EPA RESPONSE
TO EXTERNAL PEER REVIEW COMMENTS

on the

DRAFT AQUATIC LIFE AMBIENT WATER QUALITY CRITERIA FOR ALUMINUM - 2017

July 2017

Office of Water
U.S. Environmental Protection Agency
Washington, DC

TABLE OF CONTENTS

1 INTRODUCTION

EPA submitted its draft *Aquatic Life Ambient Water Quality Criteria for Aluminum - 2017*, for contractor-led independent, external peer review from February 24th to March 13th, 2017. The external peer reviewers provided their independent responses to EPA's charge questions. This report documents the EPA's response to the comments provided to EPA.

This report presents the five peer review charge questions and individual reviewer comments (verbatim) in Sections 2.1 through 2.4. Section 2.5 presents additional comments provided by one reviewer. New information (e.g., references) provided by reviewers is presented in Section 3. EPA separated each reviewer's comments by charge question into distinct topics and responded to each topic individually, and also indicated how the draft aluminum criteria approaches may be revised in response to peer reviewer comments.

1.1 BACKGROUND

EPA establishes national recommended Ambient Water Quality Criteria (AWQC) under the Clean Water Act (CWA). Aquatic life criteria address the CWA goals of providing for protection and propagation of fish and shellfish. Section 304(a)(1) aquatic life criteria serve as recommendations to states and tribes by defining ambient water concentrations that will protect against unacceptable adverse ecological effects to aquatic life from exposure to pollutants in water. Once EPA publishes final §304(a) recommended water quality criteria, states and authorized tribes may adopt these criteria into their water quality standards to protect designated uses of water bodies. As required by the CWA, EPA periodically reviews and revises §304(a) AWQC to ensure they are consistent with the latest scientific information. In support of this mission, EPA is working to update water quality criteria to protect aquatic life from aluminum in freshwater environments.

The current aluminum criteria are from 1988 and applied to pH range 6.5 – 9.0 and across all hardness and dissolved organic carbon (DOC) ranges. The approach described in the 2017 draft aluminum AWQC document is a multi-linear regression (MLR) model using pH, hardness, and DOC, with interaction terms between these parameters for freshwater aluminum criteria.

1.2 PEER REVIEWERS

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

The EPA contractor provided reviewers with instructions, the draft 2017 aluminum criteria document, the aluminum criteria model calculator (an Excel spreadsheet), and the charge to reviewers prepared by EPA. Reviewers worked individually to develop written comments in response to the charge questions.

1.3 REVIEW MATERIALS PROVIDED

- Draft Aluminum AWQC (Draft National Recommended Ambient Water Quality Criteria for Aluminum - 2.24.17.docx)
- Aluminum criteria model calculator (Aluminum MLR Criteria Calculator_2.9.17.xlsx)

1.4 CHARGE QUESTIONS

1. Please comment on the overall clarity of the document and construction as it relates to the derivation of each criterion.

2. Please comment on the technical approach used to derive the draft aluminum criteria; is it logical, does the science support the conclusion, and is it consistent with the protection of freshwater life from acute, chronic, and bioaccumulative effects? Please provide specific comments.

3. Please comment on the data used to derive the revised criteria, including data adequacy/comprehensiveness, and the appropriateness of the data selected and/or excluded from the derivation of the draft criteria. Are there other relevant data that you are aware of that should be included? If so, please provide the reference.

4. Are the derived criteria appropriately protective of commercially and recreationally important species, and of ecosystems overall?

2 EXTERNAL PEER REVIEWER COMMENTS AND EPA RESPONSES, ORGANIZED BY CHARGE QUESTION

The following tables list the charge questions submitted to the external peer reviewers, the external peer reviewers' comments regarding those questions (broken into distinct topics), and EPA's responses to the peer reviewers' comments. EPA revised the 2017 draft considering the external peer review comments, and noted in the table how/where the document was edited.

2.1 CHARGE QUESTION 1

1. Please comment on the overall clarity of the document and construction as it relates to the derivation of each criterion.

Reviewer	Comments	EPA Response	Revision Location in 2017 Draft Aluminum AWQC
Reviewer 1	Overall, the document is well constructed and reasonably clear in its presentation.	Thank you for your comment.	No edits needed.
Reviewer 2	I would first like to commend the authors for producing this draft report on the Water Quality Criteria (WQC) for aluminum. It is a complex topic and the authors incorporate very recent scientific approaches which improves the reports technical basis and its application to states and other users of the criteria. The Environmental Protection Agency and the Office of Water, in particular, have a successful history of improving the quality and utilization of our nation's waters and while others may forget turbid, oily, lakes and streams with detergent bubbles floating on them, I do not, and I thank you for continuing your important role in the protection of people and the environment in which they live.	Thank you for your comments.	No edits needed.
Reviewer 2	The purpose of the document, in the Executive Summary, is listed as improving scientific knowledge. I would suggest adding – and to increase the utilization by states, tribes, and other users. Developing the criteria using the multiple linear regression (MLR) approach will use water chemistry parameters that the vast majority of the states currently measure in their water monitoring programs. The relevance of these draft criteria will be very high because the states and tribes can easily apply them in their water quality programs.	Thank you for your comments and suggestion. Text was added to the document. EPA decided to use an empirical MLR approach in this draft aluminum criteria update due to the relative simplicity and transparency of the model, and the decreased number of input data on water chemistry needed to derive criteria at different sites.	Section 2.3.1 (formerly, Section 2.4.1)
Reviewer 2	In the Executive Summary, I would suggest a paragraph on the history of utilization of the previous aluminum WQC (or others) by the states to show how useful the documents are at protecting our Nation's water.	Thank you for your suggestion. However the text was not edited.	No edits.
Reviewer 2	Overall the clarity of the writing is very good. The document organization follows the format of the ecological risk assessment framework and that process is familiar to the readers of the document.	Thank you for your comments.	No edits needed.

Reviewer	Comments	EPA Response	Revision Location in 2017 Draft Aluminum AWQC
Reviewer 2	There is not enough data to conduct WQC for estuaries and marine environments, however discussion sections continue through the document. I would suggest stating that there is not enough data and then combine the further text and data on estuary/marine systems into an appendix.	Thank you for your suggestion, but to be consistent with other AWQC for different pollutants, the estuarine/marine sections were retained throughout the document.	No edits.
Reviewer 2	I think the reader needs to brought along with the progression of the use of the biotic ligand model (BLM) in the copper WQC to the current selection of the MLR approach. An expansion of the text on page 13 would benefit the report. A discussion on the development and application of the full BLM, which could include text on the collaborative effort to produce the BLM approach and its utilization of aquatic chemistry data. Implementation issues concerned the amount of water quality data required to run the full BLM has been a significant for the states and tribes. Then text on potential solutions to the implementation issues could include a discussion of analysis of the partial BLM approach and analysis using the MLR. This text would help the reader understand the amount of effort, in coordination with the scientific community, that went into the development of a more relevant approach.	Thank you for your suggestion. Additional text was drafted to explain the difference between the biotic ligand model (BLM) and MLR approaches.	Section 5.3.5 (new section)
Reviewer 3	Overall the draft criteria document is well organized and clearly written. It contains a satisfying level of breadth and depth in analyzing the results of a comprehensive literature review and successfully performing a critical analysis of a large peer-reviewed data set concerning aluminum risk to aquatic ecosystems. The MLR model approach and the thoroughness of the input data selection/rejection are strengths of this criteria development document.	Thank you for your comment.	No edits needed.
Reviewer 3	**Specific Document Section Comments** **Executive Summary** The Executive Summary has sufficient detail to provide readers with the regulatory premise, background, critical concepts, derivation and modeling approaches and the criteria conclusions presented in the document. I would suggest "getting to the point" in the opening paragraph of the Executive Summary with the following edit: "This update establishes a freshwater criteria magnitude that is affected by total hardness, pH, and DOC and expands on the toxicity database to include those studies beyond the pH range of 6.5-9.0. The criteria results are presented in a series of look-up tables in Appendix K."	Thank you for your suggestion. Additional text was added for clarity.	Executive Summary

Reviewer	Comments	EPA Response	Revision Location in 2017 Draft Aluminum AWQC
Reviewer 3	**Introduction and Background** This is a well-developed and informative section. The regulatory premise and application of criteria are discussed in clear exposition. The active web link to National Recommended Water Quality Criteria - Aquatic Life Criteria Table will be helpful in establishing context to some readers, especially non-experts, who may access this document with limited background knowledge or as their first introduction to water quality criteria.	Thank you for your comments.	No edits needed.
Reviewer 3	**2 Problem Formulation** This section establishes cause and framework for the document in a concise manner.	Thank you for your comment.	No edits needed.
Reviewer 3	**2.1 Overview of Stressor Sources and Occurrence** This section is a well-cited review. The analysis is complete and representative. The supporting information and in context explanations (i.e. dissolved Al samples) are helpful without being pedantic.	Thank you for your comments.	No edits needed.
Reviewer 3	**2.2 Stressors of Concern** The treatment of 2009 total recoverable metal differences of aquatic Al concentrations and the potential for "conservative" risk assessment is balanced and offers a fair treatment of potential for bias from a change in sampling and analysis methodology. The potential for bias is managed in the selection of published controlled exposure studies cited in the draft document that uses the same dose construction, sampling, and analysis approaches.	Thank you for your comments.	No edits needed.
Reviewer 3	**2.3 Environmental Fate and Transport** Aluminum environmental chemistry is complex, and Figure 2 is very helpful in demonstrating this complexity to the reader.	Thank you for your comment.	No edits needed.
Reviewer 3	**2.4 Mode of Action and Toxicity** This section is a well-developed analysis and introduces the reader to the aquatic chemistry challenges of Al, as it relates to exposure studies. It establishes the context for study criteria to be included in model inputs. The section is comprehensive in exploring the practical challenges of performing controlled Al dosing studies. The discussion of pH effects and limits for study inclusion is a strength.	Thank you for your comments.	No edits needed.

Reviewer	Comments	EPA Response	Revision Location in 2017 Draft Aluminum AWQC
Reviewer 3	**2.4.1 Water Quality Parameters Affecting Toxicity** This section introduces BLM and MLR modeling approaches with sufficient depth and breadth and is well cited. The MLR approach is well documented and clearly discussed in this section. The in-review manuscripts of Cardwell 2017, Gensemer 2017, Brix 2017 and DeForest 2017 are a critical foundation needed to support the draft document. Substantive changes to the final peer-reviewed papers will need to be reflected in changes to this draft criteria document. Because of the entwined fate of those papers and this draft document, acceptance for publication with minor changes is a hopeful outcome of that review process. My read of the DeForest 2017 manuscript suggests it is a strong work.	EPA anticipates that the submitted manuscripts will be available to the public when the draft AWQC is released for public review. Should these manuscripts receive significant changes that would affect the criteria, the AWQC document will be updated to reflect such changes.	No edits.
Reviewer 3	**2.5 Measurement Endpoints** This section establishes a context for the approach. The outline of process and procedures is clearly developed and clearly written with a sufficient and satisfying level of explanation for a wide range of readers.	Thank you for your comments.	No edits needed.
Reviewer 3	**2.5.1 Overview of Toxicity Data Requirements** The use of bullets in this section is effective to outline an overview of data requirements.	Thank you for your comment.	No edits needed.
Reviewer 3	**2.6 Measures of Effect** The section does a solid job in explaining the data and study research search methodology that is used to determine test acceptability the development of the criteria.	Thank you for your comment.	No edits needed.
Reviewer 3	**2.6.1 Acute measures of effect** The section solidly introduces the reader to fundamental terms and resources found in the draft document and appendices.	Thank you for your comment.	No edits needed.
Reviewer 3	**2.6.2 Chronic measures of effect** The section solidly introduces the reader to fundamental terms and resources found in the draft document and appendices. Table 1 is a useful organization and summary of concepts.	Thank you for your comments.	No edits needed.

Reviewer	Comments	EPA Response	Revision Location in 2017 Draft Aluminum AWQC
Reviewer 3	**2.6.3 Aluminum Toxicity Data Fulfilling Minimum Data Requirements** The section introduces the summary of acceptable toxicity data in Table 2 and the insufficient minimum dataset requirements for estuarine/marine criterion value determination with sufficient robustness. Communicating the need for additional acute and chronic toxicity testing on estuarine/marine taxa in this section is helpful in supporting research proposals to develop that data gap.	Thank you for your comments.	No edits needed.
Reviewer 3	**2.7 Conceptual Model** The approach of conceptual models is introduced with clarity and conciseness.	Thank you for your comment.	No edits needed.
Reviewer 3	**2.7.1 Conceptual Diagram** Figure 3 and the related discussion is an excellent resource for organizing the sources, pathways, receptors and controls of environmental Al. This section and figure do not have citations, and this is a curious omission as Al is a primary resource in common use. The authors may want to explore, review and cite similar conceptual diagrams and models that may exist in other quality resources.	Thank you for raising this point. Citations were added where appropriate.	Section 2.4.1 (formerly, Section 2.7.1)
Reviewer 3	**2.8 Analysis Plan** This section is a thorough and comprehensive discussion of the background and approach for developing the data set used in modeling input and criteria development. The complexity of the topic is well treated in this exposition, and the authors present clear explanations with sufficient detail and definitions. The section is supportive of the criteria development goals of the draft document.	Thank you for your comments.	No edits needed.
Reviewer 3	**2.9 Identification of Data Gaps and Uncertainties for Aquatic Organisms** This section presents a fair and transparent treatment of data gaps and uncertainties in Al criterion development. Assumptions are clearly and openly discussed. The supporting discussion of management of data gaps and assumptions is convincing and support of the final criteria development. The explicit identification of data gaps in estuarine/marine species and for plants support the research community in their project proposals to help close those data gaps. As a minor edit, the third line from the bottom on page 30 has an extra space inserted.	Thank you for your comments. The typo you mentioned was corrected.	Section 5.3 (formerly, Section 2.9)
Reviewer 3	**3 Effects Analyses** **3.1 Acute Toxicity to Aquatic Animals** The comprehensive data sets of Appendix A and B are introduced.	Thank you for your comment.	No edits needed.

Reviewer	Comments	EPA Response	Revision Location in 2017 Draft Aluminum AWQC
Reviewer 3	**3.1.1 Freshwater** This section explicitly, and in detail, explores the available freshwater Al toxicity data. The use of tables and figures helps organize the complex available resources. The critical analysis and discussion of the limitations and usefulness of available studies are well done and supported by references to criteria development standards. The clarifying sentence: "That is, EPA's acute freshwater criteria are not fixed values, but equations based on bioavailability and hence toxicity of aluminum under different water chemistry conditions (see Appendix K for additional criteria values)" helps communicate the new Al guidance in the draft document.	Thank you for your comments.	No edits needed.
Reviewer 3	**3.1.2 Estuarine/Marine** This section discusses the data gaps and limitations that sufficiently support a finding of inability to calculate an estuarine/marine criterion.	Thank you for your comments.	No edits needed.
Reviewer 3	**3.2 Chronic Toxicity to Aquatic Animals** This section introduces Appendix C that comprehensively lists acceptable chronic toxicity data.	Thank you for your comment.	No edits needed.
Reviewer 3	**3.2.1 Freshwater** This section explicitly, and in detail, explores the available freshwater Al chronic toxicity data. The use of tables and figures helps organize the complex available resources. The critical analysis and discussion of the limitations and usefulness of available studies are well done and supported by references to criteria development standards. Many of the cited studies are recent and very high quality, addressing the anticipated data needs for developing an Al water quality criterion. The exploration of MDR missing the third family in phylum Chordata is a well-developed, transparent and satisfying approach to missing data. The detailed abstracting of the studies in this section is a useful approach to establishing the scientific basis for consideration of a diverse range of research studies as data resources for criterion development.	Thank you for your comments.	No edits needed.
Reviewer 3	**3.2.2 Estuarine/Marine** The section contains the explicit data reference of Appendix D for no estuarine/marine chronic toxicity data.	Thank you for your comment.	No edits needed.

Reviewer	Comments	EPA Response	Revision Location in 2017 Draft Aluminum AWQC
Reviewer 3	**3.3 Bioaccumulation** This is a useful section, especially as it relates to fish accumulation and Al dietary intake potential. While there is absent or limited data, the section is important for completeness.	Thank you for your comments.	No edits needed.
Reviewer 3	**3.4 Toxicity to Aquatic Plants** This too is a useful section, and while there is absent or limited data, the section is important for completeness.	Thank you for your comments.	No edits needed.
Reviewer 3	**3.5 Summary of National Criteria** **3.5.1 Freshwater** The summary approach, including example calculations and an example look-up table, with further reference to the full Appendix K look-up tables, is a good approach to introduce the reader to the National Criteria of the draft document. The approach is clear and concise, with good readability.	Thank you for your comments.	No edits needed.
Reviewer 3	**4 Effects Characterization** **4.1 Effects on Aquatic Animals** **4.1.1 Freshwater Acute Toxicity** This section explicitly, and in detail, explores the available freshwater Al acute toxicity effects data. The critical analysis and discussion of the limitations and usefulness of available studies are well done and supported by references to criteria development standards. The justifications for inclusion or rejection of specific study data appear reasonable and well supported in the document discussion.	Thank you for your comments.	No edits needed.
Reviewer 3	**4.1.2 Freshwater Chronic Toxicity** This section explicitly, and in detail, explores the available freshwater Al acute toxicity effects data. The critical analysis and discussion of the limitations and usefulness of available studies are well done and supported by references to criteria development standards. The justifications for inclusion or rejection of specific study data, for example, pH or study duration, appear reasonable and well supported in the document discussion.	Thank you for your comments.	No edits needed.
Reviewer 3	**4.1.3 Freshwater Field Studies** The discussion is useful for completeness of criteria background development and observation of need for Al management, especially in consideration of historical acidification of some surface waters in the eastern US.	Thank you for your comment.	No edits needed.

Reviewer	Comments	EPA Response	Revision Location in 2017 Draft Aluminum AWQC
Reviewer 3	**4.1.4 Estuarine/Marine Acute Toxicity** The discussion is useful for completeness of criteria background development, and correlation/comparison with other Appendix B and I data**.**	Thank you for your comment.	No edits needed.
Reviewer 3	**4.1.5 Estuarine/Marine Chronic Toxicity** The discussion is useful for completeness of criteria background development, and correlation/comparison with other Appendix D and I data.	Thank you for your comment.	No edits needed.
Reviewer 3	**4.1.6 Bioaccumulation** The discussion of effects of Al bioaccumulation is well executed with good breadth and depth of the available literature. There is limited published work available, and the analysis appears complete.	Thank you for your comments.	No edits needed.
Reviewer 3	**4.2 Effects on Aquatic Plants** With reference to Appendix E, this section summarizes and reviews toxic effect on plants with available citations.	Thank you for your comment.	No edits needed.
Reviewer 3	**4.3 Protection of Endangered Species** The summary statement that the 2017 criterion is protective of endangered species is an important early summary statement.	Thank you for your comment. Please note that the EPA was asserting that the data for the endangered species that have been tested indicated that criteria would be expected to be protective of these species; the comment was limited to the available data only. There are many untested endangered species for which no data are available, and thus no judgement was made on that point in this criteria document, as you note in the following comment.	No edits needed.
Reviewer 3	**4.3.1 Key acute toxicity data for listed fish species** This section summarizes available data in a complete and clear manner. It is an important observation that there is no acceptable acute toxicity data for endangered or threatened estuarine/marine aquatic fish species.	Thank you for your comments.	No edits needed.
Reviewer 3	**4.3.2 Key chronic toxicity data for listed fish species** This section summarizes available data, a single study, in a complete and clear manner. It is an important observation that there is no acceptable chronic toxicity data for endangered or threatened estuarine/marine aquatic fish species.	Thank you for your comments.	No edits needed.

Reviewer	Comments	EPA Response	Revision Location in 2017 Draft Aluminum AWQC
Reviewer 3	**4.3.3 Concerns about Federally listed endangered mussels** Available studies on Federally listed endangered mussels are reviewed with clarity and completeness.	Thank you for your comment.	No edits needed.
Reviewer 3	**4.4 Comparison of 1988 and 2016 Criteria Values** The title of this section should be "Comparison of 1988 and 2017 Criteria Values." This comparative analysis is important and useful to the overall document. The assumptions and approaches to setting the data in Table 11 (e.g. hardness at 100 mg/L) are scientifically reasonable and defendable. Statements concerning the relative level of protection, for example, brook trout, are an important qualifier in support of the MLR model approach that is the basis for criteria development.	Thank you for your comments and highlighting the typographical error. The section was re-titled appropriately.	Section 5.5 (formerly, Section 4.4)
Reviewer 3	**5 References** The literature supporting the draft document represents a complete and exhaustive review of the available peer-reviewed science.	Thank you for your comment.	No edits needed.
Reviewer 4	a) The document is very well prepared. The communication, documentation and clarity of the draft criteria is dramatically improved compared to the 1985 criteria document.	Thank you for your comment.	No edits needed.
Reviewer 4	b) One criticism though is in the reuse of some material from the 1985 document. This results in some very old references in the initial parts of this new draft criteria document. The initial sections (up to about page 12) draws heavily from the older document and could do with insertion of some of the new references which are detailed in later sections of the draft criteria document. In particular the first paragraph of section 2.2 and the paragraph right before section 2.4.1 are very much like the 1985 document	Thank you for your suggestions. More recent citations were added to Section 2 *Problem Formulation*, and text was moved to introduce newer topics earlier.	Section 2.2 (formerly, Section 2.3) Section 2.3 (formerly, Section 2.4)

Reviewer	Comments	EPA Response	Revision Location in 2017 Draft Aluminum AWQC
Reviewer 4	c) There should be an aluminum chemistry section. This can be an appendix and does not necessarily need to be integrated into the text. As it is written now so many "random" aspects of aluminum chemistry are introduced as the document progresses. It would be nice to have one central place discussing aluminum chemistry in some details. Emphasizing aluminum speciation and reaction kinetics.	Thank you for your suggestions. Text was added and moved to provide a clearer picture of aluminum chemistry in the main body of the document. Additionally, text was drafted that provided sources of other documents and publications available for additional detail should the reader want more information. A new Appendix was not drafted.	Section 2.2 (formerly, Section 2.3)
Reviewer 4	d) In terms of more specific communication issues, detailed comments are included as an appendix to this review.	Thank you for including these additional items.	No edits needed.
Reviewer 5	The document very clearly explained the derivation of each criterion, including the toxicity data and the water chemistry parameters used. The Multiple Linear Regression (MLR) model for determining Al criteria uses equations based on pH, hardness, and dissolved organic carbon (DOC). Therefore, the criteria would vary in a site-specific manner. The ways in which these water chemistry variables modify Al toxicity, specifically including changes in Al speciation, is very well explained.	Thank you for your comments.	No edits needed.

2.2 CHARGE QUESTION 2

2. Please comment on the technical approach used to derive the draft aluminum criteria; is it logical, does the science support the conclusion, and is it consistent with the protection of freshwater life from acute, chronic, and bioaccumulative effects? Please provide specific comments.

Reviewer	Comments	EPA Response	Revision Location in 2017 Draft Aluminum AWQC
Reviewer 1	The criteria are based primarily on multiple linear regression models that use pH, hardness and DOC as modulators of 7-day toxicity in two models – a fish (Pimephales promelas) with biomass as the endpoint, and a crustacean (Ceriodaphnia dubia) with reproduction as an endpoint. The best fit models for each species were different from each other which poses many questions: a) pH, hardness and DOC are described as important modulators of bioavailability. Do the different models then mean that bioavailability differs between species as a function of their physiology and not as a function of Al speciation? This is extremely important to articulate because the chemistry and bioavailability of Al is complex. Species also vary widely and the physiology of Al uptake is not very well understood comparatively across species.	Thank you for your comment. EPA agrees that this creates additional uncertainty; therefore text was added to Section 5.3 *Identification of Data Gaps and Uncertainties for Aquatic Organisms*. The uncertainties associated with the model are future research areas that need to be investigated.	Section 5.3.6 (formerly Section 2.9)
Reviewer 1	b) Do we know that the model based on C. dubia fits other invertebrates well? Do we know that the model based on P. promelas fits other fish well? Why are we so willing to make this assumption? Shouldn't this be tested first?	Thank you for your comment. EPA agrees that this creates additional uncertainty; therefore text was added to Section 5.3 *Identification of Data Gaps and Uncertainties for Aquatic Organisms*. The uncertainties associated with the model are future research areas that need to be investigated.	Section 5.3.6 (formerly Section 2.9)
Reviewer 1	c) Do we know that the models fit equally well for acute and chronic toxicity?	Thank you for your comment. EPA agrees that this creates additional uncertainty; therefore text was added to Section 5.3 *Identification of Data Gaps and Uncertainties for Aquatic Organisms*. The uncertainties associated with the model are future research areas that need to be investigated.	Section 5.3.6 (formerly Section 2.9)

Reviewer	Comments	EPA Response	Revision Location in 2017 Draft Aluminum AWQC
Reviewer 1	This is extremely important to articulate the uncertainty of the assumption that the models apply to other species based on gross taxonomy (vertebrates vs invertebrates).	Thank you for your comment. EPA added text to Section 5.3 *Identification of Data Gaps and Uncertainties for Aquatic Organisms*. The uncertainties associated with the model are future research areas that need to be investigated.	Section 5.3.6 (formerly Section 2.9)
Reviewer 1	P. 15: "In the final MLR model, predicted EC20s were within a factor of two of observed values for 95% of the tests. I assume that this means for the tests performed with P. promelas and C. dubia, most observations came within a factor of 2 of the model predictions that were built from the same data. More information would be useful here. Are predictions generally biased in an over-protective or under-protective way? If the uncertainty for any specific prediction is a factor of 2, shouldn't this uncertainty be built into the criteria? Should an interspecific uncertainty factor be applied as well?	Additional text was drafted that clarifies the sentence. Also additional detail on model performance and trends in predictions were added. Thank you for your suggestion of an uncertainty factor, but EPA will not be pursuing one at this time. The uncertainties associated with the model are future research areas that need to be investigated and these uncertainties were highlighted in Section 5.3.6.	Section 2.7.1 (formerly Section 2.4.1) Section 5.3.6 (formerly Section 2.9)
Reviewer 1	P. 3: Al associated with sediments are "unavailable to aquatic organisms"? This is a rather broad sweeping statement to make.	The sentence was deleted.	Section 2.1
Reviewer 1	P. 9: The temperature-solubility issue is completely ignored in the model. If a 15 degree change is equal to a whole pH unit, and pH drives the model, shouldn't there be a temperature term in the model?	As stated in Section 2.7.1, "Although many factors might affect the results of toxicity tests of aluminum to aquatic organisms (Sprague 1985), water quality criteria can quantitatively take into account only factors for which enough data are available to show that the factor similarly affects the results of tests with a variety of species." Unfortunately no empirical data is available to include temperature in the MLR model.	No edits.

Reviewer	Comments	EPA Response	Revision Location in 2017 Draft Aluminum AWQC
Reviewer 1	P. 10: "aluminum has also been shown to increase respiration, and thereby energy demands"… but an increase in respiration is a cause for excluding data from consideration?	No, an increase in respiration is not a cause for excluding data from consideration. However, these other endpoint effects are used qualitatively to support toxicity data compiled for existing species to derive the criteria. While some of these data may be used in effects characterization, only assessment endpoints specific to the survival, growth and reproduction to aquatic organisms are used quantitatively. The specific sentence you reference is from Herrmann and Andersson (1986) which is an "Unused study" (Appendix J) because the dilution water is stream water that was not characterized.	No edits.
Reviewer 1	P. 30: "Application of water only laboratory toxicity tests to develop water quality criteria to protect aquatic life" fails on many fronts and among the reasons that EPA is considering modernizing the 1985 guidelines. This paragraph is not entirely truthful.	An objective of the 1985 Guidelines is to provide for the development of criteria that are applicable to a variety of field conditions. Laboratory toxicity tests are conducted with organisms that are determined to be fundamentally healthy and with methods that meet a consistent set of standards. This facilitates evaluation of test acceptability and the development of criteria that are not impacted by testing artifacts and support national criteria application. This approach is consistent with internationally-recognized and broadly applied approaches for developing effects analyses for toxicants for both	No edits.

Reviewer	Comments	EPA Response	Revision Location in 2017 Draft Aluminum AWQC
		ecological and human health, relying on such reproducible laboratory data because they are designed to be as free from confounding influences as possible, in order to permit for robust, unconfounded consideration of risk for a given chemical, and relative risk across chemicals. States, tribes, and other end users can then consider site-specific conditions and variables in the development of standards that are applicable to their specific end use, such as the application to a particular water body or region. Criteria are generally derived primarily considering laboratory water-based exposures using procedures that are consistent with the 1985 Guidelines for some of these reasons. In addition, generally good agreement has been reported for microcosm studies/whole effluent toxicity test results with corresponding field observed effects (Clements and Kiffney 1996; Clements et al. 2002; Norberg-King 1986). EPA uses and considers field and micro/mesocosm data in criteria development when available.	

Reviewer	Comments	EPA Response	Revision Location in 2017 Draft Aluminum AWQC
Reviewer 1	P. 30: "…Only chronic data are used in the model, and application to acute data normalization assumes that the same relationships are present". This is a big assumption. Such an assumption would generally fail for BLM models based on chronic data applied to acute data, and to acute BLMs applied to chronic data for other metals. Why should we make this assumption here?	Thank you for your comment. EPA agrees that this creates additional uncertainty. Text was added to Section 5.3 *Identification of Data Gaps and Uncertainties for Aquatic Organisms*. The uncertainties associated with the model are future research areas that need to be investigated. Because aluminum chronic criteria values are lower than the acute criteria values, we expect chronic criteria to be the main foci in regulating application of the model.	Section 5.3.6 (formerly Section 2.9)

Reviewer	Comments	EPA Response	Revision Location in 2017 Draft Aluminum AWQC
Reviewer 1	P. 60: "Comparison of 1988 and 2016 (sic) criterion values": This paragraph is troubling. The new criteria values significantly weaken protection of aquatic ecosystems by allowing higher Al concentrations in most situations than the 1988 criteria. To suggest otherwise is disingenuous.	The draft 2017 AWQC for aluminum includes consideration of water chemistry parameters (pH, hardness and DOC) and reflect the current state-of-the-science on the effects of water chemistry on bioavailability and hence toxicity. The previous 1988 criteria were not derived considering water chemistry conditions, and were applied across a broad pH range despite using a specific, narrow pH test range. Additionally, the criteria were based on the very limited toxicity data available nearly 30 years ago, at the time of derivation. Therefore, based on new data and current scientific understanding regarding metal bioavailability, the1988 criteria were likely overly conservative under a number of water chemistry conditions. While the new 2017 AWQC magnitudes for aluminum may be numerically higher in certain conditions, the draft criteria do not weaken protection of aquatic ecosystems. By using the current science, application of the MLR-based aluminum criteria should more accurately yield the level of protection intended to protect and maintain aquatic life uses. If you are aware of any additional data please provide it so that it can be incorporated if acceptable.	No edits.

Reviewer	Comments	EPA Response	Revision Location in 2017 Draft Aluminum AWQC
Reviewer 2	Overall, the approach is technically sound. The vast improvement of the MLR method, which includes key water chemistry parameters, over those used for the 1988 aluminum criteria, will enable states, tribes, and local governments to assess potential aluminum toxicity at any aquatic site based on its unique aquatic chemistry.	Thank you for your comments.	No edits needed.
Reviewer 2	Section 2.2. The justification for the use of total recoverable aluminum is a complex topic which requires a decision to support the criteria development process. Further details on why the authors of the cited toxicity studies thought the dissolved concentrations were relatively level as total aluminum concentrations and toxicity increased would be good. Also, are these studies the exception? Do most toxicity studies show a correlation between dissolved concentrations of aluminum and toxicity? The toxic forms of aluminum should be in both of the dissolved and total recoverable values.	Text was added to indicate that dissolved concentrations do not correlate well with toxicity and therefore only total aluminum concentrations are appropriate measures of effect. The studies cited are not the exception.	Section 2.6.2 (formerly Section 2.2)
Reviewer 2	Section 2.3. At the start of this section the inclusion of text on acid rain issues in the 1980s and the building awareness of aluminum toxicity to aquatic biota would be informative. The discussion and the Figure 2 provide the reader with important information on the environmental chemistry, fate and transport of aluminum.	Text was added that included a brief discussion about the history of acid rain and aluminum toxicity in affected systems.	Section 2.2 (formerly Section 2.3)
Reviewer 2	Section 2.4. Increasing the range of pH in toxicity studies included in the criteria to 5.0 – 9.0 is a very good decision. Toxic forms of aluminum increase at a pH lower than 6.5.	Thank you for your comments.	No edits needed.

Reviewer	Comments	EPA Response	Revision Location in 2017 Draft Aluminum AWQC
Reviewer 2	Section 2.4.1. The discussion on the BLM method and the use of MLR is important for the reader. I think text should be added that the data requirements (ten parameters) were difficult for many state monitoring programs to address. The MLR approach uses data that almost all states already collect, so the MLR would be much more likely to be used and provide increase scientific rigor for the protection of aquatic species from aluminum. A discussion comparing the output of the methods would be good. The 2017 publications cited in this section are still (as of this writing) in review. A text box or text on key points from Brix et al., 2017 and DeForest et al., 2017 and why they support the use of the MLR would help the reader. Some limits of the DeForest study are mentioned on page 30, but It would be good to have that discussion earlier in the document and all in one place. I commend the EPA for their awareness of these studies and the cooperation with these authors to include their research.	Text was added to Section 5.3 *Identification of Data Gaps and Uncertainties for Aquatic Organisms* to include some of the uncertainties associated with the two approaches. Text was moved to provide clarity in presentation. Additional text was added to provide clarification of why the MLR approach was selected. EPA decided to use an empirical MLR approach in this draft aluminum criteria update due to the relative simplicity and transparency of the model, and the decreased number of input data on water chemistry needed to derive criteria at different sites.	Section 2.3.1 (formerly, Section 2.4.1) Section 5.3 (formerly Section 2.9)
Reviewer 2	Once the discussion focuses on the methods and results from DeForest, I would start a new section or subsection of the report. This text is the technical support for your document.	Thank you for your comment. A new subsection was created.	Section 2.7.1 (formerly Section 2.4.1)
Reviewer 2	Section 2.5. On page 17, when you first mention assessment endpoints, please state that they are listed in Table 1. On page 16, add a web address for Stephan et al., 1985 or as it is also cited, EPA, 1985.	Table 1 was cited earlier in Section 2.5 *Assessment Endpoints*. A hyperlink for Stephan et al. 1985 was added to the citation in Section 7 *References*.	Section 2.5 (new section) Section 7 (formerly Section 5)
Reviewer 2	Section 2.6.2. This section briefly discusses use of the effects concentration (EC20) to protect populations from long term chronic effects of a toxicant. If population models or a discussion relating the draft criteria to field studies are not going to be included in the document, I would recommend using the wording, "inferred protection of populations… ". Adding a population model as an additional line of evidence would be informative if the required biological and effects data are available for a species of interest.	Thank you for your suggestions. No Section 2.6.2 text edits were made, however, additional field studies were included in Section 5.1.3 that describe the aluminum effect levels to different aquatic organisms.	No edits.

Reviewer	Comments	EPA Response	Revision Location in 2017 Draft Aluminum AWQC
Reviewer 2	Section 2.7. The conceptual model section is well written and important for the user to understand the various exposure pathways and which ones may be most important.	Thank you for your comment.	No edits needed.
Reviewer 2	Section 2.8. The analysis plan is well presented. The use of a species sensitivity distribution (SSD) for the four most sensitive species and using the 5th percentile of that distribution is conservative. The decision to use a factor of 2 is more conservative and is based on a 1978 Federal Register notice. This process is supposed to protect 95% of the species in an aquatic community. Conventional use of a SSD requires at least 7 species, why is the use of the four most sensitive species proper? A conventional SSD and using an EC5 percentile is predicted to be protective of 95% of a given community (Greenberg et al., 2013). Why not use all of the data? From the text in the section, it seems that the 1978 FR notice data is a method to get to near control test organism survival percentages. I don't think that relates to protection of 95% of the species in an aquatic community. If you want to use the factor of 2, it could be stated as a policy decision.	The 1985 Guidelines recommendations were followed in that the four Genus Mean Acute Values (GMAVs) closest to the 5th percentile are used to estimate the Final Acute Value (FAV). This approach is supported by the EPA Science Advisory Panel (SAP) tasked with evaluating the merits of the "Use of SSDs to Estimate HC5 using Varying Amounts of Data" (U.S. EPA 2012).

U.S. EPA. 2012. Comparative Effects Methodology Developed by the Office of Pesticide Programs and the Office of Water. SAP Minutes No. 2012-02. FIFRA Scientific Advisory Panel Meeting Held January 31 to February 2, 2012 at One Potomac Yard, Arlington, Virginia. U.S. Environmental Protection Agency, Washington, DC. 77 pp., https://www.epa.gov/sites/production/files/2015-06/documents/013112minutes.pdf).

In essence, CWA acute criteria are based on a sensitivity distribution (SD) comprised of GMAVs, calculated from species mean acute values (SMAVs) for | No edits. |

Reviewer	Comments	EPA Response	Revision Location in 2017 Draft Aluminum AWQC
		all acceptable data that are available for the chemical. GMAVs are then rank-ordered by sensitivity from most sensitive to least sensitive, with the cumulative probability calculated for each based on the total number (N) of GMAVs. The final acute value (FAV) is determined by regression analysis using a log-triangular fit based on the four most sensitive genera (reflected as GMAVs) in the data set to interpolate or extrapolate (as appropriate) to the 5th percentile of the distribution represented by the tested genera. The SAP noted that the log-triangular distribution is a favored choice due to the more uniform representation provided by GMAVs and more uniform in its representation of the ranges of concentrations in populations across taxa.

The 5th percentile is usually determined by interpolation when N is greater than 20, and by extrapolation when N is less than 20. The SAP also noted that increasing the sample size will simultaneously decrease both the bias and the uncertainty of the HC5 estimate. If there are 59 or more GMAVs (N $\geq$59), the four GMAVs closest to the 5th percentile of the distribution are used to calculate the FAV.

Dividing the FAV by a factor of two to | |

Reviewer	Comments	EPA Response	Revision Location in 2017 Draft Aluminum AWQC
		derive a Criterion Maximum Concentration (CMC) is the standard approach used by the Agency to derive its 304(a) acute criterion recommendations, consistent with the 1985 Guidelines. The FAV is a statistical estimate of the 5th percentile of a set of LC_{50}s. The LC_{50} is defined as the concentration that kills 50% of the exposed organisms. Thus, by definition, the FAV, as defined in the 1985 Guidelines, is a concentration that would be lethal to 50% of organisms with a sensitivity greater than 95% of genera. Since the FAV is a concentration that may affect 50 percent of the 5th percentile or 50 percent of a sensitive species, this value was not considered to be protective of that percentile or that species. Therefore, per the 1985 Guidelines, to derive the CMC EPA divides the FAV by a factor of 2 with the intention of defining a concentration that will not affect the majority of organisms. The rationale for adjusting the FAV to derive the CMC is explained in item 6 on page 17 of the 1985 Guidelines. The basis for this adjustment factor is an analysis of data from 219 acute toxicity tests showing that the mean concentration lethal to 0-10% of the test population was 0.44 times the LC_{50} or the LC_{50} divided by 2.27. The data and	

Reviewer	Comments	EPA Response	Revision Location in 2017 Draft Aluminum AWQC
		analysis on which the 2.27 value is based is described in the Federal Register on May 18, 1978 (43 FR 21506-21518). Best professional judgment was used to round the FAV "adjustment factor" of 2.27 to 2 in revisions of the Guidelines that occurred subsequent to the 1987 Federal Register notice. The use of the factor became the final approach in the1985 Guidelines, under which criteria continue to be derived.	
Reviewer 2	Section 3.1.1. Once again, the factor of 2. Maybe state that it is intended to protect 95% ….	The text was edited as suggested; "intended to protect 95 percent of the species." Again, we note that the factor of 2, as stated above, is not associated with determining the 95th centile, but with reducing mortality in the 95th centile from 50% to approximately control levels, as noted above.	Section 3.1.1
Reviewer 2	Section 3.2. The discussion of the chronic toxicity studies was a very good summary. I would add at the beginning of the section that EC20 values were used for primarily sublethal endpoints.	The chronic endpoints were related to survival, growth, and reproduction including population size, biomass, and young/adult emergence as identified in Appendix C.	Section 3.2
Reviewer 2	Section 3.3. The authors could consider adding bullets on bioconcentration/biomagnification issues for inorganic metals from the EPA's Framework for Metals Risk Assessment (US EPA, 2007). An example bullet: • Trophic transfer can be an important route of exposure for metals, although biomagnification of inorganic forms of metals in food webs is generally not a concern in metals assessments.	Thank you for your suggestion. The bullet you suggested was added to the document as text.	Section 3.3

Reviewer	Comments	EPA Response	Revision Location in 2017 Draft Aluminum AWQC
Reviewer 2	Section 3.5. Table 10 is a good presentation of the effects of pH and hardness on the criteria. It would be helpful to the reader to explain why when the pH is above 7.5, and higher, that the criteria values are lower as the hardness concentrations increase.	Model performance was added to the document to highlight these trends and uncertainties.	Section 2.7.1 (formerly Section 2.4.1) Section 5.3.6 (new section)
Reviewer 2	Section 4. The discussion of aquatic toxicity studies that had non-standard experimental designs or results was good to include in this document. The results are discussed and the reasons for the lack of inclusion in the datasets are presented. A given regulator could utilize those studies if their inclusion met their needs.	Thank you for your comment.	No edits needed.
Reviewer 2	Section 4.3. In the discussion on endangered species, it might be good to discuss habitat in the U.S. where a species would be most at risk, e.g. acidic lakes with low hardness and dissolved organic carbon (DOC). An overlap of at risk habitat with an endangered species range would be informative. This suggestion may exceed the objectives of this report, but could be considered in more focused inter-agency studies.	Thank you for your suggestions. Unfortunately, these topics are beyond the scope of this document and will not be addressed in the draft aluminum AWQC.	No edits.
Reviewer 2	Section 4.4. The comparison of 1988 criteria and the 2017 draft criteria should be briefly discussed in the Executive Summary.	Thank you for your suggestion, text was added to the Executive Summary.	Executive Summary
Reviewer 3	Overall the technical approach used to derive the draft Al criteria is logical and supportive of the criteria data found in the look-up tables of Appendix K. Within the constraints of available peer-reviewed data, the technical approach is consistent with the protection of freshwater life from acute, chronic, and bioaccumulative effects. Specific comments related to a critical analysis of each section of the draft document appear above.	Thank you for your comments.	No edits needed.
Reviewer 4	Technical comments are itemized below: a) Selection of an MLR approach should be justified. Why not a BLM approach? I have no problem with an MLR approach but it is a bit of a compromise compared to a full BLM approach so the reason for selecting this method should be given. Ideally highlighting how the MLR method is as protective as a "full BLM" would be. If MLR is used for practical reasons I think that is fine as a justification, provided that it is clearly shown that there is no loss of protection.	Thank you for your suggestion. Additional text was drafted to explain the difference between the BLM and MLR approaches, and the selection of the MLR.	Section 2.3.1 (formerly, Section 2.4.1) Sections 5.3.5 and 5.3.6 (new sections)

Reviewer	Comments	EPA Response	Revision Location in 2017 Draft Aluminum AWQC
Reviewer 4	b) I think aluminum speciation is not fully addressed. In particular, solid versus dissolved Al. I am currently reviewing a manuscript by Robert Santore et al. (Windward) that does this very well (submitted to Environ. Toxicol. And Chem. "Development and application of a biotic ligand model for predicting chronic toxicity of dissolved and precipitated aluminum to aquatic organisms"). I suspect that the authors of this report are familiar with that work but if not they should contact the authors to preview a copy. This draft criteria document over simplifies the chemistry of aluminum by lumping particulate and dissolved aluminum into a total measure of aluminum. I think this is fine – based on reading the Santore draft paper – but the draft criteria document does not convince me it is fine. Different modes of exposure (particulate versus soluble) and not even really hinted at in the draft criteria document.	Additional text was added to provide clarification of this ambiguity and application to field exposures.	Section 2.6.2 (formerly Section 2.2) Section 5.3.3 (new section)
Reviewer 4	c) Selection of total Al as the independent variable needs to be more fully justified (see comment above) but also it needs to be made clear if this is in fact what was done. It seems that the discussion on page 3 and 4 about the operational definitions of total and dissolved is used in the rest of the document, in particular, for the numerical criteria. This needs to be made clear. As mentioned in the Appendix (to this review), aluminum speciation needs to be defined more clearly. Total aluminum could include particulate refractory minerals, weathering product minerals (i.e., clay such as kaolinite) or it could include reactive solids such as amorphous gibbsite which can exchange back and forth solid to dissolved phase. Should these really all be included in one "box". This might be ok for lab studies where only amorphous gibbsite occur as solids – but what about field?	Additional reference materials were added to the document to provide the reader sources for more specific information.	Section 2.2 (formerly Section 2.3)
Reviewer 4	d) Kinetics of "land scale" processes compared to acute and chronic tests. The criteria is based on lab-based studies but will be implemented with real samples. In acute tests the time scales are very short but in the field longer timescales will control the total aluminum. Not geologic timescales but longer than laboratory tests.	Thank you for your comments.	No edits.

Reviewer	Comments	EPA Response	Revision Location in 2017 Draft Aluminum AWQC
Reviewer 4	e) Bottom-line the proposed criteria will be protective for aquatic life in an acute and chronic sense. The science is logical but needs just a bit more justification (as detailed in my points above). My personal bias is that I would prefer a BLM approach but I completely understand the practical needs that make MLR an attractive and possibly more likely to be adopted option.	Thank you for your comments.	No edits needed.
Reviewer 5	The technical approach used to derive the draft Al criteria is logical and scientifically supported. The MLR model has been validated using *Ceriodaphnia dubia* and *Pimephales promelas* with greater than 86% success at predicting Al toxicity values. Bioaccumulation of Al is most observed in the lower trophic levels, as specifically stated in the document.	Thank you for your comments.	No edits needed.

2.3 CHARGE QUESTION 3

3. Please comment on the data used to derive the revised criteria, including data adequacy/comprehensiveness, and the appropriateness of the data selected and/or excluded from the derivation of the draft criteria. Are there other relevant data that you are aware of that should be included? If so, please provide the reference.

Reviewer	Comments	EPA Response	Revision Location in 2017 Draft Aluminum AWQC
Reviewer 1	The compilation of references is thorough and robust. The reasons for excluding data are generally reasonable, though in a data-limited world, it may be more appropriate to look for ways to include data rather than excluding it. Non-standard but biologically important endpoints should probably be retained rather than excluded, but I don't think they would change criteria values in this case.	The 1985 Guidelines recommends using specific endpoints (see Section VI. E of the Guidelines). More specifically, these "Other data" are used qualitatively to support toxicity data compiled for existing species to derive the criteria. While some of these data may be used in characterization, data deemed unacceptable are not used in criteria derivation. The Guidelines assessment endpoints are specific to the survival, growth and reproduction of aquatic organisms.	No edits.
Reviewer 1	It is apparent that there was a big push from industry to generate data to change this criteria, and industry is well within their rights to do so. There is no accompanying effort by EPA to ground truth or validate the industry supplied model that drives this criteria.	Thank you for your comment. EPA disagrees with your assertion. EPA examined the MLR model developed by DeForest et al. (2017), through internal examination of the data and the model. EPA also compared the output of the model to outputs of the aluminum BLM (Santore, 2016 model), and to independent empirical data (not included in the model development). All of these "ground truthing" efforts returned acceptable results. EPA thus concluded that the model was acceptable for application in criteria development.	Section 5.3.6 (formerly Section 2.9)

Reviewer	Comments	EPA Response	Revision Location in 2017 Draft Aluminum AWQC
		Additionally, the MLR model of DeForest et al. (2017) has been published in a high quality, peer-reviewed journal, and as such was subjected to independent external peer review by expert scientists selected by the journal. The uncertainties associated with the model are future research areas that do need to be investigated, as with most models.	
Reviewer 2	The presentation of the aquatic toxicology data on aluminum was very comprehensive. The rational for what data were used for acute and chronic calculations for WQC was clear. The presentation and discussion of the data that were not utilized, in appendices I and J, certainly addresses past complaints about transparency and bias in selecting studies to include in analysis. A reader will know why a given study was or was not selected for criteria development.	Thank you for your comment.	No edits needed.
Reviewer 2	As I stated above, after explaining that there is not enough data to calculate WQC for estuarine/marine biota, moving all of the text on that topic from the main body of the report into an appendix would improve the flow of the document.	Thank you for your suggestion, but to be consistent with other AWQC for different pollutants, the estuarine/marine sections were retained throughout the document.	No edits.
Reviewer 2	In Figure 2 and in the paragraph on page 9 above the figure, the text states that aluminum in the water column at pH 7.0 is almost all in the insoluble form of aluminum hydroxide. It seems that there should be a more dramatic shift in criteria values moving from pH 7 to pH 5.5. In table 10 at a hardness concentration of 100 ppm and DOC of 1.0 ppm, the criteria at pH 5.5 is six fold lower than the value at pH 7.0. I wonder if the difference could be greater. In section 2.4.1, page 16, the process for normalization of the data to pH 7.0 is explained, however, the models are built from studies with a pH ranging from: pH 6.14 – 8.0, pH 6.3 – 8.1, and pH 6.0 – 8.0. These models extrapolate criteria below pH 6.0. Perhaps further validation of the models would be useful with data from similar studies that have data at a pH below 6.0. The authors, in section 2.9, identified other data needs and resulting uncertainty issues building from the DeForest et al. (2017) manuscript.	Thank you for your comment. EPA agrees that additional data below pH 6 would be useful. Text was added to Section 5.3 *Identification of Data Gaps and Uncertainties for Aquatic Organisms*.	Section 5.3.6 (formerly Section 2.9)

Reviewer	Comments	EPA Response	Revision Location in 2017 Draft Aluminum AWQC
Reviewer 2	The chronic brook trout study by Cleveland et al. (1989) is a driver for the chronic aluminum WQC. More discussion of the study could include more information on the statistical analysis of endpoints and the presentation of endpoints that were not significantly different from controls. It is noted that EC20 values were used for primarily sublethal endpoints.	The most sensitive acceptable endpoint (growth, survival or reproduction) is used for each study. Typically AWQC documents do not expand on those endpoints that are not different from the control. Selection of EC_{20} for chronic criteria for non-bioaccumulative chemicals is consistent with EPA's recent criteria. The chronic brook trout study by Cleveland et al. (1989) was re-examined in depth by EPA for this draft 2017 criteria document, and the acceptable endpoints were carefully selected to capture the best available science. The criteria document contains a new, more extended discussion of this key study.	Section 3.2.1
Reviewer 2	On page 17, data requirement from the EPA 1985 report are discussed. It would be informative for the reader to include a table of those data requirement. Since 1985, data quality has become a more studied topic and issues of experimental design, chemical analysis, statistical analysis, replication, etc. have become more important to include in toxicity publications.	The 1985 guidelines minimum data requirements are listed in bulleted form on page 19 in the draft criteria document. Table 2 on page 26 identifies which data requirements were met with quantitative data	No edits.
Reviewer 2	Section 3.4. The bioconcentration discussion was well presented. I suggested previously that some broader policy statements on the bioconcentration of inorganic metals be included. I would also recommend that the discussion in 4.1.6 be combined with the discussion in this section.	Thank you for your suggestions. The sections were not combined, but the bullet you recommended above on this topic was included.	Section 3.3
Reviewer 2	Section 4.1.3. The inclusion of field data in the report is a very good step to relate the criteria development process to observed field effects.	Thank you for your comment.	No edits needed.

Reviewer	Comments	EPA Response	Revision Location in 2017 Draft Aluminum AWQC
Reviewer 3	The data qualified in the derivation of the revised criteria are the result of a complete and thorough analysis through 2013. The data set is comprehensive, and acceptance or exclusion judgments or assumption are transparent, linked to formally published guidance, and explicitly discussed. Exploration of the 47 citations retrieved from a "Web of Science" database search of aluminum aquatic toxicity, for 2013 to the present did not yield any new controlled Al exposure studies. Specific comments related to the data used to derive the revised criteria of each section of the draft document appear above.	Thank you for conducting an additional search.	No edits needed.
Reviewer 4	a) Data seems very comprehensive and the data standards are very high. The rejected datasets are still discussed in the text so the document has a very balanced view.	Thank you for your comment.	No edits needed.
Reviewer 4	b) I am not aware of any other data sets that should be included.	Thank you for your comment.	No edits needed.
Reviewer 5	There were more available, acceptable data used to derive freshwater acute and chronic criteria in the 2017 Al document as compared to that used in the 1988 document. For acute values, the 2017 Al document included 21 species (12 invertebrates, eight fish species, and one frog species) in 19 genera as opposed to the 15 species (eight invertebrates and seven fish species) in 14 genera that were included in the 1988 Al document. These data fulfilled the minimum data requirements (MDR). For chronic values, the 2017 document included 11 species (seven invertebrates and four fish species), whereas, the 1988 document included only three species (two invertebrates and one fish species). The third family in phylum chordata is missing in the chronic database; however, "Other Data" was used to fulfill the missing MDR group. Data from a study using a wood frog was included in the chronic database for this purpose. I agree that this is the more scientifically defensible route to directly determine the final chronic value rather than using both the final acute value and the final Acute-Chronic Ratio (FACR) to estimate the final chronic value. All of the freshwater data were normalized to a pH of 7, hardness of 100 mg CaCO3/L and DOC of 1 mg/L for comparison. The criteria in the 2017 Al document are determined based on those site-specific parameters (pH, hardness, and DOC).	Thank you for your comments.	No edits needed.

Reviewer	Comments	EPA Response	Revision Location in 2017 Draft Aluminum AWQC
	New acceptable, acute estuarine/marine data is available for five families representing five species in the 2017 Al document, as compared to no data before; however, the database still does not meet the minimum of eight families necessary to fulfill the MDR. One data gap identified is for estuarine/marine chronic Al toxicity, where there were no acceptable data. The data selected to be included in the databases were, in my opinion, appropriate and thorough. For comparison purposes, studies were also discussed (data provided in Appendices) that were missing some/all water chemistry parameters (i.e. pH) necessary for inclusion in the database.		

2.4 CHARGE QUESTION 4

4. Are the derived criteria appropriately protective of commercially and recreationally important species, and of ecosystems overall?

Reviewer	Comments	EPA Response	Revision Location in 2017 Draft Aluminum AWQC
Reviewer 1	This is a difficult question to answer because the application of the MLM (peer reviewer typo) models to other species has not been validated. Nor has the model transferability between chronic and acute effects been validated. What is particularly concerning, is that the publication cited as being the basis for this criterion is a submitted manuscript that at this time of writing and has not even completed the peer-review process. The model looks promising, but to use it without doing the cross species validation, and without chronic to acute validation is pre-mature.	Thank you for your comment. EPA agrees that this creates additional uncertainty; therefore text was added to Section 5.3 *Identification of Data Gaps and Uncertainties for Aquatic Organisms*. The uncertainties associated with the model are future research areas that need to be investigated. The principle that bioavailability affects both chronic and acute toxicity is anticipated to be a reasonable assumption, consistent with numerous aquatic toxicity studies available in the open literature. EPA anticipates that the submitted manuscripts will be available to the public when the draft AWQC is released for public review. Should these manuscripts receive significant changes then the AWQC document will be updated to reflect relevant changes.	Section 5.3.6 (formerly Section 2.9)

Reviewer	Comments	EPA Response	Revision Location in 2017 Draft Aluminum AWQC
Reviewer 1	If adopted, in practice, much will depend on the assumptions that state-level entities apply to the stability of pH, DOC and hardness in the systems that these criteria are intended to protect. Are they assumed to be stable over time? Are they based on a single measurement in a site specific context? Is a state wide mean to be used for these important toxicity modifiers? Does the normalization of the model to conditions that are minimally toxic leave open the possibility that appropriate conversions based on pH, DOC and hardness may not occur?	Previous implementation guidance that EPA developed for other criteria (i.e., Freshwater Cu BLM) would also apply. Of particular importance is estimating water chemistry in natural waters. In 2016, EPA issued a draft document providing ecoregional data for ions and DOC across the nation, and further subdivided this data by stream order. The underlying data were collected from federal empirical databases. (web link: https://www.epa.gov/sites/production/files/2016-02/documents/draft-tsd-recommended-blm-parameters.pdf). The document was originally developed to support data gathering for BLM model implementation, but can be used for other models that rely on water chemistry data.

EPA plans to update this document and issue a final document to support state application of criteria based on water chemistry where states have difficulty collecting adequate data. Additionally, EPA plans to develop additional technical assistance documents to support application of bioavailability - based water quality criteria for metals. | No edits. |
| Reviewer 2 | I think that the derived criteria are appropriately protective of aquatic biota and aquatic ecosystems. | Thank you for your comment. | No edits needed. |

Reviewer	Comments	EPA Response	Revision Location in 2017 Draft Aluminum AWQC
Reviewer 2	Compared to the 1988 criteria, the 2017 draft criteria are based on a larger dataset which provides more confidence in the criteria development process. The utilization of the aquatic chemistry of aluminum in the development of the criteria is a major improvement which recognizes the complex aspects of aluminum toxicity to aquatic biota. Also, the progression of tools to assess metal aquatic chemistry and toxicity from the BLM to MLR provides a method that states and tribes can easily use with aquatic monitoring data they already collect. At the end of the document, a new section (4.5 or 5.0) entitled conclusions should be added that states these major benefits of the draft 2017 criteria.	Thank you for your suggestion. However, no text edits were made. The major differences/benefits of the 2017 update are now highlighted in the Executive Summary and in Section 5.5 (Comparison of 1988 and 2017 Criteria Values).	No edits.
Reviewer 2	For most aquatic systems in the U.S., with a pH between 6.5 – 8.0 (or so) aluminum should not be a significant stressor of aquatic biota. I think a broad discussion of the regions of the county where the aquatic chemistry of streams and lakes would indicate where aluminum toxicity could be a significant stressor to biota would be informative to the reader. In addition, issues of background concentrations of aluminum in low pH, hardness, and DOC systems could have aluminum concentrations above the proposed criteria and tolerance or adaptation may have occurred in these systems over time.	Based on empirical data, regions of the U.S. with low pH, low DOC, and low hardness, would have low recommended aquatic life criteria for aluminum based on bioavailability. Tolerance to aluminum in specific water types is possible, and states and tribes may modify criteria for site-specific waters (Allin and Wilson 2000; Tietge et al. 1988).	No edits.
Reviewer 2	Section 2.8, page 27. I have concerns on the discussion of the acute criterion where the final acute value (FAV) is divided by 2 to protect 95% of the species in a representative aquatic community from acute effects. The protection of 95% of the species in a community from acute effects is a policy goal and it should be stated as such. The dataset of no observable effects concentration (NOEC) values which are then divided by a safety factor of two is a management decision and I don't see how dividing by two is related to a 95% protection goal. It seems a SSD where 95% of the species are protected would be a more defensible approach. On page 45, it appears that the chronic criteria are developed using the MLR with no safety factor. As you move toward finalization of the aluminum WQC, "May The Force be with you."	EPA notes that the goal of protecting approximately 95% of taxa in aquatic ecosystems and the approach of dividing the acute 5th percentile LC_{50} by 2 are two separate issues. Safety factors were not applied to the EC_{20}, the chronic endpoint selected for use for aluminum.\n\nThe 1985 Guidelines recommendations were followed in that the four GMAVs closest to the 5th percentile are used to estimate the FAV. Because aquatic ecosystems can tolerate some stress and occasional adverse effects, protection of	No edits.

Reviewer	Comments	EPA Response	Revision Location in 2017 Draft Aluminum AWQC
		all species at all times and places is not deemed necessary. If acceptable data are available for a large number of appropriate taxa from an appropriate variety of taxonomic and functional groups, a reasonable level of protection will probably be provided if all except a small fraction of the taxa are protected, unless a commercially or recreationally important species is very sensitive. The small fraction was set at 0.05 because other fractions resulted in criteria that seemed too high or too low in comparison with the sets of data from which they were calculated. This approach is consistent with the broad international use of the HC5 from a sensitivity distribution in ecological assessments Dividing the FAV by a factor of two to derive a CMC is the standard approach used by the Agency to derive its 304(a) acute criterion recommendations, consistent with the 1985 Guidelines. The FAV is a statistical estimate of the 5th percentile of a set of LC_{50}s based on a sensitivity distribution. Thus, by definition, the FAV, as defined in the 1985 Guidelines, is a concentration that would be lethal to 50% of organisms with a sensitivity equal to or greater than 95% of genera, thus this value cannot be	

Reviewer	Comments	EPA Response	Revision Location in 2017 Draft Aluminum AWQC
		considered to be protective of genera in that or lower percentiles (i.e., very sensitive taxa). The rationale for adjusting the FAV by dividing by 2 to derive the CMC (acute criterion) is explained in item 6 on page 17 of the 1985 Guidelines. The basis for this adjustment factor is an analysis of data from 219 acute toxicity tests showing that the mean concentration lethal to 0-10% of the test population was 0.44 times the LC_{50} or the LC_{50} divided by 2.27. The data and analysis on which the 2.27 value is based is described in the Federal Register on May 18, 1978 (43 FR 21506-21518). Best professional judgment was used to round the FAV "adjustment factor" of 2.27 to 2 in revisions of the Guidelines that occurred subsequent to the 1987 Federal Register notice. The use of the factor became final EPA guidance in the 1985 Guidelines. The Final Chronic Value (FCV) has no safety factor: Chronic values obtained from toxicity tests already correspond to a minimal level of effect. For aluminum this was selected as the EC_{20}. Because the aluminum chronic criterion is derived directly from the chronic Genus Sensitivity Distribution, no additional ratios are applied, in accordance with the Guidelines.	

Reviewer	Comments	EPA Response	Revision Location in 2017 Draft Aluminum AWQC
Reviewer 3	Yes, the derived criteria appropriately protective of commercially and recreationally important species, and of ecosystems overall. The 2017 draft criteria look-up tables present a research-based improvement on the 1988 criteria. Specific comparative examples such as brook trout in the draft document section 4.4, offer support for this conclusion. Quantifying the effects of water chemistry (pH, DOC, and Hardness) for aluminum risk to aquatic species is an improvement to meeting the goals of the CWA from the 1988 approach, helping states, tribes, and the regulated community maintain environmental quality and healthy aquatic ecosystems.	Thank you for your comments.	No edits needed.
Reviewer 4	a) Yes the document clearly demonstrated that criteria will be protective to commercially and recreationally important species and in fact to endangered species. The revised criteria are an improvement over the existing criteria.	Thank you for your comments.	No edits needed.
Reviewer 4	b) In terms of ecosystem protection overall the field studies that are discussed make it pretty clear that effects were observed for higher aluminum then would be recommended by this draft criteria.	As stated previously, extrapolation from the laboratory to the field is a scientifically valid and protective approach for aquatic life criteria development. In certain field situations, however, additional factors may be present that affect the bioavailability/ toxicity of aluminum to aquatic organisms. Thus, depending on the concentrations of specific key variables that influence aluminum toxicity (e.g., pH, total hardness and DOC), the aluminum effect levels observed in the cited field studies are higher than those recommended in this document at specific chemical conditions (pH of 7, hardness of 100 mg/L as $CaCO_3$ and DOC of 1.0 mg C/L).	No edits.
Reviewer 5	The database for acute and chronic Al data in freshwater and estuarine/marine systems is comprehensive; however, data gaps have been identified. Data for endangered freshwater mussels or closely related species are lacking. For these	EPA collaborated with USGS to solicit new data generation for freshwater mussels to serve as surrogates for	No edits needed.

Reviewer	Comments	EPA Response	Revision Location in 2017 Draft Aluminum AWQC
	organisms the derived criteria may not be protective. Using current available data, I believe the derived criteria will be appropriately protective of most commercially and recreationally important species and of ecosystems.	endangered mussels, due to concerns over potential effects of aluminum on endangered mussels. (e.g., see citations for Wang et al. in criteria document.) As noted in the draft 2017 aluminum criteria document, while the 96-hr LC_{50} juvenile test on fatmucket mussel (*Lampsilis siliquoidea*) failed to elicit an acute 50% response at the highest concentration, the 28-day biomass normalized SMCV ranked as the third most sensitive genus in the dataset. However, the SMCV was more than 2-fold greater than that of the most sensitive species, Atlantic salmon, and 2.7 times higher than the freshwater criterion. Thus, the chronic criterion is expected to be protective of this and related species. The fatmucket tested is not a threatened and/or endangered species, but the genus *Lampsilis* contains several listed species with a wide distribution across the United States. Additional testing on endangered mussel species, or closely related surrogates, would be useful to further examine the potential risk of aluminum exposures to endangered freshwater mussels.	

2.5 OTHER COMMENTS PROVIDED

Reviewer	Comments	EPA Response	Revision Location in 2017 Draft Aluminum AWQC
Reviewer 4	**Appendix: detailed "editorial" comments.** a) Chemistry section is a bit sparse. In particular colloidal and particulate and kinetics. If you measure total Al does it also include minerals like feldspar and weathering minerals like kaolinite? Then how does that compare to freshly precipitated amorphous gibbsite?	Thank you for your suggestion. Text was added and moved to provide a clearer picture of aluminum chemistry. Additionally, text was drafted that provided sources of other documents and publications available for additional detail should the reader want more information.	Section 2.2 (formerly Section 2.3)
Reviewer 4	b) A table of acronyms at the start of the document would be beneficial. Also making sure that each acronym is defined before first use (for example DOC in the executive summary (page vi) is not defined).	Thank you for your suggestions. A new section was added and DOC was defined in the Executive Summary.	New section (Acronyms) added to the upfront material.
Reviewer 4	c) For many comparisons in the text the criteria value for water chemistry of pH 7, DOC=1mg C/L and 100 mg CaCO3 hardness is used. I realize many more possible values are reported in the look up tables at the end of the document but why were these conditions selected for comparisons in the text? (such as the comparisons to the "old" criteria document). pH 7 is not exactly a common pH and DOC of 1 mg C/L is a pretty low value.	Text was added to provide clarity why these conditions were selected. These specific values were chosen to represent pH, hardness and DOC levels simply as examples found in the environment. Appendix K shows recommended criteria for pH ranges 5-9, DOC ranges 0.5 – 5 mg/L, and hardness ranges 25-400.	Section 2.7.1 (formerly Section 2.4.1)
Reviewer 4	d) Page vii there is a space missing 4 lines up from the bottom; "1985Guidelines" should be "1985 Guidelines".	Thank you for highlighting this typographic error; it was corrected.	Executive Summary
Reviewer 4	e) Page 2, "… yet very rarely in the elemental state …". The use of the word "yet" is a strange choice. And I'm not sure Aluminum is ever found in its elemental state in natural systems.	The sentence was edited for clarity.	Section 2.1
Reviewer 4	f) Paragraph at the start of section 2.1 (page 2) does not mention aluminum in clays or soil or sediment. Just aluminum in rocks and that is complexed by some inorganic and organic ligands. Also, particulate (suspended) aluminum and colloidal and polymeric forms should be mentioned. These forms are referred to later in the document so they should be introduced here in the overview section.	Information was added to this section.	Section 2.1

Reviewer	Comments	EPA Response	Revision Location in 2017 Draft Aluminum AWQC
Reviewer 4	g) Paragraph starting at the bottom of page 2 starts with ``Aluminum enters the environment …'' is a strange opening. Aluminum is everywhere in the environment. The authors point out in the text how abundant aluminum is in the earth's crust. So I think the authors need to be specific here – what forms of aluminum are "entering" the environment? Does this mean dissolved aluminum? The document is written as if aluminum is like other metal contaminants and aluminum is very different.	Sentence edited to, "enters aquatic environment."	Section 2.1
Reviewer 4	h) Paragraph starting at the bottom of page 2 says both natural and anthropogenic sources of aluminum are responsible for a "meaningful amount found in the aquatic environment". I am not clear what meaningful amount means in this context.	The sentence was edited to, "Aluminum enters the aquatic environment from both natural and anthropogenic sources with natural sources typically dominating occurrence (Lantzy and MacKenzie 1979)."	Section 2.1
Reviewer 4	i) Paragraph starting at the bottom of page 2 might be an appropriate place to mention aluminum solubility and Ksp (unless a separate section on chemistry is adopted per my suggestion above). This is an important concept in nature and a really important concept in the toxicity experiments.	Several sources had conflicting Ksp values for $Al(OH)_3$ so we did not add this information.	No edits.
Reviewer 4	j) Page 3 around the middle, aluminum is also used in wastewater treatment to remove phosphorus	Information was added to the document.	Section 2.1
Reviewer 4	k) Page 3 around the middle of the page, it is not clear what is meant by ``diversely abundant''.	The sentence was deleted.	Section 2.1
Reviewer 4	l) Top of page 4, colloidal and particulate aluminum species should be clearly discussed. The current text does explain soluble speciation (i.e., complexation) but fails to recognize solid speciation. There is a large difference between a particle of feldspar or kaolinite from freshly precipitated aluminum hydroxide. Also polymeric species.	Thank you for your suggestion. Text was added to provide a clearer picture of aluminum chemistry. Additionally, text was drafted that provided sources of other documents and publications available for additional detail should the reader want more information. Also the text was added to Section 5.3 to address the uncertainty of measuring total aluminum in the natural environment.	Section 2.2 (formerly Section 2.3) Section 5.3.3 (formerly Section 2.9)

Reviewer	Comments	EPA Response	Revision Location in 2017 Draft Aluminum AWQC
Reviewer 4	m) This is a bit of a technical question but why does the minimum detection limit for dissolved concentration differ from the minimum detection limit of total aluminum? (bottom of page 4).	The paragraph was updated and edited whereby the specific reference to a detection limit was removed. Different methodology and matrices involved with each analyte will often yield different detection limits.	Section 2.1
Reviewer 4	n) Figure 1 gives a map of wells but there is no Figure showing a map of the surface water samples discussed.	Thank you for your suggestion. We tried to identify a map with this information, but unfortunately no map is available.	No edits.
Reviewer 4	o) Below Figure 1 on page 5 in reference to marine and estuarine systems the recent paper by Angel et al. should be cited. It gives specific measured values. This paper is also useful because it points out the dynamics of aluminum in toxicity test solutions – it specifically considers the timescale and chemistry of exposure solutions. [Geochemical controls on aluminium concentrations in coastal waters. Brad M. Angel, Simon C. Apte, Graeme E. Batley and Lisa A. Golding. Environmental Chemistry 13(1) 111-118 http://dx.doi.org/10.1071/EN15029].	Thank you for highlighting this additional publication. Information from the paper was added to the document.	Section 2.1
Reviewer 4	p) Bottom of page 6, ``higher acid extractable aluminum'' should be explained what this term means. Also, this highlights the need for solid (particulate) aluminum speciation measurements. There is a difference between geologic minerals and transient hydroxide precipitates.	The sentence was edited for clarity.	Section 2.1
Reviewer 4	q) Middle of page 7 is the first mention of polymeric and colloidal forms of Al and also the first mention of Al sorbed to clay. These concepts need to be mentioned sooner – again, a separate appendix on Al chemistry should be considered.	Thank you for your suggestion. Text was added and moved to provide a clearer picture of aluminum chemistry. Additionally, text was drafted that provided sources of other documents and publications available for additional detail should the reader want more information. A new Appendix was not drafted.	Section 2.2 (formerly Section 2.3)
Reviewer 4	r) Bottom of page 7 there is a mention of carbonate precipitates of aluminum. I'm not sure these exist. A citation should be given.	Text was edited for clarity.	Section 2.6.2 (formerly Section 2.2)

Reviewer	Comments	EPA Response	Revision Location in 2017 Draft Aluminum AWQC
Reviewer 4	s) Middle of page 8 is the first mention of measuring aluminum speciation (Driscoll reference). There is a large body of literature from the 1980s on aluminum speciation and although this is old literature it should be reviewed. The authors use this one Driscoll citation to downplay the usefulness of aluminum speciation methods which define fractions such as monomeric and polymeric aluminum. It is true that these methods are operationally defined but perhaps this section could include more substantial literature review before rejecting them out of hand? Lumping all particulate aluminum into a single "box" as is proposed with a single total aluminum measurement has its own limitations.	Thank you for your suggestion. Text was added to provide a clearer picture of aluminum chemistry. Additionally, text was drafted that provided sources of other documents and publications available for additional detail should the reader want more information. Text was added to indicate that dissolved concentrations do not correlate with toxicity and therefore only total aluminum concentrations are appropriate measures of effect. The studies cited are not the exception.	Section 2.2 (formerly Section 2.3) Section 2.6.2 (formerly Section 2.2)
Reviewer 4	t) Sentence in the middle of the first complete paragraph on page 8 is out of place. "Aluminum toxicity is important in freshwater and marine environments due to both anthropogenic sources …" Just kind of comes out of no where.	The sentence was deleted.	Section 2.6.2 (formerly Section 2.2)
Reviewer 4	u) Section 2.3 it is not clear if this section is referring to soluble aluminum transport or also including particulate aluminum transport.	This section is a general discussion about the fate and transport of aluminum in the aquatic environment.	No edits.
Reviewer 4	v) Page 9 near the top of the page says that at neutral pH aluminum is nearly insoluble – this should be quantified. The Ksp of aluminum hydroxide allows clear estimation of the solubility limits of aluminum.	Several sources had conflicting Ksp values for $Al(OH)_3$ so we did not add this information.	No edits.
Reviewer 4	w) Figure 2 is a good figure but another figure showing precipitation equilibria would give a more balanced view of aluminum speciation. Also the final sentence of the figure caption is not in bold face font.	Thank you for your suggestion. However additional figures are not included. Text was drafted that provided sources of other documents and publications available for additional detail should the reader want more information.	No edits.
Reviewer 4	x) The list of species at the bottom of page 10, top of page 11, is a repetitive list but does not mention the same species as the previous listing of species (i.e., first paragraph of section 2.1.).	Sentence was edited for clarity.	Section 2.2 (formerly Section 2.3)

Reviewer	Comments	EPA Response	Revision Location in 2017 Draft Aluminum AWQC
Reviewer 4	y) Are all filters created equal? The authors refer to 0.45 μm filtration throughout the document but do not specify filter media. Different media can lead to different answers. Probably beyond the level of detail this document is intended for but I mention this incase the authors want to consider.	The 0.45 μm filtration reference used throughout the document is to distinguish between total and dissolved aluminum. Indication the specific type of filter media is beyond the scope of this document. As stated in EPA Method 1669 (U.S. EPA 1996), the recommended filter is a 0.45 μm Gelman Supor 12175, or equivalent.	No edits.
Reviewer 4	z) The first complete paragraph on page 12 goes into more details on aluminum chemistry. These details could be included earlier in the more general discussion paragraphs. In particular time is mentioned as a variable here. This is significant in aluminum chemistry and needs to be highlighted earlier. The Angel manuscript mentioned above does a good job of showing aluminum kinetics on the timescale of toxicity testing.	Text was added and moved to provide a clearer picture of aluminum chemistry. Additionally, text was drafted that provided sources of other documents and publications available for additional detail should the reader want more information. Also text added was added on aluminum kinetics as it relates to toxicity testing.	Section 2.2 (formerly 2.3) Section 2.3 (formerly 2.4)
Reviewer 4	aa) The mention of aluminum flocs coprecipitating nutrients in the middle paragraph of page 12 is a well known phenomena. It is in fact how chemically mediated phosphorus removal works in wastewater treatment. Also these examples (and this text) is taken from the 1985 criteria document. Some more recent examples would be better.	Additional references were added where appropriate.	Section 2.1 (formerly Section 2.2) Section 7 (formerly Section 5)
Reviewer 4	bb) Page 13 when the Appendices are referred to it would be nice to give the title of the appendix. That would really help the reader (I noticed this in a later section when the title was given in reference to the appendix.). Specifically here though it would be nice to include some of the numbers and not rely on the reader to go and read the Appendix to find the number for themselves.	Thank you for your suggestion. The title of each appendix was added where appropriate.	Various locations

Reviewer	Comments	EPA Response	Revision Location in 2017 Draft Aluminum AWQC
Reviewer 4	cc) Second to last paragraph on page 13 refers to MLR models as less complex than BLM models. I think I would also emphasize that MLR models are not mechanistic. They are referred to as bioavailability models and it is semantics but since there is no mechanistic aspect to an MLR I would prefer if they were not referred to as bioavailability models. It is true MLR models take toxicity mitigating factors into account so maybe some reference to that would be better?	Text was added to clarify that the MLR models are not mechanistic bioavailability models, like the BLM. The authors (DeForest et al. 2017) refer to the MLR model as a bioavailability model.	Section 2.7.1 (formerly 2.4.1)
Reviewer 4	dd) When MLR models are discussed on page 14 (bottom paragraph) it is not clear if total aluminum (as defined in this document) is the independent variable or if other forms of measured aluminum where tested.	Text was edited to provide clarity that the MLR models are based on total aluminum effect concentrations.	Section 2.7.1 (formerly 2.4.1)
Reviewer 4	ee) Last paragraph of page 14, are pH, hardness and DOC really independent variables? More independent than say pH and alkalinity but these aren't really independent variables. Harder water could have more alkalinity (hardness cations are strong base cations after all) and high DOC waters are often acidic because of organic acids.	They water quality parameters (pH, hardness and DOC) were treated as independent variables for model development. However, the MLR model does contain interaction terms for the parameters.	No edits.
Reviewer 4	ff) The MLR paragraph at the bottom of page 14 really emphasizes the importance of pH, DOC and hardness. These variables are mentioned in the earlier text but I think could be emphasized more strongly. So many other things like sulfate, phosphate, … are mentioned as well. The reader should be convinced that these are the best variable choices.	These parameters (pH, DOC, and hardness) are emphasized throughout the document.	No edits.
Reviewer 4	gg) Page 15 6 lines down, "more simple" should be "simpler"	Text was edited as suggested.	Section 2.7.1 (formerly 2.4.1)
Reviewer 4	hh) Page 15, only R2 values are given, p values should also be given.	The individual p values for each model term are available in the DeForest et al. (2017). There is no real overall p value for each model.	No edits.
Reviewer 4	ii) Section 2.7.1 repeats things that have been mentioned before. It almost seems that the earlier sections drew heavily from the 1985 document and then this text was created for this current document. The end result is repetitive though. I prefer how things are presented in section 2.7.1 and I really like Figure 3.	Thank you for your suggestion. Text was moved and edited for clarity.	Section 2.4.1 (formerly section 2.7.1)
Reviewer 4	jj) A figure similar to Figure 3 could be developed showing aluminum speciation. Or Figure 3 could include some indication of aluminum speciation.	Thank you for your suggestion. However no text edits were made.	No edits.

Reviewer	Comments	EPA Response	Revision Location in 2017 Draft Aluminum AWQC
Reviewer 4	kk) Page 30 middle paragraph argues how laboratory toxicity tests can be extrapolated to the field. This is reasonable for other metals but some caveats should be given for aluminum. The slow kinetics and the difference between geological aluminum materials suspended in natural water and the transient precipitates formed during toxicity testing should be discussed. I agree that lab-based toxicity testing is the best we can do for now but longer term field-based aluminum toxicity tests could be very useful. (such as section 4.1.3).	Additional discussion of uncertainties regarding the criteria and laboratory to field exposures was added.	Section 5.3 (formerly Section 2.9)
Reviewer 4	ll) Table 3 and elsewhere when SMAV values (or other aluminum values are referred to) it should be made clear if this is referring to total or dissolved aluminum. I think it is pretty much always total aluminum but readers may want to just skip to the tables so this should be made clear in the table footnotes. Figure 4 does specifically mention that it is total aluminum.	Tables, figures and text were edited to provide additional clarity that the criteria are based on total aluminum concentrations.	Various locations
Reviewer 4	mm) Section 4.1.3 ends by stating that the study exceeded the minimum acceptable guidelines for tests of this duration. It is not clear what this means.	Sentence was edited for clarity to: "However, mortality of control fish in both the *in-situ* and laboratory exposures exceeded the minimum 80 percent survival acceptable guideline for tests of this duration."	Section 5.1.3 (formerly Section 4.1.3)
Reviewer 4	nn) Section 4.1.3 are there any field based fish studies? I know that aluminum mobilized by environmental acidification was implicated in fish kills. Are there any systematic studies in this regard? The authors repeat other acidic test results (top of page 50 for example).	Additional fish studies were added.	Section 5.1.3 (formerly Section 4.1.3)
Reviewer 4	oo) Throughout section 4 aluminum concentrations are often referred to (i.e., effects concentrations). Are these always total aluminum values? This applies to the summary tables at the end of the document as well – maybe some blanket statement would help avoid any ambiguity here. I am not sure if all accepted papers measured total aluminum so the authors need to clarify this.	Tables, figures and text were edited to provide additional clarity that the criteria are based on total aluminum concentrations.	Various locations
Reviewer 4	pp) Page 59 refers to the appendix by name (Appendix J reference). This is a nice way to do this and would make the rest of the document more readable if this convention was adopted throughout.	Thank you for your suggestion. The title of each appendix was added where appropriate.	Various locations

Reviewer	Comments	EPA Response	Revision Location in 2017 Draft Aluminum AWQC
Reviewer 4	qq) Page 62 says that MLR accounts for the bioavailability of the metal. It is semantics but I think it is overstating things to state that MLR account for bioavailability. There is no mechanism in an MLR.	Text was added to clarify that the MLR models are not mechanistic bioavailability models, like the BLM. The authors (DeForest et al. 2017) refer to the MLR model as a bioavailability model.	Section 2.7.1 (formerly 2.4.1)

3   REFERENCES CITED BY REVIEWERS AND EPA RESPONSES

Allin, C.J and R.W. Wilson. 2000. Effects of pre-acclimation to aluminium on the physiology and swimming behaviour of juvenile rainbow trout (*Oncorhynchus mykiss*) during a pulsed exposure. Aquat. Toxicol. 51(2): 213-224.

Angel, B.M., S.C. Apte, G.E. Batley and L.A. Golding. 2015. Geochemical controls on aluminium concentrations in coastal waters. Environ. Chem. 13(1): 111-118. http://dx.doi.org/10.1071/EN15029.

Clements, W.H. and P.M. Kiffney. 1996. Validation of whole effluent toxicity tests: Integrated studies using field assessments, microcosms, and mesocosms. In: D.L. Grothe, K.L. Dickson and D.K. Reed-Judkins (*Eds.*). *Whole effluent toxicity testing: an evaluation of methods and prediction of receiving system impacts*. Pensacola, FL., Society of Environmental Toxicology and Chemistry (SETAC), p. 229-244.

Clements, W.H., D.M. Carlisle, L.A. Courtney and E.A. Harrahy. 2002. Integrating observational and experimental approaches to demonstrate causation in stream biomonitoring studies. Environ. Toxicol. Chem. 21(6): 1138-1146.

DeForest, D.K., K.V. Brix, L.M. Tear and W.J. Adams. 2017 (Manuscript). Multiple Linear Regression (MLR) models for predicting chronic aluminum toxicity to freshwater aquatic organisms and developing water quality guidelines. Environ. Toxicol. Chem. (submitted).

Greenberg, M.S., I. Schoeters, R.S. Wentsel, D.W. Charters, I.A. Mitchell and B. Zajdlik. 2014. Regulatory considerations for the potential development and application of metal clean-up values. Integr. Environ. Assess. Manag. 10: 401-414.

Norberg-King, T. and D. Mount. 1986. Validity of effluent and ambient toxicity tests for predicting biological impact, Skeleton Creek, Enid, Oklahoma. Environmental Research Laboratory, Office of Research and Development, U.S. EPA. Duluth, MN. EPA/600/8-86/002.

Stephan, C.E., D.I. Mount, D.J. Hansen, J.H. Gentile, G.A. Chapman and W.A. Brungs. 1985. Guidelines for deriving numerical national water quality criteria for the protection of aquatic organisms and their uses. PB85-227040. National Technical Information Service, Springfield, WA. Available online at: https://www.epa.gov/sites/production/files/2016-02/documents/guidelines-water-quality-criteria.pdf.

Tietge, J.E., R.D. Johnsona and H.L. Bergman. 1988. Morphometric changes in gill secondary lamellae of brook trout (*Salvelinus fontinalis*) after long-term exposure to acid and aluminum. Can. J. Fish. Aquat. Sci. 45: 1643-1648.

U.S. EPA. 2007. Framework for metals risk assessment. EPA 120/R-07/001. March 2007. Office of the Science Advisor. U.S. Environmental Protection Agency. Washington, DC. 172 pp.

U.S. EPA. 1996. Method 1669: Sampling Ambient Water for Trace Metals at EPA Water Quality Criteria Levels. U.S. Environmental Protection Agency, Office of Water, Engineering and Analysis Division, Washington, D.C. 35 pp.

Wisconsin DNR. 1996. Analytical detection limit guidance & laboratory guide for determining method detection limits. Wisconsin Department of Natural Resources, Laboratory Certification Program, April 1996. PUBL-TS-056-96. 33 pp.