Minutes of the QA/QC Workgroup Meeting

June 20, 2012

Attendees: Murage Ngatia, Don Guy, Bill Templin, Dave Bosworth, Cindy Garcia, Steve San Julian, Travis Brown. On the phone: Perry LeBouf, Scott McReynolds, Kelley Pepper, Bruce Agee

1. Introductions
2. Update for New Members
    a. All docs are on the DES shared drive. Krista will work on getting the DIRWM and DSIWM members access with DES's LAN administrators.
3. Outliers Presentation – Murage (Power Point presentation on shared drive)
    a. Outliers may also be called anomalies because the word "outlier" has a negative connotation. Outliers maybe valid or invalid data.
    b. Grubbs definition and method is widely used in lab environmental data outlier detection.
    c. How to detect outliers: can use the Grubbs' method. However, it assumes a normal distribution which is not common with environmental data. There are other methods to use.
    d. How to deal with outliers: remove them but don't delete them. You can also use robust methods that do not use the mean and standard deviation because they are highly affected by outliers.
    e. Keep track of instrument maintenance. If there are outliers in your data, they may indicate your instrument needs maintenance.
    f. CDEC's main problem is highlighted with outliers: it is real time so there is no QC done on the data. This is why IRWM uses Hydstra and uploads QCeed data to the WDL.
    g. O&M uses Logger Net and its Split scripting module to detect outliers, remove the data, and upload to CDEC. This came about because people were worried about uploading raw data onto CDEC.
    h. Dave B. will send Murage the info on marking outliers in continuous turbidity and chlorophyll data so it can be distributed.
4. Hydstra Presentation – Dave
    a. Staff in the Water Quality Evaluations Unit (WQES) in NCRO use slope criteria to find potential outliers in their continuous turbidity and chlorophyll data. These slope criteria are not firm in that WQES staff also use their best professional judgment to determine if a data point is unreliable or not. The slope criteria are arbitrary but hard to exceed. When a data point is determined to be an unreliable outlier, WQES staff mark them with an unreliable quality code in hydstra. The hydstra software does not delete the marked data, but this data won't be uploaded to the WDL.
    b. How do you know the outlier is really an anomaly? Know your station and its surroundings! If WQES staff suspect that the instrument has been interfered with such as a fish living near the instrument or the occurrence of submerged aquatic vegetation

then this should be used as further evidence to mark a data point as an unreliable outlier.

c. The program allows for Excel data uploading so you don't have to do everything manually in Hydstra.
d. The WQES methods for marking unreliable outliers is for time-series data collected at stationary stations, not data collected at different depths and locations. However, everyone can try it and tweak it for their use.
e. Hydstra can be used to view, archive, import, and QC data. You can graph, chart, plot, and summarize data. Custom reports are available as well.
f. The cost is approximately $600 per seat to use. Eric Senter in DSIWM/DIRWM is the contact.

5. Rotating Meeting Chair
    a. The next meeting chair is scheduled to be Bill Burkhard. Murage will see if he is available. If not, Don Guy will present.