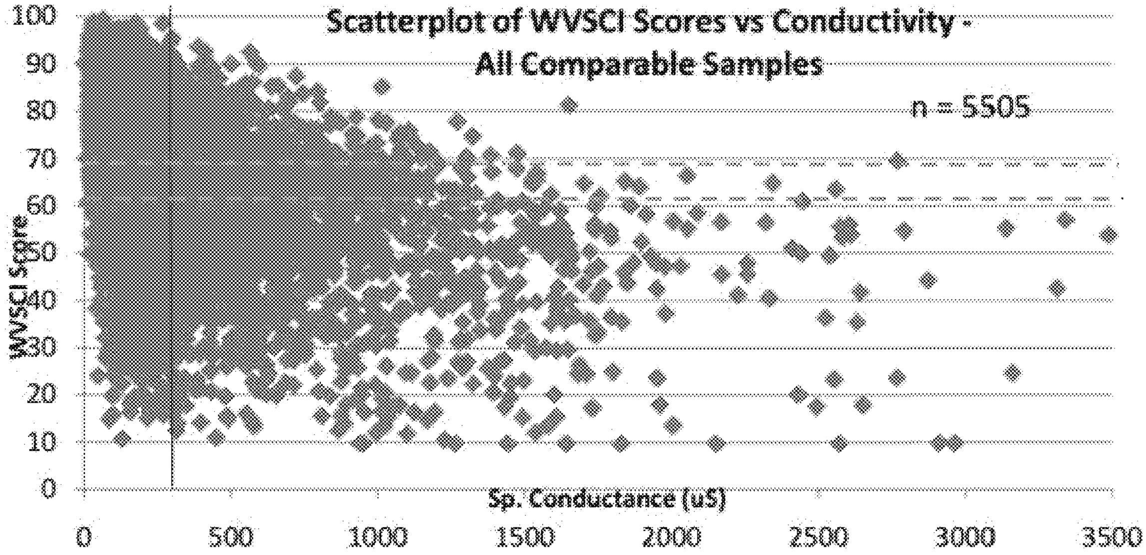


Figure 2-1: Plot of West Virginia Stream Condition Index (WVSCI) scores as a function of conductivity (from WVDEP 2010). Red line indicates approximate CCCs for Ecoregions 69 and 70. Orange line indicates 303(d) listing threshold, and green line indicates acceptable ecological integrity value.

2.2.2 Functional Feeding Groups

Although the benthic macroinvertebrate community may differ across ranges of conductivity, as discussed in Section 3.2.2.3, the ecological functionality of these different communities are retained throughout the relevant range. A biological community can be considered ecologically functional if all key functional feeding groups (FFGs) (*e.g.*, carnivores, omnivores, grazers, etc.) are present (Cummins et al., 2005). GEI conducted an analysis (GEI 2010) to evaluate whether changes in FFG balance occurs across the range of conductivity. For a sequence of conductivity values ranging from 100 to 10,000 µS/cm, we determined the FFG composition for taxa that would be present above each value using the 163 genera that were used in the EPA Benchmark Document. For example, for the conductivity value of 400 µS/cm, all taxa with an XC_{95} value less than 400 µS/cm were removed from the data set (as though extirpated). We then determined the relative proportion of each functional feeding group in the remaining taxa. We concluded that the proportional abundance of FFGs within the regional pool of taxa was nearly constant through approximately 2,500 µS/cm. Omnivores, which represent less than 2% of taxa in most locations, became effectively absent above 5,000 µS/cm. (Figure 2-2). The first major FFG to become absent with increasing conductivity, was the filter-collectors, but this occurred only when conductivity values exceeded 10,000 µS/cm.

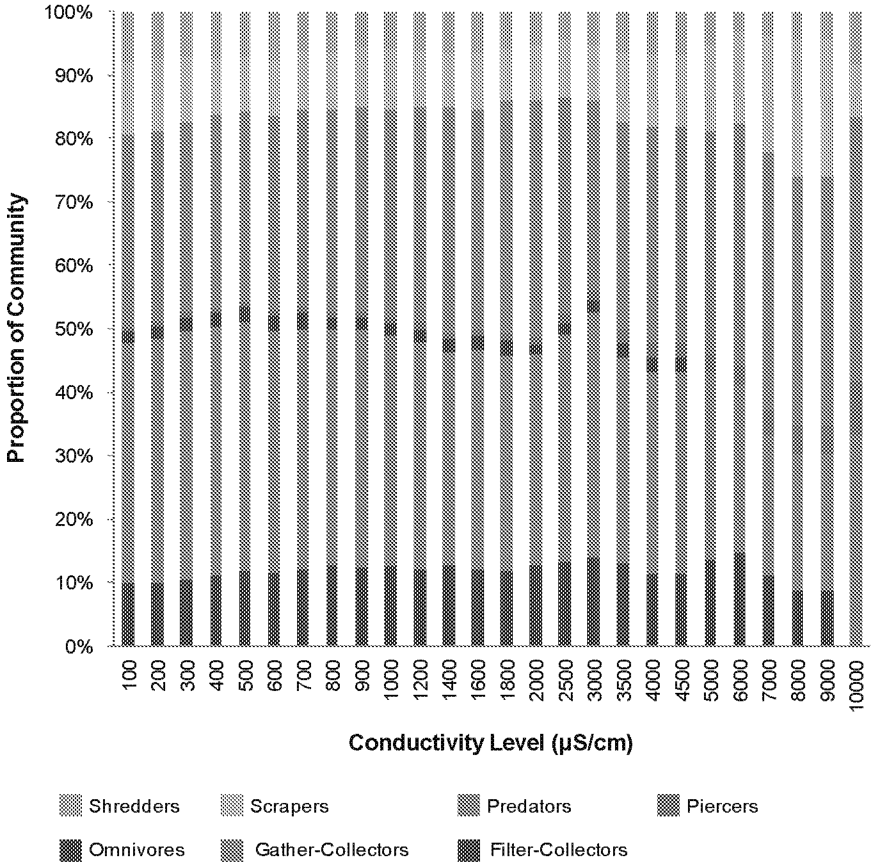


Figure 2-2: Proportion of generic richness by functional feeding group within the regional taxa pool at varying conductivity levels. All genera with an XC_{95} less than the conductivity level are considered to be unavailable. Note that the x-axis is not evenly divided.

2.2.3 Measures of Effect

The measures of effect in Section 2.6.2 of the EPA Draft Conductivity Criteria are based on the 5th percentile (HC_{05}) of the distribution of XC_{95} values. EPA recognizes that not all species within a macroinvertebrate genus will necessarily exhibit the same physiological sensitivity to SC (footnote 2, page 2-22), and even that specific genera and/or their resulting XC_{95} values may only serve as indirect "surrogates" of actual sensitivity in field populations (much like SSDs for more traditional criteria). However, as we show in Section 3.2.3.1, differences in XC_{95} values for the *same* taxa in different ecoregion show substantial variability which stretches the credibility of these assumptions. At the very least, this variability represents an untested assumption upon which their proposed SC criteria are based.

2.3 Selection of a Field-Based Method

Given our concerns noted above, it is premature for EPA to use a field-based method for derivation of SC criteria using the methods described in the EPA Draft Conductivity Criteria. With respect to the Problem Formulation section of these draft criteria, we note significant concerns with how the 95% protection goals of aquatic life protection criteria can possibly supported by the XCD method, and that the assessment endpoints and measures of effect would benefit from an analysis of other alternatives that may achieve the same goals. Finally, as we have noted in previous review of the conductivity benchmark (GEI 2010, and 2012), we also do not agree with all aspects of EPA's causality analysis suggesting that the many confounding factors which control the presence vs. absence of macroinvertebrate genera have been addressed. And as noted in Section 3.2.3.1 below, EPA fails to heed their own advice to conduct new causality analyses for ecoregions outside of 69 and 70 for which their original analysis was done. Therefore, we consider the issue of causality to be as yet unresolved.

3. Comments on Analysis Plan

The Analysis Plan section of the EPA Draft Conductivity Criteria describes the numeric methods for derivation of the Criterion Continuous Concentration (CCC, often referred to as "chronic criteria") and Criterion Maximum Exposure Concentration (CMEC, which presumably is equivalent to what is often referred to as "acute criteria) for SC. These methods are to be used in ecoregions in which sufficient macroinvertebrate and SC data are available. In regions for which sufficient paired data are not available, methods are presented to estimate SC criteria for "new" areas using either background matching or the background-to-criteria regression method. In the sections below, we summarize our primary technical concerns and comments regarding these numeric methods for derivation of SC criteria.

3.1 General Comparison to 1985 Guidelines

While the EPA states that this Draft Conductivity Criteria methodology is modeled after the 1985 Guidelines, the use of field benthic macroinvertebrate community data as opposed to individual species laboratory toxicity data represents a significant technical departure from this guidance. As detailed in Stephan et al. (1985), the 1985 Guidelines provide a very structured approach for developing water quality criteria for the protection of aquatic life. The 1985 Guidelines method relies on data collected by conducting laboratory toxicity tests, because as stated "…it is not feasible to determine national criteria by conducting such field tests…" (Stephan et al. 1985).

Following the 1985 guidelines, a literature review is conducted and the available acute and chronic data are evaluated, as well as all test conditions under strictly controlled conditions. There are specific test acceptability criteria which must be considered before a toxicity study is considered valid and acceptable for criteria development. Some of these requirements include: use of appropriate control and dilution water, acceptable control survival, technical grade test material, whether or not species were fed during the test, test renewal type and frequency, age of organisms at test initiation, and test duration. If the specific criteria for each study is not met, the data should not be used for criteria development. Field-based criteria differ significantly from the recommendations in the 1985 Guidelines as none of these test acceptability conditions can be determined or controlled.

One example of how these acceptability conditions could affect the field based criteria is when you evaluate the use of rare taxa in the EPA Criteria Document. As we noted in our review of the EPA Benchmark Document (GEI 2010 and 2012), EPA could have controlled for the effects of rare taxa by including in their SSD only those genera that had a high capture probability in the reference sites. In discussing criteria development, the 1985 Guidelines (Stephan et al. 1985) stated that "data should usually be rejected if they are from… tests in which too many organisms in the control treatment died or showed signs of stress of disease...." EPA (2011) considered a 1% collection probability in reference sites to be acceptable, but a 1% survival rate in a laboratory test would clearly not be acceptable. Use of taxa with low capture probability should be excluded if the 1985 Guidelines were more closely followed.

The method recommended in the EPA Draft Criteria document, which was also used in EPA (2011) has the appearance of being based on the 1985 Guidelines primarily because it used the 5th percentile of an SSD as the basis for mathematical derivation of the criteria value. As we have noted in our previous reviews of the EPA Benchmark Report (GEI 2010 and 2012), an SSD represents the response of individual taxa to a toxicant as a distribution with respect to exposure and is a widely used statistical approach for derivation of regulatory aquatic life criteria worldwide (Posthuma et al. 2002, Stephan et al. 1985). It is implicitly assumed that if the exposure level of a stressor or pollutant is kept below the 5th percentile of the SSD, at least 95% of tested aquatic species (or their surrogates) composing the distribution will be protected. In this respect, EPA's conductivity data analysis nominally follows the standard methodology in aggregating species responses to SC to genera and using interpolation to estimate the chosen percentile.

However, the method described in the draft document and in EPA (2011) differs significantly from the 1985 Guidelines in that the points in the SSDs consist of extirpation coefficients (XCs) rather than median lethal concentrations (LC_{50}s) or chronic response values (e.g., EC_{20}s) from exposure to a single chemical in controlled laboratory studies. The XC is defined by EPA as the level of exposure above which a genus is "effectively absent" from waterbodies in a region—i.e., the 95th percentile of the distribution of a calculated "probability of occurrence" of a genus with respect to conductivity, or XC_{95} (EPA 2003 and 2011). The EPA Draft Conductivity Criteria goes on to derive the CCC on the basis of the 5th percentile (i.e., the hazard concentration 5th percentile; HC_{05}) of all of the available XC_{95} data for a particular ecoregion.

Our primary concern is that this is a fundamentally incorrect application of an SSD approach. In particular, and as we discussed in Section 2 of this report, the resulting CCC derived using this method *does not* represent a SC concentration that would protect 95% of the species with respect to SC. Rather, the CCC identifies the lowest possible SC concentration higher than undisturbed "background" SC that is associated with subtle changes in the presence/absence of select macroinvertebrate genera. This is a fundamentally inappropriate interpretation of the 1985 Guidelines and its goal to derive aquatic life protection criteria that are protective of <u>all</u> but 5% of the most sensitive species to a pollutant or stressor.

3.2 Concerns Regarding Derivation of the CCC

3.2.1 Establishing the Data Set

One significant problem specific to this field-based approach is that the data that will be used to calculate SC criteria are likely not going to have been collected using an experimental design appropriate for EPA's usage of the data.

Although as we noted above in our discussion of the West Virginia multimetric index, the methods used to collect and enumerate macroinvertebrates are appropriate for general comparisons of overall community health using standard metrics (e.g., taxonomic diversity or abundance, e.g., EPA 2000), they are not appropriate for making inferences about the presence or absence of any specific taxa, particularly rare taxa. The use of data collected with a study design inappropriate for the end use of the data leads to uncertainty and inaccurate results.

For example, GEI conducted an analysis of subsampling that is a required step in the field sampling method and model assumptions used in EPA 2011 to determine how much these factors affected the field-based conductivity benchmark (Roark et al. 2013; Figure 3-1). The conclusions reached in that evaluation apply to the draft EPA SC document as well, since the recommended methods are similar: 1) subsampling does affect the HC_{05} calculation – larger subsamples would result in higher HC_{05} values; 2) the number of bins used and weighting of the cumulative distribution function for conductivity also affects the HC_{05}; 3) the modeling decision to only use genera that occurred in ≥ 25 samples and at least one reference location reduced the number of taxa included in the derivation of the benchmark from about 500 taxa to only 163, which has a significant effect on the HC_{05}.

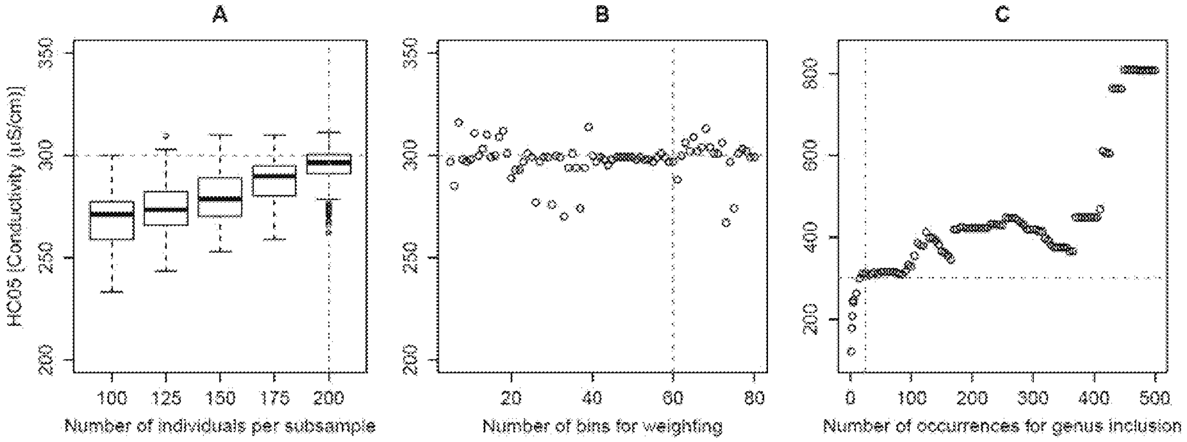


Figure 3-1: Effect of subsampling (A), binning and weighting (B), and number of stations required for inclusion of genera (C) on the final conductivity benchmark value. Vertical and horizontal dashed lines in each panel mark the x-axis value elected, and the conductivity benchmark proposed, respectively, as per Cormier et al. (2013) and USEPA (2011). In panel A, n=250 per group; dark bars represent median; boxes bound 1st and 3rd quartiles; whiskers extend to the most distant point with 1.5x the height of the box. (From Roark et al. 2013)

There is another concern with developing extirpation coefficient and species sensitivity distributions using field based count data. Specifically, the relative abundance of any one genus is not factored into the presence/absence benchmark approach. Thus, single individuals are afforded the same weighting as multiple individuals in the same sample. This is tenuous when extirpation of a genus is largely pinned on the presence or absence of a single individual. As noted above, the benthic invertebrate processing approach that utilizes subsampling of the entire sample can greatly affect the outcome of distributions. In the case of data from Ecoregion 50 in northeastern Minnesota, for the 20 most sensitive genera, single individuals represented from 7% to 63% of their respective occurrences with an average of *40% of the data being used to develop extirpation coefficients and species sensitivity distributions based on genera represented by a single organism in a sample* (Cormier 2016). And as noted below, those presumptively "sensitive" genera exhibited little relation to conductivity. With such a significant portion of the data comprised of single individuals, sampling and laboratory processing bias is likely significantly influencing the outcome of the use of presence/absence data.

3.2.2 Use of XCD Method

As has been noted in our previous reviews of the EPA Benchmark Report (GEI 2010 and 2012) it is important to recognize that the XC_{95} is not the same as an effect concentration that would typically be derived from a laboratory toxicity test (e.g., LC_{50} for acute or EC_{20} for chronic or equivalent). The mathematical construct behind the XC_{95} is quite different from that of the EC_{20}. The XC_{95} is not equivalent to a concentration at which 95% of the individuals of a genus would demonstrate an adverse effect in a controlled laboratory toxicity test with a single toxicant, nor is it equivalent to a concentration at which there is a 95% probability of absence from a field location. EPA (2011) used an empirical cumulative distribution function (ECDF) to determine the XC_{95} for each of the 163 selected genera in the EPA Benchmark Report. Simply described, the conductivity values for each sample in which a genus is present are ranked from low to high, and the XC_{95} is determined as, "the conductivity value below which 95% of the observations [in fixed-count subsamples] of the genus occur and above which only 5% occur" (EPA 2011). While this is an accurate description of the value represented by the 95th percentile of the ECDF, neither the EPA Draft Conductivity Criteria nor the EPA Benchmark Report (EPA 2011) present convincing evidence that this is an appropriate metric to quantify extirpation or to use in an SSD.

The actual relationship of the XC_{95} to the likelihood of presence or absence is unclear. For example, for any taxon, additional samples with low conductivity values well below any proposed conductivity effect concentration (e.g., 100 µs/cm) would not alter the calculated XC_{95}, because only samples with the taxon present are used in the ECDF (except for the calculation of the weights), yet the existence of additional sites where the taxon was absent should be reflected as a difference in the probability of detection or of the inference of extirpation. Yet, due to the ECDF-based estimate, additional absence data do not change the XC_{95}. In contrast, within the data set used to estimate extirpation for each taxon, each instance that a genus is incorrectly inferred to be absent based on a fixed-count subsample, the XC_{95} is incrementally reduced, which also reduces the HC_{05}. The net effect of all incorrect inferences of absence across all genera may significantly bias the HC_{05} value downward, whereas there is no error (i.e., false presence) that affects a move of the HC_{05} in the positive direction.

3.2.2.1 XC_{95} Inconsistencies

One of the major conceptual issues with the XCD-based CCC derivation method is the assumption that the absence of any benthic invertebrate genera is solely due to a conductivity level that exceeds the *physiological limits* of that genus. However, as we have noted in our prior comments to the underlying EPA Benchmark Report (GEI 2010 and 2012; Roark et al. 2013) and in Section 3.2.3.1 of this document, there are many factors that contribute to the absence of benthic invertebrates from a stream sample, such as interspecific competition, habitat suitability, other stressors (i.e., metals and sedimentation), or simply sampling and sample processing bias, which were not addressed in the document.

However, if the premise is true that absence is due to conductivity's effects on a genus' physiological limits, then any particular genus' XC_{95} should be relatively consistent across ecoregions. As a demonstration of how the XC_{95} can vary across ecoregions, GEI recently reviewed EPA's white paper and corresponding data regarding the development of a conductivity benchmark for Ecoregion 50 (Northern Lakes and Forest) in Minnesota (Cormier 2016; also Appendix D of the EPA Draft Conductivity Criteria). The XC_{95}s derived from the data set from Ecoregions 69 & 70 in West Virginia (EPA 2011), which was originally used to develop the EPA conductivity benchmark and XC_{95} methodology described in the EPA Conductivity Criteria document, were compared to the XC_{95}s for Ecoregion 50 (Minnesota) to provide some insight to the presumed physiological limits and/or conflicting limits previously identified for some taxa.

When considering the entire taxa list, most genera rankings greatly changed between ecoregions, putting the concept of conductivity (hence "physiological limits") as the prime reason for presence/absence into significant doubt. In fact, when considering the 20 most sensitive taxa based on their XC_{95} ranking for each ecoregion (i.e., 50 vs. 69 &70), there are only two genera – *Leptophlebia* and *Epeorus* – that are common to both lists, strongly indicating there is not a universally expressed relationship between presence of genera and conductivity.

There were also noticeable differences in extirpation coefficients (Figure 3-2) for genera that were common to both datasets, suggesting that any conclusion of the cause for presence/absence is not actually tied to physiological limits or any purported relationship to "conductivity." Rather, it is simply an artifact of the EPA Conductivity Criteria methodology, or the many other ecological factors other than SC that affect the frequency of occurrence (or absence) of invertebrates.

As case in point, the most sensitive genus found in Ecoregions 69 & 70 was the genus *Lepidostoma* (XC_{95} = 121 µS/cm). However, in Ecoregion 50, the genus *Lepidostoma* was actually one of the least sensitive genera (XC_{95} = 1,527 µS/cm). The difference between these two extirpation coefficients represents a positive *1,162% difference* (Figure 3-2, left panel Rank 1, right panel Rank 121). Even if the SC sensitivity of individual species within the genus may be expected to differ (e.g., Footnote 2; EPA 2016), it is difficult to imagine that species within the same genus would exhibit such a large magnitude of XC_{95}s values across two different ecoregions if conductivity is the primary cause of presence/absence of the genus.

Similarly, the third most sensitive genus in Ecoregion 50 – *Rhyacophila* (XC_{95} = 254 µS/cm) – was actually fairly tolerant of conductivity (XC_{95} >1,890 µS/cm) in Ecoregions 69 and 70; a negative 87% difference (Figure 3-2, right panel Rank 3).

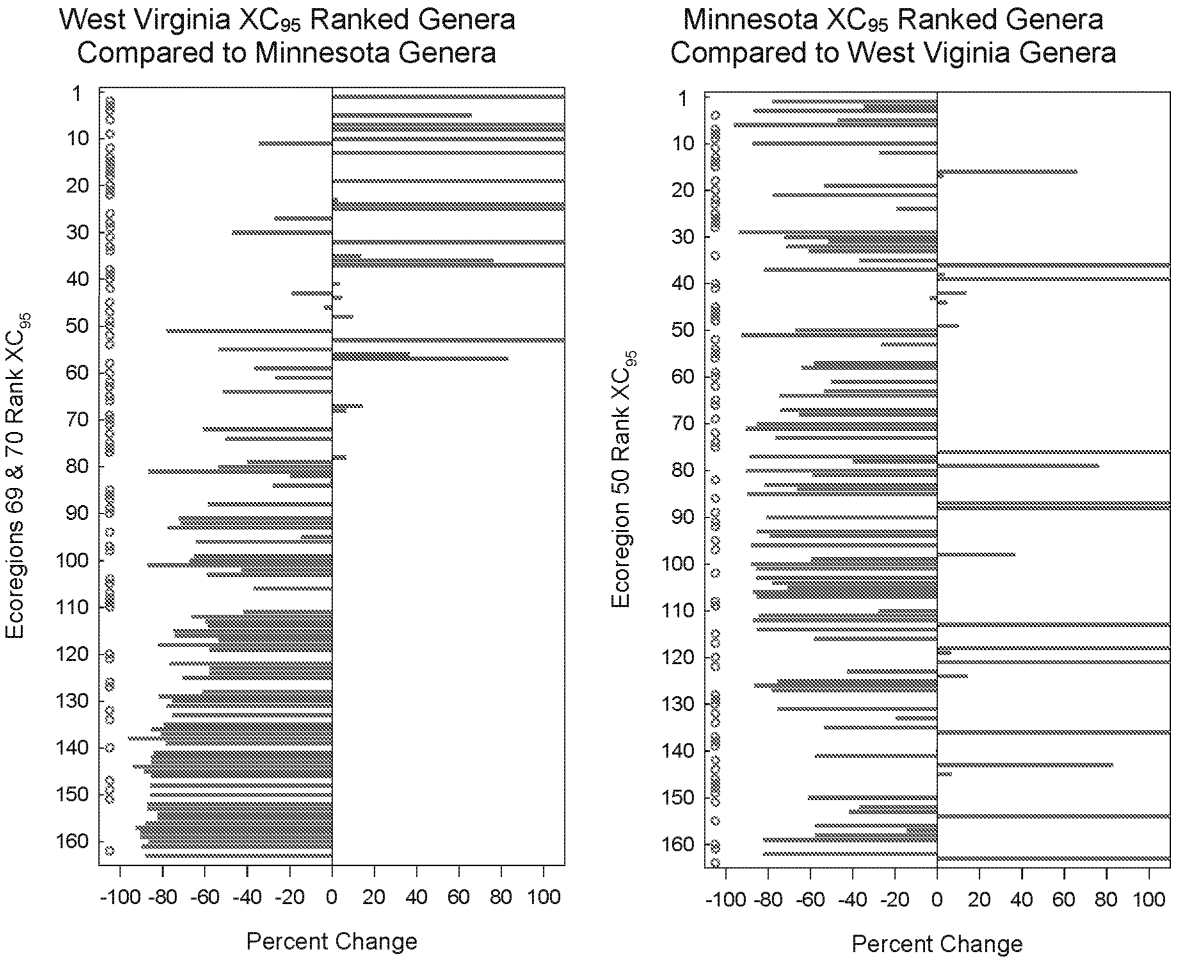


Figure 3-2: The percent change in extirpation coefficients for common genera from Ecoregions 69 & 70 compared to Ecoregion 50. Rank 1 is the smallest XC_{95} value for each ecoregion. Open circles denote unique genera to each ecoregion. Positive percent change values are truncated at 110%, because the maximum percent change was 1,162%. Percent Change = [(MN XC_{95} - WV XC_{95}) / WV XC_{95}] × 100.

These substantial differences in XC_{95} values within a genus across ecoregions highlight our concern that genera characterized as being sensitive to conductivity in one ecoregion may in fact not be sensitive to conductivity in another ecoregion. Such large variability in the purported physiological limits of "sensitive" genera raises considerable uncertainty regarding the applicability of conductivity as a determinant of the frequency of occurrence (or absence/extirpation) of a given taxa. This is an extremely important issue that puts into question the entire approach used in EPA's Benchmark Report and needs to be fully investigated by EPA before any conductivity "criterion" is adopted on a more national basis.

While arguments have been presented that differing extirpation coefficients for the same genus would be expected because the physiological limits of individual species within a genus may be different (EPA 2016; Footnote 2), that is an untested hypothesis and should be a concern with how a national level criterion approach is applied to site-specific water quality conditions.

However, we do have data from species within the genus *Boyeria* which exemplify this issue from the Ecoregion 50 and 69 & 70 data sets. All three ecoregions are represented by two common species *Boyeria grafiana* and *Boyeria vinosa* as well as an unidentified species. Given the commonality of the taxa at the species level it would be expected that extirpation coefficients would be similar for species within the genus – yet there was a difference of over 7,000 µS/cm in their XC_{95} values, again significantly challenging the premise that conductivity is the primary factor that affects the distribution of this genus.

3.2.2.2 Taxonomic Responses to Conductivity – Preference or Tolerance

As we've noted in prior reviews of the EPA Benchmark Document (GEI 2010 and 2012), and worth repeating here, one of the primary underlying principles governing the use of an SSD to derive biological thresholds is that all of the organisms represented in the distribution exhibit a negative response to an increase in the stressor in question (Posthuma et al. 2002, Stephan et al. 1985). As we pointed out in our original reviews (GEI 2010 and 2012), three types of stressor-responses are recognized by EPA (2011):

- decreasing probability of observation with increasing conductivity,
- increasing probability of observation with increasing conductivity, and
- optimal or "bell-curve" probability of observation with increasing conductivity.

In addition to these three stressor-response profiles, a fourth type not recognized by EPA—but frequently observed in their dataset— is characterized by basically no response or a bimodal, "inverse optimal" response to conductivity (Figure 3-3).

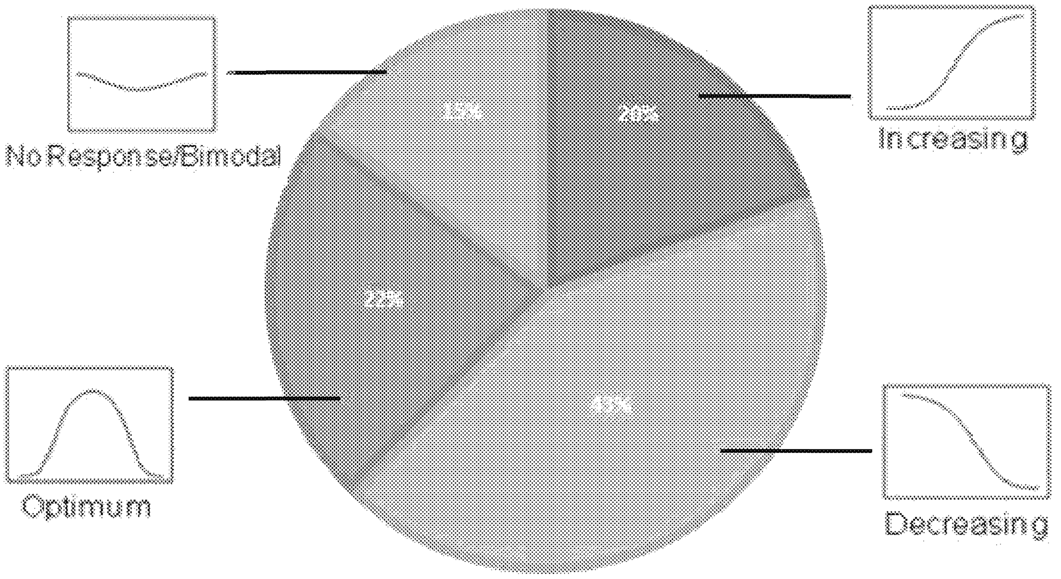


Figure 3-3: Percentage of genera with different types of stressor-response profiles with respect to conductivity and probability of capture (based on data from EPA 2016).

If EPA wishes to use stressor-response data for criteria development, a consistent "dose-response" should be apparent, where each of the taxa would respond negatively (e.g., decreased "probability of occurrence") to increased conductivity. A graphical representation of this type of response (i.e., the stressor-response profile) would resemble the "decreasing" conductivity responses for *Ephemerella* in Figure 3-4 (EPA 2016; Appendices A and B), where the y-axis shows the response (i.e., probability of observing) and the x-axis shows the concentration of the stressor that is presumably inducing that response (i.e., conductivity). In this case, to be protective, it appears that conductivity should always be *below* roughly 300-400 µS/cm. Note that roughly 43% of the taxa in EPA's original analysis of data from Ecoregions 69 and 70 follow this pattern (Figure 3-3).

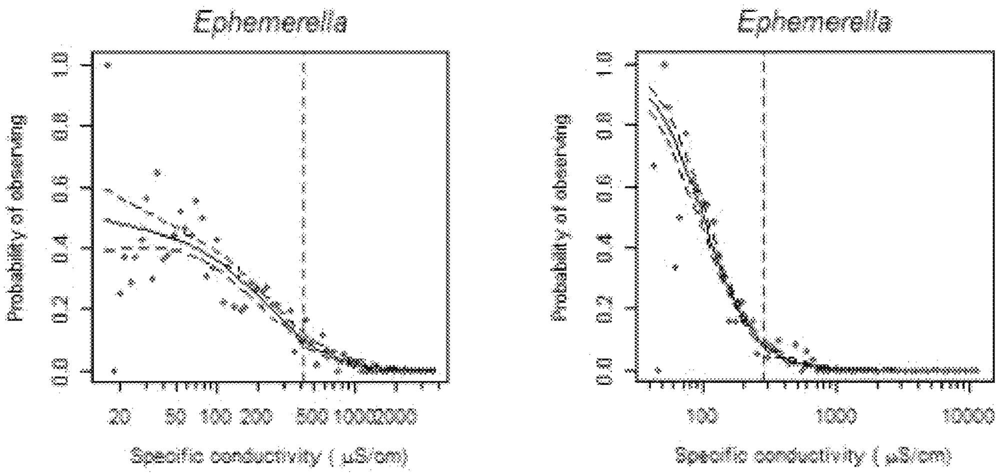


Figure 3-4: *Ephemerella* probability of observing versus conductivity (Appendices A and B from EPA 2016). Ecoregion 69 - left panel, ecoregion 70 – right panel.

Again, the approach to building an SSD used by EPA would *only be valid* if all of the organisms incorporated into the SSD respond similar to *Ephemerella*, since it assumes a protective level set at the lower end of the distribution (i.e., where organisms are more sensitive) will also protect all of the species at the upper end of the distribution (i.e., where organisms are less sensitive). *This is simply not the case with the field-based data used in the EPA Draft Conductivity Criteria.*

In fact, approximately 20% of the genera in the original EPA analysis from Ecoregions 69 and 70 exhibit "positive" stressor-response profiles (Figure 3-3), as exhibited by *Hemerodromia* in Figure 3-5 (EPA 2016; Appendices A and B) – a direct contradiction to the "decreasing" stressor-response profile required to be used in an SSD. And if conductivity is the reason for this response profile and we wish to protect this organism from extirpation, under this circumstance, conductivity would always need to be *above* roughly 200 µS/cm.

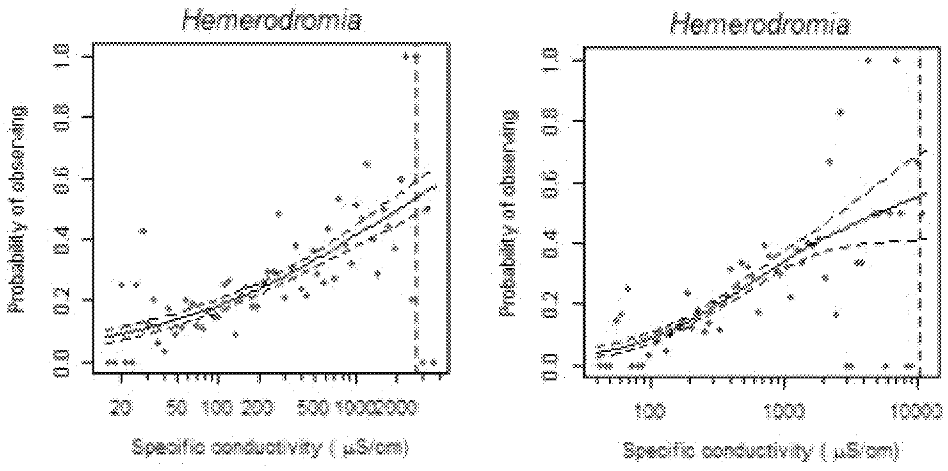


Figure 3-5: ***Hemerodromia*** **probability of observing versus conductivity (Appendices A and B from EPA 2016). Ecoregion 69 - left panel, ecoregion 70 – right panel.**

Roughly 22% of the taxa exhibit a *negative response* at both the low and high end of the range of conductivity levels, but a positive response in the middle of that same range (Figure 3-3). This "optimum" bellcurve-type of stressor-response profile would resemble that shown in the *Caenis* panels in Figure 3-6 (EPA 2016; Appendices A and B). Here, in order to accurately calculate a "protective value", it would theoretically be necessary to calculate two XC_{95} values, with one being >100 µS/cm <u>and</u> the other being <1200 µS/cm for Ecoregion 69 and one <2,500 µS/cm for Ecoregion 70; i.e., one threshold at the low and one at the high end of the x-axis where the positive and negative responses are observed.

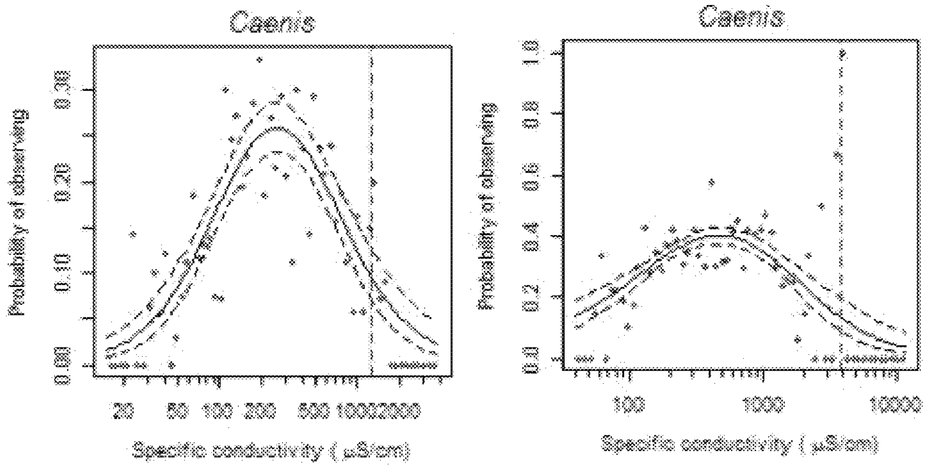


Figure 3-6: ***Caenis*** **probability of observing versus conductivity (Appendices A and B from EPA 2016). Ecoregion 69 - left panel, ecoregion 70 – right panel.**

A few of the taxa exhibit the inverse of that curve, responding *positively* at both the low and high end of the range of conductivity levels, but poorly in the middle of that same range (Figure 3-3). This type of "bimodal" stressor-response profile would resemble that shown in the *Dolophilodes* panel in Figure 3-7 (EPA 2016; Appendix B). Here, it would theoretically be necessary to again calculate two XC_{95} values with one *less than* roughly 100 µS/cm and the other *greater than* roughly 7,000 µS/cm; i.e., two thresholds bracketing the middle of the x-axis to capture the range where negative responses are observed. Of course, this "stressor response profile" may be an artifact that results from the use of non-independent stressor analysis curves generated using data collected using field-based methods, rather than an actual response to conductivity (GEI 2010 and 2012).

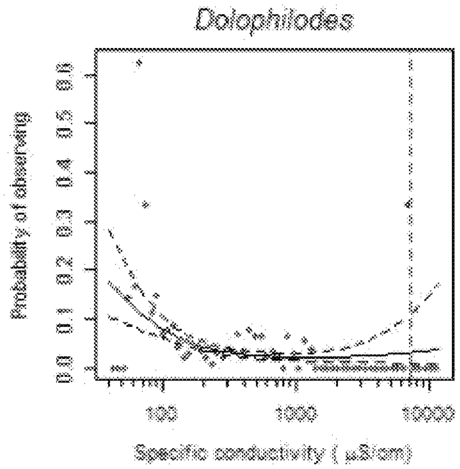


Figure 3-7: *Dolophilodes* probability of observing versus conductivity (Appendix B from EPA 2016).

A number of the taxa (roughly 15%) in the Ecoregions 69 and 70 datasets actually show very little response to conductivity at all, as shown in the "No-response" *Tvetenia* stressor-response profile in Figure 3-8 (EPA 2016; Appendices A and B), making it difficult to identify any kind of effect concentration. In fact, as noted in our analysis of the Ecoregion 50 dataset earlier, many of the "most sensitive" genera showed this non-response profile.

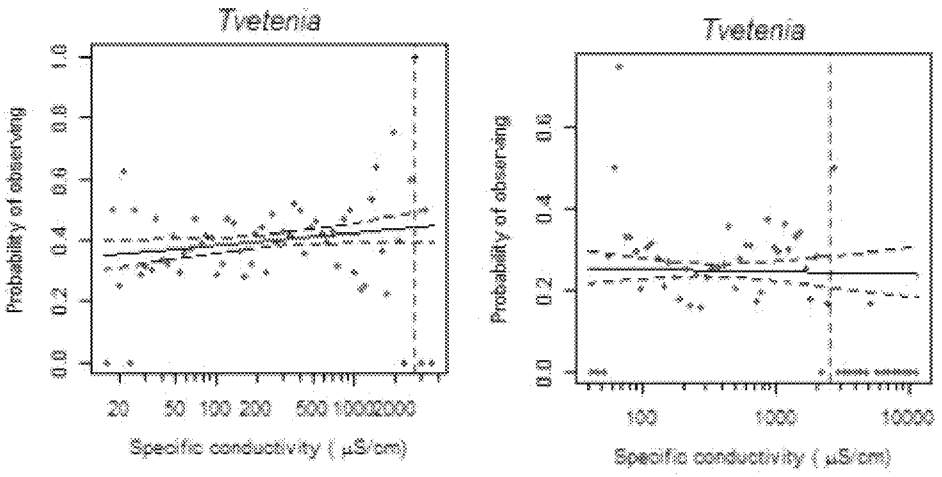


Figure 3-8: ***Tvetenia*** **capture probability versus conductivity (Appendices A and B from EPA 2016). There does not appear to be an Optimal Conductivity Range for this genus. Ecoregion 69 - left panel, ecoregion 70 – right panel.**

Thus, if we accept the premise that physiological limits with respect to conductivity are responsible for the presence/absence of benthic invertebrate taxa (i.e., extirpation), thresholds more appropriately based on the full range of capture probability/stressor-response profiles would be as follows:

Response Profile	Possible Threshold
Negative response (*Ephemerella*):	< 300-400 µs/cm
Positive response (*Hemerodromia*):	> 200 µs/cm
Bell-curve response (*Caenis*):	> 100 and < 1,200-2,500 µs/cm
Bimodal response (*Dolophilodes*):	< 100 and > 7,000 µs/cm
No response (*Tvetenia*):	no threshold necessary

As noted in our original reviews (GEI 2010 and 2012), there is simply no way to reconcile these widely conflicting stressor-responses into a single SC criteria concentration that would actually be "protective" of 95% of the taxa. More importantly, if one accepts that the capture probability data in Appendices A and B (EPA 2016) represent actual stressor-response relationships, the criteria values in case studies I and II of approximately 300 µS/cm could actually be interpreted as *not protective*, for a large percentage of the organisms. This is examined further in the next section.

3.2.2.3 Protectiveness of Criteria/Benchmark Calculated Using the XCD Method

To further evaluate the protectiveness of using the XCD and accompanying SSD method for development of field-based SC criteria, we reexamined the data used in the original EPA Benchmark Document in the context of data screening tools used by EPA and how those tools affect the premise of protection of 95% of the taxa given the conflicting response profiles noted above and the potential need to create multiple HC_{05} values if we truly want to protect 95% of the taxa.

If we accept, for the sake of discussion, the oversimplification that conductivity is the causative agent behind the presence and absence (i.e., "capture probability") of taxa, it becomes apparent that most taxa have not only an upper limit to conductivity (i.e., an extirpation concentration, per EPA) but a lower limit as well. Using the same approach described in the EPA Draft Conductivity Criteria, we can develop a minimum conductivity that would afford protection to the 95% of taxa that would be extirpated by low conductivity. Specifically, we estimated the 95th percentile of the weighted cumulative distribution function (CDF) of XC05 (instead of 5th percentile of weighted CDF of XC_{95}) using the same 163 taxa that were used to develop the EPA Conductivity Benchmark to address both ends of the "sensitivity distribution" if all the various response curves are taken into account. The resulting lower bound on conductivity is 130 µS/cm, suggesting that as long as conductivity is above 130 µS/cm not more than 5% of species will be extirpated by the effect of low conductivity.

Figure 3-9 displays the outcome of this analysis. The values for each taxon are represented on a horizontal gray line, with the median conductivity (i.e., the XC50) for each taxon depicted by red circles, the minimum conductivity of each taxon depicted by blue circles, and the maximum conductivity of each taxon depicted by green circles. The 5th and 95th percentiles of these conductivity ranges are indicated by the 'x' and '+' symbols located to the left and right of the median, respectively. Note that the 95th percentile (+) is equivalent to the EPA's XC_{95}. As shown in Figure 3-9, a range of conductivity from 130 to 300 µS/cm, highlighted in yellow, would be acceptable to 90% of taxa evaluated – equivalent to protecting the 95th percentile from both ends of the distribution. Again, if conductivity is the causative agent for the distribution of these genera, then to meet the goal of 95% protection, we have to protect the full range of responses…including those genera sensitive to higher conductivity as well as those that are sensitive to lower conductivity.

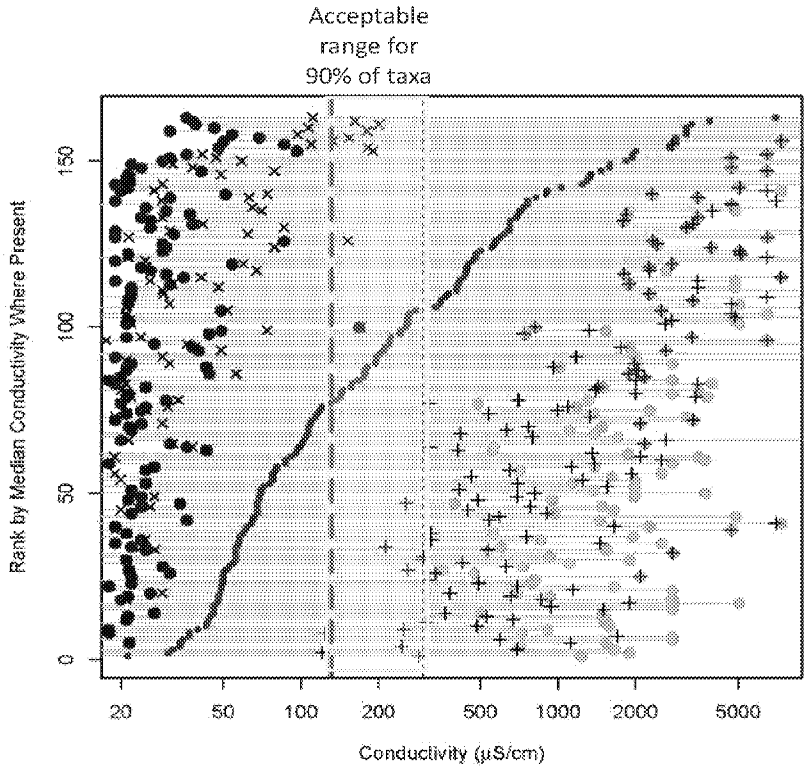


Figure 3-9: Preferred conductivity range for 163 of the taxa included in the 2011 EPA benchmark for Ecoregions 69 and 70. The x-axis is on a logarithmic scale.

Recognizing that the 163 taxa included in EPA's Benchmark analysis represent only a third of the taxa in Ecoregions 69 and 70 in the WVDEP water analysis database (WABbase) database, we ran a similar analysis to that described above using all of the approximately 500 taxa in Ecoregions 69 and 70 in the WABbase when the benchmark was developed. Using all taxa available, we noted that there are many taxa that are never found in locations with conductivity as low as 300 µS/cm.

Regardless, the analysis with the full list of genera indicates that conductivity would need to be *greater than* 1130 µS/cm to protect 95% of the taxa from the effects of low conductivity. Conversely, conductivity would also have to be *less than* 120 µS/cm to protect 95% of taxa from the effects of high conductivity (Figure 3-10 [Note that the symbols and colors presented in Figure 3-10 adhere to the descriptions of those presented previously in Figure 3-9]). This effectively results in the "Impossible Zone," identified in Figure 3, which contrasts with the "Acceptable Range" identified in Figure 2. We called this the "Impossible Zone" because with all 500 taxa included, the minimum conductivity required to protect 95% of taxa from low conductivity is greater than the maximum conductivity required to protect 95% of taxa from high conductivity. Again accepting the oversimplification that conductivity is the determinant of species presence/absence, it becomes readily apparent that conductivity preference varies widely among taxa, and that, when we include all 500 taxa, there is no single conductivity value or range that will protect 95% of taxa – the apparent goal of the EPA draft criterion.

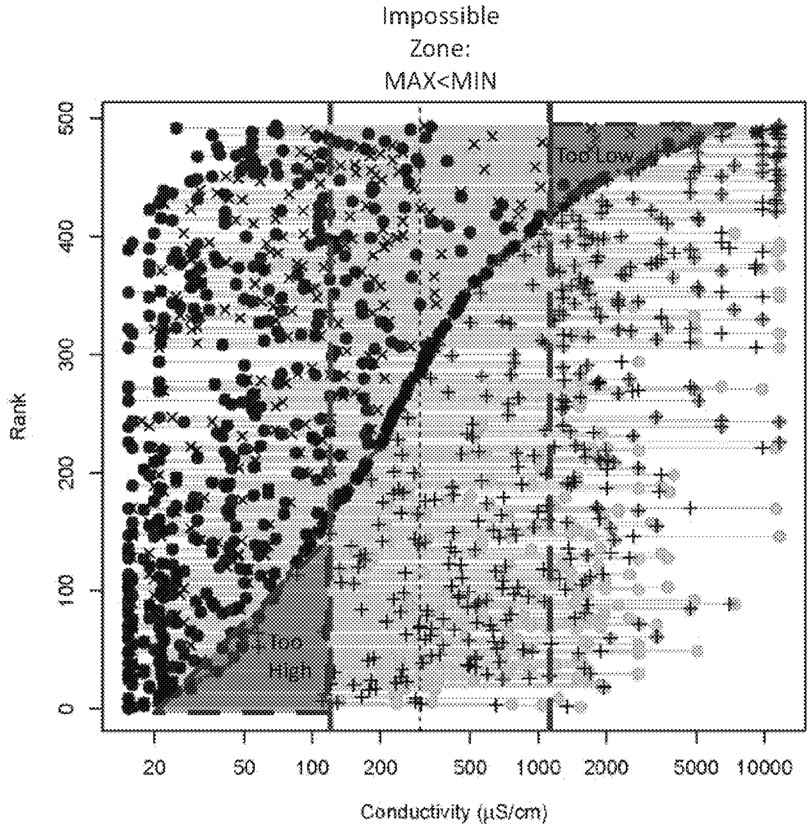


Figure 3-10: Preferred conductivity range for the nearly 500 taxa included in the WABbase used for the EPA benchmark for Ecoregions 69 and 70. The x-axis is on a logarithmic scale.

Conductivity tolerance/preference, as a surrogate summary metric for total ionic content of water, is evidently more akin to temperature tolerance/preference than to that of traditional contaminants such as metals or pesticides. Just as there are taxa (*e.g.*, specific species of fishes) that prefer warm waters and those that prefer cool waters, there are taxa that appear to prefer low conductivity and those that appear to prefer higher conductivity. Of course, like temperature, there is little doubt that extreme conductivity, low or high, has the potential to exceed the physiological limits of all taxa.

Finally, despite the capture probability graphs suggesting an upper and lower bound on the range of conductivity preference for many taxa, we do not agree that conductivity is necessarily the direct causative agent for the presence and absence of taxa in any particular location. Ionic tolerance may play one role in establishing which species inhabit a site, but other chemical, physical, and biological factors discussed in the preceding section likely have equally important roles.

3.2.3 *Use of the Background-to-Criterion Regression Method*

Section 3.7 and Appendix D in the EPA Draft Conductivity Criterion document discuss the development of the Background-to-Criterion Regression Model (B-C Model) that is intended to be applied to areas where there is insufficient water chemistry and biological data to calculate ecoregion-specific taxa level extirpation coefficients and hazardous concentrations. However, as is clear from our analyses above, the premise that salt-intolerant genera only occur within habitats where the lowest background conductivity occur and that genera are extirpated in habitats with higher conductivity is not well supported by the variable XC_{95} values observed for genera common among ecoregions. As such, development of the regression analysis using such variable XC_{95} values is fundamentally flawed or is, at best, an untested hypothesis.

3.2.3.1 Discussion of top 10 most sensitive taxa

The EPA Draft Conductivity Criteria further highlights the key issue of variable stressor-response profiles when expanding the approach to multiple ecoregions. As discussed herein there are multiple profiles that characterize the probability of capturing a single individual representative of a taxonomic element based on the range of SC observed within an ecoregion. The methodology imposes the extirpation coefficient regardless of the taxonomic element's (genus level) preference/response to the environmental condition. Furthermore, EPA assumes that the taxonomic list does not need to exhibit the same responses to conductivity for the calculation of the HC_{05}, and this is evident in the top ten list of taxa for selected ecoregions.

Seven ecoregions were selected from the list of 24 ecoregions used in the EPA's development of the Background to Criteria Regression model, and included Arizona/ New Mexico – Ecoregion 23, Idaho – Ecoregion 17, Illinois – Ecoregions 54 & 72, Minnesota – Ecoregion 50, Mississippi – Ecoregion 65, and West Virginia – Ecoregion 69 (note that the values for genera from Ecoregion 69 in the EPA Draft Conductivity Criteria differ from those in the original benchmark document as additional data were added; EPA 2011). Ecoregions were selected to represent different ecotypes from the West, Midwest, and Eastern portion of the U.S. Benthic invertebrate data were paired with measurements of conductivity for the respective sample location and the taxonomic list was filtered for taxa that were not present in at least 25 samples for an ecoregion per EPA's approach. For each ecoregion, the XC_{95} values were calculated using the methodology described in Section 3.1.2 of EPA 2016 while capture probabilities were based on the observed range of conductivity with 60 bins created for each ecoregion. The derivation of the HC_{05} resulted in a value for each ecoregion that was contained in the range of XC_{95}s for top ten list of taxa (Table 3-1) and corroborated results presented in Table 2-D of the EPA Draft Conductivity Criteria.

Using these data, the top ten "most sensitive" taxa (i.e., genera with the lowest resulting XC_{95} values) from each ecoregion were identified and capture probability/response-curves developed. Each top ten list of the most sensitive taxa for each ecoregion contains a mix of decreasing or no response/bimodal profiles with varying degrees of occurrence as well as a mixture of the other profiles identified earlier (Appendix A – Figure A-1). For example, in Idaho Ecoregion 17, *Stempellina* exhibits the smallest XC_{95} value of 175 µS/cm yet exhibits no apparent response to conductivity. In Arizona Ecoregion 23, taxa such as *Amphinemura*, *Lepidostoma*, and *Zaitzevia* all exhibit decreasing responses to the range of conductivity values, but these taxa also present conflicting information in their respective stressor-response profiles. All three taxa reveal patterns in capture probabilities that are either zero (0.00) or one (1.00) on the lower range of conductivity, which stems from the left skewed distribution for conductivity (Appendix A– Figure A-1). The left skewed distribution results in a number of conductivity bins that contain a small number of sites, thus presence of merely one individual is sufficient to indicate a 100% chance of finding the taxon given the imposed distribution on the range of conductivity values observed in an ecoregion. The relatively high number of zero probability findings could be an artifact of a skewed distribution or simply the taxon was not collected even though it may have been present.

Other patterns in the response-profiles indicate that taxa can be relatively rare across the range of environmental conditions. For example, in Minnesota Ecoregion 50, *Larsia*, *Paraponyx*, and *Campeloma* all exhibit poor capture probabilities, generally less than 20%, through the conductivity range of approximately 50 to 300 µS/cm and do not show a clear response to conductivity (Appendix A– Figure A-1). Furthermore, these taxa were observed in a small number of samples just above the ≧ 25 sample cut-off ranging from 25 to 47 samples out of the 734 samples comprising the Ecoregion 50 dataset. The absence of a well-defined response to conductivity raises concern regarding the appropriateness of calculating "extirpation coefficients" when the rarity of a taxon is not fully vetted with respect to habitat preferences, substrate conditions, or flow characteristics, much less a lack of significant response to conductivity. Similarly poor capture probabilities were observed for Illinois Ecoregion 54 where the vast majority of taxa exhibited no definitive response to conductivity.

Next, the same "top ten" taxa from these ecoregions were summarized alphabetically by genus with their sensitivity rank and XC_{95} value (Table 3-1). Note that since there was little overlap in "top ten" genera across ecoregions, we also included the corresponding sensitivity rank for the genus in the other ecoregions where it was found. This top ten list further highlights the contradictory nature of assumed physiological sensitivities to conductivity across ecoregions – contradictory, of course, only if we accept the premise that the responses to conductivity is a measure of physiological tolerance. While the species within a genus may differ across ecoregions; the premise behind a genus level response to a toxicant is that all individuals should exhibit similar sensitivity level (Stephan et al. 1985). And, indeed, there are similarities for certain genera across ecoregions. For example, the XC_{95}s for *Epeorus* range from only 200 to 499 µS/cm for four of the seven selected ecoregions (Table 3-1). Yet, for the majority of taxa, there are vastly different XC_{95}s across ecoregions. For example, the XC_{95}s for *Isonychia* ranged from 149 to 1,830 µS/cm, while *Rhyacophilia* ranges 10-fold across ecoregions, from 201 to 2,054 µS/cm. These differences in the genus' extirpation coefficients across multiple ecoregions provide further evidence, in addition to the earlier examples, that puts into doubt that conductivity is the causative factor in the presence/absence of these genera.

Table 3-1: Listing of the top ten genera for seven selected ecoregions along with the corresponding rank for the same taxon when found in the other ecoregions (see text for explanation). Taxa rank and XC_{95} value are provided (Rank / XC_{95}).

Genera	AZ, 23	ID, 17	IL, 54	IL, 72	MN, 50	MS, 65	WV, 69
Agapetus	3 / 200	18 / 424					
Agnetina					9 / 317		
Alloperla							5 / 261
Ameletus	11 / 297	83 / 962					7 / 295
Amphinemura	5 / 252					98 / 563	36 / 722
Anthopotamus			4 / 808				
Baetisca					174 / 1,998	10 / 149	41 / 918
Basiaeschna				7 / ,1160	114 / 1,412		
Bezzia					134 / 1,527		6 / 294
Campeloma					10 / 318		
Caudatella		3 / 233					
Centroptilum			10 / 849	4 / 1,070	171 / 1,998		50 / 1,163
Cinygmula	4 / 207	79 / 907					16 / 366
Conchapelopia			64 / 1,600	1 / 918	100 / 1,353	74 / 442	9 / 323
Dolophilodes		9 / 333			1 / 198		103 / 2,768
Doroneuria		4 / 233					
Dromogomphus			2 / 755	51 / 1,830		43 / 325	
Drunella	1 / 168	63 / 780					19 / 423
Endochironomus			65 / 1,600	3 / 1,040	55 / 867		
Epeorus	18 / 350	25 / 499			2 / 200		8 / 302
Ephemera			1 / 723		46 / 719		15 / 360
Ephemerella	23 / 420	41 / 687			45 / 719	3 / 111	17 / 388
Ephoron			6 / 814				
Erpobdella			88 / 1,960	2 / 943			
Eurylophella					14 / 397	9 / 149	29 / 574
Glossosoma	9 / 259	30 / 600			39 / 650		61 / 1,725
Glutops		5 / 254					
Heptagenia			28 / 1,071	14 / 1,260	53 / 867		10 / 343
Hesperophylax	8 / 259						
Hexatoma	66 / 820	42 / 687			26 / 533	2 / 92	105 / 2,768
Isonychia	54 / 800		13 / 908	49 / 1,830	54 / 867	8 / 149	54 / 1,270
Larsia					7 / 317		
Lepidostoma	6 / 258	98 / 1,345			96 / 1,353		11 / 357
Leptophlebia					16 / 422		1 / 191

Genera	AZ, 23	ID, 17	IL, 54	IL, 72	MN, 50	MS, 65	WV, 69
Macromia			3 / 770	90 / 2,551	67 / 882	13 / 156	
Megarcys		7 / 280					
Menetus			46 / 1,283	8 / 1,190			
Mooreobdella			87 / 1,960	5 / 1,110			
Musculium				6 / 1,110			
Narpus	2 / 191	48 / 725					
Neophylax		8 / 283			78 / 1,134		18 / 388
Oxyethira					119 / 1,447	1 / 79	
Paracloeodes			5 / 810	59 / 2,020			
Paralauterborniella			43 / 1,264	114 / 2,620		6 / 135	
Paraleptophlebia	17 / 340	38 / 653			154 / 1,998		4 / 256
Paraperla		6 / 265					
Paraponyx					8 / 317		
Procloeon			9 / 832	61 / 2,020	31 / 568		28 / 556
Progomphus			8 / 822	68 / 2,160		25 / 221	
Prosimulium		67 / 907				7 / 135	31 / 606
Pseudocloeon			21 / 935	10 / 1,205	170 / 1,998		
Pseudolimnophila						4 / 114	33 / 660
Pycnopsyche					120 / 1,447	5 / 123	3 / 232
Remenus							2 / 223
Rhithrogena		10 / 367					
Rhyacophila		62 / 780			3 / 201	108 / 607	76 / 2,054
Serratella	36 / 606	23 / 482			4 / 251		26 / 500
Somatochlora				9 / 1,194	141 / 1,594		
Stempellina		1 / 174			81 / 1,134		27 / 529
Sweltsa	10 / 273	33 / 635					64 / 1,756
Triaenodes			7 / 821		22 / 502		
Trissopelopia					6 / 295		
Visoka		2 / 203					
Xenochironomus					5 / 293		
Zaitzevia	7 / 259	95 / 1,345					

The B-C Model characterizes the relationships between HC_{05} values and the "Background" conductivity conditions for each ecoregion as based on the 25th percentile of measured concentrations. While the least squares regression model does not require predictor variables to meet the assumptions of normality, the use of skewed data in the calculation of XC_{95} values and characterizing genera conductivity response profiles is a concern. Benthic invertebrate and water quality data often are collected in a clumped and uneven way, thus it is common to have data with high number of observations aggregated at a particular part of a gradient, while another part of the gradient is relatively under-represented. This results in skewed distributions which is apparent in the conductivity data (Appendix A – Figure A-2). The weighted cumulative distribution function is used to account for this discrepancy but the skewed data may have unintended consequences on the calculation of the XC_{95}. Genera observed to occur in the tails of the distribution, whether left or right skewed typically express greater capture probabilities due to the small number of sites (i.e., <5) observed. Thus, the response profiles provide the allure that a genus may always be present at low conductivity or extirpated at higher conductivities when it is possible that other factors are affecting the presence/absence of a genus.

3.3 Issues with Deriving the CMEC

Most aquatic life criteria consist of two criterion magnitudes, or concentrations: the Criterion Maximum Concentration (CMC, or "acute criterion") and the Criterion Continuous Concentration (CCC, or "chronic criterion"). For the EPA Draft Conductivity Criteria, EPA correctly points out that the data do not generally exist to directly evaluate what SC level might protect aquatic life from acutely toxic (i.e., short-term) exposures. Instead, EPA suggests derivation of a Criterion Maximum Exposure Concentration (CMEC) which is the 90th percentile of SC observations at site with water chemistry conditions such that they meet the CCC. The CCC used to determine whether or not SC criteria are met at a site are defined as the annual geometric mean SC values from sampling stations that meet the CCC for that region.

While this 90th percentile-based CMEC may indeed represent a less stringent SC criterion than the CCC, this is a largely arbitrary percentile selection that has little basis on a SC concentration that would protect from acute exposures to sensitive taxa at relevant times of year. Nor does it account for the variable stressor-response profiles, so would not be protective of the overall invertebrate community. While EPA does evaluate the limited data that do exist to evaluate maximum SC concentrations that precede observations of salt-intolerant taxa, they admit that even for relatively large data sets (e.g., ecoregions 69 and 70), only modest amounts of such data exist. Therefore, we suggest that until such data can be made available, EPA should not propose derivation of CMECs for SC.

3.4 Discussion of Causation

We agree with EPA that an evaluation of causality is critically important in any scientific endeavor, particularly when proposing the use of a field-based method using macroinvertebrate community structure data that is subject to a potentially high degree of confounding. Any chemical or biological variables that are correlated with either conductivity or the biotic response may confound the presumed relationship between conductivity and biological impairment. In its original Conductivity Benchmark EPA acknowledges that plausible confounding factors likely exist, and recommends use of the method described in the EPA Benchmark Report to analyze confounding factors. However, in the EPA Benchmark Report they concluded that the influence of confounding factors is not *strong enough* to prevent use of the conductivity benchmark (EPA 2011). As we have previously commented (GEI 2010 and 2012), we do not agree with this conclusion; the confounding factors require a more in-depth analysis to evaluate whether or not conductivity *alone* is a strong enough indicator of adverse changes to allow for its use in derivation of a regulatory criteria.

For the EPA Draft Conductivity Criteria, however, no new causality analysis has been conducted. Rather, EPA only choses to suggest that "it is good practice" to further evaluate the performance of the XCD model. We agree, but are very concerned that EPA has not chosen to conduct this analysis. Given the observations we present in Section 3.2.3.1 of this report, the high amount of variability in XC_{95} values for the same taxa between different ecoregions suggest that alternate explanations for the presence vs. absence of several taxa critically need to be evaluated. This concern also extends to use of the background-correction method because it still relies on XCD-based CCC values from ecoregions from which a causal analysis has not yet been conducted.

4. Discussion of Using Fish as an Alternative Assessment Endpoint

The previous sections highlighted recommendations for revisions to the document and criteria. In this section we present a discussion of Appendix G: Using an Alternative Assessment Endpoint (Species of Fish) and areas which need reconsideration. While we may not have specific recommendations on how to address the issues with this section, it should be reexamined to ensure the science supporting the decisions made to develop the conductivity criteria are solid.

Appendix G discusses the extent to which EPA feels the ecoregional criteria for benthic macroinvertebrates are protective of fish. EPA utilized a combined (composite) data set for stream fish from Ecoregions 67-70 in their assessment. A composite data set had to be used due to lesser amounts of fish data in comparison to invertebrates.

GEI notes that the fish data were aggregated across multiple sampling programs, states, and ecoregions. While EPA made clear their rationale for this, the approach is inherently problematic due to variability in fish survey/sampling techniques, likely wide variability in stream size, and assessment approaches utilized by different states and regulatory programs. Additional concerns related to EPA's Appendix G approach are summarized below.

Fish community sampling conducted by state and federal agencies is typically not done to characterize *every* single species in a population. Rather, the goal of most fish sampling efforts, particularly "rapid bioassessment" type protocols that were utilized for most of the fish community data EPA relied on in its assessment, is to obtain fish samples that provide a snapshot of the overall fish community (Barbour et al. 1999). This hopefully captures the majority of the species present at the time of sample. In other words, these methods are not developed for assessing which species might specifically be considered "present" or "absent" (i.e., extirpated) in a given water body. For example, Pennsylvania Department of Environmental Protection's 2013 protocol states that "The objective is to acquire a *representative* (emphasis added) sample of the fish population in a wadeable stream or river by sampling all physical stream habitats in relative proportion to their availability. The collected sample will contain most of the species in the stream at the time of sampling in numbers proportional to their actual abundances." (Botts 2013). This approach will bias against fish species with low abundance or attributes that make them more difficult to capture through standard fish sampling techniques (e.g., cryptic coloration, small size, highly-specific habitat requirements, etc.). In addition, electrofishing is regarded as standard equipment for fish community surveys in most wadeable and headwater streams. However, standard fish community survey approaches are not designed to ensure capture of 100% of species in a surveyed area. Cryptic species and smaller fish are more likely to not be detected via electrofishing (Zale et al. 2013).

While Appendix G notes that the observed, and presumed, tolerance to conductivity may be due to the probability of capturing and enumerating fish, additional analysis is needed. This describes one of the main issues with using this method on fish. Fish are highly mobile, with some species exhibiting long-range (e.g., dozens to hundreds of miles) movement patterns that driven by many factors including, but not limited to reproductive cycles, food availability, thermal and flow regimes, etc. The fish survey data are generally "snapshots" of a fish community on a given day not an in depth assessment of the kinds of data that can reliably determine "extirpation," particularly on the basis of presumed physiological tolerance to SC.

EPA's evaluation/characterization of reference data for fish was not consistent with the approach taken for benthic macroinvertebrates. Specifically, the fish data set "contained an uncertain number of reference sites; but there are at least 134 sites with >90% forest cover which are more likely to be representative of good to high quality stream systems than those with less forest cover." Such a "qualitative characterization" of reference sites seems inadequate, particularly when it is inconsistent with the approach utilized for benthic macroinvertebrates.

The problem formulation section earlier in the EPA Conductivity Criteria suggests that indirect effects to fish, such as prey (benthic invertebrate) abundance, are possible. Yet, Appendix G does not provide any further examination on this issue, making this a critical but untested statement by EPA. This is particularly of concern because the XCD method inherent to the SC criteria does not take organism abundance into account.

Appendix G also states that "*the purpose of this assessment is to determine the sensitivity of fish relative to macroinvertebrates…* " (page G-5). It concludes to have demonstrated that fish are either directly or indirectly affected by increased SC based on results that show "*XC_{95} values for fish fall within the range of XC_{95} values calculated for benthic macroinvertebrates.*" (page G-33). This statement is provided without any of the accompanying effort to address correlation or causation or confounding or any level of uncertainty in the analysis of this broad fish analysis. Without undertaking any such effort, there is no way to determine whether XC_{95} similarities are functionally significant, as EPA surmises, or instead whether this is purely a coincidence.

GEI's also evaluated EPA's Appendix G fish data file (from the docket) titled "Combined Less". Through this review GEI observed multiple data-related issues that are worrisome:

- Most fish community survey sites were sampled only once. This is of concern given that fish are highly mobile, with distributions that can vary seasonally and with their reproductive cycles, and as such, extirpation may not be conclusively determined from only one survey.

- Table G-1 in the "Combined Less" data file points out that the majority of water quality parameters had sample sizes substantially less than the 3,277 values for SC. In other words, matched water quality and fish community parameters were available for far less than 3,277 sites for which SC values were available. Specific examples include:
  - Habitat – 801 values (24%)
  - Alkalinity – 995 values (30%)
  - Bicarbonate, sulfate, calcium – 1,014 values (30%)
  - Hardness – 1,488 values (45%)

- Pennsylvania Fish and Boat Commissions's (PaFBC's) data set provided a significant amount of the fish community and water quality data in Appendix G with 2,101 sites being surveyed, as compared to all other data sets. Of the 42,336 records in EPA's fish data file, 17,100 (40%) were from PaFBC. Appendix G does not appear to have evaluated the extent to which the PaFBC data may be driving the Appendix G SC evaluation for fish.

- PaFBC's data set also seems to lack presumably important data. Specifically, the PaFBC data set lacks:
  - Fish abundance numbers
  - Habitat data (i.e., there are no qualitative habitat assessment scores in the PaFBC data set)
  - Fish size data is also missing, which presumably could be used to evaluate for overall health and status, i.e., multiple year classes present, etc.

- 1,326 PaFBC records show hatchery fish (brook, brown, rainbow trout) from hundreds of sites. The presence of hatchery trout can alter fish community dynamics. As an example, Vincent (1987) reported 49% decrease in wild brown trout number and biomass in a Montana stream following three years of rainbow trout stocking. Fish stocking with legal (or near legal) sized fish also increases the abundance of larger-sized predatory fish in a system, which could reduce abundance (and presence) of prey species. Thus, abundance of fish species that are already in low abundance for any factor could be decreased further, or result in extirpation, following stocking.

- A brief evaluation of PaFBC sites sampled multiple times showed potentially significant variability in the fish assemblage despite relatively consistent conductivity numbers. For example:

- PaFBC Site ID PaFBC0043 (Monogahela River watershed) was visited twice (May and September) in 2004. Location verified via GPS coordinates.
  - May visit: SC = 125 µS/cm, total fish taxa = 8
  - Sept visit: SC = 111 µS/cm, total fish taxa = 24
- PaFBC Site ID PaFBC1339 (Allegheny River watershed) was visited twice: July 1991 and July 2008. Location verified via GPS coordinates.
  - July 1991 visit: SC = 780 µS/cm, total fish taxa = 9
  - July 2008 visit: SC = 1,033 µS/cm, total fish taxa = 15 (despite higher conductivity).

To summarize, the examples above clearly display significant variability in the fish community despite relatively consistent or even increasing SC values. This underscores the importance of multiple fish community site visits, which are not available for the majority of fish community data that are included in the "Combined Less" Appendix G fish community data file.

Additionally, EPA conducted no ground-truthing of the data set. This could have been done by simply identifying sites where there are paired fish, invertebrate, and SC data and comparing the taxa assemblages against the XC_{95}s to generate "predicted taxa assemblage" vs. "observed taxa assemblage". This evaluation could likely have been done using the GPA coordinates in the data files.

5. Conclusions

Our key concerns and comments with respect to the EPA Draft Conductivity Criteria are summarized below:

- The CCC derived using the methods described in the EPA Draft Conductivity Criteria does not represent a SC concentration that would protect 95% of the species with respect to SC. This is a fundamentally inappropriate interpretation of the 1985 Guidelines and its goal to derive aquatic life protection criteria that are protective of all but 5% of the most sensitive species to a pollutant or stressor.

- We are concerned that the XCD method effectively sets conductivity criteria concentrations very close to natural background concentrations, which is not consistent with the broad ecological integrity goals of the Clean Water Act. No consideration to diversity or abundance or other straightforward methods of evaluating aquatic community structure, even though diversity is a well-established means of evaluating benthic invertebrate community health and structure.

- There are substantial differences in XC_{95} values within a genus across ecoregions; with genera characterized as being sensitive to conductivity in one ecoregion not sensitive to conductivity in another ecoregion. Such large variability in the purported physiological limits of "sensitive" genera raises considerable uncertainty regarding the applicability of conductivity as a determinant of the frequency of occurrence (or absence/extirpation) of a given taxa.

- If EPA wishes to use stressor-response data to build an SSD for criteria development, a consistent stressor-response is required, where all genera respond negatively (e.g., decreased "probability of occurrence") to increased SC. Yet, there are widely conflicting stressor-responses across genera used by EPA and there is simply no way to reconcile these into a single SC criteria concentration that would be protective of 95% of the invertebrate community.

- We do not agree that conductivity is necessarily the direct causative agent for the presence and absence of taxa in any particular location. Just as there are taxa (e.g., specific species of fishes) that prefer warm waters and those that prefer cool waters, there are taxa that appear to prefer low conductivity and those that appear to prefer higher conductivity.

- When the capture probability/response-curves for the top ten "most sensitive" taxa from each ecoregion are compared, there is little overlap in "top ten" genera across ecoregions. Additionally, individual general often exhibit significantly different XC_{95} values across ecoregions. These differences in the genus' extirpation coefficients across multiple ecoregions provide further evidence that conductivity is not the causative factor in the presence/absence of these genera.

- EPA suggests derivation of a Criterion Maximum Exposure Concentration (CMEC) set at the 90th percentile of SC observations at sites with water chemistry conditions such that they meet the CCC. While this 90th percentile-based CMEC may indeed represent a less stringent SC criterion than the CCC, this is an arbitrary percentile selection that has little direct relationship to a SC concentration that would protect from acute exposures to sensitive taxa at relevant times of year. In addition, as with the CCC, it does not take into consideration the variable stressor-response profiles, so would not be protective of 95% of the invertebrate community. Therefore, EPA should not propose a method for derivation of CMECs for SC.

- EPA states that an evaluation of causality is critically important in any scientific endeavor, particularly when proposing the use of a field-based method using macroinvertebrate community structure data that is subject to a potentially high degree of confounding. We are very concerned that EPA has not chosen to conduct such an analysis here.

6. References

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 Invertebrates, and Fish, 2nd Edition. EPA 841-B-99-002, U.S. Environmental Protection Agency, Washington, DC.

Bukantis, R. 1998. *Rapid Bioassessment Macroinvertebrate Protocols: Sampling and Sample Analysis SOPs.* Working Draft. Montana Department of Environmental Quality; Planning, Prevention, and Assistance Division, Helena, MT.

Botts, W. 2013. *Wadeable semi-quantitative fish sampling protocol for streams*. Bureau of Clean Water. Pennsylvania Department of Environmental Protection. Harrisburg, PA.

Cormier, S.M. 2016. *An Evaluation of a Field-based Aquatic Life Benchmark for Specific Conductance in Northeast Minnesota*. National Center for Environmental Assessment – Cincinnati, Ohio. Office of Research and Development, U.S. EPA. February 2016.

Cummins KW, Richard M, and Andrage PCN. 2005. The use of invertebrate functional groups to characterize ecosystem attributes in selected streams and rivers in south Brazil. *Studies on Neotropical Fauna and Environment* 40(1): 69-89.

GEI Consultants, Inc. (GEI). 2010. *Technical Review: A Field-based Aquatic Life Benchmark for Conductivity in Central Appalachian Streams, DRAFT EPA report.* Report for National Mining Association, Washington, DC.

GEI Consultants, Inc. (GEI). 2012. *Technical Review: A Field-based Aquatic Life Benchmark for Conductivity in Central Appalachian Streams-2011.* Report for National Mining Association, Washington, DC.

Jessup, B. 2010. Recalibration of the Macroinvertebrate Multi-Metric Index for Colorado. Prepared for Colorado Department of Public Health and Environment, Water Quality Control Division, Monitoring Unit. Denver, CO, and EPA Region 8. Prepared by Tetra Tech, Inc.

Jessup, B. and J. Gerritsen. 2002. Stream macroinvertebrate index. Pp. 29-74 *in* Grafe, C.S. (ed.). Idaho Small Stream Ecological Assessment Framework: An Integrated Approach. Idaho Department of Environmental Quality, Boise, ID.

Johnson, B.L. and M.K. Johnson. 2015. Review of *An Evaluation of a Field-based Aquatic Life Benchmark for Specific Conductance in Northeast Minnesota*. Prepared for Water Legacy. November 2015.

Omernik, JM. (1987) Ecoregions of the conterminous United States. Ann Assoc Am Geograph 77:118–125.

Omernik, JM. (1995) Ecoregions: a spatial framework for environmental management. In: Davis, WS; Simon, TP; (eds). Biological assessment and criteria: tools for water resource planning and decision making. [Pp 49–62]. Boca Raton, FL: Lewis Publishers.

Posthuma, L., G. W. Suter II, and T. P. Traas (eds.). 2002. *Species Sensitivity Distributions in Ecotoxicology*. Lewis Publishers, Boca Raton, FL.

Roark, S. A., C. F. Wolf, G. D. DeJong, R. W. Gensemer, and S. P. Canton. 2013. Influences of Subsampling and Modeling Assumptions on the US Environmental Protection Agency Field-Based Benchmark for Conductivity. *Integrated Environmental Assessment and Management* 9:533-534.

Royer, T.V. and C.A. Mebane. 2002. River macroinvertebrate index. Pp. 3-1-3-21 *in* Grafe, C.S. (ed.). Idaho River Ecological Assessment Framework: An Integrated Approach. Idaho Department of Environmental Quality, Boise, ID.

Stephan, C. E., D. I. Mount, D. J. Hansen, J. H. Gentile, G. A. Chapman, and W. A. Brungs. 1985. *Guidelines for Deriving Numerical National Water Quality Criteria for the Protection of Aquatic Organisms and Their Uses*. PB-85-227049. U.S. Environmental Protection Agency, Office of Research and Development, Duluth, MN.

U.S. Environmental Protection Agency (EPA). 1991. *Technical Support Document for Water Quality-based Toxics Control*. EPA/505/2-90-001, U.S. Environmental Protection Agency, Washington, DC.

U.S. Environmental Protection Agency (EPA). 1992. *Framework for Ecological Risk Assessment*. EPA/630/R-92/001, U.S. Environmental Protection Agency, Washington, DC.

U.S. Environmental Protection Agency (EPA). 2000. *A Stream Condition Index for West Virginia Wadeable Streams*. Prepared by Tetra Tech, Inc., for U.S. EPA Region 3.

U.S. Environmental Protection Agency (EPA). 2011. *A Field-Based Aquatic Life Benchmark for Conductivity in Central Appalachian Streams*. EPA/600/R-10/023F, U.S. Environmental Protection Agency, Washington, DC.

U.S. Environmental Protection Agency (EPA). 2016. Public Review Draft: *Field-Based Methods for Developing Aquatic Life Criteria for Specific Conductivity*. EPA-822-R-07-010, U.S. Environmental Protection Agency, Washington, DC.

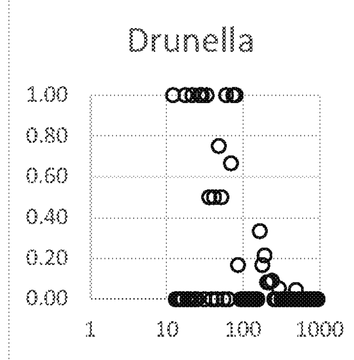
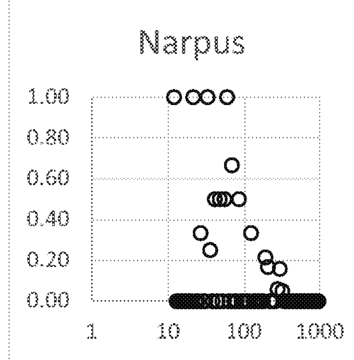
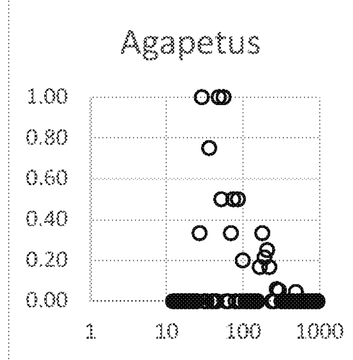
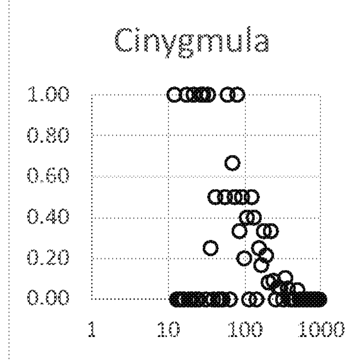
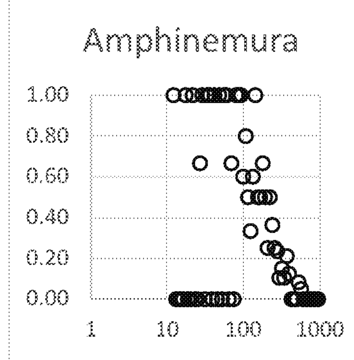
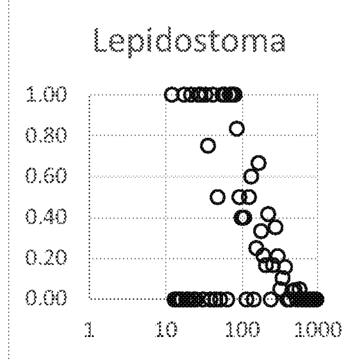
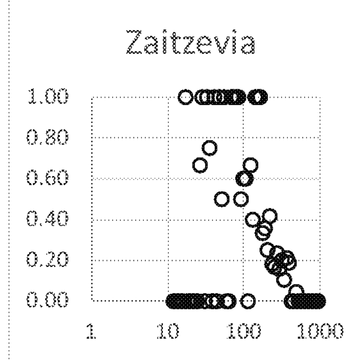
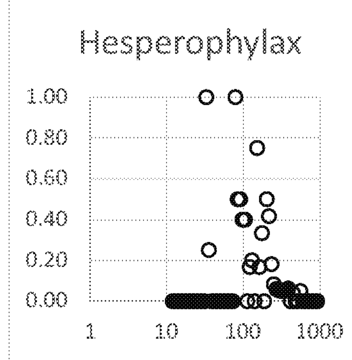
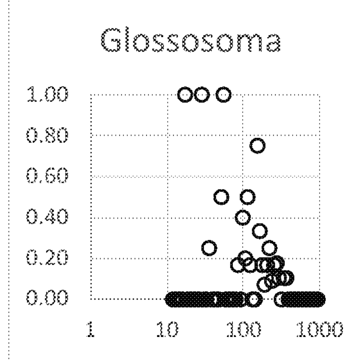
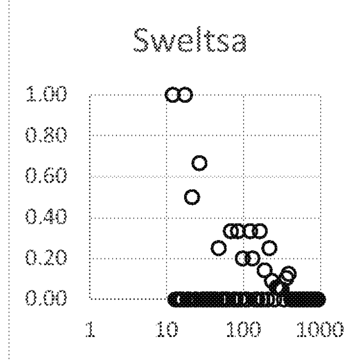
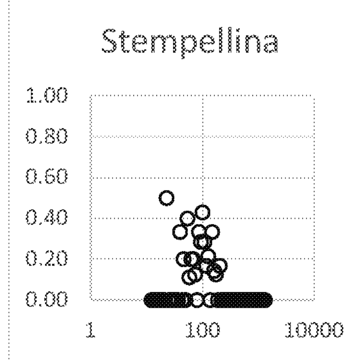
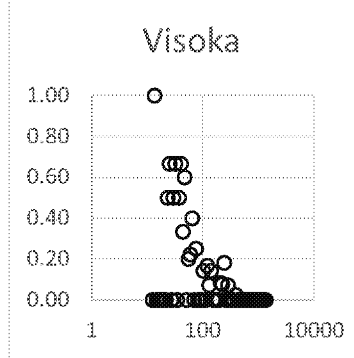
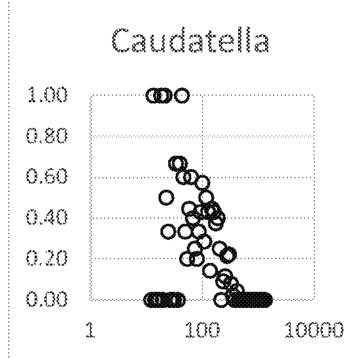
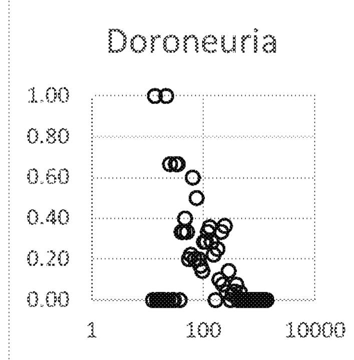
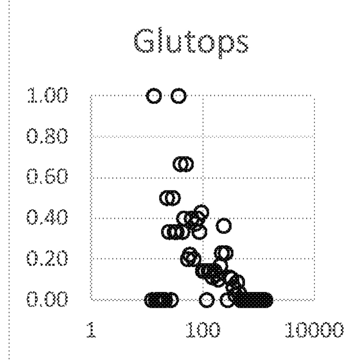
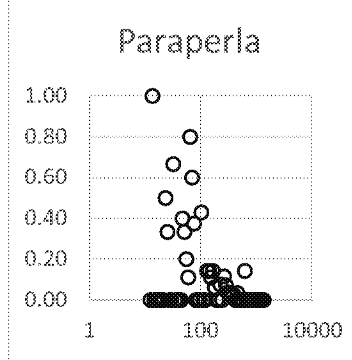
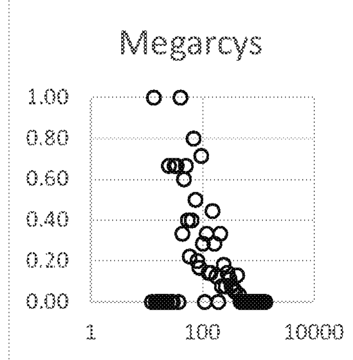
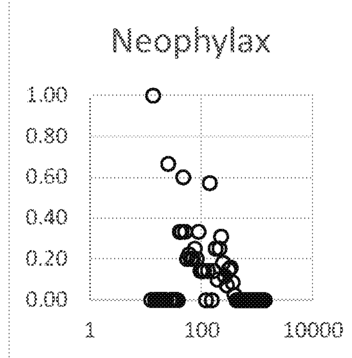
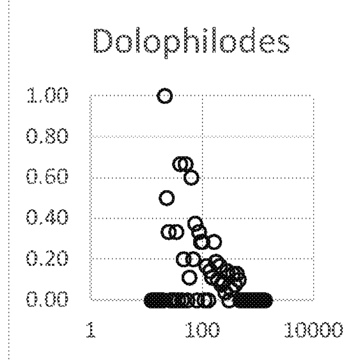
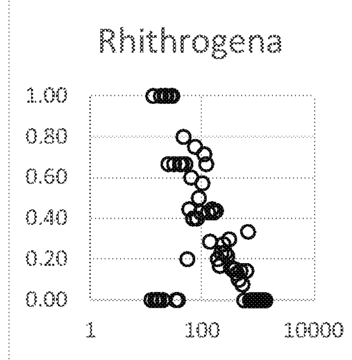
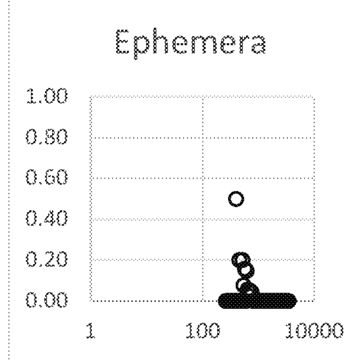
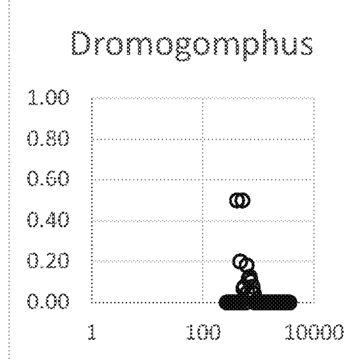
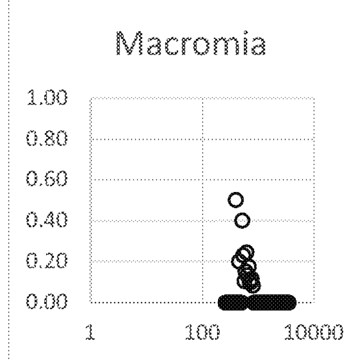
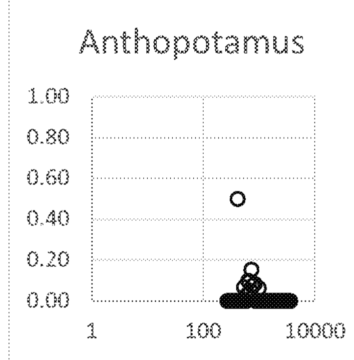
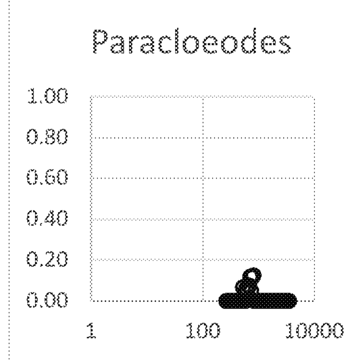
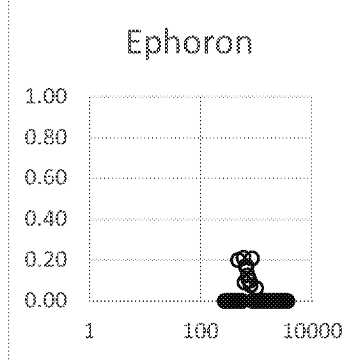
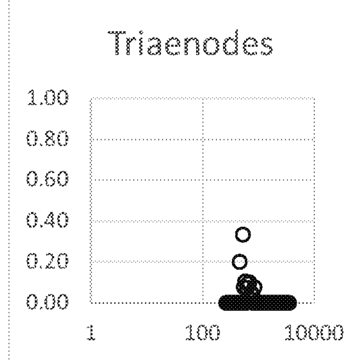
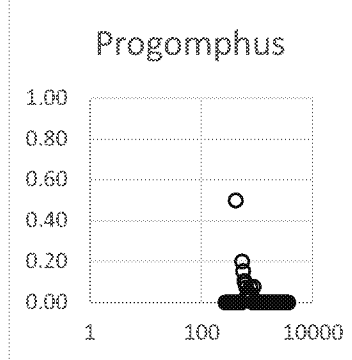
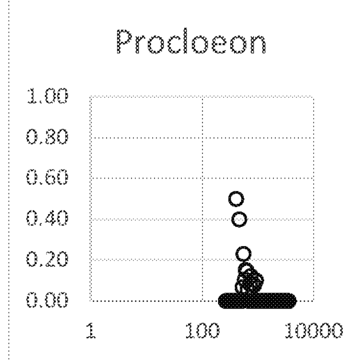
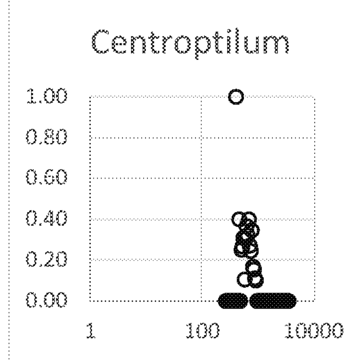
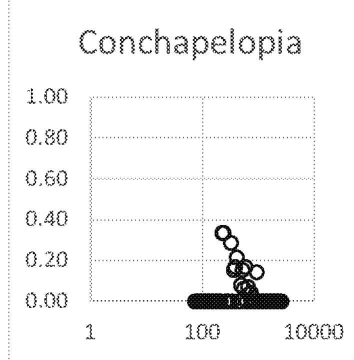
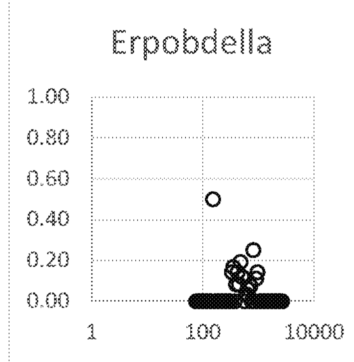
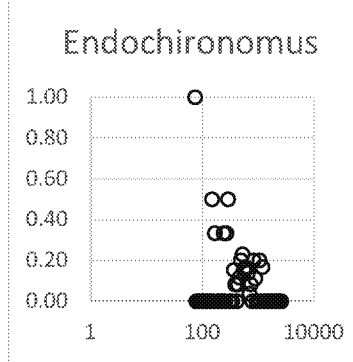
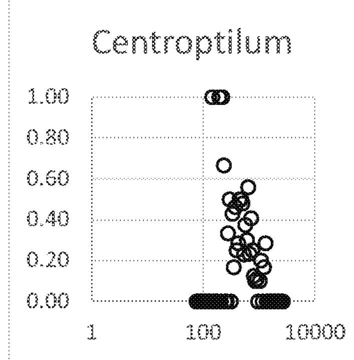
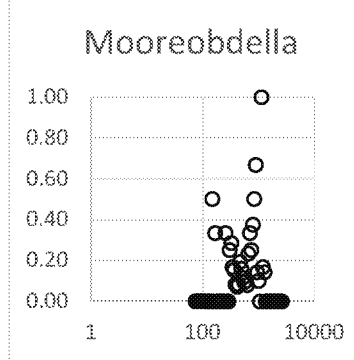
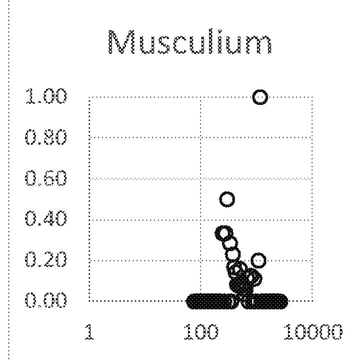
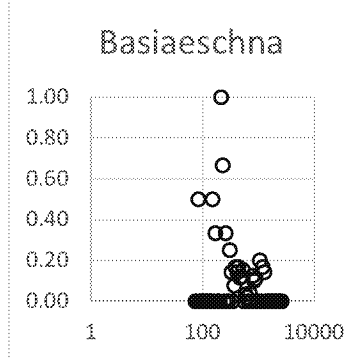
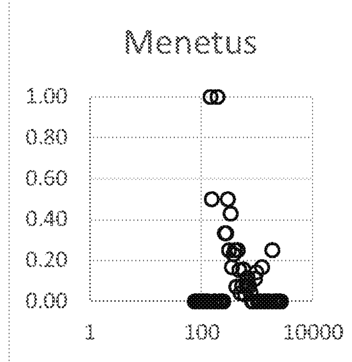
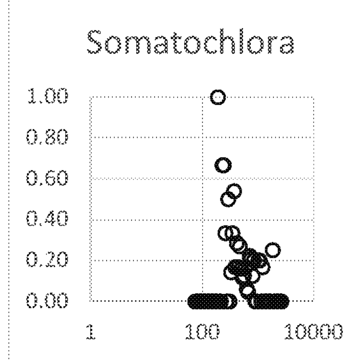
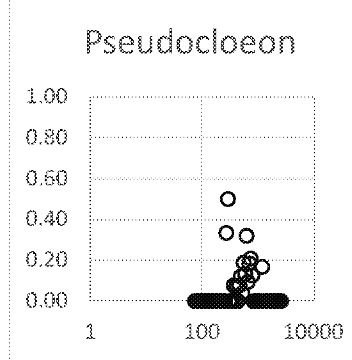
Vincent, E. R.. 1987. Effects of Stocking Catchable-Size Hatchery Rainbow Trout on Two Wild Trout Species in the Madison River and O'Dell Creek, Montana. *North American Journal of Fisheries Management* 7:91-105.

West Virginia Department of Environmental Protection (WVDEP). 2010. *Justification and Background for Permitting Guidance for Surface Coal Mining Operations to Protect West Virginia's Narrative Water Quality Standards, 47 C.S.R. 2 §§ 3.2e and 3.2.i.* Release Date: August 12, 2010.

Zale, A. V., D. L. Parrish, and T. M. Sutton (eds.). 2013. *Fisheries Techniques*, 3rd edition. American Fisheries Society, Bethesda, MD.

Appendix A Additional Figures

Figure A-11:Capture probability profiles for the top ten list of genera for the selected ecoregions. Genera are ranked from 1 to 10 with number 1 exhibiting the lowest XC$_{95}$ value for the respective ecoregion. X-axis is specific conductivity in µS/cm, and y-axis is capture probability.

State / Ecoregion	1	2	3	4	5	6	7	8	9	10
AZ 23 n = 97 HC$_{05}$ 252 µS/cm	 Drunella	 Narpus	 Agapetus	 Cinygmula	 Amphinemura	 Lepidostoma	 Zaitzevia	 Hesperophylax	 Glossosoma	 Sweltsa
ID 17 N = 110 HC$_{05}$ 265 µS/cm	 Stempellina	 Visoka	 Caudatella	 Doroneuria	 Glutops	 Paraperla	 Megarcys	 Neophylax	 Dolophilodes	 Rhithrogena
IL 54 n = 124 HC$_{05}$ 821 µS/cm	 Ephemera	 Dromogomphus	 Macromia	 Anthopotamus	 Paracloeodes	 Ephoron	 Triaenodes	 Progomphus	 Procloeon	 Centroptilum
IL 72 n = 116 HC$_{05}$ 1,110 µS/cm	 Conchapelopia	 Erpobdella	 Endochironomus	 Centroptilum	 Mooreobdella	 Musculium	 Basiaeschna	 Menetus	 Somatochlora	 Pseudocloeon

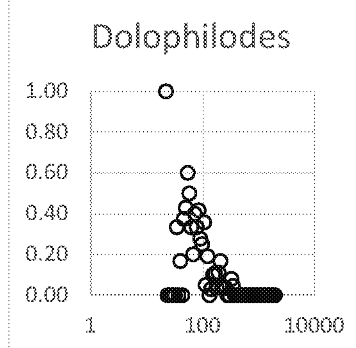
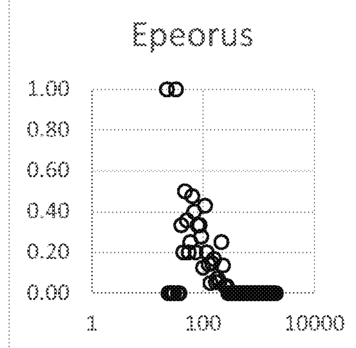
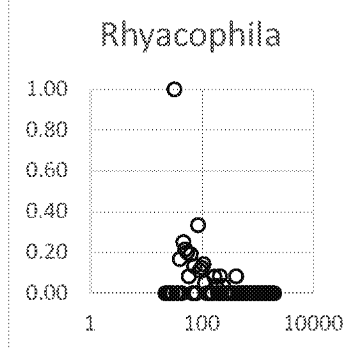
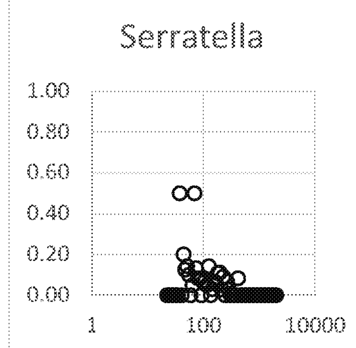
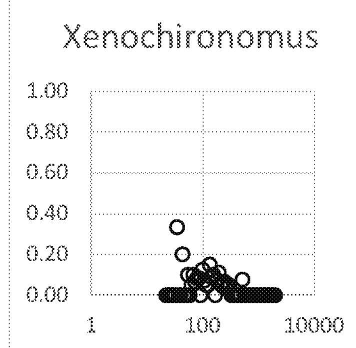
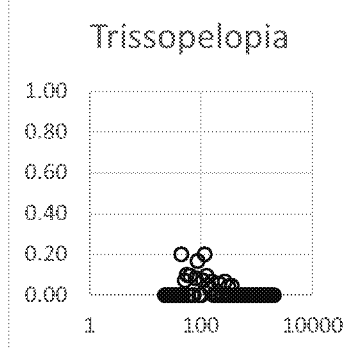
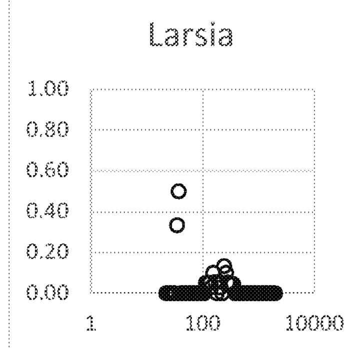
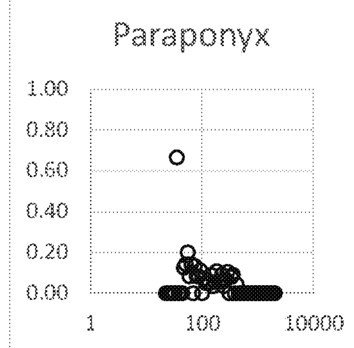
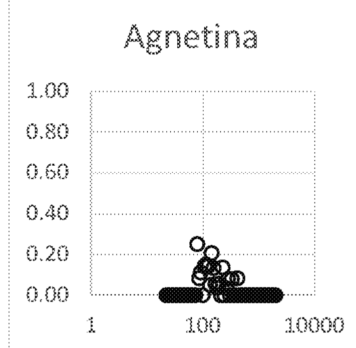
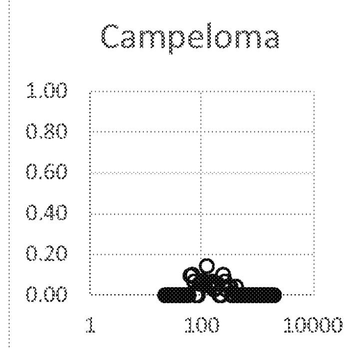
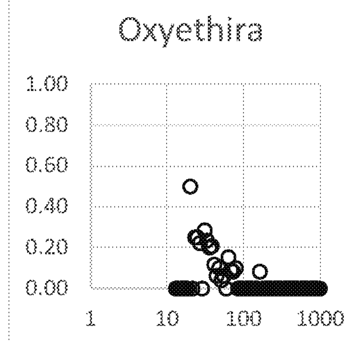
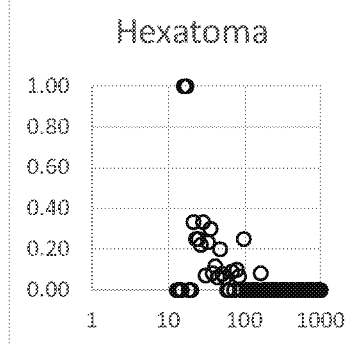
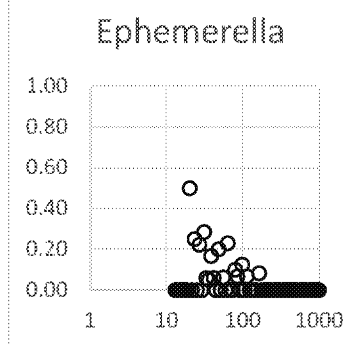
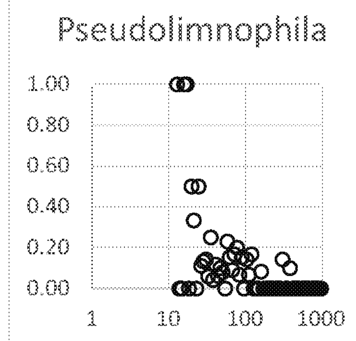
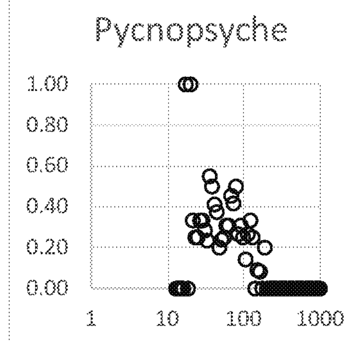
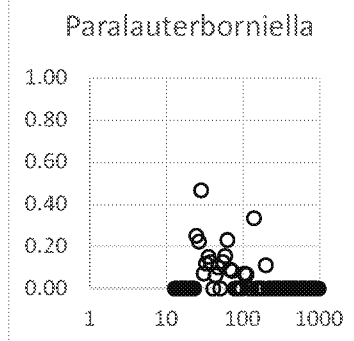
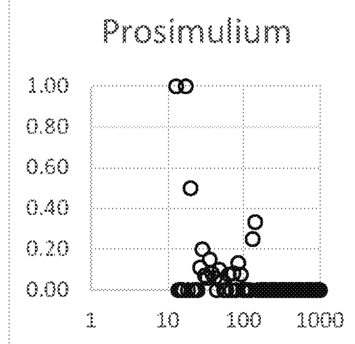
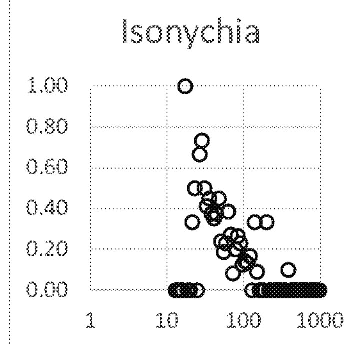
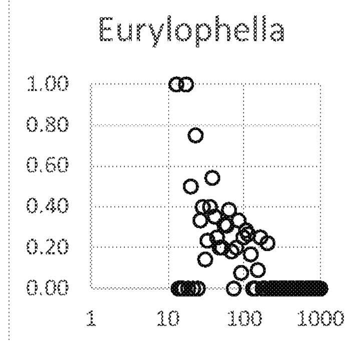
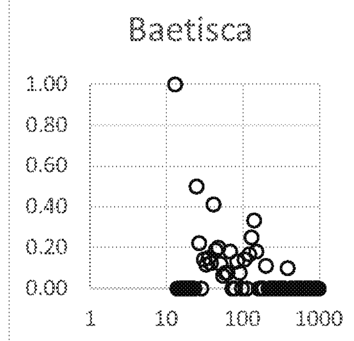
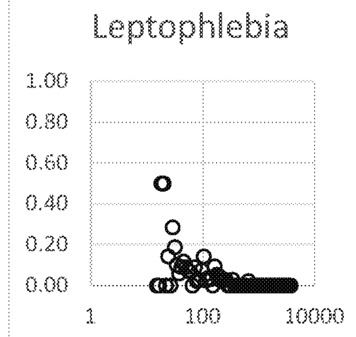
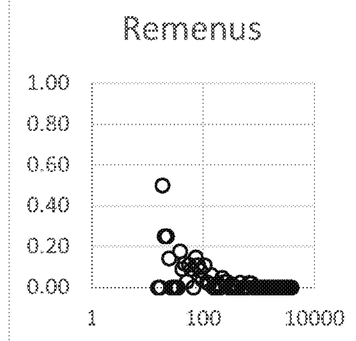
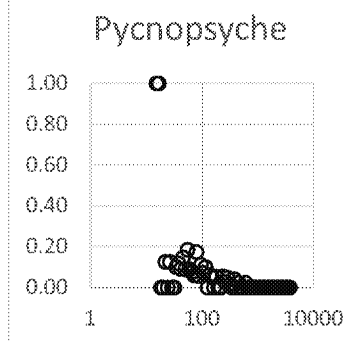
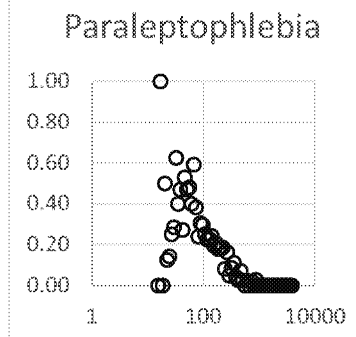
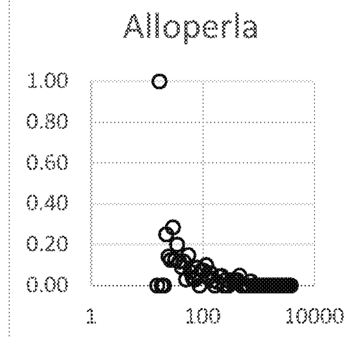
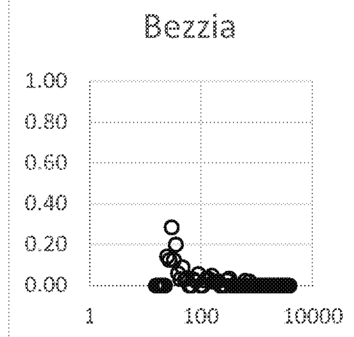
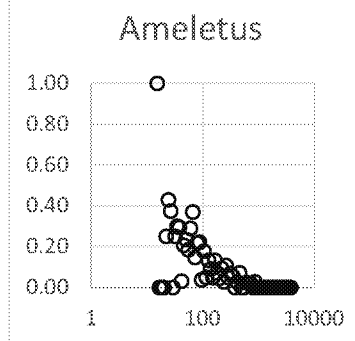
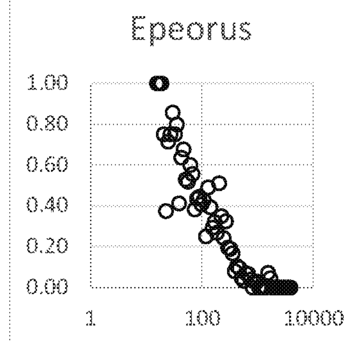
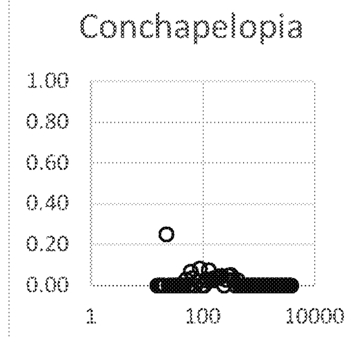
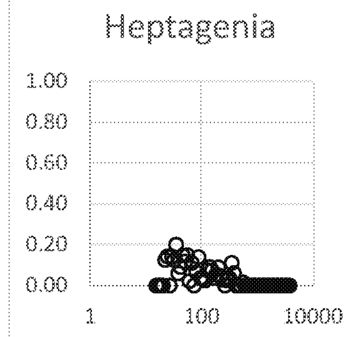
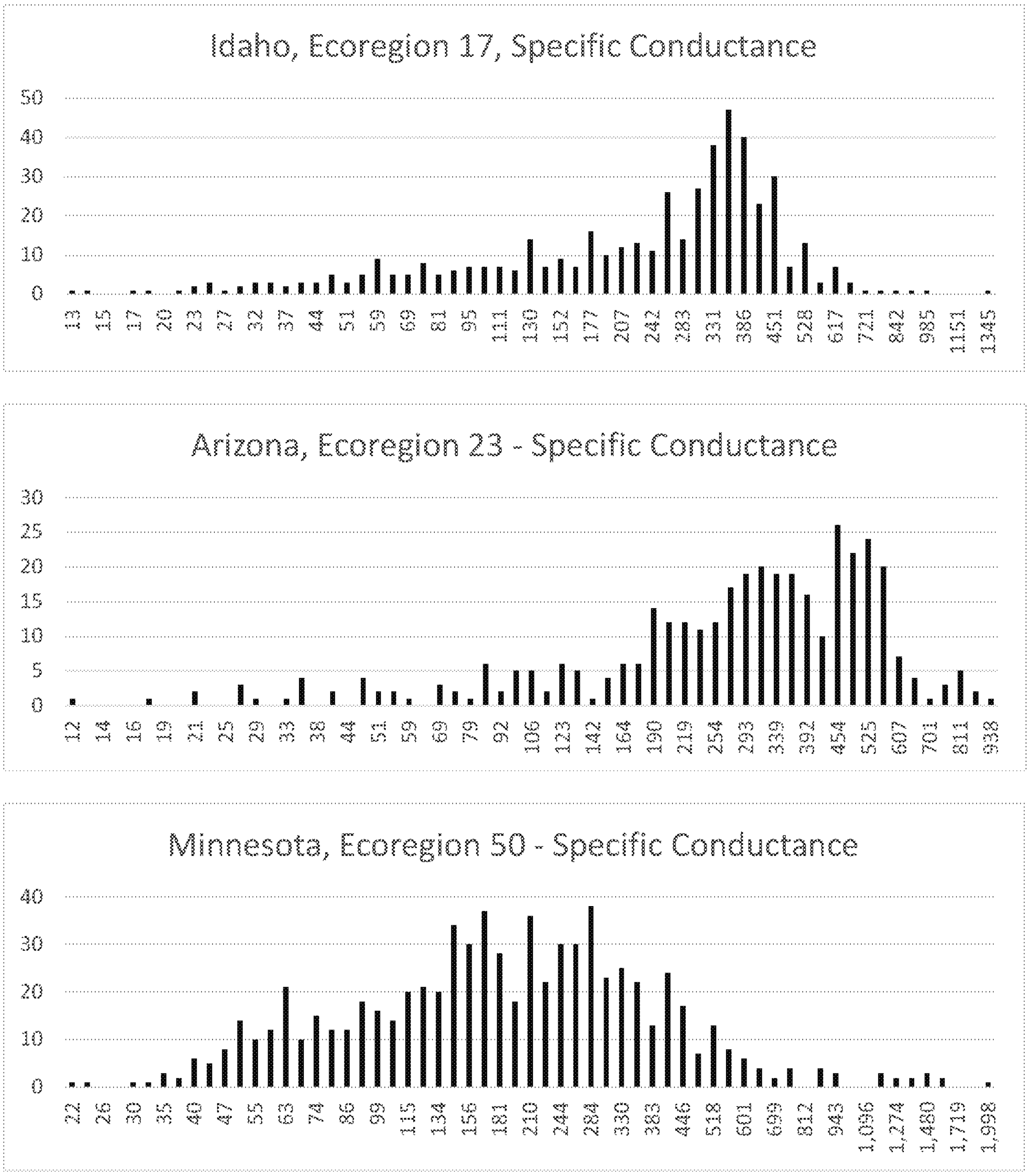
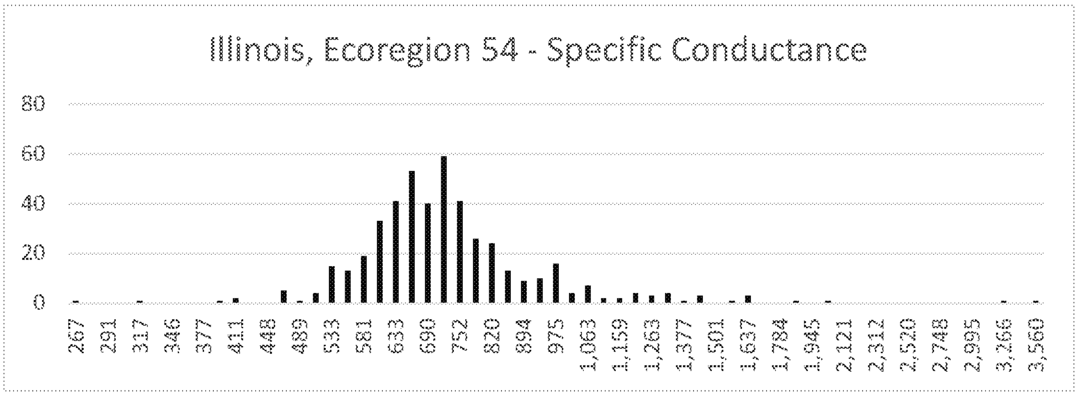
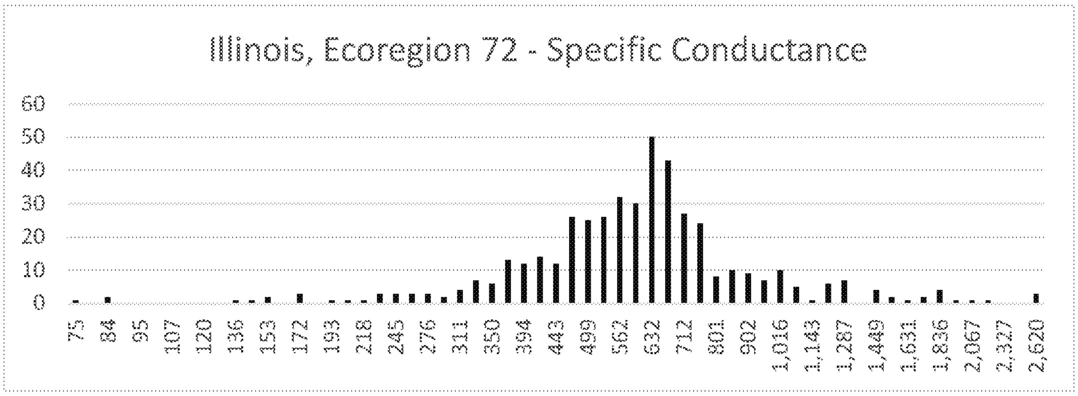
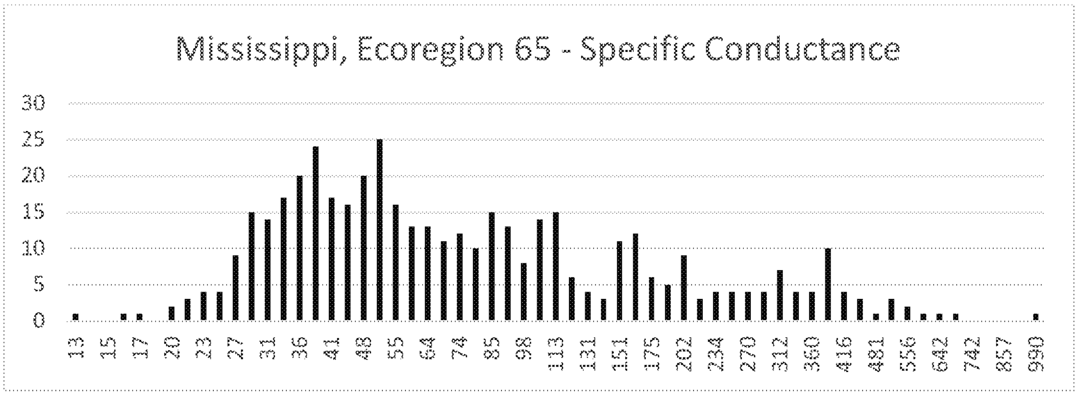
State / Ecoregion	1	2	3	4	5	6	7	8	9	10
MN 50 n = 176 HC$_{05}$ 317 µS/cm	 Dolophilodes	 Epeorus	 Rhyacophila	 Serratella	 Xenochironomus	 Trissopelopia	 Larsia	 Paraponyx	 Agnetina	 Campeloma
MS 65 n = 115 HC$_{05}$ 135 µS/cm	 Oxyethira	 Hexatoma	 Ephemerella	 Pseudolimnophila	 Pycnopsyche	 Paralauterborniella	 Prosimulium	 Isonychia	 Eurylophella	 Baetisca
WV69 n = 142 HC$_{05}$ 302 µS/cm	 Leptophlebia	 Remenus	 Pycnopsyche	 Paraleptophlebia	 Alloperla	 Bezzia	 Ameletus	 Epeorus	 Conchapelopia	 Heptagenia

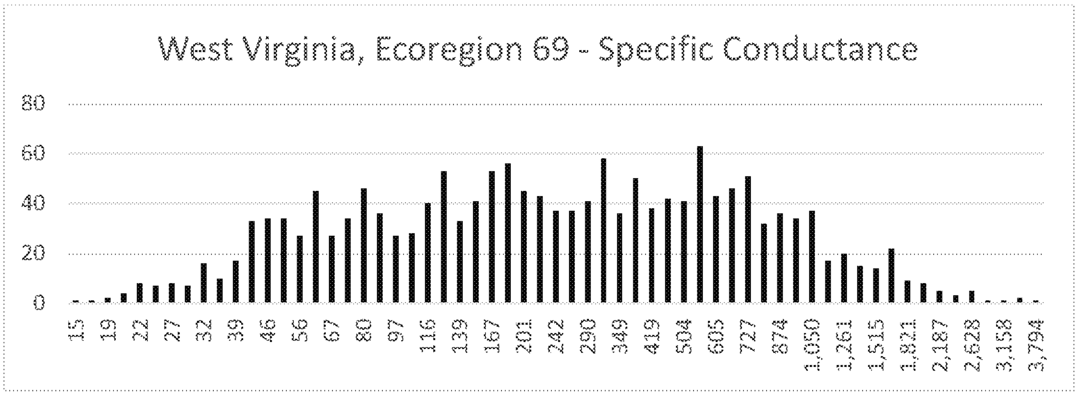
Figure A-2: Histograms of conductivity for selected ecoregions. X-axis is specific conductivity in µS/cm, and y-axis is number of observations per bin.




Illinois, Ecoregion 54 - Specific Conductance
80
60
40
20
0
267
291
317
346
377
411
448
489
533
581
633
690
752
820
894
975
1,063
1,159
1,263
1,377
1,501
1,637
1,784
1,945
2,121
2,312
2,520
2,748
2,995
3,266
3,560


Illinois, Ecoregion 72 - Specific Conductance
60
50
40
30
20
10
0
75
84
95
107
120
136
153
172
193
218
245
276
311
350
394
443
499
562
632
712
801
902
1,016
1,143
1,287
1,449
1,631
1,836
2,067
2,327
2,620


Mississippi, Ecoregion 65 - Specific Conductance
30
25
20
15
10
5
0
13
15
17
20
23
27
31
36
41
48
55
64
74
85
98
113
131
151
175
202
234
270
312
360
416
481
556
642
742
857
990


West Virginia, Ecoregion 69 - Specific Conductance
80
60
40
20
0
15
19
22
27
32
39
46
56
67
80
97
116
139
167
201
242
290
349
419
504
605
727
874
1,050
1,261
1,515
1,821
2,187
2,628
3,158
3,794

Appendix B WVDEP 2010

Release Date: August 12, 2010

west virginia department of environmental protection

Justification and Background for Permitting Guidance for Surface Coal Mining Operations to Protect West Virginia's Narrative Water Quality Standards, 47 C.S.R. 2 §§ 3.2.e and 3.2.i

PURPOSE

The West Virginia Department of Environmental Protection ("DEP") adopts this Justification and Background for its "Permitting Guidance for Surface Coal Mining Operations to Protect West Virginia's Narrative Water Quality Standards" (the "Guidance"). The Guidance is intended to facilitate compliance with applicable statutory and regulatory requirements and to provide reasonable means of effectuating the intent of the narrative criteria, as well as to enforce the mandate of the Clean Water Act ("CWA") that every permit contain effluent limitations that reflect the practicable pollution reduction a state can achieve.[1]

The Guidance was developed in accordance with the West Virginia Water Pollution Control Act ("WVWPCA"), which states that "the public policy of the State of West Virginia to maintain reasonable standards of purity and quality of the water of the State consistent with (1) public health and public enjoyment thereof; (2) the propagation and protection of animal, bird, fish, aquatic and plant life; and (3) the expansion of employment opportunities, maintenance and expansion of agriculture and the provision of a permanent foundation for healthy industrial development."[2]

As it must, the Guidance also recognizes the intent of the West Virginia Legislature, which has formally resolved as follows:

- That any interpretation and implementation of West Virginia's narrative water quality standards is the responsibility of the West Virginia Department of Environmental Protection;
- That the requirements of the narrative criteria are met when a stream (a) supports a balanced aquatic community that is diverse in species composition; and (b) contains appropriate trophic levels of fish (in streams with sufficient flows to support fish populations); and (c) the aquatic community is not composed only of pollution tolerant species or

---

[1] *American Paper Institute, Inc. v. United States Environmental Protection Agency*, 996 F.2d 346, 349 (D.C. Cir., 1993)

[2] W. Va. Code § 22-11-2(a).

the aquatic community is composed of benthic invertebrate assemblages sufficient to perform the biological functions necessary to support fish communities within the assessed reach (or, if the assessed reach has insufficient flows to support a fish community, in those downstream reaches where fish are present); and

- That interpretation of West Virginia's narrative water quality standards must faithfully balance the protection of the environment with the need to maintain and expand opportunities for employment, agriculture, and industry as set forth in the Legislature's statement of public policy as contained in the West Virginia Water Pollution Control Act.[3]

BACKGROUND

West Virginia has had primacy of the NPDES program since 1982 and has narrative water quality standards that predate its NPDES primacy. These criteria are found in West Virginia's *Code of State Rules*, which states, in pertinent part, "No significant adverse impact to the chemical, physical, hydrologic, or biological components of aquatic ecosystems shall be allowed."[4]

In light of its goals to advance, wherever attainable, water quality that provides for recreation and the protection and propagation of fish, shellfish, and wildlife,[5] and to assure that surface mining operations are conducted so as to protect the environment,[6] DEP reviewed its NPDES permitting and compliance assessment protocols vis-à-vis West Virginia's narrative water quality standards and solicited public comment regarding these issues. As a result, DEP adopts the Guidance, which describes the procedures DEP will implement in the development of NPDES permits for the coal mining industry. These new procedures shall take effect immediately. In light of the changing nature of the policy concerns addressed herein, this document is intended to be dynamic and will likely be modified in the future as technology and best management practices develop and improve.

While DEP appreciates EPA's recent effort to assist the states in interpreting their various narrative water quality standards, DEP finds that the Guidance is the more appropriate approach for West Virginia for several reasons. First, it involves subject matter uniquely within DEP's expertise and special knowledge. Further, while this document specifically addresses concerns related to the mining industry, it is designed to be adapted in the future to address all discharges to water bodies that will cause, or that have the reasonable potential to cause or contribute to, excursions from water quality standards. Finally, it does not use an overbroad, generic criterion (i.e. conductivity) to set unattainable limits, but instead identifies specific pollutants that can be managed through the inclusion of appropriate whole effluent toxicity ("WET") monitoring and/or limits and best management practices ("BMPs") in NPDES permits, where there is reasonable potential to cause or contribute to excursions from water quality criteria. If the

---

[3] H.C.R. 111 (2010 Regular Session).
[4] 47 C.S.R. 2 § 3.2.i
[5] *See* 33 U.S.C. § 1251(a)(2)
[6] *See* 30 U.S.C. § 1202(d)

applicant cannot demonstrate, by means of its chemical and biological monitoring and the control measures outlined in the plans it will submit with its application, that it does not have reasonable potential ("RP") to cause or contribute to an excursion above the narrative criteria, the permit writer should treat new or expanded discharges as if they have RP and include WET limits in the permit, in accordance with 40 C.F.R. § 122.44(d)(1)(v). Alternatively, if the operator identifies toxic pollutants that can be regulated through the use of numeric limits, DEP will put a regulatory control number for those pollutants in the operator's permit.

PROTECTION OF THE AQUATIC ECOSYSTEM

As stated above, the narrative water quality criteria set out in 47 C.S.R. 2 § 3.2.i prohibits the introduction of wastes that cause significant adverse impact to the chemical, physical, hydrologic or biological components of aquatic ecosystems. These criteria are valid components of West Virginia water quality standards that have been properly promulgated by the West Virginia Legislature and approved by the EPA. The phrase "significant adverse impact" is not defined in the CWA or the WVWPCA, the regulations promulgated thereunder or in any literature or guidance published by the EPA. DEP has determined that "significant adverse impact" is more than a change in the numbers or makeup of the benthic macroinvertebrate community in a segment of a water body downstream from a point source discharge. It is, instead, a material decline in the overall health of an aquatic ecosystem.[7] A goal of the CWA and the WVWPCA is to protect the aquatic ecosystem as a whole; it is a holistic standard that requires a holistic approach to ecosystem assessment. In contrast to numeric water quality criteria, which can be applied by analysis of samples of water taken at any discharge or monitoring point in a stream, compliance with a standard that protects the aquatic ecosystem must be assessed in the broader area comprising the ecosystem. An ecosystem does not exist at a single point and, accordingly, its health cannot be assessed at a single point.

The Pond-Passmore Study, upon which EPA relied in the development of its guidance on this subject, concludes that West Virginia's narrative standard is violated by surface coal mining operations based on the Study's application of two biologic assessment tools, the West Virginia Stream Condition Index ("WVSCI") and the draft Genus Level Index of Most Probable Stream Status ("GLIMPSS"), to samples of benthic macroinvertibrate life taken from these streams. This conclusion is flawed for two reasons. First, West Virginia does not use the draft GLIMPSS in its assessment of the biologic health of State streams. Second, these tools are just that – tools. They are not stand-alone determinants of compliance with the narrative standard. Any application of these assessment tools in determining compliance with the narrative standard must faithfully apply the language of the standard itself, which prohibits significant adverse impacts on the chemical, physical, hydrologic or biological components of the aquatic ecosystem. Thus, DEP's Guidance follows long-standing EPA guidance, which indicates that biosurveys cannot fully characterize an entire aquatic community and its many attributes, and accordingly suggests that "State standards should contain biological criteria that consider various components (e.g.

---

[7] An aquatic ecosystem is a dynamic complex of plant, animal, and microorganism communities and their non-living environment interacting as a functional unit within water. *See*, Coweeta Long Term Ecological Research "Glossary of Terms."

algae, invertebrates, fish) and attributes (measures of structure and/or function) of the larger aquatic community."[8]

Through implementation of the Guidance, DEP continues its existing practice of using WVSCI in addition to consideration of other factors affecting the aquatic ecosystem to enforce its narrative water quality standards. By way of background, WVSCI was developed for EPA by national experts to assess biological integrity in West Virginia's waterways through "careful measurement of the natural aquatic ecosystem and its constituent biological communities,"[9] including the evaluation of benthic macroinvertebrate communities. It was specifically designed for assessment of the biological component of the 47 C.S.R. 2 § 3.2.i narrative criteria and has been used as a tool in developing the Impaired Streams List ("303(d) List") and the TMDLs resulting therefrom for almost a decade.[10] WVSCI acknowledges that "[i]t is the responsibility of West Virginia's [Department] of Environmental Protection to maintain and protect the ecosystem health of the state's waters[,]" and "[i]n keeping with the Clean Water Act and technical guidance from USEPA, DEP developed water quality standards for the protection of ecosystem health."[11]

DEP's Guidance is the appropriate methodology for implementing West Virginia's narrative water quality standards, because it is consistent with the Federal Regulations regarding establishing limitations, standards, and other permit conditions for NPDES programs, and it incorporates a holistic approach to ecosystem assessment and protection. The CWA's implementing regulations require WET testing and limits when the State finds that a discharge has RP to cause or contribute to excursions from water quality standards.

> [W]hen the permitting authority determines . . . that a discharge causes, has the reasonable potential to cause, or contributes to an in-stream excursion above a narrative criterion within an applicable State water quality standard, the permit must contain effluent limits for whole effluent toxicity. Limits on whole effluent toxicity are not necessary where the permitting authority demonstrates in the fact sheet or statement of basis of the NPDES permit . . . that chemical-specific limits for the effluent are sufficient to attain and maintain applicable numeric and narrative State water quality standards.[12]

WET testing allows flexibility where appropriate (e.g. allowing time to collect additional data for RP determination to supplement limited data sets) and is consistent with DEP's policy that

---

[8] EPA's *Policy on the Use of Biological Assessments and Criteria in the Water Quality Program* (May 1991) ("1991 Policy")

[9] A Stream Condition Index for West Virginia Wadeable Streams, March 28, 2000 (Rev. July 21, 2000) ("Stream Condition Index").

[10] However, a stand-alone WVSCI score has never been the sole determinant of compliance or non-compliance with the narrative standard. This is because WVSCI scores are influenced by many factors (e.g. habitat, geology, and pH).

[11] Stream Condition Index

[12] 40 C.F.R. § 122.44(d)(1)(v)

permittees develop robust monitoring plans with the intention of identifying any causative pollutants and adjusting their methods of operation so that those problems may be remedied before the aquatic community suffers a significant breakdown.

WVSCI considers various components (e.g. algae, invertebrates, fish) and attributes (measures of structure and/or function) of the larger aquatic community. "Because biological integrity is a strong indicator of overall ecological integrity, it can serve as both a meaningful goal and a useful measure of environmental status. . . ."[13] Based on the 5th percentile of reference values, the current WVSCI score that indicates the integrity of a benthic macroinvertebrate community in West Virginia's wadeable streams is 68.0. The threshold for inclusion on the 303(d) List has historically been 60.6. That value subtracts a precision estimate from the 5th percentile of reference values, and its historical use was intended to take into account sampling error and to aid DEP in allocating its resources so as to avoid misclassifying non-impaired waters as impaired. WVSCI and its application in the 303(d) listing process are consistent with methodologies implemented to assess protection of aquatic ecosystems by all of West Virginia's neighboring states.

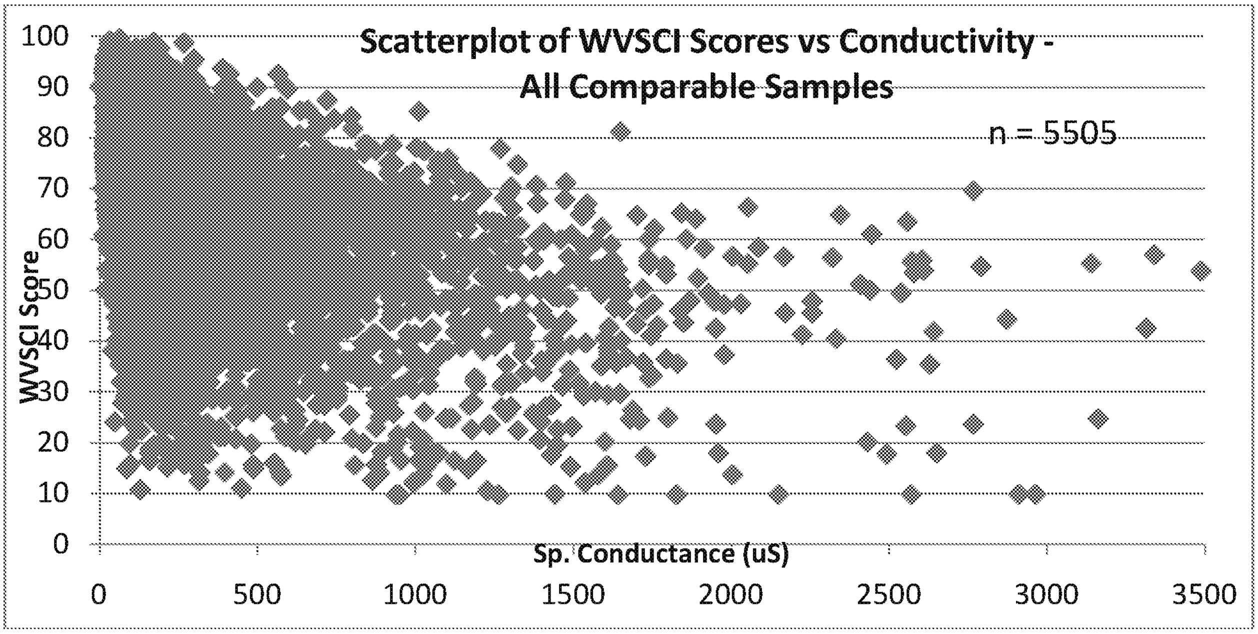
CAUSATIVE POLLUTANTS / PROTECTIVE THRESHOLDS

EPA has recently set a numeric limit on conductivity at 500 µS/cm, finding that 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 adverse impacts that may rise to the level of exceedances of narrative state water quality standards.[14] However, DEP's data shows that more than a simple conductivity measurement is necessary to determine the health of a stream. As proof that a number for specific conductance is an inappropriate gauge, FIGURE 1 below illustrates that a stream can have a low level of specific conductance and a WVSCI score firmly within the range for impairment; conversely, a stream can have a high level of specific conductance and a WVSCI score that indicates the stream is above the threshold for impairment. WVSCI scores are affected by many factors: habitat, other uses of the stream and the surrounding land, other pollutants unrelated to conductivity (e.g. fecal coliform), *inter alia*. Certain stream reaches simply cannot attain a "good" WVSCI score because of those factors.

---

[13] 1991 Policy

[14] EPA's *Detailed Guidance: Improving EPA's Review of Appalachian Surface Coal Mining Operations under the Clean Water Act, National Environmental Policy Act, and the Environmental Justice Executive Order* (April 1, 2010) ("April 1 Memo")

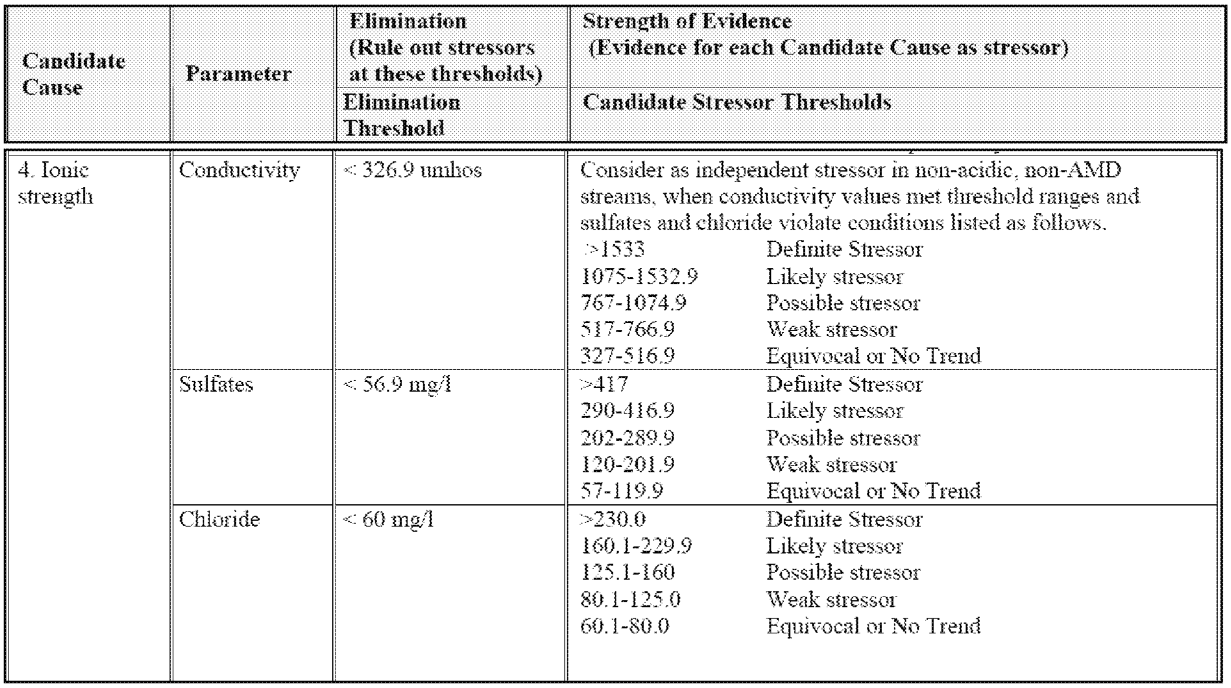


The Pond-Passmore Study found a shift in the benthic macroinvertibrate community downstream from mining activity, but did not otherwise correlate this finding with any significant or adverse impairment of the ecosystem. Where the only impacts to this component of the ecosystem are diminished numbers of certain genera of mayflies, without evidence that this has had any adverse impact of any significance on the rest of the ecosystem, the State cannot say that there has been a violation of its narrative standard. Various scientific studies and evaluations performed by DEP indicate that lowered biological condition is associated with increased ionic strength, but scientists remain less than certain about the specific causative pollutant(s) and the concentration(s) responsible for impairment. Additional uncertainty is present in correlative studies, because the effects of increased ionic strength cannot be completely distinguished from the effects of other stressors that often co-occur (e.g. organic enrichment, sedimentation). In fact, most available information attempts to relate biological condition to a surrogate parameter, such as specific conductance.

Because conductivity represents the combined concentrations of all different dissolved ions, each with potential varying toxic effects, regulation solely via an indicator such as specific conductance is not the best way to protect against excursions from narrative standards. For example, the elevated dissolved pollutants most commonly associated with mining discharges are sulfate and bicarbonate alkalinity. EPA has not published national recommended aquatic life protection criteria for those pollutants. Similarly, chloride, for which West Virginia has adopted EPA's recommended numeric aquatic life protection water quality criteria, may also be present in some cases. But because chloride seldom exists in the absence of sulfates or alkalinity, singular control of chloride cannot be expected to resolve all ionic stress.

DEP has performed a correlative evaluation of benthic condition and specific conductance. This evaluation suggests that native aquatic life is protected at various values and ranges of specific conductance. This finding supports the basic scientific principle that correlation is not cause and effect. Even though the DEP evaluation applied various filters to the evaluated dataset to address complicating factors listed above, the biological condition of a stream may be different from the condition predicted by specific conductance. In situations such as these, where DEP has determined that it is infeasible to calculate a numeric effluent limit to implement a narrative water quality standard, DEP will include in the permit appropriate WET limits and BMPs to control or abate the discharge of pollutants, in accordance with 40 C.F.R. § 122.44(k)(3).

DEP routinely identifies biological stressors when developing TMDLs for biologically impaired waters. Stressor identification employs a strength-of-evidence approach that considers multiple information sources. Researchers evaluate water quality monitoring data, physical habitat data, field notes, and the composition of the biological assemblage concurrently to identify significant stressors. DEP's most recent stressor identification protocols, as used in the EPA-approved TMDL process, include the guidelines shown in FIGURE 2 below for evaluating water chemistry to determine if ionic strength is a significant stressor:



Candidate Cause	**Parameter**	**Elimination (Rule out stressors at these thresholds)**	**Strength of Evidence (Evidence for each Candidate Cause as stressor)**	
		Elimination Threshold	**Candidate Stressor Thresholds**	
4. Ionic strength	Conductivity	< 326.9 umhos	Consider as independent stressor in non-acidic, non-AMD streams, when conductivity values met threshold ranges and sulfates and chloride violate conditions listed as follows.	
			>1533	Definite Stressor
			1075-1532.9	Likely stressor
			767-1074.9	Possible stressor
			517-766.9	Weak stressor
			327-516.9	Equivocal or No Trend
	Sulfates	< 56.9 mg/l	>417	Definite Stressor
			290-416.9	Likely stressor
			202-289.9	Possible stressor
			120-201.9	Weak stressor
			57-119.9	Equivocal or No Trend
	Chloride	< 60 mg/l	>230.0	Definite Stressor
			160.1-229.9	Likely stressor
			125.1-160	Possible stressor
			80.1-125.0	Weak stressor
			60.1-80.0	Equivocal or No Trend

Based on FIGURE 2, it is clear the EPA limits of 300 – 500 µS/cm established in the April 1 Memo are far more stringent than what it has long approved for West Virginia's TMDL process. As shown above, conductivity in the 300 – 500µS/cm range is "Equivocal or No Trend" as a stressor. Conductivity does not even become a "Likely Stressor" of a stream under this EPA-approved approach until it reaches three to five times these limits: 1075-1532.9 µS/cm. This is additional support for the State's conclusion that reliance on the single surrogate of specific conductance to implement and/or enforce the State's narrative water quality standards is improper. It also demonstrates that EPA's proposed limits are too narrowly focused on a single parameter and single aquatic species to determine the health of the impacted watershed.

Only the West Virginia Legislature can adopt a numeric water quality standard for conductivity (or any other pollutant); DEP has no authority to immediately or unilaterally implement numeric standards. Through adoption of H.C.R. 111, the West Virginia Legislature has given DEP direction as to how it should implement its narrative water quality standards. Even if the Legislature does adopt a numeric standard for conductivity, DEP cannot implement it until after it is approved by the EPA. Based on the loose and questionable causal relationship between conductivity and stream impairment, it remains unclear whether EPA would approve such a numeric limit. EPA's duly promulgated regulation endorses establishment of WET limits where, as here, a state is unable to use a limit for a surrogate parameter. DEP can implement new permitting controls based on the agency's best professional judgment of actions necessary to protect the State's waters using its narrative criteria, with follow-up monitoring and contingencies for unsatisfactory outcomes. Thus, DEP is protecting against excursions from its narrative water quality standards by establishing WET limits and verifying impacts to a stream (or lack thereof) by requiring an extensive, comprehensive monitoring plan for the entire watershed.