TASK ORDER 51 UNDER
CONTRACT EP-C-12-029

PEER REVIEW OF EPA'S DRAFT DOCUMENT "APPLICATION TO NEW AREAS"

PEER REVIEW SUMMARY REPORT

July 29, 2015

Submitted to:
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
Attn: Luis Cruz
Cruz.Luis@epa.gov

Submitted by:
Eastern Research Group, Inc.
110 Hartwell Avenue
Lexington, MA 02421

CONTENTS

1.0 INTRODUCTION

This report documents the results of an independent peer review of the draft, *Application to New Areas*, developed by the U.S. Environmental Protection Agency (EPA) to describe the estimation of an HC_{05} based on a model of background specific conductivity. Eastern Research Group, Inc. (ERG, a contractor to EPA) organized this review and developed this report. Section 2 of this report provides the individual reviewer comments for each charge question, as well as a summary of these comments. New information provided by the reviewers is presented in Section 3. Appendices A and B provide, respectively, the charge to reviewers and the complete set of comments submitted by each reviewer. EPA subsequently edited this report to remove the names of reviewers so that comments are not attributed to specific individuals.

1.1 Background

EPA's Office of Water is charged with protecting ecological integrity and human health from adverse anthropogenic, water-mediated effects, under the purview of the Clean Water Act (CWA) and Safe Drinking Water Act. Elevated conductivity has been shown to impact aquatic life designated uses in a range of freshwater resources. Elevated conductivity is associated with multiple anthropogenic sources, including discharge from wastewater treatment facilities, ground water infiltration affected by climate change, surface mining, oil and gas exploration, runoff from urban areas, and discharge of agricultural irrigation return waters, among others. EPA, through its authority under CWA Section 304(a), developed a draft *Recommended Field-based Method for States to Develop Ambient Aquatic Life Water Quality Criteria for Conductivity,* a method that states can use to derive field-based ecoregional ambient aquatic life criteria for specific conductivity (SC), a measurement of ionic strength or concentration. This draft document underwent a contractor-managed external letter peer review in late 2014. Based on concepts presented and reviewed in the draft *Recommended Field-based Method for States to Develop Ambient Aquatic Life Water Quality Criteria for Conductivity,* as well as external peer reviewer comments, EPA developed additional methodology for inclusion in the draft recommended method. EPA required a peer review of this additional methodology presented in the draft document entitled, *Application to New Areas*.

This draft presents a model that estimates a hazardous concentration level (HC_{05}) for specific conductivity based on background conductivity to protect aquatic life uses. Because large, paired chemical and biological data sets are not available for all ecoregions in the U.S., it would be useful to have a way to identify an SC Criterion Continuous Concentration (CCC) by another method. The draft characterizes the relationship between background SC and the presence of genera that fill low SC niche space. The draft shows how a HC_{05} can be estimated from a model of 19 ecoregions that model the relationship using ecoregional background for SC as the independent variable.

1.2 Peer Reviewers

ERG searched, identified, and selected five reviewers who met the technical selection criteria provided by EPA and who had no conflict of interest in performing the review:

- **David Buchwalter, Ph.D.**: Associate Professor, Department of Biological Sciences, North Carolina State University.
- **Yong Cao, Ph.D.**: Stream Ecologist/Associate Program Leader, Illinois Natural History Survey and Adjunct Associate Professor, Department of Natural Resources and Environmental Science, University of Illinois at Urbana-Champaign.
- **Bruce K. Hope, Ph.D.**: Private Consultant in Environmental Toxicology and Risk Assessment, Oregon.
- **Marion R.L. Maas, Ph.D.**: Private Consultant, Nebraska.

- **Raymond P. Morgan II, Ph.D.**: Professor of Aquatic Science, Appalachian Laboratory, University of Maryland Center for Environmental Science.

ERG provided reviewers with instructions, the review document, the charge to reviewers (Appendix A) prepared by EPA, and supporting background materials as described in the charge. Reviewers worked individually to develop written comments in response to the charge questions. After receiving reviewer comments, ERG summarized reviewers' responses to the charge questions, noting areas of agreement and disagreement, where relevant (Section 2).

2.0 SUMMARIES ORGANIZED BY CHARGE QUESTION

This section presents summaries of reviewer comments, organized by charge question. Each summary is followed by a table presenting the individual reviewer comments by charge question. Individual comments are provided exactly as submitted by reviewers (see Appendix B for reviewer submissions).

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.)

Four reviewers thought that the ecological theory was presented clearly, though in one case, the reviewer specified that the presentation was adequate because the review document is an addition to a larger document that explains the ecological theory in more depth. The fifth reviewer thought the theory was not clearly presented, and noted that the use of "niche," "range," and "limits" was confusing. This reviewer suggested that EPA use a simple Gaussian curve to illustrate the relationship between these terms.

Two reviewers commented that the ecological theory was adequately linked to the B-C model. One of them thought that the B-C model, when implemented, would serve as an important tool to support better evaluation of water quality in the U.S. The other suggested that EPA add some basic ionic physiology (e.g., a gill physiology molecular biology approach) to support the methods, as well as an examination of the evolution of salt-intolerant North American species The three other reviewers did not think that the section sufficiently linked this theory to the B-C model. One of them noted that no ecological theory is presented to show why there should be a consistent relationship between the low end of an ecoregion's exposure concentration (EC) values and extirpation concentration (XC95) values to support the B-C model. Another pointed out that the discussion switches suddenly from species to genera without an explanation.

A reviewer noted that Figure 1-1 should be changed or expanded, explaining that the realized niche in this example is apparently limited by the availability of low EC waters and not by the organism's physiology. The reviewer suggested including a second example in which the probability of observation shows a bell curve. Another reviewer noted that the labels and caption in Figure 1-1 need revision.

A reviewer noted that this section seems to conflate two very different scenarios: (1) the distribution of specific conductivity (SC) niches within an ecoregion which serve as a species "filter," and (2) SC change (increase) within a given system, causing extirpation of species once the SC exceeds species' tolerance. Also, he noted, the discussion assumes that species in a given genus have comparable SC tolerances, which may skew genus level XCs in an upward, less protective direction.

A reviewer suggested that the discussion mention anthropogenic influence on specific conductivity, as well as the effects on an organism living in conditions near their outer limits.

Reviewer	Comments
Reviewer 1	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 XC95 values to support the B-C model. 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. Line 4: is it appropriate to describe field populations of species as "test organisms". 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. 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." 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. P. 1-2, line 6: suggest changing "is supported by" to "primarily the result of" P. 1-2, line 12-13: This last sentence is awkward and doesn't make sense to me. 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?
Reviewer 2	The basic ecological theory and relationship to the model is NOT presented clearly. The use of the terms, niche, range, limits, is confusing. 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. *(Note: The following comments were provided as annotations to the draft document.)* *Page 1-1, line 29, referring to "other habitat requirements":* Such as? *Page 1-1, line 31, referring to "The lower and upper ranges of SC":* Confusing. Do you mean upper and lower limits of SC? *Page 1-1, line 34, referring to "the lowest limit of the available niche (see Figure 1-1)":* Lowest limit or lower limit? Limits are the high or low boundary. More importantly, the lowest level of SC in a region is NOT necessarily the lower limit as a species might do well in pure water although such water does not occur in nature. *Page 1-1, line 35, to page 1-2, line 1, referring to "they remove part of the available SC niche and the realized niche decreases":* No, increased SC DOES not remove part of the available SC niche, but part of available habitats. By definition, the SC niche of a species is the range of SC where the species can sustain. It has nothing to do with the level of SC in any particular area. Similarly, the realized SC niche does not decrease because of increased SC, but the overall ecological niche realized or occupied habitats decreases. Re-write to clarify the confusions. *Page 1-2, line 1, referring to "niche":* replace with "level". *Page 1-2, line 5, referring to "represent sink habitats":* Any data support? Also, this might be the case if the occurrence of a genus really rapidly decline to a very low level with increasing SD. Many genera did not show this type of responses in EPA 2015 report.

	Page 1-2, line 7, referring to "niches that are defined in part": Here you obviously refer to the overall ecological niche (multi-dimensions), not SD niche only. Be specific on which type of niche you refer to throughout the text. *Page 1-2, line 12, referring to "greater":* "upper" seems more appropriate here. *Page 1-3, Figure 1-1, referring to "Approximately 5% of observations of a taxon occurs in sink habitats where a population cannot persist without immigration":* Be clear that this is an assumption, even a reasonable one, but not an observation or conclusion. Modify.
Reviewer 3	• 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. • page 1-1, line 7. "primary" Not clear what this is. SD method? • 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. • page 1-2, line 5. Why the sudden switch from species to genera? • page 1-2, line 10. Presumably XC95 as shown in Figure 1-1? Please clarify. • 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? • 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 XC95 but you can't tell that from the figure.
Reviewer 4	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. 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. 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, relative to marine, so that is the condition to which the biota is adapted." b) Further, I find that this same sentence, lines 22-24, almost a contradiction of the preceding sentence. c) I suggest inserting "or introduced" in line 27, pg. 1-1, reading: "A species may not exploit its full tolerance range because competitor or introduced species … " d) 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.. e) 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. f) 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." g) 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.
Reviewer 5	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*). There may also be a need to address 'sink habitats' in the text (P1-2, l5). 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.**

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)

Four reviewers noted that the use of the B-C regression model to estimate a criterion was scientifically defensible. One reviewer emphasized that although background SC varies considerably throughout the country, the strong relationship between background SC and HC_{05} in the model proves the process is statistically sound. This reviewer cited several examples to illustrate the document's scientific defensibility and the method's expert development. Another reviewer noted that the B-C regression model is obviously robust enough to produce an acceptable criterion continuous concentration (CCC) for a region because of the strong association of SC and HC_{05}. This reviewer did note that the statement that the B-C method can be used in smaller scales than the L III regions would need ground-truthing. One reviewer suggested that EPA present the model closer to be beginning of the document.

Two reviewers commented that it was difficult to follow the logic of the section. One noted that the terms used in this section lack specificity and the equation parameters lack parameter definitions. This reviewer pointed out multiple examples of terms, parameters, and text needing clarification. The other reviewer suggested that the document provide more robust model validation and was concerned with the potential biases introduced when the predicted SC for HC_{05} was transformed to a real sensitivity distribution (SD) level. He also suggested that EPA make the model more generally applicable by selecting ecoregions at random (currently, the ecoregions chosen are heavily clustered). This reviewer questioned the biological relevance of the 95 percent confidence interval in the context of the background-matching approach; he was concerned that this criterion might not be strict enough. He also questioned the use of bootstrapping, rather than a conventional method, to estimate the confidence limits of a parameter.

One reviewer had minor concerns about the estimation of model input values (e.g., SC XC_{95} values and minimally disturbed background for each ecoregion), but noted that trained modeling analysts should be able to produce the estimates with the same standardization of process and accuracy for all eco-regions.

Reviewer	Comments
Reviewer 1	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. Is the term "SD" ever defined in the document? P. 2-9, line 32: rare/salt intolerant taxa do not contribute to the SD? Don't these species drive the HC05 estimate? This is confusing. 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". P. 2-10, lines 10-13 seem to really weaken the argument. P. 2-10 line 12: should spell out Prediction Interval (PI). Be aware that the r reported in the instructions to reviewers is 0.87 and in the text it is reported as 0.93.
Reviewer 2	The strong B-C regression is very interesting and its use for estimate SC HC_{05} is scientific defensible. 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. 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. 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. *(Note: The following comments were provided as annotations to the draft document.)* *Page 2-4, line 25 to line 27, referring to "usually this has a minor effect because only observations of occurrences rather than presence/absence or abundance is used in the calculations":* 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. *Page 2-4, line 31 to line 33, referring to "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.":* 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 HC05. Also, how often the 95% CI overlapped among the 19 Level-3 ecoregions? *Page 2-6, line 1 to line 2, referring to "However, care should be taken": How??*

	Page 2-7, line 10 to line 12, referring to "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.": Support needs for this general statement. Also, "the full distribution" of what? *Page 2-7, line 18 to line 23, referring to "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.":* 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? *Page 2-8, line 35 to line 36, referring to "In such cases, site-specific SC criteria may be considered to protect low-SC streams.":* How? May you refer to some existing studies? *Page 2-9, line 28, referring to "SDs":* Is this defined in other part of the whole report? *Page 2-9, line 30 to line 33, referring to "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.":* 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? *Page 2-9, line 34, referring to "niche":* Not "genus"? *Page 2-9, line 35, referring to "evidence":* evident? *Page 2-10, line 1, referring to "salt-intolerant species":* 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. *Page 2-10, line 21, referring to "prediction":* Not confidence limit? *Page 2-11, 5th line in the paragraph below Figure 2-2, referring to "10% PL":* What is the justification for this particular threshold? Also, 10% PL is upper or lower limit? And Why? *Page 2-12, first line on page, referring to "be":* to be *Page 2-12, 3rd to 5th lines from top of page, referring to "because the data is from the region and it is within the 95% PL of the model and the regional data used to generate the HC_{05} is more relevant than from other ecoregions":* Why is this argument not applicable to the case above? And why the argument for the case above not applicable here? Confusing. *Page 2-12, 6th and 7th lines on page, referring to "because relatively small paired data sets are too variable to precisely calculate and HC_{05} and is likely to be over protective":* This argument is vague. How does sample size become an issue here, but not in the two cases

	above? Why is the field-based HC05 here necessarily over-protective? Any biological or statistical reason? This whole section needs to be re-written. *Page 2-12, line 19, referring to "(2-2)" at end of equation:* Where does this equation come from? If you did not develop it, give a citation. Any assumption for this equation, such as normality? If yes, is the assumption met?
Reviewer 3	• 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. • 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? • 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? • page 2-4, line 34. "ecoregional" Is this different than simply SC background? If this is yet another new term, please define. • 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. • page 2-6, line 3. Define "ecoregional criteria." • page 2-6, line 5. Define "subregional." • 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. • 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. • page 2-6, line 34. Influenced, disturbed, anthropogenically affected? Can you use consistent terminology? • page 2-7, line 8. Least versus minimally? It seems you are attaching specific, different meanings to these terms. If so, please clarify. • page 2-8, lines 19-22. This discussion is supposed to be about differences in background yet here you skip ahead to criteria development. • 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.

	• 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? • 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. • 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. • 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. • 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…" • 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. • 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. • 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. • 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. • 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. • 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.
Reviewer 4	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. 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. 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 25^{th} centile to estimate background SC because it is less susceptible to error than the 10^{th} might also be relevant. 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. 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 HC05 as the empirical data; b) the estimation of HC_{05} only for ecoregions with a background <600 uS/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.
Reviewer 5	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.

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)

Three reviewers responded that the case example demonstrated the method. One noted that the case example included excellent descriptions of the ecoregion and dataset characteristics and that the use of EPA and USGS datasets allowed for model verification. Another commented that the B-C model derivation made the process for deriving the CCC using pre-existing SC values from background information clear. Two other reviewers had positive comments, including that the case study was laid out well, and that is was the most informative section about the method.

Reviewers had several suggestions about how EPA could improve this section. One reviewer noted that the decision to use the EPA rather than the USGS estimate of the 25th centile of SC was arbitrary. This reviewer suggested that EPA derive the CCC outcome using the USGS dataset. If this was reasonably close to the CCC outcome from the EPA dataset, that would reassure readers. If not, that would raise red flags about the method. This reviewer also recommended that this section be more explicit about whether data are paired chemistry/biology samples or just chemistry.

One reviewer sought confirmation that the USGS and EPA datasets could be used together in the statistical package, because they were collected over different time periods and frequencies. This reviewer also suggested that EPA explain differences in the method for calculating centiles in the two datasets and the reason for the difference in SC distributions for the two datasets.

A reviewer suggested that EPA justify the use of a 10 percent prediction limit (PL) and the adoption of the field-based or predicted HC_{05} or lower limits of 10 percent PL. Another reviewer suggested pulling out all sites with a SD greater than 0.25 to determine what drove the variation of SD values in this set of sites. A third reviewer suggested that the section could be written with more clarity, including providing more specific definitions of parameters.

Reviewer	Comments
Reviewer 1	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 25th 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.

	P. 3-15, line 7: more details of how ionic composition was evaluated would be useful. P. 3-16, line 3: spell out CMEC. 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. 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. P. 3-18, line 21-22: this would indicate that the background for this region is towards the extreme end of the model. >550 uS/cm could mean anything greater than this value, should an estimate be presented here without the > sign? Table 3-2: the last row of data does not make intuitive sense to me. What are the units? 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 SO4 and Mg concentration between the 2 datasets. p. 3-22 line 9-10: is this t the justification for using the EPA estimate of the 25th centile rather than the USGS estimate?
Reviewer 2	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. I also made several minor comments in the documents. *(Note: The following comments were provided as annotations to the draft document.)* *Page 3-22, Figure 3-4, referring to figure caption:* Does the box represent 25th and 75th percentiles? Clarify. *Page 3-23, line 23, referring to "102.870925 µS/cm = 742.890 µS/cm":* 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).
Reviewer 3	• 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. • 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. • 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. • page 3-16, lines 6-9. This seems redundant with discussion in Section 3.1.3.1. • page 3-22, lines 9-10. This statement is both unclear and possibly circular. Please re-write this whole paragraph for clarity. • 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. • page 3-23, line 17. Using "may be" suggests that this step is optional, which it does not appear to be? • page 3-23, line 29. Isn't 743 the actual value, not the log of the value? Please clarify. • 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. • 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. • 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. • page 3-27, line 2. Average or geomean? Please clarify. • 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.
Reviewer 4	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. I have the following comments and/or questions on Section 3: a) 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? b) 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 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. c) 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. d) 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.
Reviewer 5	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?). I like the B-C model derivation (P3-23) – 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. 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?). Figure 3-6 (P3-27) 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!

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)

Four reviewers responded that the presentation of the derivation of the B-C regression model and prediction of an HC_{05} from minimally-disturbed background was reasonably clear. One of these reviewers noted that the work described in Section 4 had been done carefully, logically, and thoroughly, and was impressed with many aspects of this section—for example, that the data requirements EPA used for estimating XC_{95} reflected carefully selected guidelines, that EPA quality assured the state datasets, that guidelines were established to minimize the chance of error, and that excellent work had been done comparing the results of the three methods of background SC estimation. The fifth reviewer noted that the model appeared adequate despite being based on only 19 data points, but that the text lacked clarity and that it was hard to follow EPA's rationale for some of the assumptions inherent in the model.

Reviewers had various suggestions for improvement. A reviewer noted that the B-C regression approach will be robust and highly useful because XC_{95} can be confidently established for the top 5 percent most sensitive genera. This reviewer suggested using a symbol to indicate the confidence level of the XC_{95} estimation shown in the Figure 4-3 sensitivity distribution plots to avoid giving the misleading impression that XC_{95} estimated from field data are the same among genera when in fact they differ greatly in precision and accuracy among genera. This reviewer also noted that the ecoregions are not randomly distributed, but rather highly clustered, and wondered whether this issue could affect model performance. The reviewer asked whether EPA will refine the model when HC_{05} values become available for more ecoregions, and suggested including some form of validation for the models, such as Jackknife validation.

Two reviewers commented on the large spread of SC values presented in Table 4-1. One noted that the large difference in values indicates a high degree of uncertainty associated with the 25^{th} centile estimate. The other was concerned with the spread in the SC in some of the ecoregions in Table 4-1 and suggested that EPA determine what drove the large spread in these estimates. A reviewer suggested that the abrupt change in base-flow SC at state lines may result from differences in data collection in different states, an explanation that was not considered in this section.

Reviewer	Comments
Reviewer 1	In general, the presentation is reasonably clear. 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 HC05s shift as a

	function of background conductivities (Fig. 4-4). 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 25th centile estimates? Can a 25th 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 25th centile estimate (x axis of the entire model). Language in the legend of table 4-4 is very difficult to interpret.
Reviewer 2	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. 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. *(Note: The following comments were provided as annotations to the draft document.)* *Page 4-30, line 15, referring to "for 19 ecoregions":* 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? *Page 4-30, line 28, referring to "Taxonomic identification to genus":* Sampling effort? 300- or 500 counts? *Page 4-34, line 3, referring to "Ecoregional data sets":* What is this dataset? From USGS gauge stations? *Page 4-39, line 1 to line 10, referring to "The regression model using the estimated Background SCs from the 19 ecoregional data sets yielded the strongest model (r = 0.93) compared to the smaller probability survey (r = 0.73) and the base-flow model (r = 0.62). However, removal of the estimate of Ecoregion 19, improves the model for the probability survey data set (r = 0.88). The model with the strongest r-value, the B-C model using the 25th centile background SC and HC_{05} values for the 19 ecoregions is shown in Figure 4-4. Each point on the graph represents the relationship between Background SC for a level III ecoregion and the HC_{05} for that ecoregion. The 10% PL from the mean regression line was calculated in order to identify when there is a 90% probability that an HC_{05} derived by the field-SD method for a new ecoregion would be equal to or greater than the y-coordinate of the lower prediction limit.":* Some form of validation is needed for the models. R^2 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).
Reviewer 3	• General: Technically, the model appears adequate despite being based on only 19 data points. 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. • page 4-30, line 4. Which method is this? The background method? Please clarify. • 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 1000. Please clarify this seeming discrepancy. • page 4-31, line 14. This appears to be your definition of "minimally." If so, it needs to be provided earlier in this document. • 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. • page 4-32, table. Add "ecoregional" to column 3 for clarity. • page 4-34, line 10. "Patterns" are very hard to see in Figure 4-1. How did you assess the similarity of patterns here? • 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? • page 4-31, Figure 4-1 caption. "Similar trends" are not readily apparent from these figures. • page 4-36, line 6. Why? Was this discussed elsewhere? If so, please provide a reference. • page 4-36, line 12. This figure says nothing about HC values. • page 4-37, Figure 4-3 caption. Where is HC in this figure? • page 4-38, Table 4-3, 6th column. "Empirical" is confusing. This appears to be the 25th centile of just the ecoregional data set. Please clarify. Where did the other two (probability, base-flow) figure in here? • 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. • 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. • 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." • 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?
Reviewer 4	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. 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 25th centile was used to estimate the background SC: 1) the 25th centile is less susceptible to error than the 10th centile, and 2) the 25th centile yielded a slightly stronger statistical model with HC_{05} than using the 10th 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.

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)"?

b) 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.

c) 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.

Reviewer 5	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). 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. 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. 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.

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.)

One reviewer commented that the sections did not flow in a logical sequence and suggested introducing and discussing one method, then another method, and ending with the case study.

A reviewer reiterated his concern that including the EPA and USGS datasets in the case study fosters uncertainty that accurate 25th centile estimates are driving the model. He suggested including a quantitative assessment to analyze how CCC estimates would vary if the 25th centiles were over- or under-estimated by various amounts. The reviewer suggested that the document include a description of how the Level III ecoregions are delineated.

Three reviewers referred to suggestions they provided in response to the previous charge questions. One reviewer recommended that EPA include an explanation of the Nebraska Sand Hills are unique and significantly different from the rest of the general region.

Reviewer	Comments
Reviewer 1	• 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. The regression (Fig. 4-4) is impressive and the application is pretty straightforward. • 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). • 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.
Reviewer 2	See my suggestions above.
Reviewer 3	• 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. • 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.
Reviewer 4	I recommend for you to review my responses to Questions 1 – 4 as I listed suggestions for specific items 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. 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.
Reviewer 5	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.

2.6 Other Comments Provided

The preceding sections summarized and provided comments reviewers made in response to the charge questions. This section presents additional comments made by some reviewers.

Reviewer	Comments
Reviewer 1	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.
Reviewer 3	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.
Reviewer 4	a) 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. b) 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. 2. 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. c) The same problem lies with the map in Figure 3-2, pg. 3-19. The red symbols should be "triangles" rather than "diamond". 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.
Reviewer 5	P1-1, 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**. P1-3 – In figure caption, rephrase sentence 3 – it is clumsy and not well constructed to

	convey meaning. P1-3 – Period missing at end of figure caption. P3-16, L6-15 - 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. P3-20 – P3-21 – 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)? **You cannot average pH values**. P4-32, Table 4-1 – In the probability survey 25th centile, watch significant figures – round! **So, check all significant figures in all tables**.

3.0 NEW INFORMATION PROVIDED BY REVIEWERS

This section compiles all new information that some reviewers provided in their individual comments to the charge questions in Section 2.0.

Reviewer	Charge Question	New Information
Reviewer 2	**2.3**	Duan N. 1983. Smearing estimate: a nonparametric retransformation method. *Journal of the American Statistical Association*, **78**, 605-610).
Reviewer 2	**2.4**	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

APPENDIX A

CHARGE TO REVIEWERS

Technical Charge to External Peer Reviewers

Contract No. EP-C-12-029
Task Order 51
June 2015

External Letter Peer Review of EPA's Draft Document, "Application to New Areas"

BACKGROUND

The draft document, "Application to New Areas," describes a method to estimate a hazardous concentration (HC_{05}) based on a model of natural background specific conductivity (SC) and calculated HC_{05} values from 19 ecoregions. It is based on concepts described in EPA's 2011 report, *A Field-Based Aquatic Life Benchmark for Conductivity in Central Appalachian Streams* (EPA Conductivity Benchmark Report) and the recently externally peer-reviewed draft EPA method, *Recommended Field-based Method for States to Develop Ambient Aquatic Life Water Quality Criteria for Conductivity* (*Recommended Field-based Method*)(draft, not yet published).

In the previously reviewed draft *Recommended Field-based Method,* the SC HC_{05} was shown to be applicable throughout an ecoregion with a similar background SC. This method is referred to as the background matching approach and was favorably reviewed. However, there remained uncertainty whether this method could apply more generally and there was no model for predicting an HC_{05} for a range of possible background SC conditions in areas without large data sets.

Because a genus will rarely occur where the natural regional background exceeds the upper tolerance limit for that genus, the lower possible tolerance limit of genera in a region is defined by the natural background. Theoretically, there will be a regular relationship between natural background SC and the SC that affects 5% of genera. To test this, SC HC_{05} values were estimated and modeled with the background SC for 19 level III Ecoregions in the United States. Background SC was estimated at the 25th centile of all sampled sites in each ecoregion. A least squares regression of HC_{05} values against background yielded a strong statistical relationship (r = 0.87). The relationship between the salt-intolerant genera and background SC demonstrates that salt-intolerant genera occupy the lowest available SC niche in an ecoregion. Also, the new model allows prediction of an HC_{05} for ecoregions with background between 20 and 600 µS/cm. EPA suggests that the regression model makes it possible to use SC background in an ecoregion to predict the level that is expected to extirpate organisms adapted to low SC in an ecoregion, i.e., to predict the HC_{05} from background.

In the recently reviewed *draft Recommended Field-based Method,* two case studies were presented that have abundant data from which to derive SC criteria using the field-based method and to determine applicability to biologically unsampled areas using the background-matching approach. Such data sets are not available for all ecoregions in the U.S. and a method to identify a reasonable SC criterion without access to a sufficiently large ecoregion-specific data set would be useful. This draft document describes the relationship between specific natural background SC and salt-intolerant genera and an example application for the Northwest Great Plains. The estimation of an HC_{05} from background is presented in the draft document for review as a recommended approach for developing water quality criteria for those ecoregions lacking sufficient data to develop criteria from a regional data set.

PEER REVIEW MATERIALS

- **Draft Peer Review Document:**
  - Conductivity Methodology - Application to New Areas-PeerReviewDrft20150622.docx

  This draft document is anticipated to be incorporated into the draft: *Recommended Field-based Method for States to Develop Ambient Aquatic Life Water Quality Criteria for Conductivity* (draft, not yet published). Much of the present document, "*Application to New Areas*," is comprised of content that has already been externally peer reviewed and is provided within the peer review draft as background material that the peer reviewers may wish to consider for context.

  For this reason, each question in this charge is meant to focus the reviewer on the new components that are in need of review (the section number and corresponding page numbers are provided in charge question below).

- **Background Materials:**
  - ANALYSIS PLAN Chapter 3 excerpt.pdf
  - Appendix C 20150511.pdf
  - Appendix D 20150508 Finalsmc.pdf
  - CASE STUDY II.pdf

  The above appendices and sections from the recently peer-reviewed draft method are cited in the draft peer review document and therefore are being provided as background materials for reviewers to reference as needed. All background materials are meant for clarity and context and it is not necessary to provide comments on them.

CHARGE QUESTIONS

The section number and corresponding page numbers within each question below are the new components that are in need of review.

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. 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; Note: Sections 2.1 and 2.2.6 are for information only and have already been reviewed).

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, Note: 3.16 is for information only and has already been reviewed, and Sections 3.2 to 3.4 follow previously reviewed methods)

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

Note: pages 4-41 to 4-46 are comprised of tables that are calculated from the model and previously published background values and are provided for completeness).

5. Please provide any suggestions on ways to strengthen either the development of the B-C model or the example case study? (Section 4)

APPENDIX B

INDIVIDUAL REVIEWER COMMENTS

COMMENTS SUBMITTED BY

Reviewer 1

Peer Review of EPA's Draft Document, "Application to New Areas"

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.)**

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 XC95 values to support the B-C model.

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.

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

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.

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."

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.

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

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

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?

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; Note: Sections 2.1 and 2.2.6 are for information only and have already been reviewed.)

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.

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

P. 2-9, line 32: rare/salt intolerant taxa do not contribute to the SD? Don't these species drive the HC05 estimate? This is confusing.

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".

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

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

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

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, Note: 3.16 is for information only and has already been reviewed, and Sections 3.2 to 3.4 follow previously reviewed methods.)

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 25th 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.

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

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

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.

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.

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

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

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 SO4 and Mg concentration between the 2 datasets.

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

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 Note: pages 4-41 to 4-46 are comprised of tables that are calculated from the model and previously published background values and are provided for completeness.)

In general, the presentation is reasonably clear.

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 HC05s shift as a function of background conductivities (Fig. 4-4).

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 25th centile estimates? Can a 25th 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 25th centile estimate (x axis of the entire model).

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

5. Please provide any suggestions on ways to strengthen either the development of the B-C model or the example case study? (Section 4.)

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

General comment:

- Our approach to protecting humans involves both epidemiological and mechanistic approaches to understanding health outcomes. When mechanisms and observations support eachother, 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.

COMMENTS SUBMITTED BY

Reviewer 2

Peer Review of EPA's Draft "Application to New Areas"

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.)

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

(Note: The following comments were provided as annotations to the draft document.)

Page 1-1, line 29, referring to "other habitat requirements": Such as?

Page 1-1, line 31, referring to "The lower and upper ranges of SC": Confusing. Do you mean upper and lower limits of SC?

Page 1-1, line 34, referring to "the lowest limit of the available niche (see Figure 1-1)": Lowest limit or lower limit? Limits are the high or low boundary. More importantly, the lowest level of SC in a region is NOT necessarily the lower limit as a species might do well in pure water although such water does not occur in nature.

Page 1-1, line 35, to page 1-2, line 1, referring to "they remove part of the available SC niche and the realized niche decreases": No, increased SC DOES not remove part of the available SC niche, but part of available habitats. By definition, the SC niche of a species is the range of SC where the species can sustain. It has nothing to do with the level of SC in any particular area. Similarly, the realized SC niche does not decrease because of increased SC, but the overall ecological niche realized or occupied habitats decreases. Re-write to clarify the confusions.

Page 1-2, line 1, referring to "niche": replace with "level".

Page 1-2, line 5, referring to "represent sink habitats": Any data support? Also, this might be the case if the occurrence of a genus really rapidly decline to a very low level with increasing SD. Many genera did not show this type of responses in EPA 2015 report.

Page 1-2, line 7, referring to "niches that are defined in part": Here you obviously refer to the overall ecological niche (multi-dimensions), not SD niche only. Be specific on which type of niche you refer to throughout the text.

Page 1-2, line 12, referring to "greater": "upper" seems more appropriate here.

Page 1-3, Figure 1-1, referring to "Approximately 5% of observations of a taxon occurs in sink habitats where a population cannot persist without immigration": Be clear that this is an assumption, even a reasonable one, but not an observation or conclusion. Modify.

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; Note: Sections 2.1 and 2.2.6 are for information only and have already been reviewed.)

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

(Note: The following comments were provided as annotations to the draft document.)

Page 2-4, line 25 to line 27, referring to "usually this has a minor effect because only observations of occurrences rather than presence/absence or abundance is used in the calculations": 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.

Page 2-4, line 31 to line 33, referring to "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.": 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 HC05. Also, how often the 95% CI overlapped among the 19 Level-3 ecoregions?

Page 2-6, line 1 to line 2, referring to "However, care should be taken": How??

Page 2-7, line 10 to line 12, referring to "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.": Support needs for this general statement. Also, "the full distribution" of what?

Page 2-7, line 18 to line 23, referring to "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.": 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?

Page 2-8, line 35 to line 36, referring to "In such cases, site-specific SC criteria may be considered to protect low-SC streams.": How? May you refer to some existing studies?

Page 2-9, line 28, referring to "SDs": Is this defined in other part of the whole report?

Page 2-9, line 30 to line 33, referring to "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.": 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?

Page 2-9, line 34, referring to "niche": Not "genus"?

Page 2-9, line 35, referring to "evidence": evident?

Page 2-10, line 1, referring to "salt-intolerant species": 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.

Page 2-10, line 21, referring to "prediction": Not confidence limit?

Page 2-11, 5th line in the paragraph below Figure 2-2, referring to "10% PL": What is the justification for this particular threshold? Also, 10% PL is upper or lower limit? And Why?

Page 2-12, first line on page, referring to "be": to be

Page 2-12, 3rd to 5th lines from top of page, referring to "because the data is from the region and it is within the 95% PL of the model and the regional data used to generate the HC_{05} is more relevant than from other ecoregions": Why is this argument not applicable to the case above? And why the argument for the case above not applicable here? Confusing.

Page 2-12, 6th and 7th lines on page, referring to "because relatively small paired data sets are too variable to precisely calculate and HC_{05} and is likely to be over protective": This argument is vague. How does sample size become an issue here, but not in the two cases above? Why is the field-based HC05 here necessarily over-protective? Any biological or statistical reason? This whole section needs to be re-written.

Page 2-12, line 19, referring to "(2-2)" at end of equation: Where does this equation come from? If you did not develop it, give a citation. Any assumption for this equation, such as normality? If yes, is the assumption met?

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, Note: 3.16 is for information only and has already been reviewed, and Sections 3.2 to 3.4 follow previously reviewed methods.)

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. I also made several minor comments in the documents.

(Note: The following comments were provided as annotations to the draft document.)

Page 3-22, Figure 3-4, referring to figure caption: Does the box represent 25th and 75th percentiles? Clarify.

Page 3-23, line 23, referring to "102.870925 µS/cm = 742.890 µS/cm": 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).

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 Note: pages 4-41 to 4-46 are comprised of tables that are calculated from the model and previously published background values and are provided for completeness.)

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. 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.

(Note: The following comments were provided as annotations to the draft document.)

Page 4-30, line 15, referring to "for 19 ecoregions": 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?

Page 4-30, line 28, referring to "Taxonomic identification to genus": Sampling effort? 300- or 500 counts?

Page 4-34, line 3, referring to "Ecoregional data sets": What is this dataset? From USGS gauge stations?

Page 4-39, line 1 to line 10, referring to "The regression model using the estimated Background SCs from the 19 ecoregional data sets yielded the strongest model (r = 0.93) compared to the smaller probability survey (r = 0.73) and the base-flow model (r = 0.62). However, removal of the estimate of Ecoregion 19, improves the model for the probability survey data set (r = 0.88). The model with the strongest r-value, the B-C model using the 25th centile background SC and HC_{05} values for the 19 ecoregions is shown in Figure 4-4. Each point on the graph represents the relationship between Background SC for a level III ecoregion and the HC_{05} for that ecoregion. The 10% PL from the mean regression line was calculated in order to identify when there is a 90% probability that an HC_{05} derived by the field-SD method for a new ecoregion would be equal to or greater than the y-coordinate of the lower prediction limit.": Some form of validation is needed for the models. R^2 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).

5. Please provide any suggestions on ways to strengthen either the development of the B-C model or the example case study? (Section 4.)

See my suggestions above.

COMMENTS SUBMITTED BY

Reviewer 3

Peer Review of EPA's Draft "Application to New Areas"

OVERVIEW

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.

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.)**

   - 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.
   - page 1-1, line 7. "primary" Not clear what this is. SD method?
   - 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.
   - page 1-2, line 5. Why the sudden switch from species to genera?
   - page 1-2, line 10. Presumably XC95 as shown in Figure 1-1? Please clarify.
   - 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?
   - 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 XC95 but you can't tell that from the figure.

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; Note: Sections 2.1 and 2.2.6 are for information only and have already been reviewed.)**

- 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.
- 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?
- 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?
- page 2-4, line 34. "ecoregional" Is this different than simply SC background? If this is yet another new term, please define.
- 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.
- page 2-6, line 3. Define "ecoregional criteria."
- page 2-6, line 5. Define "subregional."
- 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.
- 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.
- page 2-6, line 34. Influenced, disturbed, anthropogenically affected? Can you use consistent terminology?
- page 2-7, line 8. Least versus minimally? It seems you are attaching specific, different meanings to these terms. If so, please clarify.
- page 2-8, lines 19-22. This discussion is supposed to be about differences in background yet here you skip ahead to criteria development.
- 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.
- 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?
- 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.

- 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.
- 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.
- 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…"
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.

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, Note: 3.16 is for information only and has already been reviewed, and Sections 3.2 to 3.4 follow previously reviewed methods.)

- 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.
- 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.
- 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.

- page 3-16, lines 6-9. This seems redundant with discussion in Section 3.1.3.1.
- page 3-22, lines 9-10. This statement is both unclear and possibly circular. Please re-write this whole paragraph for clarity.
- 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.
- page 3-23, line 17. Using "may be" suggests that this step is optional, which it does not appear to be?
- page 3-23, line 29. Isn't 743 the actual value, not the log of the value? Please clarify.
- 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.
- 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.
- 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.
- page 3-27, line 2. Average or geomean? Please clarify.
- 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.

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 Note: pages 4-41 to 4-46 are comprised of tables that are calculated from the model and previously published background values and are provided for completeness.)

- General: Technically, the model appears adequate despite being based on only 19 data points. 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 25^{th} centile). In addition, apparently shifting terminology and inconsistent labeling of table columns detracts from its overall clarity.
- page 4-30, line 4. Which method is this? The background method? Please clarify.
- 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 1000. Please clarify this seeming discrepancy.
- page 4-31, line 14. This appears to be your definition of "minimally." If so, it needs to be provided earlier in this document.
- 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.
- page 4-32, table. Add "ecoregional" to column 3 for clarity.
- page 4-34, line 10. "Patterns" are very hard to see in Figure 4-1. How did you assess the similarity of patterns here?

- 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?
- page 4-31, Figure 4-1 caption. "Similar trends" are not readily apparent from these figures.
- page 4-36, line 6. Why? Was this discussed elsewhere? If so, please provide a reference.
- page 4-36, line 12. This figure says nothing about HC values.
- page 4-37, Figure 4-3 caption. Where is HC in this figure?
- page 4-38, Table 4-3, 6th column. "Empirical" is confusing. This appears to be the 25th centile of just the ecoregional data set. Please clarify. Where did the other two (probability, base-flow) figure in here?
- 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.
- 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.
- 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."
- 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?

5. Please provide any suggestions on ways to strengthen either the development of the B-C model or the example case study? (Section 4.)

- 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.
- 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.

COMMENTS SUBMITTED BY

Reviewer 4

Peer Review of EPA's Draft "Application to New Areas"

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.)**

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.

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.

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, relative to marine, so that is the condition to which the biota is adapted."

b) Further, I find that this same sentence, lines 22-24, almost a contradiction of the preceding sentence.

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

d) 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..

e) 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.

f) 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."

g) 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.

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; Note: Sections 2.1 and 2.2.6 are for information only and have already been reviewed.)

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.

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.

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 25^{th} centile to estimate background SC because it is less susceptible to error than the 10^{th} might also be relevant.

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.

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 HC05 as the empirical data;

b) the estimation of HC_{05} only for ecoregions with a background <600 uS/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.

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, Note: 3.16 is for information only and has already been reviewed, and Sections 3.2 to 3.4 follow previously reviewed methods.)

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.

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

a) 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?

b) 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

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.

c) 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.

d) 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.

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 Note: pages 4-41 to 4-46 are comprised of tables that are calculated from the model and previously published background values and are provided for completeness.)

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.

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 25th centile was used to estimate the background SC: 1) the 25th centile is less susceptible to error than the 10th centile, and 2) the 25th centile yielded a slightly stronger statistical model with HC_{05} than using the 10th 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.

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)"?

b) 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.

c) 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.

5. Please provide any suggestions on ways to strengthen either the development of the B-C model or the example case study? (Section 4.)

I recommend for you to review my responses to Questions 1 – 4 as I listed suggestions for specific items 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.

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.

Editing / Typos

a) 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.

b) 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.

   2. 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.

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

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.

COMMENTS SUBMITTED BY

Reviewer 5

Peer Review of EPA's Draft "Application to New Areas"

1. **The rationale for developing the Background-to-Criterion (B-C) model is based on the availability of specific conductivity (SC) 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.)**

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*).

There may also be a need to address 'sink habitats' in the text (P1-2, l5). 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.

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; Note: Sections 2.1 and 2.2.6 are for information only and have already been reviewed.)**

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.

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, Note: 3.16 is for information only and has already been reviewed, and Sections 3.2 to 3.4 follow previously reviewed methods.)

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?).

I like the B-C model derivation (P3-23) – 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.

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?).

Figure 3-6 (P3-27) 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!

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 Note: pages 4-41 to 4-46 are comprised of tables that are calculated from the model and previously published background values and are provided for completeness.)

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). 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.

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.

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.

5. Please provide any suggestions on ways to strengthen either the development of the B-C model or the example case study? (Section 4.)

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.

Other Comments:

P1-1, 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**.

P1-3 – In figure caption, rephrase sentence 3 – it is clumsy and not well constructed to convey meaning.

P1-3 – Period missing at end of figure caption.

P3-16, L6-15 - 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.

P3-20 – P3-21 – 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)? **You cannot average pH values**.

P4-32, Table 4-1 – In the probability survey 25th centile, watch significant figures – round! **So, check all significant figures in all tables**.