



MassDEP

**Massachusetts Department of Environmental Protection
Bureau of Water Resources
Division of Watershed Management
Watershed Planning Program**

STANDARD OPERATING PROCEDURE

Water Quality Data Validation Procedures
for LABORATORY DATA

CN 56.62

June 2025–June 2027

Revised: 6/12/2025

**Prepared and
Approved by:**

/s/ Kari Winfield

Kari Winfield, Environmental Analyst

Date: 6/12/2025

Approved by:

[Signature]

Jasper Sha, QA Analyst

Date: 6/17/25

Approved by:

[Signature]

Richard Chase, Section Chief, Data Management
and Water Quality Assessment

Date: 7/18/25



List of Revisions

Revision Date	Revision	Pages #s	CN/ (Old CN if applicable)	Initials
10/4/2018	Original			RFC/KW
6/12/2025	Entire document			KW/SF



Table of Contents

I	Applicability	4
II	Overview	4
III	Prerequisites for Initiating Procedures.....	4
IV	Assumptions	5
V	Roles and Responsibilities	5
V	Data Validation Procedures for Laboratory Data	6
APPENDIX A: Validation Decision Criteria.....		14
APPENDIX B: Reporting rules for Laboratory Analytes		16
APPENDIX C: Manual EDD Creation for non-WES Laboratories		18
APPENDIX D: QC3 Review Guidelines for Laboratory Data		21
APPENDIX E: Symbols and Qualifiers Used for WPP Laboratory Data		22



I Applicability

These procedures apply to water quality samples collected by MassDEP's Watershed Planning Program (WPP) monitoring staff for laboratory analysis at selected sites (streams, rivers, lakes and ponds throughout the State). Water quality samples may be collected as "grab" samples using water bottles, or with equipment such as tygon tubing for integrated samples or Van Dorn bottles. Analyses depend on the individual project's goals, but may include nutrients, bacteria, microbiology, chloride, personal care products, PFAS, metals, physico-chemical (color, hardness, turbidity, solids), etc. Resulting analytical data are provided to WPP as Electronic Data Deliverables (EDDs) from the Wall Experimental Station (WES) state laboratory, WPP's laboratory, and/or contract laboratories via their individual Laboratory Information Management Systems (LIMS). While these data are "final" from a laboratory perspective, they undergo further review by WPP to evaluate data quality and completeness.

II Overview

These procedures document WPP's approach to organizing, processing, validating, and finalizing water quality data from laboratories using automated and manual processes applied to individual-year data sets by WPP data management and quality assurance (QA) staff. EDDs are compiled for all laboratories, compared to WPP field sheet records for completeness, and then reviewed against acceptance criteria, based on Quality Assurance Project Plan (QAPP) data quality objectives. "Technical" QC checks on laboratory data include field duplicate precision, field blank "hits", frequency of QC sample collection per survey "trip", violations of analytical holding times, non-adherence to field and/or laboratory methods, tidal influences, non-representativeness, analytical accuracy issues, and other factors affecting data quality. Based on checks against acceptance criteria, data are either accepted "as is", qualified or censored. Following technical QC review, data are provided to Principal Investigators for further review prior to finalization using standardized output formats.

III Prerequisites for Initiating Procedures

The following information is required in order to initiate these procedures:

- 1) Raw EDDs from laboratory LIMS systems
- 2) Final laboratory data reports
- 3) Copies of sample Chains of Custody (COCs)
- 4) Proofed electronic field sheet files (entered in relational database, such as the Water Quality Database "WQD" or EDGE/EQuIS)
- 5) Validation Decision Criteria (accept vs. qualify vs. censor) (see **Appendix A**)
- 6) WPP reporting rules for the data year (rounding rules, significant figures) (see **Appendix B**)
- 7) Final WPP site or "station" information (waterbody names, location descriptions, and geo-references)



- 8) Program software/languages (MS Windows based): MS Excel, MS Access, Visual Basic for Applications (VBA), Visual Basic Script (VBScript)

Working files are managed using MS Excel (.xlsx or .xlsm, if macro-enabled). Data management and QA staff should be well-versed in MS Excel, VBA and VBScript coding, and process documentation

IV Assumptions

These procedures have the following underlying assumptions:

- 1) Proofed field sheet files are accurate (based on 100% data entry QC), until shown to be inaccurate. Any required changes to field sheet information are documented and transferred to other databases as needed.
- 2) Laboratory EDDs are accurate until shown to be inaccurate (including that reported laboratory data have met internal laboratory QC acceptance limits). Any required changes are documented as part of the validation process but are not coordinated with the lab—the WPP QC4 data files supercede the laboratory-provided data.

V Roles and Responsibilities

Field Staff are responsible for conducting water quality surveys (under pre-defined project names) where discrete probe measurements are collected, following appropriate SOPs for field methods. They are responsible for accurately filling out field sheets that track probe serial numbers, dates/times of measurements, sample IDs (OWMIDs), and noting any field issues either on the field sheet directly or by emailing the Data Manager, Quality Assurance Officer, and/or Field & Lab Operations Coordinator. Each **Principal Investigator** in charge of a water quality project should submit Probe Request forms to the Field & Lab Operations Coordinator in advance of the monitoring season for planning and calibration purposes.

The **Field & Lab Operations Coordinator** is responsible for preparing probes and loggers for field use, calibrating probes prior to surveys, and completing probe request forms with serial numbers assigned to projects for specific dates/locations. Data files from data loggers are also downloaded and stored in a commonly accessible file-sharing space

The **Data Manager** compiles field sheet information from appropriate data bases, runs automated scripts to process/check probe data files, and prepares files for review by the QA Officer.

The **QA Officer** performs manual quality assurance checks on probe data and overrides automated decisions where necessary.



V Data Validation Procedures for Laboratory Data

1) Laboratory EDD compilation (manual) Role: QA Officer

- a) When all surveys have been completed and the resulting laboratory data received for the data year-set, export final WinLIMS data from the WES laboratory in .csv format. Refer to “QuickGuide WinLIMS data download- June 2021” for viewing and downloading results or to checking the approval status of data.
- b) Assemble all “other” (non-WES) lab data EDDs for the year-set. The EDDs can be from the WPP lab, contract labs, EPA lab, etc. (as applicable).
- c) For data from the WPP lab, check that the electronic files have been fully checked against the printed bench sheets for any transcription errors before starting EDD compilation.
- d) Compile non-WES EDDs (see **Appendix C** for full directions):
 - i) Compile all EDDs from “other” labs into one Master EDD file using the standard EDD template format (EXCEL). Save to: DEP BWR\DATA\Raw Data\yyyy\labs. Rename files to current date when new data are added.
 - ii) Review master EDD file for date errors, dilution rate vs. result errors, dilution rate vs. MDL errors, etc. and make corrections.

2) Produce QC1 combined EDD (automated/manual) Role: Data Manager

- a) Run VB script **EDD2LIMS.wsf** (.wsf = Windows script host file) to create a standardized output file format (automated)
 - i) Combines Master EDD file with the extracted WinLIMS DATA file to create one ALL-LAB-DATA file (EXCEL)
 - ii) Combined file is automatically saved in \Data\laboratory_QA\YYYY\LIMS-EDD Data\combined, where YYYY is the 4-digit year being processed.

3) Review ALL-LAB-DATA file to ensure consistency (manual) Role: Data Manager

- a) Erroneous differences due to syntax (e.g., remove spaces as needed so that same data are made equivalent) (this is done in the code)
- b) SampleID (aka OWMID)
 - i) must start with the 7-character OWMID (XX-XXXX) in order to be properly matched to the field sheet meta data OWMID
 - ii) any characters preceding the OWMID must be omitted, trailing characters (e.g. for bottle groups) are OK
 - iii) any character substituted for the dash must be corrected
 - iv) after the field sheet cross referencing is completed, the reviewer has another chance to catch any unmatched OWMIDs, correct the combined ALL-LAB-DATA file and re-run the code
- c) Remove non-project data



- i) Special projects or regional data without field sheet meta data
 - ii) Analyte-specific data not being validated or for other projects (e.g. fish toxics data)
 - iii) Laboratory QC data or other audit samples
 - d) Add missing values to ANALYSISTIME column if available through the lab report OR correct the format of the date/time columns
 - e) Correct blank cells
 - f) Where corrections are needed, MAKE CHANGES TO THE INDIVIDUAL WinLIMS and/or EDD MASTER FILES (not the combined file) and repeat step 2a)
- 4) Cross-reference the ALL-LAB-DATA file with the field sheet meta data (Automated/Manual) Role: Data Manager
- a) Run VB code **LabDataXREF.wsf** (automated)
 - i) Matches OWMIDs from ALL-LAB-DATA file to OWMIDs from electronic field sheet meta data file for the data year being processed
 - ii) Generates list of “Unmatched FS OWMIDs”—those OWMIDs in electronic field sheet not found in the laboratory data set
 - iii) Generates list of “Unmatched Lab Cust Sample Num”—those OWMIDs in the ALL-LAB-DATA file not found in the electronic field sheet file
 - iv) Creates hold time shell with laboratory codes, analytes, methods from ALL-DATA-LAB file
 - v) Creates NOSAMP template for adding missing samples (for entire bottle group or for individual analytes)
 - b) Reconcile “unmatched” OWMIDs from field sheet or ALL-LAB-DATA file (manual)
 - i) Identify any data not to be validated and move to separate Excluded Data worksheet (e.g. special projects, non-WPP data, laboratory QC or other audit data)
 - c) Perform additional error checks
 - i) Check that sample type is correct for laboratory samples (verify against paper field sheet as needed)
 - ii) Check that collection date reported by lab matches field sheet start (i.e. sampling) date
 - iii) Check that collection time reported by lab matches field sheet start (i.e. sampling) time
 - iv) Check that analysis date/time reported by lab occurs after field sheet sampling date/time
 - d) Make necessary changes to the field sheet metadata file, the LIMS extract and/or the EDD master file as needed; track changes made in the QC1 Comments column in the ALL-LAB-DATA file or individual lab EDD files
 - e) Rerun program in 2a) and then 4a) for each year as necessary until issues have been completely addressed
- 5) Following creation of QC1 file and error reconciliation, populate Hold Time file, Units file, and NO SAMP file for year being processed (manual/automated) Role: Data Manager/QA Officer
- a) Hold time file (manual):
 - i) Save a copy of the hold time criteria file with today’s date
 - ii) Open the hold time template file located in the QC1 folder, copy and paste the “hold time shell” worksheet to the hold time workbook on a new sheet labeled with the data year



- iii) Populate the right hand columns with the final parameter name, final method, bottle group, hold times (most data can be copied from the prior year's sheet or from earlier sheets if necessary)
 - b) Units file (manual):
 - i) Save a copy of the units file with today's date
 - ii) Select UNIT column in the QC1 file and copy and paste to a new blank worksheet in the Units file
 - iii) Select newly pasted column, click on the "Data" tab on the MS Ribbon, select "Remove Duplicates" and select "my data has headers" in the resultant pop-up box, click OK
 - iv) Add column for standardized units spellings called "UNITS_WPP"
 - v) Populate UNITS_WPP column with standardized final units
 - vi) Rename worksheet to year being processed, e.g. "2021"
 - c) NOSAMP file (automated/manual)
 - i) For samples not collected (as indicated in field sheet meta data file), fill in all bottle groups (using full name indicated in hold time file Bottle Group column) on "OWMID-Analytes" sheet (comma-separated)
 - ii) Run VBA macro called **add_missing_data** (automated)
 - (1) Creates a sheet called "Year_lab_NOSAMP_QC1" containing cross-referenced field sheet meta data and lab chemistry data for pre-identified missing lab samples
 - (2) Adds analytes based on bottle group indicated on "OWMID-Analytes" sheet using hold time file
 - iii) Remove any analytes not needed from "Year_lab_NOSAMP_QC1" sheet
 - d) Update the above worksheets/files as needed when new data are added to ALL-LAB-DATA file during the OWMID reconciliation process
- 6) Produce QC2 file for QA review (automated) Role: Data Manager
- a) Run VB code **LabDataQA.wsf**
 - i) Generates *Year_laboratory_QC2.xlsm* Excel file based on QC1 file
 - ii) Creates *Year_laboratory_QC2* working sheet with columns added for
 - (1) final parameter and method (from Hold Time file)
 - (2) final units (from Units file)
 - (3) qualifiers
 - iii) Generates automated decisions ("ACCEPT", "QUALIFY" and "CENSOR") for the "basic-4" QC issues: field blank contamination (b), field duplicate precision (d), frequency of QC samples (f), and analytical sample holding time violations(h) (see **Appendix A** for acceptance criteria)
 - iv) Adds any missing sample data by analyte (from separate NOSAMP file)
 - b) Review automated decisions for b, d, f, and h qualifiers and make changes as needed to the "b Result", "d Result", "f Result", and "h Result" (using Accept, Qualify, or Censor keywords)
- 7) Review results in QC2 file (manual) Role: QA Officer
- a) General notes:
 - i) Automated-QC decisions are conservative by design and based on set criteria; use BPJ to re-evaluate individual automated decisions. Justify any manual changes by adding comments (with initials and date) on *each* row changed to the *QC2 review comments* [column EO] in the working file.
 - ii) "Trip effects": in general if a QC sample (field duplicate or blank) is censored, the other results for that parameter in the same field trip (same Field Sheet Log # listed in columns



[D] and [AY]) get qualified. If the QC sample is only qualified, the other trip samples remain “accept”.

- iii) Extending “trip-effect” censoring: for QC problems that indicate a broader problem (e.g. consistent lab problems with an analyte on a given day) consider extending the qualification or censoring to all the associated samples.
- b) Review and edit the **Year_laboratory_QC2** file, making any changes to the “Year_laboratory_QC2 working” tab. Highlight changes, and document reasoning in QC2 comments column; include your initials and the date in *each* comment
 - i) **Review automated decisions/edit as needed** for: hold time (“h”) [DU], field blank contamination (“b”) [DX], field duplicate precision (“d”) [EA], and frequency of QC samples (“f”) [EC]. Make changes by overwriting decisions for these fields to Accept, Qualify, or Censor as appropriate. Considerations:
 - (1) Hold times for samples that were frozen (for analytes that are not routinely frozen per the SOP) can be recalculated manually to not include the time frozen in the HT calculation. HT exceedances are calculated as $((QA-HoldTime/QA-HoldTimeMax)-1)*100$. If necessary, consider greater flexibility for analytes that are relatively stable (TN, TP, NO₃/NO₂, and chloride). e.g. when WES was having significant problems with meeting nutrient analysis HTs in 2016, relaxed decision points were used in order to retain some data.

For bacteria samples where the HT exceedance is less than about 2 hours, consider changing the auto-decision to *qualify*; literature supports allowing longer hold times for non-regulatory E. coli samples.
 - (2) Duplicates near the MDL: “d” decisions may be changed from *censor* to *qualify* for duplicate results near the MDL, based on a “low-number-effect” (high RPDs for low value numbers). DQOs for some analytes include absolute differences (as documented in the WPP Programmatic QAPP), for other analytes use BPJ.
 - (3) Project-specific QC frequencies: Generally frequency decisions are based on 1 QC per day or ~10% of the day’s samples. For some projects, for example lake sampling where only two lakes are visited per day, lower frequency of QC samples may be acceptable by design. Check with project coordinators about the project design.
 - ii) **Review Secchi depth (SECCHI)** and associated meta-data SECTIME, SECVIEW, SECBOT, and MAXDEPTH (columns AM – AQ) for lake field sheets. Add qualifiers as needed in **Secchi Qualifiers** column (EK) and populate **Secchi Decision** column (EL). Because Secchi Depth applies to all the analytes for a given OWMID, apply any Secchi-depth qualifies/decision to all analyte rows associated with OWMID. For river field sheets, where there are no Secchi readings, fill in “N/A” for all Secchi Decisions (all Secchi Decisions must have something filled in). Check the following:
 - (1) If the Secchi disk hit bottom (SECBOT = yes), the reported Secchi result may be an underestimate of the true condition (if “>” is not used); qualify or censor “a.”
 - (2) If Secchi depth is \geq max depth, qualify the reading “e” (not theoretically possible).
 - (3) Check **Secchi duplicate readings** on *the original field sheets* (duplicate readings are not recorded in the spreadsheet):



accept duplicate readings within 10% RPD
10 – 20% RPD (qualify “d”)
>20% RPD (censor “d”)

(4) Check for readings made outside the approx. “solar noon” time (+/- 3 hours) of 10am-4pm (qualify “m”).

- iii) **Populate a/e/j/m/p/r/t decision columns with initial “accept” for all data.** These will be adjusted as needed based on further review.
- iv) **Transfer lab qualifiers** [column CO] and review other laboratory internal QC data, lab comments, etc. that relate to data quality, such as lab QC outside acceptance limits for accuracy (“a”). Most lab qualifiers can be transferred directly to WPP data using the same or similar WPP qualifier symbol, e.g. various versions of “j” qualifiers from the lab (J1, J2 etc., would be carried forward as “j”. Review both the laboratory EDDs and official reports (usually in PDF format), to ensure that qualifiers have been carried forward from the PDF report to the EDD and applied to all results from that analysis batch; add if needed. (Note: Alpha Analytical does not automatically carry forward all their qualifiers.)
- v) **Apply “e” quals/censors for results not theoretically possible.** Examples include: “Total” vs. “Partial” results for split samples (e.g., fecal coliform < E. coli (a fecal coliform organism), dissolved fraction > total, Secchi depth > max. depth, etc. Use BPJ for qualify vs. censor.
- vi) **Apply “t” quals/censors.** Review location information for coastal samples to determine potential for tidally-influenced data. Check site conditions in ArcGIS or Google Maps/Earth and consult with the project coordinator for more information. Use BPJ for qualify vs. censor.
- c) The Final Decision [column EN] will be automatically populated once all the individual “decision” columns are filled (h, b, d, f, a, e, j, m, p, r, and t) and will update with any changes. If a Final Decision is blank when you’ve finished, check that all the decisions have something selected.
- d) **Review fieldsheet comments /apply decisions as needed.**
 - i) Fieldsheet comments [column AT] often indicate real or potential effects to data quality, such as tidal conditions (t), lack of adherence to method (m), sample preservation issues (p) and lack of representativeness (r).
 - ii) In addition to fieldsheet comments, brackish or saline conditions can be verified using attended probe conductivity data, if available. Use BPJ for qualify vs. censor.
 - iii) **newFlowStat:** check field sheet comments for any indications of stagnant or dry conditions; change the newFlowStat if necessary.
- e) **Misc. communications** (e.g., e-mail): review emails and other communications related to data quality. Emails are filed with the raw data for the year (e.g. [WPP_WQData_2005-2022.mdb](#))
- f) **Review any field audits /apply decisions as needed.** Field audits may reveal issues related to data quality, such as adherence to sample collection procedures. Use BPJ for qualify vs. censor.
- g) **Review external laboratory audit results/apply decisions as needed** for potential accuracy (a), precision (d) or other problems. These are typically QC samples (Certified Reference Materials



from ERA Waters) prepared by WPP and sent to the lab(s) or purchased Proficiency Testing (PT) studies sent to the lab by a 3rd party. Use BPJ for qualify vs. censor.

- h) **Remove any QC sample data** remaining that was not previously removed during QC1 data processing. Place removed data rows into the “excluded data” folder.
 - i) Review each qualifier column for **3 or more qualifiers**. This may indicate the need for censoring if the datum was not already (automatically) censored.
 - j) **Correct newResult column [DL] as needed**, based on any field sheet comments (e.g. switched samples), obvious outliers or erroneous results, etc. If the result is “Missing”, it may be able to be recovered. Make a note in the newResultFlag column [DM] for any changes.
 - k) **Review auto-trip-effect decisions** (censoring of QC sample(s) results in minimum qualification of associated trip samples). Note: If manual changes are made to decisions for QC samples (e.g., from censor to qual), subsequent additional manual changes may be necessary to the associated trip samples (change to accept or leave as qualified).
 - l) **Check “final decision” column for completeness** (all rows are auto-filled), including confirming that “No Result” decisions are based on missing data.
 - m) **Review all CENSORED data** for final acceptability.
 - n) **Review and update laboratory QC “read me” files** (both QC2 and QC4 read me) from the most recent version.
 - o) **Update the year-specific Reporting Rules and Hold Time file (separate files through 2021, combined in 2022)**; copy from previous year and update. Method numbers and MRL/MDLs are laboratory-specific (and can be checked in official data reports and SOPs from the labs). Hold times are lab and method-specific and may be defined in the method SOP or confirmed with the lab. Reporting levels (number of decimals reported to) and rounding rules are defined by WPP.
- 8) Finish QC2 Review and generate final files for Principal Investigator review (manual/automated)
Role: Data Manager
- a) Check that all QC2 changes by QA Officer have been documented with comment in the QC2 Comment column (with initials and date) and highlight the actual change
 - b) Review reporting rules file and add any new rules to VBA macro **FinalizeLabData**
 - c) Check for any new columns in the QC2 file format (typically due to additional bottle groups in the field sheet meta data file) and adjust variables as needed in VBA macro **FinalizeLabData**
 - d) Run VBA macro **FinalizeLabData** (automated):
 - i) Applies formatting and reporting rules (see **Appendix B** and **Appendix E**)
 - ii) Adds station information (Watershed, Water Body, Station Description, Mile Point, Latitude, Longitude) from separate Excel file
 - iii) Creates worksheets called: “Year_lab_QC2 final all data” (all sample types), “Year_lab_QC2 final routine” (routine samples only), “Year_lab_QC2 final field blanks” (field blank or trip blank samples only), “Year_lab_QC2 final duplicates” (samples paired as duplicates)
 - iv) Creates worksheets called: “Year QC2 lab data summary” (summarizing result counts by project and analyte group/sample type, “Year QC2 lab qualifier summary” (summarizing result counts by analyte, QC status (accept, qualify, censor), and qualifier for each project)
 - v) Creates individual project files for each project



- (1) Adds QC2 “read me” tab imported from separate Excel file
- (2) Adds “lab data - main”, “lab data – duplicates”, and “lab data- field blanks” sheets with final data formatting
- (3) Adds “lab data summary” sheet with qualifier counts by analyte
- e) Review “YYYY_laboratory_QC2 final all data” sheet and “YYYY QC2 lab data summary” sheet, making changes to the working sheet or the macro if necessary (DO NOT make changes to the final sheets, these are meant to be based solely on the data as presented on the working sheet and these sheets get overwritten when the macro is re-run)
 - i) Check for blank cells: the only columns that can have blanks are Mile Point, Secchi Depth qualifiers, Result Qualifiers, and QC OWMID
 - ii) Check for “rounding not applied” results
 - iii) Check for “<” or “>” that require manual rounding (in the working sheet) (e.g., <0.02000)
 - iv) Check that reporting rules have been applied to all data
 - v) Check final data file for center alignment/left justified cells (all)
 - vi) Random QC checks (optional)
- f) Run visual basic script **LabData_QC3.wsf** (automated)
 - i) Generates separate QC3 file based on the QC2 file format to track QC3 review changes
- g) Provide files by project to Principal Investigators for QC3 review

9) QC3 Review Process (manual) Role: Principal Investigators for each project

- a) Staff principal investigators review QC2 FINAL files for issues/errors (see **Appendix D** for additional detail)
 - i) Examine project file for completeness (are all data that were collected present; if not, why? Do any data appear to be missing?)
 - ii) Are data presented reasonable based on field conditions and professional judgment (e.g., are there outlier data that are not “real”)?
 - iii) Examine project file for any errors and inaccuracies you may find
 - iv) Other “problems”
- b) Staff principal investigators submit their QC3 review comments to QA Officer

10) Review comments received during QC3 review (manual) Role: QA Officer/Data Manager

- a) Compile staff QC3 (project-level) comments in a separate sub-folder under “QC3”.
- b) QA Officer makes necessary changes to “Year_laboratory_QC3 lab working” sheet in the QC3 Excel file, including QC3 changes sheet (where needed, for general changes that apply to all projects) and QC3 Comments column
- c) In collaboration with Data Manager, write a response to the QC3 comments for the Principal Investigators.
- d) Update the QC4 “read me” Excel file for changes in processes (by data year).

11) Produce final QC4 data files (automated) Role: Data Manager

- a) Once all changes have been made, make sure there are no blank cells in content-required fields, then run the macro “**FinalizeLabData**” to generate the QC4 files



- b) Resolve any errors and re-run macro if needed (overwriting previous files). Once the macro is finished, review files to make sure formatting is correct

12) Check QC4 data files (Manual) QA Officer

- a) Look for any missing information
- b) Filter the data columns to look at range of reported values, missing values, or other issues
- c) Review the summary of qualifiers



APPENDIX A: Validation Decision Criteria

WPP Validation Criteria for Field QC results on Laboratory Data

FIELD BLANKS			
<MDL	=MDL	>MDL	>>MDL
accept	accept	b (qualify)	b (censor, BPJ and/or >2X RDL)
NOTES: 1) If "ND", then OK (show as "<MDL value" (if available), then "<RL value"; if no MDL or RL, flag for error check 2) If <X, OK 3) "Trip Effect": If field blank censored, qualify all other same-analyte samples for that survey trip			

QC FREQUENCY		
FB and FD (both collected)	FB or FD (only one type collected)	No field QC collected on crew trip
accept	accept	f (qualify)
NOTES: 1) In general due to most programmatic QC samples meeting acceptance limits, censoring for lack of QC samples for any given trip is avoided; use BPJ in cases of chronic disregard of QC sampling to decide if censoring for such an extreme case is warranted.		

HOLDING TIME			
<HT	=HT (based on one decimal place)	>HT	>>HT
accept	accept	h (qualify)	h (censor, BPJ and/or >HT+10%)
NOTES: 1) Use hold time lookup table for specific year being validated 2) If no analysis date AND time, at least qualify 3) If no analysis time, interpolate to add times using 5 min increments (if start and/or end known), OR add 5 min sequentially onto last time provided for each sample with unknown time within a batch (to master LIMS/EDD file). Adding 1/2 to 1 hour for lunch not needed (not that critical) 4) All times using 24 hour time format			



FIELD DUPS (bacteria)															
<= 50 CFU/MPN				51-500 CFU/MPN				501-5000 CFU/MPN				>5000 CFU/MPN			
+/- 50CUF or <30% RPD	=30% RPD	>30% RPD	>>30% RPD	+/- 50CUF or <20% RPD	=20% RPD	>20% RPD	>>20% RPD	<10% RPD	=10% RPD	>10% RPD	>>10% RPD	<5% RPD	=5% RPD	>5% RPD	>>5% RPD
accept	accept	d (qualify)	d (censor, BPJ and/or 2X DQO (60%))	accept	accept	d (qualify)	d (censor, BPJ and/or 2X DQO (40%))	accept	accept	d (qualify)	d (censor, BPJ and/or 2X DQO (20%))	accept	accept	d (qualify)	d (censor, BPJ and/or 2X DQO (10%))

FIELD DUPS			
<20% RPD (or < absolute difference for specific analytes)	=20% RPD	>20% RPD	>>20% RPD
accept	accept	d (qualify)	d (censor, BPJ and/or 2X DQO (40%))
NOTES: 1) For "<X" results, use "X" in RPD calculation (if "0" result, use MDL (or minimal if no MDL provided) value for calculation and result 2) "Trip Effect": If field duplicates censored, qualify all other same-analyte samples for that survey trip. BPJ used to determine if trip samples should also be censored. 3) Manually review all qualify and censor auto-decisions for low number effect (LNE); edit decisions using BPJ as necessary 4) For bacteria data, a separate table is used and RPDs are calculated using log-10-transformed data.			
FIELD DUPS (Secchi)			
<10% RPD	=10% RPD	>10% RPD	>>10% RPD
accept	accept	d (qualify, BPJ)	d (censor, BPJ)
NOTES: 1) Manually review field sheets for duplicate Secchi readings, manually calculate RPDs, and apply qualifiers as needed. 2) Applicable when duplicate Secchi readings are required by SOP and available for evaluation			



APPENDIX B: Reporting rules for Laboratory Analytes

Each lab that WPP employs may have different reporting conventions and there may be errors and/or inconsistencies in data presented in EDDs. As a result, WPP has established standard reporting rules for data deliverables, which ensure consistent data presentation with regard to issues such as significant figures, measurement units, formatting conventions and rounding rules. The majority of these rules are automated via VB code and/or Excel macros.

- 1) Rounding to remove insignificant digits is performed as needed based on Standard Methods, 21st Edition guidance. Rounding procedures are also applied to data to account for appropriately decreasing resolution as concentration ranges increase.
- 2) Where dilutions have occurred at the lab, the corresponding increases in the method detection limit and upper quantitation limit are typically accounted for in the lab result, but not always (e.g., if a lab reports “ND” rather than “<10” for a 10X dilution with a MDL of 1; in this case the translation of “ND” equates to a “<10” result, not “<1”).
- 3) In cases where different lab methods are used for the same analyte resulting in different resolution capabilities, the data are reviewed to ensure that the application of standard reporting rules is appropriate.
- 4) Non-detect results shown as “ND” are transformed to less than the numeric method detection limit for the analyte in question (e.g., <0.005).
- 5) By convention for field duplicate results, the lower OWMID# is taken as the primary result for reporting and the higher OWMID# is made the duplicate result, which is only presented in the context of quality control (precision).
- 6) For bacteria results of “TNTC” (too numerous to count), the result is transformed to greater than the upper quantitation limit (UQL) (e.g., >1600 CFU/100 mls.) defined in the original EDD.
- 7) Non-censored but qualified results are shown with the qualifier symbol immediately adjacent (to the right in a separate column) to the numeric result.
- 8) The data elements for final data tables include the following metadata fields:

RIVERS	LAKES
Project	Project
Watershed	Watershed
SARIS_PALIS_CAMIS	SARIS_PALIS_CAMIS
Water Body	Water Body
Unique ID	Unique ID
Station ID	Station ID
Station Description	Station Description
Station Type	Station Type
Mile Point	
Latitude (dec - Degrees)	Latitude (dec - Degrees)
Longitude (dec - Degrees)	Longitude (dec - Degrees)
Field Sheet Log	Field Sheet Log



RIVERS	LAKES
Field Sheet Type	Field Sheet Type
Sample OWMID	Sample OWMID
QC OWMID	QC OWMID
QC Type	QC Type
Sample Date	Sample Date
Sample Time	Sample Time
Flow Condition	Lake Level
	Max Depth (meters)
	Sample Depth (meters)
	Relative Sample Depth
	Secchi Depth (meters)
	Secchi Depth Qualifiers
Analyte	Analyte
Units	Units
Result	Result
Result Qualifiers	Result Qualifiers
Analysis Method	Analysis Method
Special Notes	Special Notes

- 9) All analytes are shown in one column called “Analyte”. This format differs from probe data tables where each individual analyte is provided in separate columns.
- 10) Sample depths for river samples are not provided since most of these samples are collected just beneath the water surface (approx. 0.2 meters). Fieldsheet comments indicating collection at depths significantly different from the assumed depth should be noted in the Special Notes field.



APPENDIX C: Manual EDD Creation for non-WES Laboratories

Summary: Manual steps to aggregating and formatting the EDDS (electronic data deliverable) of results from non-WES laboratories. The final format for the non-WES EDD should be consistent with the format WPP requests from external labs. Columns are as listed below. (use previous year's non-WES EDD as a reference).

		Format	Required
LabID	Laboratory Name	General	Yes
LabSNum	Laboratory Sample Number	General	Yes
FieldSampNum	Field/Client Sample Number	General	Yes
Analyte/Characteristic	Analyte	General	Yes
Sample Fraction	Fraction associated with analyte	General	Yes
Result	Result value	Text	Yes*
LabQual	Laboratory Qualifier	General	Conditional
ResComm	Result Comments	General	Conditional
Units	Analyte/Characteristic Units	General	Yes
MDL	Minimum detection level	Text	Yes*
RDL	Reporting detection limit	Text	Yes*
UQL	Upper Quantification Limit	Text	Conditional*
Analytical Method	Analytical Method	General	Yes
AnalDate	Analysis Date	Date	Yes
AnalTime	Analysis Time	Time	Yes
SiteLocator	Site or Station locator information	General	optional
CollectDate	Sample Collection Date	Date	optional
CollectTime	Sample Collection Time	Time	optional

* Must use text field and report result to the correct number of Sig Figs

STEPS

- 1) Find the data: [RawData](#) > yyyy > Labs
Labs may include any contract labs used that year, EPA, Alpha Analytical (may be contracted through WES), and WPP's internal lab (chlorophyll, E. coli, color, and turbidity).
- 2) **Copy** the data to a working directory. **Leave the original raw data unaltered.**
- 3) Combine multiple files for each lab:
 - a) If the results from one lab are submitted as multiple identically formatted spreadsheets, combine the spreadsheets, then adjust the columns and sort out the QC data as needed.
 - b) The "Get and Transform Data" options in Excel offer various ways to do this. See this helpful video: <https://www.youtube.com/watch?v=Oy8WBZCGY1Y> or Google for others. OR
 - c) Use VBA code **mergefiles** to append files from each laboratory (as long as files have the same format, i.e. column layout and worksheet names)
- 4) Adjust the columns and filter out any laboratory QC data as needed. Copy any lab QC data to a separate sheet in the workbook.



- a) **EPA** submits nutrient data with the results and lab QC combined. The sample results are listed as analysis names: Total Suspended Solids in Water, Combined Nitrate & Nitrite, Total Nitrogen, Total Phosphorus in Water as P, Ammonia and N, and Ortho Phosphate in Water as P. The QC results have analysis names including LFB, matrix spike, laboratory media blank, etc.
- b) **Alpha** Analytical submits their laboratory QC in a separate tab with each file.
 - i) It is useful to compile both the results and the lab QC (in separate tabs) to facilitate review of the QC data.
 - ii) Also review the QC results from Alpha's PFD'ed results for any consistent problems (detections in blanks, poor lab duplicate agreement, poor spike recoveries) and any lab qualifiers that were not carried forward into their EDD submission.
- c) For **WPP** results, the QC results also need to be calculated and qualifiers assigned to result batches (if this has not already been done). Because there are a limited number of sheets, it is practical to do the aggregation, sorting of QC data, and calculations manually. Qualifiers "d" (lab precision), "a" (accuracy), and/or "b" (lab blank) should be assigned based on analysis of the lab QC sample results. See Appendix A and the table below.
 - i) Chlorophyll (e.g. [EDD 2022 WPPchlorophyll-COMPILED&EDITED.xlsx](#)).
 - ii) Turbidity and Color (note that color was not analyzed in 2023): (example combined file [TC22-all-batches.xlsm](#)).
 - iii) E. coli (example: [EDD 2022 WPPecoli-COMPILED&EDITED.xlsx](#))

DQOs from WPP's Programmatic QAPP for WPP lab analyses:

Analyte	Method	Range	MDL	MRL	Accuracy	Precision
Chlorophyll a (ug/L)	EPA 445.0 Mod.	0 - 100	0.1	1.0	75-125%	2.0 ug/L or 20%
Color (true) CU	SM 2120C	0 – 500	2	5	80- 120%	<50 CU, 10 CU >50 CU, 20%
Turbidity (NTU)	SM 2130B	1-100	0.2	0.5	95 – 105%	20%
E. coli (MPN/100ml)	SM 9223B	0 – 2420 (undiluted)	1	1	<RDL for negative	+/- 50 CFUs, OR For Log ₁₀ data: <30% (<50 CFU) <20% (50-500 CFU) <10% (500-5000 CFU) < 5% (>5000 CFU)



- 5) Check that all rows have an OWMID identifier. (Talk to Kari about whether to manually change incorrectly formatted OWMIDS; e.g. missing a dash, combined with other information, or missing.)
- 6) Check for obvious outliers (e.g. missing or misplaced decimal point) or typos, especially in the WPP lab data.
- 7) Leave the analyte names as submitted. These will be corrected later in the data processing steps.
- 8) Finally, combine the various lab EDDs into one "master" EDD for all non-WES labs.



APPENDIX D: QC3 Review Guidelines for Laboratory Data

During the QC2 stage, data not meeting WPP validation criteria or otherwise suspect have been either qualified or censored. For QC3-level review, the general goal is to try to avoid having to make changes to the data after they are “published”. The current QC3 process is basically an external check on the data processing and QC2 review. Questions to ask include, but are not limited to:

- 1) Is the dataset **complete**? Are there missing data that are not represented in the QC3 export?
- 2) Are there **errors in the QC2 technical assessment**, or was something missed (e.g., omissions)?
- 3) Do there appear to be any **inexplicable outliers or unreasonable data** that are probably not real?
- 4) Are there any indications that any of the data are not **representative** of the station and site conditions? Is there reason to believe, for example, that one or more samples may have been collected improperly for a specific reason or in error (e.g., in a mixing zone of an outfall, in a backwater pool)? We did look for these potential situations, but the project leads may have information to add to the evaluation.
- 5) Is there **anything “broken” in the data files** (e.g., a feature that does not/did not work)?

For QC3 Reviewers: If you do have comments and/or suggested edits, please DO NOT MAKE CHANGES TO THE DATA FILES, but provide written comments to me so we can document any changes that may be made as a result of your reviews.

Upon completion of the review, any required edits are made by QC2 staff. At that point, the data are considered final (QC4-level) and moved to the “QC4 data” sub-folder, where they will remain. Once published, the data will be at QC5-level (but the final data table will still reside in the QC4 data location).



APPENDIX E: Symbols and Qualifiers Used for WPP Laboratory Data

The following data qualifiers or symbols are used in the MassDEP/WPP data validation process for qualified and censored LABORATORY DATA. Decisions regarding censoring vs. qualification for specific, problematic data are made based on a thorough review of all pertinent information related to the data.

WPP General Symbols:

“ ## ” = Censored data (i.e., data that has been discarded for some reason; check qualifier symbol for cause(s)).

“ ** ” = Missing data (i.e., data that should have been reported, but were not for any reason other than no water).

“ -- ” = No data (i.e., data not collected nor intended)

“ ^^ ” = No water (i.e., a special case of missing data due to dry/no water conditions)

WPP Laboratory Data Qualifiers:

“ a ” = accuracy as estimated at WES Lab via matrix spikes, PT sample recoveries, internal check standards and lab-fortified blanks did not meet project data quality objectives identified for program or in QAPP.

“ b ” = blank Contamination in lab reagent blanks and/or field blank samples (indicating possible bias high and false positives).

“ d ” = precision of field duplicates (as RPD) did not meet project data quality objectives identified for program or in QAPP. Batched samples may also be affected.

“ e ” = not theoretically possible. Specifically, used for bacteria data where colonies per unit volume for e-coli bacteria > fecal coliform bacteria, for lake Secchi and station depth data where a specific Secchi depth is greater than the reported station depth, and for other incongruous or conflicting results.

“ f ” = frequency of quality control duplicates did not meet data quality objectives identified for program or in QAPP.

“ h ” = holding time violation (usually indicating possible bias low)

“ j ” = ‘estimated’ value; can be used for lab-related issues where certain lab QC criteria are not met and re-testing is not possible (as identified by the WES lab only). Also used to report sample data where the sample concentration is less than the ‘reporting’ limit or RDL and greater than the method detection limit or MDL (MDL < result < RDL). Also used to note where values have been reported at levels less than the mdl. Also used for estimated ranges based on known metadata.



" m " = method SOP not followed, only partially implemented, or not implemented at all, due to complications with sample matrix (e.g. sediment in sample, floc formation), lab error (e.g. cross-contamination between samples), additional steps taken by the lab to deal with matrix complications, lost/unanalyzed samples, use of expired reagents and missing data.

" p " = samples not preserved per SOP or analytical method requirements.

" r " = data may not be representative due to circumstances and/or conditions at the time of sampling, including the possibility of "outlier" data.

" s " = field sheet recorded data were used to accept data (i.e., not data electronically recorded in a data logger or in cases where data logging is not possible (e.g., single-probes)).

" t " = tidal influence likely (not indicative of freshwater flow)