Appendix L: Assessment of Compliance with the Mercury Water Quality Objectives

This appendix describes proposed methods that can be considered for assessing progress towards attainment with the mercury water quality objectives pertaining to human and wildlife consumption of fish and the COMM, WILD, and RARE beneficial uses. The "Water Quality Control Policy for Developing California's Clean Water Act Section 303(d) List" (SWRCB 2004) (Listing Policy) prescribes the methods for adding or removing a water body from the Clean Water Act Section 303(d) list of waters not meeting water quality standards. The Listing Policy provides that a situation-specific weight of evidence approach may be used to list a water segment as not meeting water quality standards if other Listing Policy factors do not result in the listing of the water segment. Likewise, the Listing Policy allows a situation-specific weight of evidence approach when the delisting factors contained in the Listing Policy do not result in the delisting of a water segment but information indicates attainment of standards. This appendix provides an alternative situation specific weight of evidence approach that can be considered for either adding or removing a water body from the section 303d list.

The first section discusses the methodology for consistent sample collection and analysis, including fish species, sizes, and sampling locations for comparison with the Mercury Water Quality Objectives. Section L.2 evaluates various methods to assess attainment of water quality standards in reservoirs, including binomial distribution, numeric averaging, and statistical averaging using the upper 90th confidence limit of the mean. Section L.3 describes a proposed method to determine compliance with the water quality objectives using a situation specific weight of evidence approach that is consistent with the Listing Policy to list or delist a water body.

L.1 Sample Collection Methods for the Mercury Water Quality Objectives

L.1.1 *Trophic Level and Sizes for the Mercury Water Quality Objectives*

Sport Fish Water Quality Objective

The fish species sampled for comparison with the Sport Fish Objective will vary depending on the fish species in a specific reservoir. The Sport Fish Objective is for top trophic level (TL) species. "Top trophic level fish species" is defined as the species of fish within a reservoir that feed at the top of an aquatic food chain. When trophic level (TL) 4 fish species are present, as is the case for most reservoirs, these fish constitute the "top trophic level" species. Some reservoirs have a shortened food chain and do not support TL4 species. Reservoirs where the top trophic level fish are TL3 are typically high-elevation reservoirs that support rainbow and brook trout as the catchable fish species.

For evaluating compliance with the Sport Fish Objective, monitoring should include representative fish species in TL4 if the objective is for TL4 fish and representative species in TL3 if the objective is in TL3 fish only (no TL4 fish present in the reservoir). A sample is considered either an analytical result from individual fish tissue or a composite of tissue from several fish. Sample sets for comparison with the Sport Fish Objective shall include a range of fish TL3 fish between 150 to 500 millimeters (mm) in total length and TL4 fish between 200 to 500 mm in total length. The objective applies to the wet weight concentration in skinless fillet.

TL4 fish are primarily piscivorous and feed at the top of the aquatic food web. TL3 fish consume TL2 organisms (zooplankton, benthic invertebrates, and some small fish). Table 2.1 identifies fish commonly caught in reservoirs and their trophic level (see section 2.2.1).

Prey Fish and CA Least Tern Objectives

For assessing attainment of the Prey Fish Objective and the Prey Fish Objective for California Least Tern, the objective applies to the wet weight concentration in whole fish between 50 to 150 mm in total length for the Prey Fish Objective and fish less than 50 mm for the Prey fish Objective for the CA Least Tern Objective (see sections 2.2.2 and 2.2.3). A sample is considered either an analytical result from individual fish tissue or a composite of tissue from several fish.

L.1.2 *Sampling Location for the Mercury Water Quality Objectives*

To ensure samples are representative of overall reservoir conditions, sampling locations should not exclude locations know to be affected by local mercury sources (e.g., mercury or gold mines). For monitoring contaminants in fish, the Surface Water Ambient Monitoring Program has developed criteria for determining the number of sampling locations needed in reservoirs with small, medium, and large surface areas (Bonnema 2007). Note that while fish samples are typically reported as being sourced from a precise location, use of a boat or electroshocking equipment may result in fish collected over a wide area.

If a reservoir is sampled at more than one location, data should be combined and analyzed for the reservoir as a whole for assessing compliance with the objectives. This analysis is done by combining data from different locations within a reservoir for calculation of the average fish mercury concentration. Samples from different locations should be evaluated together because fish and consumers can use more than one area of a reservoir.

L.1.3 *Sample Averaging Period and Frequency of Collection for the Mercury Water Quality Objectives*

An averaging period describes the period of time during which risk due to exposure is assessed. For methylmercury, the harmful effects being addressed by the sport fish and small fish objectives are caused by chronic exposure. For the Mercury Water Quality Objectives, the averaging period is one calendar year. Fish samples collected within one calendar year must be grouped in a set for comparison with the mercury objective. Samples collected more than one year apart may not be grouped in a single sample set for comparison with the applicable mercury water quality objective(s).

The frequency of sample collection may be one or more times during the averaging period. For example, analytical results from sport fish collected in May, August, and October within the same calendar may be averaged together. Methylmercury concentrations in sport fish result from methylmercury intake over time. Although aqueous methylmercury concentrations may vary by season within reservoirs, this variation is muted within fish because methylmercury is retained with a fish's body for months (USEPA 2010 pg. 57). Sampling for comparison with the sport fish objective can occur throughout a calendar year.

Within a calendar year, sampling for comparison with the prey fish objectives is limited to coincide with the breeding period of sensitive wildlife species (Eagles-Smith et al. 2014). For the Prey Fish Objective, the time period for sample collection is limited to February 1 through July 31 unless site-specific information indicates another appropriate breeding period for the species of concern. For the California Least Tern Objective, the time period for collection of prey fish is April 1 through August 3. If site-specific information exists indicating another breeding period for any of these species, the time period of collection should be altered to reflect that breeding period and provide appropriate protection for the species of concern.

Periodic fish sampling is recommended to demonstrate continued attainment of the objectives. Because sampling frequency likely depends on availability of resources, sampling on a particular time schedule is not required. A sampling frequency of 10 years or less is recommended.

L.1.4 *Stocked Fish*

Fish that are stocked should be included in sample collection for assessing attainment of the Sport Fish Objective if creel surveys or other information is provided showing that stocked fish are consumed by people and wildlife species of concern. The rationale for including stocked and consumed fish in compliance evaluations is that methylmercury in stocked fish contributes to the totals of methylmercury intake by people and wildlife. Where stocked fish are important for consumption, methylmercury intakes would be overestimated if stocked fish are not included in assessment of attainment.

L.1.5 *Analysis Method and Detection Limit*

Mercury in fish samples can be quantified as total mercury, rather than methylmercury. As noted in section 1.3, mercury exists almost entirely in the methylated form in fish. Per sample cost is greater for methylmercury than mercury. Thus, most fish mercury data is reported as total mercury. There is a slight level of protection gained by assuming the entire concentration of mercury reported is methylmercury.

Fish tissue should be sampled in accordance with USEPA's Standard Method 7473 (USEPA 2007) and the State Water Board's Surface Water Ambient Monitoring Program, Bioaccumulation Oversight Group (SWAMP /BOG) protocol (Bonnema 2014). The method quantitation limit of 0.009 mg/kg wet weight for prey fish and 0.012 mg/kg wet weight for sport fish (see Table 10.1). Fish samples should be collected and processed using the procedures above or by an alternative procedure only if approved in advance by staff of the State Water Board or a Regional Water Board.

L.2 Statistical Analysis of Monitoring Data

This section provides a brief overview and evaluation of methods to evaluate fish methylmercury levels. Methods evaluated here include a weighted average, the binomial distribution, arithmetic mean, and the upper 90th confidence limit of the average. The water quality objectives are specified as average concentrations in fish tissue.

L.2.1 *Weighted Average with Composite Samples*

Fish may be processed for sampling as individuals or composites. Often, multiple fish of the same species are collected and analyzed as a composite sample. Using a weighted average of composite samples is an option for determining compliance with the water quality objectives.

In order to reduce per-sample analytical costs, tissue from two or more fish are sometimes composited into a single sample prior to analysis. In comparison with composites, analysis of samples from individual fish provides more information about the ranges and trends in methylmercury concentrations between individual fish. However, if composites are used, it is recommended that composites be comprised of the same species of fish with a consistent quantity to more accurately compare ranges and trends.

Composite samples must adhere to criteria detailed by the Surface Water Ambient Monitoring Program (Bonnema 2007) or a similarly rigorous protocol to be representative of the fish population being monitored. In particular, the smallest fish in the composite must be at least 75% of the total length of the largest fish (USEPA 2000c pgs. 6-19). Composites must be comprised of fish of the same species collected at a single location (i.e., same reservoir).

For statistical analysis, if composites are to be averaged with concentration data from individual fish, the composite average may be weighted by the number of fish in the sample. Calculation of a weighted average is described in Appendix K and further in the USEPA's Guidance for Implementing the 2001 Methylmercury Criterion (USEPA 2010 pg. 61-63).

L.2.2 *Binomial Distribution*

In California, the binomial distribution is applied to determine whether water bodies do or do not attain water quality standards, in accordance with the Listing Policy. Application of the binomial distribution consists of grouping individual or composite samples of fish into averaged sample sets then counting the number sample sets that exceed a water quality standard (see Figure L.1). For example, if two or more sample sets each containing between two and 24 fish samples exceed the standard, then the water body does not attain the standard and is added to the section 303(d) list. Conversely, to demonstrate attainment of a water quality standard, at least 28 sample sets must be obtained, with no more than two sample sets containing between 28 and 36 fish samples exceeding the standard in order to remove the water body from the section 303(d) list.

L.2.3 *Arithmetic Mean*

The mercury water quality objectives are specified as average methylmercury concentrations in fish tissue within a calendar year. Use of an average (arithmetic mean) is consistent with other efforts at the State and federal levels that evaluate environmental, particularly tissue, samples to

assess mercury risk to consumers. The USEPA's guidance for assessing attainment of fish tissue criteria for methylmercury is written with the expectation that data will be analyzed as averages (USEPA 2010, pg. 59 and 61-63).

The USEPA guidance provides for direct comparison of the average to the criterion for small sample sets that do not support rigorous statistical testing. (Additionally, the guidance provides for use of the t-test for comparison of large, statistically robust data sets to the criterion. As a practical matter, most California reservoir fish data sets are small.) The California Office of Environmental Health Hazard Assessment (OEHHA) typically calculates the arithmetic mean of a set of fish mercury concentration data for use in development of a fish consumption advisory (OEHHA 2008 pg. 59; OEHHA 2009, pg. 3; OEHHA 2014).

"Average methylmercury concentration" is defined as the arithmetic mean calculated from a set of individual values of methylmercury concentrations in fish tissue. (As described in Chapter 1, it is acceptable to use laboratory analytical results reported for total mercury rather than methylmercury.) The first step to determine assessment towards attaining an objective is to collect and group individual methylmercury concentration values for location (i.e., whether one or more locations in a reservoir are appropriate) and date (i.e., each calendar year). The next step is to calculate the arithmetic mean being sure to use proper significant figure protocols (i.e., sum the concentration, round the sum to the least number of decimal places in the data set, then divide by the count, and no rounding needed because the count is an exact number).

Calculation of the arithmetic mean from a set of methylmercury concentration values is a statistical treatment that produces a number that can be compared directly with the objective. Moreover, the number can be compared directly with a data set collected in a different year, thus supporting trend analysis, for use in assessing progress in reducing fish methylmercury levels and progress towards meeting the objective.

However, the numeric average does not describe the certainty with which the population mean is being estimated. Confidence limits are one statistical approach that can be employed to describe certainty associated with estimates.

L.2.4 *Upper 90th Confidence Limit of the Arithmetic Mean [UCL$_{90}$]*

As described in the previous section, use of an arithmetic mean is supported by USEPA's guidance for assessing attainment of fish tissue criteria for methylmercury and OEHHA's practice. For fish tissue data, using the average tissue concentration is appropriate because the risks of mercury are assessed from exposures that occur over time. However, typical reservoir fish data sets have very small sizes (i.e., commonly 9, and generally no more than 30 samples). Large sample sets of 30 individual fish often used in scientific research because of variance in fish MeHg levels (Slotton et al. 2007). Comparatively, SWAMP BOG aims to collect 11 largemouth bass to ensure statistically robust data set. For the purpose of this program, a minimum of nine samples is required. This number is appropriate when sampling events are limited or are not successful, but still provides statistical confidence. These are very small numbers in comparison with reservoir fish populations, and therefore likely do not provide a robust estimate of the true reservoir average. Use of a statistical method that describes the certainty associated with the estimated average would provide a more robust estimate of the true reservoir average.

Comparison with the upper confidence limit of the arithmetic mean takes into account the greater uncertainty that a small sample set is representative of the population and the variability of that data set. Maintaining the water quality objective over time takes into account that fish methylmercury levels vary from year-to-year in response to varying environmental factors. Therefore if the weight of evidence approach is used for either adding or removing a reservoir from the Clean Water Act section 303(d) list, staff proposes that the UCL_{90} of the dataset be used for comparison with the water quality objectives (see Staff Report sections 2.2, L.3.1 and L.3.2). To apply the UCL_{90}, samples are collected and grouped consistent to methods specified in the Mercury Water Quality Objectives (see sections L.1.1 through L.1.3).

The equation to calculate the UCL_{90} is the following:

$$UCL_{90} = \bar{x} + 1.645\left(\frac{\sigma}{\sqrt{n}}\right)$$

Where: $\bar{x}$ = calendar-year mean of samples

σ=standard deviation

$\sqrt{n}$= square root of number of samples

The UCL_{90} is to be rounded to one significant figure and compared directly to the numeric value of the objective. No statistical procedure in addition to calculation of the UCL_{90} is required.

L.3 Site-Specific Weight of Evidence Approach for Listing and Delisting

Clean Water Act section 303(d) requires states to compile a list of waters containing pollutants at levels that exceed protective water quality standards. For California, the State Water Board uses the Listing Policy for this identification process. This report does not modify or supersede the Listing Policy, but does identify an alternative process that can be considered under a weight of evidence approach when evaluating listing and delisting factors for reservoir fish methylmercury. If application of the Listing Policy Factors in sections 3.5 or 4.5 does not result in the listing or delisting of a water body, this section describes an alternative assessment procedure to determine if reservoirs achieve or do not achieve the water quality objectives using the weight of evidence approach described in the Listing Policy Factors 3.11 and 4.11.

L.3.1 *Determination of Water Quality Standards Non-attainment for Methylmercury in Reservoirs (Listing)*

California Listing Factors are provided in section 3 of the Listing Policy. Methylmercury is a bioaccumulative pollutant and the data are for concentrations in fish tissue. Hence, the first listing factor to consider is Factor 3.5, Bioaccumulation of Pollutants in Aquatic Life Tissue.

The Listing Policy's Factor 3.5 directs that data be evaluated in comparison with a pollutant-specific evaluation guideline. Fish mercury data would be compared to the water quality objectives for sport fish (150 – 500 mm TL3 and 200 – 500 mm TL4), prey fish (50 – 150 mm), and CA least tern (< 50 mm).

Typically, fish are collected infrequently and on one date. Listing Policy section 6.1.5.6 requires that when "multiple samples are collected at a single location on the same day, the measurement shall be combined and represented by a single resultant value". Therefore if many fish were collected all on one date, the Listing Policy would require all of that data be combined (averaged), resulting in one data result. However, at least two data results are needed for determination of the attainment of standards using the binomial distribution. Thus, in many cases there would be no assessment due to insufficient data from infrequent monitoring.

When a listing decision cannot be made using the Listing Policy Factors 3.1 – 3.10, Factor 3.11 provides for an evaluation of Situation-Specific Weight of Evidence Listing Factors. This Listing Factor states:

> When all other Listing Factors do not result in the listing of a water segment but information indicates non-attainment of standards, a water segment shall be evaluated to determine whether the weight of evidence demonstrates that a water quality standard is not attained.

In situations where Listing Policy Factors 3.1 – 3.10 do not result in a listing determination, but the data suggests the water body may not meet the standards (for example, only one data result exists and it exceeds water quality criteria), data must be evaluated by Listing Policy Factor 3.11.

Situation-Specific Weight of Evidence Approach: Figure L.1 provides a flow chart for the proposed weight of evidence assessment of reservoir fish mercury objectives. Data grouped according to the Statewide Water Quality Objectives (see sections L.1.1 – L.1.3) would be compared with the applicable numeric objective using the statistical average (upper 90th confidence limit of the average [UCL_{90}]) described in section L.2.4. At least nine samples are required using this approach. A sample is considered either an analytical result from individual fish tissue or a composite of tissue from several fish. Reservoirs would be placed on the section 303(d) list if the upper 90th confidence limit of the arithmetic mean exceeds the applicable water quality objective. Accordingly, one set of samples containing nine or more individual or composite fish and grouped by the Mercury Water Quality Objectives (see sections L.1.1 – L.1.3) that exceed either the sport or prey fish objectives is sufficient to determine that the reservoir should be placed on the section 303(d) List.

L.3.2 *Determination of Water Quality Standards Attainment for Methylmercury in Reservoirs (Delisting)*

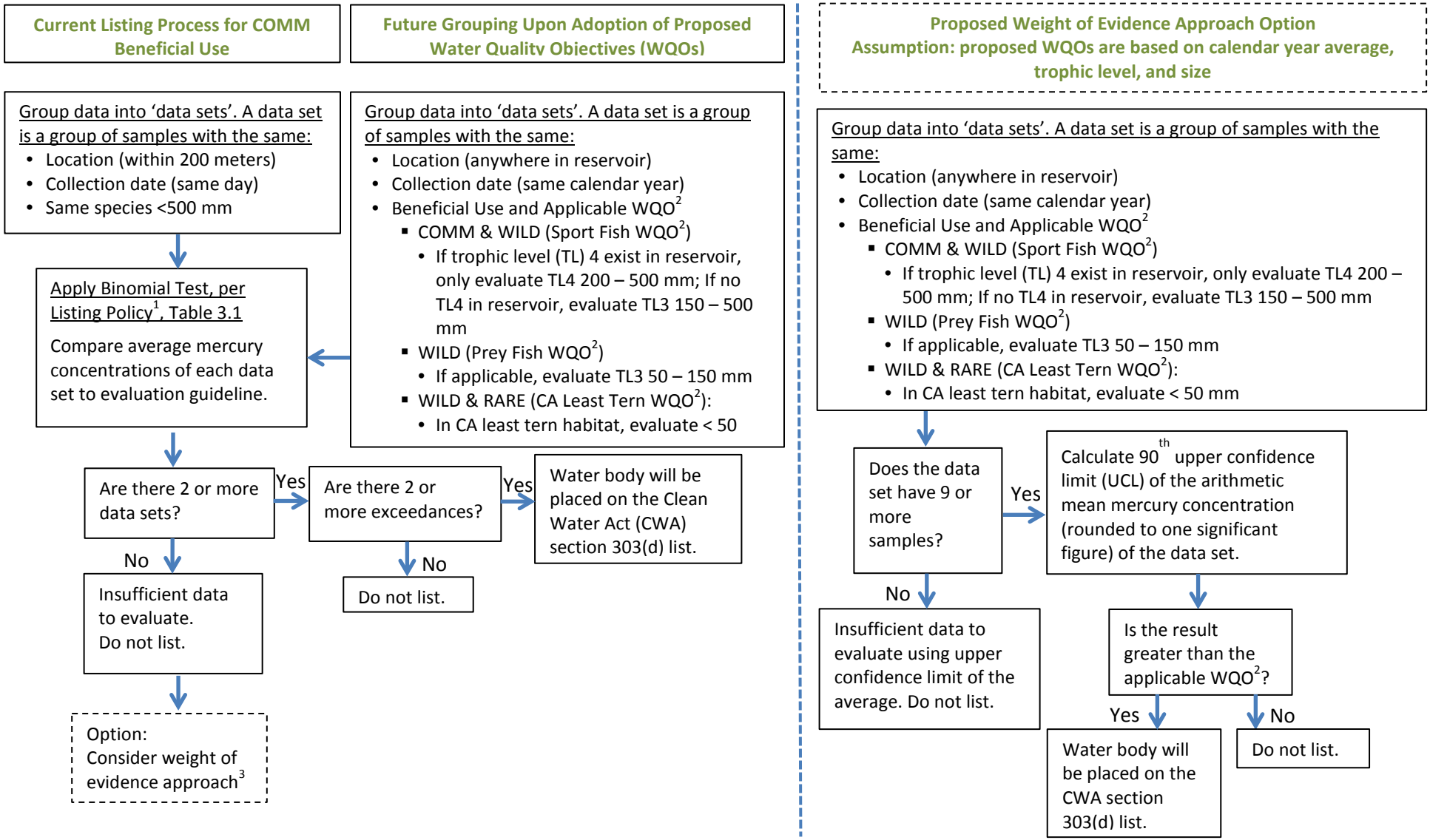
Delisting means removing waters from the section 303(d) List. California Delisting Factors are provided in section 4 of the Listing Policy. Methylmercury is a bioaccumulative pollutant, and the data are for concentrations in fish tissue. Hence, the first delisting factor to consider is Listing Policy Factor 4.5, Bioaccumulation of Pollutants in Aquatic Life Tissue. Listing Policy Factor 4.5 refers to the use of the binomial distribution method, described in Listing Policy Factor 4.1. However, Listing Policy Factor 4.1 indicates that the binomial distribution cannot be used to support a delisting with a sample size less than 28.

Since the data in any one averaging period are aggregated and counted as one sample, it would require 28 annual (calendar-year) sample events to demonstrate compliance with an

objective that is based on annual averages (i.e., 28 years of data collected annually). Because of the extensive amount of time required to make a delisting evaluation, staff propose the data be evaluated according to Listing Policy Factor 4.11, Situation-Specific Weight of Evidence Delisting Factor and compared to the numeric sport and prey fish objectives. Given the 28 years need to make an assessment determination, it is anticipated that most fish data will be evaluated by this situation-specific weight of evidence approach.

Figure L.1 in this appendix provides a flow chart showing application of the Listing Policy and the weight of evidence assessment for the reservoir fish mercury objectives. A reservoir may be determined to meet mercury standards when individual averages, using the upper 90th confidence limit of the mean method (described in section L.2.4), are equal to or less than the applicable water quality objective(s) over three consecutive annual (calendar-year) sampling events. Consecutive annual sampling events do not need to occur during consecutive calendar years, but they do need to occur in succession. Evaluation must occur within 10 years and with no intervening exceedance of an objective.

Figure L.1: Flow Chart for Weight of Evidence Assessment of Reservoir Fish for Mercury Objectives

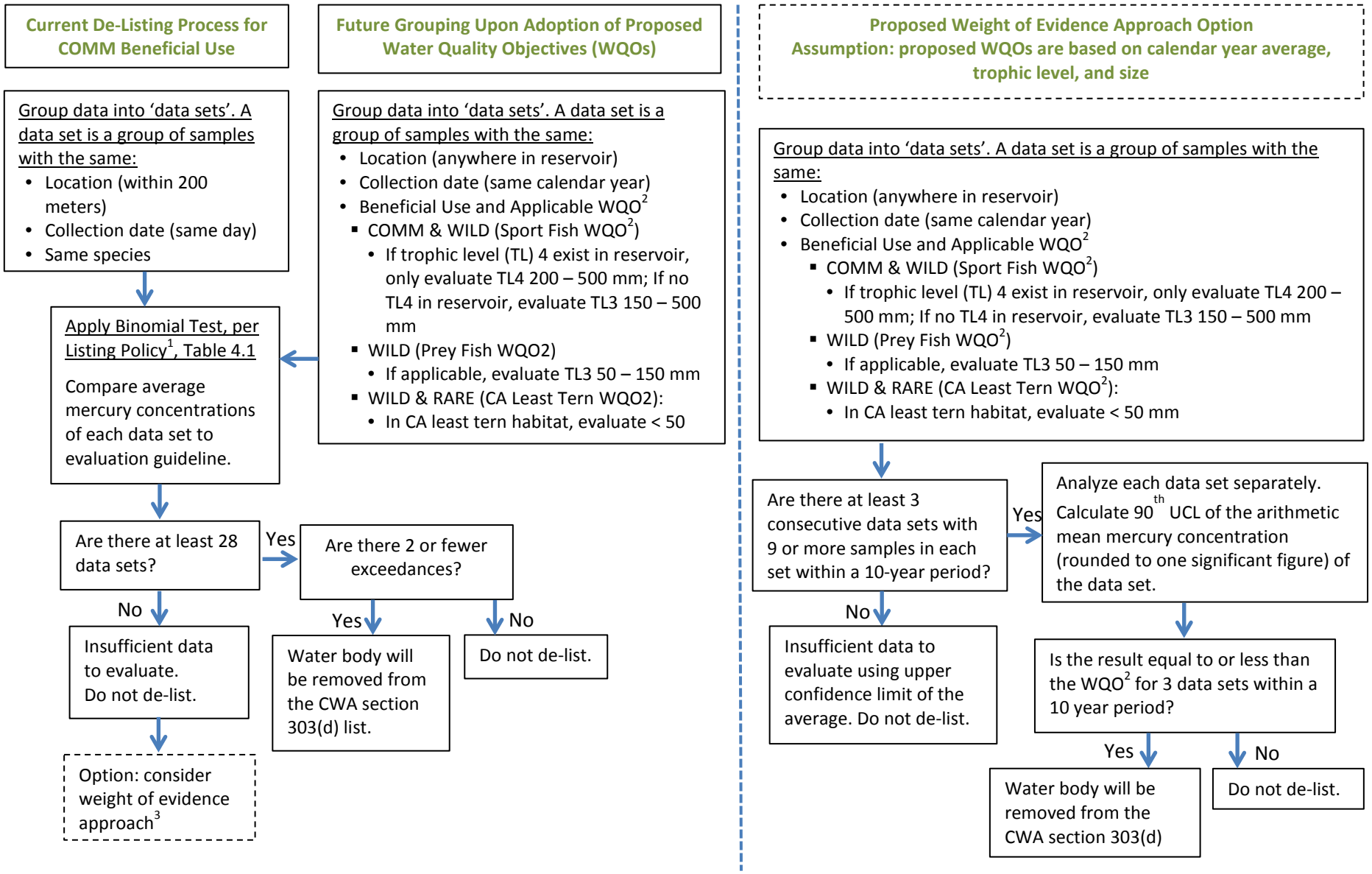


[1] State Water Resources Control Board Water Quality Control Policy for Developing California's Clean Water Act Section 303(d) List (SWRCB 2004).

[2] See Figure 2.1 for applicable Water Quality Targets (objectives) to evaluate.

[3] A situation-specific weight of evidence approach may be used when all other Listing or Delisting Factors do not result in the listing or delisting of a water segment, but information indicates non-attainment or attainment of standards. Examples of weight of evidence approaches used by Regional Water Boards:

- Region 5, 2014 weight of evidence: Group data by location (within 200 meters), collection date (same day), and species <500 mm. If only 1 data set and it exceeds criteria, evaluate individual sample results. List if 2 samples exceed criteria.
- Region 1, 2012 weight of evidence: Assess all samples individually and apply binomial test. List if 2 samples exceed criteria.



[1] State Water Resources Control Board Water Quality Control Policy for Developing California's Clean Water Act Section 303(d) List (SWRCB 2004).

[2] See Figure 2.1 for applicable Water Quality Targets (objectives) to evaluate.

[3] A situation-specific weight of evidence approach may be used when all other listing or delisting factors do not result in the listing or delisting of a water segment, but information indicates non-attainment or attainment of standards. Examples of weight of evidence approaches used by Regional Water Boards:

- Region 5, 2014 weight of evidence: Group data by location (within 200 meters), collection date (same day), and species <500 mm. If only 1 data set and it exceeds criteria, evaluate individual sample results. List if 2 samples exceed criteria.
- Region 1, 2012 weight of evidence: Assess all samples individually and apply binomial test. List if 2 samples exceed criteria.