United States Environmental Protection Agency

Office of Water
Mail code 4304T

EPA-822-R-21-005
August 2021

# Ambient Water Quality Criteria to Address Nutrient Pollution in Lakes and Reservoirs



**Office of Science and Technology**

**Office of Water**

**U.S. Environmental Protection Agency**

**Washington, DC**

## Notices

This document has been drafted and approved for publication by the Health and Ecological Criteria Division, Office of Science and Technology, Office of Water, United States Environmental Protection Agency, and is approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

Cover photo by Dana Thomas, 2019

# Contents

Executive Summary ..... viii

1 Introduction and Background ..... 1

2 Problem Formulation ..... 3

2.1 Management Goals ..... 3

2.2 Assessment Endpoints and Risk Metrics ..... 4

2.2.1 Aquatic Life Use ..... 5

2.2.2 Recreational Use ..... 8

2.2.3 Drinking Water Source ..... 9

2.3 Risk Hypotheses ..... 9

2.4 Analysis Plan ..... 10

3 Analysis ..... 11

3.1 Data ..... 11

3.1.1 Biological Data ..... 13

3.1.2 Chemical Data ..... 14

3.1.3 Dissolved Oxygen and Temperature Profiles ..... 14

3.1.4 Mapped Data ..... 15

3.2 Stressor-Response Models ..... 15

3.2.1 Zooplankton Biomass ..... 16

3.2.2 Deep Water Hypoxia ..... 23

3.2.3 Microcystin Concentration ..... 40

3.2.4 Phosphorus-Chlorophyll *a* ..... 48

3.2.5 Nitrogen-Chlorophyll *a* ..... 58

3.3 Duration and Frequency ..... 62

4 Characterization ..... 65

4.1 Other Measures of Effect and Exposure ..... 65

4.2 Incorporating State Data ..... 67

4.3 Existing Nutrient-Chlorophyll *a* Models ..... 67
4.4 Limitations and Assumptions ..... 68
4.5 Deriving State-Specified Criteria ..... 70
5 References ..... 71
Appendix A: State Case Study: Chlorophyll *a*-Microcystin ..... A-1
A.1 Data ..... A-1
A.2 Statistical Analysis ..... A-1
A.3 Results ..... A-3
A.4 Criteria Derivation ..... A-7
Appendix B: State Case Study: Chlorophyll *a*–Hypoxia ..... B-1
B.1 Data ..... B-1
B.2 Statistical Analysis ..... B-1
B.3 Results ..... B-2
B.4 Criteria Derivation ..... B-5
Appendix C: State Case Study: Total Nitrogen–Chlorophyll *a* ..... C-1
C.1 Data ..... C-1
C.2 Statistical Analysis ..... C-1
C.3 Results ..... C-2
C.4 Criteria Derivation ..... C-3
Appendix D: Operational Numeric Nutrient Criteria ..... D-1
Appendix E: Example Workflow for Deriving Lake Nutrient Criteria ..... E-1
E.1 Identify Designated Uses ..... E-1
E.2 Compile Data ..... E-1
E.3 Management Decisions ..... E-2
E.4 Accounting for Interannual and Sampling Variance ..... E-6
E.5 Incorporating State Data in National Models ..... E-7

# Figures

Figure 1. Conceptual model linking increased nutrients to aquatic life use (*Source*: US EPA 2010a). 5

Figure 2. Conceptual model linking increased nutrient concentrations to public health endpoints. 9

Figure 3. Simplified conceptual model showing pathways selected for analysis. 10

Figure 4. Schematic of network of relationships for modeling zooplankton biomass. 16

Figure 5. Relationships between measured biovolume, Chl *a*, and estimated seasonal mean phytoplankton biovolume. 20

Figure 6. Estimated relationships between zooplankton and Chl *a* for lakes > 8 m deep.. 21

Figure 7. Locations of NLA lakes used to fit the $DO_m$ model. These lakes were designated as dimictic and were stratified at the time of sampling. 24

Figure 8. Illustrative examples of depth profiles of temperature, temperature gradient, and DO. 25

Figure 9. Schematic of hypoxia model 26

Figure 10. Chl *a* vs. $DO_m$. $DO_m$ values 29

Figure 11. Demers and Kalff (1993) predicted stratification day vs. model mean estimate 30

Figure 12. Relationships between individual predictors and $DO_m$, holding other variables fixed at their mean values 31

Figure 13. Model predicted $DO_m$ vs. observed $DO_m$. *Open circles*: individual samples; *solid line*: 1:1 relationship. 31

Figure 14. Contours of modeled mean lake temperature computed at the overall mean elevation and mean sampling day. 33

Figure 15. Relationship between lake temperature and sampling day (*left panel*) and elevation (*right panel*) 33

Figure 16. Days of the year that mixed layer temperatures decrease below the critical temperature for cool-water species 35

Figure 17. Days of the year that mixed layer temperatures decrease below the critical temperature for cold-water species. 35

Figure 18. Simplified DO profile used to compute threshold for $DO_m$. *Open circle*: the targeted condition of DO at 5 mg/L, 30 cm below the thermocline. 37

Figure 19. Effects of other predictors on Chl *a* criteria. 38

Figure 20. Schematic showing relationship between different variables predicting microcystin. 41

Figure 21. Modeled relationships for the microcystin model. 44

Figure 22. Example of derivation of Chl *a* criterion to protect recreational uses based on targeted microcystin of 8 µg/L and exceedance probability of 1%. 45

Figure 23. Probability of the indicated number of excursions observed in 10-day assessment windows during a 100-day season for different single day exceedance probabilities. 48

Figure 24. Schematic representation of compartment model for TP. $P_{diss}$: dissolved P; Chl: Chl *a*; Turb: total turbidity; $Turb_{np}$: turbidity attributed to nonphytoplankton sources 49

Figure 25. Turbidity vs. Chl *a*. *Solid line*: the limiting relationship between Chl *a* and turbidity when contribution of allochthonous sediment is negligible. 53

Figure 26. Relationship between $Turb_{np}$, $P_{diss}$, and lake depth. *Open circles*: mean estimate of parameter value in each of 30 lake depth classes. 53

Figure 27. Ecoregion-specific values of $\log_e(d_1)$, P bound to nonphytoplankton suspended sediment. 54

Figure 28. TP versus $Turb_{np}$ and Chl *a*. 55

Figure 29. Example of deriving TP criteria for a Chl *a* target of 10 µg/L for data from one ecoregion (Southeastern Plains) 56

Figure 30. Variation in the concentration of N bound in phytoplankton among Level III ecoregions at the overall mean Chl *a* = 9.3 µg/L. Gray scale shows N concentrations in µg/L. 59

Figure 31. Variation in DON among Level III ecoregions at an overall mean DOC = 5.6 mg/L. Gray scale shows N concentrations in µg/L. ..... 60

Figure 32. TN-DIN vs. Chl *a* and DOC. ..... 61

Figure 33. Illustrative example of deriving TN criteria for a Chl *a* target of 10 µg/L for one ecoregion (Southeastern Plains) ..... 62

Figure 34. Variation in the relationship between Chl *a* and cyanobacterial-relative biovolume among states. ..... A-4

Figure 35. Comparison of Chl *a*/cyanobacterial-relative biovolume relationships in Iowa. ..... A-5

Figure 36. Comparison of predicted relationship between Chl *a* and microcystin for the state-national model (*left panel*) and a model using only Iowa state data (*right panel*). ..... A-6

Figure 37. Microcystin and Chl *a* measurements in Iowa. ..... A-7

Figure 38. Observed $DO_m$ vs. Chl *a*, sampling day, DOC, and depth below the thermocline. ..... B-3

Figure 39. Estimated first day of stratification for Missouri lakes (*left panel*) and NLA lakes (*right panel*). ..... B-4

Figure 40. Model coefficients estimated for models for Missouri data, NLA data, and combined data. ..... B-4

Figure 41. Relationships between day of year and $DO_m$ for six Missouri lakes. ..... B-5

Figure 42. Relationship between Chl *a* and $DO_m$ in an illustrative lake with depth below thermocline at 10m, DOC at 1.6 mg/L, and 130 days after spring stratification. ..... B-6

Figure 43. Chl *a* vs. TN-DIN in Iowa. ..... C-2

Figure 44. Chl *a* vs. TN-DIN in Beeds Lake, Iowa. ..... C-3

Figure 45. Lake-specific criteria derivation using combined Iowa-NLA model for two different lakes in Iowa. ..... C-4

Figure 46. Example distribution of 10 TP measurements ..... D-2

Figure 47. Example of defining an operational criterion magnitude. ..... D-4

# Executive Summary

Excess nutrients can degrade the condition of water bodies worldwide; in lakes and reservoirs (hereafter, referred to only as "lakes" unless noted otherwise), the effects of excess nitrogen (N) and phosphorus (P) may be particularly evident. High levels of nutrient loading commonly stimulate excess growth of phytoplankton, which can limit the recreational use of lakes. Overabundant phytoplankton also increase the amount of organic matter in a lake, which, when decomposed, can depress dissolved oxygen (DO) concentrations below the levels needed to sustain aquatic life. In extreme cases, the depletion of DO causes fish kills. Nutrient pollution can stimulate the excess growth of nuisance phytoplankton, such as cyanobacteria, which can produce cyanotoxins that are toxic to animals and humans. Elevated concentrations of cyanotoxins can reduce the suitability of a lake for recreation and as a source of drinking water.

Numeric nutrient criteria provide an important tool for managing the effects of nutrient pollution by providing nutrient goals that ensure the protection and maintenance of designated uses. The United States (U.S.) Environmental Protection Agency (EPA) published recommended numeric nutrient criteria for lakes and reservoirs in 2000 and 2001 for 12 out of 14 ecoregions of the conterminous United States. Those criteria were derived by analyzing available data on the concentrations of total nitrogen (TN), total phosphorus (TP), chlorophyll *a* (Chl *a*), and Secchi depth.

Scientific understanding of the relationships between nutrient concentrations and deleterious effects in lakes has increased since 2001, and standardized, high-quality data collected from lakes across the United States have become available. In this document, EPA describes analyses of these new data and provides models from which numeric nutrient criteria can be derived. The criterion models replace the recommended numeric nutrient criteria of 2000 and 2001 and are provided in accordance with the provisions of Section 304(a) of the Clean Water Act (CWA) (Title 33 of the *United States Code* [U.S.C.] § 1314(a)) for EPA to revise ambient water quality criteria from time to time to reflect the latest scientific knowledge. CWA Section 304(a) water quality criteria serve as recommendations to states and authorized tribes for defining ambient water concentrations that will protect against adverse effects to aquatic life and human health. The ecological and health protective responses on which the criterion models are based were selected by applying a risk assessment approach to explicitly link nutrient concentrations to the protection of designated uses.

The criterion models are nonregulatory, and this document describes the models and their inputs. States may use the recommended models to derive candidate nutrient criteria for each applicable designated use and, after demonstrating that the criteria protect the most sensitive designated use, adopt the criteria into their state standards. States may also modify the criteria to reflect site-specific conditions or establish criteria based on other scientifically defensible methods (Title 40 of the *Code of Federal Regulations* [CFR] 131.11(b)). The updated recommended CWA Section 304(a) nutrient criteria for lakes do not compel a state to revise current EPA-approved and adopted criteria, total daily maximum load nutrient load targets, or N or P numeric values established by other scientifically defensible methods. As part of their triennial review, if a state uses its discretion to not adopt new or revised nutrient criteria based on these CWA Section 304(a) criterion models, then the state shall provide an explanation when it submits the results of its triennial review (40 CFR 131.20(a)).

Following the risk assessment paradigm, EPA first defined water quality management goals for numeric nutrient criteria, and then defined assessment endpoints and metrics that are associated with achieving these goals and are sensitive to increased nutrient concentrations. The water quality management goals are articulated as designated uses in Section 101(a)(2) of the CWA (33 U.S.C. § 1251) (i.e., the protection and propagation of fish, shellfish, and wildlife [aquatic life] and recreation in and on the water). Another common designated use for lakes is to serve as drinking water sources. Excess loads of nutrients can lead to excessive growth of phytoplankton that can adversely impact designated uses in different ways, described below as assessment endpoints and metrics. EPA modeled stressor-response relationships using these endpoints and metrics to derive recommended numeric nutrient criterion models (Table 1).

Table 1. Summary of designated uses and associated measures of effect and exposure

| Designated use | Assessment endpoint | Risk metric | Applicability |
|---|---|---|---|
| Aquatic life | Zooplankton biomass | Rate of change of zooplankton biomass relative to phytoplankton biomass | All lakes |
| Aquatic life | Cool- and cold-water fish | Daily depth-averaged DO below the thermocline | Seasonally stratified lakes with cool- or cold-water fish |
| Recreation | Human health | Microcystin concentration to prevent liver toxicity in children | All lakes |
| Drinking water | Human health | Microcystin concentration to prevent liver toxicity in children | All lakes |

For aquatic life, EPA identified two assessment endpoints. The first endpoint is zooplankton biomass, and the risk metric is the slope of the relationship between zooplankton and phytoplankton biomass, which quantifies the degree to which energy produced by phytoplankton at the base of the food web is transferred to zooplankton and subsequently to higher trophic levels. When excess nutrients are available, phytoplankton biomass can increase at rates that exceed the capacity of zooplankton to consume. The threshold for the risk metric is one in which the rate of change of zooplankton biomass relative to phytoplankton biomass is approximately zero. This condition describes a lake in which the biomass of grazing biota (i.e., zooplankton) does not increase with increases in food (i.e., phytoplankton), and primary production at the base of the food web is weakly linked to production at higher trophic levels. This endpoint applies to all lakes in the conterminous U.S.

The second aquatic life endpoint is health and survival of cool- and cold-water fish, and the risk metric is the DO concentration in deep water that protects against mortality of these fish. Excess nutrients typically increase primary productivity, which then increases the amount of organic matter in a lake. Then, in the deep waters of a lake, DO is consumed as this organic matter is decomposed, leading to hypoxic and anoxic conditions. The risk metric is defined more specifically as the daily DO concentration, calculated as a depth-averaged value below the thermocline, which can be reduced to concentrations insufficient to support some fish species during the critical period of the summer when they require deep, cold waters to escape high temperatures at shallower depths. This endpoint applies to seasonally stratified lakes harboring cool- and cold-water fish.

For recreational uses and drinking water sources, the assessment endpoint is human health. For recreational uses, EPA selected the threshold for the risk metric as the concentration of microcystin associated with adverse effects on children (specifically, liver toxicity) from incidental ingestion of water during recreation. When excess nutrients are available, phytoplankton communities can shift toward a greater abundance of cyanobacteria that can release cyanotoxins, and microcystins are the most commonly monitored and measured freshwater cyanotoxin in the U.S. The threshold for the risk metric is 8 micrograms per liter (µg/L), based on recently published national recommendations for human health recreational water quality criteria and swimming advisories for cyanotoxins (US EPA 2019). For the drinking water use, EPA selected the threshold for the risk metric as the concentration of microcystins

associated with adverse effects on children resulting from oral exposure to drinking water (0.3 µg/L), consistent with the health advisory for microcystins (US EPA 2015b). The microcystin concentration from the health advisory applies to finished drinking water; however, EPA is aware that states and authorized tribes express water quality standards for protecting drinking water sources as either protecting the ambient source water before treatment or after treatment. The ability of treatment technologies to remove microcystin is too variable for EPA to set a national recommendation for an ambient source water concentration that would yield a protective concentration after treatment. If a state or authorized tribe applies the health advisory standard to drinking water after treatment, then they can account for the expected treatment in their facilities and select a higher microcystin concentration in the ambient source water that would result in the targeted microcystin concentration in the finished drinking water.

Data used in this analysis were collected in EPA's National Lakes Assessment (NLA), which sampled lakes across the conterminous U.S. in 2007 and 2012. Most of the sampled lakes were selected randomly so the resulting data represent the characteristics of the full population of lakes in the conterminous U.S. At each lake, standardized protocols were used to collect extensive measurements of biotic and abiotic characteristics.

This document describes statistical stressor-response models that relate Chl *a* concentrations to each of the risk metrics and that relate TN and TP concentrations to Chl *a*. A hierarchical Bayesian network is specified for each model to represent the effects of different variables on the relationship of interest. For example, microcystin is related to cyanobacteria biovolume, which is then linked to Chl *a* concentration. The models directly represent the processes that govern the relationships of interest and facilitate the use of other data sets in conjunction with data from EPA's NLA. When coupled with the targets for each response, the models provide candidate Chl *a*, TN, and TP criteria recommendations that states may then use with state risk management decisions to demonstrate they are protective of different designated uses. For lakes with multiple use designations, the states shall adopt criteria from these candidate criteria that protect the most sensitive use.

Models provided in this document are based on national data, but states often collect extensive data during routine monitoring. Incorporating local data into the national models can refine and improve the precision of the stressor-response relationships. In the appendices of this

document, EPA describes three case studies in which state monitoring data have been combined with national data, yielding models that can be used to derive recommended numeric nutrient criteria that account for both unique local conditions and national, large-scale trends.

# 1 Introduction and Background

Nutrient pollution, or the excess loading of nitrogen (N) and phosphorus (P), can degrade the conditions of water bodies worldwide, and in lakes the effects of nutrient pollution are often most evident. One visible consequence of nutrient pollution in lakes and reservoirs (hereafter, referred to only as "lakes" unless noted otherwise) is cultural eutrophication, an increase in primary productivity and algal abundance that increases the amount of organic matter in a water body (Smith et al. 2006, Smith and Schindler 2009). Decomposition of organic matter reduces dissolved oxygen (DO) concentrations in the water column, especially in deeper waters under stratified conditions. These hypoxic conditions are inhospitable to most aquatic species and reduce their ability to survive within a particular lake (Jones et al. 2011, Scavia et al. 2014).

Nutrient pollution also favors the growth of undesirable, nuisance phytoplankton (e.g., cyanobacteria), some of which produce cyanotoxins (Paerl and Otten 2013). Many species of cyanobacteria are superior competitors for light compared to other phytoplankton. Hence, in lakes with nutrient pollution, cyanobacteria can dominate by reducing the light available to other phytoplankton (Carey et al. 2012). A number of other mechanisms, including superior uptake rates for carbon dioxide and an ability to migrate vertically in the water column, also may explain the frequent occurrence of cyanobacteria dominance in eutrophic systems (Dokulil and Teubner 2000). Cyanobacteria dominance can interfere with the designated uses of a lake because cyanobacteria not only can form unsightly and odorous surface scums (reducing the aesthetic appeal of the lake for recreation) (Paerl and Ustach 1982), but also can produce cyanotoxins that can limit the use of the lake as both a source of drinking water and for recreation (Cheung et al. 2013). Many species of cyanobacteria are also less palatable than other phytoplankton to grazing organisms, and so, increases in cyanobacterial abundance can alter lake food webs and reduce the efficiency with which energy from primary production is transferred to higher trophic levels (Elser 1999, Filstrup et al. 2014a, Heathcote et al. 2016).

Nutrient pollution in lakes and resulting adverse environmental effects are widespread in the United States (U.S.). Nutrient pollution occurs in lakes of different sizes, in catchments with varying land uses, and in different climates. The U.S. Environmental Protection Agency (EPA) has long recognized the effects of nutrient pollution and has recommended that states and authorized tribes (hereafter, "states"), acting under their Clean Water Act (CWA) authorities,

adopt numeric nutrient criteria as one way to facilitate the management of these effects. A state's numeric nutrient criteria (1) provide nutrient goals to protect and maintain the designated uses of a water body (Title 33 of the *United States Code* [U.S.C.] § 1313(c)), (2) provide thresholds that allow the state to make accurate water quality assessment decisions (33 U.S.C. § 1313(d)), and (3) provide targets for restoration of water bodies that can guide waste load allocation decisions (33 U.S.C. § 1313(d)). To assist states and authorized tribes in deriving numeric nutrient criteria, EPA has published a series of technical support documents on methods for deriving criteria for lakes and reservoirs (US EPA 2000a), streams and rivers (US EPA 2000b), wetlands (US EPA 2008), and estuaries and coastal waters (US EPA 2001). A technical support document on using stressor-response relationships for deriving numeric nutrient criteria has also been published (US EPA 2010a). In 2000 and 2001, under its authority described in Section 304(a) of the CWA (33 U.S.C. § 1314(a)), EPA issued 12 documents that provided recommended numeric nutrient criteria for lakes, streams, and rivers in different ecoregions of the U.S. These criteria were derived by using available monitoring data to estimate the concentrations of total nitrogen (TN) and total phosphorus (TP) that were expected to occur in least-disturbed reference water bodies in different nutrient ecoregions.

In accordance with the provisions of Section 304(a) of the CWA, which directs EPA to revise ambient water quality criteria from time to time to reflect the latest scientific knowledge, EPA is issuing revisions to numeric nutrient criteria recommendations for lakes based on analyses of newly available, national-scale data and reflecting advances in scientific understanding of the relationship between excess nutrients and adverse effects in lakes. The criterion recommendations are models that generate numeric nutrient criteria based on national data and state risk management decisions. State data, if available, can be incorporated into the national criterion models to compute relationships that more accurately represent local conditions. In deriving these models, EPA uses a risk assessment framework (Norton et al. 1992; US EPA 1998, 2014) to identify assessment endpoints that relate directly to the water quality management goals for U.S. lakes specified by the CWA and that are sensitive to increased concentrations of N and P. Then, EPA uses stressor-response analysis to estimate relationships between increased N and P (estimated by measurements of TN and TP) and different risk metrics directly linked to the assessment endpoints (US EPA 2010a). National criterion models are provided for both TN and TP as the simultaneous control of both nutrients provides the most effective means of controlling the deleterious effects of nutrient pollution

(US EPA 2015a, Paerl et al. 2016). These stressor-response criterion models replace the ecoregion-specific nutrient criteria recommended previously for lakes,[1] which were derived using a reference distribution approach.

The remaining sections of this document are organized broadly according to the steps of risk assessment: (1) problem formulation, (2) analysis, and (3) characterization. The purpose of this document is to provide the technical details underlying the estimation of relationships between increased nutrient concentrations and different responses, as well as details regarding the derivation of numeric nutrient criterion recommendations using the national models. States may use these criterion models to derive candidate nutrient criteria and, after demonstrating that the criteria protect designated uses, adopt the criteria into their state water quality standards. States may also modify the criteria to reflect site-specific conditions or establish criteria based on other scientifically defensible methods (40 CFR 131.11(b)). For waters with multiple use designations, the state shall adopt criteria from the candidate criteria that support the most sensitive designated use (40 CFR 131.11(a)(1)). Water quality standards adopted by states are subsequently subject to review by EPA, pursuant to Section 303(c) of the CWA (33 U.S.C. § 1313(c)).

# 2 Problem Formulation

Problem formulation is the first step in a risk assessment, in which "…the problem is defined, and a plan for analyzing and characterizing risk is determined" (US EPA 1998). More specifically, during problem formulation, the management goals are articulated, assessment endpoints and risk metrics are selected, and conceptual models are specified that link stressors with assessment endpoints.

## 2.1 Management Goals

EPA focused on protecting uses that reflect management goals articulated in Section 101(a)(2) of the CWA (33 U.S.C. § 1251), which include maintaining conditions so different water bodies support aquatic life use (i.e., providing for the protection and propagation of fish, shellfish, and

[1] Ecoregional nutrient criteria for lakes and reservoirs (EPA 822-B-00-007, EPA 822-B-01-008, EPA 822-B-01-009, EPA 822-B-01-010, EPA 822-B-00-008, EPA 822-B-00-009, EPA 822-B-00-010, EPA 822-B-00-011, EPA 822-B-00-012, EPA 822-B-00-013, EPA 822-B-00-014, and EPA 822-B-01-011)

wildlife), recreation (i.e., providing for recreation in and on the water), and use of the water body as a source of drinking water. Under the CWA, it is a state's responsibility to designate uses for its waters, and many states have designated uses that provide for aquatic life and recreation uses. Some states have also designated waters as sources of drinking water. EPA focuses on aquatic life, recreation, and drinking water source because they represent uses that are particularly sensitive to increased concentrations of N and P. States can derive candidate nutrient criteria for each of the applicable designated uses in their lakes and, by comparing these criteria, identify the most sensitive use. Water quality criteria adopted by states for waters with multiple use designations must support the most sensitive use (40 CFR 131.11(a)).

## 2.2 Assessment Endpoints and Risk Metrics

The next step in problem formulation is to define assessment endpoints that can be used to quantify attainment of the management goals. Each of the management goals expressed in terms of different designated uses was associated with different assessment endpoints. Protection of recreational uses and drinking water sources pertains to public health rather than ecological health, and hence, the assessment endpoint is human health for these two designated uses. For aquatic life, the procedures of ecological risk assessment were followed to select assessment endpoints defined as "explicit expressions of the actual environmental values that are to be protected" (US EPA 1998). Three considerations guided the selection of these endpoints: ecological relevance, susceptibility to the stressor of interest (i.e., increased nutrient concentrations in the present case), and relevance to management goals.

After selecting the assessment endpoints, EPA developed conceptual models that represented current understanding of the linkages between increased N and P concentrations and effects on the assessment endpoint and management goals (Figure 1). The conceptual models were used to select specific risk metrics that quantified key steps along the causal path linking increased N and P concentrations to deleterious effects on aquatic life and public health. The final selections for the criterion recommendations were also influenced by the availability of data at the continental spatial scales considered in this analysis. These risk metrics were used as the response variables in stressor-response analysis. For a narrative description of the conceptual model, refer to *Using Stressor-Response Relationships to Derive Numeric Nutrient Criteria* (US EPA 2010a).

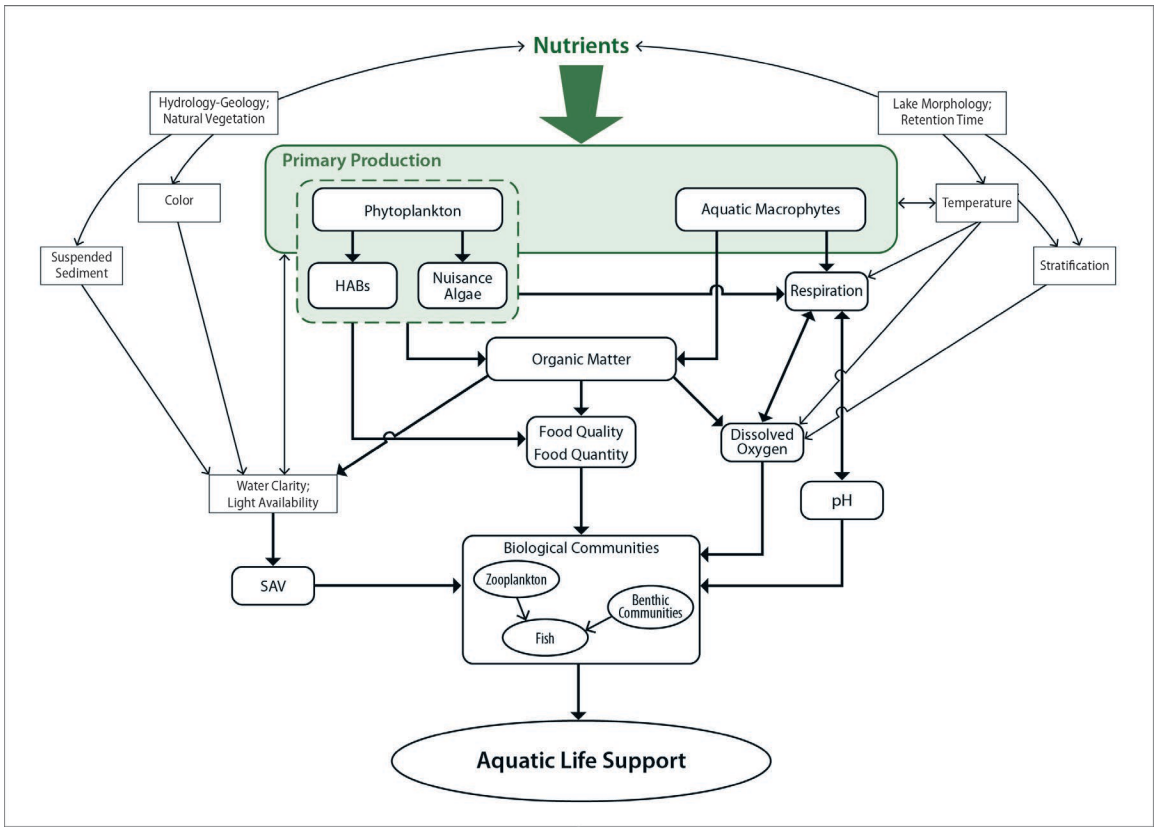


Figure 1. Conceptual model linking increased nutrients to aquatic life use (*Source*: US EPA 2010a).

### 2.2.1 Aquatic Life Use

Nutrient pollution and eutrophication can affect the health of the lake biological community via many pathways (see Figure 1). As discussed earlier, increased nutrients typically stimulate primary productivity and increase the amount of organic matter in a lake. Decomposition of the organic matter depletes the DO in the water, reducing the suitability of deeper waters as habitat for fish and invertebrates (Cornett 1989). Increased production and respiration also can increase the range of acidity throughout the day-night cycle in some lakes (Schindler et al. 1985), reducing the suitability of shallow waters as habitat for certain species. Increased algal biomass also reduces water clarity, and the reduction in light availability limits the depths at which submerged aquatic vegetation can persist (Phillips et al. 2016). Reduced water clarity can also shift fish assemblage composition away from species that depend on sight for foraging (De Robertis et al. 2003). Further, high nutrient concentrations favor the growth of cyanobacteria, which are less palatable to grazing species than other phytoplankton, altering the food web of the lake (Haney 1987).

EPA selected two assessment endpoints to characterize the health of aquatic life in lakes: (1) zooplankton biomass, which is applicable to all lakes, and (2) the survival and growth of cool- and cold-water fish in seasonally stratified lakes. For the second endpoint, EPA selected depth-averaged DO concentration as the risk metric. In seasonally stratified lakes with cool-water fish, criteria based on zooplankton biomass and DO can be compared, and the more stringent criterion is applied to ensure that aquatic life is protected. Collectively, the two assessment endpoints provide a broad assessment of the health of the lake biological community. Data were also available for each endpoint, and each endpoint quantified well-studied effects of nutrient pollution.

#### *2.2.1.1 Zooplankton biomass*

The rate of change of zooplankton biomass compared to the rate of change of phytoplankton biomass quantifies changes in the shape of biomass pyramids in lakes (Elton 1927). Biomass pyramids provide a graphical depiction of the amount of biomass at different trophic levels, and typically, the biomass of primary producers (at the bottom of the pyramid) exceeds the biomass of herbivores and carnivores at successively higher levels of the pyramid. In lakes, the ratio of herbivore biomass (i.e., zooplankton) to primary producer biomass (i.e., phytoplankton) (Z:P) has been observed to decrease along eutrophication gradients (Leibold et al. 1997). Reasons for the decreasing trend in Z:P have been the subject of some debate, much of which centers on the relative importance of top-down versus bottom-up food web effects. For zooplankton, top-down forces consist mainly of the effects of planktivore fish consuming zooplankton biomass (Jeppesen et al. 2003) and bottom-up forces include changes in the quantity and quality of the phytoplankton assemblage on which zooplankton feed (Filstrup et al. 2014a). With excess nutrients, one particularly important bottom-up mechanism is the decrease in the edibility of the phytoplankton assemblage associated with the increased dominance of cyanobacteria with increasing levels of eutrophication. Laboratory studies demonstrate that the lack of highly unsaturated fatty acids in cyanobacteria negatively affects the growth rates of a common zooplankton species (Daphnia) (Demott and Müller-Navarra 1997, Persson et al. 2007). Field observations (Müller-Navarra et al. 2000) and microcosm experiments (Park et al. 2003) have added further support for this finding. Many cyanobacteria also present physical challenges to grazers, collecting in colonies or filaments that are too large to be consumed (Bednarska and Dawidowicz 2007), or surrounding themselves with gelatinous sheaths (Vanni 1987). Altered

elemental stoichiometry and, hence, nutritional quality of phytoplankton under different levels of eutrophication may also influence zooplankton biomass (Hessen 2008).

The *rate* of change of zooplankton biomass with respect to increasing phytoplankton biomass (ΔZ/ΔP) provides an informative measure of the effects of eutrophication on food web function for the purposes of informing the derivation of numeric nutrient criteria (Yuan and Pollard 2018). This rate of change can be thought of as the slope of the relationship between Z and P. In most lake food webs, any increase in the basal resources (i.e., phytoplankton biomass) would be expected to be associated with a corresponding increase in the biomass of consumers of those resources (i.e., zooplankton biomass), and the slope between Z and P would be positive. In eutrophic lakes, however, increases in phytoplankton biomass often are not associated with an increase in zooplankton biomass, and the slope (ΔZ/ΔP) approaches zero (Leibold et al. 1997, Hessen et al. 2006, Heathcote et al. 2016). Based on this observation, EPA used the rate of change in zooplankton biomass relative to changes in phytoplankton biomass (ΔZ/ΔP) as a measure of the effect of excess nutrients on lake food webs.

#### *2.2.1.2 Dissolved oxygen*

Excess nutrients typically increase primary productivity, which increases the amount of organic matter in a lake. Then, DO is consumed as the organic matter is decomposed, leading to hypoxic and anoxic conditions (see Figure 1). Low concentrations of DO limit the extent to which habitat is available to fish and zooplankton (Colby et al. 1972, Tessier and Welser 1991, Vanderploeg et al. 2009), and oxygen availability is a key determinant of the quality and quantity of habitat available to aquatic biota in many lakes (Evans et al. 1996). Although hypoxia occurs naturally in a small number of systems (Diaz 2001), anthropogenic nutrient loads have greatly increased the occurrence of hypoxia worldwide (Jenny et al. 2016). Deoxygenation of lake water typically begins near the lake bottom and proceeds to shallower depths over the summer, especially in stratified, relatively deep lakes, where the replenishment of DO from surface mixing is restricted (Cornett 1989, Wetzel 2001). Therefore, an increasing proportion of the deeper waters of a lake can become uninhabitable for certain organisms over the course of the summer (Molot et al. 1992). Exclusion of deeper waters as viable habitat, in particular, can disproportionately affect particular species of adult and juvenile fish (Lienesch et al. 2005).

Another strong determinant of the available habitat for fish and zooplankton is water temperature. Summer brings a longer photoperiod and more intense solar insolation, which

increases water temperatures near the surface of many lakes to levels harmful to certain species (Ferguson 1958, Eaton and Scheller 1996). The viable habitat for cool- and cold-water species, in particular, can be restricted by surface warming (Jacobson et al. 2010, Arend et al. 2011). In contrast to deoxygenation, warming starts at the surface of the lake and proceeds to deeper depths over the course of the summer. Therefore, certain species of fish are "squeezed" between increasing temperatures at shallow depths and decreasing DO at deeper depths (Coutant 1985, Stefan et al. 1996, Lee and Bergersen 1996, Plumb and Blanchfield 2009), requiring them to choose between suboptimal temperatures or oxygen (Arend et al. 2011). Under those conditions, the metalimnion and the upper edge of the hypolimnion can provide an important refuge, and even a thin layer of cool water with sufficient DO can provide an important habitat for supporting fish health through the warmest summer days. Because they often can tolerate lower DO concentrations than fish, zooplankton can retreat to deeper depths of the hypolimnion to escape fish predation, but are also limited ultimately by low DO concentrations (Tessier and Welser 1991, Stemberger 1995).

Based on these considerations, the mean concentration of DO below the thermocline was identified by EPA as an appropriate metric for assessing risks to cool- and cold-water fish in seasonally stratified lakes. In those lakes during the summer, the availability of cool-water habitat is constrained by deep water DO concentrations, and so, this risk metric links increased nutrient concentrations to deleterious effects affecting the survival of fish and zooplankton in deep lakes.

### 2.2.2 Recreational Use

EPA selected the concentration of cyanotoxins as the risk metric linking increased nutrients to the suitability of lake water for primary and secondary contact recreation. Increased nutrient concentrations and an attendant increase in cyanobacterial abundance can increase concentrations of cyanotoxins (Figure 2), which cause adverse effects on the health of people exposed to the water (US EPA 2019). One of the most commonly occurring types of cyanotoxins in freshwaters is microcystins (based on available data). To protect recreational uses of lakes, EPA identified microcystin concentration as the best risk metric because of the availability of National Lakes Assessment (NLA) data (US EPA 2010b) and because microcystin thresholds for recreational exposures have recently been published (US EPA 2019).

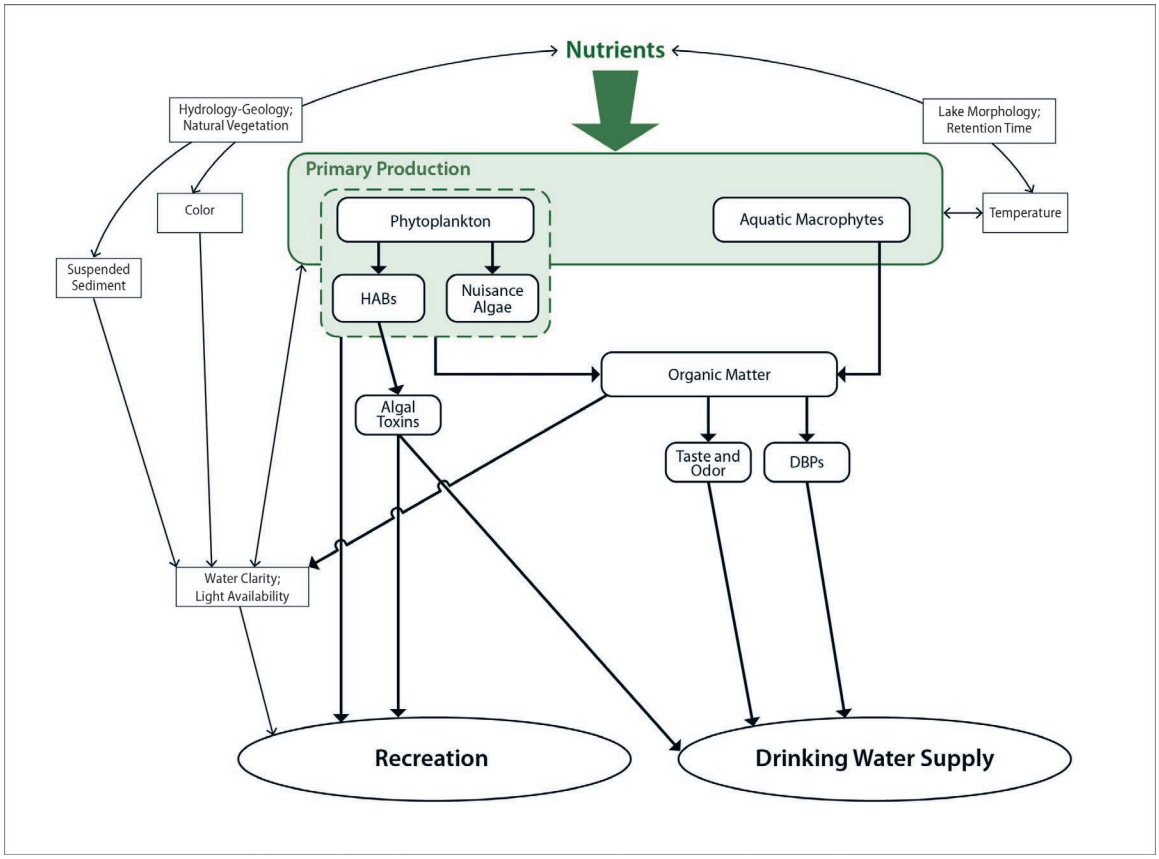


Figure 2. Conceptual model linking increased nutrient concentrations to public health endpoints.

### 2.2.3 Drinking Water Source

Increased nutrient concentrations and an attendant increase in cyanobacteria can increase concentrations of cyanotoxins, which are toxic when consumed at certain concentrations and quantities (see Figure 2) (Chorus 2001, Stewart et al. 2008, US EPA 2015b). As was done for recreational use, EPA selected microcystin in lake source water as the risk metric for the drinking water use.

## 2.3 Risk Hypotheses

EPA specified risk hypotheses for each of the selected assessment endpoints. Based on a survey of available literature, EPA concluded that increased concentrations of N and P increase the risk to both ecological and human health (Figure 3). For aquatic life, the risk hypotheses consist of the pathway in which increased nutrient concentrations increase phytoplankton biovolume (measured as chlorophyll *a* [Chl *a*]). Then, as phytoplankton biovolume and associated biomass increase, the relationship between zooplankton biomass and phytoplankton biomass ($\Delta Z/\Delta P$)

changes so that increases in phytoplankton biomass are no longer associated with increases in zooplankton biomass, and increases in primary production at the base of the lake food web are not transferred to higher trophic levels. For the case of deep water DO concentrations, increased phytoplankton biomass increases organic matter in the lake, which, when decomposed, consumes DO (Walker 1979). The decreased concentrations of DO then affect lake aquatic life. The risk hypotheses for both recreation and drinking water source designated uses state that increased nutrient concentrations increase the biovolume of cyanobacteria and concentrations of microcystin.

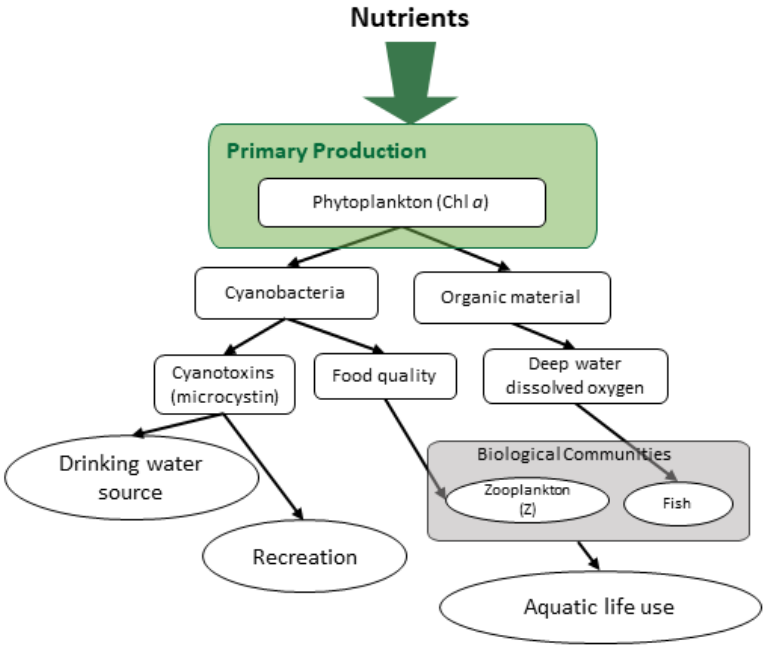


Figure 3. Simplified conceptual model showing pathways selected for analysis.

## 2.4 Analysis Plan

An analysis plan outlines the scope, approaches, and methods to be used in the risk assessment. Here, the analysis plan consists of acquiring appropriate data and estimating relationships between phytoplankton biovolume and each of the risk metrics as well as between N, P, and phytoplankton biovolume. The critical measurement in all these relationships is Chl *a*, which is closely associated with phytoplankton biovolume. Stressor-response analysis was applied to available data to estimate relationships between nutrient concentrations and different risk metrics. Because Chl *a* concentration is the critical parameter for all risk metrics, EPA developed

different stressor-response models associating Chl *a* concentration with each of the risk metrics (i.e., zooplankton biomass (ΔZ/ΔP), deep water DO concentration, and microcystin). The models then yielded candidate criteria for Chl *a* corresponding to each of the risk metrics (and their associated endpoints). A candidate Chl *a* criterion is defined as a protective concentration associated with a particular risk metric and/or with a particular designated use. The final criterion for a lake is then derived by comparing candidate criteria and selecting the criterion that protects the most sensitive use. N and P are estimated in field measurements as TN and TP, and so, EPA developed models relating TN and TP concentrations to Chl *a* concentrations that can translate each of the different Chl *a* criteria into recommended TN and TP criteria.

Because different risk metrics have been identified for each of the three designated uses, these risk metrics lead to the derivation of different recommended numeric nutrient criteria. In general, a state's water quality criteria for any single lake would need to protect the most sensitive use (i.e., the state should select the most stringent numeric nutrient criteria across all relevant uses and endpoints) (40 CFR 131.11(a)(1)).

# 3 Analysis

Because stressor-response analyses for each of the risk metrics differed substantially from one another, most of this section is organized by models for the different risk metrics—zooplankton biomass, deep water hypoxia, and microcystin—followed by models relating TN, TP, and Chl *a*. Because the same data were used to fit each of these models, all the data used in the analyses are discussed first.

## 3.1 Data

EPA analyzed data collected in the NLA in the summers (May–September) of 2007 and 2012 to support the derivation of recommended numeric nutrient criteria. The NLA data were collected from a random sample of lakes from the continental U.S. In 2007, lakes with surface areas larger than 4 hectares and, in 2012, lakes larger than 1 hectare were selected from the contiguous U.S. using a stratified random sampling design (US EPA 2012c). The final data set was supplemented by a small number of hand-picked lakes identified as being less disturbed by human activities (US EPA 2010b). The additional lakes were included to increase the number of least-disturbed lakes for which data were available, and by helping ensure the full range of conditions was

sampled, data from the additional lakes were expected to improve the accuracy of the estimated stressor-response relationships. The overall sampling design of the NLA was synoptic, but 10% of sampled lakes were randomly selected and resampled on a different day after the initial visit. The timing of the second visit varied among lakes, but on average, the second sample was collected approximately 46 days after the first. Approximately 20% of the lakes were sampled in both 2007 and 2012. The sampling day of the year was recorded for each visit and used in subsequent analyses to account for temporal changes in deep water DO concentration. Overall, data from approximately 1,800 different lakes are included in the data set, but the specific number of samples used to estimate each stressor-response relationship varies slightly based on data available at each lake. The specific number of samples is provided in the subsequent discussion of each model. Quality assurance project plans for the NLA are available in separate publications (US EPA 2009, 2012b).

During each visit to a selected lake, an extensive suite of abiotic and biological variables was measured. Only brief details on sampling protocols are provided here regarding the parameters used to derive these criterion models; more extensive descriptions of sampling methodologies are available in the NLA documentation (US EPA 2007, 2011). A sampling location was established in open water at the deepest point of each lake (up to a maximum depth of 50 meters [m]) or in the mid-point of reservoirs. In 2012, an additional sampling location for collection of microcystin, algae, and Chl *a* data was established in the littoral zone approximately 10 m away from a randomly selected point on the shoreline.

At the open water site, a vertical, depth-integrated methodology was used to collect a water sample from the photic zone of the lake (to a maximum depth of 2 m). Multiple sample draws were combined in a rinsed, 4-liter (L) cubitainer. When full, the cubitainer was gently inverted to mix the water, and an aliquot was taken as the water chemistry sample. That subsample was placed on ice and shipped overnight to the Willamette Research Station in Corvallis, Oregon. A second aliquot was taken to use in characterizing the phytoplankton community and was preserved with a small amount of Lugol's solution. A Secchi depth measurement was also collected at this site. Two zooplankton samples were collected with vertical tows for a cumulative tow length of 5 m using fine- (50-micrometer- [-µm-]) and coarse- (150-µm-) mesh Wisconsin nets. In lakes at least 7 m deep, one 5-m deep tow was collected with each mesh. In

shallower lakes, vertical tows over shorter depths were combined to reach the cumulative tow length of 5 m.

At the littoral zone site, two grab water samples were collected 0.3 m below the surface where the lake was at least 1 m deep using a 2-L brown bottle. The first sample was split into two subsamples: one subsample for quantifying algal toxin concentration and the second subsample preserved with a small amount of Lugol's solution and used to characterize the phytoplankton community. The second grab sample collected with the 2-L bottle was used to quantify Chl *a* concentration.

### 3.1.1 Biological Data

Phytoplankton biovolume from the field samples was measured in the laboratory. Samples collected from both open water and littoral zone locations were examined by taxonomists, who identified at least 400 natural algal units to species under 1,000× magnification. Observations were aggregated and abundance was calculated as cells per milliliter. In each sample, the dimensions of the taxa that accounted for the largest proportion of the observed assemblage were measured and used to estimate biovolume. Biovolumes of the most abundant taxa were based on the average of measurements from at least 10 individuals, while biovolumes of the less abundant taxa were based on somewhat fewer measurements. The biovolume was reported as cubic micrometers per milliliter ($\mu m^3/mL$) (US EPA 2012a), which was converted to cubic millimeters per liter ($mm^3/L$). Approximately 5% of the phytoplankton samples were randomly selected and reidentified and measured by a second taxonomy laboratory. These reidentified samples provided a basis for estimating laboratory measurement error. Biovolume measurements were converted to biomass using a density of 1 gram per milliliter (g/mL) (Holmes et al. 1969).

Zooplankton samples from the coarse- and fine-mesh net tows were processed separately. In each sample, zooplankton specimens were examined and counted under 100–1,000× magnification in discrete subsamples until at least 400 individuals were identified. In the coarse-mesh net samples, all taxa were identified and enumerated. In the fine-mesh net, only "small" taxa were identified and enumerated (Cladocera less than 0.2 millimeters [mm] long, copepods less than 0.6 mm long, rotifers, and nauplii). Zooplankton abundance was estimated based on the volume of sampled lake water used to identify the targeted count of 400 individuals. Measurements of at least 20 individuals were collected for dominant taxa (i.e., taxa

encountered at least 40 times in the subsample); at least 10 individuals were measured for taxa encountered from 20 to 40 times; and at least 5 individuals were measured for rare taxa (encountered less than 20 times in the subsample). Zooplankton biomass estimates were based on existing length and width relationships (Dumont et al. 1975, McCauley 1984, Lawrence et al. 1987). Estimates from the coarse- and fine-mesh samples were added to yield a single zooplankton sample per lake visit.

### 3.1.2 Chemical Data

For both 2007 and 2012 data, TN, nitrate-nitrite ($NO_x$), ammonia, and TP concentrations; true color, dissolved organic carbon (DOC) concentration, turbidity, and acid-neutralizing capacity (ANC) were measured in the laboratory from the open water sample at prespecified levels of precision and accuracy (US EPA 2012a). Typical laboratory methods included persulfate digestion with colorimetric analysis for TN and TP, nephelometry for turbidity, comparison to a calibrated color disk for true color, and automated acidimetric titration for ANC. To measure Chl *a* concentration, 250 mL of lake water was pumped through a glass fiber filter in the field and quantified in the laboratory to prespecified levels of precision and accuracy. Examples of lower reporting limits include 20 µg/L for TN, 4 µg/L for TP, and 0.5 µg/L for Chl *a*.

Microcystin sample processing began with three sequential freeze/ thaw cycles to lyse cyanobacteria (Loftin et al. 2008). Processed samples were filtered using 0.45 µm polyvinylidene difluoride membrane syringe filters and stored frozen until analysis. The concentration of microcystin in the filtered water sample was measured with an enzyme-linked immunosorbent assay (ELISA) using an Abraxis kit for Microcystin-ADDA, which employs polyclonal antibodies that are unique to microcystins and other similar compounds. The binding mechanism of the Microcystin-ADDA assay is specific to the microcystins, nodularins, and their congeners; therefore, results from that assay could include contributions from any compound within the ADDA functional group (Fischer et al. 2001). The minimum reporting level for the assay was 0.1 µg/L as microcystin-LR.

### 3.1.3 Dissolved Oxygen and Temperature Profiles

At the deepest point of each lake (or in the midpoint of reservoirs), a multiparameter water quality meter was used to measure profiles of DO concentrations, temperature, and pH at a minimum of 1-m depth intervals (see Section 3.2.2 for an examples of depth profiles). Profiles in

lakes less than 3 m deep were sampled at 0.5-m depth intervals. Water temperatures were converted to estimates of water density (Jones and Harris 1992), and density gradient was estimated between all available depths below 0.5 m as the difference in density between two successive measurements divided by the difference in the depths of the two measurements. Temperature gradients were computed with the same approach. Samples collected in the uppermost 0.5 m were excluded to limit the effects of surface warming on the gradient calculations.

#### 3.1.4 Mapped Data

Lake physical characteristics, including lake surface area, geographic location (latitude and longitude), elevation, lake catchment area, and lake perimeter, were estimated from mapped data. From these characteristics, the following composite variables were calculated: (1) the drainage ratio, which is defined as the ratio of catchment area to lake surface area and characterizes the degree to which the lake catchment influences the lake; (2) the shoreline development, which is defined as the ratio between the perimeter of the lake and the perimeter of a circle with the same area as the lake and characterizes the geometric complexity of the lake shore; and (3) the lake geometry ratio, which is defined as $area^{0.25}$/depth, or the ratio between fetch and lake maximum depth, and has been shown to differentiate lakes that stratify seasonally (low values of the geometry ratio) from lakes that are polymictic (Gorham and Boyce 1989, Stefan et al. 1996). Variables quantifying the mean annual precipitation and mean annual air temperature at the lake location were extracted from 30-year averaged climatic data (Daly et al. 2008).

### 3.2 Stressor-Response Models

Stressor-response models estimate relationships between environmental stressors (e.g., increased nutrient concentrations) and responses, or risk metrics. In this section, stressor-response models for zooplankton biomass, deep water hypoxia, and microcystin concentration are described. The stressor in these models is phytoplankton biovolume, quantified as Chl *a*. Models estimating relationships between TN, TP, and Chl *a* are also described.

### 3.2.1 Zooplankton Biomass

Zooplankton occupy a key link in the lake food web, and the relationship between phytoplankton biovolume and zooplankton biomass provides insights into changes in the food web as lake productivity increases.

#### *3.2.1.1 Statistical analysis*

EPA specified a Bayesian network model to estimate the relationship between summer mean phytoplankton and zooplankton biomasses (Figure 4). A Bayesian network provides a unified framework for modeling the cascading relationships between different measurements and propagates estimation errors and model uncertainty correctly throughout the model (Qian and Miltner 2015; Yuan and Pollard 2018).

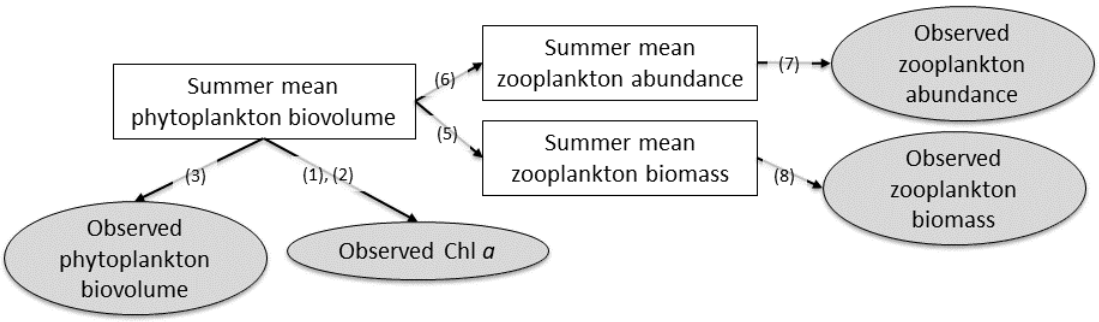


Figure 4. Schematic of network of relationships for modeling zooplankton biomass. *Gray-filled ovals*: available observations*; other nodes*: modeled parameters*; numbers in parentheses* refer to equation numbers in the text.

The first set of relationships in the network estimated seasonal mean phytoplankton biovolume based on measurements of Chl *a* concentration and of phytoplankton biovolume. The two measurements provided independent estimates of phytoplankton biovolume, each with different sources of error. Chl *a* is measured precisely from field samples, but the Chl *a* content of phytoplankton can vary depending on environmental conditions and species composition (Kasprzak et al. 2008), so that a measured Chl *a* concentration in one sample might indicate a slightly different phytoplankton biovolume than the same Chl *a* measured in another sample. Hence, Chl *a* concentration is modeled as being directly proportional to the seasonal mean phytoplankton biovolume in lake *j* ($P_j$), but the constant of proportionality, *b*, (i.e., the Chl *a*

content of phytoplankton in a sample) is allowed to vary among samples. The log-transformed version of this model equation is as follows:

$$\log(Chl_i) = \log\left(P_{j[i]}\right) + \log(b_i) \quad (1)$$

$$\log(b_i) \sim Normal(\mu_b, \sigma_b) \quad (2)$$

where the value of $b_i$ for each sample, *i*, is drawn from a single log-normal distribution characterized by a mean, $\mu_b$, and a standard deviation, $\sigma_b$. The mean value, $\mu_b$, accounts for differences in the measurement units of Chl *a* and phytoplankton biovolume; however, in the model, both Chl *a* and phytoplankton biovolume measurements are standardized by subtracting their observed mean values and dividing by their standard deviation, so $\mu_b$ is zero. The multilevel expression of the model equation allows the mean Chl *a* content of phytoplankton cells estimated for each sample to vary, but imposes the constraint that estimates of phytoplankton Chl *a* content for each sample must be drawn from a common log-normal distribution (Gelman and Hill 2007). Measurements collected at the same lake on different days and collected on the same day in the littoral zone and in the middle of the lake were used to estimate seasonal mean phytoplankton biovolume.

Direct measurements of phytoplankton biovolume provide an unbiased estimate of true phytoplankton biovolume. These direct measurements, however, are obtained by summing contributions from measurements taken from many different individual phytoplankton, each of which includes measurement error. Hence, the summed estimate of total biovolume includes a substantial amount of measurement error. That measurement error was explicitly modeled, and a second estimate of the seasonal mean phytoplankton biovolume was expressed as follows:

$$\log\left(P_{obs,i}\right) \sim Normal\left(\log\left(P_{j[i]}\right), s_1\right) \quad (3)$$

where $P_{obs,i}$ is the observed phytoplankton biovolume in sample *i*. Final model estimates of $P_{j[i]}$ were consistent with both Chl *a* and observed phytoplankton biovolume, and by combining the two measurements, the accuracy of the final estimate was maximized.

Zooplankton abundance (A) and biomass (Z) were modeled as increasing functions of seasonal mean phytoplankton biovolume (or biomass, using the conversion factor of 1 g/mL). Previous studies in oligotrophic lakes found that zooplankton biomass increased as a constant proportion of phytoplankton biomass (Rognerud and Kjellberg 1984, del Giorgio and Gasol 1995). That is,

after log-transforming, the relationship between P and Z should approach the following at low concentrations of P:

$$\log(Z) = \log(k) + \log(P) \quad (4)$$

where the slope of the relationship between log(Z) and log(P) approaches 1. In contrast, in eutrophic lakes, minimal changes in Z were observed with changes in P, and the slope between log(Z) and log(P) approached zero (Yuan and Pollard 2018). Those patterns guided the selection of the following functional form for modeling the relationship between log(Z) and log(P):

$$E[\log(Z_j)] = f_1 + f_2 \log(P_j) - f_3 q \log\left[1 + \exp\left(-\frac{\log(P_j) - c_p}{q}\right)\right] \quad (5)$$

where, in general, *E[.]* indicates the expected value of the variable enclosed in the square brackets. The coefficients $f_1$, $f_2$, $f_3$, $c_p$, and $q$ were estimated from observations of $Z_j$ and the estimated seasonal mean phytoplankton concentration, $P_j$, estimated from measurements of phytoplankton biovolume and Chl *a*. The slope of this function approaches $f_2$ at large values of P and approaches a slope of $f_2 + f_3$ at low values of P. A prior distribution for $f_2 + f_3$ was specified as a normal distribution centered at 1 with a standard deviation of 0.2, expressing the prediction (stated above) that, at low levels of phytoplankton (oligotrophic lakes), zooplankton biomass should increase as a constant proportion of phytoplankton biomass.

A similar model was specified for zooplankton abundance (A) as follows:

$$E[\log(A_j)] = a_1 + a_2 \log(P_j) + a_3 \log\left[1 + \exp\left(-\frac{\log(P_j) - c_p}{r}\right)\right] \quad (6)$$

where the parameters, $a_1$, $a_2$, $a_3$, and $r$ were estimated from the data, and the third term on the right side of the equation again introduces curvature in the fitted relationship. The change point for zooplankton abundance, $c_p$, was estimated as being the same as for zooplankton biomass because of the strong influence of abundance on total biomass. In the case of zooplankton abundance, no *a priori* assumptions about the slope of the relationship at high or low levels of phytoplankton guided the choice of parameter values.

Observed values of zooplankton abundance and biomass were then related to the estimated expected values as follows:

$$\log(A_{obs,i}) \sim Normal(\mathrm{E}[\log(A_{j[i]})], s_2) \quad (7)$$

$$\log(Z_{obs,i}) \sim Normal(\mathrm{E}[\log(Z_{j[i]})], s_3) \quad (8)$$

Similar to the model equations for phytoplankton, variability in the observations of zooplankton abundance and biomass relative to estimated mean values were modeled as log-normal distributions with standard deviations of $s_2$ and $s_3$. These error terms included contributions from temporal, sampling, and measurement error.

Because the strength of the interaction of the zooplankton assemblage with benthic resources was expected to differ between shallow and deep lakes (Benndorf et al. 2002, Scheffer and van Nes 2007) and exploratory data analysis indicated that lake depth was an important covariate (Yuan and Pollard 2018), different parameter values for $a_1$, $a_2$, $a_3$, $f_1$, $f_2$, $f_3$, and $c_p$ were estimated for each of three classes of lakes defined by depth. The curvature parameters $q$ and $r$ were fixed at 1. The number of lake classes was specified to balance between accounting for differences in lake depth and maintaining enough samples within each class to estimate relationships. Depth thresholds defining each class were selected to ensure that a similar number of samples was assigned to each class, yielding the following thresholds: less than 3.8 m, 3.8–8.0 m, and more than 8.0 m.

All model equations were fit simultaneously to data collected at each lake, including revisits on different days, and littoral and mid-lake samples for phytoplankton. Weakly informative priors were specified for all model parameters except for $f_2 + f_3$ (Gelman 2006). Weakly informative prior distributions constrain parameter estimates away from extreme values, while allowing the data to determine the estimate for each parameter. All other statistical calculations were performed with R, an open-source statistical modeling software (R Core Team 2017). Hierarchical Bayesian models were fit using the rstan library, which implements the No-U-Turn sampler, a variant of a Hamiltonian Monte Carlo sampling approach (Duane et al. 1987, Stan Development Team 2016).

#### *3.2.1.2 Results*

Data collected at a total of 1,096 lakes were available for analysis, with approximately 330 lakes assigned to each depth class. Estimated mean phytoplankton biovolume within each sample was much more strongly associated with Chl *a* concentration than with measured phytoplankton biovolume, because of the high variability associated with measured phytoplankton biovolume (Figure 5).

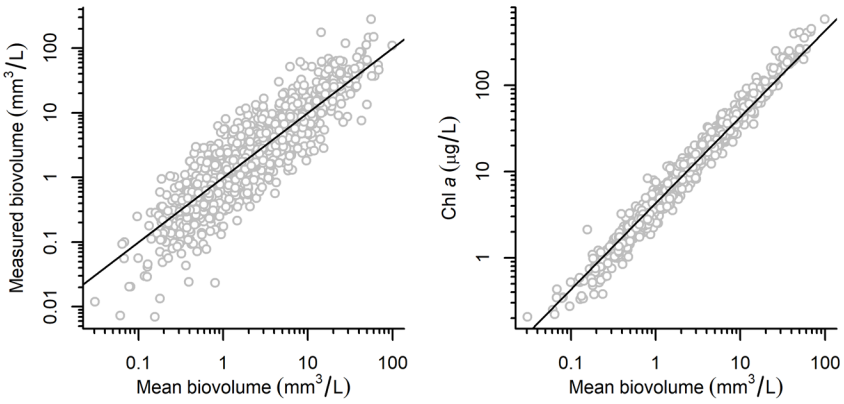


Figure 5. Relationships between measured biovolume, Chl *a*, and estimated seasonal mean phytoplankton biovolume. Solid lines: 1:1 relationship (left panel), proportional relationship (right panel).

Estimated relationships between phytoplankton biomass (as quantified by Chl *a*) and zooplankton abundance and biomass matched trends observed in the data (see Figure 6 for an example for lakes deeper than 8 m in left and middle panels, respectively). The relationship between zooplankton biomass and phytoplankton biomass also was consistent with the initial assumption that, in oligotrophic lakes with low levels of phytoplankton biomass, the slope approached 1, and in eutrophic lakes with high levels of phytoplankton biomass, the slope approached zero (right panel, Figure 6).

The models show the gradual change in the shape of the biomass pyramid along the eutrophication gradient. In oligotrophic lakes, the slope of the relationship between zooplankton and phytoplankton biomass is near 1, indicating that small increases in phytoplankton biomass are reflected in a proportional increase in zooplankton biomass. As Chl *a* increases, however, the slope decreases, and the increase in zooplankton biomass per unit of increase in phytoplankton biomass [$\log(\Delta Z)/\log(\Delta P)$] approaches zero. In eutrophic lakes, increases in phytoplankton biomass do not result in comparable changes in zooplankton biomass. These changes along the eutrophication gradient are consistent with other similar studies, as reviewed in Yuan and Pollard (2018).

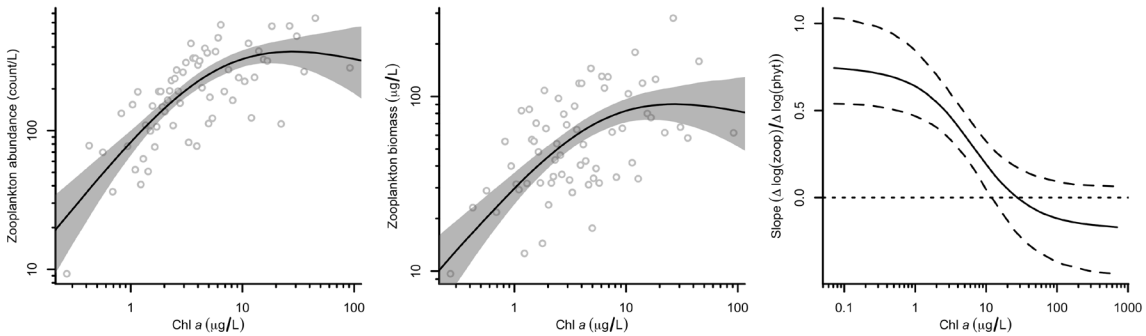


Figure 6. Estimated relationships between zooplankton and Chl *a* for lakes > 8 m deep. *Left panel*: Chl *a* vs. zooplankton abundance; *middle panel*: Chl *a* vs. zooplankton biomass; *right panel*: Chl *a* vs. slope of the relationship between zooplankton biomass and Chl *a. Solid lines*: mean relationships; *shaded areas* (*left and middle panels*): 90% credible intervals (bounded by the 5th and 95th percentiles of the distribution of possible mean relationships; *dashed lines* (*right panel)*: 90% credible interval; *open circles (left and middle panels)*: average of five samples nearest the indicated Chl *a* concentration; *dotted horizontal line* (*right panel*): one example value of threshold for deriving a Chl *a* criterion.

#### *3.2.1.3 Chl a criterion derivation*

Calculating candidate criteria for Chl *a* based on this response requires the specification of two parameters—the value of the slope between log(Z) and log(P) and the certainty level, as quantified by the credible interval (i.e., the Bayesian analog to a confidence interval).

**Considerations: Slope between log(Z) and log(P).** The threshold slope of zero between log(Z) and log(P) is the limit beyond which changes in zooplankton biomass are not associated with changes in phytoplankton biomass. This threshold identifies the point at which a disconnect between phytoplankton and zooplankton production begins and increasing primary productivity in the lake escapes zooplankton grazer control, a condition which can exacerbate blooms. EPA therefore recommends a threshold of zero as the minimum value for this parameter. Higher threshold slopes may be selected for certain types of lakes in which a higher proportion of phytoplankton is expected to be consumed by zooplankton (e.g., oligotrophic), but quantifying the appropriate value for this slope requires the collection of additional data from these lakes. Graphically, this threshold defines the horizontal line on which the Chl *a* criterion will be based (see Figure 6).

**Considerations: Certainty level.** The certainty level, as quantified by the credible interval, expresses the statistical uncertainty about the position of the mean relationship and is comparable to a confidence interval used in frequentist statistics. The percentile value selected as the certainty level (e.g., the 90% certainty level) specifies the probability that the mean

relationship is located within the estimated bounds of the corresponding credible interval (e.g., there is a 90% probability that the actual mean relationship is located within the bounds associated with the 90% credible interval). In practice, the bounds of the credible interval are calculated from percentiles of the posterior distribution of the mean relationship. For example, at any Chl *a* concentration, the criterion model provides a distribution of possible values of the slope between zooplankton and phytoplankton biomasses. The credible interval is centered at the median value of this distribution of slopes, and the bounds of the 90% credible interval are calculated as the $5^{th}$ and $95^{th}$ percentiles of this distribution.

Basing the criterion value on the lower bound of the selected credible interval provides additional assurance that the calculated criterion is protective, given the data and model uncertainty. That is, more protective criteria are based on lower percentiles of the distribution of possible mean relationships (i.e., lower bounds of the credible interval). For example, if one selects a certainty level of 90%, the corresponding bounds for the 90% credible interval are computed as the $5^{th}$ and $95^{th}$ percentiles of the distribution of estimated mean relationships. Therefore, basing a criterion on the lower bound of the 90% credible interval implies there is a 5% chance that the actual slope between log(Z) and log(P) is less than the selected threshold. That is, there is a 5% chance that the derived criterion value is greater than the concentration needed to achieve the desired condition, and therefore, a 5% chance that the criterion is under-protective. More certainty on the criterion value is achieved by selecting greater certainty levels. For example, there is a 1% chance that a criterion based on the 98% certainty level would be under-protective. In statistical hypothesis testing, convention suggests that *p*-values of 1% or 5% are statistically significant results, which can also inform the selection of the certainty level.

Selection of the certainty level as the basis for the criteria is a management decision, and a range of certainty levels (and associated credible intervals) from 50% to 99% is provided in the associated interactive tool (see below). The ecological effects of reduced grazer control of phytoplankton biomass associated with a slope threshold value of zero can be substantial and difficult to reverse; and therefore, conservative certainty levels (i.e., 90% - 99%) are recommended for this response.

**Chl *a* criterion derivation:** Illustrative criteria for Chl *a* for different management decisions are shown in Table 2. The interactive tool, which uses posterior simulation with the estimated parameter distributions, computes candidate criteria for different combinations of the slope

threshold and the certainty level (https://nsteps.epa.gov/apps/chl-zooplankton). With this tool, a user can specify the value of the slope between log(Z) and log(P), lake depth, and the certainty level, and the associated criteria and stressor-response relationship are updated to reflect those selections. The tool provides transparent information regarding the effects of different levels of certainty and can provide a useful means of engaging with stakeholders regarding the risk management decisions underlying criterion development.

Table 2. Illustrative Chl *a* criteria (µg/L) for different certainty levels and a slope threshold value of 0 for Δ(log Z)/Δ(log P). Values shown for each lake depth class.

| | Depth class | | |
|---|---|---|---|
| **Certainty level** | **< 3.8 m** | **3.2 – 8 m** | **> 8 m** |
| 90% | 51 | 21 | 13 |
| 99% | 34 | 15 | 9 |

### 3.2.2 Deep Water Hypoxia

EPA specified a model for deep water DO that represents the temporal decrease in DO during summer stratification, while accounting for differences among lakes in eutrophication status, depth, and DOC concentrations (Yuan and Jones 2020a).

#### *3.2.2.1 Data*

EPA first restricted analysis to data collected from seasonally stratified lakes because hypoxic and anoxic conditions occur more consistently during stratified conditions. Lakes were identified that were likely to be seasonally stratified by computing the lake geometry ratio. This metric approximates the relative effects of lake fetch and depth on stability of stratification, and lakes with a geometry ratio less than 3 $m^{-0.5}$ exhibit seasonal stratification (Gorham and Boyce 1989). Therefore, EPA restricted NLA data to lakes with geometry ratios less than that threshold. Lakes likely to be dimictic (i.e., mixing fully in the spring and in the fall and ice-covered in the winter) were also identified based on latitude and elevation. This classification approach adjusts the lake latitude by elevation, and then identifies lakes with *adjusted* latitudes greater than 40˚ N as dimictic (Figure 7) (Lewis 1983). Restriction to dimictic lakes allowed EPA to use a simple relationship based on annual mean temperature in the model to more accurately predict the first day of stratification, and thus, more accurately estimate model parameters (see below). The location of dimictic lakes also roughly corresponded with lakes that were likely to harbor

cool- and cold-water fish. Finally, data were restricted to samples in which temperature profiles exhibited evidence of stratification (defined as a temperature gradient of at least 1 degree Celsius per meter [°C/m]).



Figure 7. Locations of NLA lakes used to fit the $DO_m$ model. These lakes were designated as dimictic and were stratified at the time of sampling.

Mean deep water DO concentrations ($DO_m$) in the selected NLA lakes were computed from temperature and DO profiles. First, measurements collected at depths less than or equal to 0.5 m were excluded to minimize the effects of surface warming. In some profiles, duplicate measurements of DO and/or temperature were collected at each depth, and in these cases, the average was used in computations. EPA used only profiles with measurements collected from at least half of the possible 1-m increments in the final analysis.

The upper boundary of the metalimnion was identified as the shallowest depth at which the temperature gradient exceeded 1 °C/m (excluding the surface layer) (Figure 8) (Wetzel 2001). $DO_m$ for each lake profile was computed as the mean of DO measurements estimated at all 1-m increments deeper than the upper boundary of the metalimnion. That estimate of $DO_m$ necessarily includes some measurements in the metalimnion, which might increase the estimates of $DO_m$ relative to studies that can focus only on the hypolimnion. In the NLA data set, the upper boundary of the metalimnion could be determined for most profiles. In contrast, many lakes in the NLA data set were too shallow to maintain a hypolimnion with small vertical temperature gradients (Jones et al. 2011), and therefore, no approach for consistently defining the hypolimnion for all lakes was available (Quinlan et al. 2005). Furthermore, inclusion of the

metalimnion was consistent with the assumption that taxa can use this transitional region as a refuge from warmer temperatures in the mixed layer (Klumb et al. 2004). The depth of water below the thermocline was computed as the difference between the maximum depth recorded for each lake and the mean depth of the upper boundary of the metalimnion. Chl *a* and DOC measurements from each lake were also used in the analysis. Prior to statistical analysis, all measurements were standardized by subtracting their overall mean values and dividing by the standard deviation. This standardization had no effect on the final model results, but helped the Bayesian models converge more efficiently (Gelman and Hill 2007).

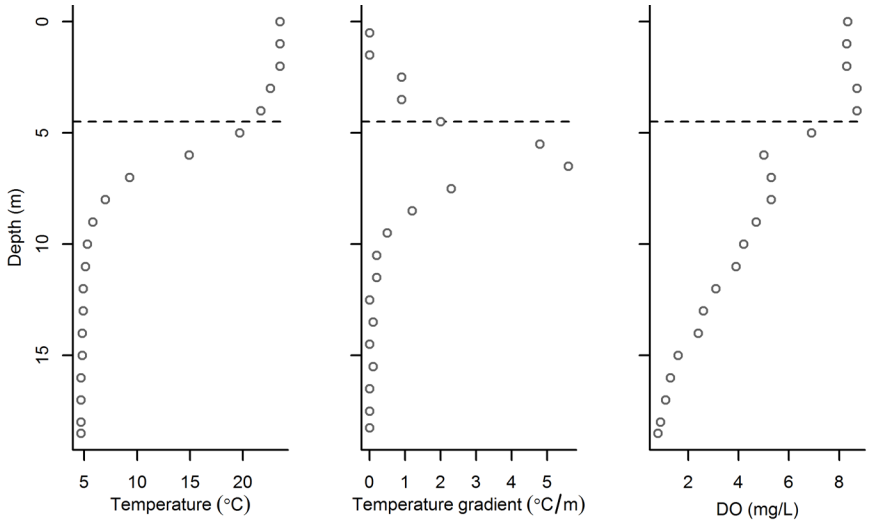


Figure 8. Illustrative examples of depth profiles of temperature, temperature gradient, and DO. *Dashed horizontal line*: estimated depth of the bottom of the epilimnion.

#### *3.2.2.2 Statistical analysis*

EPA modeled the decrease in $DO_m$ as a linear function, an approximation that is appropriate for $DO_m$ concentrations higher than approximately 2 milligrams per liter (mg/L) (Burns 1995). This threshold reflects experimental evidence indicating that the rate of decrease of hypolimnetic DO is constant at relatively high ambient concentrations of DO, but can be affected by DO concentrations near zero (Cornett and Rigler 1984). The linearly decreasing function also precludes the possibility of episodic mixing events that transport DO from shallow waters to deeper depths of the lake. In some lakes, those mixing events are rare, but in other lakes (e.g., cold polymictic lakes), they might occur frequently. In the latter group of lakes, the model predicts $DO_m$ during extended periods of still weather, and the associated criteria would protect

aquatic life in those scenarios. Below, the statistical model is first described followed by a description of the approach for addressing $DO_m$ measurements less than 2 mg/L.

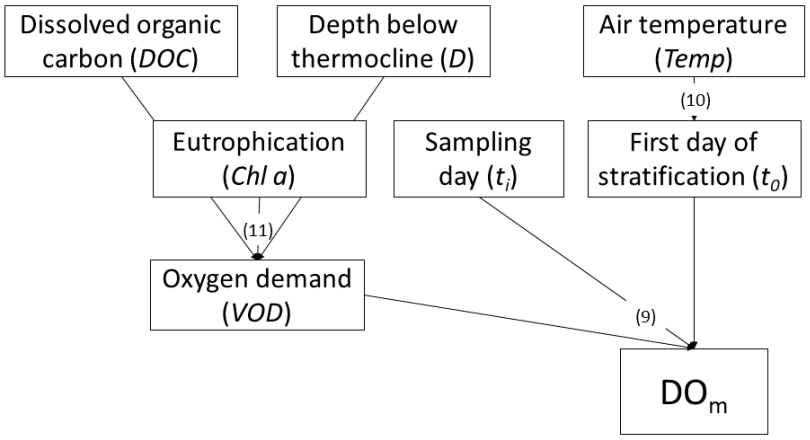


Figure 9. Schematic of hypoxia model. *Numbers in parentheses* refer to equation numbers in the text.

NLA data were fit to the following model equation:

$$E[DO_{m,i}] = DO_{0,j[i]} + VOD_{j[i]}\left(t_i - t_{0,k[i]}\right) \quad (9)$$

where $DO_{0,j[i]}$ is the value of $DO_m$ at the start of spring stratification in lake *j* corresponding to sample *i*, and volumetric oxygen demand ($VOD_j$) is the net imbalance in the volumetric oxygen budget for lake *j*, expressed as mg/L/day of DO (Burns 1995). That is, VOD estimates the rate of decrease in $DO_m$ per day. $t_i$ is the date that sample *i* is collected, and $t_{0,k[i]}$ is the date of the beginning of stratification for the lake-year *k* corresponding to sample *i*. Observed values of $DO_{m,i}$ were modeled as being normally distributed about the expected value, with a standard deviation of $\sigma_1$.

The first day of stratification ($t_0$) was not measured for any of the lakes, and the precise day on which stratification occurs for a given lake and year depends on local wind speeds, temperatures, and lake morphology (Cahill et al. 2005). Previous work in northern temperate dimictic lakes found that the first day of stratification could be modeled as a function of mean annual temperature (Demers and Kalff 1993), so EPA specified the following relationship for $t_0$:

$$t_{0,k} = b_1 + b_2 Temp_{j[k]} + e_k \quad (10)$$

where $Temp_{j[k]}$ is the mean annual air temperature at the location for lake *j* corresponding to lake-year *k*, and $b_1$ and $b_2$ are coefficients that are fit to the data. The published relationship in

Demers and Kalff (1993) provided initial estimates for $b_1$ and $b_2$, which were used to specify prior distributions for the two parameters. The error term $e_k$ is included in the model because the first day of stratification varies substantially in different years for a given lake because of differences in weather. Data published by Demers and Kalff (1993) indicated that the standard deviation of residual error for this relationship was approximately 12 days, so this value was used to specify the prior distribution for the standard deviation of $e_k$.

The initial concentration of DO at the time of stratification, $DO_0$, was also not measured for any of the lakes. Deep water temperatures in many seasonally stratified lakes are determined by temperatures prior to initiation of stratification (Hondzo and Stefan 1993), and so, deep water lake temperatures at the time of stratification were approximated as the minimum annual air temperature at the lake location. Then, the saturated DO concentration at the minimum annual air temperature provided an estimate for $DO_0$. Minimum air temperatures less than 4 degrees Celsius (˚C) were set to 4 ˚C, corresponding to water temperatures when the lake surface begins to freeze (Demers and Kalff 1993).

Lake trophic status affects VOD because increased phytoplankton production in the epilimnion increases the quantity of organic material available for decomposition in the hypolimnion and in lake sediments (Hutchinson 1938). In many lakes, allochthonous sources also provide organic matter that fuels bacterial respiration and depletes oxygen in deep lake waters (Pace et al. 2004, Kritzberg et al. 2004). VOD has also been observed to decrease with increasing hypolimnion depth, a phenomenon attributed to a weaker overall influence of sediment oxygen demand as the volume of the hypolimnion increases (Cornett and Rigler 1980, Müller et al. 2012). Based on these mechanisms, EPA modeled VOD as a linear function of the long-term mean Chl *a* concentration and depth below the thermocline in the lake. To account for the effect of allochthonous organic matter, DOC was also included as a third predictor variable for VOD (Hanson et al. 2003, Cole et al. 2011). The model equation for VOD can then be written as follows:

$$E[VOD_j] = d_1 + d_2 \log\left(Chl_{mn,j}\right) + d_3 D_{mn,j} + d_4 \log\left(DOC_{mn,j}\right) \quad (11)$$

where $d_1$, $d_2$, $d_3$, and $d_4$ are model coefficients estimated from the data; $\log(Chl_{mn,j})$ is the long-term mean of the log-transformed Chl *a* concentration lake, *j*; $D_{mn,j}$ is the mean depth of the lake below the thermocline; and $\log(DOC_{mn,j})$ is the seasonal mean of log-transformed DOC concentration in the lake. Variability in VOD across individual lakes about the mean value

estimated from the predictor variables was modeled as a normal distribution. Because Chl *a* concentrations can vary substantially over the summer in a lake, the modeling approach used with the zooplankton model provided a distribution of possible long-term mean log($Chl_{mn}$) values for each lake, given one or more instantaneous measurements of Chl *a* concentration. More specifically, seasonal mean log($Chl_{mn}$) values for different lakes were modeled as a normal distribution as follows:

$$\log(Chl_{mn,j}) \sim Normal(\mu_{Chl}, s_{Chl,1}) \quad (12)$$

Then, individual log-transformed measurements from each lake were assumed to be drawn from a normal distribution with a mean value equal to the long-term mean as follows:

$$\log(Chl_i) \sim Normal(\log(Chl_{mn,j[i]}), s_{Chl,2}) \quad (13)$$

where $Chl_i$ is the Chl *a* concentration measured in sample, *i*, associated with the mean $log(Chl_{mn})$ concentration in lake *j[i]*. (Note that Equations (12) and (13) are not shown in Figure 9.) Within-year variability of DOC and depth below the thermocline were substantially less than observed for Chl *a*, so long-term means for each of those parameters were estimated as the mean value of all available data for each lake. Weakly informative prior distributions were assigned to all parameters except for those that are noted above. Weakly informative prior distributions constrain parameter estimates away from extreme values, while allowing the data to determine the estimate for each parameter.

As noted earlier, $DO_m$ approaches zero asymptotically over time and modeling that relationship with the linear model described above would introduce biases to the model. To account for the asymptotic relationship, EPA modeled samples with $DO_m$ less than 2 mg/L with methods used for measurements that are below a known detection limit. That is, the samples were modeled as if their "true" $DO_m$ values were unknown but their maximum values were 2 mg/L (Gelman and Hill 2007). This approach retained some information inherent in a sample with $DO_m$ less than 2 mg/L (i.e., Chl *a*, lake depth, DOC, and sampling day are consistent with low $DO_m$), but allowed the use of linear relationships in the model to estimate the rate of DO depletion. More specifically, the model fits a linear trend in time to $DO_m$ observed from lakes with similar Chl *a*, DOC, and depth. By assuming that measurements of $DO_m$ less than 2 mg/L are unknown, the estimates of the linear relationship are more strongly determined by the higher $DO_m$ concentrations, and samples with $DO_m$ less than 2 mg/L exert a weak influence that is still consistent with the overall relationship. Retaining samples with $DO_m$ less than 2 mg/L in the

model prevents biases that would be introduced by considering only lakes with relatively high $DO_m$.

*3.2.2.3 Results*

A total of 477 samples collected at 381 lakes were available for analysis. $DO_m$ concentrations in 165 samples were less than 2 mg/L and were modeled as unknown values that were less than 2 mg/L. The asymptotic relationship can be seen in the plot of Chl *a* versus $DO_m$ (Figure 10), in which $DO_m$ decreases steadily up to a Chl *a* concentration of about 4 µg/L. At higher Chl *a* concentrations, the magnitude of the slope of the relationship between $DO_m$ and Chl *a* decreases and approaches zero.

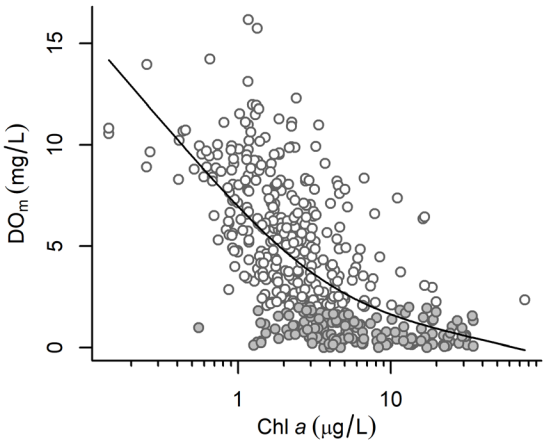


Figure 10. Chl *a* vs. $DO_m$. $DO_m$ values. *Gray-filled circles:* values < 2 mg/L; *solid line*: nonparametric fit to the data shown only to highlight the asymptotic relationship.

The majority of the estimates for the first day of stratification ranged from day 30 to day 120 (Figure 11). In most lakes, the Demers and Kalff (1993) estimate for the first day of stratification was later than the value of $t_0$ estimated by the model. This systematic difference is consistent with the fact that most of the lakes considered in Demers and Kalff (1993) were located north of the mean latitudinal location of the NLA lakes. The strong association between the Demers and Kalff (1993) estimates and the current estimates indicates that the overall formulation of the model, in which stratification day is a function of mean annual temperature, is valid.

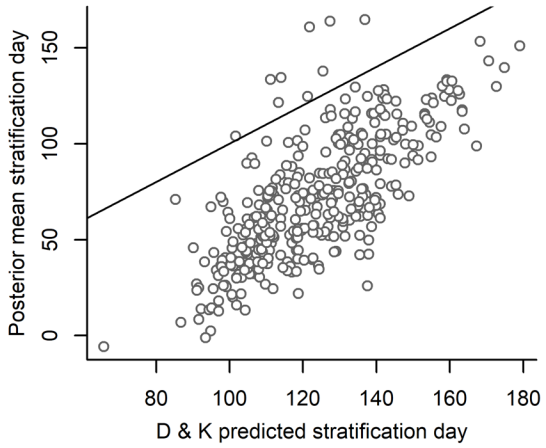


Figure 11. Demers and Kalff (1993) predicted stratification day vs. model mean estimate. *Solid line*: the 1:1 relationship.

Relationships estimated between $DO_m$ and different predictors were consistent with the hypothesized effects of each of the predictors (Figure 12). $DO_m$ decreased strongly with increases in DOC and Chl *a*, reflecting the increased organic material available in lakes with high concentration of the two parameters. Conversely, $DO_m$ increased with increasing depth below the thermocline, consistent with observations in other studies. Substantial uncertainty is associated with the relationship between $DO_m$ and day of the year, reflecting the inherent uncertainty in estimating the first day of stratification for different lakes.

The root mean square (RMS) error on model predictions for samples with $DO_m$ higher than 2 mg/L was 1.5 mg/L. RMS error is defined as the square root of the average squared difference between predicted and observed values. Slightly greater residual variability in the observations about the mean predictions were observed at high values of $DO_m$ (Figure 13).

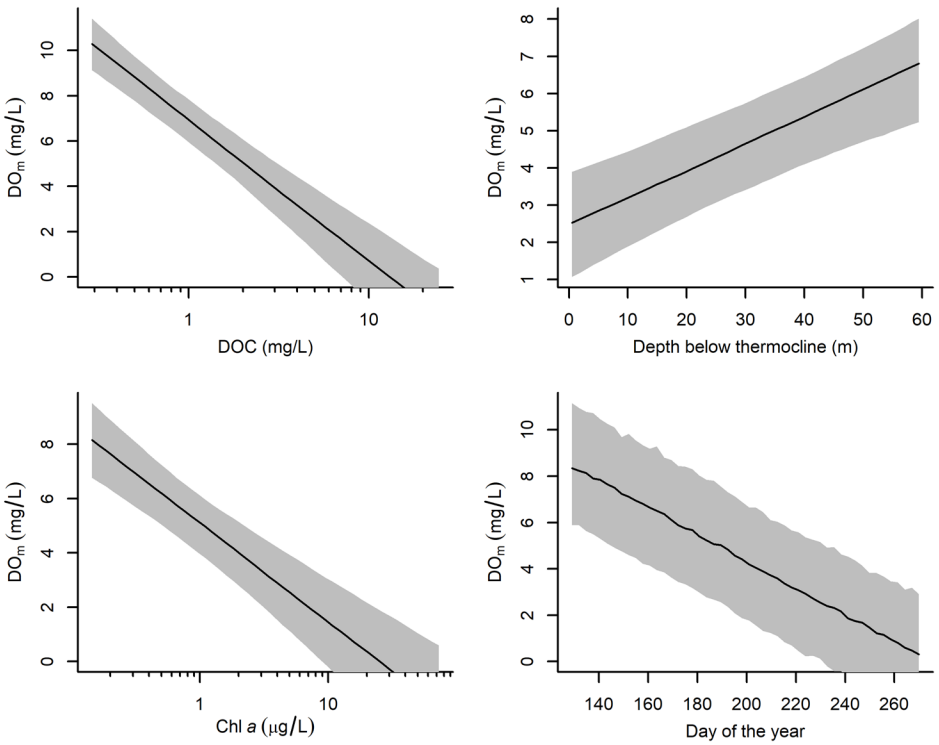


Figure 12. Relationships between individual predictors and $DO_m$, holding other variables fixed at their mean values. *Solid line*: mean relationship; *gray shading*: 90% credible intervals.

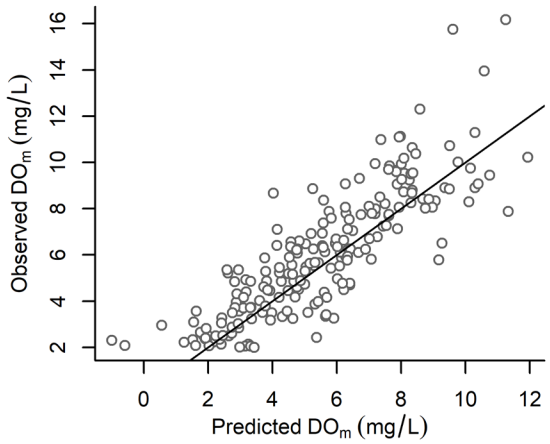


Figure 13. Model predicted $DO_m$ vs. observed $DO_m$. *Open circles*: individual samples; *solid line*: 1:1 relationship.

The statistical model described for $DO_m$ is consistent with the mechanisms of DO depletion in the deep waters of a lake, in which available DO below the thermocline is progressively depleted after the initiation of spring stratification. The estimated effects of eutrophication, DOC, and

lake depth on the rate of oxygen depletion were consistent with trends observed in other studies.

#### *3.2.2.4 Chl a criteria derivation*

As described earlier, warm temperatures in the shallow mixed layer of a lake act together with deep water hypoxia to constrain the available habitat for cool- and cold-water taxa. Therefore, to derive criteria based on deep water hypoxia, estimates of changes in water temperature over the course of the summer are required to identify periods of time during which mixed layer temperatures are too high for different taxa. Those periods of time then determine when deep water DO concentrations need to be sufficiently high to support different organisms.

Water temperature in the lake mixed layer depends on a variety of factors, including the local climate, solar insolation, lake morphology, and the day of the year (increasing in the spring and summer and decreasing in the fall). To identify temperatures in different lakes that were likely to limit available habitat for different fish, EPA first developed models to predict temperature in the shallow, mixed layer of different lakes. NLA data collected at all lakes in the conterminous U.S. were used to fit the model. At each lake, maximum temperature (excluding the top 0.5 m of the surface layer) observed in vertical profiles collected in each lake were modeled as a function of lake geographic location, elevation, and sampling day of the year with a generalized additive model (Wood 2006) of the following form:

$$E[T_i] = f_1 + f_2 Elev_{j[i]} + s(yday_i, df = 7) + s\left(Lat_{j[i]},\ Lon_{j[i]}, df = 30\right) \quad (14)$$

where E[$T_i$] is the expected value of the maximum temperature in the lake observed in sample *i*. $Elev_{j[i]}$ is the elevation of the lake, *j*, corresponding to sample *i*. The variable $yday_i$ is the day of the year that the sample was collected, and $Lat_{j[i]}$ and $Lon_{j[i]}$ are the latitude and longitude of the lake. The relationship between temperature and elevation was modeled as a simple linear relationship, characterized by two regression coefficients, $f_1$, and $f_2$. Relationships between lake temperature and sampling day and between lake temperature and location were modeled as nonparametric splines, represented in Equation 14 as *s(.)*, with the maximum degrees of freedom, *df*, as indicated. Observed values of $T_i$ were assumed to be normally distributed about the modeled expected value.

Lake temperature generally decreased with increased latitude, as would be expected (Figure 14), but deviations from that latitudinal pattern were observed on the west coast of the U.S.,

where lake temperatures were substantially lower than lakes at a similar latitude in the eastern U.S. This trend likely arises from the moderating influence of the coastal waters on air temperatures. Lake temperatures in eastern Texas and Louisiana were warmer than lake temperatures at the same latitudes elsewhere. Lake temperatures decreased with elevation, as expected, and exhibited a unimodal pattern with sampling day, with maximum temperatures occurring on average on Day 204, or July 22 (Figure 15). Overall, the model predicted lake temperature with an RMS error of 1.9 ˚C.

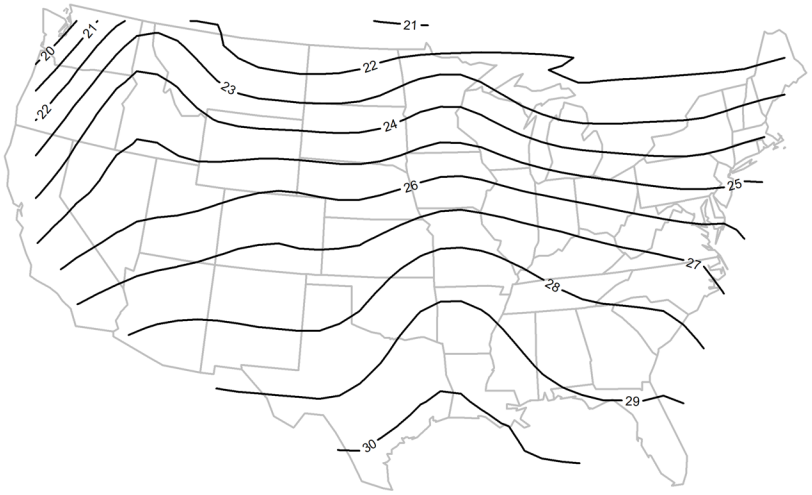


Figure 14. Contours of modeled mean lake temperature computed at the overall mean elevation and mean sampling day.

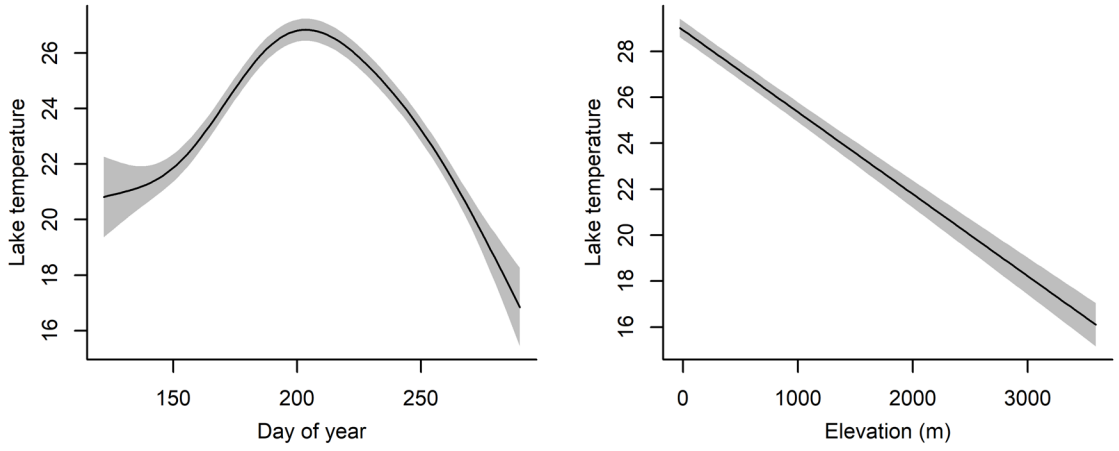


Figure 15. Relationship between lake temperature and sampling day (*left panel*) and elevation (*right panel*). Variables that are not plotted are fixed at their mean values. *Gray shading*: 90% confidence intervals; *solid lines* shows the mean relationships.

The pattern of temperature changes with time (see Figure 15) provides insight into the critical period during which the severity of deep water hypoxia can influence aquatic life in lakes. For most lakes, mixed layer temperatures increase in the spring and exceed critical temperatures for different species, at which point cool- and cold-water obligate species must move to deeper depths. Then, in the fall, decreasing mixed layer temperatures allow those species to move back to shallower waters. Models for $DO_m$ indicate that, in dimictic lakes after the onset of spring stratification, $DO_m$ decreases monotonically over time until fall turnover (see Figure 12). Therefore, the length of time between spring stratification and when mixed layer temperatures decrease below the critical temperature thresholds in the fall is a key factor for deriving a protective Chl *a* criterion.

EPA used documented temperature thresholds defined for cool- and cold-water fish as examples of critical mixed layer temperatures (Coker et al. 2001). For cool-water species, EPA identified an illustrative temperature threshold of 24 °C. Walleye, striped bass, and yellow perch are examples of lake fish that are members of that group (McMahon et al. 1984). For cold-water species, EPA identified an illustrative temperature threshold of 18 °C. Lake trout and cisco are examples of cold-water obligate species (Marcus et al. 1984, Jacobson et al. 2008). (Note: These references are only examples of the types of information that can be used to inform decisions regarding the critical temperature for different fish species and different fish life stages.) Then, given a lake's location and elevation, the lake temperature model predicts the day of the year that the mixed layer temperature would decrease below the critical temperatures. For cool-water species, mixed layer temperatures decreased below the critical temperature of 24 °C on days 210–260 (Figure 16), taking into account the fact that the dimictic lakes considered in this analysis are located in the northern half of the country (see Figure 7). Lakes in which mixed layer temperatures increased above 24 °C at some point during the year were predominantly located in the eastern U.S., as high elevations and climate in the western U.S. moderate lake temperatures. For cold-water species, mixed layer temperatures decreased below the critical temperature of 18 °C on days 220–280 (Figure 17). Temperatures in many lakes in the southeast part of the U.S. rarely decrease below the critical threshold in the summer, but those lakes also generally do not harbor cold-water fish.

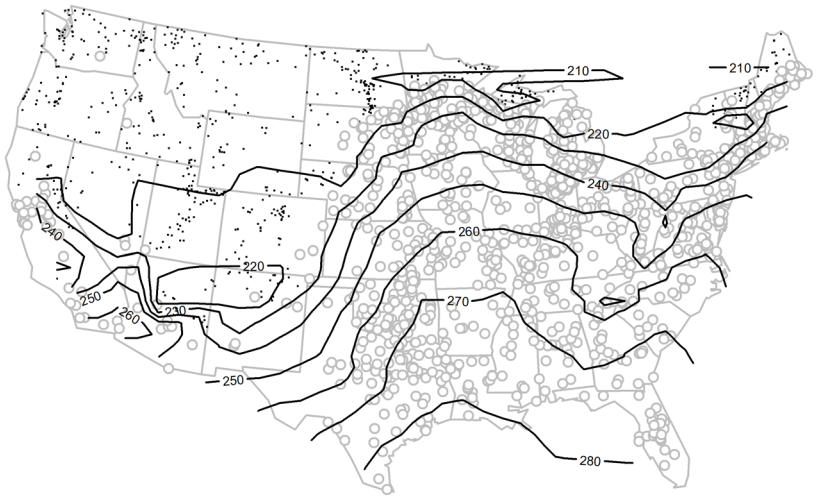


Figure 16. Days of the year that mixed layer temperatures decrease below the critical temperature for cool-water species. *Small dots*: lakes in which mixed layer temperatures never exceed 24 °C.

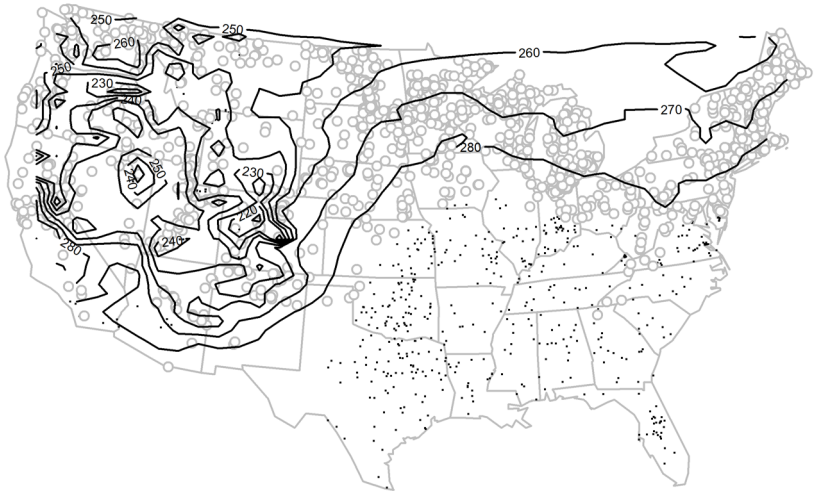


Figure 17. Days of the year that mixed layer temperatures decrease below the critical temperature for cold-water species. *Small dots*: lakes in which mixed layer temperatures do not decrease below 18 °C during the summer; *contours*: effects of large differences in elevation across lakes in the western U.S.

Criterion values for Chl *a* are calculated from the model equation for $DO_m$, rewritten here:

$$DO_m = DO_0 + [d_1 + d_2 \log(Chl_{mn}) + d_3 D + d_4 \log(DOC_{mn})](t - t_0) \quad (15)$$

Deep water DO concentrations depend not only on Chl *a* concentration, but also on the depth of the lake below the thermocline (*D*), DOC concentration ($DOC_{mn}$), and length of time that has elapsed since the establishment of stratification ($t - t_0$). A procedure for computing the day of the year, $t_{crit}$, at which mixed layer habitat is cool enough for different species to move to shallower water is also described above, highlighting the influence of lake location and elevation

as additional factors to consider. Based on these models, Chl *a* criteria for different lakes vary considerably depending on each lake's specific characteristics.

Prior to calculating a Chl *a* criterion, a threshold value for $DO_m$ must be selected. Existing EPA recommendations specify that the 7-day mean minimum DO concentration should be at least 5 mg/L to support cold-water fish (US EPA 1986). This threshold is also consistent with DO concentrations that fish have been observed to avoid in field studies (Coutant 1985, Plumb and Blanchfield 2009). A thin layer of cool water with sufficient DO provides a critical refuge for fish during the warmest periods of the year, and fish have been observed to seek out those cool water refuges. Observations of fish in warm lakes during the summer have indicated that they will congregate in cold water refuges as shallow as 30 centimeters (cm) (Coutant and Carroll 1980, Snucins and Gunn 1995, Baird and Krueger 2003, Mackenzie-Grieve and Post 2006). Hence, maintaining a DO concentration of at least 5 mg/L at a depth of 30 cm below the thermocline can provide a sufficient refuge for certain fish species and be protective of aquatic life. To convert this condition to a value of $DO_m$, EPA considered a simplified case in which DO linearly decreases from saturated conditions above the thermocline (DO = 8.4 mg/L at 24 °C) to a concentration of zero at some deeper depth (Figure 18). The linear decrease in DO is consistent with a steady-state solution of the diffusion equation, assuming a constant eddy diffusivity (Stefan et al. 1995). Based on this DO profile and the requirement that DO is 5 mg/L at 30 cm below the thermocline, an illustrative threshold value for $DO_m$ can be computed as 1.6 mg/L for a lake that is 2 m deep below the thermocline. That is, when the temperature profile is as depicted in Figure 18, depth-averaged DO computed for the water column below the thermocline is 1.6 mg/L. Other thresholds for $DO_m$ specific to different species of fish and different depths can also be calculated. For example, the threshold value for $DO_m$ for a lake that is 10 m deep below the thermocline would be 0.3 mg/L.

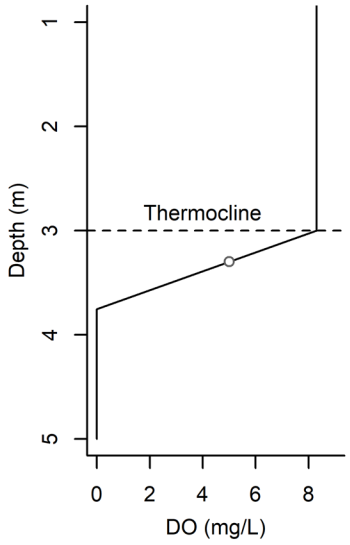


Figure 18. Simplified DO profile used to compute threshold for $DO_m$. *Open circle*: the targeted condition of DO at 5 mg/L, 30 cm below the thermocline.

The influence of different factors on Chl *a* criterion can be visualized by computing criteria at median values of all covariates and then examining changes in criteria that occur with the change in a single covariate. The relationship between Chl *a* and $DO_m$ at median values for all other covariates are shown as solid lines in each panel of Figure 19. Lakes in which covariate values differ from the medians of the data set cause changes in the candidate Chl *a* criteria. For cool-water species, using the illustrative threshold temperature, the median number of days between spring stratification and release of the temperature constraint in the mixed layer was 135 days. The 75th percentile of this day range, corresponding to lakes in warmer climates, was 151 days, whereas the 25th percentile, corresponding to lakes in cooler climates, was 116 days. When the critical window for maintaining sufficient DO in the deeper waters decreases to 116 days, the corresponding Chl *a* criterion increases to 11 µg/L, whereas in lakes in which the critical window is 151 days long, the Chl *a* criterion is 2 µg/L (left panel, Figure 19).

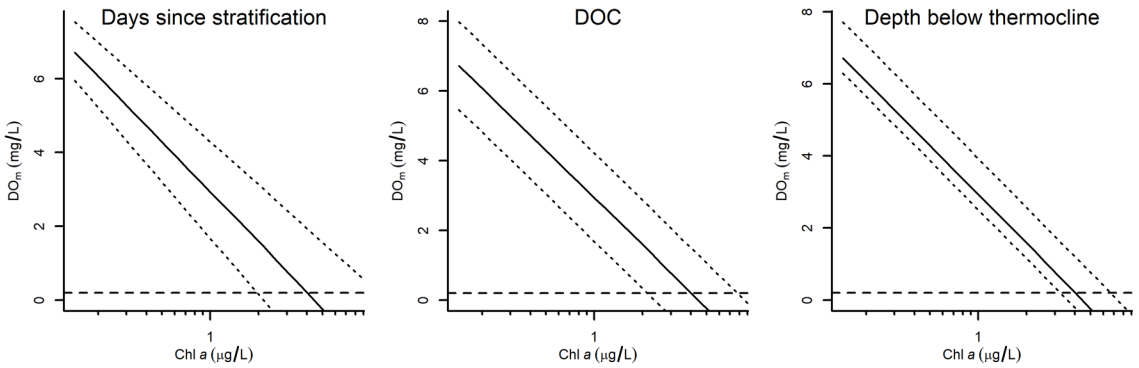


Figure 19. Effects of other predictors on Chl *a* criteria. *Solid lines*: relationship between Chl *a* and $DO_m$ at median values for all other variables; *dashed line*: $DO_m$ = 0.3 mg/L; *dotted lines*: 25th and 75th percentiles of days elapsed since stratification (*left panel*), 25th and 75th percentiles of mean DOC concentrations (*middle panel*), and depth below thermocline of 4 m and 20 m (*right panel*).

Similar ranges of criteria can be calculated for changes in DOC and the lake depth below the thermocline. The median concentration of DOC in the available data was 5 mg/L, but in lakes in which DOC is 3 mg/L (the 25th percentile of observed DOC in the data), the Chl *a* criterion increases to 8 µg/L; and in lakes in which DOC is 7 mg/L (the 75th percentile), the Chl *a* criterion decreases to 2 µg/L (middle panel, Figure 19). Finally, the median lake depth below the thermocline was 9 m. In a deeper lake, with 20 m of water below the thermocline, the Chl *a* criterion increases to 7 µg/L; but in a shallower lake, with only 4 m of water below the thermocline, the Chl *a* criterion decreases to 3 µg/L (right panel, Figure 19).

To better illustrate the possible range of criteria, EPA computed illustrative Chl *a* criteria for each of the dimictic lakes sampled in the NLA. Because those lakes represent a random sample of the population of lakes in the U.S., the resulting Chl *a* criteria are a representative distribution of criteria, providing insight into likely criteria for different types of lakes. For dimictic lakes harboring cool-water species (again, using the illustrative temperature threshold), the median Chl *a* criteria is 3.4 µg/L, and the range defined by the 25th and 75th percentiles is 1.3–10.6 µg/L. For lakes harboring cold-water species, using the illustrative temperature threshold, the median Chl *a* criterion is 1.8 µg/L, with a range of possible values extending from 1–7.6 µg/L.

In states where measurements of profiles of DO are available, these data can be readily modeled in conjunction with the national data (see Appendix B). In the example shown in Appendix B, modeling temporally resolved DO profiles from one state with the national data improved the precision of estimates of the first day of stratification. Because of this

improvement in model precision, the results of the combined state-national model are provided in the interactive criterion derivation tool.

The interactive tool used for estimating candidate Chl *a* criteria is provided at https://nsteps.epa.gov/apps/chl-hypoxia. With this tool, the user can specify lake physical characteristics that influence the relationship between Chl *a* and $DO_m$ as well as management decisions about targeted conditions that affect the criterion.

**Considerations: Lake physical characteristics.** Lake physical characteristics that are specified include the lake location (latitude and longitude) and lake elevation. That information is converted to an estimate of mean annual air temperature and, coupled with the model results, these data provide an estimate of the date of spring stratification. Note, though, that this estimate applied most accurately to dimictic lakes, and estimates of the first day of stratification for other type of lakes (e.g., polymictic) are more uncertain and may require additional data. Other lake physical characteristics that are specified are lake depth below the thermocline and average lake DOC concentration, factors that influence $DO_m$.

Water quality management decisions that influence the calculated criterion include parameters that define fish habitat (i.e., the critical maximum temperature for fish species in the lake, the threshold DO concentration, and the depth of the summer refugia) and the certainty level.

**Considerations: Fish habitat.** The critical maximum temperature for fish species in the lake is used to calculate the average day of the year that temperature constraints are released in the epilimnion. That is, the annual temperature model (see Figure 15) is used to identify the date that fish can potentially move to oxygen-rich shallower waters. The threshold DO concentration for the fish (e.g., a DO concentration of 5 mg/L for cold-water fish) and the desired minimum thickness of the refugia (e.g., 30 cm) are used to compute the targeted condition for $DO_m$. That targeted value of $DO_m$ is the minimum concentration required on the days prior to the release of temperature constraints.

**Considerations: Certainty level.** The certainty levels, as with other criteria, provide additional assurance that the calculated criterion is protective, based on the data and model uncertainty. For example, selecting the 50% certainty level implies that, at the estimated Chl *a* criterion, only 25% of predicted mean values of $DO_m$, based on the data, were less than targeted value. In statistical hypothesis testing, convention suggests that *p*-values of 1% or 5% are statistically significant results, which can also inform the selection of the percentile, but selection of the

certainty level as the basis for the criterion is ultimately a water quality management decision. More conservative certainty levels (e.g., higher percentiles) may be appropriate when applying this response in non-dimictic lakes in which the first day of stratification is more uncertain.

The interactive tool uses posterior simulation with model parameter distributions to predict the $DO_m$ on the critical day prior to a release from temperature constraints in the surface layer for different Chl *a* concentrations. These model results can be used to help derive criteria for a specified threshold $DO_m$. Samples with covariate values similar to those selected by the user are highlighted in the provided plots in the app.

### 3.2.3 Microcystin Concentration

The model for microcystin relates Chl *a* concentration to a distribution of microcystin concentrations. By specifying different targets for microcystin concentration, this model can inform the derivation of criterion values for the protection of drinking water uses or recreational uses.

#### *3.2.3.1 Statistical analysis*

A network of relationships can be specified that reflects current understanding of the linkage between lake eutrophication (as represented by Chl *a*) and increased concentrations of microcystin in individual samples (Figure 20). At the bottom of the diagram, cyanobacterial biovolume is directly associated with microcystin. Cyanobacterial biovolume is then expressed as the product of total phytoplankton biovolume and the proportion of the biovolume that is cyanobacteria (i.e., the relative biovolume of cyanobacteria), which clarifies the nature of the relationship between Chl *a* and cyanobacterial biovolume. More specifically, Chl *a* is directly proportional to phytoplankton biovolume (repeating the relationship used in the zooplankton model) (Kasprzak et al. 2008), and, as Chl *a* increases, the relative biovolume of cyanobacteria has been observed to increase (Downing et al. 2001).

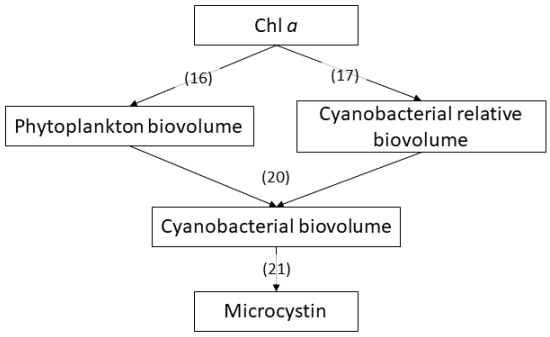


Figure 20. Schematic showing relationship between different variables predicting microcystin. *Numbers in parentheses*: refer to equation numbers in the text.

Each of the relationships in the network described above is expressed mathematically in the Bayesian network. First, phytoplankton biovolume, $P_i$, is modeled as being directly proportional to Chl *a* concentration ($Chl_i$), in sample *i*:

$$P_i = k_{c,i} Chl_i \tag{16}$$

The reciprocal of the parameter $k_{c,i}$ is the average amount of Chl *a* per unit biovolume of phytoplankton. Because the Chl *a* content of phytoplankton can vary with environmental conditions and assemblage composition, different values of this parameter are estimated for each sample, *i*. The overall distribution of the set of values for $k_{c,i}$ is assumed to be log-normal with a mean value of $\mu_k$ and a standard deviation of $\sigma_k$.

Exploratory analysis indicated that a quadratic function provided a reasonable representation of the relationship between the expected relative biovolume of cyanobacteria, $p_c$, and Chl *a* as follows:

$$E[logit(p_{c,i})] = f_1 + f_2 chl_i + f_3 chl_i^2 \tag{17}$$

where $f_1$, $f_2$, and $f_3$ are coefficients estimated from the data.

Because laboratory replicates of $P_i$ and $p_{c,i}$ were available, uncertainty associated with measuring phytoplankton and relative biovolume of cyanobacteria was estimated as follows:

$$\log(Bm_j) \sim Normal(\log(B_{i[j]}), s_1) \tag{18}$$

$$logit(pm_{c,j}) \sim Normal(logit(p_{c,i[j]}), s_2) \tag{19}$$

where $Bm_j$ and $pm_{c,j}$ are the laboratory measurements of phytoplankton biovolume and proportion cyanobacteria, respectively, and the index *j* maps replicate measurements to the corresponding estimate of the true value of the measurement for sample *i*. These laboratory replicates are assumed to be normally distributed about their respective estimates of the transformed sample means, with standard deviations of $s_1$ and $s_s$, respectively.

Cyanobacterial biovolume (*C*) can then be expressed as the product of the relative biovolume of cyanobacteria and total phytoplankton biovolume. After log-transforming, the expression is as follows:

$$\log(C_i) = \log\left(k_{c,i}\right) + \log\left(p_{c,i}\right) + \log(Chl_i) \quad (20)$$

where cyanobacterial biovolume in sample *i* is the sum of a log-transformed parameter $k_c$, the log-transformed cyanobacterial relative biovolume in the sample, and the log-transformed Chl *a* concentration.

The final component of the model relates cyanobacteria biovolume to microcystin. Initial exploration of the data indicated that microcystin increases at a rapid rate relative to cyanobacterial biovolume at high levels of cyanobacteria. At low levels of cyanobacteria, however, microcystin increases at a somewhat lower rate. To account for this change in rate, microcystin was modeled with a piecewise linear model as follows:

$$\log\left(\mu_{MC,i}\right) = g(\log\left(C_i\right)) \quad (21)$$

where the response variable in this relationship is $\mu_{MC,i}$, the estimated mean concentration of microcystin in sample *i*. The function *g*(.) is the piecewise linear function, which is characterized by four parameters: the intercept, $d_1$, and slope, $d_2$, of the first segment; the point along the gradient at which the slope changes, $c_p$; and the slope of the second segment, $d_3$.

The distribution of observed microcystin concentrations about the mean value was then modeled as a negative binomial distribution as follows:

$$MC_i \sim NB(\mu_{MC,i}, \varphi) \quad (22)$$

where $MC_i$ is the microcystin observed in sample *i* and *NB(.)* is a negative binomial distribution with overdispersion parameter, φ. Because the negative binomial distribution specifies only nonnegative integer outcomes, before fitting the model, EPA multiplied microcystin measurements by 10 and rounded to the nearest integer. Microcystin measurements below the

detection limit of 0.1 µg/L were set to zero (Yuan and Pollard 2017). Weakly informative prior distributions were assigned to all parameters in the model.

#### *3.2.3.2 Results*

A total of 2,352 observations of microcystin, cyanobacterial and phytoplankton biovolume, and Chl *a* were available from the NLA data set for analysis. Those measurements were collected from 1,116 different lakes spanning the conterminous U.S. An additional 112 samples of laboratory replicates of phytoplankton and cyanobacterial biovolume measurements were available to quantify measurement variability.

Three different relationships were estimated in the national model: (1) Chl *a* and phytoplankton biovolume, (2) Chl *a* and cyanobacterial relative biovolume, and (3) cyanobacterial biovolume and microcystin. (The relationship between phytoplankton biovolume, cyanobacterial relative biovolume, and cyanobacterial biovolume required no statistical estimation.) The observed relationship between Chl *a* and phytoplankton biovolume was accurately represented as a line with a slope equal to 1 on log-log axes (left panel, Figure 21), similar to the relationship estimated in the zooplankton model.

Cyanobacterial relative biovolume exhibited an increasing relationship with Chl *a* (middle panel, Figure 21). The quadratic functional form allowed the model to represent the steepening of the relationship at higher concentrations of Chl *a*. Mean microcystin increased with cyanobacterial biovolume (right panel, Figure 21). The slope of the relationship increased at a cyanobacterial biovolume of 1.9 $mm^3/L$, but the 90% credible interval on the location of this changepoint ranged from 0.5 to 5 $mm^3/L$. At cyanobacterial biovolumes greater than the changepoint, the slope of the mean relationship was statistically indistinguishable from 1, whereas at cyanobacterial biovolumes less than the changepoint, the slope was 0.61, with the 90% credible interval ranging from 0.51 to 0.69. Overall, the credible interval about the cyanobacteria-microcystin relationship was narrow compared to those estimated for the Chl *a*-cyanobacterial relative biovolume relationship as shown.

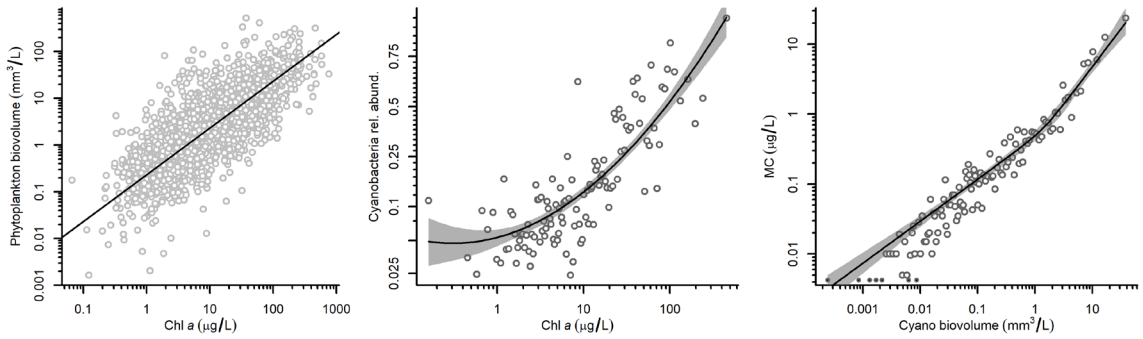


Figure 21. Modeled relationships for the microcystin model. *Left panel*: relationship between Chl *a* and phytoplankton biovolume; *open circles*: observed measurements of Chl *a* and phytoplankton biovolume; *solid line*: has a slope of 1. *Middle panel*: relationship between Chl *a* and cyanobacterial relative biovolume; *open circles*: average cyanobacterial relative biovolume in ~20 samples at the indicated Chl *a* concentration; *solid line*: estimated mean relationship; *gray shading*: 5th and 95th percentiles of the distribution of possible mean relationships (i.e., the 90% credible interval); *vertical axis*: has been logit-transformed. *Right panel*: relationship between cyanobacterial biovolume and microcystin; *open circles*: average microcystin in ~20 samples at the indicated cyanobacterial biovolume; *solid line*: estimated mean relationship; *gray shading*: 90% credible intervals about the mean relationship; *small filled circles*: Chl *a* bins in which microcystin in all samples was zero.

#### *3.2.3.3 Chl a criteria derivation*

Chl *a* criteria to protect recreational uses and drinking water sources can be derived from the estimated network of relationships by combining the model equations for total phytoplankton biomass, cyanobacterial-relative biovolume, and microcystin and the uncertainty inherent in each of the relationships (Figure 22). More specifically, based on a threshold concentration for microcystin and an allowable exceedance frequency of that threshold, Equation (22) can be used to compute the mean predicted microcystin that would be associated with these values. Then, Equations (20) and (21) can be used to calculate the Chl *a* concentration associated with this mean microcystin. This model is based on instantaneous measurements of Chl *a*, cyanobacterial biovolume, and microcystin. To relate instantaneous Chl *a* concentrations to a seasonal mean Chl *a* concentration, EPA computed the variance of Chl *a* concentrations within lakes over the summer sampling season using repeat visits included in the NLA data set. Then, the variance was used to estimate the probability of exceeding an instantaneous Chl *a* concentration, based on the seasonal mean Chl *a* concentration.

Threshold concentrations for microcystin have been published, and those targeted conditions can guide the use of the models to derive Chl *a* criteria. To protect sources of drinking water, the EPA uses a health advisory that recommends a threshold concentration for microcystin of 0.3 μg/L for preschool children less than 6 years old (US EPA 2015b). This short-term health

advisory does not include consideration of exposure to other sources (i.e., a relative source contribution term). Typically, EPA would use an ambient water quality criterion (developed under Section 304(a) of the CWA) as the endpoint for the protection of human health. The criterion would include consideration of exposure through drinking water and fish consumed, if both are designated uses for a waterbody. The criteria also provide estimates for only fish consumption. Because human health criteria are not available for microcystins, EPA recommends using the drinking water health advisory, recognizing that this threshold is an underestimate of risk for water bodies where consumption of fish is an important route of exposure.

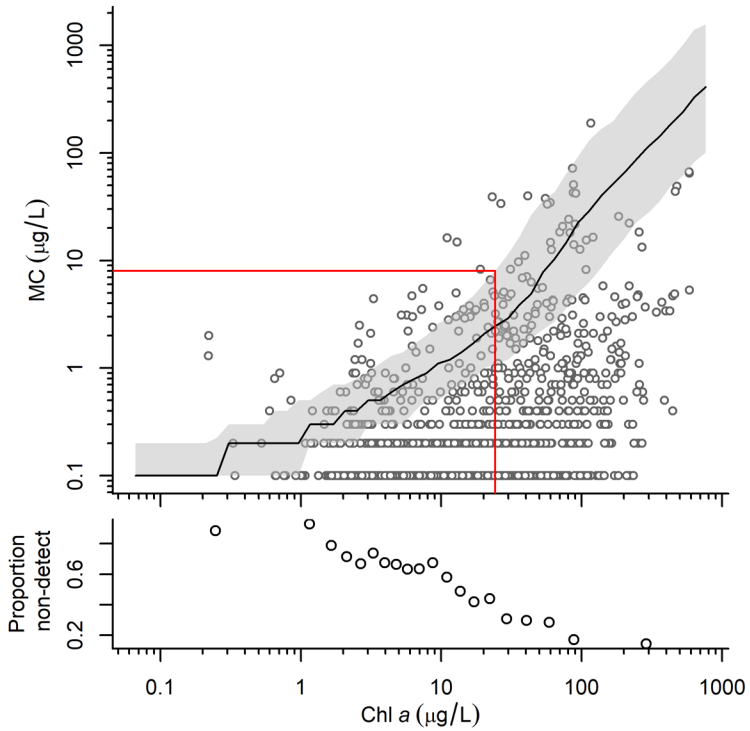


Figure 22. Example of derivation of Chl *a* criterion to protect recreational uses based on targeted microcystin of 8 µg/L and exceedance probability of 1%. *Top panel–open circles*: observed values of microcystin and Chl *a* for samples in which microcystin was greater than the detection limit; *solid line*: predicted microcystin that will be exceeded 1% of the time for the indicated Chl *a* concentration; *gray shading*: 25th and 75th percentiles of the distribution of possible mean relationships (i.e., the 50% credible interval); *solid vertical and horizontal line segments*: candidate Chl *a* criterion based on targeted microcystin. *Bottom panel*: proportion of samples for which microcystin was not detected in ~100 samples centered at the indicated Chl *a* concentration.

The drinking water health advisory applies to finished drinking water, but national ambient water quality criteria are designed to protect sources of drinking water. Human health criteria focus on water sources prior to treatment for two main reasons. First, the ability of treatment technologies to remove microcystin is too variable (Hoeger et al. 2005, Westrick et al. 2010, US EPA 2015c) for EPA to set a national recommendation for a protective ambient source water concentration that would yield a protective concentration after treatment. Second, consistent with the core statutory objective of the CWA Section 101(a) " to restore and maintain the chemical, physical, and biological integrity of the Nation's waters", the criteria are premised on the position that ambient waters should not be contaminated to a level where the burden of achieving health objectives is shifted away from those responsible for pollutant discharges and placed on downstream users such as drinking water utilities to bear the costs of upgraded or supplemental water treatment. EPA is aware that states or authorized tribes express water quality standards for protecting drinking water sources as either protecting the ambient source water before treatment or after treatment. If a state or authorized tribe applies the health advisory standard to drinking water after treatment, then they can account for the expected treatment in their facilities and select a higher microcystin concentration in the ambient source water that would result in the targeted microcystin concentration in the finished drinking water. Doing so will result in a concentration of Chl *a* in the ambient source water that will protect human health from the effects of microcystin in the finished drinking water. To protect recreational uses, EPA recommends a threshold concentration for microcystin of 8 μg/L to protect children (US EPA 2019). This threshold was based on incidental ingestion of water during recreation.

After selecting the designated use of interest, calculating the corresponding Chl *a* criterion requires two additional management decisions: selection of the allowable exceedance probability of the threshold and selection of a certainty level. These decisions are combined with a posterior simulation using the estimated distributions of the model parameters to estimate Chl *a* criteria.

**Considerations: Allowable exceedance probability**. The allowable exceedance probability can be interpreted directly in terms of environmental outcomes as the probability of observing a specified microcystin in a sample for a given seasonal mean Chl *a* concentration. For example, after accounting for model uncertainty by selecting a 50% certainty level, microcystin in lakes

with a seasonal mean Chl *a* concentration of 22 µg/L would be expected to exceed a threshold of 8 µg/L in 1% of samples (Table 3) (solid vertical line in Figure 22). With daily sampling, one interpretation of this result is that on average, microcystin would exceed the threshold 1 day out of every 100, so 1 day during the swimming season for many states would not support recreational use of the lake.

The frequency component of the recreational cyanotoxin criteria is specified in terms of excursions of the criterion magnitude during 10-day assessment windows (US EPA 2019), in which a water body is considered impaired for recreation when more than three excursions occur in 1 year. If we assume initially that daily samples are independent events, we can calculate the probability of at least one exceedance of the criterion magnitude in a 10-day window, and then compute a seasonal probability associated with greater than three excursions within all 10-day windows during a 100-day season. For example, if our single day exceedance probability is 0.05, there is a 62% chance that we will observe greater than three excursions during a 100-day season (Figure 23).

The initial assumption that daily samples of microcystin are independent from one another is conservative, as we would expect that a day with a high concentration of microcystin would be more likely to be followed by another day with high concentrations. That is, cyanobacterial blooms and the associated increases in microcystin concentration are likely to be clumped in time. This tendency for temporal autocorrelation may lower the computed seasonal probabilities, as we would expect somewhat fewer excursions during 10-day assessment windows if daily observations of elevated microcystin occur in groups. Overall, this information, or similar calculations tailored to conditions within a particular lake can further inform the selection of exceedance probability when deriving the candidate criterion.

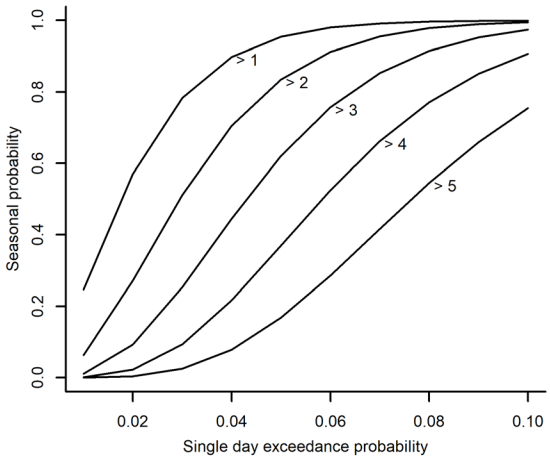


Figure 23. Probability of the indicated number of excursions observed in 10-day assessment windows during a 100-day season for different single day exceedance probabilities.

**Considerations: Certainty level.** The certainty level, as quantified by the credible interval, expresses the uncertainty in the model predictions of different exceedance probabilities. So, the shaded area in Figure 22 shows the range over which at least 50% of the possible curves would be located that describe microcystin concentrations that have a 1% probability of exceedance. Selection of a higher certainty level yields more conservative criteria in terms of model uncertainty. An interactive tool allowing the user to examine Chl *a* criteria associated with different combinations of microcystin threshold, probability of exceedance, and certainty level is available at https://nsteps.epa.gov/apps/chl-microcystin.

Table 3. Illustrative Chl *a* criteria (µg/L) for different exceedance probabilities using a 50% certainty level and a microcystin threshold of 8 µg/L**.**

| Probability of exceedance | Candidate Chl *a* criterion (µg/L) |
|---|---|
| 1% | 22 |
| 5% | 31 |
| 10% | 39 |

### 3.2.4 Phosphorus-Chlorophyll *a*

A TP measurement is composed of P contained within different compartments, including P bound in phytoplankton, P bound to suspended sediment, and dissolved P (i.e., chemically dissolved P and P bound to particles small enough to pass through a filter) (Effler and O'Donnell 2010). In many lakes, much of measured TP is associated with phytoplankton, and so,

differences in phytoplankton biomass among lakes can be associated with differences in both Chl *a* and TP, yielding a strong correlation between the two (Lewis and Wurtsbaugh 2008). In other lakes, high concentrations of suspended sediment can contribute to TP and affect observed TP-Chl *a* relationships (Jones and Knowlton 2005). When TP-Chl *a* relationships are being estimated, lakes with high concentrations of suspended sediment show low Chl:TP ratios relative to the average pattern (Hoyer and Jones 1983, Jones and Knowlton 2005).

EPA modeled the relationship between TP and Chl *a* by explicitly accounting for the contributions of different compartments to observed TP, resulting in the positions of TP and Chl *a* in the model equation being reversed from the typical model formulations. The model explained variations in TP in various compartments, rather than explaining variation in Chl *a* (Yuan and Jones 2020b).

#### *3.2.4.1 Statistical analysis*

EPA specified a model that estimates contributions to TP from different compartments, where TP is modeled as the sum of contributions from dissolved P, P bound to nonphytoplankton sediment, and P bound in phytoplankton (Figure 24).

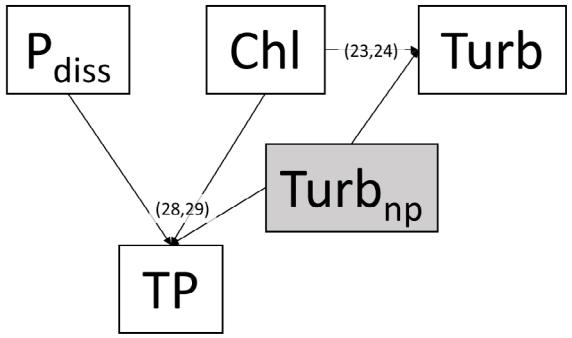


Figure 24. Schematic representation of compartment model for TP. $P_{diss}$: dissolved P; Chl: Chl *a*; Turb: total turbidity; $Turb_{np}$: turbidity attributed to nonphytoplankton sources. *Shaded box for* $Turb_{np}$: a variable inferred by the model; *numbers in parentheses*: refer to equation numbers in the text.

Direct measurements of nonphytoplankton sediment were not collected during the NLA. Instead, turbidity measurements were available that are associated with total suspended solids and include contributions from both nonphytoplankton and phytoplankton components. Because an estimate of nonphytoplankton sediment is needed to model TP, turbidity is modeled as the sum of two components: (1) turbidity that is directly associated with phytoplankton biomass, or autochthonous suspended sediment ($Turb_{aut}$) and (2) turbidity associated with all

other sources, or nonphytoplankton turbidity ($Turb_{np}$). The second component of turbidity includes turbidity associated with allochthonous sediment and sediment resuspended from the lake basin (Hamilton and Mitchell 1996). EPA modeled $Turb_{aut}$ as being directly proportional to Chl *a* (Jones et al. 2008), a measure of algal biomass and, therefore, the components of turbidity were expressed as follows:

$$E[Turb] = Turb_{np} + Turb_{aut} = Turb_{np} + fChl \quad (23)$$

where *E[Turb]* indicates that the model applies to the expected value of turbidity (Turb). The amount of $Turb_{aut}$ associated with each unit of Chl *a* is expected to vary with algal composition. For example, small phytoplankton species would tend to scatter light differently than larger species. Assuming that algal composition changes with trophic conditions (Godfrey 1982), the change in algal composition can be modeled by expressing the coefficient *f* as an unknown function of Chl *a*. Also, assuming that *f(Chl)* can be modeled as a power function ($f = bChl^m$), the product of *f(Chl)* and Chl *a* can be written as follows as $bChl^k$ without any loss of generality:

$$E[Turb] = Turb_{np} + f(Chl)Chl = Turb_{np} + bChl^k \quad (24)$$

where the exponent, *k*, is equal to *m*+1.

Exploratory analysis indicated that concentrations of $Turb_{np}$ varied with different lake characteristics, but the predictor that accounted for the most variability was lake depth. Therefore, 30 classes of lakes based on maximum depths were defined, and the value of $Turb_{np}$ within each of the classes was modeled as a log-normal distribution about a mean value specific to that depth class as follows:

$$\log(Turb_{np}) \sim Normal(\mu_{a,i},\ \sigma_a) \quad (25)$$

where $\mu_{a,i}$ is the mean value of log($Turb_{np}$) for depth class *i*, and $\sigma_a$ is the standard deviation of the distribution of individual measurements of $Turb_{np}$. The set of values for $\mu_{a,i}$ was then assumed to be drawn from a single normal distribution as follows:

$$\mu_{a,i} \sim Normal(\mu,\ \sigma_\mu) \quad (26)$$

where $\mu$ and $\sigma_\mu$ are the mean and standard deviation of this distribution. The mean distribution loosely constrains the possible values of $\mu_{a,i}$, while allowing lakes with smaller amounts of data to "borrow information" from lakes with larger amounts of data (Gelman and Hill 2007).

Finally, sampling variability for Turb was assumed to be log-normally distributed as follows:

$$\log(Turb) \sim Normal(E[\log(Turb)], \sigma_T) \quad (27)$$

where *E[log(Turb)]* is the expected value of log(Turb) expressed in Equation (24).

EPA used results from the model for turbidity simultaneously to estimate contributions to different components of TP. Recall that TP is modeled as being composed of contributions from dissolved P ($P_{diss}$), P bound to suspended sediment, and P bound to phytoplankton. Based on this assumption, the following model equation can be written:

$$E[TP] = P_{diss} + g_1 Turb_{np} + g_2 Chl \quad (28)$$

where the concentration of P bound to nonphytoplankton suspended sediment is modeled as being directly proportional to $Turb_{np}$, and P bound to phytoplankton is modeled as being directly proportional to Chl *a*. The coefficient $g_1$ quantifies the P content of $Turb_{np}$, while the coefficent $g_2$ expresses P concentration relative to Chl *a* concentration. P content is expected to vary with the level of turbidity and the composition of the phytoplankton assemblage, so, similar to the model for turbidity, the coefficients $g_1$ and $g_2$ were allowed to vary as power functions of $Turb_{np}$ and Chl *a*, respectively. So, the final model equation can be written as follows:

$$E[TP] = P_{diss} + d_1 Turb_{np}^m + d_2 Chl^n \quad (29)$$

Exploratory analysis indicated that dissolved P was associated with lake depth, so, similar to $Turb_{np}$, different values of $P_{diss}$ were estimated for each of 30 lake depth classes as follows:

$$\log(P_{diss,mn,i}) \sim Normal(\mu_{diss}, \sigma_{diss}) \quad (30)$$

where $P_{diss,mn,i}$ is the mean dissolved P concentration in lake depth class i, and the overall mean value of $\log(P_{diss,mn,i})$ is $\mu_{diss}$ with a standard deviation of $\sigma_{diss}$.

Exploratory analysis also indicated that the P associated with each unit of $Turb_{np}$ and Chl *a* (i.e., the values of the coefficients $d_1$ and $d_2$) varied most strongly with geographic location. Because of that trend, different values for these coefficients were estimated for different Level III ecoregions. Ecoregion-specific values for these parameters were assumed to be drawn from log-normal distributions as follows:

$$\log(d_{1,i}) \sim Normal(\mu_{d1}, \sigma_{d1})$$

$$\log(d_{2,i}) \sim Normal(\mu_{d2}, \sigma_{d2}) \quad (31)$$

where the index, *i*, refers to values of each parameter for different ecoregions.

Finally, sampling variation for TP was assumed to be log-normally distributed as follows:

$$\log(TP) \sim Normal(E[\log(TP)], \sigma_{TP}) \quad (32)$$

All the relationships described in this section on statistical analysis were fit simultaneously to the available data with a hierarchical Bayesian model (Stan Development Team 2016). Prior distributions for all model parameters were assumed to be weakly informative.

#### *3.2.4.2 Results*

Observations of turbidity were correlated with Chl *a*, and a distinct lower boundary in the scatter of data was evident (Figure 25). The model relationship defining this lower boundary can be computed by setting $Turb_{np}$ to zero in Equation (24). Then, after log-transforming, the equation can be written as $\log(Turb) = \log(b) + klog(Chl)$. In other words, when $Turb_{np}$ is negligibly small, the relationship between $Turb_{aut}$ and Chl *a* is a straight line in the plot of log(Chl) vs. log(Turb) (solid line in Figure 25). Deviations in sampled values above that line show the contribution of $Turb_{np}$ to the overall turbidity measurement. Mean values of *b* and *k* estimated from the model were 0.67 (0.62, 0.73) and 0.67 (0.65, 0.69) (90% credible intervals shown in parentheses). Based on the functional form that was assumed for the relationship between turbidity and Chl *a*, the contribution of phytoplankton to turbidity (i.e., $Turb_{aut}$/Chl *a*) was estimated as being proportional to $Chl^{-0.33}$. That is, as Chl *a* increases, the amount of turbidity associated with each unit of Chl *a* decreases, a trend that is consistent with a shift from small-bodied, diatom-dominated assemblages to colonies of cyanobacteria cells (Scheffer et al. 1997).

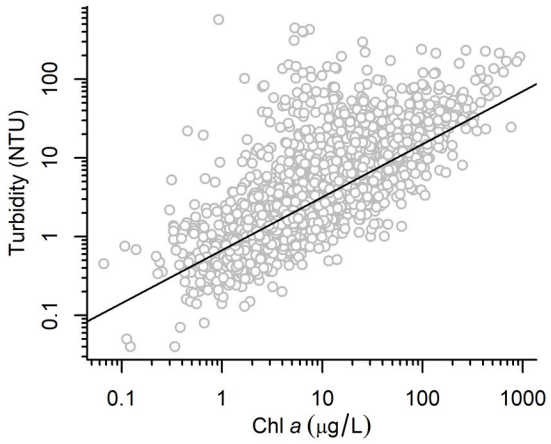


Figure 25. Turbidity vs. Chl *a*. *Solid line*: the limiting relationship between Chl *a* and turbidity when contribution of allochthonous sediment is negligible.

Estimates of $Turb_{np}$ and mean dissolved P both exhibited decreasing relationships with increasing depth (Figure 26). $Turb_{np}$ decreased from approximately 1.4 nephelometric turbidity units (NTU) in shallow lakes to nearly zero in deep lakes, while $P_{diss}$ varied from approximately 2.6 µg/L in shallow lakes to 1.6 µg/L in deep lakes. Both of these relationships are consistent with a mechanism by which fine sediment from the lake bottom is likely to be collected in surface water samples in shallow lakes. In the case of $P_{diss}$, the trend indicates that measurements of dissolved and particulate components of TP are determined by filter size and P bound to sediment fine enough to pass through the filter contributes to estimates of dissolved P.

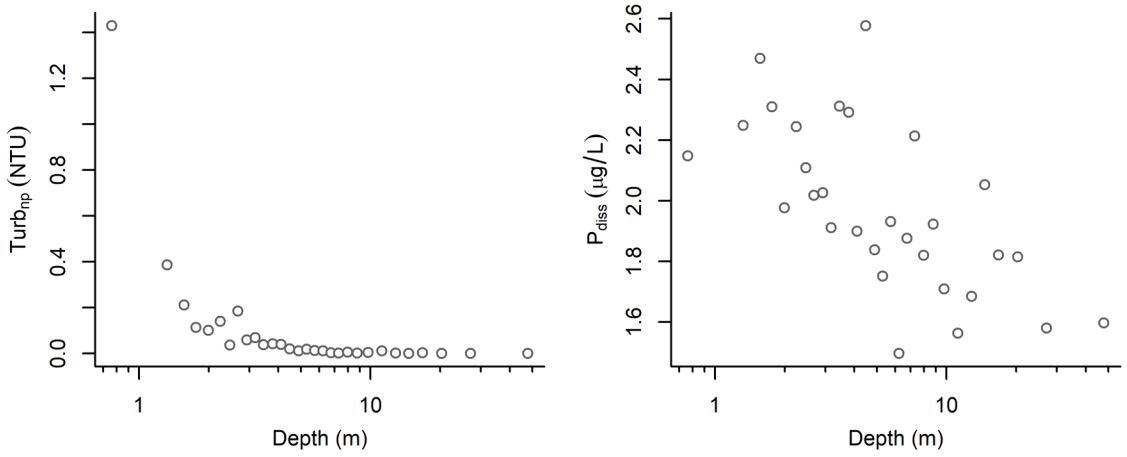


Figure 26. Relationship between $Turb_{np}$, $P_{diss}$, and lake depth. *Open circles*: mean estimate of parameter value in each of 30 lake depth classes.

The quantity of P bound to nonphytoplankton suspended sediment expressed by the coefficient $d_1$ exhibited substantial geographic variation (Figure 27). Coherent spatial patterns could be discerned in the variation of $d_1$ among different states, with relatively high levels of P content in the upper Midwest region of the country (e.g., Montana, North Dakota, and South Dakota) as well as in parts of the western mountains. Comparatively lower levels of P content were observed in the northeast region of the U.S. Mechanisms for these large-scale variations in P content are likely related to the underlying geology of soils in each region (Olson and Hawkins 2013). Values of $d_2$, the amount of P within phytoplankton, spanned a much narrower range than estimated for $d_1$, only ranging from 1.6 to 4.5 per unit of Chl *a*. The relative difference in regional variability in the coefficients indicates that spatial differences in the amount of P bound to nonphytoplankton suspended sediment account for more of the variability in TP-Chl *a* relationships than spatial differences in P within phytoplankton, and the amount of P residing in phytoplankton is relatively constant. The effects of differences in the amount of P bound to nonphytoplankton sediment can be taken into account by computing ecoregion-specific TP criteria.

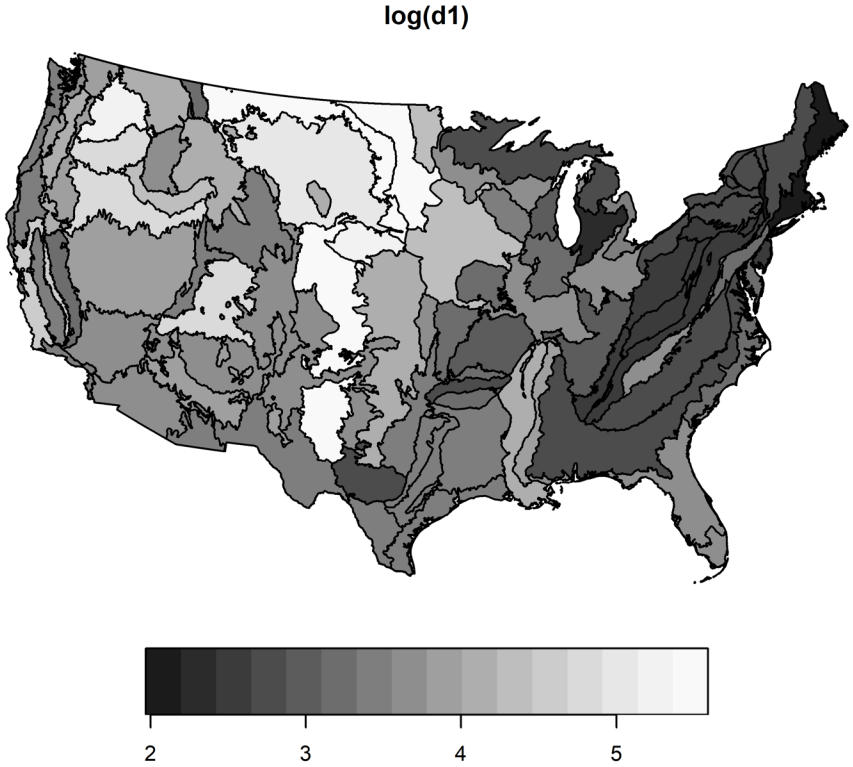


Figure 27. Ecoregion-specific values of $\log_e(d_1)$, P bound to nonphytoplankton suspended sediment.

Limiting relationships that estimate the P content of phytoplankton biomass and $Turb_{np}$ can also be calculated (Figure 28). For phytoplankton biomass, the limiting relationship is calculated by setting $P_{diss}$ and $Turb_{np}$ in Equation (29) to zero, yielding the following log-transformed relationship: $\log(TP) = \log(d_2) + n\ log(Chl)$. Different values of $d_2$ were estimated for each ecoregion, but the distribution of those values is characterized by an overall mean value of 2.5 (2.0, 3.1), while the mean value of the parameter *n* was 0.87 (0.82, 0.92). The straight line based on the two parameter values represents P associated with phytoplankton biomass, as quantified by Chl *a*, and it tracks the lower limit of the observed data (solid line, right panel, Figure 28). As a limiting relationship, one would expect that the majority of values of TP would be greater than this line indicates, but variability associated with the value of $d_2$ causes some values of TP to fall below the limit.

For $Turb_{np}$, setting $P_{diss}$ and Chl *a* to zero yields the following relationship: $\log(TP) = \log(d_1) + m\log\ (Turb_{np})$. The mean value of the coefficient $d_1$ was 31 (23, 40), and the value of the exponent *m* was 0.35 (0.32, 0.40) (left panel, Figure 28). Overall, the RMS error for predicting $\log_e$(TP) was 0.52 for the model.

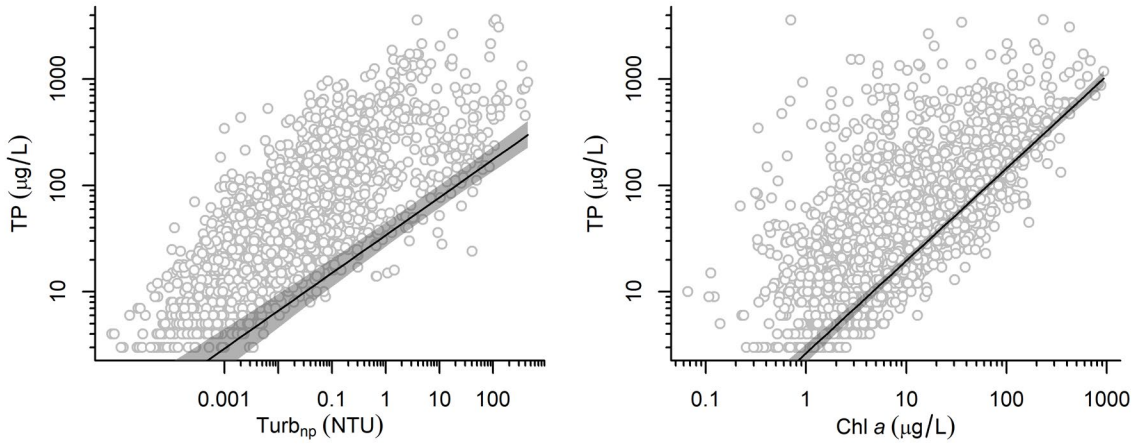


Figure 28. TP versus $Turb_{np}$ and Chl *a*. *Solid lines*: the limiting relationship between the indicated variable and TP; *gray shaded areas*: the 90% credible intervals about the mean relationship.

### *3.2.4.3 Phosphorus criteria*

Two relationships between Chl *a* and TP that can be inferred from the TP model inform the derivation of TP criteria. First, the limiting relationship between Chl *a* and TP estimated from the model quantifies the amount of P that is bound to phytoplankton (see Figure 28). This relationship predicts TP concentration in samples in which suspended sediment and dissolved P

concentrations are very low and defines the *minimum* value of TP that is associated with a targeted Chl *a* concentration. This limiting relationship can also be interpreted as the Chl *a* yield of P (Yuan and Jones 2019) and could be used to predict the change in Chl *a* that would potentially result from a change in the amount of biologically available P in the water column (Reynolds and Maberly 2002).

A second relationship between TP and Chl *a* accounts for contributions from P bound to nonphytoplankton sediment. If lake depth is specified, then the relationship estimated between lake depth and nonphytoplankton sediment can be used to estimate an average contribution to TP from these other compartments in the water column (see Figure 26). The resulting relationship then provides an estimate of the *ambient* TP concentration one would expect to observe as a function of Chl *a*, and therefore, provides a TP criterion value.

The prediction of ambient TP that accounts for contributions from nonphytoplankton sediment provides an estimate of the mean TP concentration that one would expect to observe for a given Chl *a* (Figure 29). As such, this ambient TP concentration provides a criterion. (Illustrative examples for TP criteria are shown in Table 4.) Note that contributions of $P_{diss}$ are not included in predictions of ambient TP criteria. In many lakes $P_{diss}$ is composed of more biologically available forms of P (e.g., soluble reactive P), and so, concentrations of $P_{diss}$ should be near zero in lakes in which reductions in P loading would be expected to influence phytoplankton abundance.

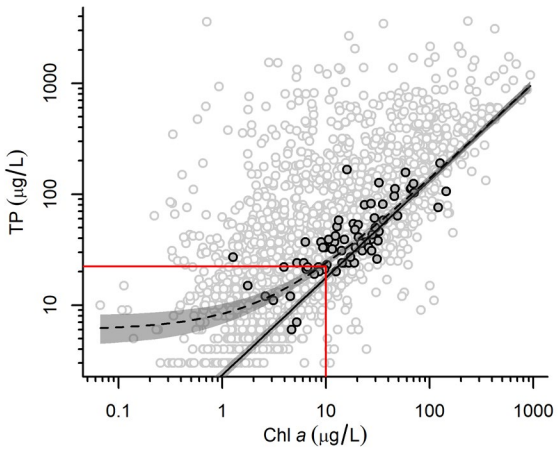


Figure 29. Example of deriving TP criteria for a Chl *a* target of 10 µg/L for data from one ecoregion (Southeastern Plains). *Open circles*: all data; *filled circles*: data from the ecoregion; *solid line*: limiting TP-Chl *a* relationship from compartment model; *dashed line*: ambient TP-Chl *a* relationship taking into account contributions from nonphytoplankton sediment for a 3-m deep lake; *solid horizontal and vertical line segments*: Chl *a* target and associated TP criterion; *shaded areas*: 80% credible intervals.

Table 4. Illustrative example of TP criteria corresponding to data shown in Figure 29. Example TP criteria for illustrative Chl *a* targets. TP criteria calculated for a 3-m deep lake.

| Chl *a* target | 80% certainty level | 50% certainty level |
|---|---|---|
| 10 µg/L | 21 | 22 |
| 15 µg/L | 28 | 30 |

The interactive tool for computing different TP criteria associated with Chl *a* is available at https://nsteps.epa.gov/apps/tp-tn-chl. This tool allows the user to specify the targeted Chl *a* concentration and the physical characteristics of the lake.

**Considerations: Targeted Chl *a* concentration**. The targeted Chl *a* concentration is the final criterion value derived for the lake(s) of interest after consideration of the different designated uses.

**Considerations: Lake characteristics.** Users can select the lake depth to be used when computing TP criteria. The coefficients $d_1$ and $d_2$ characterize the amount of P associated with nonphytoplankton suspended sediment and phytoplankton, respectively, and these coefficients vary among ecoregions (Figure 27). Therefore, users also can select an ecoregion for computing TP criteria. Data selected for an ecoregion are highlighted in the provided plots.

**Considerations: Certainty level.** Users can select a certainty level (as quantified by the credible interval) to account for the effects of model uncertainty on the calculated criteria. Basing criteria on the lower bound of the credible interval provides additional assurance that the calculated criterion is protective, given the data and model uncertainty. For example, selecting a certainty level of 50% dictates that the 25th percentile of the distribution of mean relationships is used to compute the lower bound of the envelope of possible relationships. A criterion based on this lower bound implies that only 25% of predicted TP concentrations at the selected Chl *a* concentration were less than the criterion value. In statistical hypothesis testing, convention suggests that *p*-values of 1% or 5% (corresponding to certainty levels of 99% and 95%, respectively) are statistically significant results. Those practices can also inform the selection of the percentile, but selection of the certainty level as the basis for the criterion is ultimately a water quality management decision.

### 3.2.5 Nitrogen-Chlorophyll *a*

Similar to the model for TP, each TN measurement is composed of N contained within three compartments: N bound in phytoplankton, dissolved inorganic N (i.e., nitrate, nitrite, and ammonia), and dissolved organic N (DON). Unlike the TP model, exploratory analysis indicated that the N content of inorganic suspended sediment was negligible (Yuan and Jones 2019).

#### *3.2.5.1 Statistical analysis*

Field measurements of the difference between TN and dissolved inorganic nitrogen (DIN = $NO_x$ + ammonia) were modeled as follows:

$$E[TN - DIN] = f_1 Chl^{k1} + DON = f_1 Chl^{k1} + f_2 DOC \quad (33)$$

where variations in the observations of total N minus dissolved inorganic N (*TN-DIN*) are attributed to two compartments: N bound in phytoplankton, modeled as $f_1 Chl^{k1}$ and DON. Exploratory analysis indicated that DON was closely associated with DOC, as they often originate from the same watershed sources (Berman and Bronk 2003), so the concentration of DON was modeled as being directly proportional to DOC.

As with the TP model, exploratory analysis indicated that the parameters $f_1$ and $f_2$ varied most strongly with geographic location. Because of those trends and to facilitate the use of this model with local data sets, different values of $f_1$, and $f_2$ were specified for each Level III ecoregion:

$$\log(f_{1,i}) \sim Normal(\mu_{f1}, \sigma_{f1})$$
$$\log(f_{2,i}) \sim Normal(\mu_{f2}, \sigma_{f2}) \quad (34)$$

where the parameters $\mu_{f1}$ and $\mu_{f2}$ estimate the mean values of the distribution of $f_1$ and $f_2$ while $\sigma_{f1}$ and $\sigma_{f2}$ estimate the standard deviations.

The sampling distribution of TN-DIN was assumed to be log-normally distributed as follows:

$$\log(TN - DIN) \sim Normal(E[\log(TN - DIN)],\ \sigma_{TN}) \quad (35)$$

where $\sigma_{TN}$ is the standard deviation of observed values of log(TN-DIN) about their expected value.

#### *3.2.5.2 Results*

A total of 2466 samples collected from 1875 lakes were available for analysis. Values for the coefficient, $f_1$, quantifying phytoplankton N content ranged from 11 to 43 in different ecoregions

with an overall mean value of 18.3 (14.9, 22.3). The values estimated for $f_2$ spanned a greater range among ecoregions with a minimum value of 35 and a maximum value of 103. The overall mean value of $f_2$ was 64.9 (61.0, 68.9). The broad range in values of $f_2$ indicates that strong differences exist among different locations regarding the nature of the relationships between DOC and DON. The mean value of the exponent, *k1*, was 0.90 (0.86, 0.94).

To visualize the variability in phytoplankton N among ecoregions, the concentration of N bound in phytoplankton at the overall mean Chl *a* concentration of 9.3 µg/L is mapped (Figure 30). With the exception of one high value of 320 µg/L estimated for the Sand Hills, Nebraska ecoregion, N-content of phytoplankton exhibited only small variations among ecoregions. N content ranged from 83 – 185 µg/L with a median value of 136 µg/L. Coherent spatial patterns in the N-content of phytoplankton were not evident.

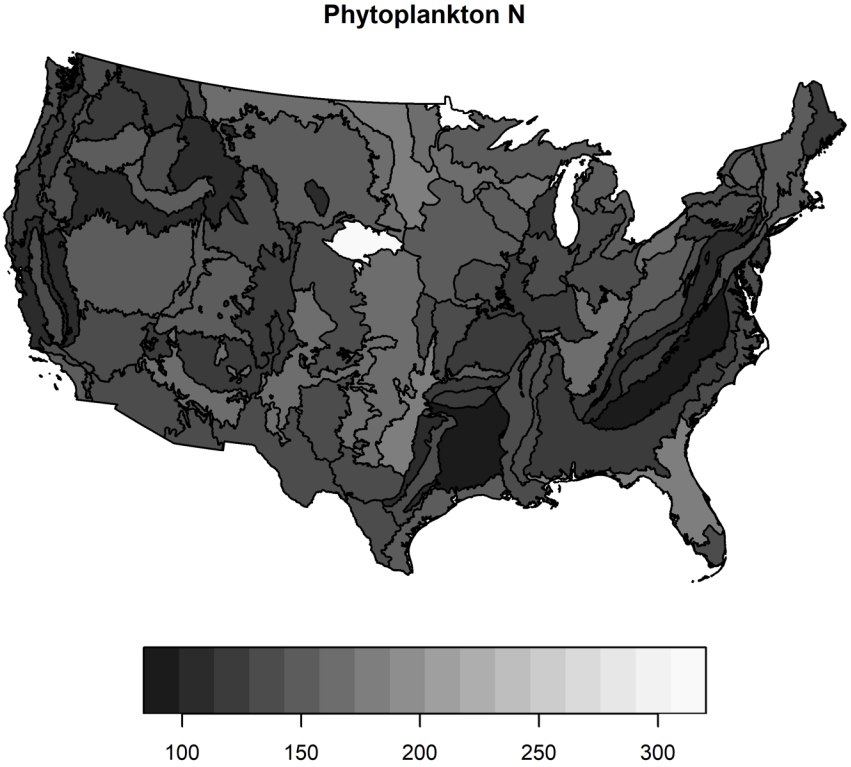


Figure 30. Variation in the concentration of N bound in phytoplankton among Level III ecoregions at the overall mean Chl *a* = 9.3 µg/L. Gray scale shows N concentrations in µg/L.

Estimated DON concentrations at the overall mean DOC concentration of 5.6 mg/L ranged from 194 – 570 µg/L with a median concentration of 365 µg/L (Figure 31). Variations in DON among ecoregions were substantially greater than observed for phytoplankton N. Spatial patterns were

also evident, with higher concentrations of DON in the upper Midwest regions of the U.S. and lower concentrations in the mountains in the western and eastern regions of the country.

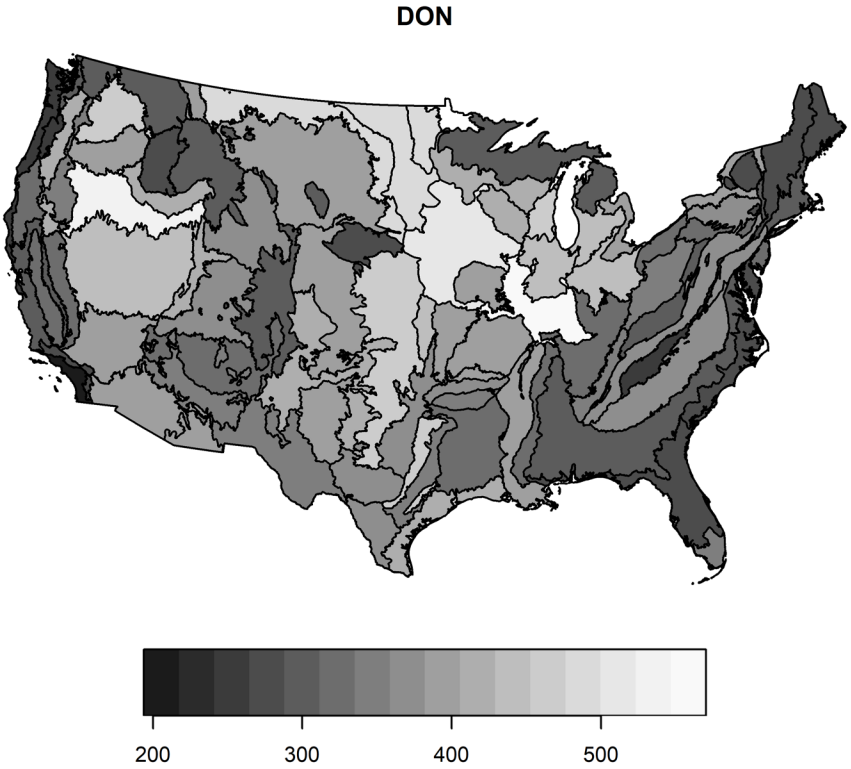


Figure 31. Variation in DON among Level III ecoregions at an overall mean DOC = 5.6 mg/L. Gray scale shows N concentrations in µg/L.

EPA calculated limiting relationships that estimate the N content of phytoplankton biomass with a procedure identical to that used for TP (Figure 32). In this case, the limiting relationship was calculated by setting the contribution from DON in Equation (33) to zero, yielding the following log-transformed relationship: $\log(TN - DIN) = \log(f_1) + k\ log(Chl)$. The straight line based on those two parameter values represents N associated with phytoplankton biomass, as quantified by Chl *a*, and it tracks the lower limit of the observed data (solid line, left panel Figure 32).

Similarly, setting DIN and Chl *a* to zero in Equation (33) yields the following limiting relationship for DON: $\log(TN) = \log(f_2) + \log(DOC)$ (solid line, right panel Figure 32). The mean value of $f_2$ indicates that, on average, the concentration of DON was 0.065 times that of DOC. Overall, the RMS prediction error for $\log_e$(TN-DIN) was 0.37.

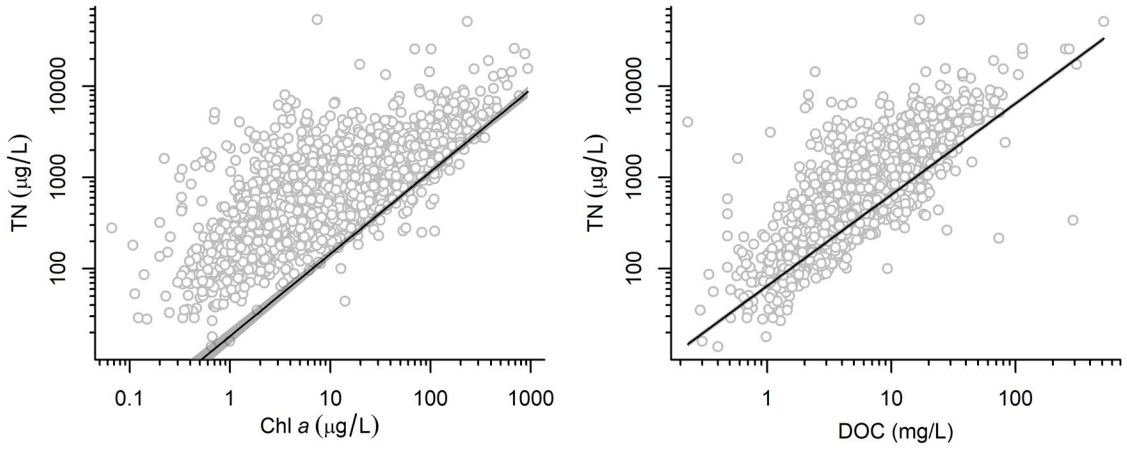


Figure 32. TN-DIN vs. Chl *a* and DOC. *Solid lines*: the limiting relationship between each variable and TN-DIN; *shaded area*: the 95% credible intervals about this mean relationship.

#### *3.2.5.3 Nitrogen criteria*

The prediction for the ambient concentration of TN-DIN accounts for the increase in TN-DIN one would expect with increased Chl *a*, but also includes contributions from DON (as estimated by DOC) in the lake. Mean predictions for TN-DIN can be computed for different values of Chl *a* that include average contributions from other sources of N in the water column. The value of this *ambient* TN-DIN concentration that is associated with a targeted Chl *a* concentration then provides a criterion for TN-DIN (Figure 33, Table 5).

Criteria for N concentrations are commonly expressed in terms of TN rather than TN-DIN. To convert a criterion for TN-DIN to a criterion for TN, the availability of DIN for phytoplankton uptake can be considered. More specifically, the components of DIN ($NO_x$ and ammonia) are easily assimilated by phytoplankton and, when excess concentrations of DIN are observed in a lake, it may indicate that factors other than N availability are limiting phytoplankton growth. Therefore, controlling phytoplankton growth by reducing available N would first require that DIN concentrations are reduced to near zero and, when that occurs, criteria expressed for TN-DIN would be the same as those for TN. Furthermore, examination of NLA data indicated that DIN concentrations were below the detection limit in 72% of samples, so TN-DIN was equivalent to TN in most samples. Hence, in most cases TN monitoring data can be assessed relative to TN-DIN predictions from the criterion models.

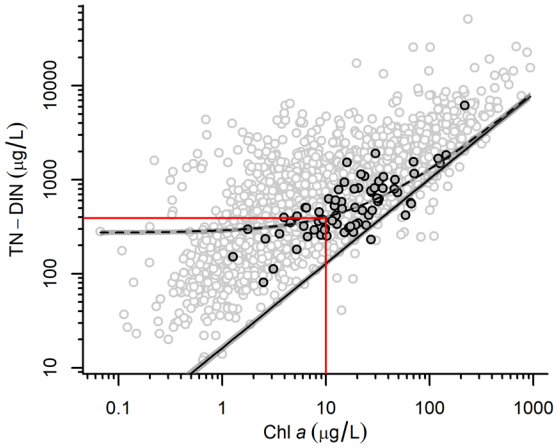


Figure 33. Illustrative example of deriving TN criteria for a Chl *a* target of 10 µg/L for one ecoregion (Southeastern Plains). *Open circles*: all data; *filled circles*: data from selected ecoregion; *solid line*: limiting TN-DIN vs. Chl *a* relationship from compartment model; *dashed line*: mean ambient TN-DIN vs. Chl *a* relationship taking into account mean DOC observed within the selected ecoregion: *shaded area*: 80% credible intervals; *horizontal and vertical solid line segments*: Illustrative Chl *a* target and associated TN criterion.

Table 5. Illustrative example of TN criteria corresponding to data shown in Figure 33.

| **Chl *a* target** | **80% certainty level** | **50% certainty level** |
|---|---|---|
| 10 µg/L | 380 | 390 |
| 15 µg/L | 440 | 450 |

The same interactive tool for computing different TP criteria also provides TN criteria associated with Chl *a* (https://nsteps.epa.gov/apps/tp-tn-chl). This tool allows the user to specify the targeted Chl *a* concentration, DOC concentration, and an ecoregion of interest. Finally, users can select the certainty level (or, credible interval) to account for the effects of model uncertainty on the calculated criteria. Data selected for an ecoregion are highlighted in the provided plots.

## 3.3 Duration and Frequency

The duration component of a water quality criterion is the length of time over which discrete water samples are averaged to assess the condition of the water body. The frequency component defines the number of times over a given time period that the specified magnitude of the criterion can be exceeded while the water body is still assessed as being in compliance with the criterion and maintaining designated uses. In conjunction with the magnitude of the criterion, these additional components define a water quality criterion.

Specification of duration and frequency components of numeric nutrient criteria is complicated by the fact that the ecological effects of elevated nutrient concentrations usually arise from a sequence of events. For example, higher nutrient concentrations increase the abundance of phytoplankton. Over time, higher abundances of phytoplankton then increase the amount of organic material in the deeper waters of a lake, and decomposition of the stored organic material can reduce the concentrations of DO. In this case, the duration and frequency of exceedance of a Chl *a* concentration in the mixed layer of the lake is related only indirectly to the ecological effect of decreased DO and the ultimate reductions in the amount of habitat for cool- and cold-water species. Contrast this example with the specification of duration and frequency of toxic pollutants, for which the length of time and frequency of exposure to the pollutant can be directly linked to effects on different organisms (e.g., mortality). A second consideration arises from the variability of environmental measurements, for which estimates of mean concentrations of Chl *a*, TN, and TP can only be estimated from a finite number of samples. So, when specifying duration and frequency components of the recommended numeric nutrient criteria, EPA considered both the timescale of the ecological responses and the statistical uncertainty in estimating mean values.

The recommended duration for Chl *a* criteria derived from the models described in this document is a growing season (typically summer) geometric mean value, consistent with the summary statistic used for Chl *a* in the stressor-response analyses. The geometric mean was selected to account for the fact that Chl *a* measurements are frequently log-normally distributed. EPA used seasonal mean Chl *a* concentrations integrated over the photic zone for analysis because timescales of ecological responses to increased nutrient concentrations are long. For example, as described earlier, some of the increase in deep water oxygen demand arises from accrual of organic material over long time periods while other oxygen demand arises from recently created organic matter that settles through the water column. Mean Chl *a* concentration in the lake is associated with mechanisms acting at both timescales, providing a measure of the average amount of organic material supplied by the photic zone. Similarly, systematic changes in zooplankton composition would be expected to occur at longer, seasonal timescales. For the microcystin model, the basic unit of analysis was an individual sample, in which the model predicted the probability of different microcystin concentrations in a sample, given the sample's Chl *a* concentration. When estimating the relationship for computing criteria, however, EPA computed probabilities of different individual Chl *a* concentrations as a function

of seasonal mean Chl *a* concentration, again linking seasonal mean Chl *a* concentration to the probability of deleterious effects. The long timescale of ecological response also supports the use of other durations that characterize average Chl *a* over long period of time. For example, in lakes with very long growing seasons, annual average Chl *a* may also be appropriate.

The unit of analysis for models relating Chl *a* to TN and TP concentrations was also the individual sample, in that TN and TP concentrations measured within a water sample were described as the sum of phytoplankton-bound N and P and other compartments in the sample containing those nutrients. Because the models are expressed as simple sums of components, each one remains applicable even if expressed in terms of seasonal averages. Hence, seasonal geometric mean Chl *a* criteria can be converted to seasonal geometric mean TN and TP criteria using the same model, and the recommended durations for TN and TP criteria are also seasonal mean values.

EPA recognizes that seasonal geometric mean concentrations of Chl *a*, TN, and TP calculated in different years can vary about their long-term means, and the frequency components of the criteria can be used to account for this variability. For example, in a year with particularly high precipitation one might observe higher than average loads of TP to downstream lakes. Similarly, in a year with longer than average periods of sunshine one might observe higher rates of accumulation of phytoplankton biomass and higher concentrations of Chl *a*. Sampling variability and within-year temporal variability also can cause large variations in estimated seasonal averages. Hence, in lakes in which long-term mean concentrations of Chl *a*, TN, and TP are below the criteria, some seasonal mean concentrations might still exceed the criterion magnitude. Variability in estimated seasonal mean concentrations can be addressed by estimating interannual and sampling variability and incorporating it into the expression of the criteria or in the assessment methods. More specifically, states may calculate adjusted criterion magnitudes that are associated with allowable frequencies of exceedance based on the observed variability of nutrient concentrations. For example, seasonal mean concentrations of TP would be expected to exceed a criterion magnitude that is equal to the long-term mean in approximately 50% of the years, whereas less frequent exceedances of a higher criterion magnitude would be expected. Appendices D and E provide examples of calculations that identify different combinations of criterion magnitudes and frequencies.

# 4 Characterization

In an ecological risk assessment, during the risk characterization phase, the results from the analysis are evaluated and the risk to ecological endpoints is described. In this section, general issues associated with applying the criterion models are considered, so a broader scope of topics is considered than is typically addressed during risk characterization. The topics include a discussion of other measures of effect and exposure that were not used in the criterion models, comparisons of current models with past models linking nutrients to Chl *a*, and limitations of the criterion models. Specific topics associated with deriving criteria in states and authorized tribes are also considered, including methods for incorporating local monitoring data into the national models, methods for developing statewide criteria, and the specification of duration and frequency components for the criteria.

## 4.1 Other Measures of Effect and Exposure

A variety of other measures of effect and exposure could be used for deriving nutrient criteria associated with each of the pathways described in Figure 1 and Figure 2. In selecting the responses for analysis, EPA considered (1) available data, (2) the current state of scientific understanding of each pathway, and (3) the degree to which a pathway and a response could be applied broadly to most lakes. For many possible measures of effect and exposure, data availability was a key consideration. For aquatic life, direct measurements of fish assemblage composition and biomass were not collected during the NLA, and the lack of those data limited the potential for considering several pathways such as evaluating alterations in fish assemblage composition because of reduced visibility. Lake benthic communities also exhibit changes along a eutrophication gradient (Vadeboncoeur et al. 2003), but none of those data were available. For recreational and drinking water source uses, the effects of other cyanotoxins (e.g., cylindrospermopsin, saxitoxin, anatoxin) might be important for certain lakes, but continental-scale data for those other cyanotoxins were not available at the time of this analysis. In certain lakes, cyanobacterial blooms have also been observed at depths below the surface layer (Jacquet et al. 2005), but observations of phytoplankton at those depths were not available. Similarly, organic matter generated by increased primary productivity can increase the concentrations of disinfection by-products during the drinking water treatment process (Graham et al. 1998, Galapate et al. 2001), and chemicals produced during blooms of certain

algal species can introduce unpleasant taste and odors to drinking water (Graham et al. 2010). However, continental-scale data pertaining to disinfection by-product precursors or taste and odor chemicals were not available.

Insufficient scientific understanding of a causal pathway also limited consideration of certain measures of effect and exposure. For example, scientific consensus is currently lacking on the precise level of cyanobacteria that is harmful to aquatic life. That information gap limited the utility of using cyanobacterial abundance as a final response measurement, despite the fact that increased cyanobacterial abundance occurs frequently with nutrient pollution (Dolman et al. 2012). (Note, however, that cyanobacterial abundance measurements quantify a key step in the model linking Chl *a* to microcystin.) Similarly, increased levels of cyanobacteria can cause rashes on people who contact the water (Pilotto et al. 1997, Zhang et al. 2015, US EPA 2015b), potentially affecting the use of a lake for recreation. However, precise quantitative relationships between the occurrence of rashes and cyanobacterial abundance are not currently available.

For certain measures of effect or exposure, data were available, but other factors limited the degree to which the response could be applied. For example, Secchi depth data were available in the NLA data set, and that measure of transparency could have informed an assessment of the aesthetic appeal of different lakes for recreation. That is, increased nutrient concentrations cause increases in the abundance of phytoplankton that reduce water clarity and decrease the aesthetic appeal of a lake (Carvalho et al. 2011, Keeler et al. 2015). Aesthetic considerations have been used by others as a basis of water quality criteria (Heiskary and Wilson 2008) and may be recommended for oligotrophic lakes in which maintaining historically high levels of water clarity is a management objective (see, for example, https://www.epa.gov/tmdl/lake-tahoe-total-maximum-daily-load-tmdl). However, the aesthetic expectations for the national population of lakes depends on geographic location (Smeltzer and Heiskary 1990), and user perception survey data at the continental scale of this analysis were not available. Similarly, reducing the frequency of phytoplankton blooms has been cited as a motivation for controlling nutrient loads (Bachmann et al. 2003), but aesthetic expectations regarding bloom frequency were not available at the national scale.

## 4.2 Incorporating State Data

State water quality managers are often interested in exploring relationships between environmental factors and biological responses using locally collected monitoring data. In many cases, leveraging knowledge from broader regional scales (e.g., national scale) can enhance local understanding. This document describes recommended numeric nutrient criterion models based on national data that link designated uses to Chl *a*, TN, and TP. The NLA data set provided a comprehensive set of measurements collected from large numbers of sites with identical protocols (US EPA 2011, Pollard et al. 2018), and the availability of consistent measures from lakes spanning broad gradients facilitated the calculation of accurate national estimates of relationships of interest. However, the number of samples is limited within the national data set that is available to estimate relationships within any single state, and uncertainty in estimating relationships specific to a single state is higher than that associated with the national models. In contrast, monitoring conducted by state agencies can yield more intensive temporal sampling over more sites, and hence, relationships estimated from those data can assist local management decisions within that state. Data collected at the state level, however, can be limited in the parameters that are measured, and the range of environmental conditions sampled is limited by conditions occurring within the state boundaries.

All the recommended criterion models described in this document are formulated to facilitate consideration of state data. State-specific values for certain coefficients in each model (e.g., Figure 34 in Appendix A) have been estimated, and local state, monitoring data can be used to refine the estimates of state-specific coefficients, while remaining consistent with national trends. Appendices A, B, and C discuss three examples of case studies in which state monitoring data have been combined with national data to refine recommended criteria. State monitoring data sets are each unique, and EPA is available to assist states in combining their monitoring data with the national models.

## 4.3 Existing Nutrient-Chlorophyll *a* Models

Empirically estimated relationships between TP and Chl *a* concentrations have provided a basis for lake water quality management for over four decades. This relationship was initially identified in Connecticut and Japanese lakes (Deevey 1940, Sakamoto 1966), and subsequently extended to a broad range of temperate lakes in the mid-1970s (Dillon and Rigler 1974, Jones

and Bachmann 1976, Carlson 1977). Those early analyses regressed Chl *a* on TP concentrations and reported similar coefficients showing the ratio of Chl:TP increased with lake trophic state. Over time, many studies have explored the veracity of that relationship and assessed sources of residual variation, testing the limits of applicability to different regions and lake types (Prairie et al. 1989, McCauley et al. 1989, Jones and Knowlton 2005, Filstrup et al. 2014b). Variations in the relationship have been attributed to differences in lake depth (Pridmore et al. 1985), TN:TP ratio (Smith 1982, Prairie et al. 1989, Molot and Dillon 1991), grazing by zooplankton and mussels (Mazumder 1994, Mellina et al. 1995), landscape characteristics (Wagner et al. 2011), and light limitation (Hoyer and Jones 1983, Knowlton and Jones 2000, Havens and Nürnberg 2004). Regional studies have evaluated the relationship as influenced by edaphic and climatic factors in locations such as Canada (Prepas and Trew 1983), Argentina, (Quirós 1990), the United Kingdom (Spears et al. 2013), and Europe (Phillips et al. 2008). Recently, lake classifications have improved the precision and accuracy of this relationship (Yuan and Pollard 2014).

As described in Sections 3.2.4 and 3.2.5, EPA reformulated the nutrient-chlorophyll models to account for variations in TP and TN, rather than in Chl *a*. The new models better account for variability in measurements of TP and TN and are consistent with an understanding of the components of TP and TN in the water column. The reformulated models cannot be directly compared with earlier studies, including those cited previously. Estimates of N and P content of phytoplankton, however, are consistent with values reported elsewhere (Yuan and Jones 2019).

## 4.4 Limitations and Assumptions

The recommended models for deriving numeric nutrient criteria are limited by the nature of the data that underlie the analysis. First, nutrient data for each lake consisted of samples collected at a single point, resulting in no information on within-lake spatial variability in nutrient concentrations being included in the analyses. Nutrient concentrations within particular lakes can vary considerably across different locations (Perkins and Underwood 2000), resulting in criteria based on samples collected at the deepest point or midpoint of the reservoir that might not be applicable to samples collected elsewhere. When deriving their criteria, states may specify assessment methodologies to collect samples from different locations in the same lake to address this issue and analyze those local data to account for spatial variability.

Similarly, nutrient and response data used in the current analysis were collected only in the summer, so monitoring data assessed with respect to these recommended criteria may also be limited to summer data. Nutrient concentrations in some lakes can vary considerably between summer and winter (Søndergaard et al. 2005), and states may specify assessment protocols to ensure that only data collected in the summer are compared with criterion concentrations. In other locations, much longer growing seasons exist, and annual average nutrient and Chl *a* concentrations are comparable to summer averages.

As noted earlier, most of the statistical criterion models described here combine the effects of spatial, temporal, and sampling variability and estimate a single value for each model that is applicable to all lakes in the data set. The components of variability, however, might differ across lakes and affect the resulting criteria. For example, spatial variability in complex, dendritic reservoirs can be much greater than in simple, circular lakes (Gloss et al. 1980). In most cases, local monitoring data can inform and potentially improve the parameter estimates both for specific locations and for groups of lakes. Segmenting lakes with complex morphology into discrete, more homogeneous areas may also help address issues with spatial variability.

The uncertainty estimated for each modeled relationship is associated with the number of samples used in the model, and consideration of sample size can affect the interpretation of the resulting criteria. For example, the number of NLA samples within a single Level III ecoregion can be small. The hierarchical structure of the model does improve the precision of model estimates in those ecoregions, but the precision of TP and TN criteria specific to ecoregions with small amounts of data could be further improved by including state monitoring data. Additional national-scale data such as that from the 2017 NLA may also be incorporated as they become available to improve model precision.

Recommended criteria based on the drinking water health advisories for microcystin incorporate some conservative assumptions that affect the final values. The recommended criteria are intended to reflect the ambient water quality conditions that protect a drinking water use before treatment. They do not, however, account for the varying levels of treatment a drinking water facility can implement to remove microcystin before generating finished drinking water, the condition of the water to which the cyanotoxin health advisories apply. As a precautionary step, a drinking water facility may implement treatment protocols that minimize the breakage of cyanobacteria cells (Chow et al. 1999, Westrick et al. 2010) which, in turn,

would minimize the release of intracellular microcystin into the treated water. EPA based the recommended models on the total microcystin present in the NLA samples, both dissolved in the water and within cyanobacterial cells, which necessitated the lysis of cyanobacterial cells prior to microcystin quantification. In contrast, breakage of cyanobacterial cells is a process that some drinking water treatments for cyanotoxins are designed to limit. Criteria based on the national models provide protective water quality conditions in the source water, but concentrations of microcystin that slightly exceed health advisory values can be further reduced in the finished drinking water through carefully engineered and operated source water treatment processes.

Recommended criteria derived using the models described here provide concentrations that, when exceeded, are associated with a loss of support for designated uses, but the models do not provide information regarding appropriate remediation actions. Indeed, among lakes in which the criteria are exceeded, appropriate remediation actions will likely differ. In some lakes, the magnitude of N loading from anthropogenic sources is small, while P loading is large, and cyanobacteria supply N to the system via fixation (Schindler et al. 2008). In those lakes, reductions in P loading might be the appropriate water quality management action. In other lakes, ample supplies of N from anthropogenic sources are available, and management actions might need to focus on reducing both N and P loading (Ferber et al. 2004). In some lakes, excess N in the form of inorganic nitrogen ($NO_x$ or ammonia) is abundant, and the presence of high concentrations of DIN might provide insights into the effects of different management interventions. For example, DIN is readily taken up by phytoplankton, so the presence of large concentrations of DIN might indicate that other factors, such as light availability, limit phytoplankton growth. In those cases, initial reductions of N loading to reduce $NO_x$ might be necessary before the effects of N control can be observed.

## 4.5 Deriving State-Specified Criteria

Criteria derived from the recommended national models vary with differences in lake characteristics (e.g., depth and ecoregion), and specifying a single set of criteria applicable to all lakes in a state might not account for those variations. Methods for deriving criteria that account for natural variations among water bodies are already available, and these methods can be applied to ensure that appropriate criteria are applied to different types of lakes. First, states can classify water bodies and derive different criteria for each class of water body. The

recommended national models facilitate the classification of lakes by providing specific insights into the factors that most affect the derivation of protective numeric nutrient criteria. Furthermore, the national models can be used to compute criteria for different lakes in a state to provide information about the types of lakes for which criterion magnitudes are most similar. For example, different recommended criteria for TP and TN are associated with different Level III ecoregions; however, among the ecoregions within one state, the difference in criterion magnitudes might be small enough to specify a single set of criteria applicable to multiple ecoregions. Second, site-specific criteria can be specified for a small number of lakes with characteristics that differ substantially from the rest of the lakes in a state. Here, too, the recommended national models provide the means of deriving these criteria for individual lakes.

# 5 References


Arend, K. K., D. Beletsky, J. V. DePinto, S. A. Ludsin, J. J. Roberts, D. K. Rucinski, D. Scavia, D. J. Schwab, and T. O. Höök. 2011. Seasonal and interannual effects of hypoxia on fish habitat quality in central Lake Erie. Freshwater Biology 56:366–383.

Bachmann, R. W., M. V. Hoyer, and D. E. C. Jr. 2003. Predicting the frequencies of high chlorophyll levels in Florida lakes from average chlorophyll or nutrient data. Lake and Reservoir Management 19:229–241.

Baird, O. E., and C. C. Krueger. 2003. Behavioral thermoregulation of brook and rainbow trout: Comparison of summer habitat use in an Adirondack River, New York. Transactions of the American Fisheries Society 132:1194–1206.

Barnett, V., and A. O'Hagan. 1997. Setting Environmental Standards: The Statistical Approach to Handling Uncertainty and Variation. Chapman and Hall/CRC, London, UK.

Bednarska, A., and P. Dawidowicz. 2007. Change in filter-screen morphology and depth selection: Uncoupled responses of Daphnia to the presence of filamentous cyanobacteria. Limnology and Oceanography 52:2358–2363.

Benndorf, Jü., W. Böing, J. Koop, and I. Neubauer. 2002. Top-down control of phytoplankton: the role of time scale, lake depth and trophic state. Freshwater Biology 47:2282–2295.

Berman, T., and D. A. Bronk. 2003. Dissolved organic nitrogen: a dynamic participant in aquatic ecosystems. Aquatic Microbial Ecology 31:279–305.

Burns, N. M. 1995. Using hypolimnetic dissolved oxygen depletion rates for monitoring lakes. New Zealand Journal of Marine and Freshwater Research 29:1–11.

Cahill, K. L., J. M. Gunn, and M. N. Futter. 2005. Modelling ice cover, timing of spring stratification, and end-of-season mixing depth in small Precambrian Shield lakes. Canadian Journal of Fisheries and Aquatic Sciences 62:2134–2142.

Carey, C. C., B. W. Ibelings, E. P. Hoffmann, D. P. Hamilton, and J. D. Brookes. 2012. Eco-physiological adaptations that favour freshwater cyanobacteria in a changing climate. Water Research 46:1394–1407.

Carlson, R. E. 1977. A trophic state index for lakes. Limnology and Oceanography 22:361–369.

Carvalho, L., C. A. Miller (nee Ferguson), E. M. Scott, G. A. Codd, P. S. Davies, and A. N. Tyler. 2011. Cyanobacterial blooms: Statistical models describing risk factors for national-scale lake assessment and lake management. Science of The Total Environment 409:5353–5358.

Cheung, M. Y., S. Liang, and J. Lee. 2013. Toxin-producing cyanobacteria in freshwater: A review of the problems, impact on drinking water safety, and efforts for protecting public health. Journal of Microbiology 51:1–10.

Chorus, I. 2001. Cyanotoxins: occurrence, causes, consequences. Springer, Berlin, Germany.

Chow, C. W. K., M. Drikas, J. House, M. D. Burch, and R. M. A. Velzeboer. 1999. The impact of conventional water treatment processes on cells of the cyanobacterium Microcystis aeruginosa. Water Research 33:3253–3262.

Coker, G. A., C. B. Portt, and C. K. Minns. 2001. Morphological and ecological characteristics of Canadian freshwater fishes. Fisheries and Oceans Canada Burlington, Ontario.

Colby, P. J., G. R. Spangler, D. A. Hurley, and A. M. McCombie. 1972. Effects of eutrophication on salmonid communities in oligotrophic lakes. Journal of the Fisheries Research Board of Canada 29:975–983.

Cole, J. J., S. R. Carpenter, J. Kitchell, M. L. Pace, C. T. Solomon, and B. Weidel. 2011. Strong evidence for terrestrial support of zooplankton in small lakes based on stable isotopes of carbon, nitrogen, and hydrogen. Proceedings of the National Academy of Sciences 108:1975–1980.

Cornett, R. J. 1989. Predicting changes in hypolimnetic oxygen concentrations with phosphorus retention, temperature, and morphometry. Limnology and Oceanography 34:1359–1366.

Cornett, R. J., and F. H. Rigler. 1980. The areal hypolimnetic oxygen deficit: An empirical test of the model: Areal hypolimnetic oxygen deficit. Limnology and Oceanography 25:672–679.

Cornett, R. J., and F. H. Rigler. 1984. Dependence of hypolimnetic oxygen consumption on ambient oxygen concentration: Fact or artifact? Water Resources Research 20:823–830.

Coutant, C. C. 1985. Striped bass, temperature, and dissolved oxygen: A speculative hypothesis for environmental risk. Transactions of the American Fisheries Society 114:31–61.

Coutant, C. C., and D. S. Carroll. 1980. Temperatures occupied by ten ultrasonic-tagged striped bass in freshwater lakes. Transactions of the American Fisheries Society 109:195–202.

Daly, C., M. Halbleib, J. I. Smith, W. P. Gibson, M. K. Doggett, G. H. Taylor, J. Curtis, and P. P. Pasteris. 2008. Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. International Journal of Climatology 28:2031–2064.

De Robertis, A., C. H. Ryer, A. Veloza, and R. D. Brodeur. 2003. Differential effects of turbidity on prey consumption of piscivorous and planktivorous fish. Canadian Journal of Fisheries and Aquatic Sciences 60:1517–1526.

Deevey, E. S. 1940. Limnological studies in Connecticut; Part V, A contribution of regional limnology. American Journal of Science 238:717–741.

Demers, E., and J. Kalff. 1993. A simple model for predicting the date of spring stratification in temperate and subtropical lakes. Limnology and Oceanography 38:1077–1081.

Demott, W., and D. Müller-Navarra. 1997. The importance of highly unsaturated fatty acids in zooplankton nutrition: evidence from experiments with Daphnia, a cyanobacterium and lipid emulsions. Freshwater Biology 38:649–664.

Diaz, R. J. 2001. Overview of hypoxia around the world. Journal of Environmental Quality 30:275–281.

Dillon, P. J., and F. H. Rigler. 1974. The phosphorus-chlorophyll relationship in lakes. Limnology and Oceanography 19:767–773.

Dokulil, M. T., and K. Teubner. 2000. Cyanobacterial dominance in lakes. Hydrobiologia 438:1–12.

Dolman, A. M., J. Rücker, F. R. Pick, J. Fastner, T. Rohrlack, U. Mischke, and C. Wiedner. 2012. Cyanobacteria and cyanotoxins: The influence of nitrogen versus phosphorus. PLoS ONE 7:e38757.

Downing, J. A., S. B. Watson, and E. McCauley. 2001. Predicting cyanobacteria dominance in lakes. Canadian Journal of Fisheries and Aquatic Sciences 58:1905–1908.

Duane, S., A. D. Kennedy, B. J. Pendleton, and D. Roweth. 1987. Hybrid monte carlo. Physics Letters B 195:216–222.

Dumont, H. J., I. V. de Velde, and S. Dumont. 1975. The dry weight estimate of biomass in a selection of Cladocera, Copepoda and Rotifera from the plankton, periphyton and benthos of continental waters. Oecologia 19:75–97.

Eaton, J. G., and R. M. Scheller. 1996. Effects of climate warming on fish thermal habitat in streams of the United States. Limnology and Oceanography 41:1109–1115.

Effler, S. W., and S. M. O'Donnell. 2010. A long-term record of epilimnetic phosphorus patterns in recovering Onondaga Lake, New York. Fundamental and Applied Limnology / Archiv für Hydrobiologie 177:1–18.

Elser, J. J. 1999. The pathway to noxious cyanobacteria blooms in lakes: the food web as the final turn. Freshwater Biology 42:537–543.

Elton, C. S. 1927. Animal Ecology. University of Chicago Press.

Evans, D. O., K. H. Nicholls, Y. C. Allen, and M. J. McMurtry. 1996. Historical land use, phosphorus loading, and loss of fish habitat in Lake Simcoe, Canada. Canadian Journal of Fisheries and Aquatic Sciences 53:194–218.

Ferber, L. R., S. N. Levine, A. Lini, and G. P. Livingston. 2004. Do cyanobacteria dominate in eutrophic lakes because they fix atmospheric nitrogen? Freshwater Biology 49:690–708.

Ferguson, R. G. 1958. The preferred temperature of fish and their midsummer distribution in temperate lakes and streams. Journal of the Fisheries Research Board of Canada 15:607–624.

Filstrup, C. T., H. Hillebrand, A. J. Heathcote, W. S. Harpole, and J. A. Downing. 2014a. Cyanobacteria dominance influences resource use efficiency and community turnover in phytoplankton and zooplankton communities. Ecology Letters 17:464–474.

Filstrup, C. T., T. Wagner, P. A. Soranno, E. H. Stanley, C. A. Stow, K. E. Webster, and J. A. Downing. 2014b. Regional variability among nonlinear chlorophyll—phosphorus relationships in lakes. Limnology and Oceanography 59:1691–1703.

Fischer, W. J., I. Garthwaite, C. O. Miles, K. M. Ross, J. B. Aggen, A. R. Chamberlin, N. R. Towers, and D. R. Dietrich. 2001. Congener-independent immunoassay for microcystins and nodularins. Environmental Science & Technology 35:4849–4856.

Galapate, R. P., A. U. Baes, and M. Okada. 2001. Transformation of dissolved organic matter during ozonation: Effects on trihalomethane formation potential. Water Research 35:2201–2206.

Gelman, A. 2006. Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). Bayesian Analysis 1:515–534.

Gelman, A., and J. Hill. 2007. Data analysis using regression and multilevel/hierarchical models. Cambridge University Press, New York, NY.

del Giorgio, P. A., and J. M. Gasol. 1995. Biomass distribution in freshwater plankton communities. The American Naturalist 146:135–152.

Gloss, S. P., L. M. Mayer, and D. E. Kidd. 1980. Advective control of nutrient dynamics in the epilimnion of a large reservoir. Limnology and Oceanography 25:219–228.

Godfrey, P. J. 1982. The eutrophication of Cayuga Lake: a historical analysis of the phytoplankton's response to phosphate detergents. Freshwater Biology 12:149–166.

Gorham, E., and F. M. Boyce. 1989. Influence of lake surface area and depth upon thermal stratification and the depth of the summer thermocline. Journal of Great Lakes Research 15:233–245.

Graham, J. L., K. A. Loftin, M. T. Meyer, and A. C. Ziegler. 2010. Cyanotoxin mixtures and taste-and-odor compounds in cyanobacterial blooms from the midwestern United States. Environ. Sci. Technol. 44:7361–7368.

Graham, N. J. D., V. E. Wardlaw, R. Perry, and J.-Q. Jiang. 1998. The significance of algae as trihalomethane precursors. Water Science and Technology 37:83–89.

Hamilton, D. P., and S. F. Mitchell. 1996. An empirical model for sediment resuspension in shallow lakes. Hydrobiologia 317:209–220.

Haney, J. F. 1987. Field studies on zooplankton-cyanobacteria interactions. New Zealand Journal of Marine and Freshwater Research 21:467–475.

Hanson, P. C., D. L. Bade, S. R. Carpenter, and T. K. Kratz. 2003. Lake metabolism: Relationships with dissolved organic carbon and phosphorus. Limnology and Oceanography 48:1112–1119.

Havens, K. E., and G. K. Nürnberg. 2004. The phosphorus-chlorophyll relationship in lakes: Potential influences of color and mixing regime. Lake and Reservoir Management 20:188–196.

Heathcote, A. J., C. T. Filstrup, D. Kendall, and J. A. Downing. 2016. Biomass pyramids in lake plankton: influence of Cyanobacteria size and abundance. Inland Waters 6:250–257.

Heiskary, S., and B. Wilson. 2008. Minnesota's approach to lake nutrient criteria development. Lake and Reservoir Management 24:282–297.

Hessen, D. O. 2008. Efficiency, energy and stoichiometry in pelagic food webs; Reciprocal roles of food quality and food quantity. Freshwater Reviews 1:43–57.

Hessen, D. O., B. A. Faafeng, P. Brettum, and T. Andersen. 2006. Nutrient enrichment and planktonic biomass ratios in lakes. Ecosystems 9:516–527.

Hoeger, S. J., B. C. Hitzfeld, and D. R. Dietrich. 2005. Occurrence and elimination of cyanobacterial toxins in drinking water treatment plants. Toxicology and Applied Pharmacology 203:231–242.

Holmes, R., R. Norris, T. Smayda, and E. J. F. Wood. 1969. Collection, fixation, identification, and enumeration of phytoplankton standing stock. Recommended procedures for measuring the productivity of plankton standing stock and related oceanic properties. Washington (DC): National Academy of Sciences:17–46.

Hondzo, M., and H. G. Stefan. 1993. Lake water temperature simulation model. Journal of Hydraulic Engineering 119:1251–1273.

Hoyer, M. V., and J. R. Jones. 1983. Factors affecting the relation between phosphorus and chlorophyll a in Midwestern reservoirs. Canadian Journal of Fisheries and Aquatic Sciences 40:192–199.

Hutchinson, G. E. 1938. On the relation between the oxygen deficit and the productivity and typology of lakes. Internationale Revue der gesamten Hydrobiologie und Hydrographie 36:336–355.

Jacobson, P. C., T. S. Jones, P. Rivers, and D. L. Pereira. 2008. Field Estimation of a Lethal Oxythermal Niche Boundary for Adult Ciscoes in Minnesota Lakes. Transactions of the American Fisheries Society 137:1464–1474.

Jacobson, P. C., H. G. Stefan, and D. L. Pereira. 2010. Coldwater fish oxythermal habitat in Minnesota lakes: influence of total phosphorus, July air temperature, and relative depth. Canadian Journal of Fisheries and Aquatic Sciences 67:2002–2013.

Jacquet, S., J.-F. Briand, C. Leboulanger, C. Avois-Jacquet, L. Oberhaus, B. Tassin, B. Vinçon-Leite, G. Paolini, J.-C. Druart, O. Anneville, and J.-F. Humbert. 2005. The proliferation of the toxic cyanobacterium Planktothrix rubescens following restoration of the largest natural French lake (Lac du Bourget). Harmful Algae 4:651–672.

Jenny, J.-P., P. Francus, A. Normandeau, F. Lapointe, M.-E. Perga, A. Ojala, A. Schimmelmann, and B. Zolitschka. 2016. Global spread of hypoxia in freshwater ecosystems during the last three centuries is caused by rising local human pressure. Global Change Biology 22:1481–1489.

Jeppesen, E., J. P. Jensen, C. Jensen, B. Faafeng, D. O. Hessen, M. Søndergaard, T. Lauridsen, P. Brettum, and K. Christoffersen. 2003. The impact of nutrient state and lake depth on top-down control in the pelagic zone of lakes: A study of 466 lakes from the temperate zone to the arctic. Ecosystems 6:313–325.

Jones, F. E., and G. L. Harris. 1992. ITS-90 density of water formulation for volumetric standards calibration. Journal of research of the National Institute of Standards and Technology 97:335–340.

Jones, J. R., and R. W. Bachmann. 1976. Prediction of phosphorus and chlorophyll Levels in lakes. Journal (Water Pollution Control Federation) 48:2176–2182.

Jones, J. R., and M. F. Knowlton. 2005. Chlorophyll response to nutrients and non-algal seston in Missouri reservoirs and oxbow lakes. Lake and Reservoir Management 21:361–371.

Jones, J. R., M. F. Knowlton, D. V. Obrecht, and J. L. Graham. 2011. Temperature and oxygen in Missouri reservoirs. Lake and Reservoir Management 27:173–182.

Jones, J. R., D. V. Obrecht, B. D. Perkins, M. F. Knowlton, A. P. Thorpe, S. Watanabe, and R. R. Bacon. 2008. Nutrients, seston, and transparency of Missouri reservoirs and oxbow lakes: An analysis of regional limnology. Lake and Reservoir Management 24:155–180.

Kasprzak, P., J. Padisák, R. Koschel, L. Krienitz, and F. Gervais. 2008. Chlorophyll a concentration across a trophic gradient of lakes: An estimator of phytoplankton biomass? Limnologica - Ecology and Management of Inland Waters 38:327–338.

Keeler, B. L., S. A. Wood, S. Polasky, C. Kling, C. T. Filstrup, and J. A. Downing. 2015. Recreational demand for clean water: evidence from geotagged photographs by visitors to lakes. Frontiers in Ecology and the Environment 13:76–81.

Klumb, R. A., K. L. Bunch, E. L. Mills, L. G. Rudstam, G. Brown, C. Knauf, R. Burton, and F. Arrhenius. 2004. Establishment of a metalimnetic oxygen refuge for zooplankton in a productive lake ontario embayment. Ecological Applications 14:113–131.

Knowlton, M. F., M. V. Hoyer, and J. R. Jones. 1984. Sources of variability in phosphorus and chlorophyll and their effects on use of lake survey data. Journal of the American Water Resources Association 20:397–408.

Knowlton, M. F., and J. R. Jones. 2000. Non-algal seston, light, nutrients and chlorophyll in Missouri reservoirs. Lake and Reservoir Management 16:322–332.

Kritzberg, E. S., J. J. Cole, M. L. Pace, W. Granéli, and D. L. Bade. 2004. Autochthonous versus allochthonous carbon sources of bacteria: Results from whole-lake 13C addition experiments. Limnology and Oceanography 49:588–596.

Lawrence, S. G., D. F. Malley, W. J. Findlay, M. A. Maclver, and I. L. Delbaere. 1987. Method for estimating dry weight of freshwater planktonic crustaceans from measures of length and shape. Canadian Journal of Fisheries and Aquatic Sciences 44:s264–s274.

Lee, W. C., and E. P. Bergersen. 1996. Influence of thermal and oxygen stratification on lake trout hooking mortality. North American Journal of Fisheries Management 16:175–181.

Leibold, M. A., J. M. Chase, J. B. Shurin, and A. L. Downing. 1997. Species turnover and the regulation of trophic structure. Annual Review of Ecology and Systematics 28:467–494.

Lewis, W. M. 1983. A revised classification of lakes based on mixing. Canadian Journal of Fisheries and Aquatic Sciences 40:1779–1787.

Lewis, W. M., and W. A. Wurtsbaugh. 2008. Control of lacustrine phytoplankton by nutrients: Erosion of the phosphorus paradigm. International Review of Hydrobiology 93:446–465.

Lienesch, P. W., M. E. McDonald, A. E. Hershey, W. J. O'Brien, and N. D. Bettez. 2005. Effects of a whole-lake, experimental fertilization on lake trout in a small oligotrophic arctic lake. Hydrobiologia 548:51–66.

Loftin, K. A., M. Meyer, F. Rubio, L. Kamp, E. Humphries, and E. Whereat. 2008. Comparison of two cell lysis procedures for recovery of microcystins in water samples from Silver Lake in Dover, Delaware, with microcystin producing cyanobacterial accumulations. Open-File Report, United States Geological Survey, Reston, VA.

Mackenzie-Grieve, J. L., and J. R. Post. 2006. Thermal habitat use by lake trout in two contrasting Yukon Territory lakes. Transactions of the American Fisheries Society 135:727–738.

Marcus, M. D., W. A. Hubert, and S. H. Anderson. 1984. Habitat Suitability Index Models: Lake Trout (Exclusive of the Great Lakes). U.S. Fish and Wildlife Service.

Mazumder, A. 1994. Phosphorus–chlorophyll relationships under contrasting herbivory and thermal stratification: predictions and patterns. Canadian Journal of Fisheries and Aquatic Sciences 51:390–400.

McCauley, E. 1984. The estimation of the abundance and biomass of zooplankton in samples. A manual on methods for the assessment of secondary productivity in fresh waters:228–265.

McCauley, E., J. A. Downing, and S. Watson. 1989. Sigmoid relationships between nutrients and chlorophyll among lakes. Canadian Journal of Fisheries and Aquatic Sciences 46:1171–1175.

McMahon, T. E., J. W. Terrell, and P. C. Nelson. 1984. Habitat Suitability Information: Walleye. Page 43. Division of Biological Services, Fish and Wildlife Service, U.S. Department of the Interior, Washington, DC.

Mellina, E., J. B. Rasmussen, and E. L. Mills. 1995. Impact of zebra mussel (Dreissena polymorpha) on phosphorus cycling and chlorophyll in lakes. Canadian Journal of Fisheries and Aquatic Sciences 52:2553–2573.

Molot, L. A., and P. J. Dillon. 1991. Nitrogen/phosphorus ratios and the prediction of chlorophyll in phosphorus-limited lakes in central Ontario. Canadian Journal of Fisheries and Aquatic Sciences 48:140–145.

Molot, L. A., P. J. Dillon, B. J. Clark, and B. P. Neary. 1992. Predicting end-of-summer oxygen profiles in stratified lakes. Canadian Journal of Fisheries and Aquatic Sciences 49:2363–2372.

Müller, B., L. D. Bryant, A. Matzinger, and A. Wüest. 2012. Hypolimnetic oxygen depletion in eutrophic lakes. Environmental Science & Technology 46:9964–9971.

Müller-Navarra, D. C., M. T. Brett, A. M. Liston, and C. R. Goldman. 2000. A highly unsaturated fatty acid predicts carbon transfer between primary producers and consumers. Nature 403:74–77.

Norton, S. B., D. J. Rodier, W. H. van der Schalie, W. P. Wood, M. W. Slimak, and J. H. Gentile. 1992. A framework for ecological risk assessment at the EPA. Environmental Toxicology and Chemistry 11:1663–1672.

Olson, J. R., and C. P. Hawkins. 2013. Developing site-specific nutrient criteria from empirical models. Freshwater Science 32:719–740.

Pace, M. L., J. J. Cole, S. R. Carpenter, J. F. Kitchell, J. R. Hodgson, M. C. Van de Bogert, D. L. Bade, E. S. Kritzberg, and D. Bastviken. 2004. Whole-lake carbon-13 additions reveal terrestrial support of aquatic food webs. Nature 427:240–243.

Paerl, H. W., and T. G. Otten. 2013. Harmful cyanobacterial blooms: Causes, consequences, and controls. Microbial Ecology 65:995–1010.

Paerl, H. W., J. T. Scott, M. J. McCarthy, S. E. Newell, W. S. Gardner, K. E. Havens, D. K. Hoffman, S. W. Wilhelm, and W. A. Wurtsbaugh. 2016. It takes two to tango: When and where dual nutrient (N & P) reductions are needed to protect lakes and downstream ecosystems. Environmental Science & Technology 50:10805–10813.

Paerl, H. W., and J. F. Ustach. 1982. Blue-green algal scums: An explanation for their occurrence during freshwater blooms. Limnology and Oceanography 27:212–217.

Park, S., M. T. Brett, E. T. Oshel, and C. R. Goldman. 2003. Seston food quality and Daphnia production efficiencies in an oligo-mesotrophic Subalpine Lake. Aquatic Ecology 37:123–136.

Perkins, R. G., and G. J. C. Underwood. 2000. Gradients of chlorophyll a and water chemistry along an eutrophic reservoir with determination of the limiting nutrient by in situ nutrient addition. Water Research 34:713–724.

Persson, J., M. T. Brett, T. Vrede, and J. L. Ravet. 2007. Food quantity and quality regulation of trophic transfer between primary producers and a keystone grazer (Daphnia) in pelagic freshwater food webs. Oikos 116:1152–1163.

Phillips, G., O.-P. Pietiläinen, L. Carvalho, A. Solimini, A. Lyche Solheim, and A. Cardoso. 2008. Chlorophyll–nutrient relationships of different lake types using a large European dataset. Aquatic Ecology 42:213–226.

Phillips, G., N. Willby, and B. Moss. 2016. Submerged macrophyte decline in shallow lakes: What have we learnt in the last forty years? Aquatic Botany.

Pilotto, L. S., R. M. Douglas, M. D. Burch, S. Cameron, M. Beers, G. J. Rouch, P. Robinson, M. Kirk, C. T. Cowie, S. Hardiman, C. Moore, and R. G. Attewell. 1997. Health effects of exposure to cyanobacteria (blue–green algae) during recreational water–related activities. Australian and New Zealand Journal of Public Health 21:562–566.

Plumb, J. M., and P. J. Blanchfield. 2009. Performance of temperature and dissolved oxygen criteria to predict habitat use by lake trout (Salvelinus namaycush). Canadian Journal of Fisheries and Aquatic Sciences 66:2011–2023.

Pollard, A. I., S. E. Hampton, and D. M. Leech. 2018. The promise and potential of continental-scale limnology using the U.S. Environmental Protection Agency's National Lakes Assessment. Limnology and Oceanography Bulletin 27:36–41.

Prairie, Y. T., C. M. Duarte, and J. Kalff. 1989. Unifying nutrient—chlorophyll relationships in lakes. Canadian Journal of Fisheries and Aquatic Sciences 46:1176–1182.

Prepas, E. E., and D. O. Trew. 1983. Evaluation of the phosphorus–chlorophyll relationship for lakes of the Precambrian Shield in western Canada. Canadian Journal of Fisheries and Aquatic Sciences 40:27–35.

Pridmore, R. D., W. N. Vant, and J. C. Rutherford. 1985. Chlorophyll-nutrient relationships in North Island lakes (New Zealand). Hydrobiologia 121:181–189.

Qian, S. S., and R. J. Miltner. 2015. A continuous variable Bayesian networks model for water quality modeling: A case study of setting nitrogen criterion for small rivers and streams in Ohio, USA. Environmental Modelling & Software 69:14–22.

Quinlan, R., A. M. Paterson, J. P. Smol, M. S. V. Douglas, and B. J. Clark. 2005. Comparing different methods of calculating volume-weighted hypolimnetic oxygen (VWHO) in lakes. Aquatic Sciences 67:97–103.

Quirós, R. 1990. Factors related to variance of residuals in chlorophyll — total phosphorus regressions in lakes and reservoirs of Argentina. Hydrobiologia 200–201:343–355.

R Core Team. 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Reynolds, C. S., and S. C. Maberly. 2002. A simple method for approximating the supportive capacities and metabolic constraints in lakes and reservoirs. Freshwater Biology 47:1183–1188.

Rognerud, S., and G. Kjellberg. 1984. Relationships between phytoplankton and zooplankton biomass in large lakes. SIL Proceedings, 1922-2010 22:666–671.

Sakamoto, M. 1966. Primary production by phytoplankton community in some Japanese lakes and its dependence on lake depth. Archiv fur Hydrobiologie 62:1–28.

Sartory, D. P., and J. U. Grobbelaar. 1984. Extraction of chlorophyll a from freshwater phytoplankton for spectrophotometric analysis. Hydrobiologia 114:177–187.

Scavia, D., J. David Allan, K. K. Arend, S. Bartell, D. Beletsky, N. S. Bosch, S. B. Brandt, R. D. Briland, I. Daloğlu, J. V. DePinto, D. M. Dolan, M. A. Evans, T. M. Farmer, D. Goto, H. Han, T. O. Höök, R. Knight, S. A. Ludsin, D. Mason, A. M. Michalak, R. Peter Richards, J. J. Roberts, D. K. Rucinski, E. Rutherford, D. J. Schwab, T. M. Sesterhenn, H. Zhang, and Y. Zhou. 2014. Assessing and addressing the re-eutrophication of Lake Erie: Central basin hypoxia. Journal of Great Lakes Research 40:226–246.

Scheffer, M., and E. H. van Nes. 2007. Shallow lakes theory revisited: various alternative regimes driven by climate, nutrients, depth and lake size. Hydrobiologia 584:455–466.

Scheffer, M., S. Rinaldi, A. Gragnani, L. R. Mur, and E. H. van Nes. 1997. On the dominance of filamentous cyanobacteria in shallow, turbid lakes. Ecology 78:272–282.

Schindler, D. W., R. E. Hecky, D. L. Findlay, M. P. Stainton, B. R. Parker, M. J. Paterson, K. G. Beaty, M. Lyng, and S. E. M. Kasian. 2008. Eutrophication of lakes cannot be controlled by reducing nitrogen input: Results of a 37-year whole-ecosystem experiment. Proceedings of the National Academy of Sciences 105:11254–11258.

Schindler, D. W., M. A. Turner, and R. H. Hesslein. 1985. Acidification and alkalinization of lakes by experimental addition of nitrogen compounds. Biogeochemistry 1:117–133.

Smeltzer, E., and S. A. Heiskary. 1990. Analysis and applications of lake user survey data. Lake and Reservoir Management 6:109–118.

Smith, V. H. 1982. The nitrogen and phosphorus dependence of algal biomass in lakes: an empirical and theoretical analysis. Limnology and Oceanography 27:1101–1112.

Smith, V. H., S. B. Joye, and R. W. Howarth. 2006. Eutrophication of freshwater and marine ecosystems. Limnology and Oceanography 51:351–355.

Smith, V. H., and D. W. Schindler. 2009. Eutrophication science: where do we go from here? Trends in Ecology & Evolution 24:201–207.

Snucins, E. J., and J. M. Gunn. 1995. Coping with a warm environment: Behavioral thermoregulation by lake trout. Transactions of the American Fisheries Society 124:118–123.

Søndergaard, M., J. P. Jensen, and E. Jeppesen. 2005. Seasonal response of nutrients to reduced phosphorus loading in 12 Danish lakes. Freshwater Biology 50:1605–1615.

Spears, B. M., L. Carvalho, B. Dudley, and L. May. 2013. Variation in chlorophyll a to total phosphorus ratio across 94 UK and Irish lakes: Implications for lake management. Journal of Environmental Management 115:287–294.

Stan Development Team. 2016. Stan Modeling Language Users Guide and Reference Manual, Version 2.14.0.

Stefan, H. G., X. Fang, D. Wright, J. G. Eaton, and J. H. McCormick. 1995. Simulation of dissolved oxygen profiles in a transparent, dimictic lake. Limnology and Oceanography 40:105–118.

Stefan, H. G., M. Hondzo, X. Fang, J. G. Eaton, and J. H. McCormick. 1996. Simulated long-term temperature and dissolved oxygen characteristics of lakes in the north-central United States and associated fish habitat limits. Limnology and Oceanography 41:1124–1135.

Stemberger, R. S. 1995. The influence of mixing on rotifer assemblages of Michigan lakes. Hydrobiologia 297:149–161.

Stewart, I., A. A. Seawright, and G. R. Shaw. 2008. Cyanobacterial poisoning in livestock, wild mammals and birds – an overview. Pages 613–637 *in* H. K. Hudnell, editor. Cyanobacterial Harmful Algal Blooms: State of the Science and Research Needs. Springer New York.

Tessier, A. J., and J. Welser. 1991. Cladoceran assemblages, seasonal succession and the importance of a hypolimnetic refuge. Freshwater Biology 25:85–93.

US EPA. 1986. Ambient Water Quality Criteria for Dissolved Oxygen. Office of Water, US Environmental Protection Agency, Washington, DC.

US EPA. 1998. Guidelines for Ecological Risk Assessment. Risk Assessment Forum, Washington, DC.

US EPA. 2000a. Nutrient Criteria Technical Guidance Manual, Lakes and Reservoirs. Office of Water, US Environmental Protection Agency, Washington, DC.

US EPA. 2000b. Nutrient Criteria Technical Guidance Manual, Rivers and Streams. Office of Water, US Environmental Protection Agency, Washington, DC.

US EPA. 2001. Nutrient Criteria Technical Guidance Manual, Estuarine and Coastal Marine Waters. Office of Water, US Environmental Protection Agency, Washington, DC.

US EPA. 2007. Survey of the Nation's Lakes: Field Operations Manual. Office of Water, U.S. Environmental Protection Agency, Washington, DC.

US EPA. 2008. Nutrient Criteria Technical Guidance Manual, Wetlands. Office of Water, US Environmental Protection Agency, Washington, DC.

US EPA. 2009. Survey of the Nation's Lakes: Quality Assurance Project Plan. Office of Water and Office of Environmental Information, Washington, DC.

US EPA. 2010a. Using stressor-response relationships to derive numeric nutrient criteria. Office of Water, U.S. Environmental Protection Agency, Washington, DC.

US EPA. 2010b. National Lakes Assessment: A Collaborative Survey of the Nation's Lakes. Office of Water and Office of Research and Development, Washington, DC.

US EPA. 2011. 2012 National Lakes Assessment. Field Operations Manual. Office of Water, US Environmental Protection Agency, Washington, DC.

US EPA. 2012a. 2012 National Lakes Assessment. Laboratory Operations Manual. U.S. Environmental Protection Agency, Washington, DC.

US EPA. 2012b. 2012 National Lakes Assessment Quality Assurance Project Plan. Office of Water, U.S. Environmental Protection Agency, Washington, DC.

US EPA. 2012c. 2012 National Lakes Assessment Site Evaluation Guidelines. Office of Water, Washington, DC.

US EPA. 2014. Framework for Human Health Risk Assessment to Inform Decision Making. Risk Assessment Forum, Washingon, DC.

US EPA. 2015a. Preventing Eutrophication: Scientific Support for Dual Nutrient Criteria. Office of Water, Washington DC.

US EPA. 2015b. Drinking Water Health Advisory for the Cyanobacterial Microcystin Toxins. Office of Water, U.S. Environmental Protection Agency, Washington, DC.

US EPA. 2015c. Recommendations for Public Water Systems to Manage Cyanotoxins in Drinking Water. Office of Water, U.S. Environmental Protection Agency, Washington, DC.

US EPA. 2019. Recommended human health recreational ambient water quality criteria or swimming advisories for microcystins and cylindrospermopsin. Office of Water, Washington DC.

Vadeboncoeur, Y., E. Jeppesen, M. J. V. Zanden, H.-H. Schierup, K. Christoffersen, and D. M. Lodge. 2003. From Greenland to green lakes: Cultural eutrophication and the loss of benthic pathways in lakes. Limnology and Oceanography 48:1408–1418.

Vanderploeg, H. A., S. A. Ludsin, S. A. Ruberg, T. O. Höök, S. A. Pothoven, S. B. Brandt, G. A. Lang, J. R. Liebig, and J. F. Cavaletto. 2009. Hypoxia affects spatial distributions and overlap of pelagic fish, zooplankton, and phytoplankton in Lake Erie. Journal of Experimental Marine Biology and Ecology 381, Supplement:S92–S107.

Vanni, M. J. 1987. Effects of nutrients and zooplankton size on the structure of a phytoplankton community. Ecology 68:624–635.

Wagner, T., P. A. Soranno, K. E. Webster, and K. S. Cheruvelil. 2011. Landscape drivers of regional variation in the relationship between total phosphorus and chlorophyll in lakes. Freshwater Biology 56:1811–1824.

Walker, W. W. 1979. Use of hypolimnetic oxygen depletion rate as a trophic state index for lakes. Water Resources Research 15:1463–1470.

Westrick, J. A., D. C. Szlag, B. J. Southwell, and J. Sinclair. 2010. A review of cyanobacteria and cyanotoxins removal/inactivation in drinking water treatment. Analytical and Bioanalytical Chemistry 397:1705–1714.

Wetzel, R. G. 2001. Limnology, Third Edition: Lake and River Ecosystems. Academic Press, San Diego.

Wood, S. N. 2006. Generalized additive models: An introduction with R. CRC Press, Boca Raton, FL.

Yuan, L. L., and J. R. Jones. 2019. A Bayesian network model for estimating stoichiometric ratios of lake seston components. Inland Waters 9:61–72.

Yuan, L. L., and J. R. Jones. 2020a. Modeling hypolimnetic dissolved oxygen depletion using monitoring data. Canadian Journal of Fisheries and Aquatic Sciences 77:814–823.

Yuan, L. L., and J. R. Jones. 2020b. Rethinking phosphorus-chlorophyll relationships in lakes. Limnology and Oceanography.

Yuan, L. L., and A. I. Pollard. 2014. Classifying lakes to improve precision of nutrient–chlorophyll relationships. Freshwater Science 33:1184–1194.

Yuan, L. L., and A. I. Pollard. 2017. Using national-scale data to develop nutrient–microcystin relationships that guide management decisions. Environmental Science & Technology 51:6972–6980.

Yuan, L. L., and A. I. Pollard. 2018. Changes in the relationship between zooplankton and phytoplankton biomasses across a eutrophication gradient. Limnology and Oceanography 63:2493–2507.

Yuan, L. L., and A. I. Pollard. 2019. Combining national and state data improves predictions of microcystin concentration. Harmful Algae 84:75–83.

Zhang, F., J. Lee, S. Liang, and C. K. Shum. 2015. Cyanobacteria blooms and non-alcoholic liver disease: evidence from a county level ecological study in the United States. Environmental Health 14:41.

# Appendix A: State Case Study: Chlorophyll *a*-Microcystin

This case study in Iowa describes chlorophyll *a* (Chl *a*) and microcystin data collected by the Iowa Department of Natural Resources (IDNR) that are combined with national data to estimate a stressor-response relationship for the state (Yuan and Pollard 2019).

## A.1 Data

Chl *a* measurements in Iowa were collected as part of an ambient lake monitoring program conducted by IDNR. Water samples were collected with an integrated water column sampler above the thermocline, when present, to a maximum depth of 2 meters (m) at the deepest point of each lake. Lake water samples were collected in the summer (May–September). An aliquot of the water sample was analyzed for Chl *a* in the laboratory by non-acidified fluorometry after filtering water samples through GF/C filters. In a separate IDNR monitoring program, microcystin concentrations are sampled regularly at swimming beaches in Iowa during the summer. This sampling effort includes state park beaches and locally managed beaches across the state. Microcystin was quantified in composite water samples collected at nine different locations on three transects spanning the swimming beach. On each transect, samples were collected at depths of 0.15, 0.5, and 1.0 m. Chl *a* and microcystin samples were matched by lake and sampling date for use in the analysis. To maximize the available data, microcystin and Chl *a* measurements collected within 1 day of each other were included as matched samples.

## A.2 Statistical Analysis

The structure of a statistical model that accommodates data collected at different spatial scales must be defined to ensure that the available data appropriately inform model estimates. Consider the case of a large national data set of approximately 1,000 samples and a state data set of approximately 50 samples. If the two data sets were pooled, the national data would dominate the state data simply because of the larger sample size, and the state data would exert a weak influence on the model. In any single state, however, only about 20 samples from the national data might be available, and we would expect the state data to dominate

estimates. Defining a hierarchical structure in the model helps ensure that each data set exerts the appropriate influence on the model results (Gelman and Hill 2007).

A second issue that arises from combining data sets is that different measurements are often collected in the different data sets. This problem is addressed in the national models by modeling a comprehensive network of relationships between different parameters to take advantage of the many different measurements available in the National Lakes Assessment (NLA) data (Qian and Miltner 2015). Then, state data sets in which only a subset of measurements were collected could still be feasibly modeled by informing specific aspects of the network.

State data from Iowa were included in the national model and inform estimates of relationships in the same network. As mentioned earlier, however, only Chl *a* and microcystin measurements were available in the Iowa state data set. To prevent over-specifying the model, EPA selected one of the relationships in the network that could be refined with data from the state. The relationship between Chl *a* and the relative biovolume of cyanobacteria relied most heavily on empirical calibration, so it was selected for refinement with state data. More specifically, the national model was revised so that model coefficients specific to each state were estimated (Equation (17)).

$$E[logit(p_{c,i})] = f_{1,k[i]} + f_{2,k[i]}chl_i + f_{3,k[i]}chl_i^2 \quad (36)$$

where different values of each of the coefficients were estimated for each state in the United States, *k*. The values of the coefficients for each state were constrained by normal distributions defined by the parameters, $\mu_f$ and $\sigma_f$. For example, the set of state-specific coefficients for $f_1$ were drawn from a single normal distribution as follows:

$$f_1 \sim Normal(\mu_{f_1}, \sigma_{f_1}) \quad (37)$$

Identical expressions can be written for the set of $f_2$ values and $f_3$ values. These distributions constrained the range of possible values so estimates of those parameters computed with relatively small sample sizes within individual states can "borrow" information from estimates computed from other states (Gelman and Hill 2007).

Iowa state data were included in the model by noting that the data should inform estimates of the coefficients only in the state of Iowa. That is, estimates of $f_1$, $f_2$, and $f_3$ from Equation (36) in Iowa are based on both the Iowa state data set and NLA data collected in Iowa. In other states,

estimates of the coefficients are based only on NLA data. The influence of Iowa state data on the national distributions of the coefficients (as characterized by $\mu_f$ and $\sigma_f$) is limited because the data affect only one element of the overall distributions of coefficients. Within the state of Iowa, however, the coefficients can be fit to maximize the predictive accuracy of the overall relationship linking Chl *a* to microcystin for both Iowa data and NLA data collected in Iowa, while remaining consistent with the range of possible values observed across all states.

One final difference in fitting the Iowa state data is that several sources of variability modeled separately in the national model (e.g., $s_1$ and $s_2$ in equations (18) and (19)) are combined into one estimate of residual variability. This combination of error terms reflects the data available from Iowa, in which no laboratory replicates or direct measurements of cyanobacterial biovolume were available. Hence, one lumped source of variability was estimated.

For comparison, a simple bivariate model was fit using only IDNR data, in which microcystin concentration was modeled as a quadratic function of Chl *a*.

## A.3 Results

A total of 556 samples of Chl *a* were measured at 28 lakes in Iowa. In some lakes, microcystin concentrations were sampled at different beaches, so 686 observations of microcystin were matched to the Chl *a* measurements.

In the revised national model with state-specific relationships between Chl *a* and the relative biovolume of cyanobacteria, coefficients varied substantially among states. Because coefficient values for quadratic relationships are not easily interpreted, the predicted mean cyanobacterial-relative biovolume at a Chl *a* concentration of 20 microgram per liter (µg/L) is plotted to visualize the range of variation among states (Figure 34). For comparison, among all the national data, mean cyanobacterial-relative biovolume was 0.18 at Chl *a* concentration of 20 µg/L. Systematic changes in cyanobacterial-relative biovolume with latitude or longitude were not evident, but some regional differences were observed. For example, cyanobacterial-relative biovolume for a Chl *a* concentration of 20 µg/L in Northeast states was generally lower than elsewhere, whereas in Midwest states, it was somewhat higher.

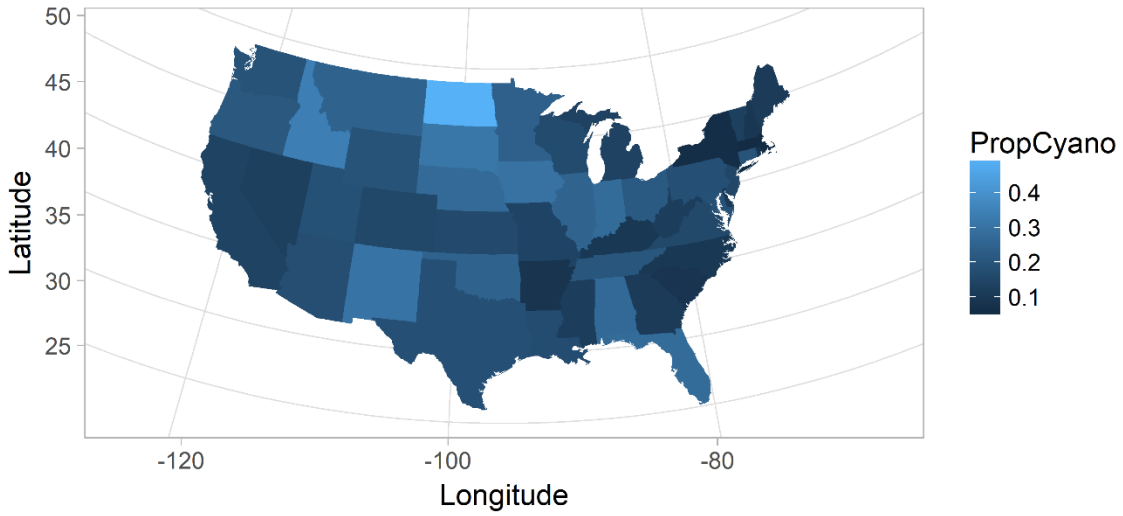


Figure 34. Variation in the relationship between Chl *a* and cyanobacterial-relative biovolume among states. PropCyano: predicted mean relative biovolume of cyanobacteria at an illustrative Chl *a* = 20 µg/L.

As described previously, the relationship between Chl *a* and cyanobacterial relative biovolume in Iowa was adjusted to maximize the accuracy of the predicted microcystin. Inclusion of Iowa data reduced the magnitude of the slope of the relationship between Chl *a* and cyanobacterial-relative biovolume but increased the intercept (Figure 35). So, higher values of cyanobacterial-relative biovolume were observed at Chl *a* concentrations less than about 10 µg/L. At higher Chl *a* concentrations, inclusion of Iowa state data did not substantively change the predicted cyanobacterial-relative biovolume. Overall, in Iowa, the estimated relationship between cyanobacterial-relative biovolume and Chl *a* was statistically indistinguishable from a constant value (Figure 35). The addition of the state data also narrowed the range of the credible intervals, as would be expected.

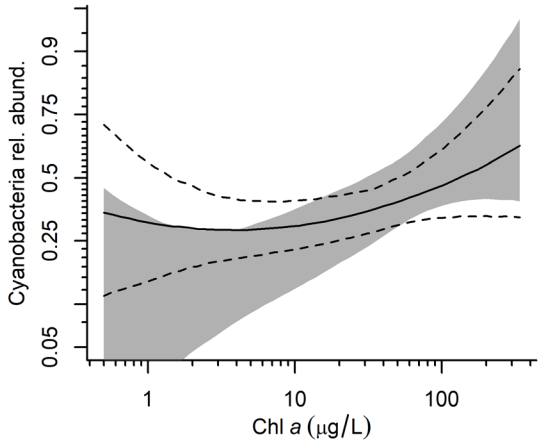


Figure 35. Comparison of Chl *a*/cyanobacterial-relative biovolume relationships in Iowa. *Filled gray*: 90% credible intervals for estimate of relationship using only NLA data collected in Iowa; *solid and dashed lines*: mean and 90% credible intervals for estimate of relationship using both Iowa state and NLA data.

The predicted mean relationship between Chl *a* and microcystin in Iowa from the state-national model closely followed the observed data (left panel, Figure 36), exhibiting a slight increase in slope as Chl *a* concentration increased. The 90% prediction intervals shown in the plot were based on the mean values of repeated random draws of 15 samples from the predicted distribution to replicate the plotted observed data. The intervals were broad and included most of the estimated mean values. The curvature observed in the simple bivariate fit between Chl *a* and microcystin using only Iowa data was opposite of that observed from the state-national model, predicting that the rate of increase in microcystin was lower at high Chl *a* concentrations than at low Chl *a* concentrations (right panel, Figure 36). The 90% prediction intervals of this fit also included most of the observed mean values, but qualitatively, the simple bivariate model did not match the observed data as closely as did the state-national model.

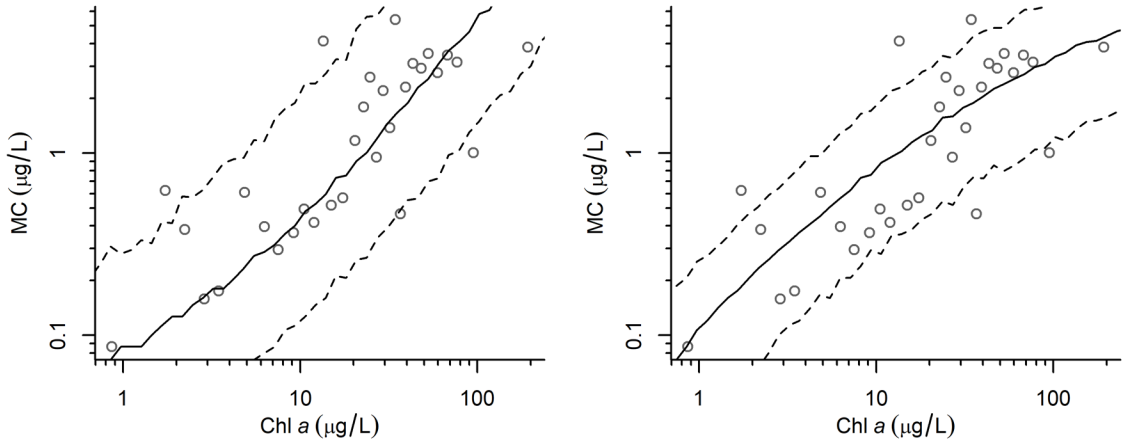


Figure 36. Comparison of predicted relationship between Chl *a* and microcystin for the state-national model (*left panel*) and a model using only Iowa state data (*right panel*). *Open circles*: average microcystin concentration computed in ~15 samples at the indicated Chl *a*; *solid lines*: mean relationship; *dashed lines*: 5$^{th}$ to 95$^{th}$ percentiles of distribution of means of 15 samples drawn from predicted distribution.

Three features inherent to the model combining state and national data are likely responsible for the improved predictions of observations in the Iowa data set. First, the network of relationships specified in the national model define a nonlinear function linking Chl *a* to microcystin that yielded a curved mean response (left panel, Figure 36). When only Iowa data are available, no information regarding the functional form of the relationship between Chl *a* and microcystin is known. Hence, it is difficult for the model to identify the correct shape of the curve. Indeed, the concavity of the mean relationship identified by the model using only Iowa data (right panel, Figure 36) was opposite of that estimated in the combined state-national model. Second, the network of relationships in the state-national model provided information regarding unobserved variables and relationships that could be used in lieu of direct observations. In this example, the relationships between Chl *a* and total phytoplankton biovolume and between cyanobacterial biovolume and microcystin were supplied by the national model. The Iowa-only model lacked the benefit of the additional information, and hence, for this model a direct relationship between Chl *a* and microcystin had to be estimated that aggregated the different causal linkages. Finally, the hierarchical structure of the national model placed constraints on the range of possible values for parameters estimated within each state. These constraints limited model parameters for the state data set to values that were generally consistent with national parameters.

## A.4 Criteria Derivation

Derivation of a recommended Chl *a* criterion based on decisions such as allowable exceedance rate, targeted microcystin, and model uncertainty follows an identical process as described for the national model. The model based on both IDNR data and NLA data yields a slightly different relationship from the model estimated from only the national data (Figure 37). Slightly greater uncertainty accompanies the estimate of the mean relationship in the Iowa-NLA model than the estimate in the NLA-only model (see Figure 22), and that uncertainty is reflected in a broader range of possible Chl *a* criteria. In the example shown in Figure 37, to maintain a maximum exceedance rate of 1% of microcystin of 8 µg/L, the Chl *a* criterion associated with the bound of the 50% credible interval was 14 µg/L.

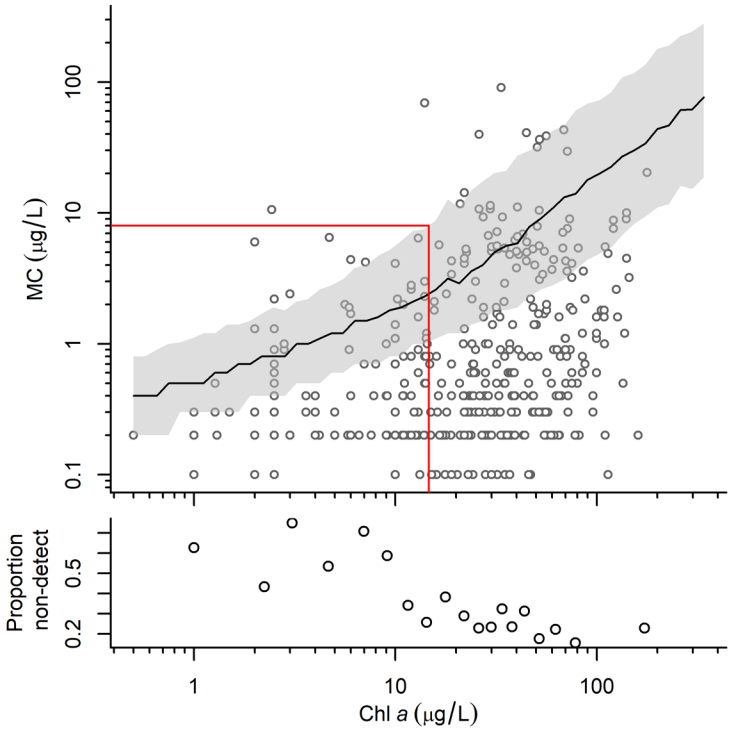


Figure 37. Microcystin and Chl *a* measurements in Iowa. *Top panel–open circles*: observed values of microcystin and Chl *a* for samples in which microcystin was greater than the detection limit; *solid line*: predicted microcystin that will be exceeded 1% of the time for the indicated Chl *a* concentration; *gray shading*: 50% credible interval about mean relationship; *horizontal and vertical line segments*: candidate Chl *a* criteria based on targeted microcystin. *Bottom panel*: proportion of samples for which microcystin was not detected in ~100 samples centered at the indicated Chl *a* concentration.

# Appendix B: State Case Study: Chlorophyll *a*–Hypoxia

This case study in Missouri describes national and state data that are combined to refine estimates of the relationship between chlorophyll *a* (Chl *a*) and deep water hypoxia. As described in Section 3.2.2, mean concentrations of dissolved oxygen below the thermocline ($DO_m$) decrease with time during the period of summer stratification. The sampling design of the NLA allowed for one visit to most of the lakes, so estimating temporal changes in deep water $DO_m$ in the national model required a space-for-time substitution. State monitoring data collected during multiple visits to a smaller number of lakes provided an opportunity to directly estimate temporal changes in $DO_m$ and to compare the relationship between eutrophication and the rate of oxygen depletion with estimates from NLA data.

## B.1 Data

The Missouri data considered in this case study were collected an average of 3–4 times per year by the University of Missouri (MU) from 1989 to 2007 as part of a statewide monitoring effort. Samples were collected near the dam for each reservoir (herein referred to as lakes for simplicity), where vertical profiles for temperature and DO concentration were measured (YSI model 51B or 550A meters). Composite water samples from a depth of approximately 0.25 meter (m) were transferred to high density polyethylene containers, placed in coolers on ice, and transported to the MU Limnology Laboratory. There, a 250-milliliter aliquot was filtered (Pall A/E) for determination of total Chl *a* via fluorometry following pigment extraction in heated ethanol (Knowlton et al. 1984, Sartory and Grobbelaar 1984). A total of 198 measurements of $DO_m$ were available for analysis, collected at 20 different lakes over 62 unique lake-year combinations.

## B.2 Statistical Analysis

The same model equations used in the national model were applied to data collected in Missouri:

$$E[DO_{m,i}] = DO_0 + VOD_{k[i]}\left(t_i - t_{0,j[i]}\right) \quad (38)$$

where $DO_0$ is the value of $DO_m$ at the start of spring stratification, volumetric oxygen demand $(VOD)_k$ is the net imbalance in the volumetric oxygen budget for lake *k* corresponding to sample *i* expressed as milligrams per liter per day of DO (Burns 1995), $t_i$ is the date that sample *i* is collected, and $t_{0,j}$ is the date of the beginning of stratification for lake-year *j*. Observed values of $DO_m$ were assumed to be normally distributed with a standard deviation of $\sigma_1$ about the expected value. Note that, like the national model, VOD is assumed to be constant for each lake, but the date of the beginning of stratification varied by year and lake. The model equation specifying the relationship between Chl *a*, dissolved organic carbon (DOC), and lake depth and VOD was the same equation used in the recommended national model (see Equation (11)). As with the national model, saturation DO concentrations at the minimum temperature in Missouri were used to set the value of $DO_0$.

The treatment of DO measurements less than 2 milligrams per liter (mg/L) in the Missouri data differed from the approach used in the NLA. From 2 to 14 measurements of $DO_m$ greater than 2 mg/L were available in the Missouri data set for each of the lake-years included in the model, so data were available to directly estimate temporal changes in $DO_m$. Because data were available at each lake before $DO_m$ approached zero, measurements of $DO_m$ that were less than 2 mg/L could be excluded without biasing the model results.

Two models were run to explore the effects of combining Missouri data with the national model. In the first model, only Missouri data were used, and in the second model, both Missouri and NLA data were used to estimate the parameter values.

## B.3 Results

The range of values spanned by each of the covariates differed between the two data sets. Missouri measurements were collected over a broader range of days than the NLA, whereas lakes sampled by the NLA covered a broader range of Chl *a* concentrations (Figure 38). Variations in DOC concentrations and depths below the thermocline were also narrower in the Missouri data than in the NLA data. Those differences in the range of observations were reflected in the strength of correlation between each covariate and $DO_m$. For Missouri, sampling day was most strongly correlated with $DO_m$, whereas for the NLA, sampling day exhibited the weakest correlation with $DO_m$. Instead, in the NLA data, Chl *a*, DOC, and the depth below the thermocline were all more strongly correlated with $DO_m$.

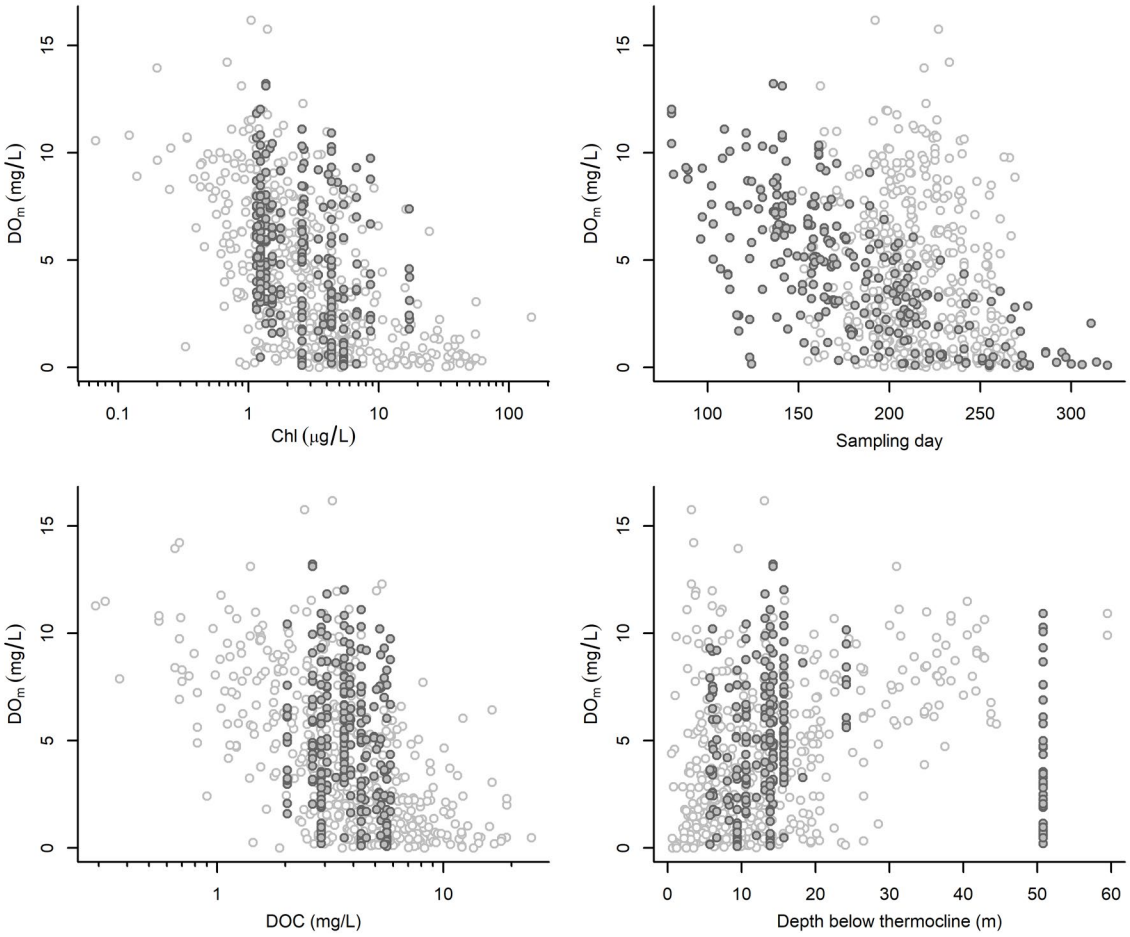


Figure 38. Observed $DO_m$ vs. Chl *a*, sampling day, DOC, and depth below the thermocline. *Open circles*: NLA data*; filled circles*: Missouri.

The first day of stratification for Missouri lakes was generally earlier than for most of the dimictic lakes considered in the national model (Figure 39), a finding that is consistent with the fact that Missouri is located at the southern end of the geographic distribution of dimictic lakes (see Figure 7). Both the Missouri-only model and the NLA-only model yielded similar estimates of the relationship between Chl *a* and VOD ($d_2$ in Equation (11)) (Figure 40), and the estimate based on the combined data sets improved further on the precision. Estimates of coefficients characterizing the relationship between VOD and depth below the thermocline ($d_3$) and DOC ($d_4$) were much more precise in the NLA-only data set than in the Missouri-only data set. Hence, the estimate based on the combined data set mainly reflects the trends in the NLA data.

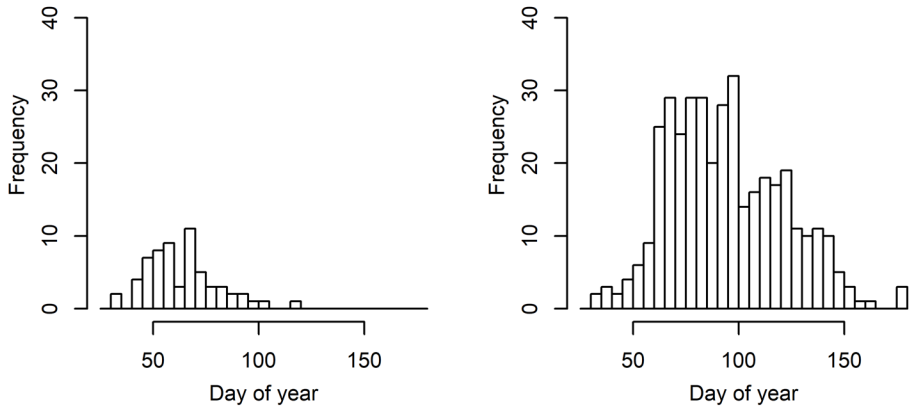


Figure 39. Estimated first day of stratification for Missouri lakes (*left panel*) and NLA lakes (*right panel*).

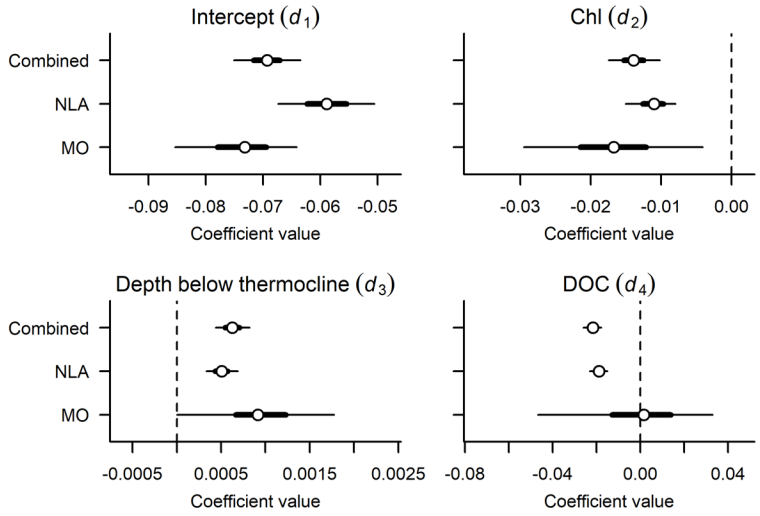


Figure 40. Model coefficients estimated for models for Missouri data, NLA data, and combined data. *Thick line segment*: 50% credible intervals; *thin line segment*: 90% credible intervals; *vertical dashed line*: coefficient value of zero.

Qualitatively, the model accurately represented the decrease in $DO_m$ over time in different lakes (Figure 41). The effects of differences in the timing of spring stratification was manifested as differences in the vertical position of each line, and in some lakes, substantial variation was observed across years.

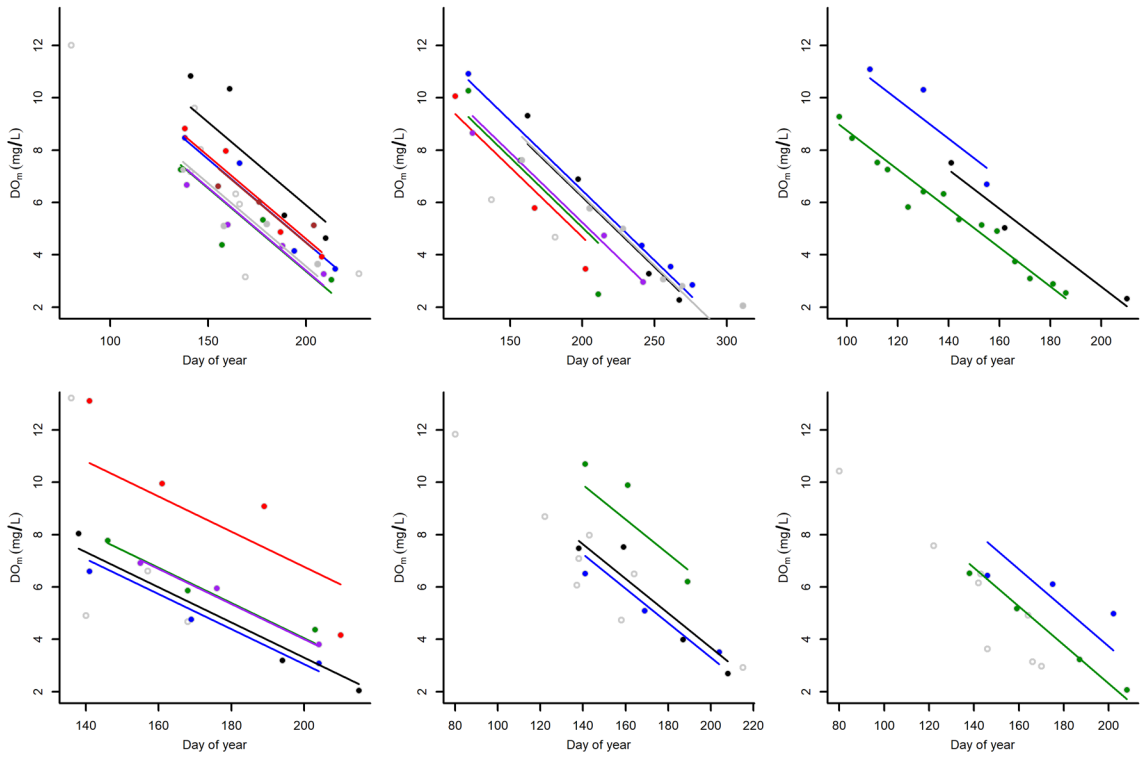


Figure 41. Relationships between day of year and $DO_m$ for six Missouri lakes. *Different line and symbol colors in each panel* correspond to data collected within different years with at least three samples. *Open gray circles*: other samples collected at each lake.

## B.4 Criteria Derivation

The utility of combining Missouri and NLA data to inform decision-making is evident when one considers the predicted relationship between Chl *a* and $DO_m$ calculated using parameter estimates from the Missouri data and from the combined Missouri-NLA data set (Figure 42). In the example shown, the relationship is calculated based on illustrative values for other covariates (depth below thermocline at 10 m, DOC at 1.6 mg/L, and time between spring stratification and sampling at 130 days). Because use of both data sets improves the precision of model parameters, the resulting mean relationship is also estimated with increased precision and a targeted Chl *a* concentration can be identified with greater confidence. In this example, the 50% credible interval for the targeted Chl *a* concentration corresponding to an illustrative threshold of $DO_m$ = 0.1 extends from 5.5–8.9 µg/L when the combined model is used. When using only Missouri data, the interval expands to 4.5–9.7 µg/L.

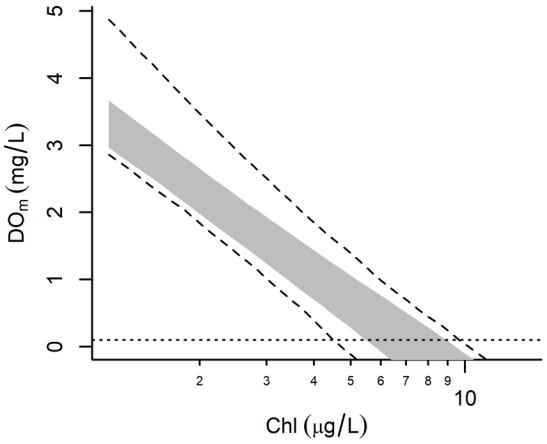


Figure 42. Relationship between Chl *a* and $DO_m$ in an illustrative lake with depth below thermocline at 10m, DOC at 1.6 mg/L, and 130 days after spring stratification. *Gray shading*: 50% credible intervals about mean relationship from combined Missouri-NLA model; *dashed line*: 50% credible intervals about mean relationship from Missouri-only model; *dotted line*: $DO_m$ = 0.1 mg/L.

# Appendix C: State Case Study: Total Nitrogen–Chlorophyll *a*

This case study in Iowa examines how combining locally collected measurements of total nitrogen (TN) and chlorophyll *a* (Chl *a*) with the national models can refine predictions calculated from these local data sets.

## C.1 Data

Data used for this case study were collected by the Iowa Department of Natural Resources (IDNR) as part of their routine monitoring program. For each lake in the data set, TN, nitrate-nitrite, Chl *a*, and dissolved organic carbon (DOC) values were measured. A total of 968 observations collected at 31 different lakes were available for analysis.

## C.2 Statistical Analysis

The same model formulation provided in Equation (33) was applied to the IDNR data, expressing TN-dissolved inorganic nitrogen (-DIN) as the sum of a phytoplankton compartment, modeled as $f_1 Chl^k$, and a dissolved organic nitrogen (DON) component, modeled as $f_2 DOC$:

$$E[TN - DIN] = f_1 Chl^k + DON = f_1 Chl^k + f_2 DOC \quad (39)$$

DOC measurements were available only at a small proportion of Iowa lakes, so EPA simplified the national model to the following form for modeling Iowa data:

$$E[TN - DIN] = f_1 Chl^k + u \quad (40)$$

where $u$ is a lake-specific constant representing the contributions of DON in each lake to observed values of TN-DIN. Recall also that, in the national model, the coefficient $f_1$ varied across states. With the IDNR data set, multiple samples were collected from each lake, so the model could be refined further to estimate a value of $f_1$ for each lake as follows:

$$\log(f_{1,j}) \sim Normal(\mu_{f1,IA}, \sigma_{f1}) \quad (41)$$

where the index, $j$, refers to different lakes, and the mean value $\mu_{f1,IA}$ is computed for data collected in Iowa.

To examine the effects of considering local state data in the context of the national model, two models were fit. In the first model, only IDNR data were used to estimate the coefficients. In the second model, relationships were fit to both the IDNR data and NLA data simultaneously. The exponent *k* was modeled as being the same in both the IDNR and NLA data, while the coefficients $f_1$ for each lake were estimated with IDNR data and NLA data collected within Iowa, and the value of $\mu_{f1,IA}$ was constrained by the national distribution among all the states in the NLA data.

## C.3 Results

Data collected during the NLA in Iowa and by IDNR spanned similar ranges of Chl *a*, TN-DIN, and DOC (Figure 43). The limiting relationship between Chl *a* and TN-DIN estimated using only IDNR data approximated the lower edge of the cloud of points (gray shading) but were estimated with more uncertainty than when estimated using both IDNR and NLA data (solid lines). The mean limiting relationships between Chl *a* and TN-DIN estimated with the two models were statistically indistinguishable from one another.

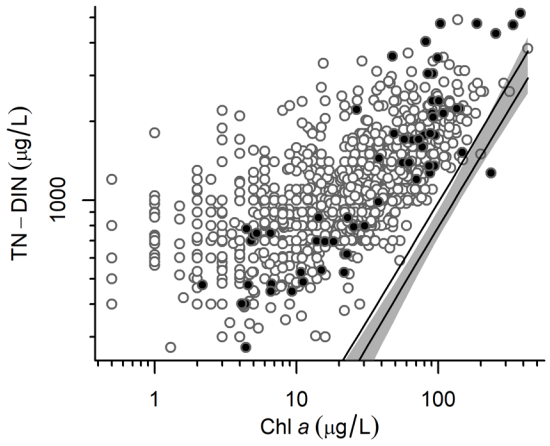


Figure 43. Chl *a* vs. TN-DIN in Iowa. *Open circles*: data collected by Iowa DNR; *filled circles*: data collected by NLA in Iowa; *solid lines*: 95% credible intervals for limiting relationships between Chl *a* and TN-DIN estimated using both NLA and IDNR data; *shaded gray area*: 95% credible intervals for limiting relationships estimated using only IDNR data.

The root mean square (RMS) prediction error of log(TN-DIN) measurements in the IDNR data was the same for the models using only IDNR data (RMS = 0.27) and the combined Iowa – NLA data (RMS = 0.27), indicating that imposing national constraints on the parameter values did not improve the accuracy of predictions at the scale of the local state data. Uncertainty about

estimates of the relationship between TN-DIN and Chl *a* for individual lakes was very similar (example shown in Figure 44), indicating that a sufficient number of samples was available for each lake to estimate the relationship without the information provided by the national model.

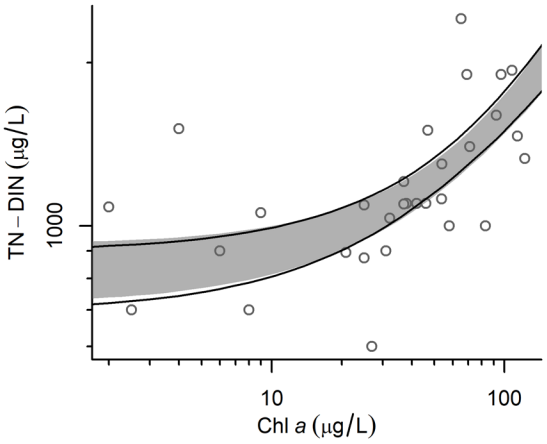


Figure 44. Chl *a* vs. TN-DIN in Beeds Lake, Iowa. *Open circles*: observed data; *gray shading*: 90% credible intervals for predicted relationship based on only IDNR data; *solid lines*: 90% credible intervals for predicted relationship using both IDNR and NLA data.

## C.4 Criteria Derivation

Because of the higher number of samples collected within each lake in the IDNR data set, unique relationships between TN-DIN and Chl *a* for each lake could be calculated, and those relationships, in turn, can be used to derive numeric nutrient criteria (Figure 45). Variations across lakes in DON and in the coefficients of the modeled relationship yield differences in the estimated relationship between TN-DIN and Chl *a*. Then, resulting TN ambient criterion differ as well. For an illustrative target Chl *a* concentration of 15 micrograms per liter (µg/L), the mean ambient TN criterion for the lake shown in the left panel of Figure 45 was 750 µg/L, while the TN criterion for the lake in the right panel was 1260 µg/L.

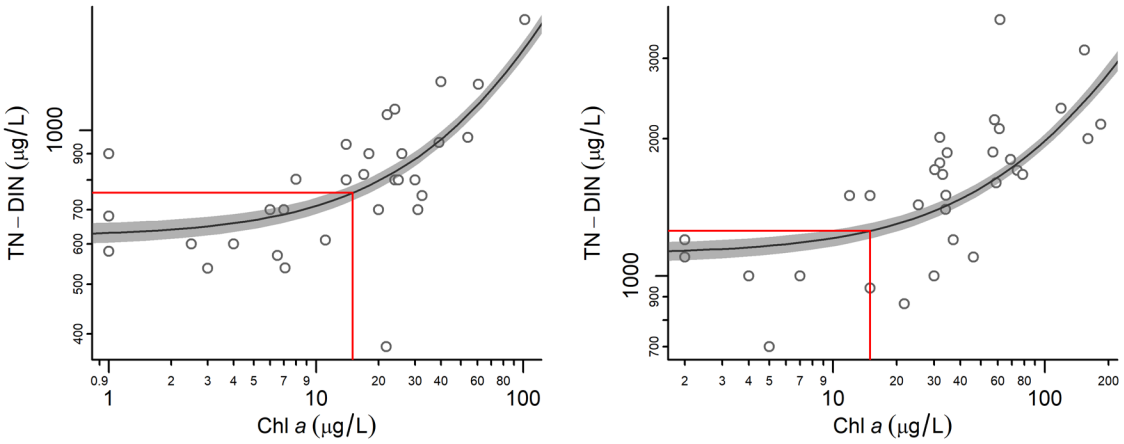


Figure 45. Lake-specific criteria derivation using combined Iowa-NLA model for two different lakes in Iowa. *Open circles*: observed values of TN-DIN and Chl *a* in Iowa for each lake; *gray shading*: 50% credible intervals about the mean relationship; *solid line*: mean relationship calculated using mean DOC concentration in lake; *horizontal and vertical line segments*: TN criterion calculation for illustrative Chl *a* target of 15 µg/L.

# Appendix D: Operational Numeric Nutrient Criteria

Operationally, chlorophyll *a* (Chl *a*), total nitrogen (TN), and total phosphorus (TP) criteria can be specified to account for the effects of sampling and temporal variability on observed mean concentrations (Barnett and O'Hagan 1997). In most cases, the condition of a lake will be assessed by examining a small number of samples and the uncertainty in the estimation of the true seasonal mean value from those data will be determined by the number of samples, the temporal variability of nutrient concentrations in the lake, and the inherent sampling variability of the measurement. By examining historical data from many different lakes, sampling variability associated with Chl *a*, TN, and TP can be estimated and "operational" criteria can be specified to account for this variability with adjusted criterion magnitudes and by adopting a frequency component that allows for some excursions of the specified magnitude.

Ambient monitoring of nutrient concentrations provides the basis for determining whether a lake complies with the specified numeric nutrient criteria. Because of logistical and resource restrictions, the number of water quality samples available at different lakes can vary from a single grab sample to weekly or monthly samples throughout the sampling season. Statewide monitoring designs also vary in how often a lake is visited in different years. For example, a typical rotating basin design might sample the same lake once every 5 years, whereas other lakes might be sampled every year. Because of the differences in the frequency of sample collection, a statistical analysis of available monitoring data might be necessary to accurately assess compliance with the numeric nutrient criteria. This appendix describes a statistical approach for deriving *operational* or *realizable* criteria magnitude, duration, and frequency components.

This document provides tools to compute numeric nutrient criteria expressed as seasonal mean values. Those criteria implicitly assumed that a large number of samples are available for characterizing the condition of each lake and that the uncertainty in the computation of the mean value is small (Barnett and O'Hagan 1997), a condition that is usually not satisfied by routine monitoring data. Operational criteria incorporate statistical uncertainty in estimating environmental conditions from a much smaller number of samples. The statistical approach

recommended here requires that one estimate the sampling and temporal variability of nutrient concentrations within lakes for which criteria are specified.

A variety of approaches are available that account for within-lake variability when defining operational criteria, but they should all be designed to consider that nutrient concentrations vary in space (e.g., at different points on a lake) and in time. Both sources of variability account for a distribution of nutrient concentrations that will arise when a lake is repeatedly sampled. For example, if a single sample of TP was collected from one lake every year, over 10 years, the distribution of values might be as shown in Figure 46, in which observed concentrations range from 30 to 80 micrograms per liter (μg/L). Given this example, the water quality management question is whether the lake complies with its specified numeric nutrient criteria. Here, if the criterion is 60 μg/L, a methodical approach for assessing compliance can enhance the utility of the criterion. This section provides one example of an approach for accounting for sampling variability and defining "operational" nutrient criteria.

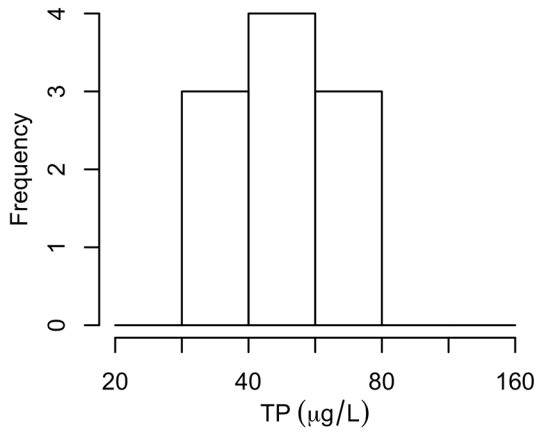


Figure 46. Example distribution of 10 TP measurements. Note that the horizontal axis is log-scaled.

Estimates of variability of measurements within lakes are needed to inform decisions on operational criteria, and those estimates can be computed from historical data. For this example, EPA analyzed TP data extracted from the Storage and Retrieval Data Warehouse (STORET) that had been collected in the summers from 1990 to 2011. From those data, lakes were identified in the United States with at least 5 years of nutrient data, yielding 25,056 samples collected from 846 different lakes. In this illustrative example, all available data were used, but screening data to identify lakes with relatively lower levels of anthropogenic nutrient

loading may be advisable. A statistical model was then used to estimate variance in nutrient measurements across different samples collected in the same year and from the same lake (within-lake variability). A model was fit to TP measurements that explicitly estimated intra-annual and interannual variability as follows:

$$\log(TP_i) = a_{j[i]} + b_{k[i]} + r_i \quad (42)$$

where $TP_i$ is measured in sample $i$ at site $j$ and in year $k$. So, observed TP in a sample is modeled as being log-normally distributed about a mean value that is the sum of an overall site mean ($a_j$) and a random effect of year ($b_k$). The random effect of year is assumed to be normally distributed with a mean value of 0 and a standard deviation of $s_{year}$, and the intra-annual variance ($r_i$) is modeled as a normal distribution with a mean of 0 and a standard deviation of $s_{sample}$. Intra-annual variance not only includes contributions from traditional sources of sampling variability (e.g., measurement uncertainty), but also includes variability that could be attributed to differences in TP concentrations among different locations in a lake and differences in TP concentrations one might observe over the course of a single sampling season. Hence, intra-annual variance was expected to differ among different lakes, so, the overall distribution of different values of $s_{sample}$ was modeled as a half-Cauchy distribution (Gelman 2006).

Fitting this model to the TP data collected from STORET yielded a mean estimate of 0.15 for the standard deviation of intra-annual variability of log(TP). Among different lakes in the data set, this value ranged from 0.10 to 0.27, so sampling variability varied substantially among the lakes in the data set. Estimating intra-annual variability from local data collected in the lake of interest would help ensure that the estimate correctly reflects variability in the lake. The mean standard deviation of inter-annual variability was 0.11.

Once intra- and inter-annual variability for the lake or lakes of interest has been estimated, this information can be combined with the criterion for that lake to estimate a distribution of nutrient concentration values that would be observed if the lake complied with the criterion. For example, if the standard deviation of the intra-annual variability of log(TP) in a particular lake is estimated as 0.15 and the inter-annual variability is estimated as 0.11, the combined variability is the square root of $0.15^2 + 0.11^2$, or 0.19. For illustrative purposes, if we assume the TP criterion for the lake is 60 µg/L, we can infer the characteristics of the cumulative distribution of single observations of TP that would be observed at the lake if it were exactly complying with its criterion (Figure 47). Then, based on this distribution, operational criteria can be derived. For

example, one might define an operational criterion that corresponds with the 10th percentile of the distribution (TP = 34 µg/L) and assert that a single TP observation below that value indicates the probability that the mean TP concentration in the lake is greater than 60 µg/L is less than 10%. That is, a lake with an observation below that threshold is likely in compliance with the criterion. Conversely, one might define a criterion at the 90th percentile of the distribution (TP = 105 µg/L) and assert that a single TP observation that exceeds that value indicates the probability that the mean TP concentration is lower than 60 µg/L is less than 10%. That is, any lakes with an observation that exceeds that threshold is likely to be out of compliance with the criterion. Different water quality management outcomes (e.g., additional sampling) could be triggered at different threshold concentrations. Also, different operational criteria can be developed depending on probabilities of error that are acceptable to environmental managers.

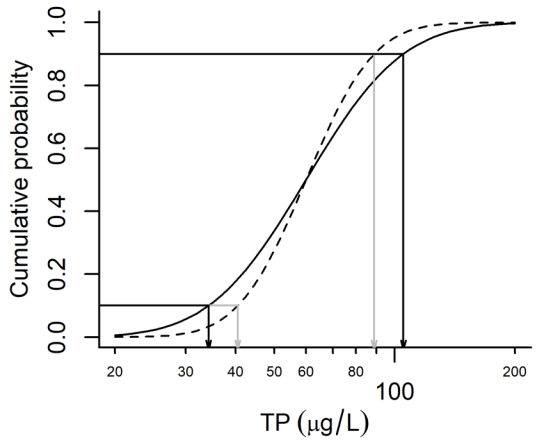


Figure 47. Example of defining an operational criterion magnitude. *Solid line*: the cumulative probability of observing a single sample TP lower than or equal to the indicated value if the true annual mean was exactly equal to the criterion (TP = 60 µg/L); *dashed line*: the cumulative probability for the average of four samples; *black arrows*: operational criteria for one sample; *gray arrows*: operational criteria associated with the mean of four samples.

This analysis also highlights the relative benefits of collecting additional samples from each lake. More specifically, the combined standard error (*s.e.*) on the estimate of a summer mean concentration is as follows:

$$s.e. = \sqrt{\frac{s^2_{sample}}{N} + s^2_{interannual}} \quad (43)$$

where $N$ is the number of samples collected, $s_{sample}$ is the sampling variability of the nutrient concentration (estimated as an intra-annual standard deviation), and $s_{interannual}$ is the inter-annual standard deviation. Hence, additional samples increase the precision with which the annual average nutrient concentration can be estimated by reducing the effect of intra-annual variability. In Figure 47, the dashed line shows the cumulative probability distribution of mean values computed using four samples. Because of the reduction in the standard error, assessments for compliance can be made with much greater confidence. The same 10% probabilities used above for single samples yield operational criteria of 41 µg/L and 89 µg/L, when applied to the case of the mean of four measurements (gray arrows in Figure 47). Information and procedures regarding the use of operational criteria in assessment might be described in a state's assessment methodology to accompany criteria specified in the water quality standards.

# Appendix E: Example Workflow for Deriving Lake Nutrient Criteria

To further assist states and authorized tribes, an example workflow for deriving numeric nutrient criteria for a lake or reservoir is illustrated in this appendix by working through the case for a hypothetical Lake XYZ. The location of the lake is selected for illustrative purposes only, and the characteristics of Lake XYZ (e.g., depth, ecoregion, location, elevation) do not correspond to an actual location in the United States. Frequently Asked Questions (FAQs) are also included in each of the steps.

## E.1 Identify Designated Uses

Designated uses for Lake XYZ are aquatic life protection and recreation. For aquatic life, a reproducing population of cool water fish are present. For recreation, swimming in the lake is a common summer activity.

## E.2 Compile Data

The maximum depth of Lake XYZ is 10 meters (m), and on average, the thermocline in the summer is located 3 m below the surface. Mean dissolved organic carbon (DOC) concentration is 5 milligrams per liter (mg/L). The lake is located at 42.5 °N, 83.5 °W at an elevation of 1500 m. The lake is located in Ecoregion 55, the Eastern Cornbelt Plains. Cool-water fish species in the lake can tolerate temperatures up to 19° Celsius (C) and dissolved oxygen (DO) down to 4 mg/L.

FAQs:

1. *How do I measure depth below the thermocline?*
   Depth below the thermocline is best estimated by examining vertical profiles of temperature and estimating the depth of the thermocline as the shallowest depth where the temperature gradient exceeds 1° C/m. Depth below the thermocline can then be calculated as the difference between maximum depth and thermocline.
2. *What do I do with multiple measurements of depth?*
   Lakes vary in depth depending on location, but to be consistent with measurements collected by the National Lakes Assessment, depth can be recorded as the maximum

depth in natural lakes and the midpoint of reservoirs. Thermocline depth varies with the day of sampling, so if data from different days are available, a seasonal average thermocline depth can be used in the models.

3. *How do I develop statewide criteria from lake-specific numbers?*

   The models provide a way to calculate criterion values based on factors that may vary among lakes in a state (e.g., depth, ecoregion, DOC). Different options are available for computing statewide criteria depending on how much these factors vary within a state, ranging from lake-specific criteria to a single set of criterion values that applies to all lakes. Analysis to examine the range of nutrient criteria associated with different lakes in a state can inform these decisions. For example, if a large range of lake depths are found in a state, and criterion values differ widely among those lakes, then specifying criteria by depth class may be appropriate. The criterion models provide a tool for directly quantifying the effects of differences in lake characteristics on criterion values.

4. *How do I derive criteria when data needed by the models are missing?*

   Data for specific parameters (e.g., DOC) may not be available for every lake, but often information regarding the distribution of parameter values within a state or ecoregion can be estimated. For example, if the mean and standard deviation of DOC in a state or ecoregion can be computed, then this distributional information can be used to specify a range of lake nutrient criteria.

## E.3 Management Decisions

For comparison, criterion values were calculated at three certainty levels: 50%, 80%, and 90%. Decisions were also required to determine the desired allowable probability of microcystin exceeding the recreational threshold of 8 micrograms per liter (µg/L), the size of the cool-water refuge, and the zooplankton slope threshold. A value of 1% was selected for the probability of exceeding the microcystin threshold, the cool-water refuge size was set at 0.3 m, and the zooplankton slope threshold was set at 0. The resulting candidate criteria are shown in the tables below (Table 6).

Table 6. Illustrative candidate Chl *a* criteria (µg/L) for Lake XYZ.

| Certainty level | Aquatic life | | Recreation |
|---|---|---|---|
| | Zooplankton biomass | Hypoxia | |
| 50% | 21 | 15 | 22 |
| 80% | 14 | 12 | 9 |
| 90% | 13 | 11 | 6 |

The tabulated values provide insights into both the most sensitive use in the lake and the uncertainty in chlorophyll *a* (Chl *a*) criteria associated with protecting that use. The candidate Chl *a* criterion for hypoxia is the lowest among uses at a certainty level of 50%, whereas candidate Chl *a* criteria for recreation are the lowest among uses for certainty levels of 80% and 90%. Among certainty levels, the range of Chl *a* criterion values is greatest for recreation, indicating that the sensitivity of certainty level to different criterion values is low relative to the other responses and uses. That is, small changes in the Chl *a* criterion are associated with small changes in the certainty level for this response. For hypoxia, the range of criterion values is narrow, indicating higher sensitivity. Here, small changes in the Chl *a* criterion are associated with relatively large changes in certainty level. Finally, the food web implications associated with lakes in which Chl *a* exceeds the level at which log(Z)/log(P) = 0 over the long term may be substantial and difficult to reverse, and therefore, the use of the 90% certainty level is recommended for this response to provide assurance that the candidate Chl *a* criterion for this response is protective. Based on these considerations, a final Chl *a* criterion of 12 µg/L might be selected. This value is slightly less than the 90% certainty level for the zooplankton biomass endpoint and corresponds to the 80% certainty level for the hypoxia endpoint. The value also corresponds with the 72% certainty values for recreation. This Chl *a* criterion value is used to compute the total phosphorus (TP) and total nitrogen (TN) criteria in Table 7.

Table 7. Criteria for TP and TN corresponding to a Chl *a* criterion of 12 µg/L

| Certainty level | TP (µg/L) | TN (µg/L) |
|---|---|---|
| 50% | 25 | 560 |
| 80% | 24 | 540 |
| 90% | 36 | 570 |

More generally, criteria could also be derived that are protective of different designated uses and applied to different lakes depending on the applicable uses, while ensuring that, for each lake, the most sensitive use is protected.

FAQs

1. *How do I select the certainty level?*

   The certainty level, or credible interval, specifies a range of possible criterion values within which the actual value lies with the specified probability. For example, the 50% certainty level implies that there is a 50% chance that the actual value is within the specified bounds (i.e., the 50% certainty level is bounded by the $25^{th}$ and $75^{th}$ percentiles of distribution of possible criteria). In other words, the selection of 50% certainty level indicates that 25% of possible criterion values lie on either side of the model's best prediction. Because the criterion is based on the lower bound of the 50% certainty level, there is a 25% chance that the derived criterion value is greater than the concentration needed to achieve the desired condition. That is, there is a 25% chance that the criterion is under-protective. More certainty on the criterion value is achieved by selecting greater certainty levels. For example, there is only a 5% chance that a criterion based on the 90% certainty level would be under-protective. The R Shiny apps provide transparent information regarding the effects of different levels of certainty and can provide a useful means of engaging with stakeholders regarding the risk management decisions underlying criterion development.

2. *How do I select a slope threshold for the zooplankton model?*

   A value of zero for the slope threshold for the zooplankton model identifies the point at which changes in zooplankton biomass are not associated with changes in phytoplankton biomass. In other words, on average there is no increase in zooplankton biomass given a corresponding increase in their food source, phytoplankton. At this point the food web in a lake may be out of balance, an indication that aquatic life use may not be supported. States and authorized tribes may opt to select a higher threshold to provide additional assurance that aquatic life is protected in their lakes.

3. *How do I select a refuge depth for the hypoxia model?*

   Limited data are available on the refuge depth that is necessary to protect cool- and cold-water fish from the combined effects of hypoxia and increased temperature. The

studies cited in the criterion document include studies conducted in lakes with both deep water hypoxia and warm surface layer temperatures, and studies conducted in lakes with cool water inflows where fish have been observed to congregate. A minimum refugia depth of 0.3 m is identified in some of these studies. But the final selection for the refuge depth may also consider other factors, including the likelihood of a prolonged stable stratification in a lake with no mixing events and the likelihood of elevated temperatures in the surface layer through the end of the summer. States and authorized tribes may opt to select a larger refuge depth to provide additional assurance that aquatic life is protected in their lakes.

4. *How do I select the exceedance probability for microcystin?*

   The exceedance probability for microcystin can be interpreted directly in terms of the environmental outcome. That is, if mean Chl *a* concentration in a lake is maintained at the calculated criterion value, then the exceedance probability can be expressed in terms of the number of days of exceedance. For example, a 1% exceedance probability suggests that on average, microcystin will exceed the threshold concentration 1 day out of 100, so on average, 1 day in the swimming season for most states would not support recreational use of the lake. Furthermore, in Section 3.2.3.3, a method is described for translating the 1-day exceedance probability to a seasonal probability for a specified number of excursions during 10-day assessment windows [as specified in the recreational cyanotoxin criteria, (US EPA 2019)], and this information can further inform selection of the exceedance probability. As with the certainty level, the R Shiny app provides a transparent means of communicating the effects of different exceedance probabilities to interested stakeholders.

5. *How do I interpret slider values that do not yield criterion values?*

   Certain combinations of slider selections do not yield numeric criteria in the R Shiny apps and instead will return an error message. An example of this scenario occurs in the zooplankton model when a high slope threshold is selected with a high certainty level. When a desired slider combination does not yield a criterion value, consult with EPA to refine the model and possibly incorporate site-specific information.

## E.4 Accounting for Interannual and Sampling Variance

Criterion values can account for sampling and temporal variability to increase the accuracy of compliance monitoring and to account for allowable exceedance frequencies of the criteria for Chl *a*, TN, and TP. Appendix D describes the statistical calculations for deriving operational criterion magnitudes and examples of those computations are shown below for the Chl *a* criterion. Based on the comparison of criterion values above, 12 µg/L was selected as the Chl *a* criterion. Intensive monitoring data at lakes similar to Lake XYZ indicate that the within-year standard deviation of $\log_e$(Chl *a*) in Lake XYZ is 0.8, while among-year standard deviation is 0.2. Given this information, the criterion magnitude can be adjusted to account for an allowable exceedance frequency. Here, to illustrate the approach, criterion magnitudes are calculated that can be exceeded once every 3 years.

To assess whether the lake is in compliance with a criterion value, we adjust the criterion magnitude as follows. First, calculate the combined standard deviation of Chl *a* measurements:

$$s_{comb} = \sqrt{s^2_{within-year} + s^2_{among-year}} = \sqrt{0.8^2 + 0.2^2} = 0.82 \quad (44)$$

For a once in 3 years exceedance frequency, calculate Chl *a* concentration that will be exceeded one-third of the time, assuming that Chl *a* measurements are log-normally distributed, with a mean value equal to the log-transformed criterion value, and a standard deviation equal to the combined standard deviation calculated above:

$$\log\left(C_{1\ in\ 3}\right) = \log\left(C\right) + z_{0.67}s_{comb} = \log(12) + 0.44(0.82) \quad (45)$$

Where $C_{1\ in\ 3}$ is the criterion value adjusted to an allowable one exceedance every 3 years, $C$ is the criterion value of 12 µg/L, $s_{comb}$ is computed above, and $z_{0.67}$ is the 67th percentile of the standard normal distribution (i.e., a 1 in 3 year frequency of exceedance). This calculation yields a value of 17 µg/L that is exceeded once every 3 years (assuming that a single sample of Chl *a* is collected each year and the mean Chl *a* concentration in the lake is equivalent to the criterion value of 12 µg/L). That is, if Lake XYZ was conforming exactly with the applicable criterion, then we would expect to see Chl *a* concentrations exceeding 17 µg/L in one-third of the years.

If several samples are combined to estimate each seasonal mean, the calculation for $s_{comb}$ is adjusted to account for this averaging as follows:

$$s_{comb} = \sqrt{\frac{s^2_{within-year}}{N} + s^2_{among-year}} = \sqrt{\frac{0.8^2}{4} + 0.2^2} = 0.6 \quad (46)$$

Note: the numerical example shows the computation when four samples are averaged. Based on this adjustment, $C_{1\ in\ 3}$ for seasonal means computed as the average of four samples is 16 µg/L.

## E.5 Incorporating State Data in National Models

Regular monitoring has been conducted at Lake XYZ and other lakes, in which a variety of water quality measurements are collected, and the state agency is interested in determining whether the state monitoring data can be combined with the national models to improve the accuracy and precision of the derived criteria. Several factors should be considered when making this decision:

1. *National models provide applicable criteria*.
   The criteria provided by the national criterion models apply to most lakes and reservoirs of the United States. The criterion models estimate relationships between measurements (e.g., Chl *a* and microcystin) that do not vary strongly with geographic location, and therefore, incorporation of state data may not appreciably change criterion values.
2. *Assess the inventory of observed data*.
   The measurements that are available in a state or tribal dataset should be compared with the variables that are included in the national models. Measurements of endpoint variables (i.e., microcystin, depth-averaged DO, zooplankton biomass, TP, and TN) and measurements of variables that account for large proportions of variability in these endpoints (i.e., Chl *a*, DOC, phytoplankton biomass, turbidity) need to be available for combining state and national data.
3. *Consider the likelihood of changes to national models*.
   Several factors determine the degree to which state data influence national relationships. Of these, two can be evaluated *a priori*. First, the number of samples collected for each parameter in the state data set should be no less than approximately 10% of the size of the national dataset. Second, relationships among variables can be more or less precise relative to the national data, depending on factors such as

differences in sampling protocols; however, state data that feature more precise relationships among variables of interest are more likely to influence the national model.

4. *Consult with EPA on statistical techniques for incorporating local monitoring data with national models*.

   Due to wide variations in the types of data that are collected, sampling protocols, and database formats, the process of combining state and national data is unique for each dataset. Therefore, EPA is ready to provide technical assistance to any state or authorized tribe that is interested in combining their data with the national criterion models as resources allow. Please contact your regional nutrient coordinator to request technical support through EPA's Nutrient Scientific Technical Exchange Partnership and Support (N-STEPS) program.