Breakout Session Report
ARM/ASR User and PI Meeting
March 16-20, 2015

Session Title: Assessing and Communicating Data Quality and Instrument Uncertainties
Session Date: Thursday, March 19, 2015
Session Time: 10:30am
Summary Authors: Jennifer Comstock, Ken Kehoe, Mike Jensen, and Shaocheng Xie

Description

ARM has a diverse suite of instruments that are used for analyses, retrieval development, and model evaluation studies. These research activities require easy access to quantitative assessment of measurement uncertainty and overall data quality. This session will be centered on discussion of best practices in both deriving and communicating measurement uncertainties and data quality. The discussion will also focus on prioritizing measurement uncertainty assessment and Value Added Product (VAP) development against alternate activities such as data visualization and analysis tools and instrument simulator development. Important user input from this session will feed directly into facility-wide prioritization efforts for the ARM program.

Main Discussion

ARM has a diverse suite of instruments that are used for analyses, retrieval development, and model evaluation studies. These research activities require easy access to quantitative assessment of measurement uncertainty and overall data quality. The objective of this session was to engage in a discussion of best practices in both deriving and communicating measurement uncertainties and data quality. Three invited talks were presented to help guide the discussion. The first presentation, led by Ken Kehoe, provided an overview of current practices within ARM for assessing and communicating data quality and uncertainty. The second presentation, led by Mike Jensen, provided an overview of current activities surrounding the assessment of measurement uncertainties for ARM instruments. The final presentation by Shaocheng Xie discussed modelers needs related to data quality and uncertainty quantification. The session concluded with discussion regarding prioritizing measurement uncertainty assessment and VAP development against alternate activities such as data visualization and analysis tools and instrument simulator development. Important user input from this session will feed directly into facility-wide prioritization efforts for the ARM program.

Key Findings

This session had extensive discussion during each presentation. Key discussion points and suggestions are listed according to topic.

Data Quality

To make data quality reports easier to access, have user/DQ notes, technical handbooks, and known issues linked to the data discovery tool. It was noted that not all VAPs currently have technical reports and not all datastreams have instrument handbooks.

Data sets should have a short summary of known issues with datastreams, clear definitions of when to trust/not trust data sets. This summary should be viewable when users are shopping for data. For retrievals, information in files about why retrievals fail is also valued.

Users are still interested in having recommended data sets.

There are multiple perspectives on data quality. For example, when performing a retrieval of a parameter—say a cloud microphysical property—if no value is reported, we need to be clear if there is no value because there is no cloud, if there was no retrieval attempted, or if there was a retrieval attempted but it was not successful. This information is critical for interpreting a datastream.

When a parameter is reprocessed, it would be very helpful to know how much the parameter changed.

Uncertainty Quantification (UQ)

Requirements for UQ include systematic and random errors on measured variables and to report these errors independently. These types of errors are needed to feed into optimal estimation retrieval frameworks. It was noted that some random error can be reduced through averaging. Other types of error cannot. Related to this – it is important to appropriately scale errors. For example, errors may be reduced when averaging to global climate model (GCM) temporal scales.

The Kalman filter framework applied by Christine Chiu was suggested as a potential method for applying UQ to retrievals.

Modelers need error distributions on retrieved quantities and diagnostic tools (i.e. NCAR AMWG) to help evaluate model simulations.

Uncertainty estimates for aircraft instruments are desired. Aircraft probe measurements are often considered direct measurements but it was pointed out that probes require retrievals to derive quantities such as particle size or liquid water content from optical or voltage based (i.e. hotwire) measurements.

Prioritization

A list of primary measurements and retrievals were provided to provoke discussion on the priorities. There was some consensus that a focus on data quality and UQ was desired over new VAP development. A small set of parameters was suggested as a start for uncertainty characterization (T, RH, U, precipitation rate for a specific set of datastreams). There seemed to be acceptance that this was a reasonable place to start. It was suggested that when thinking about the next set—we should consider the frequency of downloads from the archive.

Issues

The main issues identified were:

1. How should the quality of the data be defined?
2. How should data quality be related to uncertainty?
3. What criteria or requirements should be used to develop uncertainty estimates?
4. Providing the uncertainty in a standardized machine readable format.

Future Plans

There is a clear need to better define what the requirements are for data quality and uncertainty. Instrument mentors and VAP developers will need guidance regarding methodology and specific requirements, which will likely vary with measurement and retrieval techniques. In the near future, and implementation plan will need to be developed to help guide these efforts.

Action Items

- Identify and implement specific improvements for data quality communication including review of current documentation (e.g. technical reports and the process for keeping them up to date).
- Develop guidelines or requirements for applying UQ.
- Develop an implementation plan for an initial set of measurements/variables that are required for the high resolution modeling efforts.
- Develop a prioritized list of VAP development efforts and identify VAPs to begin applying UQ methods.
- Develop a standard for communicating data uncertainty.