

1-8-10 Science Advisory Board (SAB) Ecological Processes and Effects Committee Advisory Report

-- Do Not Cite or Quote --

This Draft is made available for quality review and approval by the chartered Science Advisory Board. This Draft does not represent EPA policy.

DATE TO BE INSERTED

The Honorable Lisa P. Jackson
Administrator
U.S. Environmental Protection Agency
1200 Pennsylvania Avenue, N.W.
Washington, D.C. 20460

Subject: SAB Review of Empirical Approaches for Nutrient Criteria Derivation

Dear Administrator Jackson:

EPA's Office of Water (OW) requested that the Science Advisory Board (SAB) review the Agency's draft guidance document titled *Empirical Approaches for Nutrient Criteria Derivation* ("Guidance"). The Guidance is one of a series of technical documents developed by OW to describe approaches and methods for developing numeric criteria for nutrients. The Guidance specifically focuses on empirical approaches for determining stressor-response relationships to derive numeric nutrient criteria. In response to the Agency's advisory request, the Science Advisory Board Ecological Processes and Effects Committee, augmented with additional experts, met on September 9 - 11, 2009 to conduct a peer review of the Guidance. OW requested that the SAB: 1) comment on the technical merit of the methods and approaches described in the Guidance; 2) suggest approaches that might be considered to improve the Guidance; and 3) offer suggestions to improve the utility of the Guidance for state and tribal water quality scientists and resource managers. The enclosed advisory report provides the advice and recommendations of the Committee.

The SAB commends EPA for addressing nutrient issues. Nutrients (nitrogen and phosphorus) are a major cause of impairment in the quality of the Nation's waters, and the SAB recognizes the importance of EPA's efforts to develop numeric nutrient criteria. The stressor-response approach is a legitimate, scientifically based method for developing numeric nutrient criteria if the approach is appropriately applied (i.e., not used in isolation but as part of a weight-of-evidence approach). We encourage the Agency to continue this important work. EPA's draft Guidance provides a primer on a limited set of statistical methods that could be used in deriving nutrient criteria based on stressor-response relationships. However, in its present form, the Guidance does not present a complete or balanced view of using the statistical methods to develop criteria. Improvements in the Guidance are needed prior to its release to make the document more useful to state and tribal water quality scientists and resource managers.

In general, we find that the scope and intended use of the Guidance should be more clearly identified. The empirical stressor-response framework described in the Guidance is one possible approach for deriving numeric nutrient criteria, but the uncertainty associated with estimated stressor-response relationships would be problematic if this approach were used as a "stand alone" method because statistical associations do not prove cause and effect. We therefore

1-8-10 Science Advisory Board (SAB) Ecological Processes and Effects Committee Advisory Report

-- Do Not Cite or Quote --

This Draft is made available for quality review and approval by the chartered Science Advisory Board. This Draft does not represent EPA policy.

1 recommend that the stressor-response approach be used with other available methodologies in
2 the context of a tiered approach where uncertainties in different approaches are recognized, and
3 weight-of-evidence is used to establish the likelihood of causal relationships between nutrients
4 and their effects for criteria derivation. In this regard, we recommend that EPA more clearly
5 articulate how this particular guidance fits within the Agency's decision-making and regulatory
6 processes and, specifically, how it relates to and complements EPA's other nutrient criteria
7 approaches, technical guidance manuals, and documents. The SAB also recognizes that methods
8 in the Guidance do not address downstream impacts of excess nutrients.

9
10 The SAB has provided many recommendations to improve the Guidance and strongly
11 recommends that they be incorporated into the final document. These recommendations focus
12 on revising the document to address: cause and effect; the utility and limitations of the statistical
13 methods and approaches in the document; the supporting analyses and data needed to correctly
14 identify predictive relationships; the need for more guidance and examples to describe when and
15 how to use various methods and approaches; linkages among designated uses and stressors; and
16 the need for a more specific and descriptive framework outlining the steps in the criteria
17 development process. Finally, the SAB strongly recommends that EPA invest in providing the
18 technical support and training needed to make the approaches and methods in the final Guidance
19 more useful to state and tribal water resource managers.

20
21 Thank you for the opportunity to review this important guidance document. The SAB looks
22 forward to receiving the Agency's response to this advisory report and stands ready to provide
23 additional advice as EPA continues to develop nutrient criteria guidance.

24
25
26 Sincerely,
27
28
29

30
31 Dr. Deborah L. Swackhamer, Chair
32 Science Advisory Board
33

Dr. Judith L. Meyer, Chair
Ecological Processes and Effects Committee

1-8-10 Science Advisory Board (SAB) Ecological Processes and Effects Committee Advisory Report

-- Do Not Cite or Quote --

This Draft is made available for quality review and approval by the chartered Science Advisory Board. This Draft does not represent EPA policy.

NOTICE

This report has been written as part of the activities of the EPA Science Advisory Board, a public advisory group providing extramural scientific information and advice to the Administrator and other officials of the Environmental Protection Agency. The Board is structured to provide balanced, expert assessment of scientific matters related to the problems facing the Agency. This report has not been reviewed for approval by the Agency and, hence, the contents of this report do not necessarily represent the views and policies of the Environmental Protection Agency, nor of other agencies in the Executive Branch of the Federal government, nor does mention of trade names or commercial products constitute a recommendation for use. Reports of the EPA Science Advisory Board are posted on the EPA website at <http://www.epa.gov/sab>.

**U.S. Environmental Protection Agency
Science Advisory Board
Ecological Processes and Effects Committee (FY 2009) Augmented for Review
of Nutrient Criteria Guidance**

CHAIR

Dr. Judith L. Meyer, Distinguished Research Professor Emeritus, University of Georgia, Lopez Island, WA

MEMBERS

Dr. Richelle Allen-King, Professor and Chair, Department of Geology, University at Buffalo, Buffalo, NY

Dr. Ernest F. Benfield, Professor of Ecology, Department of Biological Sciences, Virginia Tech, Blacksburg, VA

Dr. G. Allen Burton, Professor and Director, Cooperative Institute for Limnology and Ecosystems Research, School of Natural Resources and Environment, University of Michigan, Ann Arbor, MI

Dr. Peter M. Chapman, Principal and Senior Environmental Scientist, Environmental Sciences Group, Golder Associates Ltd, Burnaby, BC, Canada

Dr. Loveday Conquest, Professor, School of Aquatic and Fishery Sciences, University of Washington, Seattle, WA

Dr. Wayne Landis, Professor and Director, Department of Environmental Toxicology, Institute of Environmental Toxicology, Huxley College of the Environment, Western Washington University, Bellingham, WA

Dr. James Oris, Professor, Department of Zoology, Miami University, Oxford, OH

Dr. Amanda Rodewald, Associate Professor of Wildlife Ecology, School of Environment and Natural Resources, The Ohio State University, Columbus, OH

Dr. James Sanders, Director and Professor, Skidaway Institute of Oceanography, Savannah, GA

Mr. Timothy Thompson, Senior Environmental Scientist, Science and Engineering for the Environment, LLC, Seattle, WA

1-8-10 Science Advisory Board (SAB) Ecological Processes and Effects Committee Advisory Report

-- Do Not Cite or Quote --

This Draft is made available for quality review and approval by the chartered Science Advisory Board. This Draft does not represent EPA policy.

CONSULTANTS

Dr. Victor Bierman, Senior Scientist, LimnoTech, Oak Ridge, NC

Dr. Elizabeth Boyer, Associate Professor, School of Forest Resources and Assistant Director, Pennsylvania State Institutes of Energy & the Environment, and Director, Pennsylvania Water Resources Research Center, Pennsylvania State University, University Park, PA

Dr. Mark David, Professor, Natural Resources & Environmental Sciences, University of Illinois, Urbana, IL

Dr. Douglas McLaughlin, Principal Research Scientist, National Council for Air and Stream Improvement, Inc., Western Michigan University, Kalamazoo, MI

Dr. Patrick J. Mulholland, Distinguished Research Staff Member, Carbon & Nutrient Biogeochemistry Group, Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN

Dr. Andrew N. Sharpley, Professor, Department of Crop, Soil and Environmental Sciences, Division of Agriculture, University of Arkansas, Fayetteville, AR

SCIENCE ADVISORY BOARD STAFF

Dr. Thomas Armitage, Designated Federal Officer, U.S. Environmental Protection Agency, Washington, DC

1 **U.S. Environmental Protection Agency**
2 **Science Advisory Board**
3 **BOARD**
4
5

6 **CHAIR**

7 **Dr. Deborah L. Swackhamer**, Professor and Charles M. Denny, Jr., Chair in Science,
8 Technology and Public Policy and Co-Director of the Water Resources Center, Hubert H.
9 Humphrey Institute of Public Affairs, University of Minnesota, St. Paul, MN
10

11
12 **SAB MEMBERS**

13 **Dr. David T. Allen**, Professor, Department of Chemical Engineering, University of Texas,
14 Austin, TX
15

16 **Dr. Claudia Benitez-Nelson**, Associate Professor, Department of Earth and Ocean Sciences and
17 Marine Science Program, University of South Carolina, Columbia, SC
18

19 **Dr. Timothy Buckley**, Associate Professor and Chair, Division of Environmental Health
20 Sciences, College of Public Health, The Ohio State University, Columbus, OH
21

22 **Dr. Thomas Burke**, Professor, Department of Health Policy and Management, Johns Hopkins
23 Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD
24

25 **Dr. Deborah Cory-Slechta**, Professor, Department of Environmental Medicine, School of
26 Medicine and Dentistry, University of Rochester, Rochester, NY
27

28 **Dr. Terry Daniel**, Professor of Psychology and Natural Resources, Department of Psychology,
29 School of Natural Resources, University of Arizona, Tucson, AZ
30

31 **Dr. George Daston**, Victor Mills Society Research Fellow, Product Safety and Regulatory
32 Affairs, Procter & Gamble, Cincinnati, OH
33

34 **Dr. Costel Denson**, Managing Member, Costech Technologies, LLC, Newark, DE
35

36 **Dr. Otto C. Doering III**, Professor, Department of Agricultural Economics, Purdue University,
37 W. Lafayette, IN
38

39 **Dr. David A. Dzombak**, Walter J. Blenko Sr. Professor, Department of Civil and Environmental
40 Engineering, College of Engineering, Carnegie Mellon University, Pittsburgh, PA
41

42 **Dr. T. Taylor Eighmy**, Vice President for Research, Office of the Vice President for Research,
43 Texas Tech University, Lubbock, TX
44

1-8-10 Science Advisory Board (SAB) Ecological Processes and Effects Committee Advisory Report

-- Do Not Cite or Quote --

This Draft is made available for quality review and approval by the chartered Science Advisory Board. This Draft does not represent EPA policy.

Dr. Elaine Faustman, Professor, Department of Environmental and Occupational Health Sciences, School of Public Health and Community Medicine, University of Washington, Seattle, WA

Dr. John P. Giesy, Professor and Canada Research Chair, Veterinary Biomedical Sciences and Toxicology Centre, University of Saskatchewan, Saskatoon, Saskatchewan, Canada

Dr. Jeffrey Griffiths, Associate Professor, Department of Public Health and Community Medicine, School of Medicine, Tufts University, Boston, MA

Dr. James K. Hammitt, Professor, Center for Risk Analysis, Harvard University, Boston, MA

Dr. Rogene Henderson, Senior Scientist Emeritus, Lovelace Respiratory Research Institute, Albuquerque, NM

Dr. Bernd Kahn, Professor Emeritus and Associate Director, Environmental Radiation Center, School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA

Dr. Agnes Kane, Professor and Chair, Department of Pathology and Laboratory Medicine, Brown University, Providence, RI

Dr. Nancy K. Kim, Senior Executive, New York State Department of Health, Troy, NY

Dr. Catherine Kling, Professor, Department of Economics, Iowa State University, Ames, IA

Dr. Kai Lee, Program Officer, Conservation and Science Program, David & Lucile Packard Foundation, Los Altos, CA

Dr. Cecil Lue-Hing, President, Cecil Lue-Hing & Assoc. Inc., Burr Ridge, IL

Dr. Floyd Malveaux, Executive Director, Merck Childhood Asthma Network, Inc., Washington, DC

Dr. Lee D. McMullen, Water Resources Practice Leader, Snyder & Associates, Inc., Ankeny, IA

Dr. Judith L. Meyer, Distinguished Research Professor Emeritus, Odum School of Ecology, University of Georgia, Lopez Island, WA

Dr. Jana Milford, Professor, Department of Mechanical Engineering, University of Colorado, Boulder, CO

Dr. Christine Moe, Eugene J. Gangarosa Professor, Hubert Department of Global Health, Rollins School of Public Health, Emory University, Atlanta, GA

1-8-10 Science Advisory Board (SAB) Ecological Processes and Effects Committee Advisory Report

-- Do Not Cite or Quote --

This Draft is made available for quality review and approval by the chartered Science Advisory Board. This Draft does not represent EPA policy.

Dr. Eileen Murphy, Manager, Division of Water Supply, New Jersey Department of Environmental Protection, Trenton, NJ

Dr. Duncan Patten, Research Professor, Department of Land Resources and Environmental Sciences, Montana State University, Bozeman, MT

Dr. Stephen Polasky, Fesler-Lampert Professor of Ecological/Environmental Economics, Department of Applied Economics, University of Minnesota, St. Paul, MN

Dr. Stephen M. Roberts, Professor, Department of Physiological Sciences, Director, Center for Environmental and Human Toxicology, University of Florida, Gainesville, FL

Dr. Amanda Rodewald, Associate Professor, School of Environment and Natural Resources, The Ohio State University, Columbus, OH

Dr. Joan B. Rose, Professor and Homer Nowlin Chair for Water Research, Department of Fisheries and Wildlife, Michigan State University, East Lansing, MI

Dr. Jonathan M. Samet, Professor and Flora L. Thornton Chair, Department of Preventive Medicine, University of Southern California, Los Angeles, CA

Dr. James Sanders, Director and Professor, Skidaway Institute of Oceanography, Savannah, GA

Dr. Jerald Schnoor, Allen S. Henry Chair Professor, Department of Civil and Environmental Engineering, Co-Director, Center for Global and Regional Environmental Research, University of Iowa, Iowa City, IA

Dr. Kathleen Segerson, Professor, Department of Economics, University of Connecticut, Storrs, CT

Dr. V. Kerry Smith, W.P. Carey Professor of Economics, Department of Economics, W.P. Carey School of Business, Arizona State University, Tempe, AZ

Dr. Herman Taylor, Professor, School of Medicine, University of Mississippi Medical Center, Jackson, MS

Dr. Barton H. (Buzz) Thompson, Jr., Robert E. Paradise Professor of Natural Resources Law at the Stanford Law School and Perry L. McCarty Director, Woods Institute for the Environment, Stanford University, Stanford, CA

Dr. Paige Tolbert, Associate Professor, Department of Environmental and Occupational Health, Rollins School of Public Health, Emory University, Atlanta, GA

1-8-10 Science Advisory Board (SAB) Ecological Processes and Effects Committee Advisory Report

-- Do Not Cite or Quote --

This Draft is made available for quality review and approval by the chartered Science Advisory Board. This Draft does not represent EPA policy.

Dr. Thomas S. Wallsten, Professor and Chair, Department of Psychology, University of Maryland, College Park, MD

Dr. Robert Watts, Professor of Mechanical Engineering Emeritus, Tulane University, Annapolis, MD

SCIENCE ADVISORY BOARD STAFF

Dr. Angela Nugent, Designated Federal Officer, U.S. Environmental Protection Agency, Washington, DC

TABLE OF CONTENTS

1. EXECUTIVE SUMMARY	xi
2. INTRODUCTION.....	1
3. RESPONSE TO CHARGE QUESTIONS.....	3
3.1. Charge Question 1. Improving the utility of the Guidance.....	3
3.2. Charge Question 2. Selecting stressor and response variables	9
3.3. Charge Question 3. Approaches to demonstrate the distribution of and relationships among variables.....	14
3.4. Charge Question 4. Methods for assessing the strength of the cause-effect relationship	18
3.5. Charge Question 5. Statistical methods to analyze the data	21
3.6. Charge Question 6. Evaluating the predictive accuracy of stressor-response relationships	30
3.7. Charge Question 7. Evaluating candidate stressor-response criteria	36
4. REFERENCES.....	42

1. EXECUTIVE SUMMARY

EPA's Office of Water (OW) requested that the Science Advisory Board (SAB) conduct a peer review of Agency's draft guidance document, *Empirical Approaches for Nutrient Criteria Derivation* (the "Guidance"). The Guidance was developed by OW to provide information for state and tribal water resource managers on empirical stressor-response approaches for developing numeric nutrient criteria. In response to the Agency's advisory request, the SAB Ecological Processes and Effects Committee reviewed the Guidance. To augment the expertise on the Committee for this advisory activity, several additional scientists with specific knowledge and expertise in assessing the effects of nutrient enrichment in aquatic systems also participated in the review.

EPA's Office of Water develops ambient water quality criteria that serve as guidance to states and tribes for adoption of water quality standards. The water quality standards include designated uses, such as aquatic life protection and recreation, and criteria that define levels of water quality variables protective of the designated uses. Because nutrients (nitrogen and phosphorus) are a major cause of impairment in the quality of the Nation's waters, state adoption of numeric nutrient criteria in water quality standards has been a high priority for OW. To assist the states and tribes in developing numeric nutrient criteria, OW published technical guidance manuals for developing nutrient criteria for lakes and reservoirs (U.S. EPA, 2000a), rivers and streams (U.S. EPA 2000b), estuaries and coastal marine waters (U.S. EPA, 2001), and wetlands (U.S. EPA, 2008). These technical guidance manuals focus primarily on describing a reference condition approach for deriving criteria from distributions of nutrient concentrations and biological responses in minimally disturbed reference waterbodies. Other basic analytical approaches for nutrient criteria derivation recognized in the manuals include mechanistic modeling (i.e., predicting the effects of changes in nutrient concentrations using site-specific parameters and equations that represent ecological processes), which EPA intends as the subject of a later document, the stressor-response approach (discussed in the Guidance and considered in this advisory report), and the application and/or modification of established nutrient/algal thresholds. The stressor-response approach involves quantifying the relationship between nutrient concentrations and biological response measures related to the designated use of a waterbody.

The Guidance outlines a five-step process for developing numeric nutrient criteria. It describes data analysis methods and approaches that could be used in each of these steps. Step one involves the use of exploratory analysis and data visualization tools to select variables that appropriately quantify the stressor (i.e., excess nutrients) and the response. Step two involves the use of conceptual models, existing literature, and other methods to assess the strength of the relationship represented in the stressor-response linkage. Step three involves the use of various statistical methods to analyze data, estimate stressor-response relationships, and identify thresholds that may be used to derive water quality criteria. Step four involves the evaluation of estimated stressor-response relationships (including validation of predictive performance for a stressor-response model, and selecting a model that best represents the data). Step five involves evaluating candidate nutrient criteria by predicting conditions that might be expected after implementing different criteria. The Guidance contains five sections, each addressing one of the

1 proposed steps in the criteria development process. In its charge to the SAB, EPA requested that
2 the Committee comment on the methods and approaches described in each section of the
3 Guidance, recommend other approaches that might be considered, and offer suggestions to
4 improve the utility of the Guidance for state and tribal water quality scientists and resource
5 managers. In its responses to the charge questions, the Committee provides comments and
6 recommendations to improve the Guidance and assist EPA in its efforts to support the
7 development of numeric nutrient criteria.

8
9 *General comments on the Guidance*

10
11 The Committee recognizes the importance of EPA's efforts to support numeric nutrient
12 criteria development and encourages the Agency to continue this important work. In addition,
13 we recognize the stressor-response approach as a legitimate, scientifically based method for
14 developing numeric nutrient criteria if it is appropriately applied (i.e., not used in isolation but as
15 part of a tiered weight-of-evidence approach as discussed here). The draft Guidance provides a
16 primer on a limited set of statistical methods that could be used in deriving numeric nutrient
17 criteria. However, we find that improvements in the Guidance are needed prior to its release to
18 make the document more useful to state and tribal water quality scientists and resource
19 managers.

20
21 In general, we find that the scope, limitations, and intended use of the Guidance should be
22 more clearly described. The Guidance addresses only one type of "empirical" approach for
23 derivation of numeric nutrient criteria (i.e., the stressor-response framework). As illustrated in
24 many of the examples in the Guidance, considerable unexplained variation can be encountered
25 when attempting to use empirical stressor-response approaches to develop nutrient criteria. The
26 final Guidance should clearly indicate that such unexplained variation presents significant
27 problems in the use of this approach. Further, the final document should clearly state that
28 statistical associations may not be biologically relevant and do not prove cause and effect.
29 However, when properly developed, biologically relevant statistical associations can be useful
30 arguments as part of a weight-of-evidence approach (further discussed in Section 3.3,
31 recommendation #7 of this advisory report) to criteria derivation. Therefore, the final Guidance
32 should provide more information on the supporting analyses needed to improve the basis for
33 conclusions that specific stressor-response associations can predict nutrient responses with an
34 acceptable degree of uncertainty. Such predictive relationships can then be used with
35 mechanistic or other approaches in a tiered weight-of-evidence assessment including cause and
36 effect relationships to develop nutrient criteria.

37
38 Tiered environmental assessment is iterative. The initial assessment is the simplest (e.g.,
39 minimal ecosystem specific data) and most conservative, and thus, will not always provide
40 sufficient certainty for decision-making. Cause and effect relationships would be inferred but
41 not demonstrated; only a few lines of evidence would be available and the corresponding
42 uncertainty high. At the highest tier, there would be several lines of evidence and factors that
43 would confound the prediction of effects, such as other stressors or the morphology of the
44 waterbody, and these need to be understood and considered. Successive tiers involve more
45 focused (e.g., specific for particular ecosystem types) investigations, based on the results of the

1 previous tier. Data needs are relatively low at the initial tier, but increase at successive tiers;
2 however, uncertainty also reduces at successive tiers, and sources of uncertainty become better
3 understood. It is important to explicitly describe and consider uncertainty in the criteria
4 development and decision-making process. Weight-of-evidence typically determines the tier at
5 which uncertainty has been reduced sufficiently for informed management decision-making.
6

7 The Committee also recommends that EPA more clearly articulate how the Guidance fits
8 within the Agency's decision-making and regulatory processes and, specifically, how it relates to
9 and complements EPA's other nutrient criteria technical guidance manuals and documents. As
10 further discussed in the response to Charge Question 1, numeric nutrient concentration criteria
11 and load-response models should be considered as two different approaches for accomplishing
12 the goal of controlling excessive nutrient loadings. In addition, the Committee notes that the
13 methods in the Guidance do not address the problem of excess nutrient enrichment downstream
14 from waters for which the criteria are being developed. There is a need for methods to address
15 this problem (one of which could be load-response modeling) and it should be clearly stated that
16 this is beyond the scope of the current guidance document.
17

18 *Charge Question 1. Improving the utility of the Guidance for state and tribal water quality*
19 *scientists and resource managers*
20

21 *What suggestions do you have that will improve the utility of the draft document, Empirical*
22 *Approaches for Nutrient Criteria Derivation, for State water quality scientists and resource*
23 *managers to derive numeric nutrient criteria based on stressor-response relationships?*
24

25 The Committee finds that improvements in EPA's Guidance are needed to make the
26 document more useful to state and tribal water quality scientists and resource managers and to
27 ensure against inadvertent misuse. In this regard, as previously mentioned, the scope,
28 limitations, and intended use of the document should be more clearly identified.
29

- 30 • In the Guidance, and the Agency's related technical manuals, EPA should more clearly
31 address the importance of: 1) establishing linkages among designated uses and measured
32 responses, stressors and measures of stressors; and 2) relating measures of stressors directly
33 to deleterious effects on designated uses.
34
- 35 • The Committee finds that the Guidance: 1) should provide a more specific and descriptive
36 framework outlining the steps in the criteria development process. Figure 1 of this advisory
37 report illustrates EPA's proposed framework for developing nutrient criteria and the SAB
38 recommendations for revision of the framework; 2) must be detailed and sophisticated
39 enough to ensure statistical rigor, but additional support must also be provided by EPA to
40 help users meet the technical demands of the methods; 3) should more clearly express the
41 caveats and limitations of the statistical methods and approaches in the document, in
42 particular the fact that statistical correlations do not establish cause and effect; 4) should
43 contain more technical guidance and examples to describe when and how to use various
44 methods and approaches; and 5) should provide additional guidance on data requirements for
45 application of the statistical methods and approaches.

- 1
- 2 • *Charge Question 2. Selecting stressor and response variables*
- 3

4 *Section 1 of the draft guidance document reviews how to select the variables that appropriately*
5 *quantify the stressor (i.e., excess nutrients) and the response (e.g., chlorophyll a, dissolved*
6 *oxygen, or a biological index). Please comment on whether the factors to consider described in*
7 *Section 1 of the draft document are appropriate for selecting response variables that are*
8 *sensitive to nutrients and related to measures of designated uses.*
9

10 In Section 1 of the Guidance, EPA discusses factors to consider when selecting the stressor
11 and response variables. The Committee finds that EPA should strengthen the Guidance by
12 including additional material.
13

- 14 • The examples in the Guidance rely heavily on taxa richness as a response variable. Some
15 rationale as to how this variable relates to a designated use should accompany these
16 examples. The coupling of response variables to designated uses must be clear and the
17 rationale explained. Further, the Guidance could be strengthened considerably by
18 presentation of examples showing strong nutrient-response relationships with response
19 variables that are clearly linked to designated uses.
20
- 21 • The Committee notes that co-limitation by both nitrogen and phosphorus may be common in
22 many systems and regions. Therefore, the use of multivariate or data stratification
23 approaches may be needed to identify nutrient-response relationships.
24
25

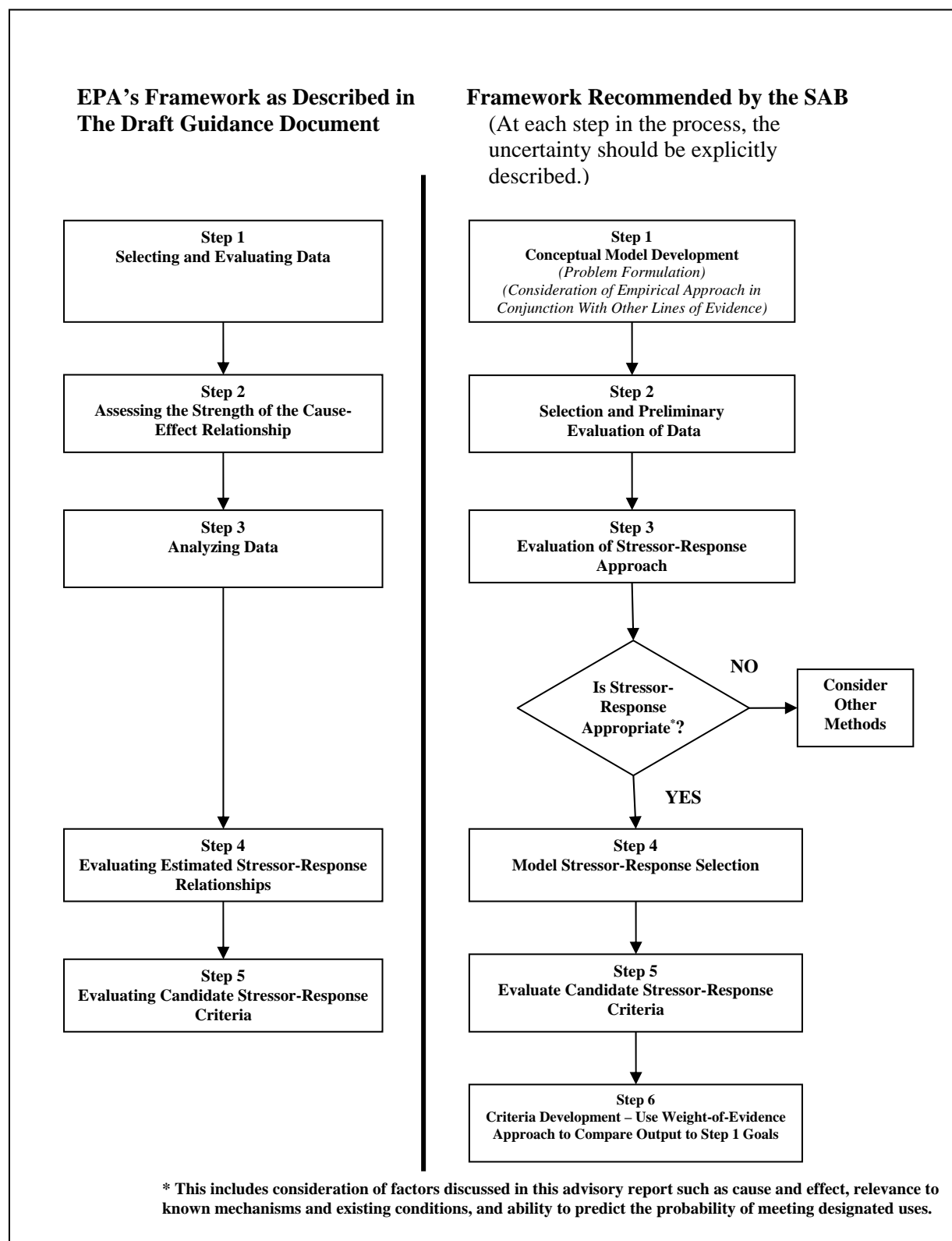


Figure 1. EPA's Framework for Developing Nutrient Criteria Based on Stressor-Response Relationships and SAB Recommendations for Revision

- The Guidance should provide more information on the data needed to characterize other stressor and constraint variables (e.g., high dissolved organic carbon versus low dissolved organic carbon lakes, shaded versus unshaded streams) which are critical for applying multivariate techniques or for stratification/classification of univariate nutrient-response relationships.
- The Guidance focuses on total nitrogen and total phosphorus as the primary nutrient stressor variables. Additional consideration should be given to inorganic forms of these nutrients because these forms are the most immediately biologically available.
- The Guidance focuses on nutrient-response pathways driven by autotrophic processes (nutrients directly control algal growth and excessive amounts of algae impair systems through indirect effects on dissolved oxygen, food web changes, and aesthetics). The Committee notes that nutrients can also directly control heterotrophic microbes (bacteria and fungi) and indirectly control decomposition of organic matter. This should be more fully discussed in the Guidance.
- The Guidance provides inadequate discussion of the temporal/spatial aspects of data needed to develop relevant stressor-response relationships. The Guidance should discuss the conditions under which mean/median or maximum/minimum values of stressor and response variables might be more appropriate than discrete instantaneous measurements for developing stressor-response relationships. The use of time series data to describe specific systems should also be addressed. Although such guidance may be provided in various system-specific technical manuals (e.g., U.S. EPA 2000a, b), a summary synthesis of the major points in these earlier documents should be included in the Guidance.

Charge Question 3. Approaches to demonstrate the distribution of and relationships among variables

Section 1 outlines methods to visualize available data. Please comment on the effectiveness of the following approaches described in the document (listed below) to demonstrate the distribution of and relationships among variables.

- a) Basic data visualization techniques*
- b) Maps*
- c) Conditional probability*
- d) Classifications*

Section 1 of the Guidance discusses exploratory data analysis, and presents several methods for demonstrating the distribution of and relationships among variables. In Subsections 1.2 - 1.6 several basic plotting techniques are presented. This is followed by a description of conditional probability analysis (a statistical approach for summarizing how changes in nutrient concentrations are associated with the probability of waterbodies attaining their designated uses).

1-8-10 Science Advisory Board (SAB) Ecological Processes and Effects Committee Advisory Report

-- Do Not Cite or Quote --

This Draft is made available for quality review and approval by the chartered Science Advisory Board. This Draft does not represent EPA policy.

1 The Committee finds that the discussion of exploratory data analysis would be more effective if
2 Section 1 of the Guidance were reorganized and expanded.

- 3
- 4 • As further discussed in the response to Charge Question 3, Subsections 1.2 - 1.6 of the
5 Guidance should be reframed as a separate major section on exploratory data analysis, which
6 should follow another separate major section on problem formulation. The material in
7 Subsection 1.1 (selection of stressor-response variables) should be moved to a later section of
8 the document.
- 9
- 10 • The Guidance should stress that exploratory data analysis, including data visualization,
11 should be conducted prior to inferential statistical analyses of potential stressors and
12 responses. The objectives of exploratory data analysis should be to better understand the
13 system of interest and to maximize accuracy and minimize variability of subsequently
14 derived stressor-response relationships.
- 15
- 16 • Additional methods for exploratory data analysis should be discussed in the Guidance. These
17 additional methods should include: use of summary statistics, time series plots at fixed points
18 in space; longitudinal plots at fixed points in time; bubble plots; Pearson and other
19 correlation analyses; and maps that show temporal (monthly, seasonal, inter-annual) as well
20 as spatial patterns.
- 21
- 22 • Clear guidance is needed on when and how to use the statistical methods and visualization
23 techniques presented in the document. The strengths and limitations of the methods should
24 also be clearly identified. It would be useful to show several case examples that range from
25 state-wide to local and data-rich to data-poor; and exemplify different types of aquatic
26 ecosystems (e.g., headwater streams, large rivers, lakes and estuaries). Examples should note
27 the strengths, limitations, assumptions and uncertainties that must be considered when using
28 the methods to explore and visualize the data. These examples should demonstrate how
29 nutrients can be identified as significant stressors in the presence of multiple stressors and
30 habitat factors that may affect the resident communities.
- 31
- 32 • Subsection 1.6 of the Guidance (examination of stressor-response distributions across
33 different classes, e.g., ecoregions) should be expanded. The subsection should discuss
34 additional data analysis and examples for different spatial classifications (e.g., ecoregions,
35 states, watersheds, systems of interest), different waterbody types (e.g., streams, rivers, lakes,
36 estuaries) and other important physical and chemical characteristics of systems that could
37 affect the applicability of the nutrient criteria.
- 38

39 *Charge Question 4. Methods for assessing the strength of the cause-effect relationship*

40
41 *Section 2 of the draft guidance document describes methods for assessing the strength of the*
42 *cause-effect relationship represented in the stressor-response linkage. Please comment on*
43 *whether the draft guidance document adequately describes how conceptual models, existing*

1 *literature, and empirical models can be used to assess how changes in nutrient concentration are*
2 *likely to cause changes in the chosen response variable.*

3
4 Section 2 of the Guidance provides a summary of how the strength of candidate stressor-
5 response pairings from step 1 can be assessed. The Committee recommends a number of
6 improvements in this section.
7

- 8 • It is appropriate to use conceptual models and existing literature as the scientific basis to
9 assess how changes in nutrient concentrations might affect response variables. However
10 additional discussion of conceptual model selection, with specific examples, would be
11 helpful. As illustrated in Figure 1 of this advisory report and further discussed in the
12 response to Charge Question 7, the Committee recommends that development of the
13 conceptual model occur in the problem formulation step, early in the process of criteria
14 development.
15
- 16 • Structural Equation Modeling (SEM) is discussed in the Guidance as a method for exploring
17 nutrient-ecosystem response. The Committee finds that use of SEM should be more fully
18 explained. Clear examples of its use should be provided.
19
- 20 • The Guidance discusses the use of Propensity Score Analysis (PSA) to estimate stressor-
21 response relationships. PSA seems to be useful for sorting out groups that share covariates
22 but may have unique nutrient characteristics. Such sorting could lead to a clearer
23 understanding of how nutrients function amid multiple covariates. The example of PSA in
24 the Guidance appendix is helpful, but further explanation of how to interpret the results of the
25 analysis is needed. An analysis such as PSA should be discussed in a later section of the
26 Guidance because it is a tool for analyzing data (Section 3 of the Guidance) rather than
27 supporting potential relationships.
28
- 29 • It is not clear why EPA did not include information obtained from mechanistic models in
30 Section 2 of the Guidance. Because mechanistic models can integrate information on the
31 interactions of major ecosystem processes to derive quantitative estimates of effects, they
32 should be discussed as a means to interpret the stressor-response relationship.
33

34 *Charge Question 5. Statistical methods to analyze the data*
35

36 *Section 3 of the draft guidance document outlines statistical methods to analyze the data to*
37 *estimate stressor-response relationships. Please comment on the appropriateness of the methods*
38 *outlined in the document (listed below) for describing stressor-response relationships associated*
39 *with nutrient pollution. What approaches would you recommend that could effectively address*
40 *indirect pathways of adverse effects? What recommendations do you have to address the effects*
41 *of confounding variables and uncertainty in the estimated relationships?*
42

43 a) *Simple linear regression*

44 b) *Quantile regression*

- c) *Logistic regression*
- d) *Multiple linear regression*
- e) *Non-parametric changepoint analysis*
- f) *Discontinuous regression models*

Section 3 of the Guidance describes a number of statistical methods for analyzing data to estimate stressor-response relationships. The Committee provides comments addressing the appropriateness of statistical methods for estimating stressor-response relationships.

- Methods described in the Guidance are generally appropriate for estimating stressor-response relationships in support of conceptual models. However, as further discussed in the response to Charge Question 5, more careful consideration of confounding variables is necessary to maximize the potential for stressor-response relationships to reflect cause and effect between nutrient concentrations and ecological responses. The Guidance should be revised to state this more definitively and better assist its users in achieving this objective.
- Those charged with using stressor-response methodology may require additional technical support to use the methods in the Guidance.
- EPA should provide guidance regarding what the Agency would consider to be a sufficiently strong stressor-response relationship for use in criteria development, and how to interpret statistical significance with “low” R^2 values. The Committee also notes that uncertainty must be identified and quantified for all methods and at all stages of the process.

Charge Question 6. Evaluating the predictive accuracy of stressor-response relationships

Section 4 of the draft guidance document describes how to evaluate the predictive accuracy of estimated stressor-response relationships. Please comment on the appropriateness of approaches in Section 4 of the guidance document and factors to consider in evaluating and comparing different estimates of the stressor-response relationships and selecting those most appropriate for criteria derivation.

The Committee provides comments on the appropriateness of approaches discussed in Section 4 of the Guidance and the factors to consider in evaluating and comparing different estimates of stressor-response relationships in order to select those most appropriate for criteria development. Overall, the Committee finds that this section of the Guidance lacks the detail provided in other sections and needs improvement.

- A clear framework for statistical model selection is needed. This framework should include:
 - 1) an assessment of whether analyses indicate that the stressor-response approach is appropriate; 2) selection criteria to evaluate the capability of models to consider cause/effect and direct/indirect relationships between stressors and responses; 3) consideration of model relevance to known mechanisms and existing conditions; 4) establishment of biological relevance; and 5) ability to predict probability of meeting designated use categories.

- The concept of “validation” as presented in Subsection 4.1 of the Guidance is inconsistent with other EPA guidance (U.S. EPA, 2009b) on development, evaluation, and application of models. Model corroboration (sensu “validation”) and uncertainty analysis should both be part of model evaluation and selection. These activities should be directed and informed by pre-established data quality objectives. Additional guidance is also needed on: data set specification and stratification; a suite of validation techniques (e.g., random or non-random held-out data, independent data, resampling/Monte Carlo); and appropriate quantitative levels of goodness-of-fit and uncertainty measures.
- With regard to validation, the Committee recommends that nutrient criteria should result from a tiered weight-of-evidence approach based on the application of multiple empirical approaches and consideration of multiple response variables as appropriate. The nutrient criteria values that may be determined, after considering validation and uncertainty, may vary significantly from technique to technique or from response variable to response variable. EPA should provide greater guidance on how to assign numeric criteria when a range of responses among analyses/models results in different values.

Charge Question 7. Evaluating candidate stressor-response criteria

Section 5 of the draft guidance document describes how to evaluate the candidate stressor-response criteria. An approach is outlined for predicting conditions that might result after implementing different nutrient criteria. Please comment on uncertainties that would remain if water quality criteria for nutrients were based solely on estimated stressor-response relationships and in what ways other information/analysis would help address and possibly reduce this uncertainty?

Section 5 of the Guidance describes how to evaluate candidate numeric nutrient criteria. The Committee provides comments on uncertainties associated with deriving candidate water quality criteria. We also recommend improvements in the Guidance to help address and reduce uncertainty.

- The Guidance describes approaches that use a data-mining exercise to demonstrate a possible cause-effect relationship for the nutrient-ecosystem response. However, as further discussed in the response to Charge Question 7, the document does not address or partition the inherent critical uncertainties associated with the stressor-response approach. We note that these uncertainties can be extremely large (e.g., several orders of magnitude). To address these uncertainties, the Guidance should better document the physical, chemical and biological variables comprising the morphological relationships (e.g., habitat, spatial, and temporal) that define the aquatic system of interest, and which may be important in modifying the relationship between nutrient concentrations (both nitrogen and phosphorus) and observed endpoints. These factors may dominate the cause-effect pathway and should be documented so that uncertainty in the relationship between nutrient concentrations and measured endpoints can be reduced.

1-8-10 Science Advisory Board (SAB) Ecological Processes and Effects Committee Advisory Report

-- Do Not Cite or Quote --

This Draft is made available for quality review and approval by the chartered Science Advisory Board. This Draft does not represent EPA policy.

- 1 • The Guidance should indicate that, at the start of the initial problem formulation exercise, a
2 realistic cause-effect conceptual model must be developed, and that the model should include
3 those factors that are likely to contribute most to the change in the response variable for the
4 specific region/system of interest. Then data analyses can be used to evaluate which of the
5 factors, or combination of factors, caused the observed change in the response variable.
6
- 7 • There is considerable uncertainty in linkage of the response variables discussed in the
8 Guidance to the Clean Water Act goals of drinkable, swimmable, and fishable waters. The
9 recommended response variables in the Guidance should be directly linked to these Clean
10 Water Act goals.

2. INTRODUCTION

EPA's Office of Water (OW) requested that the Science Advisory Board (SAB) conduct a peer review of the Agency's draft guidance document, *Empirical Approaches for Nutrient Criteria Derivation* (the "Guidance"). The Guidance was developed by EPA's Office of Water to provide information for water resource managers on the scientific foundation for using empirical approaches to describe stressor-response relationships for developing numeric nutrient water quality criteria. The SAB Ecological Processes and Effects Committee (Committee) met on September 9th-11th, 2009 to review the Guidance. To augment the expertise on the Committee for this advisory activity, several additional scientists with specific knowledge and expertise in assessing the effects of nutrient enrichment in aquatic systems also participated in the review. This report transmits the advice of the Committee.

EPA's Office of Water is charged with protecting aquatic life, wildlife, and human health from adverse water-mediated effects of anthropogenic pollutants. In support of this mission, OW develops ambient water quality criteria that serve as guidance to states and tribes for adoption of water quality standards. State and tribal water quality standards include designated uses, such as aquatic life protection and recreation, and criteria that define levels of water quality variables protective of the designated uses. Because nutrients (nitrogen - N and phosphorus - P) are a major cause of water quality impairment in the Nation's waters, state adoption of numeric nutrient criteria into water quality standards has been a high priority for OW. The Office of Water has stated that numeric nutrient water quality standards are important because they can: support development of nutrient related Total Maximum Daily Loads (TMDLs); provide targets for nutrient trading programs; and make it easier to write National Pollutant Discharge Elimination System (NPDES) permits, evaluate the success of nutrient runoff minimization programs, and measure environmental progress.

To assist states and tribes in developing numeric nutrient criteria, OW published peer reviewed technical guidance for developing such criteria for lakes and reservoirs (U.S. EPA, 2000a), rivers and streams (U.S. EPA 2000b), estuaries and coastal marine waters (U.S. EPA, 2001), and wetlands (U.S. EPA, 2008). These technical guidance documents focus primarily on a reference condition approach for deriving nutrient criteria from distributions of nutrient concentrations and biological responses in minimally disturbed reference waterbodies. Other basic analytical approaches for nutrient criteria derivation identified in prior guidance documents include mechanistic modeling (i.e., predicting the effects of changes in nutrient concentrations using site-specific parameters and equations that represent ecological processes), the stressor-response approach, and the application and/or modification of established nutrient/algal thresholds. The stressor-response approach involves quantifying a relationship between nutrient concentrations and biological response measures related to the designated use of a waterbody. In the Guidance, EPA states that, when developing nutrient criteria, the strengths and characteristics of each analytical approach should be carefully considered with respect to data availability and designated use protection needs.

1 The Guidance outlines a five-step process for developing numeric nutrient criteria. Step one
2 involves selecting variables that appropriately quantify the stressor (i.e., excess nutrients) and the
3 response. The Guidance describes various techniques for exploratory data analysis to understand
4 the properties of different variables and visualize data. These techniques include histograms,
5 box and whisker plots, quantile-quantile plots, cumulative distribution plots, scatter diagrams,
6 and spatial mapping. Step two involves assessing the strength of the relationship represented in
7 the stressor-response linkage. The Guidance discusses the use of conceptual models, existing
8 literature, and empirical models to assess the degree to which changes in nutrient concentration
9 are likely to cause changes in a chosen response variable. Step three involves analysis of data to
10 estimate stressor-response relationships and identify thresholds that may be used to derive
11 criteria. The Guidance describes a number of statistical methods for analyzing data to estimate
12 stressor-response relationships. These methods include linear regression, logistic regression,
13 quantile regression, non parametric changepoint analysis, and discontinuous regression
14 modeling. Step four involves evaluating the stressor-response relationships (including validation
15 of predictive performance for a stressor-response model and selecting a model that best
16 represents the data). Step five involves evaluating candidate stressor-response criteria. The
17 Guidance outlines an approach for predicting conditions that might be expected after
18 implementing different nutrient criteria and selecting a value to optimize resource protection.
19 The Committee was asked to comment on the scientific and technical merit of the methods and
20 approaches discussed in the Guidance and to offer suggestions to improve the usefulness of the
21 document to state and tribal water quality scientists and resource managers.

22
23 The Committee recognizes the importance of EPA's efforts to support numeric nutrient
24 criteria development and encourages the Agency to continue this important work. In addition,
25 we recognize the stressor-response approach as a legitimate, scientifically based method for
26 developing numeric nutrient criteria if it is appropriately applied (i.e., not in isolation). The draft
27 Guidance provides a primer on a limited set of statistical methods that could be used in deriving
28 nutrient criteria based on stressor-response relationships. However, the Committee finds that
29 improvements in the Guidance are needed prior to its release to make the document more useful
30 to state and tribal water quality scientists and resource managers.

31
32 In general, we find that the scope, limitations, and intended use of the Guidance need to be
33 more clearly described. The Guidance addresses only one type of "empirical" approach for
34 derivation of numeric nutrient criteria (i.e., the stressor-response framework). In this regard, we
35 strongly recommend that EPA more clearly articulate how the Guidance fits within the decision-
36 making and regulatory processes and, specifically, how it relates to and complements EPA's
37 other nutrient criteria technical guidance manuals and documents. As illustrated in the data
38 analysis examples in the Guidance, a large degree of unexplained variation can be encountered
39 when attempting to use empirical stressor-response approaches to develop nutrient criteria. The
40 final Guidance should clearly indicate that such unexplained variation can present significant
41 problems in the use of this approach. Further, *the final document should clearly state that*
42 *statistical associations may not be biologically relevant and do not prove cause and effect.*
43 When properly developed, statistical associations can be useful in supporting cause and effect
44 arguments as part of a weight-of-evidence approach (further discussed in Section 3.3,
45 recommendation #7 of this advisory report) to criteria development. Therefore, the final

Guidance should provide more information on the supporting analyses needed to improve the basis for conclusions that specific stressor-response associations can predict nutrient responses with an acceptable degree of uncertainty. Such predictive relationships can then be used, with mechanistic or other approaches, in a tiered weight-of-evidence assessment to develop nutrient criteria. Tiered environmental assessment is iterative. The initial assessment is the simplest (e.g., minimal ecosystem specific data) and most conservative, and thus will not always provide sufficient certainty for decision-making. Cause and effect relationships would be inferred but not demonstrated; only a few lines of evidence would be available and the corresponding uncertainty high. At the highest tier, there would be several lines of evidence and factors that would confound the prediction of effects, such as other stressors or the morphology of the waterbody, and these need to be understood and considered. Successive tiers will involve more focused (e.g., specific for particular ecosystem types) investigations, based on the results of the previous tier. Data needs are relatively low at the initial tier, but increase at successive tiers; however, uncertainty also reduces at successive tiers, and sources of uncertainty become better understood. It is important to explicitly describe and consider uncertainty in the criteria development and decision-making process. Weight-of-evidence typically determines the tier at which uncertainty has been reduced sufficiently for informed management decision-making.

In our responses to the charge questions we have recommended specific revisions to improve various sections of the Guidance before it is published. These recommendations focus on: modifying the framework of the Guidance to make it more specific and descriptive (as illustrated in Figure 1 of this report); providing additional information on conditions under which the stressor-response framework may apply; more clearly expressing the caveats, limitations, and data requirements associated with the approaches presented in the Guidance; providing additional information and examples showing when and how to use methods and approaches described in the document; and providing more detailed and descriptive guidance on the use of statistical methods and additional support from EPA to help users meet the technical demands of the methods.

3. RESPONSE TO CHARGE QUESTIONS

In the responses to each of the charge questions, the Committee has listed key findings and comments as bullets. These findings are followed by the Committee's key recommendations.

3.1. Charge Question 1. Improving the utility of the Guidance

What suggestions do you have that will improve the utility of the draft document, *Empirical Approaches for Nutrient Criteria Derivation*, for State water quality scientists and resource managers to derive numeric nutrient criteria based on stressor-response relationships?

The Committee was asked to offer suggestions to improve the usefulness of the Guidance to state and tribal water quality scientists and resource managers for deriving numeric nutrient criteria based on stressor-response relationships. In this regard, we find that the following

1 improvements in EPA's Guidance are needed. A number of these findings are further discussed
2 in responses to the other charge questions.

3
4 *Findings concerning improving the utility of the Guidance*

- 5
- 6 • The scope, limitations, and intended use of the Guidance should be more clearly identified.
7 The Guidance addresses only one possible approach (i.e., the stressor-response framework)
8 for derivation of numeric nutrient criteria. The Guidance would be more useful if it: 1)
9 expanded upon the utility of the mechanistic modeling approach for understanding stressor-
10 response relationships and the reference condition approach for criteria derivation; 2) more
11 clearly articulated how it relates to EPA's other published nutrient criteria guidance; 3)
12 explained the linkages among designated uses, stressors, measures of stressors, and the
13 deleterious effects of stressors on designated uses; 4) explained that the Guidance does not
14 address "downstream" effects of nutrients; and 5) acknowledged other factors that have
15 appeared to limit state progress toward developing nutrient criteria (e.g., lack of data and
16 technical expertise, insufficient resources, or other factors).
17
 - 18 • Substantial revision of the document is needed to facilitate identification of the most
19 scientifically defensible approaches to deriving numeric nutrient criteria. The Committee
20 emphasizes that understanding the causative link between nutrient levels and impairment is
21 necessary in order to assure that managing for particular nutrient levels will lead to desired
22 outcomes. As further discussed below, the stressor-response framework in the Guidance may
23 often not be the most appropriate approach for deriving numeric nutrient criteria.
24
 - 25 • Substantial revision of the document is needed to increase its usability while reducing the
26 likelihood of misuse. The Committee finds that the Guidance would be more useful if it: 1)
27 provided a more specific and descriptive framework outlining the steps in the criteria
28 development process (a specific example is illustrated in Figure 1 of this report); 2)
29 contained more technical guidance and examples to describe when and how to use various
30 methods and approaches in the document and ensure statistical rigor (with additional support
31 provided from EPA to help users meet the technical demands of the methods); 3) more
32 clearly expressed the caveats and limitations of the statistical methods and approaches in the
33 document; and 4) provided additional guidance on data requirements for application of the
34 statistical methods and approaches.
35
 - 36 • The absence of a direct causative relationship between stressor and response is one of the
37 most serious issues raised by the Committee. Without a mechanistic understanding and a
38 clear causative link between nutrient levels and impairment, there is no assurance that
39 managing for particular nutrient levels will lead to the desired outcome. There are numerous
40 empirical examples where a given nutrient level is associated with a wide range of response
41 values due to the influence of habitat, light levels, grazer populations and other factors. If the
42 numeric criteria are not based upon well-established causative relationships, the scientific
43 basis of the water quality standards will be seriously undermined.
44

This Draft is made available for quality review and approval by the chartered Science Advisory Board. This Draft does not represent EPA policy.

- Numeric nutrient concentration criteria and load-response models should be considered as two different approaches for accomplishing the goal of controlling excessive nutrient loadings. EPA has put forth the reference condition approach, the empirical stressor-response approach, and mechanistic modeling as basic analytic approaches for development of numeric nutrient criteria. However, the way in which EPA used results from mechanistic models to develop nutrient load reduction goals for the Gulf of Mexico (Mississippi River/Gulf of Mexico Watershed Nutrient Task force, 2008), and the way in which it is currently using mechanistic models for nutrient and sediment TMDLs for Chesapeake Bay, does not involve development or use of numeric nutrient criteria. The reason is that these mechanistic models (Scavia et al., 2004; Cerco and Noel, 2004) are load-response models, not empirical stressor-response models, and hence they obviate the need for numeric nutrient criteria because they directly link nutrient loads to response variables that represent water quality impairments (e.g., dissolved oxygen, chlorophyll, water clarity and acreage of submerged aquatic vegetation). This reasoning applies not only to mechanistic models but can also apply to empirical models. Turner et al. (2008) and Hagy et al. (2004) developed empirical statistical models for hypoxia in the Gulf of Mexico and Chesapeake Bay, respectively. Both of these models are load-response models and neither involves numeric nutrient concentrations. Further support for this reasoning can be found in Carleton et al. (2005), an EPA study designed to demonstrate the use of mechanistic models to develop nutrient criteria. In fact, in the two examples presented in this study, mechanistic models were actually used as load-response models and not to develop ambient nutrient concentration criteria.

Key recommendations concerning identification of the scope, limitations, and intended use of the document

As a consequence of the Committee's discussion and the findings listed above, we provide the following recommendations for revising the Guidance

1. EPA should specify how the Guidance is to be used in combination with other EPA nutrient criteria technical guidance manuals. In the preamble, the Guidance should clearly state that the contents represent one of several possible approaches (i.e., the stressor-response framework in the Guidance, mechanistic modeling, reference condition, and the application and/or modification of established nutrient/algal thresholds) that should be considered when deriving numeric nutrient criteria, and expand upon the utility of considering all approaches in a tiered weight-of-evidence approach before deciding on a particular course of action. In this regard, the Guidance should indicate that numeric nutrient concentration criteria and load-response models should be considered as two different approaches for accomplishing the goal of controlling excessive nutrient loadings. To provide additional information on other approaches, EPA should consider appending to the document relevant portions from earlier guidance manuals.
2. EPA should more clearly articulate how the Guidance fits within the decision-making and regulatory processes and, specifically, how it relates to and complements EPA's nutrient criteria technical guidance manuals and other EPA technical documents. Outlining the

1 fundamental principles that underlie the use of stressor-response relationships and
2 providing background information on water quality impairments (e.g., causes and types
3 of impairments, types of designated uses) might provide a useful context. Including a
4 clearer description of how water-use designations influence the derivation of empirically-
5 derived nutrient criteria might be considered as well. Considering the number and
6 usefulness of other EPA-developed processes and recommendations, the authors should
7 consider how they might improve the integration of this document with other EPA
8 efforts. For example, the Guidance would benefit by incorporating the problem
9 formulation stage that is part of the Ecological Risk Assessment process (see Figure 1 of
10 this advisory report).

- 11
12 3. In the Guidance, EPA should address the importance of: 1) establishing linkages among
13 designated uses, measured responses, stressors, and measures of stressors; and 2) relating
14 measures of responses directly to deleterious effects on designated uses. We agree with
15 the statement in the Florida Department of Environmental Protection's letter of
16 September 4, 2009 (letter from Daryll Joyner, Florida Department of Environmental
17 Protection to Thomas Armitage, Designated Federal Officer, EPA Science Advisory
18 Board Staff Office) indicating that the "most scientifically defensible strategy for
19 managing nutrients within the range of uncertainty is to verify a biological response prior
20 to taking a management action." This risk/performance-based approach to setting
21 nutrient criteria is evident not only in Florida's program, but also in those developed by
22 California and Maine (Florida Department of Environmental Protection, 2009; Maine
23 Department of Department of Environmental Protection, 2009; McLaughlin and Sutula,
24 2007). Those risk-based linkages are not addressed in either the Guidance or EPA's
25 Nutrient Criteria Technical Guidance documents for Rivers (2000), Lakes/Reservoirs
26 (2000), and Estuaries (2001).
27
- 28 4. In the Guidance, EPA should emphasize that the document does not address downstream
29 effects of nutrient enrichment, which are intended to be the focus of a separate future
30 document. Load-response models may prove useful in addressing downstream effects.
31 The Committee has some reservations about addressing downstream effects in a separate
32 document because fragmentation of the guidance documents will increase the likelihood
33 that each will be used in isolation and potentially provide misleading results.
34
- 35 5. In the Guidance, EPA should acknowledge key factors that have appeared to limit state
36 progress toward developing nutrient criteria. It is the Committee's understanding that
37 one of the key aims of the Guidance is to accelerate State progress toward adopting
38 numeric nutrient criteria. Because a variety of issues (such as limited availability of data
39 and technical expertise, insufficient resources, and expense) are likely responsible for
40 slow progress, the Guidance may not sufficiently remedy the underlying problems and
41 therefore not facilitate state numeric nutrient criteria adoption. A more thorough
42 exploration of the underlying reasons for slow progress would help EPA more directly
43 address specific issues that impede progress.
44

1 *Key recommendations concerning identification of the most scientifically defensible approaches*
2 *to deriving numeric nutrient criteria*
3

- 4 6. In the Guidance, EPA should recommend that users consider alternative conceptual and
5 methodological approaches in cases where such approaches may be needed to account for
6 complex problems associated with nutrients. The problem of eutrophication is complex,
7 involving multiple causal variables, multiple response variables, and feedbacks among
8 the variables (e.g., plants increase in response to nutrients then, in turn, those nutrients
9 are provided a second time as plants decay). Moreover, response variables can be at
10 multiple levels - primary response variables (e.g., plants), secondary response variables
11 (e.g., dissolved oxygen [DO], pH), and tertiary response variables (e.g., fish,
12 macroinvertebrates). A change in a response variable is unlikely to be satisfactorily
13 described by changes in a single “causal” variable (e.g., total nitrogen [TN] or total
14 phosphorus [TP]). The Committee suggests that developing conceptual models/diagrams
15 (more detailed and accurate than shown in Figure 10 of the Guidance) to illustrate
16 linkages and feedbacks between nutrients and response variables would be a useful
17 approach to capture ecological complexity and better construct the conceptual
18 framework.
19
- 20 7. In the Guidance, EPA should explicitly acknowledge the conditions under which the
21 stressor-response relationship applies. For example, the stressor-response relationship is
22 relatively strong and well-established in lakes and reservoirs as opposed to streams and
23 rivers where the relationship is more complex and influenced by many factors (e.g.,
24 shading, sediment, flow regime). In cases where the relationship is not the most
25 appropriate lens through which the problem should be viewed, the user could be directed
26 to other approaches that might better fit the problem. Other documents referenced above
27 (e.g., Florida nutrient guidance document) provide useful examples. The Guidance
28 would benefit from addition of an inset “red-flag” text box that lists circumstances or
29 system characteristics that would alert the user to the need to consider approaches other
30 than stressor-response. This box also might caution the user about mixed systems that
31 have been highly modified and are not easily classified. Likewise, these caveats should
32 also include explicit recognition that the most appropriate criteria may depend upon
33 contexts of the waterbody (e.g., shaded versus open canopy streams), as was done in
34 Florida’s guidance document. Searching for a single statewide criterion might obscure
35 important relationships.
36
- 37 8. The Committee suggests that EPA consider the following two key questions as the
38 Agency selects variables to develop numeric criteria: 1) which measures will allow
39 detection of impairment of designated uses? and 2) is the relationship sufficiently strong
40 to determine a management or regulatory target (i.e., a criterion) to ensure that the
41 designated use is protected? In certain cases, the most appropriate numeric criterion may
42 not be a particular concentration level of a nutrient. Moreover, the stressor-response
43 framework is but one approach for developing numeric nutrient criteria, and often it may
44 not be the most appropriate. Because this concern cuts across all recommendations and

1 approaches included in the Guidance, and also cuts across all charge questions, it must be
2 addressed.

3
4 *Key recommendations to increase the usability of the Guidance and reduce the likelihood of*
5 *misuse*

- 6
7 9. EPA should consider modifying the steps that provide the framework of the Guidance.
8 The Committee suggests that the steps in the framework should be more specific and
9 descriptive. An example is illustrated in Figure 1 of this advisory report. Two important
10 aspects of the example in Figure 1 currently are missing from the Guidance: conceptual
11 model development should be the first step in the process, and the framework should
12 contain an explicit step to determine whether the stressor-response relationship is
13 appropriate.
14
- 15 10. EPA should revise the Guidance to include more detailed and descriptive information on
16 the use of the statistical methods discussed in the document. In addition, EPA should
17 provide additional support to help users meet the technical demands of the methods. The
18 Committee finds that the current draft of the Guidance is written for a user with
19 considerable statistical expertise that may or may not be possessed by state water
20 agencies. This potential mismatch has two serious potential consequences. First, the
21 Guidance will not be helpful if it cannot be easily used by state/tribal water scientists, and
22 second, the recommended methods could be misused and/or misapplied if not sufficiently
23 understood by the user. As a corollary, the Guidance could specify the level of expertise
24 needed by potential users. Correctly identifying the level of expertise of the anticipated
25 users and providing detailed and descriptive information for them is perhaps the most
26 critical step in the continued development and refinement of the Guidance. As part of
27 this process, EPA needs to outline a relatively straightforward process that the users can
28 follow to employ the methods described, and provide technical support for their use.
29
- 30 11. In the Guidance, EPA should more clearly express the caveats and limitations of the
31 approaches presented. In this regard, the following issues are of greatest concern to the
32 Committee: a) The approaches presented in the Guidance are correlative and do not
33 demonstrate causation. b) Many water quality problems are site-specific and
34 confounding variables likely exist. c) As further discussed in the responses to charge
35 questions 2, 3 and 5, there are limitations associated with the retrospective approaches
36 that are the primary focus of the Guidance, and also shortcomings associated with the
37 multivariate techniques presented in the document. In particular, EPA should better
38 identify potential confounding variables and other latent variables that may affect the
39 response.
40
- 41 12. The Guidance should be revised to include additional information (i.e., technical
42 guidance) and more examples showing when and how to use different approaches
43 presented in the document, the advantages and limitations of each approach, the
44 underlying assumptions and data requirements, appropriate interpretations of statistical
45 results, and how to best parameterize the statistical models. This “how-to” information

could take a number of forms, including keys, inset boxes, and appendices. Users must be given additional information that provides a clear understanding of why and under which conditions they should consider any particular approach. Related to this, the Committee recommends that the Guidance contain additional examples of the methods described in the document. Specific topics that might be included in this technical guidance include: how to modify the approaches in order to derive site-specific criteria, how to identify thresholds, use of weight-of-evidence approaches, and how to handle censored values. EPA also could include an appendix that lists other sources of assistance (e.g., Regional Technical Assistance Groups [RTAGs]), and methodological resources). Organization of the document and current section headings also could more clearly identify the steps involved in the suggested empirical approaches. It would also be helpful to incorporate case studies that apply datasets typical of what most states have. These case studies could highlight decision points in the process of criteria derivation. The use of a single case study across all the various approaches suggested in the document would be particularly helpful.

13. The document should better address data requirements (including data acquisition and data quality requirements). Without providing guidelines on data requirements, the potential for applying techniques to inappropriate or inadequate datasets is great. The Committee recommends casting this discussion in terms of data quality objectives (DQOs) and therefore suggests the following process: 1) state the problem; 2) identify the decision; 3) identify inputs to the decision; 4) define the study boundaries; 5) develop a decision rule; 6) specify tolerable limits on decision errors; and 7) optimize the design for obtaining data.

3.2. Charge Question 2. Selecting stressor and response variables

Section 1 of the draft guidance document reviews how to select the variables that appropriately quantify the stressor (i.e., excess nutrients) and the response (e.g., chlorophyll a, dissolved oxygen, or a biological index). Please comment on whether the factors to consider described in Section 1 of the draft document are appropriate for selecting response variables that are sensitive to nutrients and related to measures of designated uses.

Section 1 of the EPA Guidance reviews factors to consider when selecting stressor and response variables for empirical derivation of numeric nutrient criteria. The Committee finds that EPA should strengthen the Guidance by including additional material to address the points discussed below. Although the current version of the Guidance addresses some of these points, we recommend including additional examples and revisions to further develop various parts of the text as discussed below.

Findings on selecting response variables

- 1 • Although the Guidance states that response variables should be coupled to designated uses,
2 the Committee finds that this point needs additional elaboration. Some response variables
3 described in the Guidance are clearly related to designated uses (e.g., DO) but the linkage of
4 other responses to designated uses is less obvious or not as well supported scientifically (e.g.,
5 macroinvertebrate species richness). Despite the importance of DO and the fact that a large
6 number of waterbodies are impaired due to low DO concentrations, none of the examples in
7 the Guidance include DO as a response variable. This is a significant omission that needs
8 correcting. The Committee notes that appropriate response variables are also highly
9 ecosystem specific. For example, chlorophyll concentrations are often more clearly related
10 to designated uses for lakes than streams. While response variables for single taxa (e.g.,
11 salmon) may be tightly related to designated use, multimetric variables (macroinvertebrate
12 indices, index of biotic integrity [IBI]) may be more powerful for integrating the response to
13 nutrients at the community or ecosystem level. The Guidance would be strengthened by
14 including more discussion relating ecosystem type and potential response variables to the
15 designated uses (a table with some accompanying text might be an effective way to do this).
16
- 17 • Conceptual model development should be required and should be incorporated early in the
18 process of criteria development (see Figure 1 of this advisory report). Conceptual models are
19 an important component in selection of response variables. Any stressor-response
20 relationship used in criteria development must have ecological relevance (based on
21 ecological understanding of the system) that can be readily explained and defended as
22 discussed in step two in the Guidance. Conceptual models based on past empirical and
23 experimental studies are important for identifying the mechanisms responsible for responses
24 and effectively communicating this linkage. In the framework suggested by the Committee
25 (Figure 1), developing the conceptual model is the first step in the process.
26
- 27 • The Guidance would be strengthened considerably by presentation of examples illustrating a
28 strong nutrient-response relationship and, as previously mentioned, clear linkage of the
29 response variable to a designated use. It is important to clearly present the rationale for such
30 linkage. Some of the examples in the Guidance illustrate relationships with very low R^2 and
31 response variables that are not clearly related to designated use.
32
- 33 • In the Guidance, further discussion of potential response variables appropriate for nutrient
34 effects on detritus-based systems is warranted (e.g., how macroinvertebrate populations
35 dependent on detritus may respond). The Guidance focuses on nutrient-response
36 relationships driven by autotrophic processes (nutrients directly control algal growth,
37 excessive amounts of which impair systems through indirect effects on DO, food web
38 changes, and aesthetics). However, nutrients can also directly control heterotrophic microbes
39 (bacteria, fungi) and indirectly control decomposition of organic matter. Excessive nutrient
40 levels could produce large microbial growths or alter food webs in detritus-based ecosystems
41 (e.g., many streams). Studies in the literature are cited, but examples using relevant response
42 variables (e.g., shredder macroinvertebrate biomass, leaf breakdown rate) would be useful.
43

44 *Findings on stressor and related variables*

- 1
2 • In the Guidance, more discussion is needed to outline and provide advice on the rationale for
3 selecting variables that should be included in data collection to allow: 1)
4 classification/stratification of data prior to evaluation of stressor-response relationships (e.g.,
5 development of different criteria for different strata of systems); and 2) use of multivariate
6 approaches to separate the influence of nutrients from other stressors (e.g., sediments, light
7 regime, toxics). Stratification/classification is a particularly important issue for defining
8 nutrient stressor-response relationships for streams where other factors can impose
9 significant constraints on the effects of excess nutrients on designated uses. For example,
10 nutrient-chlorophyll relationships may not be observed in highly shaded (forested) streams,
11 but may be significant in open-canopy streams. Similarly, nutrient-chlorophyll relationships
12 may be weak in high gradient streams but much stronger in low-gradient streams. For lakes,
13 nutrient-chlorophyll relationships may be much different for highly-colored (high dissolved
14 organic carbon [DOC]) versus clear (low DOC) systems.
15
- 16 • Single variable stressor-response relationships (e.g., those derived using the simple linear
17 regression approach discussed in the Guidance) that explain a substantial amount of variation
18 are likely to be uncommon for most aquatic ecosystems (in particular, streams). Multivariate
19 approaches (multiple regression, structural equation modeling [SEM], etc.) may be needed to
20 identify nutrient effects. These approaches require data on other potential stressors or
21 constraining variables. Multivariate approaches may also be useful early in the analysis to
22 determine whether nutrient effects are significant relative to other stressors and constraints
23 and whether/how to pursue the nutrient effects using simple univariate regressions, perhaps
24 after stratification of systems.
25
- 26 • The Guidance focuses primarily on TN and TP as the primary nutrient stressor variables.
27 Some consideration should be given to inorganic N and inorganic P because these forms are
28 the most immediately biologically available. In addition, it is easier to measure the inorganic
29 forms of N and P and more and/or better data may be available for these forms. This is
30 particularly true for ammonium and nitrate versus TN, but perhaps less so for P.
31
- 32 • In many regions N and P are often co-limiting to plants and microbes and stressor-response
33 relationships based on only one nutrient are weak. Nevertheless, nutrients (N and P) may be
34 the primary factor controlling productivity/biomass. There have been several recent papers
35 arguing for management of N and P in combination rather than singularly (Lewis and
36 Wurtsbaugh, 2008; Conley et al., 2009; Paerl, 2009). This would suggest development of
37 multivariate stressor-response relationships (e.g., multiple regression) that include both N
38 and P as independent variables.
39
- 40 • A basic conceptual problem concerning selection of nutrient concentrations as stressor
41 variables (as illustrated in the Guidance) is that nutrient concentrations directly control only
42 point-in-time, point-in-space kinetics, not peak or standing stock plant biomass. Plant
43 biomass is driven by nutrient supply rates (i.e., nutrient mass loads). Ambient nutrient
44 concentrations are not necessarily good surrogates for nutrient mass loads. Relationships

1 between nutrient mass loads and ambient nutrient concentrations are highly system-specific
2 and depend on many factors including inflows, hydrology, bathymetry, sediment-water
3 exchanges and chemical-biological processes. Consequently, there may be many systems for
4 which nutrient concentrations will not be appropriate stressor variables. For such systems it
5 may be more appropriate, and scientifically defensible, to use site-specific mechanistic
6 models incorporating loading to determine the nutrient controls required to attain designated
7 uses.

8
9 *Findings on temporal/spatial aspects of data*

- 10
- 11 • The Guidance provides little discussion regarding the temporal/spatial aspects of data needed
12 to develop relevant stressor-response relationships. For example, the document could be
13 strengthened by providing additional material to address the following questions. “Under
14 what conditions might the use of mean/median or maximum/minimum values of stressor and
15 response variables be more appropriate than discrete instantaneous measurements?” “Are
16 there instances when the use of temporally out-of-phase stressor and response data are most
17 appropriate (e.g., the widely recognized relationship between spring nutrient concentration
18 and summer maximum chlorophyll concentration in lakes)?” “How can time series or
19 longitudinal data in specific systems be used to develop more generalized stressor-response
20 relationships?” Although such guidance may be covered in the various system specific
21 technical manuals (USEPA 2000a, 2000b, 2001), a summary/synthesis of the major points of
22 these earlier documents should be included in the empirical approaches document.
23
 - 24 • The Guidance could be strengthened by including a discussion of the importance of
25 considering “data bias” in interpreting the stressor-response relationships. This discussion
26 should focus on how “data bias” (i.e., limits on data representativeness) might affect
27 predictive performance and uncertainty in stressor-response relationships. Uncertainty
28 imposed by model assumptions should also be discussed. Specifically, additional guidance is
29 needed with regard to interpretation of data from particular environments (e.g., a set of lake
30 data from a particular region) and its appropriateness (or lack thereof) for describing
31 conditions more broadly. It would be helpful to include in the Guidance examples
32 illustrating databases that would be “ideal” or appropriate for each empirical model
33 presented. For example, information could be provided to indicate whether a conceptual
34 model for considering nutrient criteria might be best approached using: seasonal data; data
35 from shaded versus unshaded streams; data from Wadeable streams versus big rivers; and/or
36 long versus short term averages of data describing the stressor or the response.
37
 - 38 • It would be useful to include in the Guidance some discussion of how nutrient recycling and
39 other feedbacks influence stressor-response relationships. For example, the Guidance could
40 be strengthened by addressing the following questions. “How does recycling contribute to
41 variability and uncertainty in stressor-response relationships?” “Are there variables that can
42 be used in stressor-response relationships to account for recycling?”
43

44 *Key recommendations concerning selection of variables to appropriately quantify the stressor*
45 *and response*

1
2 The Committee provides the following key recommendations to address the findings above
3 and strengthen Section 1 of the Guidance.
4

- 5 1. The Guidance should be revised to elaborate upon the coupling of response variables to
6 designated uses and the importance of ecological relevance of the stressor-response
7 relationship. Examples should be included to further illustrate this important point. The
8 examples should show strong nutrient-response relationships. The Guidance should be
9 revised to include at least one example for DO as a response variable. Ideally, each method
10 should include an example for streams/rivers and an example for lakes. If empirical stressor-
11 response relationships are not appropriate or workable for DO in lakes, then the Guidance
12 should state this specifically and recommend other approaches, for example, site-specific
13 mechanistic models. There are a large number of waterbodies that are impaired by low DO
14 and the draft Guidance is silent on this important nutrient-related problem.
15
- 16 2. The Guidance should be revised to include discussion of potential response variables
17 appropriate for assessing nutrient effects on detritus-based systems.
18
- 19 3. The Guidance should be revised to include more discussion and advice concerning selection
20 of variables and data needed to allow:
21
 - 22 – Classification/stratification of data prior to evaluation of stressor-response
23 relationships (e.g., development of different criteria for different strata of systems).
24
 - 25 – Use of multivariate approaches to separate the influence of nutrients from other
26 stressors (e.g., sediments, light regime, toxics). In general, the importance of
27 multivariate stressor-response relationships and tools for multivariate approaches
28 should be further discussed in the final Guidance.
29
- 30 4. The Guidance should consider including inorganic N and inorganic P as nutrient stressor
31 variables because these forms are the most immediately biologically available.
32
- 33 5. The basic conceptual problem associated with selecting nutrient concentrations as stressor
34 variables should be addressed in the Guidance (i.e., nutrient concentrations directly control
35 only point-in-time, point-in-space kinetics, not peak or standing stock plant biomass).
36
- 37 6. The Guidance should be revised to include discussion of:
38
 - 39 – The temporal/spatial aspects of data needed to develop relevant stressor-response
40 relationships. (e.g., are there instances when the use of temporally out-of-phase
41 stressor and response data are most appropriate?)
42
 - 43 – How “data bias” (e.g., data from different types of systems) might affect predictive
44 performance and uncertainty in stressor-response relationships.

- 1
2 – How nutrient recycling and other feedbacks influence stressor-response relationships.
3

4 **3.3. Charge Question 3. Approaches to demonstrate the distribution of and**
5 **relationships among variables**
6

7 **Section 1 outlines methods to visualize available data. Please comment on the**
8 **effectiveness of the following approaches described in the document (listed below)**
9 **to demonstrate the distribution of and relationships among variables.**
10

11 a) Basic data visualization techniques

12 b) Maps

13 c) Conditional probability

14 d) Classifications
15

16 Section 1 of EPA's Guidance discusses exploratory data analysis and presents several
17 methods for demonstrating the distribution of and relationships among variables. Several basic
18 plotting techniques are presented in Subsections 1.2 – 1.6 of the document. This is followed by a
19 description of conditional probability analysis (a statistical approach for summarizing how
20 changes in nutrient concentrations are associated with the probability of waterbodies attaining
21 their designated uses). The Committee was asked to comment on the effectiveness of the
22 methods presented in this section of the Guidance.
23

24 The Committee notes that the response to Charge Question 3 necessarily overlaps with
25 responses to other charge questions, particularly those that focus on identifying stressor-response
26 relationships and conducting statistical analyses. We emphasize that visualization of data is of
27 secondary importance if the data and statistical methods being visualized are inappropriate,
28 because the visualization in itself suggests authenticity. Furthermore, the exploratory data
29 analysis, including visualization, should be conducted prior to inferential statistical analyses of
30 potential stressors and responses. The objectives of exploratory data analysis should be to better
31 understand the system of interest and to maximize the accuracy and minimize the variability of
32 the subsequent stressor-response relationships. The Committee finds that discussion of
33 exploratory data analysis in the Guidance would be more effective if the document were
34 reorganized and expanded to address the following points.
35

- 36 • The Guidance would be more effective if exploratory data analysis were included by itself in
37 a separate section of the document following a major section on problem formulation
38 (corresponding to the Framework in Figure 1 of this advisory report).
39
40 • Additional methods for exploratory data analysis should be described in the Guidance. These
41 additional methods should include: the use of summary statistics; time series plots at fixed
42 points in space; longitudinal plots at fixed points in time; bubble plots; Pearson and other
43 types of non-parametric correlation analyses; and maps that show temporal (monthly,
44 seasonal, inter-annual) as well as spatial patterns.

- 1
2 • Clear guidance is needed for identifying when and how the statistical methods and
3 visualization techniques should be used. The strengths and limitations of the methods should
4 also be identified. It would be useful to show several case examples that range from state-
5 wide to local and data-rich to data-poor, and exemplify different types of aquatic ecosystems
6 (e.g., headwater streams, large rivers, lakes and estuaries). Examples should note the
7 strengths, limitations, assumptions and uncertainties that must be considered when using the
8 methods to explore and visualize the data, and subsequently develop the criteria. These
9 examples should demonstrate how nutrients can be identified as significant stressors when
10 multiple stressors and habitat factors are present that may affect the resident communities.
11
 - 12 • The discussion in Subsection 1.6 of the Guidance (examination of stressor-response
13 distributions across different classes, e.g., ecoregions) should be expanded. The subsection
14 should discuss additional data analysis and contain examples for different spatial
15 classifications (e.g., ecoregions, states, watersheds, systems of interest), different waterbody
16 types (e.g., streams, rivers, lakes, estuaries) and other important physical and chemical
17 characteristics that could affect the applicability of the nutrient criteria.
18
 - 19 • The examples provided in the Guidance generally do not demonstrate a strong nutrient
20 stressor linkage to beneficial use impairment. The stream examples show very weak
21 correlations that have high levels of uncertainty, and lump data from distinctly different
22 ecosystems where multiple factors in addition to nutrients will contribute to biotic responses.
23
 - 24 • All of the statistical and visualization methods discussed in Subsections 1.2 -1.6 of the
25 Guidance can be effective but they should be presented and used in a combined, weight-of-
26 evidence approach because they each involve exploring the data in different ways.
27
 - 28 • The Committee emphasizes the importance of choosing the biological endpoints (i.e.,
29 response variables) that respond specifically to nutrients. We note that responses of benthic
30 indices can be related to many types of stress. We question why periphyton would not be a
31 better receptor to measure.
32
 - 33 • The Committee suggests that field-based species sensitivity distributions (SSDs) may be
34 useful for nutrient criteria development. We note that SSDs have been used effectively in
35 recent publications for establishing guidelines (or refuting them) for contaminants,
36 temperature, and salinity (Hickey, 2008; Leung et al., 2005).
37
- 38 The Committee also notes the following technical edits and corrections needed in the Guidance.
- 39 a. Clarify that macroinvertebrate richness is plotted in examples in Subsection 1.3.
 - 40 b. The Guidance (p. 7) states that “variables are equally weighted” yet only one variable is
41 plotted in each box plot. A better statement would be: “One limitation for box plots is
42 that all of the samples are equally weighted.”
 - 43 c. Explain probability survey design and data smoothers or provide references.

- 1 d. Figure 7 is very confusing to those unfamiliar with scatterplot matrices; some additional
2 explanation regarding how to “read” the horizontal and vertical axes of each graph in the
3 matrix would help. Suggested wording: “For each scatterplot, its x-axis is the variable
4 stated in the column in which the graph appears. Its y-axis is the variable stated in the
5 row in which the graph appears.”
6

7 *Key recommendations regarding methods for demonstrating the distribution of and*
8 *relationships among variables*
9

10 As discussed above, the Committee recommends that EPA restructure and revise the
11 Guidance to strengthen discussion of methods for demonstrating the distribution of and
12 relationships among variables. The following key recommendations are provided.
13

- 14 1. The Committee recommends that the Guidance be clarified by reframing Subsections 1.2
15 through 1.6 as a separate major section on exploratory data analysis. These subsections
16 should follow another separate major section on problem formulation (see Figure 1), and
17 the material in Subsection 1.1 (selecting stressor and response variables) should be
18 moved to later section(s) of the document.
19
- 20 2. The Guidance should be revised to include additional methods for exploratory data
21 analysis. These additional methods should include: the use of summary statistics; time
22 series plots at fixed points in space; longitudinal plots at fixed points in time; bubble
23 plots; Pearson and nonparametric correlation analyses; and maps that show temporal
24 (monthly, seasonal, inter-annual) as well as spatial patterns.
25
- 26 3. Subsection 1.6 of the Guidance should be expanded to include additional examples of
27 different spatial classifications. Specifically, the classification subsection of the
28 Guidance (Subsection 1.6) should be expanded with data analysis examples for different
29 spatial classifications (e.g., ecoregions, states, watersheds, systems of interest), different
30 waterbody types (e.g., streams, rivers, lakes, estuaries) and other important characteristics
31 that will affect the applicability of the nutrient criteria. These characteristics could
32 include, but should not be limited to, stream order, flow, velocity, canopy cover,
33 dissolved oxygen, reference condition trophic status, channel width and depth.
34
- 35 4. The Guidance should be revised to clarify, early in the document, that there are many
36 useful statistical and visualization methods that are not presented and which may be
37 useful. The more common/well accepted methods could be listed in a table with
38 references. It may also be useful to mention methods that are inappropriate. With each
39 method the associated strengths, limitations, assumptions and uncertainties should be
40 noted to better guide the user.
41
- 42 5. Several case examples of exploratory data analysis should be included in the Guidance.
43 These examples should illustrate cases ranging from national to local in scope, and data-
44 rich to data-poor, with guidance on how best to explore and visualize the data.
45

6. The Guidance should contain additional information concerning statistical assumptions associated with various methods. Some guidance should be presented, as in other EPA documents (e.g., U.S. EPA, 2006a; U.S. EPA 2006b), to address the importance of ensuring that statistical assumptions are not violated and that adequately trained statisticians, in concert with experienced aquatic ecologists and environmental modelers, evaluate the data. An example could be included to show how overly simplistic statistical analysis could not identify a relationship that became evident after complex/advanced analysis. The Committee notes that CProb 1.0, EPA's tool for conditional probability analysis was developed with the R language and environment for statistical computing. The Committee questions whether R, an open-source freeware product that is becoming very popular, is completely acceptable, in the sense that there are many R-macros in use that remain to be properly "vetted." There should be some level of assurance that the recommended R-products have been properly vetted (e.g., CProb 1.0).
7. The Guidance should contain a quantitatively based weight-of-evidence framework using multiple methods and then combining them into figures and tables for visualization. Multiple statistical methods on one dataset do not equate to a reasonable weight-of-evidence that significantly reduces uncertainty. Rather, the weight-of-evidence should involve different assessment methods (e.g., different datasets, different biological endpoints, measures of habitat, etc.). This premise has been embraced by other EPA programs and the scientific community (Adams, 2003; Burton et al. 2002; Chapman, 2007; Chapman et al., 2002; Collier, 2003; Cormier et al., 2010; Fox, 1991; Linder et al., 2010; Linkov et al., 2009; Suter et al., 2002; Suter et al., 2010; U.S. EPA, 2000c; Weed, 2005; Wickwire and Menzie, 2010).
8. The Guidance should contain a discussion of how the stressor/response variables to be used are linked to one another in space and time for further analysis. There is no mention of this in Subsection 1.1 of the Guidance. The Committee questions whether it should be assumed that stressor/response measurements always occur at the exact same time and locations. It is also important to ensure that high flow events have been measured. It is well established that most nutrient loading occurs during high flows. Therefore, the influence of seasonality and smaller-scale temporal dynamics (e.g., storm events) and the importance of linking stressor and response variables with these factors should be at least noted in the Guidance.
9. The Guidance should discuss the use of modeled data (e.g., land use characterization, hydrology, surface runoff, receiving water quality) for estimating nutrient concentrations/exposures. The pros and cons associated with the use of such data should be briefly mentioned. There are a number of EPA-supported models that have been widely used and documented in recent years (e.g., HSPF, BASINS, QUAL2K, WASP, AQUATOX, and Chesapeake Bay WQSTM). Some of these are integrated watershed models designed to represent inflows and non-point source runoff loads. Typically, they are used as a "loading engine" for a receiving water quality model. Receiving water quality models describe load-response relationships for exposures (ambient nutrient

1 concentrations) and effects (e.g., plant biomass, zooplankton, dissolved oxygen), and
2 response parameters that represent use impairment. Some receiving water quality models
3 can address multiple stressors. For example, they can include N, P and silicon as
4 potentially limiting nutrients, sediment (suspended solids) and its influence on
5 underwater light attenuation, incident solar radiation, temperature, and grazing pressure.
6 It is possible to use these water quality models to describe exposure (in terms of ambient
7 nutrient concentrations) but in the absence of empirical data, this would not be
8 scientifically defensible.

9
10 10. The Committee recommends that EPA re-evaluate many of the figures in the Guidance
11 (e.g., 4-8, 13-16, 21, 25, and 26). These figures show widely varying data that
12 demonstrate weak relationships.

13
14 11. The Committee recommends that the Guidance be revised to clearly indicate the
15 statistical assumptions and uncertainties that should be taken into consideration when
16 using methods described in the document. Some of the methods are complex and their
17 descriptions lack transparency. Guidance should be provided to ensure that states and
18 other users have an understanding of the data requirements and limitations, the associated
19 statistical assumptions, and uncertainties.

20
21 12. The document should contain a discussion of ways to examine the independent and
22 interactive effects of the variables to be considered in deriving numeric nutrient criteria
23 (i.e., provide a menu of options to examine independent and interactive effects).
24 Statistically, there are several well known ways to address additional contributing
25 variables, such as total suspended solids (TSS). One way would be to use a multiple
26 regression model or analysis of covariance (ANCOVA). This would be a valuable
27 approach, as the additional variables are to be treated as continuous variables, and
28 interaction terms could be added to see if the effects of TN/TP were dependent on levels
29 of TSS, which would be expected, particularly for TP. If one treats the additional
30 variables as factors then an analysis of covariance (ANCOVA) model would be most
31 appropriate. For example, if there were a TSS threshold of interest, a relationship could
32 be established between an invertebrate endpoint and nutrient levels above and below a
33 critical TSS threshold. This would allow one to examine independent and interactive
34 effects.

35
36 13. The Guidance should mention the potential benefits of using proxy variables in an initial
37 approach for exploratory analysis of data trends. For example, variable data sets that are
38 easier and more practical to obtain, such as more generic point/nonpoint source loadings
39 or commonly sampled stressor/response variables, might be used as proxy variables for
40 exploratory analysis of data trends. This is briefly mentioned in Subsection 3.1 of the
41 Guidance (auxiliary model), but such an approach could also be useful for selecting
42 stressor/response variables early in the process (Section 1).

43
44 **3.4. Charge Question 4. Methods for assessing the strength of the cause-effect**
45 **relationship**

Section 2 of the draft guidance document describes methods for assessing the strength of the cause-effect relationship represented in the stressor-response linkage. Please comment on whether the draft guidance document adequately describes how conceptual models, existing literature, and empirical models can be used to assess how changes in nutrient concentration are likely to cause changes in the chosen response variable.

Section 2 of the Guidance provides a summary of how the strength of tentative stressor-response pairings from step 1 can be assessed. Certainly, as indicated in the Guidance, conceptual models and existing literature can be used to support relationships that will be explored with the statistical analysis that follows. At this stage of the analysis, stressor-response relationships for which there is no reasonable conceptual model or literature to explain the underlying mechanisms would be of limited value for setting criteria. Such relationships should be set aside. The Committee finds that the Guidance should be improved by incorporating revisions to address the following points.

- Section 2 of the Guidance does not address the strength of the stressor response relationship, but rather support for the stressor-response relationship that is to be explored statistically. “Support” for the stressor response relationship, rather than “strength” of the relationship, would be a better term to use in this section of the Guidance, because strength refers to the “tightness” of the statistical association between stressor and response. Use of the term “support” would, therefore, be less confusing to the user.
- It is not clear why information from mechanistic models was not included in Section 2 of the Guidance. Because mechanistic models can integrate information on the interactions of major ecosystem processes to derive quantitative estimates of effects, they too should be discussed as a possible way of supporting the stressor-response relationship.
- Additional discussion of conceptual model selection (with specific examples) would be helpful. There are many ways to select a conceptual model and various model selection criteria that could be applied. An expanded discussion of these issues could help provide further background for a user of the document. Specific examples could be followed in later sections with discussion of statistical approaches to analyze the strength of the potential cause-effect relationships. In other words, EPA could provide an example from beginning to end that a user could follow from step to step.
- One important aspect of finding support for stressor-response pairings is that without formal training and practical experience in the sciences, especially biological and ecological disciplines, it is difficult to fully understand the complex relationships that may be identified. The Guidance should state the level of statistical and ecological expertise needed to use the document.

- 1 • Structural equation modeling (SEM) and Propensity Score Analysis (PSA) are techniques
2 that can be used to organize and evaluate relationships between nutrients and response
3 variables when extensive data are available. SEM might be more useful in tracing pathways
4 (it is also called path analysis) of cascades that are initiated by excess nutrients than in
5 defining criteria candidates. A relevant example of SEM is really needed in the Guidance if
6 this approach is to be considered by users. PSA, on the other hand, seems to be useful for
7 sorting out groups that share covariates but may have unique nutrient characteristics. Such
8 sorting could lead to a clearer understanding of how nutrients function amid multiple
9 covariates. The example of PSA in the Guidance appendix is helpful, but further explanation
10 of how to interpret the results of the analysis is needed. An analysis such as PSA might
11 really belong in a later section of the document, as it is used for data analysis rather than
12 supporting potential relationships.
13
- 14 • A reasonable way to assess nutrient effects might be to split datasets (through PSA, principal
15 components analysis, and/or cluster analysis) to enable a system-specific analysis (or analysis
16 of a small groups of sites). Given the many factors that affect streams and rivers, system-
17 specific analysis really provides an assessment of whether altering nutrient concentrations
18 would have the desired effect on the biotic communities present. Possible factors to consider
19 in splitting data for streams and rivers might include, for example, stream order, flow,
20 velocity, canopy cover, dissolved oxygen, bottom type, channel width, habitat, and depth.
21
- 22 • Experimental validation of causal relationships between nutrient and response variables
23 should be approached with caution. The final method discussed on page 17 of the Guidance
24 is experimental validation of causal relationships between selected nutrients and response
25 variables. The Committee notes that this approach could be helpful *in situ* and there are
26 examples of this (Benstead et al., 2009; Cross et al., 2006; Cross et al., 2007; Greenwood et
27 al., 2007; Peterson et al., 1985; Slavik et al., 2007; Stockner and Shortreed, 1978), but
28 mesocosm or laboratory experiments are of limited use in validating causal relationships
29 between nutrient and response variables. For example, Hill and Fanta (2008) and Hill et al.
30 (2009) showed in Oak Ridge National Laboratory artificial streams how P and light interact.
31 This type of work provides fundamental data on how stream algae respond to P and light, and
32 supports basic conceptual models of this relationship. These and previous studies have
33 shown that, under controlled conditions it takes very little P to maximize algal growth given
34 high light and this fundamental relationship could be applied to any stream in the U.S.
35 However, the relationship is often not observed in datasets because other factors such as
36 bottom substrate, turbidity, canopy cover, hydrology, or depth limit algal production.
37 Therefore, caution must be used in applying a relationship from a subset of data to all data
38 from systems that do not have the same or similar conditions.
39

40 *Key recommendations concerning methods for assessing the strength of the cause-effect*
41 *relationship represented in the stressor-response linkage*
42

43 In light of the comments and findings discussed above, the Committee provides the following
44 key recommendations to improve Section 2 of the guidance.
45

1. Section 2 of the Guidance would be more appropriately titled “Assessing Support for the Potential Cause-Effect Relationship.”
2. Mechanistic models should be discussed in the Guidance as one way of supporting the stressor-response relationship.
3. The discussion of conceptual models should be expanded to address various criteria for model selection, and additional examples should be included.
4. The level of statistical and ecological expertise needed to use the Guidance should be stated.
5. Structural Equation Modeling (SEM), offered as an alternative model for exploring nutrient-ecosystem response, should be more fully explained with clear examples.
6. Further explanation of how to interpret the results of propensity score analysis (and additional examples) should be included in the Guidance.
7. Experimental validation of causal relationships between nutrient and response variables should be approached with caution because a number of factors can affect the response of a system to nutrient enrichment.

3.5. Charge Question 5. Statistical methods to analyze the data

Section 3 of the draft guidance document outlines statistical methods to analyze the data to estimate stressor-response relationships. Please comment on the appropriateness of the methods outlined in the document (listed below) for describing stressor-response relationships associated with nutrient pollution. What approaches would you recommend that could effectively address indirect pathways of adverse effects? What recommendations do you have to address the effects of confounding variables and uncertainty in the estimated relationships?

- a) Simple linear regression
- b) Quantile regression
- c) Logistic regression
- d) Multiple linear regression
- e) Non-parametric changepoint analysis
- f) Discontinuous regression models

The Committee notes that EPA’s draft Guidance appropriately states that numeric nutrient criteria should be based on predictive stressor-response relationships so that changes in the level of stressor variables will result in predictable ecosystem responses. However, based on examples presented in the draft document and elsewhere, a large degree of unexplained variation can be encountered when attempting to use empirical stressor-response approaches to establish criteria. The final Guidance needs to clearly indicate that such unexplained variation can present

1 a significant problem to this method of developing numeric criteria. Further, the final document
2 should emphasize that statistical associations may not be biologically relevant and do not prove
3 cause and effect. However, when properly determined, statistical associations can be very useful
4 in supporting a cause and effect argument as part of a weight-of-evidence approach to criteria
5 development. To this end, the final document should provide greater detail on the
6 implementation of statistical procedures and development of other supporting information to
7 minimize the degree of unexplained variation and maximize the potential for the empirical
8 stressor-response approach to result in useful numeric nutrient criteria. EPA should also provide
9 guidance on the strength of stressor-response relationships needed to support criteria
10 development using an empirical stressor-response approach. Further, because nutrients are
11 essential elements, the application of statistical methods must consider both nutrient deficiency
12 and excess. Clear links between response variables and designated uses are needed to ensure that
13 both of these possible impairment types are addressed. The Committee provides the following
14 findings and comments concerning the appropriateness of statistical methods in the Guidance,
15 approaches to address indirect pathways of adverse effects, and ways to address the effects of
16 confounding variables and uncertainty in the estimated relationships.

17
18 *Findings on appropriateness of listed statistical methods*
19

- 20 • The Guidance represents a substantial step forward in describing statistical methods that can
21 be used in deriving nutrient criteria based on stressor-response relationships, but more
22 information is needed to describe supporting analyses necessary for application of the
23 methods. The six methods identified in the Guidance generally provide appropriate options
24 for describing stressor-response relationships that may be sufficiently predictive to support
25 setting numeric nutrient criteria. As many examples in the draft document illustrate, there is
26 likely to be considerable variability in stressor-response nutrient relationships and, thus, in
27 the predicted outcome or response to both target setting and response to mitigation efforts.
28 Therefore, the document must provide more information on the supporting analyses needed
29 for each method to correctly identify useful predictive relationships, and acknowledge that
30 the use of these statistical methods alone cannot provide sufficient evidence of a cause-effect
31 relationship.
- 32
33 • The use of non-parametric change point analysis and discontinuous regression analysis must
34 be associated with biological significance and the designated uses to be protected by numeric
35 nutrient criteria. As stated previously, response variables must be associated with designated
36 uses in all cases. This has implications for the use of non-parametric change point analysis
37 (nCPA) and discontinuous regression in criteria development. The Guidance indicates that,
38 because these procedures may identify breakpoints in nutrient responses that can serve as
39 criteria thresholds, the methods may be used when designated use thresholds are not
40 available. However, although these methods may be able to identify and characterize
41 breakpoints, such breakpoints may not necessarily have any biological significance, nor will
42 they necessarily be related to designated uses that are to be protected by numeric nutrient
43 criteria. Use of these methods must be associated with designated uses.
44

- 1 • The statistical methods in the Guidance require careful consideration of confounding
2 variables before being used as predictive tools. For example, the appropriate use of bivariate
3 regression methods requires additional efforts through classification or other means to
4 minimize the influence of other potential causal variables so that an acceptable level of
5 confidence in the predictive power of the relationship can be achieved. Without such
6 information, nutrient criteria developed using bivariate methods may be highly inaccurate.
7 Multiple linear regression is an appropriate way to incorporate covariates into a single
8 analysis, although predictive power using this procedure must also be evaluated carefully.
9
- 10 • As previously noted, because plant biomass is driven by nutrient supply rates (mass loads), a
11 potential conceptual problem exists with the selection of nutrient concentration (often used in
12 the Guidance) as a stressor variable. This problem illustrates the importance of careful
13 characterization of confounding variables. Nutrient concentrations control only point-in-
14 time, point-in-space kinetic rates, not peak or standing stock plant biomass. Plant biomass is
15 driven by nutrient supply rates (mass loads). Furthermore, nutrient concentrations may not
16 be direct surrogates for nutrient mass loads. Relationships between nutrient mass loads and
17 ambient nutrient concentrations are highly system-specific and depend on many factors.
18 Consequently, in some circumstances, statistical methods alone will not adequately account
19 for the influence of confounding variables and reduce uncertainties. In other words, the
20 Committee anticipates situations in which stressor-response statistical analysis may not lead
21 to a scientifically justified endpoint.
22
- 23 • In order to be scientifically defensible, empirical methods must take into consideration the
24 influence of other variables. On page 22 of the Guidance, the authors acknowledge that
25 factors co-varying with TP concentrations may explain a portion of the 61% of the variation
26 in log chlorophyll a concentrations apparently attributable to log TP concentrations. This
27 presents a critical challenge in the use of empirical methods as a means of establishing
28 numeric nutrient criteria because it means that controlling TP concentrations may have no
29 potential to yield reductions in chlorophyll a concentrations. Thus, in order to be
30 scientifically defensible, empirical methods must take into consideration the influence of
31 other variables.
32
- 33 • It is important to discuss strength-of-relationship concerns and how results of empirical
34 approaches should be interpreted in the context of criteria development. Figure 13 on page
35 24 of the Guidance provides an illustration of the challenges facing the users of simple linear
36 regression (SLR) and other empirical approaches. In this case, total macroinvertebrate
37 species richness was regressed against total N concentrations obtained from EPA
38 Environmental Monitoring and Assessment Program (EMAP) West Xeric region streams.
39 Overall, total species richness declines with increasing TN concentration in these stream
40 data. Applying SLR to log-transformed data yields a statistically significant slope
41 $-3(\log(\text{TN}))$ at $p < 0.001$. However, a large degree of scatter remains, as indicated by the R^2
42 value of 0.19. A TN “candidate criterion” of 320 ug/L is obtained by finding the point of
43 intersection of an assumed designated use total species richness threshold of 40 and the mean
44 regression line $\log(\text{TN}) = \sim 2.5$. Unfortunately, the points where the lower and upper 90%

This Draft is made available for quality review and approval by the chartered Science Advisory Board. This Draft does not represent EPA policy.

prediction interval lines cross a species richness threshold of 40 cover a TN concentration range from about $\log(\text{TN}) = 1.25$ to $\log(\text{TN}) = 4$ based on inspection of Figure 13. This corresponds to a TN concentration range of 16 ug/L to 10,000 ug/L. It is important to understand the management consequences of this considerable uncertainty. Also, the fact that the relationship in Figure 13 is both statistically significant (i.e., some trend is evident) and has a low $R^2 = 0.19$ (much scatter also exists) presents an opportunity to discuss strength-of-relationship concerns and how such results should be interpreted in the context of criteria development.

- As previously discussed, relationships for streams may be more complex than for lakes and must account for multiple stressors/conditions and/or stream ‘types’ or conditions, and then be applied appropriately. For example, a stratified approach that considers attributes known to be important for a particular environment (lake, stream, estuary) such as canopy, habitat, etc., should be considered. It is also important to deal with both N and P simultaneously and to consider inorganic N and dissolved P. An exercise in Section 3 of the Guidance illustrates the relationship between chlorophyll a and TP in lake water. This is perhaps the easiest and most well known example of stressor-response in natural waters, and specifically in lakes. This relationship is less certain in streams because they are more heterogeneous than lakes. The Guidance also inappropriately assumes that only nutrients affect taxa. The functionality of aquatic food chains is not solely dependent on one type of biota, sediment type, or single nutrient concentration. There are multiple stressors affecting receptors in a number of ways, over the landscape and watershed in question. Confounding variables are not sufficiently addressed in the Guidance. As previously discussed, approaches that address multiple factors, such as a stratified (or hierarchical) approach that considers other attributes known to be important (e.g., canopy, habitat, multiple nutrients) should be considered.
- The Guidance could be improved by replacing many examples that provide low explanatory power. Concerns include examples with very low R^2 indicating low explanatory power and incomplete description of large uncertainty. These examples indicate that variables other than TP or TN have a greater impact on response, which implies that reducing TP or TN may not have the desired effect. Helpful examples could include: one with a response variable indirectly associated with a designated use; and one from a state where a Secchi depth is used as a criterion for water quality (otherwise Subsection 3.1, paragraph 2 sounds extremely vague).
- Parametric (e.g., Pearson) and nonparametric (e.g., Spearman’s rank, Kendall’s tau) correlation analyses can assist in identifying the influence of confounding variables, but these methods are not specifically mentioned in the Guidance. Both of these types of analyses would be helpful in exploratory data analysis.
- The Guidance lacks sufficient discussion of the importance of variable selection and data characteristics to ensure useful implementation of the statistical procedures. In addition to its incomplete treatment of confounding variables, the Guidance lacks sufficient discussion of the importance of variable selection and data characteristics to ensure useful implementation

1 of the statistical procedures. Many of the nonparametric procedures rely upon bootstrap
2 procedures to obtain confidence intervals. This underscores the importance of using a
3 probability sampling procedure. The implications of different sample sizes should also be
4 more fully discussed. The Guidance states that an advantage of using quantile regression
5 (QR) is that it can provide direct estimates of percentiles of a distribution of Y values at
6 given X values, which may be better estimates of these values than provided by SLR when
7 the assumptions of SLR are not met. Uncertainty associated with estimating extreme
8 quantiles from "small" sample sizes is appropriately identified in the Guidance as a concern
9 for QR. However, small sample size is likely to present considerable challenges to any
10 nutrient criteria development approach, and the Guidance should provide a discussion of how
11 the amount of data may affect the utility of empirical stressor-response approaches.
12

- 13 • In the Guidance, more information must be provided regarding regression assumptions,
14 limitations, and diagnostic procedures. Although the Guidance should not be expected to
15 provide the same level of detail on the implementation of statistical procedures contained in a
16 statistics textbook, more information must be provided regarding regression assumptions,
17 limitations, and diagnostic procedures. The appropriateness of the regression methods will
18 depend on the assumptions and use restrictions of each method. Although the document
19 discusses many of the important assumptions, it would be helpful for this information to be
20 clearly summarized in a table. The table could include headings for each method such as use,
21 inherent assumptions, and specific remarks. In addition, the importance of regression
22 diagnostic procedures should be emphasized. Examples and specific references to additional
23 sources of information should be provided. This could include evaluating data with and
24 without outliers or unusual values.
25
- 26 • More guidance is needed on the interpretation of results from the listed regression
27 procedures. For example, how does one decide whether the results of quantile regression are
28 adequate for criterion development? In the discussion of logistic regression (p. 28, last
29 paragraph), nothing is said about whether the coefficients in this analysis are significantly
30 different from zero, or about the proportion of total deviance accounted for by the regression.
31 For multiple linear regression (p. 31) a reference (e.g., Kutner et al. 2004) is needed for
32 Akaike and the other methods listed in the third paragraph of the page.
33
- 34 • The role of, and options for, data transformations should receive considerably more
35 discussion in the Guidance. Data transformation may be appropriate in the development of
36 stressor-response relationships using regression analysis, but this topic (including the
37 associated back-transformation of slope estimates and confidence intervals to yield criteria)
38 should be more carefully developed. In reading the document, one wonders when the log-
39 transformation should be used to establish linear relationships or whether curvature that may
40 be present in raw data (with no transformation) should be characterized. In addition, the
41 document does not describe the range of data transformations that may be appropriate,
42 instead focusing only on the log-transformation. For example, regarding the nCPA presented
43 in Figure 24, would the analysis give the same result if it were based on TP data that were not
44 log transformed? It is not clear in the Guidance when to apply a linear method to
45 transformed data or a changepoint or discontinuous regression method to untransformed data.

1 As a start, a table like Table 6.5, "Linearizing Transformations" in Weisberg (1985, p. 142)
2 could be included in the Guidance, along with some explanation. Finally, "back-
3 transformation" has the potential to introduce bias into the criterion value if done incorrectly,
4 and this topic should be treated more completely to minimize that potential.
5

- 6 • The Guidance appropriately points out that regression relationships should generally not be
7 used to project conditions beyond the range of conditions used to develop the relationships.
8
- 9 • The Guidance is silent on how and when the results of multiple statistical procedures may be
10 integrated to support numeric criteria as an alternative to selecting "the best" model in
11 situations where a clearly preferred model does not emerge from the analysis. Rather than
12 presenting the statistical techniques strictly as alternatives, the document could describe how
13 these procedures can complement each other and provide a more robust picture of what an
14 appropriate criterion should be. For example, a linear regression whose residuals appear to
15 show the presence of curvature might also be evaluated with nCPA to evaluate the range of
16 stressor values over which the curved response occurs. Model averaging (Burnham and
17 Anderson, 2002) is recommended for use with multiple regression when slight changes in the
18 data lead to different final models.
19
- 20 • The Guidance provides a limited list of the statistical methods that could be explored to yield
21 useful criteria. If a data set includes censored values, maximum likelihood estimation can
22 provide an alternative to bivariate or multivariate linear regression that avoids the need to
23 substitute values such as one-half the detection limit for nondetects. In addition, parametric
24 multivariate methods including principal components analysis (PCA), discriminant function
25 analysis, cluster analysis, and others may also provide a useful means of incorporating
26 covariates in a stressor-response relationship. PCA may be used to describe a group of
27 correlated variables through a single equation. A number of nonparametric linear regression
28 approaches are also available, including the family of Kendall tests available from the U.S.
29 Geological Survey (Helsel and Hirsch, 1992; Helsel et al., 2006)
30
- 31 • A key and an associated appendix of case studies should be included in the Guidance to
32 explain the appropriate use of statistical methods and inherent assumptions and uncertainties.
33 Since choice of method(s) will depend on the nature of the data being modeled and on the
34 underlying assumptions, it would be useful to include in the Guidance some kind of key
35 giving an explanation of "which method to use when," with the inherent required
36 assumptions and uncertainties associated with each method. Better use of case studies (from
37 lakes, streams, estuaries) in an appendix could help show "why one approach works in a
38 particular situation and another does not." One case study should estimate the stressor-
39 response relationship when the data form a "wedge-shaped" scatterplot, a pattern commonly
40 observed in nutrient stressor-response relationships.
41
- 42 • Statistical rigor is essential to the development of scientifically defensible criteria. Simplistic
43 application of approaches in the Guidance can lead to stressor-response relationships with
44 poor predictive power and result in inappropriate numeric nutrient criteria. Therefore, EPA

will need to provide support to states in the training and use of these statistical methods. As previously stated, the use of bivariate methods (including nCPA) must involve a careful examination of potentially confounding variables to develop support for a predictive relationship. In order to properly evaluate the predictive power of empirical stressor-response relationships, uncertainties associated with each method used must be identified and quantified. Simulated data sets designed to contain specific properties that may be encountered by users of the Guidance could help communicate how these statistical procedures behave over a variety of data set characteristics (e.g., a range of uncertainty in the regression slope).

- The need for statistical rigor applies to both the strength and the form of the relationship among variables (i.e., evaluating the presence of curvature in a stressor-response relationship). The Guidance should describe the goal of data analysis as one of characterizing not only the strength of relationship but also its form, and the evidence supporting conclusions about both. This is particularly relevant when deciding to use nCPA or discontinuous regression to characterize a relationship. A more complete approach should be presented to test the hypothesis that a true data threshold exists.
- Perhaps most importantly, EPA should provide guidance on the degree of relationship (indicated by R^2 , residuals analysis, and other evidence) needed to establish sufficiently predictive stressor-response relationships. At a minimum, EPA should describe how to address the important question of “when is the evidence insufficient to support using a empirical stressor-response approach?” One suggestion is to better incorporate the EPA data quality objectives process into the Guidance (see U.S. EPA, 2009c).

Findings on indirect pathways

- The Committee notes that, with respect to approaches used to address indirect pathways of adverse effects, the Guidance currently does not contain a clear definition of the term “indirect pathway.” One definition follows in part from the caption of Figure 10 in the Guidance:

“Simplified diagram illustrating the causal pathway between nutrients and aquatic life use impacts. Nutrients enrich both plant/algal as well as microbial assemblages, which lead to changes in the physical/chemical habitat and food quality of streams. These effects directly impact the insect and fish assemblages. The effects of nutrients are influenced by a number of other confounding factors as well, such as light, flow, and temperature.”

This description appropriately indicates that nutrient concentrations directly impact plant/algal and microbial communities and indirectly impact insect and fish assemblages through impacts on plant/algal and microbial communities. As discussed previously, a challenge in using empirical approaches is establishing sufficient evidence to support conclusions of cause and effect so that relationships with adequate predictive power can be developed. The farther removed the response variables are from immediate responses of variations in nutrient concentrations, the more difficult it may be to demonstrate a useful

1 degree of predictive power. Guidance on the acceptable degree of uncertainty, and/or the
2 desired level of predictive power, may help users of the Guidance identify useful
3 relationships whether or not pathways are direct or indirect. On the other hand, empirical
4 methods alone are unlikely to effectively address indirect pathways of adverse effects. This
5 requires appropriate conceptual and mechanistic models, adequate site-specific data, and
6 experienced professional judgment.

7
8 *Findings on confounding variables and uncertainty*
9

- 10 • As previously discussed, exploratory data analysis that includes classification of data by
11 similarities in confounding variables prior to the evaluation of stressor-response relationships
12 may improve the predictive power of the relationships if sufficient data are available.
13 Incorporation of confounding variables in a multiple regression is also appropriate.
14
15 • Because uncertainty in the appropriate criterion value cannot be eliminated, it is prudent to
16 evaluate the potential consequences of varying degrees of uncertainty in a stressor-response
17 relationship on the resulting criteria and management objectives. This may be accomplished
18 in part through the use of the EPA data quality objectives (DQO) process or a similar
19 approach.
20
21 • References should be provided to direct the reader to more information on regression
22 diagnostics including leverage statistics and information on influential points. This would
23 assist the user in addressing uncertainties associated with these values. (One useful textbook
24 is Kutner et al., 2004; there are many others.)
25
26 • The Guidance should emphasize the importance of careful pairing of potential stressor and
27 response variables. Uncertainty in a stressor-response relationship may be increased if
28 incompatible data types are paired. For example, combining a seasonal average chlorophyll
29 a concentration calculated from multiple samples with a TP concentration obtained from a
30 single grab sample could introduce considerably more uncertainty than if both variables
31 represent seasonal averages. There are places in the guidance document where measured
32 values are presented without a clear description of the spatial or temporal components that
33 the value represents (on p. 22, for example, 15 ug/L chlorophyll a is presented as a threshold
34 between mesotrophic and eutrophic conditions without indicating the applicable averaging
35 period). The Guidance should consistently include such information in its descriptions of
36 various components of the threshold identification and criteria-setting process.
37

38 *Key recommendations concerning statistical methods in the Guidance*
39

40 The Committee provides the following key recommendations to address the comments and
41 findings presented above.
42

- 43 1. In the Guidance, EPA must provide more information on the supporting analyses needed
44 for each statistical method to correctly identify useful predictive relationships, and

- 1 acknowledge that the use of these statistical methods alone cannot provide sufficient
2 evidence of a cause-effect relationship.
3
- 4 2. The Guidance should indicate that response variables must in all cases have biological
5 relevance and be associated with designated uses.
6
- 7 3. The Guidance should emphasize that use of the statistical methods requires careful
8 consideration of confounding variables before the methods can be used as predictive
9 tools. As discussed above, further information on how to address confounding variables
10 should be included in the document.
11
- 12 4. The Guidance should contain additional discussion of the potential consequences of
13 varying degrees of uncertainty in a stressor-response relationship on the resulting criteria
14 and management objectives. This may be accomplished in part through the use of the
15 EPA DQO process or a similar approach.
16
- 17 5. The Guidance should contain more information on approaches that address multiple
18 factors, such as a stratified (or hierarchical) approach that considers other attributes
19 known to be important such as canopy, habitat, multiple nutrients, etc.
20
- 21 6. EPA should consider replacing the examples in the Guidance that provide low
22 explanatory power.
23
- 24 7. As discussed above, the Guidance should contain additional specific information (or
25 guidance on where to find it) on:
26
- 27 – The use of parametric (e.g., Pearson) and nonparametric (e.g., Spearman's rank,
28 Kendall's tau) correlation analyses.
 - 29 – The importance of variable selection (including careful pairing of stressor and
30 response variables) and data characteristics to ensure useful implementation of the
31 statistical procedures.
 - 32 – Regression assumptions, limitations, and diagnostic procedures.
 - 33 – Interpretation of results from the listed regression procedures.
 - 34 – The role of, and options for, data transformations.
 - 35 – How and when the results of multiple statistical procedures may be integrated to
36 support numeric criteria.
 - 37 – An appendix of case studies to explain the appropriate use of statistical methods
38 and inherent assumptions and uncertainties.
39
- 40 8. The Committee recommends that EPA consider providing technical support and training
41 to states and tribes to assist them in the use of the statistical methods in the Guidance.
42

9. The Guidance should describe the goal of data analysis as one of characterizing not only the strength of relationship but also its form, and the evidence supporting conclusions about both.

10. The Committee emphasizes that EPA should provide guidance on the degree of relationship (indicated by R^2 , residuals analysis, and other evidence) needed to establish sufficiently predictive stressor-response relationships for numeric nutrient criteria development.

3.6. Charge Question 6. Evaluating the predictive accuracy of stressor-response relationships

Section 4 of the draft guidance document describes how to evaluate the predictive accuracy of estimated stressor-response relationships. Please comment on the appropriateness of approaches in Section 4 of the guidance document and factors to consider in evaluating and comparing different estimates of the stressor-response relationships and selecting those most appropriate for criteria derivation.

Overall, the Committee notes that Section 4 of the Guidance lacks the detail provided in other sections and, as discussed below, needs improvement. The Committee finds that this section is particularly important because it addresses the reliability or “validity” of the approaches considered. The Guidance should provide information to help managers decide which criteria derivation approach to use (e.g., analysis of best fit by regression or some other means). These are important decisions and additional guidance on how to select the best tools would be helpful. If the proposed methods yield inaccurate results, this could lead to inappropriate or ineffectual solutions to comply with Clean Water Act goals. The Committee provides the following findings and comments in response to Charge Question 6.

- The Committee finds that a clear framework and criteria for statistical model selection is needed in the Guidance. This framework should include a set of decision tools and criteria used not only to determine which model fits best, but also to decide whether the stressor-response approach to criteria development is appropriate. Model selection criteria should include:
 - Capability of model to consider cause-effect and direct-indirect relationships between stressor and response;
 - Biological relevance;
 - Relevance to known mechanisms and existing conditions; and
 - Capability of model to predict probability of meeting designated use categories.

Findings on model validation

- More detail is needed in Subsection 4.1 of the Guidance to describe model validation techniques. In the Guidance there is limited discussion of validation of empirically derived stressor-response relationships. This is a critical component. Validation can be defined as

1 demonstrating the accuracy of the model for a specified use. Within this context, accuracy is
2 the absence of systematic and random error - in ecology they are commonly known as
3 trueness and precision, respectively. All models are by their nature incomplete
4 representations of the system they are intended to model but, in spite of this limitation,
5 models can be useful. Many discussions of mathematical modeling discriminate between
6 model confirmation (i.e., plausible, worthy of belief) and model verification (i.e., shown to
7 be true). Given the nature of the environmental stressor and response data, such stressor-
8 response models cannot be fully validated. EPA should provide much more detailed
9 validation guidance including four components:

- 10 – *Conceptual validation* concerns the question of whether the model accurately
11 represents the environmental system. This is largely qualitative and requires
12 consideration of the strength of the cause-effect relationships. To consider whether
13 the empirical model assumptions are credible, a conceptual model of factors affecting
14 the stressor-response relationship should be developed. For each of the proposed
15 methods, guidance should be provided with examples showing the mechanistic
16 reasoning behind the cause-effect assumptions and the direct-indirect responses of the
17 stressor and response variables. This should be supported by some experimental
18 evidence relevant to the context in which it is used (e.g., data needs appropriate for
19 lakes may be different than for streams). For each application of the empirical model,
20 experimental or observational data in support of the principles and assumptions
21 should be presented and discussed.
- 22 – *Algorithm validation* concerns the translation of model concepts into mathematical
23 formulae. It addresses questions such as: "Do the equations represent the conceptual
24 model?" "Under which conditions can simplifying assumptions be justified?" "Is
25 there agreement among the results from use of different methods (e.g., different
26 response variables) to solve the model?" For ecological stressor-response models,
27 these questions relate to the adequacy of the empirical models themselves for
28 describing the effects of nutrient enrichment on aquatic life.
- 29 – *Functional validation* concerns checking the model against independently obtained
30 observations. For this type of validation the Guidance recommends using additional
31 empirical observations (an alternative experimental dataset). However, this requires
32 more information than is usually available, and results may not be expected to be the
33 same from one dataset to another given the heterogeneity of environmental systems.
34 Such data cannot truly validate the stressor-response model per se, but may produce
35 valuable insights. Guidance is needed to answer questions such as: "what are the
36 minimum data requirements for validation?" and "if one is working with a limited
37 dataset, how does one consider the tradeoffs between using more data in the original
38 analysis and reserving data for validation?"
- 39 – *Software validation* concerns the implementation of mathematical formulae in various
40 computer software. This validation takes into consideration the possible effects of
41 software-specific factors on the model output (e.g., with regard to precision). For
42
43

- 1 example, problems have been documented with regard to performing statistical
2 analyses with some spreadsheet programs or open source codes.
3
- 4 • The Committee finds that the concept of “validation” as presented in Subsection 4.1 of the
5 Guidance is inconsistent with other EPA guidance (U.S. EPA, 2009a) on development,
6 evaluation, and application of models. In EPA’s other modeling guidance, model evaluation
7 includes model corroboration, and sensitivity and uncertainty analyses. Model corroboration
8 is defined as quantitative and qualitative methods for evaluating the degree to which a model
9 corresponds to reality. In practical terms, this is the process of “confronting models with
10 data.” In some disciplines, this process has been referred to as validation. EPA prefers the
11 term “corroboration” because it implies a claim of usefulness and not truth. The Committee
12 finds that this is not just a semantic distinction and we recommend that Subsection 4.1 of the
13 Guidance be revised so that it is consistent with U.S. EPA (2009a).
14
 - 15 • The use of data quality objectives (DQOs) should be discussed in Subsection 4.1 of the
16 Guidance. The DQOs should be established at the beginning of the criteria development
17 process (i.e., Guidance step one) but they can also be used to evaluate the potential stressor-
18 response models (Guidance step four). The discussion of DQOs should address levels of
19 uncertainty, Type I and Type II error rates, and the extent to which each model can predict
20 the probability of meeting designated use categories .
21
 - 22 • In Subsection 4.1, more detailed guidance should be provided on the use of randomly or non-
23 randomly selected data sets to help address questions about how much data should be held
24 out of the original analysis to adequately support the validation process. Subsection 4.1 is
25 intended to describe how to validate “the predictive performance of different models.”
26 Recommended approaches include: a) collecting new samples; and b) holding out a subset of
27 the original data from the analysis. Reserved samples may be selected randomly or non-
28 randomly. Authors of the Guidance appropriately note that a potential problem with using
29 random subsetting is that the covariance structure of the data is likely to be the same, so that
30 this approach may not provide an independent test of the predictive power of a relationship.
31 As stated in the Guidance, reserving a non-random subset may be a useful alternative. Some
32 discussion of the relative size of calibration and validation data sets is warranted.
33
 - 34 • The concept of “best fit” needs elaboration in the Guidance. Best fit is based on the
35 assumptions made and the model developed and, as previously discussed, there may be
36 considerable uncertainty even if a model is thoroughly and carefully developed.
37 Assumptions that are incorrect or incomplete will lead to erroneous criteria. Authors of the
38 Guidance understand this, and state that relationships can be confounded by unsampled or
39 unmodeled factors. This statement is true and it should be more fully discussed, and perhaps
40 given much greater weight in each section. EPA should consider whether each example in
41 the Guidance should be accompanied by a discussion of possible confounding issues and
42 what might be missing. The concept of uncertainty, its effect on model results, and ways to
43 at least understand the level of uncertainty are not fully described in the Guidance.
44

This Draft is made available for quality review and approval by the chartered Science Advisory Board. This Draft does not represent EPA policy.

- 1 • The Guidance should contain additional information to assess the closeness of root-mean-
2 square predictive error (RMSPE). The RMSPE as presented on p. 42 of the Guidance is a
3 well-recognized measure of how well a statistical model does in predicting response values
4 from given stressor values. Figure 27 of the Guidance gives an example where the RMSPE
5 for the calibration dataset was 0.28, while the RMSPE for the held-out validation data (from
6 a particular State) was 0.27. Many would agree that those two RMSPEs are "close." But it is
7 necessary to answer the question, "how close is close?" No further statements appear in the
8 Guidance about how to assess the closeness of two RMSPEs. Comparing 0.28 with 0.27 in a
9 single example does not help users of the Guidance extend this example to their own data
10 sets. It might be possible to take a bootstrap approach with regard to the calibration data set
11 to derive an actual distribution of values for the calibration RMSPE against which the
12 RMSPE of the validation data set could be compared. The Guidance does not address this.
13 In addition, it is appropriate to characterize fit quality using other information such as R^2 ,
14 residuals analysis, regression results, etc.
15
- 16 • With regard to validation, nutrient criteria should result from weight-of-evidence from the
17 application of multiple empirical approaches considering multiple response variables and
18 other approaches as appropriate. The nutrient criteria values determined after considering
19 validation and uncertainty may vary significantly from technique to technique or from
20 response variable to response variable. The Committee suggests that EPA consider the range
21 of responses and concordance among analyses/models and, as stated previously, establish
22 linkage between response variables and designated use categories. The Guidance should
23 discuss model averaging and should recommend considering the range of responses as a
24 measure of overall utility of the empirical approach. In addition, the Guidance should more
25 strongly advocate decision making based on weight-of-evidence from multiple empirical and
26 other approaches.
27

28 *Findings on qualitative assessment of the uncertainty of the estimated stressor-response*
29 *relationship*
30

- 31 • The Committee finds that Guidance Subsection 4.2 (addressing uncertainty) is too brief.
32 Given the importance of this cross-cutting issue, a section on uncertainty is needed for each
33 of the steps outlined in the Guidance, and uncertainty should be summarized at the end of the
34 document.
35
- 36 • Subsection 4.2 of the Guidance should address both qualitative and quantitative estimates of
37 uncertainty. Given reasonable expectations for data availability and inevitable limits on the
38 conceptual understanding of complex environmental systems, the Guidance should discuss
39 both qualitative and quantitative estimates of uncertainties. The Committee notes that an
40 explicit accounting of uncertainty is critical.
41
- 42 • Validity of the space-for-time substitution assumption can be supported by analysis of long-
43 term stressor-response data for selected data-rich sites. Subsection 4.2 of the Guidance states
44 that all stressor-response models estimated from cross-sectional or synoptic data must also

1 invoke the assumption that spatial differences in sites can be substituted for temporal
2 differences without a substantial degradation of model accuracy (i.e., the space-for-time
3 substitution). As the Guidance states, a good way to provide support for the validity of this
4 assumption is to analyze long-term stressor-response data for selected data-rich sites.
5

- 6 • As previously discussed, the Guidance should contain additional information about the
7 importance of considering “data bias” in interpreting the stressor-response results with regard
8 to predictive performance and uncertainty, and also the importance of uncertainty imposed
9 by model assumptions. Additional guidance is needed on to how to interpret data from a
10 particular environment (e.g., a data set based on lake data) and its appropriateness (or lack
11 thereof) for describing conditions more broadly. It would be helpful to include in the
12 Guidance examples of databases that would be “ideal” or appropriate for each empirical
13 model presented. For example, would the conceptual model for considering nutrient criteria
14 be ideally approached using seasonal data, data from shaded versus unshaded tributaries, data
15 from wadeable streams versus big rivers, and/or long versus short term averages of data
16 describing the stressor or the response?
17

18 *Findings on selection of the stressor-response model* 19

- 20 • The Committee notes that Subsection 4.3 of the Guidance should discuss grounding models
21 in reality through use of prior knowledge. A great deal is known about the effects of
22 nutrients on aquatic systems, and the relationships between variables should reflect that
23 knowledge. All models should be evaluated to determine whether they make sense
24 biologically (e.g., is the range of data used appropriate?; are they mechanistically sound?).
25
- 26 • Subsection 4.3 of the Guidance could be improved by providing a more detailed discussion
27 of how to decide when to use each method to model stressor-response relationships, and the
28 advantages/disadvantages associated with each method. Table 1 on page 44 of the Guidance
29 is not sufficient for this purpose. It would be beneficial to provide a case study using a single
30 data set to demonstrate the comparison of a range of model choices.
31
- 32 • The Committee notes that the stated objective of Subsection 4.3 in the Guidance,
33 “demonstrating how to select a stressor-response model using the response variable that best
34 represents the data,” is not the same as the goal of Section 4, “evaluating the predictive
35 accuracy of estimated stressor-response relationships.” Confidence in predictive accuracy
36 should be the primary consideration in model selection. Further, while it may ultimately be
37 necessary to select a single model, one should also understand the significance to criteria
38 derivation of selecting among reasonable alternative models or the effect of model averaging
39 when a single most appropriate model cannot clearly be identified.
40
- 41 • In Subsection 4.3 of the Guidance, more detail should be provided in the discussion of
42 conditions under which the last two methods, non-parametric changepoint analysis (nCPA)
43 and discontinuous regression, should be applied (other than simply stating that they should be
44 used when a direct designated use impairment threshold is unavailable). In addition, the

1 Committee notes that a curved response: 1) may or may not be real; 2) may or may not signal
2 an impaired designated use; and 3) may or may not be indicated at all by the data. Further, a
3 curved response may be modeled by one of the linear methods after transformation.
4

- 5 • The Committee notes that linear stressor-response functions may not provide high levels of
6 accuracy for nutrient criteria development. Six different methods are summarized in Table 1
7 of Subsection 4.3. The first four methods all assume that the stressor-response function can
8 be modeled sufficiently as a linear model or a generalized linear model. It is unlikely that
9 linear stressor-response functions can ever achieve high levels of accuracy across the many
10 different confounding variables and the many different physical, chemical and biological
11 characteristics of specific sites.
12

13 *Key recommendations concerning evaluating the predictive accuracy of estimated stressor-*
14 *response relationships*
15

16 As a consequence of the findings presented above, the Committee provides the following key
17 recommendations.
18

- 19 1. The Guidance should be revised to provide a clear framework for statistical model selection.
20 This framework should include a set of decision tools and criteria used not only to determine
21 which model fits best, but also whether the stressor-response approach to criteria
22 development is appropriate.
23
- 24 2. The Guidance should be revised to provide much more detailed model validation guidance.
25
- 26 3. Subsection 4.1 of the Guidance (Model validation) should be revised to:
27
 - 28 – Make it consistent with other EPA guidance (U.S. EPA, 2009a) on development,
29 evaluation, and application of models.
30
 - 31 – Provide more detailed information on the use of randomly or non-randomly selected
32 data sets to help address questions about how much data should be held out of the
33 original analysis to adequately support the validation process.
34
 - 35 – Elaborate upon assumptions and uncertainties in “best fit” determinations, and in
36 particular provide additional information to assess the closeness of root-mean-square
37 predictive error (RMSPE).
38
 - 39 – State that nutrient criteria should result from a weight-of-evidence approach based on
40 the application of multiple empirical approaches considering multiple response
41 variables as appropriate.
42

4. Subsection 4.2 of the Guidance should be revised to provide an expanded discussion of uncertainty. This section should address both qualitative and quantitative estimates of uncertainty as well as data bias.
5. Subsection 4.3 of the Guidance should be revised to:
 - Address grounding models in reality through use of prior knowledge.
 - Provide a more detailed discussion on how to decide when to use each method for modeling stressor-response relationships, and the advantages/disadvantages associated with each method.
 - Provide more detail regarding the conditions under which the last two methods, non-parametric changepoint analysis (nCPA) and discontinuous regression, should be applied.
 - Address inaccuracies associated with linear stressor-response functions.

3.7. Charge Question 7. Evaluating candidate stressor-response criteria

Section 5 of the draft guidance document describes how to evaluate the candidate stressor-response criteria. An approach is outlined for predicting conditions that might result after implementing different nutrient criteria. Please comment on uncertainties that would remain if water quality criteria for nutrients were based solely on estimated stressor-response relationships and in what ways other information/analysis would help address and possibly reduce this uncertainty.

Section 5 of the Guidance is an important part of the document because selection of criteria has environmental, social, and economic consequences. We provide the following comments and findings in response to Charge Question 7.

Findings on recognizing uncertainty

- As previously discussed, the Guidance does not address or partition inherent critical uncertainties in the stressor-response approach. The Guidance describes approaches that use a data-mining exercise to demonstrate a possible cause-effect relationship for the nutrient-ecosystem response. However, the document does not address or partition inherent critical uncertainties in the stressor-response approach which, as demonstrated in examples in the Guidance and in public presentations given to the Committee, can be extremely large (e.g., several orders of magnitude). Because of the demonstrated uncertainties, prediction from an empirical stressor-response model for a specific system of interest cannot always be interpreted as an accurate prediction of future conditions.

- Uncertainty also results from climatic or other environmental conditions under which studies were conducted. In addition to uncertainties documented in the Guidance and in the public presentations to the Committee, uncertainty also results from the climatic or other environmental conditions under which empirical studies were conducted and response models developed. Studies conducted over relatively limited conditions (e.g., wet or dry years) or short-term periods (e.g., base flows, summer) are unlikely to provide the robust response relationships required for criteria development.

Findings on reducing uncertainty

- A major uncertainty inherent in the Guidance is accounting for factors that influence biological responses to nutrient inputs. For criteria that meet EPA's stated goal of "protecting against environmental degradation by nutrients," the underlying causal models must be correct. Habitat condition is a crucial consideration in this regard (e.g., light [for example, canopy cover], hydrology, grazer abundance, velocity, sediment type) that is not adequately addressed in the Guidance. Thus, a major uncertainty inherent in the Guidance is accounting for factors that influence biological responses to nutrient inputs. Addressing this uncertainty requires adequately accounting for these factors in different types of waterbodies.
- Uncertainty in the water quality criteria for nutrients could be reduced by obtaining data from well-designed site-specific monitoring programs. If "water quality criteria for nutrients were based solely on estimated stressor-response relationships," a critical overall uncertainty would be understanding where, within the range of probabilities, a single waterbody to which the criteria are applied will fall. This, in effect, is uncertainty in the space-for-time assumption discussed in the Guidance. That is, if the criterion nutrient concentration developed using an approach involving data from multiple locations is exceeded, will the predicted response and designated use impairment occur at a single location of interest? This type of uncertainty can be reduced by obtaining data from well-designed site-specific monitoring programs. Such monitoring would focus on obtaining specific information on the variability in stressor and response variables and important covariates with a goal of better defining the interactions of multiple variables on the attributes affecting the designated uses of a waterbody. Measurement of actual biological responses would be appropriate, emphasizing variables that respond most directly to changes in nutrient concentrations. These are typically measures of primary productivity or primary producers, or water chemistry changes such as DO and pH. Where necessary, such data may be used to develop computer simulation models specific to the system of interest that can facilitate forecasting of stressors and associated responses.
- Numeric nutrient criteria developed and implemented without consideration of system specific conditions (e.g., from a classification based on site types) can lead to management actions that may have negative social and economic and unintended environmental consequences without additional environmental protection. The Committee emphasizes the importance of not only recognizing but also making allowance in the Guidance for conditions specific to the system of interest so that the resulting science allows the best management

1 decisions to be made. In this regard, as previously discussed, we recommend use of a tiered
2 weight-of-evidence approach to criteria development. Weight-of-evidence is typically used
3 to determine the tier at which uncertainty has been reduced sufficiently for informed
4 management decision making.
5

- 6 • The Guidance can be used to develop numeric nutrient criteria in a tiered, weight-of-evidence
7 assessment using appropriately modified EPA approved procedures together with other
8 approaches that address causation. Large uncertainties in the stressor-response relationship
9 and the fact that causation is neither directly addressed nor documented indicate that the
10 stressor-response approach using empirical data cannot be used in isolation to develop
11 technically defensible water quality criteria that will “protect against environmental
12 degradation by nutrients.” The Guidance can, however, be used in a tiered, weight-of-
13 evidence assessment (using appropriately modified U.S. EPA-approved procedures, e.g.,
14 EPA’s Causal Analysis/Diagnosis Decision Information System [CADDIS]), (U.S. EPA,
15 2009b).
16
- 17 • EPA should consider addressing the use of probabilistic modeling (using the distribution of
18 data in the model and re-sampling or simulating a new distribution) to better determine
19 significant stressor-response relationships. For instance, a statistically significant stressor-
20 response relationship can be derived that may represent only a small portion of the variability
21 in the data. Relying solely on this relationship would result in a tremendous amount of
22 uncertainty in the final criterion developed. A good example of this is Figure 14 (p. 25) of
23 the Guidance, which shows a statistically significant model that explains only 5% of the
24 variation in the data - meaning that 95% of the variation is not explained by the model.
25 Guidance on model selection is critical to reducing uncertainty. The selection of target
26 numeric criteria as outlined in the Guidance is enhanced by the attempt to predict post-
27 implementation conditions. However, the example used in Figures 29 and 30 of the
28 Guidance is confusing as it appears that the values are re-projected using one criterion value
29 (log TP=2) and the prediction analysis is made (i.e., that all 8 of the sites would still exceed
30 the criterion) using a different value (log TP=1.6).
31

32 *Findings on criteria application and monitoring for assessment*

33

- 34 • The approach presented in Section 5 of the Guidance should be revisited and possibly
35 replaced. It appears to be highly sensitive to the way that individual data points located
36 above a response threshold are distributed around the regression line. For example, in
37 Figures 30 and 31 of the Guidance, near the intersection of TP and chlorophyll a targets and
38 candidate criteria, more than half of the data points fall above the regression line which
39 reflects the best fit to all the data. Projecting back to lower TP concentrations for each of
40 these individual data points would force a lower TP criterion than would be the case if the
41 data were actually normally distributed around the regression line. In other cases, there may
42 be a "cluster" of data points below the regression line, and the back-projected TP criterion
43 would be higher than if all data points were distributed randomly about the regression line.
44

- 1 • The Guidance does not adequately address the important issue of continued monitoring and
2 assessment for adaptive management. With regard to application of numeric nutrient criteria,
3 Section 5 of the Guidance discusses comparison of predicted and observed data to evaluate
4 response(s), along the lines of adaptive targets. This intrinsically implies that continued
5 monitoring and assessment of concentration versus biological response is taking place.
6 While this is a good idea in principle, it is not clear from the Guidance that this is to be done,
7 how it is to be done, or at what scale it should be done. This is important because it relates to
8 the issue of measuring changes in indicators of biological response as nutrient inputs are
9 reduced to waterbodies. It is unclear how hereditary or legacy losses or inputs of N and P to
10 waterbodies will be considered and accounted for in such an empirical approach. This begs
11 the next set of questions facing water resource managers who establish targets for nutrient
12 loss reduction: “if no water quality improvement or indicator biological response is seen, are
13 the targets/criteria too high or are legacy nutrient inputs increasingly significant
14 contributors?” and “how long does it take dynamic ecosystems and watersheds to respond to
15 changing nutrient inputs?”
16
- 17 • The Guidance should address a number of questions to clarify how the evaluation of
18 candidate stressor-response criteria will occur, presumably through monitoring. These
19 questions include the following:
 - 20 – While a sound monitoring program will be essential, what form will this take?
 - 21 – At what level in time and space will monitoring be established to evaluate criteria?
 - 22 – Where, when, and how will samples be collected to establish a long-term monitoring
 - 23 program to clearly define and measure candidate response(s) to any changes in
 - 24 management and stressor inputs, as predicted by nutrient criteria?
 - 25 – How will monitoring be conducted to give a whole watershed assessment,
 - 26 considering all nutrient sources and stressors that are contributing spatially and
 - 27 temporally?
 - 28 – How will continued legacy stressor inputs (N and P) be distinguished from
 - 29 management change-related decreases? Internal recycling of nutrients can mask
 - 30 water quality improvements brought about by nutrient loss reductions resulting from
 - 31 land management changes.
 - 32
- 33 • The direct and indirect effects of best management practices should be captured in setting
34 numeric nutrient targets and evaluating responses to target reductions. Implementation of
35 practices to decrease nutrient losses or inputs to surface waters (i.e., best or beneficial
36 management practices [BMPs]) can influence other factors that will affect biological
37 response to nutrient loadings. For instance, riparian buffers are effective at removing
38 sediment and sediment-bound nutrients (particularly P), as well as removing N by uptake and
39 denitrification. However, they also provide shade and will influence stream water
40 temperature and thereby the stressor-response relationship. Such interactions should be
41 addressed in nutrient criteria development. In addition, the use of buffers, for example, will
42 influence the size of particulates or sediment in a stream or river that may affect the benthic

1 population dynamics or species diversity. These direct and indirect effects and complexities
2 should be captured in target setting and the evaluation of response to achieving target
3 reductions.
4

5 *Key Recommendations in response to Charge Question 7*
6

7 The Committee provides the following key recommendations to address the comments and
8 findings above.
9

10 *Key Recommendations with regard to recognizing uncertainty*
11

- 12 1. The Guidance needs to clearly indicate that the empirical stressor-response approach does not
13 result in cause-effect relationships; it only indicates correlations that need to be explored
14 further. For example, the words “cause-effect” should be removed from the title of Step two.
15
- 16 2. The Guidance should address partitioning the uncertainty among the various factors that are
17 involved in the stressor-response relationship for the specific region/system of interest.
18 Some variables may be irrelevant to the hypothesized model for that system.
19
- 20 3. The Guidance should better document the physical, chemical and biological variables
21 comprising the relationships (e.g., habitat, spatial, and temporal) that define the aquatic
22 system, and which may be important in modifying the relationship between nutrient
23 concentrations and observed endpoints. These factors need to be well documented so that the
24 uncertainty in the relationship between nutrient concentrations and measured endpoints can
25 be reduced.
26

27 *Key recommendations with regard to conceptual models and uncertainty description/analysis*
28

- 29 4. The Guidance should caution users about potential problems associated with using the
30 overall regression to predict conditions that might result after implementing different nutrient
31 criteria.
32
- 33 5. EPA should consider addressing the use of probabilistic modeling to better determine
34 significant stressor-response relationships.
35
- 36 6. The Guidance should address uncertainty resulting from climatic or other environmental
37 conditions under which studies were conducted.
38
- 39 7. EPA should avoid using “biased” databases (i.e., that do not contain the range of data
40 necessary to fully characterize a system of interest) to develop stressor-response
41 relationships.
42
- 43 8. When cross-sectional data are used to develop empirical models, the ranges of values for
44 stressors and responses in the cross-sectional data should fully encompass not only the

1-8-10 Science Advisory Board (SAB) Ecological Processes and Effects Committee Advisory Report

-- Do Not Cite or Quote --

This Draft is made available for quality review and approval by the chartered Science Advisory Board. This Draft does not represent EPA policy.

current conditions in systems of interest, but also the predicted values for the stressors and responses corresponding to removal of the designated use impairment.

9. The Committee recommends predicting conditions that might result after implementing different nutrient criteria and testing these conditions on specific data-rich systems of interest.
10. The Committee recommends that EPA frame uncertainty according to the following key issues:
 - What are the goals of the decision makers (e.g., what are the designated uses and when are they impaired?), and what amount of certainty is required to make that decision?
 - Are the mechanisms of the cause-effect relationship understood and are they reflected in the types of measurements recommended?
 - Do the variables measured reflect the goals of the Clean Water Act? In the examples presented in Section 5 of the Guidance species richness or chlorophyll a are not clearly linked to the stated goals (fishable, swimmable waters, etc).
 - Does the analysis tool reflect a known cause-effect relationship and does it allow an understanding of the process?
 - What are the *a priori* criteria to be met by the data? This must be established to make it possible to tell when the data cannot support the decision making process.

4. REFERENCES

- Adams, S.M. 2003. Establishing causality between environmental stressors and effects on aquatic ecosystems. *Human and Ecological Risk Assessment* 9:17-35.
- Benstead, J.P., A.D. Rosemond, W.F. Cross, J.B. Wallace, S.L. Eggert, K. Suberkropp, V. Gulis, J.L. Greenwood, and C.J. Tant. 2009. Nutrient enrichment alters storage and fluxes of detritus in a headwater stream ecosystem. *Ecology* 90:2556-2566.
- Burnham K.P. and D.R. Anderson. 2002. *Model Selection and multimodel Inference: A Practical Information-Theoretic Approach*. Springer-Verlag, NY, 488 pp.
- Burton, G.A., Jr., P.M. Chapman, and E.P. Smith. 2002. Weight-of-evidence approaches for assessing ecosystem impairment. *Human and Ecological Risk Assessment* 8:1657-73.
- Carleton, J.N., M.C. Wellman, A.S. Donigian, J.C. Imhoff, J.T. Love, R.A. Park and J.S. Clough. 2005. Nutrient Criteria Development with a Linked Modeling System: Methodology Development and Demonstration Case Studies for Blue Earth, Rum and Crow Wing Rivers, Minnesota. EPA-823-R-05-003. U.S. Environmental Protection Agency, Office of Water and Office of Science and Technology, Washington, DC.
- Cerco, C.F. and M.R. Noel. 2004. The 2002 Chesapeake Bay Eutrophication Model. EPA 903-R-04-004. U.S. Environmental Protection Agency, Region III, Chesapeake Bay Program Office, Annapolis, MD, and U.S. Army Corps of Engineers, Engineer Research and Development Center, Vicksburg, MS.
- Chapman, P.M. 2007. Determining when contamination is pollution – weight-of-evidence determinations for sediments and effluents. *Environment International* 33:492-501.
- Chapman, P.M., B.G. McDonald, and G.S. Lawrence. 2002. Weight-of-evidence frameworks for sediment quality and other assessments. *Human and Ecological Risk Assessment* 8:1489-1515.
- Collier, T.K. 2003. Forensic ecotoxicology: Establishing causality between contaminants and biological effects in field studies. *Human and Ecological Risk Assessment* 9:259-266.
- Conley, D.J., H.W. Paerl, R.W. Howarth, D.F. Boesch, S.P. Seitzinger, K.E. Havens, C. Lancelot and G.E. Likens. 2009. Controlling eutrophication: nitrogen and phosphorus. *Science* 323:1014-1015.
- Cormier, S.M., G.W. Suter, and S.B. Norton. 2010. Causal characteristics for ecoepidemiology. *Human and Ecological Risk Assessment* 16 (in press).
- Cross, W.F., J.B. Wallace, A.D. Rosemond, and S.L. Eggert. 2006. Whole-system nutrient enrichment increases secondary production in a detrital-based ecosystem. *Ecology* 87:1556-1565.

1-8-10 Science Advisory Board (SAB) Ecological Processes and Effects Committee Advisory Report

-- Do Not Cite or Quote --

This Draft is made available for quality review and approval by the chartered Science Advisory Board. This Draft does not represent EPA policy.

- 1 Cross, W.F., J.B. Wallace, and A.D. Rosemond. 2007. Nutrient enrichment reduces constraints
2 on material flows in a detritus-based food web. *Ecology* 88:2563-2575.
3
- 4 Florida Department of Environmental Protection. 2009. Draft Technical Support document:
5 Development of Numeric Nutrient Criteria for Florida Lakes and Streams. Standards and
6 Assessment Section, Tallahassee, FL [Available at:
7 <http://www.dep.state.fl.us/water/wqssp/nutrients>]
8
- 9 Fox, G.A. 1991. Practical causal inference for ecoepidemiologists. *Journal of Toxicology and*
10 *Environmental Health* 33:359-373.
11
- 12 Greenwood, J.L., A.D. Rosemond, J.B. Wallace, W.F. Cross, and H.S. Weyers. 2007. Nutrients
13 stimulate leaf breakdown rates and detritivore biomass: bottom-up effects via heterotrophic
14 pathways. *Oecologia* 151:637-649.
15
- 16 Hagy, J.D., W.R. Boynton, C.W. Keefe and K.V. Wood. 2004. Hypoxia in Chesapeake Bay,
17 1950-2001: Long-term change in relation to nutrient loading and river flow. *Estuaries* 27(4):634-
18 658.
- 19 Helsel D.R. and R.M. Hirsch. 1992. *Statistical Methods in Water Resources*. Elsevier, NY, 522
20 pp. [Available online at:
21 http://www.practicalstats.com/aes/aes/AESbook_files/HelselHirsch.PDF]
22
- 23 Helsel, D.R., Mueller, D.K., and Slack, J.R. 2006. Computer program for the Kendall family of
24 trend tests: *U.S. Geological Survey Scientific Investigations Report* 2005–5275, 4 p.
25
- 26 Hickey, G.L. 2008. Making species salinity sensitivity distributions reflective of naturally
27 occurring communities: using rapid testing and Bayesian statistics. *Environmental Toxicology*
28 *and Chemistry* 22:2403-2411.
29
- 30 Hill, W.R. and S.E. Fanta. 2008. Phosphorus and light colimit periphyton growth at subsaturating
31 irradiances. *Freshwater Biology* 53:215-225.
32
- 33 Hill, W.R., S.E. Fanta and B.J. Roberts. 2009. Quantifying phosphorus and light effects in stream
34 algae. *Limnology and Oceanography* 54:368-380.
35
- 36 Kutner M.H., C.J. Nachtsheim, J. Neter and W. Li. 2004. *Applied Linear Statistical Models*, 5th
37 Edition. McGraw-Hill Irwin, NY, 1396 pp.
38
- 39 Lewis, W.M., Jr. and W.A. Wurtsbaugh. 2008. Control of lacustrine phytoplankton by nutrients:
40 erosion of the phosphorus paradigm. *International Review of Hydrobiology* 93:446-465.
41
- 42 Leung, K.M.Y., A. Bjorgesester, J. Gray, W.K. Li, G.C.S. Lui, Y. Wang, and P.K.S. Lam. 2005.
43 Deriving sediment quality guidelines for field-based species sensitivity distributions.
44 *Environmental Science and Technology* 39:5148-5156.

1-8-10 Science Advisory Board (SAB) Ecological Processes and Effects Committee Advisory Report

-- Do Not Cite or Quote --

This Draft is made available for quality review and approval by the chartered Science Advisory Board. This Draft does not represent EPA policy.

- 1
2 Linder, S.H., G. Delclos, and K. Sexton. 2010. Making causal claims about environmentally-
3 induced adverse effects. *Human and Ecological Risk Assessment* 16 (in press).
4
5 Linkov, I., D. Loney, S. Cormier, F.K. Satterstrom, and T. Bridges. 2009. Weight-of-evidence
6 evaluation in environmental assessment: Review of qualitative and quantitative approaches.
7 *Science of the Total Environment* 401:5199-5205.
8
9 Maine Department of Department of Environmental Protection. 2009. *Nutrient Criteria for*
10 *Fresh Surface Waters*.
11 http://www.maine.gov/dep/blwq/rules/Other/nutrients_freshwater/index.htm . [Accessed on
12 October 25, 2009]
13
14 McLaughlin, K. and M. Sutula. 2007. *Developing Nutrient Numeric Endpoint and TMDL Tools*
15 *for California Estuaries: An Implementation Plan*. Southern California Coastal Water Research
16 Project Technical Report 540. Costa Mesa, CA [Available at:
17 ftp://ftp.sccwrp.org/pub/download/DOCUMENTS/TechnicalReports/540_CA_NNE_PhaseII.pdf
18]
19
20 Mississippi River/Gulf of Mexico Watershed Nutrient Task Force. 2008. *Gulf Hypoxia Action*
21 *Plan 2008 for Reducing, Mitigating, and Controlling Hypoxia in the Northern Gulf of Mexico*
22 *and Improving Water Quality in the Mississippi River Basin*. Washington, DC.
23
24 Paerl, H.W. 2009. Controlling eutrophication along the freshwater-marine continuum: dual
25 nutrient (N and P) reductions are essential. *Estuaries and Coasts* 32:593-601.
26
27 Peterson, B.J., J.E. Hobbie, A.E. Hershey, M.A. Lock, T.E. Ford, J.R. Vestal, V.L. McKinley,
28 M.A.J. Hullar, M.C. Miller, R.M. Ventullo, and G.S. Volk. 1985. Transformation of a tundra
29 river from heterotrophy to autotrophy by addition of phosphorus. *Science* 229:1383-1386.
30
31 Scavia, D., D. Justic and V.J. Bierman, Jr. 2004. Reducing hypoxia in the Gulf of Mexico:
32 Advice from three models. *Estuaries* 27(3):419-425.
33
34 Slavik, K., B. J. Peterson, L. A. Deegan, W.B. Bowden, A. E. Hershey, and J. E. Hobbie. 2004.
35 Long-term responses of the Kuparuk river ecosystem to phosphorus fertilization. *Ecology* 85(4):
36 939-954.
37
38 Stockner, J.G. and K.R.S. Shortreed. 1978. Enhancement of autotrophic production by nutrient
39 addition in a coastal rainforest stream on Vancouver Island. *Journal of the Fisheries Research*
40 *Board of Canada* 35:28-34.
41
42 Suter, G.W., II, S.B. Norton, and S.M. Cormier. 2002. A methodology for inferring the causes of
43 observed impairments in aquatic ecosystems. *Environmental Toxicology and Chemistry* 21:1101-
44 1111.
45

1-8-10 Science Advisory Board (SAB) Ecological Processes and Effects Committee Advisory Report

-- Do Not Cite or Quote --

This Draft is made available for quality review and approval by the chartered Science Advisory Board. This Draft does not represent EPA policy.

- 1 Suter, G.W., II, S.B. Norton, and S.M. Cormier. 2010. The science and philosophy of a method
2 for assessing environmental causes. *Human and Ecological Risk Assessment* 16 (in press).
3
- 4 Turner, R.E., N.N. Rabalais and D. Justic. 2008. Gulf of Mexico hypoxia: alternate states and a
5 legacy. *Environmental Science and Technology*. 42:2323-2327.
6
- 7 U.S. EPA. 2000a. *Nutrient Criteria Technical Guidance Manual: Rivers and Streams*. EPA-822-
8 B-00-001. U.S. Environmental Protection Agency, Washington, DC.
9
- 10 U.S. EPA. 2000b. *Nutrient Criteria Technical Guidance Manual: Lakes and Reservoirs*. EPA-
11 822-B00-001. U.S. Environmental Protection Agency, Washington, DC.
12
- 13 U.S. EPA. 2000c. *Stressor Identification Guidance document*. EPA/822/B-00/025, U.S. EPA
14 Office of Water, Washington, DC.
15
- 16 U.S. EPA. 2001. *Nutrient Criteria Technical Guidance Manual: Estuarine and Coastal Marine*
17 *Waters*. EPA-822-B-01-003. U.S. Environmental Protection Agency, Washington, DC.
18
- 19 U.S. EPA. 2006a. *Data Quality Assessment: A Reviewers Guide (QA/G-9R)*. EPA/240/B-06/002.
20 U.S. Environmental Protection Agency, Washington, DC.
21
- 22 U.S. EPA. 2006b. *Data Quality Assessment: Statistical Tools for Practitioners (QA/G9s)*.
23 EPA/B-06/003. U.S. Environmental Protection Agency, Washington, DC.
24
- 25 U.S. EPA. 2008. *Nutrient Criteria Technical Guidance Manual: Wetlands*. EPA-822-B-08-001,
26 U.S. Environmental Protection Agency, Washington, DC.
27
- 28 U.S. EPA. 2009a. *Guidance on the Development, Evaluation, and Application of Environmental*
29 *Models*. EPA/100/K-09/003. Office of the Science Advisor, Council for Regulatory
30 Environmental Modeling, U.S. Environmental Protection Agency, Washington, DC.
31
- 32 U.S. EPA. 2009b. *CADDIS: Helping Scientists Identify the Causes of Biological Impairments*.
33 <http://cfpub.epa.gov/caddis/> [Accessed September 15, 2009]
34
- 35 U.S. EPA. 2006c. *Quality Management Tools - Systematic Planning*.
36 <http://www.epa.gov/quality1/dqos.html> [Accessed November 11, 2009]
37
- 38 Weed, D.L. 2005. Weight-of-evidence: A review of concept and methods. *Risk Analysis*
39 25:1545-57.
40
- 41 Weisberg, S. 1985. *Applied Linear Regression*, 2nd Edition. John Wiley & Sons, New York, 324
42 pp.
43
- 44 Wickwire, T. and C.A. Menzie. 2010. The causal analysis framework: Refining approaches and
45 expanding multidisciplinary applications. *Human and Ecological risk Assessment* 16 (in press).