DRAFT

DWM

Standard Operating Procedure

**File Processing and Data Validation**

**for UNATTENDED Water Quality Probe Data**

CN 56.5

Contents

[I Applicability 2](#_Toc330980470)

[II Overview 2](#_Toc330980471)

[III Prerequisites for Initiating Procedures 2](#_Toc330980472)

[IV Assumptions 3](#_Toc330980473)

[V. Unattended Probe Data Validation Procedures (DRAFT) 4](#_Toc330980474)

[VI. Data Validation Reports 8](#_Toc330980475)

[APPENDIX A: Checklist for Processing & Validation of Unattended Data 8](#_Toc330980476)

[APPENDIX B: Out-of-Water Analysis Procedures 11](#_Toc330980477)

[APPENDIX C: Decision Criteria 14](#_Toc330980478)

[APPENDIX D: VBScript and VBA Code Modules 16](#_Toc330980479)

[Separate Raw Data Files 16](#_Toc330980480)

[Process Raw Data Files 19](#_Toc330980481)

[Post-Processed Status File Messages 24](#_Toc330980482)

[Correcting File Formats 27](#_Toc330980483)

[Automated Validation and Processing 29](#_Toc330980484)

[QC Summary Workbook: Explanation of Worksheets 39](#_Toc330980485)

[Apply Final QC2 Review Decisions 42](#_Toc330980486)

[Calculate OWMID Statistics 44](#_Toc330980487)

[Generate QC2 FINAL and QC3 OWMID Files 50](#_Toc330980488)

[Create Station Files 56](#_Toc330980489)

[Calculate Station File Statistics 60](#_Toc330980490)

[Generate QC4 OWMID Files 64](#_Toc330980491)

[Generate QC4 Station Files 70](#_Toc330980492)

[APPENDIX E: Statistics 76](#_Toc330980493)

[Definitions 76](#_Toc330980494)

[Notes 78](#_Toc330980495)

[Examples 78](#_Toc330980496)

# I Applicability

These procedures apply to unattended probe data generated by DWM using multiprobe sondes and temperature-only data loggers. Each year, DWM monitoring staff deploys pre-programmed probes with internal data logging capabilities in streams, rivers, lakes and ponds throughout the State. These unattended probes collect water quality data at pre-set intervals primarily for dissolved oxygen and temperature (pH and specific conductance can also be collected). Probe deployment durations range from 2-3 days to 3-4 months, depending on need and circumstances. (For QC purposes, attended probe readings are collected “side-by-side” with the deployed units to allow comparison between readings.) On an approximate monthly basis, recorded data files from unattended 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 collected via unattended fashion (programmed to record at set intervals and deployed in-situ for long-term data collection). The steps consist of both automated and manual processes. Essentially, raw probe data files that have previously been uploaded to DWM’s network are assembled, checked for completeness, pre-processed to link up with fieldsheet metadata, trimmed to delete erroneous data at the beginning and end of each file, checked against field and lab QC data, and then reviewed against acceptance criteria. Data are either accepted “as is”, qualified or censored. At the completion of these steps, data are provided to project managers for further review prior to finalization using standardized output formats. These procedures are implemented by DWM’s data processing and quality assurance staff and are currently applied to individual year data sets (e.g., all 2008 files).

# III Prerequisites for Initiating Procedures

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

1. Raw probe data files (unattended) (either .xmd, .csv, or .xls/.xlsx formats)
2. Lab QC summary reports
3. Proofed electronic fieldsheet files
4. Final (QC4) attended (field QC) data files
5. Acceptance criteria and decision matrices (accept vs. qualify vs. censor) for each analyte
6. Program software/languages: MS Excel, Visual Basic for Applications, VBScript

Working files are managed using MS Excel (.xlsx or .xlsm). Data management/validation staff should be well-versed in MS EXCEL, Visual Basic for Applications (VBA) and VBScript 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, applied to the master “flat” .xlsx file (e.g. MetaData2008.xlsx), and transferred to other databases as needed.
2. Validation decisions applied to a parameter in a trimmed unattended file (e.g., D.O. censored) apply to all data for that parameter in that unattended file (exceptions include specific data points that are missing, for example, due to logger/sonde power losses or malfunctions, or that are censored, for example, due to out of water or tidal conditions).
3. Any decision applied to dissolved oxygen (D.O.) applies automatically to DO percent saturation (DOsat) and any decision applied to specific conductance applies automatically to total dissolved solids (TDS).
4. If field QC data using the attended QC probe (during setup and pickup of the deployed probe) were not collected right next to (side-by-side) the deployed probe, then it is assumed that the ambient sampling volume is completely mixed.
5. 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. Similar suspicions arise for unattended dissolved oxygen data collected using membrane-based DO probes and in-situ for longer than one week. (Optic DO probes are generally assumed to perform better for deployment periods longer than one week).
6. Validation of the unattended data is contingent on prior validation of the attended QC data, which is used to finalize the former. The SOP for processing and validation of the attended data is here: [W:\DWM\Data\process\_SOPs\attended](W:\\DWM\\Data\\process_SOPs\\attended)
7. All deployments resulting in final data are assumed to be representative of ambient surface water quality conditions for the duration of the deployment, unless qualified with “r” (representativeness) or “t” (tidal).
8. All deployments resulting in final data are assumed to have been generated using approved methods.

# V. Unattended Probe Data Validation Procedures (DRAFT)

1. Gather all raw data files (MANUAL)
   1. File location: place files in Data Pre-Process folder for appropriate year and logger/sonde type
      1. Multiprobe files (Hydrolab or YSI)
         1. W:\DWM\Data\unattended-QA\2008\**Data Pre-Process\Multiprobe**
      2. Temperature logger files
         1. W:\DWM\Data\unattended-QA\2008\**Data Pre-Process\Temperature**
   2. File formats:
      1. Multiprobe files
         1. should be comma-delimited (.xmd, .csv) or Excel (.xlsx, .xls)
         2. must follow the native Hydrolab format (.xmd) in terms of header placement, column spacing, and presence of individual date and time columns (*not* dates presented on single rows, as with the .txt files)
         3. must include the OWMID in the header block after “Log File Name:”
      2. Temperature logger files
         1. should be comma-delimited (.csv) or Excel (.xlsx, .xls)
         2. must include the date and time in a single column which later gets split into separate Date and Time columns
         3. must include the logger serial number in the header block (after “Plot Title:” or “Serial Number:”) (the serial # gets matched to field sheet meta data to identify the OWMID for the file)
   3. Correct file formats as necessary
      1. data sheet names for all file types must be identical to filename (minus the file extension) (this is because when a .csv file is opened with Excel and saved as an Excel file, the data sheet name is automatically taken from the file name)
      2. insert date and time columns (as described above, in 1.b.) (see Correcting File Formats, Appendix D)
2. Process raw data files (AUTOMATED/MANUAL)
   1. Run VBScript: **UnAtt\_SeparateFiles.wsf** (for 2005 and 2006 only) to separate overlapping files
   2. Run VBScript: **UnAtt\_ProcessRawFiles4.wsf** to create a standardized output file format
      1. Output files (.xlsx format) are created in Data Post-Process folder for year (YYYY) being processed (W:\DWM\Data\unattended-QA\YYYY\**Data Post-Process**)
      2. Review log file YYYY\_PPStatus.txt (where YYYY is the data year) for errors (see Appendix D, PPStatus.txt), reconcile errors, and re-run VBScript until no further errors occur
      3. Move air-temperature deploy files to DO NOT USE folder
3. Process standardized data files and apply automated validation steps (initial QC2 level) (AUTOMATED)
   1. Run VBScript: **UnAtt\_TrimQA\_24.wsf**
   2. Output files (.xlsx, .xlsm format) are created in QC2 folder for year (YYYY) being processed (W:\DWM\Data\unattended-QA\YYYY\**QC2**) (see Appendix D for more detail on output file formats)
4. Review initial QC2 files (MANUAL)
   1. Individual QC2 files named after OWMIDs (e.g., 10-0048\_QC2\_3-14-2012.xlsm)
      1. Look for anomalous points on the graphs
      2. Look for untrimmed files (start or end of file shows spikes)
      3. Look at QC Summary sheet for field QC, lab QC, and overall (combined field and lab QC) decision for each analyte
   2. Field sheet meta data and e-mail correspondence with Principal Investigators
      1. Review Comments and FSComments fields for any notes related to errors, probe malfunction, out-of-water situations, etc.
      2. Review e-mail for information that would cause data to be suspect (probe misplacement, out-of-water, low flow, etc.) and warrant analyte qualification or censoring
   3. QC2 Summary workbook (e.g., 2010\_Unattended\_QC2\_Summary\_3-14-2012.xlsm)
      1. Review Lab QC Summary sheet
         1. Make sure that there are no flags for “match already exists for OWMID” in yyyy\_Unattended Calculate Statistics\_Status.txt file
      2. Review Unattended-Attended OWMID Sets sheet for
         1. SmpTypName not equal to In-situ: QC Cal Check or In-Situ: Un-attended
         2. More than 1 Attended QC Type = “Drop off” or “Pick Up” for each FSLOG (this likely indicates a startdate error on the field sheet)
      3. Review Flags worksheet for potential errors
         1. Time errors
         2. Analyte thresholds (for threshold values used in the code, see Flags worksheet explanation, Appendix D, Automated Validation and Processing)
         3. Analyte spikes (for spike definitions (analyte value/time) used in the code, see Flags worksheet explanation, Appendix D, Automated Validation and Processing)
      4. Review Analytes worksheet for potential errors
         1. File Gear Matches FS Gear? If No, then manually fix sondeID in the field sheet file and/or the data files as necessary (changes may require re-running Unatt\_TrimQA)
      5. Review QC Decision Summary worksheet
         1. make notes about changes to be applied to individual rows of each data file in the “Actions Needed” column (DO NOT make changes to individual rows yet; QC2 final decisions/qualifiers must first be applied (see step 5 below))
         2. Add Final QC2 decisions for each analyte (DO, T) and add any additional qualifiers other than “i” (to be applied to entire column for each analyte)
         3. Identify the duplicate sample OWMIDs NOT to be included in the final dataset
            1. Add “X” to the “Dupes Not 2 Use” column
5. Apply final QC2 review decisions to unattended files (AUTOMATED)
   1. Run VBA subprocedure: **apply\_qcstatus** (located within the QC2 Summary workbook)
      1. Enable macros (this may require re-opening the QC2 Summary workbook to see the enable macros button on the ribbon bar)
      2. Go to Developer tab🡪Macros button🡪select **apply\_qcstatus** and hit Run
      3. Macro will:
         1. Check the QC2 Decision Summary worksheet for errors (such as capitalized or unrecognized qualifiers, blank decisions, and QUALIFY or CENSOR decisions without qualifiers, etc.)
         2. Open individual QC2 files and apply the QC2 final decisions/qualifiers to:
            1. QC2 Decision Summary sheet
            2. Data sheet (named after the OWMID)
6. Edit individual unattended files for row-by-row censoring decisions (MANUAL)
   1. For cases of missing attended QC data at probe drop off or pick up:
      1. Manually trim the files by deleting the appropriate rows at the beginning and/or end of each file (changes must be made to the OWMID data sheet, the OWMID backup sheet, the graphs-stats sheet, and the final sheet)
   2. Apply row-by-row censoring or missing data decisions to the analyte QC Decision column and add the associated qualifiers to the analyte qualifier column (MISSING or NO DATA in the qualifier column overrides any values in the QC decision column for each analyte; any other qualifiers present with the MISSING keyword are ignored)
7. Review all auto and manually applied decisions (MANUAL)
8. Calculate OWMID statistics (AUTOMATED)
   1. Run VBScript: **UnAtt\_CalcStats\_8.wsf**
   2. Statistics are calculated for each OWMID file and added to the Summary Statistics sheet and the Daily Statistics sheet in the OWMID file, and the Statistics sheet in the QC summary workbook
   3. See Appendix E for description of statistics
9. Generate QC2 FINAL files and QC3 Review Files (AUTOMATED)
   1. Run VBScript: **UnAtt\_QC3\_V6.wsf**
   2. QC2 FINAL files are created for each OWMID along with the summary file, organized by project name folder; QC2FINAL folder is located in the original QC2 Review folder and holds all project folders (e.g. W:\DWM\Data\unattended\_QA\2005\QC2\2005UQC2\_4-9-2012\2005UQC2\_FINAL\_5-10-2012)
   3. QC3 Review files are copies of the QC2 Review files and stored in a separate QC3 folder (e.g. W:\DWM\Data\unattended\_QA\2005\QC3\2005UQC3\_5-10-2012)
10. Copy the QC2 FINAL project folders and contents to [W:\DWM\data\QC2 data](file:///W:\DWM\data\QC2%20data) folder for staff QC3 review (MANUAL)
11. QC3 Review Process (MANUAL)
    1. Staff principal investigators review QC2 FINAL files for issues/errors
    2. Staff principal investigators submit their QC3 review comments to QC Officer (Richard Chase)
    3. QC Officer (or other validation staff) records changes needed to individual QC3 files in the QC3 summary workbook, QC2 QC3 Decisions sheet, in the QC3 Actions Needed and QC3 Review columns
    4. Apply QC3 Review changes
       1. to QC3 summary workbook, QC2 QC3 Decision sheet, Final DO and T decision columns and their associated qualifier columns, if needed
       2. to the individual QC3 files
          1. if final decisions have changed
             1. change individual cells by filtering for previous value (don’t fill down because there may be row by row censored or missing values)
          2. if individual rows/cells have changed
             1. change individual cells by filtering for previous value (don’t fill down because there may be row by row censored or missing values)
12. Recalculate OWMID statistics (AUTOMATED) (if required)
    1. Statistics must be recalculated if:
       1. Individual data files have changed (censored or missing cell changes only) for the year in question
    2. Run vbscript: **Unatt\_CalcStats\_8.wsf**
13. Create station files (AUTOMATED)
    1. Run vbscript: **Unatt\_CreateStationFiles\_V7.wsf**
    2. Station files are compiled for each UniqueID using the associated OWMID data files (for multiprobe files only)
       1. Data sheet becomes UniqueID sheet with each OWMID file added by date (earliest to latest)
          1. Rows are added to fill in discontinuous dates/times using NO DATA qualifier
       2. Final sheet is named after UniqueID with each OWMID file added by date (earliest to latest)
          1. Rows are added to fill in discontinuous dates/times using NO DATA qualifier and “- -“ in analyte data columns
       3. Graphs-stats sheet is named after UniqueID with each OWMID file added by date (earliest to latest)
          1. Rows are added to fill in discontinuous dates/times using NO DATA qualifier
          2. Standards are added to blank cells
       4. Daily Statistics sheet includes all daily statistics by OWMID, ordered by date
          1. Rows are added to fill in discontinuous dates using “- -“ (NO DATA)
       5. Station Statistics sheet is blank
       6. Graphs show all multiprobe OWMIDs for the station in question for the data year and are renamed with the UniqueID in place of the OWMID
14. Calculate station statistics (AUTOMATED)
    1. Run vbscript: **Unatt\_CalcStationStats.wsf**
    2. Statistics that could not be calculated directly from the OWMID statistics are computed for all multiprobe OWMIDs represented by the UniqueID (e.g. overall sampling period statistics, such as average, min, max, median, standard deviation, interquartile range)
    3. those statistics that can be calculated directly from OWMID file statistics (time/duration statistics, mean of daily statistics) are added to the QC2 or QC3 Summary workbook Station Statistics sheet
15. Generate QC4 FINAL files (AUTOMATED)
    1. Make sure there is a read me file for unattended QC4 data
    2. OWMID files
       1. Run vbscript: **Unatt\_QC4.wsf**
    3. Station files
       1. Run vbscript: **Unatt\_Stations\_Finalize.wsf**
16. Copy the QC4 FINAL project folders and contents to [W:\DWM\data\QC4 data (FINAL)](file:///W:\DWM\data\QC4%20data%20(FINAL)) folder (MANUAL)
    1. OWMID files
       1. copy files from each project folder to a subfolder called “Unattended OWMID Files” under the corresponding project folder name
    2. Station files
       1. Copy files from each project folder to a subfolder called “Unattended Station Files (Multiprobe only)” under the corresponding project folder name

# VI. Data Validation Reports

For each year-set of data that has been validated, a brief summary report is produced by DWM QA Analyst to document what took place. These reports are given Control Numbers (CN) and placed on DWM’s network drive as a reference: [w\dwm\SOP](file:///\\dep.govt.state.ma.us\enterprise\Worcester-Workgroup\DWM\SOP).

# APPENDIX A: Checklist for Processing & Validation of Unattended Data

*This abbreviated list summarizes the general steps taken (in order) to produce final UNATT data.*

* Validate ATTENDED probe data to QC4 status, including addition of “manual” temperature QC OWMIDs (The attended data are field QC records for UNATTENDED). See CN 56.X for detailed procedures.
* Verify availability of all complete and final (QC’d) supporting files:
  + FS metadata
  + UNATTENDED probe files
  + Lab QC records (for unattended)
  + Current reference files (e.g., reporting rules) to verify coding is correct
  + QC4 attended probe data files
* Assemble all FS comments and email communications relating to UNATTENDED data quality, to identify the specific cases where “manual” (non-automated) changes may be needed, due to circumstances affecting data quality
* RUN VBscript file “***UNATT\_ProcessRawFiles4.wsf***”: Processes all unattended files to standardize file format, link to fieldsheet information (sonde/serial #) and incorporate the OWMID into the post-processed filename.
* Using PPstatus.txt output files, resolve error messages as needed. Possible corrections include, but are not limited to, removing data files from the PreProcess folder to “DO NOT USE” folders (e.g., air-temperature deploys), re-exporting data files, adding/deleting rows in data files, changing header format in data files, changing worksheet names in data files, changing time format in data files and changing fieldsheet metadata. See Table X for more specific information.
* Re-run VBscript file “***UNATT\_ProcessRawFiles4.wsf***” to confirm lack of “errors” and move files to Data\_PostProcess folder.
* RUN VBscript file “***TrimFilesforQA\_11.wsf***”: ………
* -
* -
* -
* Identify all files in which Out-of-Water (OOW) conditions may have occurred. Use standard procedure (Appendix B) to analyze the identified files with the objective of censoring OOW data blocks. Record draft decisions in QC2 summary file.
* Manually review auto-flagged files to concur or recommend corrections based on BPJ. Possible corrections include, but are not limited to, censoring data in specific rows (e.g., spikes), applying alternate qualify/censor decisions based on fieldsheet comments, email communication or other information, or other. Record draft auto and manual decisions in QC2 summary file.
* Following concurrence on recommended “manual” changes between two validation staff, re-run QC2 summary file macro to apply any new auto-decisions
* Revise data files manually based on manual decisions from the QC2 summary file. First, “enable macros” for each data file, then make the changes. Save edited files, and record in the QC2 summary file that manual changes have been made to the data files.
* Generate files for QC3 project-level review…
* Compile QC3 comments and review them for validity and concurrence on recommended changes. Make necessary changes to the individual unattended files. Document edits and non-edits in the QC2 summary sheet using the QC3 headings.
* If QC3 edits were made, rerun code (…) to regenerate statistics…
* Generate QC4 files (final)…

# 

# APPENDIX B: Out-of-Water Analysis Procedures

**Data Validation Procedure for Manual Review of Continuous Temperature-only Data Files Observed to have been OOW (Out of Water):**

Note: Consistent with DWM’s inclusive data validation approach, no pre-determinations are made with respect to a-priori exclusion of water quality data due to applicability issues such as intermittency, as they relate to SWQS and/or assessment processes. In other words, any/all data collected are subject to validation during QC2/QC3, regardless of how the data might be used or not used. Also, once the data are censored, they cannot be used at a later date.

1. At a minimum, any data subject to OOW observation(s) shall be qualified with “r” due to uncertainty surrounding the OOW/IW analysis and what actually occurred in the field.
2. Using the fieldsheet metadata file and any other available information, identify all UNATT data files where the OOW condition was observed. Assume that files which do not have indications of OOW condition are OK in that regard.
3. View file and FS pickup QC to ascertain if OOW was falsely checked. If pickup QC was taken and it closely matches the same-time data in the file, this may indicate that the probe was not in fact OOW (i.e., fieldsheet error). Normally, field QC @ pickup would not be taken if logger was OOW.
4. For each flagged file, determine daily max/min (range) for all water temperature data in the file via pivot table analysis in Excel.
5. Acquire air temperature daily min/max (range) data via NCDC (verified and QC’d) for the climate station nearest the water quality station @ <http://cdo.ncdc.noaa.gov/qclcd/QCLCD>. Other air temperature stations can also be accessed to enhance the analysis (e.g., [http://www.wunderground.com/weatherstation/ListStations](http://www.wunderground.com/weatherstation/ListStations.asp?selectedState=MA&selectedCountry=United+States)),but use non-QC’d air temperature data with caution. Download daily ASCI summary for approp. air temp data station and desired year/month. Copy to Excel worksheet; text to columns (comma delimited). Note: Air temps at the sampling site may differ from air temps at the NCDC site due to micro-climes, OOW condition, etc. at the sampling site.
6. Copy daily climate data for each applicable month (spanning deployment period) to separate “OOW analysis” worksheet. Calculate daily “temperature buffers” as the difference in daily maximums between air/water plus the difference in daily minimums between air/water. Very low buffering (e.g., <5 deg. C) can indicate an OOW situation.
7. Insert charts of daily max’s and daily min’s (air vs. water) for visual comparisons. Expand visual chart for more detailed view. Note: do not base decisions on one chart only; look at both max and min charts to inform decision. Charts showing daily minimums for air and water can be especially useful in showing trends.
8. Based on buffer calculations, chart analyses and other information as available, use BPJ to make conservative censoring determinations from OOW-observation-date back to date of the last station visit for which there was no OOW observation, or further back based on BPJ. Use missing data determinations (for sondes based on non-deployment due to dry conditions) as additional information for temp logger OOW condition. Mark the proposed-to-be-censored rows by shading. Conservatively censor complete days (not partial). If the analysis is largely inconclusive and the data look generally suspect with respect to associated air temps (e.g., intermittent wetting/drying), then censor the entire file. It is advisable to reach BPJ conclusions by consensus of at least two staff, preferably the project lead (good knowledge of the sites and deployment surveys) and a data/QA staff person.
9. **If same-site air deploy data available (preferred)**: Where air temperature deploy data files are available for the same site, use the 30 min. interval data for direct comparison of air vs. water temps, assuming shading/sunlight reaching water-deploy when exposed similar to that of air deploy. Use buffer calculations and charts to make BPJ determination as indicated above (do not base decision solely on one chart or on just the calculations). When air and water temperatures are very closely matched, it is highly probable that the logger was OOW.
10. **If tidal situation**: Review charts for patterns and compare to historical, station-specific tides to make BPJ decision. <http://tidesandcurrents.noaa.gov/tide_predictions.shtml>
11. **Optional**: Compare non-censored temp logger data to same-site/same time multiprobe sonde temperatures, where available, to confirm non-censored data is most probably “real” or In-Water (IW). Maximum and average differences between sonde temp and logger temp should be < approx. 1.0 deg C, and preferably <0.6 deg. C.
12. Save “OOW analysis” worksheet in the approp. QC2 work area. The worksheet must clearly show the proposed data to be censored and a brief text description of the rationale behind the decision.
13. Transfer QC2 decisions for each OOW file to “manual changes” area for use by VB code

**Data Validation Procedure for Manual Review of Continuous DO/T Data Files Observed to have been OOW (Out of Water):**

Note: Consistent with DWM’s inclusive data validation approach, no pre-determinations are made with respect to a-priori exclusion of water quality data due to applicability issues such as intermittency, as they relate to SWQS and/or assessment processes. In other words, any/all data collected are subject to validation during QC2/QC3, regardless of how the data might be used or not used.

1. At a minimum, any data subject to OOW observation(s) shall be qualified with “r” due to uncertainty surrounding the OOW/IW analysis and what actually occurred in the field.
2. Using the fieldsheet metadata file and any other available information, identify all UNATT data files where the OOW condition was observed. Assume that files which do not have indications of OOW condition are OK.
3. **If SC data available** (even if not calibrated, data can be used as approximated), note where SC values approach “0” indicating OOW. Censor complete-day data blocks as needed.
4. If SC data are not available in the file, look for obvious anomalies in the DO and/or temperature data from the time it was observed OOW back. Examples include:
5. large jumps in values
6. blank data rows
7. “0” values
8. low DO minima early a.m. one day followed by much higher DO values at early a.m. the next day
9. Chart 30 minute data over time to view data graphically. This can help identify spikes, trends, etc. (e.g., flat temperature curve usu. Indicates IW buffering, delayed mins/maxs for temp when IW vs. expected times for air). Note: DO in air is usu. 7-9 mg/l depending on rel humidity/temp, so if <6 or >10, this can indicate probe is probably IW.
10. **If same-site air deploy data available (preferred)**: Where air temperature deploy data files are available for the same site, use the 30 min. interval data for direct comparison of air vs. water temps, assuming shading/sunlight reaching water-deploy when exposed similar to that of air deploy. Use buffer calculations and charts to make BPJ determination as indicated above (do not base decision solely on one chart or on just the calculations). When air and water temperatures are very closely matched, it is highly probable that the logger was OOW.
11. View precipitation data for deploy period to look for:
12. very large storm events that can move probes out of the water and onto the banks
13. little/no rain conditions for extended period before deploy that possibly led to stream level recession
14. other…
15. **If tidal situation**: Review charts for patterns and compare to historical, station-specific tides to make BPJ decision. <http://tidesandcurrents.noaa.gov/tide_predictions.shtml>
16. Optional: Chart daily maxs and mins vs. that for air temperature (as for temp loggers) to look for buffering and to graphically view data. Weather station data: <http://cdo.ncdc.noaa.gov/qclcd/QCLCD>; [http://www.wunderground.com/weatherstation/ListStations](http://www.wunderground.com/weatherstation/ListStations.asp?selectedState=MA&selectedCountry=United+States).
17. Where diurnal curve appears shifted in time or offset from expected pattern (e.g., DOmins in afternoon/early evening and DOmaxs in early morning)… ???
18. Save “OOW analysis” worksheet in the approp. QC2 work area. The worksheet must clearly show the proposed data to be censored and a brief text description of the rationale behind the decision.
19. Transfer QC2 decisions for each OOW file to “manual changes” area for use by VB code.

# APPENDIX C: Decision Criteria

|  |  |
| --- | --- |
| **QC qualifiers** |  |
| code | description |
| A | Accept |
| Q | Qualify |
| C | Censor |
| none | not available |

|  |  |  |  |
| --- | --- | --- | --- |
| **type of QC** | | | |
| code | description |  |  |
| F | Field QC | (average of individual differences between [ATTdropoff vs. UNATT] and [ATTpickup vs. UNATT]; OR pickup OR dropoff QC) | Include intermediate QC in the avgs, if applicable |
| L | Lab QC | (using post-check data) |  |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Dissolved Oxygen (DO) or % Oxygen Saturation (DOsat)** | | | | | |
| Decision matrix | |  |  |  |  |
|  |  | FA | FQ | FC | F None |
|  | limits | < 0.5 | 0.5-1.0 | >1.0 | N/A |
| LA | <0.2 | A | Q | Q | Q |
| LQ | 0.2-0.5 | Q | Q | C | Q |
| LC | >0.5 | Q | C | C | C |
| L None | N/A | A | Q | C | C |
|  |  |  |  |  |  |
| DO status compared to associated T status: | | | |  |  |
|  |  | DO | | |  |
|  | status | A | Q | C |  |
| T | C | Q | Q | C |  |
| \* additional constraint on DO | | | |  |  |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Temperature (T)\*** | |  |  |  |  |  |  |  |
| Decision matrix | |  |  |  |  |  |  |  |
|  |  | FA | FQ | FC | F None |  |  |  |
|  | limits | < 0.3 | 0.3-0.6 | >0.6 | N/A |  |  |  |
| LA | <0.3 | A | Q | Q | Q | <--default for no lab QC | | |
| LQ | 0.3-0.6 | Q | Q | C | Q |  |  |  |
| LC | >0.6 | Q | C | C | C |  |  |  |
| L None | N/A | A | Q | C | C |  | | |
| \*assumes lab QC program is being implemented |  |  |  |  |  |  |  |  |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Specific Conductance (SC) and TDS** | | | | | |
| Decision matrix | |  |  |  |  |
|  |  | FA | FQ | FC | F None |
|  | limits | < 2% | 2-5% | > 5% | N/A |
| LA | < 2% | A | A | Q | A |
| LQ | 2-5% | A | Q | Q | Q |
| LC | > 5% | Q | Q | C | C |
| L None | N/A | A | Q | C | C |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
| SC/TDS status compared to associated T status: | | | | |  |
|  |  | SC/TDS | | |  |
|  | status | A | Q | C |  |
| T | C | Q | Q | C |  |
| \* additional constraint on SC/TDS | | |  |  |  |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **pH** |  |  |  |  |  |  |  |  |
| Decision matrix | |  |  |  |  |  |  |  |
|  |  | FA | FQ | FC | F None |  |  |  |
|  | limits | <0.2 | 0.2-0.4 | >0.4 | N/A |  |  |  |
| LA | <0.2 | A | A | Q | A | <--default for no lab QC | | |
| LQ | 0.2-0.4 | A | Q | Q | Q |  |  |  |
| LC | >0.4 | Q | Q | C | C |  |  |  |
| L None | N/A | A | Q | C | C |  | | |
| \*assumes lab QC program is being implemented |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
| pH status compared to associated T status: | | | | |  |  |  |  |
|  |  | pH | | |  |  |  |  |
|  | status | A | Q | C |  |  |  |  |
| T | C | Q | Q | C |  |  |  |  |
| \* additional constraint on pH | | |  |  |  |  |  |  |

# APPENDIX D: VBScript and VBA Code Modules

## Separate Raw Data Files

**General Purpose:** To separate overlapping “raw” (pre-processed) unattended files downloaded from multiprobe data loggers (for 2005 and 2006 data years); for 2005 and 2006 data years, sondes were programmed to stop collecting at 5 pm, however, for redeployed sondes this meant that part of the second site’s data was stored in the first file; overlapping files are labeled with the same download date and sonde letter, with a 1 indicating the first file (first site or OWMID) and a 2 indicating the second file (second site or OWMID); the first file contains data pertaining to the second file: file 1 can be left as is (later trimming of the file will remove this extra information), but file 2 must be amended by adding data collected on the same start date from file 1

**Filename:** UnAtt\_SeparateFiles.wsf

**Code Language:** VBScript

**Input Files:**

* Data files downloaded from multiprobe sondes, stored in the appropriate folder under Data\_PreProcess for the data year in question
* Multiprobe files:
* must be comma delimited (.csv or .xmd) or Excel (.xls or .xlsx).
* If YSI or text formats are used, the files must be converted to .csv or saved as .xlsx with the same formatting as the hydrolab files (i.e. may require insertion of analyte comment columns and Date in its own column not on each row).

**Output Files:**

* Original copies of files collected at site 2 (redeploy site) are moved to a subfolder of Data\_PreProcess called “overlappingfiles\_DO\_NOT\_USE”
* YYYY\_FileOverlap.xlsx
  + Tracks each file name in Data\_PreProcess folder, the base file name, the file order (1 or 2), the number of files in the set with the same base file name, the starting and ending rows of data, starting and ending dates/times in the file, the target start and end dates/times to be copied from file 1 to file 2, # of new rows added to file 2, whether the original file was kept or deleted from Data\_PreProcess, whether the new file (.xlsx) was kept or deleted from the “overlappingfiles\_DO\_NOT\_USE” folder
* YYYY\_SeparateFileStatus.txt
  + Tracks error messages and code progress
    - Unrecognized File Format
    - Incorrect Data Worksheet Name
* Amended files for site 2 are saved under the original name but as .xlsx (e.g., YYMMDDS1.xmd becomes YYMMDDS1.xlsx, where YY is the file download year, MM is the download month, DD is the file download day, S is the sonde letter code (1 or 2 characters)) in the Data\_PreProcess folder

**Approximate Run Time:** 10 minutes, varies by data year and network traffic/PC speed

**Code Outline:**

1. announce to user what the program does and give them the chance to back out
2. user enters data year to process
3. check for presence of directories
   1. user selects Data\_PreProcess folder for data year entered in II.
   2. Code checks for existence of “overlappingfiles\_DO\_NOT\_USE” folder, if it doesn’t exist, create it
   3. Create status file for tracking errors (YYYY\_SeparateFileStatus.txt)
4. check for presence of files in the selected data folder
   1. announce to user if there are no files or, if files are found, the number found
5. Create a new workbook called “YYYY\_FileOverlap.xlsx” in the “overlappingfiles\_DO\_NOT\_USE” folder that tracks all files in the Data\_PreProcess folder
6. Loop through all files in Data\_PreProcess
   1. Extract file extension, base file name (YYMMDDS, download date of file + sonde letter code), file order (1 or 2)
   2. Write values to FileOverlap.xlsx workbook
7. Sort worksheet by base file name and file order
8. Loop through file names on worksheet, count the number with the same base name, write to worksheet
9. copy files where base file name is the same to overlap directory
10. open the copied files
    1. save as .xlsx in the overlap directory
    2. extract the data worksheet name
    3. write file errors to YYYY\_SeparateFileStatus.txt
    4. get max # of rows and columns from each file
    5. find the Date column
    6. get the start and end rows of data (using cells that have dates)
    7. convert dates to standard format
    8. get start date/time and end date/time for each file
    9. overwrite dates and times in file with standard date/time formats
    10. copy overlapping data block from file 1 to file 2
        1. find row where start date from file 2 in file 1 begins (usually midnight)
        2. find row where start time from file 2 in file 1 is located (minus 1 row to avoid duplicate data rows)
        3. copy rows from start date to startdate/time in file 1 and insert at beginning of file 2
        4. add # of rows copied to FileOverlap.xlsx
11. delete or transfer files
    1. if only 1 file with base file name, then do nothing (no overlap file)
    2. if count of base file name is > 1 (overlap files exist), then
       1. if no rows were added, then do nothing
       2. if rows were added and file order =1, then
          1. delete .xlsx and original file from overlap directory
       3. if rows were added and file order = 2, then
          1. copy .xlsx file from overlap directory to Data\_PostProcess directory
          2. delete .xlsx file from overlap directory
          3. delete original file from Data\_PreProcess directory

## Process Raw Data Files

**General Purpose:** To process “raw” (pre-processed) unattended multiprobe and temperature logger files, such that the file format is standardized, and to link the sonde ID or serial number of the logger to the OWMID found in the electronic field sheet for the data year in question

**Filename:** UnAtt\_ProcessRawFiles4.wsf

**Code Language:** VBScript

**Input Files:**

* Data files downloaded from multiprobe or temperature loggers, stored in the appropriate folder under Data\_PreProcess for the data year in question
* Multiprobe files:
* must be comma delimited (.csv or .xmd) or Excel (.xls or .xlsx).
* If YSI or text formats are used, the files must be converted to .csv or saved as .xlsx with the same formatting as the hydrolab files (i.e. may require insertion of analyte comment columns and Date in its own column not on each row).
* Temperature files: water temperature only (not air-deploys)
* must be either comma delimited (.csv) or Excel (.xls, or .xlsx).

**Output Files:**

* YYYY\_PPStatus.txt
  + log file showing error messages in input file processing (stored in Data\_PostProcess folder for each data year); see table below for explanation of error messages
* PPXX-XXXX\_origname.xlsx
  + post-processed data files stored in Data\_PostProcess for each year (file naming convention = prefix of PP for post-processed + the 7 digit OWMID (XX-XXXX) + underscore + original file name + .xlsx)

**Approximate Run Time:** 2 hours, varies by data year and network traffic/PC speed

**Code Outline:**

1. User enters data year for analysis (YYYY)
2. Check for presence of directories and files
   1. Input files
      1. Files located in Data-PreProcess folders (Multiprobe and Temperature folders)
      2. Field sheet meta data file (MetaDataYYYY.xlsx, where YYYY is the data year)
   2. Create output directory called Data-PostProcess, if it doesn’t exist
3. set up text file for recording progress, issues (YYYY\_PPStatus.txt)
4. set up second text file for recording date corrections (for purposes of code debugging)
5. open each file in Data-PreProcess directory, get headers, reformat file
   1. process multiprobe files
      1. get multiprobe file count
      2. for each file extension type, open the file, save it as temp .xlsx file in the Data-PostProcess folder and close it
      3. call VBScript subroutine to process multiprobe formatted files (**processHydFile.vbs)**
         1. open the temp.xlsx file
         2. check for data worksheet with the same base name as original PreProcess file
            1. if worksheet name is incorrect, write error to YYYY\_PPStatus.txt and exit subroutine
         3. check for OWMID after “Log File Name:” header
            1. if none found, then write error to YYYY\_PPStatus.txt and exit subroutine
            2. if OWMID doesn’t contain dash, then add one
         4. match OWMID to field sheet meta data file
            1. filter worksheet “Combined Meta Data” for OWMID

if no match found, write error to YYYY\_PPStatus.txt and exit subroutine

if more than one match found, then write error to YYYY\_PPStatus.txt and exit subroutine

if one match found, then get SmpTypName

if SmpTypName is not equal to “In-situ: Unattended”, then store error for later output to YYYY\_PPStatus.txt

* + - 1. get time interval for file from header “Interval”, convert to minutes (see 15 below)
      2. get sondeID from first row of data file, if available
      3. check for presence of at least one valid date format in “Date” column
         1. if no dates found, write errors to YYYY\_PPStatus.txt and exit subroutine
      4. identify starting row for data (first row from “Date” header that contains a valid date format) and ending row for data (working backwards from last row in file to first row that contains a valid date format)
      5. delete extra rows before starting data row and after ending data row
      6. check for correct time format in “Time” column
         1. if at least one bad time format found, write errors to YYYY\_PPStatus.txt and exit subroutine
      7. insert columns at start of worksheet and write OWMID and SondeID to these new columns
      8. create new workbook with name "PP" & OWMID & "\_" & base file name & ".xlsx"
      9. find location of column headers in temp.xlsx file and copy these columns to new PP workbook using a standard column order
         1. if headings not found, store as messages to be later written to YYYY\_PPStatus.txt
      10. convert dates and times to standardized format
          1. delete rows where date and/or time is “bad” (identified as -1 by date/time conversion subroutines), usually due to the presence of text (e.g., “Power failure”)
          2. store # of bad rows deleted to be later written to YYYY\_PPStatus.txt
      11. if time interval was not found in file header, then calculate the median time interval (in minutes) between points in the file using date + time values
          1. store time interval as message to be later written to YYYY\_PPStatus.txt
      12. check for correct time interval between adjacent points in the file
          1. compare ratio of delta time between points to the file time interval

if ratio < 0, then write error message to YYYY\_PPStatus\_Dates.txt

if ratio is > 1, then determine the # of missing rows and add them; insert dates and times only on new rows

if ratio = 1, then no new rows needed between adjacent pair of data points

* + - * 1. if relevant, write # of rows added or whether ratio < 0 to YYYY\_PPStatus.txt
      1. add standard column headers
      2. check for occurrence of “Power loss” in original file
         1. if “Power loss” found at least once in file, store message for later writing to YYYY\_PPStatus.txt
      3. save new PP workbook, close temp.xlsx workbook
      4. write stored messages/errors to YYYY\_PPStatus.txt
  1. process temperature logger files
     1. get temperature logger file count
     2. for each file extension type, open the file , save it as temp.xlsx file in the Data-PostProcess folder and close it
     3. call VBScript subroutine to process temperature-logger formatted files (**processTFile.vbs**)
        1. open the temp.xlsx file
        2. check for data worksheet with the same base name as original PreProcess file
           1. if worksheet name is incorrect, write error to YYYY\_PPStatus.txt and exit subroutine
        3. check for logger serial number after “Plot Title:” or “Serial Number:” header
           1. if not found, then write error to YYYY\_PPStatus.txt and exit subroutine
        4. get time interval for file from header “Interval”, convert to minutes (note: this is not a typical header in temperature files: see 15 below)
        5. check for presence of at least one valid date/time format in “Date/Time” column
           1. if no dates/times found, write errors to YYYY\_PPStatus.txt and exit subroutine
        6. identify starting row for data (first row from “Date/Time” header that contains a valid date/time format) and ending row for data (working backwards from last row in file to first row that contains a valid date/time format)
        7. delete extra rows before starting data row and after ending data row
        8. check for presence of “Logged” on any rows
           1. count # of rows where “Logged” appears
           2. delete rows where Logged appears in any column
        9. match serial # and date to field sheet meta data file
           1. extract date from starting row of data (note no longer needed)
           2. filter worksheet “Combined Meta Data” for SmpTypName = “In-situ: Unattended”, get visible range and turn off filter
           3. filter worksheet “Combined Meta Data” for serial # (SNSampGear column) and get visible range
           4. intersect visible ranges from b and c to determine which in-situ unattended records overlap with the logger serial #

if no match found, write error to YYYY\_PPStatus.txt and exit subroutine

if more than one match found, then write error to YYYY\_PPStatus.txt and exit subroutine

if one match found, then get OWMID

* + - 1. insert columns at start of worksheet and write OWMID and SondeID to these new columns
      2. create new workbook with name "PP" & OWMID & "\_" & base file name & ".xlsx"
      3. find location of column headers in temp.xlsx file and copy these columns to new PP workbook using a standard column order
         1. if headings not found, store as messages to be later written to YYYY\_PPStatus.txt
      4. split date/time column into separate date and time columns
         1. delete rows where date and/or time is “bad” (identified as -1 by date/time conversion subroutines), usually due to the presence of text (e.g., “Power failure”)
         2. store # of bad rows deleted to be later written to YYYY\_PPStatus.txt
      5. get temperature units from header
         1. if units = “F” (Fahrenheit), then convert temperature values to Celsius and store message for later writing to YYYY\_PPStatus.txt
      6. if time interval was not found in file header, then calculate the median time interval (in minutes) between points in the file using date + time values
         1. store time interval as message to be later written to YYYY\_PPStatus.txt
      7. check for correct time interval between adjacent points in the file
         1. compare ratio of delta time between points to the file time interval

if ratio < 0, then write error message to YYYY\_PPStatus\_Dates.txt

if ratio is > 1, then determine the # of missing rows and add them; insert dates and times only on new rows

if ratio = 1, then no new rows needed between adjacent pair of data points

* + - * 1. if relevant, write # of rows added or whether ratio < 0 to YYYY\_PPStatus.txt
      1. add standard column headers
      2. check for occurrence of “Power loss” in original file (note: this shouldn’t be an issue with temperature files)
         1. if “Power loss” found at least once in file, store message for later writing to YYYY\_PPStatus.txt
      3. save new PP workbook, close temp.xlsx workbook
      4. write stored messages/errors to YYYY\_PPStatus.txt
  1. find unmatched field sheet meta data OWMIDs (OWMIDs that were entered into the WQD database, but which have no associated data files)
     1. loop through files in Data-PostProcess folder, get OWMID from file name, store OWMIDs in an array
     2. loop through field sheet OWMIDs, comparing against data file OWMIDs
        1. Field sheet OWMIDs and associated meta data are written to YYYY\_PPStatus.txt as unmatched
        2. Total counts (# OWMIDs in field sheet, # OWMIDs matched to data files, # OWMIDs not matched to data files) are written to YYYY\_PPStatus.txt

### Post-Processed Status File Messages

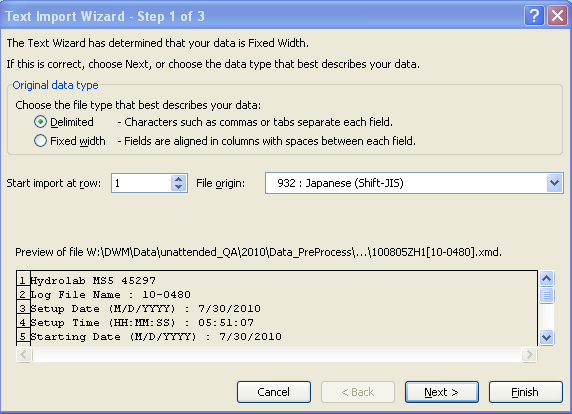
Explanation of messages found in file PPStatus.txt, located in the unattended\_QA “year” folder under Data\_PostProcess, and common solutions are found in the table below. All corrections should be made to the electronic field sheet or the data files in the Data\_PreProcess folder for the year in question. In some cases it may be necessary to delete or remove files from the PreProcess folder (for example, if some files are found to be redundant, test, empty, or for incorrect years, etc.) After corrections are made, the code will be re-run and the error message/data reconciliation process will continue until the messages are resolved.

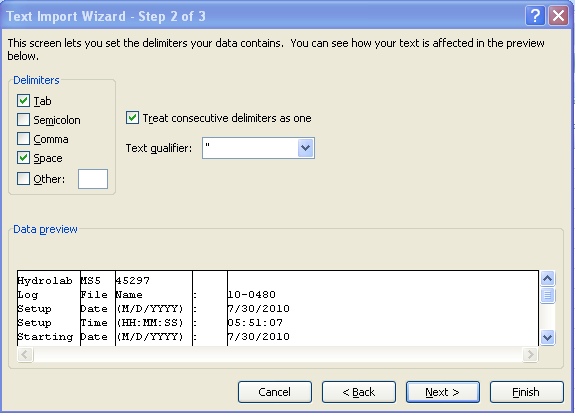
| **Error Messages/Analyte Headings Not Found** | **Applies to File Type** | **Potential Problem(s)** | **Potential Solution(s)** |
| --- | --- | --- | --- |
| Added XX new rows (where XX is a number) | Multiprobe, Temperature Logger | N/A (text is informational only) | N/A (text is informational only) |
| Deleted XX bad rows (where XX is a number) | Multiprobe, Temperature Logger | N/A (text is informational only) | N/A (text is informational only) |
| Header Not Found | Multiprobe, Temperature Logger | Multiprobe file doesn’t contain Log File Name header followed by OWMID; Temperature file doesn’t contain Plot Title or Serial Number header followed by serial number | Add the header row above where the data starts to data pre-process file; re-export file from lab PC |
| Incorrect Data Worksheet Name | Multiprobe, Temperature Logger | Worksheet name must be the same name as the file | Change data worksheet name of pre-process data file to match file name |
| Incorrect Sample Type (SmpTypName) on Field Sheet | Multiprobe | Field sheet SmpTypName field shows a value other than “In-situ: Un-attended” | Change SmpTypName on electronic field sheet for OWMID in question to “In-situ: Un-attended” |
| Incorrect Time Format | Multiprobe | Time column in file appears as HHMMSS (without colon separators) | Insert colons between hours, minutes, and seconds (HH:MM:SS) |
| Interval between rows < 0.5 x File Time Interval | Multiprobe, Temperature Logger | Date/Time between adjacent rows is incorrect (e.g. previous row shows 7/31/10 11:30:00 PM and current row shows 7/31/10 12:00:00 AM), OR time interval between adjacent rows is less than the file time interval (previous row 7/31/10 10:00 AM, current row 7/31/10 10:01 AM, next row 7/31/10 10:30 AM, file time interval is 30 minutes) | Check that dates that were manually filled in (for example text files that were converted to .xlsx) are correct (in example: previous row should be 7/30/10 11:30:00 PM and current row remains 7/31/10 12:00:00 AM); manually delete rows where time interval is not correct (in example: current row 7/31/10 10:01 AM should be deleted) |
| More than 1 match found in field sheet file for OWMID | Multiprobe | OWMID has been duplicated in field sheet | Check field sheet file against paper field sheets, make correction to field sheet as needed |
| More than 1 match found in field sheet file for Serial Number | Temperature Logger | Serial number of temperature logger has been duplicated in field sheet for SmpTypName = “In-situ: Unattended” | Check field sheet file against paper field sheets, make correction to field sheet as needed |
| No Match found in field sheet file for OWMID | Multiprobe | OWMID of multiprobe is incorrect or missing in data file; OWMID is missing from field sheet file; OWMID must be listed after the heading **Log File Name:** | Check file in Data\_PreProcess folder; re-export file from lab PC or manually correct file so that it may be post-processed; if probe was not deployed, OWMID may be missing on electronic field sheet (add new record to field sheet); OWMID may have been incorrectly entered into datalogger file (check for number transposition) |
| No Match found in field sheet file for Serial Number | Temperature Logger | serial number of temperature logger is incorrect or missing in first row of data file; serial number must be listed after the heading **Plot Title:** or **Serial Number:** | Check file in Data\_PreProcess folder; re-export file from lab PC or manually correct file so that it may be post-processed; if probe was not deployed, serial number may be missing on electronic field sheet (add as needed to SNSampGear field) or OWMID may not have been entered (add new record to field sheet) |
| Power loss | Multiprobe | N/A (text is informational only) | N/A (text is informational only) |
| Specific analyte name is listed | Multiprobe, Temperature | Analyte was not sampled or programmed; analyte name was misspelled | check the file in Data\_PreProcess folder; if heading is there, check spelling, correct spelling in file if necessary; if heading is not there, but you suspect it should be, re-export file from the lab computer or check lab calibration book to see what analytes were calibrated for on the sonde |
| Temperature converted from degrees F to C | Temperature Logger | Temperature recorded in raw data file was in Fahrenheit | No changes needed; alert lets user know that temperature was automatically converted from Fahrenheit to degrees C |
| Time Interval (MIN) = XX (where XX is a number) | Multiprobe, Temperature | N/A (text is informational only) | N/A (text is informational only) |
| Unrecognized Date Format/Empty File | Multiprobe, Temperature Logger | Date/Time format is incorrect or data file is empty (multiprobe files must list date and time in separate columns; temperature probe files must list date and time in same column) | Change Date and Time format of file in Data\_PreProcess folder; re-export file from lab PC |
| Unrecognized file format | Multiprobe, Temperature Logger | Raw data file does not have the correct extension (.xmd, .csv, .xls, .xlsx) | Re-export file from lab PC with correct file extension and format or open file and save with correct extension |

### Correcting File Formats

To correct the format of a .txt file or an .xmd file that was saved as space-delimited text:

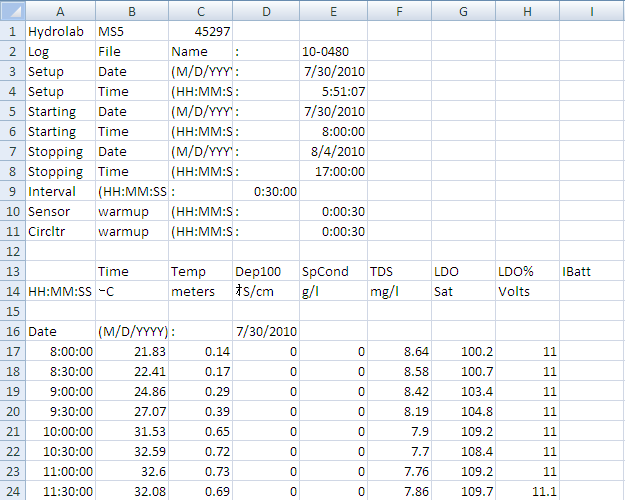
1. Open Excel, then open the .xmd file. Choose delimited in the window that pops up, click Next, then choose tab and space delimited.



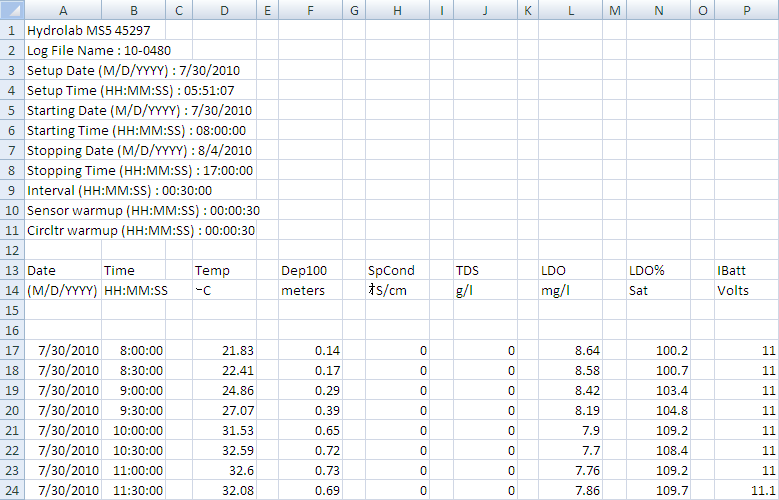


1. Click Next again and Finish. Data will now be separated into individual columns. Check for misalignment of the headings and re-align with the correct data columns.
2. Delete the upper block of cells that contain the sonde ID and OWMID, open the original file in NotePad or WordPad, copy and paste this same header block into column A.
3. The starting date will be in its own row: insert a date column and move the date header into the cell before Time, also move the date format to the cell below the date header. Insert the dates for each block of time from 0 to 23:30 (using the appropriate file time interval) for each day (do not fill down with the date, Excel won’t be able to figure out where the breaks are in each day).
4. The analyte comment columns are absent in the .txt version of the hydrolab files: insert the extra columns after Time (not Date) and each analyte/battery column.

How the file first appears:



How the file should look:



## Automated Validation and Processing

**General Purpose:** Trimming OWMID files, linking to field sheet and sample meta data, determining field and lab QC decisions, and applying initial analyte QC2 decisions based on the appropriate decision matrix for combining lab and field QC information (described in Appendix C)

**Filename:** UnAtt\_TrimQA\_24.wsf

**Code Language:** VBScript

**Input Files:**

* PPXX-XXXX\_origname.xlsx
  + post-processed data files stored in Data\_PostProcess for each year (file naming convention = prefix of PP for post-processed + the 7 digit OWMID (XX-XXXX) + underscore + original file name + .xlsx)

**Output Files:**

* XX-XXXX\_QC2\_MM-DD-YYYY.xlsm or .xlsx
  + QC2 files stored in the QC2 folder named after the code run (yyyyUQC2\_MM-DD-YYYY, where yyyy is the data year, MM is the month, DD is the day, and YYYY is the year the code was run)
  + .xlsm files contain unattended data. The xlsm extension indicates that the file contains VBA macros. The macros will:
    - prevent changes to QC2 files prior to applying final QC2 review decisions (accept/qualify/censor) for analytes (see Appendix D: Apply Final QC2 Review Decisions)
    - check for completeness/errors when manually editing the QC2 files (after final QC2 review decisions have been applied)
  + .xlsx files are placeholders for OWMIDs without post-processed data files. No post-processed file indicates that there was no raw data file to begin with or that the raw file was set aside as unusable during the file reconciliation process (see Appendix D: Process Raw Data Files)
* yyyy\_Unattended\_QC2\_Summary\_MM-DD-YYYY.xlsm
  + QC2 summary file used as the master file for recording QC2 review comments and applying final QC2 review decisions, for storing summary information about the OWMID files (for example: analytes present, statistics, flags for error checking, unattended-attended OWMID sets, etc.)
  + .xlsm indicates that the file contains a VBA macro. The macro (apply\_qcstatus) applies the final QC2 review decisions to the individual OWMID files (see Appendix D: Apply Final QC2 Review Decisions)

**Approximate Run Time:** 1-3 hours

**Code Outline:**

1. Announce to user what is about to happen and give user the chance to back out
2. User enters data year for analysis
3. check for presence of directory and files
   1. user navigates to the location of the field sheet summary file (MetaDatayyyy.xlsx, where yyyy is the data year) and selects it
   2. user navigates to the location of the QC4 attended data file
   3. Copy attended and field sheet files to QC2 subfolder
   4. post-processed file directory
   5. user navigates to the location of the lab QC file and selects it
   6. VBA modules
      1. For insertion into QC2 summary workbook
         1. FinalizeUnattData\_QC2Review
      2. For insertion into OWMID files
         1. Unatt\_QC2\_checkstatus.bas
         2. Unatt\_QC2\_update\_datareporting.bas
         3. Unatt\_QC2\_datasheet.cls
         4. Unatt\_QC2\_ThisWorkbook.cls
         5. Unatt\_QC2\_copy\_to\_backup.bas
         6. Unatt\_QC2\_QCSumsheet.cls
   7. Unattended probe template file (for read me sheet of QC2 summary workbook, explaining worksheets)
   8. Create directories
      1. QC2 folder (W:\DWM\Data\unattended-QA\yyyy\QC2)
      2. Subfolder for code run in the QC2 folder (named yyyyUQC2\_MM-DD-YYYY)
   9. Create text file for recording issues (yyyy\_Unattended Trim Files and QA\_Status.txt)
4. Create QC2 summary workbook
   1. define worksheets
      1. File Process Info
      2. Analytes Present
      3. Flags
      4. Unattended-Attended OWMID Sets
      5. Statistics
      6. QC2 Decision Summary
      7. Lab QC Summary
   2. Define columns
   3. Add headers
   4. Apply formatting
   5. Copy read me sheet from template workbook
5. open field sheet summary workbook
   1. get column locations from “Combined Meta Data” sheet
   2. copy columns to QC Summary Unattended-Attended OWMID Sets worksheet
      1. clear and apply formatting to Unattended-Attended OWMID Sets worksheet
      2. create temporary column to store check for row deletions
      3. loop through Sample Type column and mark rows with X for deletion if not In-situ: Un-attended or In-situ: QC Cal Check
      4. delete rows marked by an X and delete temporary column
      5. get used range of data
      6. convert dates and times to standard formats
   3. count QC sondes for each unattended OWMID (using FSLOG) on Unattended-Attended OWMID Sets sheet and determine which is drop off, pick up, intermediate, or unknown (determined by StartDate); write results to Unattended-Attended OWMID Sets sheet
6. Copy unattended OWMIDs to flags sheet in QC Summary workbook
   1. Sort Unattended-Attended OWMID Sets by SmpTypName
   2. Filter by “In-situ: Un-attended”
   3. Copy unattended OWMID, FSLOG, Gear, SondeID, UniqueID, Project, Duplicate, Duplicate OWMID, StartDate, and StopDate to Flags sheet
   4. Get used range of data on Flags sheet
7. Loop through OWMIDs on Flags sheet and match to post-processed file
   1. Loop through files in post-processed file directory
      1. If match found, write post-processed file name to PP File column on Flags sheet
      2. If match not found, write “NO FILE” to PP File column on Flags sheet
8. Copy OWMID, field sheet columns, and PP File column from Flags sheet to Analytes Present sheet, Lab QC Summary sheet, and QC2 Decision Summary sheet
9. Open Status file from post-processing of files to check for power loss issues
   1. Open yyyy\_PPStatus.txt for yyyy data year as comma delimited Excel file
   2. Find occurrence of current PP file name
   3. Check for “Power loss” text flag and if found, set powerflag = “Yes”
   4. Check for “Time Interval” text flag and if found, get the file time interval
   5. Write the power loss flag to Flags worksheet and Time Interval to Analytes Present sheet
   6. Find info about Post-processing of files (VBScript code name, Start and End Date/Times, Raw Data File Directory) and write to File Process Info
   7. Close yyyy\_PPStatus.txt
10. Determine analytes present in workbook
    1. Sort Unattended-Attended OWMID Sets sheet by FSLOG, SmpTypName, and StartDate
    2. Define standard columns on data worksheets from post-processed files
    3. Get OWMID, PP File, SondeID from Analytes Present sheet
       1. If PP File = “NO FILE”, then write zeroes/No Data to Analytes Present sheet
       2. If PP File exists, then open file
          1. Get used range of data, write # of rows to Analytes Present sheet
          2. Loop through each analyte header column
             1. Filter for non-blank cells and get count
             2. Filter adjacent “SS” comment column for blank cells and get count
             3. Calculate percent of analyte column that contains data without comment (comment symbols imply no data, no calibration, etc.)

Create text flag for “DATA IN ALL ROWS”, “SOME DATA”, or “NO DATA”

* + - * 1. Filter analyte column for zero values and get count
        2. Calculate percentage of zeros in analyte column
        3. Write percentages of zeros and percentages of values, and text flag, to Analytes Present sheet
      1. Get sondeID from post-processed file and write to Analytes Present sheet
      2. Check whether Field sheet sondeID matches sondeID from file and write Yes/No to Analytes Present sheet
      3. Close PP file

1. Match Lab QC file info to OWMIDs on Lab QC Summary sheet
   1. Open Lab QC file
      1. Get used range
      2. Get specific column headings (Year, Cal Data, Post-Cal Date, Project Name, SondeID, Parameter, QC Result)
      3. Copy header row to Lab QC Summary sheet (after previously copied columns)
   2. Sort Lab QC Summary file by year, project name, cal date
   3. Filter Lab QC Summary file by data year
   4. Loop through each data year row in lab QC file
      1. get calibration info from columns in XI.a.ii. above
      2. Filter Lab QC Summary sheet by project name found in Lab QC file
      3. Filter Lab QC Summary sheet for all StartDates between the Cal Date and Post-Cal Date found in Lab QC file
      4. Filter Lab QC Summary sheet for the SondeID found in Lab QC file
      5. Get count of filtered rows
         1. If count = 0, then write “No match found for lab QC info” to program text file
         2. If count > 0 and project exists on Lab QC Summary sheet, then write “Match already exist for OWMID” to program text file
         3. If count > 0 and project name is blank, then write “Match found for OWMID” and row number to program text file and lab QC info to Lab QC Summary sheet
   5. Close Lab QC file
   6. Fill out any blank QC Result cells on Lab QC Summary sheet
      1. Get used range of Lab QC Summary sheet after adding calibration info
      2. Filter Lab QC Summary sheet by blank QC Status column (indicating no match was found to Lab QC file) and get count
      3. If count of blanks > 0 then set blank QC results = Accept
   7. NOTES:
      1. Lab QC file only logs calibration problems
      2. No match found for OWMIDs to lab QC file are automatically assumed to be “Accept” (i.e. without problems); this assumes that the lab QC file is accurate and doesn’t contain omissions
      3. All lab QC temperature decisions are automatically Accept for “i” qualifier
2. Prep before creating unattended QC2 review files
   1. Define standard columns for QC2 workbooks, main data sheet
   2. Loop through each row on Flags sheet
      1. Get OWMID, PP File, and Gear
      2. Define QC2 file extension based on PP File existence
         1. If PP file = “NO FILE”, then file extension = .xlsx
         2. If PP file exists, then file extension = .xlsm
      3. Define QC2 file name and path
      4. write QC2 workbook name to QC2 File column on Flags, Analytes Present, QC2 Decision Summary, and Lab QC Summary sheets
      5. get time interval of file from Analytes Present sheet
   3. if PP file = “NO FILE” then
      1. create blank QC2 workbook
      2. insert “NO DATA for OWMID” on sheet named after OWMID
   4. if PP file exists then create QC2 file (proceed to step XIII.)
3. create unattended QC2 files
   1. create QC2 workbook with sheets:
      1. data sheet named after OWMID
      2. Summary Statistics
      3. Daily Statistics
      4. QC Summary
      5. Field QC
      6. Lab QC
   2. Open PP file and copy data to data sheet columns in QC2 file
   3. Insert columns for Qualifiers and QC Decisions on data sheet
      1. Columns to the right of analyte columns (“comment” columns) become qualifier columns
      2. New columns inserted before qualifier columns become QC Decision columns
   4. Combine date and time into 1 column following time column
   5. Get attended OWMIDs associated with the current unattended OWMID
      1. Find current unattended OWMID on Unattended-Attended OWMID Sets sheet
      2. Filter by FSLOG, then by “In-situ: QC Cal Check”
      3. If no rows found then write to .txt file “NO QC rows found for OWMID”
      4. If rows found then
         1. Copy attended OWMIDs to field QC sheet in QC2 file
         2. Open attended (QC3) workbook
         3. On attended QC4 final data sheet, find each attended OWMID and copy info to field QC sheet (data value, QC Decision, qualifiers, etc.), including row number
         4. If match not found write “NO MATCH FOUND in ATTENDED QC3 FILE” for OWMID listed on field QC sheet
   6. Trim unattended file based on Attended QC info
      1. Find “Pick up” on Field QC sheet
      2. Find “Drop off” on Field QC sheet
      3. Get Date and Time for Pick up and Drop off attended QC
      4. Record whether missing dates/times exist on Flags sheet
      5. Find first and last dates and times on unattended data sheet
      6. Store unattended dates and times in arrays
      7. Find unattended row that matches attended drop off date/time
         1. If attended drop off QC equals unattended date/time, then get next unattended date/time row
         2. If attended drop off QC is earlier than a given unattended date/time, then use current unattended date time row
         3. If attended drop off QC is later than a given unattended date/time use previous unattended date time row
         4. If no match found, use first unattended date/time row; record as “Unattended starts after attended QC” as Yes on Flags sheet
      8. Find unattended row that matches attended pick up date/time
         1. If attended pick up QC equals unattended date/time then use previous unattended date/time row
         2. If attended pick up QC is earlier than a given unattended date/time, then use previous unattended date/time row
         3. If attended pick up QC is later than a given unattended date/time, then use the current unattended date/time row
         4. If no match found, use last unattended date/time row; record as “Unattended stops earlier than attended QC” as Yes on Flags sheet
      9. Delete rows after unattended row match to pick up QC
      10. Delete rows before unattended row match to drop off QC
   7. Get new unattended start and end dates and compare to field sheet start and stop dates
      1. If dates differ then write “Yes” to Flags sheet (File start date doesn't match FS or File stop date doesn't match FS)
   8. Add data to LabQC worksheet
      1. Filter Lab QC Summary sheet by current OWMID
      2. Copy header and row to Lab QC sheet
         1. Add qualifiers flags for missing and no data
            1. NO DATA

Filter qualifier column for non-blank values and change existing symbols to “NO DATA”

Filter analyte column for text values and change analyte cells to blank and qualifier cells to “NO DATA”

* + - * 1. MISSING

Filter analyte columns for blanks and qualifier column for blanks; add “MISSING” to qualifier column cells

* 1. Determine QC status for each analyte
     1. Rebuild arrays that store unattended date/times after trimming occurred
     2. Sort field QC sheet by date and get attended QC info
     3. Check each attended QC for tidal (“t”) qualifier
     4. Check each attended QC value for no water (“^^”)
     5. Combine attended dates and times
     6. Get attended QC type and get closest unattended row
        1. Drop off QC = first unattended row
        2. Pick up QC = last unattended row
        3. Intermediate QC = unattended row with closest date time match
     7. Get unattended values for each QC type
     8. Calculate time differences between unattended and attended points for each QC type (if either value is missing write -99999 to ‘QC Summary’ sheet
     9. Calculate absolute delta analyte values between unattended and attended points for each QC type (if either value is missing write -99999 to ‘QC Summary’ sheet)
     10. store results in an 2-D array
     11. calculate average delta values for each analyte based on values stored in the 2-D array (if time difference is > 30 minutes or if any delta value = -99999, then don’t include in average)
     12. determine field QC decisions for each analyte based on average delta values (**see Appendix C: Decision Criteria**;Note:if average delta cannot be calculated, then analyte field QC Decision = “NONE”)
     13. get lab QC decision for DO from lab QC sheet (T decision is automatically Accept)
     14. determine overall QC decision for each analyte based on combination of lab and field QC decisions (**see Appendix C: Decision Criteria**)
     15. write decisions to QC Summary sheet in QC2 file and to QC Decision sheet in summary workbook
  2. if any attended QC had tidal qualifiers write Yes to Flags sheet, Potential Tidal Influence column
  3. if any attended QC had no water symbol, then add Yes to Flags sheet, no water for attended QC column
  4. add “i”qualifier/QC decisions for each analyte to data sheet
  5. add standards for graphing to data sheet for all dates/times
  6. create final data sheet
     1. make a copy of main data sheet and rename as “XX-XXXX final”, where XX-XXXX is the 7-digit OWMID
     2. add symbols to final worksheet
        1. for DO, replace low values with <0.2
        2. for DOsat, replace low values with <2
        3. for any analyte with CENSOR in QC Decision column, replace data value with "##" (this must come first otherwise missing data cells are no longer flagged)
        4. for any analyte with MISSING in qualifier column, replace data value with "\*\*" and set qualifier cell to blank
        5. for any analyte with NO DATA in qualifier column, replace data value with "--" and set qualifier cell to blank
        6. set data changes flag to false (to prevent macros in OWMID files from running when file is saved)
  7. create graphs-stats sheet
     1. Copy final sheet and rename as “XX-XXXX graphs-stats”, where XX-XXXX is the 7-digit OWMID
     2. Replace non-numeric data values to allow for graphing and future statistical calculations
        1. Replace “\*\*”, “--“, and “##” cells with blanks
        2. Replace <0.2 and <2 with 0.2 and 2, respectively
  8. Generate charts
     1. Determine minimum date for x-axis
     2. Determine maximum # of days in file (# of rows \* time interval of file)
     3. Determine x-axis units to use based on max # of days
        1. if max days > 49, Major unit = 30 days, minor unit = 1 day
        2. if max days < 49, Major unit = 7 days, minor unit = 0.25 day
     4. create separate T, DO, DOsat sheets (skip DO, DOsat sheets if Gear = Temperature Logger)
     5. define x and y data ranges for charts, including standards
     6. add data series to charts and format charts
  9. add list for manual changes to QC Decision columns on data sheets
     1. create lists sheet with choices for pull-down list
     2. define named range (“QCResult”) based on this list of values
     3. apply named range to data sheet analyte QC Decision columns
  10. create copies of main data sheet
      1. copy main data sheet and rename as “XX-XXXX backup”, where XX-XXXX is the 7-digit OWMID; backup sheet is used by VBA code to check for changes to main data sheet during manual editing of file
      2. copy main data sheet and rename as “XX-XXXX init”, where XX-XXXX is the 7-digit OWMID; initial sheet is preserved as the initial QC2 data sheet with original automated decisions (in case we need to revert back to this initial state)
  11. import VBA modules
      1. Unatt\_QC2\_checkstatus.bas
      2. Unatt\_QC2\_updatereporting.bas
      3. Unatt\_QC2\_copy\_to\_backup.bas
      4. Unatt\_QC2\_datasheet.cls
      5. Unatt\_QC2\_thisworkbook.cls
      6. Unatt\_QC2\_QCSumsheet.cls
  12. Protect data sheet and set protect sheet flag to true
  13. Set changes flag to false on final sheet
  14. Save and close QC2 file

1. Finish up with QC Summary workbook
   1. QC Decision Summary
      1. Add “N/A” to QC Decision columns where PP file = “NO FILE”
      2. Add “i“ qualifier to qualifier columns if QC Decision = “QUALIFY” or “CENSOR”
      3. Change text color in final columns to red (to indicate these are the unedited final decisions)
      4. Create list worksheet for Accept/Qualify/Censor/N/A, add named range “QCResult”, apply named range list to analyte final decision column
   2. Flags sheet
      1. Calculate deltas and slopes between analyte points in QC2 files
         1. Open each QC2 file (skipping PP file = “NO FILE”)
         2. Write final QC2 Decisions temporarily to “QC Summary” sheet to prevent macros from running
         3. Unprotect main data sheet
         4. Add headers for delta analyte values and slopes
         5. Get file time interval from Analytes Present sheet
         6. Add formulas to calculate delta between adjacent analyte values
         7. Add formulas to calculate slopes (defined as delta value/file time interval) between adjacent analyte points
      2. Determine if analyte threshold values and spikes have been tripped
         1. Filter for blank analyte values to identify data gaps
         2. Filter battery voltage, temperature, pH, SpC, DO for values above or below threshold values (see Appendix D: QC Summary Workbook: Explanation of Worksheets for threshold values); count # of cells
         3. Filter slope columns for temperature, pH, SpC, DO for values greater than spike values (see Appendix D: QC Summary Workbook: Explanation of Worksheets for threshold values); count # of cells
      3. Close QC2 files
         1. Delete temporary QC2 decisions from QC Summary sheet
         2. Autofilter data sheet, go to “A1”, freeze top row
         3. Protect data sheet and set protect sheet flag to true
         4. Set changes flag to false on final sheet
         5. Save and close QC2 file
      4. Write counts of analyte spikes and threshold exceedances, and whether data gaps exist, to Flags sheet
   3. Import VBA module “FinalizeUnattData\_QC2Review.bas” for adding final QC2 decisions to QC2 files
   4. Write program end time to File Process Info
   5. Save and close summary workbook

### QC Summary Workbook: Explanation of Worksheets

**File Process Info**

Stores information pertaining to code runs for all automated steps in the unattended data validation process, including brief process description, VBA/VBScript code name, start and end times of run

**Flags worksheet explanation**

Worksheet shows flags useful in checking for data that need to be censored or qualified, missing data, files that weren’t properly trimmed due to date/time errors; values determined after post-processed files were trimmed

| **Column Label** | **Data Type** | **Description** |
| --- | --- | --- |
| **Power Loss** | Yes/No | pre-processed file indicated power loss in one or more rows |
| **Data gaps** | Yes/No | post-processed file shows data gaps where rows were inserted (usually associated with power loss in file) |
| **T, etc. Threshold Tripped** | Numeric | Count of # of times (data rows) where analyte threshold was tripped (see table below for threshold values) |
| **T spikes** | Numeric | Count of # of times (data rows) where spikes occurred (see table below for spike values); note that the delta values are calculated between the current and next row, so the total # of delta values equals the total data rows - 1 |
| **missing drop off QC time** | Yes/No | Indicates whether drop off attended QC was missing a time (so no trimming could occur, no delta value could be computed between unattended and attended point) |
| **missing pick up QC time** | Yes/No | Indicates whether pick up attended QC was missing a time (so no trimming could occur, no delta value could be computed between unattended and attended point) |
| **Unattended starts after attended QC** | Yes/No | Indicates cases where the unattended file start date/time occurs after the attended QC date/time |
| **Unattended stops earlier than attended QC** | Yes/No | Indicates cases where the unattended file stop date/time occurs before the attended QC date/time |
| **Potential tidal influence** | Yes/No | If any of the attended QC samples had a "t" qualifier for any analyte, then potential tidal influence for the unattended file is indicated |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Analyte** | **Threshold Values** | **Delta value** | **Units** | **Time interval (minutes)** | **Spike (Delta/time) value** | **Units** |
| Batt | <6.4 | N/A | volts | 30 | N/A | N/A |
| T | <5, >30 | 1 | Degrees C | 30 | 0.033 | degrees C/min |
| DO | <1, >12 | 1 | mg/l | 30 | 0.033 | mg/l/min |
| pH | <5, >8 | 0.5 | SU | 30 | 0.017 | SU/min |
| SpC | <20 | 10 | µS/cm | 30 | 0.333 | µS/cm/min |

**Analytes sheet explanation**

Percentages of analytes present in the OWMID files are determined prior to trimming post-processed files

|  |  |  |
| --- | --- | --- |
| **Column Label** | **Data Type** | **Description** |
| **File Gear Matches FS Gear?** | Yes/No | Checks whether file sondeID (File Gear) matches field sheet sondeID (SNSampGear) |
| **Avg. Time Interval** | Numeric | Average time interval (in minutes) between points in the data file |
| **Total Data Rows** | Numeric | Total # of data rows (prior to trimming) |
| **T, pH, DO, SpC, Dosat, TDS** | Text | DATA IN ALL ROWS (100% of rows contain data), SOME DATA (< 100% but > 0% of rows contain data), NO DATA (0 % of rows contain data); note that zeros are included in the count of data points |
| **T, etc. Percent** | Numeric (percent) | For the analyte listed, percent of file that contains data, calculated as # of rows where a value is present (excluding rows with blanks or indicators of no data in SS column) divided by total # of data rows |
| **T, etc. Zeros** | Numeric (percent) | For the analyte listed, percent of file that contains zeros, calculated as the # of rows where a zero is present (excluding rows with blank values or indicators of no data in SS column) divided by the total # of data rows |

**QC Decision Summary sheet explanation**

Worksheet holds all of the initial automated analyte decisions for “i” qualifier, the QC2 review comments and recommended manual changes, the final QC2 review decisions and any additional qualifiers

|  |  |  |
| --- | --- | --- |
| **Column Label** | **Data Type** | **Description** |
| **Avg Delta T** | Numeric | Average delta between the unattended and attended QC points, value determines Field QC Decision |
| **T Field QC Decision** | Text | Accept, Qualify, Censor (based on average delta value between unattended and attended) |
| **T Lab QC Decision** | Text | Accept, Qualify, Censor (based on decision provided by Jeff Smith during lab calibration of sondes) |
| **T Inital QC2 Decision**  **(i Qualifier)** | Text | Accept, Qualify, Censor (based on decision matrix between lab and field QC decisions) |
| **T Final QC2 Decision**  **(i Qualifier)** | Text | Placeholder for QC2 review adjustments to initial QC decision for "i" qualifier |
| **Avg Delta DO** | Numeric | Average delta between the unattended and attended QC points, value determines Field QC Decision |
| **DO Field QC Decision** | Text | Accept, Qualify, Censor (based on average delta value between unattended and attended) |
| **DO Lab QC Decision** | Text | Accept, Qualify, Censor (based on decision provided by Jeff Smith during lab calibration of sondes) |
| **DO Inital QC2 Decision**  **(i Qualifier)** | Text | Accept, Qualify, Censor (based on decision matrix between lab and field QC decisions) |
| **DO Final QC2 Decision**  **(i Qualifier)** | Text | Placeholder for QC2 review adjustments to initial QC decision for "i" qualifier |

**Statistics sheet**

Worksheet containing all of the station and water body information for each OWMID, in addition to statistics for each analyte (see Appendix E for more detail on the statistics); worksheet is not populated until the Create OWMID Statistics module has been run.

**Station OWMIDs**

Worksheet containing all of the OWMIDs used to build the station files. Worksheet is not created until the Create Station Files module has been run. This sheet is required for running the station statistics code.

**Station Statistics sheet**

Worksheet containing all of the station statistics, including the station and water body information for each station. Stations files are defined by UniqueID and only include data from multiprobe OWMIDs. Worksheet is not created until the Create Station Statistics module has been run.

**Lab QC Summary**

Worksheet shows the OWMID and associated lab QC decisions (matching involves the calibration date range, the sondeID, and the project name); note that only qualified or censored lab QC decisions were summarized in the lab QC file, so all other OWMIDs automatically get Accept for DO and DOsat

**Unattended-attended OWMID Sets**

Worksheet that shows an abbreviated set of metadata columns with rows relating only to the unattended-attended data sets (linked by common FSLOG numbers); used by the code to identify the pick up, drop off, and intermediate attended field QC OWMIDs to use for trimming the files and for determining the “i” qualifier automated decisions

## Apply Final QC2 Review Decisions

**General Purpose:** To apply final QC2 review decisions for each analyte, and the associated qualifiers, to the OWMID files based on the final decisions listed on the QC2 Decision Summary worksheet in the QC2 Summary workbook for the data year being validated; macro is run from within the QC2 Summary workbook

**Filename:** FinalizeUnattData\_QC2Review.bas

**Macro name:** apply\_qcstatus

**Code Language:** VBA

**Input Files:**

* input files are the individual OWMID files (e.g. XX-XXXX\_QC2\_MM-DD-YYYY.xlsm, where XX-XXXX is the 7 digit OWMID and MM, DD, and YYYY are the month, day, and year, respectively, that the file was generated)

**Output Files:**

* same as the input files

**Approximate Run Time:**  10-15 minutes, depending on network speed, # of files to be processed for the data year, and size of original files (temperature logger files take longer to process than multiprobe files)

**Code Outline:**

1. tell user what the code does and ask if they want to continue
   1. WARNING: after the final QC2 decisions have been applied by running this code, and after any edits have been made to individual OWMID files, DO NOT run this code again, otherwise row-by-row changes will be overwritten
2. write start time to File Process Info sheet
3. check for correct number of QC2 files
   1. count # of files in same directory as the QC2 Summary workbook
   2. get # of OWMIDs from QC2 Decision worksheet
   3. compare counts and tell user if they don’t match
4. check for blank QC decision cells for each analyte
   1. tell user if blank decision cells exist
5. check for cases where decision is ACCEPT but qualifiers exist other than MISSING or NO DATA
   1. tell user if errors exist
6. check for correct qualifier format (capitalization, spacing) and invalid qualifier codes in each analyte qualifier column
   1. define acceptable qualifiers (c, i, m, r, s, t, u)
   2. change qualifiers to lower case (i listed first automatically changes to "I")
   3. check for invalid qualifiers
   4. check for duplicated qualifier symbols (like rr or r, t, r)
   5. tell user row #s and analyte where invalid or duplicated qualifier codes were found
7. apply final qc decisions by looping through files listed on QC2 Decision Summary sheet
   1. skip cases where Post-Processed File = NO FILE
   2. open each QC2 file
   3. unprotect data sheet to allow changes to occur in the future and in this macro
   4. write final QC2 decisions/qualifiers for each analyte to QC Summary sheet in OWMID file
      1. set protect sheet flag to false to allow manual edits to data sheet after this macro is run
   5. check for changes to analyte decisions and qualifiers between initial and final decisions listed on QC2 Decision summary sheet
      1. if no change from initial decision/qualifiers for all analytes, then
         1. move on to next OWMID
      2. if changes exist for any analyte in file, then move on to step 7f.
   6. write final QC2 decisions and qualifiers for each analyte, if applicable, to data sheet
      1. check individual rows for NO DATA or MISSING flag in the qualifier column
         1. if these flags exist, then add new qualifiers after NO DATA or MISSING (e.g. MISSING, i, r, u)
      2. Copy data sheet to BACKUP worksheet
      3. Copy data sheet to final sheet
      4. Add symbols for reporting to final sheet
         1. for DO, replace low values with <0.2
         2. for DOsat, replace low values with <2
         3. for any analyte with CENSOR in QC Decision column, replace data value with "##" (this must come first otherwise missing data cells are no longer flagged)
         4. for any analyte with MISSING in qualifier column, replace data value with "\*\*" and set qualifier cell to blank
         5. for any analyte with NO DATA in qualifier column, replace data value with "--" and set qualifier cell to blank
         6. set data changes flag to false (to prevent macros in OWMID files from running when file is saved)
      5. Copy final sheet to graphs-stats sheet
         1. Replace “\*\*”, “--“, and “##” cells with blanks to allow graphing and for future statistical calculations
         2. Replace <0.2 and <2 with 0.2 and 2, respectively, to allow graphing and for future statistical calculations
      6. freeze top row
      7. Go to cell A1
      8. Filter data sheet
      9. Save file and close
   7. Write macro end time to File Process Info

## Calculate OWMID Statistics

**General Purpose:** To calculate statistics for multiprobe (short-term) and long-term continuous temperature data; statistics are calculated for individual QC2 or QC3 files and added to the appropriate worksheets within the files; statistics are also added to the QC2 or QC3 summary workbook; for more information on the statistics calculated see appendix E

**Filename:** UnAtt\_CalcStats\_8.wsf

**Code Language:** VBScript

**Input Files:**

* the individual OWMID QC2 or QC3 files (e.g. XX-XXXX\_QC2\_MM-DD-YYYY.xlsm, where XX-XXXX is the 7 digit OWMID and MM, DD, and YYYY are the month, day, and year, respectively, that the file was generated; and QC2 is replaced by QC3 if QC3 files are being processed)

**Output Files:**

* same as input files, with statistics added to Daily Statistics and Summary Statistics worksheets, as well as to the Statistics worksheet in the appropriate QC summary workbook
* YYYY\_Unattended Calculate Statistics.txt
  + Status file for recording issues encountered during processing

**Approximate Run Time:** 1.5 – 3.5 hours, depending on number of temperature logger files and network/PC speed

**Code Outline:**

1. announce to user what code does and give user opportunity to back out
2. user enters data year for analysis
3. user enters QC level to process (QC2 or QC3)
4. check for presence of directories/files
   1. user selects station ID and water body summary file
      1. open file, present error message to user if worksheet called Station\_WB is not found, quit
   2. user navigates to data file directory for QC level and data year entered above
      1. check for OWMID files in directory
      2. present user with error message if no files found
   3. extract year, month, day from subdirectory name in IV b.
      1. check for presence of QC summary workbook (with same date in name as subdirectory)
      2. present user with error message if QC summary workbook not found, quit
   4. create text file for recording issues (YYYY\_Unattended Calculate Statistics.txt)
5. open QC summary workbook
   1. check for presence of required worksheets (Statistics, Analytes Present, QC Decision Summary)
      1. if Analytes Present or QC Decision Summary worksheets don’t exist, then present error message to user and quit
      2. if Statistics sheet doesn’t exist, then create it otherwise clear contents from previous runs
   2. define columns on Statistics sheet and add column headings
   3. write when statistics module was started on File Process Info sheet
   4. copy unattended OWMIDs and other data from Analytes Present sheet to Statistics sheet
      1. remove previous autofilters otherwise not all cells will be copied
      2. define columns on Analytes Present sheet
      3. sort Analytes Present sheet by OWMID
      4. define destination addresses on Statistics sheet for copying (includes OWMID, project name, FSLOG, UniqueID, Gear type, SondeID, Post-Processed file, QC2/QC3 file name, Duplicate, Duplicate OWMID, and Time Interval)
   5. copy FlowStat from added worksheet to Statistics sheet
   6. copy DupesNot2Use column from QC Decision summary sheet
      1. remove previous autofilters
      2. define columns on QC Decision sheet
      3. sort QC Decision sheet by OWMID
      4. define destination addresses on Statistics sheet
   7. get used range of Statistics sheet, including max row
6. find columns in station/WB workbook
   1. station ID
      1. for 2005/2006, use STAID column
      2. for other years use STAID\_YYYY column where YYYY is the data year
      3. present error to user if correct station ID column not found and quit
   2. other columns include: Watershed, Saris/Palis/Camis, Waterbody name, description, mile point, latitude, longitude, uniqueID
7. Loop through OWMIDs on Statistics sheet and calculate statistics
   1. Define columns on data sheet in OWMID file, on Daily Statistics sheet, and on Summary Statistics sheet
   2. Get info about OWMID (file name, gear type, file time interval, UniqueID, FSLOG, Project, FLowstat)
   3. Find UniqueID for OWMID in question in station/WB workbook
      1. Write values to Statistics sheet in QC Summary workbook
      2. Write “NO MATCH FOUND” if UniqueID not found
   4. Skip files where PPFile = “NO FILE”
   5. Open OWMID file
      1. Define worksheet objects
      2. Check for presence of Daily Statistics and Summary Statistics sheets
         1. If not present, then create them
   6. Write station/waterbody info to OWMID file, Summary Statistics sheet
   7. Write FlowStat to Summary Statistics sheet in OWMID file
   8. Get FILE START TIME and FILE END TIME, TOTAL COUNT, and TOTAL DEPLOYMENT DURATION from OWMID graphs-stats sheet
   9. Format Daily Statistics sheet and add headings
   10. Calculate daily statistics
       1. Loop through each day on graphs-stats sheet, store date cells in a range
       2. filter for non-blank temperature points (all dates), store as a range
       3. filter for non-blank DO points (all dates), store as a range
       4. filter for non-blank DOsat points (all dates), store as a range
       5. check for complete day (# of rows\* time interval)/(60\*24)of values
          1. If less than complete day, then set daily statistics to blanks
          2. if a complete day, then calculate statistics
       6. Temperature statistics
          1. Intersect date range with temperature range
          2. Compare # of temperature points to # of rows for a complete day
             1. If points are missing, then skip temperature calcs (set equal to blank)
             2. If all points are present, then continue with calculations
          3. Calculate DAILY MIN, DAILY MAX, DAILY MEAN
          4. Calculate DAILY AMOUNT OF TIME > WQS (WQS = 29.4, 28.3, and 20 degrees C)
             1. Interpolate the starting date/time of T exceeding WQS
             2. Interpolate the ending date/time of T exceeding WQS
             3. Calculate the difference between the starting and ending time of each excursion above WQS
             4. Add all excursions for the day
       7. DO statistics
          1. Intersect date range with DO range
          2. Compare # of DO points to # of rows for a complete day
             1. If points are missing, then skip DO calcs (set equal to blank)
             2. If all points are present, then continue with calculations
          3. Calculate DAILY MIN, DAILY MAX, DAILY MEAN
          4. Calculate DAILY AMOUNT OF TIME < WQS (WQS = 3, 4, 5, and 6 mg/L)
             1. Interpolate the starting date/time of DO below WQS
             2. Interpolate the ending date/time of DO below WQS
             3. Calculate the difference between the starting and ending time of each excursion below WQS
             4. Add all excursions for the day
       8. Write statistics to Daily Statistics sheet
   11. Apply formatting and headings to Summary Statistics sheet in OWMID file
   12. Calculate Summary Statistics: basic statistics
       1. Temperature
          1. Define temperature range for file (based on graphs-stats sheet)
          2. Count non-blank cells = OBSERVED COUNT
             1. Skip calculations if all cells are blank
          3. Calculate AVG, STDEV, MIN, MAX, MEDIAN, IQR from T range
          4. Count blanks and calculate OBSERVED DEPLOYMENT TIME from T range
       2. DO
          1. Define DO range for file (based on graphs-stats sheet)
          2. Count non-blank cells = OBSERVED COUNT
             1. Skip calculations if all cells are blank
          3. Calculate AVG, STDEV, MIN, MAX, MEDIAN, IQR from DO range
          4. Count blanks and calculate OBSERVED DEPLOYMENT TIME from DO range
       3. DOsat
          1. Define DOsat range for file (based on graphs-stats sheet)
          2. Count non-blank cells = OBSERVED COUNT
             1. Skip calculations if all cells are blank
          3. Calculate AVG, STDEV, MIN, MAX, MEDIAN, IQR from DOsat range
          4. Count blanks and calculate OBSERVED DEPLOYMENT TIME from DOsat range
   13. Calculate Summary Statistics: Amount of Time and Max Duration >< WQS
       1. Temperature where WQS = 29.4, 28.3, and 20 deg C
          1. Skip calculations if all cells in T range are blank
          2. For each data point, determine if point is blank (B), above WQS (Y), or below WQS (N) and store flags in array
          3. Determine where the T curve exceeds the WQS
             1. interpolate the date/time where each excursion begins and ends
             2. blank cells automatically set the start or end time of an excursion
             3. calculate the amount of time between the start and end of the excursion
             4. store each excursion in an array
             5. determine the length of each continuous excursion (uninterrupted by blank cells) and store in an array
             6. calculate the AMOUNT OF TIME > WQS by adding all of the excursion times
             7. calculate the MAX DURATION > WQS by finding the maximum of the continuous excursion times
       2. DO where WQS = 3, 4, 5, and 6 mg/L
          1. Skip calculations if all cells in DO range are blank
          2. For each data point, determine if point is blank (B), above WQS (Y), or below WQS (N) and store flags in array
          3. Determine where the DO curve is below the WQS
             1. interpolate the date/time where each excursion begins and ends
             2. blank cells automatically set the start or end time of an excursion
             3. calculate the amount of time between the start and end of the excursion
             4. store each excursion in an array
             5. determine the length of each continuous excursion (uninterrupted by blank cells) and store in an array
             6. calculate the AMOUNT OF TIME < WQS by adding all of the excursion times
             7. calculate the MAX DURATION < WQS by finding the maximum of the continuous excursion times
   14. Calculate Summary Statistics based on Daily Statistics
       1. Temperature
          1. Skip calculations if all daily values are blank
          2. From Daily Statistics sheet, define ranges of DAILY MIN, DAILY MAX, DAILY MEAN, DAILY AMOUNT OF TIME > WQS (where WQS = 29.4, 28.3, and 20 deg C)
          3. Calculate the MEAN OF THE DAILY MIN, MEAN OF DAILY MAX, MEAN OF DAILY MEAN, AVG DAILY AMOUNT OF TIME > WQS (where WQS = 29.4, 28.3, and 20 deg C)
       2. DO
          1. Skip calculations if all daily values are blank
          2. From Daily Statistics sheet, define ranges of DAILY MIN, DAILY MAX, DAILY MEAN, DAILY AMOUNT OF TIME < WQS (where WQS = 3, 4, 5, and 6 mg/L)
          3. Calculate the MEAN OF THE DAILY MIN, MEAN OF DAILY MAX, MEAN OF DAILY MEAN, AVG DAILY AMOUNT OF TIME < WQS (where WQS = 3, 4, 5, and 6 mg/L)
   15. Calculate 7-day statistics (Temperature only)
       1. Check that there are at least 7-days in the file before proceeding
       2. Loop through each of the DAILY MAX T values on the Daily Statistics sheet
          1. Check that current day is at least 4th day from end or beginning of date range (rolling average defined here as average of 7 days centered on the current day)
          2. Check that there are no blank cells in the current 7-day interval
          3. Calculate the 7-DADM for current day
          4. Write values to Daily Statistics sheet
       3. Loop through each of the DAILY MEAN T values on the Daily Statistics sheet
          1. Check that current day is at least 4th day from end or beginning of date range (rolling average defined here as average of 7 days centered on the current day)
          2. Check that there are no blank cells in the current 7-day interval
          3. Calculate the 7-DADA for current day
          4. Write values to Daily Statistics sheet
   16. Calculate 7-day summary statistics (Temperature only)
       1. Check that there are at least 7-days in the file before proceeding
       2. Define the range of 7-DADA values on the Daily Statistics sheet
       3. Calculate the MWAT as the maximum of all 7-DADA values
   17. Add NO DATA symbol (“--“) to blank cells on Daily Statistics sheet
   18. Write statistics to QC summary workbook
   19. Write statistics to Summary Statistics sheet in OWMID workbook
8. Write program end time to File Process Info sheet

## Generate QC2 FINAL and QC3 OWMID Files

**General Purpose:** To produce final QC2 unattended data files for QC3 principal investigator review with all symbols, reporting rules, and other formatting applied; and to produce QC3 files that can be edited if necessary as a result of QC3 review comments in preparation of finalizing the files to QC4 status

**Filename:** UnAtt\_QC3\_V6.wsf

**Code Language:** VBScript

**Input Files:**

* QC2 OWMID files (XX-XXXX\_QC2\_MM-DD-YYYY.xlsm, where XX-XXXX is the OWMID, MM is the month, DD is the day, and YYYY is the year that the file was generated)

**Output Files:**

* QC2 FINAL OWMID files (XX-XXXX\_QC2FINAL\_MM-DD-YYYY.xlsm, where XX-XXXX is the OWMID, MM is the month, DD is the day, and YYYY is the year that the file was generated); files are output to project folders (e.g. Projectname\_yyyy\_UQC2\_FINAL\_MM-DD-YYYY, where MM is the month, DD is the day, and YYYY is the year that the file was generated, and yyyy is the data year)
* QC2 FINAL Summary workbooks for each project (Projectname\_yyyy\_unattended\_QC2FINAL\_MM-DD-YYYY.xlsx, where MM is the month, DD is the day, and YYYY is the year that the file was generated, and yyyy is the data year); files are output to project folders; files contain a sheet for statistics and a sheet for qualifier summary counts for all OWMIDs in the project for the data year
* QC3 OWMID files (XX-XXXX\_QC3\_MM-DD-YYYY.xlsm, where XX-XXXX is the OWMID, MM is the month, DD is the day, and YYYY is the year that the file was generated); files are output to a subfolder of the QC3 folder for the data year (e.g. W:\DWM\Data\unattended-QA\yyyy\QC3\yyyyUQC3\_MM-DD-YYYY\, where yyyy is the data year; and MM is the month, DD is the day, and YYYY is the year that the file was generated)

**Approximate Run Time:** 15 – 60 minutes

**Code Outline:**

1. announce to user what code does and give user opportunity to back out
2. user enters data year for analysis
3. check for presence of directories/files
   1. user navigates to QC2 OWMID file directory for data year entered above
   2. user navigates to location of QC2 unattended “read me” file and selects it
      1. open workbook and check for presence of “read me” worksheet
      2. if worksheet not found, present error to user and quit
      3. find rows for “META DATA NETWORK LOCATION”, “NETWORK FILE LOCATION”, and “SPECIFIC METHODOLOGY” (for later population of hyperlinks to files/folders)
   3. get path to “MetaDatayyyy.xlsx” file, where yyyy is the data year (hard-coded unless folder not found, then user navigates to the file and selects it) for populating “META DATA NETWORK LOCATION” on read me sheet
   4. get path to network location of QC4 files (hard-coded unless folder not found, then user navigates to correct location) for populating “NETWORK FILE LOCATION” on read me sheet
   5. get path to process SOPs directory (hard-coded unless folder not found, then user navigates to correct location)for populating “SPECIFIC METHODOLOGY” on read me sheet
   6. check for files in QC2 directory
      1. if no files present, present user with error message and quit
      2. if files present, present user with # of files found
   7. extract year, month, day from QC2 folder name in III a.
      1. check for presence of QC summary workbook (with same date in name as QC2 file directory)
      2. present user with error message if QC summary workbook not found, quit
   8. create folder for QC2 final files (yyyyUQC2\_FINAL\_MM-DD-YYYY, where MM is the month, DD is the day, and YYYY is the year the folder is created; yyyy is the data year; stored as a subfolder of the QC2 OWMID file directory selected in III a.)
   9. create AllProjects folder for storing AllProjects QC Summary file (AllProjects\_yyyy\_UQC2\_FINAL\_MM-DD-YYYY, where where MM is the month, DD is the day, and YYYY is the year the folder is created, and yyyy is the data year; stored as a subfolder of the QC2 final folder in III h.)
   10. define name of AllProjects QC2 Summary workbook
   11. create folder for storing QC3 review files (yyyyUQC3\_MM-DD-YYYY, where MM is the month, DD is the day, and YYYY is the year the folder is created, and yyyy is the data year; stored as a subfolder of the QC3 directory)
   12. define name of QC3 Summary workbook
   13. create text file for recording issues (YYYY\_Unattended Create QC3 Files.txt) in QC3 folder created in III k.
4. create QC3 files
   1. skip text (status) files
   2. check for presence of attended data file and meta data files and copy those
   3. define QC3 file names, copy and save QC2 files with QC3 names in the QC3 subfolder
   4. open QC2 Summary workbook and save as QC3 summary workbook in the QC3 subfolder
   5. open QC3 Summary workbook
      1. add QC3 review columns to QC Decision Summary sheet
      2. rename QC Decision sheet as “QC2 QC3 Decision Summary”
      3. rename QC2 file column to “QC3 File” on Statistics sheet
5. create Allprojects QC2 Final summary workbook from QC2 (“working”) summary workbook
   1. check for presence of required worksheets (Statistics)
      1. present user with error message if sheet doesn’t exist and quit
   2. set Statistics worksheet object and get used range
   3. define columns on Statistics sheet
   4. check that statistics have been run by comparing # of values in a statistics column against the # of QC3 files listed on the sheet
      1. present user with error message, if no statistics found or if # of statistics doesn’t match # of files, and quit
6. create QC2 Final project folders
   1. sort Statistics sheet in QC Summary workbook by Project, Watershed, and File Start Date
   2. correct cases where Connecticut misspelled (2008 only)
   3. get unique project names from Project column and store in array
      1. correct names for unwanted or illegal characters that can’t be used in folder names
   4. create project folders in QC2 directory in III h.
7. Apply reporting rules to FlowStat and STAID columns in AllProjects QC4 Summary workbook
   1. Apply symbols for NO WATER (“^^”) and MISSING (“\*\*”) Flowstat values
   2. Apply NO DATA (“--”) to all 2005 FlowStat cells (wasn’t a field then)
   3. Apply NO DATA (“--”) to blank STAID cells
8. Create new sheet in Allprojects QC2 Summary workbook for qualifier counts
   1. New sheet name = “OWMID Qualifier Summary”
   2. Define columns on qualifier sheet
   3. Copy select columns from Statistics sheet to Qualifer sheet (Project-related, station-related, waterbody-related, OWMID-related, and Total Count of rows per file)
   4. Add column headers to qualifier sheet
9. Create QC2 final OMWID files from QC2 working OWMID files
   1. Define columns in OWMID files for data sheets, Daily Statistics sheet, and Summary Statistics sheet
   2. Loop through OWMIDs listed on Statistics sheet in Allprojects QC2 summary file
      1. Get QC2 working file name, OWMID, Gear type, Flowstat, STAID, project
      2. Define QC3 file from QC2 file name
      3. Overwrite QC2 file name in QC3 Summary workbook , Statistics sheet, with QC3 file name
      4. Match current project to project folder using projects stored in array
      5. Define QC2 final OWMID file name and path (based on project folder name)
         1. If no project folder match found, place file in Allprojects folder
      6. Check for Post-Processed file = “NO FILE”
      7. Check for Duplicate OWMIDs to skip (DupesNot2Use column = “X”)
      8. Open QC2 working OWMID file if not identified as a duplicate to skip
         1. Save QC2 working file as QC2 final file
            1. NO FILE workbooks are saved as QC2 FINAL without further processing
            2. All other files saved as QC2 temp files (.xlsm) (otherwise Excel presents dialog box for saving files with VBA code as .xlsx that interrupts the code)
            3. Write QC2 final file name to Allprojects QC2 Summary file
         2. Delete code modules from QC2 temp file
         3. Delete unneeded worksheets from QC2 temp file (main data sheet, init, backup, lists, QC Summary, Field QC, Lab QC)
         4. Save QC2 temp file as QC2 final file, close temp file and delete it
         5. Open QC2 final file
         6. Define worksheet objects and used ranges
         7. Add read me sheet
            1. Add hyperlinks to files/paths on read me sheet (see III c-e.)
            2. Add date of finalization
         8. Apply reporting rules to worksheets (final data, graphs-stats, Daily Statistics, and Summary Statistics)
            1. Significant figures
            2. Remove extra “NO DATA” or “MISSING” text from analyte qualifier columns on final data sheet and graphs-stats data sheet
            3. NOTE: all symbols (censored (##), missing (\*\*), and NO DATA (--) have been previously applied to the final data sheet during QC2 review
         9. Add qualifier counts to AllProjects QC2 summary workbook
            1. Filter and count each analyte data column on final sheet for

Missing (“\*\*”)

No data (“--“)

Censored (“##”)

* + - * 1. Filter and count each analyte qualifier column for

Blanks (# accepted is calculated as # blanks - # missing - # no data)

Non-blanks (# qualified only = # tot qualified - # censored)

Specific qualifier symbols (i, m, r, t, u)

* + - * 1. Write counts to Allprojects QC2 summary file, Qualifer sheet
        2. NOTES:

# accepted + # qualified + # censored = observed count; observed count + # missing + # no data = total count

Sum of qualifier symbol counts = # qualified + # censored

* + - 1. Adjust graphs (marker and line colors, gridlines, marker size)
      2. Apply formatting to data sheets (final, graphs-stats, Daily Statistics, and Summary Statistics)
         1. Rename columns
         2. Set font type, size, weight, color
         3. Align cells and set fill colors
         4. Adjust column widths and apply text wrapping
         5. Apply cell borders
         6. Correct cases where Connecticut misspelled (2008 only) on Summary Statistics sheet
         7. Autofilter, freeze top row and go to cell “A1” (final, graph-stats, Daily Statistics sheet only)
      3. Delete extra analyte columns from data sheets
      4. Protect worksheets
      5. Select Read me sheet and go to cell “A1”
      6. Set file attribute to read only
      7. Save and close QC2 final file

1. Format AllProjects QC2 Summary workbook
   1. Apply reporting rules and formatting to Statistics sheet
      1. Rename columns
      2. Set font type, size, weight, color
      3. Align cells and set fill colors
      4. Adjust column widths and apply text wrapping
      5. Apply cell borders
      6. Apply significant figures
      7. Add “--“ to statistics for NO FILE rows
   2. Apply reporting rules and formatting to Qualifier sheet
      1. Rename columns
      2. Set font type, size, weight, color
      3. Align cells and set fill colors
      4. Adjust column widths and apply text wrapping
      5. Apply cell borders
      6. Apply autofilter
      7. Apply significant figures
      8. Add “--“ to statistics for NO FILE rows
   3. Add read me sheet
      1. Add hyperlinks to files/paths on read me sheet (see III c-e.)
      2. Add date of finalization
2. Split AllProjects QC2 summary workbook into project-specific QC2 summary files
   1. Define project file name and path (from project folder names stored in array)
   2. Create new workbook
   3. Copy read me sheet to project workbook
      1. Add hyperlinks to files/paths on read me sheet (see III c-e.)
      2. Add date of finalization
   4. Add Statistics sheet
      1. Filter AllProjects QC2 summary file Statistics sheet for project name
      2. Copy rows from Allprojects QC2 summary file to project workbook
      3. Delete duplicate OWMIDs from project workbook (by filtering for DupesNot2Use = “X”)
      4. Delete unneeded columns (DupesNot2Use, Duplicate OWMID, Duplicate, PPFile, SondeID)
      5. Freeze top row and go to cell “A1”
      6. Redefine used range after deleting columns and autofilter
      7. Protect sheet
   5. Add Qualifier Summary sheet
      1. Filter AllProjects QC2 summary file Qualifier sheet for project name
      2. Copy rows from Allprojects QC2 summary file to project workbook
      3. Delete duplicate OWMIDs from project workbook (by filtering for DupesNot2Use = “X”)
      4. Delete unneeded columns (DupesNot2Use, Duplicate OWMID, Duplicate, PPFile, SondeID)
      5. Freeze top row and go to cell “A1”
      6. Redefine used range after deleting columns and autofilter
      7. Protect sheet
   6. Delete extra sheets from project workbook
   7. Select read me sheet and go to cell “A1”
   8. Save and close project workbook
   9. Set file attribute to read-only
3. Apply final changes to Allprojects QC2 summary workbook
   1. Statistics sheet
      1. Delete duplicate OWMIDs from project workbook (by filtering for DupesNot2Use = “X”)
      2. Delete unneeded columns (DupesNot2Use, Duplicate OWMID, Duplicate, PPFile, SondeID)
      3. Freeze top row and go to cell “A1”
      4. Redefine used range after deleting columns and autofilter
      5. Protect sheet
   2. Qualifier Summary sheet
      1. Delete duplicate OWMIDs from project workbook (by filtering for DupesNot2Use = “X”)
      2. Delete unneeded columns (DupesNot2Use, Duplicate OWMID, Duplicate, PPFile, SondeID, Station/Waterbody columns)
      3. Freeze top row and go to cell “A1”
      4. Redefine used range after deleting columns and autofilter
      5. Protect sheet
   3. Select read me sheet and go to cell “A1”
   4. Save and close Allprojects QC2 workbook
   5. Set file attribute to read-only
4. Reopen QC2 FINAL summary and QC3 Summary workbooks
   1. Add start and end time for code run to File Process Info sheet
   2. Save and close

## Create Station Files

**General Purpose:** To create station files (identified by UniqueID) from unattended multiprobe data files for individual OWMIDs

**Filename:** UnAtt\_CreateStationFiles\_V7.wsf

**Code Language:** VBScript

**Input Files:**

* QC2 or QC3 OWMID files that have been edited for final decisions (the QC review or working files) and where the OWMID statistics have been calculated
* QC2 or QC3 Summary workbook

**Output Files:**

* Station files named with the UniqueID, the QC level, and the creation date (e.g. W0168\_QC3\_MM-DD-YYYY.xlsx, where W0168 is the 5 character UniqueID, starting with W and followed by 4 numbers; QC2 or QC3 is the QC level; and MM is the month, DD is the day, and YYYY is the year program was run)
* Station files are stored in a subfolder of the parent directory (e.g. W:\DWM\Data\unattended\_QA\2007\QC3\2007UQC3\_5-10-2012\2007UQC3\_Stations\_7-5-2012, where 2007UQC3\_3-10-2012 is the parent folder containing the OWMID files that are used to build the station files, and 2007UQC3\_Stations\_7-5-2012 is the folder where the station files are created)

**Approximate Run Time: 15 – 60 minutes, depending on the number and size of the OWMID files**

**Code Outline:**

1. Announce to user what is about to happen and allow them to back out
2. Get year for analysis from user
3. Get QC level to process from user (entry must be either QC2 or QC3)
4. Check for presence of directories and files
5. Check for OWMID files to process in the directory related to the QC level indicated in step 3
6. Open the QC Review summary workbook and check for presence of specific worksheets needed to run the code (note: the QC review workbook stored in the OWMID directory retains information about the station files—a new workbook is not created)
   1. Create sheet called “Station OWMIDs” if it doesn’t exist
   2. Create sheet called “Station Statistics” if it doesn’t exist
   3. Required sheets include: Analytes, Flags, QC2 DecisionSummary/QC2 QC3 Decision Summary, and Statistics
7. Prepare QC Review summary workbook for station file creation:
   1. define worksheet objects, worksheet columns, remove autofilters, and get used data ranges on each sheet
   2. clear Station Statistics and Station OWMIDs sheets if they were already present
   3. add headers to Station Statistics sheet
   4. copy info from other sheets to the Station OWMIDs sheet
      1. from Analytes sheet copy OWMID, FSLOG, UniqueID, Gear Type, SondeID, Project Name, Sample Type, Duplicate OWMID, File Time Interval, Post processed and QC-level file names
      2. from QC Decision Summary sheet copy DupesNot2Use column
      3. from Flags sheet copy StartDate
      4. copy headers from existing sheets
      5. define used data range after copying
   5. sort Station OWMIDs sheet by UniqueID, then by StartDate
8. Create station files
   1. Define columns on data sheets in OWMID files
   2. Get current UniqueID from Station OWMIDs sheet, filter sheet by this value, get first and last row for this UniqueID, then loop through each OWMID
      1. Get current OWMID, current Gear, DupesNot2Use value, and Post-Processed file name
         1. If PPfile = “NO FILE” then skip the OWMID (no data to add to station file)
         2. If Gear = “Temperature Logger” then skip the OWMID (T logger files are not included in station files)
         3. If DupesNot2Use = “X”, then skip the OWMID (only one duplicate is included in final dataset; Duplicates not used are pre-identified as part of QC Review process)
      2. Open first OWMID file
         1. Define station file name as Unique ID followed by current QC level followed by month, day, year of program run (e.g. W2128\_QC3\_7-6-2012.xlsx; note: all station files are .xlsx—no macros are included because these files are not intended to be edited)
         2. Get time interval of file from QC Summary workbook and store in an array
         3. Save first OWMID file as temp file in station file directory and close file
         4. Open temp file
         5. Delete code modules from temp file
         6. Delete unneeded sheets (“QC Summary”, “init”, “backup”, “Field QC”, “Lab QC”)
         7. Rename data sheets using UniqueID (includes main data sheet, and final and graphs-stats sheets)
         8. Store last row of file in an array
         9. Delete named range for lists (“QC Result”) to eliminate Excel dialog box from popping up during code run
         10. Save temp file as official station name, delete temp file
      3. Open remaining OWMID files (if applicable)
         1. Get time interval of new file and store in array
         2. Re-open station workbook
            1. Get used data ranges of all worksheets (except Summary Statistics)
         3. Open new OWMID workbook
            1. Get used data ranges of all worksheets (except Summary Statistics)
            2. Delete named range (“QC Result”) from new OWMID workbook
         4. Copy data from OWMID workbook to station workbook (main data sheet, final sheet, graphs-stats sheet)
         5. Store last row of data in an array
         6. Save and close station file
      4. After looping through all OWMIDs in filter, and if station file comprises more than one OWMID, then:
         1. Insert blank rows between OWMID data blocks to create a continuous time series on each data sheet (main data sheet, final sheet, and graphs-stats sheet)
            1. fill in OWMID and SondeID columns with values from OWMID data block preceding the new rows
            2. fill in date, time, and date/time columns with new values calculated from the first data row of the OWMID data block following the new rows and working back to the last row of the OWMID data block preceding the new rows (using time interval of preceding data block)
            3. add “NO DATA” to analyte qualifier columns of new rows
            4. for graphs-stats sheet

insert values of standards for graphing in new rows

* + - * 1. for final sheet

insert NO DATA symbol “--“ in analyte data columns of new rows

* + - 1. insert blank rows between OWMID data blocks on the Daily Statistics sheet
         1. fill in OWMID and SondeID columns with values from OWMID data block preceding the new rows
         2. fill in date column with new values calculated from the first data row of the OWMID data block following the new rows and working back to the last row of the OWMID data block preceding the new rows (whole days only; no time interval needed)
         3. add “NO DATA” symbol (“--“) to all statistics columns for new rows
      2. change graphs to accommodate new rows
         1. determine minimum and maximum x-values, major unit to use based on total # of days in file
         2. for each chart, delete existing data series and re-add using new ranges for x and y data
      3. add named range for lists back as “QC Result”
      4. save and close station file
    1. if a station file exists for the current UniqueID, then
       1. re-open station file
       2. clear Summary Statistics sheet in station file
       3. rename charts with UniqueID
       4. write info to Station Statistics sheet in summary workbook
          1. determine if station file contains mixed time intervals and write result
          2. write string containing list of OWMIDs used to build station file
          3. copy applicable project, watershed, station columns from Statistics sheet in QC Summary workbook
       5. save and close station file

1. write info about code run to File Process Info sheet in Summary workbook
2. save and close summary workbook

## Calculate Station File Statistics

**General Purpose:** To calculate statistics for multiprobe (short-term) data for multiple OWMIDs (where data are located in station files, associated by UniqueID, for a given sampling year only); statistics are added to the appropriate worksheets within the station files and are also added to the QC summary workbook; see Appendix E for description of statistics

**Filename: UnAtt\_CalcStationStats.wsf**

**Code Language:** VBScript

**Input Files:**

* Station files named with the UniqueID, the QC level, and the creation date (e.g. W0168\_QC3\_MM-DD-YYYY.xlsx, where W0168 is the 5 character UniqueID, starting with W and followed by 4 numbers; QC2 or QC3 is the QC level; and MM is the month, DD is the day, and YYYY is the year program was run)
* Requires that statistics have been run on OWMID files for same QC-level (Daily Statistics sheet created during Station File creation)

**Output Files:**

* Same as input files, with statistics added to Daily Statistics and Summary Statistics sheets in staton file, and to the QC Summary workbook for the same QC-level

**Approximate Run Time:** 5-15 minutes

**Code Outline:**

1. announce to user what code does and give user opportunity to back out
2. user enters data year for analysis
3. user enters QC level to process (QC2 or QC3)
4. check for presence of directories/files
   1. user navigates to station data file directory for QC level and data year entered above
      1. check for station files in directory
      2. present user with error message if no files found
   2. get parent directory of folder selected in IV a. (where the QC Summary file is stored)
   3. extract year, month, day from parent directory name in IV b.
      1. check for presence of QC summary workbook (with same date in name as parent directory)
      2. present user with error message if QC summary workbook not found, quit
   4. create text file for recording issues (YYYY\_Unattended Calculate Station Statistics.txt)
5. open QC summary workbook
   1. check for presence of required worksheets (Statistics, Analytes Present, Station Statistics, File Process Info)
      1. if worksheets don’t exist, then present error message to user and quit
   2. define columns on Statistics sheet
   3. define columns on Station Statistics sheet and add column headings
6. Loop through UniqueIDs on Station Statistics sheet and calculate statistics
   1. Define columns on data sheet in OWMID file, on Daily Statistics sheet, and on Summary Statistics sheet
   2. Get info about UniqueID (file name, Project)
   3. Open station file
      1. Define worksheet objects
      2. Check for presence of Daily Statistics sheet
         1. If not present, then present user with error message
         2. If present, get used range
      3. Check for presence of Summary Statistics sheet
         1. If not present , then present user with error message
         2. If present, clear sheet
   4. Copy station/waterbody info from QC Summary workbook to Summary Statistics sheet
   5. Get STATION FILE START TIME and STATION FILE END TIME, TOTAL STATION FILE COUNT, and TOTAL STATION FILE DURATION from Station File graphs-stats sheet
   6. Apply formatting and headings to Summary Statistics sheet in Station file
   7. Calculate Summary Statistics: basic statistics
      1. Temperature
         1. Define temperature range for file (based on graphs-stats sheet)
         2. Count non-blank cells = OBSERVED COUNT
            1. Skip calculations if all cells are blank
         3. Calculate AVG, STDEV, MIN, MAX, MEDIAN, IQR from T range
         4. Count blanks
      2. DO
         1. Define DO range for file (based on graphs-stats sheet)
         2. Count non-blank cells = OBSERVED COUNT
            1. Skip calculations if all cells are blank
         3. Calculate AVG, STDEV, MIN, MAX, MEDIAN, IQR from DO range
         4. Count blanks
      3. DOsat
         1. Define DOsat range for file (based on graphs-stats sheet)
         2. Count non-blank cells = OBSERVED COUNT
            1. Skip calculations if all cells are blank
         3. Calculate AVG, STDEV, MIN, MAX, MEDIAN, IQR from DOsat range
         4. Count blanks
   8. Calculate Summary Statistics from OWMID files: Amount of Time and Max Duration >< WQS
      1. Get the OWMID string used to build station file, parse into individual OWMIDs, store OWMIDs in an array
      2. Temperature
         1. Skip calculations if all T cells are blank
         2. Loop through OWMIDs in array
            1. Find current OWMID in QC Summary workbook, Statistics sheet
            2. Get values for AMOUNT OF TIME > WQS (where WQS = 29.4, 28.3, and 20 deg C) and store in separate arrays
            3. Get values for MAX DURATION > WQS (where WQS = 29.4, 28.3, and 20 deg C) and store in separate arrays
            4. Replace any “--“ NO DATA values with zero in the arrays
            5. Calculate AMOUNT OF TIME > WQS (where WQS = 29.4, 28.3, and 20 deg C) for the station file, by adding array values in VI.h.ii.b.
            6. Calculate MAX DURATION > WQS (where WQS = 29.4, 28.3, and 20 deg C) for the station file, by getting the max values from the arrays in VI.h.ii.c.
      3. DO
         1. Skip calculations if all DO cells are blank
         2. Loop through OWMIDs in array
            1. Find current OWMID in QC Summary workbook, Statistics sheet
            2. Get values for AMOUNT OF TIME < WQS (where WQS = 3, 4, 5, and 6 mg/L) and store in separate arrays
            3. Get values for MAX DURATION < WQS (where WQS = 3, 4, 5, and 6 mg/L) and store in separate arrays
            4. Replace any “--“ NO DATA values with zero in the arrays
            5. Calculate AMOUNT OF TIME < WQS (where WQS = 3, 4, 5, and 6 mg/L) for the station file, by adding array values in VI.h.iii.b.
            6. Calculate MAX DURATION < WQS (where WQS = 3, 4, 5, and 6 mg/L) for the station file, by getting the max values from the arrays in VI.h.iii.c.
   9. Calculate Summary Statistics based on Daily Statistics
      1. Erase “--“ NO DATA symbols on Daily Statistics sheet prior to computing average daily statistics
      2. Temperature
         1. Skip calculations if all daily values are blank
         2. From Daily Statistics sheet, define ranges of DAILY MIN, DAILY MAX, DAILY MEAN, DAILY AMOUNT OF TIME > WQS (where WQS = 29.4, 28.3, and 20 deg C)
         3. Calculate the MEAN OF THE DAILY MIN, MEAN OF DAILY MAX, MEAN OF DAILY MEAN, AVG DAILY AMOUNT OF TIME > WQS (where WQS = 29.4, 28.3, and 20 deg C)
      3. DO
         1. Skip calculations if all daily values are blank
         2. From Daily Statistics sheet, define ranges of DAILY MIN, DAILY MAX, DAILY MEAN, DAILY AMOUNT OF TIME < WQS (where WQS = 3, 4, 5, and 6 mg/L)
         3. Calculate the MEAN OF THE DAILY MIN, MEAN OF DAILY MAX, MEAN OF DAILY MEAN, AVG DAILY AMOUNT OF TIME < WQS (where WQS = 3, 4, 5, and 6 mg/L)
   10. Calculate 7-day statistics (Temperature only)
       1. Check that there are at least 7-days in the file before proceeding
       2. Loop through each of the DAILY MAX T values on the Daily Statistics sheet
          1. Check that current day is at least 4th day from end or beginning of date range (rolling average defined here as average of 7 days centered on the current day)
          2. Check that there are no blank cells in the current 7-day interval
          3. Calculate the 7-DADM for current day
          4. Write values to Daily Statistics sheet
       3. Loop through each of the DAILY MEAN T values on the Daily Statistics sheet
          1. Check that current day is at least 4th day from end or beginning of date range (rolling average defined here as average of 7 days centered on the current day)
          2. Check that there are no blank cells in the current 7-day interval
          3. Calculate the 7-DADA for current day
          4. Write values to Daily Statistics sheet
   11. Calculate 7-day summary statistics (Temperature only)
       1. Check that there are at least 7-days in the file before proceeding
       2. Define the range of 7-DADA values on the Daily Statistics sheet
       3. Calculate the MWAT as the maximum of all 7-DADA values
   12. Add NO DATA symbol (“--“) to blank cells on Daily Statistics sheet
   13. Write statistics to QC summary workbook
   14. Write statistics to Summary Statistics sheet in station workbook
7. write program start and end time on File Process Info sheet

## Generate QC4 OWMID Files

**General Purpose:** To produce final QC4 unattended data files with all symbols, reporting rules, and other formatting applied, as well as QC Summary project workbooks containing statistics and qualifier summary counts for all OWMIDs in a project

**Filename:** UnAtt\_QC4\_V3.wsf

**Code Language:** VBScript

**Input Files:**

* QC3 OWMID files (XX-XXXX\_QC3\_MM-DD-YYYY.xlsm, where XX-XXXX is the OWMID, MM is the month, DD is the day, and YYYY is the year that the file was generated)

**Output Files:**

* QC4 OWMID files (XX-XXXX\_QC4FINAL\_MM-DD-YYYY.xlsm, where XX-XXXX is the OWMID, MM is the month, DD is the day, and YYYY is the year that the file was generated); files are output to project folders (e.g. Projectname\_yyyy\_UQC4\_FINAL\_MM-DD-YYYY, where MM is the month, DD is the day, and YYYY is the year that the file was generated, and yyyy is the data year)
* QC4 Summary workbooks for each project (Projectname\_yyyy\_unattended\_QC4FINAL\_MM-DD-YYYY.xlsx, where MM is the month, DD is the day, and YYYY is the year that the file was generated, and yyyy is the data year); files are output to project folders; files contain a sheet for statistics and a sheet for qualifier summary counts for all OWMIDs in the project for the data year

**Approximate Run Time:** 1-2 hours

**Code Outline:**

1. announce to user what code does and give user opportunity to back out
2. user enters data year for analysis
3. check for presence of directories/files
   1. user navigates to QC3 OWMID file directory for data year entered above
   2. user navigates to location of QC4 unattended “read me” file and selects it
      1. open workbook and check for presence of “read me” worksheet
      2. if worksheet not found, present error to user and quit
      3. find rows for “META DATA NETWORK LOCATION”, “NETWORK FILE LOCATION”, and “SPECIFIC METHODOLOGY” (for later population of hyperlinks to files/folders)
   3. get path to “MetaDatayyyy.xlsx” file, where yyyy is the data year (hard-coded unless folder not found, then user navigates to the file and selects it) for populating “META DATA NETWORK LOCATION” on read me sheet
   4. get path to network location of QC4 files (hard-coded unless folder not found, then user navigates to correct location) for populating “NETWORK FILE LOCATION” on read me sheet
   5. get path to process SOPs directory (hard-coded unless folder not found, then user navigates to correct location)for populating “SPECIFIC METHODOLOGY” on read me sheet
   6. check for files in QC3 directory
      1. if no files present, present user with error message and quit
      2. if files present, present user with # of files found
   7. extract year, month, day from QC3 folder name in III a.
      1. check for presence of QC summary workbook (with same date in name as QC3 file directory)
      2. present user with error message if QC summary workbook not found, quit
   8. create folder for QC4 final files (yyyyUQC4\_FINAL\_MM-DD-YYYY, where MM is the month, DD is the day, and YYYY is the year the folder is created; yyyy is the data year; stored as a subfolder of the QC3 OWMID file directory selected in III a.)
   9. create AllProjects folder for storing AllProjects QC Summary file (AllProjects\_yyyy\_UQC4\_FINAL\_MM-DD-YYYY, where where MM is the month, DD is the day, and YYYY is the year the folder is created, and yyyy is the data year; stored as a subfolder of the QC4 final folder in III h.)
   10. define name of AllProjects QC4 Summary workbook
   11. create text file for recording issues (YYYY\_Unattended Create QC4 Files.txt) in QC4 final folder in III h.
4. create Allprojects QC4 summary workbook from QC3 summary workbook
   1. check for presence of required worksheets (Statistics)
      1. present user with error message if sheet doesn’t exist and quit
   2. set Statistics worksheet object and get used range
   3. define columns on Statistics sheet
   4. check that statistics have been run by comparing # of values in a statistics column against the # of QC3 files listed on the sheet
      1. present user with error message, if no statistics found or if # of statistics doesn’t match # of files, and quit
5. create QC4 project folders
   1. sort Statistics sheet in QC Summary workbook by Project, Watershed, and File Start Date
   2. correct cases where Connecticut misspelled (2008 only)
   3. get unique project names from Project column and store in array
      1. correct names for unwanted or illegal characters that can’t be used in folder names
   4. create project folders in QC4 directory in III h.
6. Apply reporting rules to FlowStat and STAID columns in AllProjects QC4 Summary workbook
   1. Apply symbols for NO WATER (“^^”) and MISSING (“\*\*”) Flowstat values
   2. Apply NO DATA (“--”) to all 2005 FlowStat cells (wasn’t a field then)
   3. Apply NO DATA (“--”) to blank STAID cells
7. Create new sheet in Allprojects QC4 Summary workbook for qualifier counts
   1. New sheet name = “OWMID Qualifier Summary”
   2. Define columns on qualifier sheet
   3. Copy select columns from Statistics sheet to Qualifer sheet (Project-related, station-related, waterbody-related, OWMID-related, and Total Count of rows per file)
   4. Add column headers to qualifier sheet
8. Create QC4 final OMWID files from QC3 OWMID files
   1. Define columns in OWMID files for data sheets, Daily Statistics sheet, and Summary Statistics sheet
   2. Loop through OWMIDs listed on Statistics sheet in Allprojects QC4 summary file
      1. Get QC3 file name, OWMID, Gear type, Flowstat, STAID, project
      2. Match current project to project folder using projects stored in array
      3. Define QC4 OWMID file name and path (based on project folder name)
         1. If no project folder match found, place file in Allprojects folder
      4. Check for Post-Processed file = “NO FILE”
      5. Check for Duplicate OWMIDs to skip (DupesNot2Use column = “X”)
      6. Open QC3 OWMID file if not identified as a duplicate to skip
         1. Save QC3 file as QC4 file
            1. NO FILE workbooks are saved as QC4 FINAL without further processing
            2. All other files saved as QC4 temp files (.xlsm) (otherwise Excel present dialog box for saving files with VBA code as .xlsx)
            3. Write QC4 name to Allprojects QC4 Summary file
         2. Delete code modules from QC4 temp file
         3. Delete unneeded worksheets from QC4 temp file (main data sheet, init, backup, lists, QC Summary, Field QC, Lab QC)
         4. Save QC4 temp file as QC4 final file, close temp file and delete it
         5. Open QC4 final file
         6. Define worksheet objects and used ranges
         7. Add read me sheet
            1. Add hyperlinks to files/paths on read me sheet (see III c-e.)
            2. Add date of finalization
         8. Apply reporting rules to worksheets (final data, graphs-stats, Daily Statistics, and Summary Statistics)
            1. Significant figures
            2. Remove extra “NO DATA” or “MISSING” text from analyte qualifier columns on final data sheet and graphs-stats data sheet
            3. NOTE: all symbols (censored (##), missing (\*\*), and NO DATA (--) have been previously applied to the final data sheet during QC2 and QC3 review
         9. Add qualifier counts to AllProjects QC4 summary workbook
            1. Filter and count each analyte data column on final sheet for

Missing (“\*\*”)

No data (“--“)

Censored (“##”)

* + - * 1. Filter and count each analyte qualifier column for

Blanks (# accepted is calculated as # blanks - # missing - # no data)

Non-blanks (# qualified only = # tot qualified - # censored)

Specific qualifier symbols (i, m, r, t, u)

* + - * 1. Write counts to Allprojects QC4 summary file, Qualifer sheet
        2. NOTES:

# accepted + # qualified + # censored = observed count; observed count + # missing + # no data = total count

Sum of qualifier symbol counts = # qualified + # censored

* + - 1. Adjust graphs (marker and line colors, gridlines, marker size)
      2. Apply formatting to data sheets (final, graphs-stats, Daily Statistics, and Summary Statistics)
         1. Rename columns
         2. Set font type, size, weight, color
         3. Align cells and set fill colors
         4. Adjust column widths and apply text wrapping
         5. Apply cell borders
         6. Correct cases where Connecticut misspelled (2008 only) on Summary Statistics sheet
         7. Autofilter, freeze top row and go to cell “A1” (final, graph-stats, Daily Statistics sheet only)
      3. Delete extra analyte columns from data sheets
      4. Protect worksheets
      5. Select Read me sheet and go to cell “A1”
      6. Set file attribute to read only
      7. Save and close QC4 final file

1. Format AllProjects QC4 Summary workbook
   1. Apply reporting rules and formatting to Statistics sheet
      1. Rename columns
      2. Set font type, size, weight, color
      3. Align cells and set fill colors
      4. Adjust column widths and apply text wrapping
      5. Apply cell borders
      6. Apply significant figures
      7. Add “--“ to statistics for NO FILE rows
   2. Apply reporting rules and formatting to Qualifier sheet
      1. Rename columns
      2. Set font type, size, weight, color
      3. Align cells and set fill colors
      4. Adjust column widths and apply text wrapping
      5. Apply cell borders
      6. Apply autofilter
      7. Apply significant figures
      8. Add “--“ to statistics for NO FILE rows
   3. Add read me sheet
      1. Add hyperlinks to files/paths on read me sheet (see III c-e.)
      2. Add date of finalization
2. Split AllProjects QC4 summary workbook into project-specific QC4 summary files
   1. Define project file name and path (from project folder names stored in array)
   2. Create new workbook
   3. Copy read me sheet to project workbook
      1. Add hyperlinks to files/paths on read me sheet (see III c-e.)
      2. Add date of finalization
   4. Add Statistics sheet
      1. Filter AllProjects QC4 summary file Statistics sheet for project name
      2. Copy rows from Allprojects QC4 summary file to project workbook
      3. Delete duplicate OWMIDs from project workbook (by filtering for DupesNot2Use = “X”)
      4. Delete unneeded columns (DupesNot2Use, Duplicate OWMID, Duplicate, PPFile, SondeID)
      5. Freeze top row and go to cell “A1”
      6. Redefine used range after deleting columns and autofilter
      7. Protect sheet
   5. Add Qualifier Summary sheet
      1. Filter AllProjects QC4 summary file Qualifier sheet for project name
      2. Copy rows from Allprojects QC4 summary file to project workbook
      3. Delete duplicate OWMIDs from project workbook (by filtering for DupesNot2Use = “X”)
      4. Delete unneeded columns (DupesNot2Use, Duplicate OWMID, Duplicate, PPFile, SondeID)
      5. Freeze top row and go to cell “A1”
      6. Redefine used range after deleting columns and autofilter
      7. Protect sheet
   6. Delete extra sheets from project workbook
   7. Select read me sheet and go to cell “A1”
   8. Save and close project workbook
   9. Set file attribute to read-only
3. Apply final changes to Allprojects QC4 summary workbook
   1. Statistics sheet
      1. Delete duplicate OWMIDs from project workbook (by filtering for DupesNot2Use = “X”)
      2. Delete unneeded columns (DupesNot2Use, Duplicate OWMID, Duplicate, PPFile, SondeID)
      3. Freeze top row and go to cell “A1”
      4. Redefine used range after deleting columns and autofilter
      5. Protect sheet
   2. Qualifier Summary sheet
      1. Delete duplicate OWMIDs from project workbook (by filtering for DupesNot2Use = “X”)
      2. Delete unneeded columns (DupesNot2Use, Duplicate OWMID, Duplicate, PPFile, SondeID , Station/Waterbody columns)
      3. Freeze top row and go to cell “A1”
      4. Redefine used range after deleting columns and autofilter
      5. Protect sheet
   3. Select read me sheet and go to cell “A1”
   4. Save and close Allprojects QC4 workbook
   5. Set file attribute to read-only
4. Reopen QC3 summary workbook
   1. Add start and end time for code run to File Process Info sheet
   2. Save and close

## Generate QC4 Station Files

**General Purpose:** To produce final QC2 or QC4 unattended multiprobe station files with all symbols, reporting rules, and other formatting applied, as well as QC Summary project workbooks containing statistics and qualifier summary counts for all stations in a project

**Filename:** UnAtt\_Stations\_Finalize\_V3.wsf

**Code Language:** VBScript

**Input Files:**

* QC2 or QC3 Station files (e.g. W0168\_QC3\_MM-DD-YYYY.xlsx, where W0168 is the 5 character UniqueID, starting with W and followed by 4 numbers; QC2 or QC3 is the QC review level of the files; and MM is the month, DD is the day, and YYYY is the year program was run)

**Output Files:**

* QC2 or QC4 final station files (e.g. W0168\_QC4FINAL\_MM-DD-YYYY.xlsm, where W0168 is the 5 character UniqueID ; QC2 or QC4 is the final QC level; and MM is the month, DD is the day, and YYYY is the year that the file was generated); files are output to project folders (e.g. Projectname\_yyyy\_UQC4\_FINAL\_Stations\_MM-DD-YYYY, where MM is the month, DD is the day, and YYYY is the year that the file was generated; UQC2 or UQC4 is the final QC level; and yyyy is the data year)
* QC2 or QC4 final Summary workbooks for each project (Projectname\_yyyy\_unattended\_QC4FINAL\_MM-DD-YYYY.xlsx, where MM is the month, DD is the day, and YYYY is the year that the file was generated, and yyyy is the data year); files are output to project folders; files contain a sheet for statistics and a sheet for qualifier summary counts for all stations in the project for the data year

**Approximate Run Time:** 15-30 minutes

**Code Outline:**

1. announce to user what code does and give user opportunity to back out
2. user enters data year for analysis
3. user enters QC level of files to process (QC2 or QC3 review)
4. check for presence of directories/files
   1. user navigates to QC station file directory for QC level and data year entered above
   2. user navigates to location of QC2 or QC4 unattended “read me” file and selects it
      1. open workbook and check for presence of “read me” worksheet
      2. if worksheet not found, present error to user and quit
      3. find rows for “META DATA NETWORK LOCATION”, “NETWORK FILE LOCATION”, and “SPECIFIC METHODOLOGY” (for later population of hyperlinks to files/folders)
   3. get path to “MetaDatayyyy.xlsx” file, where yyyy is the data year (hard-coded unless folder not found, then user navigates to the file and selects it) for populating “META DATA NETWORK LOCATION” on read me sheet
   4. get path to network location of QC4 files (hard-coded unless folder not found, then user navigates to correct location) for populating “NETWORK FILE LOCATION” on read me sheet
   5. get path to process SOPs directory (hard-coded unless folder not found, then user navigates to correct location)for populating “SPECIFIC METHODOLOGY” on read me sheet
   6. check for files in QC2 or QC3 directory
      1. if no files present, present user with error message and quit
      2. if files present, present user with # of files found
   7. get parent directory for station file directory selected in IV a.
   8. extract year, month, day from QC2 or QC3 station folder parent directory (from IV. g.)
      1. check for presence of QC summary workbook (with same date in name as parent directory)
      2. present user with error message if QC summary workbook not found, quit
   9. create folder for QC2 or QC4 final files (yyyyUQC4\_FINAL\_Stations\_MM-DD-YYYY, where MM is the month, DD is the day, and YYYY is the year the folder is created; yyyy is the data year; QC2 or QC4 is the final QC level; stored as a subfolder of the QC2 or QC3 station file directory selected in IV a.)
   10. create AllProjects folder for storing AllProjects QC Summary file (AllProjects\_yyyy\_UQC4\_FINAL\_MM-DD-YYYY, where where MM is the month, DD is the day, and YYYY is the year the folder is created; and yyyy is the data year; QC2 or QC4 is the final QC level; stored as a subfolder of the QC4 final folder in III h.)
   11. define name of AllProjects QC2 or QC4 Summary workbook
   12. create text file for recording issues (YYYY\_Unattended Finalize Station Files\_Status.txt) in QC2 or QC4 final folder in IV i.
5. create Allprojects QC2 or QC4 Final summary workbook from QC2 or QC3 summary workbook
   1. check for presence of required worksheets (Station Statistics)
      1. present user with error message if sheet doesn’t exist and quit
   2. set Statistics worksheet object and get used range
   3. define columns on Statistics sheet
   4. check that statistics have been run by comparing # of values in a statistics column against the # of QC2 or QC3 files listed on the sheet
      1. present user with error message, if no statistics found or if # of statistics doesn’t match # of files, and quit
6. get min start time and max end time of files for standardizing data range of x-axis on graphs
7. create QC2 or QC4 final project folders
   1. sort Statistics sheet in QC Summary workbook by Project, Watershed, and File Start Date
   2. correct cases where Connecticut misspelled (2008 only)
   3. get unique project names from Project column and store in array
      1. correct names for unwanted or illegal characters that can’t be used in folder names
   4. create project folders in QC2 or QC4 directory in IV i.
8. Apply reporting rules to STAID column in AllProjects QC2 or QC4 Final Summary workbook
   1. Apply NO DATA (“--”) to blank STAID cells
9. Create new sheet in Allprojects QC2 or QC4 Final Summary workbook for qualifier counts
   1. New sheet name = “Station Qualifier Summary”
   2. Define columns on qualifier sheet
   3. Copy select columns from Statistics sheet to Qualifer sheet (Project-related, station-related, waterbody-related, and Total Count of rows per file)
   4. Add column headers to qualifier sheet
10. Create QC2 or QC4 final station files from QC2 or QC3 review files
    1. Define columns in station files for data sheets, Daily Statistics sheet, and Summary Statistics sheet
    2. Loop through UniqueIDs listed on Statistics sheet in Allprojects QC2 or QC4 Final summary file
       1. Get QC2/QC3 file name, UniqueID, STAID, project
       2. Match current project to project folder using projects stored in array
       3. Define QC2/QC4 Final station file name and path (based on project folder name)
          1. If no project folder match found, place file in Allprojects folder
       4. Open QC2/QC3 Station file
          1. Save QC2/QC3 file as QC2/QC4 temp file (.xlsm) (otherwise Excel present dialog box for saving files with VBA code as .xlsx)
          2. Delete unneeded worksheets from QC2/QC4 temp file (main data sheet, lists)
          3. Save QC2/QC4 temp file as QC2/QC4 final file, close temp file and delete
          4. Open QC2/QC4 final file
          5. Write final QC2/QC3 name to Allprojects QC2/QC4 Final Summary file (Statistics and Qualifier sheets)
          6. Define worksheet objects and used ranges
          7. Add read me sheet
             1. Add hyperlinks to files/paths on read me sheet (see III c-e.)
             2. Add date of finalization
          8. Apply reporting rules to worksheets (final data, graphs-stats, Daily Statistics, and Summary Statistics)
             1. Significant figures
             2. Remove extra “NO DATA” or “MISSING” text from analyte qualifier columns on final data sheet and graphs-stats data sheet
             3. NOTE: all symbols (censored (##), missing (\*\*), and NO DATA (--) have been previously applied to the final data sheet during QC2 and QC3 review
          9. Add qualifier counts to AllProjects QC4 summary workbook
             1. Filter and count each analyte data column on final sheet for

Missing (“\*\*”)

No data (“--“)

Censored (“##”)

* + - * 1. Filter and count each analyte qualifier column for

Blanks (# accepted is calculated as # blanks - # missing - # no data)

Non-blanks (# qualified only = # tot qualified - # censored)

Specific qualifier symbols (i, m, r, t, u)

* + - * 1. Write counts to Allprojects QC4 summary file, Qualifer sheet
        2. NOTES:

# accepted + # qualified + # censored = observed count; observed count + # missing + # no data = total count

Sum of qualifier symbol counts = # qualified + # censored

* + - 1. Adjust graphs (marker and line colors, gridlines, marker size, xmin and xmax on x-axis)
      2. Apply formatting to data sheets (final, graphs-stats, Daily Statistics, and Summary Statistics)
         1. Rename columns
         2. Set font type, size, weight, color
         3. Align cells and set fill colors
         4. Adjust column widths and apply text wrapping
         5. Apply cell borders
         6. Correct cases where Connecticut misspelled (2008 only) on Summary Statistics sheet
         7. Autofilter, freeze top row and go to cell “A1” (final, graph-stats, Daily Statistics sheet only)
      3. Delete extra analyte columns from data sheets
      4. Protect worksheets
      5. Select Read me sheet and go to cell “A1”
      6. Set file attribute to read only
      7. Save and close QC4 final file

1. Format AllProjects QC2/QC4 Final Summary workbook
   1. Apply reporting rules and formatting to Statistics sheet
      1. Rename columns
      2. Set font type, size, weight, color
      3. Align cells and set fill colors
      4. Adjust column widths and apply text wrapping
      5. Apply cell borders
      6. Apply significant figures
   2. Apply reporting rules and formatting to Qualifier sheet
      1. Rename columns
      2. Set font type, size, weight, color
      3. Align cells and set fill colors
      4. Adjust column widths and apply text wrapping
      5. Apply cell borders
      6. Apply autofilter
      7. Apply significant figures
   3. Add read me sheet
      1. Add hyperlinks to files/paths on read me sheet (see IV c-e.)
      2. Add date of finalization
2. Split AllProjects QC2/QC4 FINAL summary workbook into project-specific QC2/QC4 FINAL summary files
   1. Define project file name and path (from project folder names stored in array)
   2. Create new workbook
   3. Copy read me sheet to project workbook
      1. Add hyperlinks to files/paths on read me sheet (see IV c-e.)
      2. Add date of finalization
   4. Add Statistics sheet
      1. Filter AllProjects QC4 summary file Statistics sheet for project name
      2. Copy rows from Allprojects QC4 summary file to project workbook
      3. Delete unneeded columns
      4. Freeze top row and go to cell “A1”
      5. Redefine used range after deleting columns and autofilter
      6. Protect sheet
   5. Add Qualifier Summary sheet
      1. Filter AllProjects QC4 summary file Qualifier sheet for project name
      2. Copy rows from Allprojects QC4 summary file to project workbook
      3. Delete duplicate OWMIDs from project workbook (by filtering for DupesNot2Use = “X”)
      4. Delete unneeded columns
      5. Freeze top row and go to cell “A1”
      6. Redefine used range after deleting columns and autofilter
      7. Protect sheet
   6. Delete extra sheets from project workbook
   7. Select read me sheet and go to cell “A1”
   8. Save and close project workbook
   9. Set file attribute to read-only
3. Apply final changes to Allprojects QC2/QC4 Final summary workbook
   1. Statistics sheet
      1. Delete unneeded columns
      2. Freeze top row and go to cell “A1”
      3. Redefine used range after deleting columns and autofilter
      4. Protect sheet
   2. Qualifier Summary sheet
      1. Delete unneeded columns
      2. Freeze top row and go to cell “A1”
      3. Redefine used range after deleting columns and autofilter
      4. Protect sheet
   3. Select read me sheet and go to cell “A1”
   4. Save and close Allprojects QC2/QC4 Final Summary workbook
   5. Set file attribute to read-only
4. Reopen QC2/QC3 summary workbook
   1. Add start and end time for code run to File Process Info sheet
   2. Save and close

# APPENDIX E: Statistics

## Definitions

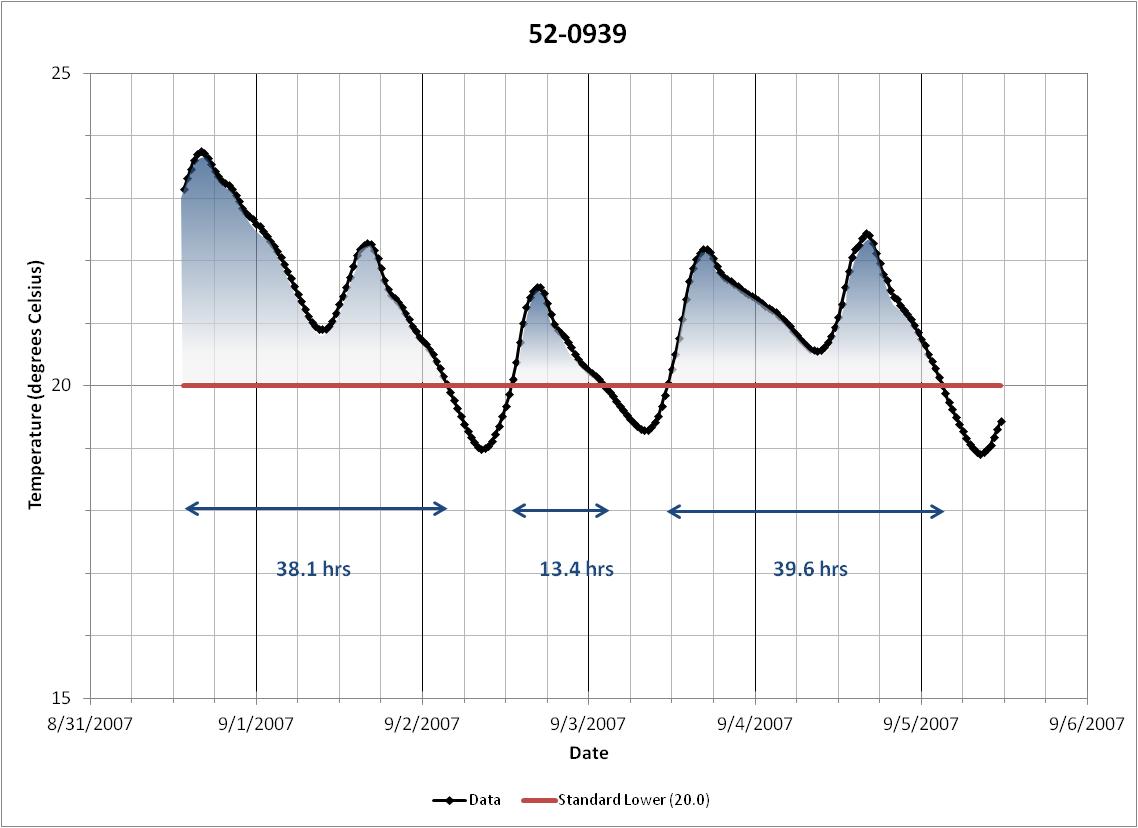
| **Statistic** | **Description** |
| --- | --- |
| Total Deployment Duration | Total amount of time from the first line of data to the last line of data in the trimmed file. (Units = Hours) |
| Observed Deployment Time | Total amount of time from the first line of data to the last line of data in the trimmed file, minus any data gaps and censored data. Reported separately for T, DO, and DOsat based on the observed number of data values. (Units = Hours) |
| File Time Interval | The pre-programmed time interval between data points in the file (Units = Minutes) |
| Avg | Average of all the usable data points for each analyte in the file (analyte units, e.g. deg. C, mg/L, %) |
| SD | Standard Deviation: Indicates how much variation or "dispersion" there is from the mean value (a low SD indicates that the data points tend to be very close to the mean; a high standard deviation indicates that the data are spread out over a large range of values). (Units are dimensionless) |
| Min | Minimum: The lowest numeric value of all the usable data points in the file (analyte units, e.g. deg. C, mg/L, %) |
| Max | Maximum: The highest numeric value of all the usable data points in the file (analyte units, e.g. deg. C, mg/L, %) |
| Median | The numerical value separating the upper half of all the usable data from the lower half, found by ranking all the observations from lowest value to highest value and picking the middle one (if there is an even number of observations, then there is no single middle value; the median is then usually defined to be the mean of the two middle values). (analyte units, e.g. deg. C, mg/L, %) |
| IQR | Interquartile Range: The distance between the 1st quartile (25th percentile) and 3rd quartile (75th percentile), or essentially the range of the middle 50% of the data. (analyte units, e.g. deg. C, mg/L, %) |
| Total Count | Total number of data points in the trimmed file (from the first line of data to the last). |
| Observed Count | Total number of usable data points for each analyte in the file (excluding censored/missing values) |
| Amount of Time ><WQS | Total number of hours > or < WQS for each analyte in the file (based on interpolated time intervals), excluding missing and censored points. For T, total amount of time above the WQS, where WQS included 20, 28.3, and 29.4 deg C. For DO, total amount of time below the WQS, where WQS included 3, 4, 5, and 6 mg/L. (Units = Hours) |
| Max Duration ><WQS | For all the usable data for each analyte in the file, the highest amount of continuous hours > or < WQS (based on interpolated time intervals) (note: continuous hours are defined by consecutive data values above or below the WQS; the duration > or < WQS may begin or end at a blank datum). (Units = Hours) |
| Avg Daily Amount of Time ><WQS | The average of each analyte's DAILY AMOUNT OF TIME >< WQS. For cases where there is only one complete day to average, the average equals the DAILY AMOUNT OF TIME. (analyte units, e.g. deg. C, mg/L) |
| MWAT | Maximum Weekly Average Temperature: The average temperature in the warmest week for water temperatures (rolling 7-day average using 7 complete days), i.e. the maximum of the 7-day rolling averages (7-DADA) (Units = deg. C) |
| Mean of the Daily Mean | The average of the 24-hour-daily means (DAILY MEAN) for each analyte (analyte units, e.g. deg. C or mg/L) |
| Mean of the Daily Min | The average of the 24-hour-daily minimums (DAILY MIN) for each analyte (analyte units, e.g. deg. C or mg/L) |
| Mean of the Daily Max | The average of the 24-hour-daily maximums (DAILY MAX) for each analyte (analyte units, e.g., deg C. or mg/L) |
| Daily Amount of Time >< WQS | For each analyte, the total number of hours within each complete day where datum < or > WQS. (Units = Hours) |
| Daily Mean | For each analyte, the average of all the usable data points within each complete day (Units are same as analyte units) |
| Daily Min | For each analyte, the lowest numeric value of all the usable data points within each complete day (Units are same as analyte units) |
| Daily Max | For each analyte, the highest numeric value of all usable points in each complete day (Units are same as analyte units) |
| 7-DADM | 7-Day Avg of the Daily Max: The rolling average of the highest numeric value on each day and three complete days before and three complete days after (7 complete days total used in calculation). Also called the "mean of the daily maximum temperature over a 7 day period". Computed from DAILY MAX values for temperature only. (Units = deg C) |
| 7-DADA | 7-Day Avg of the Daily Avg : The rolling average of the mean value for each day and three complete days before and three complete days after (7 complete days total used in calculation). From these calculations of weekly averages, the maximum weekly average temperature or MWAT can be derived. Computed from DAILY MEAN values for temperature only. (Units = deg C) |

## Notes

|  |
| --- |
| *1 Lower limits for DO (0.2 mg/L) and DOsat (2 %) were used in place of actual values below these limits during statistical calculations due to the inherent uncertainty in the sensors at these low values.* |
| *2 Daily statistics reported using the "--" symbol were either not planned to be calculated or represent cases where a complete day of an analyte's data was not available due to censored/missing data blocks or due to starting or ending the deployment part way through the day.* |
| *3 If desired, % of time >< WQS can be calculated by dividing the AMOUNT OF TIME >< WQS by the OBSERVED DEPLOYMENT TIME.* |
| *4 The threshold lines shown on the charts do not necessarily apply to the station sampled in each data file (i.e., they are provided for general reference purposes only). For applicable standards, please check the Mass SWQS.* |
| *5 Daily statistics calculated from analyte values are based on complete days only, where a complete day is defined from midnight to 11:30 PM (or other logged time interval) with no censored or missing data points; for daily statistics calculated as times >< WQS, a complete day is defined from midnight on the current date (t = 0) to midnight (t = 24 hours) on the next date, with the statistic recorded on the current date.* |

## Examples

1. **Temperature: Amount of Time > WQS and Maximum Amount of Time > WQS**

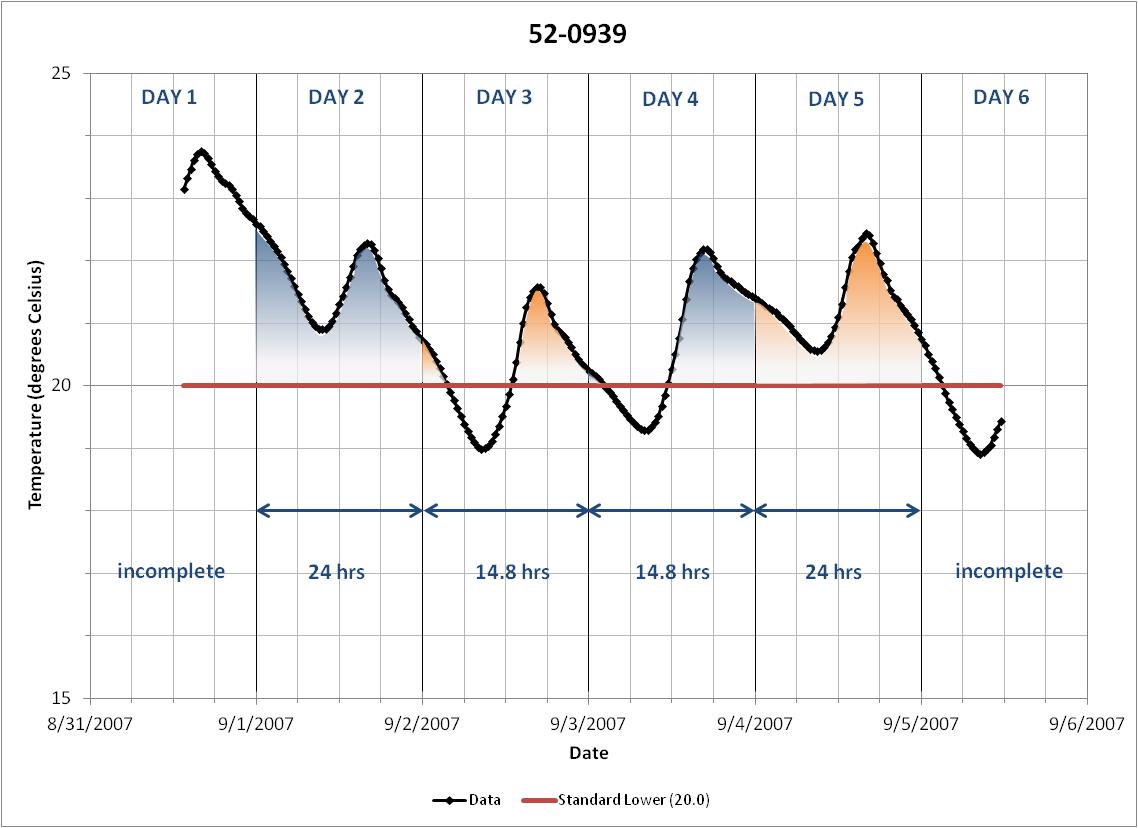


The number of temperature excursions above the water quality standard (20 deg C) is 3, at 38.1 hours, 13.4 hours, and 39.6 hours in length.

**Amount of Time > 20 deg. C** = 38.1 + 13.4 + 39.6 = 91.1 hours

**Max Duration > 20 deg. C** = 39.6 hours

1. **Daily Amount of Time > WQS**



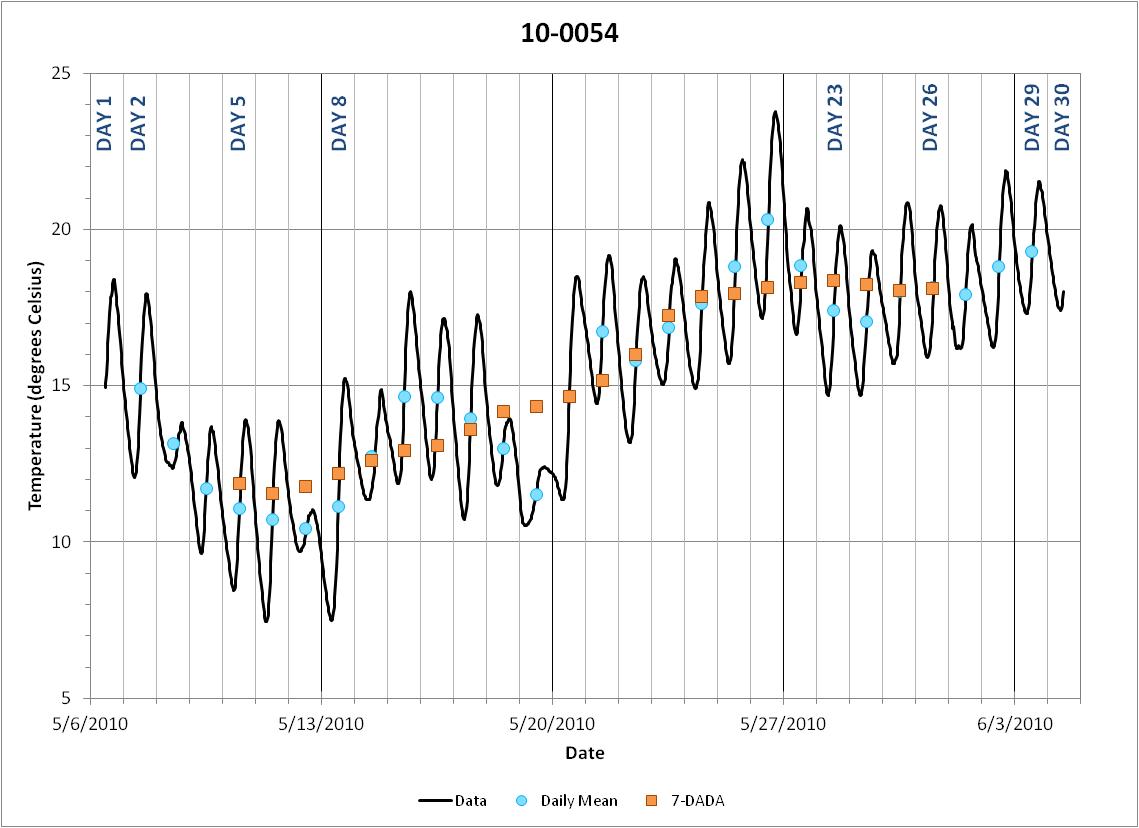
**Daily Amount of Time > 20 deg. C**

Day 1 and Day 6 are excluded because they are not complete.

Days 1 through 4 are complete with the amount of time above the standard at 24 hours, 14.8 hours, 14.8 hours, and 24 hours in length.

**Avg Daily Amount of Time > 20 deg. C** = (24 + 14.8 + 14.8 + 24)/4 = 19.4 hours

1. **Daily Mean, 7-Day Average of the Daily Average, and Maximum Weekly Average Temperature**



| **Day** | **Date** | **Daily Mean** | **7-DADA** |
| --- | --- | --- | --- |
| 1 | 5/6/2010 | incomplete |  |
| 2 | 5/7/2010 | 14.9 |  |
| 3 | 5/8/2010 | 13.2 |  |
| 4 | 5/9/2010 | 11.7 |  |
| 5 | 5/10/2010 | 11.1 | 11.9 |
| 6 | 5/11/2010 | 10.7 | 11.6 |
| 7 | 5/12/2010 | 10.4 | 11.8 |
| 8 | 5/13/2010 | 11.2 | 12.2 |
| 9 | 5/14/2010 | 12.8 | 12.6 |
| 10 | 5/15/2010 | 14.7 | 13.0 |
| 11 | 5/16/2010 | 14.6 | 13.1 |
| 12 | 5/17/2010 | 14.0 | 13.6 |
| 13 | 5/18/2010 | 13.0 | 14.2 |
| 14 | 5/19/2010 | 11.5 | 14.3 |
| 15 | 5/20/2010 | 14.7 | 14.7 |
| 16 | 5/21/2010 | 16.7 | 15.2 |
| 17 | 5/22/2010 | 15.8 | 16.0 |
| 18 | 5/23/2010 | 16.9 | 17.3 |
| 19 | 5/24/2010 | 17.6 | 17.9 |
| 20 | 5/25/2010 | 18.8 | 18.0 |
| 21 | 5/26/2010 | 20.3 | 18.1 |
| 22 | 5/27/2010 | 18.9 | 18.3 |
| 23 | 5/28/2010 | 17.4 | 18.4 |
| 24 | 5/29/2010 | 17.1 | 18.3 |
| 25 | 5/30/2010 | 18.0 | 18.0 |
| 26 | 5/31/2010 | 18.2 | 18.1 |
| 27 | 6/1/2010 | 17.9 |  |
| 28 | 6/2/2010 | 18.8 |  |
| 29 | 6/3/2010 | 19.3 |  |
| 30 | 6/4/2010 | incomplete |  |

**Daily Mean**

Day 1 and Day 30 are incomplete, so daily means could not be computed for these days.

**7-DADA**

The 7-day rolling average is computed for a given day from the daily mean values (complete days only) starting three days before the date in question. For example, because Day 1 is incomplete, the first 7-DADA is computed for Days 2 through 8 and is recorded on Day 5. The next 7-DADA is computed for Days 3 through 9 and is recorded on Day 6. As a result of the way the 7-day average is calculated and because Day 30 is incomplete, the last 7-DADA value is recorded on the 5th day from the end of the survey (Day 26).

**MWAT**

The maximum 7-DADA value of 18.4 deg. C was recorded on Day 23.

