- QAQC working group 3
  - real time data
    - range checks
    - missing value codes
    - minimum could be nothing
  - issue: is it responsible to release unapproved data?
    - end users
      - technical and logistical use
    - can watermark graphs and not release data
    - people can pick numbers off scans
    - but we can't be responsible for how people might misuse data
      - our responsibilities are to tell them what we did to check
      - they are responsible for their use
  - good practice
    - varies depending on scale/number of measurements
    - some common ones
      - range/min/max
        - definition is subjective
        - can get from sensor manual
          - that deals with impossible, not unlikely
    - easy to use tools may help
  - levels?
    - work in context of individual systems - but do not necessarily apply across systems
    - usually
      - level 0 raw
      - level 1 calibrated
      - level 2 gap filled
    - can put things in metadata detailing
    - lots of domain choices
  - running averages - lengths
    - specific to datasets
    - can calibrate on existing data
  - redundant sensors would help
    - can use nearby stations
    - multiple loggers also help
    - 3 is ideal number
      - increases costs
      - don't all need to be of equal quality/resolution
    - also increases reliability
    - multiple sensors of different types
  - seven tests are pretty good but don't apply to everything
    - how conservative should you be?
      - may cause more harm than good if strip out interesting points
    - ultimately you are responsible for your choices
  - are methods for doing GENERIC QA - but no easy to use tools

- tools?
  - lots of candidates
  - Matlab/GCE toolbox
  - Could have some R libraries
  - statistical languages/programming languages
    - code can be put in metadata
  - kepler
    - can share agents
  - web service-based approaches
  - proprietary software
    - Campbell RTMC VistaVision
      - and other graphical software
    - LabView
  - good to have toolkit - that can be customized for data
- education and training
  - need training workshops
- gap filling
  - is it nobler to gap fill?
    - very useful but also dangerous
    - need to make clear what was done
  - use of nearby stations
  - lots of other potential models
  - from internal temp sensor could regress to main
    - also weir to wier
    - snow models
    - sometimes with boundary techniques
  - backfilling may be a SCIENCE problem - not a IM question
  - Borer ESA Bulletin paper
  - what would you want from a gap filling model
    - lots of tools - lots of approaches
    - if anything too many tools
  - documentation critical
    - should report level of uncertainty
      - can create "fake" gaps to test fit
      - also regression reports
- clearinghouse for tools and approaches would help
- will we ever be able to get away from provisional data?
  - even NEON will use human eyes as last test
  - timing depends on availability of human
  - maximizing automation can help
  - automated QC can AID human - guide human vision
  - human won't work for 20,000 datasets
- would be good to see some tools widely used.....
  - test them on exemplar datasets
    - take good data and corrupt it
    - would also be useful for testing

        - seven criteria are a good start
        - could be used with educational material in training using different tools
        - also exemplar data documentation
    - Action items
        - what should go in best practices document?
        - QA procedures applied must be documented -esp. if NONE
        - Tools - people are eager to use
            - I'd like to see some comparisons and then workshops to select
            - getting a list of on pager on tools
                - can do survey once you have short list
                - articles in DataBIts
            - example analyses in different tools
        - gap filling tools
            - proof is in pudding - test exemplar datasets using different methods
                - what is criteria for which is best
                - want UNCERTAINTY - confidence intervals
                - Mean square error
                - could be a publication comparing methods
            - ameriflux network might be resource - they have done alot
                - they use "gold standard" dataset
                - they published in agricultural meteorology