EPA Response to External Peer Review Comments

on

EPA's Draft Document, *Application to New Areas*

(Analysis in Appendix D of EPA's Public Review 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

This document has been reviewed in accordance with U.S. Environmental Protection Agency policy and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

ANA	Application to New Areas report
B-C	background-to-criterion
CCC	criterion continuous concentration
CI	confidence interval
CMEC	criterion maximum exposure concentration
EPA	Environmental Protection Agency
GAM	generalized additive model
HC	hazardous concentration
HU	hydrologic units
SC	specific conductivity
SD	sensitivity distribution
PL	prediction limit
PI	prediction interval
RFM	Recommended Field-based Method
SAB	Science Advisory Board
USGS	United States Geological Survey
XC	extirpation concentration

1. INTRODUCTION

The U.S. Environmental Protection Agency (EPA) submitted its *Draft Application to New Areas* (a supplemental analysis supporting its Draft *Field-Based Methods for Developing Aquatic Life Criteria for Specific Conductivity*) for contractor-led, independent, external peer review. The external peer reviewers provided their independent responses to EPA's charge questions to the contractor who developed a final report dated July 29, 2015.[1] 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 method developed by the EPA to estimate a hazardous concentration (HC_{05}) based on a model of background specific conductivity.

This report presents the 5 peer review charge questions and summaries of individual reviewer comments (verbatim) in Section 2. Each summary is followed by a response from EPA to the individual reviewer comments by charge question. Please note that some sections refer to the extracted draft sent for review and page numbers do not track with the current document.

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:

- **David Buchwalter, Ph.D.** Associate Professor, Department of Biological Sciences, North Carolina State University
- **Yong Cao, Ph.D.** Illinois Natural History Survey, University of Illinois at Urbana-Champaign
- **Bruce K. Hope, Ph.D.** Independent Consultant
- **Marian R.L. Maas, Ph.D.** Independent Consultant
- **Raymond P. Morgan II, Ph.D.** Appalachian Laboratory, University of Maryland Center for Environmental Science

The EPA contractor provided reviewers with instructions, the review document, 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] Eastern Research Group (ERG). 2015. Peer Review of EPA's Draft Document "Application to New Areas," Peer Review Summary Report. Dated July 29, 2015. Task Order 51 under Contract EP-C012-029.

2. SUMMARIES ORGANIZED BY CHARGE QUESTION

2.1. THE RATIONALE FOR DEVELOPING THE BACKGROUND-TO-CRITERION (B-C) MODEL IS BASED ON THE AVAILABILITY OF SPECIFIC CONDUCTIVITY NICHES FOR SALT-INTOLERANT GENERA. IS THE BASIC ECOLOGICAL THEORY AND RELATIONSHIP TO THE MODEL PRESENTED CLEARLY? IS THE RATIONALE SUFFICIENT, AND IF NOT, WHAT OTHER RATIONALES SHOULD BE ARTICULATED? (SECTION 1, PAGES 1-1 TO 1-3.)

2.1.1. Reviewer #1

Comment: In general, the material on pages 1-1 to 1-3 provides a construct for understanding how salinity tolerances, realized niches, and the environmental availability of SC niches determine species distributions. These ideas are reasonably well presented (but see specific comments below), but they don't provide a strong linkage to or rationale for the specifics of the B-C model. Only the last two sentences of P. 1-2 provide a direct link to the B-C model (but the second of those 2 sentences is really not clear). There is no ecological theory presented as to why there should be a consistent relationship between then low end of an ecoregion's EC values and XC_{95} values to support the B-C model.

> **Response:** The reviewer indicates that there should be more explanation of the link between an available specific conductivity (SC) habitat and SC tolerance limits (niche) of species and genera and the background-to-criterion (B-C) model. Section 2.4 in the public review draft document describes the general underpinning of the evolutionary basis for different tolerance limits to SC. The section was edited as suggested by this and other reviewers to make more clear the evolutionary basis for the B-C model. A sentence was added that directs the readers to Appendix D Figure D-3, which shows that in ecoregions with increasingly higher background SC, the lower tolerance limit begins at near background and the upper tolerance limit of the 5% most salt-intolerant genera increases the HC_{05} for each ecoregion accordingly.

Comment: Figure 1-1 should either be changed or expanded. The arrow on the X-axis labeled as "lower tolerance limit" is clearly in error as it corresponds with specific conductivities associated with high Probabilities of Observing. In this example, the realized niche is apparently limited by the availability of low EC waters, and not by the biology/physiology of the organism. If more sites in that region had very low ECs, you might expect the realized niche to expand based on the Probability of Observing data shown here. Perhaps it would be useful to include a second example of a genus where the Probability of Observing shows a bell or "n"- shaped curve—with the genus limited by both lower and upper EC values. In such a plot, it would be appropriate to indicate the lower tolerance limit. It is not appropriate to suggest a lower tolerance limit on the figure shown.

> **Response:** The figure legend (Figure 2-2 in public review draft document) was edited as follows:

“A species’ (or genus’) realized niche is defined by its lower and upper limits of occurrence. In this case, the lower tolerance limit is less than or equal to the lowest specific conductivity (SC), which is the lower limit of occurrence. The XC_{95} represents the upper tolerance limit. Approximately 5% of observations of a taxon are assumed to occur in sink habitats where a population cannot persist without immigration from source habitats. A species or genus optimum is the environmental condition most easily tolerated both physiologically and competitively and can be estimated by the conditions where the taxon is most often observed. The optimum SC may be estimated at the maximum probability of observing the taxon from a generalized additive model, shown here to be the minimum SC. The example involves the genus *Ephemerella* which is comprised of several species of mayflies.”

Comment: Line 4: is it appropriate to describe field populations of species as “test organisms”.

Response: This phrase was removed from the public review draft document.

Comment: Throughout the text: The term “range” is used numerous times in the text and in different ways that may cause some confusion to the readers. At times, range is used to indicate the level of SC (e.g., line 31), whereas other times it is used in relation to the span or width of SCs that describe a species realized niche regardless of the actual levels of SC (e.g., line 29–30), or the range of SC’s found in the ecoregion. I would try to avoid this confusion.

Response: The term range was edited to clarify these different concepts (Section 2.4 in public review draft).

Comment: P. 1-2, Line 1: “When the available SC niche is greater than the tolerance range of a species, it can no longer persist and is extirpated”. I think I understand what is trying to be conveyed here but I believe this sentence could be made clearer in several ways. In the first part of the sentence “available SC niche is greater” seems to be referring to the level of SC and not with breadth of the niche……this should be more explicit. The latter part of the sentence that refers to species persistence and extirpation implies that the species previously occurred there and has been removed because of SC change. This is very different from the idea that the particular species never occurred in a place because the SC is outside of its tolerance range. Perhaps something like—“When the available SC niche has been altered such that it no longer overlaps with the tolerated SC range of a given species, the species in question can no longer persist and is extirpated.”

Response: Section 2.4 in the public review draft document was edited as suggested by this and other reviewers to make more clear the evolutionary basis for the B-C model. The text was edited as follows:

“The range of SC conditions varies in natural aquatic systems. Species do not occur where the SC is lower or higher than their SC tolerance. The lowest SC in a freshwater system, therefore, is the lowest possible limit of the potential SC niche (see lower tolerance limit in Figure 2-2). When mineral salts are added to an aquatic system, SC increases, part of the potentially habitable SC niche space is lost, and the size of the realized niche for species adapted to low SC decreases. When the SC is above the physiological tolerance of a species due to natural or anthropogenic causes, it does not persist and the species is extirpated.”

Comment: Along these lines, there seem to be two separate themes that appear to be conflated. On one hand, there is attention paid to the distribution of SC niches within an ecoregion serving as a species "filter". Taxa with generally tolerating only lower SC conditions are excluded from systems with higher SCs. The other theme is SC change (increase) within a given system and the extirpation of species from it due to SCs exceeding the species' tolerance. These are two very different scenarios but they are thematically interwoven throughout.

> **Response:** Section 2.4 in the public review draft document was edited as suggested by this and other reviewers to make more clear the evolutionary basis for the B-C model.

Comment: P. 1-2, line 6: suggest changing "is supported by" to "primarily the result of"

> **Response:** Edited as suggested (Section 2.4 of public review draft).

Comment: P. 1-2, line 12-13: This last sentence is awkward and doesn't make sense to me.

> **Response:** Section 2.4 in the public review draft document was edited as suggested by this and other reviewers to make more clear the evolutionary basis for the B-C model. The text was edited as follows:
>
> "The relationship between ambient SC levels and SC tolerances of species that are present has at least two practical implications. First, it is inappropriate to set criteria below natural background. Second, the lower limit for any XCD in any given ecoregion cannot be lower than the natural background of the ecoregion. In practical terms, this shifts the origins of field-XCDs and their 5th centiles toward higher SC (graphically to the right) as the background SC increases. Hence, when XCDs from regions of low to high natural background are simultaneously plotted on the same graph the curves progress to the right. (For an example, see the XCDs in Appendix D, Figure D-3). Therefore, the background SC of an ecoregion is strongly associated with a predictable extirpation of 5% of species or genera. This relationship between background SC and the proportion of extirpation can be used to predict the SC that will extirpate 5% of species or genera in an ecoregion solely based on ecoregional background (see Section 3.7.2 and Appendix D, Figure D-4)."

Comment: Another general comment: Throughout pp. 1-3, the discussion centers around species. However, the models are based on genera. I understand that this is a practical necessity because we can't put species names on everyone. However, there is a theoretical assumption that species within the same genus should be similar. Is it important to have at least have some cursory understanding of how variable species within a given genus are? A genus such as *Ephemeralla* that has many species is likely to have some "spread" in physiological tolerances. We know from some of our work that they range widely in their trafficking of calcium for example. What are the implications for genus level lumping in species rich genera? Does this skew genus level XCs in an upward, less protective direction?

> **Response:** The variations among species within a genus are discussed in the public review draft document (Footnote 3):

“Conventionally, species have been aggregated to the genus level. However, effect levels may be different for species within a genus due to niche partitioning afforded by naturally occurring causal agents such as ionic strength (Remane, 1971; Suter, 2007). Hence, an apparently salt-tolerant genus may contain both salt-intolerant species and tolerant species. Analyses with fish species indicate that the range of XC_{95} values within a genus can be quite broad and the empirical genus-level XC_{95} tends to represent the mean XC_{95} of the species in the data set (see Appendix G).”

2.1.2. Reviewer #2

Comment: The basic ecological theory and relationship to the model is NOT presented clearly. The use of the terms, niche, range, limits, is confusing.

Response: Section 2.4 in the public review draft document was edited as suggested by this and other reviewers to make more clear the basic ecological theory and relationship to the B-C model. The text was edited as follows:

“The range of SC conditions varies in natural aquatic systems. Species do not occur where the SC is lower or higher than their SC tolerance. The lowest SC in a freshwater system, therefore, is the lowest possible limit of the potential SC niche (see lower tolerance limit in Figure 2 2). When mineral salts are added to an aquatic system, SC increases, part of the potentially habitable SC niche space is lost, and the size of the realized niche for species adapted to low SC decreases. When the SC is above the physiological tolerance of a species due to natural or anthropogenic causes, it does not persist and the species is extirpated.

The upper tolerance limit of a species is estimated by its XC_{95} (see Figure 2-2). Extirpation is the depletion of a population of a species to the point that it is no longer a viable resource or is unlikely to fulfill its function in the ecosystem (U.S. EPA, 2003). The occurrences of benthic invertebrate species at locations with a SC greater than their XC_{95} value are believed to represent sink habitats (Pond et al., 2014). Sink habitats are those locations where occurrence of species is primarily the result of immigration from locations with low SC termed source habitats from which immigrants originate. They are “sinks” in the sense that immigrants have low success in establishing sustainable populations in those locations.

These phenomena have practical application. The proportion of species or genera extirpated as a result of increased SC in an ecoregion can be determined and is the basis for the XCD method.

Several other predictions can be made from niche theory. Species with niches that limit them to low SC water are not expected to occur where low SC water does not occur. The source of high SC could be natural or due to anthropogenic inputs (Cormier et al., 2012, Coffey et al., 2014). For example, in an ecoregion lacking streams <400 µS/cm, any species with an upper tolerance limit <400 µS/cm SC would not be expected to occur because there is no habitat for them. As a corollary, where there is a low SC habitat in an ecoregion, species tolerant to low SC will occur.

> The relationship between ambient SC levels and SC tolerances of species that are present has at least two practical implications. First, it is inappropriate to set criteria below natural background. Second, the lower limit for any XCD in any given ecoregion cannot be lower than the natural background of the ecoregion. In practical terms, this shifts the origins of field-XCDs and their 5th centiles toward higher SC (graphically to the right) as the background SC increases. Hence, when XCDs from regions of low to high natural background are simultaneously plotted on the same graph the curves progress to the right. (For an example, see the XCDs in Appendix D, Figure D-3). Therefore, the background SC of an ecoregion is strongly associated with a predictable extirpation of 5% of species or genera. This relationship between background SC and the proportion of extirpation can be used to predict the SC that will extirpate 5% of species or genera in an ecoregion solely based on ecoregional background (see Section 3.7.2 and Appendix D, Figure D-4)."

Comment: The authors should use a simple Gaussian curve to illustrate the relationships among the three terms. The term niche is also used either as the range of SD suitable for a species or as the multidimensional environmental space suitable (overall ecological niche). The authors need to specify which one they refer to each time (see more comments in the word document). Otherwise, the rational appears sufficient.

> **Response:** Regarding Figure 2-2, the Gaussian curve is not the curve that is seen at the lower SC portion of observation distributions of salt intolerant genera (invertebrates) or species (fish). In general the left side of the Gaussian curve is absent because the organisms lower limit is defined by the background conductivity for the region. The original figure was retained.
>
> The document is concerned with describing and explaining the distribution of taxa with respect to SC, and so the niche discussed is the SC niche.

Comment: In Figure 1-1, the reviewer highlighted the sentence, "Approximately 5% of observations of a taxon occurs in sink habitats where a population cannot persist without immigration," and commented, "Be clear that this is an assumption, even a reasonable one, but not an observation or conclusion. Modify."

> **Response:** Figure legend 2-2 was edited and now reads, "…Approximately 5% of observations of a taxon are assumed to occur in sink habitats where a population cannot persist without immigration from source habitats…"

2.1.3. Reviewer #3

Comment: I reviewed the entire document—including those portions that were flagged as "already been reviewed"—simply to maintain continuity and context. Overall, I found many parts of the document (including the already reviewed portions) to not be written very clearly and thus being hard to follow. There is some redundancy in the text and it almost seems as if different sections were written by different people at different times and not edited together for consistency, brevity, and clarity. The section headings are almost all too long and wordy. The same can be said for many of the figure captions, which sometimes do not align with the text, contain extraneous material, or are not helpfully explanatory. The entire text needs to be edited for clarity and consistency.

General: The ecological theory is pretty straightforward—SC is one defining dimension of the niche hypervolume—but this section does not explain the linkage between this theory and the B-C model. The model is only mentioned once (page 1-1, line 11) and there is no apparent connection between that one mention and the ecological discussion that occurs in Section 1.1. Perhaps this linkage is implicit in some of the references cited but it needs to be articulated explicitly and clearly here as well.

> **Response:** Section 2.4 in the public review draft document was edited as suggested by this and other reviewers. See responses above to Reviewers 1 and 2 (2.1.1 and 2.1.2).

Comment: page 1-1, line 7. "primary" Not clear what this is. SD method?

> **Response:** Primary was deleted and "field-XCD" was inserted.

Comment: page 1-1, line 34. The distinction between the realized SC niche (actually inhabited niche) and the available SC niche (niche available to be inhabited?) is not clear here.

> **Response:** Section 2.4 in the public review draft document was edited as suggested by this and other reviewers. See responses above to Reviewers 1 and 2 (2.1.1 and 2.1.2).

Comment: page 1-2, line 5. Why the sudden switch from species to genera?

> **Response:** The text in Section 2.4 was changed to species for basic theoretical discussions. Genus was originally interspersed in the paragraph because some evidence was based on genus level analysis.

Comment: page 1-2, line 10. Presumably XC_{95} as shown in Figure 1-1? Please clarify.

> **Response:** After revisions, Figure 1-1 became Figure 2-2 in the public review draft document. The suggested change was made and a call out was added to Figure 2-2 after the abbreviation XC_{95}.

Comment: page 1-2, lines 12-14. It is not at all clear what you are trying to say here. Please clarify. Can you illustrate what you mean here with a figure? And, again, why the switch from species to genera?

> **Response:** Section 2.4 in the public review draft document was edited as suggested by this and other reviewers. See responses above to Reviewers 1 and 2 (2.1.1 and 2.1.2). Also, the text in Section 2.4 was changed to species for basic theoretical discussions. Genus was originally interspersed in the paragraph because some evidence was based on genus level analysis.

Comment: page 1-3, Figure 1-1. There could be a better alignment between this caption and the corresponding labels on the figure. This would make the caption more explanatory. Also, the range of "optimum" needs to be more clearly delineated, maybe with before and after arrows? Presumably it extends from the LTL to XC_{95} but you can't tell that from the figure.

Response: See response to Reviewer 1 in Section 2.1.1.

2.1.4. Reviewer #4

Comment: Yes, the basic ecological theory and its relationship to the model is adequate, especially since this is an addition to the larger document in which more attention was given to the ecology. It would be duplicating previous text to expand to any great extent the ecological aspects in this additional section. Of course, since my interest area is ecology, I would have liked to see more information in this area but I understand that this is a likely natural bias towards my own interests.

Response: Thank you for your comment.

Comment: It is presented clearly and scientifically correct. I believe that Section 1 addresses the rationale for developing the Background-to-Criterion (B-C) model sufficiently and correctly. The choices of 5th and 10th centiles are well explained and appropriately determined. Rationale presented on pg. 2-9, lines 33−36, and pg. 2-10, lines 1−5, is well done, clear and correct.

Response: Thank you for your comment.

Comment: I have some instances in which I either have a question or a comment. These are as follows:

a) Lines 15−22, pg. 1-1, describes tolerance and ecological aspects well. However, I suggest in line 23 that "relative to marine" (or something similar) be inserted, reading: "However, in most of the United States, freshwater habitats have low concentrations of dissolved ions, <u>relative to marine</u>, so that is the condition to which the biota is adapted."

Response: In Section 2.4 of the public review draft document the suggested text was added:

"Aquatic species inhabit nearly pure water, estuarine and marine conditions, hypersaline pools, and everything in between (Remane, 1971; Potapova and Charles, 2003; Potapova, 2005; Berra, 2007). In most of the United States, freshwater habitats have very low concentrations of dissolved ions relative to marine systems, so that is the condition to which most freshwater biota are adapted"

Comment: Further, I find that this same sentence, lines 22−24, almost a contradiction of the preceding sentence.

Response: See edits above to preceding sentence. The point here is that organisms have evolved to very low ionic concentrations in freshwater, not just less than marine water.

Comment: I suggest inserting "or introduced" in line 27, pg. 1-1, reading: "A species may not exploit its full tolerance range because competitor <u>or introduced</u> species … "

Response: This sentence in Section 2.4 of the public review draft document reads: "A species may not exploit its full tolerance range because competitor species are better suited for a particular ionic concentration or for other ecological reasons such as predation, parasitism, and habitat requirements." Introduced species was not added to this sentence because they competitors and stressors in themselves.

Comment: What are "other habitat requirements" in line 29, pg. 1-1? Might be good to list them (if there aren't too many) rather than leaving the reader wondering.

Response: This text "other habitat requirements" was removed from the public review draft document.

Comment: Anthropogenic influence on specific conductivity (SC) is not mentioned and although I understand that the document is addressing background conditions, I believe that it should at least be addressed here in lines 25–30, pg. 1-1. Species live and survive in conditions which are within their tolerance levels, which in all reality consist of both background and anthropogenic pollutant contributions. This at least needs to be acknowledged.

Response: For the purposes of this section, the text focuses on the evolution of a range of tolerances with an emphasis on defining the low specific conductivity (SC) niche. It is not about adaptation and tolerance to pollution and is independent of sources of ions. Anthropogenic sources are discussed in Section 2.2 of the public review draft. Also, there is this statement in Section 2.4: "When mineral salts are added to an aquatic system, SC increases, part of the potentially habitable SC niche space is lost, and the size of the realized niche for species adapted to low SC decreases. When the SC is above the physiological tolerance of a species due to natural or anthropogenic causes, it does not persist and is extirpated."

Comment: Suggest an additional sentence to follow line 30, pg. 1-1, or in pg. 1-2, for a bit more ecological explanation, such as: "All species try to occupy those niches which allow optimum physiological functioning. If conditions are at their outer limits, species are less able to grow, reach adulthood, reproduce and thrive. Abundant and robust populations are less likely to exist."

Response: In Section 2.4 of the public review draft a sentence was added: "At the extremes of their physiological tolerance, species are less able to develop, grow, and reproduce."

Comment: I can't help but wonder where is that dividing line between tolerable SC levels, and the levels which increase stress and thus, diminished populations? And, how much does pH, temperature, ammonia, and other pollutants accentuate or facilitate high SC levels' impact on aquatic organisms? Development of the Background-to-Criterion model is clearly a very necessary first step in establishing water quality standards for SC, and in moving towards better understanding of all aspects of this important parameter for streams and rivers. Its implementation throughout the country will be a major tool in better evaluation of the country's water quality. Future paths will need to look at the impacts of pollutants on SC and the measuring of those impacts.

Response: Thank you for your comment. Cumulative stress decreases an organism's tolerance to other stressors. Diminished populations can be determined from the point of decline of the generalized additive model (GAM) plots, but are not used in the method. Many genera decline in occurrence at specific conductivity (SC) levels below the HC_{05}.

2.1.5. Reviewer #5

Comment: For the "Application to New Areas' report (ANA), I agree that the basic ecological theory and its relationship to the model is presented clearly and the overall rationale is basically correct, with the accurate statement that salt-intolerant species will occupy a realized niche (Figure 1-1) dependent on the ionic composition of the water body (as well as the other chemical-biological factors of the water body in question). However, the true definition of salt-intolerant genera and their species complexes may be difficult**—**I believe that I made this point before in my review of the Recommended Field-based Method (RFM), or perhaps should have. Although both the RFM and ANA are well-written, well-documented and provide support for the ecological basis for defining the HC_{05}, perhaps there is a need to bring in some basic ionic physiology, especially a gill physiology - molecular biology approach, from well-established fish physiologists (Steve McCormick and Dan Evans to name just two) to support the methods, as well as some examination of the evolution of the salt-intolerant species in North America (e.g., *Marshall, W.S. 2002. Na^+, Cl^-, Ca^{2+} and Zn^{+2} transport by fish gills: retrospective review and prospective synthesis. Journal of Experimental Zoology 293:264−283*).

Response: The physiology is briefly reviewed in the description of possible mechanisms of action in Section 2.5.1 of the public review draft document. The suggested reference (Marshall 2002) was added to Section 2.5.1.

Comment: There may also be a need to address 'sink habitats' in the text (P 1-2, L 5). Normally, in most ecological work that I am familiar with, the term is 'source-sink' dynamics. In the report, the authors only address the function of sink, without much explanation of this important ecological concept of a source. Readers not having a basic understanding of this important ecological mechanism may be slightly misled.

Response: Section 2.4 in the public review draft document has been edited to be more explicit regarding sources and sinks. For example: "The occurrences of benthic invertebrate species at locations with a SC greater than their XC_{95} value are believed to represent sink habitats (Pond et al., 2014). Sink habitats are those locations where occurrence of species is primarily the result of immigration from locations with low SC termed source habitats from which immigrants originate. They are "sinks" in the sense that immigrants have low success in establishing sustainable populations in those locations."

2.2. PLEASE COMMENT ON THE CLARITY AND SCIENTIFIC DEFENSIBILITY OF THE METHOD THAT USES THE B-C REGRESSION MODEL TO ESTIMATE A CRITERION. IT MAY BE HELPFUL TO SEE THE B-C MODEL IN FIGURE 4.4. (SECTIONS 2.2−2.2.7, PAGES 2-9 TO 2-14)

2.2.1. Reviewer #1

Comment: The model is strong and should be shown closer to the front of the document. The presentation is pretty clear and defensible. I don't have the statistical background to comment on equation 2-2.

> **Response:** Thank you for your comment. The development of the model is explained in detail in Appendix D. A copy of the model was added to the methods sections to aid the reader in visualizing the text.

Comment: Is the term "SD" ever defined in the document?

> **Response:** Yes, it is defined earlier in the document and in the glossary. The term has been changed to avoid confusion with "standard deviation." It is now termed an extirpation concentration distribution (XCD).

Comment: P. 2-9, line 32: rare/salt intolerant taxa do not contribute to the SD? Don't these species drive the HC_{05} estimate? This is confusing.

> **Response:** The most salt-intolerant taxa for a region do occur near the 5^{th} centile of the sensitivity distribution (SD). The use of the word "rarer" was deleted.
>
> The text was clarified in Section 3.7.2: "In an ecoregion with low background SC, very salt-intolerant taxa are represented. In an ecoregion with a moderate background SC, taxa with an XC_{95} greater than the moderate background are likely to survive and contribute to the XCD, whereas salt-intolerant taxa with XC_{95} values less than the moderate background are not likely to contribute to the XCD. As XCDs are developed for ecoregions with increasing background SC levels, each XCD begins at a higher background SC, and thus the most salt-intolerant genera in an ecoregion occur at progressively higher SC levels."
>
> Note: to avoid confusion between sensitivity distribution (SD) and standard deviation, the distribution was more distinctly named "extirpation concentration distribution" (XCD) throughout the document.

Comment: P. 2-9, line 34: add the language to the sentence ending on this line ….“than do the most salt-intolerant taxa in regions with lower SC levels”.

> **Response:** Section was edited as noted in previous response.

Comment: P. 2-10, lines 10−13 seem to really weaken the argument.

Response: Section 3.7.2 was edited as follows:

"Individually, many of the 24 HC_{05} values used to develop the B-C method have not been subject to analyses needed for development of a CCC and should be considered as provisional. For example, the HC_{05} estimates used in this model were not supported by full confounding analyses, as is described in the *EPA Benchmark Report* (U.S. EPA, 2011a). However, the true HC_{05} value is expected lie between the upper and lower 50% prediction limit (PL). Values in Appendix D, Table D-3 are provisional with a good degree of confidence owing to the larger sample sizes (>60 samples) used to estimate background SC. Table D-4 lists ecoregions with background estimates based on modest survey data sets (N = 20-60 samples) and would benefit from additional sampling to confirm the calculated background SC and the calculated HC_{05}. Table D-5 lists ecoregions where the data set may represent fewer than 25% minimally affected streams and therefore are protective of aquatic life in least disturbed streams. Table D-6 lists ecoregions that may not be served by the B-C method because the ionic mixture is likely to be different (e.g., chloride dominated), the estimated natural background SC exceeds the range of the model, and/or there were fewer than 20 samples available. In all cases, EPA recommends using the largest data set possible to estimate background SC, understanding and accounting for areas with different (higher or lower) background SC, and performing independent calculations to derive HC_{05} values."

Comment: P. 2-10 line 12: should spell out Prediction Interval (PI).

Response: The first instance was identified for all abbreviations and they are spelled out.

Comment: Be aware that the r reported in the instructions to reviewers is 0.87 and in the text it is reported as 0.93.

Response: One value is an r^2 and the other is an r. The r^2 has been removed and only r is shown because this is not a test of significance but of strength of the association and this will minimize confusion.

2.2.2. Reviewer #2

Comment: The strong B-C regression is very interesting and its use for estimate SC HC_{05} is scientific defensible.

Response: Thank you for your comment.

Comment: However, the procedure illustrated in Fig. 2-2 is quite confusing. The authors need to explain why 10% prediction limit (or confidence limit?) is chosen to decide whether field-based or predicted HC_{05} or predicted lower 10% limit of HC_{05} should be adopted.

Response: The choice of which prediction limit to use is a matter of judgment. Section 3.7.2.1 now reads: "(3) If the field-XCD estimate is below the lower 50% PL, then the lower 50% PL is recommended as the HC_{05} (see eq 3-5). Because the XCD is calculated from a smaller data set, it may be overly protective and is more uncertain than the modeled results which indicate that 75% of HC_{05} values from areas with a similar background SC are estimated to be greater than that value. Also, the lower 50% PL is also recommended when there are fewer than 200 paired biological samples because there is no XCD for comparison. The SC data and the B-C model is used to estimate the HC_{05}."

Comment: I am also concerned about the lack of model validation and the potential biases introduced when the predicted SC for HC_{05} (log-scaled) is transformed back to real SD level. I provided two references that help the authors to address these issues.

Related to above comment, the reviewer commented on the document:

"Some form of validation is needed for the models. R2 most likely drops to some extent when applied to new datasets. Because only 19 data points, it may be not sensible to set 1/3 of sample aside for validation. I suggest conducting a Jackknife validation (see Olden, J.D., Jackson, D.A., and Peres-Neto, P.R. 2002. Predictive models of fish species distributions: A note on proper validation and chance predictions. T. Am. Fish. Soc. 131: 329−336)."

Response: The validation of the B-C model is described in Section D.4.1. of the public review draft document. The leave-one-out-cross-validation (LOOCV) method was used rather the one suggested by the reviewer and so the suggested reference was not included.

Comment: The reviewer commented on the document,

"Back-transformation can introduce bias in linear modeling. Methods are available to correct it (Duan N. 1983. Smearing estimate: a nonparametric retransformation method. Journal of the American Statistical Association, 78, 605−610)."

Response: As long as the log transformation was not rounded, the effect of back-transformation is small. No change was made because it increased the potential for error by users of the smearing method recommended by the reviewer.

Comment: In addition, the 19 ecoregions used for B-C regression are apparently not selected at random. These are heavily clustered in space (e.g., 4 in Minnesota, 3 in Idaho, and 3 in West Virginia). May this issue compromise the general applicability of the model? It will be really nice to refine the model based on more and spatially-balanced ecoregions, and validate it properly.

The reviewer commented on the document:

"Those 19 ecoregions are not randomly distributed, but highly clustered. Not sure how much this issue can affect the performance of the B-C model, but worth attention. Also, does EPA plan to refine the model when HC_{05} becomes available for more ecoregions?

Response: EPA has added more data sets into the B-C model in the public review draft which now includes 24 data sets from 23 ecoregions. All science is subject to revision and this may occur in the future.

Comment: The reviewer highlighted, "usually this has a minor effect because only observations of occurrences rather than presence/absence or abundance is used in the calculations" and commented in the text that, “This is a bold statement. Any data/citations support it? If the sampling method also differ in sampling effort or fixed-count or sampling gear, it would also certainly affect whether a given taxon will be observed. Revise.”

> **Response:** This text was deleted from the public review draft. See Section 3.1.1 in the public review draft for information on establishing the data set.

Comment: The reviewer highlighted the sentence, “If the 95% confidence interval (CI) of the background SC of the ionic mixture of the new area overlaps with the 95% CI of the background in the original area, the original criterion is considered applicable’ and commented “I understand 95% confidence intervals commonly used in statistics, but what is its biological relevancy? This criterion might be not strict enough. It would be nice to test what confidence intervals are sufficient to have the same or similar levels of HC_{05}. Also, how often the 95% CI overlapped among the 19 Level-3 ecoregions?”

> **Response:** The background-matching approach requires that a new area be compared with background from an area with an HC_{05} derived from a large data set. EPA’s recommended approach to compare the 95% confidence interval (CI) is a standard statistical approach and is reasonable given the variability within the estimation of background SC. The 95% CI background SC are narrow and therefore using a smaller CI would result in few matches between areas with similar HC_{05} values. The 95% CI was retained for this method.

Comment: The reviewer highlighted the phrase, “However, care should be taken,” and commented “How??”

> **Response:** The text in Section 3.7.1 was edited as follows: “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.”

Comment: The reviewer highlighted the phrase, “But, in general, estimation based on a random sample of the region tends to yield a more accurate estimate of current background when there are sufficient data to characterize the full distribution,” and commented, “Support needs for this general statement. Also, “the full distribution” of what?”

> **Response:** The previous references support this statement (NYSDEC, 2000; TDEC, 2000; U.S. EPA, 2000a). At the end of the sentence, “SC in the region” was substituted for “chemical parameter.”

Comment: The reviewer highlighted the sentences, “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 of 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,” and followed with a comment “A brief justification is needed for using Bootstrapping rather than a conventional method. Is this because the sample is too small or frequency distribution is not normal?”

> **Response:** The sample is one data set with no variance. Bootstrapping is a commonly used and accepted statistical method in such a situation. No change was made.

Comment: The reviewer highlighted the sentence, "In such cases, site-specific SC criteria may be considered to protect low-SC streams." And commented, "How? May you refer to some existing studies?"

> **Response:** The public review draft 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 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.

Comment: The reviewer highlighted the sentence, "In an ecoregion with a moderate background SC, species with an XC_{95} greater than the background are likely to survive and contribute to the SD, whereas salt intolerant species are rarer and are less likely to contribute to the SD, and commented, "Difficult to follow. Why are salt-intolerant species necessarily rarer? I thought that those sensitive species/genera are particularly important for establishing HC_{05}. In contrast, XC_{95} for salt-tolerant species are hard to determine as their occurrences do not decline or only decline slowly with increasing SC. Am I missing something here?"

> **Response:** The text in Section 3.7.2 was edited as follows: "In an ecoregion with low background SC, very salt-intolerant taxa are represented. In an ecoregion with a moderate background SC, taxa with an XC_{95} greater than the moderate background are likely to survive and contribute to the XCD, whereas salt-intolerant taxa with XC_{95} values less than the moderate background are not likely to contribute to the XCD. As XCDs are developed for ecoregions with increasing background SC levels, each XCD begins at a higher background SC, and thus the most salt-intolerant genera in an ecoregion occur at progressively higher SC levels."

Comment: The reviewer commented. "Confusing. Do you rank the tolerance of all genera across the nation? If so, the statement holds, but it is irrelevant. This is because you define HC_{05} for each ecoregion based on the taxa present in the ecoregion. So, one always has regionally salt-intolerant species and their absolute salt-tolerance is not important. Clarify.

> **Response:** The text in Section 3.7.2 was edited to be more clear: "The B-C regression method shown in Figure 3-10 was derived using independent data sets from 24 ecoregions (see Appendix D). First, SC XC_{95} values were estimated. From these, 24 genus-level XCDs were constructed and HC_{05} values derived. Those HC_{05} values were regressed against the estimate of background SC for each ecoregion. In an ecoregion with low background SC, very salt-intolerant taxa are represented. In an ecoregion with a moderate background SC, taxa with an XC_{95} greater than the moderate background are likely to survive and contribute to the XCD, whereas salt-intolerant taxa with XC_{95} values less than the moderate background

are not likely to contribute to the XCD. As XCDs are developed for ecoregions with increasing background SC levels, each XCD begins at a higher background SC, and thus the most salt-intolerant genera in an ecoregion occur at progressively higher SC levels."

2.2.3. Reviewer #3

Comment: General: There is a lack of specificity in the terms used that makes it hard to follow what you're doing here. There is also a lack of definition in all of the equations and their associated parameters that is confusing. Why use generic terms like "X" and "Y" when the parameters have meaning specific to this model? The details of the equations are in Section 3 (the case study) and not here in Section 2 (where they are introduced). It is not clear exactly what manipulations you are performing until the reader reaches Section 3.

> **Response:** The terms are defined below the equations (e.g. see Equation 3-4). The stressor (i.e. specific conductivity, SC) has been added in some places to be more clear.

Comment: page 2-4, line 8. You need to define what you mean by "minimally" here, which can otherwise be highly subjective. What is the cutoff or threshold for this?

> **Response:** Minimally affected condition is defined in the glossary of the public review draft as "the physical, chemical, and biological habitat found in the absence of significant human disturbance (Stoddard et al., 2006). Contrast with 'least disturbed condition.'"

Comment: page 2-4, line 18. What is the relationship of a "defined" area to a "new" or "original" area? This this a new term or just imprecise use of an existing one?

> **Response:** Text in Section 3.7.1 was edited as follows: "In this discussion, the phrase, original area, refers to the geographic area from which the data are obtained to develop SC criteria using the XCD method. The phrase, new area, refers to a geographic area within the same Level III ecoregion that was not represented in the criterion derivation data set. When applying field based SC criteria developed with data from the original area to a new area, the background SC levels and the ion composition should be similar in both areas. For instance, the example criteria are derived with data for streams where the ionic mixture is dominated on a mass basis by ($[SO_4^{2-}] + [HCO_3^-]) > [Cl^-]$."

Comment: page 2-4, line 34. "ecoregional" Is this different than simply SC background? If this is yet another new term, please define.

> **Response:** Because the area of interest may be smaller than an ecoregion, the adjective ecoregional had been used here. The word "ecoregional" was deleted from the public review draft.

Comment: page 2-5, Figure 2-1 caption. This was "minimally influenced" in above text. You should maintain the same terminology throughout (so chose either influenced or disturbed) for clarity. Also, as noted above, a definition of "minimally" is required.

Response: To maintain a clear distinction, the terminology used by Stoddard et al. (2006) was modified from “minimally disturbed” to “minimally affected” and least disturbed was retained in the public review draft. Terms were globally edited.

Comment: page 2-6, line 3. Define “ecoregional criteria.”

Response: Text in Section 3.7.1 was edited as follows: “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.”

Comment: page 2-6, line 5. Define “subregional.”

Response: Subregional is an area smaller than a Level III ecoregion.

Comment: page 2-5, lines 5-6. This did not seem like an exhaustive list of the factors that may cause areas to differ. If it is intended to be exhaustive, please say so.

Response: The phrase indicates that only the primary factors are listed and the references are provided. The text in Section 3.7.1 reads: “Portions of the same ecoregions in different political jurisdictions are expected to have similar characteristics with respect to the primary factors that control background SC (Hem, 1985, Griffith, 2014, Olson and Hawkins, 2012, see Section 2.1). These primary factors are underlying geology, physiography, and climate; secondary factors include soils and vegetative cover (Olson and Hawkins, 2012; Griffith, 2014; Hem, 1985).”

Comment: page 2-6, line 21. Perhaps this is explained in detail elsewhere, but what is the definition of “different?” Does it always involve a preference for data sets dominated by sulfate and bicarbonate or was that just for this example? You seem clearer about this in later sections.

Response: Qualitative difference is meant, i.e., waters dominated by different salts. For example, the case examples focus on sulfate and bicarbonate dominated waters $([HCO^{3-}] + [SO_4^{2-}]) > [Cl^-]$ (in mg/L) rather than chloride dominated waters $[Cl^-] > ([HCO^{3-}] + [SO_4^{2-}])$ (in mg/L).

The edited text in Section 3.7.1.1 reads: “The first step in the background matching approach is to assemble the data sets from sampled sites that are distributed across the full range of SC conditions in the new area. All else being equal, the larger the data set, the more reliable the estimate of background SC. Next, sites with qualitatively different ionic mixtures are removed from the data set. In the example case studies, chloride dominated sites are removed from the data set so that background SC is estimated only for sites dominated by sulfate and bicarbonate (i.e., $([HCO^{3-}] + [SO_4^{2-}]) > [Cl^-]$ in mg/L).”

Comment: page 2-6, line 34. Influenced, disturbed, anthropogenically affected? Can you use consistent terminology?

Response: The term anthropogenically was removed.

Comment: page 2-7, line 8. Least versus minimally? It seems you are attaching specific, different meanings to these terms. If so, please clarify.

> **Response:** To maintain a clear distinction, the terminology used by Stoddard et al. (2006) was modified from "minimally disturbed" to "minimally affected" and least disturbed was retained in the public review draft. Terms were globally edited.

Comment: page 2-8, lines 19−22. This discussion is supposed to be about differences in background yet here you skip ahead to criteria development.

> **Response:** The text was edited in Section 3.7.1.4 of the public review draft: "If the estimated background SC and the ion composition is similar is higher or lower in the new area compared to the original area, two possible causes should be considered. First, differences may be due to natural geological factors (e.g., higher SC due to salt springs, lower SC due to glacial melt, or other differences due to natural geological features) or to climatological factors. In these situations, criteria for the new area can be developed using either the XCD method or the B-C regression method. Second, differences may be due to widespread anthropogenic changes that have increased the apparent background (e.g., due to irrigation, agriculture, impervious surfaces, resource extraction, or acid deposition, etc.). In this second situation, the criteria developed for the original area may or may not be applicable to the new area.
>
> To distinguish between these two possibilities, the cause of the apparent background can be evaluated in a weight of-evidence analysis based on geology, land use, ionic signatures and known inputs, historical and recent trend analysis, atmospheric sources, discontinuities in background across political boundaries, identification of high and low SC anomalies, stream size and connectivity, data set characteristics, sampling methods, and biological evidence of past and present observation of susceptible genera (see Appendix C)."

Comment: page 2-10, lines 7−8. It would be better if this section about the model followed immediately after this one. Both should be before the case study (i.e., application) section.

> **Response:** The figure of the background-to-criterion (B-C) model (Figure 3-10 in public review draft) is presented in the main document and in Appendix D.

Comment: page 2-10, line 25. You have now mentioned tolerance limits, confidence intervals, prediction intervals, and prediction limits without clarifying whether they all have specific uses in this process or are just being mentioned. Why the PL? And why the 95% PL here but the 10% PL in the following figure and text?

> **Response:** The 95% prediction interval (PI) was edited in Section 3.7.2: "The central tendency of a regression model is more robust than any single measurement. For the purpose of model development, data requirements were relaxed relative to those for calculating a HC05 using the field-XCD method (i.e., fewer than 90 genera across 500 sites) (see Appendix D for a description of data requirements for the B C method). Individually, many of the 24 HC05 values used to develop the B C method have not been subject to analyses needed for development of a CCC and should be considered as provisional. For example, the HC05 estimates used in this model were not supported by full confounding analyses, as is

described in the EPA Benchmark Report (U.S. EPA, 2011a). However, the true HC05 value is expected lie between the upper and lower 50% prediction limit (PL). Values in Appendix D, Table D 3 are provisional with a good degree of confidence owing to the larger sample sizes (>60 samples) used to estimate background SC. Table D-4 lists ecoregions with background estimates based on modest survey data sets (N = 20-60 samples) and would benefit from additional sampling to confirm the calculated background SC and the calculated HC05. Table D-5 lists ecoregions where the data set may represent fewer than 25% minimally affected streams and therefore are protective of aquatic life in least disturbed streams. Table D-6 lists ecoregions that may not be served by the B-C method because the ionic mixture is likely to be different (e.g., chloride dominated), the estimated natural background SC exceeds the range of the model, and/or there were fewer than 20 samples available. In all cases, the EPA recommends using the largest data set possible to estimate background SC, understanding and accounting for areas with different (higher or lower) background SC, and performing independent calculations to derive HC05 values."

Comment: page 2-11, Figure 2-2. It would help if the boxes in this figure were lettered or numbered and those letters/numbers were then referenced in the text below.

Response: The four boxes at the bottom of Figure 3-11 were numbered.

Comment: page 2-11, Figure 2-2 and accompanying text. Text and figure should match better. In the figure, what happens after you develop >500 paired data? You need to add some arrows out of this box and some explanatory text here. This figure should match Figure 3-5.

Response: Figure 3-11 was redrawn to make each step separate.

Comment: page 2-11, last sentence. Based on the figure and the text, it's not clear what these two possible values are and where you compare them. Rather it seems you choose between three possible HC_{05} values: mean B-C modeled, Field-SD derived, or the 10% PI of the B-C modeled HC_{05}. This discrepancy needs to be substantially clarified.

Response: The text was edited in Section 3.7.2.1: "Where the background is less than 626 μS/cm and the waters have a similar ion composition to those used to derive the model, the B C method can be used (see Figure 3 11). Where there are >200 but <500 sites with paired biological and SC data, HC_{05} values are derived using the field-XCD method and compared to the mean and lower 50% PL of the the B C model. These values are compared to select the appropriate HC_{05} as follows. (1) If the field XCD HC_{05} (see eq 3 1) is greater than the mean B C modeled HC_{05} (see eq 3 4), then the mean B C modeled HC05 is recommended (see eq 3 4) as a conservative approach to account for uncertainty associated with a smaller data set. (2) If the field XCD HC_{05} is between the mean B C modeled HC_{05} and the lower 50% PL, then the field XCD HC_{05} is recommended because the XCD from measured data from the region is more likely to represent the region than the more general B-C model. (3) If the field XCD estimate is below the lower 50% PL, then the lower 50% PL is recommended as the HC_{05} (see eq 3 5). Because the XCD is calculated from a smaller data set, it may be overly protective and is more uncertain than the modeled results which indicate that 75% of HC_{05} values from areas with a similar background SC are estimated to be greater than that value. Also, the lower 50% PL is also recommended when there are fewer than 200 paired biological samples because there is no XCD for comparison. The SC data and the B-C model is used to estimate the HC_{05}. (4) Where the background SC is greater than 626 µS/cm, the range of the model is exceeded, and it is recommended that data be collected to derive the HC_{05} using the XCD method (see Section 3.1)."

Comment: page 2-12, 1st paragraph (no line numbers). This appears in the figure and thus needs to be included and explained here. The text itself could also be written more clearly. For example, "If the field-SD HC_{05} is between the 10% PL and mean HC_{05} calculated with the B-C model, then…"

Response: See previous response.

Comment: page 2-12, 1st paragraph, last sentence: It is not clear what you were trying to say here. Are these different PLs or the same? Are you using the upper or lower elsewhere? Are you switching to the lower here only? Please clarify.

Response: See previous response.

Comment: page 2-12, line 15. Are you using PL interchangeably with PI? If so, PI is the more common term. Presumably "y hat" is the HC_{05} since, based on the text, that seems the only term for which you need a PI. Yes? If so, specify your parameters exactly.

Response: When the prediction interval (PI) is calculated both the upper and lower prediction limit (PL) is defined. For the background-to-criterion (B-C) model the mean and lower 50% PI is used to determine the criterion continuous concentration (CCC). The upper 50% PI may be used provisionally as the criterion maximum exposure concentration (CMEC) if there is insufficient data to calculate a CMEC from chemical data. "y hat" is the log 10 of the mean predicted HC_{05} from the model. The definition was added to the list of symbols below equation 3-5.

Comment: page 2-12, table. This table and the same table in Section 3 need to match and both need to be corrected for their deficiencies with respect to clearly defining and labeling parameters.

Response: The definitions for the equation in Section 3.7.2.3 were updated as recommend as follows:

Symbol	Explanation	Example from the 23 Ecoregion B-C Model
$\hat{y}$	Log 10 of the mean predicted HC_{05}	variable differs for each case
n	number of samples	n= 24
α	Alpha error rate for prediction interval (desired confidence level)	50% prediction interval (α = 0.5)
t_{n-2}	Student's t value at specified confidence level (alpha, α) and n-2 degrees of freedom	for 50% prediction interval (α = 0.5), $t_{(1-0.5)/2,24-2}$ = 0.686

S_y	Residual standard error of prediction (standard deviation)	S_y = 0.11
SS	Sum of square of *x* deviation from their mean, SS = $\sum_{i=1}^{n}(x_i - \bar{x})^2$	*SS* = 4.21
$\bar{x}$	Mean *x* values used in the model generation	$\bar{x}$ = 2.15
x°	A new log 10 background (*x*) value for a new prediction interval	SC value differs for each case
PL	Upper and lower prediction limits of mean predicted $\tilde{\hat{y}}$	SC value differs for each case

Comment: page 2-13, line 17. In the text and figures above, you seem to be applying the PI/PL only to the HC_{05} and not to the SC, so it's not clear what's happening here. Are you saying that the X term in Equation 2-1 is really the 10% PI of the background SC? Please clarify.

> **Response:** Text was edited in Section 3.7.2.4: "When there are insufficient paired data, background SC is used to calculate the lower 50% PL of the mean HC_{05} which is rounded to two significant figures to yield the CCC (see eq 3-5). This result is shown in the box 3 in Figure 3-11. The estimated CCC at the lower 50% PL for 62 ecoregions can be found in Tables D-3, D-4, and D-5 in Appendix D."

Comment: page 2-14, line 4. This is very confusing (and doesn't match the figure) in that you really always have three possible choices, not two.

> **Response:** Detailed edits were made to more clearly describe Figure 3-11, the decision tree for calculating and applying a hazardous concentration of the 5th centile of a hazardous concentration (HC_{05}). The terminus of each decision path is called out in the summary immediately following the figure in Section 3.7.2.1.

Comment: page 2-14, Section 2.2.6. No details of this calculation are provided here. The calculation of the CMEC is shown in Section 3 but none of the terms are described or defined there.

> **Response:** The details for the criterion maximum exposure concentration (CMEC) are described in a section that underwent a separate peer review. The call-out to Section 3.2 points to that section. No change required.

2.2.4. Reviewer #4

Comment: The strength of the association between background SC and HC_{05} (r=0.93) is one of the most important reasons for the strength of the defensibility of this method. This says that there is a strong association between them, and thus, strong scientific defensibility for the B-C regression model to establish a criterion. And as stated, pg. 2-9, lines 26-27, "gives the importance of the concentrations of major ions in defining the tolerance of species", this relationship between background SC and the HC_{05} is the foundation of the premise, and will be reliable, applicable and calculable country-wide. Although background SC varies considerably, the strong relationship between background SC and HC_{05} in the model proves the process is statistically sound.

Response: Thank you for your comment.

Comment: If there is any area that remotely would be held to question would be in the arena of estimation of input values. For example: "SC XC_{95} values were estimated", and "estimates of minimally disturbed background for each ecoregion" (pg. 2-9, line 31−33). However, trained modeling analysts—who are familiar with this process and the data (and ideally, a familiarity with field conditions/measurements)—should be able to produce the estimates with the same standardization of process and accuracy for all ecoregions.

Response: Thank you for your comment.

Comment: The discussion on pg. 2-10, lines 8−16, informs of the "relaxing" of data requirements and although initially a concern of mine, acceptability is made with the reminder that, "The central tendency of a regression model is more robust than any single measurement,", line 8, pg. 2-10. Nevertheless, a short description of what "relaxing of the data requirements" consists of should be briefly addressed here in the early discussions. I see it also appears on pg. 4-30. Additionally, use of the 25th centile to estimate background SC because it is less susceptible to error than the 10th might also be relevant.

Response: The method defines the data set requirements, so no change was made at this location. The relevant section in Appendix D Section D.2.1 of the public review draft reads:

"State data sets from 48 ecoregions were considered. Five requirements were included:

- A minimum of 200 sampled sites for both biology and SC. This sample size is less than what is used to derive a stand-alone region-specific criterion. However, it is a sufficient sample size to derive HC_{05} values that are used as a group to represent a general relationship between the occurrence of salt-intolerant genera and regional SC levels represented in the data set.
- Taxonomic identification to genus of all individuals or a minimum sample or subsample of 100 individuals;
- Information about the ionic composition of streams in the region indicating that on a mass basis ($[HCO_3^-] + [SO_4^{2-}]) > [Cl^-]$, and $([Ca^{2+}] + [Mg^{2+}]) > ([Na^-] + [K^-])$;
- At least some samples with SC greater than 1,000 µS/cm to ensure a range of exposures; and
- More than 5% of XC_{95} values have 95% confidence bounds on the genus generalized additive model that intersect zero occurrence (see Section 3.1.2.1); that is, extirpation occurs in the sampled SC range."

Comment: I found the discussion in 2.2.1 to be helpful (Figure 2-2 less helpful) in understanding the material since statistics are not my area of expertise, and since I have limited capability in this area, I'm not able to provide comment on the formulae for calculating the lower and upper 10% prediction limits.

Response: Thank you for your comment.

Comment: I have some additional comments in support of the scientific defensibility of the document. EPA has shown caution and expert development of the method by these following examples:

a) the base-flow modeled SC estimates are not recommended for use in the B-C model because they are not as strongly correlated with HC_{05} as the empirical data;

b) the estimation of HC_{05} only for ecoregions with a background <600 µS/cm to avoid extrapolation beyond the modeled data;

c) avoidance of derivation of a CCC with a method that has not been verified;

d) use of the 10% PL for an ecoregion's SC background to calculate an HC_{05} when there is insufficient paired biological data (<200 paired);

e) when estimating the HC_{05} from the field-SD method and the B-C model, the lower of the two estimates is the one used unless the field-SD estimate is less than the 10% PL from the B-C method; the document states correctly that a relatively small paired data set is too variable;

f) use of the field-SD method to check the B-C model results when there are >200 and <500 paired biological and SC data.

Response: Thank you for your comment.

2.2.5. Reviewer #5

Comment: First, the B-C regression model is obviously robust enough to produce an acceptable CCC for a region, essentially because SC and HC_{05} had a very strong association which did not really surprise me (I certainly like r values greater than 0.90 for modeling purposes). I am not sure of the validity of the statement that the B-C method can be used at smaller scales than the L III ecoregions—here, there would need to be some ground-truthing, or perhaps some exploratory analyses with L IV ecoregions (or even smaller areas) across a broad geographical scale. I like Figure 2-2—it really is good for showing the logical pathway for deriving an HC_{05}. Nice way of explaining a difficult concept, especially with the supporting formulas in 2.2.2 and 2.2.3.

Response: Thank you for your comment.

2.3. PLEASE COMMENT ON THE CASE EXAMPLE FOR ECOREGION 43. PLEASE COMMENT ON WHETHER THE CASE EXAMPLE CLEARLY DEMONSTRATES THE METHOD. (SECTION 3, PAGES 3-15 TO 3-26)

2.3.1. Reviewer #1

Comment: In general, the case example is reasonably well laid out. The inclusion of the USGS data muddies the water a bit, because the reader is drawn to differences between the datasets and their associated estimates of the 25^{th} centile of SC. The decision to use the EPA estimate (492) rather than the USGS estimate (567) comes off as somewhat arbitrary. Would it make sense to provide the CCC outcome had the USGS data been used instead? Given that the USGS dataset is reasonably robust, it would be reassuring to readers if the CCC outcome from the USGS dataset was reasonably close to the CCC outcome from the EPA dataset. If they are not reasonably close, this would raise significant red flags about the method.

> **Response:** The EPA-survey and USGS data sets were combined for Case Studies III and IV and appropriate caveats were added.

Comment: P. 3-15, line 7: more details of how ionic composition was evaluated would be useful.

> **Response:** Details regarding the ionic composition were described in parts of the main methods document (public review draft) that underwent a separate external peer review. For the case studies in the public review draft, sites dominated by chloride were removed prior to analysis.

Comment: P. 3-16, line 3: spell out CMEC.

> **Response:** The first instance was identified for all abbreviations and they are spelled out.

Comment: Table 3-1: It should be clear what this table is presenting. Are these paired chemistry/biology samples? If that is the case, there appears to be more the 200 sites which would contradict the text above. What data are contained for the # of samples? From later reading I deduce that it appears to be chemistry samples, but it should be explicit in the table legend.

> **Response:** This case study uses only the chemistry data for the region. The legend for Table 6-1 in the public review draft reads, "EPA and U.S. Geological Survey (USGS) chemistry data sets included in this study."

Comment: P 3-17 line 2−8: be explicit about what the data are….is this chemistry only? Perhaps it should be made clear in section 3.1.1 somewhere what the estimate of the 25th centile of SC was (492), because it is presented clearly for the USGS data in the following section.

> **Response:** Section 6.1.5 "Comparison of Background Specific Conductivity (SC) Estimates" was added to the public review draft: "In this case example, the area weighted average predicted mean base flow SC from the Olson-Hawkin's model in Ecoregion 43 is 489 µS/cm. Raw values for predicted mean flow for 10 digit HUs are shown as box plot in Figure 6-4. The 25th centile of the EPA-survey data set is 483 µS/cm. The USGS SC data set has a slightly narrower overall range and mid-range of values resulting in a slightly higher quartile SC than the EPA survey data set (564 µS/cm) (see Figure 6-4 and Table 6-3). When the USGS data set is combined with the EPA-survey data set, the 25th centile (542 µS/cm) is also greater than the predicted mean baseflow of 489 µS/cm (see Table 6-4, Figure 6-4). Therefore, this background SC estimated from the combined EPA-USGS data set is least disturbed.
>
> The 25th centile of the combined EPA-survey and USGS data set was used to calculate the HC05 following the decision tree described Section 3.7.2 and Figure 3-11. Because the

lower quartile ranges from 85 to 564 µS/cm, additional analysis is recommended for streams known to have low SC regimes."

Comment: P. 3-18, line 21−22: this would indicate that the background for this region is towards the extreme end of the model. >550 µS/cm could mean anything greater than this value, should an estimate be presented here without the > sign?

Response: Edited as in previous comment.

Comment: Table 3-2: the last row of data does not make intuitive sense to me. What are the units?

Response: A footnote was added to several tables (e.g. Table 6-4): [a] $HCO_3^- + SO_4^{2-}/Cl^-$ greater than 1 indicates the mixture is dominated by $HCO_3^- + SO_4^{2-}$.

Comment: Should tables 3-2 and 3-3 have a consistent order of analytes for easier comparisons? How should the reader consider the differences between datasets? For example, there is significant disagreement in SO_4 and Mg concentration between the 2 datasets.

Response: Units were reordered in Tables 6-2 and 6-3. The differences between data sets may be due to any of a number of reasons. The United States Geological Survey (USGS) data set is a targeted sampling and was not designed to measure geospatial background, but it is useful for characterizing overall trends.

Comment: p. 3-22 line 9−10: is this the justification for using the EPA estimate of the 25th centile rather than the USGS estimate?

Response: A new section was created with results of background estimates distinct from interpretation and the use of those values: Section 6.1.5 "Comparison of Background Specific Conductivity (SC) Estimates." See response above.

2.3.2. Reviewer #2

Comment: The case study clearly demonstrate the method except it is subject to some of my comments on Question-2 above, particularly on the choice of 10% PL and the reasoning for adopting the field-based or predicted HC_{05} or lower limits of 10% PL.

Response: The choice of which prediction limit to use is a matter of judgment. Section 3.7.2.1 now reads: "(3) If the field XCD estimate is below the lower 50% PL, then the lower 50% PL is recommended as the HC_{05} (see eq 3-5). Because the XCD is calculated from a smaller data set, it may be overly protective and is more uncertain than the modeled results which indicate that 75% of HC_{05} values from areas with a similar background SC are estimated to be greater than that value. Also, the lower 50% PL is also recommended when there are fewer than 200 paired biological samples because there is no XCD for comparison. If the field-XCD estimate is below the lower 50% PL, then the lower 50% PL is recommended as the HC_{05} (see eq 3-5) because the field-XCD calculated from small data sets may be overly protective and is more uncertain than the modeled results."

Comment: I also made several minor comments in the documents.

Why is this argument not applicable to the case above? And why the argument for the case above not applicable here? Confusing. This argument is vague. How does sample size become an issue here, but not in the two cases above? Why is the field-based HC_{05} here necessarily over-protective? Any biological or statistical reason? This whole section needs to be re-written.

> **Response:** The arguments are different because the two approaches rely on different inferences. The background-matching approach compares background of two areas and infers a similar HC_{05} calculated from a robust data set. The 95% CIs are narrow for background SC and a smaller CI would result in few matches between areas with similar HC_{05} values.
>
> Whereas, the application of the B-C approach compares HC_{05} values as they relate to a prediction limit from a regression model using background SC. The text was revised to clarify the rationale for the recommendations to use either the calculated HC_{05} from a small data set, the B-C modeled HC_{05}, or the lower bound of the 50% PL.
>
> Please note that after the external peer review, the section was rewritten and a different PL was selected by a workgroup of EPA scientists because it was found that using the lower 10% PL could result in estimated HC_{05} values below background in some instances. The 50% PI was adopted instead of the 80% PI, i.e. the 10% lower PI.
>
> The text was edited in Section 3.7.2.1 as follows: "Where the background is less than 626 µS/cm and the waters have a similar ion composition to those used to derive the model, the B-C method can be used (see Figure 3-11). Where there are >200 but <500 sites with paired biological and SC data, HC_{05} values are derived using the field-XCD method and compared to the mean and lower 50% PL of the the B-C model. These values are compared to select the appropriate HC_{05} as follows. (1) If the field-XCD HC_{05} (see eq 3-1) is greater than the mean B-C modeled HC_{05} (see eq 3-4), then the mean B-C modeled HC_{05} is recommended (see eq 3-4) as a conservative approach to account for uncertainty associated with a smaller data set. (2) If the field-XCD HC_{05} is between the mean B-C modeled HC_{05} and the lower 50% PL, then the field-XCD HC_{05} is recommended because the XCD from measured data from the region is more likely to represent the region than the more general B-C model. (3) If the field-XCD estimate is below the lower 50% PL, then the lower 50% PL is recommended as the HC_{05} (see eq 3-5). Because the XCD is calculated from a smaller data set, it may be overly protective and is more uncertain than the modeled results which indicate that 75% of HC_{05} values from areas with a similar background SC are estimated to be greater than that value. Also, the lower 50% PL is also recommended when there are fewer than 200 paired biological samples because there is no XCD for comparison. The SC data and the B-C model is used to estimate the HC_{05}. (4) Where the background SC is greater than 626 µS/cm, the range of the model is exceeded, and it is recommended that data be collected to derive the HC_{05} using the XCD method (see Section 3.1)."

Comment: The reviewer highlighted Fig. 3-4 (now figure 6-4) and asked, "Does the box represent 25^{th} and 75^{th} percentiles? Clarify."

Response: Yes, all box plots in the document are the same and are described in the glossary.

Comment: The reviewer refers to the equation to calculate a prediction interval and comments, "Where does this equation come from? If you did not develop it, give an citation."

Response: The following reference was added to Section 3.7.2.3: Zaiontz, C. 2014. Real Statistics Using Excel: http://www.real-statistics.com/regression/confidence-and-prediction-intervals/

Comment: Any assumption for this equation, such as normality? If yes, is the assumption met?

Response: EPA performed a regression diagnosis on residuals to check for normality and equal variances. Normal probability and residual plots indicate no skewness/deviation or unbalanced residuals.

2.3.3. Reviewer #3

Comment: General: Although this section is not written with great clarity, it does (unlike Section 2) detail the operation of the equations, particularly the numerous log transforms. However, like Section 2, it defines the parameters in generic ways and not with respect to this specific context (i.e., "X" is simply not as informative as "SC"). Overall, however, this case study was much more informative as to how all of this goes together than were the sections both preceding and following it.

Response: Thank you for your comment.

Comment: page 3-15, line 10. It's not clear how you got from the 95% CI to the 25th centile or whether the 95% CI was even used here. Please clarify.

Response: The 95% confidence intervals (CI) are not used here, but other reviewers wanted uncertainty bounds for all calculated values. The method for calculating the CI using bootstrapping is described in Section 3.1.3.1. No change required.

Comment: page 3-15, lines 15−17. This text here makes it clear that there are three ways to do this, something which earlier text (Section 2.2.1) does not. You need to make all of the text consistent when describing these three possibilities.

Response: Revisions were made in Section 3.7.2.1 of the public review draft document to improve clarity and readability. The three possible outcomes referred to in the comment are described in text and depicted in a flow diagram.

Comment: page 3-16, lines 6−9. This seems redundant with discussion in Section 3.1.3.1.

Response: The text is redundant but was retained in deference to previous reviewers who wanted the case studies to be able to stand alone. No change was made.

Comment: page 3-22, lines 9−10. This statement is both unclear and possibly circular. Please re-write this whole paragraph for clarity.

> **Response:** A new section was created with results of background estimates distinct from interpretation and the use of those values: Section 6.1.5 “Comparison of Background Specific Conductivity (SC) Estimates.” See response above.

Comment: page 3-23, line 3. There is information here about the calculation (e.g., the log conversions) that should have been explicitly presented in Section 2.

> **Response:** Edits were made to in numerous locations in the public review draft indicating when log was used in calculations or back transformations to perform an operation. For example see Sections 3.7.2.2, 3.7.2.3, 4.2.1, 4.2.2, 5.2.1, 6.1.6, 6.1.7, etc. and many other places where the use of log10 in a calculation may have been ambiguous.

Comment: page 3-23, line 17. Using “may be” suggests that this step is optional, which it does not appear to be?

> **Response:** The word “is” was substituted for the word “may” as suggested in Section 6.1.6. The change is underlined: “The predicted value Y <u>is</u> converted from log base ten to a number that is the modeled HC_{05} for that region.”

Comment: page 3-23, line 29. Isn’t 743 the actual value, not the log of the value? Please clarify.

> **Response:** Log10 transformations were identified throughout the document for better clarity. In this particular instance log10 was added to the preceding text, and in equation 6-2, a distinction was made between the log predicted HC_{05} and the predicted HC_{05}.

Comment: page 3-25, Section 3.1.5. This table has all of the same flaws as the similar table in Section 2. You are also very unclear about when you are and are not using logs.

> **Response:** Log values are called out in Section 6.1.7 of the public review draft. The edited sections are highlighted in red.

Symbol	Explanation	Example from the B-C Model
$\hat{\bar{y}}$	Log10 of mean predicted HC_{05}	2.87 µS/cm, $\log_{10}$ of 743 µS/cm
n	Number of samples	n = 24
α	Alpha error rate for prediction interval (desired confidence level)	50% prediction interval (α = 0.5)
t_{n-2}	*t*-value at specific level (alpha, α) and degrees of freedom ($n-2$) of interval	For 50% prediction interval (α = 0.5), $t_{(1-0.5)/2,24-2}$ = 0.686
S_y	Residual standard error of prediction (standard deviation)	S_y = 0.11
SS	Sum of square of *x* deviation from their mean, SS = $\sum_{i=1}^{n}(x_i - \bar{x})^2$	*SS* = 4.21
$\bar{x}$	Mean *x* values used in the model generation	$\bar{x}$ = 2.15
x°	*x* value for a new prediction interval	Log10 542 µS/cm = 2.73
PL	Upper and lower prediction limits of mean predicted $\hat{\bar{y}}$	Calculated in eq 6-4

Comment: page 3-26, line 11. Is this a geometric mean? You don't say so here but do say so in Section 3.2. Make these sections consistent.

> **Response:** Yes, this is an annual geometric mean. Section 6.1.8 was corrected in the public review draft.

Comment: page 3-26, line 30 (and equation). What are the terms in this calculation and their basis? Provide a reference for this calculation either here or in Section 2 but preferably in both.

> **Response:** Section 6.1.8 of the public review draft describes how the maximum exposure concentration (CMEC) was calculated for this case study using chemical data. The general method to develop a CMEC is described in Section 3.2 as stated in the case study: "The example calculation of the CMEC for Ecoregion 43 is shown below using eq 3-2 from Section 3.2."

Comment: page 3-27, line 2. Average or geomean? Please clarify.

> **Response:** The opening paragraph of Section 6.2 was edited as suggested in tracked changes provided by the reviewer. It reads: "The case example for Ecoregion 43 includes an annual geometric mean (i.e., CCC) and a 1-day mean (i.e., CMEC), which are not to be exceeded more than once in 3 years on average (see Table 6-6). Both of these distinct expressions of the example SC criteria should be met in order to adequately protect aquatic life. On average, freshwater animals are protected if the annual geometric mean SC concentration in flowing waters does not exceed 620 µS/cm and the 1-day mean does not exceed 1,700 µS/cm more than once every 3 years on average (see Table 6-6)." Also, in Section 6.1.8, the

sentence was clarified, "The example CMEC (see Table 6-5) rounded to two significant figures yields a CMEC of 1,700 µS/cm for Ecoregion 43. If this level is not exceeded, where the annual geometric mean SC <620 µS/cm, 90% of the observations are expected to be less than the CMEC."

Comment: page 3-28, Sections 3.2 and 3.3. Information in Section 3.3 is redundant and could be summarized in one section—Section 3.2.

Response: The section was redundant and was removed.

2.3.4. Reviewer #4

Comment: The case example for Ecoregion 43 was well done and clearly demonstrates the method. Excellent description of the ecoregion and data set characteristics. Use of the lower 10% PL for when there were <200 paired SC and biological data for development of the example HC_{05} is a well-taken and "cautious" approach. The use of the two data sets, the EPA and the USGS, provided significant data points and allowed for verification—all more possible than if there were fewer data available. The map figures were very helpful and Figure 3-5 (Process and decision path case example for Ecoregion 43) was explanatory and helpful in seeing that the chosen process is concise and logical. All material on pg. 3-18 was especially informative and procedures were wisely implemented. It is a well-done case study.

Response: Thank you for your comment.

Comment: I have the following comments and/or questions on Section 3:

Two different data sets were used, the EPA probability samples, and the USGS mixed sampling. It was stated that both were used in the statistical package R: the EPA survey for the characterization of ion concentrations, water chemistry and then the CCC, and, the USGS dataset to "calculate the CMEC" (line 3, pg. 3-16). Further it states that the USGS dataset served to verify the ionic mixture. Since these are two different data sets—collected by different agencies, over different time periods and at different frequencies—is there confidence that they can be used together in the statistical package?

Response: Because the data sets were designed to collect information for different purposes, the EPA compared the data sets for spatial representation (Figure 6-2) and range of SC that was measured (Figure 6-4). The EPA chose to combine the data sets to obtain more complete coverage of the area and addressed uncertainties associated with this decision in the Section 6.2.

Comment: On pg. 3-20, centiles from the two datasets were calculated differently, as noted here in the following:

in the EPA survey data, "Centiles are calculated using each sample observation rather than the mean of site measurements." Line 4−5.

And—

in the USGS dataset, “Centiles are calculated using the geometric mean of site measurements” Line 11−12

> **Response:** In Section 6.1.4.1 the sentence was edited that describes the EPA-survey data set: “Centiles were calculated using each sample observation because most measurements were single grab samples.”
>
> In the case of the United States Geological Survey (USGS) data, there are multiple measurements over the course of a year, hence a geometric mean was taken and compared to the maximum. In the EPA survey data were from one time grab samples, so a site mean could not be calculated. However, the text was clarified in Section 6.1.4.2: “Unlike the EPA-survey data set, the USGS data set are targeted sites of interest rather than probability samples. Also, in some cases, there are multiple measurements from the same site, and other sites are autocorrelated with downstream sampling locations. Therefore, the distribution of sites in this data set is not necessarily representative of Ecoregion 43 in its entirety; however, the data set can be used to define the overall pattern of SC for the ecoregion because it contains samples in areas not represented in the EPA-survey data set. Centiles were calculated using the geometric mean of site measurements (except pH and the ionic ratio) and were qualitatively compared to the probability-based EPA-survey data set.”

Comment: I think it might be wise to clarify the reason for these two different processes for those of us who are limited in our ability for calculations and statistics.

a) In Figure 3-4, pg. 3-22, the box plots of SC distributions for the EPA Survey and the USGS Survey data sets, shows the EPA data with a broader distribution. No specific explanation was provided other than stating that the EPA data had a “slightly broader mid-range of values and a more skewed distribution”. Could a contributing factor for the difference be attributed to the EPA data’s collection occurring through three different survey programs (Table 3-1)? Objectives, methodology and site selection could possibly be slightly different for each of these surveys. In contrast, the USGS data was collected under one overall continuous sampling program over many years. But, the more likely reason is that the USGS data set is so much larger, and with size usually comes a tighter mid-range and a less skewed distribution. My point for all of this is that perhaps this should be briefly addressed more than it has been.

> **Response:** A few edits were made in the case study to make some of these points. For example, Section 6.1.5 compares the different background SC estimates in more detail, and text was added to Figure legend 6-4: “The USGS data set captures a slightly narrower midrange of values possibly owing to the targeted sampling and the mean values rather than the single measurements in the EPA sample.”

Comment: In Figure 3-5, pg. 3-24, third box down: “Using weight of evidence, is different SC background in new area naturally caused?” Shouldn’t this be addressed, regardless of whichever method/pathway is used? It is not shaded gray. I would think that if some of the data points have been influenced by anthropogenic contributions, there is danger of having more than just background levels reflected in the database, yielding inaccurate products in the estimating and modeling. Perhaps this has been factored in already, and/or I have not interpreted the flow chart correctly. If so, please disregard these comments.

Response: Figure 6-5 has had several corrections and been made consistent with the process described in the case study. The method recommends an evaluation of similar background across the ecoregion, however, in this case example, background was not assessed using the weight of evidence described in Appendix C. The 25th centile estimated (542 µS/cm) background was compared with the mean modeled base flow (489 µS/cm), therefore the example likely represents least disturbed conditions

2.3.5. Reviewer #5

Comment: This was a very nice approach using a case example, with key supporting figures. (My one gripe with Tables 3-2 and 3-3 would have been to match the order of analytes between tables—it's a little difficult to bounce back-and-forth. Also, why the '_' in the tables for Mg, K and Na?).

Response: The analytes in Tables 6-3 and 6-4 were reordered for easier comparison as suggested by the reviewer. The underscore lines were removed.

Comment: I like the B-C model derivation (P 3-23)—this makes it clear to readers the process for deriving the CCC using only pre-existing SC values from background information. There are some graphical techniques that I may have used to more closely evaluate the EPA and USGS data, but it would be difficult for many readers to understand this process and the analysis, so let's leave the tables to stand alone.

Response: Thank you for your comment.

Comment: Again, Figure 3-5 is excellent (I have always liked P&D charts). In concert with the text, it illustrates how to effectively utilize existing data for ecological analyses (and here is my gripe, there is more data out there that may be used for these types of analyses but who has the time and resources to pull this together?).

Response: Thank you for your comment.

Comment: Figure 3-6 (P 3-27) leads to a question as to why there are a number of sites with a SD greater than 3.0. I might pull out all sites with a SD greater than 0.25 and see what is driving the variation in this set of sites. Is there some driver (soil, geology, etc.) in Ecoregion 43 that forces the SD outside of the 0.25 boundary, or is there some baseflow mechanisms occurring? Also, the fitted line is difficult to see with all the data points—increase the thickness of fitted line!

Response: The dots on Figure 6-6 were made smaller so that the line was more visible and black rather than red was used for the fitted line.

The additional analysis suggested by the reviewer is interesting, but not necessary for the purpose of this report. The analysis was not performed at this time.

2.4. PLEASE COMMENT REGARDING THE CLARITY OF THE DERIVATION OF THE B-C REGRESSION MODEL AND PREDICTION OF AN HC_{05} FROM MINIMALLY-DISTURBED BACKGROUND. (SECTION 4, PAGES 4-30 TO 4-46)

2.4.1. Reviewer #1

Comment: In general, the presentation is reasonably clear.

Response: Thank you for your comment.

Comment: P. 4-31 line 7: background is not indicated on this plot. Perhaps the language should indicate that SD's are not identical across regions (Fig. 4-3) and that HC_{05}s shift as a function of background conductivities (Fig. 4-4).

Response: The term distinct was added to the methods section and the figure legend for Figure D-3 was edited: "**The lower 30% of the 24 extirpation concentration distributions (XCDs).** Hazardous concentration (HC_{05}) is the specific conductivity (SC) at the 5^{th} centile of each XCD (horizontal dashed line). The ecoregion legend and XCDs are ordered from lowest (NC66) to highest (IL72) HC_{05} and roughly from lowest to highest background SC. Untransformed SC values shown on log10 scaled x axis."

Comment: It is unclear how the reader is to interpret differences between the values presented in table 4-1. Should base flow area weighted means be generally significantly higher than 25^{th} centile estimates? Can a 25^{th} centile estimate be generated from the base flow model? This table makes me think that there is a high degree of uncertainty associated with the 25^{th} centile estimate (x axis of the entire model).

Response: Table 4-1 was updated to Table D-3 in the public review draft. Different sampling designs will yield different estimates of background. However, the B-C model is inherently consistent because the estimate of the background SC and the HC_{05} are from the same data set. The modeled (not measured) natural base flow SC was used to gage whether the 25^{th} centile SC estimated for the dataset is more likely to represent natural background or a disturbed system. Similarly, the state data set was compared with the probability data set to gage any effect from a non-random dataset used to calculate 25^{th} centiles. On average, the 25^{th} centile SCs estimated from the 24 Level III ecoregional state data sets were similar to those from the EPA survey data set ($r = 0.88$). Where there is more than 50 µS/cm difference in the 25^{th} centile between the state and EPA-survey data, a notation has been added to Tables D-3, D-4, and D-5.

Text in D.6.1 was added to increase clarity: "Background SC was estimated at the 25^{th} centile of the available state and EPA survey data sets and evaluated as being representative of minimally affected or least disturbed background for the ecoregion. Background represented minimally affected when the 25th centile was less than the mean natural base flow SC estimated from a geophysical model (Olson and Hawkins, 2012). Background was characterized as least disturbed when the 25^{th} centile measured from the field was greater than the mean natural base flow SC estimated from a geophysical model. The predicted mean natural base flow SC was calculated using geology, climate, soil, vegetation, topography, and other factors calibrated with reference sites (Olson and Hawkins, 2012)."

Comment: Language in the legend of table 4-4 is very difficult to interpret.

Response: New tables (D-3 through D-6) were prepared that sorted ecoregions based on data attributes.

2.4.2. Reviewer #2

Comment: The derivation of the B-C regression model and prediction of an HC_{05} from minimally-disturbed background are clearly described. However, my comments and suggestions for Question 2 are also applicable here. I also like to make another general comment. The B-C regression approach will be robust and highly useful given that XC_{95} can be confidently established for top 5% most sensitive genera.

Response: Thank you for your comment.

Comment: In the EPA 2015 report, many or even most genera examined did not rapidly decline to a low probability level. Although this may not affect the ultimately establishing of HC_{05}, but sensitivity distribution plots, as shown in Fig 4-3, are misleading. This is because XC_{95} estimated based on field data differ greatly in precision and accuracy among genera, but is presented as if they are the same. I suggest using a size symbol (circle or triangle) to indicate the confidence level of XC_{95} estimation.

Response: The suggested circle-triangle recommendation was made and can be seen on most of the extirpation distribution (XCD) plots in the public review draft.

Comment: The reviewer highlighted the term Ecoregional data sets and commented in the document, "What is this dataset? From USGS gauge stations?"

Response: In Section D.3.1 the text was revised for clarity: "The B-C model was developed using paired background SC and HC_{05} values, listed in Table D-2, that were calculated as described for the XCD method (see Section 3.1). HC_{05} values were calculated from XCDs from the 24 data sets. Background SC was estimated at the 25th centile from these same data sets. A linear least squares regression was used to develop the B-C model using the log of background SC as the independent variable and the log of HC_{05} as the dependent variable. The 50% prediction interval (PI) was also calculated." United States Geological Survey (USGS) data sets were not used for this analysis.

2.4.3. Reviewer #3

Comment: General: Technically, the model appears adequate despite being based on only 19 data points.

Response: Thank you for your comment. Note that the B-C model in the public review draft is based on 24 state data sets from 23 Level III ecoregions.

Comment: Here again, however, the text lacks clarity and it is hard to follow your rationale for some of the assumptions inherent in the model (e.g., the 25th centile). In addition, apparently shifting terminology and inconsistent labeling of table columns detracts from its overall clarity.

Response: Specific edits are described below as called for individually by the reviewer.

Comment: page 4-30, line 4. Which method is this? The background method? Please clarify.

Response: Section D.1 was edited: "Not all areas of the country have sufficient water chemistry and biological data to derive criteria for specific conductivity (SC) by the field method of calculating extirpation concentrations (XC_{95}) and hazardous concentrations (HC_{05} values) from an XC_{95} distribution (XCD) (see Section 3.1.3.). For such cases, the U.S. Environmental Protection Agency (EPA) is providing alternative methods that geographically extend results of the primary XCD method. This appendix describes a method to estimate criteria for new areas with different background SC using a background to criterion (B-C) model."

Comment: page 4-30, line 31. In Figure 2-2, you have a cut-off of 600 to stay within the bounds of the model. Yet here you apparently did not include data below 1,000. Please clarify this seeming discrepancy.

Response: The reviewer refers to an unclear bullet in Section D.2.1. The requirement is to include data greater 1,000 µS/cm as well as below, but at least some in the higher range. It has been edited to read, "Some samples with SC greater than 1,000 µS/cm to ensure a range of exposures."

Comment: page 4-31, line 14. This appears to be your definition of "minimally." If so, it needs to be provided earlier in this document.

Response: Minimally affected and least disturbed are defined in the glossary and earlier in the document as recommended. The 25th centile was used as an estimate. In other cases a different centile may be used. In Section D.3.1, edits were made noting that the background-to-criterion (B-C) model undoubtedly includes both minimally affected and disturbed conditions,

"The state data sets may represent either minimally affected or least disturbed (Stoddard, 2006) background SC for an ecoregion. In this document, natural background SC is defined as the range of ionic concentrations naturally occurring in waters that have not been substantially influenced by human activity. Minimally affected background is defined as the physical, chemical, and biological habitat found in the absence of significant human disturbance. Least disturbed background refers to the best available physical, chemical, and biological habitat conditions given the present state of the landscape.

Background SC was estimated at the 25th centile of the available state and EPA survey data sets and evaluated as being representative of minimally affected or least disturbed background for the ecoregion. Background represented minimally affected when the 25th centile was less than the mean natural base flow SC estimated from a geophysical model (Olson and Hawkins, 2012). Background was characterized as least disturbed when the 25th centile measured from the field was greater than the mean natural base flow SC estimated from a geophysical model. The predicted mean natural base flow SC was calculated using geology, climate, soil, vegetation, topography, and other factors calibrated with reference sites (Olson and Hawkins, 2012)."

Comment: page 4-31, line 20. Least versus minimally disturbed. You seem to be fashioning a fairly loose (i.e., poorly defined) connection between disturbance and SC. A little more discussion—beyond simply a reference—would be helpful here.

Response: See previous response.

Comment: page 4-32, table. Add "ecoregional" to column 3 for clarity.

Response: All data used are ecoregional. For clarity the header was changed to background-to-criterion (B-C) model data sets in Table D-1.

Comment: page 4-34, line 10. "Patterns" are very hard to see in Figure 4-1. How did you assess the similarity of patterns here?

Response: The three maps were removed. The legend for Figure D-6 was simplified to point out the general patterns. "Predicted natural base flow for stream specific conductivity (SC). Mountainous regions tend to have lower background SC, the central plains and arid lands tend to have the highest, and other ecoregions have intermediate background SC."

Comment: page 4-34, lines 28−29. Is this the only reason SC could go up in drier months? Could not increased evaporation alone lead to higher SC even absent anthropogenic events?

Response: There are many natural and anthropogenic causes for SC levels and changes in SC and this was not the focus of the document, so this text was removed.

Comment: page 4-31, Figure 4-1 caption. "Similar trends" are not readily apparent from these figures.

Response: The three maps were removed. The legend for Figure D-6 was simplified to point out the general patterns. "Predicted natural base flow for stream specific conductivity (SC). Mountainous regions tend to have lower background SC, the central plains and arid lands tend to have the highest, and other ecoregions have intermediate background SC."

Comment: page 4-36, line 6. Why? Was this discussed elsewhere? If so, please provide a reference.

Response: The section was edited for clarity. "The XCDs for the 24 data sets are shown in Figure D-3. Only the lower 30% is shown because many of the XC_{95} values in the upper portion of the XCDs are greater than the calculated value and are not defined except in a relative sense. Plotting the lower 30% more clearly shows the region of the XCD containing the HC_{05}.

As the background SC increases, the XCDs for the ecoregions plot further to the right at higher SC (see Figure D-3). Similarly, the HC05 identified from the intercept of each XCD with the 5th centile line, increases as background SC increases. The genera that contribute most directly to each HC_{05} are the most salt intolerant genera in each ecoregion."

Comment: page 4-36, line 12. This figure says nothing about HC values.

Response: The caption now explains the relation of the figure to hazardous concentration (HC) values. See the next response.

Comment: page 4-37, Figure 4-3 caption. Where is HC in this figure?

Response: The legend for Figure D-3 was edited to read: "Figure D-3. The lower 30% of the 24 extirpation concentration distributions (XCDs). Hazardous concentration (HC05) is the specific conductivity (SC) at the 5^{th} centile of each XCD (horizontal dashed line). The ecoregion legend and XCDs are ordered from lowest (NC66) to highest (IL72) HC_{05} and roughly from lowest to highest background SC. Untransformed SC values shown on log10 scaled x axis."

Comment: page 4-38, Table 4-3, 6^{th} column. "Empirical" is confusing. This appears to be the 25^{th} centile of just the ecoregional data set. Please clarify. Where did the other two (probability, base-flow) figure in here?

Response: All the tables were reorganized and new column headers clearly call out the state and EPA-survey data set and the mean base flow model."

Comment: page 4-39, Figure 4-4. There is an R-squared value (0.87) in the corner of the figure that does not match any of the r-squared values discussed in the text.

Response: Figure 4-4 is now Figure D-4 in the public review draft. Only the r value (not R-squared) is shown in the public review draft because this is not a test of significance but of strength of the association.

Comment: page 4-40, lines 15−18. You need to keep it clear that you first developed the B-C model (which just equates SC and HC) and then used it to derive a CCC.

Response: Edits were made to the introduction section (D.1) to improve clarity and readability.

Comment: page 4-40, lines 19−22. Perhaps this is detailed elsewhere (if so, reference?) but, if not, you need to provide the reader with more explanation for how you went from the B-C model to these CCCs [which you do in Section 3]. You also need to explain (or reference) "corrected."

Response: Edits were made to the introduction section (D.1) to improve clarity and readability.

Comment: page 4-41, Table 4-4. This is an overly wordy and cumbersome table title. Everything after the first sentence should be moved to a table footnote. "Values in bold" is not clear. Bold italic is minimum of what? How does this relate to the range shown? What is the purpose of including blank rows in this table?

> **Response:** New tables (D-3 through D-6) were prepared that sorted ecoregions based on data attributes and included edits to improve clarity.

2.4.4. Reviewer #4

Comment: I believe that Section 4—the presentation of the development of background to criterion regression model and the development of estimating the minimally disturbed background SC—has been clearly presented. The work has been done carefully, logically and with thorough process. It was understandable and I appreciated the extra effort that seemed to be put into this section.

> **Response:** Thank you for your comment.

Comment: The following are some of the lines and/or concepts of which I particularly like:

a) Many areas of the country, indeed, do not have sufficient water chemical and biological data sets to derive criteria for SC (line 4-5, pg. 4-30). Many state agencies don't have comprehensive monitoring programs and particularly the paired biological and SC data sets. The development of the background-to-criterion regression model method will facilitate the establishment of SC criteria for those ecoregions (and ultimately, states) who haven't complete data sets.

b) The requirements for the 19 ecoregions which met the data set requirements for estimating XC_{95} were listed in the bullets in lines 28−31, pg. 4-30. These reflect carefully selected guidelines.

c) I appreciate it is mentioned that state data sets have been quality assured (lines 34−35, pg. 4-30 and lines 1−2, pg. 4-31). While this is always generally assumed, it is good to see it confirmed in any document that deals with data and its usage.

d) Good to see the discussion which tells why the 25^{th} centile was used to estimate the background SC: 1) the 25^{th} centile is less susceptible to error than the 10^{th} centile, and 2) the 25^{th} centile yielded a slightly stronger statistical model with HC_{05} than using the 10^{th} centile, lines 15−19, pg. 4-31. These illustrate the care taken by the authors to establish guidelines which provided less chance of error to enter the picture.

e) Good to see Table 4-2, "Predictor variables for the conductivity model". I was very interested in seeing these listed. It seems complete.

f) Excellent work in comparing the three methods of background SC estimation: 1) the 19 Level III Ecoregional data sets, 2) probability surveys, and 3) values calculated using a base-flow model (lines, 3−10, pg. 4-34)—and finding that the three methods yielded similar patterns of background SC. It gives significant defensibility to this body of work.

g) Figure 4-1, A, B, and C, pg. 4-35, is excellent and the discussion on pg. 34, lines 13−20, goes over it well. Altogether the text and the figures are very informative and helpful.

h) Similarly, I am glad for the discussion in lines 21−35, pg. 4-29. I concur that estimates of SC during drier months tend to have higher stream SC with some anthropogenic components. Thus the results of model estimates of base-flow only will likely be different from those in which both surface and base-flow are modeled.

i) Page 4-39 is well done. Both the figure (Figure 4-4) and the text (lines 1−10) are informative, and provide strong defensibility for the paper's regression model using the estimated Background SCs from the 19 ecoregional data sets. There is clarity and reasonableness. For me, the method is a wisely developed model and should meet with approval by water quality professionals.

j) The values in Table 4-4, pg. 4-42 (High Plains, Central Great Plains), pg. 4-43 (Northwestern Great Plains, Nebraska Sand Hills, Western Corn Belt Plains) and pg. 4-44 (Central Corn Belt Plains) are in accord with data of which I am familiar, either through collection of samples myself, reviewing of states' databases or analyzing specific watersheds' water quality monitoring data. SC can reach relatively high levels in the central and western Corn Belt and Plains. Your data reflects that accurately.

Response: Thank you for your comment.

Comment: I have two questions and a concern for Section 4.

a) What is the name of the model referred to in line 27, pg. 4-31, "a model that predicts natural base-flow water chemistry using geology, climate, soil, vegetation, topography, and other factors calibrated with reference sites (Olson and Hawkins, 2012)"?

Response: In Section D.3.1 and throughout the document, this model is now referred to as the predictive mean natural base-flow model or base flow model. For example, the text identified by the reviewer is now in a separate section: "The predicted mean natural base-flow water chemistry is calculated using geology, climate, soil, vegetation, topography, and other factors calibrated with reference sites (Olson and Hawkins, 2012). This base-flow model used a random forest method to generate predictions for natural base-flow SC values in streams."

Comment: Has the term "area-weighted" been explained in the document? It is used in the following places: line 29, pg. 4-31; Table 4-1, pg. 4-32; line 8, pg. 4-33, and line 17, pg. 4-34. My memory fails me in that it may have been described in the previously reviewed *Recommended Field-based Method for States to Develop Ambient Aquatic Life Water Quality Criteria for Conductivity*. If not, it might be good to explain how the weighting was done.

Response: The area-weighted method is no longer used. The method for predicting natural mean base-flow SC was revised using the 2016 released NHDPlusV2 catchments that allowed the predictive base-flow model to accumulate and generate full catchment summaries rather than averaging USGS 10 digit hydrologic units after weighting them based on geographical area.

Comment: Lastly, I believe that the abrupt changes in base-flow SC seen at political boundaries (state lines) may be equally as much a result of differences between states' data collections in all respects rather than …"artifacts from geological coverages", although GIS layers and attributes may indeed fail to match-up. When base-flow SC abruptly changes at the border between two states, it is likely because one state has more complete data collection in general, and water quality in particular, than the other state due to greater attention to the distant sites. States vary considerably in the amount of data acquired, and often streams and rivers in outlying, low populated regions of the state receive far less attention (monitoring) than areas closer to the state capitol, populated areas, and less rough terrain. Some state borders lie in distant, low populated areas, and those areas often receive less priority (funding for data collection) by both the state department of environmental quality and the state legislature. Speaking from experience, Iowa's 305(b) and 303(d)/TMDL water quality and biological monitoring programs sample far fewer stream reaches in the western third of Iowa which has lower population than the center and eastern thirds, and does nothing on the Missouri River and very little on its tributaries. Thus, if one state samples more thoroughly on a river system which passes through to the neighboring state, it is not surprising if the two states' databases have distinctly different estimates of SC. One state may have little, or even no sampling data, from which to draw if asked to provide data.

> **Response:** EPA redid the analysis using the 2016 release of the NHDPlusV3 stream delineation and thus the artifact from using hydrologic units is gone. See Figure D-4.

2.4.5. Reviewer #5

Comment: I have no problems with either the clarity of the B-C model derivation or the HC_{05} prediction, especially since there is strong written support and justification. For Table 4-1, I compared some of my calculations for a few of the Eastern U.S. ecoregions, and there was generally good agreement among my estimates using two distinct approaches, although one of mine was based closely on the 25th centile (or percentile).

> **Response:** Thank you for your comment.

Comment: My one concern was the spread in the BC in some of the other ecoregions in Table 4-1. Is the spread due to a sample-size problem? I think there should be some effort by EPA to determine what drove the large spread in the SC estimates.

> **Response:** The correlation between the state data sets and the EPA survey data sets was good. On average, the 25th centile SCs estimated from the 24 Level III ecoregional state data sets were similar to those from the EPA survey data set (r = 0.88). The predicted, area weighted mean, base flow SCs tended to be similar to the mid range 25th centile SC values but were higher at the lower levels and lower at the higher levels compared to the background measured in either the 24 ecoregional State data sets (r = 0.74, n = 24) or the EPA survey data sets (r = 0.80, n = 71). Where there is more than 50 µS/cm difference in the 25th centile between the state and EPA-survey data, a notation has been added to Tables D-3, D-4, and D-5.

Comment: Figure 4-2 tells it all, and gives the reader a good sense of variation in the US for SC—a nice little piece of work. Will probably be used in textbooks in the future, if there are still textbooks.

Response: Thank you for your comment.

Comment: Figure 4-3 is a little chunky. It is difficult to compare among the 19 ecoregions—too hard to distinguish line types and colors. Perhaps if one would cut it off at 0.20, it would focus on those parts of the distributions that are most interesting.

Response: Figure D-5 was redrawn with thicker lines. The 0.30 was retained.

2.5. PLEASE PROVIDE ANY SUGGESTIONS ON WAYS TO STRENGTHEN EITHER THE DEVELOPMENT OF THE B-C MODEL OR THE EXAMPLE CASE STUDY? (SECTION 4.)

2.5.1. Reviewer #1

Comment: As stated above, I think including the EPA and USGS datasets in the case study fosters uncertainty that accurate estimates of the 25th centile are driving the model.

> **Response:** See previous response. The data sets were combined.

Comment: The regression (Fig. 4-4) is impressive and the application is pretty straightforward.

> **Response:** Thank you for your comment.

Comment: It might be useful to show a quantitative assessment how much the CCC estimates would vary if the 25th centiles were over- or under-estimated by different amounts in a given new area. (It could be easily done from the fig 4-4 regression).

> **Response:** The change in the HC_{05}/criterion continuous concentration (CCC) given a change in background is the slope of the line. No additional analyses were performed.

Comment: Perhaps this information is included in another portion of the document, but I would like to see language describing how the Level III ecoregions are delineated. P. 2-5 line 10 mentions "similar considerations" to a list of SC determinants, but the reader should understand how the regions were determined without having to go to the Omernik document.

> **Response:** The factors involved in delineating the Level III ecoregions are too long to reproduce here for all 85 ecoregions. The basic factors used in delineation are repeated in several locations in the document. For example, "There are 85 Level III ecoregions in the contiguous United States (Omernik, 1995, 1987) characterized by living and nonliving components of the region, including geology, physiography, vegetation, animal life, climate, soils, water quality, and hydrology."

Comment: Our approach to protecting humans involves both epidemiological and mechanistic approaches to understanding health outcomes. When mechanisms and observations support each other, we have the most confidence. Our efforts to protect ecological health in the face of salinization are exclusively epidemiological in nature. Some basic science is much needed and would strengthen our understanding of salinity-tolerance relationships among species. For example, we have very little basic understanding of how ion toxicity works, particularly in aquatic insects. Are ion rich matrices toxic because of competitive exclusion of physiologically important ions? Are ion rich matrices toxic because insects self-poison the hemolymph because of excessive ion uptake? Are ion rich matrices toxic because most insects are most energetically efficient in relatively more dilute waters? We know for example that ionic matrices with similar conductivities but different ionic compositions are differentially toxic to the same aquatic organisms (see Kunz et al, 2013 for example). There is no acknowledgement that SC-species responses can be strongly affected by ionic composition beyond the crude lumping of anions as descriptors of water type. I generally applaud the efforts that EPA has taken to make use of field data in light of our lack of fundamental understanding of this topic, but I think we could better interpret field data if we had a better understanding of how and why species respond differently to salinity.

> **Response:** Thank you for your comment.

2.5.2. Reviewer #2

Comment: See my suggestions above.

> **Response:** Comments not included above are transcribed from the tracked changes and comments in the text.

Comment: In Section D- Data Sets, the reviewer highlighted, “Taxonomic identification to genus and commented in the document, “Sampling effort? 300- or 500 counts?”

> **Response:** The bulleted requirement for inclusion was expanded in Section D.2.1 to read: “Taxonomic identification to genus of all individuals or a minimum subsample of 100 individuals."

2.5.3. Reviewer #3

Comment: The major sections do not flow in a logical sequence (for example, the model is explained after its use in the case study). You are introducing one method then part of the other, then adding a case study for a CCC, and then coming back to further explanation of second method. It would be much better if you introduced and discussed one method and then the other and put the case study at the end.

> **Response:** The main methods (Section 3) and the case studies are intended to be stand-alone documents. The development of the B-C model is a supporting analysis.

Comment: I’ve included a mark-up of the whole document which shows the exact locations of the above comments and also includes lesser comments of an editorial nature, adoption of which might improve the document’s clarity.

> **Response:** Almost all of the editorial comments were accepted and are not replicated here.

2.5.4. Reviewer #4

Comment: As I progressed through the questions. Also, since *Application to New Areas* will be part of the larger previous document, some of the fuller descriptions and explanations are in it and it’s unnecessary to repeat them in this document.

> **Response:** Thank you for your comment.

Comment: One suggestion which I would like to mention here is that a brief note as to the nature of the Nebraska Sand Hills should be included. (I realize they are not included in the ecoregion but they border it immediately to the south.) There is no mention of that area being any different than the rest of the general region, pg. 3-15, yet the Sand Hills are an exceptionally unique area and significantly different from all of the land area surrounding them and to the north. The Sand Hills are comprised of wind deposited sand, blowouts, and arid vegetation intermingled with wet meadows, rich hay fields, grassed pastures, high water tables, and natural lakes. There is only one other site in the world with the wind-deposited sand hill geology like this and it is in China. And since the Sand Hills clearly stand-out as different from their surroundings in the maps of Figure 4-1 and Figure 4-2, pgs. 4-35 and 4-36, respectively, it would add to the document to explain why.

> **Response:** EPA agrees the Sand Hills stand out as one of the lower conductivity areas in the Central Plains. However, EPA did not expand this discussion.

Editing/Typos

Comment: Last line in top paragraph just above the title for Section 2.2.2 (line numbers are absent here): the word "and" should be deleted.

> **Response:** Edited as suggested.

Comment: In Figure 3-1's description, pg. 3-17, there are two errors:

1. A space is needed after the comma in the third line down.

> **Response:** Edited as suggested.

Comment: In the same paragraph under Figure 3-1, the red symbol on the map should be "triangles" rather than "diamond". Diamonds have already been designated green, and the symbol which is red on the map is in the shape of a triangle.

> **Response:** Edited as suggested.

Comment: The same problem lies with the map in Figure 3-2, pg. 3-19. The red symbols should be "triangles" rather than "diamond".

> **Response:** Edited as suggested.

Comment: I believe that *Application to New Areas* has been well done and will provide significant assistance to the determination of specific conductivity criterion throughout the states. I appreciate this opportunity to provide peer review of the document and wish it success.

> **Response:** Thank you for your comment.

2.5.5. Reviewer #5

Comment: Throughout the review (1-4), I have made some suggestions on other analyses that could be undertaken. I commend all on worked on this task for their hard work. It will be interesting over the next few years to see how SC criteria play out.

> **Response:** Thank you for your comment.

Comment: 6. Other Comments:

P 1-1, L 26—Vander Laan et al. 2013 is not in references. So, if one reference is missing (a warning signal to me in my past life as a journal editor), was the report thoroughly checked for citations, and dates, etc., not only in the bibliography but also in the text? **If not, do it**.

> **Response:** All references were checked and cross-checked. The Vander Laan et al. reference is cited.

Comment: P 1-3—In figure caption, rephrase sentence 3—it is clumsy and not well constructed to convey meaning.

> **Response:** The caption for Figure 2.2. was edited to read: "A species' (or genus') realized niche is defined by its lower and upper limits of occurrence. In this case, the lower tolerance limit is less than or equal to the lowest specific conductivity (SC). The XC_{95} represents the upper tolerance limit. Approximately 5% of observations of a taxon are assumed to occur in sink habitats where a population cannot persist without immigration from source habitats. A species or genus optimum is the environmental condition most easily tolerated both physiologically and competitively and can be estimated by the conditions where the taxon is most often observed. The optimum SC may be estimated at the maximum probability of observing the taxon from a generalized additive model, shown here. The example is of the genus *Ephemerella* comprised of several species of mayflies."

Comment: P 1-3—Period missing at end of figure caption.

> **Response:** The figure caption's last sentence ends with a period.

Comment: P 3-16, L 6−15—Although I commented on this earlier, I wonder if there were any analyses based on the ionic mixtures used in the B-C regression. Since chloride dominated samples were removed from the data set, I wonder what would happen if an analysis was extended to examine the range of sulfate and bicarbonate dominated samples in relation to chloride. For example, what if bicarbonate and sulfate exceeded chloride by just a small amount, or what would the analyses look like if it exceeded it by a large amount, or other levels in between? It appears that there are some data sets (e.g., the USGS sets) where this analysis could be done.

> **Response:** These analyses were considered and some pilot studies were performed. At this time, the case studies use field data to demonstrate how to apply the methods described in this document to derive example criteria for SC for flowing waters contaminated by calcium, magnesium, sulfate, and bicarbonate ions and to a lesser extent to sodium, potassium, sulfate, and bicarbonate ions but not for flowing waters polluted by chloride-dominated salts. EPA

chose to limit these examples to a well-defined ionic mixture because the data for chloride dominant sites are less abundant.

Comment: P 3-20−P 3-21**—**I would be a little careful doing any kind of statistics on ratios (bicarbonate plus sulfate/chloride in Tables 3-2 and 3-3)**—**this is akin to doing statistics on diversity indices or other ecological indices (see Sokal and Rohlf, Zar, Green, Krebs, etc.). I also wonder if pH was converted to hydrogen ion concentrations for stats and then converted back (calculation of pH SD or SE is a bit of a bear, but 95% CI is adequate)? **<u>You cannot average pH values</u>**.

Response: No analyses were done with ionic ratios or pH. They were used strictly to identify sites dominated by chloride salts and sites with pH <6.

Comment: P 4-32, Table 4-1**—**In the probability survey 25th centile, watch significant figures**—**round! **<u>So, check all significant figures in all tables</u>**.

Response: Figures were not rounded until the CCC or CMEC was rounded to 2 significant figures. By not converting to significant figures for intermediate values, it allows EPA to check for anomalies and avoid rounding errors. Some rounding was necessary for describing calculations with log 10 transformed values, but were not used in actual calculations.