EPA Response to External Peer Review Comments

on

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

December 2016

U.S. Environmental Protection Agency
Office of Water, Office of Science and Technology
Health and Ecological Criteria Division
1200 Pennsylvania Avenue, NW
Washington, DC 20460

DISCLAIMER

TABLE OF CONTENTS

1. INTRODUCTION 1
1.1. PEER REVIEWERS:...1
2. PEER REVIEW COMMENTS TABLES 2
TABLE 1. GENERAL IMPRESSIONS 3
TABLE 2. RESPONSE TO CHARGE QUESTIONS 1−3: *DATA SET CONSIDERATIONS*..15
Charge Question 1: 15
Charge Question 2: 24
Charge Question 3: 32
TABLE 3. RESPONSE TO CHARGE QUESTIONS 4−8: *EXAMPLE CRITERIA CALCULATIONS*..38
Charge Question 4: 38
Charge Question 5: 46
Charge Question 6: 53
Charge Question 7: 56
Charge Question 8: 59
TABLE 4. RESPONSE TO CHARGE QUESTIONS 9−12: GEOGRAPHIC APPLICABILITY..68
Charge Question 9: 68
Charge Question 10: 79
Charge Question 11: 82
Charge Question 12: 87
TABLE 5. RESPONSE TO CHARGE QUESTIONS 13−14: SUPPORTING INFORMATION: *FIELD-BASED HC_{05} FOR FISH IN APPALACHIAN STREAMS (SEE APPENDIX G)*.....................................94
Charge Question 13: 94
Charge Question 14: 101
TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT 103
Reviewer 1 103
Reviewer 2 113
Reviewer 4 129
Reviewer 5 142

1. INTRODUCTION

The U.S. Environmental Protection Agency (EPA) submitted its *Draft Field-based Methods for Developing Aquatic Life Criteria for Specific Conductivity*[1] for contractor-led, independent, external peer review in October 2014. The external peer reviewers provided their independent responses to EPA's charge questions. This report documents the EPA's responses to the comments provided by five expert peer reviewers. The review focused on the clarity of the technical document and validity of the draft field-based method and case studies.

This report presents the 13 peer review charge questions and individual reviewer comments (verbatim) in the tables in Section 2. EPA separated each reviewer's comments by charge question into distinct topics and responded to each topic individually, and also indicated how the draft conductivity method was revised in response to peer reviewer comments.

1.1. PEER REVIEWERS:

An EPA contractor identified and selected five reviewers who met the technical selection criteria provided by EPA and who had no conflict of interest in performing this review:

Yong Cao, Ph.D.
Illinois Natural History Survey
University of Illinois at Urbana-Champaign

Marian R.L. Maas, Ph.D.
Independent Consultant

Raymond P. Morgan II, Ph.D.
University of Maryland Center for Environmental Science
Appalachian Laboratory

Edward T. Rankin, M.S.
Midwest Biodiversity Institute

Carl E. Zipper, Ph.D.
Department of Crop and Soil Environmental Sciences
Virginia Polytechnic Institute and State University

The EPA contractor provided reviewers with instructions, the review document (including appendices), the charge questions for reviewers prepared by EPA, and supporting reference materials as described in the charge. Reviewers worked individually to develop written comments in response to the charge questions.

[1] At the time of the peer review the document was titled, "*Draft Recommended Field-based Method for States to Develop Ambient Aquatic Life Water Quality Criteria for Conductivity*."

2. PEER REVIEW COMMENTS TABLES

TABLE 1. GENERAL IMPRESSIONS

Reviewer	Comment	EPA Response
Reviewer 1	It is great to see U.S. EPA developing a field-based method for establishing aquatic-life criteria for conductivity, an increasingly important stressor for freshwater ecosystems.	Thank you for your comment. EPA agrees that stress from elevated ionic concentration as measured by specific conductivity has been shown to cause significant adverse effects on a range of freshwater ecosystems across the United States.
Reviewer 1	The effects of abiotic (habitat, flow, water chemistry) and biotic factors (e.g., competition and predation) on the responses of a taxon to increased conductivity are complex and poorly understood. It is therefore sensible to use a field-based approach, rather than lab tests, to derive the criterion.	Thank you for your comment. EPA agrees that a field-based approach is appropriate for developing specific conductivity criteria. See Section 2.7 in the public review draft.
Reviewer 1	I also believe that genera are the best choice of taxonomic units because the sensitivities of species from the same genus are often similar, and the identifications at this level are normally more accurate and less costly than at the species level. Clumping taxonomic data to the genus level also increases the number of taxon occurrences and makes more taxa available in a region for establishing a conductivity criterion.	Thank you for your comment. EPA agrees that it is appropriate to use a genus level method for benthic macroinvertebrates. However, species-level taxonomic identification can be used when it is available and the number of species is sufficient for constructing a sensitivity distribution (see Appendix G of the public review draft for an example that uses fish species).
Reviewer 1	Overall, the document is well written. However, I have several major concerns, particularly on the concept and measure (XC_{95}) of taxon extirpation and associated statistical analysis. The vague and inconsistent relationship between XC_{95} and extirpation appears to have compromised the rigor of the process of criterion development.	EPA has added text cautioning against the use of the sensitivity distribution (SD) for predicting the proportion of genera extirpated at the upper end of the SD where the XC_{95} are undetermined (> values) (see Section 3.1.3 in the public review draft).

TABLE 1. GENERAL IMPRESSIONS (continued)		
Reviewer	**Comment**	**EPA Response**
Reviewer 1	I am also worried about how specific case studies in the document are used to justify extrapolation of a conductivity criterion developed for one region to another region.	EPA disagrees with the reviewer's concern about extrapolation and has added new analyses to the public review draft which show that SC criteria can be predicted based on the natural history of resident aquatic organisms and background SC (see Sections 6 and 7 and Appendix D of the public review draft). This method described in Appendix D was externally peer reviewed separately in 2015. The explanation of the ecological theory supporting the method has been expanded and a model illustrates the relevance of the concept of niche space to the occurrence of salt-intolerance in naturally very dilute waters. The new analyses show that even when regions are thousands of miles apart and separated by areas with higher background SC, available low SC niche space is occupied by salt-intolerant genera, and the HC_{05} can be predicted based on background SC.
Reviewer 1	In addition, some terms (e.g., probability of capture) need to be more clearly defined, and equations need to be constructed in a standard format.	EPA added more explanation in Section 3.1.2.1. Equations have been revised and are now in standard format.
Reviewer 2	It was a pleasure to review the U.S. EPA draft document, *Recommended Field-based Method for States to Develop Ambient Aquatic Life Water Quality Criteria for Conductivity*. As a biologist who has worked both in the field and in managerial capacity for water quality monitoring and bioassessment, the area of specific conductivity has long been more or less overlooked. This is largely due to two reasons: 1) most city, watershed and state monitoring programs don't quite know what to do with conductivity data, and 2) the role of conductivity in its effect on aquatic and benthic organisms is not well understood. The need for a criterion is basic to the improvement in these areas.	Thank you for your comment. EPA agrees that ionic stress is an important water quality issue affecting aquatic life and specific conductivity is the appropriate measure of the ionic mixture for the draft field-based methods. See Section 2.7 in the public review draft.

TABLE 1. GENERAL IMPRESSIONS (continued)

Reviewer	Comment	EPA Response
Reviewer 2	This document's clear and strong guidance in providing a method for development of conductivity criterion and for a method to make it applicable to adjoining regions is an immensely valuable new and long over-due tool for monitoring programs. The document provides considerable information on the effects of high conductivity levels on macroinvertebrates and fish, and provides strong, data-supported rationale for its approaches and methods. Very large data sets, paired analyses, and strong/reliable/widely used statistical models were used in all of the analytical processes. The biological information in all sections was especially accurate, thorough, and clearly written. It was a pleasure to read those sections and to learn new information.	Thank you for your comment. No response needed.
Reviewer 2	The statistical material was less clear for me, but that is more of a deficiency on my part than that of the document. In that regard, perhaps more explanation of several of the calculation processes could be provided, and a full, working example of each (probably placed in the Appendices), would be helpful for water quality staff who have limited statistical training. With such examples, staff could follow the step-by-step process. I realize this might be viewed by the authors as somewhat of an unnecessary effort, but I believe it would help the document's usability by a greater number of staff with varying backgrounds and knowledge base.	EPA plans to prepare a step-by-step technical support document to assist users. This document is expected to be published simultaneously with the final version of the field-based methods document.
Reviewer 2	Thank you for the opportunity to review this excellent document. The document is well done and its conclusions correct. I believe it will be a valuable guidance for development of much needed criterion for conductivity. It will provide an immensely important function in the improvement of water quality and ecological health for the nation's streams and rivers.	Thank you for your comment. No response needed.

TABLE 1. GENERAL IMPRESSIONS (continued)

Reviewer	Comment	EPA Response
Reviewer 3	I was impressed by the overall depth and breadth of this very well prepared report by EPA (in my opinion it is one of the best technical reports, within my research areas, ever prepared by EPA). Obviously, this report has already gone through a rigorous internal EPA review (by many people who I know professionally) and by EPA contract support, as well as review by several other people who I also know and respect for their work. There is no question that the internal EPA technical workgroup contributed strongly to the report quality, again many people that I know professionally and respect.	Thank you for your comment. No response needed.
Reviewer 3	Having worked with acid mine drainage, acid rain and numerous stream chemistry studies (as well as a few other lotic and lentic projects where conductivity was measured, besides a past life in estuarine and marine ecosystems measuring salinity), I am quite familiar with the strengths and vagaries of this very important measurement in both the field and laboratory.	No response needed.
Reviewer 3	Also, I was pleased to see most of the key, but rather ancient, papers cited (e.g., the 1985 Hem paper), indicating that the literature review was excellent (although the key Hem 1982 paper was missing). However, there were also some recent, rather significant papers missing, but that may be due to the timing of the report preparation. No matter how hard you try, there are always supportive papers that may be missed in any literature review.	The reviewer did not provide a list of papers except the Hem 1982 reference, which has been added. The Hem 1982 paper is referenced in the more complete 1985 paper. Because it was redundant, EPA did not cite both. Hem. 1982, Conductance: a collective measure of dissolved oxidation: U.S. Geological Survey Water-Supply Paper ions, in Minear, R. A., and Keith, L. H., eds., Water 1667-A, 64 p. analysis, v. I, Inorganic species pt. 1: New York, Academic Press, p. 137-161. EPA has updated the list of references in the public review draft with more recent studies on physiology of ionoregulation, mesocosm studies, and field studies of recovery, for example.

TABLE 1. GENERAL IMPRESSIONS (continued)

Reviewer	Comment	EPA Response
Reviewer 3	One concern was the redundancy of writing throughout the report (clarity was excellent for most sections), where general concepts appear to be often restated within some sections of the main report. I don't think that you always need to restate the obvious throughout every section of the main body of the report. However, with the potential wide array of future readers, some writing redundancy may be helpful. Here, my advice would be to have an outside professional editor (e.g., someone from Academic Press, Science, etc.) review the report structure and make recommendations to streamline this effort.	The language was retained so that Sections 3 and 4 and Appendices A and B can stand alone and serve as separate reference documents for states and other stakeholders. No change was made.
Reviewer 3	I like the Level III ecoregion approach, a method that I am using in some of my own work.	EPA agrees that the Level III ecoregion approach is generally appropriate for developing field-based SC criteria in order to take into account natural differences in background SC. However, there may be situations where it is not appropriate to apply criteria derived for the ecoregion to a particular stream reach. For example, naturally lower or higher concentrations of ions may occur due to subecoregional differences such as cross boundary influences, glacial melt, salt springs, highly soluble rock, or other natural sources. In such cases, SC criteria could be developed on a smaller scale provided there are sufficient data to do so. See Section 3.7.1 in the public review draft.
Reviewer 3	My only concern here would be if there are applicable and robust data sets for each of the 85 Level III ecoregions. The report certainly uses a very robust, regional database to develop criteria and to examine the statistical techniques needed to develop conductivity criteria within an ecoregion, and adjoining ecoregions. It will be really interesting to see how the States and Tribes respond to this report, as well as Congress.	Large data sets are not available for all 85 Level II ecoregions for de novo derivation of criteria; however, some ecoregions may have similar background SC, climate, and geology and be appropriate for combining data sets. In addition, EPA has added a regression model so that the HC_{05} can be predicted from background SC (see Appendix D in the draft document). The method described in Appendix D was externally peer reviewed separately in 2015.

TABLE 1. GENERAL IMPRESSIONS (continued)

Reviewer	Comment	EPA Response
Reviewer 4	Overall, this is a well-written, scientifically sound paper which lays out the technical approach and underpinnings for deriving conductivity benchmarks for aquatic life for streams using field-derived measures of water chemistry and ionic strength measures with co-currently collected measure of aquatic macroinvertebrate response at the genus level of taxonomy.	Thank you for your comment. No response needed.
Reviewer 4	My major issues are related to the actual application of these results to protect aquatic life uses in State water quality standards and, particularly, under tiered aquatic life use frameworks. Application of this method in states, like Ohio, that have tiered aquatic life uses, could result in benchmarks that are overprotective of the baseline warmwater aquatic life use, but could also be under-protective of exceptional (EWH) uses. Fortunately, I think the methodology presented here can take these factors into account. For example, for "EWH" streams there may be a more restrictive suite of species that occur in these waters and exclusion of more tolerant forms could drive the XC_{95} a bit lower. In contrast, for "WWH" streams, the most sensitive species may not occur frequently enough in those streams to "drive" the XC_{95} benchmarks and the more common sensitive species might result in a less stringent, but perhaps more attainable benchmark.	On a site-specific basis, the example criteria developed using the draft methods could be adjusted or recalculated to protect important species, highly valued aquatic communities, or specially protected waters.
Reviewer 4	Another key issue that may influence the derivation of criteria with this method is the definition of reference conditions. I think some more discussion of reference conditions as in Stoddard et al. (2006) would be helpful, I will delve into these comments in more detail in the charge questions below.	Definitions have been added to the glossary and the text was expanded. The definitions proposed by Stoddard et al. (2006) were used with minor modifications.
Reviewer 4	I do think this paper provides a solid technical basis for deriving benchmarks using field derived SSDs and the derivation of HC_{05} and XC_{95} values.	Thank you for your comment. No response needed.

TABLE 1. GENERAL IMPRESSIONS (continued)

Reviewer	Comment	EPA Response
Reviewer 4	My comments focus on the need to deal with some of the application issues surrounding these benchmarks. The ability of a State to use this methodology will be related to the quality of their monitoring, assessment, and water quality standards programs (Yoder and Barbour, 2009). Reference to the critical elements that monitoring programs should have to accomplish this methodology should be a part of this document. Again, more detail will be provided below.	In response to this comment EPA included a reference to EPA's technical assistance document, *Biological Assessment Program Review: Assessing Level of Technical Rigor to Support Water Quality Management* (U.S. EPA, 2013), which describes critical technical elements of a robust biological assessment program (e.g., taxonomic resolution, sample collection, sample representativeness, sample processing, data management, and professional review) (see Section 3.1.1.3 in the draft method).
Reviewer 5	For the most part, the document is well written—although with some exceptions that I have noted below. I was impressed with the thoroughness of the presentation. I found most of the substantive information in the document to be accurate. I did find some errors, and I differ with some of the interpretations.	Thank you for your comment. See responses below.
Reviewer 5	I find language used to express essential concepts as problematic, especially the manner in which the term "conductivity" is used to express what is more widely described as "specific conductance". If EPA persists with its current use of the term "conductivity" within the context of the program proposed, the term "conductivity" would then have two different meanings: electrical conductivity (the raw measure) and the 25°C-temperature-corrected value. This result can only breed confusion. I strongly encourage EPA to use the words "conductivity" and "specific conductance" (SC) in accord with well-recognized and widely used precedents and practice.	EPA used the abbreviation SC as suggested by the reviewer. This document uses specific conductivity to measure ionic concentration rather than as an electrical property of water. As ionic concentration increases, SC increases. Both specific conductivity and specific conductance are often used synonymously in the open literature indicating normalization or measurement at 25°C. Conductivity is a property of water expressed as µS/cm. Conductance of a sample or electrical component is measured as S. All measurements in this document refer to specific conductivity measured as µS/cm at 25°C.

TABLE 1. GENERAL IMPRESSIONS (continued)

Reviewer	Comment	EPA Response
Reviewer 5	Given the current status of scientific knowledge concerning major-ion effects on aquatic biota in the Appalachian coalfield, I see the primary method presented by the document and illustrated by Case Studies 1 and 2 as generally adequate, as a temporary measure, for describing specific conductance (SC) levels that will be protective of 95% of benthic macroinvertebrate taxa in Appalachian coalfield streams. I say "temporary" given the lack of scientific certainty concerning the precise nature of the stressor that is causing the effects that are being observed so widely. I expect that the stressor will be identified with greater certainty, eventually, and that will allow more precise targeting by regulatory and other management actions.	EPA agrees that the method is adequate and the Agency will consider new science as it becomes available. EPA has added recent citations for artificial stream studies with artificial salt solutions which have shown effects from the major ions, which are measured as SC (e.g. Clements et al., 2016; Nietch, 2014).
Reviewer 5	I do have some technical concerns which are described in the responses. Many of my technical responses are focused on what I see as inadequate biological confirmation of results which are obtained from analysis of large datasets. While the conclusions derived from such analysis would likely be considered as adequate if expressed with appropriate caveats in the context of academic studies, these results are proposed for application to individual situations of widely varying circumstances as a regulatory program.	It is not clear to EPA what exactly the reviewer is referring to by "biological confirmation of results." EPA assumes the reviewer is referring to the criterion maximum exposure concentration (CMEC) method, an analysis based on stream chemistry data, which is supported by biological validation. Specifically, the ten most salt-intolerant genera were rarely present where the CMEC was exceeded, thereby providing biological validation (see Appendices A and B in the public review draft).

TABLE 1. GENERAL IMPRESSIONS (continued)

Reviewer	Comment	EPA Response
Reviewer 5	Clearly, aquatic communities are depressed in waters influenced by mining throughout the Appalachian coalfield. Given that numerous studies have found close associations between elevated SC/major ions from mining and alterations of benthic macroinvertebrate community metrics, it is reasonable to expect that some effect on water by surface coal mining is playing a major role. Given the number of studies that have found negative associations between elevated SC and benthic macroinvertebrate conditions in the Appalachian coalfield, and the lack of relevant studies that have failed to find such effect, release of SC/major ions has to be considered as prime suspect. However, the direct causal agent—i.e., the precise combination of ions and/or SC-associated stressor such as, perhaps, specific ion combinations or ratios, mining-induced hydrologic alterations, or other unstudied factors—is not known. Hence, it is my view that any public policy actions taken should recognize that the science defining causation is not yet settled.	Definitive artificial stream studies with artificial salt solutions have shown effects from the major ions (Clements et al., 2016; Nietch, 2014) which are measured by SC. Furthermore, EPA developed a causation assessment which was favorably reviewed by the EPA Science Advisory Board (SAB) (EPA 2011a) and peer reviewers of the journal *Environmental Toxicology and Chemistry* (Cormier and Suter, 2013; Cormier et al., 2013).
Reviewer 5	Hence, I see the following statement from the document's forward as appropriate: "State and tribal decision makers would retain the discretion to adopt approaches on a case-by-case basis that differ from those described in this draft document, even if the method in this document is issued under CWA section 304(a)."	EPA's document provides draft methods to assist states and tribes in the development of water quality criteria and other tools to protect aquatic life from effects of elevated ionic concentration as measured by specific conductivity (SC) in flowing waters. States and tribes planning to develop water quality criteria for SC may consider using alternative, scientifically defensible methods. While this document reflects EPA's assessment of the best available science regarding ambient concentrations of SC in flowing waters that protect aquatic life, it is not a regulation and does not impose legally binding requirements on EPA, states, tribes, or the regulated community, and might not apply to a particular situation based upon the circumstances.

TABLE 1. GENERAL IMPRESSIONS (continued)

Reviewer	Comment	EPA Response
Reviewer 5	I have been conservative in my interpretations due to recognition that the document's content has the potential to become a regulatory program. Given the consequences of the potential regulatory actions that may be based on HC_{05} values derived as described here, ensuring those values' validity across the full range of resources targeted for application requires additional biological confirmation. The HC_{05} values are being defined by SC values associated with small numbers of taxa ("limit-defining taxa"). Those values should be checked by conducting additional analysis to determine if the limit-defining taxa occur throughout the resources being proposed for application. For example, the ecoregions used for the examples of Chapters 5 and 6 extend over considerable distances in the north-south direction, and climatic differences can be expected to occur throughout such ecoregions. Do the limit-defining taxa occur only within one portion of the ecoregion, or throughout? Such logic can also be applied across stream orders, and across other dimensions that define the extent of water resources proposed for application. Given the potential consequences of a regulatory program, I see the additional assurances that would be provided by such confirmations as prudent and essential.	While the draft document reflects EPA's assessment of the best available science regarding ambient concentrations of SC in flowing waters that protect aquatic life, it is not a regulation and does not impose legally binding requirements on EPA, states, tribes, or the regulated community, and might not apply to a particular situation based upon the circumstances. The geographic distributions of genera were mapped during the development of the case studies and for different stream orders. Maps of the distributions of genera are available in Appendix E of the draft document.

TABLE 1. GENERAL IMPRESSIONS (continued)

Reviewer	Comment	EPA Response
Reviewer 5	I have concern with the document's attempt to nationalize an approach developed in Appalachia where issues concerning elevated SC are well studied. Given the lack of understanding that concerns the causal mechanism, given that ions at issue are released to waters due to both natural processes and anthropogenic activities, and given that relationships of aquatic biota to SC/major ions are not well documented in other areas of the country (at least to my knowledge), I am unable to reach the conclusion that the proposed method—and its reliance on either SC or [HCO_3^- + SO_4^{2-}] as a measurement endpoint—could be implemented without unanticipated problems in other areas of the country.	Based on laboratory, field, and mesocosm studies (described in the draft document and in Cormier et al., 2013 and EPA 2011a), the mechanism of action appears to be interference with ion regulation and more recent studies continue to explore the variation among species. Natural versus anthropogenic concentrations can be distinguished empirically and by geophysical models. Ionic characterization has been modeled (e.g., Griffith et al. 2015, Olson and Hawkins, 2012). Twenty-three ecoregions were analyzed to better characterize the relationship between aquatic biota and changes in major ion concentration measured as SC. These analyses are provided in Appendix D of the public review draft. These analyses show that when SC increases above background a predictable number of genera are extirpated.
Reviewer 5	I reach my conclusions concerning the method's adequacy reluctantly due to several related concerns: • Concern for the effect that a water quality criterion for (SC) would have communities throughout the Appalachian coalfield and the people who live there, given the historic and recent importance of coal mining as an economic activity that brings money into region. The economic and human effects of recent coal-mining declines in these communities are severe, and implementation of a ~300 µS/cm water quality criterion would continue that trend.	While the draft document reflects EPA's assessment of the best available science regarding ambient concentrations of SC in flowing waters that protect aquatic life, it is not a regulation and does not impose legally binding requirements on EPA, states, tribes, or the regulated community, and might not apply to a particular situation based upon the circumstances.

TABLE 1. GENERAL IMPRESSIONS (continued)

Reviewer	Comment	EPA Response
Reviewer 5	• Concern for "regulatory equity." or a lack thereof in this case. As I understand Clean Water Act implementation procedures elsewhere in the U.S., such as urban, agricultural, and residential areas: Regulatory procedures intended to enforce maintenance of 95% of reference taxa in local streams and rivers are not in place, as the multimetric indices that are commonly used for biomonitoring and bioassessment are developed on a different basis.	It is unclear what the reviewer means by "regulatory equity." While the draft document reflects EPA's assessment of the best available science regarding ambient concentrations of SC in flowing waters that protect aquatic life, it is not a regulation and does not impose legally binding requirements on EPA, states, tribes, or the regulated community, and might not apply to a particular situation based upon the circumstances.
Reviewer 5	I express these concerns with expectation that a ~300 µS/cm criterion, if established as a firm limit in the Appalachian coalfield, would fail to incentivize further development and implementation of the mining and reclamation technologies that are intended to reduce mining environmental impacts and improve environmental restoration—the incentive would be to shut the mines down.	While the draft document reflects EPA's assessment of the best available science regarding ambient concentrations of SC in flowing waters that protect aquatic life, it is not a regulation and does not impose legally binding requirements on EPA, states, tribes, or the regulated community, and might not apply to a particular situation based upon the circumstances.
Reviewer 5	I also have concern for environmental quality in the Appalachian coalfields; that concern is informed by recognition that regional ecosystems are among the richest (biologically) and well-preserved non-tropical ecosystems on the face of this earth; and that the scales of mining operations and mining effects are large and growing.	EPA's draft field-based methods are not limited to Appalachian ecoregions. EPA tailored these methods to enable derivation of ecoregional criteria for specific conductivity anywhere in the United States.
Reviewer 5	With all of that said: I have reviewed the document objectively and have endeavored to provide my professional and technical opinions without bias.	No response needed.

TABLE 2. RESPONSE TO CHARGE QUESTIONS 1−3: *DATA SET CONSIDERATIONS*

Charge Question 1: Matrix Characterization. The ionic composition of water samples represented in the Case Study data sets was dominated by the cations calcium (Ca^{2+}) plus magnesium (Mg^{2+}) and the anions bicarbonate (HCO_3^-) plus sulfate (SO_4^{2-}) ions (see Sections 4.1.3, 5.1.3, and 6.1). The Case Study example criteria are derived for an ionic mixture dominated on a mass basis by $[SO_4^{2-}] + [HCO_3^-] > [Cl^-]$. Please comment on when it is appropriate to remove samples from the data set (e.g., ionic mixtures not represented in the data set, or based on physiological rationales). Is it more appropriate to use all the data and note the conditions that are represented by the data set used to derive the criterion? Please comment on adequacy of the discussions and data analyses provided prior to deriving the Case Study example criteria for $[SO_4^{2-}] + [HCO_3^-] > [Cl^-]$ on a mass basis and estimating background conductivity to assess geographic applicability (e.g., are different or no data exclusion thresholds more appropriate?)

Reviewer	Comment	EPA Response
Reviewer 1	If NaCl in stream water is from natural sources, it would be appropriate to exclude those samples dominated by Cl. However, if it is clearly from human activities, such as road de-icing, exclusion of the samples will make the conductivity criterion derived not applicable to NaCl contamination, a major stressor in streams of the snow zone. It seems to make sense to include all sampling sites where $[SO_4^{2-}] + [HCO_3^-]$ is **naturally** greater than $[Cl^-]$.	EPA recommends performing a case-specific assessment when the ionic composition differs from the example cases presented in the public review draft.
Reviewer 2	I believe it is appropriate to remove samples (data) from the data set which might move the results from reflecting the true condition. The more “types” of data included in a database, outliers, for example, the more general/less specific will be the results—and therefore, less accurate.	EPA avoided unnecessary removal of samples except where specifically described and justified. See Section 3.1.1 of the public review draft.
Reviewer 2	The question for this study was “how to derive example criteria for conductivity for flowing waters dominated by <u>calcium, magnesium, sulfate and bicarbonate ions</u>” (pg. xvi), and <u>not</u> for flowing waters dominated by <u>chloride</u> ions. All of these ions are predominant throughout the study’s ecoregions because of the geology, physiography, vegetation, animal life, climate, soils, water quality, and hydrology found here. However, calcium, magnesium, sulfate and bicarbonate come from weathering of limestone and dolomite (the geological composition of this region) and are the ions which have the greatest impact on specific conductivity which is the intent of this study. Although chloride ions are also prevalent, the decision to exclude chloride anions is logical and appropriate.	As noted by Reviewer 1, the stressor of concern is the ionic mixture as measured by SC not the composition of the natural water. The highlighted text has been clarified to read, “but not for flowing waters polluted by chloride-dominated salts.”

TABLE 2. RESPONSE TO CHARGE QUESTIONS 1−3: DATA SET CONSIDERATIONS (continued)

Charge Question 1: Matrix Characterization

Reviewer	Comment	EPA Response
Reviewer 2	Additionally, the decision to exclude sample sites with <6 pH is also probably wise, although this is perhaps less definitive. Acidity directly affects conductivity by causing calcium and magnesium to become more mobile with decreasing pH, thus having a clear role in conductivity levels. But the level of its effect and its associated variables—such as temperature—would then also need to be considered, increasing the study's data needs and broadening the question. On the converse, acidic conditions do exist in waters of this geographical region because of anthropogenic influences such as urban stormwater runoff, surface mining runoff, gas/oil extraction waste water, and aerial deposition. And from this standpoint only, there might be adequate justification for its inclusion. However, since <6 pH waters in this study were not large in number, the decision to include or exclude could go either way. Would their inclusion have had much influence? A basic rule of thumb for most scientific studies is "the more specific the testing or measuring, the more specific and accurate will be the results".	The comment reflects a concern for characterizing all causal relationships between water quality and biota. However, the EPA's concern is with characterizing the causal relationship between the dissolved ions and the biota. Therefore, low pH is treated as a confounder, not as a contributing cause. Text was added in Section 3.1.1.2.6 of the document, "There are two common means for reducing the influence of confounders. First, sites with a confounder can be removed from the data set, thus reducing its influence on the XC_{95} estimates and XCD model. For example, EPA removed samples with low pH in the case study examples (see Appendices A.2.3 and B.2.3 of the public review draft). Secondly, the effect of a confounder can be minimized by normalizing the influence of a confounder with appropriate weighting. EPA used this approach to assess the influence of temperature and season in the case study examples (see Appendices A.2.3 and B.2.3). Removing samples from the data set can reduce the number of species or the range of exposures of the stressor of interest, thus affecting the reliability of the estimates. Therefore, it is important to evaluate whether the manipulation of the data set improves the accuracy of the HC_{05}. Each case is different, and professional judgment is required."

TABLE 2. RESPONSE TO CHARGE QUESTIONS 1−3: DATA SET CONSIDERATIONS (continued)

Charge Question 1: Matrix Characterization

Reviewer	Comment	EPA Response
Reviewer 2	Toxicities of ions differ, and keeping the data collection and the subsequent analyses limited to the four ions ensures data results free of the additional variables inherently associated with any additional ions. Any field-based study should limit its parameters of study for this reason. Samples from waters with only the same ionic composition will yield the most representative and accurate results.	EPA agrees that consideration of ionic composition is important when developing aquatic life criteria. See Section 2.2 of the public review draft.
Reviewer 2	The authors point out (pg. 2-11) that the relative concentration of bicarbonate is pH dependent, and that the dominant form of the ion in soil is bicarbonate at circumneutral pH. This gives further justification for limiting collection of samples to waters with >6.0 pH	EPA agrees that exclusion of sites with pH<6 is appropriate because it is a likely confounder. See Section 3.1.1.2.6 in the public review draft.
Reviewer 2	The authors have done an excellent job in discussing the many factors and general background information in Sections 1, 2, and 3. The discussion in these sections is valuable for the reader, and is thorough and clear in its presentation. The background information is presented objectively and will be helpful and adequate for state water quality staff.	Thank you for your comment. No response needed.

TABLE 2. RESPONSE TO CHARGE QUESTIONS 1−3: DATA SET CONSIDERATIONS (continued)		
Charge Question 1: Matrix Characterization		
Reviewer	**Comment**	**EPA Response**
Reviewer 2	If there is concern about the merits of keeping or excluding data, I recommend the question be directly discussed in the two introductory chapters. Even though the authors have discussed the reasons why they excluded chlorine and acidic conditions (actually multiple times throughout the document), a table that straightforwardly addresses the pros and cons could be included. List the pros and cons for excluding chloride and sites with <6.0 pH is my recommendation.	A comment has been added to the introduction regarding the pros and cons for excluding sites. Text was added in Section 3.1.1.2.6 of the document, “There are two common means for reducing the influence of confounders. First, sites with a confounder can be removed from the data set, thus reducing its influence on the XC_{95} estimates and XCD model. For example, EPA removed samples with low pH in the case study examples (see Appendices A.2.3 and B.2.3). Secondly, the effect of a confounder can be minimized by normalizing the influence of a confounder with appropriate weighting. EPA used this approach to assess the influence of temperature and season in the case study examples (see Appendices A.2.3 and B.2.3). Removing samples from the data set can reduce the number of species or the range of exposures of the stressor of interest, thus affecting the reliability of the estimates. Therefore, it is important to evaluate whether the manipulation of the data set improves the accuracy of the HC_{05}. Each case is different, and professional judgment is required.”
Reviewer 2	I agree with the inclusion of all other data, i.e., impaired and high quality streams, all stream sizes, and sampling from all seasons.	Thank you for your comment. No response needed.

TABLE 2. RESPONSE TO CHARGE QUESTIONS 1−3: DATA SET CONSIDERATIONS (continued)

Charge Question 1: Matrix Characterization

Reviewer	Comment	EPA Response
Reviewer 3	I like the approach for using the ionic basis of: $[SO_4^{2-}] + [HCO_3^-] > [Cl^-]$ to develop the initial conductivity criteria. This step alone eliminates any problems due to the potential effect of road salts, especially throughout the Appalachians and the Eastern Seaboard in general. It would be an interesting exercise to run the same analyses with no sample exclusion, and then do a comparison of XC_{95} and HC_{05} for only a few selected sensitive genera and some important benthic assemblages (e.g., EPT). I would assume that these would not be too time consuming, but may be worthwhile if there are ecoregions with lower sample sizes than the very rich data set employed in this report.	There have been many studies indicating that different ion mixtures have different effects. Hence, as a precaution, EPA excluded chloride-dominant samples from the analysis. EPA estimated an HC05 with the full data set and because there are very few sites dominated by chloride ions there was no meaningful effect. In the fish assessment, none of the sites were dominated by Cl. The data sets dominated by chloride ions are too small to perform other analyses.

TABLE 2. RESPONSE TO CHARGE QUESTIONS 1−3: DATA SET CONSIDERATIONS (continued)

Charge Question 1: Matrix Characterization

Reviewer	Comment	EPA Response
Reviewer 3	I would be a little concerned with any fall samples collected during an extreme drought period. If one assumes the normal two-component groundwater mixing model for eastern ecoregions, there is the possibility that a severe drought could result in over 95% of the stream flow coming from deep groundwater, and would represent an anomalous case for stream chemistry (successive years of drought may also be a very strong stressor on aquatic biota). It may be best to exclude any sample pairs (biota X conductivity) collected where gaged stream flows in a watershed, or a series of watersheds, dropped to below the 5th percentile of long-term flow records. Also, any exceptionally high-water events (greater than 99th percentile or perhaps 100−500 year storm events) may need to be considered if they occurred in the year before sampling. Benthic assemblage recovery (as cited in the report using the classic paper by Wallace, 1990) may take more than one year, depending on the species complex present in the stream and nearby refugia. Over my career, I learned quickly that there is no such thing as a normal year, and benthic and fish field collections need to be correlated with antecedent climatic conditions (e.g., temperature, flow, etc.).	The reviewer describes a situation where drought or flood years can affect the relative proportion of ground and surface flow. The implication is that the ionic composition and the biota would change. If the reviewer is concerned about these events changing the composition of the ionic mixture, this is not believed to be a problem because sites dominated by chloride were few, and they were removed from the data set prior to analysis. If the reviewer is concerned that these events change the concentrations of ions (i.e., the SC), that is part of the season-to-season and year-to-year variability, which the EPA acknowledges. If the reviewer is concerned that droughts and floods may remove organisms, the method is resistant to this influence because it uses presence, not abundance or presence/absence. Droughts and floods do not affect the ability to measure SC where a genus is present. It may reduce the overall probability of observing a genus in the region due to floods and droughts, but the XC_{95} value is not weighted by a genus's absolute or relative abundance in the region so removal by floods and droughts would not be expected to have much if any effect on the estimation of XC_{95} values for SC so long as some non-drought years are included.
Reviewer 3	Not being very familiar with the water chemistry of western streams, I believe it may be important to think seriously about any exclusionary criteria for these lotic systems. However, I know that some of the mid-western and western states have good data bases with which to run the same analyses as done for ecoregion 69.	The reviewer reiterates that the composition of the pollutant ionic mixture needs to be taken into account and may differ in different parts of the country. Characterizing the ionic composition of natural waters and the pollutant are part of the method.

TABLE 2. RESPONSE TO CHARGE QUESTIONS 1−3: DATA SET CONSIDERATIONS (continued)

Charge Question 1: Matrix Characterization

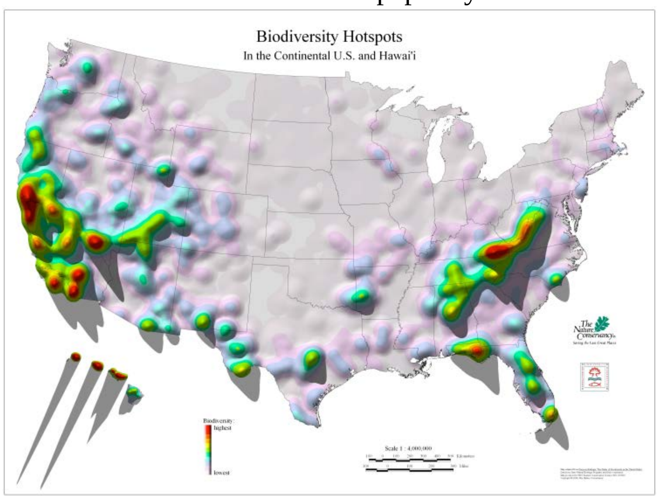
Reviewer	Comment	EPA Response
Reviewer 4	I have no problem with removing sites from the analyses where different ionic mixtures are likely to confound results (e.g., $[Cl^-] > [SO_4^{2-}] + [HCO_3^-]$) or where other stressors (e.g., pH <6) may also contribute to confounded results. I do have some question on whether there could some other confounding caused by: 1) natural variation in conductivity at "reference" sites, and 2) variation in conductivity along a gradient of sites that may be considered "reference" in the sense of "least impacted" conditions vs. "minimally disturbed" in the sense described in the paper by Stoddard et al. (2006). 	The reviewer indicates that reference sites may have different ionic composition. Although, this is true in some cases, the criterion is for the pollutant ionic mixture which is defined. The reviewer is also concerned with variation in terminology with respect to reference. The terminology has been adopted from Stoddard et al., 2006. The terms minimally affected, and least disturbed are used. That issue is now specifically addressed in Appendix C of the public review draft.

TABLE 2. RESPONSE TO CHARGE QUESTIONS 1−3: DATA SET CONSIDERATIONS (continued)		
Charge Question 1: Matrix Characterization		
Reviewer	**Comment**	**EPA Response**
Reviewer 4	It is clear that there are clearly natural biodiversity “hotspots” in the ecoregions examined here (see example Nature Conservancy map). Some of these hotspots may partly be remnants of where biodiversity has been minimally disturbed by human activities, but some are where there is a combination of natural features (e.g., habitat, gradient, elevation, water chemistry) that combine to maximize biodiversity. My concern is that these natural “hotspots” may well be driving the XC_{95}/HC_{05} value particularly when aquatic life use potential is defined by a single aquatic life use, and therefore a single benchmark is derived. The effect of a single benchmark is that it may be under-protective of the most unique “hotspots” but overprotective of more typical habitats. I will address this comment more specifically below.	The reviewer expresses a concern regarding the effect of biodiversity hotspots, but does not explain why hot spots would be biased with respect to salt tolerance. SD models represent the distribution of salt-intolerance of communities, regardless of the number of taxa, so the occurrence of a large number of species is not inherently biasing. The reviewer’s concern may be more complex. It might be hypothesized that biodiversity hot spots would influence the results if the additional species are uncommon (<25 occurrences) and if uncommon species were biased in their salt-intolerance. We addressed this hypothesis by performing an experiment with the data. The requirement for calculation of an XC_{95} of at least 25 occurrences of a genus in a data set restricted those genera that were included in the SD to common genera. More than 300 genera did not meet the 25 minimum and were not included in the SD. However, during the development of the method originally described in EPA 2011, the maximum SC levels at which these rarer genera were observed were calculated as surrogate XC_{100} values. These values spanned the full SD indicating that the model with genera with >25 occurrences reasonably represented the full complement of genera. Because they were not biased to either low or high tolerance to SC, the HC_{05} was not substantially changed by including their XC_{100} values in the SD. More fundamentally, the method minimizes the potential to apply a criterion to ecoregions with different biodiversity by recommending derivation within a defined ecoregion.

TABLE 2. RESPONSE TO CHARGE QUESTIONS 1−3: DATA SET CONSIDERATIONS (continued)		
Charge Question 1: Matrix Characterization		
Reviewer	**Comment**	**EPA Response**
Reviewer 5	Data from sites with elevated TDS/SC but with ionic composition that differs from the dominant ion matrix (i.e., dominated by Ca^{2+}, Mg^{2+}, HCO_3^-, SO_4^{2-}), should be excluded from the datasets used for the analysis, as described by the document. Scientific literature is clear in demonstrating that the ionic composition of TDS influences the SC/TDS concentration at which toxic effects are observed (Mount et al., 1997). Scientific literature is also clear in documenting that that Ca^{2+}, Mg^{2+}, HCO_3^-, and SO_4^{2-} are the predominant dissolved ions in most Appalachian coal-surface-mine influenced waters (Bryant et al., 2002; Pond et al., 2008; Fritz et al., 2010; Timpano et al., 2011; Agouridis et al., 2012; Bernhardt et al., 2012; Lindberg et al., 2012; Wood and Williams, 2013; Pond et al., 2014; Sena et al., 2014).	EPA agrees with the reviewer that sites dominated by pollutants of chloride salts should be excluded.

TABLE 2. RESPONSE TO CHARGE QUESTIONS 1−3: DATA SET CONSIDERATIONS (continued)		
Charge Question 2: Catchment Size. All data from the example criterion data set that met selection criteria were included in the analyses used to derive the Case Study example criteria regardless of stream size. The confounding analysis in the EPA Benchmark Report and additional analyses provided in Section 3.6.2 (Waterbody Type) of the current draft document indicated no scientific reason to exclude data from streams with large catchment areas (>155 km^2) primarily because sensitive genera were documented in these large streams, background conductivity estimates were sufficiently similar, and the ionic mixture was the same (dominated by sulfate plus bicarbonate anions). Do the analyses and discussions provided in the aforementioned section provide adequate support for the decision to include all samples regardless of catchment size? If not, please describe additional analyses and/or discussions needed or identify any shortcomings in the current analyses and/or discussions.		
Reviewer	**Comment**	**Response**
Reviewer 1	The analyses and discussions provided are adequate for the decision to include all stream samples regardless of catchment size. But, can the criteria developed be applied to great rivers, like Mississippi, Ohio, and Colorado rivers? These rivers support very different aquatic fauna, likely fewer sensitive genera, but some unique ones. If no large-river samples are included, could the criterion derived protect those unique taxa? Or, may the criterion be over-protective of large rivers?	These great rivers were not included in the analyses and therefore additional verification and analyses may be warranted beyond what was done in the case study derivation which did include large rivers. Text in Section 3.6.2 in the public review draft was added to make that point clear: "However, professional judgment is warranted when applying the example criteria to streams crossing ecoregional boundaries and stream catchments draining >1,000 km^2, because they are less well represented in the data sets (see Figure 3-8 in the public review draft). For example, great rivers such as the Ohio and Mississippi Rivers were not represented in the data set, and they cross many ecoregional boundaries."

TABLE 2. RESPONSE TO CHARGE QUESTIONS 1−3: DATA SET CONSIDERATIONS (continued)		
Charge Question 2: Catchment Size.		
Reviewer 2	It is excellent that all stream types and sizes were included in the data sets for the case studies, especially the smaller and intermittent streams. Smaller streams, both perennial and intermittent, are where valuable macroinvertebrate habitat is most often found. These are likely to have the appropriate streambed composition, rocks and logs for colonization, leaf litter, bank overhangs, and freedom of siltation—which are all crucial for macroinvertebrate life cycles, population abundance and diversity. So often only the larger, perennial streams and rivers are studied. The authors are “right on” when they point out that discharge from headwaters, intermittent and even ephemeral streams ultimately affect downstream stream reaches and rivers. This is often not understood or realized fully by program managers, who are not well versed in stream ecology, and policy makers. Additionally, the authors make an important point in that many macroinvertebrate taxa often use temporary streams for at least a portion of their life cycle. Much of my experience in stream ecology and water quality has been with the smaller streams and it is my belief that their value to the river system and its taxa cannot be over-emphasized. I thank the authors for their recognition of this.	EPA agrees with the reviewer and recommends analyzing the effect of catchment size on the model and documenting the decision, rationale, and supporting analyses for applicable water body types for SC criteria derived using this method. See Section 3.6.2 in the public review draft.
Reviewer 2	Exclusion of data from the larger catchment areas is, however, worthy of a little discussion here. The authors present four good reasons for not excluding them: 1) sensitive genera were found in the larger rivers; 2) inclusion of data from larger rivers did not significantly change the magnitude of the hazardous concentration; 3) Analysis of 3115 sites with drainage areas up to 17,986 sq km showed a very weak (a very weak, indeed!) correlation of conductivity and drainage area; and 4) background conductivity estimates for drainages > 155 sq km were within confidence bounds for establishment of background values.	EPA agrees with the reviewer’s characterization of the reasons supporting inclusion of data from all stream sizes for Case Studies I and II described in the public review draft.

TABLE 2. RESPONSE TO CHARGE QUESTIONS 1−3: DATA SET CONSIDERATIONS (continued)		
Charge Question 2: Catchment Size.		
Reviewer 2	However, the EPA's Benchmark Report initial exclusion of larger streams—because sampling methods might differ for non-wadeable streams—has substantial merit. Sampling methods are indeed different for the larger rivers, and large river sampling requires greater resources (time, staff, boat/equipment) and, therefore, also happens less frequently. Collected macroinvertebrates in larger rivers can be low in numbers as well—due probably to a combination of factors: manmade channel morphology changes, river velocity too high, fewer colonization sites, poor habitat, deposition of sediment, anthropogenic contaminants, and difficulty in sampling at greater depths and velocities. Thus, more variability likely exists in data for macroinvertebrate databases for large rivers. However in this study, sensitive taxa were documented in the larger rivers, so perhaps collection methods and expertise in sampling has improved, but perhaps more importantly, these rivers are likely of higher quality than those here in the Midwest of which I am familiar and which are heavily impacted by agriculture.	EPA agrees with the reviewer's description of issues in large rivers. EPA sampling methods are different (Flotemersch 2006). Also, most large rivers have been chemically and physically modified. The analysis for Ecoregions 69 and 70 demonstrate that some high quality (low conductivity) large rivers still exist in the area. However, this may not be the case in all areas and so EPA provided methods for assessing criteria applicability of to larger systems. These analyses appear in sections 3.6.2. EPA recommends analyzing the effect of catchment size on the model and documenting the decision, rationale, and supporting analyses for applicable water body types for SC criteria derived using this method. See Section 3.6.2 in the public review draft. Flotemersch, J. E., J. B. Stribling, and M. J. Paul. 2006. Concepts and Approaches for the Bioassessment of Non-wadeable Streams and Rivers. EPA 600-R-06-127. U.S. Environmental Protection Agency, Cincinnati, Ohio.
Reviewer 2	In conclusion, the authors have provided good discussion and support for the decision to include all samples regardless of catchment size. A bit more discussion as I have presented here might be helpful but probably is not necessary.	Thank you for your comment. No response needed.
Reviewer 2	Lastly, I wish to reiterate the value of data from intermittent and ephemeral streams. These small streams provide irreplaceable habitat for macroinvertebrates, invertebrates, amphibians, aquatic/wet terrestrial species of all kinds. Their loss has been significant through ditching and tiling in agriculture, diverting and damming for irrigation, and in placing into underground pipes in urban development.	EPA agrees that these types of streams are ecologically important. Available information from the open literature indicates that many of the macroinvertebrate taxa persist in intermittent and perennial channels, albeit at different densities and for varying amounts of time. See Section 3.6.2 of the public review draft.

TABLE 2. RESPONSE TO CHARGE QUESTIONS 1−3: DATA SET CONSIDERATIONS (continued)		
Charge Question 2: Catchment Size.		
Reviewer 3	My personal opinion, and scientific bias, is to use only data from wadeable streams—this is the critical field design driver for stream assessment with EPA, and many of the eastern States and NGOs. EPA, along with many States, did a lot of work on developing such protocols to assure that there was robust physical, chemical, and biological data collected in order to make non-biased estimates of many important parameters. Indeed, many key biotic and habitat metrics were developed based solely on wadeable streams. Also, 1st through 3rd order streams may constitute 70−90% of stream km in an ecoregion, with larger streams (4th to 12th order) representing less than 10−30% of stream km.	EPA disagrees that the field-based method should be limited to wadeable streams. Samples used in the analyses in the public review draft were collected using wadeable stream methods which were independent of stream order. Stream size did not substantially change the HC_{05}, and 25 of the 30 most salt-intolerant taxa were collected from rivers with large drainages. EPA recommends analyzing the effect of catchment size on the model and documenting the decision, rationale, and supporting analyses for applicable water body types for SC criteria derived using this method. See Section 3.6.2 in the public review draft.
Reviewer 3	If one follows the River Continuum Theory, the 1st through 3rd (and perhaps some small 4th) order streams are where the real action is, and that the larger streams and rivers (large 4th to 5th and higher) start to reflect a major change in both ecological structure and function. OK, so one may collect some benthic organisms (genus may be the same, but probably different species) in the larger order streams that would also be found in lower order streams. However, stream processes in the larger order streams are so different I feel it would be difficult, and unjustifiable, to use this approach. Obviously, EPA would welcome this opportunity to be able to set conductivity criteria for large aquatic ecosystems (large stream and rivers), especially in light of the NPDES permits, etc.	EPA has chosen to limit the cases to the stream sizes included in the data set and to allow flexibility in the method to accommodate large differences between stream order and flow characteristics of different types of streams. Text was added to Section 3.6.2: "For example, great rivers such as the Ohio and Mississippi Rivers were not represented in the data set and they cross many ecoregional boundaries." However, the applicability of the method does not depend on "stream processes" or "where the action is." The method is based on protecting 95% of the aquatic community. Analyses showed that the few large rivers (based on drainage area, not Strahler order) that still have low SC have salt-intolerant taxa; 25 of the 30 most salt intolerant genera based on derived XC_{95} values for Ecoregions 69 and 70 (see Appendix E) were documented in these larger rivers (see also Appendix B in U.S. EPA, 2011a).

TABLE 2. RESPONSE TO CHARGE QUESTIONS 1−3: DATA SET CONSIDERATIONS (continued)		
Charge Question 2: Catchment Size.		
Reviewer 4	My concern with this discussion is not so much with stream size as important variable, but other natural classification issues and some anthropogenic changes that might have occurred from human habitation and land disturbance that are not acute or readily controllable and are within a definition of "least impacted" streams. For example, in the mountainous regions of the WAP ecoregion in West Virginia for example, the relief has led to land uses (e.g., forestry, park, light agricultural, low density residential) that result in more highly forested (>90%) reference conditions. Along the edge of the WAP ecoregion in Ohio for example, the relief is more variable and farming and some other land use changes are somewhat more intense. "Least impacted" reference sites are much less likely to be ">90% forested." This broaches the important question of whether a single benchmark or multiple benchmarks to match tiered uses are more appropriate.	On a site-specific basis, the example criteria developing using the draft methods could be adjusted or recalculated to protect important species, highly valued aquatic communities, or specially protected waters.
Reviewer 5	My discussion below assumes that reference streams are low-order, small-drainage-area streams.	Reference streams were not defined by stream order for the methods because some reference streams were very large. See response below.
Reviewer 5	No, the document does not provide adequate basis for including all observations, regardless of stream size. There should be a stream size cutoff, and EPA should provide guidance on an appropriate cutoff. One factor in defining the stream-size range appropriate for the analysis concerns reference sites. As stated on page 2-1, line 23−24: "Genera that are not observed at reference sites … are excluded from the data set." Therefore, only streams of size classes where community compositions can be documented as being similar to those at reference sites should be included; or, only taxa found to be both occurring at reference sites and as characteristic of the higher-order streams should be included.	The reviewer recommends EPA provide a river size cut-off. EPA does indicate in the public review draft that the tested range of drainage area is the limit of the reported values, but notes that the method may also be appropriate in larger systems. Some reference streams are very large; for example, 25 of the 30 most salt-intolerant genera occurred in the New River in WV, drainage area 11,800 to 17,985 square kilometers. Secondly, the derived HC05 is not based on a particular community of taxa, but on a general model of the proportion a taxa affected by different concentrations of ions. See also response below.

TABLE 2. RESPONSE TO CHARGE QUESTIONS 1−3: DATA SET CONSIDERATIONS (continued)		
Charge Question 2: Catchment Size.		
Reviewer 5	There is a large volume of scientific literature supporting the understanding that aquatic communities and community compositions vary by stream size (e.g., Vannote et al., 1980). Scientific literature documents the taxonomic differences that occur between in the river continuum which extends from headwater (low-order) streams and the higher-order streams commonly known as rivers. For example, Grubauch et al. (1996) refer to the "rapid faunal replacement" that occurs in the mid-order reaches of an Appalachian river continuum; and they cite other studies with similar findings.	EPA evaluated the effect of stream size on the HC_{05} in the SC benchmark (EPA 2011a) and the current document (see Section 3.6.2). The analysis shows a very weak correlation between specific conductivity and drainage area and supports inclusion of data from all stream sizes in the data set for example criteria derivation. HC_{05} is consistent across stream size. Communities do change from place to place, large to small, warm to cold, but if the natural background is low, then salt-intolerant species occur in them. New analyses in Appendix D were added to improve applicability methods and demonstrated species adaptation to low SC systems.
Reviewer 5	The proposal to include both large-stream and small-stream (headwater stream, low-order stream) observations in the analysis dataset is not well supported by the logic in the paragraph starting on page 3-31, line 15.	EPA evaluated the effect of stream size on the HC_{05} in the SC benchmark (EPA 2011a) and the current document (see Section 3.6.2). The analysis shows a very weak correlation between specific conductivity and drainage area and supports inclusion of data from all stream sizes in the data set for example criteria derivation. HC_{05} is consistent across stream size. Communities do change from place to place, large to small, warm to cold, but if the natural background is low, then salt-intolerant species occur in them. New analyses in Appendix D were added to improve applicability methods and demonstrated species adaptation to low SC systems.

TABLE 2. RESPONSE TO CHARGE QUESTIONS 1−3: DATA SET CONSIDERATIONS (continued)		
Charge Question 2: Catchment Size.		
Reviewer 5	The first argument cited by the paragraph concerns Ephemeroptera taxa in large streams and cites Appendix B of U.S. EPA (2011a) which states that Ephemeroptera occur at lesser richness in large streams with elevated SC than in large streams with low SC. This fact, in and of itself, is peripheral to the logic proposed by this document which concerns frequencies of occurrence by reference-site taxa. Appendix B (U.S. EPA, 2011) does not document that the relevant Ephemeroptera taxa—those occurring in larger streams with low SC but not occurring in larger streams with high SC—are taxa that also occur at reference sites.	There are more salt-intolerant Ephemeroptera genera in rivers with low SC than at high SC as shown in the assessment of causation (Appendices A and B of the Benchmark Report, EPA 2011a; Cormier et al., 2013). No Ephemeroptera were excluded from the analysis because they all occurred at reference sites. The list of non-reference genera are provided in the EPA 2011a report and in the case studies.
Reviewer 5	Even if it did, that additional fact would not provide full support because it does not document that the taxa composition high-SC high-order streams are altered in a manner that exceeds the 5%-of-reference-taxa loss threshold. If both high- and low-order streams are to be included in the analysis dataset, only taxa observed as characteristic of both high- and low-order streams should be considered in the analysis. If conducting such analysis, the finding that a given taxon occurring at reference sites is also characteristic of high-order streams, should be based on more than a single occurrence by following the logic of the so-called extirpation concentration defined as the 95th percentile of capture probability and not as the maximum SC for observed occurrence. This precaution is justified by these organisms' mobility.	Very few genera were excluded from the analysis because they did not occur at reference sites, and their exclusion had little effect on the HC_{05} (Cormier, S. M., 2015. Field-based Methods for Developing Water Quality Benchmarks Invited Expert Meeting on Revising U.S. EPA's Guidelines for Deriving Aquatic Life Criteria 14-16 September 2015, Arlington, VA). EPA has added the information that 25 of the 30 salt-intolerant genera are present in large rivers. The SD is a model of the proportion of genera affected by increasing SC, not of any particular set of genera. Suter and Cormier (2013) showed that genera with low XC_{95}s are present in streams with low SC and absent in streams with high SC regardless of drainage area. The work called for by the reviewer was previously published by EPA (2011a); this material was provided as background material to all reviewers.
Reviewer 5	The document (paragraph starting on page 3-31, line 15) also states that "conductivity and drainage area are very weakly correlated" within the areas studied. This fact is not of direct relevance to the argument that biological data from rivers and headwater streams should be mingled within datasets that are analyzed using the methods described. The use	EPA did not use information from reference sites except to identify genera that never occurred at them or to confirm estimates of natural background SC. The background SC for rivers was presented to show that

TABLE 2. RESPONSE TO CHARGE QUESTIONS 1−3: DATA SET CONSIDERATIONS (continued)		
Charge Question 2: Catchment Size.		
	of "background" SC (i.e. 25th percentile) to approximate reference condition for large streams does not help the logic, in my view; "background" SC and "reference condition" are different concepts, with the "reference condition" concept as more restrictive.	rivers do not necessarily have higher SC than lower order streams.
Reviewer 5	The document also states that "Inclusion of the data from large streams did not significantly change the magnitude of the HC_{05}". That statement is supported by citing Suter et al. (2011), which is a conference presentation and not a peer-reviewed manuscript that is accessible to reviewers, potential regulatory commenters, etc.	The text reads in section 3.6.2 on page 3-33, "Inclusion of the data from large streams did not significantly change the magnitude of the HC_{05} (289 µS/cm) compared to the HC_{05} without data from larger systems (295 µS/cm) (Suter et al., 2011)." The reference does refer to an EPA cleared presentation which was never prepared as a manuscript and is available in another EPA cleared publication that is available on the web (Cormier 2015), but the work can also be considered as being reported here in print for the first time. Therefore, we have removed the citation.
Reviewer 5	Most importantly: The method proposed by this document is novel, as admitted by the authors. However, when applied to headwater streams in coalfield areas, it is being applied in a context where numerous studies have found altered benthic macroinvertebrate communities in low-order streams influenced by major ions discharged by coal surface mines; and no peer-reviewed studies I am aware of have found the opposite to be occurring. A comparable body of supporting science does not exist for the higher-order, high-drainage-area streams.	The reviewer comments that effects are well established for small streams but that there is not a comparable literature for large rivers. EPA has shown that some large rivers have been affected by increased ionic inputs based on this study and U.S. EPA (2011).

TABLE 2. RESPONSE TO CHARGE QUESTIONS 1−3: DATA SET CONSIDERATIONS (continued)

Charge Question 3: Seasonality. The data sets used in Case Study I and II did not employ weighting to account for seasonal effects. While the vast majority of samples were taken once on an annual basis, further analyses indicated that the effects of seasonality on the example criteria were minor (see Sections 4.1.3 and 5.1.3). Do the analyses employed for seasonal effects and corresponding results adequately support the decision not to weight for season? If not, please describe additional analyses and/or discussions needed or identify any shortcomings in the current analyses and/or discussions.

Reviewer	Comment	Response
Reviewer 1	Annual samples, particularly those collected in summer, likely miss some or many sensitive insect genera. However, as long as sampling time is NOT correlated with conductivity (e.g., sampling high-conductivity sites early, but low-conductivity sites later), this source of error is probably minor compared with other sources, such as selection of sites, sampling variability, and the temporal variation of conductivity. I would examine the correlation between site conductivity and sampling date (Julian Day).	The reviewer concurs with the method. The proposed analysis regarding Julian Day is presented in Appendices A and B and called out in Section 4.4.3 and 5.1.3 of the public review draft. Sampling date had a minor effect on HC_{05}.
Reviewer 2	The data of conductivity concentrations show that they do vary by season. This was addressed by comparing hazardous concentration values by season. “Due to the similarity at the low end of the sensitivity distribution (SD) between spring HC_{05} and HC_{05} of the full dataset” (pg.4-11), it was determined to use all data regardless of month. I question why this wasn’t also done for the fall (especially October) data? Granted, February—April exhibited the most noticeable change but October was significant as well. In Ecoregion 70 from the Watershed Assessment Branch database, September stood-out because it had significantly higher conductivity values (pg. 5-7), as did April with definitely lower values (pg. 5-8), although not as extensive as October’s. The box plot on pg. 5-9 for Ecoregion 70 shows the apparent seasonal variation of July−October.	Three analyses were done: "spring," "summer-fall" and the "full year" data sets. These are shown in Appendix A.2 which is noted at the bottom of the paragraph. September and October were included in the summer fall sample. Some grouping was needed in order to do the analyses. Samples were grouped by when high and low SC periods occur, and the spring group is also the period when salt intolerant taxa occur in samples.

TABLE 2. RESPONSE TO CHARGE QUESTIONS 1−3: DATA SET CONSIDERATIONS (continued)		
Charge Question 3: Seasonality		
Reviewer 2	Reference sites, however, are stated to have conductivity levels "generally low and similar throughout the year although slightly higher in August, September and October," (pg. 5-12). I think it is more than just "slightly" higher! On pg. 5-7, Ecoregion 70, the September values are so much higher that it is difficult for me to understand that the September data doesn't adversely skew the results. Perhaps separate CCC and CMEC for the September timeframe is reasonable. Since my area of expertise is not statistics, I am not really able to investigate this myself and will rely on the authors' determination that seasonal differences do not require weighting, and that the seasonal differences do not alter the results to any great degree.	It appears that the reviewer confused the plot of the probability samples with the reference samples. The reference sites Figures 4-3 and 5-3 have lower relative variability when the y-axis scale is taken into account compared to the data sets with ionic inputs (e.g., see Figure 5.2). A call out was added to note the scale of the axes in an attempt to reduce confusion.
Reviewer 2	On pg. 5-6, the conductivity background for Ecoregion 70 is <200 µS/cm December−June, and >200 µS/cm July−October. This seems to be enough of a distinction that perhaps all data need to be divided into two sets, one containing the December−June data and the second, the July−October data. It would seem that this would be sufficient rationale to have this separation but I am presenting this more as a question than a statement.	Weighting made no difference in the HC_{05} because the primary influence is presence of univoltine genera that are only collected in the spring because they are too small to be collected through standard macroinvertebrate monitoring methods at other times of the year. The information is presented in Appendices A and B in the public review draft.

TABLE 2. RESPONSE TO CHARGE QUESTIONS 1−3: DATA SET CONSIDERATIONS (continued)		
Charge Question 3: Seasonality		
Reviewer 2	As a side personal note: Here in the Midwest we have distinct seasons, and many parameters clearly show this in their values. I am accustomed to looking at the seasonal data and its use in planning for monitoring programs and watershed recovery plans. Having this specificity of data is more informative for these purposes than "lumping" or weighting of the data because it provides greater insight as to pollutant sources and causal relationships. For state staff, determination of sources of impairment is usually the overall objective and is frequently difficult to ascertain. Having a clear understanding of what is happening each month (when there is monthly data available) helps to provide insight. With that noted, I fully realize the objectives for those purposes and the objectives for this study are different. But it may be of value to the authors to understand how state WQ staff usually look at data and use it.	Thank you for your comment. EPA has provided a method for evaluating the probability that the criterion would be exceeded based on an annual maximum. It may be possible for states wanting higher resolution to perform the same analysis on a monthly or even shorter time frame.
Reviewer 2	Additionally, with these comments in mind, I must also add that I prefer limiting the amount of weighting when working with a dataset. On pg. 3-18 it is stated that if "the weighted HC_{05} overlap the confidence bounds of un-weighted HC_{05}, the un-weighted model is accepted." This seems to be a logical and accurate decision. Further, it states that in general, "the use of unweighted SDs is easier and requires fewer data points." I agree. Where weighting and manipulating the data can be reasonably minimized, I believe it should be. A balance must be made in the need for normalizing, scaling and weighting and the loss of variations that reflect the actual conditions.	The reviewer recommends using weighting only if necessary. In the examples, weighting by date did not improve the model and thus, EPA did not use weighting in the model.
Reviewer 2	Also pgs. 3-16−3.18, the three approaches to seasonality are given. It is well done.	Thank you for your comment. No response needed.
Reviewer 3	First, see part of the response to question 1. I would be careful including data for extreme stream flow conditions, which may occur in spring (high flows), summer (possible hurricanes), and fall (drought). Care should also be taken to examine any unusual antecedent conditions within watersheds to be studied. In our regional work, we	Because the presence of genera rather than absence in waters with different SCs is the driver of the model, the loss of genera during extreme events is not likely to influence the model. They are simply recorded as absent. Extreme flows do not affect the ability to measure SC

TABLE 2. RESPONSE TO CHARGE QUESTIONS 1–3: DATA SET CONSIDERATIONS (continued)

Charge Question 3: Seasonality

	needed to delete a few 1st and 2nd order sites due to extreme high flow conditions in the previous year that affected two subwatersheds in our study area.	where a genus is present. It may reduce the overall probability of observing a genus in the region, but the XC_{95} value is not weighted relative to other genera if a genus is more or less abundant compared to other genera in the region. Removal by antecedent conditions would not be expected to have much if any effect on the estimation of XC_{95} values for SC.
Reviewer 4	Since the analyses indicated the effects of seasonality are minor, I have no problem with how the paper dealt with this issue.	Thank you for your comment. No response needed.
Reviewer 4	Given the pattern in conductivity in some of the datasets where there are higher values in the late summer (e.g., August−September), a period that corresponds with typical lowest monthly flow periods, it might be of use to discuss how this might influence monitoring for compliance with any derived criteria. For example, the paper talked about a monthly weighting of conductivity values to determine the effect of seasonality on the criteria. If a State only collects data during a summer period (e.g., Aug/Sep) should the values be adjusted to the annual geometric mean to determine whether benchmarks are exceeded?	Some text was added to Section 3.1.1.2 which indicates that seasonal sampling is needed to include sensitive taxa in the SD: "For example, samples taken only in dry seasons when SC tends to be higher would likely bias results toward less sensitive genera and to maximum SC exposures rather than an annual average." It may not be appropriate to develop criteria from summer only sampling because sensitive genera are in a life stage that are unlikely to be collected and therefore would not meet the requirement of having 25 occurrences for inclusion in the SD.
Reviewer 5	I do not agree with the "not weight for season" decision. Research at Virginia Tech (Boehme, 2013; Boehme et al., 2013) has demonstrated that composition of benthic macroinvertebrate samples from coalfield headwater streams varies seasonally, both in reference streams and in streams with elevated TDS originating from mining sources. Other research demonstrates seasonal differences in response by a multimetric index to contemporaneous SC in both reference streams and those affected by elevated SC/TDS (Timpano et al., 2011), meaning that community composition differs by season. Also, the document itself demonstrates clearly that SC in non-reference streams varies by season (Figures 4-2, 4-4, 5-2, and 5-4). Hence, I do not see scientific support	EPA agrees that sampling methods collect larger instars and therefore different genera are collected in different seasons depending on their natural history of emergence and number of reproductive cycles per year or multiple years. It is important to note that the genera are present in the stream year round, but they are collected at different times. The salt-intolerant taxa tend to be captured in the spring and thus EPA has performed several analyses to ensure that streams with higher SC are included in the spring sample. Furthermore, EPA has evaluated the effect of including summer samples and the

TABLE 2. RESPONSE TO CHARGE QUESTIONS 1−3: DATA SET CONSIDERATIONS (continued)		
Charge Question 3: Seasonality		
	for analysis using methods described in the document of data sets that mingle samples from different seasons without a seasonality check, such as a check to determine if limit-defining taxa are seasonal.	results of spring only and all year are about the same. EPA provided options for weighting if states, tribes and territories want to include weighting, but weighting had little or no effect in the cases presented in the draft document. The decision to weight is left as an option, but so far has not been necessary nor advisable for the example cases. The natural histories of aquatic invertebrates are published in a variety of sources. The species are not seasonal, only their collection is seasonal due to size. This was made clear in the draft document.
Reviewer 5	Section 3.14 describes a seasonality check procedure that compares spring, summer, and "all year" samples. That section defines spring as March−June. Elsewhere, the document describes "summer" as July−October, an unorthodox definition of that season. Does that definition of "summer" also apply in Section 3.14, and in the Case Study 1 and 2 seasonal analyses (Figures A-7 and B-7)? Seasonal definitions should be stated clearly.	The reviewer felt that the use of summer/fall or spring were ambiguous terms. EPA has made changes to document so that seasons are referred to as distinct periods, such as lower SC season 1 (March−June) and higher SC period, season 2 (July−October) or any other season that is appropriate for an analyst to use for their region.
Reviewer 5	Seasonal HC_{05} values were developed for Case Studies 1 and 2; spring and summer HC_{05} values are similar for Case Study 1 (Figure A-7) but not for Case Study 2 (Figure B-7). On page 5-12 (Case Study 2), the document states: "In the final assessment, due to the similarity at the low end of the genus sensitivity distribution (SD) between the spring HC_{05} and the HC_{05} based on the full data set, the example ecoregional criteria were derived using all available data, regardless of the time of year they were collected." Based on Figure B-7, I do not see the seasonal HC_{05} values as similar.	In Figure B-7, the all year HC_{05} is slightly less than the spring only HC_{05}. Because all-year data HC_{05} is lower than the spring HC_{05}, the spring value is not biased relative to the all-year value, as would be expected if seasonal differences were influential. The slightly lower all-year HC_{05} is due to the inclusion of more genera in the SD because the data set is larger, not due to seasonality. EPA has added the HC_{05} values for the spring and all-year data sets.

TABLE 2. RESPONSE TO CHARGE QUESTIONS 1−3: DATA SET CONSIDERATIONS (continued)		
Charge Question 3: Seasonality		
Reviewer 5	In conclusion, I see no justification for a procedure that would mingle data from all seasons with no seasonality check or adjustment for the data's seasonal distribution. The Case Study 2 results justify the need for consideration of season. The fact that both community composition and water quality vary by season demonstrate that seasons should be considered separately in HC_{05} development.	Community composition is not a valid argument as explained above. The option and method for adjusting for seasonality is provided if an analyst chooses to use it and is illustrated in two examples in Appendices A and B.

TABLE 3. RESPONSE TO CHARGE QUESTIONS 4−8: *EXAMPLE CRITERIA CALCULATIONS*		
Charge Question 4: Criterion Continuous Concentration (CCC): Please comment on the clarity of the method to derive the XC_{95} and HC_{05} (see Section 3.1, Deriving a CCC).		
Reviewer	**Comment**	**Response**
Reviewer 1	The method of deriving HC_{05} is straight forward and clearly described.	Thank you for your comment. No response needed.
Reviewer 1	However, the method used to estimate the extirpation threshold (XC_{95}) is confusing and problematic. XC_{95} considers neither the direction of response of a genus to increased conductivity nor the relative frequency of the taxon ("probability of capture" in the text), two key factors for inferring extirpation. Therefore, it appears not possible to establish any consistent and meaningful relationship between XC_{95} of a genus and its extirpation.	The reviewer states that the method to derive the XC_{95} is confusing. Taxa that are not declining at the calculated XC_{95} are designated as greater-than values, not as actual estimates of the SC at which they are extirpated. These genera are more salt tolerant and do not significantly impact the HC_{05}. The relative frequency of a taxon was not used because if depends on accurate measurement of both presence and absence. Absence is the lack of observation; it is not directly measured. A genus may be present but not observed due to season, sampling method, or insufficient sampling. In contrast, presence is observed and directly measured. The basis of the XC_{95} value is the presence of a taxon at a SC level which indicates it can survive at that ionic concentration. Presence is the better measurement endpoint for deriving an XC_{95} because it can be directly measured. Clarifying text was added.

TABLE 3. RESPONSE TO CHARGE QUESTIONS 4–8: *EXAMPLE CRITERIA CALCULATIONS* (continued)

Charge Question 4: Criterion Continuous Concentration (CCC)

Reviewer	Comment	Response
Reviewer 1	The authors use a GAM model to refine XC_{95}. That is helpful for those genera negatively affected by conductivity, but the threshold of extirpation for those genera positively or neutrally responding to increased conductivity over the range observed still remains indefinable. For example, XC_{95} of *Cheumatopsyche* (A-29) is estimated to be >3140 µs/cm (A17), while the genus reached its highest "probability of capture" at this conductivity level. Even with a qualifying designation of ">", is this estimate really meaningful? The same designation (>) is also given to those genera that have very different response curves, such as *Cheumatopsyche* and *Leuctra* in Fig. 3-1.	The GAM models are not used to refine the XC_{95}; they are used to caution users that some values are either approximate or greater than the listed value. The assignation of greater than and approximate values does not affect the HC_{05}. These qualifications draw attention to the uncertainty of some XC_{95} values for the benefit of researchers who may choose to use the XC_{95} values for other uses, for example to compare field (e.g., Kunz et al., 2013) and toxicity test results or diagnostically (e.g., Coffey et al., 2014). As noted, the XC_{95} of *Cheumatopsyche* is undetermined having a greater than value and no effect on the HC_{05} except to contribute to total size of the denominator. No change required.
Reviewer 1	When the values of XC_{95} for genera that substantially differ in occurrence frequency and response to conductivity are treated equally, the SD curve is no longer interpretable and potentially misleading, at least in my opinion.	Essentially, the reviewer is pointing out a general issue with SDs in which those taxa on the left of the X-axis are salt-intolerant and those to the right are less salt-intolerant and may not have fully defined XC_{95} values. The reviewer is correct that the entire curve is not a predictive model. However, the entire SD curve is not used to define the HC_{05} – only the lower portion of the curve is used. The 5th centile does not occur in the range of the SD where the XC_{95} values are uncertain. EPA has added the following text to figure legend 3-1: "XC_{95} values that were defined as greater than values are indicated by triangles." Also in section 3.1.2 the text reads: "The assignations of greater than (>) and approximately (~) does not affect the HC_{05}. They are provided to alert users of the uncertainty of the XC_{95} values for other uses such as comparison with toxicity test results or with results from other geographic regions."

TABLE 3. RESPONSE TO CHARGE QUESTIONS 4–8: *EXAMPLE CRITERIA CALCULATIONS* (continued)		
Charge Question 4: Criterion Continuous Concentration (CCC)		
Reviewer	**Comment**	**Response**
Reviewer 1	Two options might be worth considering. First, presumably one can appropriately determine extirpation thresholds (i.e., XC_{95} without > designation) for more than >10% of the genera. If so, he/she may put all other genera in a single category, "indefinitely high". The authors may then use the first group of genera to define HC_{05}. Second, the authors can look at how many genera declined down to ≤1% of the highest "probability of capture" in the max-conductivity bin in GAM models. If more than 10% or 20%, as in their case studies, they should be able to easily determine HC_{05}, leaving out the idea of XC_{95} entirely.	The reviewer suggests two options that have not been vetted in the scientific community. Use of extirpation thresholds was rejected because when the full range of exposures are evaluated, all species have a unimodal distribution so using only genera without designations would be entirely dependent on the range of exposures which varies among data sets and regions (see GAMs for fish species). The second suggestion suffers from regression to the mean and the range of the exposures in the data set as described above. EPA has chosen to retain the method reviewed and approved by the SAB that uses the 5th centile of affected taxa, an assessment endpoint consistent with Agency guidance for aquatic life criteria development.
Reviewer 2	As my knowledge base is centered on biological aspects of rivers and streams rather than statistics, several of my comments will be limited in this regard. I am listing the various thoughts which I had as I went through Section 3.1:	Thank you for your comment. No response needed.
Reviewer 2	• The inclusion of both high quality and impaired sites is correctly done. This provides a well-represented database, covering all levels of conditions and taxa, and at all times of the year. This reflects the variability that will exist because of seasons, habitats, and the effects of manmade influences in the river basin which affects the ionic composition.	Thank you for your comment. No response needed.
Reviewer 2	• The step-by-step explanation on pg.3-1 is helpful. More of this could be done to increase understanding of the calculation processes used in this document.	EPA plans to prepare a step-by-step technical support document to assist users. This document is expected to be published simultaneously with the final version of the field-based methods document.

TABLE 3. RESPONSE TO CHARGE QUESTIONS 4–8: *EXAMPLE CRITERIA CALCULATIONS* (continued)

Charge Question 4: Criterion Continuous Concentration (CCC)

Reviewer	Comment	Response
Reviewer 2	• Improve clarity by explaining how the actual weighting and cumulative distributional function is done on pg. 3-1.	EPA plans to prepare a step-by-step technical support document to assist users. This document is expected to be published simultaneously with the final version of the field-based methods document.
Reviewer 2	• Good explanation on pg. 3-2.	Thank you for your comment. No response needed.
Reviewer 2	• Figure 3-2, pg. 3-4: Gives good general process flow. Is it possible that an actual mathematical example could follow along with each step?	EPA plans to prepare a step-by-step technical support document to assist users. This document is expected to be published simultaneously with the final version of the field-based methods document.
Reviewer 2	• The bullets on pg. 3-5 are thorough and give good support to adequacy of data. Sample size discussion is well done. Sensitivity analysis, which includes a representative proportion of sensitive genera, is well done. Having 90−120 genera and 500−800 sites are large numbers, and are seen throughout this document. This is excellent. It strengthens the development of the criterion, its applicability, and the justification of the concentrations determined. If only all studies could have such numbers!	Thank you for your comment. No response needed.
Reviewer 2	• Bootstrapping needs to be described more fully (for non-statistical readers). While the paragraph on pg. 3-7 is probably adequate for many, there are a considerable number of state agency or other watershed staff who have minimal statistical backgrounds. A few additional paragraphs detailing/giving examples of such exercises as bootstrapping would make the document more usable by the large range of agency staff.	The reviewer recommends more explanation for calculations. An additional paragraph was added regarding bootstrapping.

TABLE 3. RESPONSE TO CHARGE QUESTIONS 4–8: *EXAMPLE CRITERIA CALCULATIONS* (continued)		
Charge Question 4: Criterion Continuous Concentration (CCC)		
Reviewer	**Comment**	**Response**
Reviewer 2	• In reference to pg. 3-9, lines 9-21, care must be taken to avoid too many repeated macroinvertebrate samplings in the same place over the course of a year. Repeated sampling is disruptive to the habitat and can diminish the taxa at the site. Unlike fish species, macroinvertebrates are less mobile, and, if young stages are removed, there may be fewer adults at the sites especially if there are no other small streams in the vicinity to repopulate.	To caution against oversampling, EPA added the following sentence, "Annual sampling is generally sufficient and avoids damage to the habitat and stream biota."
Reviewer 2	• I would like to see a little more specificity in describing sampling methods. Is there assurance that there was a standardized field sampling protocol observed for all biological sampling? It is important that all sampling crews used the same techniques. It is more of a problem between jurisdictions (states, cities, or private organizations which do monitoring) but can also occur within an agency. It is vital that, for example, an equal number of sweeps of the catch net are made at each site, or, the same number of individual samples comprise a composite.	Standardized methods are needed when the measured attribute of two or more treatments or sites are compared. That is not the case in this analysis. It is the presence of a species in water of a particular SC that is relevant for determining their ability to survive at that ionic concentration. Failure to observe, even when present but not collected, has little or no effect on the XC_{95} or HC_{05}. Variation in the sampling method would affect the HC_{05} only if it were biased with respect to SC; that is, if sampling were performed by one method at sites where salt-intolerant taxa occurred and with another method where they were absent. This method is very different from methods that are designed to compare the biological composition among sites. Bias could be introduced if the sampling effort was different at different SC, but that was not the case. Quality assurance of taxonomic identification does need to be addressed if there are differences in the nomenclature in different data sets. The unimportance of this issue is illustrated by the fact that similar results were obtained in WV and KY despite their use of different protocols for sampling and enumeration.

TABLE 3. RESPONSE TO CHARGE QUESTIONS 4−8: *EXAMPLE CRITERIA CALCULATIONS* (continued)

Charge Question 4: Criterion Continuous Concentration (CCC)

Reviewer	Comment	Response
Reviewer 2	• I see that my thoughts in the above bullet is addressed on the next page (pg.3-16), lines 8−15.	The section on pg. 3-16 relates primarily to season of sampling, not sampling method.
Reviewer 2	• The use of different protocols by different organizations and agencies is a very real concern to any large database that has merged several smaller data sets. It probably is one of the biggest and most pervasive problems. The importance of initial training, repeated review throughout the monitoring season, and dedicated adherence to the field sampling quality control document can't be overemphasized.	Standardized methods are needed when the measured attribute of two or more treatments or sites are compared. That is not the case in this analysis. We are constructing an empirical model not performing a hypothesis test. It is the presence of a species in water of a particular SC that is relevant for determining their ability to survive at that ionic concentration. Failure to observe, even when present but not collected, has little or no effect on the XC_{95} or HC_{05}. This is very different from methods that are designed to compare the biological composition among sites. Bias could be introduced if sampling effort was different at different exposures, but that was not the case in the data sets used in the examples. Quality assurance of taxonomic identification does need to be addressed if there are differences in nomenclature in different data sets. Although training is important, this response to comment document is only responding to comments on the scientific underpinnings of the criteria methodology.
Reviewer 2	The authors have (gratefully) recognized this problem and have provided how to address this: by comparing all-year HC values from one region to that of another comparable region. If the datasets have a large number of data points, I believe this would be an acceptable way to handle this.	EPA agrees that sampling uncertainty can best be evaluated by comparing the results of independent studies. One estimate of that uncertainty may be provided by comparing the all-year HC05 values derived from the region for which criteria is being derived to another comparable region. Even if data are obtained in different areas by different agencies using different laboratory processing protocols, the HC05 values may be similar.

TABLE 3. RESPONSE TO CHARGE QUESTIONS 4–8: *EXAMPLE CRITERIA CALCULATIONS* (continued)

Charge Question 4: Criterion Continuous Concentration (CCC)

Reviewer	Comment	Response
Reviewer 2	• I'm not sure that I fully understand the third approach to seasonal variability in Section 3.1.4 (Assessing Seasonality, Life History, and Sampling Methods). However I do believe that the authors have done well in going step-by-step in their presenting of the third approach.	Thank you for your comment. The method is described in greater clarity in Section 3.1.4, and an example is provided in Appendix A in the public review draft.
Reviewer 3	I really like this approach, since these are well developed exposure-response relationships at the genus level, assuming that any species within the genus would share a similar response (well-known for many fish genera exposures to numerous stressors). The entire sequence of CCC analysis and the derivation of the CCC for conductivity are very well-presented in Figure 3-2, as well as in the text. Also, the example in Figure 3-1 is good, giving the reader an example of how to derive the HC_{05} of a genus sensitivity index—not a particularly easy concept to grasp unless one has some background in bioassay statistical techniques.	Thank you for your comment. No response needed.
Reviewer 3	One analytical - statistical comment: There have been a series of papers in recent years by King and Baker who use the TITAN model to examine stressor relationships with biota. It may be beneficial to explore this model to estimate conductivity-response as a check on the CCC.	The method proposed for criterion development has been supported by EPA's Science Advisory Board (SAB). Voss, King, and Bernhardt (2015) provide a comparison between the Titan and the EPA 2011a methods and found them to give similar results, with Titan generating a benchmark that was somewhat lower (Methods in Ecology and Evolution. 2015.6: 795-805). Prior to selection of the method to use for the EPA 2011 report, EPA compared the proposed method with other methods and found them to be similar. Comparisons with other methods were not included in this document so that it was clear which method EPA was recommending for the development of criteria for SC.

TABLE 3. RESPONSE TO CHARGE QUESTIONS 4–8: *EXAMPLE CRITERIA CALCULATIONS* (continued)		
Charge Question 4: Criterion Continuous Concentration (CCC)		
Reviewer	**Comment**	**Response**
Reviewer 3	I generally found the approach to derive the XC_{95} and HXC05 relatively easy to understand, with perhaps a more step-by-step on how to calculate the weighted CFD values. Was this done using Excel, R, or some other application?	The calculations were performed using R. EPA plans to prepare a step-by-step technical support document to assist users. This document is expected to be published simultaneously with the final version of the field-based methods document.
Reviewer 4	The methods for deriving the values are described clearly, especially when viewed in association with the examples presented.	Thank you for your comment. No response needed.
Reviewer 5	However, it is not quite clear what the CCC is intended to be within the context of a potential regulatory program. The CMEC is described as the maximum concentration likely to occur at a site where water quality satisfies the CCC 90% of the time, yet the CCC is also described elsewhere as a geometric mean. Which is it?	The CCC and CMEC are two different values. The CCC is an expression of the chronic exposure over a yearly cycle, hence it is a geometric mean. The CMEC is a maximum exposure concentration, that is, the highest concentration that could occur and the site could still meet the chronic criterion during that sampling year. The current text reads: "The CMEC is estimated as the highest SC level that may occur and attain the annual geometric mean SC, i.e., the CCC (annual geometric mean), 90% of the time." Examples are provided in Chapters 4 and 5.

TABLE 3. RESPONSE TO CHARGE QUESTIONS 4–8: *EXAMPLE CRITERIA CALCULATIONS* (continued)		
Charge Question 5: Criteria Maximum Exposure Concentration (CMEC). The CMEC is the maximum concentration that occurs while meeting the CCC 90% of the time. Does the analysis to derive this maximum exposure concentration (using the subset of data available with temporal resolution requirements described in Section 3.2, Deriving a CMEC), characterize the maximum concentration that will result in meeting the CCC 90% of the time, and is it reasonable to expect it to be a protective upper limit for sites in the data set? What are the strengths and weaknesses of the approach described in Section 3.2 to derive upper limits for the HC_{05} values?		
Reviewer	**Comment**	**Response**
Reviewer 1	Confidence levels (e.g., 90% or 95%) can be estimated only if the frequency distribution of data is known (e.g., normal). Did the authors check the data distribution before using Eq. 3-2? Is the critical value used here for normal distribution? If confirmed, the method is reasonable.	(1) The histogram of the data set was examined and found to approximate a normal distribution. (2) It was expected that the SC followed a log normal distribution, as observed in previous studies. (3) EPA examined the unequal representation of the data from different sites and found minimum influence of unequal variances. Therefore, the total variances within the specified range of SC was used where the unequal representation of sample sizes/variances were minimized. However, to address this reviewer's comment, EPA also performed a Shapiro's normality test for the whole dataset, close to 5000 samples, the *p*-value <0.000001. To test the effect of the large sample size, 100 random samples from the original large data were similarly analyzed and none were significantly non-normal.
Reviewer 2	***a) Does the analysis to derive the maximum exposure concentration (with temporal resolutions) characterize the maximum concentration that will result in meeting the CCC 90% of the time?*** I can only provide comment in a limited manner. The annual geometric mean is appropriate for comparing different values and finding a central tendency or typical values for a set of numbers. It normalizes the ranges and removes the effect of large differences so that no one particular range of values dominates the weighting. This is appropriate for the intent of the calculations in this section/document.	Thank you for your comment. No response needed.

TABLE 3. RESPONSE TO CHARGE QUESTIONS 4–8: *EXAMPLE CRITERIA CALCULATIONS* (continued)		
Charge Question 5: Criteria Maximum Exposure Concentration (CMEC)		
Reviewer 2	However, because I am not proficient in this, I am less sure of the maximum condition at any given station can be established by incorporating among-station and within-station variability. To achieve this, wouldn't the sampling sites and their particular data points need to be central in tendency and not exhibit values at the further reaches of the ranges? How was the 90% determined—review of that for the reader would be helpful.	The intent is to derive a CMEC that could be used in combination with the CCC. It is not intended to predict the maximum at any particular location, only whether any particular site is likely to meet the CCC. Examples are given in Chapters 4 and 5 in the public review draft. EPA has added a second call out to the location of the examples. "The steps involved in selecting, characterizing and analyzing SC (chemistry) sampling data to derive a CMEC for flowing waters in the study area are described below (see Figure 3-6) and example derivations in Sections 4.2.2 and 5.2.2."
Reviewer 2	***b) Is it reasonable to expect it to be a protective upper limit for sites in the data set?*** Yes, I think it is appropriate for determining the upper limit. 90% is definitely a protective level. Indeed, there will likely be certain interests in watersheds who will contend that this is too stringent. However, based on the sensitive genera and maximum exposure concentrations found in this document, the data (and thus, the rationale) for establishing these levels is very strong and definitive. Using the paired analyses (daily measurements of conductivity paired with macroinvertebrate sampling) is a very strong statistical test and widely used in biological and environmental studies.	Thank you for your comment. No response needed.
Reviewer 2	***c) What are the strengths and weaknesses of the approach described in Section 3.2 to derive upper limits for the HC_{05} values?*** As I have mentioned the annual geometric mean and paired analyses are strong points. The subset of frequently sampled sites is a critical element. It would seem to me that it would be important that these are clearly representative of the majority of the sites, or does the annual geometric mean make this an unnecessary concern?	EPA has added the following text to Section 3.2: "As with the derivation of the CCC, a range of exposures that leads to adverse effects on the most salt-intolerant taxa needs to be represented in the data set and there needs to be assurance that there is no bias in the sampling within that range."

TABLE 3. RESPONSE TO CHARGE QUESTIONS 4−8: *EXAMPLE CRITERIA CALCULATIONS* (continued)

Charge Question 5: Criteria Maximum Exposure Concentration (CMEC)		
Reviewer 2	The sampling of at least six times is also an important feature. I fully support the use of six times per year per site. I would increase the *n* to two in the spring (March−May) and two in the fall (Aug−Oct) and leave the remaining two for one in the summer and one in the winter. Greatest changes occur in the spring and fall months and therefore each warrant another sampling event to help capture this variability. Even with six samples, standard deviation will likely be high, especially if there are considerable differences between the sites, and even within the sites if weather, etc. are quite variable. Lastly, as much as one would like to have repeat sampling six times/site, it is often beyond the budget of many state 305(b), 303(d), and TMDL programs. Perhaps federal support can be made available for state criterion development.	Representative sampling times may vary geographically. Sampling should be done to ensure that salt-intolerant genera are represented in the SD.
Reviewer 2	The flow chart in Figure 3-6 is helpful for overall process steps. But perhaps a working example of this could be placed in the Appendix and referenced here. I think having an example would be especially helpful to state water quality staff.	Examples are available in Chapters 4 and 5 of the public review draft. EPA plans to prepare a step-by-step technical support document to assist users. This document is expected to be published simultaneously with the final version of the field-based methods document.
Reviewer 2	In keeping with the above, I would suggest greater description of LOESS and a full example. Although such processes as LOESS and bootstrapping are familiar to tacticians and to those who conduct these analyses regularly, many workers in the field of water quality programs haven’t as much familiarity.	EPA plans to provide the R code and other tools to states, tribes and territories upon finalization of the field-based methods document. LOWESS was briefly defined in Figure legend 3.6. Bootstrapping explanations were added to the text to the Figure legend 3.5. In general, these methods require a bit of study provided by statistical texts.
Reviewer 3	Similar to my comments for Question 4, the sequence for determining the CMEC is very well described in Figure 3-6. I feel that the derivation of both the CEC and the CMEC are very robust, in part because of the availability of rich ecoregion data sets. I like the fact that there is careful trimming of the data set, followed by examining for unequal variances and for estimating Type I errors	Thank you for your comment. No response needed.

TABLE 3. RESPONSE TO CHARGE QUESTIONS 4‒8: *EXAMPLE CRITERIA CALCULATIONS* (continued)		
Charge Question 5: Criteria Maximum Exposure Concentration (CMEC)		
	(there are often models published that do not perform these simple tests).	
Reviewer 4	This approach seems reasonable, however it seems that further empirical analyses of the consequences of this approach would be useful. For example, for sites that are achieving some biological benchmarks (e.g., IBI, ICI) what is the frequency that these are considered impaired based on the CCC and/or CMEC? Again my concern with single criteria rather than tiered criteria has some consequences with use of both of these benchmarks. An example of using tiered criteria and calculating CCCs and CMECs for both would be useful.	An example of the suggested analysis was provided for the CCC in U.S. EPA (2011). A more direct approach which states could use would be to compare the proportion of sites exceeding the CCC or CMEC in a region. An example is provided in Appendix C in the public review draft.

TABLE 3. RESPONSE TO CHARGE QUESTIONS 4–8: *EXAMPLE CRITERIA CALCULATIONS* (continued)		
Charge Question 5: Criteria Maximum Exposure Concentration (CMEC)		
Reviewer 5	The logic for the CMEC derivation (Section 3-2) is not presented. Where did this equation come from, and where is the supporting logic? Has the validity of the proposed approach been checked using laboratory bioassays, or with any other method that uses measured data? If so or if not, that should be stated clearly.	The formula is a routine formula for estimating the 90% confidence interval. If the CCC is met, the biota is protected even if there are occasionally higher SC than the CCC. The CCC is based on an annual average with a definable range. The derivation is a conventional statistical analysis of empirical data. EPA checked these values using in-stream measurements from the maximum SC in the summer and the occurrence of each genus in the following spring. This validated that the chemistry only method was reasonable when both biological and chemical data are not sufficient for direct analysis. Those results have been added to the case examples in Appendix A and B. Section 3.2 has been updated now that data is available for both methods. For greater clarity text was edited to read: “The CMEC analysis described here estimates the 90th centile of observations at sites with water chemistry regimes for sites meeting the CCC. It is not directly estimated from paired biological and water chemistry during acute exposures. However, if sufficient data are available (e.g., daily measurements of SC paired with macroinvertebrate sampling), a protective criterion maximum concentration could be estimated from the maximum concentration in a year prior to the observation of salt intolerant genera at a site. An example of this type of analysis is provided in Appendices A.3 and B.3, but such data sets are rare.”
Reviewer 5	The assumption underlying the CMEC calculation appears to be that the CMEC is defined as a maximum concentration that is likely to occur at a site that satisfies the CCC (estimated as the HC_{05}) 90% of the time. If one wishes to estimate a maximum concentration that is likely to occur at a site that satisfies the HC_{05} 90% of the time, one must know the temporal distribution for the target variable—SC in this case. It appears that the CMEC	“Using only water chemistry measurements and a previously determined CCC, a CMEC is estimated as the highest SC level that may occur and is likely to attain the annual geometric mean SC, i.e., the CCC (annual geometric mean), in 90% of observations.” The CMEC is not the likely maximum, rather it estimates the 90th centile of observations at sites with water chemistry regimes for sites meeting the CCC. For example, SC

TABLE 3. RESPONSE TO CHARGE QUESTIONS 4−8: *EXAMPLE CRITERIA CALCULATIONS* (continued)		
Charge Question 5: Criteria Maximum Exposure Concentration (CMEC)		
	equation has been derived assuming that SC will vary in time independently and as a normal distribution. Has this been demonstrated with field data? Others have noted that water quality data rarely vary normally and are often autocorrelated (Helsel and Hirsch (2002), see Chapter 12 for temporal analysis).	is expected to be greater in the summer and early autumn than during other times of the year in the West Virginia case example. The relationship used to derive the CMEC is based on temporal variation.
Reviewer 5	The proposed site selection procedure for the CMEC derivation is not adequate. The proposed procedure requires: At least 6 samples over a given year. A minimum of one sample in the spring (low conductivity, March−June), and one sample in the summer (high conductivity, July−October) are included to capture temporal variability. Desirable changes are:	See responses to specific comments in boxes below.
Reviewer 5	To remove the specific date designations from the second bullet, if the document goes forward with an intent for national application. Certainly, both high-concentration and low-concentration periods should be represented; but these periods may vary by time of year among regions, and among years (based on climate variability) for any given region.	EPA has edited the document to be applicable more broadly as suggested by the reviewer. The following text was inserted into Section 3.1.4: "Both high-concentration and low-concentration periods should be represented when salt-intolerant genera are collected in order to ensure that the tolerated range is evaluated. These periods may vary by time of year among regions, and among years (based on climate variability) for any given region." The following text was added to Section 3.2: “A minimum of one sample during the low conductivity season (e.g., March−June in Appalachia), and one sample in the high SC season (e.g., July−October in Appalachia) may be sufficient to capture temporal variability.”
Reviewer 5	To add an additional criterion: that remaining samples should be evenly distributed throughout the year. If remaining samples are clustered within a given time of year, they will not be	The following text was added, “The preferred data set would have multiple SC measurements evenly distributed throughout the year. A minimum of one sample during the low SC season

TABLE 3. RESPONSE TO CHARGE QUESTIONS 4−8: *EXAMPLE CRITERIA CALCULATIONS* (continued)		
Charge Question 5: Criteria Maximum Exposure Concentration (CMEC)		
	representative of the SC variability that occurs throughout the year; and, hence, would not be suitable for estimating a CMEC using statistical procedures.	(e.g., March−June in Appalachia), and one sample in the high SC season (e.g., July−October in Appalachia) may be sufficient to capture temporal variability. As with the derivation of the CCC, a range of exposures that leads to adverse effects on the most salt intolerant taxa needs to be represented in the data set and there needs to be assurance that there is no bias in the sampling within that range."
Reviewer 5	Also concerning CMEC: What are the units for the Y axes for Figure 4-9 and 5-9? Presumably, the Y axis (standard deviation) is expressed as $\log_{10}$ SC, is that right? Whatever it is, it should be stated either in the axis label or in the figure caption. Also: If I understand the axis correctly, those numbers look quite low to me—I suggest they be checked.	The label for the y-axis was added to Figures 4-9 and 5-9. All values were rechecked.
Reviewer 5	Also, the CMEC concept is not clearly defined by the document. For example, page xviii (Executive Summary) states "Below the CMEC, sites are expected to meet the CCC 90% of the time; i.e., a conductivity level that is protective of acutely toxic exposures for 95% of macroinvertebrate genera." This sentence is not written correctly. Similarly, the Glossary defines the CMEC as "In this document, the CMEC is the conductivity level at which the CCC is met 90% of the time." I think I understand what is meant by these sentences, but the language is not clearly stated.	EPA edited the sentence on xviii and in the Glossary to read "The CMEC is estimated at the 90th centile of observations at sites with water chemistry regimes meeting the CCC."
Reviewer 5	As an overall comment: I find that logic that underlies the CMEC as thinly supported, considering that its purpose is regulatory program development. The logic being applied here is statistical, not biological; and no biological data are presented as confirmation of results derived from statistical analyses.	It is true that the CMEC derivation is a statistical analysis of empirical field data and used only logic and water chemistry data throughout the year. However, the method was biologically validated. The analysis assessed the occurrence of a genus at a site after experiencing the CMEC. The results show that salt-intolerant genera are rarely present above the CMEC. This validation was added to Appendices A and B.

TABLE 3. RESPONSE TO CHARGE QUESTIONS 4−8: *EXAMPLE CRITERIA CALCULATIONS* (continued)

Charge Question 6: Duration. Please comment on the adequacy of the description and justification supporting the duration of the CCC (one year) and CMEC (one day) (see Section 3.3, Estimation of Criteria Duration)? What additional key published studies or publicly available scientific reports exist that may be useful in this discussion?

Reviewer	Comment	Response
Reviewer 1	The description and justification appear adequate. I am not aware of additional publications.	Thank you for your comment. No response needed.
Reviewer 2	This approach relies "directly on paired in-situ measurements of conductivity and benthic macroinvertebrate assemblage composition," pg. 3-20, a very reliable analysis test. Macroinvertebrates are indeed exposed to quite different conductivity levels throughout the year.	Thank you for your comment. No response needed.
Reviewer 2	The authors are quite correct that with only annual sampling, "it may be difficult to determine precisely how long conductivity levels can be above the CCC before extirpation…" (pg. 3-21, line 12−13) occurs. I would say it is <u>most</u> difficult and next to impossible to tell from one sample. Sampling only once is the reality, however, of many state bioassessment sampling programs. Nothing is better than having repeated (in the field) sampling for each site. Depending upon only one sample per year is what state programs would like to avoid but in many cases, it is all that they have. So from this standpoint, the approach seems to take this into consideration and makes sense. Lastly, lines 12−13 appear to support the argument of using only one sample/year as the basis to determine duration of CCC and CMEC. In general I believe that the authors have worked hard to provide adequate description and justification for the duration of the CCC and CMEC. The description and justification for the approach on pg. 3-22 to 3-23, line 1−16 and lines 1−16 is excellent. This is very well done.	Thank you for your comment. No response needed.

TABLE 3. RESPONSE TO CHARGE QUESTIONS 4−8: *EXAMPLE CRITERIA CALCULATIONS* (continued)		
Charge Question 6: Duration		
Reviewer 2	On a side note, is there a tag or footnote which could indicate that a data point(s) represents only one sampling per year? This would distinguish it from mean values from sites which have multiple sampling times during a year, thus allowing for all data to a dataset to be used, and yet allow the reader to know that some data are single data points and others are mean/geometric means. Seems this would be in the best interest for states wherein multiple databases are being used for criterion development or even just a single database which has some sites with only one sampling per year and some which have multiple samples. It is preferable to have more samples when possible, but it would be easy for state budget-cutters to limit sampling to just one sample per year if that is all it takes to establish criterion development. "Why sample more if only one is needed?" Further, those interests who oppose water quality criterion in general ("infringement on private property rights", "over-regulation for agriculture", "costly programs for cities", ...) would use the "only one sample per year" to justify their opposition to the criterion's validity. The argument will be that there isn't enough data and therefore the criterion is not based on "good science."	The reviewer recommends that we indicate when a value is based on a single measurement versus several and provide text to describe the benefits of multiple samples. EPA added the following text: "The preferred data set would have multiple SC measurements evenly distributed throughout the year. A minimum of one sample during the low SC season (e.g., March−June in Appalachia), and one sample in the high SC season (e.g., July−October in Appalachia) may be sufficient to capture temporal variability. As with the derivation of the CCC, a range of exposures that leads to adverse effects on the most salt intolerant taxa needs to be represented in the data set and there needs to be assurance that there is no bias in the sampling within that range."
Reviewer 3	Description and justification are more than adequate to support both CCC and CMEC. I am always a little leery about a CCC (or any water quality criteria that is based on a yearly value), since Figure 3-7 does illustrate very well the potential for large yearly variations in stream conductivity. In one of my forested study sites, conductivity may range from 75−100 µS/cm in the spring to over 600−700 µS/cm in late summer−early fall due to the dynamics of stream flow and forest transpiration.	Thank you for your comment. No response needed.
Reviewer 4	States like Ohio and localities such as the MSDGC (collected by MBI) commonly collect biological data paired with one or more weekly continuous regimes of conductivity data (e.g., Datasonde collectors). It seems that some of these sites can be used to examine duration questions in more detail. Such datasets have hourly values of conductivity collected over 7−10 days, once or twice a summer.	The reviewer recommends that EPA analyze continuous data-sondes in order to strengthen the estimation of duration for effects. EPA does not yet have this data; therefore, no change was made to the document. EPA is responsive to new research, e.g., field exposures, artificial stream exposures.

TABLE 3. RESPONSE TO CHARGE QUESTIONS 4–8: *EXAMPLE CRITERIA CALCULATIONS* (continued)		
Charge Question 6: Duration		
Reviewer 5	Section 3.3 discusses studies that are relevant to the duration question, but none of those studies address the question directly. I am not aware of relevant studies other than those discussed by the document.	Thank you for your comment. No response needed.
Reviewer 5	Answering this question would require continuous monitoring of water quality in association frequent benthic macroinvertebrate sampling, such as the data described by Boehme (2013), Boehme et al. (2013), and Timpano et al. (2013); but I am not aware of analyses by these or other authors that address this question directly.	The reviewer recommends continuous monitoring but indicates that no studies have been done to date that directly address this question using this method. Researchers have conducted artificial stream exposures that show that benthic invertebrate are affected following exposure within less than a day (Clements, 2016; Nietch, 2014). This new research has been added to the document.

TABLE 3. RESPONSE TO CHARGE QUESTIONS 4–8: *EXAMPLE CRITERIA CALCULATIONS* (continued)		
Charge Question 7: Frequency. Please comment on the adequacy of the description and justification supporting the estimation of frequency (not to be exceeded more than once in three years on average) (see Section 3.4, Estimation of Criteria Frequency)? What additional key published studies or publicly available scientific reports exist that may be useful in this discussion?		
Reviewer	**Comment**	**Response**
Reviewer 1	The recovery of macroinvertebrates strongly depends on nearest sources. If the exposure occurs at a local scale, three years may be enough for re-colonization. However, if the exposure happens at some broad scales, three years may be not enough.	EPA agrees that three years may not be sufficient for ecosystem recovery in some cases. See Section 3.4 in the public review draft.
Reviewer 2	This is one of the best sections in the document! Descriptions and reasoning are exceptionally well done throughout this entire portion. The details and thoroughness are reflective of excellent biological knowledge on the part of the authors.	Thank you for your comment. No response needed.
Reviewer 2	I only have a couple of comments: First, I am surprised at the high level of conductivity, <960 µS/cm (pg. 3-27, line 10), before extirpation of sensitive crustaceans. This seems exceptionally high. As a general rule, crustaceans, and mollusks specifically, are front line indicators of contaminants and water quality pollutants. Because water passes through them, low pH, chemicals, and excessive suspended solids and siltation are known to affect them significantly and earlier than many other aquatic organisms.	Crustaceans initially invaded freshwater from the sea whereas insects invaded freshwater after evolving to adapt to arid terrestrial conditions. It is for this reason that it has been suggested that crustaceans have different physiological abilities to adapt to ion concentrations than some aquatic insects. Owing to the sampling methods, bivalves were rarely collected, but EPA believes that this taxonomic group is protected by criteria developed according to this method because other very sensitive genera are included, some of which begin to decline just above background SC.

TABLE 3. RESPONSE TO CHARGE QUESTIONS 4–8: *EXAMPLE CRITERIA CALCULATIONS* (continued)		
Charge Question 7: Frequency.		
Reviewer 2	Secondly, I would have liked to have seen consideration given in the causal assessment methodology (Sec. 3.5, pg. 3-28, lines 15–20) of the relationship with "other known stressors such as metal toxicity, streambed erosion and siltation, and eutrophication." These conditions do contribute as stressors, often co-exist during times of high conductivity, and seem to compound effects.	Causation as described in Section 3.5 is for general causation (i.e., does SC cause extirpation rather than did SC cause extirpation in this stream?) and the model is assessed for its characterization of the responsiveness to SC compared to other stressors. For more additional types of evidence concerning potential confounding by other stressors see Suter and Cormier (2014), and U.S. EPA (2011), Appendix B. The recommended method for assessing potential confounders is described in Section 3.1.1.2.6 and analyses are illustrated in Appendices A and B. This work was referenced but not repeated.
Reviewer 2	I know from experience that during rain events and urban stormwater runoff (with increased suspended solids and accompanying high turbidity values), that conductivity also can substantially increase. A causal relationship seems to me to exist between increased turbidity and increased conductivity. This is not really addressed in the document. Do the increased conductivity values during rain events come exclusively from ions associated with concrete weathering, industrial runoff, fertilizer runoff, or, is the increase also coming from the suspended eroded soil particles (and their attachments)?	The reviewer asks about SC increases associated with storm events. EPA has chosen to provide a method that relies on sampling that often is not tied to storm events, because sampling is more commonly and safely done outside of peak flow. In the case studies, the effect of sediment was evaluated by removing poor habitat sites and recalculating the HC_{05}. EPA has added text in several places in the document on this topic recognizing that SC patterns may vary with geography, climate, and source of ions.
Reviewer 3	This section is well written and uses two classic papers (Niemi et al. and Wallace) to illustrate recovery rates in benthic organisms (insects primarily) from stressors. Recovery in stream fishes is not as clear since there may be multiple physical stressors that create long-term problems after water quality remediation (e.g., AMD), especially for lithophilic spawners. Generally, this section is highly supportive of CCC and CMEC.	Thank you for your comment. No response needed.

TABLE 3. RESPONSE TO CHARGE QUESTIONS 4–8: *EXAMPLE CRITERIA CALCULATIONS* (continued)		
Charge Question 7: Frequency.		
Reviewer 4	I think the discussion of the estimation of frequency is reasonable (not to be exceeded more than once in three years on average), but perhaps can be supplemented by some ambient analyses as another form of evidence. One suggestion might be to derive "biological stressor metrics" using the XC_{95} values. For example, I have used the most sensitive 15th percent of conductivity weighted mean values by taxa to determine taxa "sensitive to conductivity." For each site one can then generate the number of conductivity sensitive taxa present which can then be used to provide evidence that the count of sensitive taxa varies with conductivity as predicted under various duration, frequency and magnitude scenarios. This can also be used to compare potential tiered use responses to conductivity.	The reviewer notes that the discussion on frequency is reasonable. The reviewer also suggests additional research, but the example does not address work that could be used to address frequency of exposures and ecosystem recovery. EPA has also looked at the use of sensitive taxa as a modeling metric to evaluate the CMEC and summer temperatures and has added these analyses to those sections in the appendices. EPA will consider new science as it becomes available.
Reviewer 5	Comment similar to the above response to Question 6.	See above response to Reviewer 5 for Question 6.

TABLE 3. RESPONSE TO CHARGE QUESTIONS 4–8: *EXAMPLE CRITERIA CALCULATIONS* (continued)		
Charge Question 8: Alternate Measurement Endpoint. Is the example alternate measurement endpoint ($[HCO_3^- + SO_4^{2-}]$) clear and adequately supported (see Appendix F)? If not, please provide a discussion of additional data or analyses needed to support the alternative measurement endpoint. What are the benefits and weaknesses, if any, of using only two anions to describe the measurement endpoint given that ionic regulation in freshwater organisms is affected by the relative amounts of individual ions (i.e., the ionic composition)?		
Reviewer	**Comment**	**Response**
Reviewer 1	It is clear and adequately supported. This alternative endpoint explicitly identifies the stressor anions. However, it does not account for any other anions, which may be less abundant, but still significant, such as Cl^- in many freshwaters. As a result, the criterion derived would be less applicable than conductivity-based criteria. In addition, this alternative is subject to the same criticisms I made early on estimating XC_{95} and HC_{05}	The following text was added to the introduction in F-1, "This method is provided as an alternative for adopters of the method who prefer to identify the ionic constituents as a measure of exposure. This alternative method is presented as an equivalent method; however, it has several logistical disadvantages compared with using SC including cost. Also, the option to use this alternative measure does not imply that the cause of benthic invertebrate decline is only due to these two anions. Only two ions are used because they are sufficient to model the mixture; however, in other situations more or other ions would be needed to estimate exposure to an ionic mixture. When feasible, measurement of all ions in the mixture is preferred as the alternative measurement endpoint." See earlier responses regarding the reviewer's comments on XC_{95} and HC_{05}.
Reviewer 2	In general and to the best of my understanding, yes, I believe the use of an alternate measurement endpoint is written reasonably clearly and is adequately supported. In instances where I felt more description or clarity is needed, I have listed it. As in Question #4, I am going to simply list individual comments which I noted as I went through Appendix F:	Thank you for your comment. No response needed.
Reviewer 2	• The correlation value for conductivity with the two ions is exceptionally tight (Figure 1) and provides excellent data justification for their use.	Thank you for your comment. No response needed.

TABLE 3. RESPONSE TO CHARGE QUESTIONS 4–8: *EXAMPLE CRITERIA CALCULATIONS* (continued)		
Charge Question 8: Alternate Measurement Endpoint. Is the example alternate measurement endpoint ([HCO_3^- + SO_4^{2-}]) clear and adequately supported (see Appendix F)? If not, please provide a discussion of additional data or analyses needed to support the alternative measurement endpoint. What are the benefits and weaknesses, if any, of using only two anions to describe the measurement endpoint given that ionic regulation in freshwater organisms is affected by the relative amounts of individual ions (i.e., the ionic composition)?		
Reviewer	**Comment**	**Response**
Reviewer 2	• There are an exceptionally large number of paired samples (pg. F-4, line 12)! If only all studies and monitoring programs could have such a large dataset. Distribution of sampling sites was also excellent. The large background data set and the wide range of conductivity throughout the sampling area indeed allows for sound characterization of the extirpation concentration.	Thank you for your comment. No response needed.

TABLE 3. RESPONSE TO CHARGE QUESTIONS 4–8: *EXAMPLE CRITERIA CALCULATIONS* (continued)		
Charge Question 8: Alternate Measurement Endpoint		
Reviewer 2	• I believe that seasonal variation is less with these two ions. Except for September and December, there is greater similarity in their monthly values (pg. F-10, Figure F-5) than with the four ions. However, the table on the next page shows considerable variability. Nevertheless, the text says there was enough similarity on the low end of the genus sensitivity distribution to allow for criterion development. Is comparison of just the low end of the sensitivity distribution adequate? Certainly avoidance of extirpation for the most sensitive of the genera is the "goal" of the criterion, but do concentrations for moderately sensitive species, or, the extent of the ranges, have some role and should be discussed?	The reviewer raises the question of the emphasis on the salt-intolerant genera. EPA is particularly interested in factors that would affect these genera because the salt-intolerant genera have the greatest influence on the HC_{05}. If one is interested in the magnitude of the effect (e.g., percent declining) in genera above the HC_{05}, this can be estimated by enumerating the genera with declining GAM plots provided in Appendix A and B and dividing by the total number of genera. Cormier et al. (2013) published some examples and Voss et al. (2015) provide a similar analysis.
Reviewer 2	Two areas of which more description and information might be helpful for the reader:	See responses below.
Reviewer 2	1) pg. F-13, Figure F-6: Advantages of using log 10 to weight values	A sentence was added to Figure F-13: "Because the distribution of the exposure data spans three orders of magnitude a log scale base 10 was used."
Reviewer 2	2) LOWESS—pg. F-16, lines 3-6	A brief description of LOWESS was added: "A LOWESS (locally weighted regression smoothing spline) estimates a line for a scatter plot by iteratively calculating many nonparametric regression models using local approximations from neighboring points."
Reviewer 2	**The second part of Question #8:** There is a very close correlation with the two ions. They are prevalent and widely distributed in the ecoregions. They have similarity on the low end of the genus sensitivity distribution—thus functioning in the statistical analyses similarly to the four ions.	No response needed.
Reviewer 2	However, disadvantages of the two ions might include:	The following text was added to the introduction in F-1, "This method is provided as an alternative for

TABLE 3. RESPONSE TO CHARGE QUESTIONS 4–8: *EXAMPLE CRITERIA CALCULATIONS* (continued)		
Charge Question 8: Alternate Measurement Endpoint		
	Measurement of individual ions is more costly and time consuming and most sampling/monitoring programs measure for conductivity routinely, even the installed in-field monitoring instruments can give continuous readout on conductivity value. Conductivity measurement is easy, quick, and inexpensive. The four ions comprising conductivity measurements are equally as widespread in distribution or perhaps more so than the two ions. Most monitoring programs only do conductivity because of limits on budgets. Are the four ions less affected by low pH values?	adopters of the method who prefer to identify the ionic constituents as a measure of exposure. This alternative method is presented as an equivalent method; however, it has several logistical disadvantages compared with using SC including cost. Also, the option to use this alternative measure does not imply that the cause of benthic invertebrate decline is only due to these two anions. Only two ions are used because they are sufficient to model the mixture; however, in other situations more or other ions would be needed to estimate exposure to an ionic mixture. When feasible, measurement of all ions in the mixture is preferred as the alternative measurement endpoint.”
Reviewer 3	Just a very general comment to start with in this response to question 8. In our water quality laboratory, we generally do a complete cation and anion scan since these are easy on an ion chromatograph. Although primarily interested in Ca and Mg, we also analyze for K and Na, and have found these to be important cations in some streams. The anion scan is important in that it also gives a few other ions that appear sometimes in our study streams, although we do nutrient scans on other instruments due to sensitivity and detection limits.	As mentioned by other reviewers, most states, tribes, and regions do not have the resources to analyze all ions. The alternative method is intended to be as inclusive as possible. EPA agrees that a full analysis of all ions would be optimal, and has added that statement to the text for those that may be able to afford the analyses. The following text was added to the introduction in F-1, “This method is provided as an alternative for adopters of the method who prefer to identify the ionic constituents as a measure of exposure. This alternative method is presented as an equivalent method; however, it has several logistical disadvantages compared with using SC including cost. Also, the option to use this alternative measure does not imply that the cause of benthic invertebrate decline is only due to these two anions. Only two ions are used because they are sufficient to model the mixture; however, in other situations more or other ions would be needed to estimate exposure to an ionic mixture.

TABLE 3. RESPONSE TO CHARGE QUESTIONS 4–8: *EXAMPLE CRITERIA CALCULATIONS* (continued)		
Charge Question 8: Alternate Measurement Endpoint		
		When feasible, measurement of all ions in the mixture is preferred as the alternative measurement endpoint.”
Reviewer 3	So, based on the discussion in Appendix F, I would be very comfortable with using the alternative measurement endpoint, but only as a last resort if the water quality data is not adequate for a data set (meaning no measurement of Ca, Mg or Cl). There is not much difference in the slopes of Figure F-1 (c) and F-1 (d), although there is less scatter in F-1 (d) with the addition of Cl. (Note: why wasn't a test for equality of slopes performed or did I miss it somewhere in this section?).	The reviewer accepts the alternative endpoint, but also suggests some added caveats; see previous comment. The inclusion of Cl gives a tighter fit but as noted in the text, Cl is often not measured. Furthermore, if the data had been restricted to sites that contained any amount of Cl, the data set would have been significantly smaller. Text was added in section F-1: “However, a requirement to use HCO_3^-, SO_4^{2-} plus Cl^- would have greatly reduced the size of the data set and the number of genera that could be assessed with little gain in precision.”
Reviewer 4	The correlation between conductivity and the alternate measurement endpoint ($[HCO_3^- + SO_4^{2-}]$) is so strong that I think most users (e.g., States) will focus on conductivity given its cost and ability to cheaply monitor it continuously. Because of this I did not analyze this as closely as some other parts of the report, but it seems to result in a similar type of benchmark.	No response needed.
Reviewer 5	The alternate endpoint is clear and is supported by the scientific information that is available at this time. However, additional investigations are warranted as a means of thoroughly documenting the appropriateness of $[HCO_3^- + SO_4^{2-}]$ as a biotic condition indicator that would provide information other than that which is provided by SC.	The following text was added to the introduction in F-1, “This method is provided as an alternative for adopters of the method who prefer to identify the ionic constituents as a measure of exposure. This alternative method is presented as an equivalent method; however, it has several logistical disadvantages compared with using SC including cost. Also, the option to use this alternative measure does not imply that the cause of benthic invertebrate decline is only due to these two anions. Only two ions are used because they are sufficient to model the mixture; however, in other

TABLE 3. RESPONSE TO CHARGE QUESTIONS 4–8: *EXAMPLE CRITERIA CALCULATIONS* (continued)		
Charge Question 8: Alternate Measurement Endpoint		
		situations more or other ions would be needed to estimate exposure to an ionic mixture. When feasible, measurement of all ions in the mixture is preferred as the alternative measurement endpoint."
Reviewer 5	It is clear that HCO_3^- and SO_4^{2-} are the two dominant anions in most Appalachian coal-surface-mine influenced waters (Pond et al., 2008; Timpano et al., 2011; Agouridas et al., 2012; Pond et al., 2014; Sena et al., 2014). Since numerous studies have found elevated SC to be closely associated with benthic macroinvertebrate community alterations and taxa losses, it seems quite reasonable to use [HCO_3^- + SO_4^{2-}] as a measurement endpoint—although no more reasonable than use of SC itself. However, I see this relationship as a reflection of the geochemical processes that drive ion release from the mine spoils and not necessarily as a causative indicator. For that matter, the sum of Ca^{++} and Mg^{++}, which are typically the two dominant cations, could also be used as a measurement endpoint to the same effect.	The reviewer indicates that the ions are an indication of geochemical processes and suggests that the two ions may not act alone as the cause of benthic invertebrate decline. EPA agrees and has added a sentence to that effect: "Also, the option to use this alternative measure does not imply that the cause of benthic invertebrate decline is only due to these two anions." The cations were not used because there is less evidence for mechanism of action.
Reviewer 5	I do not see support for an argument that [HCO_3^- + SO_4^{2-}] would be a "better" endpoint than SC; I do not presume it to be more or less representative as an indicator of the "actual toxicant" because the actual toxicant or toxicants is/are unknown.	EPA agrees and has added a sentence to that effect. The following text was added to the introduction in F-1, "This method is provided as an alternative for adopters of the method who prefer to identify the ionic constituents as a measure of exposure. This alternative method is presented as an equivalent method; however, it has several logistical disadvantages compared with using SC including cost. Also, the option to use this alternative measure does not imply that the cause of benthic invertebrate decline is only due to these two anions. Only two ions are used because they are sufficient to model the mixture; however, in other situations more or other ions would be needed to estimate exposure to an ionic mixture. When feasible,

TABLE 3. RESPONSE TO CHARGE QUESTIONS 4–8: *EXAMPLE CRITERIA CALCULATIONS* (continued)		
Charge Question 8: Alternate Measurement Endpoint		
		measurement of all ions in the mixture is preferred as the alternative measurement endpoint."
Reviewer 5	I find the role of HCO_3^- in the observed phenomena to be quite puzzling. HCO_3^- is often elevated (relative to ecoregion 69 reference levels) in an adjacent ecoregion (Ridge and Valley, 67) due to natural conditions. Therefore, why would similar HCO_3^- levels contribute to benthic macroinvertebrate impairments in the coalfields? Are the taxa that different? Is it possible that the ratio of HCO_3^- to other ions present acts as an ecotoxicological influence? The fact that Mount et al. (1997) found HCO_3^- to be more directly associated with lab-test organism toxicity than most of the other ions studied does support a potential ecotoxicological role for HCO_3^-. However, Mount et al. (1997) also found Mg^{2+} to be associated with those toxicities; and scientific literature (e.g., Pond et al., 2008 and 2014, and other studies) demonstrates that Mg^{2+} is also quite elevated in mining-influenced high-SC streams; and Mg^{2+}/Ca^{2+} ratios are often altered in mining.	EPA agrees that the toxicological role of HCO_3^- and the issue of natural background of different ions is an unanswered question worthy of additional study, but lack of deeper mechanistic knowledge does not change the empirical relationship between increased ionic concentrations and benthic invertebrate decline.

TABLE 3. RESPONSE TO CHARGE QUESTIONS 4–8: *EXAMPLE CRITERIA CALCULATIONS* (continued)		
Charge Question 8: Alternate Measurement Endpoint		
Reviewer 5	I suggest that EPA investigate benthic macroinvertebrate status in mine-influenced SC > 300 µS/cm waters where SO_4^{2-} concentrations are quite low and HCO_3^- is the predominant anion to determine if biotic condition is such waters is consistent with expectations based on studies to date. The current proposed document would apply HC_{05} levels as criteria in such waters but it is not clear that such inclusion is warranted. Such waters have not been represented in any of the existing studies that associate SC levels with biological effects. Mine-spoil leaching studies (Agouridis et al., 2012; Daniels et al., 2013; Sena et al., 2014) indicate that SO_4^{2-} concentrations decline with progressive leachings more rapidly than SC/TDS, suggesting that HCO_3^- remains as an important solution component and becomes the dominant anion. Hence, one would expect effluents from aging mine-spoil fills constructed with non-pyritic spoils to approach a condition: SC/TDS remains elevated and HCO_3^- concentrations are also elevated (relative to reference) but SO_4^{2-} concentrations have declined substantially from the elevated levels that characterize leachates from fresh mine spoils to a concentration much lower than [HCO_3^-]. To my knowledge, biological effects of such waters have not been studied. If my understanding of mine spoil geochemistry is correct: Frequencies of occurrence by such waters will increase with time as the existing stock of mine-spoil fills that have been constructed throughout central Appalachia age and their leachate chemistries change.	The relative effects of different ions are long term research efforts being conducted in EPA in laboratory, mesocosm, and field studies. Contrary to the reviewer's comments, HCO_3^- has been implicated as a key anion in chloride ion regulation whereas sulfate has not been shown to be the toxic except indirectly (e.g, Griffith 2016, Bradley 2009, Evans 2008). In those studies, work in the last few months have indicated that sodium ions rather than sulfate ions as the primary toxic elements [Griffith 2016, other unpublished studies presented at the November 2016 Society of Environmental Toxicology and Chemistry (SETAC) Annual Meeting by Souchek, Mount, and others]. The reviewer is correct that reducing sulfate is nevertheless beneficial because it forms strong acids in water and increases rock demineralization. Mine spoil studies were not included unless they also examined biological effects. In particular, EPA examined papers on long term effects, both chemical and biological after cessation of mining. In these studies, increased conductivity and biological effects persisted below valley-fills for very long times; some have not changed in 30 years. EPA welcomes additional research that clarifies the science and will update documents as new information becomes available.

TABLE 3. RESPONSE TO CHARGE QUESTIONS 4−8: *EXAMPLE CRITERIA CALCULATIONS* (continued)		
Charge Question 8: Alternate Measurement Endpoint		
		Evans, DM; Zipper, CE; Donovan, PF; Daniels, WL. (2014) Long-term trends of specific conductance in waters discharged by coal-mine valley fills in central Appalachia, USA. J Am Water Res Assoc 50(6):1449−1460. Price, JE; Zipper, CE; Jones, JW; Franck, CT. (2014) Water and sediment quality in the Clinch River, Virginia and Tennessee, USA, over nearly five decades. J Am Water Res Assoc 50(4):837−858. DOI: 10.1111/jawr.12219. Pond, GJ; Passmore, ME; Pointon, ND; Felbinger, JK; Walker, CA; Krock, KJ; Fulton, JB; Nash, WL. (2014) Long-term impacts on macroinvertebrates downstream of reclaimed mountaintop mining valley fills in Central Appalachia. Environ Manage 54(4):919−933. Available online at http://www.ncbi.nlm.nih.gov/pubmed/24990807.

TABLE 4. RESPONSE TO CHARGE QUESTIONS 9−12: GEOGRAPHIC APPLICABILITY		
Charge Question 9: General: Is the process clearly described for assessing geographic applicability of conductivity criteria to a new area (see Section 3.6, Assessing Geographic and Waterbody Applicability)? If not, please provide suggested additional description or clarifications. Is the process a reasonable application of the recommendations made by the SAB for geographic extrapolation (see Section 3.6 and Appendix D)? Do the discussions and data analyses (to determine similarity of ionic matrix composition and estimated background conductivity) provided in these sections adequately support applicability of existing criteria to a new area with a similar ionic signature?		
Reviewer	**Comment**	**Response**
Reviewer 1	The process is clearly described, but it does not seem to fully address the concerns of SAB. To apply the conductivity criteria (HC_{05}) developed for one region to another, we need to make sure both water chemistry and macroinvertebrate fauna to be comparable. The authors did a good job for assessing water-chemistry comparability. However, their responses to SAB's other comments do not seem adequate. The authors are correct that SD does not require the same set of genera. However, it does require that the distribution of XC_{95} (at least for sensitive genera, tolerant ones are not used anyway) in the new area is similar to in the original region. Say, two regions share all genera, but if the new region happens to have more highly-sensitive genera that meet the minimum sample size (25 samples) than in the original region, its HC_{05}, if derived, may be lower, and thus the original HC_{05} would be less protective. The authors need to address the importance of biological comparability.	The reviewer indicates that the treatment and description of similar water chemistry were well assessed and that the SD does not require the same set of genera. The reviewer suggests that it be noted that more salt-intolerant genera may be present in a new region compared to those analyzed in the original area and may result in a lower HC_{05}. Additional analyses with 24 other data sets, demonstrates that the HC_{05} can be predicted from background SC and this demonstrates that different genera living in different ecoregions with similar background SC exhibit similar salt-intolerance. This analysis and method was added to the document and detailed in Appendix D subsequent to the review and after a separate external peer review in June 2015. It suggests that the Low-SC niche space in different regions is filled by similar numbers of salt-intolerant genera.

TABLE 4. RESPONSE TO CHARGE QUESTIONS 9−12: GEOGRAPHIC APPLICABILITY (continued)

Charge Question 9: General: Is the process clearly described for assessing geographic applicability of conductivity criteria to a new area (see Section 3.6, Assessing Geographic and Waterbody Applicability)? If not, please provide suggested additional description or clarifications. Is the process a reasonable application of the recommendations made by the SAB for geographic extrapolation (see Section 3.6 and Appendix D)? Do the discussions and data analyses (to determine similarity of ionic matrix composition and estimated background conductivity) provided in these sections adequately support applicability of existing criteria to a new area with a similar ionic signature?

Reviewer	Comment	Response
Reviewer 1	Yes, water quality criteria of EPA established based lab tests is applicable across the nation or multiple regions. This is because a standard lab procedure is used (test species and experimental setting). Here, we do not have a standard set of genera with the same occurrence frequencies and same environmental conditions to derive a universal HC_{05}. Extrapolating conductivity criteria to beyond the original region may be risky even if water chemistry is comparable, as I argued above.	The field-SD method models communities that are representative of other areas with similar background SC. Because of concerns about spatial extrapolation beyond an ecoregion, the EPA developed a new method based on the similarity of salt-intolerance in waters with similar background SC. Also, EPA has developed a general model that can be applied to any ecoregion in the nation. This method is in Appendix D and was externally peer reviewed in June 2015.
Reviewer 2	**First part of Question #9:** Geographic applicability is approached by the background-matching method. The background conductivity of the original area and the new area should be similar. Also the ionic mixture for the background should be the same. The authors have described the elements of this very well and I believe this will be the most exacting and appropriate way to apply criterion to a new region—a vitally important facet to ensure use by all states and for a conductivity criterion to be widely implemented.	Thank you for your comment. No response needed.

TABLE 4. RESPONSE TO CHARGE QUESTIONS 9−12: GEOGRAPHIC APPLICABILITY (continued)

Charge Question 9: General: Is the process clearly described for assessing geographic applicability of conductivity criteria to a new area (see Section 3.6, Assessing Geographic and Waterbody Applicability)? If not, please provide suggested additional description or clarifications. Is the process a reasonable application of the recommendations made by the SAB for geographic extrapolation (see Section 3.6 and Appendix D)? Do the discussions and data analyses (to determine similarity of ionic matrix composition and estimated background conductivity) provided in these sections adequately support applicability of existing criteria to a new area with a similar ionic signature?

Reviewer	Comment	Response
Reviewer 2	**Second part of Question #9:** Yes. The only area I would question would be the SAB's recommendation that "consideration be given to the species composition of stream communities, which might be different in different states . . ." (pg. D-2, line 15−17). I interpret this to mean a direct, species by species comparison. However, the authors of this document have used a taxonomic sensitivity distribution model which doesn't do this, rather, it looks at a set of species/genera and how the communities in general respond to a stressor. I believe they have provided satisfactory support for their choice. My tendency towards the SAB's recommendation is because my experience lies with species' inventories and direct counts for abundance and diversity as compared to reference streams. This preference also goes back to whether one prefers to "lump" data or "split" data**—**a long known philosophical debate among biologists!	Thank you for your comment. To further illustrate this point after this review, EPA developed 24 data sets with different background SC. Where the background SC was similar, the estimated HC05 was similar for distinct communities thousands of miles apart. A least square linear regression yielded an r = 0.93. This work has been added as Appendix D. In an independent assessment using EPA's field based method in a river basin in China with a similar background as in Ecoregion 69, the estimated HC05 was also in the 300 µS/cm range (Zhao et al 2016). The emphasis on community responses rather than responses of particular species is a long-standing EPA policy and is the basis for water quality criteria derivation since at least 1985. Zhao, Q; Jia, X; Xia, R; Lin, J; Zhang, Y. 2016. A field-based method to derive macroinvertebrate benchmark for specific conductivity adapted for small data sets and demonstrated in the Hun-Tai River Basin, Northeast China. Environmental Pollution. 216: 902-910. http://dx.doi.org/10.1016/j.envpol.2016.06.065
Reviewer 2	**Third part of Question #9:** Yes, I believe it does. This question is similar to the first part of #9 and I really don't have anything additional to add.	No response needed.

TABLE 4. RESPONSE TO CHARGE QUESTIONS 9–12: GEOGRAPHIC APPLICABILITY (continued)

Charge Question 9: General: Is the process clearly described for assessing geographic applicability of conductivity criteria to a new area (see Section 3.6, Assessing Geographic and Waterbody Applicability)? If not, please provide suggested additional description or clarifications. Is the process a reasonable application of the recommendations made by the SAB for geographic extrapolation (see Section 3.6 and Appendix D)? Do the discussions and data analyses (to determine similarity of ionic matrix composition and estimated background conductivity) provided in these sections adequately support applicability of existing criteria to a new area with a similar ionic signature?

Reviewer	Comment	Response
Reviewer 3	**First Comment:** I wonder if the general statement could be made that the process would be applicable for any Ecoregion III level embedded within any Ecoregion II level. This seems logical to me, since the original development of ecoregions was designed to address similarities in geology (and other parameters) that would translate to similar, but not identical, stream ionic concentrations. In the examples in Section 3.6, the ecoregions are adjoining so there may be a very high probability that stream chemistries may be similar.	As a result of this peer review, EPA developed two additional methods, the background matching approach for within a Level III ecoregion and another method to assess applicability of criteria derived for one ecoregion to another. These two methods were externally peer reviewed in June 2015.
Reviewer 3	**Second Comment:** A good test may be to examine two Ecoregion III level watersheds that are not contiguous, or at a large distance from each other, or perhaps two watersheds close to each other and two distant from each other.	EPA has compared results with 24 independent data sets and found that background SC predicted the HC_{05} for ecoregion thousands of miles apart. The analyses and additional case studies in Ecoregion 43 (Northwestern Great Plains) and Ecoregion 4 (Cascades) were added to the document.
Reviewer 3	**<u>Sidebar</u>**: In regard to section 4.1.1, I calculated background conductivity using the Y-intercept method developed by Dodds (for estimating background nutrients in mid-western streams) for some 152 probability-based stream sites that we sampled in Ecoregion 69 over the years. My estimate of background conductivity was 82 μS/cm and the estimate in the report was 80 μS/cm. I was pleased that these two estimates were in close agreement, especially since the analytical approaches were quite different. It is not that the Dodds' technique is so great (unless there is an adequate sample size), but similar background estimates indicate that the EPA approach for estimating	The reviewer validated the SC background in Ecoregion 69 with an independent data set and an alternate method of calculation. The results were similar. EPA also independently validated the model.

TABLE 4. RESPONSE TO CHARGE QUESTIONS 9−12: GEOGRAPHIC APPLICABILITY (continued)		
Charge Question 9: General: Is the process clearly described for assessing geographic applicability of conductivity criteria to a new area (see Section 3.6, Assessing Geographic and Waterbody Applicability)? If not, please provide suggested additional description or clarifications. Is the process a reasonable application of the recommendations made by the SAB for geographic extrapolation (see Section 3.6 and Appendix D)? Do the discussions and data analyses (to determine similarity of ionic matrix composition and estimated background conductivity) provided in these sections adequately support applicability of existing criteria to a new area with a similar ionic signature?		
Reviewer	**Comment**	**Response**
	background conductivity is consistent with other potentially useful statistical techniques.	
Reviewer 4	To derive criteria as proposed, the process for assessing general geographic applicability of conductivity is fine. As I discussed above, derivation of benchmarks under a tiered series of aquatic life use may need to accommodate modifications to the derivation approach. The geographic applicability approach generally compares whether the range/variability in conductivity in background conductivity is similar between regions. I am not sure it addresses conditions where a subset of streams may have uniquely (and predictably) lower conductivity that need to be considered separately.	A sentence has been added to Section 3.7.1: "However, there may be situations where it is not appropriate to apply criteria derived for the ecoregion to a particular stream reach. For example, naturally lower or higher concentrations of ions may occur due to subecoregional differences such as cross boundary influences, glacial melt, salt springs, highly soluble rock, or other natural sources."
Reviewer 4	Unfortunately, this is somewhat confounded with accurately identifying "background" conditions, particularly in the Ohio region of ecoregion 70. I think this paper would be well served to be placed within the conceptual framework of the Biological Condition Gradient framework (Davies and Jackson, 2006) and the reference site framework of Stoddard et al. (2006) For example, Appendix D defines "Background conductivity as the range of ionic concentrations naturally occurring in the environment that has not been influenced by human activity." Several paragraphs later, in weighing lines of evidence, it asks: "Are conductivity values at natural background (least-disturbed) sites similar in the new area compared to the original area?" As defined by Stoddard et al. (2006), least impacted is the best	EPA reviewed the document to ensure that the term "background" is consistently defined and used. Definitions were added to the glossary, and the text was expanded. The definitions proposed by Stoddard et al. (2006) were used with minor modifications. For the most part, minimally affected is preferred for background SC.

TABLE 4. RESPONSE TO CHARGE QUESTIONS 9−12: GEOGRAPHIC APPLICABILITY (continued)		
Charge Question 9: General: Is the process clearly described for assessing geographic applicability of conductivity criteria to a new area (see Section 3.6, Assessing Geographic and Waterbody Applicability)? If not, please provide suggested additional description or clarifications. Is the process a reasonable application of the recommendations made by the SAB for geographic extrapolation (see Section 3.6 and Appendix D)? Do the discussions and data analyses (to determine similarity of ionic matrix composition and estimated background conductivity) provided in these sections adequately support applicability of existing criteria to a new area with a similar ionic signature?		
Reviewer	**Comment**	**Response**
	available physical, chemical and biological habitat conditions given today's state of the landscape. With a naturally occurring measure such as conductivity, this definition can be important. Is the single benchmark or recommended criteria a reflection of "minimally disturbed" (site condition in the absence of significant human disturbance) or of least impacted conditions? Tiered uses allow a State to recognize that a subset of sites may approach minimally impacted and the associated criteria can form a baseline to protect that level of condition. Conversely, if a State has another class of sites with an appreciably higher level of acceptable development across the landscape and these sites are still considered least impacted (and these cannot be managed in a way to reduce the conductivity footprint), then different criteria for certain stressors may be applicable. For nonpoint sources of pollutant the CWA talks about controlling stressors that can be feasibly addressed with best management practices.	
Reviewer 4	One way to more closely examine the influence of tiered aquatic life uses and tiered water quality criteria would be construct a number of human disturbance indices or gradients (e.g., Bryce et al., 1999; Wang et al., 2008) and then relate them back to well-founded biological condition gradient exercises that classify sites into six ranges of biological condition based on definitions for ten components of biological condition (Davies and Jackson, 2006). This has been done for many States. It may be that the species that comprise the upper tiers of the BCG (e.g., 1−2) could well be the ones that drive the selection of the XC_{95} value, and absence of	EPA expects that additional research in the future will explore these and other questions related to ion toxicity to aquatic organisms.

TABLE 4. RESPONSE TO CHARGE QUESTIONS 9−12: GEOGRAPHIC APPLICABILITY (continued)

Charge Question 9: General: Is the process clearly described for assessing geographic applicability of conductivity criteria to a new area (see Section 3.6, Assessing Geographic and Waterbody Applicability)? If not, please provide suggested additional description or clarifications. Is the process a reasonable application of the recommendations made by the SAB for geographic extrapolation (see Section 3.6 and Appendix D)? Do the discussions and data analyses (to determine similarity of ionic matrix composition and estimated background conductivity) provided in these sections adequately support applicability of existing criteria to a new area with a similar ionic signature?

Reviewer	Comment	Response
	more tolerant taxa from reference sites would drive the conductivity benchmark lower. Conversely, these species may occur in too few reference sites at lower tiers that represent “least impacted” conditions (e.g., BCG tier 3−4).	
Reviewer 4	Appendix C, in describing how to use weight-of-evidence to examine the geographic applicability, did not examine the different aquatic life use streams across this ecoregion (particularly EWH vs WWH). Table C-7 indicates that there were no reference sites; however, biocriteria for WWH streams in this ecoregion is based on the 25th percentile of “least impacted” reference sites (EWH based on 75th percentile of sites statewide). Analysis of conductivity values at reference sites and by aquatic life use would be important evidence for this analyses.	Exploratory analysis of some Ohio sites with a "least impacted" reference designation was not necessarily associated with high ICI scores or good water quality or habitat scores. So, reference sites were not compared to any other piece of evidence.
Reviewer 4	Base flow seemingly is an important variable not considered other than in a general way in this appendix. Elevated conductivity values at sites in Appendix C in August and September correspond with the lowest estimate monthly average flows by month in Ohio (USGS ungagged model output). Some more explicit consideration of local flow influences on ionic strength would be useful. It is my experience that local base flows in headwater streams can vary considerably within across this region.	The document already included text on base-flow and the influence of natural history on time of sampling. Text was added to 3.7.1.2.: “In particular, the data set should not be biased toward seasonal extremes by sampling only during seasons of freshets or droughts.”

TABLE 4. RESPONSE TO CHARGE QUESTIONS 9−12: GEOGRAPHIC APPLICABILITY (continued)		
Charge Question 9: General: Is the process clearly described for assessing geographic applicability of conductivity criteria to a new area (see Section 3.6, Assessing Geographic and Waterbody Applicability)? If not, please provide suggested additional description or clarifications. Is the process a reasonable application of the recommendations made by the SAB for geographic extrapolation (see Section 3.6 and Appendix D)? Do the discussions and data analyses (to determine similarity of ionic matrix composition and estimated background conductivity) provided in these sections adequately support applicability of existing criteria to a new area with a similar ionic signature?		
Reviewer	**Comment**	**Response**
Reviewer 5	For the most part, the process is clearly described generally, especially when the Chapter 3 description is viewed in association with the Case Study 3 example. However, the process is not fully supported as a reasonable process for regulatory development as described. Certainly, the described process might be used by resource managers in the "new" ecoregion to inform management decisions, but regulatory development would be a different and more serious application.	While the draft document reflects EPA's assessment of the best available science regarding ambient concentrations of SC in flowing waters that protect aquatic life, it is not a regulation and does not impose legally binding requirements on EPA, states, tribes, or the regulated community, and might not apply to a particular situation based upon the circumstances.
Reviewer 5	The process described depends upon a "matching" of background SCs for two regions. The document's glossary defines the term "background conductivity" as representing the "conductivity for a region that occurs naturally and not as the result of human activity" so that is clear.	Text was edited and now uses the definition in Stoddard et al (2006) for minimally affected.

TABLE 4. RESPONSE TO CHARGE QUESTIONS 9−12: GEOGRAPHIC APPLICABILITY (continued)

Charge Question 9: General: Is the process clearly described for assessing geographic applicability of conductivity criteria to a new area (see Section 3.6, Assessing Geographic and Waterbody Applicability)? If not, please provide suggested additional description or clarifications. Is the process a reasonable application of the recommendations made by the SAB for geographic extrapolation (see Section 3.6 and Appendix D)? Do the discussions and data analyses (to determine similarity of ionic matrix composition and estimated background conductivity) provided in these sections adequately support applicability of existing criteria to a new area with a similar ionic signature?

Reviewer	Comment	Response
Reviewer 5	A deficiency in the Geographic Applicability description concerns the process for determining a background SC when using a distribution of water quality data from a given region. As noted by the document: "the 25^{th} centile is conventionally used … However, when land cover modification (or other anthropogenic disturbance) is pervasive, selection of a centile lower than 25% may be justifiable." The fact that the 25^{th} percentile is based on assumptions and not on rigorous analysis should be noted. The document references U.S. EPA (2000a), but U.S. EPA (2000a) only asserts that a percentile within the range of 5^{th} to 25^{th} percentile can be used to represent a reference value without citing supporting studies or analyses. As stated by U.S. EPA (2000a), page 4-8: "Both the 75^{th} percentile for reference streams and the 5^{th} to 25^{th} percentile from a representative sample distribution are only recommendations. The actual distribution of the observations should be the major determinant of the threshold point chosen." As far as I can tell, the decision concerning which centile to select as a "background SC" indicator is being left to judgment, yet the outcome of this process could be significant as a determinant of the "new" ecoregion's HC_{05} according to the process described here. A decision which centile to select can make a big difference in the resulting background estimate. In regional databases I have available for analysis, the 25^{th} percentiles of SC distributions differ from the 5^{th} percentiles by multiples ranging from 3 to 10.	EPA provided a weight-of-evidence approach that uses multiple types of evidence when the reliability of the background estimate is uncertain or inconsistent with scientific expectations. In the public review draft document, EPA has also added a model that relies on geochemistry, climate, and vegetative cover as another check of base-flow conductivity. Also in the public review draft document, the uses of different estimates of background are described and the 25^{th} centile from a probability based survey sample is recommended unless the area is shown to be affected by anthropogenic disturbance. In any case, estimation of background will depend in part on professional judgment.

TABLE 4. RESPONSE TO CHARGE QUESTIONS 9–12: GEOGRAPHIC APPLICABILITY (continued)		
Charge Question 9: General: Is the process clearly described for assessing geographic applicability of conductivity criteria to a new area (see Section 3.6, Assessing Geographic and Waterbody Applicability)? If not, please provide suggested additional description or clarifications. Is the process a reasonable application of the recommendations made by the SAB for geographic extrapolation (see Section 3.6 and Appendix D)? Do the discussions and data analyses (to determine similarity of ionic matrix composition and estimated background conductivity) provided in these sections adequately support applicability of existing criteria to a new area with a similar ionic signature?		
Reviewer	**Comment**	**Response**
Reviewer 5	One of the authors of this document has published a peer-reviewed study (Griffith, 2014) that derives 25th percentiles SC distributions for ecoregions from throughout the U.S.; that study describes a method summarized in the abstract as "followed EPA methods to estimate reference values" but does not describe the resulting values as "background". In fact, the author states that "Much discussion exists in the literature as to whether estimates like mine are true estimates of background or at least current reference conditions …".	In the cited work, Griffith provided the centiles based on probabilistic sampling which shows areas of lower and higher conductivities across the country in a descriptive report. Other centiles are also reported in his paper. Almost all of the estimates are based on fewer than 60 samples and so should be taken as illustrative rather than a final vetted background, regardless of the centile that might be selected.
Reviewer 5	The process for assessing geographic applicability of conductivity criteria to a new area is not fully supported. What is missing is a biological confirmation. Are benthic macroinvertebrate communities for the two ecoregions in streams of similar sizes comprised of similar taxa? Do the limit-defining taxa also occur in the "new" ecoregion? A method for evaluating biological data as a means of answering such questions should be described. If background conductivity, background ionic signature, and benthic macroinvertebrate taxa for the two ecoregions are all similar and the limit-defining taxa are present in both ecoregions, it would be reasonable to consider applying an HC_{05} value developed for one ecoregion to another.	Taxa occupy low ionic niches where those environments are available. EPA developed 24 SD from 23 ecoregions across the country. In every case, salt-intolerant genera with low XC_{95} values were present at the lowest available SC habitats. A model background (measured at the 25th centile) and HC_{05} values for 23 ecoregions (24 data sets) and the model was strong ($r = 0.93$). This shows that although particular genera may not have the same XC_{95} in every ecoregion, the general model of community salt-tolerance is very similar across ecoregions. This analysis and method was added to the document and detailed in Appendix D subsequent to the review and after a separate external peer review in June 2015.

TABLE 4. RESPONSE TO CHARGE QUESTIONS 9−12: GEOGRAPHIC APPLICABILITY (continued)		
Charge Question 9: General: Is the process clearly described for assessing geographic applicability of conductivity criteria to a new area (see Section 3.6, Assessing Geographic and Waterbody Applicability)? If not, please provide suggested additional description or clarifications. Is the process a reasonable application of the recommendations made by the SAB for geographic extrapolation (see Section 3.6 and Appendix D)? Do the discussions and data analyses (to determine similarity of ionic matrix composition and estimated background conductivity) provided in these sections adequately support applicability of existing criteria to a new area with a similar ionic signature?		
Reviewer	**Comment**	**Response**
Reviewer 5	The underlying assumption of the Geographic Applicability process, as described, is as follows: If background SC estimates for two regions are similar, then sensitivity of regional taxa to elevated SC will be similar as well. The document cites no studies to support the validity of that assumption.	EPA has modeled background (measured at the 25th centile) and HC_{05} values for 23 ecoregions using 24 data sets. The model was strong (r = 0.93) and now complements the background matching approach. The analysis shows that the HC_{05} can be predicted from background SC. These analyses confirm that when the SC is low, there are salt-intolerant genera. In addition to this analysis, basic ecology supporting the association is described and citations are provided.

TABLE 4. RESPONSE TO CHARGE QUESTIONS 9−12: GEOGRAPHIC APPLICABILITY (continued)		
Charge Question 10: Geographic Applicability To A New Area Within An Ecoregion: Please comment regarding the clarity of the process described for assessing geographic applicability of field-based conductivity criteria to locations within the same ecoregion that are outside the geographic bounds of the parent data sets (see Section 3.6). Do the Case Study analyses (see Sections 4.3 and 5.3) adequately support the application of the derived example criteria within those ecoregions? If not, please describe why and any additional data and analyses needed.		
Reviewer	**Comment**	**Response**
Reviewer 1	The process described is clear (Section 3.6), and the case study analyses support the limited extrapolation. However, I am not clear why one would not include samples from the whole ecoregion in the criterion development at the first place. Even if one ecoregion includes streams in more than one state, it appears much easier to combine raw data from all states involved, standardize them (e.g., sub-sample size), and then develop the region-wide criteria, than to rely on extrapolation.	EPA agrees with the reviewer that a composite data set is a reasonable option. However, for the case studies EPA chose to analyze the WV Ecoregion 69 and 70 data sets separately in order to demonstrate the applicability method should the need arise, as pointed out by Reviewer 3. Furthermore, comparable paired SC and benthic invertebrate data were not available from some of the areas in each ecoregion. In the fish example, combined data sets for fish were used in order to increase the size of the data set thus illustrating both options.
Reviewer 2	The authors have been meticulous in setting the parameters for the study and then clearly describing in this document the process for assessing geographic applicability to locations within the same ecoregion but outside of the parent data sets. First, as they stated on pg. 3-32, most streams in an ecoregion tend to have a similar conductivity regime and ionic composition of dissolved salts. This is generally true, but exceptions do occur, and they wisely caution to have care when applying the example ecoregion criteria to any one particular stream reach. Specific changes in rock composition or feeder streams with springs can alter a particular reach. Good job in recognizing this.	Thank you for your comment. No response needed.

TABLE 4. RESPONSE TO CHARGE QUESTIONS 9−12: GEOGRAPHIC APPLICABILITY (continued)		
Charge Question 10: Geographic Applicability To A New Area Within An Ecoregion: Please comment regarding the clarity of the process described for assessing geographic applicability of field-based conductivity criteria to locations within the same ecoregion that are outside the geographic bounds of the parent data sets (see Section 3.6). Do the Case Study analyses (see Sections 4.3 and 5.3) adequately support the application of the derived example criteria within those ecoregions? If not, please describe why and any additional data and analyses needed.		
Reviewer	**Comment**	**Response**
Reviewer 2	Regional background conductivity is defined well on pg. 3-33, lines 7−13. Continuing, they clearly point out that for a data set from one geographic area to be applicable to another similar area, there needs to be: a) similar background conductivity levels and ion composition, and b) a comparison of the confidence intervals of the background data set of the new area to those of the original area; confidence bounds for background estimated from the example criterion data set overlapped with the confidence bounds for the background of the rest of Ecoregion 69. The weight-of-evidence assessment for applicability of criteria to the new area adds much to the soundness of the approach. This validation of background specifically for the Ohio portion of Ecoregion 70 was done with a weight-of-evidence. A weight-of-evidence process is something which state staff are accustomed to doing and thus can identify easily with this and conduct it.	Thank you for your comment. No response needed.
Reviewer 2	Excellent discussion on pg. 3-38, lines 14−25, of causes of and considerations when the background conductivity is greater in the new area than in the original area. I appreciated seeing such a complete listing; well thought-out and applicable. The summary of 3.6.3.4 is also well done. The discussion on pg. 5-18, Section 5.3 showed further rationale for the reliability of this process: using first through fourth-order streams (thus maintaining some uniformity of catchment size), extensive data sets, probability-based designs, methods comparable across the assessments and QA/QC. The approach has been well done.	Thank you for your comment. No response needed.

TABLE 4. RESPONSE TO CHARGE QUESTIONS 9–12: GEOGRAPHIC APPLICABILITY (continued)		
Charge Question 10: Geographic Applicability To A New Area Within An Ecoregion: Please comment regarding the clarity of the process described for assessing geographic applicability of field-based conductivity criteria to locations within the same ecoregion that are outside the geographic bounds of the parent data sets (see Section 3.6). Do the Case Study analyses (see Sections 4.3 and 5.3) adequately support the application of the derived example criteria within those ecoregions? If not, please describe why and any additional data and analyses needed.		
Reviewer	**Comment**	**Response**
Reviewer 2	In Section 4.3, the utilization of the background-matching approach for geographical applicability was effective in Ecoregion 69 as well as in Section 5.3 and Ecoregion 70. The new portion was estimated at the 25th percentile, comparing with the background conductivity estimates of the original set. All chloride-dominated samples were removed before estimating background conductivity, thus keeping the same ionic mixture for the new area the same as the example criteria. The importance of keeping data inputs all of the same "category" for a quality comparison assessment is more valuable and fundamental to good statistics than satisfying an approach that believes all data should be included. Thus, this answers previous questions of whether there should be exclusion of particular ions.	Thank you for your comment. No response needed.
Reviewer 3	Based on the analyses presented, I do not have any problem with the process. Indeed, the case study analyses in Sections 4.3 and 5.3 do support the criteria application. Basically, I feel that the parent data set is a training set, and the conductivity estimates outside of this set should be well within statistical bounds.	Thank you for your comment. No response needed.
Reviewer 4	The same comments I provided above apply to this charge question.	See response to Charge Question 9 above.
Reviewer 5	Answer is similar to that provided to Question 9 above: The assumption here is that taxa comprising benthic macroinvertebrate communities within the smaller area are similar to those that occur within the larger area. This validity of this assumption should be verified before extending criteria developed in one area to another, regardless of ecoregion boundaries.	As explained previously, taxonomic similarity is not required, only similar salt tolerances and our analysis confirmed that salt-intolerances are similar when the backgrounds are similar. EPA has developed a model that reliably predicts HC_{05} values for ecoregions from background. This model, which was externally peer reviewed in June 2015, demonstrates that it is possible to extend criteria from one area to another.

TABLE 4. RESPONSE TO CHARGE QUESTIONS 9−12: GEOGRAPHIC APPLICABILITY (continued)		
Charge Question 11: Geographic Applicability To A New Area In Another Ecoregion: Please comment regarding the clarity of the applicability analysis for the background-matching approach described in Section 3.6.3 and illustrated in Section 6. Do the data and analyses adequately support the application of the example criteria to other areas? If not, please describe why and any additional data and analyses needed.		
Reviewer	**Comment**	**Response**
Reviewer 1	The applicability analysis is clearly described, but again I am not fully convinced that the approach is sufficient for the reason described in my comments on Question 9. The case study appears to support the approach, but the new ecoregion (68) is just next to the original one (69). The result may differ substantially if the new ecoregion is further away and associated with very different benthic fauna. It is difficult to generalize the effectiveness of this background-matching approach based on this special case study.	Additional analyses with other data sets thousands of miles away from each other yield similar HC_{05} values when the background conductivity is similar. EPA has modeled background (measured at the 25th centile) and HC_{05} values for 23 ecoregions using 24 data sets. The model was strong ($r = 0.93$) and now complements the background matching approach. The analysis shows that the HC_{05} can be predicted from background SC. These analyses confirm that when the SC is low, there are salt-intolerant genera. In addition to this analysis, basic ecology supporting the association is described and citations are provided. The case study in Section 6 for Ecoregion 68 was replaced with two examples from ecoregions in the West. These analyses and results have been added to the document.
Reviewer 2	As in Question #10 and within-ecoregion, the analysis for the background-matching approach for geographic applicability to a new area in another ecoregion was well presented. The monitoring and sampling procedures (Alabama, Kentucky, Tennessee) are very clearly described, well-defined, specific, and a pleasure to read. The Results in Section 6.2 are presented point by point. It is helpful to have these points in paragraphs 1−4 on pg. 6-10−6-11.	Thank you for your comment. No response needed.
Reviewer 2	Confidence intervals are greater in Ecoregion 70 than the other two regions (pg.6-12, Table 6.4). Perhaps reasons should be given for this. The difference is quite notable.	The reason that confidence intervals are greater for ecoregion 70 is unknown. Greater land use modification is a possible cause. EPA does not have the data or analyses to definitively answer this

TABLE 4. RESPONSE TO CHARGE QUESTIONS 9−12: GEOGRAPHIC APPLICABILITY (continued)		
Charge Question 11: Geographic Applicability To A New Area In Another Ecoregion: Please comment regarding the clarity of the applicability analysis for the background-matching approach described in Section 3.6.3 and illustrated in Section 6. Do the data and analyses adequately support the application of the example criteria to other areas? If not, please describe why and any additional data and analyses needed.		
Reviewer	**Comment**	**Response**
		question. No doubt many aspects of this work will be a source of new research in years to come.
Reviewer 2	Also in this table it is mentioned that the WABbase data set for Ecoregion 69 included samples without genus identification, meaning that identification was carried just to the Family level. Although it is always better to be able to key down to genus, this is not unusual. This is a problem for stream monitoring/sampling programs and will probably only get worse as fewer individuals are training in entomology.	The WABbase data set did include taxa that were not identified to genus, but they were not used in the analysis. Because only water chemistry was analyzed, sites were included where biological identification was at family or greater level. However, Section 6 was replaced with two new examples, so this was not noted in this location.
Reviewer 2	Verification of applicability of example criterion from an original ecoregion to a new ecoregion using the background-matching method was done by independently estimating the HC 05. Verification is an important step—a necessary "hurdle"—and when it is successful the method can be confirmed to be reliable and its use can proceed. I compliment the authors for presenting the first demonstration of successfully applying a criterion to a new region. This is a major step in expanding agencies' ability to establish a criterion for a parameter which has a significant role in the health of aquatic organisms.	Thank you for your comment. No response needed.

TABLE 4. RESPONSE TO CHARGE QUESTIONS 9–12: GEOGRAPHIC APPLICABILITY (continued)		
Charge Question 11: Geographic Applicability To A New Area In Another Ecoregion: Please comment regarding the clarity of the applicability analysis for the background-matching approach described in Section 3.6.3 and illustrated in Section 6. Do the data and analyses adequately support the application of the example criteria to other areas? If not, please describe why and any additional data and analyses needed.		
Reviewer	**Comment**	**Response**
Reviewer 2	I do have concerns about the need for about 500 samples in order to achieve consistent results with the HC_{05} derivation (pg. 6-13). The need for a large data set is well understood, but it may be difficult for some entities to have that many samples.	As mentioned in the text, the number of samples needed varies depending on the sampling methods, range of exposure, and number of genera identified. However, criteria can be developed without a de novo derivation by using the background matching approach or using background and a new regression model as described in the public review draft document.
Reviewer 2	The mention that there were different sampling methods, and that some methods tend to collect different types of genera, is appreciated. While it would be better to have uniform sampling methods, in reality that doesn't always happen. The authors tried to restore confidence that a large variety of taxa were nevertheless represented. It might be wise to recheck the methods/protocols used, verifying that each method was used about equally throughout the data set.	This method depends on presence and is therefore not strongly affected by sampling method unlike sampling for biological composition or abundance. Furthermore, samples are not being compared to one another, the information is used separately for each genus to develop an XC_{95}. If a genus is observed, that is relevant regardless of the method used to observe it. If there was a bias in the range of exposures sampled, i.e., one method sampling mostly low conductivity and another method sampling mostly high conductivity sites, this might affect the result. However, in the example cases, a wide range was sampled by each method. Text was added to an existing statement to highlight this issue.
Reviewer 2	The authors were very specific and thorough in their description on pg. 6-13. The applicability was well presented.	Thank you for your comment. No response needed.
Reviewer 2	As I have mentioned previously, my expertise lies in the biological aspects rather than the statistical analyses. With this in mind, it would be helpful to have more information about bootstrapping and an example by which one could follow. This would be an Appendix supplement I realize, but I think it would be useful to staff in watershed programs.	A technical support tool for calculation may be prepared and more explanation of bootstrapping was added to the current document.

TABLE 4. RESPONSE TO CHARGE QUESTIONS 9−12: GEOGRAPHIC APPLICABILITY (continued)		
Charge Question 11: Geographic Applicability To A New Area In Another Ecoregion: Please comment regarding the clarity of the applicability analysis for the background-matching approach described in Section 3.6.3 and illustrated in Section 6. Do the data and analyses adequately support the application of the example criteria to other areas? If not, please describe why and any additional data and analyses needed.		
Reviewer	**Comment**	**Response**
Reviewer 3	I think this section was well written, and that the data and analyses do support the application of conductivity criteria to other areas.	Thank you for your comment. No response needed.
Reviewer 4	I think the background matching analysis is a sound approach for comparing applicability to another ecoregion. Again, I have the same caveat about potentially doing this in a tiered use framework.	Thank you for your comment. No response needed.
Reviewer 5	The background matching concept is clearly described, generally, but certain details are murky.	See EPA response below.
Reviewer 5	As mentioned above, the process for defining a “background SC” should be more explicit**—**if it is to be used as an essential component of a criterion definition process as described by the document.	EPA has edited the text in Section 3.7.1 and in Appendix C of the draft document to more clearly describe the weight-of-evidence approach and definition of background SC.
Reviewer 5	One detail in the Case Study 3 example is not clear. Multiple data sets were used for Case Study 1 and for Case Study 2. A background SC estimate (±CI) was derived for each of these datasets. Then, Case Study 3 uses a single background SC estimate from each of the two case studies for the “matching” analysis but it does not state a clear rationale for selection. More specifically: the value described as the “the example Criterion data set” (as described on page 4-21) was used to represent Case Study 1 (94 µS/cm; 95% CI 86−101 µS/cm); however, the background SC estimated for the corresponding data set	Example case study 3 was replaced with an example using the B-C regression method after a separate peer review in June 2015. EPA recommends using overlapping 90% CI of the criterion data set and data sets from the rest of the region. Flow charts of the processes were updated for clarity in Section 3.7. EPA also recommended a weight-of-evidence assessment if there is uncertainty or ambiguous results. An example was provided in Appendix C of this review draft.

TABLE 4. RESPONSE TO CHARGE QUESTIONS 9−12: GEOGRAPHIC APPLICABILITY (continued)

Charge Question 11: Geographic Applicability To A New Area In Another Ecoregion: Please comment regarding the clarity of the applicability analysis for the background-matching approach described in Section 3.6.3 and illustrated in Section 6. Do the data and analyses adequately support the application of the example criteria to other areas? If not, please describe why and any additional data and analyses needed.

Reviewer	Comment	Response
	("example Criterion derivation data set" as described on page 5-22) was not used to represent Case Study 2; rather, the "WABbase data set, probability sample subset" (147 µS/cm; 95% CI 136−159 µS/cm) was used to represent Case Study 2. What is the rationale for deciding which background SC estimate to use in procedures such as the Case Study 3 example?	
Reviewer 5	Also, it is not clear to me how the 95% CIs are derived for the background SC estimates. Maybe that procedure is described somewhere in the document but I am not finding it.	EPA added the following text in section 3.7.1.2: "The confidence interval for a background SC estimate can be calculated using a bootstrapping technique. Bootstrapping is a statistical resampling technique that is often used in environmental studies to estimate confidence limits of a parameter. This bootstrapping application involves randomly resampling the original water chemistry data set 1,000 times with replacement, storing the 1,000 data sets, calculating the background for each data set, and then estimating the 95% CI for the mean of the set of 1,000 background values generated by the bootstrapping procedure. This is similar to the procedure described in Section 3.1.3.1."
Reviewer 5	In my view, the underlying rationale for use of the background matching approach alone to establish HC_{05} values is not clearly supported (as stated above in response to question 9). Justification of the process would require biological confirmation; no process for biological confirmation is described.	EPA performed analyses demonstrating that if a SC niche is available, there will be a taxa that fills it. Specifically, additional analyses with other data sets thousands of miles away from one another yield similar HC_{05} values when the background SC is similar. These analyses were added to the document.

TABLE 4. RESPONSE TO CHARGE QUESTIONS 9–12: GEOGRAPHIC APPLICABILITY (continued)

Charge Question 12: Applicability To Ephemeral Streams: In their 2011 review of the EPA Benchmark Report, the SAB indicated that because the data used to derive the benchmark were collected from perennial streams, the empirical relationship between conductivity and genera occurrence likely would be applicable to perennial and intermittent streams, but not to ephemeral streams. In preparing the current draft document, EPA found several publications that indicate that some aquatic organisms on which the Case Study example criteria are based do occur in ephemeral streams and that these organisms are critical to these headwater systems. EPA also believes it appropriate to include ephemeral waters as applicable water bodies for field-based conductivity criteria in order to ensure protection of aquatic communities in downstream intermittent or perennial waters. Therefore, EPA considers the field-based method applicable to all types of flowing waters, including perennial, intermittent, and ephemeral streams. Do you believe that this recommendation is well supported by Section 3.6.2 (Waterbody Type), including the publications it cites? Are you aware of any additional published studies or publically available scientific reports or data (e.g., paired chemical and biological sampling) relevant to this issue?

Reviewer	Comment	Response
Reviewer 1	Ephemeral streams typically support highly-mobile taxa (e.g., beetles) and taxa of short life cycles (e.g., some chironomids). They may collectively share most genera with perennial streams, as shown by Grubbs (2010), but the occurrence frequencies and abundance of most shared taxa are most likely to be much lower. My experience is that not many sensitive genera live in temporal streams. As a result, these streams may be "over-protected" by a HC_{05} established for perennial streams. However, I agree that they are important components of stream networks, but should be protected by separate conductivity criteria. Temporal streams have recently attracted much research interest. Following references are relevant:	EPA agrees that ephemeral streams are important components of stream networks and that sensitive taxa occur in them.
Reviewer 1	Lake, P.S. 2011. *Drought and aquatic ecosystems: effects and responses*. Chichester, UK. Wiley-Blackwell.	This book deals with extreme conditions in Australia and therefore was not cited in the draft methods document. It focuses on effects of temperature, lack of water and very little on increased ionic stress. For the most part it deals with a NaCl rather than bicarbonate and sulfate-dominated mixtures.

TABLE 4. RESPONSE TO CHARGE QUESTIONS 9−12: GEOGRAPHIC APPLICABILITY (continued)

Charge Question 12: Applicability To Ephemeral Streams: In their 2011 review of the EPA Benchmark Report, the SAB indicated that because the data used to derive the benchmark were collected from perennial streams, the empirical relationship between conductivity and genera occurrence likely would be applicable to perennial and intermittent streams, but not to ephemeral streams. In preparing the current draft document, EPA found several publications that indicate that some aquatic organisms on which the Case Study example criteria are based do occur in ephemeral streams and that these organisms are critical to these headwater systems. EPA also believes it appropriate to include ephemeral waters as applicable water bodies for field-based conductivity criteria in order to ensure protection of aquatic communities in downstream intermittent or perennial waters. Therefore, EPA considers the field-based method applicable to all types of flowing waters, including perennial, intermittent, and ephemeral streams. Do you believe that this recommendation is well supported by Section 3.6.2 (Waterbody Type), including the publications it cites? Are you aware of any additional published studies or publically available scientific reports or data (e.g., paired chemical and biological sampling) relevant to this issue?

Reviewer	Comment	Response
Reviewer 1	Steward, A.L., D. von Schiller, K. Tockner, J.C. Marshall, and S.E. Bunn 2012. When the river runs dry: human and ecological values of dry riverbeds. *Frontiers in Ecology and the Environment* **10**: 202−209	This paper does not address how species use wetted ephemeral streams. It deals with the functional value of dry river beds primarily in dryland Australia. This citation was not included in the draft document.
Reviewer 1	Williams, D.D. 2006. *The biology of temporary waters*. New York, NY: Oxford University Press.	This is an introductory text and primarily deals with ephemeral ponds rather than streams or other flowing waters. This citation was not included in the draft document.
Reviewer 2	Yes, the discussion in Section 3.6.2 well supports the field-based method for applicability to ephemeral streams. The support is well defined by the publications cited. It clearly discusses that macroinvertebrates are found in intermittent and ephemeral streams. Grubbs' (2010) research provides excellent quantifying results. In my experience, I believe the abundance and diversity of macroinvertebrates in ephemeral and intermittent streams is far greater than larger rivers which (in this part of the country) have nearly all been channelized, straightened, trees/brush removed, island and meanders removed, streambeds laden with silt, and hydrologically modified.	Thank you for your comment. No response needed.

TABLE 4. RESPONSE TO CHARGE QUESTIONS 9−12: GEOGRAPHIC APPLICABILITY (continued)

Charge Question 12: Applicability To Ephemeral Streams: In their 2011 review of the EPA Benchmark Report, the SAB indicated that because the data used to derive the benchmark were collected from perennial streams, the empirical relationship between conductivity and genera occurrence likely would be applicable to perennial and intermittent streams, but not to ephemeral streams. In preparing the current draft document, EPA found several publications that indicate that some aquatic organisms on which the Case Study example criteria are based do occur in ephemeral streams and that these organisms are critical to these headwater systems. EPA also believes it appropriate to include ephemeral waters as applicable water bodies for field-based conductivity criteria in order to ensure protection of aquatic communities in downstream intermittent or perennial waters. Therefore, EPA considers the field-based method applicable to all types of flowing waters, including perennial, intermittent, and ephemeral streams. Do you believe that this recommendation is well supported by Section 3.6.2 (Waterbody Type), including the publications it cites? Are you aware of any additional published studies or publically available scientific reports or data (e.g., paired chemical and biological sampling) relevant to this issue?

Reviewer	Comment	Response
Reviewer 2	I found the short discussion of the various adaptations to survive temporary dry periods (pg. 3-20, lines 27−30) to be exceptionally helpful. Very seldom is this addressed or even widely known, however, it is a significant fact among many of the taxa of ephemeral streams. It was a pleasure to see this included. And the continuing discussion of the use of upstream temporary streams for part of their life cycle (pg. 3-30−3-31) is also accurate and equally important to include. This fact, and the documented presence of the “vast majority (91 out of 108) of macroinvertebrate taxa were observed in both the perennial and temporary channels” (Grubbs, 2010), provides strong rationale for the applicability of field-based method/criterion for conductivity to ephemeral streams. Upstream water quality conditions affect lower reaches’ aquatic life and the exposure to harmful levels of conductivity (and all other contaminants as well).	Thank you for your comment. EPA agrees that criteria derived using this method are applicable to flowing fresh waters, including ephemeral, intermittent, and perennial streams (see Section 3.6.2).

TABLE 4. RESPONSE TO CHARGE QUESTIONS 9−12: GEOGRAPHIC APPLICABILITY (continued)

Charge Question 12: Applicability To Ephemeral Streams: In their 2011 review of the EPA Benchmark Report, the SAB indicated that because the data used to derive the benchmark were collected from perennial streams, the empirical relationship between conductivity and genera occurrence likely would be applicable to perennial and intermittent streams, but not to ephemeral streams. In preparing the current draft document, EPA found several publications that indicate that some aquatic organisms on which the Case Study example criteria are based do occur in ephemeral streams and that these organisms are critical to these headwater systems. EPA also believes it appropriate to include ephemeral waters as applicable water bodies for field-based conductivity criteria in order to ensure protection of aquatic communities in downstream intermittent or perennial waters. Therefore, EPA considers the field-based method applicable to all types of flowing waters, including perennial, intermittent, and ephemeral streams. Do you believe that this recommendation is well supported by Section 3.6.2 (Waterbody Type), including the publications it cites? Are you aware of any additional published studies or publically available scientific reports or data (e.g., paired chemical and biological sampling) relevant to this issue?

Reviewer	Comment	Response
Reviewer 2	As I've mentioned in my response in Question #2, intermittent and ephemeral streams (even "often-wet" depressions in fields, wet meadows and pasture drainages, wet areas in riparian corridors or nearby river valleys, etc.) provide habitat for macroinvertebrates—at least for a portion of their life cycle. The value of these small streams and temporary wet areas as habitat for many taxa has not been appreciated nor understood by many property owners, developers, and policy decision-makers. The decision to include ephemeral streams in this criterion development is especially important and gratefully appreciated by biologists such as myself. The information provided here in this section also provides strong rationale for the current debate on "navigable waters" regulations.	Thank you for your comment. EPA agrees that criteria derived using this method are applicable to flowing fresh waters, including ephemeral, intermittent, and perennial streams (see Section 3.6.2).
Reviewer 2	Please refer back to my response for Question #2 for my other previous comments.	See response above.
Reviewer 3	My basic comment is that we are not doing enough to protect either intermittent or ephemeral streams, and I would recommend that EPA take a stronger stance on these important characteristics of the watershed. It is hard enough to protect 1st order streams in the United States, but trying to gain protection for zero order streams is almost impossible. Consequently, and where applicable, any paired analyses with conductivity and benthic organisms would be beneficial to support the importance of ephemeral	Thank you for your comment. No response needed.

TABLE 4. RESPONSE TO CHARGE QUESTIONS 9−12: GEOGRAPHIC APPLICABILITY (continued)

Charge Question 12: Applicability To Ephemeral Streams: In their 2011 review of the EPA Benchmark Report, the SAB indicated that because the data used to derive the benchmark were collected from perennial streams, the empirical relationship between conductivity and genera occurrence likely would be applicable to perennial and intermittent streams, but not to ephemeral streams. In preparing the current draft document, EPA found several publications that indicate that some aquatic organisms on which the Case Study example criteria are based do occur in ephemeral streams and that these organisms are critical to these headwater systems. EPA also believes it appropriate to include ephemeral waters as applicable water bodies for field-based conductivity criteria in order to ensure protection of aquatic communities in downstream intermittent or perennial waters. Therefore, EPA considers the field-based method applicable to all types of flowing waters, including perennial, intermittent, and ephemeral streams. Do you believe that this recommendation is well supported by Section 3.6.2 (Waterbody Type), including the publications it cites? Are you aware of any additional published studies or publically available scientific reports or data (e.g., paired chemical and biological sampling) relevant to this issue?

Reviewer	Comment	Response
	streams.	
Reviewer 4	I would examine Ohio EPA's Primary Headwater Assessment data. It focused on streams generally less than about a square mile and includes ephemeral as well as perennial and interstitial streams. They have collected conductivity data and have sites in the WAP ecoregion as part of their studies. Ohio University also collected primary headwater stream data as part of a study in the vicinity of the Portsmouth nuclear facility that might be of use. In a neighboring ecoregion (Interior Plateau), MBI has data from around 100 or so streams around Hamilton Co., although many have urban impacts.	The Ohio EPA's Primary Headwater Assessment data were included in the case study of applicability to ecoregion 70 in Ohio. The data set included sensitive genera. EPA did not have access to either of the other two mentioned data sets.
Reviewer 5	I answer this question with two caveats: (1) I am not trained in aquatic ecology and have no professional experience dealing with the subject of this question, and (2) I am well aware of legal issues concerning "Waters of the United States," and that the role of ephemeral streams within that framework is at issue; in answering this question, I take no position on that issue.	No response required.

TABLE 4. RESPONSE TO CHARGE QUESTIONS 9−12: GEOGRAPHIC APPLICABILITY (continued)		
Charge Question 12: Applicability To Ephemeral Streams: In their 2011 review of the EPA Benchmark Report, the SAB indicated that because the data used to derive the benchmark were collected from perennial streams, the empirical relationship between conductivity and genera occurrence likely would be applicable to perennial and intermittent streams, but not to ephemeral streams. In preparing the current draft document, EPA found several publications that indicate that some aquatic organisms on which the Case Study example criteria are based do occur in ephemeral streams and that these organisms are critical to these headwater systems. EPA also believes it appropriate to include ephemeral waters as applicable water bodies for field-based conductivity criteria in order to ensure protection of aquatic communities in downstream intermittent or perennial waters. Therefore, EPA considers the field-based method applicable to all types of flowing waters, including perennial, intermittent, and ephemeral streams. Do you believe that this recommendation is well supported by Section 3.6.2 (Waterbody Type), including the publications it cites? Are you aware of any additional published studies or publically available scientific reports or data (e.g., paired chemical and biological sampling) relevant to this issue?		
Reviewer	**Comment**	**Response**
Reviewer 5	Because I lack experience with the precise issue, I reviewed several articles on the topic including some cited by the document and some not. In my reading of the literature, it became clear that there is significant overlap among taxa residing in “temporary” streams (as ephemeral streams are often called in these studies) and those residing in permanent streams. Therefore, it is reasonable to expect that HC_{05} levels derived from analysis of biological data collected from low-order permanent streams will be protective of most taxa occurring in ephemeral streams.	Thank you for your comment. EPA agrees that criteria derived using this method are applicable to flowing fresh waters, including ephemeral, intermittent, and perennial streams (see Section 3.6.2).
Reviewer 5	However, because some taxa occurring in low-order permanent streams do not occur in ephemeral streams, and vice versa, it is not reasonable to expect that the HC_{05} levels derived from low-order permanent streams to achieve the goals of the document’s HC_{05} derivation method exactly. Hence, it is my view that a process for applying HC_{05} levels derived from permanent streams to ephemeral streams would require biological confirmation. One method of biological confirmation could be to verify the presence of the permanent streams’ limit-defining taxa within the ephemeral streams; other methods are also possible.	Using a taxa list published by the reviewer with Timpano et al. (2011) for headwater streams in Ecoregion 69 of Virginia, EPA estimated the proportion of genera present in those streams that would be extirpated at the proposed HC_{05} for Ecoregion 69 and compared to the effects at the reported test sites to predicted extirpation based on the SC at the site. The proportion extirpated is greater than 5% because many of the genera are salt-intolerant. This constitutes a field confirmation that genera inhabiting ephemeral streams are at least as intolerant or even more so.

TABLE 4. RESPONSE TO CHARGE QUESTIONS 9−12: GEOGRAPHIC APPLICABILITY (continued)		
Charge Question 12: Applicability To Ephemeral Streams: In their 2011 review of the EPA Benchmark Report, the SAB indicated that because the data used to derive the benchmark were collected from perennial streams, the empirical relationship between conductivity and genera occurrence likely would be applicable to perennial and intermittent streams, but not to ephemeral streams. In preparing the current draft document, EPA found several publications that indicate that some aquatic organisms on which the Case Study example criteria are based do occur in ephemeral streams and that these organisms are critical to these headwater systems. EPA also believes it appropriate to include ephemeral waters as applicable water bodies for field-based conductivity criteria in order to ensure protection of aquatic communities in downstream intermittent or perennial waters. Therefore, EPA considers the field-based method applicable to all types of flowing waters, including perennial, intermittent, and ephemeral streams. Do you believe that this recommendation is well supported by Section 3.6.2 (Waterbody Type), including the publications it cites? Are you aware of any additional published studies or publically available scientific reports or data (e.g., paired chemical and biological sampling) relevant to this issue?		
Reviewer	**Comment**	**Response**
Reviewer 5	Articles reviewed to reach this opinion include DeJong et al. (2013), del Rosario and Resh (2000), Delucchi (1988), Feminella (1996), Grubb (2010), Price et al. (2003), Stout and Wallace (2003), and Williams (1996).	All of these papers were reviewed by EPA and some were cited in the document.

TABLE 5. RESPONSE TO CHARGE QUESTIONS 13−14: SUPPORTING INFORMATION: *FIELD-BASED HC_{05} FOR FISH IN APPALACHIAN STREAMS (SEE APPENDIX G)*

Charge Question 13: Fish: The method used to derive the fish HC_{05} generally followed the same field-based method used to derive the macroinvertebrate HC_{05} described in the Analysis Plan (see Section 3) and in the original EPA Benchmark Report. However, different data sets were used in the fish analysis (see Appendix G, Section 2), and some modifications to the method were required to account for differences between fish and macroinvertebrate natural history; e.g., modification to the boot-strapped statistical approach used to characterize uncertainty in the fish XC_{95} and HC_{05} values (see Appendix G, Section 3.4). Please comment on the sufficiency of the data set and the clarity and validity of the modified method to derive the fish XC_{95} and HC_{05} values.

Reviewer	Comment	Response
Reviewer 1	The process is clearly described. I have a number of concerns. First, the fish sampling methods (sampling gears and distance) used by different agencies/programs are not detailed, but likely inconsistent. For example, if one set of samples were collected over a reach of 40-time channel width, but another over 20-time channel width, the former likely capture more species. As a result, the occurrence frequency of a species may vary with sampling method or data sources, introducing noise into the analysis. The sample comparability of the various sources needs to be evaluated. Second, the sample sub-setting based on major basins is well defensible, but much less so when based on stream size. Although many fish species prefer streams of certain sizes, they also occur in streams of different sizes. Any numerical thresholds seem arbitrary. The modification to the bootstrapping process is reasonable; however, my earlier criticisms to XC_{95} estimation, its relevancy to species extirpation and SD curves are applicable here.	The HC_{05} derivation method relies on the observation of a taxon at an SC level not on its frequency of capture at a site. Although uniform methods would be ideal, the differences would not bias the estimates unless they were biased with respect to the occurrence of tolerant and intolerant species. Fish were sampled across a wide range of SC and there is no reason to believe that sampling method introduced bias. EPA calculated the HC_{05} with and without weighting for stream size and found no statistical difference. In response to this comment, EPA changed the exclusion criteria; sites with no fish and a few sites that had erroneous units were excluded. See also EPA's earlier response to comments related to XC_{95} values in the salt-tolerant portion of the SD (> values).

TABLE 5. RESPONSE TO CHARGE QUESTIONS 13–14: SUPPORTING INFORMATION: *FIELD-BASED HC_{05} FOR FISH IN APPALACHIAN STREAMS (SEE APPENDIX G)* (continued)

Charge Question 13: Fish:

Reviewer	Comment	Response
Reviewer 2	I believe that the work done to derive a fish HC_{05} was exceptionally well done. Data sets were very large and a number of considerations which are fish-specific were incorporated. These provide validity to the modified method to derive the fish XC_{95} and HC_{05} values. The suitability of the method—to be applicable to fish as well as macroinvertebrates—is especially reflective of the quality of this work and its usability for widespread application to aquatic organisms. The following are thoughts and notes which I made as I progressed through the section.	Thank you for your comment. No response needed.
Reviewer 2	• On page G-3, the section accurately makes the connection between loss of macroinvertebrates (impacted by high conductivity levels) and the stress this puts on fish by decreased food availability. This is clearly an additional justification for a conductivity criterion—fish losses reduce the natural resource quality of the stream, reduces recreational potential, increases the number of threatened and endangered species and possible extirpation.	Thank you for your comment. No response needed.
Reviewer 2	Should other stressors be included in the above discussion? For example: poor habitat, lack of water depth diversity, high ammonia levels, high suspended solids, low DO … all are stressful to fish either directly or indirectly. This weakens fish so that the effects of other stresses, such as the ionic imbalance from high conductivity levels, are likely enhanced. In other words, should other types of stress be taken into the analyses with conductivity because of the increased level of impact there might be?	As with any pollutant, physiological and toxicological stress from multiple exposures may increase effects. A confounding analysis for pH, temperature and habitat did not affect the HC_{05} for fish or invertebrates and can be found in Appendix G. No change was made.

TABLE 5. RESPONSE TO CHARGE QUESTIONS 13−14: SUPPORTING INFORMATION: *FIELD-BASED HC_{05} FOR FISH IN APPALACHIAN STREAMS (SEE APPENDIX G)* (continued)

Charge Question 13: Fish:

Reviewer	Comment	Response
Reviewer 2	• Of the six fish species listed as relatively tolerant of elevated conductivity, pg. G-4, lines 4−7, I believe that *Micropterus dolomieu* (smallmouth bass) perhaps shouldn't be included. It is generally intolerant of pollutants and poor water quality. Even *Lepomis cyanellus* (green sunfish) prefers somewhat good conditions, even though it can be found in euthrophic waters.	The reviewer recommends removal of some species. These species meet data requirements for the method, are native to the area, and therefore EPA cannot justify their removal.
Reviewer 2	• Identification of fish to the species level is indeed the routine for fish sampling, unlike macroinvertebrates, and this does lend itself to species-level XC_{95} values.	Thank you for your comment. No response needed.
Reviewer 2	• The fish analyses used a combined data set for fish from portions of four contiguous ecoregions and seven states! Seven data sets collected between 1991 and 2009, 1,657 sampling events across 1,364 distinct sites, gives great spatial and temporal data. What an amazing quantity and variety of data! This extensive data base is difficult to find in the environmental arena. Kudos for bringing together such an excellent base from which to assess for a fish conductivity criterion.	Thank you for your comment. No response needed.

TABLE 5. RESPONSE TO CHARGE QUESTIONS 13−14: SUPPORTING INFORMATION: *FIELD-BASED HC_{05} FOR FISH IN APPALACHIAN STREAMS (SEE APPENDIX G)* (continued)

Charge Question 13: Fish:

Reviewer	Comment	Response
Reviewer 2	A concern: Reference sites were not identified in the dataset, however the document seems to imply that 134 sites, which were >90% forested, were likely such. What is the reason for not identifying, and assuring, that reference sites were included? Could the data not have been identified, perhaps as a separate grouping within the dataset? How can you be sure that there were adequate reference sites in the initial sample collections? Heavily forested does not necessarily assure that water quality will be of high quality. This exact situation happened with a monitoring program which I designed and implemented for the stream system running through Omaha, Nebraska. After careful searching, I found what appeared to be a minimally impacted small stream of which for most its length it flowed through rolling wooded hills. Although some small acreages, occasional houses, and further back, a new development were in among the hills and woods, the stream appeared to be minimally impacted and exhibited great fish and macroinvertebrate habitat. It appeared to be the perfect reference stream (there was no existing data available to use as a guide). Eventually, it became exceedingly apparent that I chose poorly as fecal coliform levels repeatedly were some of the highest of the 24 sampling sites in the system and some of the other parameters were also not appreciably better than any of the other sites.	The developers of the individual data sets did not identify reference sites. EPA assumed that, overall, sites with >90% forest in the watershed were of generally high quality and those data were included in the analysis. These data were deemed adequate for evaluating whether fish, and thus the aquatic community as a whole, would be protected by criteria developed using macroinvertebrate data alone.
Reviewer 2	• Description on pg. G-4, Lines 15−28 is very thorough; provides the reader with a detailed understanding of the area, conditions, etc. In G-7, there is good use of excluding larger catchment areas, sites with low pH and high chloride levels, and sites which were too small to support fish. This strengthens the data analyses. The discussion of fish on pg. G-10 regarding the exclusion of sites where there is question of the presence or absence of a species is also good. I am surprised to read that brook trout (*Salvelinus fontinalis*) are stocked. While they are native, they are few in number and are especially sensitive to poor water quality conditions. I believe that I disagree that they should be	Additional text has been added to explain the many uncertainties surrounding XC_{95} values associated with stocked fish. EPA added a large data set from Pennsylvania that improved the estimates especially for brook trout.

TABLE 5. RESPONSE TO CHARGE QUESTIONS 13−14: SUPPORTING INFORMATION: *FIELD-BASED HC_{05} FOR FISH IN APPALACHIAN STREAMS (SEE APPENDIX G)* (continued)

Charge Question 13: Fish:		
Reviewer	**Comment**	**Response**
	included. I have concerns about counting any of the stocked fish species because of the possibility of affecting the XC_{95} estimates. Stocking is a manmade “condition”, largely to improve recreational fishing; the expected life span is pretty much irrelevant and independent of the stream’s condition.	
Reviewer 2	The paragraph on pg. G-11, lines 9−15, is not fully clear to me. It would be helpful to have it explained a bit more fully. The following lines on that page, lines 16−31, are well done. I would say, however, if a sensitive taxon is found in a waterway in which it is unexpected (outside the distribution of that species), there is the possibility that perhaps the species range had not been accurately established originally. Or that it had expanded its range—either way, it would probably be best to check with local fish biologists before exclusion.	The reviewer suggests a more complex method to estimate distribution of species. Although the method affects the weights because some species do not occur throughout the full range, calculation without weighting had little to no effect. No change was made.
Reviewer 3	Unless I am missing something, there is no Section 3.4 in Appendix G.	All formatting was checked and made consistent with EPA standard format.
Reviewer 3	So, there are many obvious differences between benthic and fish data in assessing conductivity effects. The benthic data is at the genus level—good enough, but fishes are easily identifiable (with a few exceptions) to the species level. However, fish sampling is more time consuming so not as many samples are collected in comparison to the benthic collections. There are a number of other considerations as well discussed on pages G-1 and G-2.	No response required.
Reviewer 3	I liked the approach where fish data was lumped from four ecoregions, which resulted in a data base of over 1,437 observations. The clarity and validity of the modified method was adequate to derive the XC_{95} and HC_{05} values. I also liked the data filters that were employed for the analysis—these eliminated a lot of potential problems with the data analysis.	Thank you for your comment. No response needed.

TABLE 5. RESPONSE TO CHARGE QUESTIONS 13−14: SUPPORTING INFORMATION: *FIELD-BASED HC$_{05}$ FOR FISH IN APPALACHIAN STREAMS (SEE APPENDIX G)* (continued)

Charge Question 13: Fish:

Reviewer	Comment	Response
Reviewer 3	OK—here is where I am unhappy with the fish data. First, both rainbow trout and brown trout must be excluded from the data set. These are exotic, introduced species and even through there are established populations of these two species, that is not a good reason to include them. Many folks are trying hard to protect native species throughout the Appalachians, and including them as well as carp just does not make sense. I would follow the listings for introduced and exotic species, as found in Wiley and Hocutt, as the cut for potentially introduced species into an ecoregion. Also, just because the two trout species are recreationally important, that is not a good enough reason to include them in the analysis. If the work done by Tim King is valid, then there are only a few places where one needs to worry about brook trout introductions. After all, this is the native salmonid of the Appalachians.	EPA's longstanding practice is to include and protect U.S. nonnative resident species as well as native species. All species in the sensitivity distribution can be considered surrogates for species occurring in the field. Carp and brown trout are examples of nonnative resident species that were retained in the analysis, based on this long-standing EPA practice (see Appendix 1 of EPA's Guidelines for deriving Numerical National Water Quality Criteria for the Protection of Aquatic Organisms and Their Uses).
Reviewer 3	I like Figure G-9 very much. First, it showed the species sensitivity distributions for many fish species. More importantly, it illustrated that even within a genus, there was wide variation in the response to conductivity, e.g., *Etheostoma* and *Cottus* sp. The derivation of the fish HC$_{05}$ is excellent, but the hazardous concentration of 392 μS/cm seems a little high to me, but that may be a reflection of the species that are common in my research sites.	Thank you for your comment. No response needed.
Reviewer 4	Again, I generally found no real problem with the statistical approach and found the modified bootstrap methodology reasonable. My main concerns are related to how the regions are combined, given biogeographical differences in fish distributions across ecoregions. The argument is made that, in a manner similar to SSDs generated for toxicity testing, it is not important that the species that make up those below the HC$_{05}$ do not occur in Ohio. I am not sure this is reasonable for a natural "stressor" such as conductivity. The benchmark is driven by coldwater and rare species that do not occur in Ohio. If this analysis was conducted with Ohio ecoregion 70	These data were deemed adequate for evaluating whether fish, and thus the aquatic community as a whole, would be protected by criteria developed using macroinvertebrate data alone. In order to develop ecoregional-specific criteria based on fish, a new analysis would be needed.

TABLE 5. RESPONSE TO CHARGE QUESTIONS 13–14: SUPPORTING INFORMATION: *FIELD-BASED HC_{05} FOR FISH IN APPALACHIAN STREAMS (SEE APPENDIX G)* (continued)

Charge Question 13: Fish:

Reviewer	Comment	Response
	data alone, perhaps with a lower threshold of sample size, perhaps it is possible that other sensitive fish species would replace the most sensitive taxa in Appendix G. It is likely, however, that those sensitive taxa are inhabitants of the EWH tiered use in Ohio rather than the WWH use. Again, I think that some discussion of tiered uses is essential to how a State might apply this approach.	
Reviewer 5	I have no professional expertise or activities that concern fish. I can see that the data analysis procedure is similar. However, my scientific knowledge of fish and of environmental characteristics that influence their condition and behavior is insufficient to allow my informed opinion about this part of the document.	No response required.

TABLE 5. RESPONSE TO CHARGE QUESTIONS 13–14: SUPPORTING INFORMATION: *FIELD-BASED HC_{05} FOR FISH IN APPALACHIAN STREAMS (SEE APPENDIX G)* (continued)

Charge Question 14: Protection: Do the analyses for fish (see Appendix G) demonstrate that the Case Study example criteria (based on macroinvertebrate data) are protective of fish in those areas? If not, please describe why and any additional data and analyses needed.

Reviewer	Comment	Response
Reviewer 1	Just as the case studies for macroinvertebrates, the conductivity criterion derived here appears to be protective for fish in the study region in practice. However, EPA needs to address the lack of biological relevancy of XC_{95} to species extirpation and vague interpretation of SD curves, as I described earlier.	See EPA's response above to previous comment. Because the fish are analyzed at the species level, almost all have a clear unimodal distribution and so the interpretation of the XC_{95}s is clear for more taxa. As mentioned previously, the SD curves are only illustrative. It is the 5th centile rank order that determines the HC_{05}. No response or change required.
Reviewer 2	Yes, the analyses for the fish example criterion is protective of fish. The HC_{05} of 392 µS/cm (95% CI 256–424 µS/cm) is appropriate. Good discussion in the Results, and as I observed in Question #13, many strong attributes accompany the fish analyses.	Thank you for your comment. No response needed.
Reviewer 3	It appears the case study criteria for the benthic organisms would also be protective of fish populations. I don't think any additional data and analyses are needed. I look at this approach as a two-factor method where the benthic conductivity criteria would drive the fish protection, and perhaps *vice versa* in special situations.	Thank you for your comment. EPA agrees that the analysis in Appendix G demonstrates that the macroinvertebrate method is protective of fish in those areas.
Reviewer 3	FINAL COMMENT: I would like to have seen a little more done with assemblages, e.g., EPT, intolerants, tolerants, etc. However, the genus level for bugs and the species level for fish approaches are great, especially with the highly robust data sets.	Thank you for your comment. The field-based method does not use metrics or assemblages because they reduce the discriminatory power of the analysis. A few analyses compared the number of ephemeroptera genera and SC in an earlier work (EPA 2011) but it does not provide new information that is not also shown by analyses of effects to individual genera (invertebrates and fish) or species (fish). EPA agrees that it would be interesting to see more work in the future.

TABLE 5. RESPONSE TO CHARGE QUESTIONS 13–14: SUPPORTING INFORMATION: *FIELD-BASED HC_{05} FOR FISH IN APPALACHIAN STREAMS (SEE APPENDIX G)* (continued)

Charge Question 14: Protection: Do the analyses for fish (see Appendix G) demonstrate that the Case Study example criteria (based on macroinvertebrate data) are protective of fish in those areas? If not, please describe why and any additional data and analyses needed.

Reviewer	Comment	Response
Reviewer 4	My experience with conductivity and other ionic strength parameters is that macroinvertebrates are generally more sensitive as a group than fish, but I think because of sample size considerations the most sensitive fish are often being excluded from the analyses here. The limited distribution of certain fish species (and macroinvertebrate genera) that are often excluded may in themselves be evidence that the aquatic life potential may vary with some natural as well as anthropogenic impacts. As monitoring programs mature, samples sizes in these areas are continually growing, so I think that State water quality standards programs need to be continually exploring their databases to refine aquatic life uses and the criteria designed to protect these uses. Modification to the approach for deriving a single ecoregion criterion for all streams requires an adequate monitoring program with robust critical program elements (Yoder and Barbour, 2009) that provide the data and the capability to conduct the analyses described in this document. For example, a program needs to be capable of accurately classifying and controlling for natural features that influence biological assemblages (e.g., stream gradient, size, elevation, base flow, etc.). Tiered uses require robust data with the ability to recognize the influence of anthropogenic influences on the landscape and the ability to address what stressors may or may not be feasibly controllable.	The reviewer concurs that the invertebrate method would be protective of fish. However, the reviewer also cautions that the fish analysis may not have included the most sensitive fish due to rarity and the need to have at least 25 occurrences. EPA constructed a data set of ranges for species observed at fewer than 25 sites. The maxima of these were higher than the relevant HC_{05} derived using invertebrates, which suggests the invertebrate method is protective of fish. EPA will consider additional information as it becomes available in any future update to the field-based method.
Reviewer 5	Same answer as for number 13.	No response or change required.

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT

Reviewer 1

Page number	Line(s)	Reviewer comment	EPA response
xviii	—	On definition of XC_{95}. It is not clear what do the authors mean by "effectively absent" here.	The sentence currently reads: "First, a weighted cumulative distribution function (CDF) is developed for each genus to determine the genus extirpation concentration (XC_{95} or 95^{th} centile of the distribution of the occurrences of a genus), the level of exposure above which a macroinvertebrate genus is **effectively absent** from water bodies in a region." The preceding clause in parentheses defines the mathematical definition. The definition of extirpation in the glossary reads: "Extirpation**—**The depletion of a population of a species or genus to the point that it is no longer a viable resource or is unlikely to fulfill its function in the ecosystem." No change was made.
xix	—	First, this definition of extirpation is very different from what commonly used in the literature of conservation biology (absent from a region or regions, but still present somewhere else), and then may cause confusion.	EPA uses a definition of extirpation in this and other documents that does not require that a taxon be completely absent but does include that condition within its scope. It is the functional absence of a taxon, which includes the state of physical absence.
		Second, it is also hard to determine when a genus is no longer a viable resource or unlikely to fulfill its functions, particularly when only 200 individuals are counted from a site.	EPA does not attempt to determine whether a genus is extirpated at each individual site so the method does not depend on the adequacy of 200 individuals to reveal extirpation. Rather, the analysis determines the SC at which a genus is extirpated, based on the regional relationship between occurrences and SC.

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)

Reviewer 1

Page number	Line(s)	Reviewer comment	EPA response
1-1	8	" . .waters dominated by . . " Do the authors mean "waters naturally dominated by"? If not, one would not be able to apply the method to streams contaminated, say, by NaCl from road deicing. Clarify.	The intent is that derived example criteria in the case studies would apply to waters dominated (mg/L) by calcium, magnesium, sulfate, and bicarbonate ions. Chloride, sodium, and potassium may account for half the total concentration of salts which does include some, but not all deicing situations. Because the example criteria were developed for waters with a specific ionic composition (i.e., dominated by calcium, magnesium, sulfate, and bicarbonate ions), they are not generally applicable to waters with different ionic compositions (e.g., waters dominated by chloride anions).
1-2	11−12	If any studies/data show Ecoregion III effectively capture the natural variation of conductivity across space, cite them. If not, the authors need to justify this decision.	Griffith et al. (2014) is cited as addressing natural variation. The document also points out that some local exceptions may occur.
2-1	9−11	Here the authors set the threshold of conductivity for extirpation as the level below which 95% observations occur. Above this conductivity level, the taxon is assumed to be no longer a viable resource or unlikely to fulfill its function. However, how they actually did this is much more complex (P3-13), They can leave the details to later, but they need to give readers some idea about how it is actually done. Otherwise, one may reject their method right away because above this threshold, a genus may be still common and viable!	As the reviewer noted, EPA recognized that the full range of tolerance would not always be sampled. Therefore, the XC_{95} value is demarcated as ambiguous or undetermined if the upper SC range of the genus is not at zero occurrences. For added clarity, in Figure 3-1, for those genera whose occurrence do not decline to zero with increasing SC, the XC_{95} value is plotted differently, as a triangle, based on a later suggestion by the reviewer.
2-10	Figure 2-1	Does increase of ion concentration always lead to decline of macroinvertebrate and fish species? I thought that the relative or true abundance of tolerant species	As is indicated by the GAM plots and by the designation of the values for some species as greater than, not all increases in concentration lead to declines of the occurrence of a particular invertebrate or fish species. Abundance is not the

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)

Reviewer 1

Page number	Line(s)	Reviewer comment	EPA response
		may increase, just as shown in the case studies (B19−30). Modify.	endpoint, and is not analyzed. The Figure legend was edited to read, "Conceptual model shows hypothesized relationships among selected sources of ions and biotic responses to ionic stress by salt intolerant taxa..." Also, Section 2.4 of the public review draft document goes into detail about tolerance ranges and optima.
2-15	21−22	See my earlier comments regarding species extirpation.	See previous response.
2-17	22	The authors need a newer citation, if available.	This establishes the phenomenon has been known for a long time. The 1992 paper is more recent. It is only one of several possible mechanisms. No change was made.
2-19	1	Replace “many states” with “most states”(?)	Changed to read "...most states and tribes monitor…"
2-19	17	“Freshwater insects are among the most sensitive. .” This statement is too general. Freshwater insects differ greatly in their sensitivity to human disturbances. Most EPT species are sensitive to organic pollution, but most chironomid genera are tolerant and so are most other dipterans (true flies). Modify.	The text was rewritten as follows: “These field-based methods can be used to develop ecoregional criteria that are fully protective of aquatic life. Many freshwater insects are among the most salt-intolerant organisms relative to other taxa, including crustaceans such as crayfish and daphnids, fish, and amphibians (compare Appendices A.4 and B.4 with Appendix G of this report). Less is known about the salt-intolerance of mussels to ionic stress, but recent studies suggest that mussels in the family Unionidae are acutely sensitive to salts (Kunz et al., 2013), particularly during early (glochidia) life stages (Bringolf et al., 2007; Gillis, 2011).”
2-20	17−18	“In other words, . . .value.” This threshold may make sense if a taxon rapidly decreases with conductivity; however, it makes no sense if a taxon increases or does not change with conductivity.	Text was added, "In other words, the probability of observing a genus above its XC_{95} SC value is 0.05. XC_{95} values that are uncertain or undetermined because the genus does not decline to near-zero occurrence within the exposure range are noted and generally do not influence the hazardous concentration

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)			
Reviewer 1			
Page number	**Line(s)**	**Reviewer comment**	**EPA response**
			(HC_{05}) because their estimated XC_{95} values are greater than those genera in the 5th centile. "
		The authors also state that "In other words, the probability of 0.05 that an observation of a genus occurs above its XC_{95} conductivity value." Does this statement really hold when observations in a large bin are down—weighted in calculating XC_{95}?	The weighted and unweighted HC_{05} are similar. No change was made.
		Even if this statement is valid, it is still confusing. Readers might interpret it as that one should expect to capture a genus at 5 of 100 sites where conductivity is greater than its XC_{95}. The authors need to give a clearer and biological meaningful interpretation of XC_{95}, if they want to use this term.	The sentence in Section 2.6.2 has been edited to read: "In other words, the probability of observing a genus above its XC_{95} SC value is 0.05; i.e., if a genus is observed at 100 sites, only 5 sites would be expected to have SC above the XC_{95}." A concentration leading to the effective loss of a genus from the environment is biologically meaningful.

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)

Reviewer 1

Page number	Line(s)	Reviewer comment	EPA response
3-3	Figure 3-1	This figure is confusing. First, when 85% of the samples contained *Cheumatopsyche* in the bin with conductivity >1,000 µs/cm, how is the genus assumed to extirpate to close to do so? Second, even when the "probability of capture" of a genus declined with increasing conductivity, the taxon may be still be common at its XC_{95}, approximately 20−40% of the samples for *Stenonem* and *Leuctra*, respectively in the case study. Considering the impact detectability of a genus associated with small sample size (200 counts) and limited sampling period (once a year), the probability of occurrence could be much higher. One can argue that both genera are strong, at least far away from extirpation. It appears difficult, if possible at all, to consistently relate the extirpation of a genus to its XC_{95}. The authors addressed this issue later in pages 3-12 and 3-13, but they need to give a full treatment when interpreting the figure or when introducing XC_{95} in P 2-20.	Some XC_{95} values are derived for declining species and others have indeterminate values. Only the XC_{95} values in the 5% range affect the HC_{05}. The figure legend has been changed to read: "Example of a genus extirpation concentration distribution (XCD) depicting the proportion of genera extirpated with increasing ionic concentration measured as specific conductivity (SC)**.** Each point on the XCD plot represents an extirpation concentration (XC_{95}) value of one genus arranged from the least to the most salt-tolerant. XC_{95} values that were defined as greater than values are indicated by triangles. The 5th centile of the XCD is shown as a dotted horizontal line. The 5th centile hazardous concentration (HC_{05}) is the SC at that intercept of the XCD and the 5th centile line. In this example, the HC_{05} is 305 µS/cm."

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)

Reviewer 1

Page number	Line(s)	Reviewer comment	EPA response
3-3	Figure 3-1	I am also concerned about their use of the term, "probability of capture". In the literature, two terms are typically used to describe observations of a species, occurrence probability and detectability. The former describes the probability of a species to occur at a site (any spatial unit). The later refers to the probability of a species to be detected when it is present at a site. "The proportion of samples with a genus present at a conductivity level" could be taken as an estimate of occurrence probability only if detectability is assumed to be 1, something hard to justify. I suggest the authors to use a term like relative frequency, rather than probability of capture, which has been commonly used to refer to the % of individuals captured by a sample. In addition, the authors estimated the proportion of genus observation for a conductivity bin, rather than a conductivity level. Modify and clarify.	The method does not rely on either occurrence probability or detectability. The GAM plots are based on probability of capture; however, these values are not used to derive the XC_{95}. The GAM plots were removed from Figure 3-1 of the public review draft to reduce confusion. EPA revised the text to "Probability of Observing" and the legends were edited to indicate that West Virginia samples are based on identification of 200 individuals.
3-5	20	"background ... region;" It would be helpful to clearly state how similar conductivity among reference sites is similar enough.	Determination of background is performed in a tiered approach. If the background based on simple methods such as 25^{th} centile of sites or references sites is unclear, a weight of evidence can be used. A call out has been added: "background SC levels are similar throughout the region (see 3.7.1)…"
3-8	8−12	One major source of salinity in freshwater waters snow zone is road de-icing. The conductivity criteria described here will not be applicable to assess the impact of NaCl used for de-icing? (also see my earlier comment on this issue)	The method itself is not specific to one particular ionic mixture. However, the example case studies in the draft document apply to waters dominated by calcium, magnesium, sulfate, and bicarbonate ions. EPA does not currently have a

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)

Reviewer 1

Page number	Line(s)	Reviewer comment	EPA response
			large enough data set with NaCl-dominated sites to develop example field-based criteria for NaCl.
3-10	26−31	See my comments on the relationship between XC_{95} and extirpation earlier.	The section reads, "For each genus meeting the data-selection conditions, a CDF is constructed that is weighted to correct for any potential bias from the unequal distribution of sampling of sites across the range of logarithm10 transformed SC values. This weighted CDF represents the proportion of observations of a genus with respect to increasing exposure levels. The extirpation effect threshold for a genus is 95% of the total occurrences of the genus. The two exposure levels bracketing the 95th centile are linearly interpolated to give an XC_{95} for a genus." No change made except to clarify that the SC values were log transformed.
3-11	1	Did the authors assess how bin delineation affected CDF? The description here is a bit vague regarding how they balanced the number of bins and the size of bins. Clarify.	Yes, this was one of many sensitivity analyses that EPA performed; however, it was not judged to be sufficiently important or useful to include in the draft document. EPA has left bin selection to the analysts' best professional judgment. No change was made.
3-11	eq 3-1	"*x*" needs to be defined.	X is defined as the stressor. "xij is the stressor value in the j^{th} sample of bin *i*."
3-12	Figure 3-3	Replace "cumulative probability" with "cumulative proportion" to be consistent with the text (L 2), and avoid the issue of imperfect detectability.	Change was made globally. Cumulative Proportion for CDF and Probability of Observing for GAM are used in the public review draft.

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)

Reviewer 1

Page number	Line(s)	Reviewer comment	EPA response
	Figure 3-3	I also suggest adding a third panel figure to show a concave increasing curve like the one for Corydalus (B-42). This type of curve really means a positive response of a genus to conductivity (B-29 for the same genus). For the 2nd (Nigronia) and 3rd types of curves (Corydalus), it is not possible to relate XC_{95} to extirpation, as I argued earlier. Glad to see the authors starting to address this critical issue here. However, it needs to be fully treated much earlier. I also do not see it being a data-distribution issue, but a fundamental limitation of CDF. CDF curves of Types 2-3 are also not anomalies, but they are normal and frequent, as shown in the case studies. Revise.	This section is not intended as a review of all possible forms of CDF curves. Rather, it highlights an example of a genus with a clear XC_{95} versus one whose range does not include its upper tolerance to SC. A third plot was not added. The concern by the reviewer regarding the fundamental limitation of CDF curves is addressed by a second analysis using GAM plots as the reviewer notes in the next comment on 3-13 regarding the value of using qualifying designations.
3-13	1−7	Yes, the qualifying designation helps for understanding HC_{05}, but relating XC_{95} to extirpation remains conceptually flawed. See my comments to Charge Question 4 for possible options.	In charge question 4, the reviewer suggests two options that have not been vetted in the scientific community. Use of extirpation thresholds was rejected because when the full range of exposures are evaluated, all species have a unimodal distribution so using only genera without designations would be entirely dependent on the range of exposures which varies among data sets and regions (see GAMs for fish species). The second suggestion suffers from regression to the mean and the range of the exposures in the data set as described above. EPA has chosen to retain the method reviewed and approved by the SAB that uses the 5th centile of affected taxa, an assessment endpoint consistent with Agency guidance for aquatic life criteria development.
3-13	21−22	Replace "mean curve" with "fitted curve". Also, what is the confidence limit? 95% or 90%? Clarify.	Edited as recommended (mean retained, fitted added, 90% added).

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)

Reviewer 1

Page number	Line(s)	Reviewer comment	EPA response
3-17	1−3	A further concern is whether sampling dates/period is related to conductivity. If most high-conductivity sites were sampled in spring (March−June), but low-conductivity sites in summer, one likely underestimates the occurrences of sensitive taxa in the latter and then overestimate HC_{05} conductivity. Correlation between conductivity and sampling time can be used to identify the bias.	The seasonal conductivity regime in the example case studies is reversed from the hypothetical posed by the reviewer. The sampling in Case Studies I and II was unbiased, but care should be taken when applying the method elsewhere. Analyses are provided for adjusting for seasonality when needed and examples are given in Appendices A and B.
3-19	eq 3-2	This equation needs to be written in a standard math format as follows: $CMEC = 10^{(\overline{X} + z_{\alpha} \times \sigma_{r})}$	Reformatted as recommended.
3-21	10	" ... and often more than 4 days". Above CCC? Clarify.	Text was added. "Based on available field data, salt-intolerant macroinvertebrate genera may be exposed to a range of SC levels greater than the CCC throughout the year and often for more than 4 days (see example in Figure 3-7)."
3-19	11−14	Is "the one-tail critical value" half of the number of the standard deviation required for 90% confidence limits?	Reformatted as recommended.
		The authors also define $\overline{X}$ twice here and differently. Clarify or correct.	Reformatted as recommended.
3-26	27	"More than 90% of ... insects." This statement is too broad. In many streams, insects took less than 90% of all individuals. Add "often" or "frequently".	EPA does not know of exceptions, but changed to "often."
F-16	eq F-1	Re-write the equation in a standard math format	Reformatted as recommended.

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)

Reviewer 1

Page number	Line(s)	Reviewer comment	EPA response
G-3	7−8	This statement is too broad. Many adult insects, such as winter stoneflies, actually only move over a short distance.	Edited to read, "Fish may be absent because of limited interbasin dispersal in contrast to the winged stages of most aquatic insects which permit them to disperse among disconnected basins~~ and occur wherever water and habitat quality are suitable~~."
G-10	3	Add "hybrids" after "immature specimen"	Such fish were generally immature specimens, and identifiable mature specimens of the species were generally present in the same sample.
G-18	14	Do the authors mean a selected **minimum** size ranging from 0−60 occurrence? If so, how it can be zero? Clarify.	This sentence was removed from the public review draft.

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)

Reviewer 2

Page number	Line(s)	Reviewer comment	EPA response
vii	5−4	Space needed between “of” and “survey”	Edited accordingly.
xi	5−10	Delete the “and” before Kentucky	Edited accordingly.
xvii	Continuing paragraph	Paired analyses is a strong statistical test. Exceptional number of field samples, sites and years of sampling!	Thank you for your comment. No response needed.
xvii and 2-3	10−12	Conductivity is described for eastern and western montane ecoregions but nothing is said about the Midwest—it’s a major portion of the mid-section of the country and probably should be included.	Edited to read, "SC tends to be low in most eastern and western montane ecoregions (25th centiles of SC <200 µS/cm), intermediate in the mid-continent (200−600 µS/cm), and sometimes very high in arid areas (>600 µS/cm (Griffith, 2014))."
xviii	Executive Summary	Very well done.	Thank you for your comment. No response needed.
2-4	25	Figure 2-1 is located five pages away; could it be moved closer to pg.2-4?	The figure does represent sources described on page 2.4, but as a method it is intended to show how to make a conceptual model. No change made.
2-3 to 2-7	Section 2.2.1	Thorough; excellent overview and foundational information	Thank you for your comment. No response needed.
2-8	Table 2-1	Parentheses should encase 2012 in: Samarina (2007); Ruhl et al. (2012).	Edited accordingly.
2-11	Sections 2.2.3, 2.3, 2.5	Also excellent information; valuable for water quality staff to better understand the causes and mechanisms.	Thank you for your comment. No response needed.
3--1	29−30	Could there be a bit more information with the “weighted CDF model”?	A sentence was added: "Weighting normalizes the distribution of samples taken across the conductivity gradient."

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)

Reviewer 2

Page number	Line(s)	Reviewer comment	EPA response
3-2	1−3	An accompanying short explanation of the statistical package R would be helpful.	The following sentence was added: "R is open-source and open-access computational software that runs on Microsoft Windows, Apple MacOS, and UNIX platforms."
3-10	Section 3.1.1.3.	Specific description of the sampling methods as well as assurances that adherence to standardized sampling techniques were observed, would be nice. Perhaps sampling details are in the Appendices.	The following text was added: "Some considerations include whether standardized quantitative or semiquantitative techniques are used…"
3-16	8−15	Good recognition of the variance in sampling protocols among different agencies or monitoring groups. My concern is whether this variability can be "handled" by the process? The authors believe that it does.	The method's ability to assess variability was favorably reviewed by the SAB and by EPA statisticians not associated with this project. No change made.
3-23	Section 3.4	Well done.	Thank you for your comment. No response needed.
3-27	10	Would not extirpation for the most sensitive crustacean occur before 960 µS/cm (of <960 µS/cm)? This is a high level of conductivity and mollusks are "canary" indicators of contaminants and water soluble stressors.	Some crustaceans have different cellular physiological means of ionic regulation and this may increase their ability to adapt to higher ionic concentrations. Mollusks were rarely collected by the sampling methods used in the example case studies, but they are believed to be protected by a criterion developed from the method based on some toxicity tests of glochidia and early life stage mussels reported in the literature. Some mussels are believed to be sensitive based on loss (Zipper et al., 2014, 2016) of mussels from the Clinch and Powell River. The sentence in Section 3.4.2 reads: "Unionid mussels were not evaluated by EPA, but some field and laboratory studies suggest that Unionidae are also salt-intolerant (Price et al., 2014; Gillis, 2011; Wang et al., 2013, Kunz et al., 2013)."

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)			
Reviewer 2			
Page number	**Line(s)**	**Reviewer comment**	**EPA response**
3-30	1−8	This was new information for me and find it very interesting. I did not know that Ephemeroptera can tolerate such low pH conditions if the conductivity is high. Good information; well done throughout all of Section 3.6.	Sentence was edited to be more clear because Ephemeroptera can tolerate such low pH conditions if SC is also <u>low</u> (not high): "...high SC has a greater effect than acidity on the occurrence of Ephemeroptera."
3-32	Figure 3-8	An amazingly weak correlation! I might not have expected this, but it is clear.	Thank you for your comment. No response needed.
3-34	7	"illiustrated" is misspelled.	Edited accordingly.
4-1 to 4-27	Section 4	Figures and tables are very helpful.	Thank you for your comment. No response needed.
4-8 & 4-9	Figures 4-3, 4-4	These figures show large increases in conductivity in October. Seems that these higher values would affect the calculations for HC_{05} and the HC_{05} when simply looking at the figures. I understand the explanations given but am not 100% sure that they are complete enough for the less-trained in statistics.	On Figures 4.3 and 5.3, the following sentence was added: "Please note the smaller scale on the *y*-axis compared to Figures 4-2 and 4-4, or 5-2 and 5-3."
4-27	4−11	Clarity of the use of the one day sampling/grab sample serving for CMEC and CCC.	Text on the relevance of one day grab sample was removed from the discussion of CMEC duration in Section 3.3.
5-4	Figure 5-1	Number of and distribution of sampling sites is exceptional.	Thank you for your comment.
5-6	12−15	Wouldn't the <200 µS/cm Dec through June and >200 July through Oct provide support for the argument that there is a seasonal difference, thus calling for separation of data by seasons or seasonal weighting?	Yes, seasonal differences are considered. See EPA responses to comments under charge question 3. Additional analyses were added and text to help explain this more fully. The effect of seasonal variability of SC on the subsequent analyses were further evaluated and presented in Appendix B.

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)			
Reviewer 2			
Page number	**Line(s)**	**Reviewer comment**	**EPA response**
5-7	Figure 5-2	Seasonal variation is clearly shown for September. Difficult to understand how this wouldn't skew the results.	See EPA responses to comments under charge question 3.
5-12	1st paragraph	Personal Comment: Here in the Midwest we have distinct seasons, and water quality parameters often reflect this. Having unweighted, monthly/seasonal data is helpful to state agency staff who are trying to determine sources of pollutants and causal relationships. Determining sources of impairment are challenging and a clear understanding of what is happening each month (when there is monthly data available!), provides insight.	Thank you for your comment.
5-14	Figure 5-7	I concur with the acceptability of the hazardous concentration of 338 μS/cm but some interest groups may believe it is too stringent.	Thank you for your comment.
5-18	Line 7	Delete the second "for."	Edited accordingly.
5-20	Figure 5-10	Delete "and" in the figure's title: "…southeastern Ohio into and Kentucky"	Edited accordingly.
A-1	Figure A-1	A sentence or two describing LOWESS would be very helpful.	LOWESS was added to the glossary.
A-4	Figure A-4	Good to address other water quality parameters; informative.	Thank you for your comment.

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)			
Reviewer 2			
Page number	**Line(s)**	**Reviewer comment**	**EPA response**
A-5	1st paragraph	My initial thoughts when reading this were that the confounders listed would have an effect on conductivity, and as I've stated in my response to one of the charge questions, I do believe that additional stressors can indirectly increase the damage done by high conductivity levels. Sorting it out, however, is an immensely difficult undertaking, requiring considerable data much uncertainty. However, in this document it was determined that confounders were not an issue.	The potential confounders are all stressors that cause the loss of species, but they do not affect the model of the HC_{05}. Cumulatively, they could affect a particular stream.
A-7	7−8	"Removal of poor habitat samples from the data set had almost no effect on the SD model or HC_{05}." Based on the work of this study, this appears to be true. Unfortunately, if the removal of poor habitat doesn't affect conductivity, then those who oppose habitat restoration projects can use this as an argument in support of their position.	Poor habitat, in general, can cause of loss of species. To assure that the genus extirpation concentration distribution (XCD) model was detecting effects from SC and not a response to poor habitat, the HC05 was recalculated using the example criterion data set in which samples were removed with an RBP score <135 total, pH <6 and fecal coliform ≤400 colonies/100 mL. Removal of poor habitat and high fecal coliform samples from the data set had almost no effect on the XCD model or HC05 (see Figure A 6). With this constrained data set the HC05 was 336 µS/cm (95% confidence interval (CI) 233−351 µS/cm). The confidence interval overlaps with the HC05 for the example criterion continuous concentration (CCC) (305 µS/cm 95% CI 233−329 µS/cm). Therefore, no correction was made for habitat quality or organic enrichment.
A-11	1−7	Weighting, by its very purpose, brings a comparable 'status' to a data set with variable values and is a method by which calculations can be made. But care must be taken to correctly do the weighting to ensure	Thank you for your comment.

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)			
Reviewer 2			
Page number	**Line(s)**	**Reviewer comment**	**EPA response**
		correct representation of the data is seen in the results. I believe the authors have endeavored to do it well.	
A-13	5	Not sure if this is recalculating to make data "fit" expectations. What was the RBP score used for the first calculations?	All sites were used for the calculation of the HC_{05} and the multivariate model. The removal of sites with low RBP scores ensures that the HC_{05} for SC is not due to habitat.
A-14	Section A.3.	Confidence intervals—there are some immensely large ranges.	Thank you for your comment.
B-9	1−5	Perhaps the RBP 130 score is not the correct level to use—could it have removed too many of poor and moderately poor sites?	To assure that the genus extirpation concentration distribution (XCD) model was detecting effects from SC and not a response to poor habitat, the HC05 was recalculated using the example criterion data set in which samples were removed with an RBP score <135 total, pH <6 and fecal coliform ≤400 colonies/100 mL. The threshold of RBP <135 was selected as an upper bound on acceptable habitat by Gerritsen, et al. (2010) that also provided an adequate sample size (relevant *n* = 922). This threshold of RBP <135 represents, on average, habitat that is not pristine, but which is adequate for maintenance of biological assemblages. Removal of poor habitat and high fecal coliform samples from the data set had almost no effect on the XCD model or HC05 (see Figure A 6).
B-9	17	Paired conductivity and biological data is a time-tested statistical test in environmental research; reliable and strong.	Thank you for your comment.
Appendix B	All figures	Well done; very helpful in conveying the relationships.	Thank you for your comment.

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)

Reviewer 2

Page number	Line(s)	Reviewer comment	EPA response
B-13	2−4	Of the 13 factors that were listed as being considered for having a causal relationship between conductivity and macroinvertebrates, some of them have had only minimal or no discussion in this document. Those are: nutrients, deposited sediments, selenium, settling ponds and dissolved oxygen. Selenium and metals were addressed in Appendix G.4.5.	These stressors had little or no effect on the model and therefore were not evaluated further. These stressors were also evaluated in a previous report using weight of evidence and also were not found to be confounders of the HC_{05} (see U.S. EPA 2011, Appendix B).
Appendix C		Excellent data design, rationales, descriptions; Table C-1 very helpful.	Thank you for your comment.
C-4	Table “C-2”	The numbering of the tables is incorrect. It should be Table C-1 because it is a continuing of the table on the previous page.	Edited accordingly.
C-5	Table “C-2”	Same as above	Edited accordingly.
C-5	1−6	Helpful for understanding the information in the table.	Thank you for your comment.
C-7	4	“(see Table C-3)” should be: (see Table C-4). The numbering of the tables for the rest of the section is now ‘off’.	Edited accordingly.
C-8	Table “C-3”	Should be “Table C-2”	Edited accordingly.
C-14	Top of page	Figure “C-3” should be Figure C-2.	Edited accordingly.
		Very interesting geological information.	Thank you for your comment.
C-14	Bottom of page	Figure “C-4” should be Figure C-3.	Edited accordingly.
C-15	Figure “C-5”	Should be Figure C-4.	Edited accordingly.

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)

Reviewer 2

Page number	Line(s)	Reviewer comment	EPA response
		Strong relationship in the cumulative distribution between the Criterion data set and the Ohio data set; gives significant strength to the analyses.	Thank you for your comment.
C-16	Figure “C-6”	Should be Figure C-5.	Edited accordingly.
		Very good illustration of distributions’ overlap and the ranges overlap; Strong.	Thank you for your comment.
C-18	Figure “C-7”	Should be Figure C-6.	Edited accordingly.
C-19	4−6 and 22−23	Good points on looking at reasons for absence of sensitive species for evidence that the regions are different. However, in stating that “current conditions may not allow re-colonization,” means that habitat is poor, and this conflicts with the previous determination in the document that habitat quality doesn’t affect the analyses. Here it appears to factor-in.	The passage in question states that if sensitive species are absent, it may be due to high conductivity but it may also be due to other factors, including poor habitat. EPA has always factored habitat quality in, and it is a cause of adverse impacts in some cases. However, the analysis of confounding indicates that habitat does not account for the effects of conductivity, where conductivity is the cause. That is, habitat is not a confounder of conductivity. No change required.
C-21	6−7	The likelihood that “watersheds with >90% native vegetation are more likely to have low conductivity” are also likely to have better quality stream habitat. Indirectly this also supports the role of habitat.	The point of the passage is, as stated, that undisturbed watersheds are likely to have low conductivity. This increases the chance that background conductivity was measured in at least some sites with little anthropogenic input. This paragraph is part of a larger assessment of whether the areas are sufficiently similar with respect to background conductivity. It does not address poor habitat as a cause. As explained above, EPA acknowledge that habitat quality is a cause of adverse impacts in some cases.

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)			
Reviewer 2			
Page number	**Line(s)**	**Reviewer comment**	**EPA response**
C-22	Table C-10	Would have liked to see responses for 23−27, but this is an example of how there might be descriptions or verifications not clear or missing. Thanks to the authors for presenting it as it is.	No response required.
C-24	Table "C-4"	Should be Table C-11.	Edited accordingly.
		Very good table; informative and well presented.	No response required.
C-26	2−3	"(see Figure C-7)" should be Figure C-6.	Edited accordingly.
		"(see Figure C-8)" should be Figure C-7.	Edited accordingly.
C-26	Figure "C-8"	Should be Figure C-7.	Edited accordingly.
C-27	Figure "C-9"	Should be Figure C-8.	Edited accordingly.
C-28	C.4.1	Appreciated the descriptions of the regions.	Thank you for your comment.
C-30	References	Brady, K: … —overly bold underlining.	All formatting was checked and made consistent with EPA standard format.
		Kahneman, D. —are there pages for the book?	All formatting was checked and made consistent with EPA standard format.
F-3	Figure F-1	The tight correlations in the scatter plots are very good.	Thank you for your comment.
F-6	Table F-1	Excellent table; SO_4 + HCO_3 clearly significant.	Thank you for your comment.
F-21	Section F.5	Summary and Tables F-5 and F-6 are helpful.	Thank you for your comment.
		I wonder how the CCC of 160 mg/L and the CMEC at 300 mg/L compares with other regions around the country?	No response required.
F-22	8	"(see Figure F-10)" should be Figure F-11.	Figures renumbered throughout.

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)

Reviewer 2

Page number	Line(s)	Reviewer comment	EPA response
F-23	Figure "F-10"	Should be Figure F-11.	Figures renumbered throughout.
F-24	Figure "F-11"	Should be Figure F-12.	Figures renumbered throughout.
		I would like to see more explanation of the bootstrapping method.	Definition added to the glossary: "Bootstrapping—A statistical technique of repeated random sampling from the data set that is often used in environmental studies to estimate confidence and prediction limits of a parameter." Additional text was added to the figure on bootstrapping, Figure 3.5. There are descriptions in two sections (see Sections 3.1.1.2.1 and 3.1.3.1), and this is a basic statistical method that can be found in any textbook.
F-25	Figure "F-12"	Should be Figure F-13.	Figures renumbered throughout.
F-54	F.10	3 references should have underlining of the authors if the format is to be kept the same throughout: Barbour, MT…, Newman, MC…, and R Development Core Team	All formatting was checked and made consistent with EPA standard format.
G-1	28−29	While I understand the need for minimum sample sizes of 500−800 macroinvertebrate sample and 800−1,000 fish samples, I wonder if state agencies will be able to have that many in their databases for each ecoregion? Has there been any checking with other states to see if most can meet this?	The reviewer comments on data set size for deriving an HC_{05} de novo. As mentioned in the text, the number of samples needed varies depending on the sampling methods, range of exposure, and number of genera identified. However, where data are insufficient to derive an HC_{05} de novo, criteria can be derived by using a background matching approach or calculation from the B-C regression model. The B-C regression method was added after this review and was reviewed separately in June 2015.

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)			
Reviewer 2			
Page number	**Line(s)**	**Reviewer comment**	**EPA response**
G-2; G-3	1−30; 1−13	Clear, accurate, and helpful discussion.	Thank you for your comment.
G-20	Figure G-8	Second to last line: “and 75−80 species evaluated.” The text on page G-18, line 29, said that it is 89, not 80. And in another, the number was 87.	All formatting was checked and made consistent with EPA standard format.
G-20		Section 3.4 seems to be missing. Pages go from G.3.2 on pg. G-17 to G.4 on page G-20. If 3.4 is there, I didn’t see it.	All formatting was checked and made consistent with EPA standard format.
G-25	Figure “G-7”	Should be Figure G-11.	Figures renumbered throughout.
G-27	G.4.6	Multivariant analysis for fish was interesting, especially the finding that catchment area and habitat significantly contributed to the model.	Thank you for your comment.
G-28	4th line down	“Catchement” is misspelled: should be catchment.	Edited accordingly.
G-34	G.6	The format of entries in the Appendix G’s references is not exactly the same as in the main reference section;	All formatting was checked and made consistent with EPA standard format.
		To maintain the format here, Gerritsen, J.,… needs to have: a) semicolons b) initials follow the last name, c) uniformity in use of periods.	All formatting was checked and made consistent with EPA standard format.
G-36	3, 7,12	“Availble” is misspelled. Should be available.	Edited accordingly.
G-42	Table G-7, title	“Ecoregions observed are the ecoregions where the species was collected in the combined data set” —needs to be bold	All formatting was checked and made consistent with EPA standard format.
7-1	References	Reference Section	

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)			
Reviewer 2			
Page number	**Line(s)**	**Reviewer comment**	**EPA response**
		The entire section does not maintain one particular format. The following are problems:	All formatting was checked and made consistent with EPA standard format.
		1) The initials on authors who are not the first and last author are not uniformly handled. In picking a uniform format, I suggest placing the initials in front of the surname. And periods following the initials.	All formatting was checked and made consistent with EPA standard format.
		2) Parentheses around the year of publication or just a period? Some entries have parentheses, others do not. Some have periods, others not.	All formatting was checked and made consistent with EPA standard format.
		3) A period after the journal name—or not?	All formatting was checked and made consistent with EPA standard format.
		4) Titles in small or all capital letters?	All formatting was checked and made consistent with EPA standard format.
		5) Listing of pages referenced in books—often missing.	All formatting was checked and made consistent with EPA standard format.
		6) The agency's name followed by its abbreviation in parentheses, or, the reverse?	All formatting was checked and made consistent with EPA standard format.

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)			
Reviewer 2			
Page number	**Line(s)**	**Reviewer comment**	**EPA response**
		The following is the first author's last name on every entry that I suggest be changed to meet a standard format and have one or more of the above problems. For me to re-write each faulty entry would be too time consuming. APHA Barbour Berra Bradley Boelter Brinck Clark Cormier (2010) Dahm Duncan Dunlop (delete the second "Water Quality") Echols (2009 a) Echols (2008b) Efron Entrekin Evans (2008a) (2008b) Evans (3001) Farag Fox Godwin Gregory Griffith Haluszczak Harper	All formatting was checked and made consistent with EPA standard format.

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)

Reviewer 2

Page number	Line(s)	Reviewer comment	EPA response
		Hem Higgins Hill Hille Hitt Hopkins Hynes Jackson (2007) Jackson (2005) Kaushal -2005 Kaushal (2013)—check to see if it is now published. Kelly Kennedy (2003), (2004), (2005) Kimmel Komnick Lasier Lefebvre, O. and R. Moletta Likens (1970) Merricks Meyer Mount Mullins Newman (2000), (2001) Nelson NYSDEC Omernik (1987), (1995) Paul, M.J. and J.L. Meyer Pond (2004), (2010), (2008), (2014)	

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)			
Reviewer 2			
Page number	**Line(s)**	**Reviewer comment**	**EPA response**
		Posthuma Remane Sams—spell out USGS: U.S. Geological Survey Scanlon—add "and" just prior to the last author. Smithson Soucek Stauffer Stubblefield Suter (2007), (2001) U.S. EPA (1985), (1987), (2000a, 2000b), (2003), (2006),(2009), (2010) (2011a, 2011b, 2011c) Van Dam—add "and" just prior to last author Veil Wallace—entomol. Needs capitalizing. Werner—remove comma and add "and" between authors; Delete "Wright et al. 1993" Wood (2008) Woods (2002), (1996) Ziegler (2007), (2010) Zielinski	

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)

Reviewer 3

Page number	Line(s)	Reviewer comment	EPA response
Entire report	Entire report	Capitalize States where appropriate—eliminates the confusion between noun (States) and verb (state or states) forms. Perhaps also capitalize Tribe. Check foreword.	The words state and tribe are capitalized when it refers to a particular political entity, not when used to define a group. Case was checked for names: river, state, and tribe.
xiii		FORWORD should be FOREWORD	Edited accordingly.
xiii	8, 14, 17	Capitalize States	See response above.
2-4	15	Split ¶ after effluents (before Ionic)	Edited accordingly.
2-10	Figure 2-1	Change black font to white on right side of all three blue blocks. One never uses black on blue.	Improved figure resolution and changed color.
2-17	18	Period after al (in et al. check entire document)	Edited accordingly.
4-6	5	Figure 4.1 rather than 4.2	No change. The sentence was referring to the box plot not the map.
4-10	Figure 4-5	Cannot see data points!!!!! Faint!!!!	Resolution of image was improved.
5-11	Figure 5-5	Cannot see data points!!!!! Faint!!!!	Resolution of image was improved.

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)			
Reviewer 4			
Page number	**Line(s)**	**Reviewer comment**	**EPA response**
—	—	**References Cited**	
		Bryce, S.A., D.P. Larsen, R.M. Hughes and P.R. Kaufmann. 1999. Assessing relative risks to aquatic ecosystems: A Mid-Appalachian case study. Journal of the American Water Resources Association. 35: 1752−1688.	Not used.
		Davies, S.P. and S.K. Jackson. 2006. The biological condition gradient: a descriptive model for interpreting change in aquatic ecosystems. Ecological Applications 16(4):1251−66.	Not used.
		Stoddard, J., P. Larsen, C.P. Hawkins, R. Johnson, and R. Norris. 2006. Setting expectations for the ecological condition of running waters: the concept of reference conditions. Ecological Applications 16:1267−1276.	Reference added.
—	—	Wang, L., T. Brenden, P. Seelbach, A. Cooper, D. Allan, R. Clark Jr. and M. Wiley. 2008. Landscape Based Identification of Human Disturbance Gradients and Reference Conditions for Michigan Streams. Environ Monit Assess (2008) 141:1−17.	Not used.
		Yoder, C.O. and M.T. Barbour. 2009. Critical technical elements of state bioassessment programs: a process to evaluate program rigor and comparability. Environmental Monitoring and Assessment 150(1−4): pp 31−42.	Not used.

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)			
Reviewer 4			
Page number	**Line(s)**	**Reviewer comment**	**EPA response**
Entire Report	—	As I mentioned in my general comments I think conductivity criteria should be considered in a tiered aquatic life use framework. There is natural variation in "background" conductivity due to variation in precipitation, base flow, etc., and within the range of "least impacted" to "minimally impacted" reference sites (as defined by Stoddard et al., 2006) there are variations due to human occupation of the landscape (e.g., agriculture, residential). My fear is that the criteria may not be stringent enough for the minimally impacted regions, and too stringent for land uses that State's would not considered to be impaired, but rather to be consistent with the swimmable/fishable goals of the act. This is not a suggestion that least impaired would encompass mine-related acute impacts or other impacts that are feasibly controllable.	Thank you for your comment. On a site-specific basis, the example criteria developed using the draft method could be adjusted or recalculated to protect important species, highly valued aquatic communities, or specially protected waters.
Glossary	—	"Background"—The definition of background I think is a bit "murky" given that later in text in includes both minimally impacted and least impacted reference sites. I would also add definitions of "least impacted" and "minimally impacted" reference sites and "tiered aquatic life uses."	The draft document describes a method to develop stressor specific aquatic life criteria and uses the term background in a specific way: "Background specific conductivity –The specific conductivity (SC) in streams in a region that occurs naturally and not as the result of human activity. Background may also be characterized as a population of minimally affected sites or low SC sites using a weight of evidence." The following definitions were added: Least disturbed condition—the best available physical, chemical, and biological habitat conditions given today's state of the landscape or the least disturbed by human

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)			
Reviewer 4			
Page number	**Line(s)**	**Reviewer comment**	**EPA response**
			activities (Stoddard et al., 2006). Contrast with "minimally affected condition. Minimally affected condition—The physical, chemical, and biological habitat found in the absence of significant human disturbance (Stoddard et al., 2006). Contrast with "least disturbed condition. "Tiered aquatic life use" was not included in the definition list because it refers to a designated use of a water body and is beyond the scope of the draft methods document. On a site-specific basis, the example criteria developed using this draft method could be adjusted or recalculated to protect important species, highly valued aquatic communities, or specially protected waters.

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)

Reviewer 4

Page number	Line(s)	Reviewer comment	EPA response
1-3	—	The document talks about how the protection of 95% of genera with this method is comparable with the protection of 95% of "species" in the lab toxicity approach. Some genera are more speciose than others. Is it truly similar? Not a major comment.	The sentence was edited to read, "In their review of the EPA Benchmark Report, the EPA SAB stated that this approach provides a degree of protection comparable to or more protective than a conventional water quality criteria based on conventional chronic toxicity testing (U.S. EPA, 2011c)" (U.S. EPA, 2011b in this response to comments document). The intent is to protect most species by using the 95th centile of genera. The selection of the 95th centile as described in EPA's *Guidelines for deriving numeric National Water Quality Criteria for the protection of aquatic organisms and their uses* (U.S. EPA 1985) involved an extrapolation to a protective level from a model of 8 genera from laboratory toxicity tests. The reviewer is correct in that a speciose genus will yield an XC_{95} that represents the more tolerant of the genus. This is shown using fish data which were identified to species in Appendix G.
2-3	25−28	This supports the contention that "background" levels can vary depending on how reference sites are defined. Within an ecoregion there could be subwatersheds with lower or higher conductivity than neighboring watersheds. Because of this, the natural distribution and abundance of those taxa that are most sensitive to conductivity can vary. These differences can be due to natural or some level of anthropogenic impacts (not acute or controllable sources). A discussion of tiered uses, I think is warranted for a stressor such as conductivity or perhaps other "natural" stressors such as habitat.	The sentence reads, "Natural geologic variability among neighboring watersheds may result in differences in ionic strength of associated streams." This statement cautions that unusual situations can arise and should be considered when applying a regional criterion. This caution is expanded in the subsequent sections and repeated more explicitly in other section of the document. A relationship between background and the HC_{05} was developed and added to the document in Appendix D and was peer reviewed in June 2015.

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)			
Reviewer 4			
Page number	**Line(s)**	**Reviewer comment**	**EPA response**
			A discussion of tiered uses was not included because it is beyond the scope of the draft methods document (i.e. refers to designated uses of water bodies other than aquatic life). On a site-specific basis, the example criteria developed using this draft method could be adjusted or recalculated to protect important species, highly valued aquatic communities, or specially protected waters.
2-3; 2-12	29−30; 3−6	Because precipitation can influence conductivity (the variation in seasonal concentrations, that is higher in late summer, tends to be months when precipitation is lowest) it seems that measures of base flow may be important in resolving within regional background variation in conductivity. My experience from Ohio is that the Exceptional Warmwater Habitat streams often have high base flows. In any case, I think discussion of potentially tiering conductivity benchmarks should be discussed.	The effects from base-flow will vary with ecoregion (e.g., for some, surface flow is greatest in the summer during snow melt rather than in April and May). Text was edited as follows: "Precipitation (e.g., rain or snow melt) can also affect ionic concentration. SC increases during episodes of below-normal surface flow and decreases during periods of above-normal surface flow. Seasonal patterns can vary greatly with regional climate, with low SC associated with spring rain or during summer from snow-melt."

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)

Reviewer 4

Page number	Line(s)	Reviewer comment	EPA response
2-16	22−29	Along a gradient of stress, such as conductivity, the probability of capturing an individual of a genus decreases with increasing stress. How does sampling methodology potentially influence derivation of benchmarks. If a taxa is not collected when the sample size is small, there is some likelihood it is present when not represented in the catch. Is there a way to use methods that count many more organisms (e.g., Ohio EPA method can have abundance estimates greater than a 1,000−2,000) to determine the bias when using methods that only count 200 or 300 individuals. Thus what is considered “mortality” may be partly under sampling. Although the trend in taxa response with conductivity may be similar between methods, the actual benchmarks could change if “sensitive taxa” show up at somewhat higher conductivities.	Although not done for the express purpose of evaluating the number of individuals identified in a sample, the validation using a full count of individuals from a Kentucky data set resulted in a similar HC_{05} as a count of 200 from many more sites in the West Virginia data set (U.S. EPA 2011b). When hundreds of sites are sampled, this appears to be comparable to identifying many more individuals in fewer samples. The reliance of the method on presence does not assume that the organism is absent when not observed. The endpoint is not mortality, it is the concentration at which a genus occurs and above which it is rarely observed. It is presence at a concentration that is important. No additional analyses other than those already described in the document were performed.
3-5	20	What is considered “similar background conditions?” If there are two groups of sites, one centered on a conductivity of 150 and the other on 250 and both are considered background, is that similar enough? When do you decide that you have two groups of sites that might be within the same ecoregion but that differ enough in conductivity and taxa that two tiers are needed?	The method describes rule-based criteria to ensure transparency and consistency. A call out to the appropriate section was added here which recommends using the confidence intervals: "background SC levels are similar throughout the region (see 3.7.1.)…"

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)

Reviewer 4

Page number	Line(s)	Reviewer comment	EPA response
3-8	22−29	This is where the concept of tiered uses could be tested. If minimally impacted reference sites can be distinguished from “least impacted” then the general not observed at reference sites could differ. The BCG process that has been used in a number of states results in output that classifies sites into different BCG tiers with reference sites usually varying between tiers 1−4 depending on definition. Sites are rarely classified as BCG1, but data sets where sites identified as BCG tiers 1−2 could be compared to those defined as BCG tiers 3−4 to see how this might affect the derivation of criteria. It would certainly influence which sites occur or do not occur at reference sites.	The BCG method for developing biocriteria is not used in this method. The method is not dependent on reference sites. EPA encourages all research and periodically updates guidance and practice as new information becomes available.
3-17	18	Because of the assumption of using samples from both seasons to generate the criteria, this implies that looking for exceedences of conductivity should be based on a geometric mean value from monthly samples. For determining water quality violations of the criteria are we expected to take monthly samples including both spring and summer periods or is there a methodology to adjust the “expected” criteria if only summer samples are taken as is common for many monitoring programs. Some more specific guidance on sampling for what would be considered violations or exceedences of criteria would be useful.	Text was edited in this section: " For example, in the example cases in Sections 4 and 5, annual insects (univoltine) that emerge in the spring, although present, are less likely to be detected in the summer, when coincidently, SC levels increase in some streams (e.g., due to decreased flow). In other locations, this pattern may be different. For example, high mountain systems may be affected by melting snow pack. Seasons may shift based on latitude." And, "At least one spring (when salt intolerant taxa can be collected) and one summer macroinvertebrate sample is recommended in order to increase the likelihood that sensitive taxa will be included in the data set." EPA is not providing specific guidance on sampling for monitoring and assessment purposes.

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)

Reviewer 4

Page number	Line(s)	Reviewer comment	EPA response
3-35	4−14	This is where I think some explicit definition of reference site conditions (minimally impacted distinguished from least impacted) is advisable (sensu Stoddard et al., 2006). In addition, an independent human disturbance gradient could be used to estimate the anthropogenic footprint.	EPA used the term "minimally affected" throughout the document to more clearly distinguish it from least impacted and cited Stoddard et al. (2006).
3-35	23	Would an independent human disturbance gradient measure be useful here?	Although tiered aquatic life uses are commonly used in biocriteria, it is not the practice used for stressor specific aquatic life criteria.
3-39	—	Given the "fuzziness" of deriving background conditions I think this can be difficult. Again the concept of tiered uses is important in this context and I think there is a need for a more explicit approach to distinguish between minimally impacted, least impacted and best attainable. For example a 90% forest threshold for a watershed may not be feasible in many areas and it is an important question whether conductivity levels are actually feasibly controllable in all cases.	The proposed weight-of-evidence relies on multiple pieces of evidence, not just percentage of land cover in forest. This method does not rely on defining minimally impacted, least impacted and best attainable.
4-17	5−19	It is easier to read if the X-axis of the plot is in actual conductivity units rather than as the log of conductivity.	Figures 4-7 and 5-7 were left as log plots due to the range of exposure. Figures 4-8 and 5-8 were redrawn without logs on the new x-axis.

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)

Reviewer 4

Page number	Line(s)	Reviewer comment	EPA response
4-24	—	Is it possible that this graph argues for the existence of two tiers of expectations? For example could there be unique, less common areas where conductivity is naturally very low and the most sensitive taxa more narrowly distributed and thus rare and less likely to meet a threshold of 25 sites for a genus to occur, and then more typical sites where more sensitive taxa are not found as frequently?	No, the apparent changes in slope are due to the loss of sensitive genera from the SD. This is explained in the figure legend in greater detail: "The HC_{05} increases greatly when a taxon in the lower 5th centile is removed because they do not meet the minimum number of samples and then more slowly alternates between increasing and decreases as genera either above or below the 5th centile are removed because they do not meet the minimum number of samples (see Figure 4-11). The pattern repeats until all genera have the same XC_{95} value (not shown). To maximize the number of genera included in the XCD, a minimum of 25 occurrences was utilized."
A-13	—	General Question: The section analyzed differences of HC_{05} values when low habitat scores were eliminated as potentially confounding variables. To explore the different benchmarks that might occur under tiered uses, perhaps the RPH habitat could be used to establish “reference” cutoffs under a crude tiered use scenario (reference sites with habitat scores >160 vs. >180). Ideally this could be done with a State like Ohio where tiered uses (EWH vs. WWH) are clearly defined.	A discussion of tiered uses was not included because it is beyond the scope of the draft methods document (i.e. refers to designated uses of water bodies other than aquatic life). On a site-specific basis, the example criteria developed using this draft method could be adjusted or recalculated to protect important species, highly valued aquatic communities, or specially protected waters.

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)

Reviewer 4

Page number	Line(s)	Reviewer comment	EPA response
C-1	19	Here is an example where I think the concept of background needs to be better quantified. At a minimum it should be related to Stoddard et al. (2006) definition of minimally vs. least impacted conditions. Ideally some form of a human disturbance score can be calculated. There is scatted mention of >90% forested as a reference benchmark, but that may be hard to find even in the WAP ecoregion of Ohio. For States to apply these benchmarks I think it argues for detail discussion of tiered uses and reference or background conditions.	EPA used the term “minimally affected” throughout the document and cited Stoddard et al. (2006). The reason for defining natural background based on soil, geology, and climate is so that geographic areas can be compared. It is not used to scale metrics as are done during the development of biocriteria. Reference sites defined by states are one way to evaluate whether background SC is measured. Results of a geophysical model was added.
C-2	15−16	Clarify whether background is minimally disturbed or least impacted.	EPA used the term “minimally affected” throughout the document so that it was easier to distinguish the two terms (minimally disturbed and least disturbed) put forth by Stoddard et al. (2006). Least disturbed is used when the data indicate that background is not natural and thus may lead to results that may not be protective of aquatic life.

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)

Reviewer 4

Page number	Line(s)	Reviewer comment	EPA response
C-2	17−18	The concept of subregions and local variations in base flow, which may not be easily predictable needs to be discussed more. Base flow seems to be a very important influence on conductivity given the variation observed in conductivity by month. Conductivity is usually higher overall in late summer early fall when flow is a minimum, but base flow makes up the greatest % of stream flow. For very small headwater streams, base flows can vary substantially within a region depending on the complexity of groundwater systems and points where streams become gaining flows. In larger streams I think these likely average out within a region, but in small streams may be important and may result in difference in rare and sensitive macro taxa.	Text was added to Figure C-5 regarding similarity on a seasonal basis and in the summary Table C-7 and C-11.
C-3	Table C-1	Regional properties—need to add consideration of base flow to this table. The sandstone aquifers in SE Ohio tend to have conductivities of 450−600 which indicates as the percent of flow that consists of groundwater increases the more likely that higher conductivity will occur. May also want to see Ohio primary headwater assessment data from Ohio (streams <1 sq mi) that has conductivity data (http://www.epa.ohio.gov/portals/35/wqs/headwaters/PHWH_Compendium.pdf)	EPA included the suggested Ohio headwater data set. Background of 450−600 was not observed even with the base flow model which gave higher estimates in some HUC10s. Base flow may be a reasonable as a line of evidence, but was not added. Contributions from base flow are captured in the data set. Text was added to C.3.9: "Because this is a regional assessment of background in Ohio Ecoregion 70, some higher and lower natural SC regimes may occur and require site specific evaluation of applicability of a regionally derived criterion, but overall a single criterion would be practical."

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)

Reviewer 4

Page number	Line(s)	Reviewer comment	EPA response
C-17	Table C-7	This table indicates that reference sites do not occur in Ohio, but Ohio does have reference sites. Also, would it be difficult to get forest cover for the Ohio sites? My guess is that few approach the 90% forest cover mentioned (page C-21, line 8) for further south. This has implications for attainability and setting feasible and controllable benchmarks.	EPA did not use reference sites because Ohio EPA cautioned that reference sites were developed for a different purpose, reference for modified habitats such as mine drainage. Exploratory analysis of sites with the designation of "least impacted" reference did not necessarily achieve high ICI scores or good water quality scores. So, reference sites were not compared to any other pieces of evidence. Entry in table was edited to read "**confirmed** reference sites not available from the Ohio data set." Forest cover is just one line of evidence.
C-17	Table C-10	Again reference sites are available for Ohio.	EPA did not use reference sites because Ohio EPA cautioned that reference sites were developed for a different purpose, a tiered use associated with mined sites. Exploratory analysis of sites with the designation of "least impacted" reference designation did not necessarily achieve high ICI scores or good water quality scores. So, reference sites were not compared to any other piece of evidence. Entry in table was edited to red "**Confirmed** reference sites not available from the Ohio data set."
G-2	23−24	The comparison of species vs. genus level XC_{95} values identified that genus values represented the more tolerant species in the genus. Why wouldn't that apply to the macroinvertebrate analyses and does this suggest that the lower conductivity sites in a region are not adequately protected? Would this be resolved with tiered uses and perhaps a lower threshold that would let more rare and sensitive species into a higher tier use?	The intent is protect the vast majority of species by using the 95^{th} centile of genera. The reviewer is correct in that a speciose genus yields an XC_{95} that represents the least sensitive among the species as was shown with fish species in Appendix G.

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)

Reviewer 4

Page number	Line(s)	Reviewer comment	EPA response
G-7	4−5	Ohio fish sites are not all in sites I would call "perennial.' Although very few of the sites dry completely, many small headwater sites can occur in what I would term interstitial streams that have periods where flows in riffles are subsurface, although permanent pools remain.	Text was edited to read, "All sites were watered at the time of sampling but may be intermittent at other times."
G-7	9	Again we have the 90% forested "benchmark" for reference with little discussion of what this means.	See previous response to Reviewer 2 in Question 13.
G-20	4−6	It would be useful to see data on sites that missed the *N* = 25 cutoff in this table to see if they characterize tiered uses or local high quality sites that may be important or some unique restricted distribution that might be important.	EPA constructed a data set of ranges for fish species observed at fewer than 25 sites. The maxima of these were higher than the HC_{05} derived using invertebrates, suggesting that fish (and the aquatic community as a whole) are protected by the invertebrate-based criteria.
G-22	9	In addition to tiered uses on some states, most states characterize coldwater uses as unique from warmwater uses. Ohio for example identifies species considered as characteristic of its coldwater uses. How would removal of coldwater taxa change the HC_{05} values?	There is no evidence to date that intolerance to SC is related to temperature preference.
G-23	26−31	If coldwater benchmarks were delineated separately from warmwater benchmarks, how would this affect this conclusion?	There is no evidence to date that tolerance to SC is related to temperature preference.

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)			
Reviewer 5			
Page number	**Line(s)**	**Reviewer comment**	**EPA response**
Entire report	—	**Terminology: Extirpation Concentrations:**	
		In my view, this term is being used inappropriately. Webster defines the term "extirpate" to mean "to destroy completely." Other dictionaries have similar meanings. That is not an appropriate term here because of the way in which the document quantifies the term: 5% of observations occur at concentrations higher than the so-called extirpation concentration (expressed as XC_{95}). The capture probability figures within the document (Figs. 3-1 and 3-4) indicate that individuals of certain taxa are being observed at concentrations >2x XC_{95}.	The definition is from the published literature and is consistent with the definition noted by the reviewer (U.S. EPA, 2003). The reviewer seems to think that the XC_{95} depends on measuring extirpation at each site. That is not the case; the XC_{95} is the concentration at which the genus is observed in 5% of sites. The effect endpoint is not the regional extirpation of a genus. Rather, extirpation refers to either the absence of a taxon or the functional absence of a taxon and the XC_{95} refers to the concentration that leads to extirpation. The XC_{95} measure and method was reviewed by EPA's Science Advisory Board (SAB) who found it to be an appropriate measure. The term "extirpation" was searched in the draft document and the explanation was accurate and consistent throughout the document.

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)

Reviewer 5

Page number	Line(s)	Reviewer comment	EPA response
		Any water quality measure is highly variable in time. The method described here recommends “measurements of the agent(s) should be paired in space and time with biological sampling”. If measurements were timed to acquire samples at the point in time within any given stream where SC is at its highest point during a given genera’s life cycle, the term “extirpation concentration” would be more justifiable—but such targeting is not described by the method presented.	EPA analyzed this issue. The reviewer's proposed method is not feasible because the highest SC in a stream often occurs in the summer when many species would appear to be extirpated when they are not, owing to the fact that their life cycle makes them too small to be collected in the summer. The CCC is based on an annual average with a definable range. The CMEC derivation is a statistical analysis of empirical data. The method was verified by a different method using the occurrence of a genus at a site after the CMEC occurred, i.e., based on sampling of large late instars in the spring compared to the highest observed SC during the previous year. This verification analysis has been added to the document after an independent, external, peer review in June 2015, but it is not a recommended approach to derive the CMEC because most states are unlikely to have these annual data and it is equivalent to the statistical method for calculating the CMEC.
Entire report	—	**Terminology: Field-based Method.** In my view, a term such as “field data analysis method” would be more appropriate because the method described includes no actual field activities data collection; it relies solely on secondary data that have been obtained for other purposes.	The use of the term field data distinguishes the method from laboratory based methods and was not changed.

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)

Reviewer 5

Page number	Line(s)	Reviewer comment	EPA response
Entire report	—	**General Comments: Box Plots:** There are a number of box plots that show distributions of SC and related ions among months throughout the document (e.g., Figs 4-2, 4-3, 4-4, etc.). Given that most that most of the annual data are distributed quite unevenly among months, I would suggest displaying numbers of observations used to generate each monthly box plot.	Sample size was added to Tables 4.2 and 5.2.
		There are a number of box plots throughout the document. Suggest stating at some point quantities represented by the box plots (I presume median, 25th, and 75th percentiles for the boxes; 90th and 10th percentiles for the tails? Are all observations lying outside of the 10th and 90th percentiles represented as data points, or only some?) Suggest stating the nature of box plot representations explicitly at some point in the manuscript. The caption for the first box plot would be a logical place to do this.	Box plots are standard depictions of distributions. The description of the plot was added to the glossary on page xix. "Boxplot—A depiction of the 25th, 50th, and 75th quantiles of a distribution as a rectangle with a central line. The 2 standard deviation range is depicted as "whiskers" extending from the box. Data beyond 2 standard deviations are indicated by individual circles or dots beyond the whiskers."
xxi	—	Glossary: Suggest that the term "Reference Site" be added to the glossary, given the importance of Reference Sites to the proposed method (e.g., Section 3.1.1.2.5. Exclusion of disturbance or pollution-dependent genera: "Genera that are not observed at reference sites or are estuarine or marine organisms are excluded from the data set.")	This document does not describe methods for reference site selection and relies upon the judgement of state biologists except when "reference sites" are shown to be contaminated or physically altered, e.g., a low pH. A definition was added to the glossary: "Reference site—Sampling locations that have been identified as minimally affected or least disturbed based on land use, habitat, and water quality characteristics other than SC."

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)

Reviewer 5

Page number	Line(s)	Reviewer comment	EPA response
2-8	Table 2-1	Table should be annotated to communicate that these are examples only (i.e., this is not a comprehensive or exhaustive list.).	The title was changed to read, "Examples of ions associated with different sources."
2-15	1−2	“charged particles” are not equivalent to “dissolved ions”.	The phrase has been edited to read: “(3) it measures only dissolved ions.”
2-15	28	“Physiological Mechanisms”: My scientific background and training does not enable me to evaluate this section.	No response required.
2-19	17−20	“Freshwater insects are among the most sensitive organisms relative to other taxa, including zooplankton, fish and amphibians (see Appendix G of this report; Kennedy et al., 2004; Echols et al., 2009b; Lazorchak et al., 2011; Consbrock et al., 2011; Williams et al., 2011).” The statement is poorly supported by the references provided. Lazorchak et al. (2011), Consbrock et al. (2011), and Williams et al. (2011) are citations of conference presentations; hence, supporting documentation is not available to the public or to reviewers and, hence, are inappropriate, in my view, as a means of providing scientific support for a statement with this level of significance in a (potential) regulatory document. Kennedy et al. (2004) compared sensitivity of *Isonychia bicolor*, a mayfly, to only one other taxon, *Ceriodaphnia dubia*, which does not typically inhabit the flowing waters where *Isonychia* are generally found. Echols	The reviewer asks for better evidence that insects are among the most sensitive to dissolved ions. The strongest evidence is the field-based analyses reported in U.S. EPA 2011a, Cormier and Suter 2013, Cormier et al. 2013, papers by other researchers in the last several years (cited in the comment), and in the public review draft document. The text has been edited to read: "These field-based methods can be used to develop ecoregional criteria that are fully protective of aquatic life. *Many* freshwater insects are among the most *salt-intolerant* organisms relative to other taxa, including daphnids, fish, and amphibians (*compare Appendices A.4 and B.4 with Appendix G of this report*). ~~*Other laboratory and field studies support these findings* (Kennedy et al., 2004; Echols et al., 2009b; Lazorchak et al., 2011; Consbrock et al., 2011; Williams et al., 2011)~~. Less is known about the salt-intolerance of mussels to ionic stress, but recent studies suggest that Unionidae are acutely sensitive to salts (Kunz et al., 2013), particularly during early (glochidia) life stages (Bringolf et al., 2007; Gillis, 2011)." Compare Figure G-9 for fish species and Figures 4-7 and 5-7 for invertebrate genera. Had fish species

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)

Reviewer 5

Page number	Line(s)	Reviewer comment	EPA response
		et al. (2009b) also worked with *Isonychia bicolor*; they compared laboratory-derived toxicity values for *Isonychia* with comparable values obtained from the literature for other species (Table 1), some of which were aquatic insects; and the aquatic insect species were not the most sensitive for most measures. Appendix G derives a species sensitivity distribution (SSD) for fish (Figure G-12); visual comparison of this distribution to the benthic macroinvertebrate SSDs (Figures 4-7 and 5-7) indicates fish as more sensitive throughout most of the SC range. I am not saying the statement in question is in error as I have not looked into that topic with depth. I am saying that statement is of great significance relative to the regulatory program proposed by this document; and the statement is poorly supported as currently presented.	been proportionally less tolerant than invertebrates genera, the curve would be expected to be to the right of the invertebrate XCD.
3-5	Section 3.1.1.2	Selection and Adequacy of Data Sets: An additional selection criterion should be that observed SC levels should be well distributed over the population of streams used for the analysis, when those streams are stratified using measured characteristics.	The suggested data set selection criterion is not necessary. In those cases where factors might be important (e.g., temperature), statistical methods were provided in the document to adjust for differences in sampling density (see Appendices A and B). In actual practice, adjustment for habitat, stream temperature, drainage area and other possible stratifications had negligible effects on the HC05. The only essential stratification is natural background SC and this is addressed in section 3.6 and Appendix D. Additional analyses were added after this review to Sections 3.6 and Appendices A and B. Appendix D was also added subsequent to this review. No change was made.

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)

Reviewer 5

Page number	Line(s)	Reviewer comment	EPA response
3-6	13−14	"As a general rule of thumb, the minimum sample size to estimate an XC_{95} using this field-based method is 25 observations of the genus in the region." Suggest that this rule of thumb be investigated further to determine if minimum number of observations should be expressed alternatively as a fraction of dataset size. If an SSD dataset were to include 500 samples, 25 observations would constitute 5% of the dataset; but if the SSD dataset were to include 2500 samples, the 25 observations would constitute only 1% of the dataset. Does 25 observations of a taxon remain as an adequate number as dataset size increases?	Based on the SAB review, EPA has balanced the need for a large enough sample to estimate and XC_{95} and HC_{05} with the need to include as many genera as possible. Although more research is always welcome, the SAB concluded that a sample size of 25 was adequate for developing an XC_{95} and HC_{05}. No change was made.
4-11	7−9	"Samples collected from the WVDEP-identified reference sites indicate that conductivity levels are generally low and similar throughout the year, although slightly higher in summer/fall months of August, September, and October …" My interpretation of "slightly higher" is not consistent with its use in this sentence. Figures 4-2 and 4-4 indicates that mean SC during the 3 months listed is >2x the mean SC during most other months.	The reviewer is correct. Figure call outs were incorrect. Deleted 4-2 and 4-4.
5-12	1−3	Same comment as for page 4-11, lines 7−9.	The reviewer is correct. Figure call outs were incorrect. Deleted 5-2 and 5-4.
7-1 and forward	—	Reference formatting is inconsistent.	All formatting was checked and made consistent with EPA standard format.

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)			
Reviewer 5			
Page number	**Line(s)**	**Reviewer comment**	**EPA response**
Entire report	—	Seasonal definitions are not clear. For example, page 4-11, lines 7−9 refer to August, September, and October as "summer/fall months" while page 3-19, line 5 refers to the July−October period as "summer."	Seasonal callouts were made consistent and when appropriate refer to the specific months or a defined period.
—	—	**References Cited**	References not already cited in the text were obtained and considered for inclusion. All formatting was checked and made consistent with EPA standard format.
		Agouridis, C., P. Angel, T. Taylor, C. Barton, R. Warner, X. Yu, C. Wood. 2012. Water quality characteristics of discharge from reforested loose dumped mine spoil in eastern Kentucky. Journal of Environmental Quality 41:454−468.	
		Bernhardt E.S., B.D. Lutz, R.S. King, J.P. Fay, C.E. Carter, A.M. Helton, D. Campagna, J. Amos. 2012. How many mountains can we mine? Assessing the regional degradation of central Appalachian rivers by surface coal mining. Environmental Science and Technology 46:8115−8122.	
		Boehme E.A. 2013. Temporal dynamics of benthic macroinvertebrate communities and their response to elevated specific conductance in headwater streams of the Appalachian coalfields. M.S. Thesis, Virginia Tech.	
		Boehme E.A., S.H. Schoenholtz, C.E. Zipper, D.J. Soucek, A.J. Timpano. 2013. Benthic	

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)

Reviewer 5

Page number	Line(s)	Reviewer comment	EPA response
		macroinvertebrate community temporal dynamics and their response to elevated specific conductance in Appalachian coalfield headwater streams. P. 7−22 in: 2013 Powell River Project Research and Education Program Reports. Virginia Tech. http://www.prp.cses.vt.edu/Reports_13/Reports_13.html	
		Bryant, G., S. McPhilliamy, H. Childers. 2002. A Survey of the Water Quality of Streams in the Primary Region of the Mountaintop/Valley Fill Coal Mining, October 1999 to January 2001, U.S. EPA Region III, Wheeling, WV.	
		Daniels W., Z. Orndorff, M. Eick, C.E. Zipper. Predicting TDS release from Appalachian mine spoils. p. 275−285. In: J.R.Craynon (ed.), Environmental Considerations in Energy Production. Society for Mining, Metallurgy, and Exploration. Englewood, CO.	
		De Jong, G.D. and S.P. Canton. 2013. Presence of long-lived taxa and hydrologic permanence. Journal of Freshwater Ecology 28(2): 277−282.	The authors evaluated the types of invertebrates inhabiting ephemeral and perennial streams and suggest that there are different strategies that species use to exploit ephemeral streams including opportunistic-life histories and burrowing. They did not find that the species were diagnostic of ephemeral vs perennial systems.
		del Rosario RB, Resh VH (2000) Invertebrates in intermittent and perennial streams: is the hyporheic	In an intermittent and a perennial stream in northern California, the findings indicated that the biota of the intermittent and perennial streams were similar, which suggests that a criterion

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)

Reviewer 5

Page number	Line(s)	Reviewer comment	EPA response
		zone a refuge from drying? J N Am Benthol Soc 19:680−696	would be relevant to both systems. Hyporheic zones were not found to act as refugia in these two streams.
		Delucchi, C. M. 1988. Comparison of community structure among streams with different temporal flow regimes. Canadian Journal of Zoology 66: 579−586.	This study found similarity among 4 ephemeral streams with different flow regimes. “It appears that differences in community structure between permanent and temporary riffles are minimized by generalized adaptations of stream benthos, such as high rates of migration, drought-resistant eggs, and the tendency to take refuge in the hyporheic zone.”
		Evans D.M., C.E. Zipper, P.F. Donovan, W.L. Daniels. 2014. Long-term trends of specific conductance in waters discharged by coal-mine valley fills in central Appalachia, USA. Journal of the American Water Resources Association 50: DOI: 10.1111/jawr.12198	
		Feminella, J.W. 1996. Comparison of benthic macroinvertebrate assemblages in small streams along a gradient of flow permanence. Journal of the North American Benthological Society 15: 651−669.	
		Fritz K.M., S. Fulton, B.R. Johnson, C.D. Barton, J.D. Jack, D.A. Word, & R.A. Burke. 2010. Structural and functional characteristics of natural and constructed channels draining a reclaimed mountaintop removal and valley fill coal mine. Journal of the North American Benthological Society. 29:673−689.	

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)

Reviewer 5

Page number	Line(s)	Reviewer comment	EPA response
		Griffith M.B. 2014. Natural variation and current reference for specific conductivity and major ions in wadeable streams of the coterminous U.S. Freshwater Sciences 33: 1−17.	
		Grubaugh J.W, J.B. Wallace, E.S. Houston. 1996. Longitudinal changes of macroinvertebrate communities along an Appalachian stream continuum. Can J Fish Aquat Sci 53:896−909	
		Grubbs S.A. 2010. Influence of flow permanence on headwater macroinvertebrate communities in a Cumberland Plateau watershed, USA. Aquatic Ecology. 45: 185−195.	
		Helsel D.R., R.M. Hirsch. 2002. Statistical Methods in Water Resources. U.S. Geological Survey. Techniques of Water-Resources Investigations of the United States Geological Survey. Book 4, Hydrologic Analysis and Interpretation Chapter A3.	
		Hem J.D. 1989. Study and Interpretation of the Chemical Characteristics of Natural Water. U.S. Geological Survey, Water Supply Paper 2254.	
		Lindberg, T. T., E. S. Bernhardt, R. Bier, A. M. Helton, R. B. Merola, A.Vengosh, and R. T. Di Giulio. 2011. Cumulative impacts of mountaintop mining on an Appalachian watershed. Proceedings of the National Academy of Sciences	

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)			
Reviewer 5			
Page number	**Line(s)**	**Reviewer comment**	**EPA response**
		108:20929−20934, with online supporting information.	
		Mount, D.R., J.M. Gulley, J.R. Hockett, T.D. Garrison, & J.M. Evans. 1997. Statistical models to predict the toxicity of major ions to Ceriodaphnia dubia, Daphnia magna, and fathead minnows (Pimephales promelas). Environmental Toxicology and Chemistry 16:2009−2019.	
		Odenhimer J.L. 2013. Determining a Total Dissolved Solids Release Index from Overburden in Appalachian Coal Fields. M.S. Thesis, West Virginia University.	
		Odenheimer J., J. Skousen, L.M. McDonald. 2013. Predicting total dissolved solids release from overburden in Appalachian coal fields. In: J.R. Craynon (ed.). Environmental Considerations in Energy Production. Society for Mining, Metallurgy, and Exploration. Englewood, CO.	
		Orndorff Z.W., W.L. Daniels WL, M. Beck, M.J. Eick. 2010. Leaching potentials of coal spoil and refuse: Acid-base interactions and electrical conductivity pp 736−766. In: Barnhisel RI (ed.), Proc Am Soc Min Reclam Ann Meetings, Pittsburgh, PA. 5−11 Jun. 2010. Amer Soc Mining & Rec	

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)			
Reviewer 5			
Page number	**Line(s)**	**Reviewer comment**	**EPA response**
		Pond, G.J., M.E. Passmore, F.A. Borsuk, L. Reynolds, & C.J. Rose. 2008. Downstream effects of mountaintop coal mining: comparing biological conditions using family- and genus level macroinvertebrate bioassessment tools. Journal of the North American Benthological Society 27:717−737.	
		Pond, G.J., M.E. Passmore, N.D. Pointon, J.K. Felbinger, C.A. Walker, K.J.G. Krock, J.B. Fulton, W.L. Nash. 2014. Long-term impacts on macroinvertebrates downstream of reclaimed mountaintop mining valley fills in central Appalachia. Environmental Management. DOI 10.1007/s00267-014-0319-6	
		Price K., A. Suski, J. McGarvie, B. Beasley, J.S. Richardson. 2003. Communities of aquatic insects of old-growth and clearcut coastal headwater streams of varying flow persistence. Can J For Res 33:1416−1432	
		Sena K.L. 2014. Influence of Spoil Type on Afforestation Success and Hydrochemical Function on a Surface Coal Mine in Eastern Kentucky. M.S. thesis, University of Kentucky.	
		Sena K., C. Barton, P. Angel, C. Agouridis, R. Warner. 2014. Influence of spoil type on chemistry and hydrology of interflow on a surface coal mine in	

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)

Reviewer 5

Page number	Line(s)	Reviewer comment	EPA response
		the eastern U.S. coalfield. Water, Air, & Soil Pollution 225: 1−14.	
		Stout B., J.B. Wallace. 2003. A Survey of Eight Major Aquatic Insect Orders Associated with Small Headwater Streams Subject to Valley Fills from Mountaintop Mining. Appendix in Mountaintop Mining/Valley Fills in Appalachia. Final Programmatic Environmental Impact Statement. U.S. Environmental Protection Agency, Philadelphia, PA	
		Timpano A.J., S.H. Schoenholtz, D.J. Soucek, C.E. Zipper. 2014. Salinity as a limiting factor for biological condition in mining influenced central Appalachian headwater streams. Journal of the American Water Resources Association. DOI: 10.1111/jawr.12247	
		Timpano A.J., D. Soucek, S. Schoenholtz, C. Zipper. May 2013. Continuous conductivity monitoring for predicting macroinvertebrate community structure in coal mining-influenced streams. Society for Freshwater Science 2013 Annual Meeting, 19−23 May, Jacksonville, Florida. Abstract ID 7546.	
		Timpano A.J., S. Schoenholtz, C. Zipper, D. Soucek. 2011. Levels of dissolved solids associated with aquatic life effects in headwater streams of Virginia's Central Appalachian coalfield	

TABLE 6. SPECIFIC OBSERVATIONS ON THE DOCUMENT (continued)			
Reviewer 5			
Page number	**Line(s)**	**Reviewer comment**	**EPA response**
		region. Final report prepared for Virginia Department of Environmental Quality; Virginia Department of Mines, Minerals, and Energy; and Powell River Project. April 2011.	
		U.S. EPA (Environmental Protection Agency). (2000a) Nutrient criteria technical guidance manual: Rivers and streams. Office of Water, Washington, DC. EPA/822/B-00/002.	
		Vannote, R.L., G.W. Minshall, K.W. Cummins, J.R. Sedell, C.E. Cushing. The river continuum concept. Canadian Journal of Fisheries and Aquatic Sciences 37: 130−137.	
		Williams D.D. 1996. Environmental constraints in temporary fresh waters and their consequences for the insect fauna. J N Am Benthol Soc 15:634−6	A review of temporary freshwater and found that it is a misperception of fact that the insect faunas of temporary waters are constrained by their ephemeral nature. Seasonal success of species was commonplace. The most successful taxa had flexible life cycles, temperature-linked development, diapausing or protected eggs, and high powers of dispersal.

EPA Response References

Bringolf, RB; Cope, WG; Eads, CB; Lazaro, PR; Barnhart, MC; Shea, D. (2007) Acute and chronic toxicity of technical-grade pesticides to glochidia and juveniles of freshwater mussels (Unionidae). Environ Toxicol Chem 26(10):2086−2093.

Clements, WH; and Kotalik, C. 2016. Effects of major ions on natural benthic communities: an experimental assessment of the US Environmental Protection Agency aquatic life benchmark for conductivity. Freshwater Science Vol. 35(1) (March 2016): DOI: 10.1086/685085

Coffey, DB; Cormier, SM; Harwood, J. (2014) Using field-based species sensitivity distributions to infer multiple causes. HERA 20(2):402−432.

Cormier, SM; Suter, GW. (2013) A method for assessing causation of field exposure-response relationships. Environ Toxicol Chem 32(2):272−276.

Cormier, SM; Suter, GW; Zheng, L; Pond, GJ. (2013) Assessing causation of the extirpation of stream macroinvertebrates by a mixture of ions. Environ Toxicol Chem 32(2):277−287.

Cormier, S. M., 2015. Field-based Methods for Developing Water Quality Benchmarks Invited Expert Meeting on Revising U.S. EPA's Guidelines for Deriving Aquatic Life Criteria 14-16 September 2015, Arlington, VA). Available at: https://www.epa.gov/sites/production/files/2016-01/documents/17_cormier_wqc_20150828_for_release_secure.pdf

Gillis, PL. (2011) Assessing the toxicity of sodium chloride to the glochidia of freshwater mussels: Implications for salinization of surface waters. Environ Pollut 159(6):1702−1708.

Griffith, MB. (2014) Natural variation and current reference for specific SC and major ions in wadeable streams of the coterminous U.S. Freshw Sci 33(1):1−17.

Nietch, C; Lazorchak, J; Johnson, B; Allen, J. (2014) Evaluation of total dissolved solids and specific conductance water quality targets with paired single-species and mesocosm community exposures. Presented at Society of Environmental Toxicology and Chemistry North America 35th Annual Meeting, Vancouver, BC, CANADA, November 09−13, 2014. Available online at http://cfpub.epa.gov/si/si_public_record_Report.cfm?dirEntryId=293057&CFID=17865572&CFTOKEN=51190896

Olson, JR; Hawkins, CP. (2012) Predicting natural base-flow stream water chemistry in the western United States. Water Resour Res 48:W02504.

Johnson, BR; Weaver, PC; Nietch, CT;. Lazorchak, JM; Struewing,KA;. Funk, DH. 2015. Elevated major ion concentrations inhibit larval mayfly growth and development. Environmental Toxicology and Chemistry. Vol 34 (1) 167-172.

Kunz, JL; Conley, JM; Buchwalter, DB; Norber-King, TJ; Kemble, NE; Wang, N; Ingersoll, CG. (2013) Use of reconstituted waters to evaluate effects of elevated major ions associated with mountaintop coal mining on freshwater invertebrates. Environ Toxicol Chem 32(12):2826−2835.

Stoddard, JL; Larsen, DP; Hawkins, CP; Johnson, RK; Norris, RH. (2006) Setting expectations for the ecological condition of streams: The concept of reference condition. Ecol Appl 16(4):1267−1276.

Suter, GW, II. (2007) Ecological risk assessment, 2nd edition. Boca Raton, FL: CRC Press.

Suter, GW, II; Cormier, SM. (2013) A method for assessing the potential for confounding applied to ionic strength in Central Appalachian streams. Environ Toxicol Chem 32(2):288−295.

Timpano, A; Schoenholtz, S; Zipper, C; Soucek, D. (2011) Levels of dissolved solids associated with aquatic life effects in headwater streams of Virginia's Central Appalachian coalfield region. Report prepared for the Virginia Department of Environmental Quality; Virginia Department of Mines, Minerals, and Energy; and the Powell River Project. Prepared by Virginia Tech. Blacksburg, VA. Available online at: http://www.prp.cses.vt.edu/Research_Results/Timpano_TDSReport_2011.pdf.

U.S. EPA (Environmental Protection Agency) (1985) Guidelines for deriving numeric National Water Quality Criteria for the protection of aquatic organisms and their uses. Washington, DC: Office of Research and Development, Environmental Research Laboratories. PB85-227049. Available online at http://water.epa.gov/scitech/swguidance/standards/upload/2009_01_13_criteria_85guidelines.pdf.

U.S. EPA (Environmental Protection Agency). (2011a) A field based aquatic life benchmark for conductivity in Central Appalachian streams. EPA/600/R 10/023F. Washington, DC: Office of Research and Development, National Center for Environmental Assessment. Available online at http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=233809.

U.S. EPA (Environmental Protection Agency). (2011b) Review of field-based aquatic life benchmark for conductivity in Central Appalachian streams. Washington, DC: Science Advisory Board, Office of the Administrator. Available online at: http://yosemite.epa.gov/sab/sabproduct.nsf/0/EEDF20B88AD4C6388525785E007331F3/$File/EPA-SAB-11-006-unsigned.pdf.

U.S. EPA (Environmental Protection Agency). (2013) Biological assessment program review: Assessing level of technical rigor to support water quality management. EPA/820/R-13/001. Washington, DC: Office of Science and Technology. Available online at: http://water.epa.gov/scitech/swguidance/standards/criteria/aqlife/biocriteria/upload/2013biological_assessment.pdf.

Voss, K. A., King, R. S. and Bernhardt, E. S. (2015), From a line in the sand to a landscape of decisions: a hierarchical diversity decision framework for estimating and communicating biodiversity loss along anthropogenic gradients. Methods Ecol Evol, 6: 795–805. doi:10.1111/2041-210X.12379

Wang, N; Ingersoll, CG; Kunz, JL; Brumbaugh, WG; Kane, CM; Evans, RB; Alexander, S; Walker, C; Bakaletz, S. (2013) Toxicity of sediments potentially contaminated by coal mining and natural gas extraction to unionid mussels and commonly tested benthic invertebrates. Environ Toxicol Chem 32(1):207−221.

Zipper, CE; Beaty, B; Johnson, GC; Jones, JW; Krstolic, L; Ostby, BJL; wolfe, WJ; Donovan, P. 2014. Freshwater mussel population status and habitat quality in the Clinch River, Virginia and Tennessee, USA: A featured collection. J. Amer Water Res. Assoc. Vol 50 (4): 807-819

Zipper, CE; Donovana, PF; Jones, JW; Li, J; Price, JE; Stewart, RE. 2016. Spatial and temporal relationships among watershed mining, water quality, and freshwater mussel status in an eastern USA river. Science of the Total Environment. 541: 603-615.