

New Techniques in Accessing and Analyzing Water-Quality Data

Workshop on retrieving and evaluating data from the Water Quality Portal
presented at the NWQMC Conference – Portland, OR - May 1, 2012

- | | | |
|------|--|---------------------------------|
| 1:30 | Overview | Nate Booth &
Lorraine Murphy |
| | Example data retrieval – Big Thompson River nutrients | |
| | Sites | |
| | Map locations (KML data & Google Earth) | |
| | Results | |
| 1:45 | Participant exercise – retrieving data | All |
| 1:55 | Example data evaluation – Sites | Dave Mueller |
| | Picking sites of interest using the .kml file | |
| | Issues: duplicate IDs, inaccurate lat/long | |
| | Sorting and selecting sites (using the data filter in Excel) | |
| 2:10 | Participant exercise – Site evaluation | All |
| 2:20 | Example data evaluation – Results | Dave Mueller |
| | Selecting and organizing columns | |
| | Sorting and selecting rows (inc. combining analytes) | |
| | Issues: various names, wrong names, unknown units | |
| | Adding calculations: censored values, unit conversions | |
| 2:35 | Participant exercise – Results manipulation & evaluation | All |
| 2:45 | Demonstrate graphical data evaluation | Dave Lorenz |
| | Importing data to other software | |
| | Time series plots | |
| | Inter-site comparison (box plots) | |

Sorting and selecting sites:

1. Retrieve site data as a .csv file
2. Open file in Excel, save as a .xlsx file
3. Set column width
4. Apply data filter
5. Sort by (lat for Big T sites, since stream flows W to E)
6. Add column for downstream order
7. Deselect blanks in the “DS order” column
8. Save revised file

Working with the results:

1. Retrieve results as a .csv file
2. Open file in Excel, save as a .xlsx file
3. Set column width
4. Remove blank line between data from NWIS and data from STORET
5. Apply data filter
6. Hide columns with no data, a single value, or data of little utility
7. Add a column for downstream order
8. Using the filter, select each “MonitoringLocationIdentifier” identified in the revised site file and input the specified “DS order” number and “MonitoringLocationName” for all results
9. Clear the filter, save the revised file (1)
10. Select blanks in the “DS order” column; delete all the selected rows
11. Save the revised file (2)
12. Sort the data by site (DS order), date (ActivityStartDate), analyte (CharacteristicName), and fraction (ResultSampleFractionText)
13. Identify and delete duplicate samples (samples for the same analyte on the same date with duplicate site IDs; some duplicate calculated values, such as ammonia as NH₄, phosphate as PO₄; and some other calculated values of little interest, such as organic N)
14. Save the revised file (3)
15. Use the filter to select and delete rows with “ResultMeasure/MeasureUnitCode” that are not appropriate for the media of interest (for example, “mg/kg” for “Water”). Though not an issue with the BT data, this is common for USGS data on chemistry of suspended sediment.
16. Add a column for common analytes names

17. Use the filter to select "CharacteristicName" for similar (identical) analytes and fill in common name
18. Recheck for duplicates (e.g., when "Nutrient-nitrogen" and "Nitrate" are both reported for the same sample, they have the same result. Check the method; rename and retain/delete as needed)
19. Check methods (e.g., some "Orthohosphate" results are from EPA 365.4, which is for Total P; some of these are duplicate results for "Phosphorus". Rename and retain/delete as needed)
20. Re-sort the data by analyte (Analyte common name), site (DS order), and date (ActivityStartDate)
21. Save the revised file (4)

22. Rearrange some columns – move "ResultDetectionConditionText" and "DetectionQuantitationLimit---" near "ResultMeasureValue"
23. Check for various units for same analyte. If found, add columns for "Value" and "Units"; copy "ResultMeasureValue" into the "Value" column or compute a values in the common units as needed. Fill in the "Units" column as appropriate.
24. Add a column for "Remark". If "ResultDetectionConditionText" is "Not detected" or "*Non-detect", or if "ResultMeasureValue" is zero, "put a "<" in the "Remark" column and copy "DetectionQuantitationLimitMeasure/MeasureValue" (with appropriate units conversion, if needed) into the "Value" column.
25. Save the revised file (5)

26. Add a column for "Quantitation limit"; copy or convert the values in "DetectionQuantitationLimitMeasure/MeasureValue", as was done for "Value" (above).
27. Add a column for "Value < QL?". For all results with a quantitation limit, determine whether this limit is greater than the reported result, in "Value".