Waterbody Assessments Rule Package: Technical Support Document

Revisions to Ch. NR 102, Wis. Adm. Code (Rule No. WY-23-13)

- Water quality standards updates to dissolved oxygen and oxythermal criteria
- Waterbody assessment methods, including biological assessment thresholds and phosphorus response indicators

Revised to reflect Germane Modifications to the rule: 08-30-2021

Contributors:

Tim Asplund
Ashley Beranek
Matt Diebel
Paul Garrison
Joanna Griffin
Francisco Guerrero-Bolano
Elizabeth Haber
Mark Hazuga
Catherine Hein
Lisa Helmuth
Kevin Kirsch
Gina LaLiberte
Aaron Larson
Johanna Lathrop
Alex Latzka
John Lyons
Cory McDonald
Alison Mikulyuk
Kristi Minahan
Amanda Minks
Dan Oele
Pat Oldenburg
Mike Shupryt
Tim Simonson
Kelly Wagner
Brian Weigel
Marcia Willhite
Sarah Yang

1. Introduction

1.1 OVERVIEW

This Technical Support Document covers rule package WY-23-13, related to Waterbody Assessments, including biological assessment thresholds and phosphorus response indicators, as well as updates to dissolved oxygen-related criteria. This rule package addresses several areas related to the state's assessments of its streams, rivers, lakes and other waterbodies. It focuses largely on assessments related to the biological quality of a waterbody. The main portions of the rule package are summarized here.

Waterbody Assessments and Reporting. Every two years, under federal Clean Water Act requirements, the department assesses the state's waterbodies to determine whether they are attaining water quality standards. A new Subchapter III is proposed that codifies Wisconsin's current procedures for conducting surface water impairment assessments, including public participation opportunities and EPA approval.

Biological assessment thresholds. The most direct and commonly-applied method of measuring the quality of a waterbody is through assessing the biological communities within the waterbody—its fish, insects, plants, and algae. The proposed rule establishes biological assessment thresholds that are used to evaluate the biological health of surface waters in the state. The proposed "Waterbody Assessments and Reporting" subchapter includes the following sections related to biological assessments:

- *Narrative biological assessment thresholds.* Narrative thresholds set expectations and goals for the biological quality of these communities. They are used to measure the quality of a waterbody's biological community and to determine attainment of its designated uses. This section also generally describes the types of biological assessments that have been conducted by the department to determine whether a waterbody's aquatic community is considered healthy and attaining its designated uses or is not attaining its designated uses and should be placed on the impaired waters list (section 303 (d) list).

- *Numeric biological assessment thresholds for lakes, reservoirs, and impounded flowing waters.* Numeric thresholds set benchmarks that indicate attainment of a lake or reservoir's designated uses. Once a numeric biological assessment threshold is codified by rule, it cannot be revised unless the rule is revised. These thresholds include:
    - *Algae thresholds for Recreation and Aquatic Life.* The rule proposes algae (chlorophyll *a*) thresholds for lakes, reservoirs and impounded flowing waters. Algae levels are a top water quality concern for the public, and are a critical component of waterbody assessments to determine whether recreational goals are met. The chlorophyll *a* thresholds created in the proposed rule are the same considerations that have been used by the department to assess water quality for recreation and aquatic life uses. Two types of algae thresholds are created: the threshold to protect aquatic life uses is based on chlorophyll *a* concentrations alone, while the thresholds to protect recreation are based on the frequency of moderate algal levels, which combines a chlorophyll *a* concentration threshold with the number of days exceeding that threshold.
    - *Aquatic plant thresholds for aquatic life.* The rule includes numeric thresholds for aquatic plants in lakes and reservoirs. These thresholds indicate attainment of healthy plant communities within lakes, an important factor in lake habitat to support aquatic life.

Dissolved oxygen criteria for Aquatic Life. Revisions to the dissolved oxygen section are needed to clarify which criteria apply to different waterbody types:

- This rule specifies for which waterbodies and at which times the more protective dissolved oxygen criterion of 7.0 mg/L applies to protect fish early life stages that require higher oxygen levels. It specifies which other DO criteria apply to other waters and other time frames. Certain dissolved oxygen criteria are also relocated from ch. NR 104 to s. NR 102.04(4), so that all dissolved oxygen criteria are located in the same part of the code.
- This rule creates oxythermal criteria for two-story fisheries. These new criteria are necessary because the existing dissolved oxygen criteria are not appropriate for this relatively rare and sensitive type of coldwater fishery, comprising only .01% of Wisconsin's lakes.

Phosphorus assessment procedures using biological metrics. Statewide phosphorus criteria were promulgated in 2010. However, the criteria did not include evaluation procedures for determining attainment of the phosphorus criteria in a waterbody. This rule specifies how attainment of the phosphorus criteria is determined. It also incorporates flexibility for determining impairment due to phosphorus levels by creating a "combined assessment" approach. Under this approach, the waterbody's phosphorus concentration is reviewed in conjunction with "phosphorus response indicators"—algae and plant metrics—that specifically indicate whether the waterbody is exhibiting a biological response to phosphorus. If a waterbody exceeds the statewide phosphorus criterion (within a specified range) but does not exhibit a biological or recreational use impairment, it would not be considered impaired for purposes of section 303(d) listing.

NR 217 calculation of upstream background phosphorus concentrations. This rule includes a revision to a portion of ch. NR 217 to align the phosphorus calculation methods used to determine background phosphorus concentrations for effluent limit calculations with those delineated in proposed s. NR 102.07 (1) (a) 2. Previously, slightly different methods were used to calculate ambient phosphorus concentrations for purposes of criteria assessment and to calculate upstream background phosphorus concentrations for WPDES permit limit derivation under s. NR 217.13 (2) (d). Although these two methods yield very similar resulting phosphorus concentrations, the differences between the two methods have caused confusion and are unnecessary. The proposed procedure detailed in s. NR 102.07 (1) (a) 2, which is the method used for criteria assessment, parallels how the criteria were initially developed and will be most appropriate for both applications.

Relation of this rule to Site-Specific Criteria for Phosphorus (WT-17-12)

This rule package ties into a second rule package that is concurrently underway (WT-17-12) which creates a new chapter NR 119. The proposed ch. NR 119 establishes standard protocols for developing site-specific criteria (SSC) for phosphorus in cases where the current statewide phosphorus criteria may be over- or under-protective of a waterbody's designated uses. Development of SSC ties directly to the ability to demonstrate support of a waterbody's phosphorus response indicators and biological assessment thresholds, contained in rule package WY-23-13 and described in this document. This Technical Support Document provides a brief overview of how this rule relates to the SSC rule. The SSC rule itself does not require a Technical Support Document as it establishes a process rather than a water quality standard. Any SSC developed using that process will have its own Technical Support Document and will be evaluated for approval by EPA.

1.2 CHANGE LOG

For a quick synopsis and explanation of changes to existing codes, a "Change Log" is provided here (Figure 1). Readers may wish to refer to this resource while reviewing proposed rule revisions. It may be particularly helpful for areas of the code in which minor revisions are proposed that are not covered as part of the text of this Technical Support Document.

Figure 1. Code revisions and explanations under rule package WT-17-12, related to waterbody assessments, including biological assessment thresholds and phosphorus response indicators, and updates to dissolved oxygen-related criteria.

Code Reference	Revisions and Explanations
102.03 Definitions	Added definitions for the following: benthic, biological assessment threshold, chlorophyll *a*, Clean Water Act, confidence interval, diatom, drainage lake (relocated), impounded flowing water, macrophyte, reservoir (relocated), Section 303(d) list, seepage lake (relocated), stratified two-story fishery lake (relocated), total phosphorus (relocated).
102.04(4) Criteria for fish and aquatic life	Revised existing language on dissolved oxygen (DO) criteria as follows: (a)1 to (a)7 Specified more clearly which DO criteria apply to each waterbody type, and clarify that cold DO criteria apply to all waters where coldwater species are present, rather than waters listed in the 1980 trout book. See section 2.1 in the Technical Support Document. (am) Created a new dissolved oxygen and habitat quantity criterion for two-story fisheries, since existing DO criteria are not appropriate for supporting these waters. See section 2.2 in the Technical Support Document for a detailed description of this paragraph. (b) Repealed, as it is incorporated into (a). (d) Split into two paragraphs, (d) for toxic substances and (f) (created) for other criteria. (f) (see (d) above)
102.04(5)(b) Recreational Use. Exceptions	Updated a reference to another portion of code, necessitated by restructuring.
102.06 Phosphorus	(1) Added a reference to phosphorus assessment procedures in 102.07. (2) Revised definitions for greater clarity for stratified lake or reservoir and stratified two-story fishery lake. Relocated several definitions to 102.03 as they are also applicable to other parts of ch. NR 102. (3) and (4) Relocated phosphorus criteria for impounded flowing waters to sub. (3) with rivers and streams, since determination of the applicable P criterion for an impounded flowing water is dependent on whether it is located on a stream or on a river. (7) Repealed the note as it is now replaced by new 102.07.
102.07 Assessing phosphorus concentration	Established protocols for assessing against the phosphorus criteria. See section 5.1 of the Technical Support Document. (1) and (2) Established general assessment procedures such as data requirements and exceedance calculation methods.
Subch. III (102.50 to 102.60) Waterbody Assessments and Reporting	102.50 to 102.53 Created to outline the department's obligations under the Clean Water Act to conduct biennial assessments, which was not previously addressed in code. These sections are provided for clarity; they do not create new obligations for the department or regulated public. See section 3 of the Technical Support Document for a description of these sections.
102.54	Provided general information about biological assessment of designated uses.
102.55	Established narrative biological assessment thresholds for determining attainment of aquatic life uses. These codify general expectations and goals for the health of a waterbody's aquatic life community. This section also describes the types of

	assessment tools that are used for such determinations. See section 4.3 of the Technical Support Document for a description of this section.
102.56	Established numeric biological assessment thresholds for lakes, reservoirs, and impounded flowing waters. (1)(a) Established aquatic life use assessment thresholds for algae (as measured by chlorophyll *a*). See section 4.4.2 of the Technical Support Document for a description of this subsection. (1)(b) Established aquatic life use assessment thresholds for a lake or reservoir's aquatic plant community, based on response to overall disturbance. See section 4.4.3 of the Technical Support Document for a description of this subsection. (2) Established recreation use assessment thresholds for frequency of moderate algae levels for lakes, rivers, and impounded flowing waters. See section 4.4.1 in the Technical Support Document for a description of this subsection.
102.60	Established a combined approach for assessing attainment of phosphorus criteria. This approach creates phosphorus response indicators for streams, rivers, and lakes. Phosphorus response indicators are used in conjunction with phosphorus criteria to make impairment determinations. (2) Established lake and reservoir phosphorus response indicators for frequency of moderate algae levels, chlorophyll *a* concentrations, aquatic plant community response, and oxythermal layer thickness. See section 5.4 of the Technical Support Document. (3) Established a river and impounded flowing waters phosphorus response indicator for chlorophyll *a* concentration. See sections 5.5 and 5.6 of the Technical Support Document. (4) Established stream phosphorus response indicators for benthic algal mass and benthic diatom community response. See section 5.7 of the Technical Support Document.
217.13(2)(d)	Aligns methods used to calculate upstream background phosphorus concentrations for WPDES permit limit derivation in NR 217 with those used to calculate ambient phosphorus concentrations for purposes of criteria assessment in NR 102.07.

1.3 RULEMAKING AUTHORITY

This rule package is related to water quality standards and assessment methods for determining attainment of water quality standards. Primarily, this rule documents assessment methods that the department uses to determine attainment of a waterbody's designated uses, which are one component of a state's water quality standards. The rule also pertains to certain water quality criteria, which are another component of water quality standards. Specifically, it updates the existing dissolved oxygen criteria and establishes oxythermal criteria.

Water quality standards include "a designated use or uses for the waters of the United States and water quality criteria for such waters based on such uses" per 40 CFR 131.3(i). Designated uses describe the way Wisconsin intends for its waters to be used. Criteria are numeric or narrative statements of the quality of water that must be present to support designated uses. Wisconsin waters are each assigned four designated uses: Fish and Aquatic Life (often shortened to Aquatic Life), Recreation, Public Health and Welfare, and Wildlife.

The majority of this rule package provides assessment protocols for how the department determines attainment of designated uses. Unlike designated uses or criteria themselves, assessment methods are not federally considered to be part of water quality standards. However, they are an important component of all states' and tribes' procedures for evaluating the attainment of water quality standards for purposes of reporting on the health of waterbodies to U.S. EPA. Most states include these assessment protocols in guidance. However, Wisconsin is proposing to establish them in rule for consistency and clarity. To this end, the rule package establishes biological assessment thresholds and phosphorus response indicators to assess whether a waterbody's uses are being attained, particularly the aquatic life and recreation uses.

Additionally, the rule package adds language that clarifies Wisconsin's obligation under the Clean Water Act (CWA) to conduct waterbody assessments using the designated uses and water quality criteria every two years and to determine which waterbodies are not meeting water quality standards. It also revises existing water quality criteria for dissolved oxygen and establishes oxythermal criteria for lakes with coldwater fish.

U.S. EPA delegates the state of Wisconsin the authority and responsibility for creating and updating water quality standards, and for monitoring and assessing waterbodies. Relevant portions of the Clean Water Act and the Code of Federal Register (CFR) include the following:

Sec. 303 (d) (1) (A) of the Federal Water Pollution Control Act (Clean Water Act) requires states to develop an impaired waters list that identifies waters that are not meeting any water quality standard.

Sec. 305 (b) (1) of the Federal Water Pollution Control Act (Clean Water Act) requires states to prepare a biennial report documenting which waterbodies are attaining their designated uses.

40 CFR § 130.3 Water quality standards, defines water quality standards as setting water quality goals for a waterbody that will protect its designated uses (such as protection of fish, wildlife, recreation, and public health and welfare). Criteria will be set to protect those uses.

40 CFR § 130.4 Water Quality Monitoring, requires water quality monitoring and assessments of state waters.

40 CFR § 130.7 Total maximum daily loads (TMDLs) and individual water quality-based effluent limitations, provides additional information related to requirements for developing the impaired waters list.

40 CFR § 130.8 Water Quality Reports, requires states to submit water quality reports to EPA that include a water quality assessment of state waters.

40 CFR § 131.4 State authority. (a) States (as defined in §131.3) are responsible for reviewing, establishing, and revising water quality standards. As recognized by section 510 of the Clean Water Act, States may develop water quality standards more stringent than required by this regulation.

40 CFR § 131.11 Criteria. (a) (1) States must adopt those water quality criteria that protect the designated use. Such criteria must be based on sound scientific rationale and must contain sufficient parameters or constituents to protect the designated use. For waters with multiple use designations, the criteria shall support the most sensitive use.

Wisconsin statutes contain the following authority for the state to promulgate water quality standards and conduct assessments of waterbody condition:

Wis. Stats. § 281.11 and 281.12 grant necessary powers and establish a comprehensive program under the WDNR to enhance quality management and protection of all waters of the state. It grants the WDNR general supervision and control to carry out the planning, management and regulatory programs necessary for prevention/reduction of water pollution and for improvement of water quality.

Wis. Stat. § 281.13 grants the department authority to research and evaluate the quality and condition of the state's natural water sources.

Wis. Stat. § 281.15 mandates that the department promulgate water quality standards, including water quality criteria and designated uses. It recognizes that different use categories and criteria are appropriate for different types of waterbodies, and that the department shall establish criteria which are not more stringent than reasonably necessary to assure attainment of the designated use for the water bodies in question.

Wis. Stat. § 281.65(4)(c) and (cd) directs the department to prepare a list of waters impaired by nonpoint source pollution.

Any updates to Wisconsin's water quality standards need to be reviewed by U.S. EPA. EPA is required by section 303(c)(3) of the CWA and 40 CFR 131.21 to review new or revised water quality standards to determine whether they are consistent with the CWA and 40 CFR Part 131. EPA's review of water quality standards involves a determination of whether the state has adopted criteria that protect the designated use, and whether the state has followed its legal procedures for revising or adopting standards.

This rule largely contains assessment methods rather than water quality standards. U.S. EPA is not required to review and approve a state's assessment methods for evaluating attainment of water quality standards. Further, Wisconsin Statutes do not require a Technical Support Document for rulemaking pertaining to assessment methods. However, to provide a comprehensive record of information pertaining to the assessment methods contained in this rule, they are included in this document.

2. Water Quality Criteria Updates

This rule package contains updates to the existing dissolved oxygen criteria and creates new oxythermal criteria for a special class of ~180 lakes that contain coldwater fish, called two-story fishery lakes.

2.1 DISSOLVED OXYGEN

The existing dissolved oxygen (DO) rule language specifies that it applies to trout waters; however, non-trout coldwater species also require higher DO and therefore the language is being adjusted to cover non-trout cold waters under the Cold DO criterion. As such, the existing DO criterion for cold waters will be applied to the following Designated Uses: Cold streams (which includes further subcategories called "Natural Communities" of Cold and Cold Transition Headwater and Cold Transition Mainstem streams), Cold lakes except for two-story fishery lakes (see below), and Great Lakes. Language was also adjusted to protect early life stages of fish until they leave their gravel nests, beyond the fall spawning season. These updates support the U.S. EPA's national recommendation that cold water DO criteria apply to any cold water systems for which higher DO is necessary:

> **U.S. EPA's National Recommended Ambient WQC for DO:** *Criteria for coldwater fish are intended to apply to waters containing a population of one or more species in the family Salmonidae or to waters containing other coldwater or coolwater fish deemed by the user to be closer to salmonids in sensitivity than to most warmwater species. Some coolwater species may require more protection than that afforded by the other life stage criteria for warmwater fish and it may be desirable to protect sensitive coolwater species with the coldwater criteria.*
> (U.S. EPA Ambient Water Quality Criteria for Dissolved Oxygen (EPA 440/5-60-003), April 1986)

In the existing code, the coldwater DO criteria of 7 mg/L to protect spawning also applied to class III trout waters. However, this is proposed for removal because class III trout waters do not have naturally reproducing trout.

The existing DO criterion of 5 mg/L will apply to warmwater streams and rivers, and all lakes other than those specified above. This is consistent with the status quo.

This rule relocates certain dissolved oxygen criteria from ch. NR 104 to s. NR 102.04(4), so that all DO criteria are located in the same part of the code in ch. NR 102. The relocated criteria are the existing dissolved oxygen criterion of 3 mg/L for limited forage fish waters and 1 mg/L for limited aquatic life waters, diffuse surface waters, and wastewater effluent channels.

Existing introductory language in NR 102.04(4) provides an exception for natural conditions. This exception may be applied to waters with either higher or lower than typical natural DO concentrations. Note that natural conditions for DO fluctuate over a 24 hour diurnal cycle, with highest DO in late afternoon, and lowest just before dawn.

2.2 OXYTHERMAL HABITAT FOR TWO-STORY FISHERY LAKES

Two-story fishery lakes have coldwater fish species, but their requirements differ significantly from coldwater streams and from other lakes. Because coldwater fishes have specific DO and temperature needs that occur within a narrow vertical habitat range, their criteria combine oxygen and temperature measurements and are called oxythermal layer criteria.

Definition of Two-Story Fishery Lake

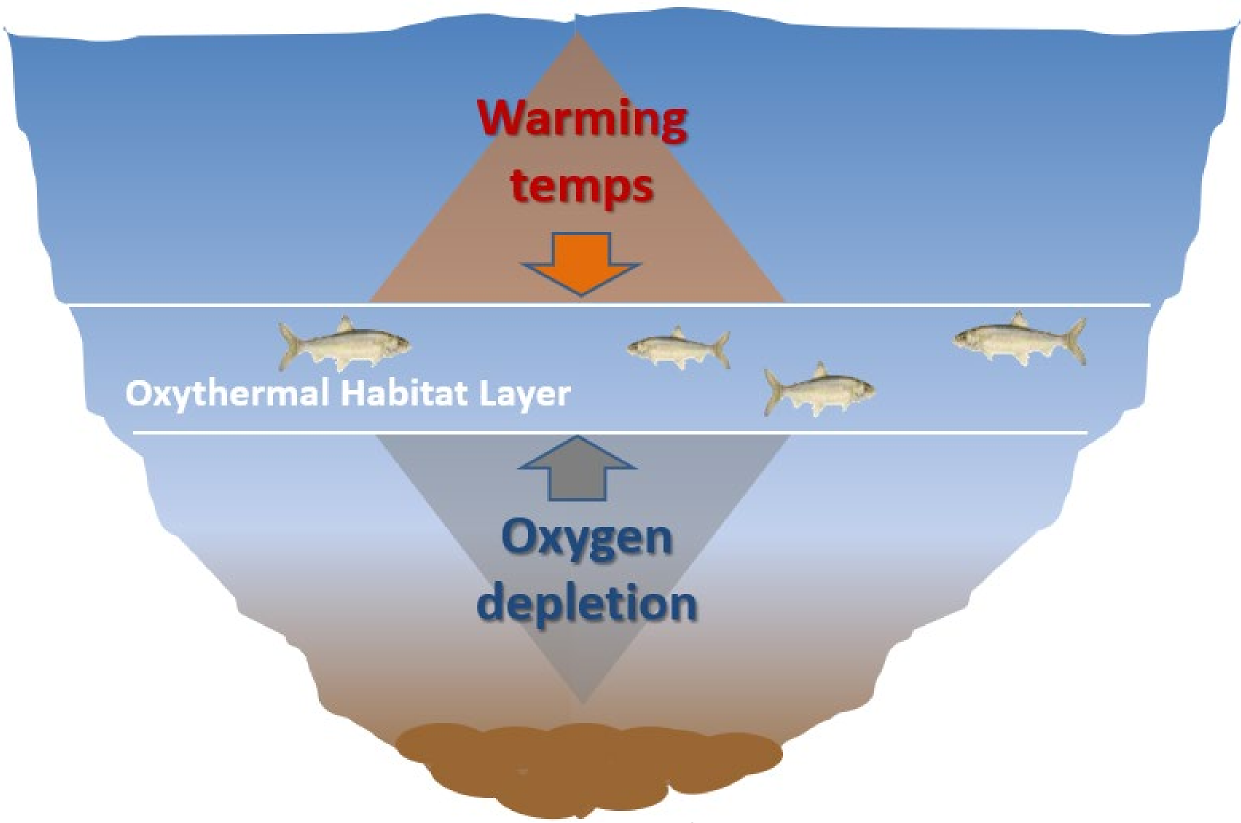
A two-story fishery lake is defined in the code as a lake greater than 5 acres in size that is always stratified in the summer, with the potential for an oxygenated hypolimnion, that has documentation since 1975 of a population of cold water fish species such as cisco, whitefish, or trout that is sustained through natural reproduction or long-term active stocking with year-to-year survival. This definition is revised slightly from the existing definition to provide additional clarity.

Habitat quantity concepts for two-story fishery lakes

For most cold waters of the state, the dissolved oxygen metric used to determine support of the fishery is a DO concentration of 6 mg/l. However, for two-story fisheries, a DO concentration alone is not the best way to represent the habitat characteristics needed to support the fishery. Cisco, whitefish and other coldwater fishes need a band of water that has both cold enough temperatures and high enough oxygen for them to survive. At the beginning of summer, the entire water column usually has both, but by the end of summer, temperatures in the surface water may be too warm and the DO near the bottom may be too low, squeezing the fish into a narrow band along the thermocline where they can survive (Figure 2). Therefore, a measure that represents the overall quantity of suitable habitat by combining both DO and temperature is a more useful metric for assessing support of the two-story fishery.

The concepts and methods used in the oxythermal layer approach are described in Lyons, et. al (2017), in which cisco data from Wisconsin two-story fishery lakes are analyzed and a metric called "cisco layer thickness" is developed. The same methods were then applied to whitefish and lake trout to develop a set of criteria that are suitable for all three coldwater species found in Wisconsin's two-story fishery lakes. Earlier work done in Minnesota developed a similar measurement called "TDO3" for their two-story fishery lakes. TDO3 is a vertical measurement of the water temperature (T) at which the dissolved oxygen (DO) concentration is 3.0 mg/l. WDNR's work built from that concept and applied a variation of this method whereby a certain *quantity* of habitat (at least 1 meter of depth) is required which attains an appropriate DO (WDNR used 6 mg/L rather than 3) and temperature. This has the advantage of requiring a certain depth of habitat to ensure survival, rather than establishing a single point at which the criterion must be met as in the TDO3 approach.

Figure 2. In late summer, coldwater fish can live only in the band in which there is sufficient dissolved oxygen and cool enough temperatures, termed the oxythermal layer.



Calculating species' DO and thermal requirements

Each fish species has specific oxygen and thermal ranges suitable to its survival. WDNR assessed species information is from recent (2011-2015) data from a majority (~155) of Wisconsin's two-story fishery lakes combined with research done in Minnesota. The oxythermal criterion for each species is based on the species' upper temperature limit and a protective DO limit of 6 mg/L. The data assessed indicated the following:

- Cisco, whitefish, and lake trout can survive oxygen levels between 3-5 mg/L, but this level is sub-optimal and may reduce growth and survival. A DO of 3 mg/L is their lower oxygen limit for a 24-hour period. WDNR selected a minimum DO of 6 mg/L for the oxythermal criteria as a level that would sustain coldwater fish populations, consistent with the DO criteria of 6 mg/L for other coldwaters. Minnesota has used a DO concentration of 3 mg/L for their two-story fishery metric, but because this type of criteria is to be applied particularly at periods of peak stress and the thermal component of the criteria is based on the species' upper temperature limits, it was determined that a more protective oxygen level was needed to maintain a healthy population and prevent fish kills.
- The upper temperature limit for cisco is 73°F (22.8°C) (i.e., cisco will begin to die if exposed to temperatures above this limit for more than a few days). Their ideal range is ~39-63°F (~4-17°C), with an optimal temperature of ~48°F (9°C) (i.e. when given a choice, most cisco are found at this temperature if the DO is above 3 mg/L).
- The upper temperature limit for whitefish is ~66°F (~19°C). Their ideal range is ~39-52°F (~4-11°C), with an optimal temperature of ~39°F (~4°C).
- The upper temperature limit for lake trout is 57°F (14°C). Their ideal range is ~39-50°F (~4-10°C).

Setting criteria values

A lake's maximum acceptable temperature and acceptable DO are set based on the species that are expected to be present and reproducing within that lake. In the proposed rule language below, the temperatures shown in subdivision paragraphs a to d represent the maximum acceptable temperature at a DO of 6 mg/L for the species indicated. The thermal thresholds were selected to represent the point at which mortality begins to occur but most fish survive. In addition to appropriate temperature and DO characteristics, a two-story lake must have a minimum quantity of suitable habitat: a band of water of at least 1 meter that is at/below the indicated temperature and at/above the indicated DO. For example, a lake that has just an inch of water within the suitable ranges would not support coldwater species and would be considered impaired.

NR 102.04 (4) (am) *Oxythermal layer thickness for two-story fishery lakes.* 1. 'Criteria.' A two-story fishery lake shall maintain, during its period of summer stratification, an oxythermal layer of at least 1 meter in thickness that maintains both a dissolved oxygen concentration of at least 6 mg/L and a maximum temperature of the following:

a. For a two-story fishery lake with lake trout, 57° F or less.

b. For a two-story fishery lake with whitefish but not lake trout, 66° F or less.

c. For a two-story fishery lake with cisco but not whitefish or lake trout, or that the department manages for brook, brown, or rainbow trout, 73°F or less.

d. For a two-story fishery lake with multiple coldwater fish species, the applicable criterion under a. to c. is that for the lake's species requiring the lowest temperature.

Assessing attainment of the oxythermal layer criteria

To measure a lake's available volume of habitat, vertical temperature and DO profiles are taken in the deep part of the lake while the lake is stratified, at least monthly from July to September (earlier samples may be useful). Multiple profiles are typically needed to account for variability, both during the summer season and across years. A minimum of three years is recommended. To analyze, plots are made of both temperature vs. depth and DO vs. depth, and the vertical extent of the depth profile is determined at which the DO is 6 mg/L or above and the temperature is at the specified threshold or below. The depth of available habitat is then compared to the criterion. During any given year, if at any point the applicable criterion is not met, that year is an exceedance year. If any two or more years within the most recent 10-year period are exceedance years, the lake is not attaining the water quality criterion and would be listed as impaired on the section 303(d) list.

3. Waterbody Assessments

3.1 DOCUMENTING CLEAN WATER ACT OBLIGATIONS

Under the Clean Water Act, all states are required to conduct waterbody assessments and impaired waters listing; these are submitted to the U.S. EPA every two years. However, Wisconsin codes do not currently contain any reference to these obligations. The rule includes a newly proposed Subchapter III, Waterbody Assessments and Reporting. The first few sections of this subchapter (ss. NR 102.50 to 102.53) document and codify Wisconsin's assessment and listing process, in a generalized manner. They describe two specific types of assessments that are required under the Clean Water Act: statewide condition assessments and individual waterbody assessments (including the 303(d) list). They also establish requirements for public participation and recognize the U.S. EPA's approval process. The documentation of these general obligations is not meant to necessitate any specific changes to how these assessments are currently conducted.

The remaining portions of proposed Subchapter III (beginning at s. NR 102.54) include more specific information about assessment protocols, and these are described in detail throughout sections 4 and 5 of this Technical Support Document. The department's protocols for assessing waterbodies and listing impaired waters are supplemented with a guidance document titled "*Wisconsin's Consolidated Assessment and Listing Methodology*" (WisCALM), which is updated every two years. This guidance document would still be used for more detailed protocols than those that are codified.

3.1.1 Variability and confidence intervals

This subchapter also contains a section on sample variability. For certain types of assessments, a site may exhibit a wide variability in samples collected. The subchapter establishes a process for using confidence intervals to determine when additional samples are needed to bolster the dataset before making an impairment determination.

This use of confidence intervals (CI) is a statistical approach to assess stream data against the applicable water quality criterion or assessment threshold. Within this rule package, it is applied to both phosphorus and chlorophyll *a* assessments, and is available to use with other parameters as appropriate. Use of an 80% CI has several benefits, including:

- It clarifies the confidence in the mean or median value of a small number of samples;
- It maximizes sampling efficiency, only requiring additional samples if the results are unclear (a small percentage of the time);
- It reduces both false positives and false negatives in decision making;
- It is consistently applied for assessments across parameters (for both total phosphorus and chlorophyll *a*) and across waterbody types (for lakes, streams, and rivers); and
- It is currently in use as part of WisCALM assessment protocols and is automatically calculated.

The CI approach involves the calculation of a two-sided 80% confidence interval around the mean (for lakes) or median (for streams) of a sample dataset. The confidence interval is calculated using measures of sample size and variation to suggest, with a specified level of certainty, that the true population statistic (e.g., mean or median) falls within a specified range of values. When sample values are normally distributed, the confidence interval around the median is identical to the confidence interval around the mean. Because phosphorus and chlorophyll *a* concentrations are usually log-normally distributed, the raw concentrations for these metrics are log-transformed for the confidence interval calculation.

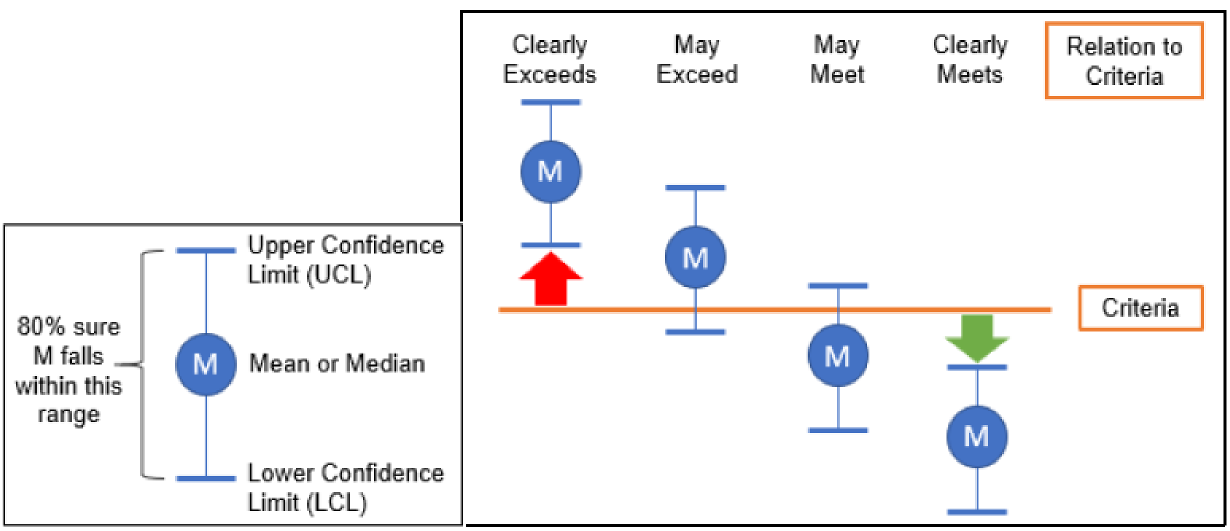
The two-sided CI has an upper and lower confidence limit (CL). The upper and lower CL are used to determine if more data are needed before making an assessment determination, as follows (Figure 3):

- If the upper CL is below the criterion, the sample dataset clearly attains the criterion/threshold. No further samples are needed.
- If the lower CL of the sample dataset from a particular site exceeds the applicable criterion/threshold, and those data were representative of normal weather and hydrology, then the corresponding site/segment is considered to be exceeding the criterion/threshold. No further samples are needed.
- If the criterion/threshold falls within the confidence interval, then more samples are needed before making an attainment determination. Typically an additional year of sampling is done to increase certainty; after that point, if results are still unclear the attainment decision is based on whether the mean or median is above or below the criterion/threshold.

By comparing only one side of the 80% CI to a criterion/threshold it provides 90% certainty that the true mean or median is above the threshold (see the "Clearly Exceeds" example in Figure 3). This is because if the 80% CI is above the criterion/threshold then the 10% uncertainty that is greater than the highest confidence limit is also above the criterion/threshold, and summing the two equals 90% confidence. Likewise, in the "Clearly Meets" example in Figure 3, there is 90% certainty that the true mean or median is below the criterion/threshold.

WisCALM 2020, section 4.5, contains the formula for the CI calculation and additional information on how the CI is used.

Figure 3. Comparison of the upper and lower confidence limit values and mean/median (M) to a criterion or threshold, to determine if additional samples are needed for an assessment determination.



4. Biological Assessment Thresholds: Assessing Overall Health

WDNR's proposed rule revisions include a new Subchapter III of NR 102 Wis. Adm. Code that will establish narrative biological assessment thresholds to be used for waterbody assessments and 303(d) listing. This code package includes several types of biologically-based assessment thresholds, which indicate the overall health of certain groups of organisms. These are expressed both in narrative format, pertaining to all waterbodies, and as numeric thresholds specific to lakes, reservoirs, and impounded flowing waters. This section focuses on biological assessment thresholds for assessing overall health. Phosphorus response indicators are a different set of biologically-based indicators discussed in section 5 that are designed to reflect impacts specifically tied to phosphorus.

Most state environmental agencies assess biological quality in their waterbodies. This can be done under several frameworks:

a) Biological thresholds used for assessment purposes can be either *narrative* or *numeric* in format, or both.
b) They may be contained in *guidance* or *promulgated* in a state's Administrative Codes (or equivalent). Some thresholds may be promulgated while others remain in guidance.
c) If promulgated, they may be codified as part of the state's *assessment methodology*, or as part of its *water quality standards* as biological criteria (biocriteria).

EPA has been working with states, including Wisconsin, for more than a decade to develop and codify biological assessment thresholds or biocriteria that fit each state's preferred framework. From the options above, Wisconsin has elected to:

a) Establish both narrative and numeric thresholds.
b) Promulgate narrative thresholds and a subset of numeric thresholds in Administrative Code, supported by additional guidance in WisCALM.
c) Establish these thresholds as part of its assessment methodology rather than as water quality standards (or criteria).

More succinctly, **Wisconsin has elected to codify narrative and numeric thresholds as part of its assessment methodology rather than as part of its water quality criteria.**

Wisconsin, like many state regulatory agencies, has a long history of using biological data to support water quality management (U.S. EPA 2011). The biological assessment thresholds presented here were developed using Wisconsin-specific data and statistical approaches appropriate for these datasets.

4.1 HOW BIOLOGICAL ASSESSMENT THRESHOLDS DIFFER FROM WATER QUALITY CRITERIA

Wisconsin has elected to propose codification of biological assessment thresholds that are separate from, but related to, its water quality criteria (Table 1). During the department's biennial assessment cycle, both water quality criteria and biological assessments are used to evaluate whether a waterbody is attaining its designated uses. This information is reported to the U.S. EPA during the biennial reporting cycle.

- **Water quality criteria** describe the quality and characteristics of the water contained within a surface waterbody. Most water quality criteria are for pollutants such as toxics or nutrients, but a few, such as dissolved oxygen (DO) or pH, describe water quality but are not pollutants. Only water quality criteria are used to set discharge permit limits or to set targets for Total Maximum Daily Load (TMDL) analyses.

- **Biological assessment thresholds** describe the condition of the living things within the waterbody, such as plants, fish, aquatic insects, and algae. They are used to determine the health of an aquatic life community and whether designated uses are supported. Aquatic life communities may be impacted by pollutants or by other factors such as physical impacts (stream bank erosion, dams), invasive species or climate change. Therefore, there are a wide range of actions that may be taken to address biological degradation, commonly including habitat restoration, watershed work, and invasive species management. Whether biological assessment thresholds are codified or in guidance, or are narrative or numeric, they are not used for setting permit limits.

Table 1. Differences between application of water quality criteria and biological assessment thresholds.

	Used to derive permit limits?	**Examples:**	**Shown on 303(d) list as:**	**Actions toward improvement:**
Water Quality Criteria (describe the water itself)	Yes	Pollutants (toxics, nutrients)	"Pollutant"	- Permit limits - TMDLs
		DO, pH	"Observed effect" of degradation	- Habitat restoration - Watershed work - Invasive species mgmt. - Site-specific criteria
Biological Assessment Thresholds (describe living things)	No	Plants, fish, insects, algae (chl)		

Note: Water quality criteria and biological assessment thresholds can be either numeric or narrative.

4.1.1 Listing "Observed effects" of biological degradation

Under EPA's section 303(d) listing process, once a waterbody is assessed for water quality criteria and/or biological assessment thresholds, the state reports any applicable "pollutants" and/or "observed effects". "Observed effects" have formerly been called "impairments" in Wisconsin; this code package proposes a change in terminology to match EPA's terms. If a waterbody is exceeding a water quality criterion for a pollutant, that is reported as a "pollutant" on the 303(d) list. In contrast, if a waterbody is found to not be attaining any one of its biological assessment metrics, whether in code or guidance, the department includes it on the 303(d) list as having "observed effects" of degradation. These observed effects are specified further as "degraded biological community", "eutrophication", or using other descriptors.

Biological metrics designed to assess overall community health are influenced by a wide range of stressors outside of specific pollutants, such as habitat loss, invasive species, and dams. Therefore, listings for these "observed effects" are not linked to specific pollutants unless a demonstration has been made that a specific pollutant is causing or contributing to the degradation. Often there is no pollutant associated with a biological listing (in which case the data field for "pollutant" is filled as "unknown" or "not applicable"). These listings for biological degradation are not addressed through permit limits, but may be addressed by many other types of restorative actions. To date, after many years of listing and addressing sites exhibiting biological degradation, the department is not aware of any regulatorily-required economic impacts of these listings outside of cases where a site-specific criterion for a pollutant was promulgated through a separate rulemaking process, as described below.

Biological assessments may intersect with water quality criteria in certain cases which are fairly uncommon. If a waterbody's water quality criteria are already being met but the department determines that the waterbody is still experiencing biological degradation due to a specific pollutant, that indicates lower levels of the pollutant are needed. In that case, a site-specific criterion for that pollutant can be proposed via rulemaking to address the degradation. Conversely, a higher criterion for a pollutant may be established for a waterbody via rulemaking if the higher amount would remain protective of the waterbody's uses. This would reset the water quality criterion for a waterbody to a different value, which could in turn affect permit limits or TMDL targets. Setting a site-specific criterion would have to go through its own rulemaking process.

4.2 NARRATIVE VS. NUMERIC BIOLOGICAL ASSESSMENT THRESHOLDS

WDNR has been assessing biological metrics over the past two decades, including fish and aquatic insects in streams and rivers, and aquatic plants and algae in lakes. Assessment protocols were incorporated into the department's *Wisconsin Consolidated Assessment and Listing Methodology* (WisCALM) guidance starting with the 2014 assessment cycle. Because this has been longstanding practice and is an important part of WDNR's work to assess the health of the state's waterways, it is appropriate to reflect this process in state Administrative Code to provide transparency for the public and clarity as to how bioassessments fit within the Clean Water Act. Under the Clean Water Act, biological assessment thresholds can be expressed as either narrative or numeric:

- **Narrative biological assessment thresholds** are a set of descriptive statements in code that express expectations and goals for the quality of aquatic biological communities. They may also provide information about the types of assessments done to evaluate biological health. Typically, narrative assessment thresholds have accompanying guidance, such as WisCALM, describing assessment protocols in more detail.

- **Numeric biological assessment thresholds** are specific numeric benchmarks at which a waterbody is determined to be attaining its aquatic life use. They are typically established for different types of waterbodies and different types of biological communities (plants, fish, insects, etc.) For instance, numeric assessment thresholds may require that a certain type of lake contain a specified (numeric) percentage of aquatic plants that are considered sensitive to disturbance in order to be considered attaining its aquatic life use. Or, for a certain stream type, a numeric threshold might say that the fish community should achieve a score of 40 or greater on a 100 point scale (the Fish Index of Biotic Integrity, or IBI). Other communities may have other types of numeric thresholds. Once a numeric biological assessment threshold is codified by rule, it cannot be revised unless the rule is revised. Numeric thresholds may also be contained in guidance instead of rule, as is the case in many states.

In this rule, the department is proposing to establish a Subchapter III containing both narrative and numeric biological assessment thresholds.

From EPA's website, https://www.epa.gov/wqc/information-bioassessment-and-biocriteria-programs-streams-and-wadeable-rivers-tabulated-format, EPA reports that a total of 38 states currently have some combination of narrative and/or numeric biocriteria promulgated. Of these, 14 states have only a narrative biocriteria statement promulgated, 18 states have a narrative statement with quantitative implementation procedures or translators, and six states have numeric biocriteria as well as narrative biocriteria. Within EPA's Region 5, which Wisconsin is part of, Ohio and Minnesota have both narrative and numeric biocriteria, and Indiana has a narrative biocriteria statement in code. The remaining Region 5 states (Illinois, Michigan) and Iowa do not have promulgated biocriteria as part of their water quality standards. They do, however, use biological assessment thresholds as part of their assessment guidance and methodology for determining whether waterbodies are attaining their designated uses and for making listing determinations.

4.3 NARRATIVE BIOLOGICAL ASSESSMENT THRESHOLDS

4.3.1 Narrative assessment statement

As described above, narrative biological assessment thresholds set goals and expectations for the quality of biological communities, and are used for determining attainment of designated uses. The primary narrative statement in the proposed code is as follows:

> **NR 102.55 (2)** NARRATIVE BIOLOGICAL ASSESSMENT THRESHOLDS. (a) The aquatic life uses under s. NR 102.04 (3), except for those specified in s. NR 102.04 (3) (d) to (e), shall be considered suitable for the protection and propagation of a balanced aquatic life community. Those uses are intended to support the growth, development, reproduction, and life cycle of the aquatic life communities for their designated aquatic life use categories, although such waters may exhibit moderate changes in aquatic life community structure due to loss of some rare native taxa or shifts in relative abundance. In determining attainment of a waterbody's designated uses, the department may compare its biological quality to the range of quality found in similar waterbodies under natural conditions. A waterbody with distinct natural characteristics that result in an aquatic life community different from or less diverse than other waters in the same use category may be considered attaining its aquatic life use if those differences are clearly related to natural characteristics.
>
> (b) A surface water that does not support a balanced aquatic life community as designated under s. NR 102.04 (3) (d) to (e) shall support its highest attainable use given its habitat and potential.
>
> (c) A surface water shall maintain at least the highest biological condition it has achieved since 1975.
>
> **Note:** Paragraphs (b) and (c) reflect federal requirements under 40 CFR s. 131.10 (g), pertaining to highest attainable uses, and 40 CFR s. 131.3 (e), specifying November 28, 1975 as the benchmark date from which to determine "existing uses" for aquatic life.
>
> **Note:** Examples of waterbodies with distinct natural characteristics are wetland-dominated streams, naturally acidic bog lakes, and ephemeral streams with only small areas of short-term refugia. Biological condition assessments should not be conducted during periods when there is insufficient water due to natural conditions to support aquatic life.

This statement reiterates the designated use requirement under NR 102.04(3) that waterbodies support the protection and propagation of a balanced aquatic life community. It then outlines key concepts used to evaluate whether a waterbody's biological communities indicate attainment of the waterbody's designated uses, including:

- a general description of the types of life stages and functions that should be supported;
- recognition that uses may still be attained even with some shifts in aquatic community structure;
- use of the comparison of a waterbody to similar waterbodies under natural conditions;
- allowance for the consideration of natural characteristics of individual sites; and
- inclusion of federal requirements related to attainment of designated uses.

4.3.2 Interpreting the narrative statement using biological metrics in guidance

The narrative statement is followed in code (at sub. NR 102.55(3)) by a description of the types of biological assessments used by the department to assess biological condition, such as IBIs or similar tools. Note that for *numeric assessment thresholds proposed for inclusion in code*, such as those for lakes, information on the assessment protocols to be used is included in the section of code describing each metric and its thresholds.

However, *narrative thresholds* may be interpreted using assessment metrics contained in guidance. This is particularly relevant for waterbody types that do not have codified numeric thresholds. For instance, assessment protocols for using fish or aquatic insects to assess streams and rivers will remain housed in the WisCALM guidance as they previously have been. For those metrics, WisCALM recommendations apply in interpreting the narrative thresholds, but as guidance these recommendations are non-binding.

Narrative biological assessment thresholds are, by their nature, flexible. At any given time or location, the most appropriate assessment protocols for a certain aquatic community would be applied. This means that as assessment protocols are improved over time, the newer protocols would be applied. This is similar to the existing approach where protocols are implemented through WisCALM guidance, where they may evolve over time to reflect the most recent scientific understanding of aquatic communities. WisCALM is updated every two years to incorporate any needed adjustments, and a public comment period is held at the start of each two-year cycle to get public feedback on any proposed revisions. Any revisions or additions to WisCALM would go through WisCALM's regular public comment period.

This section provides a brief description of biological metrics currently in use by WDNR that are not proposed for promulgation as part of this rule package, but would remain in guidance until or unless a future rule package codifies them. Primarily, these are the fish and insect thresholds for streams and rivers. WDNR is not yet proposing to codify these thresholds because it is in the process of reviewing and revising the existing IBI tools for these metrics. It is worth noting that once a bioassessment tool is implemented it rarely undergoes major revisions. For instance, the coldwater stream fish and stream macroinvertebrate biological assessment tools were developed in 1996 and 2003, respectively, and incorporated into the WisCALM protocols for waterbody assessments around 2014. During that time frame, there have not been any major revisions to the protocols. Currently, after many years of use, WDNR is evaluating whether updates to those tools are appropriate based on updated scientific knowledge. This effort may also include consolidating multiple IBI tools into a unified scoring system for simplicity of application. These revisions in guidance may be completed in time for the 2024 assessment cycle. Any proposed changes would be public noticed as part of the WisCALM guidance public notice period for that assessment cycle, which would take place in the fall of 2022. When the updates are complete, WDNR also expects to propose codification of those assessment thresholds for streams and rivers as a separate rule package. Once numeric biological thresholds are promulgated in code, any revisions to the thresholds would also need to be revised in code.

Because each of these tools are well-documented within WisCALM and/or the papers cited, they are not described in depth here. WisCALM can be accessed online at https://dnr.wi.gov/topic/surfacewater/assessments.html.

Biological assessments for streams and rivers

WDNR has a long history of assessing streams and rivers for fish and macroinvertebrates (aquatic insects). Wisconsin has made considerable investments in developing biotic indices for wadeable streams and rivers.

- **Macroinvertebrates (aquatic insects):** Three Macroinvertebrate Indices of Biotic Integrity (MIBI) are tailored to wadeable streams in specific ecoregions of Wisconsin: Driftless Area, Northern Forest, and Central/Southeast (Weigel, 2003). These subcategories are based on landscape-scale characteristics including geology, physiography, vegetation, climate, soils, land use, wildlife, and hydrology. In the Wadeable Stream MIBI, the three MIBIs tailored to each ecoregion are combined into a single Wadeable Stream MIBI scoring scale that is comparable across all ecoregions. Additionally, a separate MIBI is used for rivers (Weigel and Dimick 2011).

- **Fish:** Five Fish IBIs (FIBI) have been developed for wadeable streams based on stream size (flow) and temperature, the waterbody characteristics most closely tied to fish community composition. The five FIBIs are: coldwater (Lyons et al., 1996), cool water transitional (cool-cold and cool-warm; Lyons, 2012), warmwater (Lyons, 1992) and small stream (Lyons, 2006). These correspond to different Natural Community stream subcategories, which are correspondingly based on the temperature and flow of the waterbody. A separate Fish IBI is used for rivers (Lyons et. Al. 2001).

Biological assessments for wetlands
To date, wetland assessments have not been included in WisCALM. However, an extensive effort has been underway to develop wetland plant indicators. These could be incorporated into WisCALM biological assessments in the future if deemed appropriate.

- **Wetland plants:** Assessment tools for wetland plant communities are under development. Surveys of 1,100 least-disturbed to most-disturbed wetlands have been completed and used to set preliminary wetland condition thresholds for each commonly-occurring wetland plant community by Omernik Level 3 Ecoregion. Floristic Quality Assessment metrics calculated include weighted and unweighted mean coefficient of conservatism. When all ecoregions are surveyed, the total dataset could be analyzed to determine statewide assessment thresholds where possible.

4.3.3 WDNR's Bioassessment Program Evaluation

U.S. EPA has been working nationwide since 2002 to assess states' biological assessment and monitoring programs in support of biological thresholds and criteria. During the course of 2013-2014, the WDNR underwent a U.S. EPA program review of its bioassessment program. The review was conducted by a U.S. EPA contractor, Midwest Biodiversity Institute, and assessed a variety of program components including the state's designated uses, monitoring capacity, assessment protocols, and biological indices. The assessment culminated in a report titled "Refining State Water Quality Monitoring Programs and Aquatic Life Uses: Evaluation of the Wisconsin DNR Bioassessment Program" (MBI, 2014). Multiple WDNR program components were scored for several waterbody types: streams/rivers, lakes, and wetlands. The review recognized that WDNR is currently employing several biological indicators through its WisCALM assessment guidance. However, it emphasized EPA's national goal of establishing at least two biological assemblages per waterbody type and strongly recommended that these biological assessment thresholds and protocols be officially promulgated. The Bioassessment Program Review provided a strong endorsement for the necessity of this rule package.

4.4 NUMERIC BIOLOGICAL ASSESSMENT THRESHOLDS FOR LAKES, RESERVOIRS, & IMPOUNDED FLOWING WATERS

Numeric thresholds set benchmarks that the department uses when determining attainment of a waterbody's designated uses. This rule proposes codification of a suite of numeric thresholds for lakes and reservoirs. These include recreation thresholds for frequency of moderate algae levels and Aquatic Life use thresholds for aquatic plants and for chlorophyll *a.* Of these, only the recreation thresholds also apply to impounded flowing waters.

4.4.1 Chlorophyll *a* thresholds for Recreation use assessments

High algal levels in Wisconsin's waterbodies are a top concern for Wisconsin citizens, as evidenced through public surveys and comments received during the department's Triennial Standards Review cycles, where algae-related topics consistently ranks among the public's top priorities. Algal biomass, as measured by chlorophyll *a* concentrations, relates to objectives concerning swimming and other recreational uses of lakes as well as aquatic habitat and trophic state. It is one of the most common response metrics to assess elevated phosphorus and was one of the primary metrics used in development of Wisconsin's phosphorus criteria in 2010. Chlorophyll *a* concentrations have also been used since 2012 as part of the department's standard lake assessment protocols, as detailed in WisCALM 2020. Every two years, the department uses an automated statistical package to assess all lakes in the state with sufficient chlorophyll *a* data.

For recreational uses such as swimming, boating, and aesthetics, the assessment thresholds for chlorophyll *a* are based on the frequency of moderate algae levels, as shown in Table 2. This section describes how moderate algae levels are defined, and the selection of a frequency threshold for different lake types.

Table 2. Recreational use assessment thresholds for frequency of moderate algae levels.

Waterbody Type	Subcategory[1]: Phosphorus subcategories	Recreation Use assessment thresholds
Lakes, Reservoirs, Impounded Flowing Waters (includes cold and warm)	Impounded flowing water, Unstratified drainage, Unstratified seepage	Does not exceed 20 ug/L for more than 30% of days during the summer sampling period[2]
	Stratified drainage, Stratified seepage	Does not exceed 20 ug/L for more than 5% of days during the summer sampling period[2]
	Stratified two-story fishery	

[1] Terms used for waterbody types and subcategories are defined in s. NR 102.03. These thresholds do not apply to streams or rivers.
[2] Summer sampling period is July 15 to September 15.

4.4.1.1 Defining moderate algae levels

Lake recreational chlorophyll *a* assessment thresholds are designed to protect primary contact recreation (swimming). Since 2002, Wisconsin's citizen lake monitoring network has collected over 10,000 chlorophyll *a* samples and corresponding user perception ratings of water quality. We conducted a statistical analysis of the relationship between user perception and chlorophyll *a* concentration to help identify appropriate thresholds. This enabled us to determine a chlorophyll *a* threshold at which conditions decline to an extent that users experience decreased enjoyment, but before substantial numbers would not swim.

Citizen monitors rate the condition of each lake and their enjoyment of it on the day they sample water quality. Citizens do not know the chlorophyll *a* concentration results at the time they rank the lake condition in one of the following five categories:

1 = Beautiful, could not be any nicer
2 = Very minor aesthetic problems; excellent for swimming and boating enjoyment
3 = Swimming and aesthetic enjoyment of lake slightly impaired because of high algae levels
4 = Desire to swim and level of enjoyment of lake substantially reduced because of algae; would not swim, but boating OK
5 = Swimming and aesthetic enjoyment of lake substantially reduced because of algae levels

These rankings were used in conjunction with the chlorophyll *a* data collected on the day of the survey. As described in the statistical analysis section below, we used a logistic regression model to evaluate the relationship between the subjective perception ratings and measured chlorophyll *a* concentration (M. Diebel, WDNR, unpublished analysis, 2016). The analysis shows that subjective perception of water quality is strongly related to measured chlorophyll *a* concentration (Figure 4). We used the results to propose a definition of "moderate algae" levels of $\geq$ 20 ug/L chl *a*.

Proposed chlorophyll *a* assessment thresholds may be determined by a) identifying inflection points in the relationship at which half the users perceive a particular condition, and/or b) specifying a target frequency of lake user perception (e.g., 90% of lake users view the water as suitable for swimming). We based our proposed definition of "moderate algae" at $\geq$ 20 ug/L chl *a* on both of these factors, which provided two main findings:

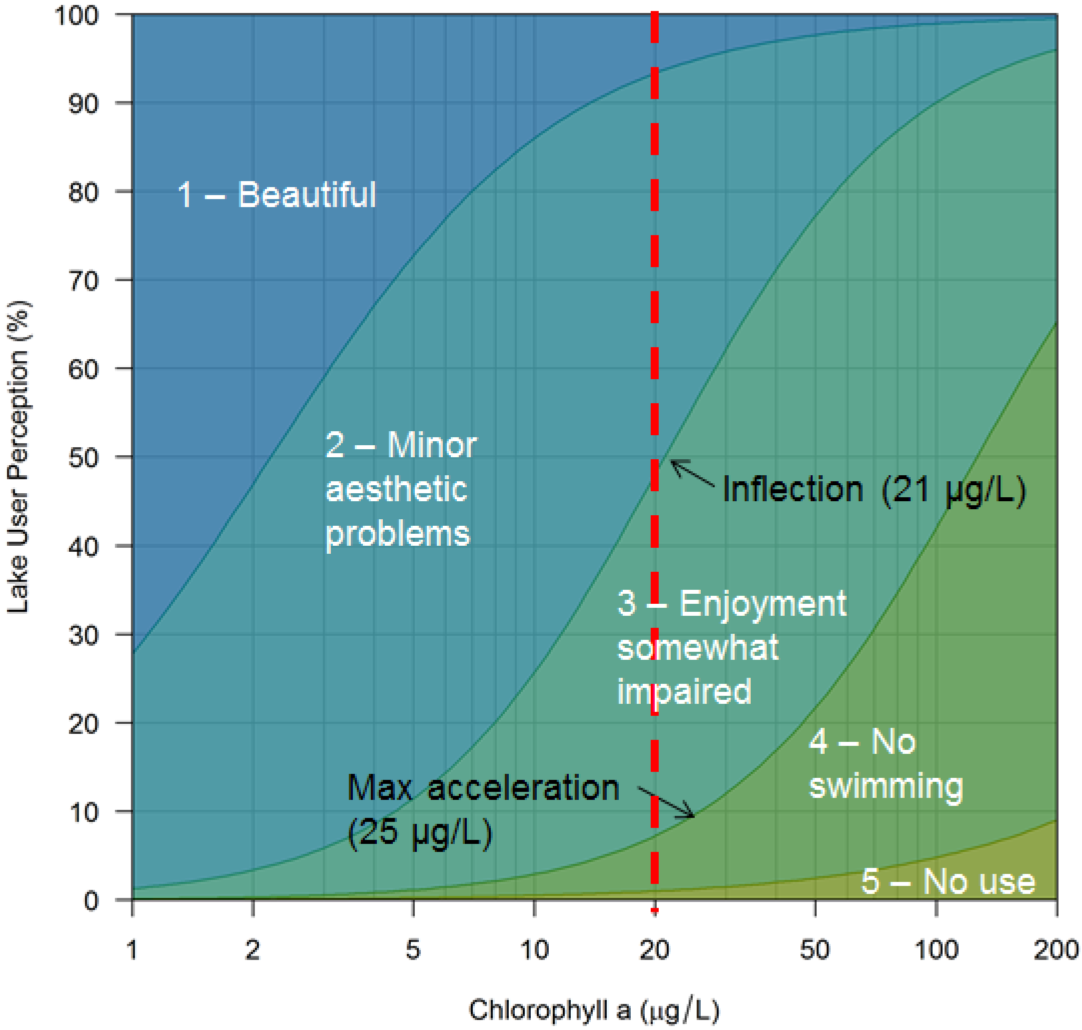
- The point at which half of users say their enjoyment is somewhat impaired due to algae is at 21 ug/L chl *a*. This is shown in Figure 4 as the inflection point for the line indicating the upper edge of category 3. Below this concentration, the majority of users indicate that algae levels do not inhibit their use of the waterbody; whereas above this concentration a majority of users experience decreased enjoyment and recreation.
- The point at which most users (over 90%) would still swim, but just before a rapid increase in the proportion of users who would *not* swim, is at 25 ug/L chl *a*. This is shown on Figure 4 as the point of maximum acceleration for the line indicating the upper edge of category 4. This establishes the threshold before the point at which swimming becomes significantly inhibited.

The inflection point for "Category 3 – Enjoyment somewhat impaired" (21 µg/L) and the point of maximum acceleration for "Category 4 – No swimming" (25 µg/L) are nearly equal, and a threshold in this range could be translated into the following narrative justification: "Half of lake users perceive some impairment to their enjoyment and recreation due to algae, but over 90% of users would still swim." Rounding these values down to 20 µg/L, the proposed assessment threshold provides a margin of safety. Importantly, once this point is exceeded, the perceived impacts of algae increase rapidly and users are much less inclined to swim.

By setting Wisconsin's frequency threshold to limit moderate algae levels of 20 ug/L chl *a*, we are also protective against "severe" and "very severe" blooms, restricting these to occur a very small percent of the time.

This analysis only evaluates instantaneous perception of water quality, not the cumulative effects of persistent algal blooms on perceived suitability for recreation. The allowable exceedance frequency is discussed in the following section.

Figure 4. Plot of fitted relationships between chlorophyll *a* concentration and Wisconsin lake user perception of water quality.



Statistical analysis

The data used in the analysis were all chlorophyll *a* samples collected from the top 2 m of the water column in Wisconsin lakes and reservoirs during the period July 8 – Sept 22 (WisCALM chlorophyll *a* assessment period) from 2002[1] to 2016. Multiple values from the same station and date were averaged, and samples without a corresponding user perception rating were excluded.

The statistical model is a set of mixed effects logistic regressions, one for each perception level, where the response variable is a binary (0/1) of that perception level or higher (i.e., worse) and the predictor is log(chlorophyll). Station ID was included as a random effect on both the intercept and slope of log(chlorophyll) to account for variance among lakes and stations in the relationship between chlorophyll *a* and user perception. The models were fit using the glmer function in the R package lme4 with the following call:

```
glmer(IMP ~ log(CHL) + (1 + log(CHL) | STATION_ID),  mdata, family=binomial, nAGQ=0,
control=glmerControl(optimizer = "nloptwrap"))
```

[1] 2002 is when the current laboratory procedure (chlorophyll *a* by fluorescence) became the standard.

Variance in the fitted relationships was assessed by selecting bootstrap sample sets with replacement and refitting the model 500 times. The control arguments in the model call were used to speed model fitting to allow use of the bootstrap procedure. These controls gave identical parameter estimates to the default controls. Models were fit for nine lake classes and for all lakes combined (see Appendix A). The following discussion of potential assessment thresholds is based on the "all lakes" model. The nine lake classes did not exhibit enough variation to warrant separation of results based on lake class.

The "all lakes" model shows that subjective perception of water quality is strongly related to measured chlorophyll *a* concentration (Figure 4). It is appropriate to plot this kind of relationship on a log-linear plot because human perception of most stimuli scales linearly with the log of the stimulus (Fechner's Law, see Smith and Perrone 1996 for evaluation of this principle to water clarity). Proposed chlorophyll *a* assessment thresholds may be identified by either a) specifying a target frequency of lake user perception (e.g., 90% of lake users view the water as suitable for swimming), or b) identifying inflection points in the relationships that signify changes in the unit response to a unit stimulus. Critical points in a logistic function are the inflection point (PI), where the slope is maximal, and the points of maximum and minimum acceleration (PAA[2] = 9% and PDA = 91% of function maximum), which are where the function breaks from its lower and upper plateaus to its growth phase (Mischan et al. 2011). In application to the user perception curves, the PI is where half of the users perceive a particular condition, and the PAA is a breakpoint, above which the rate of increase in perception is highest. The PI for "3 – Enjoyment somewhat impaired" (21 µg/L) and the PAA for "4 – No swimming" (25 µg/L) are nearly equal, and a threshold in this range could be translated into the following narrative justification: "At least half of lake users do not perceive any significant water quality problems and the water quality is suitable for swimming for the vast majority of users." Rounding these values down to 20 µg/L, the proposed threshold would provide a margin of safety. This analysis only evaluates instantaneous perception of water quality, not the cumulative effects of persistent algal blooms on perceived suitability for recreation. The allowable exceedance frequency is discussed in the following section.

Consistency with previous protocols and research

The findings of Wisconsin's user perception survey and resulting selection of a "moderate algae" level of 20 ug/L chl *a* are consistent with Wisconsin's previous thresholds and assessment protocols and earlier research done by other parties, as described below. The 2016 analysis of Wisconsin user perception data supported the continued use of this threshold.

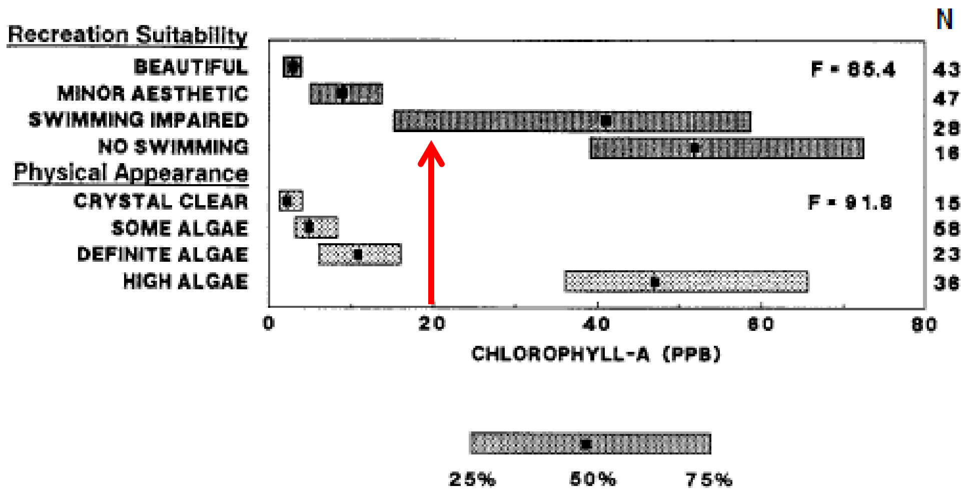
- A chlorophyll *a* threshold of 20 ug/L chl *a* was previously used by WDNR to develop Wisconsin's statewide phosphorus criteria for lakes, promulgated in 2010. During development of the statewide phosphorus criteria, the threshold of 20 ug/L chl a was based on Minnesota's work, discussed below. WDNR has also used this concentration in assessment protocols since the promulgation of phosphorus criteria in 2010.

- WDNR's definition of a "moderate algae" level directly corresponds with the Minnesota Pollution Control Agency's (MPCA's) definition of a "nuisance" algal bloom. Minnesota conducted an earlier study that surveyed user perceptions of lakes' recreational suitability and physical appearance (Heiskary and Walker, 1988). The study coupled user perceptions with simultaneously collected data on phosphorus, chlorophyll *a*, and Secchi depth. MPCA defined four algal bloom categories during their development of phosphorus criteria for Minnesota lakes: a "mild bloom" is greater than 10 ug/L; a "nuisance bloom" is greater than 20 ug/L; "severe nuisance bloom" is greater than 30 ug/L; and a "very severe nuisance bloom" is greater than 40 ug/L chl *a* (Heiskary and Wilson, 2008).

[2] In the paper cited, PAA stands for Point of Asymptotic Acceleration; PDA stands for Point of Asymptotic Deceleration. Letters of the acronym are rearranged with the A for Asymptotic last (Mischan et al. 2011).

As shown in Figure 5, a chlorophyll *a* concentration of 20 ug/L corresponds with the lower end of perceived swimming impairment, and is between a physical appearance of "definite algae" and "high algae". This study was used by both the Minnesota Pollution Control Agency and the Wisconsin DNR in setting phosphorus criteria for lakes.

Figure 5. Excerpt from Figure 3 in Heiskary and Walker's 1988 paper showing results of user perception surveys of a range of chlorophyll *a* concentrations. Interquartile ranges of measurements in each response category. Legend: N = number of observations; F = variance ration (among-group mean square/within-group mean square) derived from one-way analysis of variance on logarithmic scales.



- The threshold of 20 ug/L chlorophyll *a* is also consistent with an extensive analysis of Wisconsin lake data by Lillie and Mason, published in 1983. This analysis recommended six categories for chlorophyll *a* in relation to water clarity (Figure 6). As shown in Figure 6, a concentration of 20 ug/L chlorophyll *a* as a moderate algae level corresponds to the lower (better) end of the "Poor" category. The frequency thresholds provided here would restrict this poorer level of water quality to a given percentage of the summer.

Figure 6. Excerpted from Lillie and Mason (1983), Table 19. Apparent water quality based on chlorophyll *a* and water clarity as related to the Carlson Trophic State Index.

Chlorophyll *a* (μg/l)	Apparent Water Quality	Approximate Water Clarity Equivalent (m)	Approximate TSI* Equivalent
<1	Excellent	>6	<34
1-5	Very Good	3.0-6.0	34-44
5-10	Good	2.0-3.0	44-50
10-15	Fair	1.5-2.0	50-54
15-30	Poor	1.0-1.5	54-60
>30	Very Poor	<1.0	>60

*Based on Carlson (1977).

4.4.1.2 Frequency of moderate algae levels

This assessment protocol is based on the percent of days (frequency) during the summer sampling season that a lake experiences moderate algae levels. This approach recognizes that algal concentrations are episodic in nature and that higher levels naturally occur a certain percent of the time. Because of this episodic nature, a frequency measure is more appropriate for assessing recreational opportunity than a mean concentration over a longer timeframe.

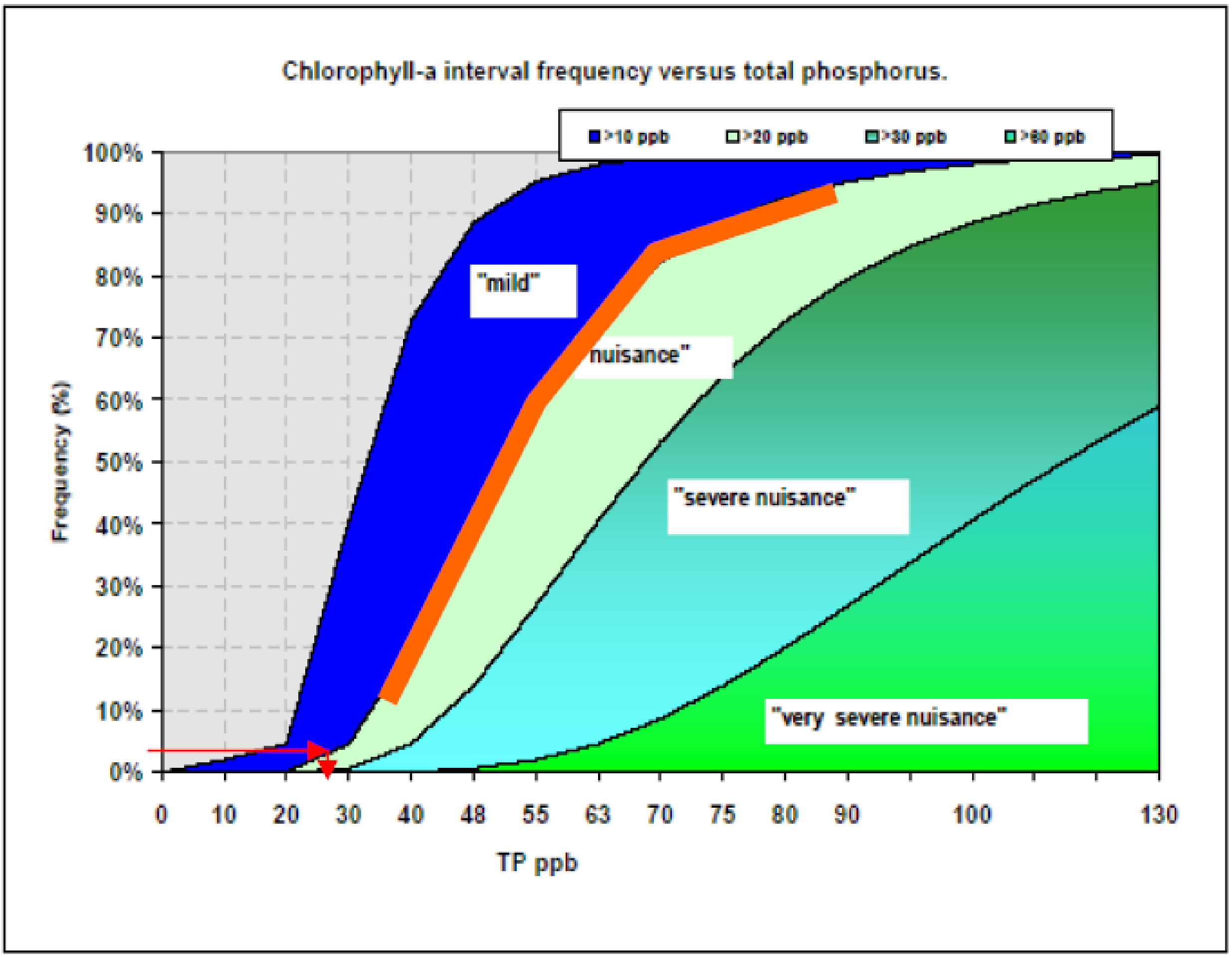
Deep lakes

The proposed recreational use assessment threshold for deep lakes is that moderate algae levels (20 ug/L chlorophyll *a*) shall not occur more than 5% of days during the summer sampling season. This threshold was one of the primary endpoints used for development of the statewide total phosphorus (TP) criteria in 2010 (Phosphorus Technical Support Document, 2010). Figure 7 was used in Wisconsin's Phosphorus Technical Support Document to show that 5% frequency corresponds to ~28 ug/L phosphorus. This was rounded up to a criterion of 30 ug/L TP for deep lakes, which should result in moderate algae less than 5% of the summer and severe blooms less than 1% of the time. The figure was originally developed from a study on user perceptions of lake recreation suitability in Minnesota (Heiskary and Walker, 1988), and was also used as a basis for Minnesota phosphorus criteria for lakes (Minnesota Pollution Control Agency, 2005).

As part of more recent efforts to develop phosphorus response indicators, we conducted a quantile regression analysis of Wisconsin's deep lakes comparing phosphorus to the frequency of moderate algae levels (Figure 8). The dataset for this analysis included all deep lakes in the state that had six or more TP and chlorophyll *a* samples (416 lakes). The 50% line of the quantile regression shows the median response of chlorophyll *a* to TP (Figure 8). The TP criterion for deep drainage lakes of 30 ug/L (the least restrictive of the phosphorus criteria for deep lakes) and the proposed chlorophyll *a* recreation assessment threshold are right at the point where chlorophyll rapidly increases with additional phosphorus. This analysis shows that a typical lake that meets the TP criterion will also meet the chlorophyll *a* assessment threshold. The chlorophyll *a* threshold will help identify the small percentage of lakes that frequently experience moderate algae levels even though they meet the TP criterion.

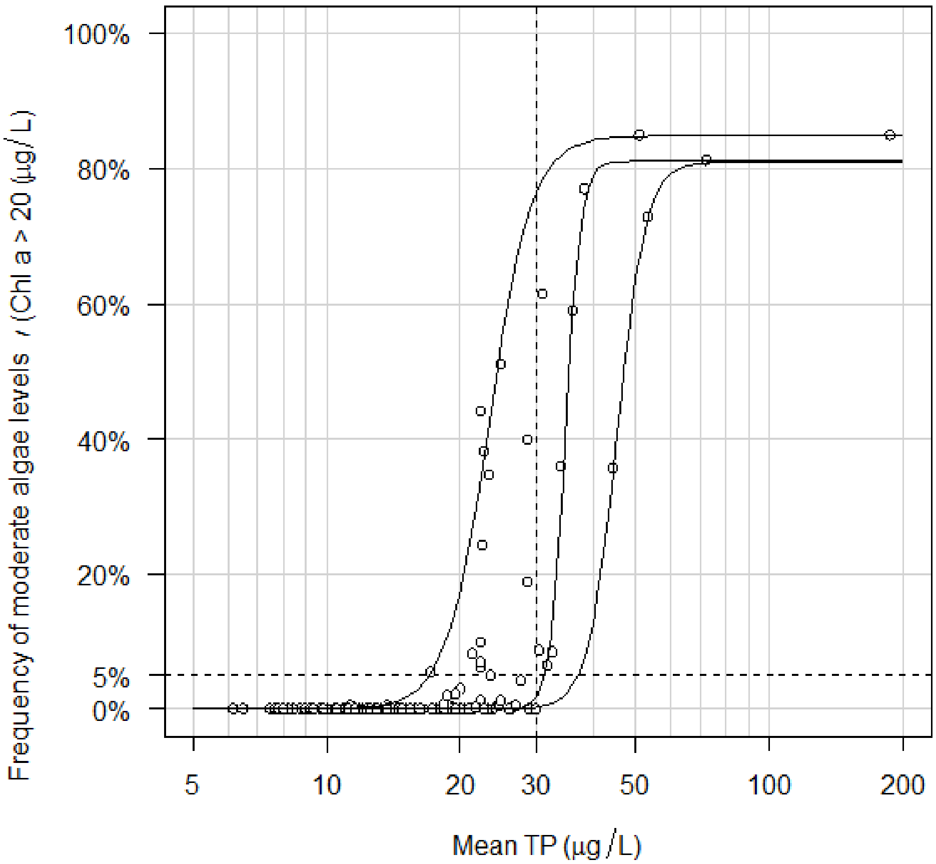
The frequency of days with moderate algae levels is not calculated as a fraction of total samples taken, but is instead estimated by fitting a distribution to all of the existing chlorophyll *a* data. For each lake, the non-central T-distribution is fit to at least 6 chlorophyll *a* concentrations and the probability of exceeding 20 ug/L chl-a is estimated from this distribution. A 90% confidence interval for frequency of moderate algae levels is also estimated, which accounts for sample size and chlorophyll *a* variability.

Figure 7. Frequency of moderate algae levels (formerly termed "nuisance" algal conditions in the 2010 Phosphorus Technical Support Document) relative to total phosphorus concentrations.



Sources: Reprinted from Wisconsin's 2010 Phosphorus Technical Support Document. Graph shows paired phosphorus and chlorophyll *a* measurements from 641 lakes as presented on page 25 of "Minnesota Lake Water Quality Assessment Report: Developing Nutrient Criteria", Third Edition, September 2005, Minnesota Pollution Control Agency. The development of the chart is described in the article by Heiskary and Walker, titled "Developing nutrient criteria for Minnesota lakes" and published in Lake and Reservoir Management 4:1-9, 1988. Arrows and highlight were added in WI's 2010 Phosphorus Technical Support Document.

Figure 8. Quantile regressions (5%, 50%, and 95%) showing relationship between phosphorus concentration and frequency of moderate algae levels (20 µg/L) in Wisconsin's deep lakes. The horizontal dashed line indicates the proposed chlorophyll *a* assessment threshold for recreation for deep lakes. The vertical dashed line indicates the established phosphorus criterion of 30 ug/L for deep drainage lakes; other phosphorus criteria for deep lakes not shown here are 20 ug/L for deep seepage and 15 ug/L for two-story fishery lakes (these lakes are also included in this dataset).



Shallow lakes

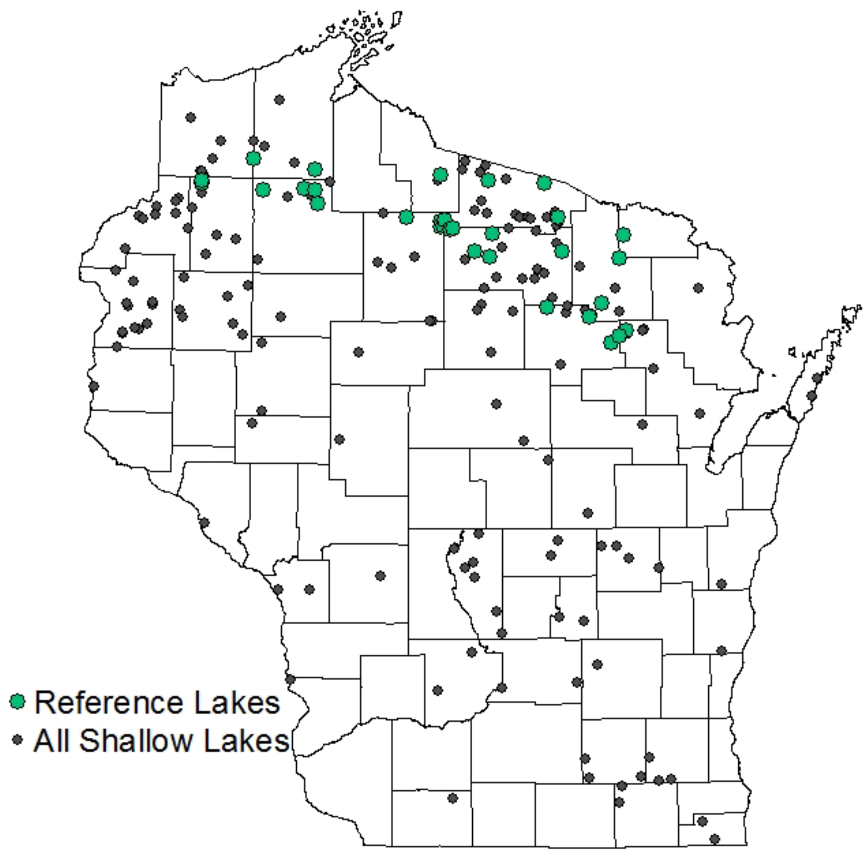
Wisconsin's statewide phosphorus criteria Technical Support Document does not specify the frequency of moderate algae levels appropriate for shallow lakes nor does Wisconsin have user perception surveys to determine a publicly 'acceptable' frequency of moderate algae for shallow lakes. Further, algal blooms are a natural and expected occurrence on many shallow lakes. Therefore, we employed a reference lake approach to determine the frequency of moderate algae levels expected in shallow lakes least disturbed by anthropogenic stressors. We then compared the frequency of moderate algae in shallow reference lakes to all shallow lakes.

We compiled a data set on all shallow lakes in Wisconsin that had at least 6 total phosphorus and 6 chlorophyll *a* samples (usually paired) from the deepest point of the lake (n=184). We analyzed the landcover (2006 National Landcover Dataset) in the entire upstream watershed of all 184 lakes and defined reference lakes as having minimal urban (Developed, Open Space; Developed, Low Intensity; Developed, Medium Intensity; Developed, High Intensity) and agricultural (Pasture/Hay; Cultivated Crops) land cover (sum of urban and agricultural land covers < 5%). The reference list based on land cover was further screened. Lakes were removed from the reference list based on regional biologists' knowledge of the lake (Table 3). The final set of 32 reference lakes are all located in northern Wisconsin, an unintended consequence of the land cover criteria used to define reference lakes (Figure 9).

Table 3. Justification for removing lakes from the shallow reference dataset even though land cover criteria were met (<5% urban and agricultural land cover in the upstream watershed).

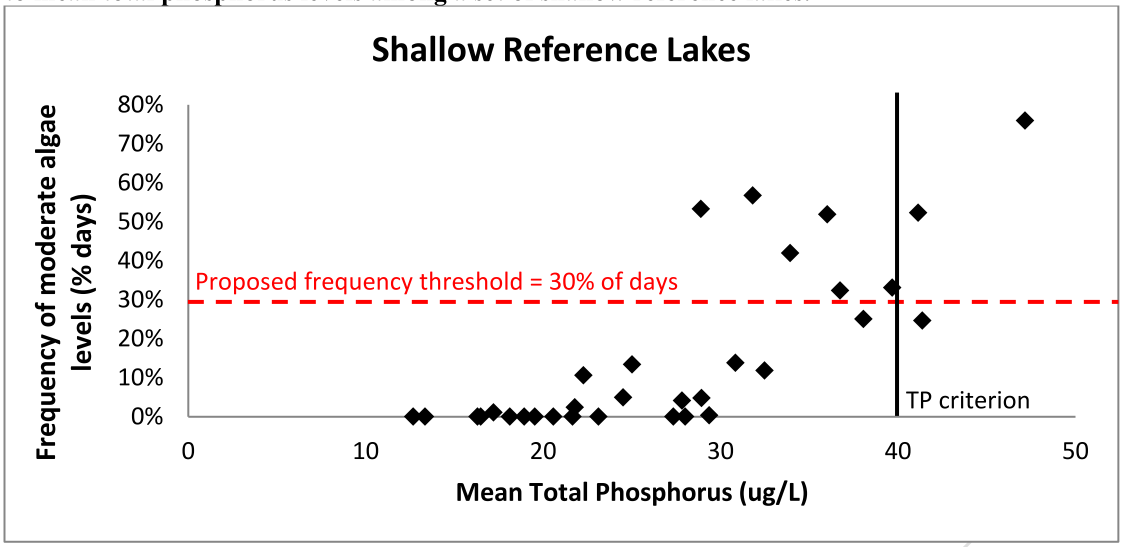
Lake Name	WBIC	Reason for removing from Reference dataset
Teal Lake	2417000	Carp reported present (carp resuspend nutrients in the sediment and shift lakes to algal dominated state)
Rolling Stone	389300	harvesting aquatic plants; impacts likely due to shoreland development and drainage from wetlands (possibly higher phosphorus soils)
Minong Flowage	2692900	major drawdown from 2013-2014
Musser Flowage	2245100	moderate development; managing for curly leaf pondweed
Cranberry	1603800	very developed, poor shoreline habitat, high runoff
Crane	388500	harvesting aquatic plants; impacts likely due to shoreland development and drainage from wetlands (possibly higher phosphorus soils)

Figure 9. Locations of reference and all other shallow lakes with at least 6 chlorophyll *a* and 6 total phosphorus samples from the deepest station.



The proposed recreational use assessment threshold for shallow lakes is that shallow lakes shall not experience moderate algae levels (20 ug/L chlorophyll *a*) more than 30% of days during the summer sampling season. This threshold was determined by calculating the 75th percentile of moderate algal frequency in all shallow reference lakes, which was 27%. Stated differently, 75% of shallow reference lakes have moderate algae levels less than 27% of the time. Given the uncertainty in selecting reference lakes, we rounded up to 30% for the shallow lake assessment threshold (Figure 10).

Figure 10. Statistically calculated frequency of summer moderate algae levels (chlorophyll *a* > 20 ug/L) in relation to mean total phosphorus levels among a set of shallow reference lakes.



Compared to all shallow lakes, shallow reference lakes show a similar relationship between mean total phosphorus and frequency of moderate algal levels (Figure 11). We observed the following patterns:

- The probability of moderate algae levels increases rapidly at 30-40 ug/L total phosphorus and remains high when TP > 40 ug/L (Figure 11). A large majority of both shallow reference lakes and all shallow lakes meeting the TP criterion experience algal blooms less than 30% of the time (Figure 12).
- Some reference lakes have moderate algae more than 30% of the time; one reference lake had moderate algae 76% of days (Figure 10). The high rates of algal blooms in a few reference lakes could be due to unknown anthropogenic stressors or could be naturally high.
- Despite high frequency of algal blooms in a handful of reference lakes, the proposed assessment threshold is not too restrictive. Most shallow lakes with moderate algae more than 30% of the time also exceeded the Aquatic Life chlorophyll *a* threshold (27 ug/L, Figure 11) and the total phosphorus criterion (40 ug/L, Figure 12).
- All lakes exceeding Aquatic Life chlorophyll *a* assessment threshold have moderate algae more than 58% of the time. This provides good separation between lakes exceeding the proposed Recreation assessment threshold for frequency of moderate algae levels versus the Aquatic Life chlorophyll *a* assessment threshold (Figure 11).

Figure 11. Probability of moderate algae levels (chlorophyll *a* >20 ug/L) in shallow reference lakes and in all shallow lakes that meet and exceed the Aquatic Life chlorophyll *a* assessment threshold (27 ug/L chlorophyll *a*).

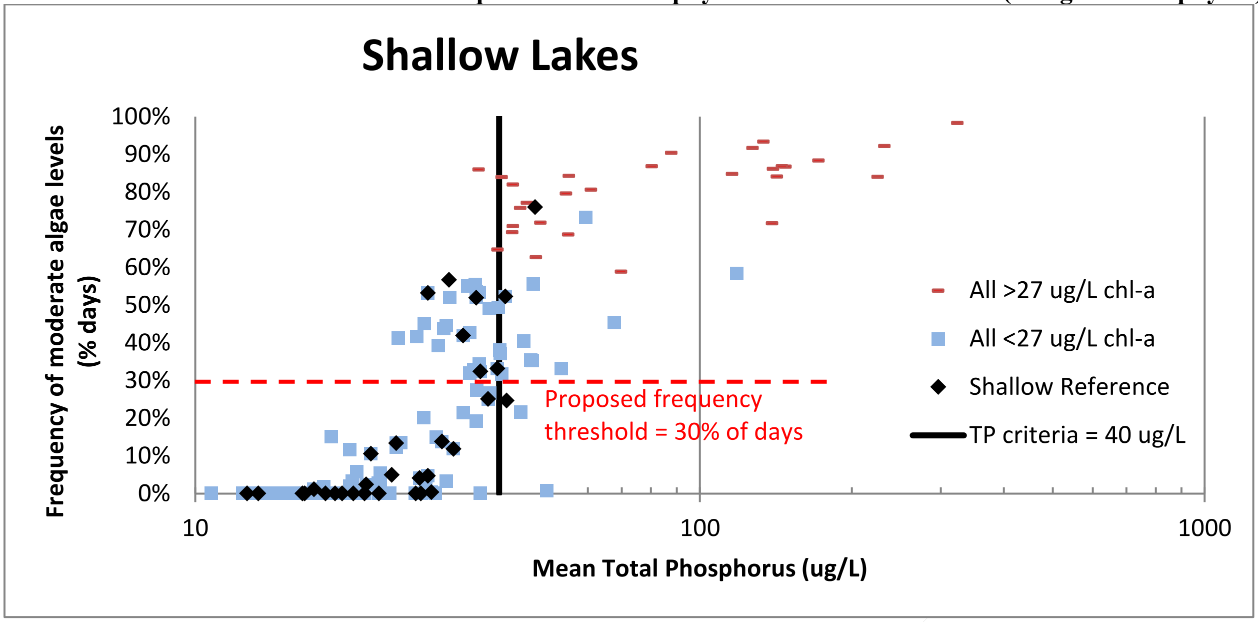
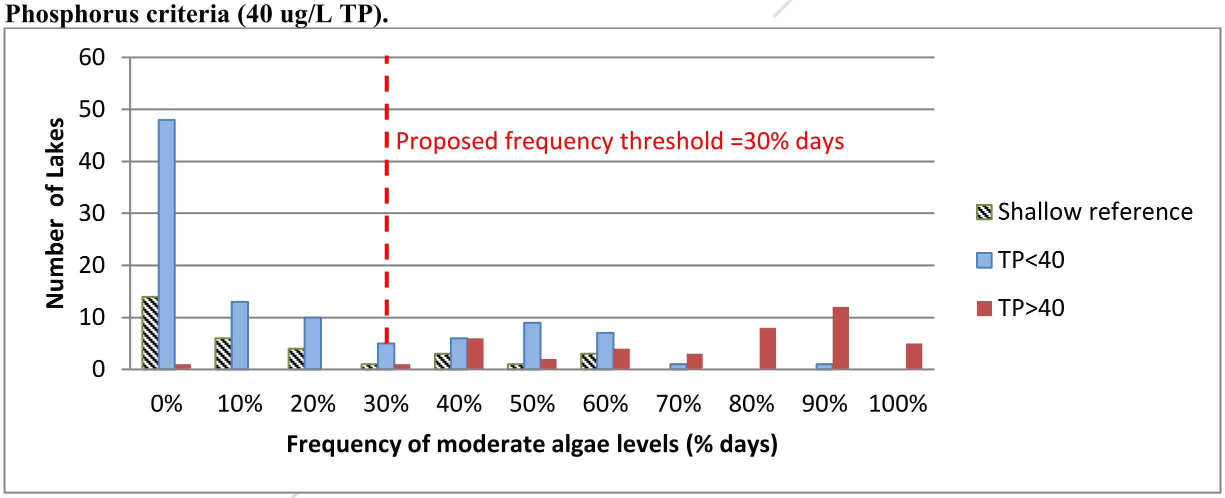


Figure 12. Number of lakes that have moderate algae levels 0 to 100% of the time. Shallow reference lakes are plotted separately from all other shallow lakes, which are divided amongst those that meet or exceed the Total Phosphorus criteria (40 ug/L TP).



The number of lakes listed as impaired for recreational chlorophyll *a* also depends on the confidence interval (CI) of the probability of chlorophyll *a* exceeding 20 ug/L. The application of the confidence interval approach is described more thoroughly in section 3.1.1. We counted the number of lakes that would clearly meet (90% confidence interval below threshold), may meet (median below threshold but 90% CI overlaps with threshold), may exceed (median above threshold but 90% CI overlaps with threshold), or clearly exceed (90% confidence interval above threshold) the recreational chlorophyll *a* assessment threshold (Table 4). Most lakes clearly meet or exceed the assessment threshold. For those with unclear assessment results (may meet or may exceed), an additional year of sampling is done to increase certainty; after that point, if results are still unclear the attainment decision is based on whether the mean is above or below the assessment threshold.

Table 4. Number of shallow lakes that will meet or exceed the recreational chlorophyll *a* assessment threshold given the 90% confidence intervals of the median probability of exceeding 20 ug/L chlorophyll *a*. Lakes are also tallied by the total phosphorus criterion (40 ug/L and the Aquatic Life chlorophyll *a* assessment threshold (27 ug/L).

Lake Group	Meet	May Meet	May Exceed	Exceed
Shallow Reference	21	3	4	4
All Shallow Lakes	68	10	21	43
TP < 40 ug/L	67	9	13	11
chl *a* < 27 ug/L	68	10	21	13
TP > 40 ug/L	1	1	8	32
chl *a* > 27 ug/L	0	0	0	30

In summary, the recreational use chlorophyll *a* assessment threshold for shallow lakes is the 75th percentile of a shallow reference lake data set and is consistent with previous WisCALM guidance. Most lakes meet the threshold and most lakes that do not also exceed aquatic life chlorophyll *a* threshold and TP criteria. This recreation assessment threshold is more stringent than the aquatic life chlorophyll *a* threshold and will thus identify some lakes showing signs of phosphorus impairment before they exceed the phosphorus criterion and/or aquatic life chlorophyll *a* threshold.

4.4.1.3 Assessment determinations using the Recreation thresholds

For the Recreational use assessment threshold, the department determines a waterbody's frequency of moderate algae levels during the chlorophyll *a* summer sampling period using the confidence interval for a percentile of a normal distribution, and uses the approach described under proposed s. NR 102.52 (2) (b) and (c) to compare that frequency (% of days) to the applicable threshold. These statistical methods for calculating the chlorophyll *a* concentrations and the number of days in the sampling season that exceed 20 ug/L chlorophyll *a* are described in detail in WisCALM 2020 in Section 4.6.

Applying this approach accounts for variability in water quality samples to determine if more samples are needed before making an assessment decision. If the determination is unclear—either "May Exceed" or "May Meet"—additional samples are required to shrink the confidence interval. Typically an additional year of sampling is done to increase certainty; after that point, if results are still unclear the attainment decision is based on whether the frequency (% of days) is above or below the threshold.

4.4.2 Chlorophyll *a* thresholds for Aquatic Life use assessments

The statewide lake phosphorus criteria specified in ch. NR 102.06 were set based on numerous factors to protect lake designated uses. As discussed above, recreational uses in lakes, primarily swimming, are impacted by algal blooms at a relatively low level of phosphorus. Aquatic life communities, particularly fish, are typically not impacted until higher levels of chlorophyll *a* are reached. This is because they are affected relatively little by rising levels of chlorophyll *a* until the lake 'flips' from a plant-dominated state to an algal-dominated state, at which point high levels of chlorophyll *a* impact visual feeding, reduce aquatic plants providing habitat, and impact availability of food sources. In order to assess support of recreational uses and aquatic life uses separately, the department is proposing chlorophyll *a* assessment thresholds for aquatic life that represent the threshold at which these respective uses are not attained.

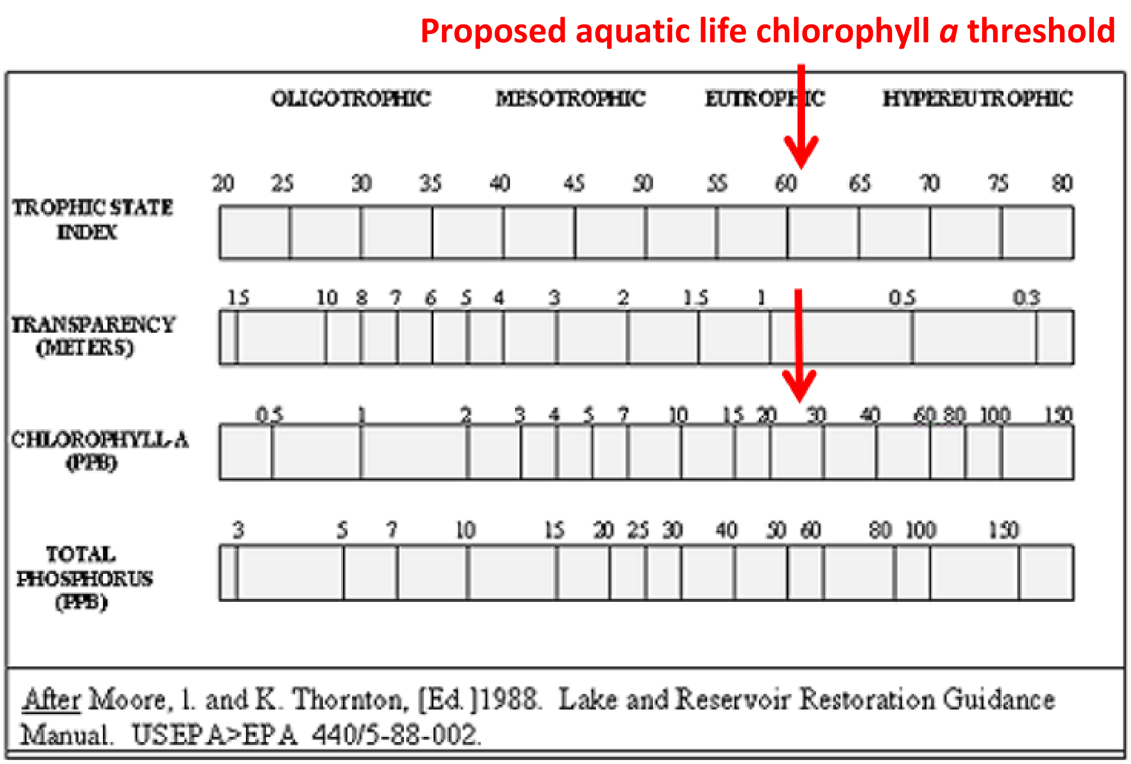
4.4.2.1 Deep and Shallow Lakes and Reservoirs

The recommended Aquatic Life chlorophyll *a* assessment threshold for all lakes and reservoirs except 2-story fishery lakes is 27 ug/L. This threshold is at the high end of eutrophic, but has not yet become hyper-eutrophic. Eutrophic lakes can support productive fisheries and it is important for Wisconsin's water quality program to allow for eutrophic conditions while protecting against nutrient pollution. An analysis of game fish in Minnesota lakes shows that yellow perch, walleye, and northern pike are most prevalent in mesotrophic and eutrophic lakes, but their occurrence precipitously declines as the TSI index climbs above 65 and approaches hypereutrophic conditions (Heiskary & Wilson, 2008).

Along the Trophic State Index (TSI) gradient, a TSI value of 63 corresponds with a concentration of 27 ug/L chlorophyll *a* and 60 ug/L total phosphorus (Figure 13, Carlson 1997). At this stage, the lake still may be restored to a clear water state, as it is before the point at which shallow lakes shift from an aquatic plant dominated to an algal dominated state (Jeppesen et al. 1990; Heiskary & Wilson, 2008). Because it is extremely difficult to shift a lake back to a plant dominated, clear water state once it has reached an algal dominated state, the threshold should be low enough to prevent this state shift.

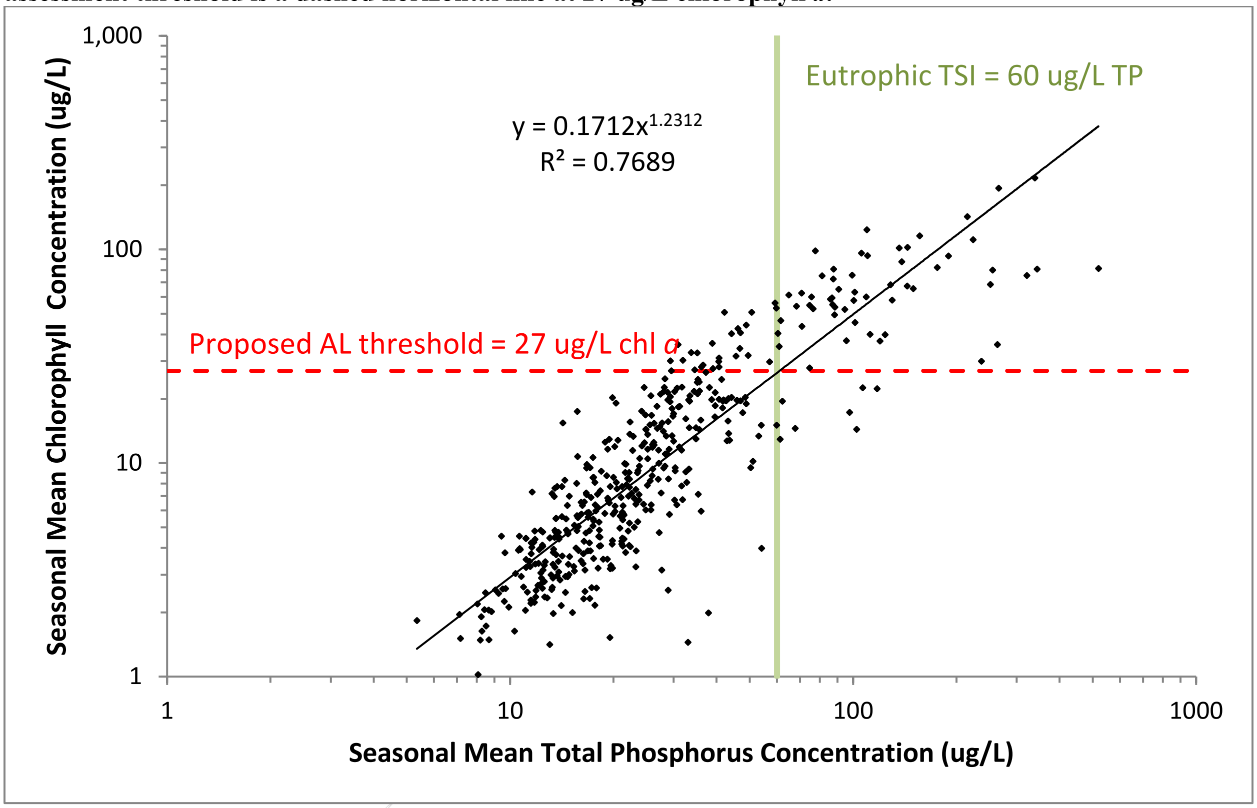
The following equations were used to equate chlorophyll *a* and TP concentrations to TSI values:
$TSI_{CHL} = 9.81 \ln(CHL) + 30.6$
$TSI_{TP} = 10 (6 - ((\ln(48/TP))/\ln(2)))$
Where: TSI = Trophic Status Index, CHL = Chlorophyll-*a* concentration (µg/L), TP = total phosphorus concentration (µg/L), ln = natural log

Figure 13. Continuum of lake trophic status in relation to Carlson Trophic State Index. The proposed biological assessment threshold of 27 ug/L chl a (63 TSI) for all lakes and reservoirs except two-story fishery lakes is at the upper end of eutrophic.



To ensure that the proposed assessment threshold based on the general TSI gradient applies to Wisconsin lakes, we fit a power relationship to mean total phosphorus and chlorophyll *a* (at lakes with at least 6 samples of chlorophyll *a* and TP taken at the deepest point of the lake). The chlorophyll *a* value at 60 ug/L TP was 27 ug/L (Figure 14). We also fit separate power relationships for shallow versus deep lakes, but decided to pool them because the relationship was the same. Thus, we confirmed that the relationship between phosphorus and chlorophyll *a* is similar in Wisconsin lakes to lakes that contributed to the Trophic State Index.

Figure 14. Mean total phosphorus (TP) and chlorophyll *a* in all lakes. The TP concentration of 60 ug/L representing a eutrophic state is shown as a green vertical line. The proposed Aquatic Life (AL) chlorophyll *a* assessment threshold is a dashed horizontal line at 27 ug/L chlorophyll *a*.



In summary, the proposed Aquatic Life chlorophyll *a* assessment threshold for deep and shallow lakes lies at the upper range of eutrophic. While initially based on the Carlson Trophic State Index, the data analysis above demonstrates that the general TSI gradient applies to Wisconsin lakes. Finally, chlorophyll *a* thresholds were previously in guidance, but will now be codified.

4.4.2.2 Two-Story Lakes

The proposed chlorophyll *a* assessment threshold for two-story fishery lakes is 8 ug/L. This threshold is designed to protect oxygenated, cold water habitat critical for coldwater fisheries by minimizing hypolimnetic oxygen depletion. Decomposing organic matter at the lake bottom is usually the main source of oxygen consumption in lakes. Oxygen depletion depends both on lake productivity and lake morphometry. The greater the productivity and the greater the ratio of sediment area to hypolimnetic water volume, the higher the hypolimnetic oxygen demand. For instance, a cone shaped lake would have greater hypolimnetic oxygen demand compared to a kettle shaped lake of equal depth.

The threshold of 8 ug/L chlorophyll *a* was chosen based on the trophic status of Wisconsin lakes with cisco populations and on supporting literature from neighboring Minnesota. A synoptic list of 136 Wisconsin inland lakes with reported populations of cisco was compiled. Each of these lakes was systematically surveyed with vertical gill nets from 2011 – 2015 to determine the current distribution and population dynamics of cisco. Since rainbow smelt can negatively impact cisco populations and confound any potential environmental relationships, 12 lakes with invasive rainbow smelt populations were excluded from the analysis. The Trophic Status Index was calculated for each lake using satellite-derived Secchi depths collected in midsummer 2011 and 2012. The TSI formula is:

$$TSI = 60 - 14.4(\ln(S))$$

where S is the Secchi depth in meters. Lakes were categorized by cisco abundance, and the TSI values at various percentiles by abundance category were examined. The goal was to choose a threshold that is sufficiently protective without listing healthy cisco lakes as impaired. We chose a TSI threshold of 51 (upper end of mesotrophic, Figure 13); this is higher than any observed TSI in lakes with high cisco abundance, it is the maximum TSI observed in lakes where cisco are moderately abundant, and it is approximately the 85th percentile of lakes with low/no cisco abundance (Table 5). Note that only lakes that historically harbored cisco are included in the "none" category. The TSI values would be much higher if all deep lakes without cisco were included. A TSI of 51 corresponds to a chlorophyll *a* concentration of 8 ug/L. Although cisco relative abundance and satellite-derived TSI are not the most precise metrics of cisco health and lake productivity, decreased cisco abundance due to cultural eutrophication has been documented in Minnesota (Jacobson et al. 2008). Investigation of a few lakes with high TSI and high cisco abundance indicated that these particular populations exhibited poor recruitment dynamics. More specific metrics like cisco production and primary production might allow for a more precise threshold in the future.

Table 5. Secchi-based Trophic Status of Wisconsin lakes with high (≥25 cisco per vertical gill net per night), medium (10 to <25), low (>0 to <10) abundance or no (0) cisco among Wisconsin lakes that historically had cisco. TSIs are listed by cisco abundance category and by percentile. For example, 10% of high abundance cisco lakes have a TSI less than 35, and the maximum TSI in high abundance lakes is 50. N lists the number of lakes in each abundance category.

Percentile	High	Medium	Low	None
10	35	34	35	37
25	37	38	37	39
50	40	41	41	43
75	44	46	46	47
80	45	46	48	50
90	47	48	53	56
Maximum	50	51	56	58
N	26	20	41	37

This assessment threshold is similar to, but slightly more conservative than what may have been derived from studies of Minnesota lakes. In Minnesota lakes, whitefish and cisco are found at TSIs up to 55-60 (Heiskary and Wilson, 2008). Jacobson et al. (2010) found that oxythermal habitat for coldwater fish in Minnesota lakes steadily decreased as total phosphorus (a measure of lake productivity) increased to about 25 ug/L. In Wisconsin lakes, TP of ~25 ug/L corresponds to ~9 ug/L chlorophyll *a* (Figure 14). After this point, lakes are often so productive that there is insufficient oxygen in the hypolimnion and metalimnion regardless of lake morphometry. Oxygenated, coldwater habitat only remains if the epilimnion is cold enough, which is rare for Wisconsin lakes during late-summer stratification. The breakpoint described in Jacobson et al. (2010) is useful for understanding the influence of productivity gradients on oxythermal habitat for coldwater fish. To ensure high quality oxythermal habitat is retained in lakes, both Wisconsin's TP criterion and chlorophyll *a* assessment threshold are set before the breakpoint described in Jacobson et al. (2010).

4.4.2.3 Assessment determinations using chlorophyll a thresholds

The statistical methods for calculating chlorophyll *a* concentrations for the Aquatic Life thresholds are described in detail in WisCALM 2020 in Section 4.5. Samples for each lake are aggregated into a "grand mean" and are compared against the threshold. Due to variability in water quality samples, the confidence interval approach as described in section 3.1.1 of this document and in WisCALM is applied to determine if more samples are needed before making an assessment decision. If the determination is unclear—either "May Exceed" or "May Meet"—additional samples are required to shrink the confidence interval. Typically an additional year of sampling is done to increase certainty; after that point, if results are still unclear the attainment decision is based on whether the mean is above or below the threshold.

4.4.3 Aquatic Plant thresholds for Aquatic Life use assessments

Lake-dwelling aquatic plants, or macrophytes, are sensitive to multiple forms of anthropogenic disturbance and can be used as a metric to signify ecological impairment (Alahuhta and Aroviita 2016, Lacoul & Freedman 2006, Wilcox 1995). Accordingly, we developed two assessment methods that evaluate the condition of a lake's aquatic plant community by relating aquatic plant abundance to anthropogenic disturbance. The first assessment method is called the Macrophyte Assessment of Condition-General (MAC-Gen) and describes overall aquatic plant community condition in response to multiple sources of anthropogenic disturbance. The second version of the method, called the Macrophyte Assessment of Condition-Phosphorus (MAC-P) is more narrowly focused, reflecting a plant community's tolerance of, and response to, phosphorus. The MAC-Gen is described in this section, and the MAC-P is discussed within the phosphorus response indicators section, 5.4.2.

Each of the lake assessment tools was developed using aquatic plant community data collected on 462 unique lakes. Surveyors employed a standardized point-intercept sampling method to estimate species abundance on a lakewide scale (Hauxwell et al. 2010, Mikulyuk et al. 2010). The MAC-Gen assessment method uses data-driven procedures to cluster plant species into three groups of species that are sensitive, moderately tolerant or tolerant to stressors related to eutrophication, population and land use. In general, when lakes are in poor condition, disturbance-tolerant plants are abundant, whereas lakes in good condition have high abundance of species that are sensitive to disturbance. Moderately tolerant plants often occur at intermediate levels of disturbance, decreasing toward either end of the disturbance gradient. This assessment method proposed by Mikulyuk et al. allows us to describe a lake's general condition using aquatic macrophytes (2017). We used a similar procedure for the MAC-P, for which we grouped species into two clusters based on their estimated upper limit of tolerance to phosphorus.

We then split lakes into four groups by region and lake type and related the abundance of each tolerance cluster to observed disturbance levels. From that, we determined thresholds that can be used to place lakes along that disturbance gradient (Table 6). Details of the procedure are outlined below.

Table 6. Aquatic plant community thresholds for lakes and reservoirs.

Lake Subcategory[1]	Macrophyte Assessment of Condition is attained if:
Northern Seepage	Moderately tolerant ≤ 64%
Northern Drainage	Tolerant ≤ 73%
Southern Seepage	Sensitive > 15%
Southern Drainage	Tolerant ≤ 50%

[1] In Table 8, northern lakes are those north of 44.84707°N latitude, and southern lakes are those south of that latitude. Thresholds have not been established for the Great Lakes.

Disturbance Thresholds

To develop plant-based disturbance thresholds, we followed the process described in Mikulyuk et al., 2017. Here we describe the process in five steps:

1. Determine the upper tolerance limit of each aquatic plant species to a variety of anthropogenic disturbance variables
2. Categorize species into one of three groups that vary in their sensitivity to disturbance
3. Calculate the frequency of occurrence of each of the three plant groups (sensitive, moderately tolerant, and tolerant) within the vegetated area of the lake
4. Split lakes into regional lake types to account for natural variation
5. For each lake type, define thresholds in the vegetated frequency of occurrence of plant tolerance groups that distinguishes lakes experiencing different levels of disturbance

Plant Tolerance to Disturbance (Step 1)

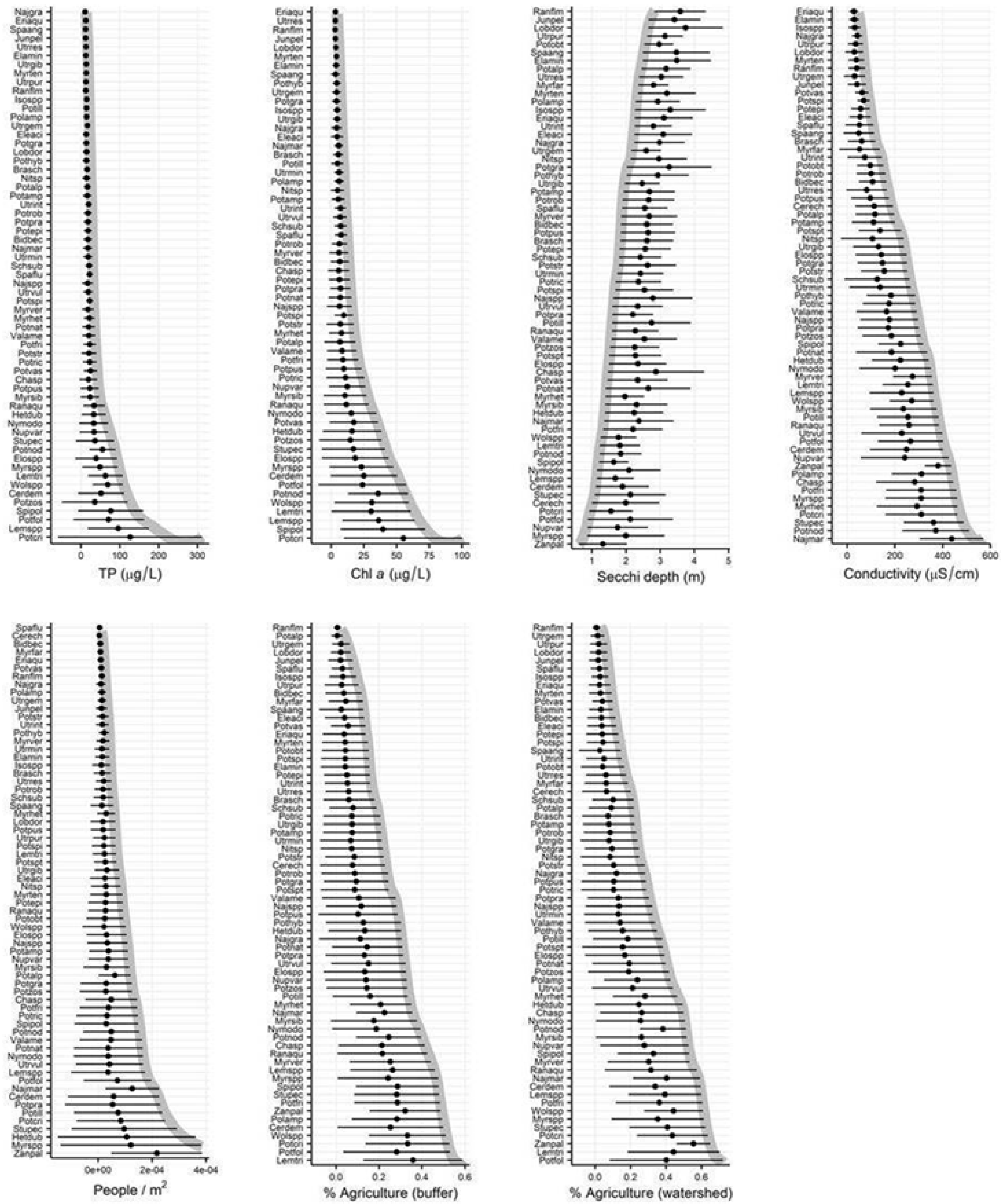
This method uses abundance patterns to estimate species-specific optimal environmental conditions and tolerance ranges. The first step is to estimate species abundance. The plant "point intercept" sampling method employed here uses a rake to collect plants at a large number of points lakewide on a grid scaled to produce more points when littoral zones are larger and when lakes have more complicated shorelines (Mikulyuk 2010). For any given lake, the sampling method produces a list of plant species that were found at each point on the grid. To characterize how widespread a given plant species is, the frequency of occurrence may be computed by taking the number of points where the species was found and dividing that by the number of sample points that occurred in areas shallow enough to support plant growth. This metric is called the littoral frequency of occurrence.

The MAC-Gen and MAC-P assume that a plant will be more commonly found in lakes that provide acceptable environmental conditions and more rarely found in lakes where environmental conditions are at the limits of a species' tolerance range. To build the general assessment, we assembled 20 disturbance variables representing human population, water quality and land cover factors. Population variables were expressed per watershed using data from the U.S. Census Bureau (2010). Water quality variables included conductivity (Papes and Vander Zanden, 2010), phosphorus, chlorophyll *a* and Secchi depth. We estimated mean summer total phosphorus, chlorophyll a and satellite-estimated Secchi depth using data drawn from the WDNR's surface water integrated monitoring system. We included samples collected from May 1 to September 1 and required at least three measurements to estimate summer means, averaging all estimates occurring within five years of a macrophyte survey. We estimated disturbed land use at the watershed and 500 m buffer scales in ArcGIS using lake and watershed polygons delineated by the WDNR (Akasaka et al., 2010; Menuz et al., 2013). We accepted the percentage of grassland occurring in the 500 m buffer as indicative of anthropogenic disturbance (e.g. lawns), but did not calculate this variable at the watershed scale where it might reflect natural conditions.

We then examined species-specific patterns in abundance across the 20 disturbance variables. We represented species abundance as the percent frequency of occurrence within a lake's littoral zone and estimated each species' optimum value for each of the 20 disturbance variables using an abundance-weighted average, holding that a higher abundance indicates more suitable conditions. We then estimated, for each species, the tolerance range for each disturbance variable using the standard deviation of abundance-weighted disturbance values. The abundance weighting step means a species' tolerance range will be wider if it tends to have abundant populations in lakes far from its optimum, and narrower if abundant populations only occur near the species' optimal environmental conditions. In general, we found that species vary in their tolerance to anthropogenic disturbance: some species do not occur when disturbance levels are high, whereas others are abundant across a wide range of conditions (Figure 15).

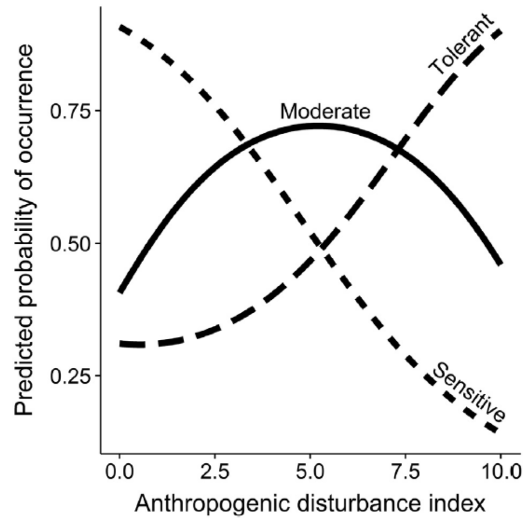
Disturbance Tolerance Groups (Step 2)
The next step was to distinguish species in a way that indicated disturbance levels. We first extracted the upper limit of the tolerance range calculated in Step 1 for each species and each disturbance variable. Next, we used a statistical technique to distinguish groups of species with shared patterns in their tolerance to disturbance using finite Gaussian mixture models (Fraley and Raftery, 2002). The best model used 7 disturbance variables to separate species into three groups that are either sensitive, moderately tolerant, or tolerant to multiple disturbance variables. Interestingly, there are morphological patterns evident across groups: disturbance-tolerant species are generally tall species with finely-dissected leaves or which are free floating; these species are less sensitive to light limitation and adapted to living in nutrient-rich waters. Floating leaf species, while tolerant of lower water clarity are less tolerant of general disturbance, which includes population-relevant metrics like urban development, potentially indicating sensitivity to shoreline development or mechanical disturbance (Radomski and Goeman, 2001). See Appendix D for a list of species that are classified as sensitive, moderately tolerant, or tolerant for use in this index.

Figure 15. Abundance-weighted average environmental optima (closed circle) and range (bars, ± 1 standard deviation) for aquatic plant species (y axis) across the 7 disturbance variables that produced the best discrimination among tolerance clusters (x axis). Aquatic plant species names are abbreviated using the first three letters of the genus and species. For example, *Najas gracillima* is abbreviated "Najgra".



Frequency of Disturbance Tolerance Groups (Step 3)
The next step involves relating the abundance of species belonging to each tolerance cluster back to an index of disturbance. First, we combined information on anthropogenic disturbance to produce a single index for Wisconsin lakes following methods outlined by Danz et al. (2007). We then reduced each group of anthropogenic disturbance variables (describing human population, water quality and land cover), to its principal components following Falcone et al. (2010) and Danz et al. (2007). We retained the orthogonal variables that explained at least 15% of observed variation and examined factor loadings to interpret each component. The seven land cover variables were reduced to two components that describe agricultural and urban land cover. The selected principle components of water quality describe nutrient enrichment and conductivity. The population variables were reduced to two components explaining population and road density. For each lake, we scored each of the six component values from 1 to 5. Values in the upper 20% of the distribution of scores were sssigned a 5, with sequentially decreasing scores assigned at the 80th, 60th, 40th, and 20th percentiles.We then added each lake's component scores and range-standardized the scores to produce a single index of anthropogenic disturbance varying from 0 (least disturbed) to 10 (most disturbed).

Figure 16. Abundance of species that are sensitive, moderately tolerant, or tolerant across an index of anthropogenic disturbance.



Next, we calculated the frequency of occurrence of each species tolerance cluster within vegetated areas by dividing the number of points occupied by at least one sensitive, moderate, or tolerant species, respectively, by the total number of vegetated points. For example, if 50 of 100 vegetated points had at least one disturbance-sensitive species (S), and 60 of the 100 vegetated points had at least one tolerant species (T), then the resulting metrics would be S = 0.50 and T = 0.60. Note that because we required at least 15 observations to determine a species' tolerance range, some plant species were not assigned a tolerance group and were removed from the analysis. Sample points having only rare species present would not count toward the number of vegetated points in the denominator.

Natural Lake Groups (Step 4)
There is a strong natural north-south gradient in geology, climate, and land use that affects plant community composition. Using a previous analysis of multi-scale patterns in aquatic plant community composition, we divided northern lakes from southern lakes at 44.84707° N latitude (Mikulyuk 2011) and conducted assessments separately for those groups. Seepage lakes (lakes with no outlets) also tend to have different natural characteristics than drainage lakes (lakes with at least one perennial stream outlet), so we also divided lakes based on hydrology. Here, reservoirs were included with drainage lakes.

Defining Thresholds (Step 5)
Finally, we used a conditional inference framework to partition lakes with similar disturbance levels into internally-consistent groups using the abundance-tolerance data calculated in step 3, above. The conditional inference procedure created the set of thresholds for northern seepage lakes, northern drainage lakes, southern seepage lakes, and southern drainage lakes depicted in Figure 6. Figure 17These decision rules were then translated into the thresholds shown in the aquatic plant condition thresholds table, Table 6, at the beginning of this section.

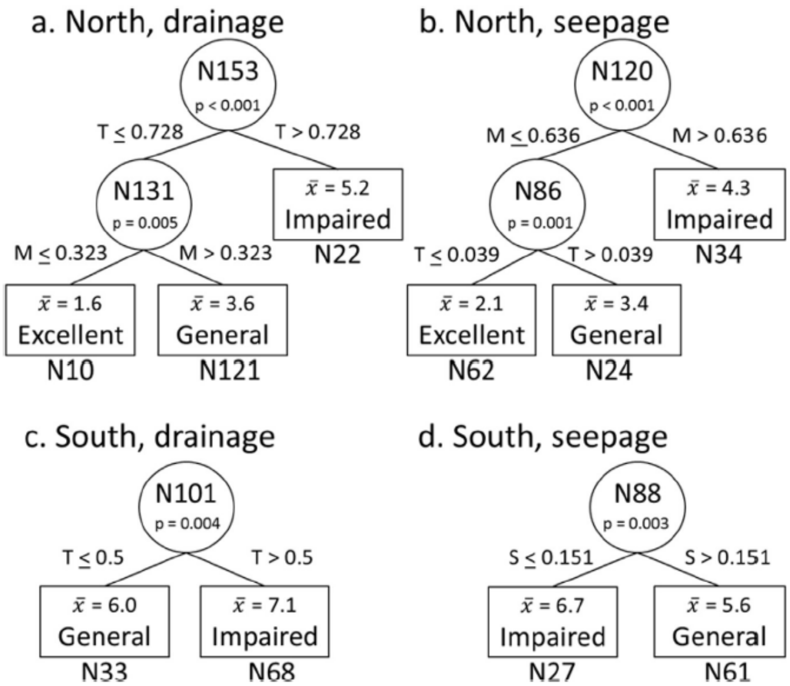


Figure 17. Conditional inference trees relating vegetated frequency of occurrence by TP-tolerance cluster to the anthropogenic disturbance index. Sample size indicated following N, p-values are printed in each node, with mean disturbance index and condition category labels in leaves. Threshold values of Sensitive (S), Moderately Tolerant (M), or Tolerant (T) vegetated frequency of occurrence are printed at each split.

5. Phosphorus Assessment Methods

This rule creates a new section, ch. NR 102.07, which defines assessment procedures for the total phosphorus criteria found in ch. NR 102.06. It contains two major components: procedures for calculating phosphorus concentrations, which reflect current protocols found in guidance, and a new component called the "combined assessment approach" which applies a set of biological phosphorus response indicators to attainment determinations.

5.1 ASSESSING PHOSPHORUS CONCENTRATIONS

The phosphorus criteria established in 2010 contained a numeric threshold but did not contain several other pieces of information that are critical to interpreting how to assess against that threshold. The information contained in this rule package provides the detail needed to apply clear assessment determinations.

The information in ch. NR 102.07, Assessing phosphorus concentration, contains data requirements for lakes and reservoirs and for flowing waters (rivers, streams, and impounded flowing waters). These specify where the criteria apply within a waterbody, the sampling period, the recommended number of samples for making an assessment determination, and the number of years over which samples may be applied for an assessment decision. Once samples are collected, it further describes the calculations necessary to compare the waterbody's phosphorus concentration to the TP criteria in ch. NR 102.06. Specifically, a lake's *mean* is compared to the criterion, and a flowing water's *median* is compared to the TP criterion (application of the mean vs. the median is discussed below). It describes the application of a confidence interval approach for determining whether more samples are necessary before making an attainment determination, referencing protocols found in proposed Subchapter III of ch. NR 102 (see section 3.1.1 of this document or WisCALM 2020, section 4.5, for further detail). This is important in cases where a small number of samples are highly variable or very close to the criteria; additional samples can provide more certainty that a waterbody's true mean or median is above or below the criterion. These protocols reflect how phosphorus assessments have been conducted under the WisCALM guidance over time.

The proposed approach also allows the department to calculate a weather-controlled mean or median TP concentration to compare against the TP criterion. The weather-controlled concentration accounts for variability over time more accurately than a site's short-term sampling data. The department runs these calculations using its Phosphorus Mixed Effects Regression (PhosMER) model, which uses the site's sampled TP data and the 30-year weather record to calculate the weather-controlled ambient concentration. PhosMER and its application is described in more detail in Appendix B. The department plans to make the PhosMER model available on its website in a format that can be easily used by external parties. To date this model applies only to stream or river data, but if it is expanded for lakes it may be used for lake assessment determinations as well.

5.1.1 Mean vs. median concentrations

For phosphorus assessments, the department uses a mean to assess the central tendency for systems that are relatively stable, such as lakes, as in these systems the mean is a better representation of the overall season. In the 2010 technical support document for the phosphorus criteria, the lakes chapter specified that the phosphorus criteria were based on mean concentrations. However, phosphorus concentrations in streams and rivers are much more variable due to storm events, and it is not unusual to have one high value out of the six required samples. In such cases using a mean could skew the results significantly even if the overall season was relatively low. The median is less influenced by extreme events than the mean, and is therefore applied in streams and rivers. The foundational paper on which Wisconsin's stream criteria were based states:

> "The USEPA developed the preliminary criteria based on median concentrations of all the data measured at each site rather than mean concentrations, because a median value represents the concentration most frequently occurring in the stream, and a statistical summary based on median values reduces the effect of outliers and values reported as less than their respective detection limits. The USEPA has provided preliminary criteria for P, N, SCHL, and turbidity for the national nutrient ecoregions and most level III ecoregions."
> (Robertson et. al., 2006)

The department applies the same protocol for chlorophyll *a* concentrations, calculating a lake or reservoir's mean and a stream, river, or impounded flowing water's median concentration to compare to the waterbody's criterion.

5.2 COMBINED ASSESSMENT APPROACH: U.S. EPA'S FOUR-PART PROCESS AND CONCEPTUAL MODEL

As part of WDNR's phosphorus assessment process, WDNR developed a suite of phosphorus response indicators. These are based on metrics that are most influenced by phosphorus, and can therefore serve as predictors of whether a waterbody is experiencing impacts due to phosphorus concentrations that are above the waterbody's phosphorus criterion.

The U.S. EPA developed a guidance document titled "Guiding Principles on an Optional Approach for Developing and Implementing a Numeric Nutrient Criterion that Integrates Causal and Response Parameters" ("Guiding Principles", USEPA 2013). Based on these principles, phosphorus response indicators should allow the state to have the capability to:

> "a) identify shifts in multiple biological assemblages (e.g., periphyton, benthic macroinvertebrates, fish) along a gradient of anthropogenic stress that can be tied to designated uses, and b) quantify the relationship between…phosphorus concentrations and measures of biological assemblage response."

The metrics selected should be sensitive to the stressor of interest (phosphorus) and should be relevant to protection of aquatic life designated uses. Measures of primary productivity and of algal assemblages are recommended as those most indicative of nutrient pollution. Higher trophic level indicators such as macroinvertebrates and fish may also be used as part of a suite of indicators, but should not be used as the sole indicator since they may not be as sensitive to phosphorus as lower-level indicators described previously. Dissolved oxygen or pH may serve as measures of ecosystem functioning.

WDNR also considered U.S. EPA's guidance titled "Using Stressor-Response Relationships to Derive Numeric Nutrient Criteria" ("Stressor-Response Guidance", USEPA, 2010) while developing phosphorus response indicators for Wisconsin. This guidance lays out a four-step process for deriving nutrient criteria and associated response indicators. The four steps are as follows (USEPA, 2010, p. ix-x):

> **1. Develop a conceptual model.** "…Conceptual models [are developed] representing known relationships between nitrogen (N) and phosphorus (P) concentrations, biological responses, and attainment of designated uses."

2. Assemble and explore data. "Data are assembled and initial exploratory analyses are performed. Variables are selected during this step that represent different concepts shown on the conceptual model, including variables that represent N and P concentrations, variables that represent responses that can be directly linked with designated uses, and variables that can potentially confound estimates of stressor-response relationships."

3. Analyze data. "…Stressor-response relationships are estimated between N and P concentrations and the selected response variables, and criteria are derived from these relationships."

4. Evaluate and document analysis. "…The accuracy and precision of estimated stressor-response relationships are evaluated and the analyses documented."

This section covers each of those four steps. Step 1, the conceptual model on which the rest of the section is based, is shown below. Because Steps 2, 3, and 4 are waterbody and metric-specific, and are often highly iterative, they are described within various portions of this section under each waterbody type and phosphorus response metric. For each metric:

- Under Step 2, Assemble and Explore Data, we selected and refined datasets for which we had robust data and a representative number of sites, and for which data were collected using well-established methods. We focused, in part, on using metrics that were used in development of the state's 2010 phosphorus criteria (WI DNR, 2010). We also briefly summarize other metrics that were considered but not selected.
- Under Step 3, Analyze Data, we used a variety of statistical approaches to visualize and analyze the data. The statistical approaches used varied based on characteristics of each dataset.
- Under Step 4, we compared the various statistical analyses with one another to verify the soundness of the analysis and determine where to set criteria thresholds. We then documented these approaches and the justification for our determinations in this Technical Support Document to ensure transparency.

In summary, this section describes when the combined assessment approach would be applied, the metrics selected as phosphorus response indicators and the justification for those selections, and briefly summarizes other metrics considered but not selected.

5.2.1 Conceptual model

WDNR developed the conceptual model shown in Figure 18 to depict commonly accepted pathways between nutrient inputs and cascading levels of responses in streams, rivers and lakes. The model includes three levels of response: primary, secondary, and tertiary, and how those responses impact Aquatic Life and Recreation Designated Uses. The U.S. EPA's "Stressor-Response Guidance" provides numerous references to publications documenting these effects (see Schindler 1974, Rosemond et al. 1993, Hill et al. 1995, Dodds and Welch 2000, Cross et al. 2006, Allen and Castillo 2007, Dodds 2007, Suplee et al. 2009). Wisconsin's conceptual diagram (Figure 18) is based primarily on the two diagrams shown in the U.S. EPA's "Stressor-Response Guidance", in EPA's Figure 2-1 for lakes and 2-2 for streams. However, we felt that a single and slightly more simplified model was accurate for representation of effects common to both lakes and streams/rivers. Wisconsin's model includes most of the same elements as depicted in the U.S. EPA's diagrams.

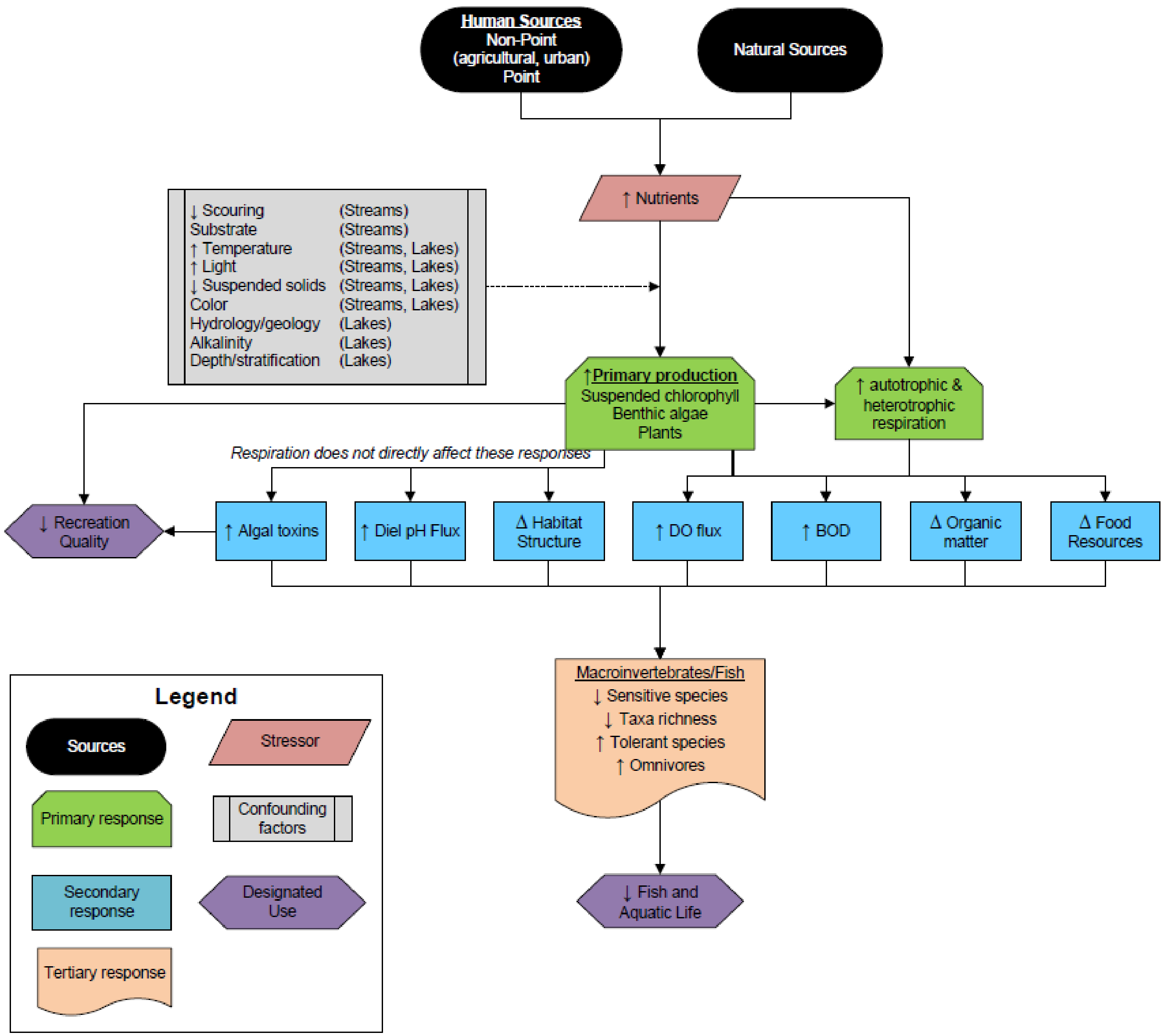
Primary response metrics
WDNR considered a variety of metrics depicted within this diagram for use as phosphorus response indicators, and concluded that the most immediate, direct and accurate measures of phosphorus response are the primary producers. The metrics selected to represent these primary response variables are benthic algal biomass and benthic diatom taxa for streams, suspended algae (chl *a*) for rivers, and aquatic plants and suspended algae (chl *a*) for lakes. Within the following sections of the section, there is discussion under each metric on how it responds to nutrient inputs and why it was selected as an appropriate indicator.

Secondary response metrics
Wisconsin already has numeric criteria representing two of the secondary response variables, dissolved oxygen concentration and pH. Because these criteria already exist, we did not include them directly as part of the phosphorus response indicators. However, oxythermal criteria is included as a phosphorus response indicator for two-story fishery lakes.

Tertiary response metrics
This rule package addresses tertiary response metrics as stand-alone biological assessment thresholds to indicate overall community health. Assessment thresholds for streams and rivers (in guidance) are based on macroinvertebrates and fish. A tertiary response metric (such as fish or insects) is not currently available for lakes, so the aquatic plant community is used as the most relevant and available assemblage-level biological metric for lakes. Additional metrics may be added in the future as they become established.

Although we assessed tertiary response metrics for use as phosphorus response indicators, the relationship was not as direct as the primary producers, and therefore we did not include them as phosphorus response indicators at this time.

Figure 18. Conceptual model demonstrating primary, secondary, and tertiary responses to nutrient inputs in lakes, streams, and rivers. Based in part on conceptual models from the U.S. EPA 2010, Figures 2-1 and 2-2.



5.3 COMBINED ASSESSMENT APPROACH: APPLICABILITY

Phosphorus response indicators are meant to be used in conjunction with the state's phosphorus criteria, as a "Combined Assessment Approach", and would be codified as such. In the combined approach, if a waterbody exceeds its applicable phosphorus criterion, but within a prescribed range, then WDNR would monitor and analyze whether the waterbody is attaining its phosphorus response indicators before making a decision to list as impaired for phosphorus. If all phosphorus response indicators are attaining the established thresholds, the waterbody would not be listed as impaired for phosphorus. If any one phosphorus response indicator is not attained, the waterbody would be listed as impaired for phosphorus.

5.3.1 Range for applying phosphorus response indicators

Phosphorus response indicators are only used if a waterbody exceeds its phosphorus criterion, but within a certain range, as shown in the rule in ch. NR 102.07 Table 10 and Table 7 here. The upper bound of a waterbody's range for applying the combined approach is defined by WDNR's existing definition of an "overwhelming exceedance" of phosphorus for each waterbody type (WisCALM 2020). If a waterbody has an overwhelming exceedance of phosphorus, it will be listed as impaired without assessing the phosphorus response indicators. However, if the waterbody's concentration lies within the range between the criterion and its overwhelming exceedance threshold, phosphorus response indicators will then be examined to determine whether the waterbody should be listed as impaired.

The definition of an overwhelming exceedance is as follows:

- Streams/rivers: the lower limit of the two-sided 80% confidence interval around the waterbody's median TP concentration exceeds the criterion by two times or more.
- Lakes: the lower limit of the two-sided 80% confidence interval around the lake's mean TP concentration exceeds the criterion by 1.5 times or more.

Therefore, for streams/rivers a combined assessment is used when a stream or river's concentration exceeds the criterion but by less than twice the criterion. For lakes a combined assessment is used when the lake's concentration exceeds the criterion but by less than 1.5 times the criterion. Streams/rivers have a wider bioconfirmation range than lakes because of their wider natural variability in phosphorus concentrations. Appendix C provides graphs for each major waterbody type (stream, river, and lake) showing the correlation between phosphorus concentrations and some of the phosphorus response indicators. The graphs show the range for applying the combined assessment and how many waterbodies falling inside this range would be considered impaired or not impaired for phosphorus.

The department will apply a confidence interval around the mean or median in making these determinations. It may use a weather-controlled mean or median phosphorus concentration if available.

Table 7. Range for applying combined assessment for total phosphorus[1]

Waterbody Type	Total Phosphorus Criterion (ug/L)	Combined Approach Range[2] (ug/L ambient total phosphorus)
Stream or its Impounded Flowing Water	75	75 to <150
River or its Impounded Flowing Water	100	100 to <200
Unstratified Reservoirs, Unstratified Drainage or Seepage Lakes	40	40 to <60
Stratified Reservoirs, Stratified Drainage Lakes	30	30 to <45
Stratified Seepage Lakes	20	20 to <30
Two-Story Fishery Lakes	15	15 to <22.5

[1]To determine whether a waterbody falls into the combined approach range, compare the lower confidence limit of the waterbody's two-sided 80% confidence interval around the mean (for lakes/reservoirs) or median (for rivers/streams) total phosphorus concentration to the ranges in the table.
[2]For streams and rivers the combined criteria range is between the applicable total phosphorus criterion and two times that criterion. For lakes, the range is between the applicable total phosphorus criterion and 1.5 times that criterion. If a waterbody has an approved site-specific phosphorus criteria, the combined criteria range for that waterbody shall be calculated using these multiplication factors.

5.4 LAKE/RESERVOIR PHOSPHORUS RESPONSE INDICATORS

Two main types of phosphorus response indicators are included in this rule package for lakes and reservoirs that are 5 acres or greater: algae (measured as suspended chlorophyll *a* concentration) and aquatic plants (macrophytes, expressed as the frequency of occurrence of macroscopic plants and algae). Additionally, for two-story fishery lakes, the oxythermal criteria apply as a phosphorus response indicator. Biological assessment based on lake water algal concentrations has been performed for years, whereas the macrophyte-based indicator was developed as part of this rule package and has now been available for several years as well.

5.4.1 Chlorophyll *a* for Aquatic Life and Recreation uses

The chlorophyll *a* assessment thresholds for recreation and aquatic life described in detail in sections 4.4.1 and 4.4.2 are also applied as phosphorus response indicators for lakes and reservoirs. Specifically, they are:

Chlorophyll *a* concentration (aquatic life thresholds):
a. Mean suspended chlorophyll *a* concentrations in lakes and reservoirs other than stratified two-story fishery lakes shall not exceed 27 ug/L.
b. Mean suspended chlorophyll *a* concentrations in stratified two-story fishery lakes shall not exceed 10 ug/L.

Frequency of moderate algal levels (recreation thresholds): A moderate algae level is defined as a chlorophyll *a* concentration of 20 ug/L or greater. If a lake, reservoir, or impounded flowing water exceeds the frequency of moderate algae levels specified in the table below during the summer sampling period, the department considers it not attaining its recreation use.

Table 8. Algae thresholds for recreational use assessment.

Waterbody Type[1]	Subcategory	Criteria for frequency of moderate algae levels
Lakes, Reservoirs, Impounded Flowing Waters (includes cold and warm)	Impounded flowing water, Unstratified drainage, Unstratified seepage	Does not exceed 20 ug/L chlorophyll *a* for more than 30% of days during the summer sampling period[2]
	Stratified drainage, Stratified seepage	Does not exceed 20 ug/L chlorophyll *a* for more than 5% of days during the summer sampling period[2]
	Stratified two-story fishery	

[1] Terms used for waterbody types and subcategories are defined in s. NR 102.03. These criteria do not apply to streams or rivers.
[2] Summer sampling period is July 15 to September 15.

5.4.2 Aquatic Plants

Aquatic plants are sensitive to nutrient enrichment, and species-specific differences in tolerance to enrichment may be used to detect impairment in natural lakes. Thus, the composition of aquatic plant communities in many cases can show impairment prior to algal indicators. Aquatic plants play stabilizing roles in lake ecosystems, supporting clear-water conditions via a positive influence on settling rates, nutrient burial and uptake. Some lakes that are enriched with nutrients will not show evidence of impairment in their ambient water dissolved phosphorus or chlorophyll *a* concentrations. However, as a lake begins to become enriched, plant community composition shifts toward more tolerant species adapted to enriched conditions. Following these principles, we developed an assessment method relating aquatic plant abundance and tolerance to total phosphorus (Macrophyte Assessment of Condition for Phosphorus, MAC-P). We developed this assessment method using the same data and following a very similar procedure as the Macrophyte Assessment of Condition (MAC) method outlined in Mikulyuk et al. (2017) and described in section 4.4.3. The MAC-P simply uses water column total phosphorus as the single disturbance measure rather than the 20 disturbance variables considered during the development of the general assessment method. The MAC-P procedure resulted in two clusters of species that differed in their estimated upper limit of tolerance to total phosphorus. Again, we split lakes into four groups by region and lake type and related the abundance of the sensitive and tolerant species clusters to observed phosphorus levels. We then determined thresholds to distinguish lakes along the total phosphorus gradient (Table 9). Details of the procedure are outlined below.

Table 9. Lake aquatic plant community phosphorus response indicator.

Subcategory: Lake Type[1]	Macrophyte Assessment of Condition for Phosphorus (MAC-P) attains if:
Northern Seepage	Phosphorus Tolerant ≤ 44.3%
Northern Drainage	Phosphorus Sensitive > 51%
Southern Seepage	Phosphorus Sensitive > 26%
Southern Drainage	Phosphorus Sensitive > 42%

[1] Northern lakes are those north of 44.84707°N latitude, and southern lakes are those south of that latitude. Seepage and drainage lakes follow the definitions in s. NR 102.03 (6h) and (1o). Seepage lakes include both stratified and unstratified seepage lakes, and drainage lakes include both stratified and unstratified drainage lakes. Plant phosphorus response indicators have not been established for Great Lakes and lakes less than 5 acres in surface area.

Phosphorus Response Thresholds

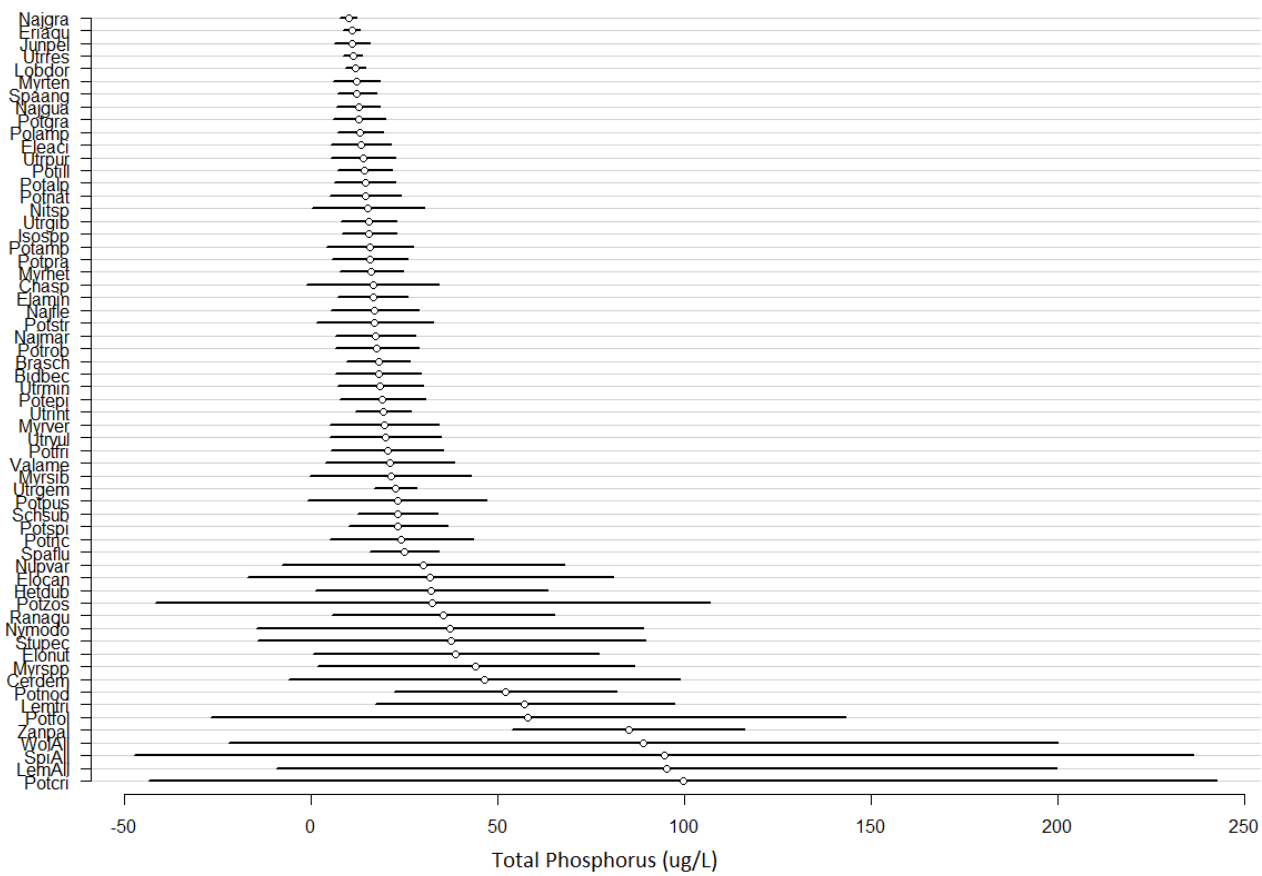
To develop plant-based phosphorus response indicators, we followed the general five-step process described in Mikulyuk et al., 2017 and in section 4.4.3. The procedures followed under each step are the same as described for the MAC-Gen in section 4.4.3 except as noted below. For each step, we present the results specific to the MAC-P tool.

Plant Tolerance to Phosphorus (Step 1)

Step 1 follows the same procedures as Step 1 for MAC-Gen except that instead of evaluating optima for a large group of disturbance variables, we only used the values for total phosphorus. This allows us to develop thresholds specific to phosphorus, which in turn enables us to evaluate the phosphorus response of a lake's plant community. Phosphorus was expressed as the mean phosphorus concentration (ug/L) in lakes with at least 3 measurements occurring from May 1 to September 1 taken at most 5 years before or after the macrophyte survey. Mean yearly concentrations were averaged when they existed for multiple years.

Figure 19 shows the TP optima and tolerance range for each species. We found that species vary in their tolerance to total water column phosphorus concentrations: some species do not occur at high phosphorus concentrations whereas others are abundant across a wide range of phosphorus conditions (Figure 19).

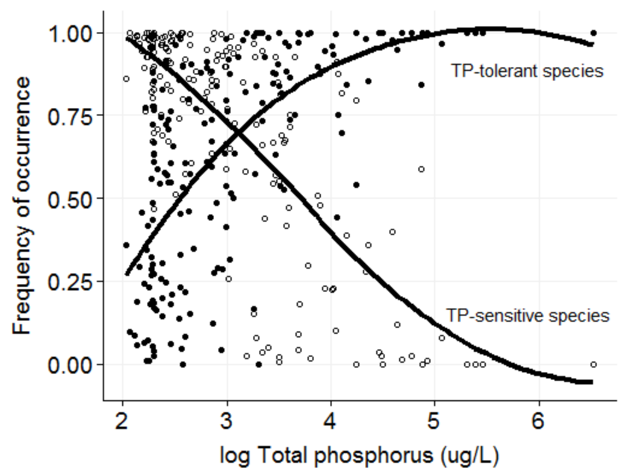
Figure 19. Abundance-weighted average optima (open circle) and range (bars, ± 1 standard deviation) of phosphorus concentrations defining the distribution of aquatic plant species. Phosphorus concentrations (µg/L) are on the x-axis and aquatic plant species are listed on the y-axis with the first three letters of the genus and then species names. For example, Najgra is *Najas gracillima*.



Phosphorus Tolerance Groups (Step 2)
As with MAC-Gen, after we determined the upper tolerance limit for each species, we used a statistical technique to group species with similar upper limits together (finite Gaussian mixture models, Fraley and Raftery, 2002). For the MAC-P, the best model divided species into two groups that are either sensitive or tolerant to phosphorus. There are again morphological patterns evident across groups: phosphorus-tolerant species are generally tall species with finely-dissected or floating leaves that are less sensitive to light limitation and adapted to living in nutrient-rich waters, whereas phosphorus-sensitive species tend to be short and compact or have wide, un-dissected leaves. See Appendix D for a list of species that are classified as phosphorus-tolerant or phosphorus-sensitive for use in this index.

Frequency of Phosphorus Tolerance Groups (Step 3)
Step 3 follows the same procedures as Step 3 for MAC-Gen. The resulting abundance of TP-tolerant and TP-sensitive species along the phosphorus gradient is shown in Figure 20.

Figure 20. Abundance of TP-tolerant and TP-sensitive species along a phosphorus gradient.



Natural Lake Groups (Step 4)
See Step 4 for MAC-Gen.

Defining Thresholds (Step 5)
Using the same procedure as Step 5 for MAC-Gen, we created a set of rules/thresholds for northern seepage lakes, northern drainage lakes, southern seepage lakes, and southern drainage lakes (Figure 21). These were then translated into the thresholds shown in the phosphorus response indicator table, Table 9 at the beginning of this section.

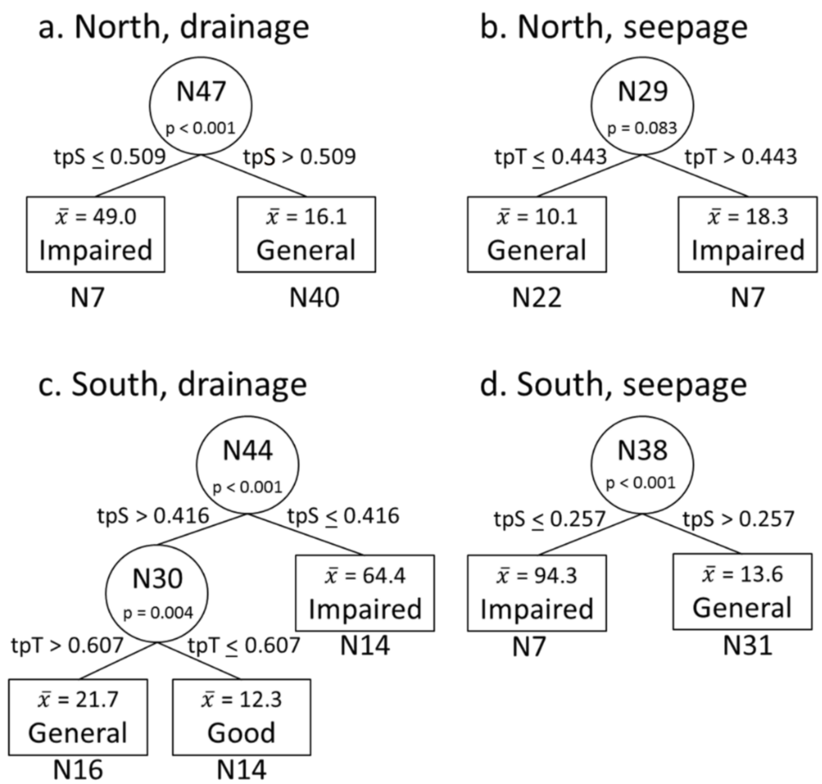


Figure 21. Conditional inference trees relating vegetated frequency of occurrence by TP-tolerance cluster to lake Total Phosphorus. Sample size indicated following N, p-values are printed in each node, with mean TP concentration and TP-condition category labels in leaves. Threshold values of TP-Sensitive (tpS) or TP-Tolerant (tpT) vegetated frequency of occurrence are printed at each split.

5.4.3 Oxythermal layer criteria

For two-story fishery lakes, the oxythermal layer thickness criteria specified in s. NR 102.04 (4) (am) also applies as a phosphorus response indicator. Elevated phosphorus can lead to oxygen depletion in lakes and reduce the habitat necessary for coldwater fish. Although phosphorus may not be the only factor affecting oxythermal habitat, if the oxythermal habitat requirement is not met in a waterbody with elevated TP levels, it is inappropriate to determine that the waterbody is not experiencing stress due to phosphorus (and not list it as impaired for TP) unless further studies indicate otherwise.

5.5 RIVER PHOSPHORUS RESPONSE INDICATORS

5.5.1 Chlorophyll *a* for Recreation use

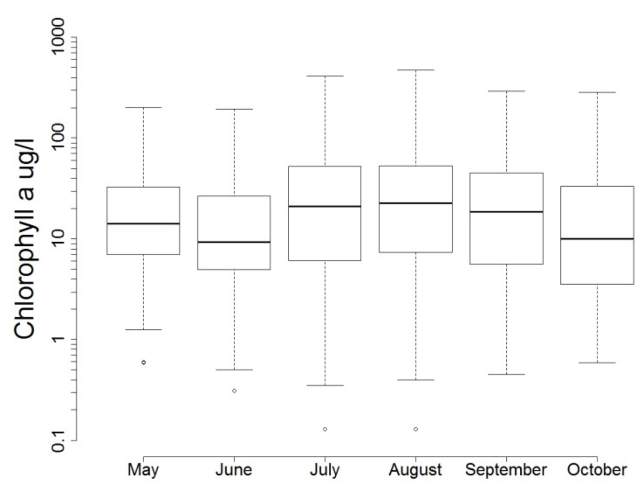
For rivers, frequency of moderate algae levels (as measured by suspended chlorophyll *a*) is established as the only statewide phosphorus response indicator at this time. Algal productivity is assessed in rivers using the same recreation use metric as for shallow lakes and reservoirs and impounded flowing waters: if a suspended chlorophyll *a* concentration of 20 ug/L is exceeded more than 30% of the summer sampling season (July 15-September 15), the waterbody would be considered impaired for phosphorus. We conducted the following analysis to examine the application of this threshold to rivers.

Wisconsin's water quality criterion for total phosphorus (TP) in nonwadeable rivers is 100 ug/L. The impacts of phosphorus in river systems vary depending on a number of factors including physical features, light availability to the water column and benthos, and phosphorus uptake pathways (i.e. benthic algae and macrophytes or phytoplankton). In river and impoundment ecosystems a common response to increased phosphorus is increased phytoplankton in the water column (measured as chlorophyll *a*), potentially reaching moderate algae levels. Suspended (sestonic) chlorophyll *a* was one of the primary indicators used in development of Wisconsin's total phosphorus criteria in 2010, which noted that nonwadeable rivers exhibited a strong correlation between total phosphorus and the amount of suspended algae as measured as chlorophyll *a* (WI DNR, 2010).

Determination of criteria thresholds

We used two datasets from nonwadeable river monitoring programs to evaluate the TP-chlorophyll *a* relationship and determine a threshold that confirms a phosphorus response. We used data from the nonwadeable river Long Term Trends (LTT) monitoring program which monitors 43 river sites across the State. Each site is sampled monthly (~2/3 of sites) or quarterly (~1/3 sites) over multiple years (http://watermonitoring.uwex.edu/pdf/level2/symposiumreferenceLTTAnnualReport2006.pdf). The LTT Rivers dataset spans rivers across gradients of size, geography, ecoregion, land use and human modifications representing the broad range of conditions seen in nonwadeable rivers across the state. We added to this dataset by including data from the *Nutrient Concentrations and Their Relations to the Biotic Integrity of Nonwadeable Rivers in Wisconsin* report by Robertson et al. (2008). This project sampled river sites over a 6 month period in 2003 from May through October. There was some overlap from the two datasets so we combined data that were from the same or proximal locations into one site for analysis.

Figure 22. Chlorophyll *a* concentrations from 2003-2013 among all months considered to determine the appropriate index period.



The first step in the analysis was to determine the appropriate index period for chlorophyll *a* in nonwadeable rivers. Chlorophyll *a* concentrations vary seasonally due to factors such as water temperature, light and nutrient concentrations. Two options already in use by WDNR for assessments include sampling TP in wadeable streams (monthly, May to October) and sampling TP and chlorophyll *a* in lakes (monthly, July-September). Using only sites that had multiple years of monthly data (n=31) we compared chlorophyll *a* values across all months and found that in nonwadeable rivers, on average, chlorophyll *a* concentrations were highest in July, August and September (Figure 22). We tested the two possible index periods to determine if there were any differences among the chlorophyll *a* values in each month over the two different index periods. For the July-September index period, there were no significant differences in chlorophyll *a* among months (ANOVA, $p=0.33$). However, there were significant differences in chlorophyll *a* among months in the May-October index period (ANOVA, $p<0.001$). Based on these findings, we decided to use July, August and September as the index period for assessing chlorophyll *a* in nonwadeable rivers. This represents both the most sensitive time period for algal response and the typical swimming period for protection of recreational uses. In addition, because there were no differences in the distribution of chlorophyll *a* among the July-September index period we were able to include sites that had only quarterly sampling events in our dataset.

Recreational uses of nonwadeable rivers are similar to shallow lakes, including boating, fishing, and swimming. Therefore, the proposed definition of moderate algae levels in rivers is the same as in shallow lakes: 20 µg/L chlorophyll *a*. To determine the acceptable frequency of moderate algae conditions in nonwadeable rivers, we plotted the estimated frequency of chlorophyll *a* > 20 µg/L during July-September against the median growing season (May-Oct) TP (Figure 23). For this analysis, we used data from all nonwadeable rivers in Wisconsin with at least 6 chlorophyll *a* and TP samples (n=49).

We used a Receiver Operating Characteristics (ROC) analysis to identify the frequency of moderate algae levels that best separates rivers that meet and exceed the TP criterion (Figure 24). This analysis plots the sensitivity (correct positive) and specificity (correct negative) rates across a range of potential thresholds. The frequency of moderate algae that best separates rivers that meet and exceed the TP criterion is in the range of 25-50%. Because the independently-determined shallow lake threshold of 30% is in this range, and for consistency, the proposed chlorophyll *a* phosphorus response indicator in nonwadeable rivers states that a waterbody within the combined assessment range will be listed as impaired for phosphorus unless it exceeds 20 ug/L chlorophyll *a* for fewer than 30 percent of days during the summer sampling period of July 15 to September 15.

Most of the rivers that exceed the TP criterion but not the chlorophyll *a* threshold are in the Driftless Area and have high turbidity, which limits algal growth. The macroinvertebrate IBI may be a more appropriate TP response indicator in these rivers. All of the sites that exceed the chlorophyll *a* threshold but not the TP criterion are on the Wisconsin or Fox Rivers, and are downstream of impoundments. TMDLs for TP in these systems will consider the hydrologic conditions that lead to high algal productivity.

Figure 23. River frequency of moderate algae levels versus growing season median total phosphorus concentration (49 Wisconsin rivers).

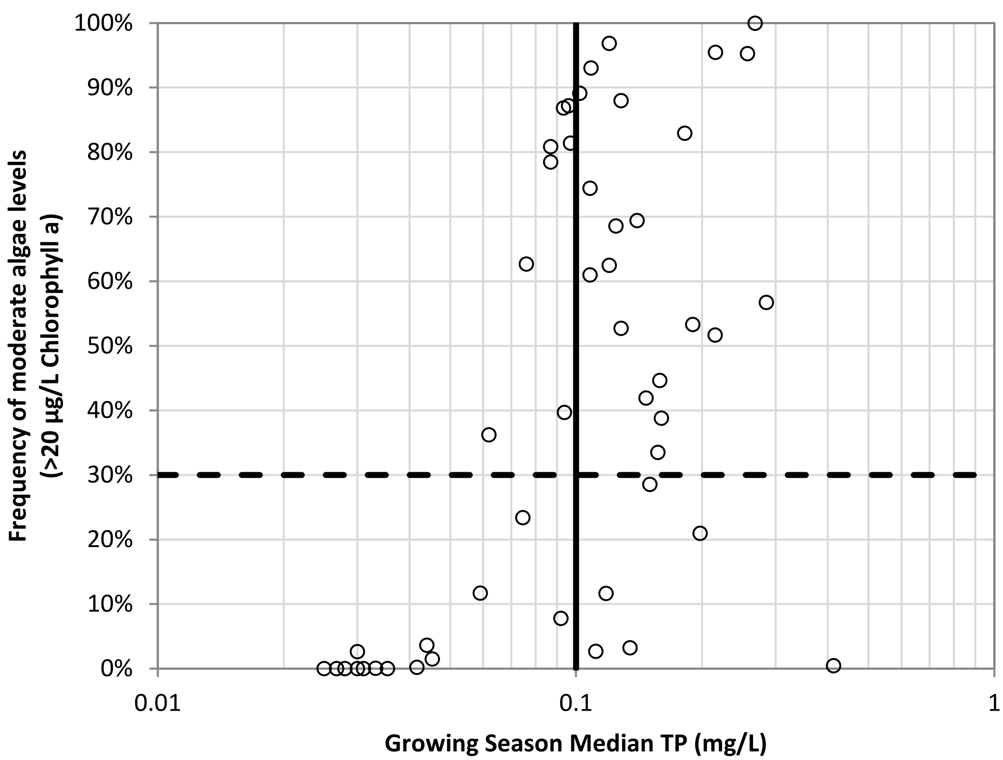
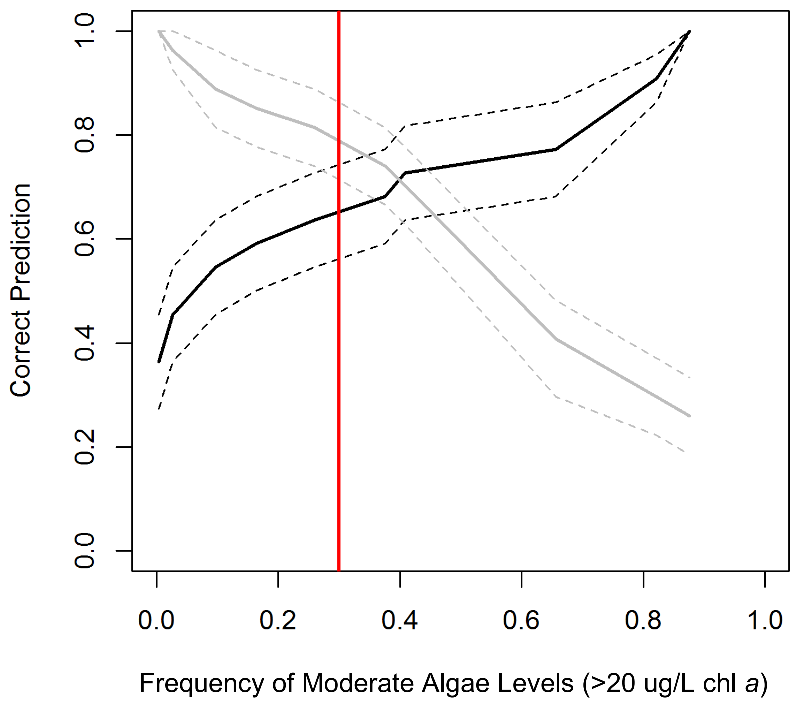


Figure 24. Plots of sensitivity (grey line) and specificity (black line) for frequency of moderate algae levels to correctly classify a river site as being above or below the TP criterion. Vertical red line indicates proposed river threshold for frequency of moderate algae levels.



5.5.2 Other metrics not selected

The following primary productivity metrics were considered for development of river phosphorus response indicators but were not selected.

- **Macroinvertebrates**. Macroinvertebrates can be a useful indicator in rivers because some rivers are too turbid to allow enough light penetration for algal growth, even if enough phosphorus is available that it would otherwise cause high algal concentrations. Macroinvertebrate communities are strongly connected to the river benthos and are influenced by increased benthic autotrophic and heterotrophic production through changes in oxygen dynamics and food and habitat quality, and for this reason could be a useful secondary indicator. The River Macroinvertebrate Index of Biotic Integrity (MIBI) had a relatively strong correlation with phosphorus (R-squared 0.31; see Weigel and Dimick, 2011 for calculations of metrics). However, for simplicity of focusing the phosphorus response indicators on primary production, we do not propose to include macroinvertebrates at this time. Nonetheless, macroinvertebrates may be a useful additional indicator in certain river systems.
- **Benthic chlorophyll *a*.** Benthic chlorophyll *a* is difficult to systematically collect in rivers because adequate substrate is usually lacking to collect a sample. Suspended chlorophyll *a* is a better river indicator.
- **Diatoms – DNI, DBI, and DPI.** The Diatom Nutrient Index (DNI), Diatom Biotic Index (DBI), and Diatom Phosphorus Index (DPI) were developed for wadeable streams and are not appropriate to apply to river sites.
- **Secchi depth.** A Secchi tube clarity reading is not required for purposes of biological assessments or the combined approach for phosphorus. Since a Secchi depth reading frequently reflects suspended

sediment as well as algae growth, a chlorophyll *a* sample is a more direct measure of biological response to phosphorus. However, a Secchi tube may be included as part of a regular sampling regimen if established by monitoring protocols to provide additional context.

- **Algal toxins.** While production of algal toxins can be a result of high TP concentrations, algal toxins are not recommended as a primary phosphorus response indicator. High algal toxins are more likely to be a problem in rivers than in streams. However, at the current time, protocols for assessing algal toxins are insufficient. An algal toxin sample may be collected and analyzed in a river if a problem is suspected, and the analysis may be used as supplementary evidence of a problem.

5.6 IMPOUNDMENT PHOSPHORUS RESPONSE INDICATORS

5.6.1 Chlorophyll *a* for Recreation use

For impounded flowing waters, suspended chlorophyll *a* is established as the only statewide phosphorus response indicator. Algal productivity is assessed in impoundments using the same metric as for shallow lakes and reservoirs and for rivers: a waterbody within the combined assessment range will be listed as impaired for phosphorus unless it exceeds 20 ug/L chlorophyll *a* for fewer than 30 percent of days during the summer sampling period of July 15 to September 15. This indicator is applied regardless of whether the impoundment is on a river or a wadeable stream.

5.6.2 Other metrics not selected

Other potential phosphorus response indicators, including benthic algal biomass, benthic diatom community structure, lake aquatic plant index, or the macroinvertebrate IBI, are generally not applicable to impounded flowing waters for two main reasons. First, the datasets used to develop these criteria did not include impounded flowing waters. Second, several characteristics of impounded flowing waters, including depth, velocity, and substrate, differ from natural lakes and free-flowing rivers enough to influence habitat conditions for plant and animal communities. However, these or other metrics may be required by the department on a case-by-case basis depending on a given site's characteristics.

5.7 STREAM PHOSPHORUS RESPONSE INDICATORS

5.7.1 Nutrient Impacts Dataset

The department used the Nutrient Impacts Dataset (Version 2) for development of stream phosphorus response indicators. To determine which stream metrics have the strongest correlation to TP concentrations, and thus which would best represent the variables in the conceptual model, WDNR assembled existing data from three different studies spanning ten years. The 197 stream sites that were used for this analysis included 171 sites from the 2001-03 wadeable stream nutrient impacts study (Robertson et al. 2006), 8 sites from WDNR's 2007-09 watershed rotation study, and 18 sites from WDNR's 2011 high N:P ratio study. The sites were selected to span the range of nutrient conditions and to minimize the correlation between total phosphorus and total nitrogen. The dataset included a variety of metrics for fish, macroinvertebrates, and diatoms, and includes sites from each Natural Community and Ecoregion. From this dataset the department determined that benthic algae had the strongest correlation with TP in wadeable streams. The dataset was further used in development of the Diatom Phosphorus Index described in this section.

5.7.2 Benthic algal biomass & diatom taxa

For streams, primary productivity can be measured in one or both of the following ways. To maximize efficiency for making assessment decisions, the "viewing bucket" method for algal biomass is recommended as the first step in assessing primary productivity. If these results are conclusive, as described below, no further analysis is required. If the results are inconclusive (mid-range scores), further analysis of the diatom community is required to determine whether the stream is exhibiting a TP response.

a. Viewing Bucket for algal biomass

A visual assessment of benthic algal biomass in streams using a quantifiable system such as a viewing bucket is an efficient and appropriate screening tool to determine whether a site clearly is, or is not, exhibiting a nutrient response. High TP can be expected to result in greater biomass and coverage of benthic algae in streams. The viewing bucket method is included in the *U.S. EPA's Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers* (Barbour et al. 1999) and is used in several states' monitoring programs.

The method will be employed during evaluation of habitat assessment transects. Benthic algal biomass will be observed and characterized on a grid with a minimum of 25 points with the viewing bucket (Figure 25). This will be done once on each of the twelve habitat transects (WDNR 2002) for a reach, staggered across the stream from left to right. Scores from each transect will then be averaged for the reach.

The assessment should be conducted during the growing season (July, Aug, or Sept) during baseflow conditions, with the first viewing bucket assessment in July or August, and second (if needed) in August or September. Because scouring during stream spate events may reduce algal biomass, sampling should be avoided within 14 to 21 days of a storm event.

Figure 25. An example viewing bucket from Rhode Island Dept. of Environmental Management. Photograph by A. Patterson.

Thresholds for algal biomass as evaluated with the viewing bucket method reflect the expectation that higher TP levels will lead to higher algal biomass. The viewing bucket scoring scale is from 0 (low biomass) to 3 (high biomass) (Table 10). If the average algal biomass score for the reach is less than 1, the stream is not impaired by TP and there is no need for further primary production assessment. If the algal biomass score is greater than 2, the stream is impaired by TP and no further assessment is necessary. If the algal biomass score is between 1 and 2, further primary production assessment via the Diatom Phosphorus Index (DPI) is needed.

Table 10. Stream benthic algal biomass phosphorus response indicator using viewing bucket method.

Benthic algal biomass, viewing bucket score (0-3)	Attainment decision	
	Aquatic Life Use	**Recreation Use**
< 1	Attained[1]	Attained
1 - 2	Inconclusive; assess benthic diatoms using DPI	
> 2	Not attained	Not attained

[1] If the mean score is <1 but 20% or more of individual transect points score a 3, a benthic diatom assessment under par. (b) is required to make an attainment determination.

The viewing bucket method can also be used to assess whether a stream is attaining its recreation use, as recreation is also impacted by algal growth. A stream's recreation use is considered attained if the viewing bucket score is at 2 or below.

b. Diatom Phosphorus Index (DPI)

Diatoms are a form of algae with a silicate shell with many species that tend to be found on stream beds or clinging as a brown substance to filamentous algae, such as Cladophora. They are found in both freshwater and marine waters and in many environments play a very substantial role in primary productivity within the system. Analysis of diatoms has been used for water quality analysis around the world. Various species have been identified as tolerant or sensitive to various stressors, including nutrients.

In development of phosphorus criteria for wadeable streams, WDNR used three indices to evaluate diatom community responses to phosphorus: the Diatom Nutrient Index (DNI), the Diatom Siltation Index (DSI), and the Diatom Biotic Index (DBI) (Robertson et al. 2006). Because these indices are primarily based on literature-derived tolerance values that are not specific to phosphorus, we decided to develop a new method that is specific to phosphorus and calibrated to Wisconsin diatom data, herein referred to as the Diatom Phosphorus Index (DPI).

The DPI is based on a statistical method called Weighted Averaging (WA; ter Braak and van Dam 1989). This method can be used to determine whether the diatom community at an assessment site resembles the community that is typically found at sites meeting the stream TP criterion. The TP criterion is based on breakpoints in the relationships between TP and diatom (and other biological) metrics, and as such represents the level of TP where the biological community changes the most.

WA estimates species-specific environmental preferences (optima) as the average value of an environmental variable (in this case, TP) where a species occurs, weighted by its relative abundance. The DPI at a site is then estimated as the weighted average of the TP optima of all the species present at that site. WA was developed to infer paleo-limnological characteristics such as pH, temperature, and TP (reviewed in Juggins and Birks 2012), and has also been used to develop a stream diatom nutrient index in New Jersey (Ponader 2007).

A WA model was developed from the Nutrient Impacts (Version 2) Dataset described above. Diatom and nutrient samples were collected in 2001-03 and 2011 using methods described in Robertson et al. (2006). Diatom samples were collected in September, and nutrient samples were collected monthly from May-Oct. Models using various subsets of nutrient samples during and prior to September were evaluated to determine whether they were better predictors of diatom community structure than the entire growing season, but the median of all six monthly samples was the best predictor. Only taxa with at least five occurrences (n=156) were used in the model development.

The WA model was fit using the WA function in the rioja package (Juggins 2014) in R. Prediction errors were estimated by leave-one-out cross-validation. The cross-validated r^2 is 0.49, which means that TP explains about half of the variation in diatom community structure among sites (Figure 26). The root mean square error of prediction (RMSEP) is 62%, which means that the average DPI differs from the measured TP by 62%. The residual variation in this relationship probably reflects sampling error in both TP and diatoms.

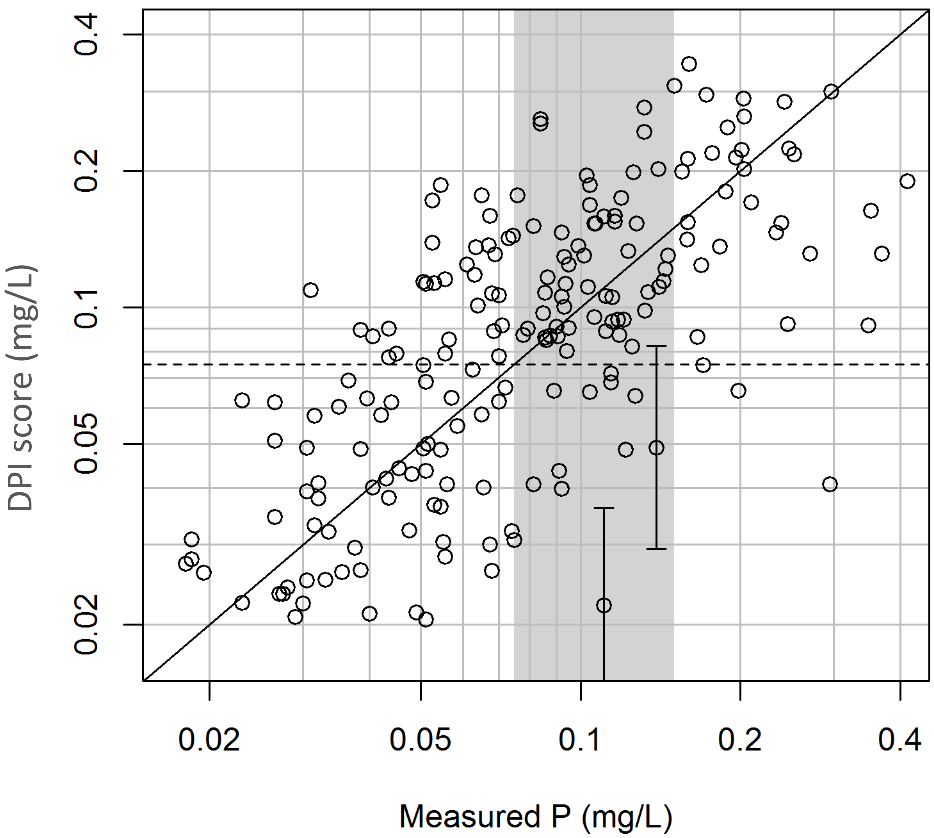
For purposes of assessing attainment of the diatom phosphorus response indicator:

- If only a single diatom survey is available from a sampling station, the department would not list a waterbody as impaired for phosphorus if it is 90% confident that the diatom community is not impaired. A bootstrapping procedure[3] was used to estimate confidence intervals around DPI values. If the upper 80% confidence limit of DPI is < 75 µg/L, we would be 90% confident that the diatom community is not impaired. Among the 68 sites in the model dataset where biological confirmation would be relevant (measured TP is 75-150 µg/L), only two would be considered not impaired through the perspective of the diatom community.
- If more than one diatom survey is available from a sampling station, the DPI scores are averaged and the mean DPI score is compared directly to the threshold of 75 ug/L without using confidence intervals. This is because the bootstrapping procedures required to calculate the confidence intervals are not practicable with more than one sample. In this case, averaging the scores addresses sample variability.

For assessment purposes, the DPI should be used only in conjunction with TP, not as a stand-alone assessment metric. It has not been shown to be sensitive to a broader range of environmental stressors than phosphorus. However, high algal viewing bucket scores may be used to list a water as impaired regardless of TP concentrations.

[3] In rioja, the predict function with sse=TRUE estimates standard errors for each site (v1), which reflect how much the inferred P varies across the bootstrapped samples.

Figure 26. Measured TP concentrations vs. DPI score from Weighted Average (WA) model (R2 = 0.49). Note log scales on both axes. Gray area is TP range where biological confirmation may be used. Error bars are 80% confidence intervals on two example points.



5.7.3 Other metrics not selected

For streams, we determined that primary production metrics were the most appropriate as phosphorus response indicators, and upper-level indicators did not add clear value to assessment of phosphorus response. We considered both macroinvertebrate and fish metrics, as described below, but determined that the relationships between these metrics and phosphorus, as assessed using currently available data, were not strong enough to include as response indicators.

The following primary productivity metrics were considered for development of stream phosphorus response indicators but were not selected.

- **Benthic chlorophyll *a*.** Benthic chlorophyll *a* is a useful metric in streams; however, physical sampling of chlorophyll *a* is highly variable in streams as growth is patchy and strongly influenced by substrate type and substrate selected for sampling (i.e. selection bias) making the development of a clear relationship with TP and an appropriate threshold difficult. Additionally, the viewing bucket method incorporates benthic algae but over a larger portion of the streambed. Though this metric will not be required in code at this time, staff are able to collect benthic algae via a rock scrape for chlorophyll *a* analysis, to be used as supplemental information and to help build WDNR's dataset on benthic chlorophyll *a*. Obtaining additional data will help the department refine benthic chlorophyll *a* thresholds in the future.

- **Suspended chlorophyll *a*.** Streams do not typically have high suspended (sestonic) chlorophyll *a* levels, so a grab sample of suspended chlorophyll *a* is not needed for stream sites. Suspended chlorophyll *a* is a more appropriate indicator for rivers.
- **Algal toxins**. While production of algal toxins can be a result of high TP concentrations, algal toxins are not recommended as a primary phosphorus response indicator. High algal toxins are very rarely a problem in streams. At the current time, protocols for assessing algal toxins are insufficient. However, an algal toxin sample may be collected and analyzed in a stream if a problem is suspected, and the analysis may be used as supplementary evidence of a problem.
- **Diatom Nutrient Index (DNI) and Diatom Biotic Index (DBI).** The weighted average Diatom Phosphorus Index (DPI) was selected over the Diatom Nutrient Index (DNI) or Diatom Biotic Index (DBI) because it shows a stronger correlation with phosphorus.

Macroinvertebrate and Fish Metrics
To support the continued development of stream nutrient criteria and phosphorus response indicators, WDNR conducted an extensive set of analyses on biotic responses to nutrients. This study is described in detail in a report titled "Evaluation of the relative effects of phosphorus and nitrogen on stream biological community structure" (Diebel 2015). One of the major conclusions of this study is that nitrogen and phosphorus have independent and statistically significant effects on the community structure of all taxonomic groups, but the strength of those effects is relatively weak compared to other environmental variables, except for the effect of P on diatoms, which is strong. In particular, both macroinvertebrate and fish communities are more strongly influenced by stream size, temperature, and conductivity than by phosphorus. These higher taxonomic groups are useful indicators of overall biological integrity in streams, but do not match the specificity of primary producers as a phosphorus response indicator.

5.8 IMPAIRED WATERS LISTING USING PHOSPHORUS RESPONSE INDICATORS

The department provided a data analysis to the External Stakeholder Committee in 2016, to provide information on the number of waterbodies that would be kept off of the section 303(d) impaired waters list using the combined approach to apply phosphorus response indicators. That analysis is summarized here. As discussed with the stakeholder committee, the percent of waterbodies that exceed the statewide P criteria but are not experiencing a biological response is small. This indicates that the statewide P criteria are set at a level that is not overly protective for most waterbodies. The following datasets contain a relatively small portion of the waterbodies in the state. As the phosphorus response indicators are applied more broadly, additional waterbodies are expected to be determined to be attaining these indicators.

From the 2016 data analysis:
Streams: There are 182 stream sites that have been evaluated for P for which diatom analysis has also been conducted. Of those 182 sites, 67 sites exceed the phosphorus criterion but are within the P range at which the combined approach can be applied. Six of these sites attained the diatom phosphorus response threshold and would therefore be removed from the impaired waters list for P or would not be listed for P when they otherwise would have been.

Rivers: There are 28 river sites that have been evaluated for P for which chlorophyll *a* data have also been assessed. Of these, 11 exceed the P criterion but are within the range at which the combined approach can be applied. Two of these attain the phosphorus response indicator for frequency of moderate algae levels, and would therefore be removed from the impaired waters list for P or would not be listed for P when they otherwise would have been.

Lakes: There are 161 lakes that have P data and also have data for the three main phosphorus response indicators: frequency of moderate algal levels (to protect recreation use), chlorophyll *a* concentration (to protect aquatic life use), and the plant phosphorus response tool (aquatic life). Of these 161 lakes, 28 exceed the P criterion but are within the P range at which the combined approach can be applied. Eight of those lakes attain all three phosphorus response indicators and would therefore be removed from the impaired waters list for P or would not be listed for P when they otherwise would have been.

6. Applying Biological Metrics to Develop Site-specific Criteria for Phosphorus

6.1 STATEWIDE PHOSPHORUS CRITERIA AND THE NEED FOR SITE-SPECIFIC CRITERIA

Concurrent with this rule package (WY-23-13, referred to here as the assessments rule), a second rule package (WT-17-12) is underway to establish a process for deriving site-specific criteria for phosphorus for individual waterbodies when needed. Rule package WT-17-12 would create a new rule, ch. NR 119, to house this process. The SSC rule package cross-references the biologically-based metrics contained in the assessments rule that are discussed within this Technical Support Document. Therefore, a short discussion is included here regarding the interplay between the biological metrics in this rule and the proposed SSC rule[4].

Wisconsin promulgated its statewide phosphorus criteria in December 2010 following the publication of ch. NR 102, Wis. Adm. Code. In reviewing statewide data trends, the department has concluded that the statewide phosphorus water quality criteria are appropriately protective in most cases. However, there may be some instances for specific waterbodies where the applicable statewide phosphorus criterion is more stringent than necessary to protect the designated uses of the waterbody in question. Alternatively, there may be some waterbodies, such as certain impounded flowing waters, that are not being adequately protected by the current phosphorus criteria. In such cases, federal and state law allow for development of site-specific criteria—criteria that are applicable only to a specific waterbody or waterbody segment, based on site-specific circumstances—which are more appropriate for individual waterbodies. After taking effect, an approved SSC becomes the applicable water quality standard for the approved waterbody or segment.

Authority for developing SSC for any substance is already contained in s. 281.15, Stats. The proposed SSC rule does not create additional authority; it establishes a process under which SSC development can be carried out. Establishment of this process will provide consistency and transparency, specifying the type of information needed to make an approvable demonstration that an SSC is appropriate for an individual waterbody.

6.2 UNDERLYING PRINCIPLE: PROTECTING DESIGNATED USES

Site-specific criteria must be set at levels that are protective of a waterbody's designated uses. In Wisconsin, the main uses associated with phosphorus are Recreation and Aquatic Life (which is further divided into several subcategories). The statewide phosphorus criteria were developed to be protective of both of these types of uses. Similarly, during development of any site-specific criteria, one of the critical goals is to select a criterion that maintains or improves protection of an individual waterbody's uses, based on the waterbody's specific ecological context and response to phosphorus. For example, some waterbodies may naturally be less sensitive to phosphorus, and can therefore assimilate more phosphorus than others without adverse impacts to their uses. Other waterbodies may be more sensitive to phosphorus

[4] The SSC rule package does not require a separate Technical Support Document because it does not consitute a new water quality standard in and of itself; it sets a process for deriving criteria. Each individual SSC developed using the new process (or any other process) would still be approved separately by EPA, along with its own analysis.

and experience biological responses and use impairments at lower levels than usually expected. In general, it may be appropriate to derive a site-specific criterion for phosphorus in either of the following scenarios:

1) The statewide phosphorus criterion is not stringent enough to protect a waterbody's designated uses. Despite the applicable statewide phosphorus criterion being met, the designated uses of a given water or waterbody segment are not attained.
   → **In this case, a *more stringent* phosphorus SSC may be needed.**

2) The statewide phosphorus criterion is more stringent than reasonably necessary to assure attainment of the designated uses for the waterbody in question and adjacent downstream waters (if applicable).
   → **In this case, a *less stringent* phosphorus SSC may be appropriate.**

It is important to note the underlying premise that a criterion may be more or less stringent but equally protective of the designated uses. The stringency needed is based on the sensitivity of the waterbody in question. A less stringent criterion may be equally protective where, due to the specific chemistry, geology, or morphology of a site, the biological community of a waterbody exhibits less sensitivity or response to phosphorus than most waterbodies. This may include areas of the state where naturally high levels of phosphorus have always existed, due to the underlying geology, and the biology is adapted to those levels. Conversely, some sites may need a more stringent phosphorus criteria because they are naturally more sensitive to phosphorus impacts.

It is also important to clarify that an SSC is a water quality standard to protect aquatic life, recreation, and other uses, rather than a compliance tool for permittees. Compliance tools for meeting phosphorus permit limits include water quality trading and adaptive management. If a permittee cannot comply with permit limits because it would cause economic hardship, an individual or multi-discharger phosphorus variance is available. A waterbody is only eligible for an SSC if an adjusted phosphorus criterion is appropriate based on the biological responses of the system.

6.3 USING BIOLOGICAL METRICS TO REPRESENT DESIGNATED USE ATTAINMENT

To determine whether a waterbody's designated uses are being met, certain biological metrics are used to indicate the ecosystem's response to phosphorus and whether uses are being impaired. The biological metrics are different for different waterbody types. Two types of biological metrics that are delineated in ch. NR 102 revisions that are described in this Technical Support Document are integral to SSC development:

1) **Phosphorus response indicators.** (proposed ch. NR 102.07) Phosphorus response indicators are based on biological metrics that are particularly responsive to phosphorus, such as algae (as measured through chlorophyll *a*) and aquatic plants. They are used to determine the effects of phosphorus within a waterbody, including attainment of phosphorus criteria and designated uses.

2) **Biological assessment thresholds.** (proposed ch. NR 102 Subch. III) Biological assessment thresholds are based on an assessment of the overall health of key biological communities, such as aquatic plants, algae, fish, or aquatic insects, which is used to determine support of aquatic life or recreation designated uses.

The phosphorus response indicators and biological assessment thresholds for each waterbody type are found in ch. NR 102 proposed Subchapter III. They are also detailed in sections 4 and 5 of this Technical Support Document. They are based on the following metrics:

- **Lakes/reservoirs:**
  - Phosphorus response indicators: suspended chlorophyll *a* (indicating algae growth, for both recreation and aquatic life uses) and aquatic plants (macrophytes), plus oxythermal habitat criteria for two-story fishery lakes
  - Biological assessment thresholds: aquatic plants and chlorophyll *a* for recreation and aquatic life uses
- **Streams:**
  - Phosphorus response indicators: benthic algal biomass and benthic diatom taxa (diatoms are a type of hard-bodied algae that grows on the substrate)
  - Biological assessment thresholds*: aquatic insects (macroinvertebrates) and fish
- **Rivers:**
  - Phosphorus response indicator: suspended chlorophyll *a*
  - Biological assessment thresholds*: aquatic insects and fish
- **Impoundments:**
  - Phosphorus response indicator: suspended chlorophyll *a*
  - Biological assessment thresholds: chlorophyll *a* for recreation use

* Denotes assessment thresholds currently in guidance, used to interpret narrative biological assessment thresholds in code.

Additional indicators may also be required to determine the health of the biotic community, and the attainment of designated uses.

For purposes of determining the appropriateness of SSC, at least two years of recent data are required for each metric to account for any temporal variability in the aquatic system. Historical data should also be analyzed if available to assess temporal variability. For a less-stringent SSC determination, the proposal must demonstrate that the proposed SSC is protective of the designated uses not only in the segment itself but also in any downstream waters. Therefore, sampling for biological metrics is required at multiple monitoring sites downstream of the SSC segment.

Once the complete dataset is obtained, modeling may be needed as part of the data analysis. Modeling techniques will need to be determined on a case-by-case basis. For instance, models such as BATHTUB are frequently be used by U.S. EPA and the department to validate appropriate lake/reservoir targets (United States Environmental Protection Agency, April 2000). Modeling is typically only available for chlorophyll *a* predictions, and would not be applied to other types of biological metrics.

6.4 APPLYING BIOLOGICAL METRICS FOR SSC DETERMINATION

As described in proposed ch. NR 119, the phosphorus response indicators and biological assessment thresholds are applied in the following ways to determine SSC eligibility:

Less stringent SSC: A waterbody or segment may be eligible for an SSC that is less stringent than the statewide phosphorus criterion in the following types of cases:

(1) The waterbody is exceeding its statewide phosphorus criterion but all of its phosphorus response indicators and biological assessment thresholds are attained. This can typically be demonstrated using only field data without modeling.

(2) If a waterbody is exceeding its statewide phosphorus criterion, and one or more of its phosphorus response indicators or biological assessment thresholds are not attained, a less-stringent SSC could be appropriate if a modeling analysis demonstrates that the phosphorus response indicators are expected to be attained if the waterbody's phosphorus concentration is sufficiently reduced to attain a proposed SSC that is less stringent than the statewide phosphorus criterion. (Example: Certain reservoirs with a statewide phosphorus criterion of 30-40 ug/L may fit in this category. For instance, a reservoir that is exceeding its statewide TP criterion of 40 ug/L with a current phosphorus level of 70 ug/L is also not attaining its chlorophyll *a* threshold. In this case, modeling may demonstrate that an SSC of 50 ug/L TP should be sufficient to attain its chlorophyll *a* threshold; it does not need to attain 40 ug/L TP to reach its biological goals.)

(3) A less stringent SSC may be appropriate if a waterbody is not attaining the statewide phosphorus criterion because the natural background phosphorus concentration is higher than the statewide phosphorus criterion.

More stringent SSC: A more stringent SSC may be appropriate in the following types of cases:

(1) The waterbody attains its statewide phosphorus criterion but does not attain one or more of its phosphorus response indicators or biological assessment thresholds. Modeling may be required to determine at what level the SSC should be set to attain its biological metrics. However, a more stringent SSC is not appropriate if a biological assessment threshold or phosphorus response indicator is not attained due to reasons other than phosphorus.

(2) A more stringent SSC may be appropriate even if a waterbody attains its statewide phosphorus criterion, phosphorus response indicators, and biological assessment thresholds in cases when it is demonstrated that a more stringent SSC than the statewide phosphorus criterion is necessary to maintain attainment of any of these indicators and the level necessary can be demonstrated through modeling.
(Example: Certain impounded flowing waters with a statewide phosphorus criterion of 100 ug/L may fit in this category. For instance, if an impounded flowing water currently has a phosphorus concentration of 50 ug/L TP and is attaining its biological metrics, a demonstration may show that an SSC of 70 ug/L TP is needed because its biological metrics will no longer be attained above that level.)

Proposed chapter NR 119 describes these processes in detail.

7. References

7.1 GENERAL REFERENCES

Allen, J.D. and Castillo, M.M. 2007. Stream Ecology: Structure and Function of Running Waters. 2nd Edition. Springer.

Dodds, W.K. 2007. Trophic state, eutrophication and nutrient criteria in streams. *Trends in Ecology and Evolution.* 22(12): 669-676.

Dodds, W.K. and Welch, E.B. 2000. Establishing nutrient criteria in streams. *Journal of the North American Benthological Society.* 19(1):186-196.

Cross, W.F., Wallace, J.B., Rosemond, A.D. and Eggert, S.L. 2006. Whole-system nutrient enrichment increases secondary production in a detritus-based ecosystem. *Ecology.* 87:1556-1565.

Hill, W.R., Ryon, M.G. and Schilling, E.M. 1995. Light limitation in a stream ecosystem: responses by primary producers and consumers. *Ecology.* 76:1297-1309.

Rosemond, A.D., Mulholland, P.J. and Elwood, J.W. 1993. Top-down and bottom-up control of stream periphyton:Effects of nutrients and herbivores. *Ecology.* 74:1364-1280.

Schindler, D.W. 1974. Eutrophication and recovery in experimental lakes: Implications for lake management. *Science.* 184:897-899.

Suplee, M.W. Watson, V., Tepley, M., McKee, H. 2009. How green is too green? Public opinion of what constitutes undesirable algae levels in streams. *Journal of the American Water Resources Association.* 45:123-140.

U.S. EPA. April 2000. Nutrient Criteria Technical Guidance Manual: Lakes and Reservoirs (1 ed.). Washington D.C: Office of Water.

U.S. EPA. July 2000. Nutrient Criteria Technical Guidance Manual: River and Streams. Environmental Protection. EPA-822-B-00-002: Office of Water.

U.S. EPA. 2010. Using Stressor-response Relationships to Derive Numeric Nutrient Criteria. Office of Water, U.S. Environmental Protection Agency. EPA-820-S-10-001.

U.S. EPA. 2013. Guiding Principles on an Optional Approach for Developing and Implementing a Numeric Nutrient Criterion that Integrates Causal and Response Parameters. Office of Water, Environmental Protection Agency. EPA-820-F-13-039. http://www2.epa.gov/sites/production/files/2013-09/documents/guiding-principles.pdf

7.2 LAKE/RESERVOIR REFERENCES

Akasaka, M., Takamura, N., Mitsuhashi, H., Kadono, Y., 2010. Effects of land use on aquatic macrophyte diversity and water quality of ponds. Freshw. Biol. 55, 909-922.

Alahuhta, J. and J. Aroviita. 2016. Quantifying the relative importance of natural variables, human disturbance and spatial processes in ecological status indicators of boreal lakes. Ecological Indicators 63, 240-248.

Carlson, R.E. 1977. A trophic state index for lakes. Limnology and Oceanography. 22:2 361--369.

Danz, N.P., Niemi, G.J., Regal, R.R., Hollenhorst, T., Johnson, L.B., Hanowski, J.M., Axler, R.P., Ciborowski, J.J.H., Hrabik, T., Brady, V.J., Kelly, J.R., Morrice, J.A., Brazner, J.C., Howe, R.W., Johnston, C.A., Host, G.E., 2007. Integrated measures of anthropogenic stress in the U.S. Great lakes watershed. Environ. Monit. Assess. 39, 631-647.

Falcone, J.A., Carlisle, D.M., Weber, L.C., 2010. Quantifying human disturbance in watersheds: variable selection and performance of a GIS-based disturbance index for predicting the biological condition of perennial streams. Ecol. Indic. 10, 264-273.

Fraley, C. and A. E. Raftery. 2002. MCLUST: Software for Model-Based Clustering, Density Estimation and Discriminant Analysis. Washington Univ Seattle Dept of Statistics. ADA459792

Hauxwell, J., S. Knight, K. Wagner, A. Mikulyuk, M. Nault, M. Porzky, and S. Chase. 2010. Recommended baseline monitoring of aquatic plants in Wisconsin: sampling design, field and laboratory procedures, data entry and analysis, and applications. Available from Wisconsin Department of Natural Resources, PUB-SS-1068 2010. Madison, WI.

Heiskary, S.A. and W.W. Walker Jr. (1988). Developing phosphorus criteria for Minnesota Lakes. Lake and Reservoir Management. 4:1, 1-9.

Heiskary, S.A. and C.B. Wilson. 2008. Minnesota’s approach to lake nutrient criteria development. Lake and Reservoir Management. 24:282-297.

Jacobson, P., T. Cross, J. Zandlo, B. N. Carlson and D. Pereira. 2012. The effects of climate change and eutrophication on cisco Coregonus artedi abundance in Minnesota lakes. Advances in Limnology. 63:417-427.

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.

Jeppesen, E., J.P. Jensen, P. Kristensen, M. Søndergaard, E. Mortensen, O. Sortkjær, K. Olrik. 1990. Fish manipulation as a lake restoration tool in shallow, eutrophic, temperate lakes 2: threshold levels, long-term stability and conclusions. Hydrobiologia, 200/201, 219 227.

Lacoul, P. and B. Freedman. 2006. Environmental influences on aquatic plants in freshwater ecosystems. Environmental Reviews. 14:89-136.

Lillie, R.A. and J.W. Mason. 1983. Limnological characteristics of Wisconsin lakes. Tech. Bull. 138. Dept. of Natural Resources, Madison, WI.

Menuz, D.R., Ruesch, A.S., Diebel, M.W., 2013. In: 1:24K Hydrography Attribution Metadata. Wisconsin Department of Natural Resources, Madison.

Mikulyuk, A., J. Hauxwell, P. Rasmussen, S. Knight, K. I. Wagner, M. E. Nault, and D. Ridgely. 2010. Testing a methodology for assessing plant communities in temperate inland lakes. Lake and Reservoir Management 26:54-62.

Mikulyuk, A., S. Sharma, et al. 2011. The relative role of environmental, spatial, and land-use patterns in explaining aquatic macrophyte community composition. Canadian Journal of Fisheries and Aquatic Sciences 68(10): 1778-1789.

Mikulyuk, A., M. Barton, J. Hauxwell, C. Hein, E. Kujawa, K. Minahan, M. E. Nault, D. L. Oele, K. I. Wagner. 2017. A macrophyte bioassessment approach linking taxon-specific tolerance and abundance in north temperate lakes. Journal of Environmental Management 199: 172-180.

Mischan, M.M., Pinho, S.Z.D. and Carvalho, L.R.D. 2011. Determination of a point sufficiently close to the asymptote in nonlinear growth functions. Scientia Agricola 68(1):109-114.

Radomski, P. and T. J. Goeman (2001). Consequences of human lakeshore development on emergent and floating-leaf vegetation abundance. North American Journal of Fisheries Management **21**: 46-61.

Smith, D.G. and Perrone, J.A. 1996. Laboratory experiments to investigate human sensitivity to changes in water clarity. Journal of Environmental Management 48(2):139-154.

Papes¸ M., Vander Zanden, M.J., 2010. Wisconsin Lake Historical Limnological Parameters 1925-2009.

Wilcox, D.A. 1995. Wetland and aquatic macrophytes as indicators of anthropogenic hydrologic disturbance. Natural Areas Journal. 15:240-248.

7.3 STREAM/RIVER REFERENCES

Barbour, M.T., J. Gerritsen, B.D. Snyder, and J.B. Strickland. 1999. Rapid bioassessment protocols for use in wadeable streams and rivers. Second Edition. EPA 841-B-99-002. Office of Water, U.S. Environmental Protection Agency, Washington, DC.

Diebel, M.W. 2015. Evaluation of the relative effects of phosphorus and nitrogen on stream biological community structure. Wisconsin Department of Natural Resources, Madison, WI. DNR PUB-WY-3200-2015-03.

Heirskary, S.A. and Walker, Jr., W.W. 1988. Developing phosphorus criteria for Minnesota lakes. *Lake and Reservoir Management.* 4(1):1-9.

Juggins, S., & Birks, H. J. B. 2012. Quantitative environmental reconstructions from biological data. In Tracking environmental change using lake sediments (pp. 431-494). Springer, Netherlands.

Juggins S. 2014. rioja package documentation. http://cran.r-project.org/web/packages/rioja/index.html.

Lyons, J. 1992. Using the index of biotic integrity to measure environmental quality in warmwater streams of Wisconsin. U.S. Forest Service Technical Report NC-149. St. Paul, Minnesota.

Lyons, J., Wang, L. and Simonsen, T.D. 1996. Development and validation of an index of biotic integrity for coldwater streams in Wisconsin. *North American Journal of Fisheries management.* 16:2, 241-256.

Lyons, J., Piette, R.R. and Niedermeyer, K.W. 2001. Development, validation, and application of a fish based Index Biotic Integrity for Wisconsin's large warmwater rivers. *Transactions of the North American Fisheries Society.* 130:1077-1094.

Lyons, J. 2006. A fish-based index of biotic integrity to assess intermittent headwater streams in Wisconsin, USA. *Environmental Monitoring and Assessment.* 122: 23-258.

Lyons, J. 2012. Development and validation of two fish-based indices of biotic integrity for assessing perennial coolwater streams in Wisconsin, USA. *Ecological Indicators.* 23:402-412.

MBI (Midwest Biodiversity Institute), 2014. Refining state water quality monitoring programs and Aquatic Life Uses: Evaluation of the Wisconsin DNR Bioassessment Program. MBI Technical Memorandum 2014-2-1.

Ponader, KC, DF Charles, and TJ Belton. 2007. Diatom based TP and TN inference models and indices for monitoring nutrient enrichment of New Jersey streams. Ecological Indicators 7:79–93.

Robertson, D. M., D. J. Graczyk, P. J. Garrison, L. Wang, G. LaLiberte, and R. Bannerman. 2006. Nutrient concentrations and their relations to the biotic integrity of wadeable streams in Wisconsin. U.S. Geological Survey Professional Paper 1722, Middleton, WI.

Robertson, D.M., Weigel, B.W. and Graczyk. 2008. Nutrient concentrations and their relations to the benthic integrity of nonwadeable rivers in Wisconsin. Professional Paper 1754. United States Geological Survey, Reston, Virginia.

ter Braak, C. J. F. & van Dam, H. 1989. Inferring pH from diatoms: a comparison of old and new calibration methods. Hydrobiologia 178:209-223.

WDNR. 2002. Guidelines for Evaluating Habitat of Wadeable Streams. Wisconsin Department of Natural Resources, Bureau of Fisheries Management and Habitat Protection. Madison, WI.

WDNR. 2010. Wisconsin Phosphorus Water Quality Standards Criteria: Technical Support Document. Unpublished white paper. Wisconsin Department of Natural Resources. Madison, WI.

WDNR, 2019. Wisconsin 2020 Consolidated Assessment and Listing Methodology (WisCALM). Wisconsin Department of Natural Resources, Bureau of Water Quality Program Guidance. Madison, WI. https://dnr.wi.gov/topic/surfacewater/assessments.html

Weigel, B.M. 2003. Development of stream macroinvertebrate models that predict watershed and local stressors in Wisconsin. *Journal of the North American Benthological Society.* 22(1):123-142.

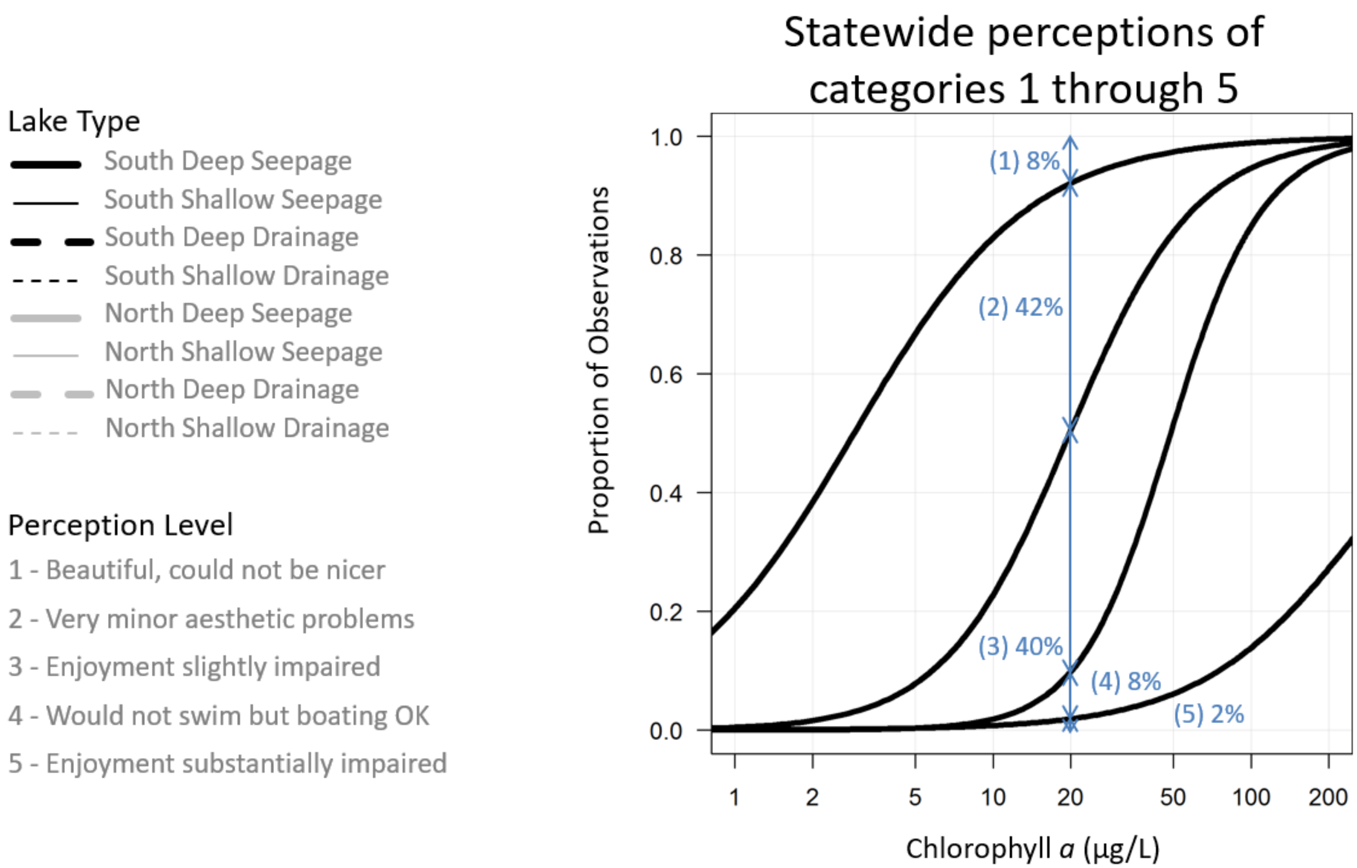
Weigel, B.M. & Dimick, J.J. 2011. Development, validation, and application of a macroinvertebrate-based Index of Biotic Integrity for nonwadeable rivers of Wisconsin. *Journal of the American Benthological Society.* 303(3):665-679

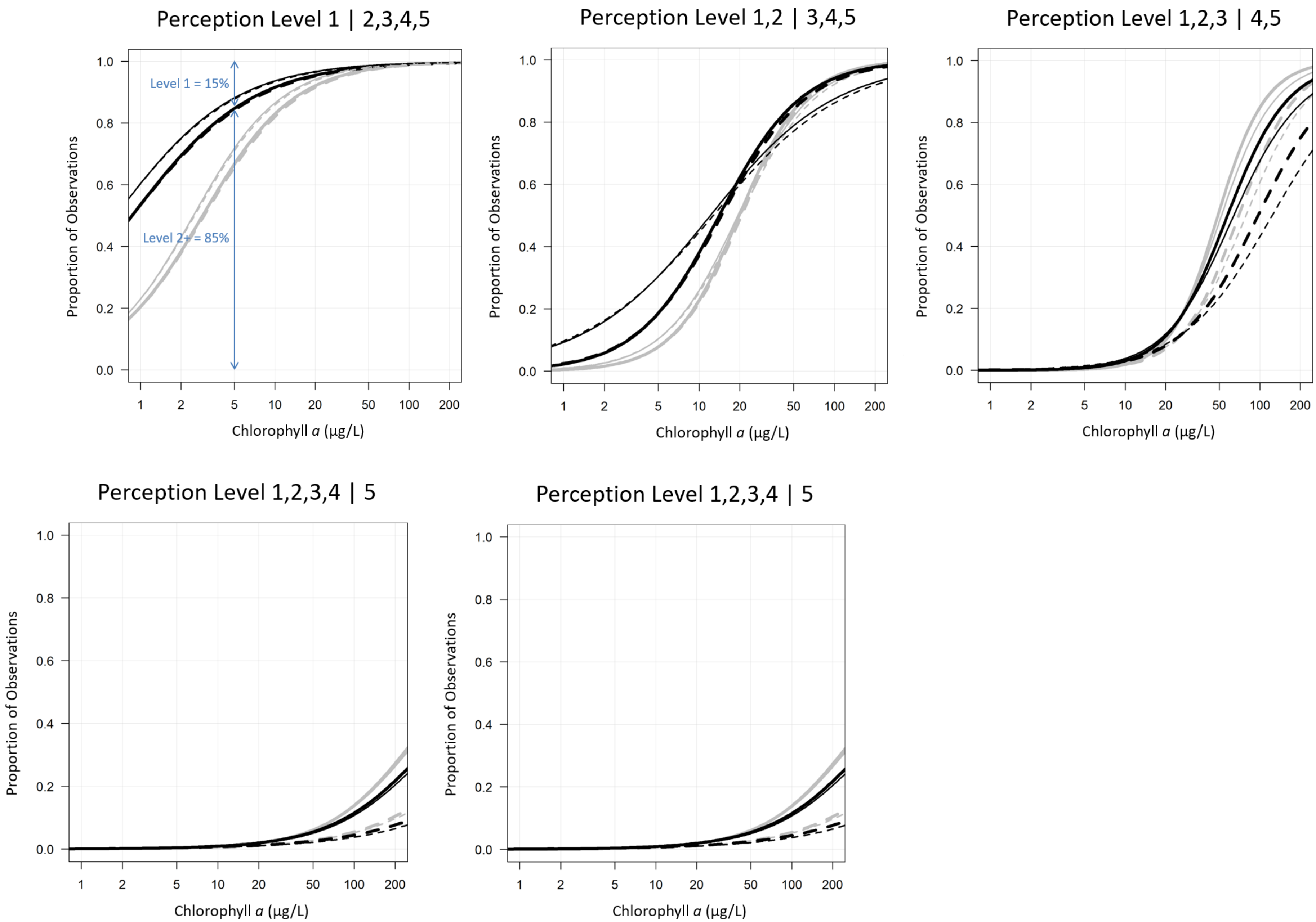
U.S. EPA. 2011. A Primer on Using Biological Assessments to Support Water Quality Management. Office of Water, Environmental Protection Agency. EPA-810-R-11-01. http://water.epa.gov/scitech/swguidance/standards/criteria/aqlife/biocriteria/upload/primer_update.pdf

Appendices

APPENDIX A. Chlorophyll *a* user perception surveys – Lake type comparison

Section 4.4.1. describes how 20 ug/L chlorophyll a was selected for use in the assessment thresholds for frequency of moderate algal levels to protect recreation in lakes, reservoirs, and impounded flowing waters. Models were fit for nine lake classes and for all lakes combined. This appendix shows the results for all lakes combined (this page) and for the nine lake classes individually (next page). The nine lake classes did not exhibit enough variation to warrant separation of results based on lake class.




Perception Level 1 | 2,3,4,5
Level 1 = 15%
Level 2+ = 85%
Proportion of Observations
Chlorophyll a (µg/L)
Perception Level 1,2 | 3,4,5
Proportion of Observations
Chlorophyll a (µg/L)
Perception Level 1,2,3 | 4,5
Proportion of Observations
Chlorophyll a (µg/L)
Perception Level 1,2,3,4 | 5
Proportion of Observations
Chlorophyll a (µg/L)
Perception Level 1,2,3,4 | 5
Proportion of Observations
Chlorophyll a (µg/L)

APPENDIX B. Weather-controlled phosphorus concentrations

Estimating long-term median total phosphorus concentrations in streams by controlling for seasonal and weather variation

Alex Latzka, Matt Diebel, Francisco Guerrero-Bolano, Kristi Minahan
11-1-2019

Motivations and technical considerations

The department is documenting its phosphorus evaluation procedures in s. NR 102.07(2)(a), Wis. Adm. Code. These procedures allow for calculating a weather-controlled phosphorus concentration to be used for assessing whether a waterbody is attaining its phosphorus criterion. The department has developed the Phosphorus Mixed Effects Regression (PhosMER) model to calculate the weather-controlled median phosphorus concentration of streams or rivers. The department plans to make the PhosMER model available on its website in a format that can be easily used by external parties. Similar tools may be developed for calculating weather-controlled lake means in the future.

Why is it important to calculate weather-controlled phosphorus concentrations?
Phosphorus samples collected during a one- or two-year period may be heavily influenced by a wet or dry year, or by the timing of sampling. Although samples are evenly spaced across the season, if these happen to fall just after storms or after a long dry spell, the mean or median of those samples may be skewed higher or lower than the waterbody's true mean or median. Similarly, the confidence intervals calculated from such samples may also be skewed and not necessarily contain the true parameter value. For a more robust assessment, i.e. less sensitive to the random fluctuations of the weather, we want to estimate the true median and the likelihood to find it within a given range regardless of temporary weather conditions.

A tool such as PhosMER uses the sampled phosphorus data and weather at a site to determine how much phosphorus is typically delivered to a specific waterbody with different levels of precipitation. This provides a long-term phosphorus-precipitation correlation for the site. Based on this correlation and the long-term weather records available for each site, the data used in the estimation of the median phosphorus concentration can be significantly expanded. Due to the statistical nature of the method, the expanded data would not introduce bias in median estimates and will help to narrow confidence intervals to provide more definitive assessments where possible.

> The current assessment methodology could be improved by the calculation of weather-controlled phosphorus concentrations in the following ways:
> - Control for weather variability.
> - Allow the use of more samples without biasing median estimates.
> - Narrow confidence intervals to provide more definitive assessments where possible.

Identifying assessment sites versus non-assessment sites
The department has minimum data requirements for assessing a waterbody's phosphorus concentration. For streams and rivers, at least six samples are required over a period of one year, to be taken monthly May to October. If a monthly sample is missed in one year, it may be made up in another year. If a waterbody has enough samples from the most recent 5-year period, data from that 5-year period will be used for the assessment. If a waterbody does not have enough data from the most recent 5-year period but it does have sufficient data in the past 10 years, then the 10-year period may be used. Sites with sufficient data within the most recent 5 or 10 years (the "assessment period") are identified as "assessment sites". The weather-controlled median phosphorus concentration of each of these sites will be compared to its phosphorus

criterion to determine whether it is attaining the criterion. Sites that do not qualify as assessment sites but have at least 6 phosphorus samples prior to that period are called “non-assessment” sites. Data from both types of sites are used within PhosMER, but for assessment sites only data from within the assessment period are used, so that the results reflect conditions only during the assessment period. For non-assessment sites the full historical dataset is used to inform the statewide phosphorus-precipitation relationships.

Estimating weather-controlled median total phosphorus with PhosMER

PhosMER stands for **Phos**phorus **M**ixed **E**ffects **R**egression and it is a statistical model that allows estimation of TP concentrations across all sampling locations in the state by using information about daily precipitation, temperature and watershed characteristics.

Mixed effects regression models are an “enhanced” version of linear regression models that overcome methodological issues that arise from the typical characteristics of water quality data such as heterogeneity, nested data, temporal correlation, and spatial correlation. Thus, a mixed effect regression model has two components: a fixed effects component and a random effects component. While the fixed effects component works exactly in the same way that a linear regression model does, the random effects component can accommodate additional sources of variability in the data that otherwise would invalidate the conclusions from a linear regression. One of these sources of variability is site-specific characteristics, that once included in the model, allow for site-specific predictions.

PhosMER fits both overall relationships between predictor variables and TP across all sites (fixed effects, such as precipitation levels), and site-specific variables affecting a site’s response (random effects). The extent to which fixed effects and random effects are applied to a site is based on the sample size and fit of the site-specific relationship. In other words, the regression equation for a site with many samples that are strongly related to sample-specific variables is largely determined by data from that site only, while the regression equation for a site with few samples that are weakly related to sample-specific variables is largely determined by the mean relationships for sites with similar site-specific variables.

PhosMER estimation of weather-controlled median total phosphorus for assessment sites is done in three steps*:

4. Identify assessment sites, and fit a mixed effects regression model by using the appropriate date ranges: only recent TP measurements from assessment sites and historical TP measurements from non-assessment sites.
5. Calculate weather-controlled TP medians using long-term records of precipitation and temperature.
6. Run 100 to 1000 iterations of PhosMER to calculate 80% confidence intervals around weather-controlled TP medians for assessment sites.

*More specific details are provided in the figures and methods section

Comparing the weather-controlled median to the phosphorus criterion

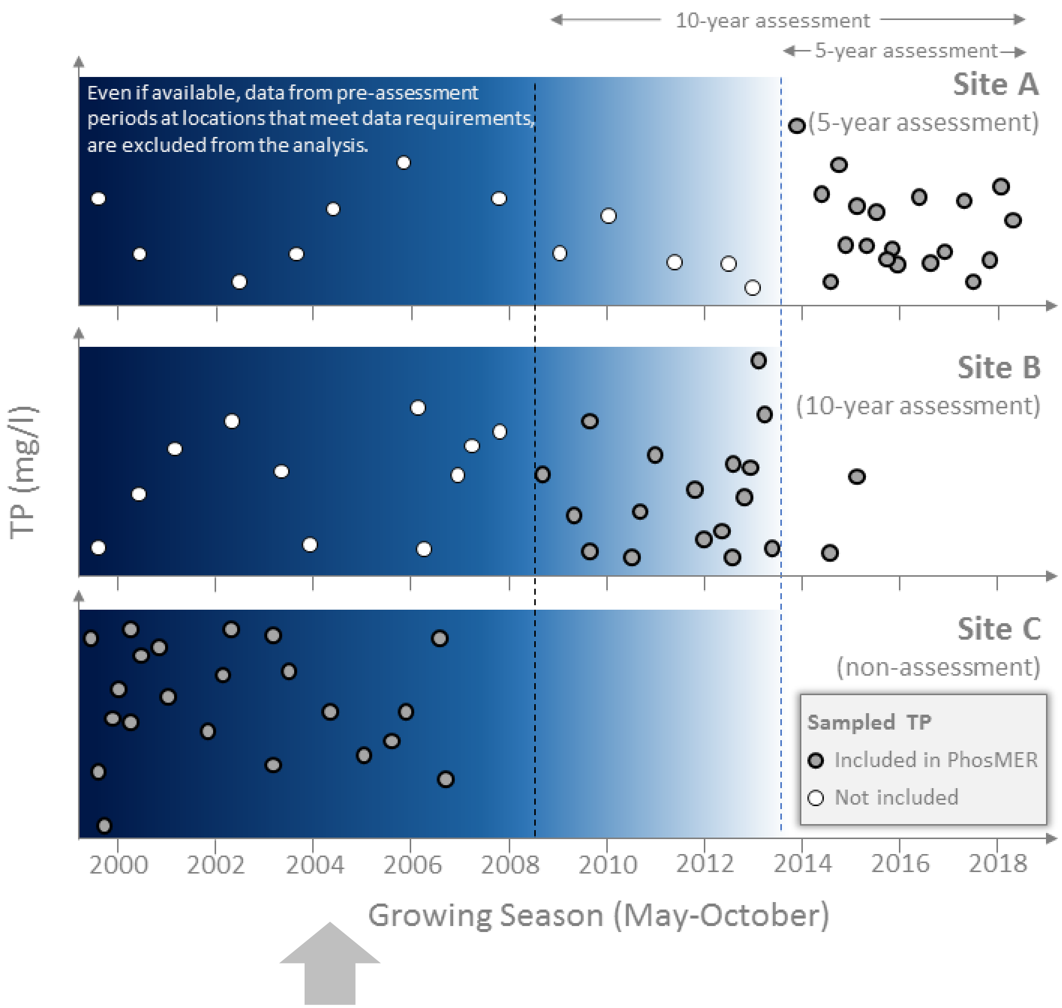
Once the weather-controlled median and 80% two-sided confidence interval are calculated using PhosMER, the results are evaluated to determine whether the waterbody’s median concentration is exceeding its phosphorus criterion. This is done by comparing the weather-controlled median and confidence interval to the applicable phosphorus criterion from s. NR 102.06, Wis. Adm. Code using the protocols specified in s. NR 102.07(2) and the confidence interval approach described in s. NR 102.52(2). Once a determination of whether the median exceeds the criterion is made, the combined assessment approach specified in s. NR 102.07(3) to (6) may also be applied if appropriate. This determines whether the waterbody is exhibiting a biological response to phosphorus and should be listed on the section 303(d) list as not attaining the phosphorus criterion.

Summary points

- Water quality samples are not always representative of the true condition of a waterbody.
- Weather variability is known to influence water quality and is relatively easy to control for.
- PhosMER provides insight into causes of phosphorus variation among sites (fixed effects) and among samples at a site (random effects).
- PhosMER's confidence intervals are likely to be more accurate and robust than those obtained from raw samples only.

(see figures and methods on the following pages)

1. PhosMER uses both recent data from assessment sites and available historical data from non-assessment sites



For 5- or 10-year assessment sites, data from the assessment period inform PhosMER about the **specific response** of TP to precipitation and temperature at that site.

Pre-assessment data from other locations is used to inform PhosMER about the **general response** of TP to precipitation and temperature across more than 500 sampling locations

Figure 1. Incorporating growing season (May – October) data in PhosMER to estimate weather-controlled median TP. Each site in the PhosMER database is classified as an assessment site (A and B) or non-assessment site (C) according to pre-specified data requirements. Data are incorporated into PhosMER in order to inform the model about: 1) average TP responses to a wide range of historical weather patterns and land use settings across all sampling locations (fixed effects), and 2) site-specific TP responses to historical weather under the predominant land use conditions during the assessment period (random effects). For assessment sites, to ensure the model will represent potential TP responses under site-specific conditions during the assessment period, historical data from these sites (even if available) are not used in the model.

2. Weather-controlled TP medians are calculated using long-term records of precipitation and temperature

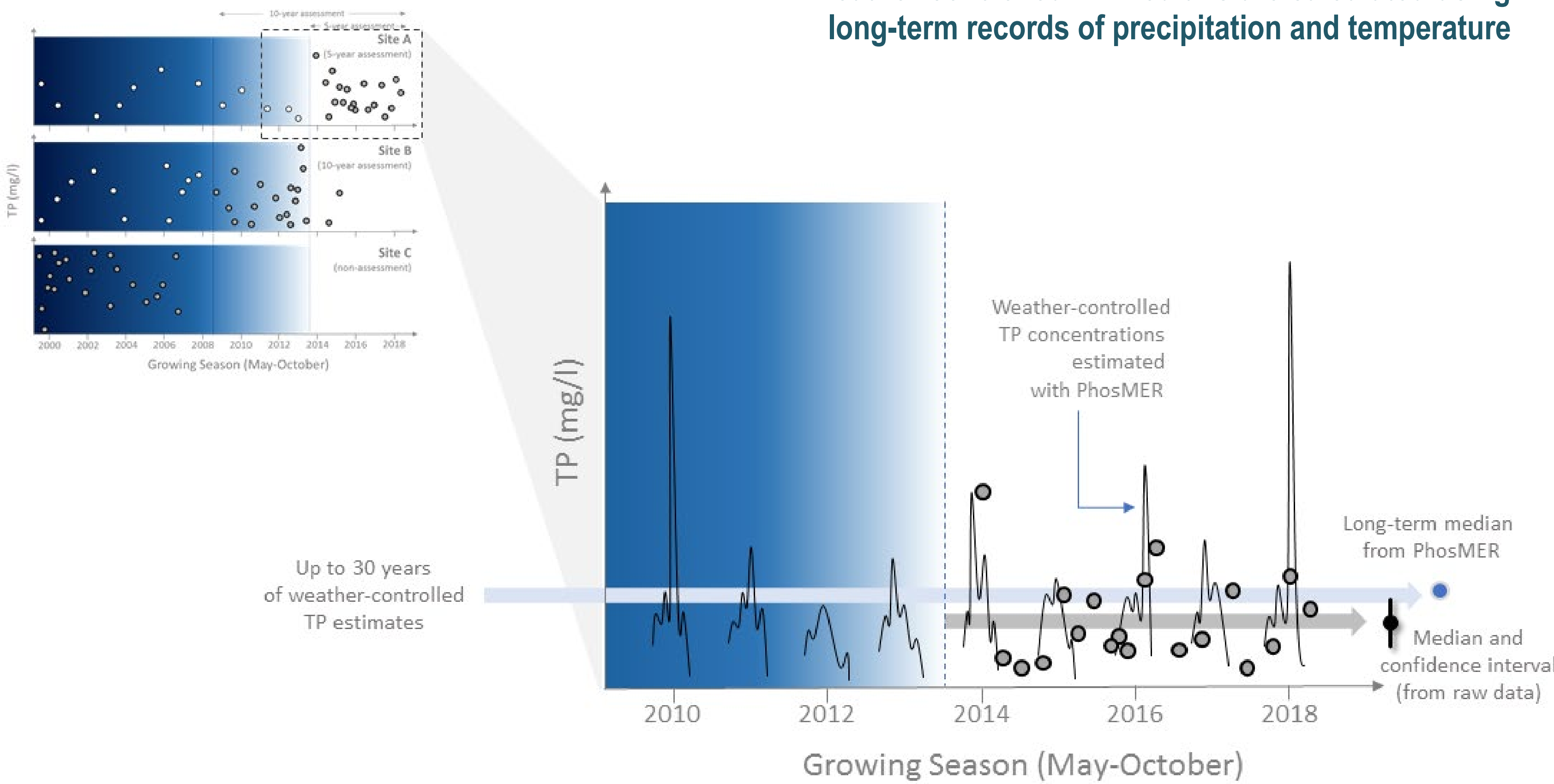
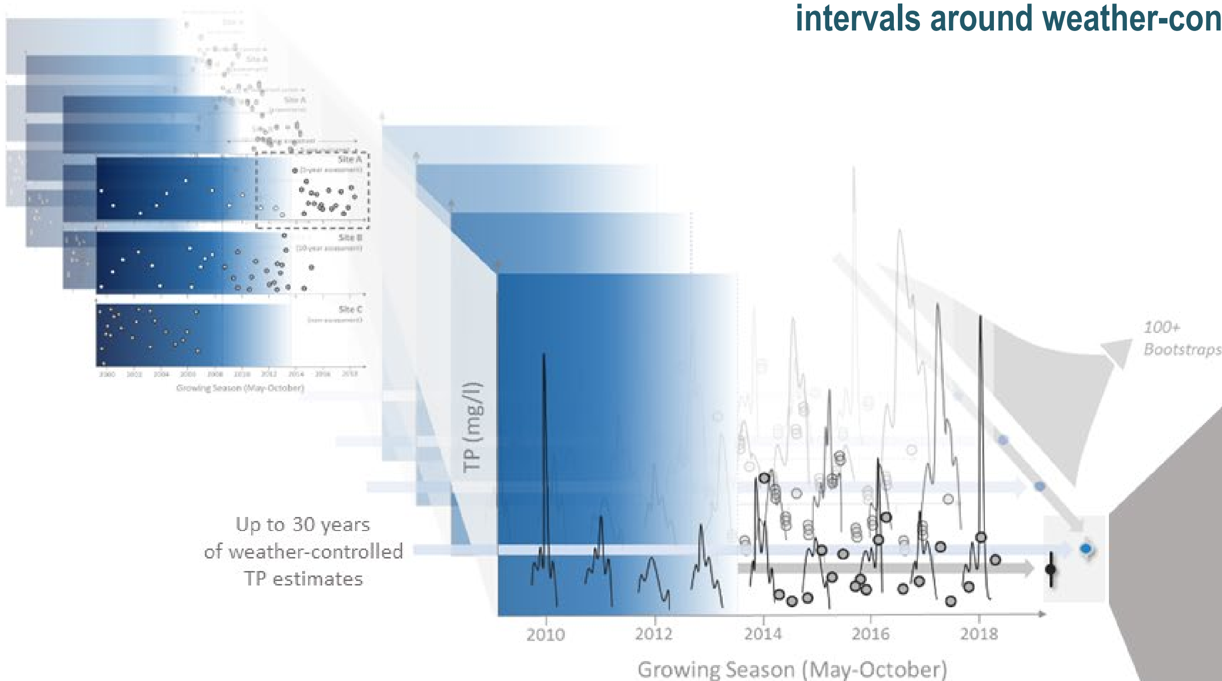


Figure 2. PhosMER estimates daily TP concentrations based on long-term weather patterns and provides a weather-controlled median TP for each site. By using a 30-year long record of daily precipitation and temperature previously geoprocessed for each site (see methods section), PhosMER estimates site-specific 30-year long records of daily TP concentrations (during the growing season). These values are averaged per site to provide a single estimate of weather-controlled TP. The dotted square in the top left highlights these predictions for an assessment site over a 9-year window (for demonstration purposes the longer historic record is truncated). Without PhosMER, the current assessment methodology will calculate a median and a confidence interval only based on the observed data that may not be representative of the true condition of the waterbody.

3. A hundred to a thousand iterations of PhosMER are used to calculate 80% confidence intervals around weather-controlled TP medians for assessment sites.



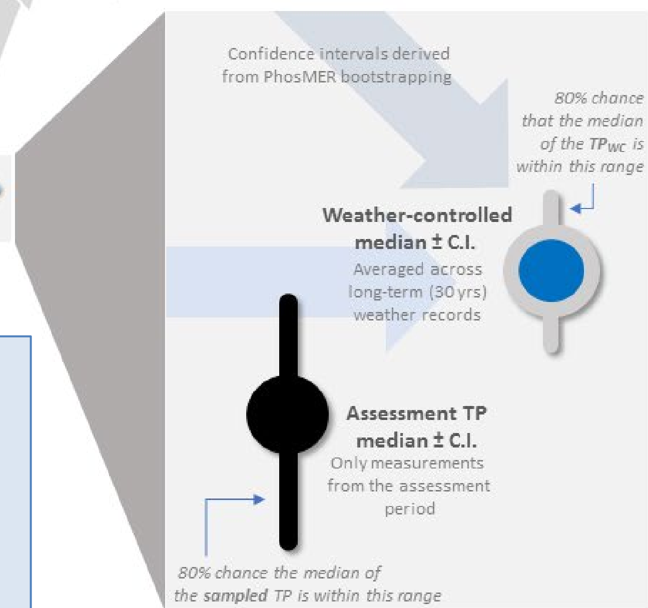


Figure 3. Between 100 and 1000 long-term time series of daily TP concentrations are calculated for each site to obtain a distribution of weather-controlled medians. PhosMER accounts for weather variability by estimating the long-term weather controlled median TP for each site. However, there is also variability associated with all the possible models that could be fitted through the available data. To account for this variability, the entire dataset is resampled with replacement and the model is refitted to produce the weather-controlled median for each site. This process, called "bootstrapping" is repeated between 100 and 1000 times to calculate a distribution of medians that will reflect the combination of the same number of models times 30 years of weather patterns. Thus, for an assessment site between 100 and 1000 weather-controlled median TP concentrations are estimated and used to build a 80% confidence interval. The weather-controlled median and confidence intervals can then be compared to the sites's phosphorus criterion to determine if it is attaining the criterion.

Methods

We constructed mixed effects models using landscape variables and daily time series of weather data, with random effects for individual sites to control for site-specific variation and enable the combination of all sites into one model. The model was built to analyze both phosphorus and total suspended solids (TSS) concentrations.

Data acquisition and processing

Water quality data
Wisconsin DNR monitors water quality in streams throughout the state, and stores water quality data from individual projects, partner agencies, etc. In general, grab samples are taken at stream sites and processed by the Wisconsin Lab of Hygiene. Data are stored in Wisconsin DNR's Surface Water Inventory System (SWIMS). We extracted data for sites on streams or impounded flowing waters between 1981 and 2014 that had phosphorus measurements that spanned at least 80% of the growing season (May to October), to allow for random effects associated with seasonality to be fitted reasonably. When multiple measurements occurred on a single day, we used the average. We also removed outlier observations identified using Grubb's test on the natural log of TP and TSS, with an alpha value of 0.01. Because our dataset spans over 30 years, in which many other changes – particularly land use change – could influence TP and TSS, we used only the most recent 10 years of data for any given site. Lastly, we removed sites with fewer than 6 observations in this time period for two reasons: 1) we wanted to include only sites with multiple observations to aid in assessing random effects, 2) 6 observations is the minimum data requirement in the assessment protocol, and 3) because there was a steep drop off in the number of sites with sufficient data if we required more than 6 observations.

Geographic predictors
We compiled catchment-scale geographic predictors from Wisconsin's 24k Hydro Layer (ftp://dnrftp01.wi.gov/geodata/hydro_va_24k/). In this database, catchment-scale variables are calculated for each stream reach, where catchments characterize the entire upstream contributing area. Predictors included catchment area, average slope, average soil permeability, percent coverage of agricultural land covers, and percent coverage of urban land covers. Land management practices such as establishment of best management practices or crop rotations are not included in the model.

Daily predictors
We joined these geographic predictors with seasonal terms and weather data for each day for each reach. We first accounted for seasonal curvature in TP and TSS trends using the sine and cosine of day-of-year as in the LOADEST model. We also calculated daily temperature and precipitation indexes from DAYMET (https://daymet.ornl.gov/), a daily gridded (1 km x 1 km) dataset that includes temperature minima and maxima and precipitation amounts. To capture the effects of variation in temperature, daily average temperatures were converted to 7-day averages, which we converted to 7-day temperature anomalies by subtracting the mean 7-day temperature for the respective day of year for a given site. These 7-day temperature anomalies (T7D) were then used as predictors in subsequent models, along with calculated antecedent precipitation indexes.

Antecedent precipitation calculation
We calculated an antecedent precipitation index for each reach-day, using DAYMET daily precipitation amounts and temperatures. We converted raw precipitation to effective precipitation by separating rain from snow and converting water in snow to snowmelt. We assumed any precipitation that occurred when the daily temperature was above 0°C to be rain and precipitation on days at or below 0°C to be snow. Snow accumulates until temperature exceeds 0°C, at which point it begins to melt. Melt rates were estimated

from air temperature and solar radiation, using equations from NOAA's Snow-17 model (http://www.nws.noaa.gov/oh/hrl/nwsrfs/users_manual/part2/_pdf/22snow17.pdf). Thus, each reach-day had estimates of direct rain and snowmelt, which were summed to represent effective precipitation. Antecedent precipitation for each day was then calculated as the weighted sum of effective precipitation over the preceding 365 days, where weights follow a time decay function. Specifically, weights were assigned to each day using an equation:

$$w = \frac{e^{-((\log(d)-m)^2/2)}}{d\sqrt{2\pi}}$$

where d is the number of days before present and m is a parameter that controls the shape of the function. Low values of m correspond to a faster and shorter response of TP or TSS to precipitation and high values correspond to a slower and longer response. Values of m were assigned to individual sites by estimating the relationship between watershed area (WSA, km2) and slope (%) (indicators of hydrologic response time), and the optimal value of m for a subset of sites with at least 12 TP or TSS samples.

(eq. 2) mTP = -0.35 + 1.14*log10(WSA) - 0.84*log(SLOPE)
(eq. 3) mTSS = 0.011 + 0.203*log(WSA)

Modeling

We fit mixed effects models using the lmer function from the lme4 package in R. The models for TP and TSS have the same structure, but different coefficients. The model formula is:

(eq. 4) log(TP/TSS) ~ PS * (SLOPE + log(WSA) + PERM + AG + URB) + SDT + CDT + T7D
+ (1 + SDT + CDT + T7D + PS + I(PS^2) | STATION_ID)

where log(TP/TSS) is the log of TP or TSS concentration, PS is the antecedent precipitation index, PERM is mean soil permeability (in/hr), AG is percent agricultural land cover, URB is percent urban land cover, and SDT and CDT, the seasonal terms, are the sine and cosine of decimal year.

We included random effects for each site where TP or TSS was measured and included in the study. Random effects included a random intercept along with random slopes for each seasonal term, temperature anomaly, precipitation index, and the square of the precipitation index. The random intercept allows for site-specific higher or lower TP or TSS concentrations, given a site's fixed effects. Sites could also presumably vary in the seasonal trend of TP or TSS concentrations, where peak concentrations could occur earlier or later. They could vary in their sensitivity of the TP or TSS response to temperature or precipitation. Finally, some sites may have nonlinear responses to precipitation. The random effect for the slope of the squared precipitation index allows sites to express responses to precipitation that increase exponentially or that peak with a moderate amount of antecedent precipitation.

APPENDIX C. Combined approach range for phosphorus assessments

The graphs below demonstrate how the combined assessment approach for phosphorus assessments is conducted. The range for combined assessments is shaded in gray. Not all waterbody types or phosphorus response indicators are shown here, but this provides an example from each major waterbody group (streams, rivers, lakes). The left-most edge of the shaded range is the phosphorous criterion for the waterbody type. Waterbodies within the shaded area are exceeding their phosphorus criterion but may have their phosphorus response indicators assessed to determine whether they should be considered impaired for phosphorus. The horizontal dashed line shows the response indicator threshold. In the examples shown here, the indicator is attained if the waterbody is below the dashed line. The indicator is not attained if the waterbody is above the dashed line.

Open circles show waterbodies that would not be listed as impaired, either because they are to the left of (below) the TP criterion or because they are within the combined range and attaining (below) the response indicator threshold. Closed circles show waterbodies that would be listed as impaired for phosphorus because they are either to the right of (above) the combined assessment range or they are within the range but not attaining (above) the indicator threshold.

As shown in these graphs, only a small number of waterbodies (for which the department currently has data) exceed phosphorus criteria while still having good response indicators. This indicates that the general phosphorus criteria are not overprotective. However, it demonstrates that there is some variability in the response of biological factors to phosphorus concentrations.

Range extent: For streams and rivers, the shaded range extends from the phosphorus criterion to 2 times the criterion: the stream range is 75-150 ug/L TP, and the river range is 100-200 ug/L TP. For lakes and reservoirs, the range is from the criterion to 1.5 times the criterion (in this example, the shallow lake range is 40-60 ug/L TP; there are several other lake criteria for different lake types). In each example, the high end of the range is set to a point above which very few waterbodies are achieving good biology (attaining their response indicators). Above the high end of the range, it is unnecessary and inefficient to apply a response indicator as there is a high level of certainty that waterbodies do not attain good biology at such high phosphorus concentrations.

Figure 27. Diatom phosphorus response indicator for streams showing the combined assessment range (shaded) from 75 to 150 ug/L TP. Open circles in the shaded area would not be listed as impaired for phosphorus because they attain (are lower than) the diatom indicator threshold (DPI score of 75 ug/L). Closed circles in the shaded area would be listed as impaired for phosphorus because they are not attaining (are higher than) the diatom indicator threshold.

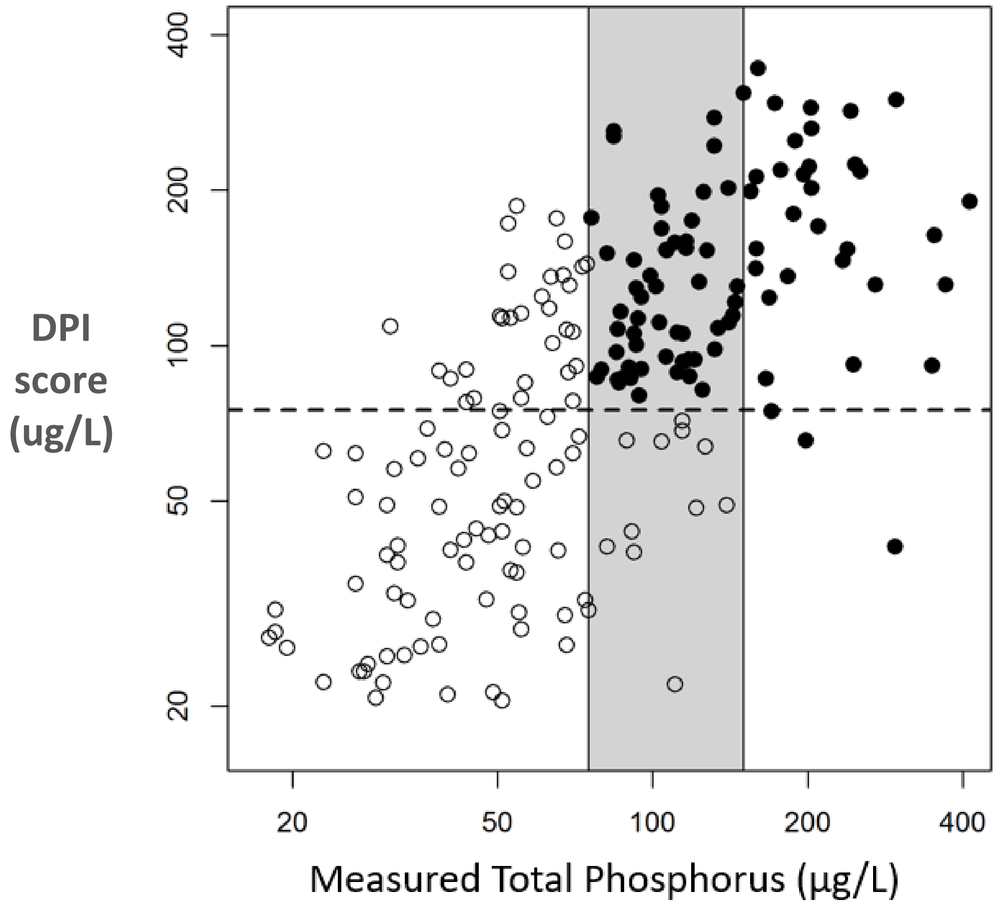


Figure 28. River chlorophyll *a* phosphorus response indicator for percent of summer days exhibiting moderate algal levels. Shows the combined assessment range (shaded) from 100 to 200 ug/L TP. Open circles in the shaded area would not be listed as impaired for phosphorus because they attain (are lower than) the algal frequency indicator threshold (30% or fewer of summer days with algal levels above 20 ug/L chlorophyll *a*). Closed circles in the shaded area would be listed as impaired for phosphorus because they are not attaining (are higher than) the algal frequency indicator threshold.

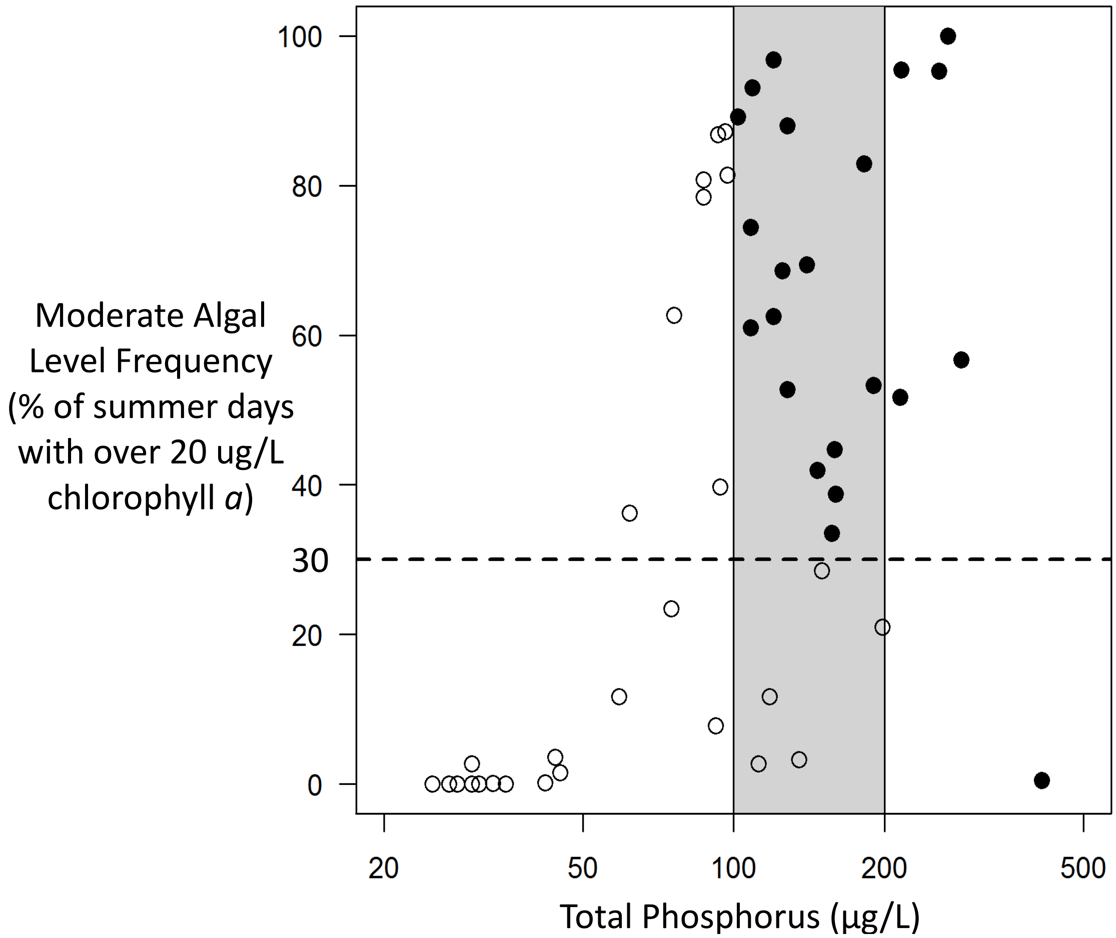
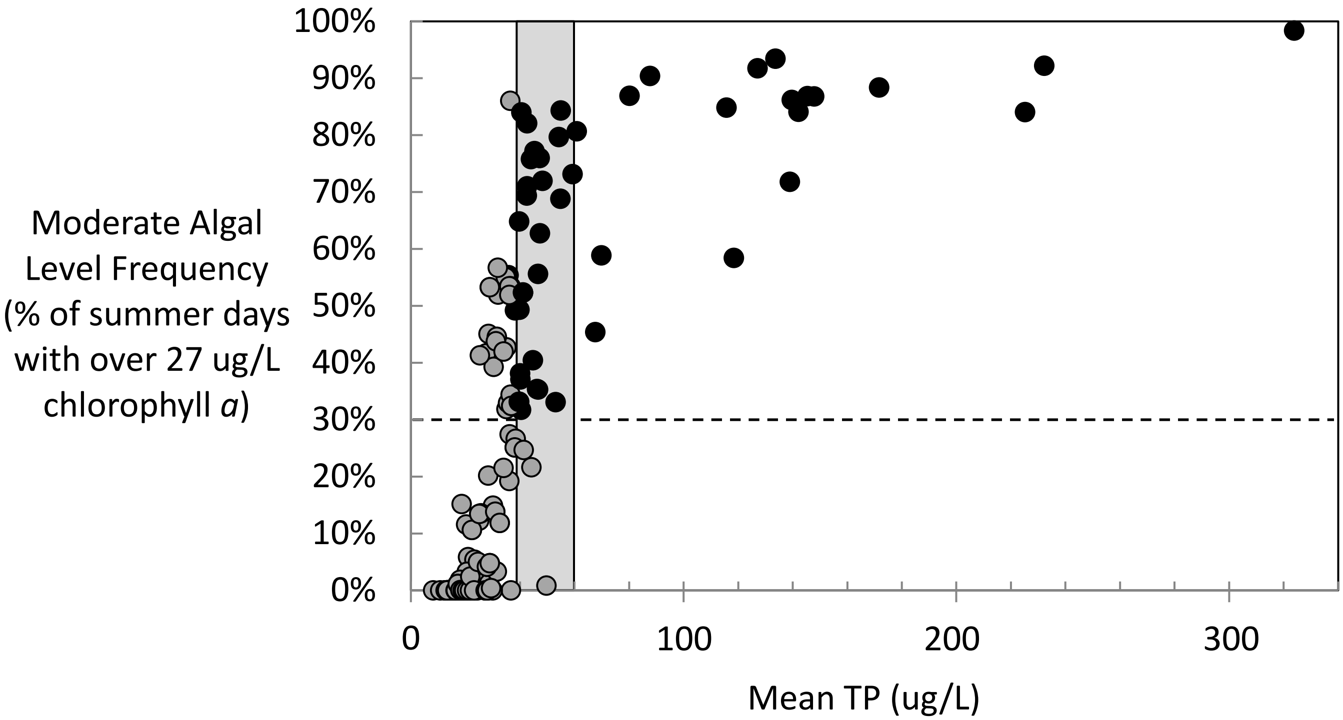


Figure 29. Shallow lake chlorophyll *a* phosphorus response indicator for percent of summer days exhibiting moderate algal levels. Shows the combined assessment range (shaded) from 40 to 60 ug/L TP. Open circles in the shaded area would not be listed as impaired for phosphorus because they attain (are lower than) the algal frequency indicator threshold (30% or fewer of summer days with algal levels above 20 ug/L chlorophyll *a*). Closed circles in the shaded area would be listed as impaired for phosphorus because they are not attaining (are higher than) the algal frequency indicator threshold.



APPENDIX D. Plant species used in the Macrophyte Assessment of Condition: MAC-General and MAC-Phosphorus.

Table 11. Macrophyte species assigned to each tolerance group for the two Macrophyte Assessment of Condition (MAC) tools. The General Macrophyte Assessment of Condition (MAC-Gen) is used to assess overall aquatic plant response to anthropogenic disturbance and has three tolerance groups: "disturbance-sensitive", "moderately tolerant to disturbance" and "disturbance-tolerant". The Macrophyte Assessment of Condition for Phosphorus (MAC-P) assesses plant response to phosphorus specifically, using two tolerance groups: "phosphorus-sensitive" and "phosphorus-tolerant". The MAC-P is used as a phosphorus response indicator as described in section 5.4.2 of this document.

Scientific Name (Synonym)	Taxonomy Level	Parent Taxon	Common Name(s)	MAC-P Tolerance	MAC-Gen Tolerance
Ceratophyllum demersum	Species	Ceratophyllales Ceratophyllaceae Ceratophyllum	Coontail	Tolerant	Tolerant
Heteranthera dubia (Zosterella dubioa)	Species	Commelinales Pontederiaceae Heteranthera	Water star-grass	Tolerant	Tolerant
Lemna minor	Species	Alismatales Araceae Lemna	Common duckweed	Tolerant	Tolerant
Lemna perpusilla	Species	Alismatales Araceae Lemna	Least duckweed	Tolerant	Tolerant
Lemna sp	Genus	Alismatales Araceae	Duckweed	Tolerant	Tolerant
Lemna turionifera	Species	Alismatales Araceae Lemna	Turion duckweed	Tolerant	Tolerant
Myriophyllum sibiricum X spicatum	Species	Saxifragales Haloragaceae Myriophyllum	Hybrid water-milfoil	Tolerant	Moderate
Myriophyllum spicatum or M sibiricum X spicatum	Species	Saxifragales Haloragaceae Myriophyllum	Eurasian water-milfoil or Hybrid water-milfoil	Tolerant	Tolerant
Potamogeton crispus	Species	Alismatales Potamogetonaceae Potamogeton	Curly-leaved pondweed	Tolerant	Tolerant
Spirodela polyrrhiza	Species	Alismatales Araceae Spirodela	Giant duckweed	Tolerant	Tolerant
Stuckenia pectinate (Potamogeton pectinatus)	Species	Alismatales Potamogetonaceae Stuckenia	Sago pondweed	Tolerant	Tolerant
Wolffia borealis	Species	Alismatales Araceae Wolffia	Northern water-meal	Tolerant	Tolerant
Wolffia brasiliensis	Species	Alismatales Araceae Wolffia	Brazilian water-meal	Tolerant	Tolerant
Wolffia columbiana	Species	Alismatales Araceae Wolffia	Common water-meal	Tolerant	Tolerant
Wolffia sp	Genus	Alismatales Araceae	Water-meal	Tolerant	Tolerant
Zannichellia palustris	Species	Alismatales Potamogetonaceae Zannichellia	Horned-pondweed	Tolerant	Tolerant
Elodea nuttallii	Species	Alismatales Hydrocharitaceae Elodea	Slender waterweed	Tolerant	Sensitive
Potamogeton pusillus	Species	Alismatales Potamogetonaceae Potamogeton	Small pondweed	Tolerant	Sensitive

Elodea canadensis	Species	Alismatales Hydrocharitaceae Elodea	Common waterweed	Tolerant	Moderate
Lemna trisulca	Species	Alismatales Araceae Lemna	Star duckweed	Tolerant	Moderate
Nuphar sp	Genus	Nymphaeales Nymphaeaceae	Pond-lily	Tolerant	-
Nuphar variegata	Species	Nymphaeales Nymphaeaceae Nuphar	Bull-head pond-lily	Tolerant	Moderate
Nymphaea odorata	Species	Nymphaeales Nymphaeaceae Nymphaea	American white water-lily	Tolerant	Moderate
Potamogeton foliosus	Species	Alismatales Potamogetonaceae Potamogeton	Leafy pondweed	Tolerant	Moderate
Potamogeton nodosus	Species	Alismatales Potamogetonaceae Potamogeton	Long-leaf pondweed	Tolerant	Moderate
Potamogeton zosteriformis	Species	Alismatales Potamogetonaceae Potamogeton	Flat-stem pondweed	Tolerant	Moderate
Ranunculus aquatilis	Species	Ranunculales Ranunculaceae Ranunculus	White water crowfoot	Tolerant	Moderate
Elodea sp	Genus	Alismatales Hydrocharitaceae	Waterweed	Tolerant	-
Najas marina	Species	Alismatales Hydrocharitaceae Najas	Spiny naiad	Sensitive	Tolerant
Persicaria amphibia (Polygonum amphibium)	Species	Caryophyllales Polygonaceae Persicaria	Water smartweed	Sensitive	Tolerant
Potamogeton hillii	Species	Alismatales Potamogetonaceae Potamogeton	Hill's pondweed	Sensitive	-
Potamogeton illinoensis	Species	Alismatales Potamogetonaceae Potamogeton	Illinois pondweed	Sensitive	Tolerant
Potamogeton praelongus	Species	Alismatales Potamogetonaceae Potamogeton	White-stemmed pondweed	Sensitive	Tolerant
Bidens beckii (Megalodonta beckii)	Species	Asterales Asteraceae Bidens	Water-marigold	Sensitive	Sensitive
Brasenia schreberi	Species	Nymphaeales Cabombaceae Brasenia	Water-shield	Sensitive	Sensitive
Elatine minima	Species	Malpighiales Elatinaceae Elatine	Small waterwort	Sensitive	Sensitive
Eleocharis acicularis	Species	Poales Cyperaceae Eleocharis	Needle spike-rush	Sensitive	Sensitive
Eriocaulon aquaticum	Species	Poales Eriocaulaceae Eriocaulon	Seven-angle pipewort	Sensitive	Sensitive
Isoetes sp	Genus	Isoetales Isoetaceae	Quillwort	Sensitive	Sensitive
Juncus pelocarpus	Species	Poales Juncaceae Juncus	Brown-fruited rush	Sensitive	Sensitive
Lobelia dortmanna	Species	Asterales Campanulaceae Lobelia	Water lobelia	Sensitive	Sensitive
Myriophyllum tenellum	Species	Saxifragales Haloragaceae Myriophyllum	Slender water-milfoil	Sensitive	Sensitive
Najas flexilis	Species	Alismatales Hydrocharitaceae Najas	Slender naiad	Sensitive	Sensitive
Najas gracillima	Species	Alismatales Hydrocharitaceae Najas	Slender water-nymph	Sensitive	Sensitive
Najas guadalupensis	Species	Alismatales Hydrocharitaceae Najas	Southern naiad	Sensitive	Sensitive
Najas sp	Genus	Alismatales Hydrocharitaceae	Naiad	Sensitive	-
Nitella sp	Genus	Charales Characeae	Nitella	Sensitive	Sensitive
Nitellopsis obtusa	Species	Charales Characeae Nitellopsis	Starry stonewort	Sensitive	Sensitive
Potamogeton amplifolius	Species	Alismatales Potamogetonaceae Potamogeton	Large-leaved pondweed	Sensitive	Sensitive

Potamogeton epihydrus	Species	Alismatales Potamogetonaceae Potamogeton	Ribbonleaf pondweed	Sensitive	Sensitive
Potamogeton gramineus	Species	Alismatales Potamogetonaceae Potamogeton	Variable-leaved pondweed	Sensitive	Sensitive
Potamogeton robbinsii	Species	Alismatales Potamogetonaceae Potamogeton	Robbins' pondweed	Sensitive	Sensitive
Potamogeton spirillus	Species	Alismatales Potamogetonaceae Potamogeton	Spiral-fruited pondweed	Sensitive	Sensitive
Potamogeton strictifolius	Species	Alismatales Potamogetonaceae Potamogeton	Narrow-leaved pondweed	Sensitive	Sensitive
Schoenoplectus subterminalis	Species	Poales Cyperaceae Schoenoplectus	Water bulrush	Sensitive	Sensitive
Sparganium angustifolium	Species	Poales Typhaceae Sparganium	Narrow-leaved bur-reed	Sensitive	Sensitive
Sparganium fluctuans	Species	Poales Typhaceae Sparganium	Floating-leaved bur-reed	Sensitive	Sensitive
Utricularia geminiscapa	Species	Lamiales Lentibulariaceae Utricularia	Hidden-fruited bladderwort	Sensitive	Sensitive
Utricularia gibba	Species	Lamiales Lentibulariaceae Utricularia	Creeping bladderwort	Sensitive	Sensitive
Utricularia intermedia	Species	Lamiales Lentibulariaceae Utricularia	Flat-leaved bladderwort	Sensitive	Sensitive
Utricularia minor	Species	Lamiales Lentibulariaceae Utricularia	Lesser bladderwort	Sensitive	Sensitive
Utricularia resupinata	Species	Lamiales Lentibulariaceae Utricularia	Northeastern bladderwort	Sensitive	Sensitive
Chara sp	Genus	Charales Characeae	Muskgrass	Sensitive	Moderate
Myriophyllum heterophyllum	Species	Saxifragales Haloragaceae Myriophyllum	Various-leaved water-milfoil	Sensitive	Moderate
Myriophyllum sibiricum	Species	Saxifragales Haloragaceae Myriophyllum	Common water-milfoil	Sensitive	Moderate
Myriophyllum verticillatum	Species	Saxifragales Haloragaceae Myriophyllum	Whorled water-milfoil	Sensitive	Moderate
Potamogeton friesii	Species	Alismatales Potamogetonaceae Potamogeton	Fries's pondweed	Sensitive	Moderate
Potamogeton natans	Species	Alismatales Potamogetonaceae Potamogeton	Floating pondweed	Sensitive	Moderate
Potamogeton richardsonii	Species	Alismatales Potamogetonaceae Potamogeton	Richardson's pondweed	Sensitive	Moderate
Utricularia vulgaris	Species	Lamiales Lentibulariaceae Utricularia	Common bladderwort	Sensitive	Moderate
Vallisneria americana	Species	Alismatales Hydrocharitaceae Vallisneria	Water-celery	Sensitive	Moderate