COLORADO
Department of Public Health & Environment

Section 303(d) Listing Methodology 2026 Listing Cycle

Left Hand Creek

Colorado Department of Public Health and Environment
Water Quality Control Division

MARCH 2024

LIST OF TABLES

SECTION 303(d) LISTING METHODOLOGY

2026 Listing Cycle

I. INTRODUCTION

Section 303(d) of the Federal Clean Water Act requires states to identify waters where effluent limitations mandated by Section 301(b)(1)(A) and Section 301(b)(1)(B) are not stringent enough to attain water quality standards. These waters are compiled into the Section 303(d) List of Impaired Waters. The Colorado Section 303(d) List identifies those waterbodies where there are exceedances of water quality standards or nonattainment of uses. This includes waters impaired as a result of nonpoint source, point source discharges or combined point source and nonpoint source contributions including natural sources. Total Maximum Daily Loads (TMDLs) are required for each listed waterbody. The Section 303(d) List is equivalent to Category 5 waters in the Environmental Protection Agency's (EPA's) *July 29, 2005 Guidance for 2006 Assessment, Listing and Reporting Requirements Pursuant to Sections 303(d), 305(b) and 314 of the Clean Water Act* (2006 Integrated Reporting Guidance), and the *September 3, 2013 Information Concerning 2014 Clean Water Act Sections 303(d), 305(b) and 314 Integrated Reporting and Listing Decisions*.

The Monitoring and Evaluation List (M&E List) identifies waterbodies where there is reason to suspect water quality problems, but there is also uncertainty regarding one or more factors, such as the representative nature of the data. The M&E List contains waterbodies that would be reflected in Category 3 of EPA's Integrated Reporting Guidance.

Waters that are on neither the Section 303(d) List nor the M&E List are:

- Attaining their uses and standards (EPA's Category 1);
    - 1a – Attaining water quality standards
    - 1b - Attaining water quality standards with TMDL
- Attaining some uses (EPA's Category 2);
- Have not been fully assessed (EPA's Category 2 or 3); or
- Impaired but do not require a TMDL for the following reasons (EPA's Category 4):
    - 4a - TMDLs have been completed but uses are not yet attained;
    - 4b - other required control mechanisms are expected to address all waterbody-pollutant combinations and will attain water quality standards in a reasonable period of time;
    - 4c - the impairment is not caused by a pollutant.

Section II of this document identifies the process that the Water Quality Control Division (division) and Water Quality Control Commission (commission) intend to follow in establishing the Section 303(d) and M&E Lists. Section III contains the listing criteria and Section IV describes how the division interprets data in order to make an attainment decision.

This document provides a framework for the determination of attainment or nonattainment of assigned water quality standards and designated uses. However, there may be site specific considerations not identified in the listing methodology that are appropriately factored into the final listing decision. Generally, the division's recommendation to list or not list a waterbody will be based upon stringent application of the listing methodology criteria, but best professional judgment (BPJ) may be applied when necessary. Parties will have the

opportunity to present mitigating evidence for the commission's consideration as part of the rulemaking hearing process.

II. 303(d) LISTING METHODOLOGY AND REGULATION # 93 PROCESS

A. Development of the Methodology

The listing methodology is reviewed and updated on a biennial basis in anticipation of 303(d) List and M&E List development. The listing methodology is revisited and revised with the intent of clarifying the division's procedures for assessing attainment of those uses and standards assigned by the commission to Colorado waters. Most often revisions or additions to the listing methodology derive from issues raised during the previous listing process.

B. Process for Adopting the Methodology

The process for formal consideration and acceptance of the listing methodology was discussed at an April 2003 commission meeting. At that time, the commission decided to convene an Administrative Action Hearing (AAH) process for adoption of the listing methodology. Since the 2004 cycle, the listing methodology has been approved in an AAH process. The following schedule is anticipated for development and finalization of the 2026 Section 303(d) Listing Methodology:

- The division's proposal will be available on the commission's website in January 2024, and be emailed to participants in the 303(d) Listing Methodology work group.
- The notice for the March 11, 2024 public hearing for consideration of approval of the listing methodology will establish a deadline of January 31, 2024 for written responsive comments on the proposed listing methodology, including any recommendations for alternative language in the document. Comments received will be posted on the commission's web site.
- The notice will also establish a deadline of February 21, 2024 for any written rebuttal comments in response to the January 31, 2024 comments. These rebuttal comments will be posted on the commission's web site.
- If the initial written comments and/or the rebuttal comments warrant revisions to the proposed listing methodology, the division will submit a revised proposal by February 28, 2024. This revised proposal will be posted on the commission's website.
- No other written materials will be accepted for this hearing except by specific permission from the commission, with written explanation as to why such materials could not have been submitted in accordance with the above deadlines.
- There will be opportunity at the March 11th, 2024 hearing for any interested persons to provide oral comments regarding the proposed listing methodology.
- At the conclusion of the March 11th, 2024 Administrative Action Hearing, the commission will modify, as necessary, and consider approving the final 2026 Section 303(d) Listing Methodology.

C. Process for Adopting the Section 303(d) and Monitoring & Evaluation Lists

The process for formal consideration and adoption of the Section 303(d) and M&E Lists was also discussed at the April 2003 commission meeting. The commission decided that the 2004 lists, and subsequent lists, would be adopted through a public rulemaking process. The following steps and anticipated schedule will be used for the adoption of the 2026 303(d) and M&E Lists:

- Any person that has a Category 4b demonstration plan that wishes the division to consider and submit to EPA must provide that information to the division by the last week in January 2025. The division will formally submit the plan to EPA by the first week in February.
- Any person that has data or other information that it wishes the division to consider in determining which water segments and parameters to propose for listing or delisting (for either the Section 303(d) List or the M&E List) must provide that information to the division by the data call deadline for the applicable regulatory basin being assessed.
    - South Platte River, Laramie River, Republican River, and Smoky Hill River Basins (Regulation #38) Data call issued in June of 2023 with a deadline of September 1st of 2023
    - Statewide data call for specific waterbody assessments - Data call issued in June of 2024 with a deadline of September 1st of 2024
- After each data submittal deadline, the division will post the preliminary data and station information on either the Water Quality Control Division website, or Google Drive website by November 1st. The segment information will be provided in a preliminary format and is subject to change during the assessment of the data.
- By the third week in November 2024, external parties should contact the division with suggestions for the 303(d) List and/or the Monitoring and Evaluation List.
- By the fourth week in November 2024, the division responds to the external parties regarding whether the segments in question will be in their proposal or not (giving external parties three weeks to develop their own proposal).
- Any person who wishes to propose the listing of water segments/parameters that may not be proposed by the division must submit any such proposal, with accompanying proposed statement of basis and purpose language, to the commission office by the third week in December 2024. Any such proposal must also include adequate information for the commission to determine that listing of the segment/parameter should be considered in the rulemaking.
- The division will also submit its proposal to the commission by the third week in December 2024.
- A draft rulemaking hearing notice, with the division's and any external proposals attached, will be prepared by the commission office, for inclusion in the commission's February 2025 meeting packets.
- The commission will review the draft notice and proposals at its February 2025 regular meeting and approve them for filing.
- The rulemaking hearing notice and proposals will be filed with the Secretary of State by February 28, 2025. The final notice and proposals will also be posted on the commission's web site by about this date.

- The rulemaking notice will include contact information for persons wishing to get more detailed information regarding data or other information supporting the listing proposals advanced by the division or other persons.
- The rulemaking hearing notice and proposal will be published in the February 2025 Colorado Register.
- Written proponent's prehearing statements will be due by the last week of February 2025.
- The notice will establish a party status deadline of the first week of March 2025.
- Responsive prehearing statements and any evidence (data and any other relevant information) regarding potential listings will be due by the third week in March. All data submitted after the data call deadline should meet data submittal criteria described in Section III.A.1.
- This March deadline for the submission of evidence (data and any other relevant information) will apply to any information from any interested persons, not just those with party status.
- The notice will provide an opportunity for the submission of written rebuttal statements, in response to the March submissions, by the forth week in April 2025.
- No new data or other new factual information will be accepted after the 3rd week of March but the rebuttal statements may contain different analyses and perspectives regarding what the submitted information shows regarding attainment and the appropriateness of listing and may include additional information solely to rebut or respond to information filed with another party's responsive prehearing statement.
- A prehearing conference will be held during the fourth week of April 2025.
- Any data or other information that is not submitted in accordance with the above deadlines will be considered in the next listing cycle.
- The commission's rulemaking hearing will be held in May 2025. At the conclusion of the hearing, the commission will consider approval of the 2026 Section 303(d) List and the Monitoring and Evaluation List as Regulation #93.

Table 1. 303(d) Listing Methodology and Regulation #93 Estimated Schedule

Topic	Important Milestone	Approximate Date
303(d) LISTING METHODOLOGY	Draft proposal deadline	Jan. 17, 2024
	Responsive comments due	Jan. 31, 2024
	Rebuttal comments due	Feb. 21, 2024
	Revised proposal due	Feb. 28, 2024
	Administrative Action Hearing	Mar. 11, 2024
CATEGORY 4b	Category 4b Plan due to division	Last week in Jan. 2025
	4b Plan due to EPA	1st week in Feb. 2025
South Platte River, Laramie River, Republican River, and Smoky Hill River Basins (Regulation #38)	Data call	1st week in Jun. 2023
	Data submittal due	Sep. 1st, 2023

Table 1. 303(d) Listing Methodology and Regulation #93 Estimated Schedule

Topic	Important Milestone	Approximate Date
Statewide data call for outstanding issues	Data call	1st week in Jun. 2024
	Data submittal due	Sep. 1st, 2024
REGULATION #93 303(d) and M&E LIST	External suggestions for list	3rd week in Nov. 2024
	Division response to suggestions	4th week in Nov. 2024
	Division and other proponents' proposals are due	3rd week in Dec. 2024
	Draft rulemaking hearing notice	Last week in Jan. 2025
	Proposal review at commission meeting	2nd week in Feb. 2025
	Rulemaking hearing, notice and proposal filed with Secretary of State	Feb. 2025
	Hearing notice and proposal published in register	Feb. 2025
	Written proponents prehearing statements due	3rd week in Feb. 2025
	Deadline to establish party status	1st week in Mar. 2025
	Responsive prehearing statements and evidence due	3rd week in Mar. 2025
	Rebuttal comments due	3rd week in Apr. 2025
	Prehearing conference	4th week in Apr. 2025
	Rulemaking hearing for 303(d) List and M&E List	2nd week in May 2025

D. Process for Removing Waterbodies from the Section 303(d) and Monitoring and Evaluation Lists

This document addresses the procedures and protocols utilized by the division in assessing information for the purpose of identifying instances of nonattainment of water quality standards and subsequently, inclusion of affected waterbodies on either the 303(d) or M&E List. In general, removal of waterbodies/pollutants from either list is subject to the same requirements as those utilized for addition to the lists. Removal from the lists is considered appropriate where new information is developed which indicates that water quality standards are being met and/or designated uses are being attained. Considerations include more recent or more accurate data (for instance, chemical data generated using clean sampling/analytical methods), more sophisticated analysis using a calibrated model, identification of deficiencies in the original assessment or changes in standards, guidance or policy.

Table 2 includes the minimum number of samples required to remove (delist) a waterbody from the 303(d) or M&E Lists.

Table 2. Number of Samples Required for Delisting by Parameter	
Topic	**# of Samples Required for Delisting***
Stream chemistry	10
Lake chemistry (303(d))	5
Lake chemistry (M&E)	3
Lake profiles	Equal or greater evidence as data used for listing
E.coli (303(d))	5 within two-month period from last two years (all other two-month periods attaining within last two years)
E. coli (M&E)	2 within two-month period from last two years (all other two-month periods attaining within last two years)
Temperature (continuous)	Equal or greater evidence as data used for listing
Temperature (discrete)	5
Fish tissue	30
MMI	1 (most recent)
Sediment	1 (TIV and %fines – most recent)
Ambient -standards assessment approach	10
*Samples must be collected within the applicable 5-year period of record.	

With the exception of *E. coli*, MMI and sediment assessments, data from the entire period of record are used to assess attainment, not simply the most recent samples. Samples indicating attainment of water quality standards should be representative of the segment or portion currently on the 303(d) or M&E List. Data should also be collected from the same season as the data that originally indicated impairment. An exception would be in the instance where data collected utilizing conventional methods is supplanted by clean methods data or where the listing decision was based upon special study results for which it is impractical to reproduce. In any case, data must be adequate to characterize current water quality conditions. Assessments demonstrating attainment of designated uses should provide documentation of a nature similar to that used to support the listing decision. Attainment of water quality standards and uses will result in removal of the waterbody, or one or more of the listed parameters, from the list.

The commission will also consider removal when *good cause* is shown. As described in EPA's **2006 *Integrated Reporting Guidance*, *good cause*** for removing a waterbody (or water body pollutant combination) from the lists includes:

- The assessment and interpretation of more recent or more accurate data in the record demonstrate that the applicable classified uses and numeric and narrative standards are being met.
- The results of more sophisticated water quality modeling demonstrate that the applicable classified uses and numeric and narrative standards are being met.
- Flaws in the original analysis of data and information led to the waterbody pollutant combination being incorrectly listed.
- Demonstration pursuant to 40 CFR 130.7(b)(1)(ii) that there are effluent limitations required by state or local authorities that are more stringent than technology based effluent limitations, required by the Federal Clean Water

Act, and that these more stringent effluent limitations will result in attainment of classified uses and numeric and narrative standards for the pollutant causing the impairment.

- Demonstration that there are other pollution control requirements required by state, local or federal authorities that will result in attainment of classified uses and numeric and narrative standards within a reasonable time. (This element is EPA's Integrated Reporting Category 4b.)
- Documentation that the state included on a previous Section 303(d) List an impaired water that was not required to be listed by EPA regulation, e.g. waters where there is no pollutant associated with the impairment (This element is EPA's Integrated Reporting Category 4c).
- Approval or establishment by EPA of a TMDL since the last Section 303(d) List. (This element is EPA's Integrated Reporting Category 4a.)
- Inappropriate listing of a water that is located within Indian lands as defined in U.S.C. § 1151: *Indian Country Defined*.
- Other relevant information that supports the decision not to include the segments on the Section 303(d) List:
    - Adoption of revised water quality standards and/or uses such that the water is now in attainment of the revised standards and/or uses;
    - Development of a new listing methodology consistent with the state water quality standards and classifications and federal listing requirements;
    - A reassessment of the data that led to the prior listing, concluding that the waterbody is no longer impaired.

Barring unforeseen circumstances, the division will only propose to revise the lists during the regularly scheduled reviews (currently biennially). Other interested persons may petition the commission at any time to request a rulemaking hearing to revise the lists (either additions or deletions). However, such a hearing will be held only upon showing that failing to either add a segment to the list or delete a segment from the list prior to the next scheduled review will result in a substantial hardship to the party or parties requesting the hearing.

E. Process for Determining Category 4b Classification

An alternative to listing an impaired segment on the state's 303(d) List is an approved Category 4b demonstration plan. A Category 4b demonstration plan, when implemented, must ensure attainment with all applicable water quality standards through agreed upon pollution control mechanisms within a reasonable time period. These pollution control mechanisms can include approved compliance schedules for capital improvements or plans enforceable under other environmental statutes (such as the Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA)) and associated regulations. A Category 4b demonstration can be used for segments impaired by point sources and/or nonpoint sources. Both the division and EPA must accept a Category 4b demonstration plan for the affected segment to be placed in Category 4b. In the event that the Category 4b demonstration plan is not accepted, the segment at issue will be included on the 303(d) List, as Category 5.

Generally speaking, the following factors will be considered necessary for Category 4b demonstration plan acceptance: (1) appropriate voluntary, regulatory or legal authority to implement the proposed control mechanisms (through permits, grants,

compliance orders for Colorado Discharge Permit System permits, etc.); (2) existing commitments by the proponent(s) to implement controls; (3) adequate funding; and (4) other relevant factors appropriate to the segment.

The following evidence must be provided as a rationale for a Category 4b demonstration plan:

1) A statement of the problem causing the impairment;
2) A description of:
    a. pollution controls to be used;
    b. how these pollution controls will achieve attainment with all applicable water quality standards;
    c. requirements under which those pollution controls will be implemented;
3) An estimate of the time needed to meet all applicable water quality standards;
4) A schedule for implementation of the necessary pollution controls;
5) A schedule for tracking progress including a description of milestones; and
6) A commitment from the demonstration plan proponent to revise the implementation strategy and pollution controls if progress towards meeting all applicable water quality standards is not shown.

Timing for proposal submittal and acceptance by the division and EPA:

- Category 4b demonstration plans should be submitted to the division by the last week in January 2025 in order for the division to submit the plan to EPA by the first week in February 2025. Parties are encouraged to work with the division well in advance of this date as states are the entity required to submit these plans to EPA.

- Acceptance from EPA must be obtained by the last week in March 2025 otherwise the division will continue to propose that the segment in question is included on the 2026 303(d) List.

- If EPA and the division accept the Category 4b Plan, the division will notify the commission and public through proposed statement of basis and purpose language in its proposal that a Category 4b demonstration plan is accepted and is appropriate for this segment.

- Category 4b demonstration plans must be included in either the proponents prehearing statement or the party's responsive prehearing statement.

- Category 4b segments will be included in Regulation #93, and will be reported to the EPA as a part of the Integrated Report.

EPA has several documents that contain additional information on Category 4b demonstration requirements, including the *2006 Integrated Report Guidance*, available at http://www.epa.gov/OWOW/tmdl/2006IRG/#documents; and *Information Concerning 2008 Clean Water Act Sections 303(d), 305(b), and 314 Integrated Reporting and Listing Decisions*, available at: http://water.epa.gov/lawsregs/lawsguidance/cwa/tmdl/2008_ir_memorandum.cfm

F. Process for Determining Category 4c Classification

In cases where the impairment is determined to be caused exclusively by pollution, that does not result in pollutant(s) levels in excess of state water quality standards, the impaired waterbody may be placed into Category 4c. As defined by the Federal Clean Water Act, pollution is "the man-made or man-induced alteration of the chemical, physical, biological and radiological integrity of water" whereas pollutants are "dredged spoil, solid waste, incinerator residue, sewage, garbage, sewage sludge, munitions, chemical wastes, biological materials, radioactive materials, heat, wrecked or discarded equipment, rock, sand, cellar dirt and industrial, municipal, and agricultural waste discharged into water." (Section 502(19) and (6)). Segments classified as Category 4c are impaired however, a TMDL will generally not be required. Examples of circumstances where an impaired waterbody segment may be placed into Category 4c include segments impaired solely due to lack of adequate flow or to stream channelization. While low flows may be a human-induced condition (i.e., a reduced volume of water) fitting the definition of pollution, lack of flow sometimes leads to the increase of the concentration of a pollutant (e.g., sediment) in a water, such that a TMDL, which may consider variations in flow, is required. Segments located below dams or stream diversions with impaired biological communities (indicated by a failing multimetric index (MMI score) not caused by the presence of a pollutant(s), may be candidates for Category 4c.

Before placing impaired waterbody segments into Category 4c, thorough monitoring and assessment needs to be performed on the segment to confirm that no pollutants are contributing to the waterbody's failure to meet water quality standards. If adequate monitoring and assessment is not performed to rule out pollutant(s) as a cause, then the impaired waterbody should be placed on the 303(d) List (Category 5).

Proposals for Category 4c should be submitted as a part of the 303(d) rulemaking process in the division's proposal or a proponent's proposal, which is attached to the notice of rulemaking. Documentation of pollutant investigations such as chemistry data, proof of impairment and support of the identified pollution source must be submitted as part of the proponent's prehearing statement. Category 4c segments will be included in Regulation #93, and will be reported to the EPA as a part of the Integrated Report.

G. Process for Carrying Over Existing Attainment Conclusions during Waterbody Resegmentation and Portioning

1. Waterbody Segmentation:

   Water quality standards, and use classifications that are applied to Colorado waterbodies are reviewed and potentially modified through the triennial review process. Modifications to standards and use classifications on a waterbody can result in re-segmentation. Waterbody re-segmentation causes "splitting" because changes result in at least two new segments. In this situation the original segment is referred to as the parent and the newly created segments are referred to as child segments.

   When waterbody parent segments are split, child segments retain the existing attainment status of the parent segment. In general, the attainment status of child segments will be retained until they are reassessed as a part of the regularly scheduled rotating basin 303(d) Listing cycle. However, if new evidence

supporting a change in attainment status is submitted through a formal external party request, new child segments can be assessed in the upcoming 303(d) Listing cycle, regardless of the basin of focus. Instructions for external requests and proposals are provided in the 303(d) Listing Methodology section II.C.

2. Attainment Portions:

During the 303(d) water quality assessment process, when differing attainment conclusions exist within a single segment, the segment may be split into two or more portions. This process is described in greater detail in section IV.C, *Portioning of Segments*. Portions of segments are referred to as Assessment Units Identifications (AUIDs). They are represented by a capital letter after the waterbody id (i.e. COUAAR01_A). When assessment portions are created or modified, the result is at least two new portions. This is the same "splitting" process used at the segmentation level. In this situation, the original assessment portion is referred to as the parent and the newly created portions are referred to as child portions.

When a parent portion is split, child portions retain the existing attainment status of the parent portion. For example, if a parent portion is already listed as impaired for sediment and new dissolved copper data exceeds standards in one portion of the stream, but attains in another, two child portions would be created from the original parent portion to represent different copper attainment conclusions. The original sediment listing would be retained on both of the newly created child portions.

In general, the attainment status of child portions will be retained until they are reassessed as part of the regularly scheduled rotating basin 303(d) Listing cycle. However, if new evidence, supporting a change in attainment status is submitted through a formal external party request, new child portions can be assessed in the upcoming 303(d) Listing cycle, regardless of the basin of focus. Instructions for external requests and proposals are provided in the 303(d) Listing Methodology section II.C.

3. Insufficient Data Considerations:

In situations where child segments have insufficient physical, biological or chemical data within the current period of record, other relevant information demonstrating that the applied attainment status is not appropriate may be considered. This other information should demonstrate that the attainment conclusion reached in one part of the segment or portion is not representative of another part. This information could include but is not limited to landscape analysis (i.e. hydrology, vegetation, soils, and elevation), underlying geology or an investigation of activities in the watershed. For example, if a parent segment is considered impaired for a specific parameter, however landscape analysis demonstrates that the cause or the source of the impairment is limited to only one of the child segments, the case could be made that the impairment status is only inherited by the child segment located in close proximity to the source. Table 4, *Guidance for Listing Decisions based on Supplemental Source Information*, provides examples of source information that may result in different attainment conclusions for child segments or portions subsequent to the splitting process. These cases will be considered on an individual basis.

In cases when the resegmentation process results in changes to the same standards for which the segment is listed, carrying the impairment over to the new child segment may not be appropriate. In these cases, the previous presumption that the segment was of 'like water quality' may no longer apply. Therefore the previous commission impairment decision should be reconsidered prior to it being applied to the child segment. If there is no current data to confirm that the existing impairment is appropriate, a reasonable historic assessment will be conducted to determine if any of the data used in the previous impairment determination(s) were collected in the newly created child segment. If none of the data used in the impairment decision(s) were collected in the new child segment, that segment will be moved from the 303(d) List to the Monitoring and Evaluation List for additional data collection. This special consideration is not applied to the splitting of attainment portions because portions all exist within the same waterbody segment and are considered to be of "like water quality".

III. LISTING CRITERIA

This listing methodology sets forth criteria that generally will be used to make decisions regarding which waters to include on the Section 303(d) List and the M&E List. However, this methodology is not adopted by the commission as a rule. The commission is not bound by the criteria set forth in the listing methodology in making individual listing decisions if they determine on a site specific basis that an alternative approach provides a more appropriate method for assessing attainment of water quality classifications and standards in a particular circumstance.

A. Existing and Readily Available Data

In determining whether data and information is existing and readily available, the division will take into account such data and information as it has utilized in the preparation of those identification processes, calculations, and models referenced in 40 CFR §130.7(b)(5)(i), (ii) and (iv) and that credible data and information presented in a readily usable format and submitted in reports provided to the division as referenced in 40 CFR §130.7(a)(5)(iii). In addition, the division will accept and take into consideration credible data and information that is submitted to the division as part of the data call process, within specified data submittal deadlines. The division will also continue to independently collect and analyze new data on a rotating basin basis as part of its triennial review efforts and will utilize such data and information in making listing determinations.

1. Data Call

In June of each year, the division issues an annual data call letter to solicit data from specific regulatory basins. The letter is sent to contacts included in the commission's distribution list for both Regulation #93 and the applicable regulatory basins (Regulations #32-#38). Stakeholders have three months to submit data in the specified format with minimal data elements, such as detection limits, GPS coordinates, lab methods, etc. This solicitation for readily available data is consistent with requirements set forth in EPA's 2006 Integrated Reporting Guidance.

Existing data, which are not provided to the division in accordance with the data call schedule, set out in Section II.C, above, will not be treated as readily

available for purposes of making listing decisions. Such information will be considered in the next listing cycle unless these data and information are provided for consideration as written testimony during the rulemaking hearing process. The division and commission strongly encourage the submittal of data during the data call, as submitting new data during the hearing process limits the time available for all stakeholders to evaluate and assess the data and requires significant duplication of efforts by the division. The commission recommends that these data and information meet the criteria provided in Appendix E and be accompanied by a document describing how the criteria are met. Following these criteria will minimize the resource demand and duplication of efforts needed to process these data outside of the division's standard practices. Commission practice stipulates new information is not admissible in rulemaking hearings after the Responsive Pre-hearing Statement deadline.

It is important that data submitted for consideration in the 303(d) List development process is in a form that is amenable to existing division data management capabilities. Chemical data submitted for consideration in the list development process should be submitted in an electronic, WQX-compatible format. Physical and biological data should be submitted in a common electronic format that can be analyzed statistically. Recommended data reporting templates will be available at the time of the data request in June of each year. The division must be consulted regarding alternate formats. Data that are submitted in hard copy or alternate electronic format will be considered subject to division resource limitations, and may not be included in the division's assessment or proposal.

2. Assessment Process

The assessment process is intended to provide continuity with similar assessments done to support the standards review process as well as to efficiently utilize division resources. The division uses a rotating basin approach, approved by EPA, for periodic standards review and coordinates water quality monitoring and assessment to support the review. The following schedule in Table 3 sets out the

relationship between basin reviews and when assessments generated by those reviews will be incorporated in the 303(d) listing process for the first time.

Table 3. Rotating Assessment and Hearing Schedule

River Basins (Regulation Number)	Division Data Collection Effort	Period of Record (POR)	Data Call/ Assessment Season	Regulation #93 Hearing
South Platte (#38)	July 2021 – June 2022	2018 through 2022	Fall 2023	May 2025
Statewide #31	July 2022 – June 2023	2019 through 2023	Fall 2024	May 2025
San Juan #34 & Gunnison #35	July 2023 – June 2024	2020 through 2024	Fall 2025	May 2027
Arkansas & Rio Grande (#32 & #36)	July 2024- June 2025	2021 through 2025	Fall 2026	May 2027
Colorado Basin (#33 & #37)	July 2025 – June 2026	2022 through 2026	Fall 2027	May 2029
* Statewide data call for additional data for Regulation #93.				

B. Credible Evidence

The water quality assessment process depends on analysis of sufficient reliable data. Listing decisions not supported by adequate data are potentially flawed. The listing criteria are intended to assure that only those waterbodies for which adequate documentation of nonattainment is available are included on the Section 303(d) List. Waterbodies for which there is evidence to suggest impairment, but for which such documentation does not meet the standards for credible evidence, will be placed on the M&E List unless good cause is shown that it should be included on the 303(d) List.

Waterbodies may be included on the Section 303(d) List based on an evaluation of biological, chemical or physical data. The division will consider proposing to list a waterbody based upon consideration of all chemical, physical, and biological information that meets established sampling, analytical and interpretive protocols. Considerations include a review of the sampling and analytical methods employed. Factors to be considered include analytical detection limits, sample size (see section III.D.5.e), spatial and temporal distribution (see section III.C.1.f), variability within the data set, and the use of clean methods. Listing is often based upon chemical data alone, subject to the data interpretation criteria identified within this document. Listing based upon biological or physical data in the absence of accompanying chemical data requires that such information clearly demonstrate use impairment. Only representative data will utilized as the basis for the listing decisions.

The following guidelines are used to evaluate the adequacy of water quality information as a basis to support listing a waterbody.

1. Data Requirements – General

Information must be available to describe the methods used for sample collection, field and laboratory analysis. Persons submitting data for

consideration during the list development process must either provide the relevant quality assurance documentation with the submittal or assure that the documentation is available for the division to review.

The party submitting the data for consideration should provide the following information accompanying their data submission:

- Written assurance that the methods and procedures specified in the Quality Assurance (QA) plan were followed.

- Any field notes, laboratory comments, or laboratory notations concerning a deviation from standard procedures, quality control, or quality assurance that affects data reliability, data interpretation, or data validity may be requested by the division.

- Statement of the analytical methods used by the laboratory, method number, detection limits, quantitation or minimum levels, if available and any quality control samples and standards necessary to properly interpret data different from that stated in the QA plan.

- If requested by the division for interpreting or validating data, any other information, such as complete field notes, photographs, climate, or other information related to flow, field conditions, etc. This information should be retained by the submitter for a period of at least five years.

- Field instruments, such as multi-parameter devices, must be operated and calibrated according to manufacturer's recommendations or other acceptable demonstrated method. Calibration information and any other documentation of accuracy may be requested by the division.

Minimum information required for each data submittal must include the following:

- Location of each sample station in latitude and longitude with the associated reference datum, e.g., North American Datum 1983, etc.
- Waterbody name and sampling location description.
- Date the sample was taken.
- Parameter or condition measured.
- Measured value.
- Unit of measurement.
- For non-detect or non-quantifiable data, the less than value associated with the method detection limits (MDL) or reporting limits (RL) (i.e., LQL, LRL, PQL, etc.).
- Method used to measure the pollutant.
- Name and contact information of the party submitting the data.

Data submittals must include precise, sufficient information on the name of the waterbody and location of the sample station to allow for accurate mapping.

2. Sampling and Analysis Plans

Chemical data should be supported by a Sampling and Analysis Plan (SAP), which identifies sampling locations, contains analytical method references, and incorporates Quality Assurance/Quality Control (QA/QC) provisions. QA/QC documentation may include references to a standard QA/QC protocol. During review of chemical data submitted for evaluation, the division may require submittal of the SAP, QA/QC protocols and the results of QA/QC efforts. The division will provide any such information to other parties upon request.

3. Toxicity Tests

In-situ bioassay test results, or other ambient toxicity test results, must demonstrate adverse effects as measured by a statistically significant response relative to a representative reference or control. Inherent variability in toxicity testing results must be adequately taken into account. Listing decisions based upon toxicity test results require that any such results be corroborated by biological information clearly demonstrating impacts to aquatic community health, composition, or productivity. Data received utilizing whole effluent toxicity (WET) methods will be considered on a case by case basis.

4. Physical and Biological Assessment

a. Physical and biological assessments must be performed in accordance with scientifically sound methodologies. All such assessments should be performed by an observer who has training and experience in performing such evaluations. Assessment reports should include a statement of the observer's qualifications and should reference the protocols utilized. Any departures from referenced protocols and methodologies should be documented and the basis for any such departure addressed. The division's recommended collection and assessment methodologies for physical and biological data are described in commission Policy 98-1, *Guidance for Implementation of Colorado's Narrative Sediment Standard in Regulation #31, Section 31.11(1)(a)(i)* and Policy 10-1, *Aquatic Life Use Attainment, Methodology to Determine Use Attainment for Rivers and Streams*. A description of the division's assessment procedure's for both aquatic life and sediment are included in sections V.E and V.F.4.

b. The division will generally accept methodologies and protocols in use by the U.S. Geological Survey, U.S. Forest Service, U.S. Bureau of Land Management, EPA, Colorado Parks and Wildlife, or others, when well documented, widely available and suitable for their intended purpose. The division's determination of the acceptability or unacceptability of any such protocol will be included in the division's discussion of data sources included in the pre-hearing statement of the Section 303(d) List.

5. Period of Record

Data collected within the dates specified in the data call for the Period of Record (POR) may be submitted for consideration as part of the assessment process. In general, the POR includes the 5 full calendar years (January through December) prior to the data call deadline. Data which are collected within the POR and meet the other credible data requirements outlined in this listing methodology will be consolidated and assessed with other data.

Submitting data during the hearing process that were collected after the initial POR for the assessment cycle is discouraged. However, if this should occur, the starting date of the POR for the assessment cycle will remain the same and the end date will be extended to include the more recently collected data. Additionally, the commission recommends these data meet the criteria for data submitted after the data call in Appendix E.

Data older than five years must meet all current data requirements and will only be considered on a case by case basis for the following reasons:

- No newer data exists for the waterbody segment/parameter or the existing data does not meet the requirements of this listing methodology;
- The data are part of a larger dataset or long term monitoring which includes data younger than five years old for the same waterbody/parameter;
- Information or rationale is provided with the data to show that the data reflects current conditions and adheres to acceptable protocols.

Data older than five years may be used when necessary to determine historical natural conditions if the data meets the QA requirements in place at the time of its collection.

6. Anecdotal Information

Anecdotal information, in the absence of chemical, physical, or biological data, will not in and of itself be adequate to support a listing decision unless such information provides clear and convincing evidence demonstrating nonattainment. Anecdotal information includes, but is not limited to fishing logs, field logs, and historical or archival documents.

7. Representative data/information

Factors to consider when determining whether or not data are representative include: spatial distribution of sampling locations within the waterbody/segment, temporal variability of the data, changes in the watershed (i.e. changes in predominant land use, presence of new discharges, source removal or remediation projects), age of the data, method detection limits, bias in sampling design, etc.

Non-representative data include data collected within the mixing zone of a discharge. Data collected during or immediately after temporary events influencing the waterbody that are not representative of normal conditions shall typically be discounted in making the listing decision. For example, scouring storm flows which lead to diminished aquatic life use or accidental spills of toxic chemicals would not be a basis upon which to list the affected segment. However, such events may be considered as a basis for listing in instances where nonattainment of standards arises from a reversible source of pollutants.

Data collected during or immediately after fires, floods or other catastrophic events will not be used to make attainment decisions if the data are not representative of conditions prior to the event or new stable conditions. When determining if an event is considered substantial enough to impact or alter the

conditions that existed prior to the event, the following factors should be considered: severity of event, size of the affected area, distance of sampling sites from the event, hydrology, geomorphic effects that include soil types and slope.

Data collected from potentially impacted waters will be assessed to determine if the water quality parameters have returned to levels reflective of the levels before the event occurred or if water quality has stabilized. In the absence of data that characterize the conditions before an event, the division will work with all available resources to try and determine those conditions. In the absence of sufficient data to analyze trends, the following general timeframes will be used to screen data after two common events, fires or floods:

- Data collected after a substantial fire event will not be considered representative for 5 years after the fire. Data may be considered representative 6 to 9 years after a substantial fire event, if there is evidence that the watershed has recovered from the effects of the fire. Data collected 10 or more years after a fire event will be considered representative unless there is evidence that the fire continues to affect watershed function.

- Data collected after a substantial storm or scouring event will not be considered representative for a minimum of 4 weeks after the event. For the assessment of macroinvertebrates (using the approved methodologies that are stated in appendix B of Policy Statement 10-1), this time frame may be extended depending on the severity of the storm, possibly until after the following spring run-off. Some of the factors that will be taken into account when determining if the data will be considered representative will be the intensity and duration of the event, the season that the event occurred, the substrate (soft vs. hard bottom streams), drought and the severity of the channel impacts, which include the loss of riparian vegetation.

Ultimately, the decision regarding whether or not data collected during or after an event are representative of normal conditions will be evaluated, in collaboration with stakeholders, on a case by case basis, as each event is unique with varying impacts. If a segment is not attaining standards but the data are determined to not be representative due to fires, floods or other catastrophic events, that segment will not be added to the 303(d) List. Instead, the segment may be added to the M&E List while more data are collected.

During the assessment of sediment and MMI/aquatic life data, there may be situations where the most recent score is failing, however the vast majority of previous scores are attaining. In these cases, extra consideration will be used to examine the representative nature of the data. If the quality or representativeness of the data is in question, the segment will be proposed for the M&E List so that additional data can be collected.

For lakes and reservoirs, if a single profile indicates impairment but numerous attaining profiles exist in the dataset and other indications of impairment are

absent, the division may place the lake on the M&E List so that additional data can be collected. If less than three profiles are available for assessment, the division will use best professional judgment to determine if the data are representative. If the quality of the data is in question, the lake or reservoir will be placed on the M&E List so that additional data can be collected.

IV. DATA INTERPRETATION

The water quality assessment process considers the numeric and narrative standards assigned to a segment, as well as the assigned use classifications. Numeric standards are identified for a given pollutant and are expressed as a threshold value or as an acceptable range of values. Determination of attainment/nonattainment of pollutant specific numeric standards is a relatively straightforward statistical process.

Narrative standards describe threshold conditions that, if exceeded, result in unacceptable water quality conditions. Narrative standards that are applied to all surface waters in Colorado address sediment, floatables, film, odor, taste, color, toxins and excessive nutrients. Narrative standards may also include site specific temperature standards as provided at section 31.7(1)(c) of the *Basic Standards,* Regulation #31. Exceedance of narrative standards is more difficult to ascertain, as there are typically no quantifiable expressions of parameter concentration or loading that result in nonattainment. It is often the impact of pollution or of a pollutant, and not the pollutant itself, which is observed.

Use classifications identify existing or potential uses of the surface water segment. These include aquatic life, water supply, recreation and agricultural uses. Specific numeric standards are attached to a given use classification. Assignment of an aquatic life use classification to a segment typically results in assignment of a related suite of numeric standards. Attainment of numeric standards serves as a surrogate measure indicating attainment of the assigned use classification. However, nonattainment of an assigned use classification, as with narrative standards, may result from causes or parameters other than those assigned numeric standards.

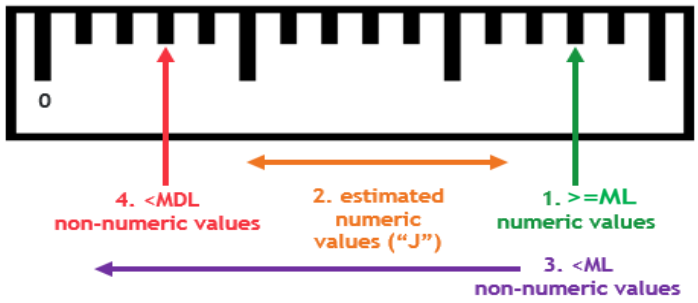
A. Detection and Reporting Limits

The reporting of analytical results is governed by detection and reporting limits of the analysis. A method detection limit (MDL) is the minimum concentration that can be measured and reported with 99% confidence that the analyte is present (as defined in 40 CFR Part 136) whereas the minimum level (ML) is the lowest concentration that an analyte can be accurately and precisely quantified according to the laboratory (Policy CW-6). The ML is typically two to ten times the MDL.

Some labs report data as a numeric value down to the ML, and then not detected or less than detection limits when the result falls below the ML (e.g., < ML). Some labs report estimated numeric values that are below the ML and greater than or equal to the MDL. These values are identified by the division with a "J" qualifier. Data that have the J qualifier have less statistical certainty as to the actual concentration, but this data can still provide information about whether a parameter is present in the sample. Therefore, any numeric values that are reported and are considered J qualified results (hereinafter J data), will not be replaced by zeros in the raw data.

Data results are reported as one of the following:

1) numeric values (>=ML)
2) estimated numeric values ("J" - <ML and >=MDL)
3) non-numeric values (<ML) with a U qualifier
4) non-numeric values (<MDL) with a DL qualifier



For assessment, non-numeric values (except for *E. coli* data), will be treated as zeros for assessment purposes. However, in some assessment scenarios, different data sets may have drastically different MDL or ML values. If values reported as not-detected for both data sets are replaced by zeros, the sample may then be unintentionally biased toward zero. To avoid this bias, it may be appropriate to omit the dataset with a higher MDL or ML from the assessment of the data.

If the dataset used for assessment purposes is comprised of J values in a high enough proportion that these estimated values may influence the assessment statistic (e.g., the 85th percentile), the data set will not be considered sufficient evidence for a 303(d) listing decision based on an exceedance of that particular standard's concentration. To determine whether the data set is sufficient evidence for a 303(d) listing, the Division will replace the J data values with zeros and compare the result to the assessment using the J data to determine whether the statistic is influenced by the J data. If replacing the J data values with zeros would still result in an exceedance of the standard, the segment may be added to the 303(d) List. Otherwise, the Division will recommend adding the waterbody to the Monitoring and Evaluation List while additional data are collected.

B. Sample Bias

Assessment techniques will be used that seek to reduce the effects of biased sampling. For example, the median of multiple samples taken within a seven day period will be used to represent that time period, and information gathered during synoptic (sampling at many locations at the same time) sampling events may be considered in a separate assessment so as not to bias the conclusions. Water quality data may be evaluated differently on a case by case basis if it is determined that data within a seven day period may not be representative of the given seven day sample period.

C. Portioning of Segments

Initially, all data submitted to the division for 303(d) assessment are evaluated by sampling location. Subsequently, the data from every station within a segment are combined for an assessment of the segment as a whole. If data are only available for a limited area of a segment, the conclusion reached for that area will be applied to the

entire segment, if the sampling area is representative of conditions that exist within the entire segment. For segments that indicate nonattainment, the division will investigate further to determine whether the impairment is widespread or limited to individual portions of the segment such as specific tributaries or reaches. Supplemental information will also be considered when determining the geographic extent of impairments. This information could include but is not limited to chemistry data, landscape analysis (i.e. hydrology, vegetation, soils, and elevation), underlying geology and an investigation of activities in the watershed. Table 4 provides guidance for listing decisions that could occur based on supplemental source information.

Table 4. Guidance for Listing Decisions Based on Supplemental Source Information

Supplemental Source Information	**Listing Decision**
No discernible source information	List entire segment
Geogenic source (example: selenium from shale formations)	List portion that shares common geogenic source formations
Suspected point or area source (example: metals from a legacy mining feature)	List portion effected by the suspected point or area source
Sediment from anthropogenic sources or erosion (example: dirt road crossings of streams)	List portion effected by the source causing sediment
Temperature impacts from diversions and dams	List portion effected by the diversion or dam
E. coli or pathogens in urban areas (example: tributaries to the South Platte River in Denver)	List entire segment
E. coli or pathogens in rural areas	If no upstream source is suspected (CAFO, septic systems), then list entire segment. If an upstream source is suspected, list only portion effected.
Other source information	List portion or entire segment – considered on a case-by-case basis

If evaluation of a data set for an entire segment does not indicate impairment, but specific location(s) within the segment consistently exceed acute or chronic standards, the specific portion of the segment may be listed. Portioning may also apply to lakes and reservoirs where sufficient data indicates impaired conditions are isolated to a specific portion of the lake. Segment portioning may also apply to those streams with MMI scores which demonstrate impairment, but not for the entire segment. Portioning for aquatic life using MMI scores will be decided on a case by case basis following Section V.E and must include representative samples within the same sampling index period.

When new assessment portions are created within segments with existing impairment classifications for other parameters, the existing impairment classifications are retained for the newly created portions. This portioning process will be consistent with language included in Listing Methodology Section II.G Process for Carrying Over Existing Attainment Conclusions during Waterbody Resegmentation and Portioning.

D. Sample size

Data sets comprised of two or three samples that indicate impairment of the chronic standard will result in placement on the M&E List except as noted for the ambient-based standards assessment approach, *E.coli* and lakes and reservoirs as shown in Table 5 below. Data sets comprised of four to nine samples where there is overwhelming evidence of nonattainment, or data sets of greater than or equal to ten samples indicating any degree of nonattainment, will result in inclusion on the 303(d) List except as noted for ambient-based standards, lakes and reservoirs.

For lakes and reservoirs, a minimum of five samples from the same location on different dates are required for the assessment of the metals and inorganic standards. If two to four samples indicate impairment, the lake is placed on the M&E List so that additional data can be collected. If the sample size is three or four and there is overwhelming evidence of impairment (see below), the waterbody will be placed on the 303(d) List. For the assessment of lake temperature and dissolved oxygen data, only a single profile is required for assessment.

For the assessment of *E. coli*, waterbody segments with data intervals containing two, three or four samples that indicate impairment of the *E. coli* standard will result in placement on the M&E List. Segments with *E. coli* data sets comprised of four samples where there is overwhelming evidence of non-attainment (see section IV.E - Overwhelming Evidence) will be placed on the 303(d) List. Data sets of five or more samples indicating any degree of nonattainment, will be added to the 303(d) List.

For the ambient-based standards assessment approach, a minimum of 10 samples are needed for conclusion of impairment (see Appendix B), except for data sets comprised of five to nine samples where there is overwhelming evidence of nonattainment will also result in a conclusion of impairment. In certain cases, where the sample size is less than five and the assessor suspects an impairment, the division may place the segment on the M&E List.

E. Overwhelming Evidence

Overwhelming evidence consists of sufficient and credible data that clearly demonstrate that a waterbody's designated uses are impaired. Overwhelming evidence is demonstrated when representative data (data that accounts for temporal and spatial variation) indicates an exceedance of numeric water quality standards by more than 50 percent in magnitude.

Table 5. Sample Size Requirements for the Assessment of Chronic Standards

Type	Sample Size (n)*	Overwhelming evidence?	List
Stream Chemistry	1	YES or NO	No Action
	2 or 3	YES or NO	M&E List
	4 - 9	NO	M&E List
	4+	YES	303(d) List
	10 or more	NO	303(d) List

Table 5. Sample Size Requirements for the Assessment of Chronic Standards

Type	Sample Size (n)*	Overwhelming evidence?	List
Lake Chemistry See below for lake nutrients	1	YES or NO	No Action
	2	YES or NO	M&E List
	3 or 4	NO	M&E List
	3+	YES	303(d) List
	5 or more	NO	303(d) List
E. coli	1	YES or NO	No Action
	2 & 3	YES or NO	M&E List
	4	NO	M&E List
	4+	YES	303(d) List
	5 or more	NO	303(d) List
Lake DO/ Temp.	1	N/A	303(d) List
Temp. - Streams	Less than 21 (equally spaced within 7 continuous days) per monitoring site	N/A	No Action
	21 or more (equally spaced within 7 continuous days) per monitoring site	N/A	303(d) List
Ambient-based standards assessment approach	Less than 5	YES or NO	No Action or M&E List (BPJ)
	5+	YES	303(d) List
	10 or more	NO	303(d) List
Aquatic Life	1	N/A	303(d) List
	2 or more	Multiple stations on same segment showing both attainment and impairment	M&E List or portion
Fish Tissue	Less than 10	YES or NO	M&E List
	10 – 30	NO	M&E List
	10 – 30	YES	303(d) List
	30 or more	YES or NO	303(d) List
Lake Nutrients (TN, TP, chl a)	Lakes - < 3 per season, for 2 or more seasons	YES or NO	M&E List
	Lakes - ≥ 3 per season; for 2 or more seasons	YES or NO	303(d) List
	DUWS (chlorophyll a) – < 5 per season for 2 or more seasons	YES or NO	M&E List

Table 5. Sample Size Requirements for the Assessment of Chronic Standards

Type	Sample Size (n)*	Overwhelming evidence?	List
	DUWS (chlorophyll a) – ≥ 5 per season for 2 or more seasons	YES or NO	303(d) List
	1 exceedance in a 5 year period	YES or NO	No Action
Stream Nutrients	TP/TN Streams – < 5 per season for 2 or more seasons	YES or NO	M&E List
	TP/TN Streams – ≥ than 5 per season for 2 or more seasons	YES or NO	303(d) List
	1 exceedance in 5 year period	YES or NO	No Action

*Samples must be collected within the applicable 5-year period of record.

F. Ambient -Based Standards

Ambient-based water quality standards are adopted where the table value standard cannot be met as a result of either natural conditions or irreversible, man induced conditions. Each ambient -standard is a site-specific characterization of existing quality derived from available representative data. To assess attainment of ambient-based standards, the division uses a statistical approach based on the concept of the confidence interval to minimize uncertainty of assessment conclusions.

If the lower confidence limit of the assessed value (e.g., 85th percentile) exceeds the standard, then the assessed concentration is significantly larger than the standard, and there is a high degree of confidence (95%) that the segment should be considered impaired. Without this statistical approach, applicable only to ambient -standards and iron, manganese and sulfate secondary water supply standards, there would be a much greater risk of incorrectly reaching an impairment decision. Because it has already been established that TVS cannot be attained due to natural or irreversible, man induced conditions, the evidence for further degradation (sufficient to warrant investment in a TMDL) should be especially compelling. Similarly, for water bodies previously determined to be impaired using the lower confidence limit approach, if the upper confidence limit of the assessed value (e.g., 85th percentile) does not exceed the standard, there is a high degree of confidence (95%) that the segment should be considered in attainment of the standard. This ensures the same degree of confidence for both impairment and attainment determinations.

Appendix B, *Assessing Attainment of Ambient-Based Water Quality Standards in Colorado* includes a detailed description of the statistical basis of this approach with examples of the assessment procedures and tables used to determine both the upper and lower confidence limits. The appendix also includes a description of other legitimate assessment methods which may be considered on a case-by-case basis, if it can be shown that the alternative approaches are more applicable or appropriate for the dataset in question.

G. Outstanding Waters

Attainment of water quality standards assigned for those segments designated as outstanding waters will be based upon the evaluation of ambient water quality characteristics and biological /physical data as described in Section V below. Attainment or maintenance and protection at their existing levels is assessed by comparison of current ambient water quality against water quality conditions at the time of designation (See ***Basic Standards and Methodologies for Surface Water***, 5 CCR 1002-31, section 31.8(1)(a)). The time of designation can usually be found in the statement of basis and statutory basis in the basin regulation for the segment in question.

H. Temporary Modifications

When temporary modifications of numeric standards have been adopted, attainment is assessed against the underlying standard, including those instances where the decision to assign a temporary modification is based specifically upon significant uncertainty as to the appropriate underlying standard (see section 31.7(3)(a)(ii)(A) of the Basic Standards).

V. DETERMINATION OF ATTAINMENT/IMPAIRMENT

Application of chemical, physical and biological information in listing determinations requires consideration of the scientific rigor of the methodologies utilized to develop any such information and the strength of that information. Rigor refers to the demonstrated validity of sampling, analytical, and assessment protocols and the availability of metadata in support of those protocols. Strength refers to the quantity of data and the extent to which such data demonstrates clear and convincing evidence of attainment or nonattainment of standards.

Availability of physical or biological data indicating use impairment may also be used to support listing when chemical data is otherwise insufficient in and of itself. Greater weight is given to data that provides direct, quantifiable documentation of impairment as opposed to data developed using surrogate indicators or parameters.

Attainment of numeric chemical standards is assessed by comparison of ambient water quality against assigned standards. Assessment of chemical data considers attainment of both chronic and acute aquatic life use based chemical standards, where both chronic and acute standards have been assigned to a given waterbody.

A. Assessment of Aquatic Life Use Based Standards

1. Chronic Standards

Attainment of chronic chemical standards, in both streams and rivers, and lakes and reservoir systems, is based upon the 85th percentile of the ranked data, except as otherwise noted below. Percentile values are calculated by ranking individual data points in order of magnitude and finding the value at which a certain percentage of the values fall above and a certain percentage fall below (for example, the value at which 15% of values fall above and 85% of values fall below for dissolved metals). Hardness based metal standards are evaluated by

comparing the 85th percentile against the assigned hardness based equation using mean hardness. Total recoverable metals are evaluated against the median value or the 50th percentile. Dissolved metals are evaluated against the 85th percentile. A waterbody is considered impaired if the standard is exceeded more than 15% or 50% of the time for dissolved and total recoverable metals, respectively. Dissolved oxygen (DO) is evaluated at the 15th percentile for streams, and a waterbody is considered impaired if the DO is below the standard more than 15% of the time. Minima pH is evaluated against the 15th percentile, maxima at the 85th percentile and a waterbody is considered impaired if either the pH is lower than the minimum 15% of the time or higher than the maximum 15% of the time.

Chronic standards can be evaluated in one of two ways:

a) A comparison of the assessment statistic for that parameter (50th, 85th percentile) to the standard, using the average hardness.
b) An evaluation of individual paired hardness and concentration assessments. The segment is considered not attaining if the paired calculations exceed more frequently than fifteen percent of the time.

In the case where both assessment approaches are conducted and the listing decision differs between the two, the paired hardness/concentration assessment decision is considered more representative.

2. Acute Standards

Acute standards are evaluated by comparison of single sample values to the assigned standard. For the assessment of metals standards, the acute table value standard is calculated for each paired hardness/concentration and attainment is determined for each data pair. In general, data indicates nonattainment of an acute standard if the standard is exceeded more frequently than once in three years. For metals, if more than one data pair exceeds the standard within a three year period, the waterbody is considered impaired. Where paired hardness and concentration data are not available, an assessment of the acute standards cannot be completed.

When determining attainment of the one-day acute standards, each site within a segment is assessed individually. If multiple samples are collected at an individual site during a single day, the average value of those samples is used. Next, the entire segment is assessed using the data from all of the sites. As was done with the individual sites, the average value of all samples collected on a single date are calculated and compared against the standard. Finally, even if the entire segment is shown to be in attainment of the acute standards but there are acute exceedances at individual sites, the segment may be portioned.

3. Temperature Standards

Numerical temperature standards are evaluated against representative instream data. Temperature varies within a reach both spatially and temporally, e.g. summer and winter. Data should be taken from a location in the stream that is representative of the reach at the time the data are collected. For example, data should not be relied upon that are taken only in locations that may be substantially warmer or cooler than the rest of the segment (e.g. backwater habitats, eddies, deep pools, or refugia).

a. Streams

i. Chronic: Attainment of the chronic numeric temperature standard is based upon a weekly average temperature (WAT) unless otherwise specified in a site-specific standard. The WAT is defined as the average of daily average temperatures over a seven-day consecutive period, with a minimum of three data points spaced equally through each day. The WAT is a simple moving average (rolling average) of the daily average temperatures over a seven day consecutive period. WAT are not to be overlapped (i.e. temperature data used in the calculation of one exceedance of an WAT will not be used in any other exceedance calculation). The maximum weekly average temperature (MWAT) is the largest WAT in the period of interest.

ii. Acute: Attainment of the acute numeric temperature standard is based upon a daily maximum (DM) water temperature unless otherwise specified in a site-specific standard. The DM is defined as the highest two-hour average water temperature recorded during a given 24 hour period. This value will be determined by calculating a rolling 2-hour average of the data record, then finding the maximum 2-hour average in a calendar day. For example, if data is collected every 15 minutes, a 2-hour average can be determined on every data point after the initial 2 hours of collection.

iii. Continuous Temperature Data: For the assessment of continuous temperature data (measurements collected by a temperature logger at a high frequency), the allowable average exceedance frequency is 1-in-3 years. The warming event is how the 1-in-3-year average exceedance frequency is applied to temperature. For data records with 4-6 years of data, an allowance will be made for one warming event within a single season (either summer or winter as defined by the temperature standards) within one year. For data records of 7 to 9 years, 2 warming events are allowed, if it is believed that data from the entire record is representative of the current condition. No warming event is allowed for data records with less than 4 years of data. The warming event is only applicable for the assessment of continuous temperature data. See Table 6 for a summary of the number of allowable warming events.

Table 6. The Number of Allowable Warming Events Based on the Number of Years of Data Assessed

Period of record evaluated	**# of warming events allowed**
Less than 4 years of data	None
4-6 years of data	One
7-9 years of data	Two

A “warming event” is defined as the maximum allowable extent of exceedances above the standard, defined in terms of degree-days (°C-days). This concept integrates both the magnitude of temperature (°C) above the standard as well as the duration (in days). The stream is allowed to exceed standards for a specific number of ‘degree-days’ within

a single season within one year. However, if the temperature exceedances of the standard at individual monitoring stations within a segment exceed the number of 'degree-days' specified in Table 7 or span multiple seasons or years, the entire segment, or the portion of the segment indicating nonattainment, will be placed on the 303(d) List as impaired for temperature. A warming event may not include days from more than one season or span multiple years in the period of record. The recurrence frequency of these warming events is limited to once every 3 years on average. See Appendix C for the technical basis used to establish degree-days for warming events and examples of the assessment procedures.

The following degree-days are allowed for Cold and Warm streams:

Table 7. Allowable Degree-days for Defining a Temperature Warming Event

	Acute	Chronic
Cold	2.4	13.5
Warm	3.8	35.5

Degree-days were derived using a biological basis and represent cumulative temperatures where growth or lethal impacts to fisheries are expected. See Appendix C. Numbers represent degrees Celsius.

To assess if one warming event has extended beyond its allowable extent, the following calculations are performed:

1. For every day the WAT or DM exceeds the standard, the difference between the WAT/DM and the standard is calculated.
2. A running total of these differences within a season within one year is compared against the 'degree-days' value in Table 7. If the running total is less than or equal to the 'degree-days' value in Table 7, the segment is not considered impaired.
3. If the running total of these differences exceeds the 'degree-days' value in Table 7 or spans multiple seasons or multiple years, then the temperature exceedances of the standard are no longer within one warming event allowance. The segment or portion of the segment is considered impaired.

In summary, the warming event, when applicable, is applied to individual monitoring stations within a segment. For data records of 4 to 6 years, an allowance will be made for one warming event in either the summer or winter. If the exceedances of the applicable temperature standard overlap multiple seasons or multiple years, this is considered more than one event for the purpose of applying the warming event allowance to determine whether the segment is considered impaired. Any data that is not covered by the warming event will be used to support a listing decision.

To delist a segment placed on the 303(d) List or M&E List due to exceedances of the temperature standard with continuous data, equal or greater evidence as the data used for the initial listing is required. All temperature measurements used for assessment purposes must be collected in accordance with or closely following the division's

temperature SOP *(WQCD – Standard Operating Procedures for the Collection of Stream Water Temperatures Utilizing the Deployment of Temperature Data Loggers).*

iv. Discrete Temperature Data: Discrete temperature measurements (measurements not collected by a temperature logger at a high frequency) may be used to assess attainment of acute temperature standards on a case-by-case basis, however, these waterbodies can only be placed on the M&E List. Discrete data cannot be used to assess attainment of chronic temperature standards unless a minimum of three equally spaced measurements are collected throughout a 24-hour period over seven consecutive days. The warming event is not applicable for the assessment of discrete temperature data.

To delist a segment that was placed on the M&E List due to exceedances of the acute temperature standard with discrete data, a minimum of 5 samples are required from the same season and conditions that indicated initial impairment. Samples should be collected in the middle afternoon, when temperatures are expected to be highest.

b. Lakes and Reservoirs
See Lake and Reservoir Section V.D.2.

4. Spawning Season Dissolved Oxygen (D.O.) Criteria

Assessment of the spawning season D.O. standard occurs through a two-step process.

- Step 1: Assess the 15th percentile of the available D.O. data and compare it to the spawning season D.O. standard. If the D.O. results are above seasonal standard, the waterbody is attaining, if the results are below the seasonal standard then we move to step 2.
- Step 2: When the 15th percentile of the D.O. data is less than the spawning season D.O. standard, the data set is divided into spawning/non-spawning season values and the 15th percentile of the spawning season data is compared to the spawning season D.O. standard.

If the 15th percentile of the D.O. data collected within the spawning season is below the spawning season standard, the waterbody is considered impaired.

The division will utilize spawning season information provided by Colorado Parks and Wildlife. In waterbodies that have spawning season D.O. standards, the default spawning season of mid-October through July is applied. Unless, specific river descriptions and spawning seasons are provided in Appendix A. Where more detailed fishery community information is available, the division will consider alternate spawning seasons as supported by such data.

For more information on the assessment of dissolved oxygen from lakes and reservoirs, see Section V.D.3.)

5. Numeric Nutrient Standards

Chlorophyll-a, total phosphorus, and total nitrogen will be assessed in waterbody segments as these standards are adopted by the commission. The anticipated standards adoption schedule for numeric nutrients is documented in Regulation #31 Section 31.17.

a. Lakes (See Section V.D.6.)

b. Rivers and Streams

For rivers and streams, the assessed total phosphorus and total nitrogen concentrations are the annual median values. When compared to the nutrient standards, these values have an allowable exceedance frequency of once in any five-year period. In cases where the median nutrient concentrations exceed the standard but where there are fewer than five samples from that year, those streams will be placed on the M&E List until additional data can be collected.

For rivers and streams, chlorophyll *a* (a measure of the areal abundance of attached algae or periphyton) is assessed during the summer season from July 1 to September 30. Assessment is based on the summertime maximum, "consistent with its foundation in a study of public responses to snapshot observations" (31.50(III)(B)). The allowable exceedance frequency is once in five years. Unlike the assessment of stream nutrient concentrations, only one sample is required for the assessment of stream chlorophyll *a*. However, the attainment of chlorophyll *a* for streams can be assessed where a representative sample can be obtained with the division's sampling protocol, which is designed for hard substrate. This protocol is titled *Standard Operating Procedures for the Collection of Periphyton Samples*, and is available upon request).

6. Listing Based on Elevated Mercury in Fish Tissue

Waterbodies are assessed for attainment of Colorado's aquatic life use (EPA's fishable goals of CWA Section 101(a)(2)) by comparing the weighted average fish tissue mercury for each species/size class to a 0.3 parts per million (ppm) threshold level. For small datasets with a large portion of the data below the detection limit, the division will substitute half the detection limit when calculating the weighted average. Those waterbodies with a weighted average fish mercury concentration for each species/size class that exceeds the 0.3 ppm threshold level will be placed on the 303(d) List. A minimum of 30 fish tissue samples (either as individual samples or composites) from each species/size class are necessary to determine impairment of a waterbody for mercury in the fish tissue. For waterbodies where the data is short of this requirement, the waterbody will be placed on the M&E List so that additional data can be collected for assessment. If the sample size is between 10 and 30 and the weighted average fish tissue mercury concentration is greater than 1.5 times the threshold level, the waterbody will be placed on the 303(d) List based on overwhelming evidence of impairment.

Those waters that are listed due to elevated levels of mercury in fish tissue may be identified as low priority (notwithstanding the provisions of Section IV.B.1 below) when the provisions applicable to EPA reporting Category 5m are satisfied (see Information Concerning 2008 Clean Water Act Sections 303(d), 305(b) and 314 Integrated Reporting and Listing Decisions, EPA, October 12, 2006). Waters

are placed in reporting Category 5m in instances where impairment is due to atmospheric deposition and where the state has a comprehensive mercury control program in place. The division will evaluate each listing arising from mercury levels in fish tissue for evidence of current and historic mining activities within the contributing watershed, for other potential industrial sources and for potential geologic influences.

B. Assessment of Agriculture and Domestic Water Supply Use Based Standards

These standards are expressed in terms of either one day or 30 day averaging periods (comparable to acute and chronic based standards, respectively) and are assessed by comparison of the percentile ranges (described in SectionV.1.a) against the standard. For assessments of standards listed as total fractions but where total species data is not available, the dissolved metals fraction is used to evaluate the standard as a conservative approach. To evaluate total standards expressed as 30 day averages, the 50th percentile of the dissolved data is assessed. For evaluation of standards expressed as one day averages, the individual dissolved values are compared to the standard. For assessment of standards expressed as one day averages, data indicates nonattainment if the standard is exceeded more frequently than once in three years.

1. Nitrate/nitrite and arsenic:
   Assessment of nitrate and/or nitrite, and arsenic water supply use based standards will consider the combined total or individual ambient concentrations for nitrate and/or nitrite and the individual ambient concentrations for arsenic. Nitrate, nitrite and arsenic standards are assessed along the entire segment for those segments where water supply standards have been adopted, regardless of whether or not there is a point of intake identified on the stream. The assessment will consider assessments and data used in permits development and will portion the impaired segment accordingly.

   Where a range is specified for the arsenic standard, waterbodies will be considered in attainment of this standard, and not included on the Section 303(d) List, so long as the existing ambient quality does not exceed the second number in the range.

2. Manganese, iron and sulfate:
   For segments with adopted iron, manganese and sulfate secondary water supply standards, the less restrictive of the following two options shall be applied as the numeric standard for assessment for the 303(d) List:

   i. Table Value Standard - Iron (Dissolved) (Fe-D) – 300ug/L
      Table Value Standard - Manganese (Dissolved) (Mn-D) - 50 ug/L
      Table Value Standard - Sulfate (SO4) – 250 mg/L

      Or

   ii. Existing quality as of January 1, 2000.

   To determine the existing quality as of the year 2000, a segment-specific library of water quality data has been created for these parameters. This 2000 water

supply data library contains data from January 1, 1995 to the most recent period of record. After a cleaning and quality control process, the division adds all data available from the Water Quality Portal for this time period to the data library every five years to coincide with the Regulation #31 assessment cycle. Additional iron, manganese, and sulfate data may be included in the data library by submitting it during the statewide data call. The submission should specify these data are for inclusion in the data library.

The division will aggregate data from 1995-1999 to characterize conditions from the year 2000. If less than ten samples are available from this time period, the period is extended using 5 year increments until at least ten data points are available. Note, data from the data library can only be considered representative of conditions from January 2000 if no new or increased sources of iron, manganese, and/or sulfate are known in the segment. If changes related to iron, manganese, and/or sulfate are known, only data collected prior to those changes can be used in assessments.

Determining the least restrictive standards for dissolved iron, dissolved manganese and sulfate and assessing current water quality requires the following steps:

1. As a screen, compare water quality data from the current assessment period to table value standards (TVS).

2. For each parameter, if table value standards are not exceeded, the segment is considered attaining water quality standards. This concludes the assessment process for these parameters. Examination of water quality as of the year 2000 is not necessary in these cases.

3. For each parameter, if table value standards are exceeded, data for the segment in which the sampling station is located is compiled from the 2000 data library.

4. For segments with at least ten samples in the 2000 data library from 1995-1999, the 85th percentile for each parameter is calculated to represent water quality as of the year 2000. This value is then compared to the table value standard and the least stringent value is used to evaluate attainment.

5. If fewer than ten samples representative of the year 2000 (1995-1999) are available for that segment, new or increased sources of iron, manganese and sulfate that may have impacted the segment after 2000 are examined. The methodology for determining new or increased sources of these parameters is described below.

6. If the source investigation demonstrates there are no new or increased sources of iron, manganese or sulfate impacting the segment, data collected subsequent to the year 2000 will be used to establish existing quality as of January 1, 2000. If the only data available for the segment is from the current period of record, it is presumed that current ambient data is representative of existing quality as of January 1, 2000

and therefore the segment is not considered impaired.

7. If the source investigation demonstrates that there have been new or increased sources of iron, manganese or sulfate along the segment, table value standards will be used to determine attainment and the segment will be considered impaired. This concludes the assessment of these parameters.

8. If the water quality as of the year 2000 value is greater than the parameter's table value standard, the method outlined in the Appendix B, *Assessing Attainment of Ambient-Based Water Quality Standards*, will be used to determine attainment. If the lower confidence limit of the assessed value (85th percentile) of the current period of record exceeds the water quality as of the year 2000 value, the segment will be considered impaired. For water bodies determined to be impaired using the lower confidence limit a delisting (attainment of the water quality standard) is appropriate when the upper confidence limit of the assessed value (85th percentile) of the current period of record does not exceed the water quality as of the year 2000 value.

Appendix B, *Assessing Attainment of Ambient-Based Water Quality Standards* in Colorado, includes a detailed description of the basis of the upper and lower confidence limit approach.

Existing quality as of the year 2000 values for iron, manganese and sulfate are not ambient-based water quality standards. However, utilizing the same upper and lower confidence limit statistical approach outlined within Appendix B, for existing quality as of the year 2000 assessment will provide some benefit. If the lower confidence limit of the current water quality exceeds the existing quality as of the year 2000 value, there is a high degree of confidence (95%) that the segment should be considered impaired. Similarly, if the upper confidence limit of the current water quality does not exceed the existing quality as of the year 2000 value, there is a high degree of confidence (95%) that the segment should be considered in attainment of the water supply standard.

Pollutant Source Investigation:

If a point source discharge that has the potential to discharge dissolved iron, dissolved manganese or sulfate, is identified along the segment or 5 miles upstream of the segment, then the permitted facility will be evaluated for new or increased levels of iron, sulfate or manganese. If the permitted facility began discharging after January 1, 2000, this will be considered a new source. For facilities discharging before 2000, effluent data (reporting on Discharge Monitoring Reports) collected from before the year 2000 to the current assessment cycle may be used to determine whether there have been increases in the discharge of these parameters. If there has been no increase, and no other new sources exist within this segment, data collected subsequent to the year 2000 will be used to establish existing quality as of January 1, 2000. Temporary discharges, such as construction dewatering permits, will not be considered in this point source evaluation.

Nonpoint sources of dissolved iron, dissolved manganese and sulfate may be considered on a case by case basis, when information indicates the existence of new or increased sources of these parameters.

Notwithstanding the above provisions regarding dissolved iron, dissolved manganese, and sulfate, the division will consider information brought forward by interested parties indicating that existing quality for these constituents in stream segments as of January 1, 2000, was affected by an unauthorized discharge with respect to which there was an enforcement action by the division or remedial action under the Comprehensive Environmental Response Compensation and Liability Act. In such circumstances, and pursuant to 31.11(6), the applicable numerical standards shall be the ambient conditions existing prior to the unauthorized discharge or the table value criteria, whichever is less restrictive.

The division will presume that an actual water supply use is present in all segments with a 'WS' standard applied to iron, manganese and sulfate in the basin regulations and will assess using the associated water supply standards. If a party disagrees with this presumption, they can provide credible evidence to the division as a part of the Regulation# 93 rulemaking process. The division will not propose 303(d) water supply impairment listings for iron, manganese and sulfate for those segments where 'actual' water supply use is not present."

Figures 1 and 2 show the steps required for the assessment of iron, manganese and sulfate water supply standards.

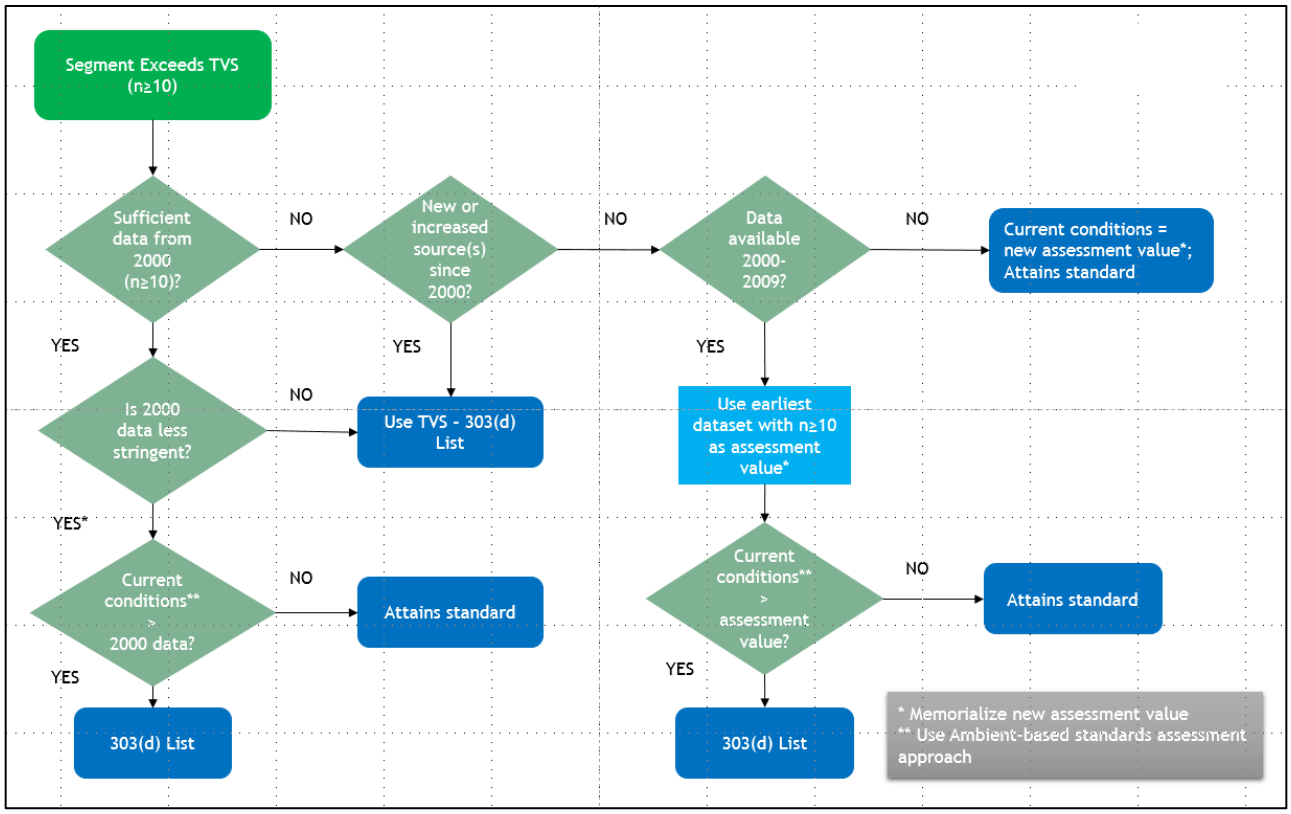


Figure 1. Flow Chart for the Assessment of Dissolved Iron, Dissolved Manganese and Sulfate Water Supply Standards

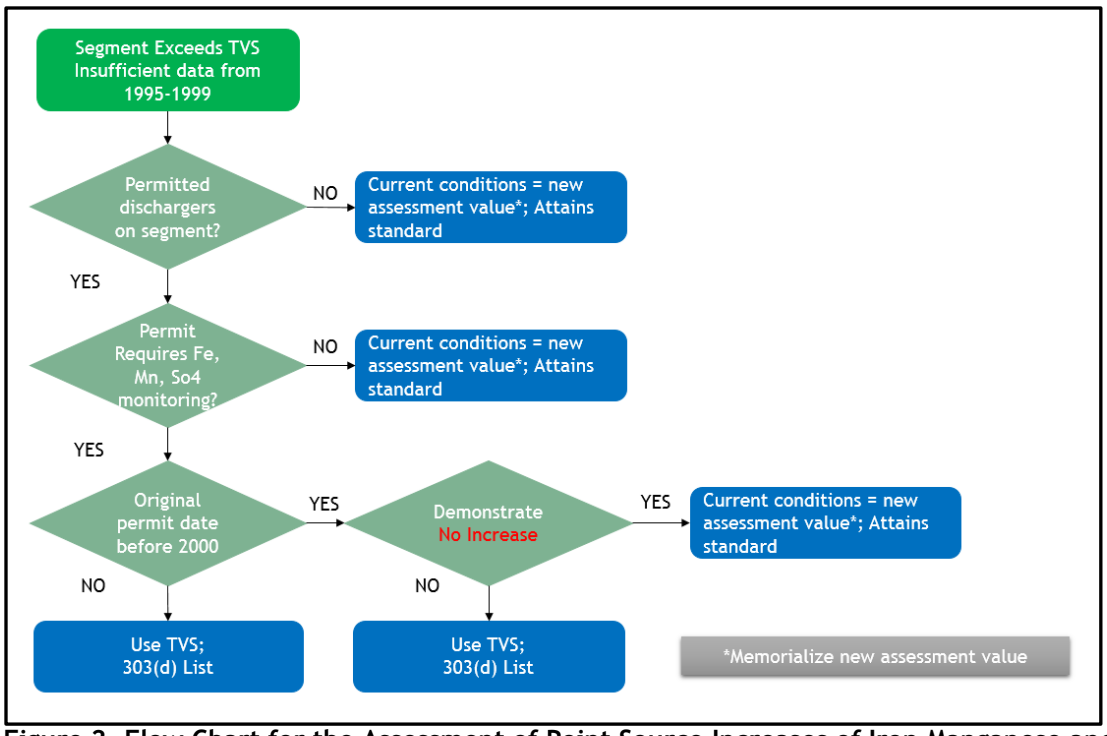


Figure 2. Flow Chart for the Assessment of Point Source Increases of Iron Manganese and Sulfate

C. Assessment of Recreational Use Based Standards

1. *E. coli* Standards

The *E.coli* standard is expressed as the geometric mean of samples collected within a two-month period. To evaluate this two-month criterion, the division calculates the geometric mean for all *E. coli* data collected within static two-month periods. The static two month periods are January and February, March and April, and so on through November and December. If samples are available for only one of the months in the period (e.g., samples available from March but not April), the geometric mean for that period is based solely on samples collected during that month.

To address sample bias (see Section IV.B), the geometric mean of E. coli samples collected on the same day at the same site, portion, or segment will be calculated. This geometric mean will be treated as a single sample.

Waterbody segments with two-month periods made up of two or three samples, after bias correction as described above, that indicate impairment of the *E. coli* standard will result in placement on the M&E List. Segments with *E. coli* data sets comprising four samples where there is overwhelming evidence of non-attainment (see section – IV.E. - *Overwhelming Evidence*) will be placed on the 303(d) List. If there are four samples with an indication of impairment but the evidence is not overwhelming, the segment will be placed on the M&E List. Data sets of five or more samples indicating any degree of non-attainment will be added to the 303(d) List.

To remove a segment from the 303(d) List, the geometric mean for at least one two-month period that includes at least 5 samples from the most recent two years of available data in the period of record must demonstrate attainment. To remove a segment from the M&E List, the geometric mean for at least one two-month period that includes at least 2 samples from the most recent two years of available data in the period of record must demonstrate attainment. For both 303(d) and M&E delistings, the two-month period must fall within the same season as the period that originally triggered the listing (see Section II.D). In addition, all other two-month periods from the most recent two years must demonstrate attainment, regardless of sample size.

In order to accurately capture representative conditions over a two-month period and to avoid the potential for bias, waterbodies will not be considered for 303(d) listing or delisting based exclusively on data from samples collected within a seven-day period. Also, to allow for calculation of the geometric mean, which requires that all contributing values be non-zero, *E. coli* data reported as less than detect (<) and/or zero (0) will be converted to a value of one (1). Similarly, reported values greater than an upper analytical threshold (>) will be converted to the applicable upper threshold value. Finally, segments with site-specific standards will be evaluated using the assessment method described here against the applicable standard. The applicable standard may change the magnitude of the standard or limit the times of year during which the standard applies.

2. Nutrient Standards

(see Aquatic Life Use Based Standards Assessment Section V.A.5)

D. Assessment of Data Collected from Lakes and Reservoirs

1. General

The sampling strategy for lakes differs from streams in important ways that affect the assessment of water quality. Typically, lakes are not sampled as often as streams because the large volume of water buffers against the short term changes in quality. In contrast, the spatial coverage within the vertical profile should generally be more comprehensive for lakes, especially where lakes are stratified in the summer.

Typically, two strategies are applied simultaneously when sampling lakes – vertical profiles and discrete samples. It is common to measure some parameters, usually temperature, DO, pH and conductance, in vertical profiles that yield measurements at closely spaced intervals from top to bottom in a lake or reservoir. Profile data are essential for defining the internal boundaries of layers that form when thermal stratification develops in the summer months, or for demonstrating that no stratification exists. The preferred sampling location for lake and reservoir profiles is in the deeper part of the lake or reservoir, most commonly in front of the dam for a reservoir.

a) Profile Data

The interpretation of profile data begins by examining the mixed layer, which is that part of a lake that is well mixed by wind action and can be expected to have relatively homogenous physical and chemical conditions. For assessment purposes, the mixed layer is evaluated by examining conditions in the upper portion of a lake. The upper portion is generally characterized within a single profile as follows:

1. Where a lake or reservoir is equal to or greater than five meters deep, measurements within a single profile are generally assessed as the average of all measurements from 0.5 meter to 2.0 meters.
2. Where a lake or reservoir is less than five meters deep, but more than 1.25 meters deep, measurements within a single profile are generally assessed as the average of all measurements from 0.5 meter to a depth equal to 40 percent of total depth.
3. Where a lake or reservoir is 1.25 meters deep or less, measurements within a single profile are generally assessed as the median of all measurements.

In a stratified lake, the upper portion is separated from a deeper, cooler layer (referred to as the lower portion) by a transition zone of rapid temperature change (thermocline).

The lower portion of a lake is assessed by averaging the measurements from one to three meters above the bottom of the lake. For example, to assess the lower portion of a lake with a maximum depth of six meters, profile measurements would be averaged from three to five meters. For lakes less than five meters deep, the lower portion is not assessed. This definition for the lower portion of a lake is only used for the purpose of pH assessment.

In cases where multiple data points along a profile are not available or feasible and only single data points are collected, a single data point from each of the upper and lower portions may be assessed against the standard, if the single data points are determined to be representative.
If multiple profiles are collected from various locations for a lake on the same day, each profile will be evaluated separately. Multiple profiles will not be averaged for assessment purposes. If the division determines impairment is isolated to an appropriate sub-segment or portion of a lake, the division may place the portion on the 303(d) List.

b) Discrete Samples
Discrete samples are used to characterize conditions at specific depths, often intended to represent a single layer. Discrete samples from lakes are analogous to grab samples taken for stream assessments. It is common to take samples from the top and bottom of each lake (which would correspond to upper and lower portions in a stratified lake) because the water quality characteristics of those two major habitat regions often diverge significantly during the growing season. It is much less common to take discrete samples from the thermocline for two reasons: it is a boundary zone with steep environmental gradients and water quality characteristics will be intermediate between those of the adjacent layers.

2. Temperature

Vertical profiles provide a record of temperature at closely spaced intervals from the top to the bottom of the water column. Unlike streams, daily fluctuation of temperature in lakes tends to be quite small. Thus, the temperature observed at each depth in the profile is assumed to be persistent on a scale of days, making it a surrogate for the weekly average temperature (WAT). The average temperature of the mixed layer (as defined in Section D.1.a) is used to assess attainment of the applicable temperature standard for the lake or reservoir being evaluated. When a lake or reservoir is stratified, the mixed layer may exceed the applicable standards, provided that adequate refuge exists in water below the mixed layer. Adequate refuge means that there is concurrent attainment of the applicable table value temperature standard (Table 1, Regulation #31) and dissolved oxygen criteria. If the refuge is not adequate because of dissolved oxygen levels, the lake or reservoir may be included on the 303(d) List as "impaired" for dissolved oxygen, rather than for temperature. In cases where temperature is exceeding at all depths in the profile, the lake may be listed as impaired for temperature.

3. Dissolved Oxygen

Assessment of dissolved oxygen (DO) within a profile of a lake or reservoir is accomplished by comparing the average of the measurements from the mixed layer, as defined above in Section D.1.a, to the applicable standard. If the average DO concentration in the mixed layer does not meet the applicable standard, an assessment of adequate refuge may be performed. Adequate refuge (Regulation #31 section 31.16 Table I footnote (9) c) is a region in the water column concurrently in attainment of DO and temperature standards. If adequate refuge is available, the lake or reservoir will not be considered impaired for DO for that date.

Fall turnover exclusion: DO may drop 1 mg/l below the criteria in the upper portion of a lake or reservoir for up to seven consecutive days during fall turnover provided that profile measurements are taken at a consistent location within the lake seven days before, and seven days after the profile with low DO. The profile measurements taken before and after the profile with low DO must attain the criteria in Table 1 (Regulation #31) in the upper portion of the lake or reservoir. The fall turnover exclusion does not apply to lakes or reservoirs with fish species that spawn in the fall unless there are data to show that adequate DO is maintained in all spawning areas for the entire duration of fall turnover (Regulation #31, Table I, Footnote 9(e)(i)).

4. pH

Data for pH often are available from vertical profiles, but the data are generally evaluated in the context of discrete samples. There are two reasons for this approach – not all sampling programs obtain pH in profiles, and pH must be determined for any discrete sample wherever ammonia is of interest. Discrete samples from the upper and lower portions are evaluated separately because they represent different habitat regions in a stratified lake. When variations in pH are driven largely by biological processes within a lake, the risks of exceedance are generally associated with high pH in the upper portion (due to

high rates of algal productivity) and low pH in the lower portion (due to high rates of decomposition).

Assessment of the pH standard for a lake is accomplished by calculating the average pH from the upper and lower portions of the lake for each profile as defined above in the profile data section (V.D.1.). The 15th and 85th percentiles of the sample averages for each portion are then compared to the minima and maximum pH standard for the determination of attainment.

Failure to attain the standard in either layer results in 303(d) listing.

5. Metals and Inorganics

These constituents are generally assessed on the basis of discrete samples (grab samples), for which the methods for data interpretation have been outlined above in the discrete samples section. For the reasons explained under the subsection on pH, discrete samples from the upper and lower portion of a lake should be assessed separately.

Failure to attain the standard in either layer results in 303(d) listing.

6. Nutrients

Nutrient concentrations in lakes and reservoirs are assessed as the seasonal average of values from the mixed layer. The mixed layer is defined in Section V.D.1, and the seasonal boundaries are defined in Section V.D.6(b) below. When samples are collected from multiple depths in the mixed layer on the same date, the median of those values will represent the assessed concentration for that date. The annual seasonal average concentrations are compared against the standards with an allowable exceedance frequency of once every five years.

a) Sample Size

The minimum sample size required to place a lake or reservoir on the 303(d) List for an exceedance of the nutrient standard is at least three representative samples per season with two exceeding seasons within a five year period. In instances where the average nutrient concentrations exceed the standard for two or more seasons but where there are fewer than three representative samples in these seasons, the lake or reservoir will be placed on the M&E List until enough additional data can be collected to determine impairment status.

Direct Use Water Supply (DUWS)

For lakes and reservoirs designated as DUWS, a minimum of five representative samples in a season (as defined in Section V.D.6(b)) are required for the assessment of chlorophyll *a*. As with other lakes and reservoirs, two or more exceeding seasons are required for placement on the 303(d) List. In instances where the average chlorophyll a concentration exceeds the standard for two or more seasons but where there are fewer than five representative samples in these seasons, the lake or reservoir will be added to the M&E List until enough additional data can be collected to determine impairment status.

b) Seasonal Boundaries

Nutrient data used for assessment of lakes and reservoirs must be collected between July 1 and September 30. For lakes and reservoirs designated as DUWS, chlorophyll *a* must be collected from March 1 through November 30. Nutrient data collected outside of the specified seasons will not be considered in the 303(d) assessment process for lakes and reservoirs.

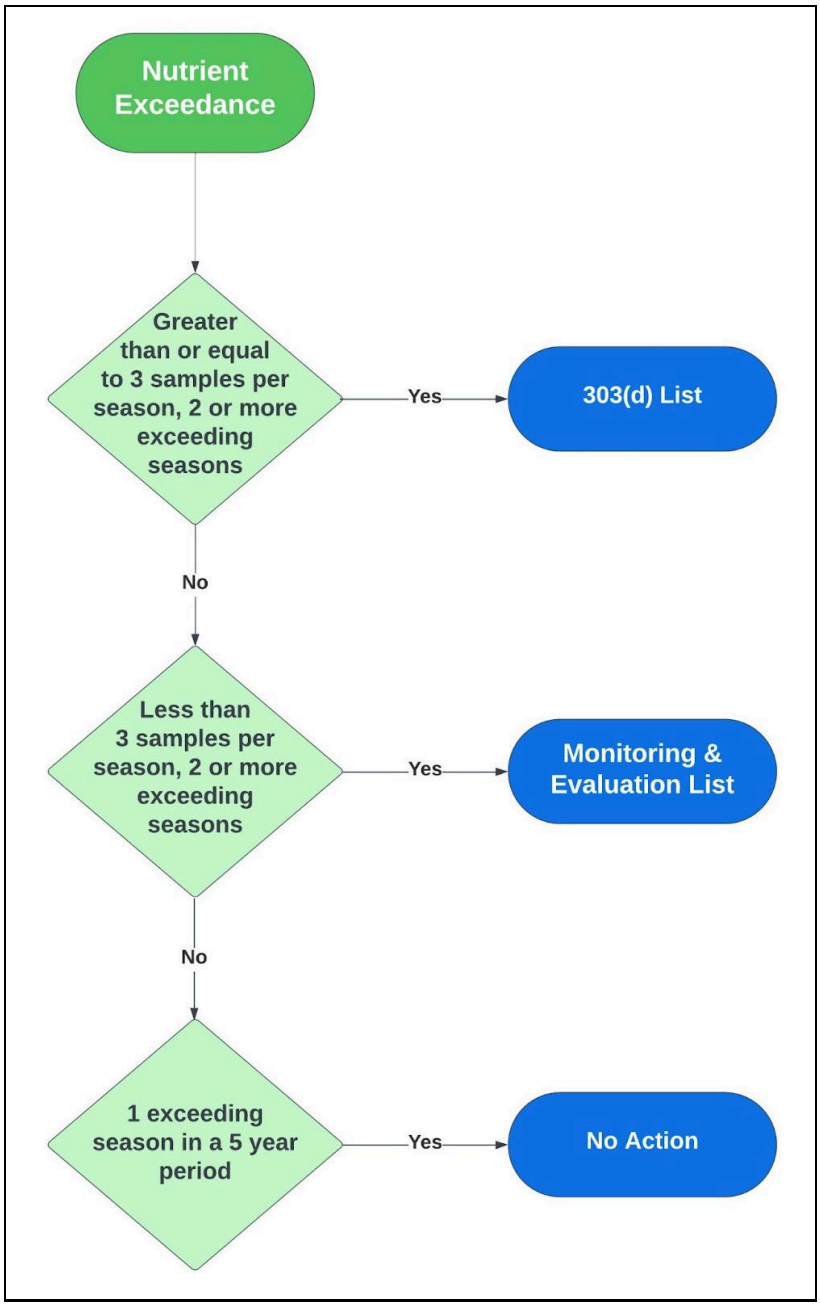


Figure 3. Flow chart for the assessment of total phosphorus, total nitrogen, and chlorophyll a for lakes and reservoirs. All samples must be collected within the current period of record and within the seasonal boundary of July 1 to September 30.

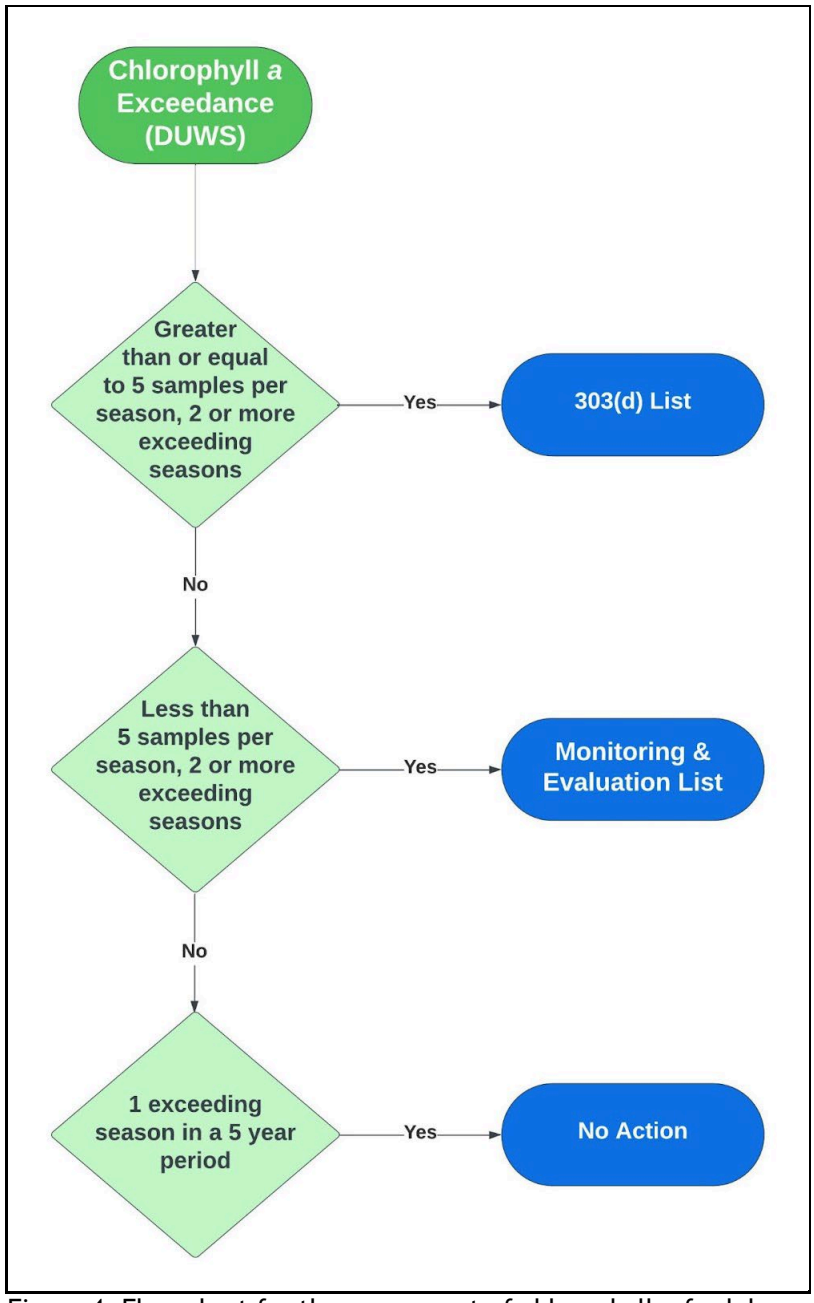


Figure 4. Flow chart for the assessment of chlorophyll a for lakes and reservoirs designated as direct use water supply (DUWS). All samples must be collected within the current period of record and within the seasonal boundary of March 1 to November 30.

7. Site Specific Standards in Control Regulations

Some lakes and reservoirs have been assigned site specific standards for nutrients (total phosphorus) and chlorophyll *a*. These site specific standards are identified in control regulations which are specific to a given waterbody. These presently include Dillon Reservoir, Cherry Creek Reservoir, Chatfield Reservoir, and Bear Creek Reservoir, which are evaluated on an annual basis for compliance with site specific standards. Usually, the period for application of site specific standards is defined as the growing season and is described in the statement of basis and purpose for that standard. For example, growing season

data are used to determine attainment with standards for phosphorus. Any determination of site specific standards attainment must be based upon application of such standards in a manner consistent with the applicable control regulation.

E. Assessment of Biological and Physical Data

The assessment of biological and physical data supports the determination of impairment or attainment of the narrative standards, as defined in section V(F), and the classified aquatic life use that are described in both Policy Statement 10-1, Aquatic Life Use Attainment, Methodology to Determine Use Attainment for Streams and Rivers (Policy 10-1), and Regulation #31, The Basic Standards and Methodologies for Surface Water (5 CCR 1002-31) (Regulation 31).

For the assessment of the biological condition of a stream or river, the Division primarily focuses on the assessment of benthic macroinvertebrates, as outlined in Policy 10-1.

For the assessment of the narrative standard, the division assesses sediment and physical data according to Policy 98-1, discussed further in section V(F)(4).

1. Policy Statement 10-1, Aquatic Life Use Attainment, Methodology to Determine Use Attainment for Streams and Rivers

The assessment of biological data, or bioassessment, is defined in Policy 10-1 as the evaluation of the biological condition of a water body using biological survey data and other direct measurements of resident biota in surface waters. A biological threshold is the establishment of numeric criteria against which the current biological condition can be evaluated. The use of bioassessments provide a means to evaluate the biological condition in relation to the goal or expected condition, like a designated use.

There are many types of biological assemblages that are available for study (i.e periphyton and fish) and all respond differently to anthropogenic stressors. A decision was made by the division to focus on the bioassessment of benthic macroinvertebrates because of a combination of their short life spans, limited mobility, and ease of standard collection and bioassessment protocols. Additionally, benthic macroinvertebrates occur in all Colorado streams that support aquatic life.

Policy 10-1 describes the bioassessment or Multi-Metric Index (MMI) tool that is used in Colorado. The MMI tool is composed of multiple indices that are calibrated for each Biotype[1]. There are three defined Biotypes in Colorado: Transition, Mountain, and Xeric/Plains. Based on the indices that were calibrated for each Biotype, the MMI tool takes the taxonomic identifications and individual counts of a sample to calculate an unitless score from 0-100. This

1 A Biotype is an aggregation of macrobenthos sites that have similar community composition. Environmental attributes (elevation, stream slope and ecoregion) common to sites within each of the Biotypes can be used to predict membership of a new site.

score is then compared to the statistically derived biological threshold for each of the Biotypes. If the MMI scores fall

below impairment thresholds, the waterbody is determined to be impaired

for the aquatic life use and listing decisions are made in accordance with the procedures outlined in section V(A), Assessment of Aquatic Life Use Based Standards.

2. Approved Data Collection Methodologies

Benthic macroinvertebrate data used for the assessment of the aquatic life use should be collected using either the kick-net or Hess sampling methodology. Alternative collection methods (i.e. Surber Sampler) may be used to assist in confirming assessment decisions, but should not be used solely to support listing decisions.

With initial versions of the MMI tool, only benthic macroinvertebrate data collected using the kick-net method were deemed appropriate for use with the tool for making attainment determinations. This requirement was established because the MMI tool was calibrated using only benthic macroinvertebrate data collected using the kick-net collection method. In 2014, a study was conducted to compare data collected using the Hess method to the kick-net method. The study demonstrated that the Hess method can produce similar MMI scores as the kick-net method if certain rules/modifications are followed. Therefore, the Commission determined that data collected using the Hess collection method can also be used with the MMI tool to calculate an MMI score for attainment/impairment determinations. A detailed description of these rules/modifications for using benthic macroinvertebrate data collected using the Hess method are included in Appendix D, Hess Method Sample Rules/Modifications.

The division recommends following the sampling protocols outlined in the Standard Operating Procedure found in Policy 10-1, Appendix B for determining the defined sampling season or index period for each biotype. The recommended standard index period for Biotypes 1 and 2 is late June through November 30th, and May 1 through November 30th for Biotype 3 due to its lower elevation.

3. Assessment of Biological (Benthic Macroinvertebrate) Data

Benthic macroinvertebrate data submitted to the Division for assessments should be provided by the data call deadline established each year as indicated in Table 1. The division reviews the submitted data to ensure the minimum data elements are included that are required for import into the Colorado Ecological Data Application System (EDAS). This data system is used to calculate MMI scores. If MMI scores are calculated by a third party, those parties are required to submit to the Division the following information in order for the division to confirm the MMI scores:

- An Excel spreadsheet with MMI scores and auxiliary metrics (HBI and SDI) by site and date*
- Master Bug Import File*

- Stations Predictors Import File*
- An Excel spreadsheet with results of the sub-sampling to a fixed 300-count*
- The query in Colorado EDAS called "Chris's Special Query to Find Benthic Data". Copy the contents of the query table into an Excel spreadsheet.

*Please note that the excel spreadsheet can be submitted with multiple tabs.

a) Biotype Assignment

Consistent with Policy 10-1, benthic macroinvertebrate sample locations determine the Biotype assignment. These assignments are based on the elevation, stream slope, and ecoregion. If uncertainty exists regarding the transitional boundaries between Biotypes, calculating a new MMI score with the macroinvertebrate data, but using the adjacent Biotype for the sample location may be used to help determine the attainment status for the segment. This additional analysis may be conducted under two circumstances:

i. At sites in Level IV Ecoregion 21c where the biotype assignment along a waterbody varies between Biotypes 1 (Transition) and 2 (Mountain) because the stream slope fluctuates above and below 0.04. This situation typically occurs when stream slopes are slightly greater than or less than 0.04 along the gradient of a waterbody resulting in varying site classifications or Biotypes.

ii. At sites that encompass the physical border between two different Level IV Ecoregions or elevation zone boundaries used in the Biotype classification. This results in a predicted site classification in one Biotype, but is narrowly adjacent to another Biotype. In such cases, sites may be represented by environmental or biological characteristics shared by more than one Biotype.

If the adjacent Biotypes are proposed to be used for the assessment, evidence must be presented to the division to show why the site is more suited for the adjacent Biotype.

b) Assessment Conclusions of Benthic Macroinvertebrate Data

If an MMI score falls in between the attainment and impairment threshold for a given Biotype, Policy 10-1 refers to this score as falling in the "gray zone". For scores that fall in the gray zone, auxiliary metrics are used to determine whether the waterbody is impaired or attaining. The two auxiliary metric scores evaluated for waterbodies that fall into the gray zone include the Hilsenhoff Biotic Index (tolerance) and the Shannon Diversity Index (diversity). Both auxiliary metric scores must pass in order for the station to be considered attaining (Table 2, Policy 10-1). If one or both auxiliary metrics fail, the station is considered impaired for the aquatic life use.

Listing decisions for the 303(d) and Monitoring and Evaluation Lists will be determined in the following manner:

- Multiple stations with MMI scores on the same segment, collected in the same calendar year, but have differing MMI scores (i.e. one passing MMI score; one failing MMI score) will be proposed for the M&E List.
- Multiple stations with MMI scores on the same segment, but collected in different calendar year sampling seasons, the most recent MMI score will be used in the listing decision.
- For high scoring waters (>56 for Biotype 1; >62 for Biotype 2; >51 for Biotype 3), a 22-point decline in the MMI score based on two successive measurements made more than 12 months apart, results in a conclusion of impairment. These segments will then be determined to be in attainment when subsequent MMI scores improve by a minimum of half of the original decline (≥11 points) assuming the scores are still greater than the attainment threshold for the Biotype.

Consistent with the listing methodology, a single, more recent and representative MMI score is sufficient to remove the segment from the 303(d) list.

The division will consider the representativeness of the benthic macroinvertebrate data before listing decisions are determined (Section III(B)(7)). Representative data require the collection and processing of samples following the previously described Approved Data Collection Methodologies by experienced professionals. Clear and convincing evidence is required to show impairment or attainment. One MMI score is considered clear and convincing evidence.

If a segment is previously listed, or proposed to be listed, for a pollutant causing an impairment of the aquatic life use, the segment will be listed for that pollutant as well as for impairment of the aquatic life use for benthic macroinvertebrates. If there is no apparent pollutant(s) causing the impairment, the impairment will be identified as "provisionally" listed. Once a segment is provisionally listed, the provisional listing process described in section V(G), Assessment Where the Cause is Unknown, will be followed.

c) Other Biological Data

Other biological data may be considered for the assessment of the aquatic life use. For other biological data to be used for assessment purposes, the division reviews the documented data sources, evaluates the validity and representativeness of the data, and considers the methodologies utilized in collecting the data. The representativeness of the data is discussed in detail in section III(B)(7).

If other biological data are determined to be valid and representative, they may be compared to the expected condition of a similar stream, river or adjacent reach. In general, the expected condition may be defined in two ways: 1) by actual conditions adjacent (upstream or downstream) to the affected reach or 2) by conditions in a comparable

stream or river located in a different watershed. When determining expected conditions, the division will review the following:

- level of disturbance
- location and description (upstream or downstream reach or within a separate watershed)
- historical conditions
- expected condition based on modeling or general expectations for highly managed systems
- other reasonable comparisons that can inform the assessment

Alternately, the expected condition may also be developed based upon biocriteria, modeling, or professional judgment. Any assessment of other biological data, except benthic macroinvertebrates, must describe the basis for defining the expected condition.

4. Policy Statement 98-1, Guidance for Implementation of Colorado's Narrative Sediment Standard

For the assessment of sediment in streams and rivers, the division follows procedures outlined in Policy Statement 98-1, Guidance for Implementation of Colorado's Narrative Sediment Standard (Policy 98-1). Listing decisions are made as discussed in section V(F)(4).

F. Assessment of Narrative Standards and Classified Uses

Impairment of narrative standards and classified uses may be supported by chemical data and/or information generated by biological and/or physical assessments. In instances where a determination of impairment is based solely upon biological and/or physical assessments, such assessments must provide clear and convincing evidence of nonattainment.

1. Aquatic Life Use

See Assessment of Biological and Physical Data Section V.E

2. Water Supply Use

For water supply uses, the division will consider chemical data, biological and/or physical assessments that provide clear and convincing evidence of non attainment. Such impairment may be demonstrated by chemical data documented at levels toxic to humans. The division will utilize commission Policy 96-2, *Human Health Based Water Quality Criteria and Standards*, in any determination of impairment based upon such information. Impairment decisions may also be supported by biological and physical data presenting overwhelming evidence of impairment due to color, taste and odor.

3. Narrative Free From Toxics Standard

In-situ bioassay, or other ambient toxicity test results which demonstrate statistically significant lethal or sub-lethal adverse effects and which are supported by biological information demonstrating adverse impacts to aquatic community health, composition, or productivity, in comparison to an appropriate

reference condition, will result in a decision of impairment. In general, interpretation of toxicity test results will conform to applicable portions of the *Implementation of the Narrative Standard for Toxicity In Discharge Permits Using Whole Effluent Toxicity (Wet) Testing* (https://www.colorado.gov/pacific/cdphe/water-quality-permitting-policies).

For lakes and reservoirs, impairment may be demonstrated where acute conditions (typically low DO levels) result in significant fish kills. Fish kills associated with accidental spills or isolated unauthorized discharges of toxics will not typically be considered a basis for listing.

4. Narrative Sediment Standards

Excessive deposition of sediment on the bottom of streams and rivers can cause harmful impacts to aquatic life such as benthic macroinvertebrates and fish, in addition to other beneficial uses. The impacts to aquatic life usually result from the loss of critical habitat for fish, aquatic invertebrates and algae. Regulation #31 includes a narrative standard that states that a waterbody should be "free from substances attributable to human-caused point source or nonpoint source discharge in amounts, concentrations, or combinations which can settle to form bottom deposits detrimental to the beneficial uses."

The division determines attainment of the statewide narrative standard by following protocols outlined in commission Policy 98-1, *Guidance for Implementation of Colorado's Narrative Sediment Standard Regulation #31, Section 31.11(1)(a)(i)*. For all state waters, the narrative standard is not in attainment when evidence demonstrates the following:

- The actual observed sedimentation condition for a specific waterbody is significantly different than the expected condition, and thus considered excess sediment
- The excess sediment is attributable to an anthropogenic source
- The excess sediment could be a detriment to a beneficial use

Policy 98-1 includes sediment thresholds that apply to rivers and streams in specific regions, as well as specific assessment methods to evaluate benthic macroinvertebrate assemblages and fish assemblages.

To evaluate the benthic macroinvertebrate assemblages, three components are examined:

- a census of the waterbody substrate and a resultant measure of the percent fines (%fines <2 mm),
- a Tolerance Indicator Value for sediment (TIV_{SED} score), and
- a review of available watershed information (watershed review).

A detailed explanation of how each component is evaluated is included in Policy 98-1. Sediment and macroinvertebrate data used to make attainment decisions must be collected within the same two week period during representative flow conditions. For a segment to be in nonattainment, a failing

TIVSED score, a failing % fines value and a watershed review are required. The watershed review must confirm the existence of anthropogenic sources of sediment and confirm that the sample site/watershed is not significantly different from the range of conditions used to establish the expected condition for the Sediment Region. Impairment decisions are not possible if only two of the three components are assessed. The TIVSED score and the % fines must be in attainment in order for the division to propose a delisting of a previously listed segment.

To evaluate fish assemblages, the percent fines (percent fines <8 mm) is measured from targeted fish spawning habitat for a given segment. If the percent fines is greater than 20 percent and the watershed review confirms that excess sediment is attributable to an anthropogenic source, the segment is considered impaired.

Only reliable and representative data will be considered in determining whether a segment is impaired. Data collected from a single representative location and time period is sufficient to make an attainment decision. Consequently, data from a single more recent representative sampling event can also result in a segment de-listing. If multiple data sets are collected from the same location within a five year period of record, the most recent representative data set is used for attainment decisions.

The extent of the impairment is determined as a part of the watershed review in the assessment process. The division considers watershed characteristics in determining whether to list only a portion of a segment consistent with methodology described in this guidance in section III.D.6.b. Determinations regarding impairment of beneficial uses other than aquatic life will be made in accordance with Section V of Policy 98-1.

G. Assessment Where the Cause is Unknown

The Federal CWA defines pollution as "the man made or man induced alteration of the chemical, physical, biological and radiological integrity of water," CWA §502(19). Pollution may result from the introduction of pollutants or from causative factors other than pollutants. Pollutants, as defined in the CWA at §502(6) include "dredged spoil, solid waste, incinerator residue, sewage, garbage, sewage sludge, munitions, chemical wastes, biological materials, radioactive materials (except those regulated under Atomic Energy Act of 1954, as amended), heat, wrecked or discarded equipment, rock, sand, cellar dirt, and industrial, municipal, and agricultural waste discharged into water." Notwithstanding the federal definition cited above, certain radiological constituents are also regulated under the state's Water Quality Control Act and are considered to be pollutants.

TMDL development is required in those instances where one or more pollutants are the cause of nonattainment. TMDLs are not required where the impairment is the result of pollution that is not a pollutant.

Waterbodies with classified uses that are impaired but where it remains unclear whether the cause of impairment is attributable to pollutants as opposed to pollution will be

provisionally listed on the Section 303(d) List. The list will include a notation identifying waterbodies that are provisionally listed.

For waterbodies that are provisionally listed on the Section 303(d) List, the division, in cooperation with other interested persons, will undertake water quality monitoring and/or other water quality studies and assessments to determine whether the cause of the impairment is a pollutant. There will be a general goal of making this determination within ten years of provisionally listing any waterbody. No TMDL will be developed for a provisionally listed waterbody unless and until it is determined that the cause of the impairment is one or more pollutants.

Provisional Listing Process

Once a segment is provisionally listed, the process to determine a cause will include the following steps:

1. Determine if the impairment is caused by a pollutant

   Once a segment is provisionally listed on the 303(d) List, the cause of impairment must be identified through additional data collection and investigation.

   If the impairment is caused by an identified pollutant, the segment will be placed on the 303(d) List as impaired for that pollutant (e.g., Cd, Fe(dis)) as well as for the applicable classified use without the provisional label. The division will proceed with development of a TMDL.

2. Determine if the impairment is caused by pollution

   If the evaluation demonstrates that the segment impairment is due to pollution, then the segment will be removed from the 303(d) List (as provisionally listed) and placed in the Integrated Report Category 4c (impairment is not caused by a pollutant) at the time of the next Section 303(d) List review cycle.

3. The cause of the segment impairment remains unknown

   If it cannot be determined that the cause is one or more pollutants or is not caused by pollution the segment will remain on the 303(d) List as provisionally impaired (e.g., aquatic life use, provisional). The cause of the impairment is to be determined within the next ten years.

A waterbody that is provisionally listed will not result in the prohibition of new or expanded discharges into the segment prior to the determination whether the impairment is caused by a pollutant.

To the extent it is suspected that a pollutant is the cause of the impairment, but the identity of the specific pollutant is not yet known, the waterbody segment will be provisionally listed. However, the fact that the waterbody is listed will not result in a prohibition of new or expanded discharges into the segment until the pollutant is identified.

H. Assessment Where the Source of the Pollutant is Natural

In cases where adequate monitoring and assessment indicate that natural conditions are the key factor of criteria exceedance(s), 303(d) listings will still be determined for impaired waterbody segments, as appropriate, without consideration of natural vs. anthropogenic causes. If natural conditions are triggering the exceedance(s), the decision is made by the commission through regulatory changes to the basin regulations in the triennial review process. Changes could involve the development of site specific standards or use removal through a use attainability analysis (UAA).

VI. PRIORITIZATION FOR TMDL DEVELOPMENT

The division must ensure that TMDLs are developed for all waterbodies and pollutants on the Section 303(d) List. Recognizing that all TMDLs cannot be completed at once and that certain risks may be greater than others, the CWA directs the division to prioritize the waters on the Section 303(d) List. The division will use the prioritized Section 303(d) List to focus resources to establish priority waters or watersheds and support the development of targeted TMDLs. Provisionally listed segments will not be prioritized for TMDL development.

A. Prioritization Objective

The objective of the prioritization on Section 303(d) list is to identify where the division and the public should focus their resources. The identification of high priority segments do not necessarily mean that the TMDLs will be developed before any lower priority segments. For some high priority segments, the development of a TMDL may be delayed due to the need for additional data collection or stakeholder outreach.

B. Assigning Priorities

Priorities defined on the 303(d) list are initially based on consideration of the severity of impairment to the use classifications for the segment. Use classifications are described in *Basic Standards and Methodologies for Surface Water,* Regulation # 31 (5 CCR 1002-8, sec. 31.13). Secondary factors can be used to modify the initial prioritization to an overall or final prioritization. Secondary factors may either elevate a waterbody into a higher priority group e.g., endangered or declining native species, public interest, administrative needs or reduce the priority ranking e.g., pace of stakeholder group development, CERCLA cleanup action in progress.

1. Severity of Water Quality Impairment

   High Priority: Non-supporting for water supply standards based on Safe Drinking Water Act primary drinking water standards (NO_2, NO_3, As) Aquatic Life Class 1 cold or warm, or Recreation Use Class E. Listings based on high levels of mercury in fish tissue.
   Medium Priority: Non-supporting for Aquatic Life Class 2 cold or warm, or Agriculture.
   Low Priority: Non-supporting for other water supply standards or Recreation Use Class P, U or N, or non-supporting for underlying standard where a temporary modification based specifically upon significant uncertainty as to the appropriate underlying standard has been adopted and the commission has

determined that there is an appropriate plan in place to resolve the uncertainty.

2. Secondary Considerations
   - Division action can support a local, regional or federal stakeholder group that is ready to move on to the next step of TMDL development or there is substantial public interest and support.
   - The waterbody is vulnerable or fragile as an aquatic habitat or there are aquatic species of special concern present.
   - The waterbody is of particular importance for recreational, economic and aesthetic uses.
   - The division can realize efficiency savings (e.g., synchronizing permits, linking segments within a watershed, availability of water quality data).
   - There are immediate programmatic needs such as waste load allocations for permits that are due to expire, ~~or~~ for new or expanding discharges, or to facilitate 319 project developments in priority watersheds.
   - There is a court ordered cleanup or CERCLA action in progress, which will change the contribution of pollutants (this consideration could reduce priority ranking).

3. Identification of Targeted TMDLs and Priority Waters or Watersheds

   It is the division's intent that TMDLs that are designated as targeted TMDLs will be completed prior to the next listing cycle, or within two years of finalizing the 303(d) List by the commission. Targeted TMDLs will most likely be included in priority waters or watersheds that are designated through a prioritization framework using multiple factors including, but not limited to, the high priority waters for TMDL development as defined in the 303(d) List. However, not all high priority listings are suitable for TMDL development within a two year window. For example, adequate data to support TMDL development is not available for all high priority listings. Conversely, waters designated as medium or low priority may be amenable to TMDL development within the next two years and therefore may be targeted for TMDL development at this point.

   TMDL development is subject to a variety of factors that are both within and beyond the division's control. These may include availability of adequate data, local or broader political concerns, new information that affects the listing decision, coordination with remedial programs such as CERCLA or Superfund, or availability of division resources. Designation of a TMDL as targeted should be considered for planning purposes, but should not be treated as a definitive division workplan commitment. The division's TMDL program workplan is updated quarterly and is available on the division website (https://www.colorado.gov/pacific/cdphe/total-maximum-daily-loads-tmdls).

VII. POLICY DETERMINATIONS

A. 2016 Listing Methodology

In March 2015, the commission made the following policy decisions as they relate to assessment methods used to determine impaired waters in Colorado.

Category 4c: The commission determined that Category 4c could be a useful tool for identifying segments that are impaired solely due to pollution and not attributable to pollutants. Segments placed in Category 4c will generally not require TMDL development but may require pollution reductions plans to address the impairment. The commission directed the division to work with interested stakeholders to develop guidance for the determination of Category 4c segments as well as guidance for future plans for restoration on these segments. Work on these issues will be presented in the 2018 303(d) Listing Methodology proposal in March 2017.

Assessment of the iron, manganese and sulfate water supply standards: The standards for iron, manganese and sulfate in the Basic Standards, Regulation #31 (circa 2015) are assessed using the least stringent standard: table value standards or the existing quality as of the year 2000. The commission clarified that when no data is available for the year 2000, current water quality should be presumed to be representative of existing water quality as of 2000 if a watershed review concludes that no additional sources of manganese, iron or sulfate have been introduced since 2000. Where there have been no changes in source contributions the segment is not impaired. Where there have been changes in source contributions, segments will be placed on the M&E List and considered high priority for site-specific standards development.

The commission also decided that in instances where the water quality representative of the year 2000 was used as the standard for listing decisions, the assessment of current conditions is conducted using the statistics used to evaluate the table value standards, for manganese and iron it is the 85th percentile and for sulfate it is the 50th percentile.

The commission established that when setting a standard for iron, manganese and sulfate using data representative of existing quality in 2000, 10 data points is typically needed to have certainty in the standard being used.

Assessment of the nitrate, nitrite and arsenic water supply standards: The commission adopted language regarding the assessment of arsenic and nitrate/nitrite based on the current standards. The commission intends to revisit this in the event that the standards are modified in the future.

B. 2018 Listing Methodology

In March 2017, the commission made the following policy decisions as they relate to assessment methods used to determine impaired waters in Colorado.

Temperature warming event – Rivers and Streams:
The commission defined the "warming event" with a method consistent with Regulation #31. The methodology was derived using literature used in the newly updated 2016 temperature database and specifies an allowable cumulative impact during this once in three years on average. The literature included information about the effects of warming events on cold water species in summer and winter, and on warm water species in summer. Although the literature lacked information about the effects of winter warming events on warm-water fish, the technical advisory committee convened by the division recommended that the method extrapolated from summer data. Future revisions may be necessary if additional information about the thermal needs of warm-water fish in winter becomes available.

The commission adopted a method which relies on the concept of 'degree-days' which integrates both the magnitude of temperatures over the standard, as well as the duration, in days, experienced by the aquatic community. A 'warming event' may include days within a season that are not consecutive but the event may not include days from more than one season in the period of record. Temperature excursions (air, low flow and shoulder season) are evaluated after the warming event is considered. If temperatures exceed the number of 'degree-days' specified, and the dates that exceed temperature standards do not have applicable excursions, the division recommends that the segment be placed on the 303(d) List as impaired for temperature when the number of allowable warming events are exceeded.

Assessment of iron, manganese and sulfate water supply standards:
Regulation 31.11(6) states:

"For all surface waters with a "water supply" classification that are not in actual use as a water supply, the water supply table value criteria for sulfate, iron and manganese set forth in Tables II and III may be applied as numerical standards only if the Commission determines as the result of a site-specific rulemaking hearing that such standards are necessary and appropriate in accordance with section 31.7."

The commission concluded that secondary water supply standards for iron, manganese and sulfate should only apply to segments where there is an 'actual' water supply use. The commission also concluded that the determination of actual water supply uses can occur in the basin standards rulemaking hearing and the Regulation 93 rulemaking processes. The commission recognized this could require significant division resources to conduct this evaluation on each segment during the assessment process. Therefore the commission asserts the division could presume actual water supply use where "WS" is applied to iron, manganese and sulfate in the basin standards tables. The commission also provided an opportunity for third parties who disagree with this presumption to provide credible evidence as a part of the 303(d) List assessment process. Where evidence indicates no 'actual' use, the division will not propose 303(d) water supply impairment listings for iron, manganese and sulfate.

The commission recommends the 'actual' water supply use evaluation would be most efficiently and effectively conducted with other use classification designations during the standards rulemaking hearing process. However, until 'actual' water supply uses can be considered independent of future potential uses, the commission supports this examination during the Regulation No. 93 rulemaking process.

C. 2020 Listing Methodology

In March 2019, the commission made the following policy decisions as they relate to assessment methods used to determine impaired waters in Colorado.

Assessment of iron, manganese and sulfate water supply standards:

The previous assessment process for iron, manganese and sulfate secondary water supply standards had the potential to result in a cycle of listing and delisting segments, when existing quality as of the year 2000 was used as the assessment value. This is because the process compared the 85th percentile of historic data (existing quality as of the year

2000) to the 85th percentile of current data. Even with waters of like quality, it would be expected that the different data sets would not have the exact same 85th percentile. This is due to the expected variation within water quality data. Therefore, impairment conclusions could be reached without an actual water quality change. The commission addressed this concern with the application of Appendix B, *Assessing Attainment of Ambient-Based Water Quality Standards*, to secondary water supply standards. This is only applicable when existing quality as of the year 2000 is the least stringent option and is used as the assessment value for determining attainment.

Process for Carrying Over Existing Attainment Conclusions during Waterbody Resegmentation and Portioning

The commission clarified how to apply use attainment status when parent waterbodies and portions are split into multiple child segments and portions. The general approach is to retain the use attainment status of the parent to the child segments or portions. Therefore, if a parent segment is considered in attainment of uses (category 1), the newly created child segments are then also considered in attainment of uses (category 1). Similarly, if a parent segment is considered impaired for a designated use (category 5), the child segments are also considered impaired for the same use (category 5). Newly created segments and portions are reassessed during the regularly scheduled basin cycle.

Much consideration was given to situations where new child segments and portions have insufficient data to support the previous attainment conclusions. The commission considered many aspects of this topic including; impacts to stakeholders, previous commission decisions, clearly defined delisting requirements, access to historic data used in previous listing decisions, and division and stakeholder resources. Consequently, when data within the current period of record is not available for a newly created segment or portion, other relevant information may be considered to evaluate whether or not the retained attainment status is appropriate. This information could include but is not limited to landscape analysis (i.e. hydrology, vegetation, soils, and elevation), underlying geology or an investigation of activities in the watershed. For example, if a parent segment is considered impaired for a specific parameter, however landscape analysis demonstrates that the cause or the source of the impairment is limited to only one of the child segments, the case could be made that the impairment status is only inherited by the child segment located in close proximity to the source.

The commission considered that when the resegmentation process results in changes to the same standards for which the segment is listed, carrying the impairment over to the new child segment may not be appropriate. In these cases, the previous presumption that the segment was of ‘like water quality’ may no longer apply. Therefore the previous commission impairment decision should be reconsidered prior to it being applied to the child segment. If there is no current data to confirm that the existing impairment is appropriate, a reasonable historic assessment will be conducted to determine if any of the data used in the previous impairment determination(s) were collected in the newly created child segment. If none of the data used in the impairment decision(s) were collected in the new child segment, that segment will be moved from the 303(d) List to the Monitoring and Evaluation list for additional data collection.

The commission did not apply the approach to assess historic data to the splitting of attainment portions because portions all exist within the same waterbody segment and are considered to be of “like water quality”.”

D. 2024 Listing Methodology

In March 2022, the commission made the following policy decisions as they relate to assessment methods used to determine impaired waters in Colorado.

Data Submitted After the Data Call

The Regulation #93 hearing process has had a longstanding issue with data being submitted after the data call deadline. This causes several challenges including a reduction in the amount of time that stakeholders have to evaluate the data, significant duplication of efforts by the division, the submissions often do not include all available representative data and it may have impacts on the assessment period of record. The commission adopted the following changes to address these issues:

1. Adopted new language strongly encouraging data to be provided by the data call deadline.
2. Adopted Appendix E, Criteria and Guidelines for Data Submitted After the Data Call, and provided guidance that new data submitted during the hearing should be accompanied by documentation of how these criteria are met.
3. Adopted language stating that the period of record end date will generally be extended to include data collected after the period of record and that data available from sources such as the Water Quality Portal should be provided for the time period that the POR was extended.

The Commission recognizes that the Criteria and Guidelines for Data Submitted After the Data Call (Appendix E) were developed and incorporated into the Listing Methodology after the 2021 Data Call for Regulations #32 and #36, Arkansas and Rio Grande Basins, was closed. While entities in these basins are still encouraged to follow the criteria as it could apply to them in the 2023 Regulation #93 hearing, the Commission intends these changes to primarily apply to data calls starting in 2022, following the adoption of the revised 2024 Listing Methodology.

E.coli Assessment Methodology Revisions

To address concerns about the complexity of the assessment methodology for *E. coli* and the stringent requirements for listing and delisting, the commission made three changes:

1. Bias Removal. The commission replaced the previous seven-day bias removal process with a one-day process. This change will make field sampling schedules more flexible without compromising the representativeness of the data. To avoid unduly biasing assessment results to one short period of time, the commission also precluded basing 303(d) listing or delisting decisions exclusively on samples collected from the same seven-day period.
2. Two-Month Periods. The commission replaced the previous methodology's rolling, 61-day assessment periods with static, two-month periods. The static, two-month approach is easier to implement, explain, and understand but neither significantly affects listing decisions nor compromises water quality protection.
3. Delisting. The commission adopted several changes to the delisting process for E. coli in order to reduce the unnecessary stringency of certain requirements that

precluded delisting despite strong evidence of attainment. First, the commission expanded the window within which samples used to delist must be collected from the same months as the samples used to list to the same season. Second, the commission reduced the minimum number of samples required to remove a portion or segment from the M&E List from five within a two-month period to two within a two-month period. Third, the commission required that (1) at least one two-month period comprising the minimum number of samples required to delist (two for M&E removals, and five for 303(d) removals) from the most recent two years of available data indicate attainment, and (2) all two-month periods for which data are available from those two years indicate attainment.

The commission reasoned that, taken together, these three changes would better capture commonly observed patterns in *E. coli* concentrations while making it easier to achieve the sample quantity required for listing determinations. They would also improve consistency among assessment determinations and simplify the overall assessment process. As a result, these changes will enhance stakeholder and public understanding without compromising listing accuracy or water quality protection.

E. 2026 Listing Methodology

In March of 2024, the commission made the following policy decisions as they relate to assessment methods used to determine impaired waters in Colorado.

Iron, Manganese, and Sulfate Data Library Update Cycle:
The commission clarified that the iron, manganese and sulfate data library will be updated on a five-year cycle to coincide with the statewide (Regulation #31) assessment cycle. Data may be provided for inclusion in the data library during the statewide assessment cycle data call process.

Lake Nutrient Assessment Clarifications:
The commission clarified the language regarding the assessment of nutrients (total nitrogen, total phosphorus, and chlorophyll *a*) for lakes and reservoirs. The commission adopted clarifications, including a reorganization of Section V.D.6 and the addition of flow charts to illustrate the assessment process. Table 5 “Sample Size Requirements for the Assessment of Chronic Standards” was adjusted to clarify the sample size requirements for nutrient assessments in lakes and reservoirs. Lakes were separated from streams, and additional rows were created for lakes and reservoirs classified as Direct Use Water Supply (DUWS). Lastly, an error in the seasonal boundary dates was corrected.

Dissolved Oxygen Clarifications and Spawning Table Updates:
The Commission clarified the dissolved oxygen spawning season language to improve readability. Specifically, the dissolved oxygen assessment instructions were separated into more clear assessment steps.

The commission updated the fish spawning table in Appendix A with more recent and precise spawning information provided by Colorado Parks and Wildlife. This replaced information from 2008. This modification also consolidated spawning, egg incubation and fry emergence phases into a single spawning season per waterbody to simplify the D.O. assessment process.

Temperature warming event – Rivers and Streams:
The commission clarified the language for temperature warming event assessment methods consistent with Regulation #31. The modifications adopted by the commission do not represent changes to the warming event assessment process. Instead, these modifications clarify the approach that has been in use since 2018. Additionally, new assessment related examples were included. Specifically, the adopted changes clarify:

- The warming event method relies on the concept of 'degree-days,' which integrates both the magnitude of temperatures over the standard, as well as the duration, in days, experienced by the aquatic community.
- The average recurrence frequency of these warming events is limited to once every 3 years on average.
- A warming event may not include days from more than one season or span multiple years in the period of record.
- Within the context of the warming event, stream temperatures are allowed to exceed standards for a specific number of degree days within a single season.
- The warming event is the application of the 1 in 3-year average exceedance frequency. WATs and DMs above the standard but within the warming event threshold still constitute an exceedance of the standard and count toward the 1 in 3-year exceedance frequency.

Ambient-Based Assessment Delistings:
For waterbodies determined to be impaired using the lower confidence limit approach within Appendix B, the commission determined that the upper confidence limit was an appropriate approach to determine if a water body is in attainment of the ambient-based standard or the existing quality as of the year 2000 assessment values for iron, manganese, and sulfate. Using the upper confidence limit ensures the attainment (delisting) determination is completed with the same degree of confidence that was used in the listing action. Because this is a delisting process, a minimum of 10 samples are required. The new tables, Supporting Tables 3 and 4, in Appendix B both contain percentile values for samples sizes that range from 10 to 100.

INDEX

APPENDIX A

Appendix A - Spawning Table

The division is defining the spawning season as a combination of the four phases, adult spawning, egg incubation, intra-gravel sac fry, and fry emergence. The table below consolidates each of these phases into spawning start and end dates for the Cold Stream Tier I and Tier II species present in the specific waterbodies. The species included in this table are Brown Trout, Rainbow Trout, various species of Cutthroat Trout and Mountain Whitefish. This information was provided by Colorado Parks and Wildlife.

Watershed	River Description	Species	Spawning Start	Spawning End
Arkansas River	Arkansas River -S Fork Ark to Badger Creek	brown trout	Oct-16	May-15
Blue River	Blue River- Straight Cr to Slate Cr	brown trout / rainbow trout	Oct-16	Jun-30
Cache La Poudre River	Cache la Poudre River- Little S. Fk to N. Fk	brown trout / rainbow trout	Oct-16	Jul-15
Cache La Poudre River	S Fk Poudre	cutthroat trout - greenback	Jun-15	Sep-30
Culebra Creek	Torcido Creek	cutthroat trout - Rio Grande	May-16	Aug-31
Eagle River	Brush Creek	brown trout / rainbow trout	Oct-16	Jun-30
Eagle River	Eagle River #1 - 3	brown trout / rainbow trout	Oct-16	May-31
Eagle River	Two Elk Creek	brown trout / rainbow trout	Oct-16	Jun-30
Fountain Creek	Bear Creek	cutthroat trout - greenback	May-16	Aug-31
Fryingpan River	Fryingpan River- Rocky Ford Cr to Roaring Fk	brown trout / rainbow trout	Oct-16	Jun-30
Gunnison River	Gunnison River - Gunnison Gorge	brown trout / rainbow trout	Nov-1	Aug-15
Gunnison River	Gunnison River- Crystal Dam to N. Fk Gunnison	brown trout / rainbow trout	Oct-16	Jun-30
Gunnison River	Gunnison River East Portal	brown trout / rainbow trout	Oct-16	Aug-31
Gunnison River	Kelso Creek (Uncompahgre Plateau)	cutthroat trout	May-1	Aug-31
Gunnison River	Taylor River- Spring Cr to East River	brown trout	Oct-16	May-31

APPENDIX A

Watershed	River Description	Species	Spawning Start	Spawning End
Lower Colorado	Bull Creek headwaters to Forest Service boundary	cutthroat trout-Colorado River	May-1	Jun-15
Lower Colorado	Carr Creek headwaters to confluence with Roan Creek	cutthroat trout-Colorado River	May-1	Jun-15
Lower Colorado	East Fork of Parachute Creek headwaters to falls	brook trout	Sep-16	Oct-31
Lower Colorado	Roan Creek	cutthroat trout	May-16	Aug-31
Lower Colorado	Roan Creek headwaters to confluence with Carr Creek	brook trout	Sep-16	Oct-31
Lower Colorado	Roan Creek headwaters to confluence with Carr Creek	cutthroat trout-Colorado River	May-1	Jun-15
Lower Colorado	West Fork of Parachute Creek headwaters to confluence with Middle Fork/Parachute creeks	brown trout	Oct-1	Nov-1
Lower Colorado River	Canyon Creek	mountain whitefish	Oct-16	Apr-15
Rio de Los Pinos	Osier Creek	cutthroat trout - Rio Grande	May-16	Aug-1
Rio Grande River	Rio Grande River- Willow Cr to S. Fk Rio Grande	brown trout / rainbow trout	Oct-1	Jul-15
Rio Grande River	South Fork of the Rio Grande-Park Cr to Beaver Cr	brown trout	Oct-1	Jun-30
Rio Grande/Mineral	Rio Grande	brown trout / rainbow trout	Oct-1	Aug-15
Roaring Fork	Cunningham Creek	brook trout / cutthroat trout-Green Lineage / cutthroat trout	Jun-1	Oct-31
Roaring Fork	Fryingpan River blw Ruedi	brown trout / rainbow trout	Oct-16	Jun-30
San Juan River	Navajo River	cutthroat trout	May-16	Sep-15
San Luis Creek	Clover Creek	brook trout	Sep-1	Apr-15
South Platte River	Middle Fork South Platte- Trout Cr to So Fk of So Platte	brown trout	Oct-1	May-31
South Platte River	South Platte- Cheeseman Dam to Horse Cr	brown trout / rainbow trout	Oct-16	Jun-30
South Platte River	St. Vrain River- S. Fork to Lefthand Cr	brown trout	Oct-16	Apr-15

APPENDIX A

Watershed	River Description	Species	Spawning Start	Spawning End
Upper Colorado	Canyon Creek	brown trout / rainbow trout / mountain whitefish	Oct-1	Jun-30
Upper Colorado	Colorado River- Fraser to Williams Fk	brown trout / rainbow trout	Oct-16	Jun-30
Upper Colorado	Grizzly Creek	brown trout / rainbow trout / mountain whitefish	Oct-16	Jun-30
Upper Colorado	Mitchell Creek	brown trout / rainbow trout	Oct-16	Jun-30
Upper Colorado	No Name Creek	brown trout / rainbow trout / mountain whitefish	Oct-16	Jun-30
Upper Colorado and Lower Colorado	Colorado River #3 - 6	brown trout / rainbow trout / mountain whitefish	Oct-16	Jun-30
Upper Colorado River	Grizzly Creek	mountain whitefish	Oct-16	Apr-15
Upper Yampa	Bear River	mountain whitefish	Oct-1	Apr-30
Upper Yampa	Fish Creek	mountain whitefish	Oct-16	Apr-15
Upper Yampa	Mad Creek	mountain whitefish	Oct-16	Apr-15
Upper Yampa	Yampa River #5	mountain whitefish	Oct-1	May-31
White River	North Elk Creek	cutthroat trout	May-16	Sep-15
White River	White River	mountain whitefish	Oct-16	Apr-30
Yampa River	Yampa River #4 - Elk River to Catamount Dam	brown trout / rainbow trout	Oct-1	Jul-15
Yampa River	Yampa River #5 - Catamount to Stagecoach Reservoir	brown trout / rainbow trout	Oct-1	Jul-15

Appendix B - Assessing Attainment of Ambient-Based Water Quality Standards in Colorado

Ambient-based water quality standards have been adopted in Colorado in limited circumstances where the table value standard cannot be met as a result of either natural conditions or irreversible, man-induced conditions. Each ambient-based standard is a site-specific characterization of *existing quality*[2] derived from "available representative data". Once an ambient-based standard has been adopted, attainment is assessed using recent, representative data.

The mechanics of setting the ambient-based standard and assessing its attainment are the same, but the characterizations are carried out with different, possibly overlapping, data sets. For dissolved metals, for example, the chronic standard is set equal to the 85th percentile of the available, representative concentration data, and the acute standard is set equal to the 95th percentile[3]. When existing quality is assessed, the 85th percentile of the available, representative concentration data in a subsequent data collection is compared to the chronic standard, and the 95th percentile is compared to the acute standard.

Assessment determines if water quality continues to meet the level of ambient quality originally characterized by the standard. In the current assessment methodology, the same quality is maintained (i.e., the standard is attained) if the assessed value does not exceed the standard. If the assessed value exceeds the standard, water quality is considered to be impaired. McBride (2005) calls this a "face value" test because it does not include consideration of sampling error.

Current assessment methodology for ambient-based standards has proven problematic. Successive assessments may yield opposite conclusions about the maintenance of existing quality. Changed assessment conclusions can have significant practical ramifications when

[2] 31.5(20) "EXISTING QUALITY" means the 85th percentile of the data for total ammonia, nitrate, and the dissolved metals, the 50th percentile for total recoverable metals, the 15th percentile for dissolved oxygen, the geometric mean for E. coli, and the range between the 15th and 85th percentiles for pH. For temperature, for the purposes of implementing the acute and chronic standard, "existing quality" is the seasonal maximum DM and WAT and which allows one warming event with a 3-year average exceedance frequency. For data records less than or equal to 3 years, existing quality is equal to the maximum WAT and DM. For data records with 4-6 years, one warming event above the standard is permitted.

[3] Concentrations corresponding to the specific percentiles are estimated from the available data using the PERCENTILE function in EXCEL. In most cases, these concentrations are interpolated values.

they cause water bodies with ambient-based standards to move on and off of the 303(d) list, as has happened in a number of cases (Table 8).

Table 8. Historical changes to listings based on ambient-based standards for selected stream segments.

Assessed	COARLA01a[4]	COARLA01b	COARLA01c	COARMA04a	COSPBD01
1998	List: Se, SO4, Fe	List: Se, SO4, Fe	List: Se, SO4, Fe	List: Se	
2002	Delist: Se, SO4, Fe	Delist: Se, SO4			
2004	List: Se	List: Se; Delist: Fe	List: Se		
2006	List: Fe				List: Se
2007				Set ambient std	Set ambient std
2008	List: SO4, Delist: Fe			Delist: Se	
2010					Set seasonal std
2012				List: Se	Delist one season

In retrospect, it should not be surprising that successive assessments of ambient-based standards could yield different conclusions even in the absence of any water quality change. Successive assessments based on the same percentile (e.g., 85th) are affected by normal variability in the available concentration data. Seasonal patterns, stochastic variation, and sampling or analytical error all contribute to that variability. Consequently, we might expect about half of the assessed values to be larger than the standard and half smaller.

When an assessment shows that an ambient-based standard is exceeded by even a small amount, the water body may be placed on the 303(d) list. Assuming no trend in ambient concentrations, it is equally likely in the next assessment cycle that the assessed value will fall below the standard. When the assessed value for a listed water body falls below the standard, the water body is removed from the 303(d) list. Thus, the examples in Table 8 of water bodies going on and off of the 303(d) list is consistent with statistical expectations for the current assessment methodology.

Having water bodies move on and off the 303(d) list creates two problems. The first problem is that it takes time and effort to develop or revise the 303(d) list. The second problem is that listing has practical ramifications for dischargers. Both problems can be addressed by adding an explicit level of confidence to assessments of ambient-based standards. The addition of a defined level of confidence would not affect the underlying definitions of existing quality or

[4] COARLA01 was split into segments 1a, 1b, and 1c in 2002.

ambient-based standards, but would establish the reliability of conclusions drawn from assessments.

Statistical Approach

Increasing the reliability of conclusions drawn from assessments of ambient-based standards is based on the statistical concept of the ***confidence interval***. The confidence interval is often viewed as the region around an estimate (i.e., the assessed concentration) within which the true concentration (i.e., the standard) is thought to be located[5]. If the confidence interval of the assessed concentration (e.g., 85th percentile) does not include the standard, then the assessed concentration is significantly different from the standard.

The width of the confidence interval, and thus the range of concentrations it spans, is determined in part by the desired level of confidence. When the level of confidence is set to 95%, for example, it means there is only a 5% probability (a 1-in-20 chance) of mistakenly concluding that the assessed concentration differs from the standard (i.e., a Type 1 error). Setting the risk of a mistake to 5% (a 1-in-20 chance) would improve the reliability of future assessments compared to the current approach.

A level of confidence other than 95% could be used, but there are tradeoffs. A higher level of confidence, such as 99%, has the advantage of reducing the risk of Type 1 errors to 1% instead of 5%, but it would also result in broader confidence intervals. Having broader confidence intervals makes it less likely that an exceedance will be identified because it is more likely that the standard will fall within the interval. Conversely, a lower level of confidence, such as 90%, yields a narrower confidence interval, but an increased risk (10% probability of a Type 1 error, instead of a 5% probability) of claiming that a segment is impaired when it is not.

For most assessments, regulators are interested only in situations where the assessed concentration is significantly larger than the standard. Thus, the null hypothesis can be defined to assume that the assessed concentration is less than or equal to the standard (i.e., H_0: Assessed concentration $\leq$ Standard); the test is one-sided. Rejection of the null hypothesis for this one-sided test means that the assessed concentration is significantly larger than the standard. In this one-sided case, a 5% probability of a Type 1 error defines the risk of claiming that a water body is impaired when it is not.

Selection of a specific statistical approach is affected by the number of assessments that the Division must undertake on a regular basis. If assessments were required only occasionally for one constituent at one site, there are a number of parametric and non-parametric tests to choose from. However, the Division must contend with about 170 ambient-based standards that have been adopted in 100 water bodies across the state. Running tests separately for

[5] McBride (2005; p. 58) explains why this simplistic view is not strictly correct. Nevertheless, it is useful for communicating the approach.

each constituent in each water body during each assessment cycle would not be practical for the Division, but may be appropriate and acceptable for parties that may have a narrower focus.

One way to keep the large workload manageable is to define the confidence interval for the assessed value in terms of percentiles rather than concentrations. Defining the confidence interval of a percentile is inherently non-parametric and well-suited to typical assessment data sets for which the distribution usually is not known in advance. Furthermore, it would be difficult, especially with small sample sizes, to validate a distribution. If a distribution cannot be assumed or validated, application of parametric methods becomes questionable.

Confidence intervals for percentiles are a function only of sample size. Tables can be developed to define confidence intervals that would be applicable to any constituent at any site for which the assessed data set consisted of the same number of measured concentrations. Thus, an assessment for zinc in one watershed and one for copper in another watershed would use the same confidence interval for percentiles as long as both sites had the same number of observations.

Assessment with percentiles is best understood with some graphical examples, and a good place to begin is with the current assessment methodology. The 85th percentile is featured in this example because it is the most common among ambient-based standards, but the concept applies equally well to the other percentiles (95th or 50th). The current methodology locates the concentration that corresponds to the 85th percentile[6] of the assessed data set and compares it to the standard. When the assessed value (85th percentile) is larger than the standard, current assessment methodology registers an exceedance (Figure 1). The magnitude of the exceedance, in terms of concentration, may be large or small, but the outcome is the same. However, the current methodology does not specify the reliability of the conclusion.

[6] The PERCENTILE function in EXCEL is used to determine the concentration that corresponds to the 85th percentile of all concentrations in the assessed data set. High concentrations correspond to high percentiles; thus, the 85th percentile is a high concentration within the assessed data set. For reasons to be explained later, the Excel function has shortcomings that should be considered in future assessments.

APPENDIX B

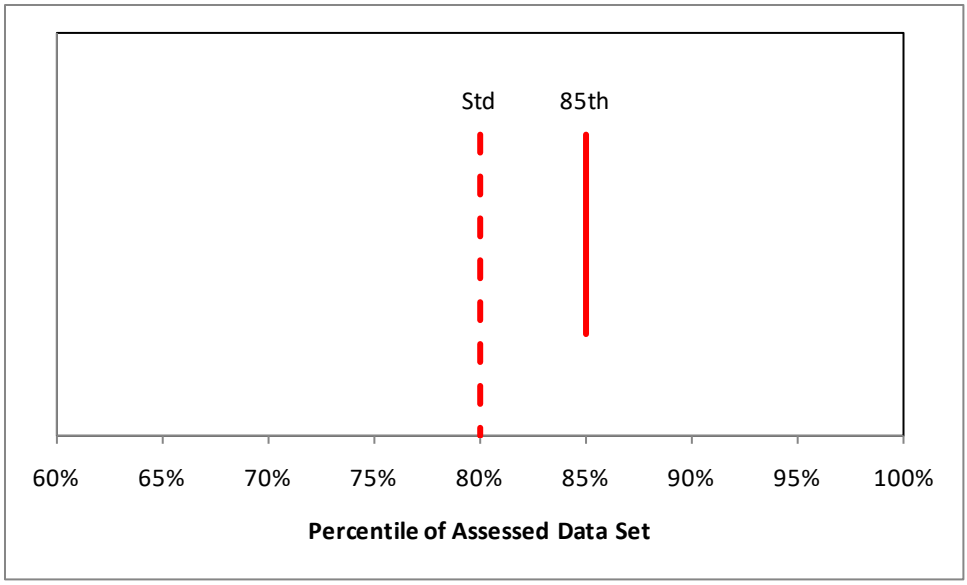


Figure 1. Graphical representation of a scenario where the 85th percentile of the assessed data (solid red line) exceeds the standard (dashed red line). For convenience in presenting the example, the concentration of the standard is represented as a percentile (80th) of the assessed data set. Current assessment methodology would interpret the result as an exceedance of the standard.

Reliability can be specified by defining a confidence interval around the 85th percentile. For example, when a confidence level of 95% is specified, the confidence interval constructed around the estimate (i.e., the 85th percentile) has a 95% probability of containing the true value (i.e., the standard). Making the right call 95% of the time is a very reliable basis for decision-making.

Building on the scenario used for Figure 1, a one-sided confidence interval is constructed for the 85th percentile of the assessed data. A statistical test is formalized with a null hypothesis (H_0) stating that the 85th percentile of the assessed data set is less than or equal to the standard. The test is one-sided to determine if the assessed value (85th percentile) exceeds the standard because assessment is focused on exceedances. The null hypothesis is rejected when the LCL exceeds the standard.

In this example, the standard falls within the confidence interval for the 85th percentile (Figure 2). Thus, the concentration corresponding to the 85th percentile of the assessed data set is not larger than the standard. The null hypothesis is not rejected, and there would be no justification for listing the water body.

APPENDIX B

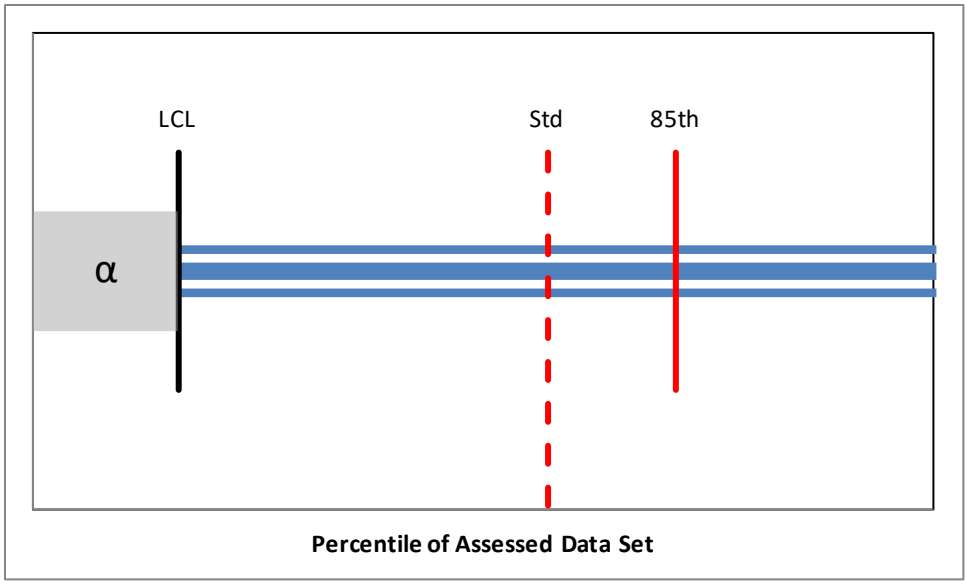


Figure 2. One-tailed confidence interval with the lower (LCL) confidence limit added to the scenario shown in Figure 1. Although the 85th percentile of the assessed data exceeds the standard, the difference is not statistically significant because the LCL of the 85th percentile does not exceed the standard. The critical region (gray region marked with alpha) extends to the left (lower percentiles) of the LCL. The confidence level for the interval is 1-α.

The scenario in Figure 2 is now changed so that the assessed value is significantly larger than the standard (Figure 3). The standard now corresponds to a low percentile of the assessed data distribution. With this scenario, the lower confidence limit (LCL) of the assessed value (85th percentile) exceeds the standard, and the null hypothesis is rejected. The assessed value is significantly greater than the standard, and the outcome would support a listing decision.

APPENDIX B

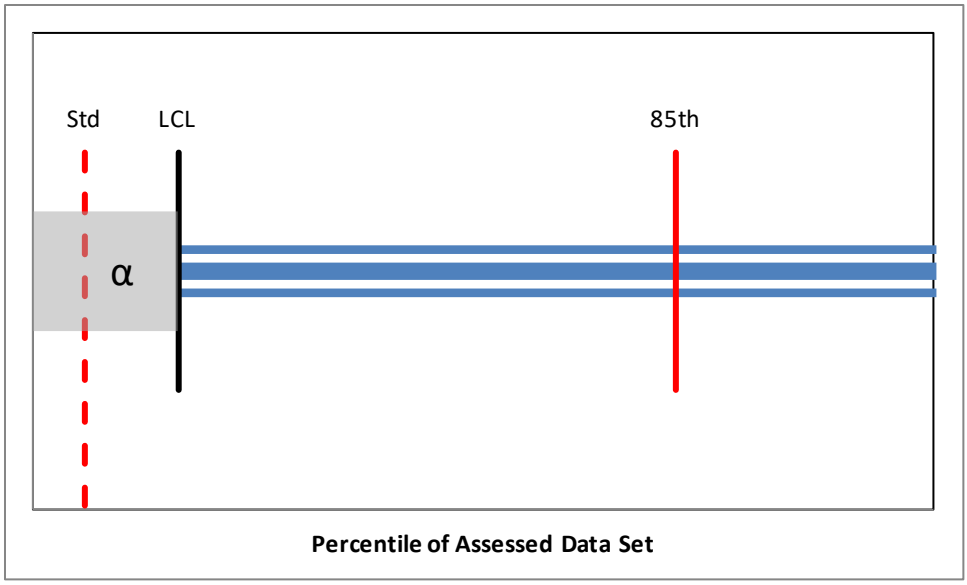


Figure 3. The scenario shown in Figure 2 is modified so that ambient concentrations are high relative to the standard. The LCL is greater than the standard, meaning that the 85th percentile of the assessed data is significantly greater than the standard.

Statistical Implementation

Defining the LCL is the key to a defensible statistical assessment of Colorado's ambient-based standards. There are a number of possible approaches (see Helsel and Hirsch 2002), but there are compelling reasons for the Division to focus on non-parametric methods, as explained below. A brief overview of common approaches provides a useful introduction to the concepts before tailoring an approach to our needs.

Overview of Methods

Variance and sample size are required for locating the LCL because the confidence interval is a statement about uncertainty. Since most environmental data sets are not normally distributed (or are too small to test for normality) a non-parametric test is preferred in most cases. A non-parametric method for locating the LCL makes no assumptions about the underlying statistical distribution of the data. Non-parametric methods for defining confidence intervals rely on the binomial distribution for defining variance. Exact and approximate non-parametric methods are available.

The Clopper-Pearson equations are used to determine exact confidence intervals for percentiles (Equation 1), but the computation is tedious, especially for large sample size. Less well-known, and much less tedious, is a direct calculation (Equation 2) using the F distribution (Leemis and Trivedi 1996), which can be evaluated in EXCEL with the FINV function.

APPENDIX B

$$\sum_{k=y}^{n} \binom{n}{k} p_L^k (1-p_L)^{n-k} = \alpha$$

Equation 1. Clopper-Pearson equation for the lower confidence interval. Probabilities are evaluated for successive proportions (p_L) until the sum reaches the desired exceedance level (α). Each proportion is calculated from the number of successes (k) out of the number of trials (n).

$$p_L = \frac{1}{1 + \frac{n-k+1}{kF_{2k,2(n-k+1),1-\alpha}}}$$

Equation 2. Exact lower confidence limit defined in terms of the F distribution; k, n, p_L, and α are defined in Equation 1. The result is identical to that obtained by iteration with Equation 1.

An exact solution for the confidence interval may sound like the ideal approach, but it is not well-aligned with Colorado's assessment needs, which would benefit from a target percentile that applies to data sets of any size. Because the binomial is a discrete distribution, assessments would logically be based on the [integer] number of samples that exceed the standard. However, the discreteness of the distribution precludes locating the LCL exactly for most sample sizes since the 85th percentile corresponds to an integer value of k only when sample size, n, is a multiple of 20[7]. When sample size is not a multiple of 20 and the confidence interval must be calculated with an integer, the resulting confidence interval would be larger than 95%. In other words, it becomes less likely that an exceedance will be identified.

Approximate methods also exist for defining confidence intervals for percentiles. In fact, many statisticians (e.g., Agresti and Coull 1998) recommend approximate methods because the exact method yields confidence intervals that tend to be too large (i.e., *exact* is something of a misnomer). A number of approximate methods have been developed for estimating confidence intervals. Historically, the Wald confidence interval has been recommended, especially when sample size is large[8]. It is also the easiest to understand.

The Wald test could be used to evaluate the null hypothesis that the estimated 85th percentile (i.e., the assessed value of a recent data set) is equal to the true 85th percentile (i.e., the standard, which characterizes existing quality). The difference between the estimate ($\hat{p}$) and the standard (p_0) is assumed to be approximately normally distributed at larger sample size. When the difference is divided by the standard error of the estimate (Equation 3), the result

[7] Given that the percentile is calculated as p=k/n, the first integer combination that delivers p=0.85 is when k=17 and n=20.

[8] A sample size of 30 or more is often regarded as large, but this rule of thumb may only be helpful where the central limit approximation is applicable. See Brown et al (2001) for a brief review of common sample size recommendations.

can be compared to standard normal deviates (z). Inverting the test yields a two-sided, 100(1-α)% confidence interval for p_0 (Equation 4).

$$z = \frac{(\hat{p} - p_0)}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}}$$

Equation 3. Wald test statistic for the difference between estimated and true values of a percentile.

$$\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

Equation 4. Inversion of the Wald test statistic to yield a confidence interval for the estimated percentile.

For many years, the Wald interval was the recommended approach for large sample size. A perusal of the literature shows that there is not much agreement now on what constitutes "small", especially where percentiles are extreme (e.g., close to zero or to 100%). At small sample size, which is common for assessments, the Wald interval tends to be too small to accurately define the 95% confidence interval. However, concerns about the performance of the Wald interval extend beyond the issue of sample size. Brown et al (2001) have shown that the Wald interval also exhibits "erratic behavior" even when sample size is large.

An alternative approximate method that seems to have broad support in the statistical literature is the Wilson, or Score Test, interval (Brown et al 2001 & 2002, Agresti and Coull 1998). It is an inversion of the score test, and the development of the equation is reviewed in Agresti and Coull (1998). The Wilson interval is proposed in preference to the Wald interval for improving assessment of ambient-based standards. The equation for the Wilson interval is somewhat intimidating (Equation 5), but it is manageable on a spreadsheet.

$$\left\{\hat{p} + z_{\alpha/2}^2/2n \pm z_{\alpha/2}\sqrt{\left[\hat{p}(1-\hat{p}) + z_{\alpha/2}^2/4n\right]/n}\right\}/(1 + z_{\alpha/2}^2/n)$$

Equation 5. The Wilson confidence interval for an estimated percentile.

Adaptation of Wilson Interval Method

The Wilson interval method is an improvement over the exact method and the Wald interval, but the discreteness issue remains. In exact and approximate methods, the target proportion ($\hat{p}$ = k/n) is formally defined by the [integer] number of "successes" (k) relative to sample size (n). As mentioned previously, most sample sizes do not match exactly the proportion (e.g., 85%) used for assessment.

The idiosyncrasies of Colorado's assessment method prompt consideration of a departure from the usual statistical approach of scoring "successes" on the basis of individual samples. In current methodology, assessment is based on locating the concentration that corresponds to the 85th percentile of the assessed data set. Usually, this step is carried out with the PERCENTILE function in EXCEL, which interpolates between measured concentrations. In other

words, the assessed value rarely matches a measured concentration from the assessed data set.

The equation for the Wilson interval requires values for $\hat{p}$, n, and z. There is no computational impediment to finding the LCL for any percentile in any data set. Hence, the equation is used to define the LCL for the three conventionally-assessed percentiles (50^{th}, 85^{th}, and 95^{th}) over a wide range of possible sample sizes (n = 5 through 100) with a one-sided, 95% confidence level (i.e., $\alpha = 0.05$). The values are summarized in **Supporting Table 1**. However, due to a quirk in the way EXCEL calculates percentiles, the values in **Supporting Table 1** must be adjusted, as explained later, when assessments are undertaken with EXCEL spreadsheets.

Translation of the LCL into an Excel-compatible Percentile

Percentiles in Excel are calculated by a method that does not match the calculations used to establish the LCLs. The computational differences are not major, but become increasingly important for small sample sizes. Insofar as the Excel function is widely used, we have incorporated an adjustment such that **Supporting Table 2** contains Excel-compatible LCLs. The basis for the adjustment and the interpretation of **Supporting Table 2** are given below.

Percentiles in Excel are set such that the smallest value is defined as 0% and the largest value is defined as 100%. The formula is p' = (k-1)/(N-1), where k is the rank of the observation and N is the number of observations in the data set (see Schoonjans et al 2011). Development of the LCL is based on the binomial distribution, which defines percentiles for each ranked observation as k/N. The largest concentration is still set to 100%, but the smallest observed concentration is 1/N rather than 0%. Thus, when the LCL is converted to a concentration, it could be smaller than the smallest observed value (and thus represent a percentile between zero and 1/N).

The difference between the two formulas affects conversion of the LCL to a concentration using EXCEL functions. Direct conversion of the LCL from **Supporting Table 1** with an Excel function would not yield the correct value. The error is very small when sample size is large, but cannot be ignored at small sample sizes. Therefore, an adjustment should be made.

The adjustment relies on simple algebra to translate the LCL from **Supporting Table 1** into LCL' that is compatible with Excel functions. In the Excel formula, where LCL' = (k-1)/(N-1), the binomial formula, LCL = k/N, is rearranged to enable substitution for k (k=LCL*N). After substituting for k, the equation for the translation is LCL' = (LCL*N-1)/(N-1). This translation is used to produce an Excel-compatible set of LCLs in **Supporting Table 2**. The threshold concentration can be derived directly with the Excel function as follows, PERCENTILE(*concentration_data*,LCL') where the assessed concentration data are in a defined range of cells.

Comparison to Exact Method

Exact and approximate methods can be compared only in those few cases where the integer values of k yield the target percentiles because the proposed method for locating the LCL is

not restricted to discrete values of k (p=k/n). For the 50th percentile, the comparison can be made for all even values of n; it shows that the exact method is more conservative, as expected, especially at small sample size. The exact and approximate methods agree within 10% when n exceeds about 20 (Figure 4.). Agreement between the two methods is generally even better for the 85th and 95th percentiles than the 50th percentile.

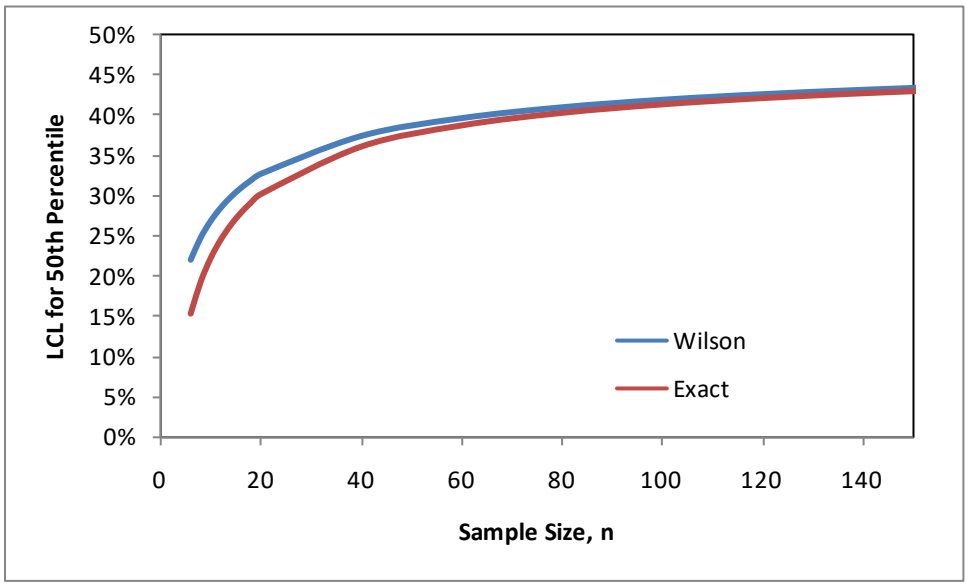


Figure 4. Comparison of LCLs defined by Wilson and exact methods for the 50th percentile. Comparison is limited to sample sizes with even numbers because the exact method can be evaluated only with integer values.

Sample Size

Sample sizes for water quality assessments tend to be small, often less than 10 samples for a segment. As is evident from **Supporting Table 1**, confidence intervals for percentiles are broader when sample sizes are small. Although LCL values can be calculated with the Wilson interval equation for virtually any sample size, there is a practical reason for avoiding very small sample sizes. When the percentile of the LCL is smaller than that of the lowest measured concentration (i.e., <1/N), it would correspond to a concentration smaller than any that were measured. For example, when only four samples are available, the LCL for the 50th percentile would be 0.182, which is smaller than 1/N (=0.25=1/4). When five samples are available, the LCL for the 50th percentile is 0.204, and this is larger than 1/N (=0.200). For the purpose of making listing decisions with ambient-based standards, at least five samples are required.

Setting a minimum of five samples for assessment of ambient-based standards differs from current assessment practice. With the latest version of Colorado's assessment methodology, a firm decision to place a segment on the 303(d) list requires more than ten samples (assuming the decision is based solely on concentration data for one constituent). With 4 to 10 samples, a listing decision must be backed up with evidence in addition to measured concentrations.

For three or fewer samples, high concentrations would at most trigger further sampling (M&E list).

Adoption of the LCL table could simplify decision-making for the listing methodology by having only two pathways related to sample size. In order to make a listing decision, there must be at least five representative[9] samples. When there are at least five samples, no additional supporting information is required because conclusions are equally reliable whether sample size is five or ten or fifty. When there are less than five representative samples, no action should be taken. If there is insistence on having a pathway to the M&E list, it should be on the basis of a table with a smaller confidence interval (e.g., 90%), which would also include a lower threshold for sample size.

Examples with Assessment Data

The current methodology for assessing attainment of ambient-based standards can be improved substantially by adding a defined level of confidence for the attainment decision. The statistical justification for the change is strong. Working through examples with real data is a good way to show that the improved approach is practical and efficient. The following examples incorporate data from the historical record for illustrative purposes, and they are not intended for reaching conclusions in the formal assessment process. Standards or segment descriptions may be changed through Commission actions (as happened with segment COARMA04a), and more recent data may be available for assessment. Nevertheless, these examples retain value for comparing old and new assessment methods for ambient-based standards.

One example is taken from Wildhorse Creek above Highway 50, which has ambient-based standards for selenium. Recent measurements (N=36; 2005-2011) of selenium concentrations serve as the assessed data set (Table 9). Most of the observed concentrations exceed table value standards (ch=4.6 ug/L; ac=18.4 ug/L) by a wide margin.

Date	Se, ug/L	Date	Se, ug/L	Date	Se, ug/L
10/13/2005	420	4/3/2007	500	7/1/2010	441
12/19/2005	496	6/5/2007	429	8/1/2010	564
2/15/2006	593	6/13/2007	95	9/1/2010	487
3/2/2006	535	9/20/2007	754	10/1/2010	479
4/26/2006	480	12/18/2007	691	11/1/2010	539
5/31/2006	362	12/1/2009	556	11/16/2010	1900
7/6/2006	9	1/1/2010	355	12/1/2010	618
9/5/2006	410	2/1/2010	646	2/1/2011	554
10/11/2006	361	3/1/2010	641	2/7/2011	1800

[9] The most recent listing methodology describes the factors to be considered when judging if data are "representative". Factors typically include spatial and temporal distribution of sampling effort, as well as analytical considerations and atypical events in the watershed.

APPENDIX B

Date	Se, ug/L	Date	Se, ug/L	Date	Se, ug/L
12/5/2006	224	4/1/2010	728	3/1/2011	607
2/20/2007	340	5/1/2010	605	4/1/2011	581
3/9/2007	531	6/1/2010	536	5/24/2011	1500

Table 9. Selenium concentrations measured in Wildhorse Creek above Highway 50.

The first step in assessing ambient-based standards by current methodology is to determine the 85th and 95th percentiles[10] of concentrations in the assessed data set. The 85th percentile concentration of the assessed data set is 680 ug/L, which exceeds the ambient-based chronic standard of 597 ug/L (Table 10). The 95th percentile concentration is 1575 ug/L, which exceeds the ambient-based acute standard of 708 ug/L. The assessed values exceed the ambient-based standards, which, by current assessment methodology, would trigger a listing.

Metric	Chronic	Acute
Ambient-based Standard	597	708
Assessed 85th Percentile	680	---
Assessed 95th Percentile	---	1575
Current outcome	Exceeded	Exceeded
Sample Size	36	36
LCL percentile (Appendix A)	0.728	0.853
LCL, Excel-compatible (App B)	0.720	0.849
LCL concentration	605	678
New outcome	Exceeded	OK

Table 10. Assessment of data for attainment of ambient-based selenium standards in Wildhorse Creek. LCL percentiles are taken from Appendices with N=36. The ambient-based standards were appropriate for segment COARMA04a at the time the samples were taken, but changes were adopted subsequently by the Commission.

The selenium data from Wildhorse Creek also are assessed with the improved methodology. Based on the sample size of 36, percentiles for chronic and acute LCLs are taken from the columns corresponding to the 85th and 95th percentiles in **Supporting Table 2**. The LCL for the 85th percentile is 0.720 (Excel-compatible value), which corresponds to a selenium concentration of 605 ug/L in the assessed data set. Therefore, the 85th percentile of the assessed data set is significantly larger than the chronic standard, which would trigger a listing by the improved methodology.

The conclusion about attainment can be properly reached based on assessment of the chronic standard alone, but the data also are assessed for attainment of the acute standard for illustrative purposes. In this case, the LCL corresponds to a selenium concentration of 678 ug/L, which is less than the acute standard. If assessment had been based solely on the acute

[10] Consistent with current methodology, threshold concentrations are determined with EXCEL's PERCENTILE function.

standard, the conclusion from the improved methodology would have been that the assessed value was not significantly larger than the acute standard, which would not trigger a listing.

A second example is taken from Big Dry Creek, which has a seasonal, ambient-based chronic standard for selenium; the acute standard remains equal to the table value standard (TVS). The example deals only with data from the Apr-Oct "season". Recent measurements (N=34; 2006-2010) of selenium concentrations serve as the assessed data set (Table 11). Some of the observed concentrations exceed the chronic TVS (4.6 ug/L), but none exceeds the acute TVS (18.4 ug/L).

Date	Se, ug/L	Date	Se, ug/L	Date	Se, ug/L
4/13/2006	8.2	6/19/2008	5.7	8/12/2010	2.3
4/12/2007	9.1	6/18/2009	2.9	9/14/2006	2.8
4/17/2008	6.4	6/10/2010	3.0	9/13/2007	2.9
4/9/2009	7.0	7/13/2006	5.5	9/11/2008	7.6
4/8/2010	5.7	7/12/2007	4.9	9/10/2009	3.7
5/11/2006	2.4	7/17/2008	2.3	9/9/2010	3.5
5/10/2007	3.4	7/9/2009	2.6	10/19/2006	5.6
5/8/2008	4.0	7/8/2010	2.3	10/11/2007	10.3
5/14/2009	7.5	8/10/2006	4.4	10/9/2008	8.1
5/13/2010	3.3	8/9/2007	7.1	10/15/2009	9.5
6/15/2006	2.9	8/14/2008	8.1		
6/14/2007	2.5	8/13/2009	5.3		

Table 11. Selenium concentrations measured in Big Dry Creek above the USGS gage (COSPBD01). Data are shown for Apr through Oct, 2006-2010.

When the selenium standards for Big Dry Creek are assessed by current methodology, the chronic standard is exceeded, but the acute standard is not (Table 12). The assessment would likely trigger a listing by current methodology. The data from Big Dry Creek also are assessed for the chronic standard with the improved methodology. The LCL for the 85th percentile is 0.716, which corresponds to a selenium concentration of 6.8 ug/L in the assessed data set. Therefore, the 85th percentile of the assessed data set is not significantly larger than the standard and would not trigger a listing. Application of the improved methodology highlights the value of using a defined level of confidence to support listing decisions.

Metric	Chronic	Acute
Ambient-based Standard	7.4	18.4
Assessed 85th Percentile	8.1	---
Assessed 95th Percentile	---	9.3
Current outcome	Exceeded	OK
Sample Size	34	
LCL percentile, Excel compatible	0.716	
LCL concentration	6.8	
New outcome	OK	

APPENDIX B

Table 12. Assessment of data for attainment of selenium standards in Big Dry Creek. The chronic standard is ambient-based, and the acute standard is TVS. Consequently, only the chronic standard is assessed with the improved methodology. The LCL percentile is taken from Appendix B with N=34.

General Comments

The proposed addition to assessment methodology for ambient-based standards offers an important statistical improvement by establishing a defined level of confidence to support impairment decisions. Obtaining the benefit of a defined confidence level comes at little additional cost because assessment requires almost no additional effort. In addition, the improved methodology retains the practical advantages of the current methodology in that it works for any constituent and requires no assumptions about the underlying statistical distribution.

The approach developed by the Division is efficient and effective for routine application, and the Division plans to apply it in the next listing cycle. Nevertheless, it is not the only approach that can provide a defined level of confidence. For example, a bootstrap approach could provide a non-parametric basis for assessment. Where a distribution can be identified, parametric options could be employed. These alternatives may be suitable where the resources are available to invest in developing a statistically-defensible approach for a particular water body.

Improving confidence in assessment decisions by any statistically-defensible method addresses important concerns about future commitment of resources where impairment is identified. By incorporating a defined level of confidence in assessments of ambient-based standards, the Division can be more certain that resources committed to TMDL development, for example, will not be wasted. A more reliable basis for listing decisions also should be well-received by stakeholders, who are affected by listing decisions (or reversals).

Increasing the statistical rigor of assessment also creates more incentive to set minimum requirements for development of ambient-based standards. Sample size and representativeness merit discussion that is beyond the scope of this assessment methodology.

Development of an improved approach for ambient-based standards invites the question of why a similar approach is not also proposed for TVS assessments. In the Division's view, there are important differences between the two kinds of standards. A TVS generally represents a physiological threshold above which concentrations threaten aquatic life. In contrast, assessment of ambient-based standards hinges on detecting degradation of water quality: Have concentrations increased significantly over "existing quality"? Thus, the Division does not recommend a change in current assessment practice for TVS.

Ambient-Based Delisting Determinations

The assessment methodology described above is a way to demonstrate that the water quality has degraded and is truly not attaining the applicable ambient-based standard. It achieves

this by deriving a lower confidence limit, based on the number of samples, for the 50th, 85th, and 95th percentiles. When that lower confidence limit is greater than the standard, there is a 95% statistical certainty that the standard was exceeded and that waterbody should be listed for non-attainment of that standard.

However, when assessing a waterbody that was added to either the 303(d) or Monitoring and Evaluation List based on the ambient-based listing methodology described in Appendix B, the use of the lower confidence limit is not appropriate. Instead, the upper confidence limit of the assessment percentile should be used. When that upper confidence limit is less than the standard (Figure 5), there is a 95% statistical certainty that the waterbody is attaining the standard and should be delisted. When the upper confidence limit is greater than the standard (Figure 6), the existing listing should be retained.

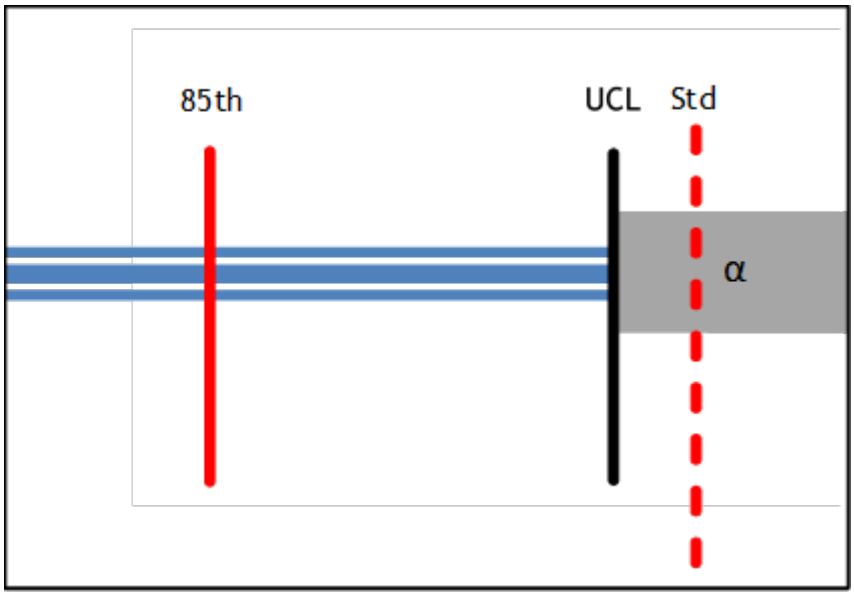


Figure 5 – The upper confidence limit of the 85th percentile is less than the standard. The waterbody should be delisted.

APPENDIX B

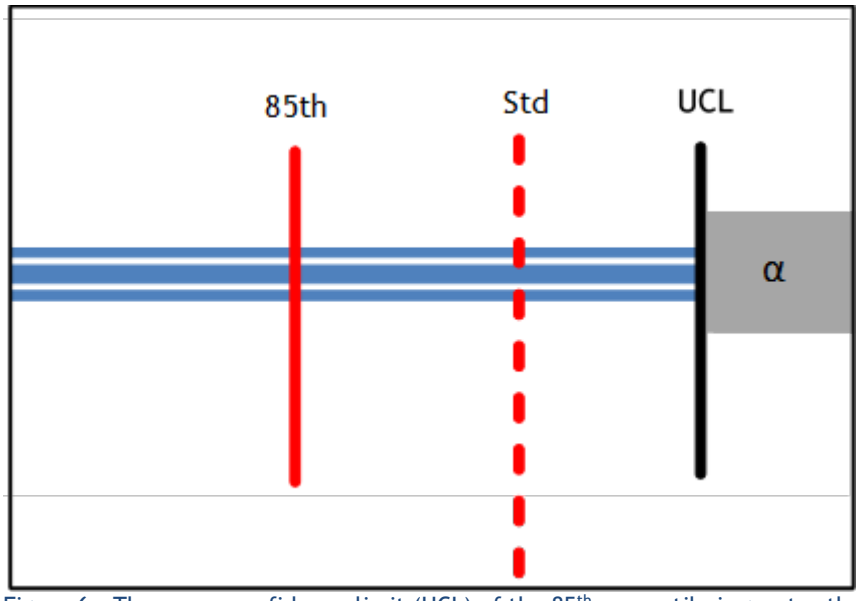


Figure 6 – The upper confidence limit (UCL) of the 85th percentile is greater than the standard. The existing listing should be retained.

As was done for the lower confidence limit, two supporting tables were created. The first of these new tables (Supporting Table 3) contains the calculated upper confidence limits, (based on equation 5 for the Wilson Interval) for a variety of sample sizes. The second new table (Supporting Table 4) contains the adjusted upper confidence limits that are compatible with the Excel PERCENTILE function. One minor difference between the two sets of supporting tables is that the number of samples in Tables 3 and 4 starts with a N of 10 because this is the minimum number of samples required for delisting actions.

References

Agresti, A and BA Coull. 1998. Approximate is better than "exact" for interval estimation of binomial proportions. The American Statistician 52(2): 119-126.

Brown, LD, TT Cai, and A DasGupta. 2001. Interval estimation for a binomial proportion. Statistical Science 16(2): 101-133 (includes appended commentary by several authors).

Brown, LD, TT Cai, and A DasGupta. 2002. Confidence intervals for a binomial proportion and asymptotic expansions. The Annals of Statistics 30(1): 160-201.

Helsel, DR, and RM Hirsch. 2002. Statistical methods in water resources. Chapter A3, Book 4 Hydrologic Analysis and Interpretation. Techniques of Water-Resources Investigations of the United States Geological Survey. http://water.usgs.gov/pubs/twri/twri4a3/

Leemis, LM and KS Trivedi. 1996. A comparison of approximate interval estimators for the Bernoulli parameter. The American Statistician 50: 63-68.

APPENDIX B

McBride, GB. 2005. Using statistical methods for water quality methods. Wiley-Interscience. 313 p.

Schoonjans, F, D De Bacquer, and P Schmid. 2011. Estimation of population percentiles. Epidemiology 22(5): 750-751.

Zar, JH. 1999. Biostatistical Analysis, 4th Edition. Prentice-Hall. 663 p. + Appendices.

APPENDIX B

Supporting Table 1: Lower confidence limits (LCLs) for three percentiles of regulatory interest (50th, 85th, and 95th) at assessed sample sizes of 5 through 100. These percentiles should not be used with the PERCENTILE function in EXCEL; the EXCEL-compatible values are presented in Appendix B.

N	$\hat{p}$ = 0.50	$\hat{p}$ = 0.85	$\hat{p}$ = 0.95		N	$\hat{p}$ = 0.50	$\hat{p}$ = 0.85	$\hat{p}$ = 0.95
5	0.204	0.482	0.588		53	0.390	0.752	0.875
6	0.221	0.514	0.625		54	0.391	0.754	0.876
7	0.236	0.540	0.654		55	0.392	0.755	0.877
8	0.249	0.561	0.678		56	0.393	0.756	0.878
9	0.260	0.579	0.698		57	0.394	0.756	0.879
10	0.269	0.595	0.715		58	0.394	0.757	0.880
11	0.278	0.608	0.730		59	0.395	0.758	0.881
12	0.286	0.619	0.742		60	0.396	0.759	0.881
13	0.292	0.630	0.753		61	0.397	0.760	0.882
14	0.299	0.639	0.763		62	0.398	0.761	0.883
15	0.305	0.647	0.772		63	0.399	0.762	0.884
16	0.310	0.654	0.780		64	0.399	0.763	0.884
17	0.315	0.661	0.787		65	0.400	0.763	0.885
18	0.319	0.667	0.793		66	0.401	0.764	0.886
19	0.323	0.673	0.799		67	0.401	0.765	0.886
20	0.327	0.678	0.804		68	0.402	0.765	0.887
21	0.331	0.683	0.809		69	0.403	0.766	0.887
22	0.335	0.687	0.813		70	0.404	0.767	0.888
23	0.338	0.692	0.818		71	0.404	0.768	0.889
24	0.341	0.695	0.821		72	0.405	0.768	0.889
25	0.344	0.699	0.825		73	0.405	0.769	0.890
26	0.346	0.703	0.828		74	0.406	0.769	0.890
27	0.349	0.706	0.832		75	0.407	0.770	0.891
28	0.352	0.709	0.834		76	0.407	0.771	0.891
29	0.354	0.712	0.837		77	0.408	0.771	0.892
30	0.356	0.714	0.840		78	0.408	0.772	0.892
31	0.358	0.717	0.842		79	0.409	0.772	0.893
32	0.360	0.719	0.845		80	0.410	0.773	0.893
33	0.362	0.722	0.847		81	0.410	0.774	0.894
34	0.364	0.724	0.849		82	0.411	0.774	0.894
35	0.366	0.726	0.851		83	0.411	0.775	0.895
36	0.368	0.728	0.853		84	0.412	0.775	0.895
37	0.369	0.730	0.855		85	0.412	0.776	0.895
38	0.371	0.732	0.856		86	0.413	0.776	0.896
39	0.373	0.734	0.858		87	0.413	0.777	0.896
40	0.374	0.735	0.860		88	0.414	0.777	0.897
41	0.376	0.737	0.861		89	0.414	0.777	0.897
42	0.377	0.738	0.863		90	0.415	0.778	0.897
43	0.378	0.740	0.864		91	0.415	0.778	0.898
44	0.380	0.741	0.865		92	0.415	0.779	0.898
45	0.381	0.743	0.867		93	0.416	0.779	0.898
46	0.382	0.744	0.868		94	0.416	0.780	0.899
47	0.383	0.745	0.869		95	0.417	0.780	0.899
48	0.385	0.747	0.870		96	0.417	0.781	0.900
49	0.386	0.748	0.871		97	0.418	0.781	0.900
50	0.387	0.749	0.872		98	0.418	0.781	0.900
51	0.388	0.750	0.873		99	0.418	0.782	0.901
52	0.389	0.751	0.874		100	0.419	0.782	0.901

APPENDIX B

Supporting Table 2: Lower confidence limits (LCLs) for three percentiles of regulatory interest (50th, 85th, and 95th) at assessed sample sizes of 5 through 100. These percentiles are compatible with the PERCENTILE function in EXCEL; see text for explanation.

N	$\hat{p}$ = 0.50	$\hat{p}$ = 0.85	$\hat{p}$ = 0.95		N	$\hat{p}$ = 0.50	$\hat{p}$ = 0.85	$\hat{p}$ = 0.95
5	0.005	0.353	0.485		53	0.378	0.748	0.873
6	0.066	0.417	0.550		54	0.379	0.749	0.874
7	0.109	0.464	0.597		55	0.380	0.750	0.875
8	0.141	0.499	0.632		56	0.382	0.751	0.876
9	0.167	0.527	0.661		57	0.383	0.752	0.877
10	0.188	0.550	0.684		58	0.384	0.753	0.878
11	0.206	0.569	0.703		59	0.385	0.754	0.878
12	0.221	0.585	0.719		60	0.386	0.755	0.879
13	0.234	0.599	0.733		61	0.387	0.756	0.880
14	0.245	0.611	0.745		62	0.388	0.757	0.881
15	0.255	0.622	0.755		63	0.389	0.758	0.882
16	0.264	0.631	0.765		64	0.390	0.759	0.882
17	0.272	0.640	0.773		65	0.391	0.760	0.883
18	0.279	0.648	0.781		66	0.392	0.760	0.884
19	0.286	0.655	0.788		67	0.392	0.761	0.884
20	0.292	0.661	0.794		68	0.393	0.762	0.885
21	0.298	0.667	0.799		69	0.394	0.763	0.886
22	0.303	0.673	0.804		70	0.395	0.763	0.886
23	0.308	0.678	0.809		71	0.396	0.764	0.887
24	0.312	0.682	0.814		72	0.396	0.765	0.888
25	0.316	0.687	0.818		73	0.397	0.766	0.888
26	0.320	0.691	0.822		74	0.398	0.766	0.889
27	0.324	0.694	0.825		75	0.399	0.767	0.889
28	0.328	0.698	0.828		76	0.399	0.768	0.890
29	0.331	0.701	0.831		77	0.400	0.768	0.890
30	0.334	0.704	0.834		78	0.401	0.769	0.891
31	0.337	0.707	0.837		79	0.401	0.769	0.891
32	0.340	0.710	0.840		80	0.402	0.770	0.892
33	0.342	0.713	0.842		81	0.403	0.771	0.892
34	0.345	0.716	0.844		82	0.403	0.771	0.893
35	0.347	0.718	0.847		83	0.404	0.772	0.893
36	0.350	0.720	0.849		84	0.405	0.772	0.894
37	0.352	0.722	0.851		85	0.405	0.773	0.894
38	0.354	0.725	0.853		86	0.406	0.773	0.895
39	0.356	0.727	0.854		87	0.406	0.774	0.895
40	0.358	0.728	0.856		88	0.407	0.774	0.895
41	0.360	0.730	0.858		89	0.407	0.775	0.896
42	0.362	0.732	0.859		90	0.408	0.775	0.896
43	0.364	0.734	0.861		91	0.409	0.776	0.897
44	0.365	0.735	0.862		92	0.409	0.776	0.897
45	0.367	0.737	0.864		93	0.410	0.777	0.897
46	0.368	0.738	0.865		94	0.410	0.777	0.898
47	0.370	0.740	0.866		95	0.411	0.778	0.898
48	0.371	0.741	0.867		96	0.411	0.778	0.898
49	0.373	0.743	0.869		97	0.412	0.779	0.899
50	0.374	0.744	0.870		98	0.412	0.779	0.899
51	0.376	0.745	0.871		99	0.413	0.779	0.900
52	0.377	0.747	0.872		100	0.413	0.780	0.900

APPENDIX B

Supporting Table 3: Upper confidence limits (UCLs) for three percentiles of regulatory interest (50th, 85th, and 95th) at assessed sample sizes of 10 through 100. These percentiles should not be used with the PERCENTILE function in EXCEL; the EXCEL-compatible values are presented in Supporting Table 4.

N	$\hat{p}$ = 0.50	$\hat{p}$ = 0.85	$\hat{p}$ = 0.95		N	$\hat{p}$ = 0.50	$\hat{p}$ = 0.85	$\hat{p}$ = 0.95
10	0.731	0.956	0.993		56	0.607	0.912	0.980
11	0.722	0.954	0.993		57	0.606	0.912	0.980
12	0.714	0.952	0.992		58	0.606	0.911	0.980
13	0.708	0.950	0.992		59	0.605	0.911	0.980
14	0.701	0.948	0.991		60	0.604	0.911	0.980
15	0.695	0.946	0.991		61	0.603	0.910	0.980
16	0.690	0.944	0.990		62	0.602	0.910	0.980
17	0.685	0.943	0.990		63	0.601	0.909	0.979
18	0.681	0.941	0.990		64	0.601	0.909	0.979
19	0.677	0.940	0.989		65	0.600	0.909	0.979
20	0.673	0.938	0.989		66	0.599	0.908	0.979
21	0.669	0.937	0.988		67	0.599	0.908	0.979
22	0.665	0.936	0.988		68	0.598	0.908	0.979
23	0.662	0.935	0.988		69	0.597	0.907	0.979
24	0.659	0.934	0.987		70	0.596	0.907	0.979
25	0.656	0.933	0.987		71	0.596	0.907	0.978
26	0.654	0.931	0.987		72	0.595	0.906	0.978
27	0.651	0.931	0.987		73	0.595	0.906	0.978
28	0.648	0.930	0.986		74	0.594	0.906	0.978
29	0.646	0.929	0.986		75	0.593	0.906	0.978
30	0.644	0.928	0.986		76	0.593	0.905	0.978
31	0.642	0.927	0.985		77	0.592	0.905	0.978
32	0.640	0.926	0.985		78	0.592	0.905	0.978
33	0.638	0.925	0.985		79	0.591	0.904	0.977
34	0.636	0.925	0.985		80	0.590	0.904	0.977
35	0.634	0.924	0.984		81	0.590	0.904	0.977
36	0.632	0.923	0.984		82	0.589	0.904	0.977
37	0.631	0.922	0.984		83	0.589	0.903	0.977
38	0.629	0.922	0.984		84	0.588	0.903	0.977
39	0.627	0.921	0.984		85	0.588	0.903	0.977
40	0.626	0.920	0.983		86	0.587	0.903	0.977
41	0.624	0.920	0.983		87	0.587	0.902	0.977
42	0.623	0.919	0.983		88	0.586	0.902	0.977
43	0.622	0.919	0.983		89	0.586	0.902	0.976
44	0.620	0.918	0.983		90	0.585	0.902	0.976
45	0.619	0.917	0.982		91	0.585	0.901	0.976
46	0.618	0.917	0.982		92	0.585	0.901	0.976
47	0.617	0.916	0.982		93	0.584	0.901	0.976
48	0.615	0.916	0.982		94	0.584	0.901	0.976
49	0.614	0.915	0.982		95	0.583	0.901	0.976
50	0.613	0.915	0.981		96	0.583	0.900	0.976
51	0.612	0.914	0.981		97	0.582	0.900	0.976
52	0.611	0.914	0.981		98	0.582	0.900	0.976
53	0.610	0.914	0.981		99	0.582	0.900	0.976
54	0.609	0.913	0.981		100	0.581	0.899	0.975
55	0.608	0.913	0.981					

APPENDIX B

Supporting Table 4 : Upper confidence limits (UCLs) for three percentiles of regulatory interest (50th, 85th, and 95th) at assessed sample sizes of 10 through 100. These percentiles are compatible with the PERCENTILE function in EXCEL; see text for explanation.

N	$\hat{p}$ = 0.50	$\hat{p}$ = 0.85	$\hat{p}$ = 0.95		N	$\hat{p}$ = 0.50	$\hat{p}$ = 0.85	$\hat{p}$ = 0.95
10	0.701	0.951	0.992		56	0.600	0.911	0.980
11	0.694	0.949	0.992		57	0.599	0.910	0.980
12	0.689	0.947	0.991		58	0.599	0.910	0.980
13	0.683	0.946	0.991		59	0.598	0.909	0.980
14	0.678	0.944	0.990		60	0.597	0.909	0.980
15	0.674	0.942	0.990		61	0.596	0.909	0.979
16	0.670	0.941	0.990		62	0.596	0.908	0.979
17	0.666	0.939	0.989		63	0.595	0.908	0.979
18	0.662	0.938	0.989		64	0.594	0.908	0.979
19	0.659	0.936	0.989		65	0.594	0.907	0.979
20	0.655	0.935	0.988		66	0.593	0.907	0.979
21	0.652	0.934	0.988		67	0.592	0.907	0.979
22	0.650	0.933	0.988		68	0.592	0.906	0.978
23	0.647	0.932	0.987		69	0.591	0.906	0.978
24	0.644	0.931	0.987		70	0.591	0.906	0.978
25	0.642	0.930	0.987		71	0.590	0.905	0.978
26	0.640	0.929	0.986		72	0.589	0.905	0.978
27	0.637	0.928	0.986		73	0.589	0.905	0.978
28	0.635	0.927	0.986		74	0.588	0.905	0.978
29	0.633	0.926	0.985		75	0.588	0.904	0.978
30	0.632	0.925	0.985		76	0.587	0.904	0.978
31	0.630	0.924	0.985		77	0.587	0.904	0.977
32	0.628	0.924	0.985		78	0.586	0.903	0.977
33	0.626	0.923	0.984		79	0.586	0.903	0.977
34	0.625	0.922	0.984		80	0.585	0.903	0.977
35	0.623	0.922	0.984		81	0.585	0.903	0.977
36	0.622	0.921	0.984		82	0.584	0.902	0.977
37	0.620	0.920	0.984		83	0.584	0.902	0.977
38	0.619	0.920	0.983		84	0.583	0.902	0.977
39	0.618	0.919	0.983		85	0.583	0.902	0.977
40	0.616	0.918	0.983		86	0.582	0.901	0.976
41	0.615	0.918	0.983		87	0.582	0.901	0.976
42	0.614	0.917	0.982		88	0.582	0.901	0.976
43	0.613	0.917	0.982		89	0.581	0.901	0.976
44	0.612	0.916	0.982		90	0.581	0.901	0.976
45	0.610	0.916	0.982		91	0.580	0.900	0.976
46	0.609	0.915	0.982		92	0.580	0.900	0.976
47	0.608	0.915	0.982		93	0.580	0.900	0.976
48	0.607	0.914	0.981		94	0.579	0.900	0.976
49	0.606	0.914	0.981		95	0.579	0.899	0.976
50	0.605	0.913	0.981		96	0.578	0.899	0.976
51	0.604	0.913	0.981		97	0.578	0.899	0.975
52	0.604	0.912	0.981		98	0.578	0.899	0.975
53	0.603	0.912	0.981		99	0.577	0.899	0.975
54	0.602	0.911	0.980		100	0.577	0.898	0.975
55	0.601	0.911	0.980					

APPENDIX C

Appendix C - Technical Basis for Defining the Temperature Warming Event

Background

In the 2016 Regulation 31 hearing, the Water Quality Control Commission (WQCC) adopted a new definition of existing quality for temperature which specifies a 1 in 3 year average recurrence frequency of a "warming event" and directed that the next 303d Listing methodology define "warming event" with a method consistent with this decision. The WQCC also directed the division to "look at the impacts of duration, multiplicity and cumulative effects" (Regulation 31.53(A)). The following definition/methodology was derived using literature used in the newly updated, 2016 temperature database and specifies an allowable cumulative impact during a warming event with a recurrence frequency of once in three years on average[1]. With input from stakeholders (including representatives from CPW, EPA, USGS, Metro Wastewater, Eagle River Water and Sanitation and the Colorado Wastewater Utilities Council) the division has proposed a method which relies on the concept of "degree-days" which integrates both the magnitude of temperatures over the standard, as well the duration, in days, experienced by the aquatic community. Because the chronic and acute standards were developed from tests on different endpoints, the workgroup developed a definition for a chronic and acute event.

Chronic Warming Event

The chronic event was derived with growth optimum studies and was calculated as the cumulative temperature above the standard at which 20% reduction in growth is expected. This 20% reduction in growth is akin to an effects concentration to 20% of the test population (EC20), used by EPA for derivation of chronic toxicant standards. The median difference between the EC20 temperature and the standard for both cold and warm fish was multiplied by 7 days to be consistent with the duration of the chronic standard. For cold tiers (I and II) 13.5 degree-days (°C) is the maximum extent of warming and for warm tiers (I, II, III) 36.6 degree-days is the maximum extent of a warming event, above which a listing decision would go forward. See Tables 1 and 2.

[1] This is specified by the recurrence interval equation **T = (n+1)/m.** Where **T** is the known recurrence interval (3 years), **n** is the number of years of the record and **m** is the magnitude ranking.

APPENDIX C

Cold Water Chronic Warming Event Derivation Table						
Species	Chronic T std	Growth rate @ chronic T std	20% reduction in growth rate	T @ 20% growth reduction	Exceedence of chronic temperature standard resulting in 20% reduction in growth rate	Reference
Mountain Whitefish	18.3	0.0652	0.0522	20.11	1.81	Brinkman et al. 2013
Rio Grande Cutthroat Trout	17.0	37.9	30.3	18.92	1.92	Ziegler et al. 2013
Westslope Cutthroat Trout	17.0	1.9	1.5	18.36	1.36	Bear et al. 2007
Rainbow Trout	18.3	1.7	1.4	19.92	1.62	Bear et al. 2007
Rainbow Trout	18.3	4.8	3.8	23.21	4.91	Hokanson et al. 1977
RBT AVERAGE					3.26	
Brook Trout	17.0	0.042	0.034	18.76	1.76	McMahon et al. 2007
Brook Trout	17.0	4.301	3.440	19.13	2.13	McCormick et al. 1972
BRK AVERAGE					1.94	
Brown Trout	18.3	4.3	3.405	20.44	2.14	Ojanguren et al 2001-test temperature data provided
Median Deviation (using RBT and BRK average)					**1.93**	
Chronic Degree Day allowance (Median Deviation x 7days)					**13.5**	

Table 13. Cold Water Chronic Warming Event Derivation. Numbers represent degrees Celsius.

Warm Water Chronic Warming Event Derivation Table							
Species	GO from lit	Chronic T std	Growth rate @ chronic T std	20% reduction in growth rate	T @ 20% growth reduction	Exceedence of chronic temperature standard resulting in 20% reduction in growth rate	Reference
White Sucker	26.9	27.5	12.900	10.320	29.45	1.94	McCormick et al. 1977
Western Mosquitofish	25	28.7	19.7	15.8	40.25	11.55	Wurtsbaugh, W.A., Cech, J.J., Jr. 1983.
Smallmouth Bass	24	28.7	1.0	0.8	30.72	2.05	Horning, W.B., II, Pearson, R.E. 1973.
Bluegill	31	28.7	3.7	2.93048	34.06	5.4	Beitinger, T., and J. Magnuson. 1979.
Bluegill	30	28.7	2.5	2.01040272	36.5	7.8	Lemke, A. 1977.
bluegill average						6.6	
spottail shiner	27.3	28.7	29.2	23.35454592	31.935	3.3	Kellogg, R.L., Gift, J.J. 1983.
channel catfish	30	28.7	559.1	447.2706662	33.746	5.1	Andrews, J.W., and R.R. Stickney. 1972.
bonytail	25.9	28.7	0.233	0.186651099	35.35	6.7	Kappenman, K. M., E. S. Cureton, J. Ilgen, M. Toner, W. C. Fraser, and G. A. Kindschi. 2012.
Median Deviation (using average for bluegill)					**5.08**		
Chronic Degree Day allowance (Median Deviation x 7days)					**35.6**		

Table 14 Warm Water Chronic Warming Event Derivation. Numbers represent degrees Celsius.

Acute Warming Event

Acute values were derived using Upper Incipient Lethal Temperatures (the temperature resulting in 50% mortality or loss of equilibrium where fish are quickly transferred to the test temperature), and converted Critical Thermal Maxima (CTM) tests. Most of the acute studies in the 2016 temperature database were performed using the CTM test where fish were slowly warmed until mortality or loss of equilibrium was observed for 50% of the test population. These CTM tests produce larger values than the UILT tests and as explained in Temperature Policy 06-1, these CTM tests can be converted to a value closer to the expected UILT values. Therefore the workgroup used UILT and converted CTM values to derive the acute warming event. The workgroup used a similar method and rationale as that employed for the chronic event calculation and used the typical (median) difference between the standard and the median UILT value (an EC50) for cold and warm communities. Currently the standard is the value that protects 95% of the community from mortality, (or in the case of the CSI, protects Cutthroat trout, an important species). The UILT may have occurred within seconds, minutes, days or weeks at that test temperature and is generally an interpolated value from a % mortality vs test temperature regression. Because the standard protects from harm and the purpose of the warming event concept is allowing for limited harm during the unsafe, once-in-three-year on average exceedance event, setting the exceedance magnitude and duration to the typical difference between the standard and cold and warm fish UILTs would be a

reasonable method to allow limited harm without allowing large-scale mortality events to the fishery. As the time to LT50 is unknown in many of these CTM and UILT studies and acute mortality can occur in a single day, allowing 1 day as the multiplier (to get from degrees to degree-days) would be a conservative assumption which would also allow for limited exceedances of the acute standard to occur over a number of days if the magnitude of those exceedances are small. See Tables 3 and 4 for derivation details.

Species COLD TIER FISH	Acute Standard	Median UILTs	Deviation from STD
cutthroat trout	21.70	24.10	2.40
brook trout	21.70	23.10	1.40
rainbow trout	24.30	25.65	1.35
brown trout	24.30	27.28	2.98
mountain whitefish	21.70	23.10	1.40
longnose sucker	24.30	26.90	2.60
mottled sculpin	24.30	29.85	5.55
CTM-->UILT Conversion		MEDIAN COLD	2.40

Table 15 Cold Water Acute Warming Derivation Table.

APPENDIX C

Species WARM TIER FISH	Acute Standard	Median UILTs	Deviation from STD
common shiner	29.0	31	2.0
orangethroated darter	29.0	32.355	3.4
Johnny darter	29.0	30.6	1.6
Northern redbelly dace	28.6	32.7	4.1
white sucker	28.6	29.3	0.7
razorback sucker	28.6	38.30	9.7
central stoneroller	28.6	36.4	7.8
finescale dace	28.6	32	3.4
longnose dace	28.6	30.6	2.0
brook stickleback	28.6	30.6	2.0
creek chub	28.6	30.8	2.2
brown bullhead	31.8	34.90	3.1
pumkinseed sunfish	31.8	34.30	2.5
fathead minnow	31.8	33.2	1.4
western mosquitofish	31.8	37	5.2
smallmouth bass	31.8	36.1	4.3
green sunfish	31.8	36.85	5.1
bluegill	31.8	34.30	2.5
spottail shiner	31.8	33.10	1.3
channel catfish	31.8	37.2	5.4
plains minnow	31.8	38.9	7.1
sand shiner	31.8	32.2	0.4
speckled dace	31.8	33.82	2.0
horneyhead chub	31.8	34.8	3.0
Southern redbelly dace	31.8	35.1	3.3
orangespotted sunfish	31.8	35.6	3.8
bigmouth shiner	31.8	35.8	4.0
plains topminnow	31.8	36.2	4.4
Arkansas Darter	31.8	36.5	4.7
rountail chub	31.8	36.5	4.7
red shiner	31.8	37	5.2
yellow bullhead	31.8	37.2	5.4
black bullhead	31.8	37.3	5.5
boneytail	31.8	38.05	6.3
plains killifish	31.8	40.45	8.7
CTM-->UILT Conversion	MEDIAN WARM		3.8

Table 16 Warm Water Acute Warming Derivation Table.

Interrupted warming events

Instances where stream temperatures exceed standards for a number of days and start the cumulative warming event count, but then drop back below the standard for a day or more followed by another increase over the standard posed a problem for this method. The technical subgroup discussed the theoretical problem of a cluster of exceedances at the beginning of a season and again at the end of a season but which cumulatively falls beneath the total degree day allowance. The subgroup voiced support of a more simple but strict implementation of "event" by allowing only one continuous string of days over the standard. However, the division believed that a temporary break in high temperatures is beneficial for the aquatic life and there is also little to no functional evidence indicating what impacts can be expected from short term recurrences of clusters of high temperature days vs one continuous warming event if the sum of degree days over the standard are equal.

Example of the calculations used to assess the duration of one warming event

APPENDIX C

For this example data record, the cold water standard is 18.3 degrees Celsius, one warming event above the standard in either the summer or winter within one year is allowed, and the maximum warming event above the chronic (WAT) standard in cold streams is 13.5 degree-days.

The following steps are used to assess one warming event:

1. For every day that water temperatures exceed the standard (starting on 7/19/2010), the difference between the water temperature and the standard is calculated (see column 'Dev. from std').
2. The sum of the differences between the standard and the stream temperature is added each day and that sum is compared to the allowable degree-days (see 'Cumulation' column). For the data record below, the cumulation of degrees exceeds 13.5 on 7/30/2010, so the temperature exceedances of the standard from 7/30/2010 to 8/4/2010 are no longer within the one warming event allowance for this site (see red cells). The sum is paused on days when attainment occurs (see grey cells).

Date	WAT	Dev. from std	Cumulation
7/19/2010	18.48	0.18	0.18
7/20/2010	18.60	0.30	0.48
7/21/2010	19.20	0.90	1.38
7/22/2010	21.40	3.10	4.48
7/23/2010	21.60	3.30	7.78
7/24/2010	21.00	2.70	10.48
7/25/2010	20.20	1.90	12.38
7/26/2010	18.90	0.60	12.98
7/27/2010	18.70	0.40	13.38
7/28/2010	18.40	0.10	13.48
7/29/2010	18.10	-0.20	13.48
7/30/2010	18.40	0.10	13.58
7/31/2010	18.90	0.60	14.18
8/1/2010	18.00	-0.30	14.18
8/2/2010	20.21	1.91	16.09
8/3/2010	19.91	1.61	17.70
8/4/2010	19.51	1.21	18.91

Applicable temperature standard is 18.3 degrees C.

Because this site includes exceedances of the temperature standard that extend beyond the one warming event (see red cells), the site would not be considered in attainment of temperature standards.

3. If water temperatures exceed the standard for multiple seasons or multiple years, then those exceedances of the standard are considered to be more than one event. For the data record below, water temperatures exceed the standard for multiple years. The exceedances of the temperature standard from 7/15/2011 to 7/19/2011 are a separate event and extend

beyond the extent of one warming event (see red cells). This example represents warming above the standard at a greater exceedance frequency than once in three years on average.

CS-II (24.3 degrees C - DM) Warming event (2.4 degree-days)			
Date	**DM**	**Dev. from std**	**Cumulation**
7/19/2010	24.489	0.19	0.19
7/20/2010	24.478	0.18	0.37
7/21/2010	24.544	0.24	0.61
7/22/2010	24.889	0.59	1.20
7/23/2010	24.467	0.17	1.37
7/15/2011	24.389	0.09	NA*
7/16/2011	24.439	0.14	NA*
7/17/2011	24.903	0.60	NA*
7/18/2011	24.399	0.10	NA*
7/19/2011	24.401	0.10	NA*
*The cumulation of degrees was not calculated because one warming event above the standard in either the summer or winter in one year is allowed for data records with 4-6 years of data. These exceedances of the standard extend beyond the allowable exceedance frequency of one warming event in three years on average.			

Because this site includes exceedances of the temperature standard that extend beyond the one warming event, the site would not be considered in attainment of temperature standards.

Appendix D - Hess Method Sample Rules/Modifications

1. **Qualifications**
    a. The rules and modifications shall only apply to data from at least three (3) Hess samples. These samples need to be processed separately (without subsampling), then data from all samples must be composited.
    b. The rules and modifications shall only apply to samples collected in Biotypes 1 and 2.
    c. The rules and modifications shall only apply to original samples greater than 360 individuals.
    d. The rules and modifications shall only apply to samples where benthic macroinvertebrate identification is conducted to the Operational Taxonomic Unit (OTU) level recommended by the division.
    e. The sum of individuals reached through the process of adding large and rare (L+R) taxa shall not exceed 360. Doing so will activate another sub-sample, which would negate the rarification process.
    f. The rules and modifications listed below shall be followed to exact specifications in order to correctly duplicate (or mimic) the process of adding L+R taxa to the laboratory picked sub-sample. Deviations from these rules will invalidate the resultant MMIs.

2. **Initial Data Preparation**
    a. Upload a list of taxa and counts, as reported by the taxonomist, into CO-EDAS using the division's Bug Import Sheet, which is a specifically formatted spreadsheet.
    b. Run sub-sample program to generate a 300-fixed count. The data is now considered rarified.

3. **Rarified Data Preparation**
    a. Locate the sub-sample data results in EDAS "*Benthics*" table. These original results are identified as "*Individuals*" while the sub-sampled results are identified as "*Ind_300*".
    b. Re-enter the 300-fixed count data, as selected by the sub-sample program, into a new bug import sheet ("*Final Import Sheet*").
    c. Follow the rules, allowable modifications, and guidelines in Section 4.

4. **Rules and Allowable Modifications**
    a. If an L+R taxon is replaced in the Final Import Sheet it shall be represented by 1 individual.
    b. If a taxon constitutes more than 0.33% of the original sample, and has been excluded, it shall be included and represented by 1 individual.

Example: An original sample has 400 organisms, including 4 individuals from Species A (1% of original sample). However, all Species A were excluded during the sub-sampling process. One (1) individual representing Species A shall be included in the Final Import Sheet.

c. Additional taxa may be added as L+R in the Final Import Sheet but they must have occurred in the original sample and had been omitted by the sub-sample program. These taxa must be larger than 1 cm during late stages of development and must be distinguishable from other similar taxa without the assistance of magnification.

Other more specific guidelines are as follows:

i. Representatives from most aquatic macroinvertebrate Orders and Families can be considered L+R if they exceed 1 cm in length and there are no other representatives from that Order or Family already included in the sample.

Example: One (1) individual from the Genus *Cheumatopsyche* shall be included as L+R only if the Family Hydropsychidae is not already represented in the sample.

ii. Members of the Family Chironomidae are not generally considered L+R. Only when there are no members of this Family in the Final Import Sheet and the sum of excluded members meets the requirements of step 4.b. should they be included.
iii. Aquatic mites shall be treated with the same guidelines that apply to Family Chironomidae.
iv. Aquatic worms and snails shall be included as L+R taxa, providing they meet the above requirements.
v. Taxa that are large enough to be considered L+R, despite the presence of other members within the same family, include Genus: *Drunella, Epeorus, Rhithrogena, Acroneuria, Claassenia, Hesperoperla, Diura, Skwala, Pteronarcys,* and *Arctopsyche*.
vi. Any taxon added as L+R must be recognizable as a unique taxon without the assistance of magnification.

5. **Post-Processing Steps**
   a. Upload the Final Import Sheet back into CO-EDAS.
   b. Verify the 300 organism sub-sample. This is required because the "*Ind_300*" is the final individual count that operates the MMI. Please see qualification 1.e. above.
   c. Calculate MMI(s).

Appendix E - Criteria and Guidelines for Data Submitted After the Data Call

In June of each year, the division issues an annual data call letter to solicit data from specific regulatory basins. Stakeholders have three months to submit data in the specified format with minimal data elements, such as detection limits, GPS coordinates, lab methods, etc. This solicitation for readily available data is consistent with requirements set forth in EPA's 2006 Integrated Reporting Guidance.

Existing data which are not brought forward through the data call may be provided for consideration during written testimony through the rulemaking hearing process. The division and commission strongly encourage submittal of data during the data call, as submitting during the hearing process limits the time available for all stakeholders to evaluate and assess the data and requires significant duplication of efforts by the division. The commission recommends that data and information provided after the data call, meet these criteria described below and be accompanied by a document describing how the criteria are met. These criteria will minimize the resource demand and duplication of efforts needed to process these data outside of the division's standard practices.

Requested elements and guidelines:

1. Communicate with the division regarding your interest in submitting data after the data call deadline as soon as possible.
2. Describe the reasons these data were not submitted during the data call.
3. Document the waterbodies and the associated waterbody segment ID's where these data were collected.
4. Data should represent the suite of analytes available from the entity submitting the data, not simply a single parameter that may influence a listing decision. Please contact the division to discuss which analytes would be informative to the assessment process.
5. Data should meet the prescribed format, and include minimum data elements described in the Listing Methodology Section III.B.1. Please contact the division for current format and minimum data elements.
6. If the data you intend to submit are outside of the period of record from the initial assessment cycle, please also include any other data available from the Water Quality Portal for the waterbody of interest for the extended period. This should include data for all waterbodies with an attainment status that may be affected by the newly submitted data. (This step is to ensure all available data is considered in the revised water quality assessment for a waterbody.)

If your entity is resource-limited and cannot complete these criteria, please reach out to the division for assistance. Data call notifications are provided to those entities who sign up for notifications on the Colorado Water Quality Control Commission website:

https://cdphe.stg.colorado.gov/wqcc-how-get-notices-upcoming-hearings

APPENDIX E

To receive data call information, select Regulation #93 and the regulatory basins (#32-#38) of interest.