QA/QC session, group 2, Christine Laney

QA/QC proceedures

Brainstorm procedures

Short term, streaming, real time

Removal / flagging of known wrong data

E.g. change of sensors, time until runs stable, cleaning cycles, etc.

Flag for technician notes

e.g. observations in the field of conditions that may have changed conditions.

Data type checking

Record completeness

Fill in missing records

No record if nothing is measured could be filled in with time stamps but empty values (or -99999). Can be flagged at beginning and end of continuous data.

Range checks

General removal/flagging of invalid data. Needs to make sure that unusual, but valid events don't get stripped.

Otherwise review flagging.

Ranges can be updated as parameter is being measured over time.

Long term min max, mean plus 2 or 3 sigma

Every site, every week/month/season

Bursts of measurements averaged and variance can be a measure of fouling

Variance

Variability of measurement over certain time, flag if variability changes.

e.g. sensor gets stuck on valid value, doesn't show any variability any more

sensitivity/precision standards

Spikes

Step

Missing data (large gaps)

Insufficient data for aggregation

Longer term checks after data have accumulated (higher level data quality, e.g. level 2)

Sensor drift

Drift correction if calibration comes back very different. Based on long term observation of how sensor drifts. May drift depending on temperature. Hardly any sensor will drift linearly.

Modeling of data

To see if this is a sensor problem or real data

Gap filling

Missing records may be filled in at this level – infill

Types of algorithms:

- Linear
- Averages (stationary of running)
- Cubic spline
- Comparable sensor

Aggregation

Data quality levels

See above

Data qualifiers

Level 1:

Avoid subjective flags (good, suspect) and use pass fail for certain tests. Plus technician documentation, but also decide for those whether the data are usable or not.

Level 2:

Gap filling – method used, drift estimates

Data Documentation

Great variability in approaches, better to document what was done

Documentation in text or automated system.

- Data processing workflow
- Data headers /variable definition,
- flag definition
- system requirements
- calibrations
- changes in sensors
- information about sensors
- expected file format/checksum
- Hardware configuration

Action items:

Book (wiki) of algorithms for data checking

Book (wiki) of algorithms for gap filling