Methodology for the Development

of the

2010 Section 303(d) List in Missouri

Missouri Department of Natural Resources
Division of Environmental Quality
Water Protection Program

Approved by the Missouri Clean Water Commission, May 6, 2009

Missouri
Department of
Natural Resources

Methodology for the Development of the 2008 Section 303(d) List

Table of Contents

I. Citation and Requirements 2
- A. Citation of Section of Clean Water Act..... 2
- B. U.S. EPA Guidance..... 2
  - Placement of Waters within the Five 305(b) Report Categories..... 3

II. Methodology Document..... 6
- A. Procedures and Methods Used to Collect Water Quality Data..... 6
- B. Identification of All Existing and Readily Available Water Quality Data Sources..... 10
- C. Data Quality Considerations..... 13
- D. How Water Quality Data is Evaluated to Determine Whether or Not Waters are Impaired..... 15
- E. 303(d) Listing Considerations..... 21

List of Appendices

Appendix A: Excerpts from EPA Guidance on Statistical Approaches..... 23
Appendix B: Statistical Considerations..... 26
Appendix C: Examples of Statistical Procedures..... 33
Appendix D: The Meaning of the Sediment Quotient and How to Calculate It..... 34

Tables

Table 1.1: Methods for Assessing Compliance with Water Quality Standards: Numeric Criteria..... 16
Table 1.2: Methods for Assessing Compliance with Water Quality Standards: Narrative Criteria..... 19
Table B-1: Description of Analytical Tools for Determining if Waters are Impaired..... 26
Table B-2 Description of Analytical Tools for Determine When Waters are No Longer Impaired..... 29
Table B-3: Effects of Type 1 Error Rates and Sample Size on Type 2 Error Rates..... 31

I. Citation and Requirements

A. Citation of Section of Clean Water Act

This document is required by revisions of rules under the Federal Clean Water Act, Section 303(d), 40 CFR 130.7, and the timetable for presenting the finished document to the United States Environmental Protection Agency (EPA) and the public is given in Part 130.10. Section 303(d) requires states to list certain impaired waters and the rules require that states describe how this list will be constructed. Missouri fulfills reporting requirements under Section 303(d), 305(b) and 314 of the Clean Water Act by the submission to EPA of an integrated report at the time the 303(d) is approved by the Missouri Clean Water Commission. In years when no integrated report is submitted, the Department of Natural Resources (Department) submits a copy of its statewide water quality assessment database to EPA.

B. EPA Guidance

In July 2003, EPA issued new guidance entitled "Guidance for 2004 Assessment, Listing and Reporting Requirements Pursuant to Sections 303(d) and 305(b) of the Clean Water Act". This guidance gives further recommendations about listing of 303(d) and other waters. In July 2005, EPA published an amended version entitled "Guidance for 2006 Assessment, Listing and Reporting Requirements Pursuant to Sections 303(d), 305(b) and 314 of the Clean Water Act." In October 2006, EPA issued a memorandum entitled "Information Concerning 2008 Clean Water Act Sections 303(d), 305(b) and 314 Integrated Reporting and Listing Decisions." This memorandum serves as EPA's guidance for the 2008 reporting cycle.

The Department is responsible for administration of the Federal Clean Water Act in Missouri. EPA regulations require that the Department describe the methodology used to develop the state's 303(d) List. This draft document should be made available to the public for review and comment. The Department should provide EPA with a document summarizing all comments received and the Department responses to significant comments. EPA's guidance recommends that the Department provide: (1) a description of the methodology used to develop the Section 303(d) List; (2) a description of the data and information used to identify (impaired and threatened) waters, including a description of the existing and readily available data and information used; and (3) a rationale for any decision for not using any existing and readily available data and information. The guidance also notes that "prior to submission of its Integrated Report, each state should provide the public with the opportunity to review and comment on the methodology." The guidelines further recommend that the methodology document include information on how interstate or international disagreements concerning the list are resolved.

Placement of Waters within the Five Categories in the 2006 EPA Assessment, Listing and Reporting Guidance

The guidance issued by EPA in 2005 recommends that all waters of the state be placed in one of five categories.

Category 1

All designated beneficial uses are fully maintained. Data or other information supporting full beneficial use attainment for all designated beneficial uses must be consistent with the state's listing methodology document. The Department will place a water in Category 1 if the following conditions are met:

- The water has physical and chemical data (at a minimum, water temperature, pH, dissolved oxygen and ammonia for streams, and total nitrogen, total phosphorus and secchi depth for lakes) and biological water quality data (at a minimum, *E. coli* or fecal coliform bacteria) that indicates attainment with water quality standards.
- The level of mercury in fish fillets or fish eggs used for human consumption does not exceed fish tissue guidelines of 0.3 mg/kg or less. Where at least three samples are available for higher trophic level species (largemouth, smallmouth and Kentucky Spotted bass, sauger, walleye, northern pike, trout, striped bass, white bass, flathead catfish and blue catfish), only those samples will be used.
- The water is not rated as "threatened".

Category 2

One or more designated beneficial uses are fully attained but at least one designated beneficial use has inadequate data or information to make a use attainment decision consistent with the state's listing methodology document. The Department will place a water in Category 2 if at least one of the following conditions are met:

- There is inadequate data for water temperature, pH, dissolved oxygen or ammonia in streams to assess attainment with water quality standards or inadequate total nitrogen, total phosphorus or secchi data in lakes.
- There is inadequate *E. coli* or fecal coliform bacteria data to assess attainment with the whole body contact recreational use.
- There is insufficient fish fillet tissue or fish egg data available for mercury to assess attainment with the fish consumption use.

Category 2 waters will be placed in one of two sub-categories.

Category 2A: Waters will be placed in this category if available data, using best professional judgement, suggests compliance with numeric water quality criteria of Tables A or B in

Missouri's Water Quality Standards (10 CSR 20-7.031) or other quantitative thresholds for determining use attainment.

Category 2B: Waters will be placed in this category if the available data, using best professional judgment, suggests noncompliance with numeric water quality criteria of Tables A or B in Missouri's Water Quality Standards, or other quantitative thresholds for determining use attainment, and this data is insufficient to support a statistical test or to qualify as representative data. Category 2B waters will be given high priority for additional water quality monitoring.

Category 3

Water quality data are not adequate to assess any of the designated beneficial uses consistent with the LMD. The Department will place a water in Category 3 if data are insufficient to support a statistical test or to qualify as representative data to assess any of the designated beneficial uses. Category 3 waters will be placed in one of two sub-categories.

Category 3A. Waters will be placed in this category if available data, using best professional judgement, suggests compliance with numeric water quality criteria of Tables A or B in Missouri's Water Quality Standards (10 CSR 20-7.031) or other quantitative thresholds for determining use attainment.

Category 3B. Waters will be placed in this category if the available data, using best professional judgement, suggests noncompliance with numeric water quality criteria of Tables A or B in Missouri's Water Quality Standards or other quantitative thresholds for determining use attainment. Category 3B waters will be given high priority for additional water quality monitoring.

Category 4

State Water Quality Standards or other criteria, as per the requirements of Table 1 of this document, are not attained, but a Total Maximum Daily Load study is not required. Category 4 waters will be placed in one of three sub-categories.

Category 4A. EPA has approved a Total Maximum Daily Load study that addresses the impairment. The Department will place a water in Category 4A if both the following conditions are met:

- Any portion of the water is rated as being in non-attainment with state Water Quality Standards or other criteria as explained in Table 1 of this document due to one or more discrete pollutants or discrete properties of the water[1], and
- EPA has approved a Total Maximum Daily Load for all pollutants causing that non-attainment.

[1] A discrete pollutant or a discrete property of water is defined here as a specific chemical or other attribute of the water (such as temperature, dissolved oxygen or pH) that causes beneficial use impairment and that can be measured quantitatively.

Category 4B. Water pollution controls required by a local, state or federal authority, are expected to correct the impairment in a reasonable period of time. The Department will place a water in Category 4B if both of the following conditions are met:

- Any portion of the water is rated as being in non-attainment with state Water Quality Standards or other criteria as explained in Table 1 of this document due to one or more discrete pollutants or discrete properties of water, and
- A water quality based permit that addresses the pollutant(s) causing the designated use impairment has been issued and compliance with the permit limits will eliminate the impairment; or other pollution control requirements have been made that are expected to adequately address the pollutant(s) causing the impairment. This may include implemented voluntary watershed control plans as noted in EPA's guidance document.

Category 4C. Any portion of the water is rated as being in non-attainment with state Water Quality Standards or other criteria as explained in Table 1 of this document, and a discrete pollutant(s) or other discrete property of the water does not cause the impairment. Discrete pollutants may include specific chemical elements (e.g., lead, zinc), chemical compounds (e.g., ammonia, dieldrin, atrazine) or one of the following quantifiable physical, biological or bacteriological conditions: water temperature, percent of gas saturation, amount of dissolved oxygen, pH, deposited sediment, toxicity or counts of fecal coliform or *E. coli* bacteria.

Category 5

At least one discrete pollutant has caused non-attainment with state Water Quality Standards or other criteria as explained in Table 1 of this document, and the water does not meet the qualifications for listing as either Categories 4A or 4B. Category 5 waters are those that are candidates for the state's 303(d) List[2].

If a designated use is not supported and the segment is impaired or threatened, the fact that a specific pollutant is not known does not provide a basis for excluding a segment from Category 5. These segments must be listed as Category 5 unless the state can demonstrate that no discrete pollutant or pollutants causes or contributes to the impairment. Pollutants causing the impairment will be identified before a TMDL study is written. The TMDL must be written within the time period allowed for TMDL development in EPA guidelines.

Threatened Waters

When a water that would otherwise be in Categories 1, 2 or 3 has a time trend analysis for one or more discrete water quality pollutants that indicates the water is currently maintaining all beneficial uses but will not continue to meet these uses before the next listing cycle, it will be considered a "threatened water." A threatened water will be treated as an impaired water and placed in the appropriate Category (4A, 4B or 5).

[2] The proposed state 303(d) List is determined by the Missouri Clean Water Commission and the final list is determined by the U.S. Environmental Protection Agency.

II. The Methodology Document

A. Procedures and Methods Used to Collect Water Quality Data

Department Monitoring

The major purposes of the Department's water quality monitoring program are:

- to characterize background or reference water quality conditions;
- to better understand daily, flow event and seasonal water quality variations and their underlying processes;
- to characterize aquatic biological communities;
- to assess time trends in water quality;
- to characterize local and regional impacts of point and nonpoint source discharges on water quality;
- to check for compliance with Water Quality Standards or wastewater permit limits;
- to support development of strategies, including Total Maximum Daily Loads, to return impaired waters to compliance with Water Quality Standards. All of these objectives are statewide in scope.

Coordination with Other Monitoring Efforts in Missouri

To maximize efficiency, the Department routinely coordinates its monitoring activities to avoid overlap with other agencies and to provide and receive interagency input on monitoring study design. Data from other sources is used for meeting the same objectives as Department sponsored monitoring. The agencies most often involved are the U.S. Geological Survey, the U.S. Army Corps of Engineers, EPA, the Missouri Department of Conservation, and the Missouri Department of Health and Senior Services. The Department also tracks the monitoring efforts of the National Park Service, the U.S. Forest Service, several of the state's larger cities, the states of Oklahoma, Arkansas, Kansas, Iowa and Illinois, and graduate level research conducted at universities within Missouri. For those wastewater discharges where the Department has required instream water quality monitoring, the Department may also use monitoring data acquired by wastewater dischargers as a condition of discharge permits issued by the department. In 1995, the Department also began using data collected by volunteers that have passed Quality Assurance/Quality Control tests.

Existing Monitoring Networks and Programs

The following list is a description of the kinds of water quality monitoring activities presently occurring in Missouri.

1. Fixed Station Network

   A. Objective: To better characterize background or reference water quality conditions, to better understand daily, flow event and seasonal water quality variations and their

underlying processes, to assess time trends and to check for compliance with Water Quality Standards.

B. Design Methodology: Sites were chosen based on one of the following criteria:

- Site is believed to have water quality representative of many neighboring streams of similar size due to similarity in watershed geology, hydrology and land use, and the absence of any impact from a significant point or discrete nonpoint water pollution source.
- Site is downstream of a significant point source or discrete nonpoint source area.

C. Number of Sites, Sampling Methods, Sampling Frequency, and Parameters:

- Department/U.S. Geological Survey cooperative network: 60 sites statewide, horizontally and vertically integrated grab sampled, six to 12 times per year. Samples are analyzed for major ions, nutrients, temperature, pH, dissolved oxygen, specific conductance and flow on all visits, two to four times annually for suspended solids and heavy metals, and for pesticides six times annually at six sites.
- Department raw water sampling of public drinking water reservoirs: nine drinking water reservoirs are sampled 4 four times per year for some commonly used agricultural herbicides.
- Department/University of Missouri-Columbia's lake monitoring network. This program has monitored about 185 lakes. About 40 lakes are monitored each year. Each lake is usually sampled four times during the summer and about 12 are monitored spring through fall for nutrients, chlorophyll, turbidity and suspended solids.
- Department routine monitoring of finished public drinking water supplies for bacteria and trace contaminants.
- Routine bacterial monitoring (typically weekly during the summer) of swimming beaches at Missouri's state parks during the recreational season by the Department's Division of State Parks.
- Monitoring of sediment quality by the Department at approximately 10 discretionary sites annually. All sites are monitored for several heavy metals and organic contaminants. A pore water sample is analyzed for ammonia, and a Microtox toxicity test is performed on the sediment.

2. Special Water Quality Studies

A. Objective: Special water quality studies are used to characterize the water quality impacts from a specific pollutant source area.

B. Design Methodology: These studies are designed to determine the contaminants of concern based on previous water quality studies, effluent sampling and/or Missouri State Operating Permit applications. These studies employ multiple sampling stations downstream and upstream (if appropriate). If contaminants of concern have significant

seasonal or daily variation, season of the year and time of day variation must be accounted for in the sampling design.

C. Number of Sites, Sampling Methods, Sampling Frequency and Parameters: The Department conducts or contracts for 10 to15 special studies annually. Each study has multiple sampling sites. Number of sites, sampling frequency and parameters all vary greatly depending on the study. Intensive studies would also require multiple samples per site over a relatively short time frame.

3. Toxics Monitoring Program

The fixed station network and many of the Department's intensive studies monitor for toxic chemicals. In addition, major municipal and industrial dischargers must monitor for toxicity in their effluents as a condition of their Missouri State Operating Permit.

4. Biological Monitoring Program

A. Objectives: The objectives of this program are to develop numeric criteria describing "reference" aquatic macroinvertebrate and fish communities in Missouri's streams, to implement these criteria within state Water Quality Standards and to continue a statewide fish and aquatic invertebrate monitoring program.

B. Design Methodology: Development of biocriteria for invertebrates and fish involves identification of reference streams in each of Missouri's 17 ecological drainage units. It also includes intensive sampling of invertebrate and fish communities to quantify temporal and spatial variation in reference streams within ecoregions and variation between ecoregions, and the sampling of chemically and physically impaired streams to test sensitivity of various community metrics to differences in stream quality.

C. Number of Sites, Sampling Methods, Sampling Frequency and Parameters: The Department has conducted biological sampling of aquatic invertebrates for many years. Since 1991, this program has consisted of standardized monitoring of approximately 55 sites twice annually. The Missouri Department of Conservation presently has a statewide fish and aquatic invertebrate monitoring program, the Resource Assessment and Monitoring Program, designed to assess and monitor the health of Missouri's stream resources. This program samples a minimum of 450 random and 30 reference sites every five years.

5. Fish Tissue Monitoring Program

A. Objective: Fish tissue monitoring can address two separate objectives. These are: (1) the assessment of ecological health or the health of aquatic biota (usually accomplished by monitoring whole fish samples); and (2) the assessment of human health risk based on the level of contamination of fish fillets or fish eggs.

B. Design Methodology: Fish tissue monitoring sites were chosen based on one of the following criteria:

- Site is believed to have water and sediment quality representative of many neighboring streams or lakes of similar size due to similarity in geology, hydrology and land use, and the absence of any known impact from a significant point source or discrete nonpoint water pollution source.
- Site is downstream of a significant point source or discrete nonpoint source area.
- Site has shown fish tissue contamination in the past.

C. Number of Sites, Sampling Methods, Sampling Frequency and Parameters:

The Department and EPA have a cooperative fish tissue monitoring program that collects whole fish composite samples[3] at approximately 12 fixed sites. Each site is sampled once every two years. The preferred species for these sites are either carp or redhorse sucker.

The Department, EPA and the Missouri Department of Conservation also sample 40 to 50 discretionary sites annually for two fish fillet composite samples. One sample is of a top carnivore such as largemouth bass, smallmouth bass, walleye or sauger. The other sample is for a species of a lower trophic level such as catfish, carp or sucker. This program occasionally samples fish eggs for certain fish species at selected locations. Both of these monitoring programs analyze for several chlorinated hydrocarbon insecticides, PCBs, lead, cadmium, mercury and fat content.

6. Volunteer Monitoring Program

Two major volunteer monitoring programs are now generating water quality data in Missouri. The first is the Lakes of Missouri Volunteer Program. This cooperative program consists of persons from the Department, the University of Missouri-Columbia and volunteers that monitor approximately 50 lakes, including Lake Taneycomo, Table Rock Lake and several lakes in the Kansas City area. Data from this program is used by the university as part of a long-term study on the limnology of midwestern reservoirs.

The second program involves volunteers who monitor water quality of streams throughout Missouri. The Volunteer Water Quality Monitoring Program is a subprogram of the Missouri Stream Team Program, a cooperative project sponsored by the Department, the Missouri Department of Conservation and the Conservation Federation of Missouri. By the end of 2006, almost 3,800 citizen volunteers had attended at least one training workshop. After the introductory class, many proceed on to at least one more class of higher level training: Levels 1, 2, 3 and 4. Each level of training is a prerequisite for the next higher level, as is appropriate data submission. Data generated by Levels 2, 3 and 4 and the new Cooperative Site Investigation Program volunteers represent increasingly higher quality assurance. Of those completing an introductory course, about 40 percent proceed to Levels 1 and 2. Eighty-two volunteers have reached Level 3 and six volunteers have reached Level 4.

[3] A composite sample is one in which several individual fish are combined to produce one sample.

The Cooperative Site Investigation Program uses trained volunteers to collect samples and transport them to laboratories approved by the Department. Volunteers and Department staff work together to develop a monitoring plan. Currently there are 11 volunteers qualified to work in the Cooperative Site Investigation Program.

Laboratory Analytical Support

Laboratories used:

- Department/U.S. Geological Survey Cooperative Fixed Station Network: U.S. Geological Survey Lab, Denver, Colorado
- Department's Public Drinking Water Reservoir Network: Department's Environmental Services Program
- Intensive Surveys: Varies, many are done by the Department's Environmental Services Program
- Toxicity Testing of Effluents: Many commercial laboratories
- Biological Criteria for Aquatic Invertebrates: Department's Environmental Services Program and University of Missouri-Columbia
- Fish Tissue: EPA Region VII Laboratory, Kansas City, Kansas and miscellaneous contract laboratories (Missouri Department of Conservation)
- Missouri State Operating Permit: Self-monitoring or commercial laboratories
- Department's Public Drinking Water Monitoring: Department's Environmental Services Program and commercial laboratories
- Other water quality studies: Many commercial laboratories

B. Identification of All Existing and Readily Available Water Quality Data Sources:

The following data sources are used by the Department to aid in the compilation of the state's 305(b) Report. Where quality assurance programs are deemed acceptable, these sources would also be used to develop the state's Section 303(d) List. These sources presently include but are not limited to:

1. Fixed station water quality and sediment data collected and analyzed by the Department's Environmental Services Program personnel.
2. Fixed station water quality data collected by the U.S. Geological Survey under contractual agreements with the Department.
3. Fixed station water quality data collected by the U.S. Geological Survey under contractual agreements to agencies or organizations other than the Department.
4. Fixed station water quality, sediment quality and aquatic biological information collected by the U.S. Geological Survey under their National Stream Quality Accounting Network and the National Water Quality Assessment Monitoring Programs.

5. Fixed station raw water quality data collected by the Kansas City Water Services Department, the St. Louis City Water Company, the Missouri American Water Company (formerly St. Louis County Water Company), Springfield City Utilities and Springfield's Department of Public Works.

6. Fixed station water quality data collected by the U.S. Army Corps of Engineers. The Kansas City, St. Louis and Little Rock Corps Districts have monitoring programs for Corps-operated reservoirs in Missouri.

7. Fixed station water quality data collected by the Arkansas Department of Environmental Quality, the Kansas Department of Health and Environment, the Iowa Department of Natural Resources, and the Illinois Environmental Protection Agency.

8. Fixed station water quality monitoring by corporations.

9. Annual fish tissue monitoring programs by the Environmental Protection Agency/Department Regional Ambient Fish Tissue Monitoring Program and the Missouri Department of Conservation.

10. Special water quality surveys conducted by the Department. Most of these surveys are focused on the water quality impacts of specific point source wastewater discharges. Some surveys are of well-delimited nonpoint sources such as abandoned mined lands. These surveys often include physical habitat evaluation and monitoring of aquatic invertebrates as well as water chemistry monitoring.

11. Special water quality surveys conducted by U.S. Geological Survey, including but not limited to:

    a) Geology, hydrology and water quality of various hazardous waste sites,

    b) Geology, hydrology and water quality of various abandoned mining areas,

    c) Hydrology and water quality of urban nonpoint source runoff in St. Louis, Kansas City and Springfield, Missouri, and

    d) Bacterial and nutrient contamination of streams in southern Missouri.

12. Special water quality studies by other agencies such as the Missouri Department of Conservation, the U.S. Public Health Service, and the Missouri Department of Health and Senior Services.

13. Monitoring of fish occurrence and distribution by the Missouri Department of Conservation.

14. Fish Kill and Water Pollution Investigations Reports published by the Missouri Department of Conservation.

15. Selected graduate research projects pertaining to water quality and/or aquatic biology.

16. Water quality, sediment and aquatic biological data collected by the Department, the Environmental Protection Agency or their contractors at hazardous waste sites in Missouri.

17. Self-monitoring of receiving streams by cities, sewer districts and industries, or contractors on their behalf, for those discharges that require this kind of monitoring. This monitoring includes chemical and sometimes toxicity monitoring of some of the

larger wastewater discharges, particularly those that discharge to smaller streams and have the greatest potential to affect instream water quality.

18. Compliance monitoring of receiving waters by the Department and EPA. This can include chemical and toxicity monitoring.

19. Bacterial monitoring of streams and lakes by county health departments, community lake associations and other organizations using acceptable analytical methods.

20. Other monitoring activities done under a quality assurance project plan approved by the Department.

21. Fixed station water quality and aquatic invertebrate monitoring by volunteers who have successfully completed the Volunteer Water Quality Monitoring Program Level 2 workshop. Data collected by volunteers who have successfully completed a training Level 2 workshop is considered to be Data Code One. Data generated from Volunteer Training Levels 2, 3 and 4 are considered "screening" level data and can be useful in providing an indication of a water quality problem. For this reason, the data is eligible for use in distinguishing between waters in Categories 2A and 2B or Categories 3A and 3B. Most of this data is not used to place waters in main Categories (1, 2, 3, 4 and 5) because analytical procedures do not use EPA or Standard Methods approved methods. Data from volunteers who have not yet completed a Level 2 training workshop do not have sufficient quality assurance to be used for any assessment purposes. Data generated by volunteers while participating in the Department's Cooperative Site Investigation Program (Section II C1) or other volunteer data that otherwise meets the quality assurance outlined in Section II C2 can be used in the Section 303(d) assessment process.

The following data sources (22-25) cannot be used rate a water as impaired (Categories 4A, 4B, 4C or 5); however, these data sources may be used to direct additional monitoring that would allow a water quality assessment for Section 303(d) listing purposes.

22. Fish Management Basin Plans published by the Missouri Department of Conservation.

23. Fish Consumption Advisories published annually by the Missouri Department of Health and Senior Services. Note: the department may use data from date source No. 9 to list individual waters as impaired due to contaminated fish tissue.

24. Self-monitoring of wastewater by cities, sewer districts and industries, or contractors on their behalf, that have significant wastewater discharges. This monitoring includes chemical and sometimes toxicity monitoring of some of the larger wastewater discharges, particularly those that discharge to smaller streams and have the greatest potential to effect instream water quality.

25. Compliance monitoring of wastewater by the Department and the Environmental Protection Agency. This can include chemical and toxicity monitoring.

The Department will review all data of acceptable quality that is submitted to the Department prior to the end of the first public notice of the draft 303(d) list. The Department reserves the right to review and use data of acceptable quality submitted after this date if the data results in a change to the assessment status of the water.

C. Data Quality Considerations

1. DNR Quality Assurance/Quality Control Program

   The Department and EPA Region VII have completed a Total Quality Management Plan. All environmental data generated directly by the Department, or through contracts funded by the Department, or EPA require a Quality Assurance Project Plan. The agency or organization responsible for collection and/or analysis of the environmental sampling must write and adhere to a Quality Assurance Project Plan approved through the Department's Total Quality Management Plan. Any environmental data generated by a monitoring plan with a Department approved Quality Assurance Project Plan is considered suitable for use in the 303(d) assessment process. This includes data generated by volunteers participating in the department's Cooperative Site Investigation Program. Under this program, the Department's Environmental Services Program will audit selected non-profit (governmental and university) laboratories. Laboratories that pass this audit will be approved for the Cooperative Site Investigation Program. Individual volunteers that collect samples and deliver them to an approved laboratory must first successfully complete Department training in proper collection and handling of samples.

2. Other Quality Assurance/Quality Control Programs

   Data generated in the absence of a Department-approved Quality Assurance Project Plan may be used to determine the 303(d) status of a water if the Department determines that the data is scientifically defensible after making a review of the quality assurance procedures used by the data generator. This review would include: (1) names of all persons involved in the monitoring program, their duties and a description of training and work related experience, (2) all written procedures, Standard Operating Procedures, or Quality Assurance Project Plans pertaining to this monitoring effort, (3) a description of all field methods used, brand names and model numbers of any equipment and a description of calibration and maintenance procedures, and (4) a description of laboratory analytical methods. This review may also include an audit by the Department's Environmental Services Program.

3. Other Data Quality Considerations

   3.1 Data Age. For assessing present conditions, more recent data is preferable; however, older data can be used to assess present conditions if the data remains representative of present conditions.

   If the department uses data to make a 303(d) List decision that predates the date the list is

initially developed by more than seven years, the Department will provide a written justification for the use of such data.

A second consideration is the age of the data relative to significant events that may have an effect on water quality. Data collected prior to the initiation, closure or significant change in a wastewater discharge, or prior to a large spill event or the reclamation of a mining or hazardous waste site, for example, may not be representative of present conditions. Such data would not be used to assess present conditions even if it was less than seven years old. Such "pre-event" data can be used to determine changes in water quality before and after the event or to show water quality time trends.

3.2 Data Type, Amount and Information Content. EPA recommends establishing a series of data codes, and rating data quality by the kind and amount of data present at a particular location (EPA 1997[4]). The codes are single digit numbers from one to four, indicating the relative degree of assurance the user has in the value of a particular environmental data set. Data Code One indicates the least assurance or the least number of samples or analytes and Data Code Four the greatest. Based on EPA's guidance, the Department uses the following rules to assign code numbers to data.

Data Code[5] One: All data not meeting the requirements of Data Code Two, Three or Four.

Data Code Two: Chemical data collected quarterly to bimonthly for at least three years or intensive studies that monitor several nearby sites repeatedly over short periods of time or at least three fish tissue samples per water body.

Data Code Three: Chemical data collected at least monthly for more than three years on a variety of water quality constituents including heavy metals and pesticides; or quantitative biological monitoring of at least one aquatic assemblage (fish, invertebrates or algae) at multiple sites, or multiple samples at a single site when data from that site is supported by biological monitoring at an appropriate control site.

Data Code Four: Chemical data collected at least monthly for more than three years that provides data on a variety of water quality constituents including heavy metals and pesticides, and including chemical sampling of sediments and fish tissue; or quantitative biological monitoring of at least two aquatic assemblages (fish, invertebrates or algae) at multiple sites.

In Missouri, the primary purpose of Data Code One data is to provide a rapid and inexpensive method of screening large numbers of waters for obvious water quality problems and to determine where more intensive monitoring is needed. In the preparation of the state's 305(b) Report, data from all four data quality levels are used.

---

[4] *Guidelines for the Preparation of the Comprehensive State Water Quality Assessments (305b) and Electronic Updates*, 1997.

[5] Data Code One is equivalent to data water quality assurance Level One in 10 CSR 20-7.050 General Methodology for Development of Impaired Waters List, subsection (2)(C), Data Code Two is equivalent to Level 2, etc.

Most of the data is of Data Code One quality, and without Data Code One data, the Department would not be able to assess a majority of the state's waters.

In general, when selecting water bodies for the Missouri 303(d) List, only Data Code Two or higher data are used, unless the problem can be accurately characterized by Data Code One data.[6] The reason is that Data Code Two data provides a higher level of assurance that a Water Quality Standard is actually being exceeded and that a Total Maximum Daily Load study is necessary. All water bodies placed in Categories 2B or 3B receive high priority for additional monitoring so that data quality is upgraded to at least Data Code Two.

D. How Water Quality Data is Evaluated to Determine Whether or Not Waters are Impaired for 303(d) Listing Purposes

Physical, Chemical, Biological and Toxicity Data

Each reporting cycle, the Department and stakeholders review and revise the guidelines for determining water quality impairment. These guidelines are shown in Tables 1.1 and 1.2 which provide the general rules of data use and assessment and Tables B-1 and B-2 that provide details about the specific analytical procedure used. In addition, if time trend data indicates that presently unimpaired waters will become impaired prior to the next listing cycle, these "threatened waters" will be judged to be impaired. Where antidegradation provisions in Missouri's Water Quality Standards apply, those provisions shall be upheld. The numeric criteria included in Table 1.1 have been adopted into the state Water Quality Standards, 10 CSR 20-7.031, and are used, as described in Table 1.1, to make use attainment decisions. For narrative criteria, the numeric thresholds included in Table 1.2 have not been adopted into state Water Quality Standards. The Department will use a weight of evidence analysis for all narrative criteria. For those analytes with numeric thresholds, the threshold values given in Table 1.2 will trigger a weight of evidence analysis to determine the existence or likelihood of use impairment and the appropriateness of proposing a listing based on narrative criteria. This weight of evidence analysis will include the use of other types of environmental data when it is available. Examples of other relevant environmental data might include biological data on fish or aquatic invertebrate animals or toxicity testing of water or sediments. When the weight of evidence analysis suggests, but does not provide strong, scientifically defensible evidence of impairment, the Department will place the water body in question in Categories 2B or 3B. The Department will produce a document showing all relevant data and the rationale for the use attainment decision. All such documents will be made available to the public at the time of the first public notice of the proposed 303(d) list. A final recommendation on the listing of a water based on narrative criteria will only be made after full consideration of all comments on the proposal.

[6] When a listing, amendment or delisting of a 303(d) water is made with only Data Code One data, a document will be prepared that includes a display of all data and a presentation of all statistical tests or other evaluative techniques that documents the scientific defensibility of the data. This requirement applies to all Data Code One data identified in Table 1 of this document.

For the interpretation of biological data, where habitat assessment data indicates habitat scores are less than 75 percent of reference or appropriate control stream scores, and in the absence of other data indicating impairment by a discrete pollutant, a waterbody judged to be impaired will be placed in Category 4C.

For the interpretation of toxicity test data, standard acute or chronic bioassay procedures using freshwater aquatic fauna such as, but not limited to, *Ceriodaphnia dubia, Pimephales promelas* or *Hyalella azteca* will provide adequate evidence of toxicity for 303(d) listing purposes. Microtox toxicity tests may be used to list a water as affected by "toxicity" only if there is data of another kind (freshwater toxicity tests, sediment chemistry, water chemistry or biological sampling) that indicates water quality impairment.

TABLE 1.1 METHODS FOR ASSESSING COMPLIANCE WITH WATER QUALITY STANDARDS USED FOR 303(d) LISTING PURPOSES: NUMERIC CRITERIA THAT ARE INCLUDED IN STATE WATER QUALITY STANDARDS, 10 CSR 20-7.031

BENEFICIAL USES	DATA TYPE	DATA QUALITY CODE	COMPLIANCE WITH WATER QUALITY STANDARDS[7]
Overall use protection (all beneficial uses)	No data. Evaluated based on similar land use/ geology as stream with water quality data.[8]	Not applicable	Given same rating as monitored stream with same land use and geology.
Any beneficial uses	No data available or where only effluent data is available. Results of dilution calculations or water quality modeling. (see ALRR p.38)	Not applicable	Where models or other dilution calculations indicate noncompliance with allowable pollutant levels and frequencies noted in this table, waters may be added to Category 3B and considered high priority for water quality monitoring.
Protection of Aquatic Life	Water temperature, pH, total dissolved gases, oil and grease.	1-4	Full: No more than 10% of all samples exceed criterion.[9] Non-Attainment: Requirements for full attainment not met.
Protection of Aquatic Life	Dissolved oxygen.	1-4	Full: No more than 10% of all samples exceed criterion.[9] Non-Attainment: Requirements for full attainment not met.
Protection of Aquatic Life	Toxic Chemicals	1-4	Full: No more than one acute toxic event in three years. No more than one exceedence of acute or chronic criterion in three years for all toxics Non-Attainment: Requirements for full attainment

[7] See section on Statistical Considerations, Table B-1 and B-2.
[8] This data type is used only for wide-scale assessments of aquatic biota and aquatic habitat for 305(b) Report purposes. This data type is not used in the development of the 303(d) List.
[9] Some sampling periods are wholly or predominantly during the critical period of the year when criteria violations occur. Where the monitoring program presents good evidence of a demarcation between seasons where criteria exceedences occur and seasons when they do not, the 10% exceedence rate will be based on an annual estimate of the frequency of exceedence.

TABLE 1.1 METHODS FOR ASSESSING COMPLIANCE WITH WATER QUALITY STANDARDS USED FOR 303(d) LISTING PURPOSES: NUMERIC CRITERIA THAT ARE INCLUDED IN STATE WATER QUALITY STANDARDS, 10 CSR 20-7.031

BENEFICIAL USES	DATA TYPE	DATA QUALITY CODE	COMPLIANCE WITH WATER QUALITY STANDARDS[7]
			not met.
Protection of Aquatic Life	Nutrients in Lakes (total phosphorus, Total nitrogen, Chlorophyll)	1-4	Full: Nutrient levels do not exceed WATER QUALITY STANDARDS. Non-Attainment: Requirements for full attainment not met.[10]
Fish Consumption	Chemicals (water)	1-4	Full: Water quality does not exceed WATER QUALITY STANDARDS Non-Attainment: Requirements for full attainment not met.
Drinking Water Supply -Raw Water.[11]	Chemical (toxics)	1-4	Full: WATER QUALITY STANDARDS not exceeded Non-Attainment: Requirements for full attainment not met.
Drinking Water Supply- Raw Water	Chemical (sulfate, chloride, fluoride)	1-4	Full: WATER QUALITY STANDARDS not exceeded . Non-Attainment: Requirements for full attainment not met.
Drinking Water Supply-Finished Water	Chemical (toxics)	1-4	Full: No MCL* violations based on Safe Drinking Water Act data evaluation procedures. Non-Attainment: Requirements for full attainment not met. NOTE: Finished water data will not be used for analytes where water quality problems may be caused by the drinking water treatment process such as the formation of Trihalomethanes (THMs) or problems that may be caused by the distribution system (bacteria, lead, copper).
Whole-Body-Contact Recreation and Secondary Contact Recreation	Fecal Coliform or *E. coli* count	1-4	Where there are at least five samples per year taken during the recreational season: Full: WATER QUALITY STANDARDS not exceeded as a geometric mean, in any of the last three years for which data is available, for samples collected during seasons for which bacteria criteria apply.[12] Non-Attainment: Requirements for full attainment not met.
Irrigation, Livestock and	Chemical	1-4	Full: WATER QUALITY STANDARDS not exceeded. Non-Attainment: Requirements for full attainment

[10] Nutrient criteria will be used in the 2010 LMD only if these criteria appear in the Code of State Regulations by the date the 2010 303(d) List is presented to the Clean Water Commission for approval.
[11] Raw water is water from a stream, lake or ground water prior to treatment in a drinking water treatment plant.
[12] A geometric mean of 206 cfu/100 ml for E. coli will be used as a criterion value for Category B Recreational Waters. Because Missouri's Fecal Coliform Standard ended December 31, 2008, any waters appearing on the 2008 303(d) List as a result of the Fecal Coliform Standard will be retained on the list with the pollutant listed as "bacteria" until sufficient E. coli sampling has determined the status of the water.

TABLE 1.1 METHODS FOR ASSESSING COMPLIANCE WITH
WATER QUALITY STANDARDS USED FOR 303(d) LISTING PURPOSES: NUMERIC CRITERIA THAT ARE INCLUDED IN STATE WATER QUALITY STANDARDS, 10 CSR 20-7.031

BENEFICIAL USES	**DATA TYPE**	**DATA QUALITY CODE**	**COMPLIANCE WITH WATER QUALITY STANDARDS[7]**
Wildlife Water			not met.

*Maximum Contaminant Level

TABLE 1.2 METHODS FOR ASSESSING COMPLIANCE WITH WATER QUALITY STANDARDS USED FOR 303(d) LISTING PURPOSES: NARRATIVE CRITERIA BASED ON NUMERIC THRESHOLDS NOT CONTAINED IN STATE WATER QUALITY STANDARDS (10 CSR 20-7.031)

BENEFICIAL USES	DATA TYPE	DATA QUALITY CODE	COMPLIANCE WITH WATER QUALITY STANDARDS[7]
Overall use protection (all beneficial uses)	Narrative criteria for which quantifiable measurements can be made.	1-4	Full: Stream appearance typical of reference or appropriate control streams in this region of the state. Non-Attainment: The weight of evidence, based on the narrative criteria in 10 CSR 20-7.031(3), demonstrates the observed condition exceed a numeric threshold necessary for the attainment of a beneficial use For example: Color: Color as measured by the Platinum-Cobalt visual method (SM 2120 B) in a water is statistically significantly higher than a control water. Objectionable Bottom Deposits: The bottom that is covered by sewage sludge, trash or other materials reaching the water due to anthropogenic sources exceeds the amount in reference or control streams by more than twenty percent. Note: Waters in mixing zones and unclassified waters which support aquatic life on an intermittent basis shall be subject to acute toxicity criteria for protection of aquatic life. Waters in the initial Zone of Dilution (ZID) shall not be subject to acute toxicity criteria.
Protection of Aquatic Life	Toxic Chemicals	1-4	Full: No more than one acute toxic event in three years. No more than one exceedence of acute or chronic criterion in three years for all toxics.[13] [14] Non-Attainment: Requirements for full attainment not met.

[13] The test result must be representative of water quality for the entire time period for which acute or chronic criteria apply. For ammonia the chronic exposure period is 30 days, for all other toxics 96 hours. The acute exposure period for all toxics is 24 hours. The Department will review all appropriate data, including hydrographic data, to insure only representative data is used. Except on large rivers where storm water flows may persist at relatively unvarying levels for several days, grab samples collected during storm water flows will not be used for assessing chronic toxicity criteria.

[14] In the case of toxic chemicals occurring in benthic sediment rather than in water, the numeric thresholds used to determine the need for further evaluation will be the Probable Effect Concentrations proposed in "Development and Evaluation of Consensus-Based Sediment Quality Guidelines for Freshwater Ecosystems" by McDonald, D.D. et al. Arch. Environ. Contam. Toxicol. 39,20-31 (2000). These - Probable Effect Concentrations are as follows: 33 mg/kg As; 4.98 mg/kg Cd; 111 mg/kg Cr; 149 mg/kg Cu; 48.6 mg/kg Ni; 128 mg/kg Pb; 459 mg/kg Zn; 561 µg/kg naphthalene; 1170 µg/kg phenanthrene; 1520 µg/kg pyrene; 1050 µg/kg benzo(a)anthracene, 1290 µg/kg chrysene; 1450 µg/kg benzo(a)pyrene; 22,800 µg/kg total polyaromatic hydrocarbons; 676 µg/kg total PCBs. Chlordane 17.6 ug/kg; Sum DDE 31.3 ug/kg; Lindane (gamma-BHC) 4.99 ug/kg. Where multiple sediment contaminants exist, the Probable Effect Concentrations Quotient shall not exceed 0.5. See Appendix D for more information on the Probable Effect Concentrations Quotient.

TABLE 1.2 METHODS FOR ASSESSING COMPLIANCE WITH WATER QUALITY STANDARDS USED FOR 303(d) LISTING PURPOSES: NARRATIVE CRITERIA BASED ON NUMERIC THRESHOLDS NOT CONTAINED IN STATE WATER QUALITY STANDARDS (10 CSR 20-7.031)

BENEFICIAL USES	DATA TYPE	DATA QUALITY CODE	COMPLIANCE WITH WATER QUALITY STANDARDS[7]
Protection of Aquatic Life	Biological	3-4	Full: For seven or fewer samples and following DNR wadeable streams macroinvertebrate sampling and evaluation protocols, 75% of the stream condition index scores must be 16 or greater. Fauna achieving these scores are considered to be very similar to regional reference streams. For greater than seven samples or for other sampling and evaluation protocols, results must be statistically similar to representative reference or control streams[15]. Non-Attainment: For seven or fewer samples and following DNR wadeable streams macroinvertebrate sampling and evaluation protocols, 75% of the stream condition index scores must be 14 or lower. Fauna achieving these scores are considered to be substantially different from regional reference streams. For more than seven samples or for other sampling and evaluation protocols, results must be statistically dissimilar to control or representative reference streams.
Protection of Aquatic Life	Toxicity testing of streams or lakes using aquatic organisms	2	Full: No more than one test result of statistically significant deviation from controls in acute or chronic test in a three-year period.[15] Non-Attainment: Requirements for full attainment not met.
Fish Consumption	Chemicals (tissue)	1-2	Full: Fish tissue levels in fillets and eggs do not exceed guidelines.[16] Non-Attainment: Requirements for full attainment not met.

Duration of Assessment Period

Except where the assessment period is specifically noted in Table 1, the time period for which data will be used in making the assessments noted in Table 1 will be determined by the data age considerations in Section II.C.3.3.1 and data representativeness considerations in Table 1 footnote 13.

[15] See Table B-1 and B-2.

[16] Fish tissue threshold levels are; chlordane 0.1 mg/kg (Crellin, J.R. 1989*, "New Trigger Levels for Chlordane in Fish-Revised Memo"* Mo. Dept. of Health inter-office memorandum. June 16, 1989); mercury 0.3 mg/kg based on "Water Quality Criterion for Protection of Human Health: Methylmercury" EPA-823-R-01-001. Jan. 2001. http://www.epa.gov/waterscience/criteria/methylmercury/merctitl.pdf; PCBs 0.75 mg/kg, MDHSS Memorandum August 30, 2006 *"Development of PCB Risk-based Fish Consumption Limit Tables*"; and lead 0.3- mg/kg (World Health Organization 1972. *"Evaluation of Certain Food Additives and the Contaminants Mercury, Lead and Cadmium"*. WHO Technical Report Series No. 505, Sixteenth Report on the Joint FAO/WHO Expert Committee on Food Additives. Geneva 33 pp. Assessment of Mercury will be based on samples solely from the following higher trophic level fish species; walleye, sauger, trout, black bass, white bass, striped bass, northern pike, flathead catfish and blue catfish.

Other Types of Information

1. Observation and evaluation of waters for noncompliance with state narrative water quality criteria. Missouri's narrative water quality criteria, as described in 10 CSR 20-7.031 Section (3), may be used to evaluate waters when a quantitative value can be applied to the pollutant (see Table 1 page 15). These narrative criteria apply to both classified and unclassified waters and prohibit the following in waters of the state:
   a. Unsightly, putrescent or harmful bottom deposits,
   b. Oil, scum and floating debris,
   c. Unsightly color, turbidity or odor,
   d. Substances or conditions causing toxicity to human, animal or aquatic life,
   e. Human health hazard due to incidental contact,
   f. Acute toxicity to livestock or wildlife when used as a drinking water supply,
   g. Physical, chemical or hydrologic changes that impair the natural biological community, and
   h. Used tires, car bodies, appliances, demolition debris, used vehicles or equipment and any solid waste as defined by Missouri's Solid Waste Law,
   i. Acute toxicity.
2. Habitat assessment protocols for wadeable streams have been established and are made in conjunction with sampling of aquatic invertebrates and the analysis of aquatic invertebrates data. The Department will not use habitat assessment data alone for assessment purposes.

E. Other 303(d) Listing Considerations

1. Adding to the Existing List or Expanding the Scope of Impairment to a Previously Listed Water

   The listed portion of an impaired water may be increased based on recent monitoring data following the guidelines in this document. One or more new pollutants may be added to the listing for a water already on the list based on recent monitoring data following these same guidelines. Waters not previously listed may be added to the list following the guidelines in this document.

2. Deleting from the Existing List or Decreasing the Scope of Impairment to a Previously Listed Water

   The listed portion of an impaired water may be decreased based on recent monitoring data following the guidelines in this document. One or more pollutants may be deleted from the listing for a water already on the list based on recent monitoring data following these same guidelines. Waters may be completely removed from the list for several reasons[17], the

[17] see, "Guidance for 2006 Assessment, Listing and Reporting Requirements Pursuant to Sections 303(d), 305(b) and 314 of the Clean Water Act". USEPA, Office of Water, Washington DC.

most common being (1) water has returned to compliance with water quality standards or (2) the water has an approved Total Maximum Daily Load study.

3. Prioritization of Waters for Total Maximum Daily Load Development

Section 303(d) of the Clean Water Act and federal regulation 40 CFR 130.7(b)(4) require states to submit a priority ranking of waters still requiring Total Maximum Daily Loads. The department will prioritize development of Total Maximum Daily Loads based on several variables including:

- severity of the water quality problem
- amount of time necessary to acquire sufficient data to develop the Total Maximum Daily Load
- court orders, consent decrees or other formal agreements
- budgetary constraints, and
- amenability of the problem to treatment

The department's Total Maximum Daily Load schedule will represent its prioritization.

4. Resolution of Interstate/International Disagreements

The Department will review the draft 303(d) Lists of all other states with which it shares a border (Missouri River, Mississippi River, Des Moines River and the St. Francis River) or other interstate waters. Where the listing in another state is different than in Missouri, the department will request the data upon which the listing in the other state is based. This data will be reviewed following all data evaluation guidelines previously discussed in this document. The Missouri list may be changed pending the evaluation of this additional data.

Appendix A

Excerpt from *Guidance for 2006 Assessment, Listing and Reporting Requirements Pursuant to Sections 303(d), 305(b) and 314 of the Clean Water Act.* July 29, 2005. USEPA pp.39-41.

G. How should statistical approaches be used in attainment determinations?

The state's methodology should provide a rationale for any statistical interpretation of data for the purpose of making an assessment determination.

1. Description of statistical methods to be employed in various circumstances:

The methodology should provide a clear explanation of which analytic tools the state uses and under which circumstances. EPA recommends that the methodology explain issues such as the selection of key sample statistics (arithmetic mean concentration, median concentration, or a percentile), null and alternative hypotheses, confidence intervals, and Type I and Type II error thresholds. The choice of a statistic tool should be based on the known or expected distribution of the concentration of a pollutant in the segment (e.g., normal or log normal) in both time and space.

Past EPA guidance, 1997 305(b) and 2000 CALM, recommended making non-attainment decisions for "conventional pollutants" – Total Suspended Solids, pH, Biochemical Oxygen Demand, fecal coliform bacteria and oil and grease – when more than 10% of measurements exceed the water quality criterion; however, EPA guidance has not encouraged use of the 10% rule with other pollutants, including toxics. Use of this rule when addressing conventional pollutants, is appropriate if its application is consistent with the manner in which the applicable water quality criterion are expressed. An example of a water quality criterion for which an assessment based on the 10% rule would be appropriate is the EPA acute water quality criterion for fecal coliform bacteria, applicable to protection of water contact recreational use. This 1976-issued water quality criterion was expressed as, "...no more than ten percent of the samples exceeding 400 CFU per 100ml, during a 30-day period. This assessment methodology is clearly reflective of the water quality criterion.

On the other hand, use of the 10 percent rule for interpreting water quality data is usually not consistent with water quality criterion expressed either as: (1) instantaneous maxima not to be surpassed at any time; or (2) average concentrations over specified times. In the case of "instantaneous maxima (or minima) never to occur" criteria use of the 10 percent rule typically leads to the belief that segment conditions are equal to or better than specified by the water quality criterion, when they in fact are considerably worse. (That is, pollutant concentrations are above the criterion concentration a far greater proportion of the time than specified by the water quality criterion). Conversely, use of this decision rule in concert with water quality criterion expressed as average concentrations over specific times can lead to

concluding that segment conditions are worse than water quality criterion, when in fact, they are not. If the state applies different decision rules for different types of pollutants (e.g., toxic, conventional, and non-conventional pollutants) and types of standards (e.g., acute versus chronic criteria for aquatic life or human health), the state should provide a reasonable rationale supporting the choice of a particular statistical approach to each of its different sets of pollutants and types of standards.

2. Elucidation of policy choices embedded in selection of particular statistical approaches and use of certain assumptions:

   EPA strongly encourages states to highlight policy decisions implicit in the statistical analysis that they have chosen to employ in various circumstances. For example, if hypothesis testing is used, the state should make its decision-making rules transparent by explaining why it chose either "meeting Water Quality Standards" or "not meeting Water Quality Standards" as the null hypothesis (refutable presumption) as a general rule for all waters, a category of waters, or an individual segment. Starting with the assumption that a water is "healthy" when employing hypothesis testing means that a segment will be identified as impaired, and placed in Category 4 or 5, only if substantial amounts of credible evidence exist to refute the presumption. By contrast, making the null hypothesis "Water Quality Standards not being met" shifts the burden of proof to those who believe the segment is, in fact, meeting Water Quality Standards.

   Which "null hypothesis" a state selects could likely create contrasting incentives regarding support for additional ambient monitoring among different stakeholders. If the null hypothesis is "meeting standards", there was no previous data on the segment, and no additional existing and readily available data and information is collected, then the "null hypothesis" cannot be rejected, and the segment would not be placed in Category 4 or 5. In this situation, those concerned about possible adverse consequences of having a segment declared "impaired" might have little interest in collection of additional ambient data. Meanwhile, users of the segment would likely want to have the segment monitored, so they can be assured that it is indeed capable of supporting the uses of concern. On the other hand, if the null hypothesis is changed to "segment not meeting Water Quality Standards": then those that would prefer that a particular segment not be labeled "impaired" would probably want more data collected, in hopes of proving that the null hypothesis is not true.

   Another key policy issue in hypothesis testing is what significance level to use in deciding whether to reject the null hypothesis. Picking a high level of significance for rejecting the null hypothesis means that great emphasis is being placed on avoiding a Type I error (rejecting the null hypothesis, when in fact, the null hypothesis is true). This means that if a 0.10 significance level is chosen, the state wants to keep the chance of making a Type I error at or below 10 percent. Hence, if the chosen null hypothesis is "segment meeting Water Quality Standards", the state is trying to keep the chance of saying a segment is impaired, when in reality it is not, under 10 percent.

An additional policy issue is the Type II errors (not rejecting the null hypothesis, when it should have been). The probability of Type II errors depends on several factors. One key factor is the number of samples available. With a fixed number of samples, as the probability of Type I error decreases, the probability of a Type II error increases. States would ideally collect enough samples so the chances of making Type I and Type II errors are simultaneously small. Unfortunately, resources needed to collect those numbers of samples are quite often not available.

The final example of a policy issue that a state should describe is the rationale for concentrating limited resources to support data collection and statistical analysis in segments where there are documented water quality problems or where the combination of nonpoint source loadings and point source discharges would indicate a strong potential for a water quality problem to exist.

EPA recommends that, when picking the decision rules and statistical methods to be utilized when interpreting data and information, states attempt to minimize the chances of making either of the following two errors:

- Concluding the segment is impaired, when in fact it is not, and
- Deciding not to declare a segment impaired, when it is in fact impaired.

States should specify in their methodology what significance level they have chosen to use, in various circumstances. The methodology would best describe in "plain English" the likelihood of deciding to list a segment that in reality is not impaired (Type I error if the null hypothesis is "segment not impaired"). Also, EPA encourages states to estimate, in their assessment databases, the probability of making a Type II error (not putting on the 303(d) List a segment that in fact fails to meet Water Quality Standards), when: (1) commonly-available numbers of grab samples are available, and (2) the degree of variance in pollutant concentrations are at commonly encountered levels. For example, if an assessment is being performed with a WQC expressed as a 30-day average concentration of a certain pollutant, it would be useful to estimate the probability of a Type II error when the number of available samples over a 30-day period is equal to the average number of samples for that pollutant in segments statewide, or in a given group of segments, assuming a degree of variance in levels of the pollutant often observed over typical 30-day periods.

Appendix B
Statistical Considerations

The most recent EPA guidance on the use of statistics in the 303(d) listing methodology document is given in Appendix A. Within this guidance there are three major recommendations regarding statistics:

- Provide a description of which analytical tools the state uses under various circumstances,
- When conducting hypothesis testing, explain the various circumstances under which the burden of proof is placed on proving the water is impaired and when it is placed on proving the water is unimpaired, and
- Explain the level of statistical significance used under various circumstances.

Description of Analytical Tools

The Tables B-1 and B-2 below describe~~s~~ the analytical tools the department will use to determine impairment (Table B-1) and to determine when listed waters are no longer impaired (Table B-2).

Table B-1. Description of Analytical Tools for Determining if Waters are Impaired

Beneficial Use	Analytes	Analytical Tool	Decision Rule/ Hypothesis	Criterion Used with the Decision Rule [18]	Significance Level
Narrative Criteria	Color (Narrative)	Hypothesis Test Two Sample, one tailed "t "Test	Null Hypothesis: There is no difference in color between test stream and control stream.	Reject Null Hypothesis if calculated "t" value exceeds tabular "t" value for test alpha	0.10
	Bottom Deposits (Narrative)	Hypothesis Test, One Sided Confidence Limit	Null Hypothesis: Solids of anthropogenic origin cover less than 20% of stream bottom where velocity is less than 0.5 feet/second.	Reject Null Hypothesis if 60% Lower Confidence Limit (LCL) of mean percent fine sediment deposition (pfsd) in stream is greater than the sum of the pfsd in the control and 20 % more of the stream bottom. i.e., where the pfsd is expressed as a decimal, test stream pfsd > (control stream pfsd)+ (0.20)	0.40

[18] Where hypothesis testing is used for media other than fish tissue, for data sets with five samples or fewer, a 75 percent confidence interval around the appropriate central tendencies will be used to determine use attainment status. Use attainment will be determined as follows: (1) If the criterion value is above this interval (all values within the interval are in conformance with the criterion), rate as unimpaired. (2) If the criterion value falls within this interval, rate as unimpaired and place in Category 2B or 3B. (3) If the criterion value is below this interval (all values within the interval are not in conformance with the criterion), rate as impaired. For fish tissue this procedure will be used with the following changes: (1) it will apply only to sample sizes of less than four and, (2) a 50% confidence interval will be used in place of the 75% confidence interval.

Table B-1. Description of Analytical Tools for Determining if Waters are Impaired

Beneficial Use	Analytes	Analytical Tool	Decision Rule/ Hypothesis	Criterion Used with the Decision Rule [18]	Significance Level
Aquatic Life	Biological Monitoring (Narrative)	For DNR Invert protocol: Binomial probability for Sample sizes 8 to 30.	Using DNR Invert. protocol: Null Hypothesis: Frequency of full sustaining scores for test stream is the same as for biological criteria reference streams.	Reject Null Hypothesis if frequency of fully sustaining scores on test stream is significantly less than for biological criteria reference streams.	0.10
		For DNR Invert protocol and sample sizes greater than 30: Direct comparison.	A direct comparison of frequencies between test and biological criteria reference streams will be made	Rate as impaired if biological criteria reference stream frequency of sustain-ing scores is more than five percent more than test stream	Not applicable
		For other biological data: An appropriate parametric or nonparametric test will be used.	Null Hypothesis, Community metric(s) in test stream is the same as for a reference stream or control streams.	Reject Null Hypothesis If metric scores for test stream are significantly less than reference or control streams.	0.1
			Other biological monitoring to be determined by type of data.		
Aquatic Life	Toxic Chemicals in Water. (Numeric)	Not applicable	No more than one toxic event, toxicity test failure or exceedence of acute or chronic criterion in 3 years.	Not applicable	Not applicable
	Toxic Chemicals in Sediments (Narrative)	Comparison of mean to PEL value.	Waters is judged to be Impaired if sample mean Exceeds 150% of PEL.[19]		
Aquatic Life	temperature, pH, total diss. gases, oil and grease, diss. oxygen (Numeric)	30 or fewer samples: Binomial probability	Null Hypothesis: No more than 10% of samples exceed the water quality criterion	Reject Null Hypothesis if the exceedence frequency is significantly more than 10%	0.10
		More than 30 samples: Percent of samples that fail to meet criterion.	If observed frequency exceeds 10%, rate as impaired.	Not applicable	Not applicable
Fish Consumption	Toxic Chemicals in water (Numeric)	Hypothesis test 1-Sided Confidence Limit	Null Hypothesis: Levels of contaminants in water do not exceed criterion.	Reject Null Hypothesis if the 60% LCL is greater than the criterion value.	0.40
Fish Consumption	Toxic Chemicals in Tissue (Narrative)	Four or more samples: Hypothesis test 1-Sided Confidence Limit	Null Hypothesis: Levels in fillet samples or fish eggs do not exceed criterion.	Reject Null Hypothesis if the 60% LCL is greater than the criterion value.	0.40

[19] Where there is convincing evidence of a healthy biological community (fish and/or aquatic invertebrate monitoring data) or convincing evidence of a lack of toxicity (two species bioassay tests of sediment elutriate water or sediment pore water), this evidence will be evaluated in conjunction with the sediment PEL data.

Table B-1. Description of Analytical Tools for Determining if Waters are Impaired

Beneficial Use	Analytes	Analytical Tool	Decision Rule/ Hypothesis	Criterion Used with the Decision Rule [18]	Significance Level
Drinking Water Supply (Raw)	Toxic Chemicals (Numeric)	Hypothesis test 1-Sided Confidence limit	Null Hypothesis: Levels of contaminants do not exceed criterion.	Reject Null Hypothesis if the 60% LCL is greater than the criterion value.	0.40
Drinking Water Supply (Raw)	Non-toxic Chemicals (Numeric)	Hypothesis test 1-Sided Confidence limit	Null Hypothesis: Levels of contaminants do not exceed criterion.	Reject Null Hypothesis if the 60% LCL is greater than the criterion value.	0.40
Drinking Water Supply (Finished)	Toxic Chemicals	Methods stipulated by Safe Drinking Water Act	Methods stipulated by Safe Drinking Water Act	Methods stipulated by Safe Drinking Water Act	Methods stipulated by Safe Drinking Water Act
Whole Body Contact and Secondary Contact Rec.	Bacteria (Numeric)	Geometric Mean	Null Hypothesis: Levels of contaminants do not exceed criterion.	Reject Null Hypothesis if the Geometric Mean is greater than the criterion value.	-Not Applicable
Irrigation & Livestock Water	Toxic Chemicals (Numeric)	Hypothesis test 1-Sided Confidence limit	Null Hypothesis: Levels of contaminants do not exceed criterion.	Reject Null Hypothesis if the 60% LCL is greater than the criterion value.	0.40
Protection of Aquatic Life	Nutrients in Lakes (Numeric)	Hypothesis test[20]	Null hypothesis: Criteria are not exceeded.	Reject Null hypothesis if 60% LCL value is more than criterion value.	0.40

[20] State nutrient criteria require at least four samples per year taken near the outflow point of the lake (or reservoir) between May 1 and August 31 for at least four consecutive years.

Table B-2. Description of Analytical Tools for Determining When Waters are No Longer Impaired

Beneficial Use	Analytes	Analytical Tool	Decision Rule/ Hypothesis	Criterion Used with the Decision Rule [19]	Significance Level
Narrative Criteria	Color (Narrative)	Same as Table B-1	Same as Table B-1	Same as Table B-1	0.40
	Bottom Deposits (Narrative)	Same as Table B-1	Same as Table B-1	Same as Table B-1	0.40
Aquatic Life	Biological Monitoring (Narrative)	DNR Invert Protocol: For 8 to 30 samples Same as Table B-1	Same as Table B-1	Same as Table B-1	0.40
		For DNR Invert Protocol For more than 30 Same as Table B-1	Same as Table B-1.	Same as Table B-1.	Same as Table B-1.
		For other biological data: Same as Table B-1.	Same as Table B-1.	Same as Table B-1.	0.40
	Toxic Chemicals in Water.	Same as Table B-1.	Same as Table B-1.	Same as Table B-1.	Same as Table B-1.
	Toxic Chemicals in Sediments	Comparison of mean to PEL value.	Water is judged to be unimpaired if sample mean does not exceed 150 % of PEL.[21]	Not applicable	Not applicable
Aquatic Life	temperature, pH, total diss. gases, oil and grease, diss. oxygen	30 or fewer samples: Same as Table B-1.	Same as Table B-1.	Same as Table B-1.	Same as Table B-1.
		More than 30 samples: Same as Table B-1.	Same as Table B-1.	Same as Table B-1.	Same as Table B-1.
Fish Consumption	Toxic Chemicals in water	Same as Table B-1.	Same as Table B-1.	Reject null hypothesis if the 60% UCL is greater than the criterion value.	0.40
	Toxic Chemicals in Tissue	Same as Table B-1.	Same as Table B-1.	Reject null hypothesis if the 60% UCL is greater than the criterion value.	0.40
Drinking Water Supply (Raw)	Toxic Chemicals	Same as Table B-1.	Same as Table B-1.	Reject null hypothesis if the 60% UCL is greater than the criterion value.	0.40
Drinking Water Supply (Raw)	Non-toxic Chemicals	Same as Table B-1.	Same as Table B-1.	Reject null hypothesis if the 60% UCL is greater than the criterion value.	0.40
Drinking Water Supply (Finished)	Toxic Chemicals,	Same as Table B-1.	Same as Table B-1.	Same as Table B-1.	Same as Table B-1.
Whole Body Contact and Secondary Contact Rec.	Bacteria	Same as Table B-1.	Same as Table B-1.	Same as Table B-1	Not applicable
Irrigation & Livestock Water	Toxic Chemicals	Same as Table B-1.	Same as Table B-1.	Reject null hypothesis if the 60% UCL is greater than the criterion value.	0.40
Protection of Aquatic Life	Nutrients in Lakes	Same as Table B-1.	Same as Table B-1.	Same as Table B-1.	0.40

Rationale for the Burden-of-Proof

[21] Where there is convincing evidence of a healthy biological community (fish and/or aquatic invertebrate monitoring data) or convincing evidence of a lack of toxicity (two species bioassay tests of sediment elutriate water or sediment pore water), sediment PEL data will not be used as the sole justification for listing a water as impaired.

Hypothesis testing is a common statistical practice. The procedure involves first stating a hypothesis you want to test, such as "the most frequently seen color on clothing at a St. Louis Cardinals game is red" and then the opposite or null hypothesis "red is not the most frequently seen color on clothing at a Cardinals game." Then a statistical test is applied to the data (a sample of the predominant color of clothing worn by 200 fans at a Cardinals game on July 12) and based on an analysis of that data, one of the two hypotheses is chosen as correct.

In hypothesis testing, the burden-of-proof is always on the alternate hypothesis. In other words, there must be very convincing data to make us conclude that the null hypothesis is not true and that we must accept the alternate hypothesis. How convincing the data must be is stated as the "significance level" of the test. A significance level of 0.10 means that there must be at least a 90 percent probability that the alternate hypothesis is true before we can accept it and reject the null hypothesis.

For analysis of a specific kind of data, either the test significance level or the statement of null and alternative hypotheses, or both, can be varied to achieve the desired degree of statistical rigor. The department has chosen to maintain a consistent set of null and alternate hypotheses for all our statistical procedures. The null hypothesis will be that the water body in question is unimpaired and the alternate hypothesis will be that it is impaired. Varying the level of statistical rigor will be accomplished by varying the test significance level. For determining impairment (Table B-1) test significance levels are set at either 0.1 or 0.4, meaning the data must show a 90% or 60% probability respectively, that the water body is impaired. However, if the department retained these same test significance levels in determining when an impaired water had been restored to an unimpaired status (Table B-2) some undesirable results can occur.

For example, using a 0.1 significance level for determining both impairment and nonimpairment; if the sample data indicate the stream had a 92 percent probability of being impaired, it would be rated as impaired. If subsequent data was collected and added to the database and the data now showed the water had an 88 percent chance of being impaired , it would be rated as unimpaired. Judging as unimpaired a water with only a 12 percent probability of being unimpaired is clearly a poor decision. To correct this problem, the department will use a test significance level of 0.4 for some analytes and 0.6 for others. This will increase our confidence in determining compliance with criteria to 40 percent and 60 percent respectively under the worst case conditions, and for most databases will provide an even higher level of confidence.

Level of Significance Used in Tests

The choice of significance levels is largely related to two concerns. The first is concerned with matching error rates with the severity of the consequences of making a decision error. The second addresses the need to balance, to the degree practicable, Type I and Type II error rates.
For relatively small databases, the disparity between Type I and Type II errors can be large. The table below shows error rates calculated using the binomial distribution for two very similar situations. Type I error rates are based on a stream with a 10 percent exceedence rate of a standard

and Type II error rates for a stream with a 15 percent exceedence rate of a standard. Note that choosing a Type I error rate of 0.05 rather than 0.10 increases an already very large Type II error rate by about 10 percent. Also note that for a given Type I error rate, the Type II error rate declines as sample size increases.

Table B-3. Effects of Type I Error Rates and Sample Size on Type II Error Rates

No. of Samples	No. Meeting Standards	Type I Error Rate	Type II Error Rate	No. of Samples	No. Meeting Standards	Type I Error Rate	Type II Error Rate
6	5	.11	.78	4	3	.05	.89
11	9	.09	.78	9	7	.05	.86
18	15	.10	.72	15	12	.05	.82
25	21	.10	.68	21	17	.05	.80
				27	21	.05	.78

Use of the Binomial Probability Distribution for Interpretation of the Ten Percent Rule

There are two options for assessing data for compliance with the ten percent rule. One is to simply calculate the percent of time the criterion value is not met and to judge the water to be impaired if this value is greater than ten percent. The second method is to use some evaluative procedure that can review the data and provide a probability statement regarding the compliance with the ten percent rule. Since the latter option allows assessment decisions relative to specific test significance levels and the first option does not, the latter option is preferred. The procedure chosen is the binomial probability distribution, for data sets up to size 30. Use of the binomial probability is difficult for larger sample sizes. And for these larger data sets impairment will be determined by making direct comparison of percent of samples not compliant with the criterion value with the ten percent guideline.

Other Statistical Considerations

Prior to calculation of confidence limits, the normality of the data set will be evaluated. If normality is improved by a data transformation, the confidence limits will be calculated on the transformed data.

Time of sample collection may be biased and interfere with an accurate measurement of frequency of exceedence of a criterion. Data sets composed mainly or entirely of storm water data or data collected only during a season when water quality problems are expected could result in a biased estimate of the true exceedence frequency. In these cases, the department may use methods to estimate the true annual frequency and display these calculations whenever they result in a change in the impairment status of a water.

For waters judged to be impaired based on biological data where data evaluation procedures are not specifically noted in Table 1, the statistical procedure used, test assumptions and results will be reported.

Appendix C
Examples of Statistical Procedures

Two Sample "t" Test for Color

Null Hypothesis: Amount of color is no greater in test stream than in a control stream. (As stated, this is a one-sided test, meaning that we are only interested in determining whether or not the color level in the test stream is greater than in a control stream.) If the null hypothesis had been "amount of color is different in the test and control streams" we would have been interested in determining if the amount of color was either less than or greater than the control stream, a two-sided test).

Significance Level (also known as the alpha level): 0.10

Data Set: Platinum-Cobalt color units data for the test stream and a control stream samples collected at each stream on same date.

Test Stream	70	45	35	45	60	60	80
Control Stream	50	40	20	40	30	40	75
Difference (T-C)	20	5	15	5	30	20	5

Statistics for the Difference: Mean = 14.28, standard deviation = 9.76, n = 7
Calculated "t" value = (square root of n)(mean)/standard deviation = 3.86
Tabular "t" value is taken from a table of the "t" distribution for 2 alpha (0.20) and n-1 degrees of freedom. Tabular "t" = 1.44.
Since calculated "t" value is greater than tabular t value, reject the null hypothesis and conclude that the test stream is impaired by color.

Statistical Procedure for Data Sets of Less than Four for Mercury in Fish Tissue

Data Set: data in ug/Kg 130, 230, 450. Mean = 270, Standard Deviation = 163.7
The 50% Confidence Interval = the sample mean plus or minus the quantity:
(0.676)(163.7)/square root 3 = 63.89. Thus the 50% Confidence Interval is 206.11 – 333.89
Since the criterion value, 300 ug/Kg, falls within this 50% Confidence Interval, this water is judged to be unimpaired by mercury in fish tissue but the waterbody is placed in Category 2B or 3B.

Statistical Procedure for Data Sets of Four or More for Mercury in Fish Tissue

Data Set: data in ug/Kg 130, 230, 450, 350, 220. Mean = 276, Std. Deviation = 124.82
The 60% Upper Confidence Limit = the sample mean plus the quantity:
(0.253)(124.82)/square root 5 = 14.12. Thus the 60% UCL is 290.12 ug/Kg.
Since the Upper Confidence Limit is less than the criterion value of 300 ug/Kg, this water is judged to be unimpaired by mercury in fish tissue.

Appendix D
The Meaning of the Sediment Quotient and How to Calculate It

While sediment criteria in the form of Probably Effect Concentrations[22] are given for several individual contaminants, it is recognized that when multiple contaminants occur in sediment, toxicity may occur even though the level of each individual pollutant does not reach toxic levels. The method of estimating the synergistic effects of multiple pollutants in sediments given in McDonald et al [10] is the calculation of a Probably Effect Concentrations Quotient. This calculation is made by dividing the pollutant concentration in the sample by the Probably Effect Concentrations value for that pollutant. These values are summed and normalized by dividing that sum by the number of pollutants.

Example: A sediment sample contains the following results in mg/kg.

Arsenic 2.5, Cadmium 4.5, Copper 17, Lead 100, Zinc 260. The Probably Effect Concentrations values for these five pollutants in respective order are 33, 4.98, 149, 128, 459.

Probably Effect Concentrations Quotient = ((2.5/33) + (4.5/4.98) + (17/149) + (100/128) + (260/459))/5 = 0.488

Based on research by McDonald (2000) 83% of sediment samples with Probably Effect Concentrations quotients less than 0.5 were non-toxic while 85% of sediment samples with Probably Effect Concentrations quotients greater than 0.5 were toxic. Based on these findings a Probably Effect Concentrations quotient greater than 0.5 will be judged to be toxic**.**

[22] Level at which harmful effects on the aquatic community are likely to be observed.