DRAFT

Standard Operating Procedure

**DWM Water Quality Probe File Processing and Validation**

**for ATTENDED PROBE DATA**

CN 56.4

Contents

[I Applicability 2](#_Toc310952320)

[II Overview 2](#_Toc310952321)

[III Pre-Requirements to Initiate Procedures 2](#_Toc310952322)

[IV Assumptions 2](#_Toc310952323)

[V Attended Probe Data Processing 3](#_Toc310952324)

[V-a APPENDIX: Outline of automation steps for each visual basic code module 8](#_Toc310952325)

[Att\_QCProcRawFiles7.wsf 8](#_Toc310952326)

[Att\_QCCheckUnmatched9.wsf 11](#_Toc310952327)

[Att\_QCApplyQualFormats10.wsf 12](#_Toc310952328)

[FinalizeAttData.bas 14](#_Toc310952329)

[Att\_QC4.wsf 15](#_Toc310952330)

[FinalizeAttData\_QC4.bas 16](#_Toc310952331)

[V-b APPENDIX: Checks on electronic field sheets/data entry 17](#_Toc310952332)

[V-c APPENDIX: Procedures for pre-processing Attended Multi-Probe Data Files (Hydrolab & YSI) 18](#_Toc310952333)

[V-d APPENDIX: Manual review of lab QC decisions 21](#_Toc310952334)

# I Applicability

These procedures apply to attended water quality probe data generated by DWM. Each year, DWM monitoring staff collect attended (i.e., not deployed for long-term data collection) probe readings for temperature, pH, specific conductance and dissolved oxygen at selected locations throughout the Commonwealth. In many cases, attended data are collected for QC purposes to compare simultaneous, “side-by-side” readings with deployed probes. On an approximate monthly basis, recorded data files from attended data loggers are uploaded to the DWM data management network.

# II Overview

These procedures document DWM’s approach to collating, processing, validating and finalizing water quality probe data files generated via attended monitoring. These procedures involve both automated and manual steps applied to individual year data sets (e.g., all 2008 files) by DWM data processing and quality assurance staff. Essentially, raw probe data files that have previously been uploaded to DWM’s network are assembled, checked for completeness, pre-processed to select “best lines” (attended files), linked to fieldsheet metadata, checked against laboratory QC data, then output to standard format for further review. Based on checks against acceptance criteria, data are either accepted “as is”, qualified or censored. Following review by the Project Manager (or other 3rd party), reports are provided to staff using standardized output formats.

# III Pre-Requirements to Initiate Procedures

The following information is required in order to initiate processing and validation steps:

1. Raw probe data files (attended only)
2. Lab QC summary reports
3. Proofed electronic fieldsheet files
4. Acceptance criteria and decision matrices (accept vs. qualify vs. censor)
5. Program Software: MS Excel, Visual Basic
6. Working file format(s): MS Excel (.xlsx),
7. Working file naming conventions: variable
8. Working folder filing convention: TBD

Database and QA staff should be well-versed in MS EXCEL, Visual Basic (VB) coding and process documentation.

# IV Assumptions

These procedures have the following underlying assumptions:

1. Proofed fieldsheet files are accurate (based on 100% data entry QC), until shown to be inaccurate. Any required changes to fieldsheet information is documented and transferred to other databases as needed.
2. Any decision applied to dissolved oxygen (D.O.) applies automatically to DO saturation (DOsat) and any decision applied to specific conductance applies automatically to total dissolved solids (TDS).
3. Indications of very low stream velocities combined with the probe stirrer in the “OFF” position or not working (as applicable) will result in suspect data.
4. Validation of the unattended data (the SOP for unattended processing and validation is located here: [w/dwm/data/unattended QA/validation/process documentation](file:///\\DEP-FP-WOR-202.dep.govt.state.ma.us\WORKGROUP$\DWM\Data\attended_QA\process%20documentation\process%20documentation)) is contingent on prior validation of the attended QC data, which is used to finalize the former.

# V Attended Probe Data Processing

1. Multiprobe data file pre-processing steps (manual)

* 1. Collect all raw attended data files, both Hydrolab and YSI, in xlsx format
  2. Add columns to allow identification of “best” line for stability in all analytes (T, DO, SpCond, pH) and to apply qualifiers for stability (“u”), making sure not to overlap new columns with any existing data or comment columns (those reserved for “no data” or no probe on board)
     1. For Hydrolab, also copy entire annotation cell into cell under “OWMID Annotation (copy)” column
  3. Identify unique data blocks
     1. Data blocks are identified by unique OWMIDs, date/time differences, or depth differences between adjacent rows in the data file
     2. Separate out data blocks
        1. For YSI, separate data blocks with same OWMID by blank lines to indicate discrete depths in a lake profile survey
        2. For Hydrolab, separate data blocks with same OWMIDs by blank lines AND add “Annotation” to column A in new blank line (missed annotations should be treated the same as lake profile surveys)
     3. Identify best line for each data block, adding Y to column labeled “’Best’ Line? (Enter Y)”
     4. For each best line, evaluate stability of each analyte (T, DO, SpCond, pH) and apply Accept/Qualify/Censor to each analyte’s qualifier column
  4. Place Hydrolab files for all years into one folder and YSI files for all years into another folder within directory AllYears

1. Collect temperature QC data (recorded on field sheet, no electronic file) (manual)
   1. Identify all temperature logger OWMIDs
      1. Filter field sheet by SmpGearName = “Temperature Logger”, make list of OWMIDs and FSLOGs
      2. Remove filters, Filter field sheet by FSLOG = IDs from 2ai
      3. Check paper field sheets
         1. Cases where OWMIDs already assigned: add any missing OWMIDs and associated field sheet info to electronic field sheet
         2. cases where no OWMIDs were assigned: generate list of new OWMIDs and manually assign starting from last available OWMID for each project; add new IDs and any associated field sheet info to electronic field sheet
         3. Create file with temperature -only QC OWMIDs and data from field sheets (date, time, depth, temperature, gear, comments) (file = **combined years raw fieldsheets + TEMP.xlsx**)
2. Processing raw files (automated)
   1. Run visual basic script called **Att\_QCProcRawFiles7.wsf** (.wsf = windows script host file), or latest version
      1. Generates AllYears\_attended\_QC1.xlsx file, containing data from all years (Hydrolab, YSI, Temperature)
      2. Adds all OWMIDs from pre-processed raw data files, including “u” qualifiers
      3. Adds all temperature QC OWMIDs from separate data file, applies “s” qualifier to “FS Data Used” column and “X” flag to indicate “NO DATA” for pH, SpCond, TDS, DO, DOsat
3. Error checking and application of selected qualifiers ( “m”) (automated/manual)
   1. Run visual basic script called **Att\_QCCheckUnmatched9.wsf** (or latest version)
      1. Enter data year to process when prompted
      2. Generates *Year*\_attended\_QC1.xlsx file, containing data from all years
      3. Matches all OWMIDs from data files to OWMIDs in electronic field sheet for year in question
      4. For logged probe data: Automatically calculates time differences between start time in water and best line time, applies “m” qualifiers per **2006-2010 Validation Decision Criteria Tables-DRAFT08-12-11.xlsx**
      5. Generates list of “Unmatched FS OWMIDs”—those OWMIDs in electronic field sheet not found in any data file
   2. Do the following checks by filtering the resulting QC1 spreadsheet (see **OWMID errors-counts by year.xlsx** file for filter instructions)
      1. Check for OWMIDs in data file with NO MATCH FOUND on field sheet
      2. Check for SmpTypName not equal to “In-situ: Attended” or “In-situ: QC-Cal Check”
      3. Check for OWMIDs on field sheet with NO MATCH FOUND in data file
      4. Check for OWMIDs with incorrect date/time in data file
      5. Check for OWMIDs with incorrect date/time in field sheet
      6. Check for duplicate OWMIDs in data file
      7. Check for time differences between data file best line and field sheet start date/time
      8. Check cases where QCType = Unknown
      9. Check cases where attended QC OWMIDs for associated unattended OWMID is not equal to 2
   3. Resolve items found in 3c) by checking electronic field sheet, paper field sheet, etc., and making changes to pre-processed data files and electronic field sheet as necessary, based on identified errors
   4. Rerun program in 2a) and then 3a) for each year as necessary until items in 4c) appear to have been completely addressed
4. Application of lab qualifiers (“i”, “c”) and final data formatting (automated/manual)
   1. Compile list of lab-related sonde issues, listed by project name, calibration date, post-calibration date, sondeID, and loggerID (see **Appendix;** lab QC file = **2005-2010SondeQC.xlsx**), where:
      1. Sondes were calibrated for particular analytes only (for later application of “NO CAL” flag and “- -“ to indicate no data)
      2. Different calibration standards were used for pH and SpCond (other than the typical ones used for a given calendar year) (for later use in applying “c” qualifier, see **2006-2010 Validation Decision Criteria Tables-DRAFT08-12-11.xlsx** )
      3. Analytes violated inaccuracy (“i”) criteria (see **2006-2010 Validation Decision Criteria Tables-DRAFT08-12-11.xlsx**)
   2. Run visual basic script called **Att\_QCApplyQualFormats10.wsf**
      1. Enter data year to process when prompted
      2. Generates *Year*\_attended\_QC2.xlsm file, including data only for year in question
      3. Creates “QC2 attended init” sheet:
         1. Adds missing data from “Unmatched FS OWMID” sheet in QC1 file, applies “M” flag to analyte-comment (“\_SS”) columns to indicate missing data
         2. Based on decisions/exceptions in lab QC file:
            1. Adds “Z” flag to analyte-comment (“\_SS”) columns to indicate no calibration
            2. Adds “c” qualifier for standards exceptions found in lab QC file or using typical standards for year in question
            3. Adds “i” qualifier
      4. Creates “QC2 attended” sheet:
         1. Compiles all qualifier information from “QC2 attended init” into a single qualifier column for each analyte
         2. Adds info from “\_SS” columns on “QC2 attended init” to new analyte qualifier column:
            1. “X” = “NO DATA”
            2. “Z” = “NO CAL”
            3. “M” = “MISSING”
         3. Adds field sheet, station, water body info based on UNIQUE\_ID field (from separate file: **rptStaidUnique\_ID\_02-24-11\_LatLongBasin.xlsx**)
      5. Imports vba module file called “FinalizeAttData.bas”, used for later data finalization
5. Review results in QC2 file for each year and finalize data for QC3 review by Principal Investigators
   1. Make necessary manual changes to “QC2 attended” sheet. (Highlight changes in green. Document manual changes on “Manual Changes” tab.)
      1. Review **FS comments for any decisions affecting data** (that were not accounted for in automation).
      2. check for **blanks** in any of the station/project info columns; analyte columns; (qualifier columns and Mile Point are the only columns that can have blanks). Date/time cells are blank or -1 if data are missing.
      3. Review each qualifier column for **3 or more qualifiers**. This may indicate the need for censoring if the datum was not already (automatically) censored.
      4. Review each data column for obviously **high/low value outliers** that may be erroneous or non-representative. If found, censor or qualify using BPJ. Example: high SpCond values that may indicate salt water conditions.
      5. For SC>1500 for coastal rivers/streams or where associated FS comments indicate reverse flow (tide flowing upstream), apply **“t” (=tidally-influenced)** to all analytes. (Note: for 2011 forward, automate this decision, if possible; will still require manual review in case sites aren’t coastal or if sites are upstream of dams, etc.)
      6. Review **NO OWMID situations where there was documented “no water”**. In order to retain information re: lack of water (in FS comment or FS “FLOWSTAT” field), apply new OWMIDs and add rows to “QC2 attended” tab. Add “NO WATER” to qualifier columns for Depth and all other analytes (vba macro automatically applies NO WATER symbol (“^^”) to data cells on “QC2 attended final” tab).
      7. Add Flow Condition column based on FLOWSTAT field (2011 forward: this step will be automated)
      8. Check all **lab QC decisions** (note 2011 forward should only require cursory checks on lab QC decisions because all lab QC decisions will be entered into the ILW files, complete with associated OWMIDs, which should eliminate most error associated with using the Sonde QC summary file; see **Manual Review of Lab QC Decisions.doc**)
      9. Make sure each analyte with “Accept” status has no qualifiers listed
      10. Make sure all analytes listed as MISSING (“\*\*”), NO DATA (“- -“), NO WATER (“^^”) have no qualifiers
      11. Make sure all cases where Flow Condition = No Water also have the NO WATER symbol (“^^”) in the data columns for all analytes
      12. Depths with values near 0 (on final sheet appear as 0 due to reporting rules): apply “i” to qualifier column for all depths with values between 0 and 0.049 (change QC Status to Qualify) (note for 2011 forward: automated decisions for depth will be applied to rounded depth values, which avoids this issue)
      13. Recover auto-censored data using **FS estimates**, when available. Example: negative depth data due to a faulty depth sensor and FS comment notes that depth was X. (optional)
      14. **Random QC checks** (optional)
   2. Make final manual changes to QC2 attended sheet based on misc. information, email correspondence and best professional judgment
   3. After all changes have been made to “QC2 attended” sheet, run vba macro called **FinalizeAttData**
      1. Enable macros
      2. Go to Developer tab🡪Macros button🡪select FinalizeAttData and hit Run
      3. Macro will:
         1. Create “QC2 attended final” sheet based on “QC2 attended” sheet
         2. Apply formatting and reporting rules (see **W:\DWM\SOP\CN 056.5 - SOP\_DWM data reporting rules\_11-9-10DRAFT.xls**)
         3. Create “Data summary” sheet with qualifier counts/percentages
         4. Create individual project files based on filtering of “Project Name” column
            1. Adds “ReadMe” tab imported from file = **attended\_probe\_template.xlsx**
            2. Adds “attended probe data” sheet
            3. Adds “Data summary” sheet
   4. Review “QC2 attended final” sheet and “Data summary” sheet
      1. Compare **stats on quals and censors** for each analyte. Weigh against expected/historical %s. (optional) Note: these stats more helpful in reviewing decision criteria.
      2. Check for blank cells (only OK for \_\_\_\_\_\_\_\_ and \_\_\_\_\_\_)
      3. Check that reporting rules have been applied to all data
      4. Check final data file for center alignment/left justified cells (all)
      5. Random QC checks (optional)
6. QC3 Review (by Principle Investigators for each project):
   1. Examine each file for completeness (are all data that were collected presented; if not, why? Does any data appear to be missing?)
   2. Are data presented reasonable based on field conditions and BPJ (e.g., are there outlier data that are not “real”)?
   3. Examine each file for any errors and inaccuracies you may find
   4. Other “problems”
7. Review comments received during QC3 review and make necessary changes to QC3 attended sheet, including documented QC3 changes sheet
8. QC4 Final Data tables:
   1. Run visual basic script **Att\_QC4.wsf** to generate QC4 files for each year
   2. Open QC4 attended file for year being reviewed, making necessary changes to “QC4 attended” sheet and documenting changes on “QC3 Review Changes” sheet
   3. Open QC4\_attended\_probe\_template.xlsx
      1. change date of release to appropriate month, year
      2. add project notes as project name in column B, with numbered list of notes in column D; enter next project on row following notes for 1st project (code will pull out project notes by relevant cell reference blocks, so don’t overlap info related to each project)
   4. Run FinalizeAttData\_QC4 macro
      1. Enable macros
      2. Go to Developer tab🡪Macros button🡪select FinalizeAttData\_QC4 and hit Run
      3. Macro will:
         1. Create “QC4 attended final” sheet based on “QC4 attended” sheet
         2. Apply formatting and reporting rules (see [DWM reporting rules\_10-20-11.xls](file:///\\DEP-FP-WOR-202.dep.govt.state.ma.us\WORKGROUP$\DWM\Data\attended_QA\process%20documentation\2006-2010%20process\DWM%20reporting%20rules_10-20-11.xls))
         3. Create “Data summary” sheet with qualifier counts/percentages
         4. Create individual project files based on filtering of “Project Name” column
            1. Adds “ReadMe” tab imported from file = **QC4\_attended\_probe\_template.xlsx**
            2. Adds “attended probe data” sheet
9. Check “QC4 attended final” data and decisions **against any historical (expedited) data validations for consistency**. Correct the historical deliverables as needed based on validated QC4 attended decisions, assuming that the QC4 process is more rigorous and results in more appropriate results than the expedited process. See RC for expedited data files.

## V-a APPENDIX: Outline of automation steps for each visual basic code module

### Att\_QCProcRawFiles7.wsf

* Required input files:
  + Pre-processed Hydrolab Excel files (placed in directory = W:\DWM\Data\attended pre-processing\AllYears\Hydrolab\_Processed)
  + Pre-processed YSI Excel files (placed in directory = W:\DWM\Data\attended pre-processing\AllYears\YSI\_Processed)
  + temperature-only QC data compiled into one Excel file (W:\DWM\Data\attended pre-processing\tempQCdata\combined years raw field sheets + TEMP.xlsx)
* Output files:
  + AllYears\_attended\_QC1.xlsx

Procedure:

1. check for presence of directory and files
2. create QC1 file to store compiled data and add headers (note: YSI and Hydrolab data are combined into one file)
3. get attended probe file names
4. open each probe data file in pre-processed directory, get identified best lines (format varies among input files so look for heading labels), copy to final file
   1. process Hydrolab files
      1. check for file prefix “PP” (pre-processed), if no prefix then add one and save file
      2. set up temporary array to hold best lines for each file
      3. get stability “u” qualifier column heading locations
      4. determine headers needed for analytes present in each data file and find location of each in file
      5. scan ahead in file to count # of best lines
         1. store bestline row # (indicated by “Y” in “Best Line?” column) from Attended raw data file in array called arBestLine
         2. store OWMID site # (after the colon for hydrolab) for each best line in array called arSite
      6. go through each pre-processed raw file again and copy data to QC1 file, matching headers
         1. add any comments from last column in data file after qualifiers to “Annotation copy” column in QC1 file
         2. check backwards from best line, until reach "Annotation" in 1st column (header = “Date”)
            1. extract OWMID from “Annotation” after colon
            2. find first date after annotation (determine starting row of data block)
         3. check forwards from best line, until reach either end of used range, or "Time", "Annotation", "Power" (for Power loss), "Recovery" in column 1
            1. get Date from previous row (last date of data block)
         4. write start date and time, end date and time to QC1 file
         5. calculate time difference between start and end date+times, write to DeltaDateTime column in QC1 file
   2. process YSI files
      1. check for file prefix PP (pre-processed), if no prefix then add one and save file
      2. set up temporary array to hold best lines for each file
      3. initialize "No Data" column values; if a header value is not found, then an X is placed in the "SS" column adjacent to the analyte column in the QC1 file
      4. get stability “u” qualifier column heading locations
      5. determine headers needed for analytes present in each data file and find location of each in file
      6. scan ahead in file to count # of best lines
         1. store bestline row # (indicated by “Y” in “Best Line?” column) from Attended raw data file in array called arBestLine
         2. store OWMID site # (after the colon for hydrolab) for each best line in array called arSite
      7. go through each pre-processed raw file again and copy data to QC1 file, matching headers
         1. add any comments from last column in data file after qualifiers to “Annotation copy” column in QC1 file
         2. check backwards from Y line, until reach a site that differs from Y line (or a blank line)
            1. get starting row of data block for site
         3. check forward from Y line, until reach a site (column labeled “Site” in YSI files) that differs from Y line, a blank, or maxrow in used range
            1. get ending row of data block for site
         4. get start date and time from start row and end date and time from end row, write to QC1 file
         5. calculate time difference between start and end date+times, write to DeltaDateTime column in QC1 file
5. apply “m” qualifier for method not followed (probe time in water) based on DeltaDateTime (see **2006-2010 Validation Decision Criteria Tables-DRAFT08-12-11.xlsx**)
   1. apply color coding to DeltaDateTime column and m qualifier columns (red = “Censor”, yellow = “Qualify”, green = “Accept”)
6. Add 2005 final summary data (Hydrolab and YSI data in separate files) to QC1 file (note: some 2006 data were included with the 2005 review)
7. correct OWMIDs in QC1 file to match field sheet format
   1. replace “?”, “\*”, “#” with “X” (hydrolab files can have stray characters after the colon that follows Annotation at…)
   2. add a dash following basin ID, if necessary
   3. change basin #s
      1. 26 = SM
      2. 24 = LB
      3. 25 = LC
8. Add temperature QC OWMIDs
   1. Open temperature QC file (see combined years raw fieldsheets + TEMP.xlsx)
   2. Filter by column “new QC OWMID” = “Y” on sheet “new Temp QC” (OWMIDS that aren’t new will be included from multi-probe data files)
   3. Go through each filtered row, correcting OWMID based on item 7 above
   4. Determine if OWMID already exists in QC1 file
      1. If Yes, then don’t add to QC1 file: original multi-probe data were kept
      2. If No, then add temperature data to QC1 file, add “X” for “NO DATA” to other analyte “SS” columns (DO, pH, SpCond, TDS, DOsat), add “s” qualifier to QC1 column called “FS Data Used” to indicate field sheet data used

### Att\_QCCheckUnmatched9.wsf

* Required input files:
  + AllYears QC1 Excel file containing best lines for all years (AllYears\_attended\_QC1.xlsx)
  + electronic field sheet meta data (file = W:\DWM\Data\RawData\*yyyy*\MetaData*yyyy*.xlsx, where yyyy is 4-digit data year)
* Output files:
  + yyyy\_attended\_QC1.xlsx (where yyyy is four-digit data year being processed)

Procedure:

1. At prompt, user enters data year to process
2. check for presence of directory and files
3. open QC1 file that contains hydrolab, ysi, and temperature-only data for all years
4. create new worksheets for unmatched OWMIDs and for unattended-QC OWMID sets
5. open field sheet file for year selected in 1 above
6. standardize “NO DATA” flags in analyte “SS” columns
   1. hide all columns in QC1 file except “SS” columns
   2. replace “\*”, “N”, “#” with “X” for “NO DATA”
   3. make all columns visible again
7. for each OWMID in QC1 file, check for match in field sheet file
   1. if match found
      1. copy probe-related field sheet meta data to QC1 attended sheet (SmpTypName, FSLOG, UniqueID, Comments, FSComment, FlowStat, Date, FSTime, StartDate, StartTime, Projname, SmpGearName, SNSampGear, SN2SampGear
      2. calculate DeltaDateTime between field sheet date/times and best line date/times (for later identification of possible field sheet errors, mismatched IDs)
   2. if match not found
      1. write “NO MATCH FOUND” under FS OWMID column on QC1 attended sheet
8. for each OWMID in field sheet file, check for matching OWMID in QC1 file
   1. if match found
      1. do nothing (see 7ai)
   2. if no match found
      1. write field sheet info to “Unmatched FS OWMID” sheet in QC1 file (note this list will be all inclusive, containing unattended and chemistry OWMIDs, in addition to unmatched attended OWMIDs)
9. generate a list of OWMIDs for Unattended and Attended QC Cal Check sets from field sheet data (used for error checking)
   1. find SmpTypName = “In-situ: Un-attended” and SmpTypName = “In-situ: QC Cal Check” (store meta data based on these fields in arrays)
   2. loop through all unattended OWMIDs in array arOWMIDU
      1. for a given unattended OWMID’s FSLOG, find all QC Cal Check OWMIDs
      2. identify drop off QC: match unattended OWMID StartDate to QC Cal Check OWMID StartDate
      3. identify pick up QC: match unattended StopDate to QC Cal Check OWMID StartDate
      4. write matching set info to “Unattended-Attended OWMID sets”
10. populate QCType from 9 on QC1 Attended sheet
    1. SmpTypName = "In-situ: Attended": “QCType” = “N/A”
    2. SmpTypName = “In-situ: QC Cal Check”: “QCType” = “Drop off” or “Pick Up” or “Unknown” (not matched or intermediate QC)

### Att\_QCApplyQualFormats10.wsf

* Required input files:
  + QC1 Excel file containing best lines for all years (yyyy\_QC1\_attended.xlsx, where yyyy is data year to process)
  + station ID file containing station and water body information (file= **W:\DWM\Data\Attended pre-processing\Validation\station\_WBinfo\rptStaidUnique\_ID\_02-24-11\_LatLongBasin.xlsx**)
  + lab QC Excel summary file for all years, Visual Basic for Applications (VBA) module (file = **W:\DWM\Data\Attended pre-processing\Validation\labQC\2005-2010SondeQC.xlsx**)
  + VBA module (W:\DWM\Data\Attended pre-processing\Validation\code)
* Output files:
  + yyyy\_attended\_QC2\_MM-DD-YYYY.xlsm (where yyyy is four-digit data year being processed, MM-DD-YYYY is the month-day-year when code was run)
  + folder containing individual Excel project files (requires running macro within the .xlsm file to finalize the data and export the project files)

Procedure:

1. at prompt, user enters year for analysis
2. check for presence of directory and files
3. at prompt, user navigates to and selects QC1 Excel file to use in analysis
   1. open file and copy information from QC1 attended file to “QC2 attended init” sheet (only for year entered in 1 above), save as QC2 workbook
4. add columns for lab qualifiers (accuracy “i", calibration range “c”); for field data used (“s” qualifier, to indicate when temperature QC data is used); for combined DO/LDO and DOsat/DO% columns to standardize DO/DOsat reporting from YSI and Hydrolab files
5. combine DO/LDO and DOsat/DO% columns into uDO\_Comb and uDOsat\_Comb columns (and associated “No Data” columns as uDO\_Comb\_SS and uDOsat\_Comb\_SS)
6. add missing OWMIDs to “QC2 attended init” sheet
   1. from QC1 file, “Unmatched FS OWMID”, go through each row until SmpTypName equals “In-situ: Attended” or “In-situ: QC Cal Check”
   2. check for existence of OWMID on QC2 attended init sheet
      1. if OWMID not found, then
         1. add OWMID as missing to QC2 attended init (add field sheet meta data from Unmatched FS OWMID sheet: FSLOG, UniqueID, SmpTypName, Comments, FSComments
         2. add “M” to analyte “SS” columns (Depth, T, pH, DO, DOsat, SpCond, TDS) to indicate “MISSING” data
      2. if OWMID found, then do nothing
7. Open lab qualifier summary file, then for attended calibrations only, match against QC2 attended init sheet by project name, calibration date range, sondeID, loggerID for year selected in 1 above
   1. Find cases where parameters marked as “cal only”
      1. apply “Z” to analyte “SS” columns indicating “NO CAL” (note an analyte may be flagged as “NO DATA”, “NO CAL”, or both; OWMIDs marked as having “MISSING” data do not get the “NO CAL” flag)
   2. loop though lab QC file again and get all lines for year in question
      1. determine which lines share common project name, sondeID, loggerID
         1. determine if dates overlap and flag these calibrations
   3. loop through lab QC file again in order to apply lab qualifiers for accuracy (“i”) and calibration range(“c”)
      1. if dates overlap, flag lab QC entry as DATE OVERLAP; do not apply any qualifiers automatically, copy entries to sheet called “Unmatched Lab QC” for manual follow up and application of qualifiers
      2. if no matches found for lab QC entry, write entry to “Unmatched Lab QC”
      3. if matches are found without date overlap
         1. apply accuracy qualifier “i” decision from “QC result” column to analyte indicated in “Parameter” column (DO, SpCond, pH) to each matching line found on “QC2 attended init” sheet
         2. cases where “Parameter” = “Post-Survey Check” (no post survey calibration check was completed): apply “i” to all analytes (DO, DOsat, SpCond, pH, TDS) except T
         3. get exceptions for standards used in calibration of SpCond and pH, determine calibration range “c” qualifier result based on qualifier decision matrix and apply to each matching line found on QC2 attended init sheet
8. Find parameters where accuracy and calibration range QC results not obtained from Lab QC file (blank values)
   1. For accuracy (“i”) qualifier, fill in blank values with “Accept” (color cells orange to identify which were automatically populated with Accept vs. those with values from lab qualifier summary file)
   2. For calibration range (“c”) qualifier, determine QC result based on typical standards used for the year in question, analyte values, and validation criteria decision tables (color cells blue to identify which values were calculated automatically vs. those added as exceptions from lab qualifier summary file)

**Typical standards used for SpCond and pH calibration**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Analyte** | **Year(s)** | **Sonde Type(s)** | **Low Std Used** | **High Std Used** | **Units** |
| **SpCond** | 2006-2009 | Hydrolab, YSI | N/A | 718 | µS/cm |
| **SpCond** | 2010 | Hydrolab | N/A | 1413 | µS /cm |
| **SpCond** | 2010 | YSI | N/A | 718 | µS /cm |
| **pH** | 2006-2010 | Hydrolab, YSI | 4.01 | 9.18 | SU |

*\*Exceptions to typical standards used for a given calendar year are those listed in the LabQC file*

1. copy info from QC2 attended init to QC2 attended
   1. copy field sheet and OWMID data
   2. combine qualifiers into individual columns by analyte, including “NO DATA”, “NO CAL”, and “MISSING” flags
   3. add QC status column for each analyte (if QC result = “censor” for at least 1 qualifier, then QC status is censor for that analyte)
   4. open station file, add station (Station ID, Description, Mile Point, Latitude, Longitude) and water body (Watershed, SARIS/PALIS/CAMIS number, Water Body) info based on Unique\_ID match
   5. convert TDS values from g/l to mg/l
2. import vba module (FinalizeAttData.bas) into QC2 attended file for data finalization

### FinalizeAttData.bas

(VBA module imported into QC2 attended file via vbscript **Att\_QCApplyQualFormats10.wsf**; VBA module contains two sub procedures: FinalizeAttData and CreateDataSummary, and one function: navfile)

* Required input files:
  + Excel template file for “read me” sheet that gets copied to each QC2 project file (W:\DWM\Data\attended pre-processing\Validation\templates\attended\_probe\_template.xlsx)
* Output files:
  + Project folder (called “Project\_Files\_MM-DD-YYYY”, where MM is month, DD is day, and YYYY is year that QC2 file was generated using **Att\_QCApplyQualFormats10.wsf**) containing individual Excel project files (named “*Project*\_YYYY\_attended\_QC2.xlsx”, where Project is the project name and YYYY is the data year)

Procedure:

1. create “QC2 attended final” worksheet based on “QC2 attended” worksheet, sorting data by project, watershed, date, time (note final sheet gets overwritten each time macro is run, project folder must be manually deleted or renamed first)
2. rename QC status columns for each analyte to “Analyte censor”, inserting actual analyte name for “Analyte” (note: these columns are later removed in order to streamline final data sheet)
3. apply reporting rules based on [DWM reporting rules\_10-20-11.xls](file:///\\DEP-FP-WOR-202.dep.govt.state.ma.us\WORKGROUP$\DWM\Data\attended_QA\process%20documentation\2006-2010%20process\DWM%20reporting%20rules_10-20-11.xls)
   1. format cells for significant figures
   2. apply <0.2 or <2 for low DO and DOsat values, respectively
   3. apply “##” to data cells for censored data
      1. if analyte is censored, change “Analyte censor” to “yes”
      2. if analyte is not censored, change “Analyte censor” to “no”
   4. apply “- -“ to data cells where qualifier column contains “NO DATA” or “NO CAL” indicating no data available (i.e. sonde not calibrated for a particular analyte, or probe to record data not on board)
      1. reset “Analyte censor” column to “no”
      2. reset qualifier column to blank
   5. apply “\*\*” to data cells where qualifier column contains “MISSING” flag
      1. reset “Analyte censor” column to “no”
      2. reset qualifier column to blank
         1. if date or time is missing (-1 or blank), apply “\*\*” to applicable data cell (date or time missing occurs only for temperature QC data obtained from paper field sheets or OWMIDs with no data in raw data files)
   6. apply “^^” to data cells where qualifier column contains “NO WATER” flag
      1. reset “Analyte censor” column to “no”
      2. reset qualifier column to blank
   7. apply “\*\*” to “MISSING” in Flow Condition column
4. apply formatting to cells (color, borders, fonts, column widths, alignment)
5. correct cases where Connecticut misspelled (from 2008 electronic field sheet)
6. remove all “Analyte censor” columns to maintain cleaner appearance on final sheet
7. create “Data summary” sheet of qualifier counts/percentages, plus counts of missing, no data, no water, censored, qualified, and accepted data by analyte
   1. summary for entire QC2 file year
   2. summary by project
8. freeze top row on QC2 attended final, autofilter all columns
9. export project files
   1. find unique project names on QC2 attended final sheet
   2. create folder to store project files (called “Project\_files\_MM-DD-YYYY”, where MM-DD-YYYY are the month-day-year when the QC2 attended file was created)
   3. copy all occurrences of project name to project file (named “Project \_yyyy\_attended\_QC2.xlsx”, where Project is the individual project name and yyyy is the data year) to “attended probe data” sheet
   4. open template file to copy "read me" worksheet to each project file
   5. add “Data summary” sheet with qualifier counts for entire project
   6. apply autofilter and freeze top row of “attended probe data” sheet
   7. select “read me” sheet before saving/exiting
   8. protect data worksheet
   9. make file read-only

### Att\_QC4.wsf

* Required input files:
  + QC2 Excel file containing data finalized for QC3 Review (yyyy\_QC2\_attended\_MM-DD-YYYY.xlsx, where yyyy is data year to process; MM, DD, YYYY are date QC2 file was generated)
  + VBA module called FinalizeAttData\_QC4.bas (W:\DWM\Data\Attended pre-processing\Validation\code)
* Output files:
  + yyyy\_attended\_QC4\_MM-DD-YYYY.xlsm (where yyyy is four-digit data year being processed, MM-DD-YYYY is the month-day-year when code was run)
  + folder containing individual Excel project files (requires running macro within the .xlsm file to finalize the data and export the project files)

Procedure:

1. At prompt, user enters year to process
2. At prompt, user navigates to QC2 folder (if necessary) and selects final QC2 file
3. Check for existence of directories and files
4. Create folder for QC4 file called YYYY\_attended\_QC4
5. Create QC4 file called YYYY\_attended\_QC4\_MM-DD-YYYY.xlsx, where MM, DD, YYYY are month, day, and year that file was created
6. Create worksheets: “Read Me”, “QC3 Review Changes”, “QC4 attended”
7. Open QC2 file and copy “QC2 attended” sheet to “QC4 attended” sheet in QC4 file
8. Apply autofilter to QC4 attended sheet, freeze top row
9. Import vba module called “FinalizeAttData\_QC4.bas”
10. Write source code/file info to “Read Me” sheet

### FinalizeAttData\_QC4.bas

(VBA module imported into QC4 attended file via vbscript **Att\_QC4.wsf**; VBA module contains two sub procedures: FinalizeAttData and CreateDataSummary, and one function: navfile)

* Required input files:
  + Excel template file for “read me” sheet that gets copied to each QC2 project file (W:\DWM\Data\attended pre-processing\Validation\templates\QC4\_attended\_probe\_template.xlsx)
* Output files:
  + Project folder (called “Project\_Files\_MM-DD-YYYY”, where MM is month, DD is day, and YYYY is year that QC2 file was generated using **Att\_QC4.wsf**) containing individual Excel project files (named “*Project*\_*yyyy*\_attended\_QC2.xlsx”, where *Project* is the project name and *yyyy* is the data year)

Procedure:

1. create “QC4 attended final” worksheet based on “QC4 attended” worksheet, sorting data by project, watershed, date, time (note final sheet gets overwritten each time macro is run, project folder must be manually deleted or renamed first)
2. rename QC status columns for each analyte to “Analyte censor”, inserting actual analyte name for “Analyte” (note: these columns are later removed in order to streamline final data sheet)
3. apply reporting rules based on [DWM reporting rules\_10-20-11.xls](file:///\\dep.govt.state.ma.us\enterprise\Worcester-Workgroup\DWM\Data\attended_QA\process%20documentation\2006-2010%20process\DWM%20reporting%20rules_10-20-11.xls)
   1. format cells for significant figures
   2. apply <0.2 or <2 for low DO and DOsat values, respectively
   3. apply “##” to data cells for censored data
      1. if analyte is censored, change “Analyte censor” to “yes”
      2. if analyte is not censored, change “Analyte censor” to “no”
   4. apply “- -“ to data cells where qualifier column contains “NO DATA” or “NO CAL” indicating no data available (i.e. sonde not calibrated for a particular analyte, or probe to record data not on board)
      1. reset “Analyte censor” column to “no”
      2. reset qualifier column to blank
   5. apply “\*\*” to data cells where qualifier column contains “MISSING” flag
      1. reset “Analyte censor” column to “no”
      2. reset qualifier column to blank
         1. if date or time is missing (-1 or blank), apply “\*\*” to applicable data cell (date or time missing occurs only for temperature QC data obtained from paper field sheets or OWMIDs with no data in raw data files)
   6. apply “^^” to data cells where qualifier column contains “NO WATER” flag
      1. reset “Analyte censor” column to “no”
      2. reset qualifier column to blank
   7. apply “\*\*” to “MISSING” in Flow Condition column
4. apply formatting to cells (color, borders, fonts, column widths, alignment)
5. correct cases where Connecticut misspelled (from 2008 electronic field sheet)
6. remove all “Analyte censor” columns to maintain cleaner appearance on final sheet
7. create “Data summary” sheet of qualifier counts/percentages, plus counts of missing, no data, no water, censored, qualified, and accepted data by analyte
   1. summary for entire QC4 file year
   2. summary by project
8. freeze top row on QC4 attended final, autofilter all columns
9. export project files
   1. find unique project names on QC4 attended final sheet
   2. create folder to store project files (called “Project\_files\_MM-DD-YYYY”, where MM-DD-YYYY are the month-day-year when the QC4 attended file was created)
   3. copy all occurrences of project name to project file (named “Project \_yyyy\_attended\_QC2.xlsx”, where Project is the individual project name and yyyy is the data year) to “attended probe data” sheet
   4. open template file to copy "read me" worksheet to each project file
   5. apply autofilter and freeze top row of “attended probe data” sheet
   6. select “read me” sheet before saving/exiting
   7. protect data worksheet
   8. make file read-only

## V-b APPENDIX: Checks on electronic field sheets/data entry

1. Make sure comment fields are not truncated due to having more than 255 characters (FSComment, Comments); re-export from Access if possible, or add complete entries back into Excel field sheets
2. Identify Temperature QC OWMIDs and assign new OWMIDs if necessary; add to electronic field sheet; enter date, time, temperature data into a separate Excel workbook
3. Check for duplicate In-situ: Un-attended
   1. make sure duplicate is not due to error in data entry (e.g. incorrect SmpTypName)
   2. if duplicate is recorded on same field sheet as original, then QC Cal checks apply to both unattended OWMIDs
   3. if duplicate is recorded on separate field sheet with no QC Cal checks, then create new QC Cal Check OWMIDs and copy original QC OWMID data to new OWMIDs; this ensures that each unattended OWMID has a pair of QC Cal Check OWMIDs (for automating trimming of unattended files; later we can decide which duplicate unattended OWMID should be kept)
4. Check for missing sonde and logger IDs for SmpTypName = In-situ: Attended and In-situ: QC Cal Check where SmpGearName = Data Sonde/Logger Combo
   1. exclude temperature loggers (SmpGearName = Data Probe (Single))—these don’t matter because temperature isn’t QC’d using loggers
   2. check that unattended OWMIDs (SmpTypName = In-situ: Un-attended) do not report logger ID for SN2SampGear (could indicate incorrect SmpTypName)
5. Include missing OWMIDs that were omitted from electronic field sheet entry because the sonde wasn’t used (usually due to no water, stagnant, low flow, or ice conditions; sometimes due to probe malfunction, battery issues, or site accessibility)
   1. look for OWMIDs already assigned on paper field sheets and add all sample info to electronic field sheet
      1. make sure OWMID hasn’t been used for another survey/field sheet
   2. assign new OWMIDs for cases where no OWMID is recorded on paper field sheet
6. Make sure OWMIDs fall under correct SmpTypName
   1. field sheets where the first digit following the basin letter code is zero are attended
   2. field sheets where the first digit following the basin letter code is 5 are deployed (In-situ: Un-attended or In-situ: QC Cal Check)
   3. make corrections as necessary
7. Check that all deployed field sheets have at least 1 un-attended OWMID and 2 QC Cal Check OWMIDs
   1. sort electronic field sheet by FSLOG, then by SmpTypName descending order
   2. apply conditional formatting where text = “un” in SmpTypName column, highlight cells a color
   3. select SmpTypName = blanks, In-situ: Un-attended, In-situ: QC Cal Check
   4. scan through file to make sure that all blanks have been addressed (should have been found by checking for missing OWMIDs) and that each field sheet has at least 1 un-attended and 2 QC Cal checks)
   5. add missing OWMIDs and the associated sample info to electronic field sheet

## V-c APPENDIX: Procedures for pre-processing Attended Multi-Probe Data Files (Hydrolab & YSI)

Objective:

The raw data files for attended multiprobes are reviewed to select the “best line” of data to associate with each OWMID. The only QC check performed at this stage is for stability of readings. For each analyte (DO,pH, SpCond and T, as applicable) in each selected “best line”, apply stability qualifier (qualify or censor) if necessary.

**Hydrolab Steps:**

1. Open a raw Hydrolab data file at: [W:\DWM\Data\attended pre-processing](file:///\\DEP-nas-BOS-002.dep.govt.state.ma.us\Workgroup$\DWM\Data\attended%20pre-processing) for year being pre-processed
2. In column A, scroll down to 1st OWMID annotation. The data contained in the rows following each OWMID# pertain to that OWMID.
3. Review the data blocks for each OWMID in order to select the “best line” (row) that appears to best represent the pH, DO, T and SpCond data (as applicable) at that station and time, in terms of stability of the data record. This will typically be within the last 3 lines. Only one row has to be selected.
4. Once you decide on the preferred “best line”, shade the entire row in green (“GOOD”) and mark the row with a “Y” for “YES” in the column called “BEST LINE?” and copy/paste the associated OWMID (column A , upper left of each data block) in the “copy annotated OWMID” column in the same row and to the right of the “best line”.
5. Once the “best line” is selected, apply the following “stability” decision criteria to each data block that has an associated field OWMID#. Use the columns to the far right in each file to choose the proper decision from the pick list for each analyte. Do not spend time calculating ranges, but visually check the value ranges against the criteria. For each analyte, a stability decision (ACCEPT, QUALIFY, CENSOR OR N/A) has to be selected. NOTE: some analytes had no probe on-board the sonde (“X” or “N”) or were not calibrated (“\*”)—these symbols represent NO DATA for that analyte and get “N/A”.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **5 minute RANGE of readings from 1st stable reading to “best line” near the end** | | |
| Analyte | ACCEPT | QUAL \*\* | CENSOR \*\* |
| T (deg. C) | <0.3 | 0.3-0.6 | >0.6 |
| DO (mg/l) | <0.4 | 0.4-0.8 | >0.8 |
| pH (SU) | <0.4 | 0.4-0.6 | >0.6 |
| SpCond (uS/cm) \* | <2% | 2-5% | >5% |
| *\* e.g., for 500 uS/cm, 2%=10uS/cm; 5%=25uS/cm)* |  | \*\* also look for very erratic, jumpy readings, and significant, continuous movement in one direction without stabilizing | |

DO sat = DO; TDS = SpCond

1. For cases where an annotation appears to have been missed (e.g. analyte data values are substantially different between adjacent rows, date/time differs between adjacent rows):
   1. Add blank line separating the two data blocks
   2. Add “Annotation” to column 1 on the row where you added the blank line
   3. If “new” data block appears to contain good data, follow above directions for determining best stable line (note: for lake surveys with missed annotations, repeat known OWMID by adding “Annotation at: OWMID” to ”Annotation Copy” column)
2. Once all “best lines” have been selected and stability qualifiers imposed as necessary, review file for completeness and errors. If satisfied, save file with same name and in same location.

**Things to look for:**

* “Outliers” at beginning, middle or end of data block. If at beginning, disregard and start with the 1st stable row. If in middle, disregard if one aberrant reading, but not if several “outliers”. If at end, do not select the last line.
* Readings that never really stabilize and change unidirectionally in big jumps (ie, not naturally-occuring variation)
* No OWMID to copy/paste

YSI Steps: *essentially the same as for Hydrolab*

1. Open a raw YSI data file at: [W:\DWM\Data\attended pre-processing](file:///\\DEP-nas-BOS-002.dep.govt.state.ma.us\Workgroup$\DWM\Data\attended%20pre-processing) for year being pre-processed
2. Review the data blocks for each OWMID in order to select the “best line” (row) that appears to best represent the pH, DO, T and SpCond data (as applicable) at that station and time, in terms of stability of the data record. This will typically be within the last 3 lines. Only one row has to be selected.
3. Once you decide on the preferred “best line”, shade the entire row in green (“GOOD”) and mark the row with a “Y” for “YES” in the column called “BEST LINE?” You do not need to copy the OWMID (as for Hydrolab)—it already exists for each line under the column labeled “Site”. For lakes files, select “best line” for each depth in the profile that contains “stable” data. The data blocks may be more like three minutes per depth, so relaxing the time criterion is OK for the depth profiles.
4. Once the “best line” is selected, apply the following “stability” decision criteria to each data block. Use the columns to the far right in each file to choose the proper decision from the pick list for each analyte. Do not spend time calculating ranges, but visually check the value ranges against the criteria. For each analyte, a stability decision (ACCEPT, QUALIFY, CENSOR OR N/A) has to be selected.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **5 minute RANGE of readings from 1st stable reading to “best line” near the end** | | |
| Analyte | ACCEPT | QUAL \*\* | CENSOR \*\* |
| T (deg. C) | <0.3 | 0.3-0.6 | >0.6 |
| DO (mg/l) | <0.4 | 0.4-0.8 | >0.8 |
| pH (SU) | <0.4 | 0.4-0.6 | >0.6 |
| SpCond (uS/cm) \* | <2% | 2-5% | >5% |
| *\* e.g., for 500 uS/cm, 2%=10uS/cm; 5%=25uS/cm)* |  | \*\* also look for very erratic, jumpy readings, and significant, continuous movement in one direction without stabilizing | |

DO sat = DO; TDS = SpCond

1. For lake surveys, the OWMID will be the same in the Site column, but depths will differ
   1. Identify the discrete depth blocks and add a blank line separating each data block with the same OWMID
   2. Follow above directions for determining stable best line
2. Once all “best lines” have been selected and stability qualifiers imposed as necessary, review file for completeness and errors. If satisfied, save file with same name and in same location.

**Things to look for:**

* “Outliers” at beginning, middle or end of data block. If at beginning, disregard and start with the 1st stable row. If in middle, disregard if one aberrant reading, but not if several “outliers”. If at end, do not select the last line.
* Readings that never really stabilize and change unidirectionally in big jumps (ie, not naturally-occuring variation)

## V-d APPENDIX: Manual review of lab QC decisions

Source file = W:\DWM\Data\attended pre-processing\Validation\labQC\2005-2010SondeQC.xlsx

Information in this file applies to qualifiers for inaccuracy (“i”) and calibration range (“c”); also “NO CAL” symbol (“--“) (same as “NO DATA” symbol) is applied for cases where less than the full suite of analytes has been calibrated for.

\*Note: only exceptions to general calibration procedures and cases where a particular analyte should be qualified or censored are listed in this file; calibrations not listed are assumed to equal “Accept” for inaccuracy and calibration range qualifier decisions

* Description of file structure
  + Headings
    - Source File
      * Original sonde QC file names compiled by year (from Jeff Smith)
    - BASIN
      * Project name recorded in lab notebooks
    - Project Name
      * BASIN translated into project name used in electronic field sheet (no year needed (code adds year from Year column))
      * separate lines are required for cases where a calibration is done for more than one basin at a time (e.g. lab notebook BASIN = S. Coastal/Taunton Deployed; 2 entries in Sonde QC file for project name = South Coastal and Taunton)
    - Year
      * Calibration year
    - CAL. DATE
      * Pre-calibration date
    - POST-CAL. DATE
      * Post-calibration date
    - SondeID
      * Sonde identification number (must match electronic field sheet format)
    - LoggerID
      * Logger identification number (must match electronic field sheet format; e.g. S1455, not 1455)
    - Attended/Unattended
      * A = In-situ: Attended or In-situ: QC Cal Check
      * U = In-situ: Un-attended
    - PARAMETER
      * Parameter to which QC information applies (enter as DO, pH, SpCond, Date/Time, Depth, Post-Survey Check)
      * Decisions for DO apply to DOsat; decisions for SpCond apply to TDS; decisions marked Post-Survey Check apply to all analytes, excluding T
      * Depth and Date/Time entries are always manually reviewed for their applicability/addition of qualifiers
      * T is not explicity listed because there are no consistent calibration checks made, records on manual thermometer checks were scarce or did not apply to the dates sondes were used in the field
      * Cases where DO listed as “Calibrated for PARAMETER only” applies to T and DOsat
    - DEVIATION FROM TABLE VALUES if ≥ 0.2
      * Jeff’s notes about differences between pre- and post-survey calibration checks (see “2006-2010 Validation Decision Criteria Tables-DRAFT08-12-11.xlsx”)
    - QC Result
      * Decisions of Qualify or Censor for the PARAMETER in question are applied with the inaccuracy qualifier (“i”); all other entries (Okay, Accept, blank) are ignored by the code
      * Decisions may need to be cleaned up for consistency and checked against parameter limits listed in data validation decision criteria tables
        + Entries where comments indicate no post-survey check was done

Copy original comment line, add “Post-Survey Check” as PARAMETER and add “Qualify” to QC Result

* + - * + Entries where comments indicate circulator was off

Copy original comment line, add “DO” as PARAMETER and add “Qualify” to QC Result

* + - Calibrated for PARAMETER only?
      * Used when a smaller subset of the full suite of analytes has been calibrated in the lab; if analytes are entered as “Y”; code will apply a “NO CAL” flag to the qualifier column of the analytes not listed; vba macro used to finalize the attended data reads any “NO CAL” flags for analytes and applies the “NO DATA” symbol (“--“) to the data value cell for final reporting
      * Enter Y = Yes if the calibration was completed only for the PARAMETER listed; Enter N = No if all analytes were calibrated for
      * Each row must have a Y or N entered in this column
      * If less than the full suite of analytes is calibrated for, then each analyte that was included in the calibration (e.g. DO and SpCond) must be listed on a separate row, not in the same cell under PARAMETER
    - Low Std used (pH, SpCond only)
      * An exception to the low standard used for pH calibration for the year in question is entered here; this column is blank if PARAMETER = SpCond
      * Both a low and high standard must be entered for pH even if only one standard is the exception
    - High Std used (pH, SpCond only)
      * An exception to the high standard used for SpCond calibration for the year in question is entered here
      * Both a low and high standard must be entered for pH even if only one standard is the exception
      * Calibration range qualifier (“c”) will be applied to pH and/or SpCond per the limits listed in the data validation decision criteria tables
    - COMMENTS
      * Jeff’s comments about the calibration, what went wrong, etc.
  + Each line pertains to a particular analyte or issue (e.g., DO, SpCond, Post-survey check, Date/Time, etc.)
  + Information for more than one decision may be represented using a single row
    - Up to 3 separate decisions may be obtained from a single row
      * QC Result = Qualify or Censor: “i” qualifier applied to analyte listed under PARAMETER
      * Low Std Used, High Std Used: “c” qualifier applied to SpCond, pH only if value falls into ranges for qualify or censor according to data validation criteria tables
      * Calibrated for PARAMETER only?: all analytes except those listed with common project name, sonde/logger IDs, and dates (recorded on separate lines in the Sonde QC file) get a “NO CAL” flag in their qualifier columns
  + Information from a single calibration that pertains to more than one analyte is listed on separate rows
* Automation issues
  + missing field sheet info may cause “unmatched” lab QC (for example, due to sonde or logger IDs that weren’t recorded on the field sheet or were recorded incorrectly)
  + lab QC decisions for cases where OWMIDs had the same project, cal dates, sondeID, and loggerID were mis-applied to “good” surveys not listed in the Sonde QC file
  + lab QC decisions overlap by date (in these cases, code did not apply any decisions listed in the Sonde QC file to the different survey data blocks; manual review required)
* Review of each Sonde QC line
  + filter Sonde QC file by “Year” = year in question and “Unattended/Attended” = “A”; copy cells to “All Lab QC” sheet
  + filter electronic FS by project, date range, sonde and logger IDs (be aware of missing sonde/logger ID issue)
  + identify which surveys go with the info listed in the lab QC sheet
    - review lab notebooks where necessary to determine if other calibrations exist for same project, dates, sonde and logger IDs
    - for cases of date overlap, OWMIDs usually have StartDate closer to cal. Date rather than post-cal date
  + Check that QC decisions have been correctly applied; make changes as appropriate to “QC2 attended”, highlighting changes in yellow; make notes of findings on “All Lab QC” sheet; in some cases, this review may require removing incorrectly applied qualifier decisions
* Helpful hints
  + Missing logger IDs in electronic field sheet and paper field sheet: look up raw data file(s) for OWMIDs in question and key the file letter code to the logger used (per lab inventory listing)
  + Missing sonde and logger IDs: obtain crew info and cross reference with Jeff’s calendar and check out sheets for probes/probe request forms
  + Watch out for incorrect logger and sonde IDs: due to the similarity in a lot of the numbers, the numbers are sometimes incorrectly recorded in the lab notebooks, on the paper field sheets, or when entering data into the electronic field sheets; check for incorrect OWMIDs, particularly if no matches are found for a given lab calibration entry
* Examples:

Unlisted “good” calibration info results in mis-application of qualifier decisions

Problem calibration listed in Sonde QC file:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Project Name** | **CAL. DATE** | **POST-CAL. DATE** | **SondeID** | **LoggerID** | **PARAMETER** | **QC Result** | **Calibrated for PARAMETER only?** | **Low Std used** | **High Std used** |
| Concord | 6/5/2006 | 6/6/2006 | 41217 | S1454 | DO |  | Y |  |  |

Unlisted calibration:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Project Name** | **CAL. DATE** | **POST-CAL. DATE** | **SondeID** | **LoggerID** | **PARAMETER** | **QC Result** | **Calibrated for PARAMETER only?** | **Low Std used** | **High Std used** |
| Concord | 6/6/2006 | 6/7/2006 | 41217 | S1454 | DO |  | N |  |  |

* 17 matches from electronic field sheet to problem calibration:
  + A010 survey, StartDate = 6/6/2006, 7 OWMIDs:
    - applies to unlisted calibration (another clue is that DO/T used for QC probes in 2006; this is an attended only survey)
    - ACTION: remove NO CAL from pH, TDS, SpCond, TDS columns for these 7 OWMIDs
  + A501 survey, StartDate = 6/5/2006, 5 OWMIDs:
    - pick up QC for survey applies to listed calibration (6/5-6/6)
    - NO ACTION
  + A502 survey, StartDate = 6/5/2006, 5 OWMIDs:
    - drop off QC for survey applies to listed calibration (6/5-6/6)
    - NO ACTION

Date overlap for two or more entries in Sonde QC file

Problem calibrations listed in Sonde QC file:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Project Name** | **CAL. DATE** | **POST-CAL. DATE** | **SondeID** | **LoggerID** | **PARAMETER** | **QC Result** | **Calibrated for PARAMETER only?** | **Low Std used** | **High Std used** |
| Conecticut | 6/3/2008 | 6/4/2008 | 41217 | S1455 | SpCond | Qualify | N |  |  |
| Conecticut | 6/4/2008 | 6/5/2008 | 41217 | S1455 | SpCond | Qualify | N |  |  |

* 16 matches from electronic field sheet to problem calibration:
  + C006 survey, StartDate = 6/3/2008, 8 OWMIDs:
    - Attended only survey applies to 6/3-6/4 entry
    - ACTION: apply “i” to SpCond qualifier column and “Qualify” to QC Status for these 8 OWMIDs
  + C500 survey, StartDate = 6/4/2008, 8 OWMIDs:
    - pick up QC for survey applies to 6/4-6/5 entry
    - ACTION: apply “i” to SpCond qualifier column and “Qualify” to QC Status for these 8 OWMIDs

Date overlap for two or more entries in Sonde QC file + unlisted calibration

Problem calibrations listed in Sonde QC file:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Project Name** | **CAL. DATE** | **POST-CAL. DATE** | **SondeID** | **LoggerID** | **PARAMETER** | **QC Result** | **Calibrated for PARAMETER only?** | **Low Std used** | **High Std used** |
| Concord | 7/14/2006 | 7/17/2006 | 41215 | S1455 | DO | Qualify | Y |  |  |
| Concord | 7/17/2006 | 7/18/2006 | 41215 | S1455 | DO |  | Y |  |  |

Unlisted calibration:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Project Name** | **CAL. DATE** | **POST-CAL. DATE** | **SondeID** | **LoggerID** | **PARAMETER** | **QC Result** | **Calibrated for PARAMETER only?** | **Low Std used** | **High Std used** |
| Concord | 7/18/2006 | 7/19/2006 | 41217 | S1454 | DO |  | N |  |  |

* 31 matches from electronic field sheet to problem calibrations:
  + A504 survey, StartDate = 7/14/2006, 12 OWMIDs:
    - Drop off QC, applies to 7/14-7/17 entry
    - ACTION: apply NO CAL to pH, TDS, SpCond, TDS columns for these 12 OWMIDs; apply “i” to DO and DOsat qualifier columns and “Qualify” to QC Status columns
  + A504 survey, StartDate = 7/17/2006, 6 OWMIDs:
    - pick up QC for survey applies to 7/17-7/18 entry
    - ACTION: apply NO CAL to pH, TDS, SpCond, TDS columns for these 6 OWMIDs
  + A505 survey, StartDate = 7/17/2006, 6 OWMIDs:
    - drop off QC for survey applies to 7/17-7/18 entry
    - ACTION: apply NO CAL to pH, TDS, SpCond, TDS columns for these 6 OWMIDs
  + A018 survey, StartDate = 7/18, 7 OWMIDs:
    - Attended only survey applies to unlisted entry for 7/18-7/19
    - ACTION: remove NO CAL from pH, TDS, SpCond for these 7 OWMIDs