Q/C Breakout Group 1 – Q/C recommendations

- Streaming data
  - NEON checks pretty complete, but subject to interpretation
    - Range min/max for sensor only, or seasonal/reasonable historic values?
    - Variance for single sensor stream or including related/redundant sensors?
  - Additional checks
    - Time step/frequency variations (clock issues)
    - Consistency of max/min, other derived values
    - Too long since calibration (depend on sensor DB)
  - Priorities
    - Must: sensor range, impossible values, timestamp inconsistency * clear errors
    - Should: highly dependent on sensor/property – should be based on best practices and domain expertise (knowledgebase)
  - Comments
    - Only logger missing value codes should be deleted from streams – even impossible values may have information
- Gap filling
  - Controversial – can seriously compromise stats, analyses and lead to misinterpretation
  - Desirable when generating summarized data, but transparency critical
  - Probably unsuitable for streaming data – much later in data cycle with expert attention

Q/C Breakout Group 1 – Qualifiers & Docs

- Qualifiers
  - Many vocabularies – desirable to harmonize, but impractical (crosswalk)
  - Good approach:
    - Rich vocabulary of fine-grained flags for streaming data – intended to guide review
    - Simpler vocabulary of flags for "final" data for public consumption
    - Different audiences may benefit from different flags
  - Certain types of qualifiers may be better as data columns
    - Method shifts, sensor shifts/sensor ids
- Documentation
  - Data level important to describe in metadata, because differences among programs
  - Q/C documentation should include methods, thresholds, assumptions
  - Most critical to document gap-filling, and flag all estimated values to allow removal
- Action items:
  - Best practices could be assembled by crowd-sourcing (wiki) gathering info on successful approaches and caveats
  - Develop tiers of flagging criteria for different classes of sensors, tied to data use