DRI QA/QC Procedures for Biscayne Aquifer Monitoring (BAM) Project.
November 2012

I. Data Flow

Data are collected from sensors at the nine stations in WCA3B using either CR1000 or CR206 data loggers (Campbell Scientific). The data collected includes water elevation, groundwater velocity and flow direction, water temperature, and specific conductance. Data files retrieved from the loggers have a "table-based" format that includes header information that identifies each data field. Each logger saves at least two types of data files. The first file contains the raw data collected at the station, typically on a 15-minute interval. The second file contains metadata including site information, probe characteristics, serial numbers, and calibration information. The metadata file is updated daily at midnight or automatically whenever the field technician updates sensor calibration information stored on the logger. Data files are retrieved by IP-based telemetry on a 15-minute interval and are saved as comma-delimited (.csv) file on the host computer at DRI.

New data present in these data files are automatically ingested into a MySQL database on a roughly two-minute interval. Within the database, data reside in three different database tables: "Raw", "QA/QC", and "Finalized". Data first enter the database and are placed in the Raw Table – this data is never modified or altered. On a 30-minute basis, new data from the Raw Table are automatically copied into the QA/QC Table. Erroneous data in this table can be modified for quality control purposes and data quality ranks assigned using a customized graphical user interface (GUI) based program. All modifications to the data are logged with user, date time, the type of change (e.g. deletion, interpolation), and the mathematical parameters used to carry out the change. New data fields can also be defined in the QA/QC Table based on mathematical formulas on one or more existing data fields. This calculated variable feature permits the use of different formulas for different time periods. Once the erroneous data have been corrected and flagged with a QC ranking code, the user then marks each data point as finalized, triggering an automatic copying of those data points from the QA/QC Table into the Finalized Table. At DRI, a Data Technician will be responsible for correcting erroneous data and for applying QC ranking codes as described below. The Data Supervisor will be responsible for checking the accuracy of the data corrections and finalizing corrections. Both the raw and finalized datasets will be available on the project website, which has a web browser-based software interface that provides for visualization, dynamic plotting, and retrieval of data.

Data within the database are backed up to an external hard drive nightly. Backup copies older than one week are deleted except for one weekly backup that is maintained for a month. Monthly backups are maintained throughout the duration of the project. Weekly and Monthly copies are also transferred to a private share on DRI's servers that are transferred to tape nightly. Tape copies reside in another building on DRI's campus and are stored for a period of one year. Copies of the original data files from the data

logger can be mirrored to a password-protected ftp server where they are available for transfer offsite (e.g. to the Jacksonville Office of USACE).

II. Recording of Field QA/QC Activities

ENTRIX personnel will complete operation and maintenance (O&M) activities at least once per month, including field verification of water surface elevation, temperature, and specific conductance. ENTRIX personnel will log onto the data logger with a portable computer and record current sensor conditions in their field notes. Water surface elevation (WSE) will be field verified. If the recorded reading is > ±0.02 ft different than the field verified value, then the "level_offset" variable in the data logger will be modified so that the recorded value matches the field verification. This change will be entered into the field notes and will be automatically recorded into the data loggers' metadata file. Specific conductance will be checked against a known standard, with the calibration check value recorded in the field notes. After several months of data collection, DRI personnel will perform a trend analysis of the specific conductance data to assess sensor performance that will be subsequently used by DRI to correct data and assign data QC ratings. Temperature values will also be field-verified while performing O&M.

QA of groundwater flowmeters will include a visual verification that the suspension rods are properly installed, oriented vertically, and oriented with respect to north. Flowmeters will be pulled from the wells and inspected for fouling at least quarterly. Spot checks of thermistors using localized temperature gradients (e.g., ice cube test) will be performed to verify proper correspondence between thermistors and the datalogger monitoring system. If a groundwater flowmeter is changed, datalogger variables containing flowmeter serial numbers and calibration coefficients will be verified. Groundwater velocity data will be reviewed in conjunction with hydraulic head data to identify possible fouling such as by sediment accumulation, or for flowmeter malfunctions. Data review will focus on time periods in which groundwater flow or direction changes occur without a corresponding change in hydraulic head associated with a precipitation event or pumped drawdown.

ENTRIX personnel will transcribe the field notes into electronic form and transmit that to both the Data Technician and Supervisor at DRI no later than one week after the site visit. The Data Technician will be responsible for transferring the electronic field notes into the database so they are viewable as "QA Field Reports" in the "Notes" feature of the project website. Field notes will be automatically parsed into their own database table and can be mirrored to the ftp site for offsite transfer.

III. Quality Control Practices

Data may be erroneous for many reasons, including electrical interference, sensor removal during site visits, instantaneous spikes or peaks, sensor drift, sensor fouling and data transmission errors. The purpose of quality control practices are to identify erroneous data, assess if a data correction should be applied, and finally to rank the

accuracy of the continuous data based on field calibration checks. This analysis will be conducted no later than one week after receiving the field report.

III.a. Identification and Correction of Erroneous Data
Data will be visually assessed by DRI for anomalies and outliers that diverge from field calibration checks and from current data trends from not only the sensor itself, but with other sensors co-located in the same well, and other sensors at appropriate nearby sites.

The following rules govern the correction of erroneous data within the QA/QC database table:

1. Missing or erroneous data due to known causes (e.g. during site maintenance):
    a. Will be interpolated if:
        i. missing or erroneous data was less than three consecutive hours in duration, and;
        ii. the trend (slope) of data during the hour preceding the questionable data is within 10% of that measured during the first hour after the questionable data.
    b. Will be removed from the data set if the duration of questionable data is longer than three consecutive hours or if the slope defined in 1.a.ii. exceeds 10% .

2. Correction of erroneous data due to sensor drift, fouling, or unknown origin:
    a. Will be conducted by interpolation if the erroneous data is:
        i. A single data point between two good data points, or;
        ii. Anomalous data consisting of two to three consecutive data points that exceed 2 times the criteria that result in a Poor quality ranking (see Section IIIb).
    b. Will otherwise be corrected utilizing the constant or variable correction approach (described below).
    c. If a malfunction occurs in up to two of the flowmeter's four thermistor pairs, flow direction and velocity will be calculated from the remaining thermistor pairs.

3. If data correction results in a change that exceeds the maximum allowable limit (Table 2), then that data will be removed from the data set. This does not apply to situations where the data is known to have been linearly shifted, such as through the application of incorrect offset values.

Table 1. Minimum criteria for correcting data. Data will be corrected only if values reported during field calibration checks and by the data logger during field checks exceed these limits.

Field Parameter	Minimum data correction criteria
Water Surface Elevation (WSE)	± 0.02 ft
Water Temperature[1]	± 0.2 °C
Specific Conductance[1]	± 5 µS/cm or ± 3%, whichever is greater[2]

[1] Based on USGS criteria (Wagner, et al., 2006).
[2] Suggested criteria for specific conductance may be modified after conducting a trend analysis conducted on several months of data.

Table 2. Maximum allowable limits for data correction.

Field Parameter	Maximum data correction criteria
Water Surface Elevation (WSE)	± 0.2 ft
Water Temperature[1]	± 2.0 °C
Specific Conductance[1]	± 30%

[1] Based on USGS criteria (Wagner, et al., 2006).

Interpolation is the process where missing or erroneous data are corrected by linearly interpolating values based on known good points before and after the questionable data. The use of interpolation will be limited to a maximum of three consecutive hours and will only be applied if the corrected data remains consistent with similar data from nearby gauges. The constant correction approach entails adding an unchanging modifier to all questionable data in a given time period. This approach would be used, for example, to correct water surface elevation data if the calibration was misread, or the stage device was not properly retuned to its pre-calibration level within the well. The variable correction approach uses a changing modifier to correct questionable data. This modifier normally starts at zero at the beginning of the time period and proportionally increases until it reaches its maximum value at the end of the time period. Variable corrections are typically used to correct for events that continue to aggregate through time, such as sensor drift and biofouling. If the start time of a variable correction cannot be specifically determined, the time of the last field calibration check will be used.

Data corrections discussed above are accomplished through the use of a GUI-based interface within the QAdjuster program. This program also stores metadata associated with the correction to the database, including time of correction, user, correction type, reason for correction, and numeric coefficients used to perform the correction. A report detailing all corrections applied to the dataset can be subsequently generated on request.

IIIb. Assignment of Quality Rating Codes.
Each data point will be assigned a quality code (Table 3) to rate the accuracy of the continuous reading relative to the field calibration checks performed in the field. There are four possible ratings (poor, fair, good, and excellent) and are assigned a numeric code of 2 to 5, respectively, in the database. For example, if the logged water surface elevation is +0.03 ft higher than the manual calibration check, then all data from the current point back to the last field check will: (a) be corrected as it exceeds the minimum criteria in Table 1, and; (b) be marked with a Fair QA code.

Table 3. Accuracy ratings of continuous water-quality records.

Field Parameter	Quality Assurance Accuracy Ratings[2]			
	Excellent(5)	**Good(4)**	**Fair(3)**	**Poor(2)**
Water Surface Elevation (WSE)	≤ ± 0.01 ft	> ± 0.01-0.02 ft	> ± 0.02-0.05 ft	> ± 0.05 ft
Water Temperature[1]	≤ ± 0.2 °C	> ± 0.2-0.5 °C	> ± 0.5-0.8 °C	> ± 0.8 °C
Specific Conductance[1]	≤ ± 3 %	> ± 3-10 %	> ± 10-15 %	> ± 15 %
Groundwater Direction & Velocity	4 thermistor pairs w/ stable baseline	3 thermistor pairs w/ stable baseline	2 thermistor pairs w/ stable baseline	2 thermistor pairs w Unstable baseline

[1] Based on USGS accuracy ratings (Wagner, et al., 2006).
[2] Numbers in parentheses refer to the numeric code assigned by QAdjuster. Internally, a code of 1 is used to denote suspect data for further evaluation and a code of 0 is assigned to data that has not yet undergone the QC process.

IIIc. Miscellaneous

The project website has a "Notes" section that will be used to record information pertinent to the data collection effort. All changes to the station's data logger programming as well as other general station maintenance are added as a "Maintenance" note type. Field site visit reports will be added as a "QA Field Report" note type. Other comments related to the quality assurance of the data will be added as a "QA" note type.

III. References

Wagner, R., R. Boulger, Jr, C. Oblinger, B. Smith. 2006. Guidelines and Standard Procedures for Continuous Water-Quality Monitors: Station Operation, Record Computation, and Data Reporting. U.S. Geological Survey Techniques and Methods 1-D3. Accessed March 3, 2010 at *http://pubs.water.usgs.gov/tm1d3*