DRAFT

Standard Operating Procedure

**DWM Data Validation Processes --- Overview**

CN 56.9

Contents

[I. Applicability 1](#_Toc338325019)

[II. Background 1](#_Toc338325020)

[III. Attended Probe Water Quality Data 2](#_Toc338325021)

[IV. Unattended Probe Water Quality Data 2](#_Toc338325022)

[V. Laboratory Data 2](#_Toc338325023)

[VI. Benthic Macroinvertebrate Data 3](#_Toc338325024)

[VII. Fish Population Data 3](#_Toc338325025)

[VIII. Benthic Algae 3](#_Toc338325026)

[IX. Data Validation Reports 3](#_Toc338325027)

[Appendix A. Data QC-Levels 4](#_Toc338325028)

# I. Applicability

The following is a summary of processes and procedures taken by DWM to validate and finalize ambient monitoring data that are used to make designated use assessment decisions, develop TMDLs and other uses. These data include physico-chemical, in-situ probe and biological data. The majority of these data are housed in formal databases, but some are not.

This overview does not address final acceptance of model output. This information is typically contained in the final Modeling Report discussing the validity of model results.

# II. Background

MassDEP-DWM-Watershed Planning Program annually monitors the quality of surface waters (rivers, streams, lakes and ponds) throughout the Commonwealth, based on a 5-year rotating basin cycle. Depending on the sampling design for each project, parameters sampled for may include ambient water quality, benthic macroinvertebrates, fish assemblages, fish tissue toxins, benthic algae, streamflow and aquatic plant communities. For more specific information regarding objectives, sampling design, data quality objectives, SOPs and quality control measures, see DWM’s programmatic QAPP. Example: [CN 365.0 - QAPP\_DWM Monitoring Program\_2010-2014 (rev.1).pdf](CN%20365.0%20-%20QAPP_DWM%20Monitoring%20Program_2010-2014%20(rev.1).pdf).

DWM monitoring activities generate raw data/metadata in the form of paper fieldsheets, logged data files and electronic data deliverables (EDDs) from state and contracted laboratories. These information sources are stored and compiled, then periodically reviewed and critiqued for completeness, accuracy, errors and representativeness. Data meeting acceptance criteria (defined in QAPPs/SOPs and/or BPJ) are determined to be final and considered usable for decision-making. For the majority of DWM data, the data are validated or otherwise finalized outside of and separately from the electronic database housing the data. Every attempt is made to finalize data within an acceptable timeframe and in conformance with approved data timeliness standards.

DWM employs different “validation” procedures for each type of data generated. Given the large quantity of data involved, many of the procedures for water quality data have been automated using VB code and MS Excel macros. The following is a brief summary of procedures for each type.

# III. Attended Probe Water Quality Data

“Attended” probe data are collected when survey staff are physically present at the site. Parameters can include pH, DO, DOsat, SC, TDS and temperature. All raw data are logged electronically, and also manually recorded on fieldsheets (as backup in case of file loss). Data files are periodically uploaded for secured network storage.

For specific procedures to validate attended probe data, see [w\dwm\Data\process SOPs\attended](../Data/process%20SOPs/attended). See CN 56.4. These water quality data are finalized from raw (QC1) through preliminary (QC2-QC3) to final (QC4-QC5). The final attended probe data are provided in MS Excel tables with associated metadata.

# IV. Unattended Probe Water Quality Data

“Unattended” probe data are collected when survey staff are NOT physically present at the site. Continuous data are logged at defined intervals (e.g., every 30 minutes) for parameters such as pH, DO, DOsat, SC, TDS and temperature. Probes are retrieved and uploaded to the network.

For specific procedures to validate unattended probe data, see [W:\DWM\Data\process SOPs\unattended](file:///W:\DWM\Data\process%20SOPs\unattended). See CN 56.5. Final unattended probe data are provided in MS Excel tables with associated metadata.

# V. Laboratory Data

Water samples collected by DWM are sent to one or more labs for analysis. Typical parameters include color, turbidity, nutrients (TP, TN, NH4-N, NO3-NO2-N), chloride, chlorophyll a, bacteria and heavy metals. Following lab analyses and internal lab data validation, labs provide data to DWM as electronic data deliverables (EDD) in a standard EDD format. These data are considered “raw” until validated by DWM. EDDs are stored in a secure network location prior to validation activities.

For specific procedures for validating laboratory water quality data, see [W:\DWM\Data\process SOPs\laboratory](file:///W:\DWM\Data\process%20SOPs\laboratory). See CN 56.6. Final lab data are provided in MS Excel tables with associated metadata.

NOTE: Fish tissue samples are also collected by DWM sent to the state lab for analyses, which may include mercury, PCBs and pesticides. Because these data need to be used as soon as possible following lab analysis by MDPH and others, these data are not validated by DWM. In this case, the data issued by the lab are final.

# VI. Benthic Macroinvertebrate Data

Benthic macroinvertebrate samples are collected by DWM to assess aquatic habitat. Samples are sorted, enumerated and identified to appropriate taxonomic level(s). Data are finalized as described in the benthic program QAPP ([CN 226.0 - QAPP\_Benthic Macroinvertebrate Program.pdf](CN%20226.0%20-%20QAPP_Benthic%20Macroinvertebrate%20Program.pdf)) by biologist’s evaluation of the results of quality control sampling and expert confirmation of split samples and voucher specimens, and best professional judgement. Final data (meeting defined data quality objectives) are used in a metrics-based scoring system. Results are typically presented in Tech Memos.

# VII. Fish Population Data

Using electroshocking techniques, fish samples are also collected by DWM to assess aquatic habitat. Live samples are enumerated and identified to species level. Data are finalized as described in the fish population program QAPP [(CN 096.0 - QAPP\_Fish Toxics Program.doc](CN%20096.0%20-%20QAPP_Fish%20Toxics%20Program.doc)) by biologist’s best professional judgement with regard to factors affecting survey data quality, such as survey efficacy (e.g., catch per unit effort) and accuracy of identifications. Final data meeting defined data quality objectives are used to evaluate the health of the fish community in relation to reference and/or expected conditions. Results are typically presented in Tech Memos.

# VIII. Benthic Algae

Periphyton (AKA benthic attached algae) samples are collected by DWM to assess aquatic habitat and impacts due to excess nutrients. In-situ measurements are taken and samples are collected for identification, density and chlorophyll a content. Biological data are finalized by biologist’s best professional judgement with respect to accuracy of identifications, adherence to SOPs and other factors potentially affecting data quality. Laboratory-based chlorophyll a data are validated as described above in Section V. Final data meeting defined data quality objectives are used to evaluate the waterbody health in relation to reference conditions, and are typically presented using Tech Memos.

# IX. Data Validation Reports

To document the methods used and the results of data validation procedures, summary data validation reports can be generated by DWM’s QA Analyst. Since water quality data are typically validated in year-sets, these reports are produced annually.

# Appendix A. Data QC-Levels

**QC Status 1:**

Raw data. Generally not suitable for use or transmission to other parties, but can be transmitted to other parties upon request provided data are sent as “DRAFT” with standard disclaimers.

**QC Status 2:**

Draft data that has been entered into the appropriate DWM electronic system or database and for which data entry QC has taken place. This stage is for technical QC review.

**QC Status 3:**

Draft data for which technical QA/QC review (e.g. QC sample results, outlier identification, comparison to project QAPP DQOs, etc.) has taken place. This stage is for project-level review.

**QC Status 4:**

Final Data. This level of data reflects project-level review by appropriate staff for reasonableness, completeness and acceptability. These data can be freely used and cited in documents without caution or caveat (reviewed and approved by all appropriate DWM staff).

The following guidelines pertain to receipt and use of QC Status 4 data:

1. When using, analyzing, presenting or transmitting QC4 data, do not make any changes affecting CONTENT, including symbols and qualifiers used, censoring decisions, etc.
2. When presenting data, provide KEY to symbols and qualifiers used.
3. See final data file “READ ME” sheets for additional information.

**QC Status 5:**

Final data in a published, citable report. The QC Status 4 guidelines stated above apply to the data contained in a report. As for QC4-level, these data have been reviewed and approved by all appropriate DWM staff.