The ARM Program Data Quality Office - A New Approach for Coordinating Data Quality Efforts

R. A. Peppler, K. L. Sonntag, and A. R. Dean
Cooperative Institute for Mesoscale Meteorological Studies
University of Oklahoma
Norman, Oklahoma

Introduction

The Atmospheric Radiation Measurement (ARM) Program Data Quality (DQ) Office was established at the University of Oklahoma in July 2000 to coordinate the continued development and implementation of a program to ensure the quality of data collected by ARM's field instrumentation.

The DQ Office has the ultimate responsibility within ARM to provide the results of DQ inspections to 1) data users so that they may make informed decisions when using ARM data in their research and to 2) ARM's site operators so that they can quickly implement corrective maintenance and troubleshooting actions to minimize the collection of unacceptable data. The results of ARM's DQ inspections are now made available to data users via the Meta Data Navigator (MDN).

The DQ Office has other responsibilities as well. These include ensuring that documentation of ARM's DQ is complete and readily available to all, participation in (with Instrument Mentors, Site Scientists, and others in the ARM Program) the process of inspecting and assessing ARM's data streams and in writing and submitting appropriate Data Quality Reports (DQRs), and working closely with relevant groups within ARM to solve issues and problems pertinent to the DQ effort. This paper describes some of the DQ activities underway in the ARM Program.

Provide the Results of Data Quality Inspections to Data Users

The Meta Data Navigator is ARM's graphical user interface for viewing both the quality assessment assigned to data and related data quicklooks. It is also a place where those involved in ARM Program DQ assessment activities can submit DQRs and thereby assign "color" quality attributes to the data stream variables. The MDN is found at http://prod.arm.gov/mdn/

The tasks of 1) incorporating what is known about data availability and quality and 2) linking this knowledge directly to individual measurements in a way that is easily identifiable by data users, such as by data flagging and metadata, were accomplished through development of the MDN. Within the MDN, a communication interface was established by color-coding data stream variables obtained from a

particular instrument with different colors indicating the availability and quality of the data on a daily aggregate scale. Pointers direct the data user to the DQ inspection method(s) used to color code each measurement. The color rules adopted include:

- White - data exist but have not been checked for quality
- Black - data are missing or do not exist
- Green - data have been reviewed for quality by various methods and are judged to be acceptable for use
- Yellow - data have been reviewed for quality by various methods and are judged to be suspect or questionable - use at your own risk
- Red - data have been reviewed for quality by various methods and are judged to be of poor enough quality to be deemed unusable

The color of the data may change over time to reflect the most current view of their quality. The entire process is automated but includes a manual override capability.

Determination of color is ultimately made by some combination of input provided by automated flagging contained within netCDF files and by manual or automated checks made by Instrument Mentors, Site Scientists, and the DQ Office, as reported in DQRs. A system is now being discussed to allow delivery of color flags seen on the MDN display, and descriptive DQR-type information, inside of a data file.

Provide the Results of Data Quality Inspections to ARM Site Operators

The DQ Office works closely with the Instrument Team and Site Scientists to immediately report irregularities in data quality to Site Operators so that corrective maintenance and troubleshooting procedures can begin. A Web interface called the Data Quality Problem Report (DQPR) was established at the Southern Great Plains (SGP) to provide a common mechanism for inputting problems and better tracking them. A similar tool called the Quality Assessment Record was developed for the North Slope of Alaska (NSA). The DQPR can found in OMIS at http://www.ops.sgp.arm.gov:591/dqpr1/default.htm. Shown below is a sample SGP DQPR entry:

DQPR#	Date	Originator	Instrument	Site	Problem
00182	12/11/2000	Karen Sonntag	Surface Meteorological Observating Station (SMOS)	E20	The wind vane appears frozen. The SMOS Current Conditions Web page indicates a southwest wind at E20. The wind direction at surrounding sites is from the northwest. Suggest PM crew take a look when on site this week.

Improve and Complete Documentation on Data Quality

The DQ Office has worked closely with the Instrument Team on a reorganization and augmentation of the Instrument Web pages to make them a true reference guide for ARM's instruments. A fundamental of quality control is a "statement of expectations." Quality is the measure of how closely something conforms to an expectation. Without an expectation, a quality assessment is not possible. Thus, the instrument pages represent ARM's statement of expectations for its instruments, the baseline against which the observations can be compared. These Web pages contain our incremental understanding of the measurement systems and their quirks and deficiencies, including common problems that have been encountered or are inherent to the measurement.

This work has seen the development of a DQ section that heretofore did not exist. This DQ section contains the following information:

- Data User Notes
- Automated Quality Control/flagging contained within netCDF files
- Instrument Mentor Quality Control Checks
- Site Scientist/DQ Office Quality Control Checks
- Value Added Procedures
- Quality Measurement Experiments
- Examples of Data
- Data Quicklooks/near real-time
- Current Health and Status

These pages represent the place for you to find specific information about the DQ for every ARM instrument. They can be found at http://www.arm.gov/docs/instruments.html

Develop a Data Quality Home Page for ARM

A home page describing ARM's DQ Program was developed and made available to the public in February 2001. It can be found at http://www.arm.gov/docs/data/dqhome.html

This page is the place where viewers can learn general information about ARM's DQ program. The site is accessible from an ARM DQ Program button located on the blue ARM menu bar. This page also refers viewers to the Instrument Web pages for the quality control (QC) details of specific instruments. With the information contained on the Instrument Web pages, the new DQ home page represents a QC Guide for the ARM Program.

Upgrade, Modernize, and Participate in Quality Inspections of ARM Data

With the SGP Site Scientist Team, the DQ Office is helping to upgrade and automate all SGP Site Scientist QC algorithms and plotting routines. This work so far includes Solar Infrared Station, SMOS, Temperature, Humidity, Winds, and Pressure System, Microwave Radiometer, Soil Water and Temperature System, Energy Balance Bowen Ratio, and Baseline Surface Radiation Network. Metrics programs now run hourly, and diagnostic plots are created daily. Reports are issued weekly.

With this modernization, the SGP Site Scientist is transitioning some responsibility for routine SGP QC checking to the DQ Office. The DQ Office participates in the process to submit DQRs when appropriate.

The DQ Office is in the initial phases of assuming the responsibility for routine QC checking of Tropical Western Pacific (TWP) data. The NSA Site Scientist Team continues to perform NSA QC activities.

Development of Health and Status Web Sites Based on Data Quality Results

With the help of the DQ Office, Chad Bahrmann of the SGP Site Scientist Team, is developing a health and status Website (updated hourly) for the SGP that provides a color table and diagnostic plots; it is based on QC results and is similar in look and feel to operationally-oriented sites for the TWP and NSA. The DQ Office will continue to assist in the development of the SGP site and participate in the augmentation of the content on the other sites. It is desired to have the same sorts of checks and diagnostic plots fueling all three health and status sites. The SGP site can be seen at http://r1.sgp.arm.gov/~sgpdq/

Work in Progress or Under Discussion

- Improvement of the MDN's DQR entry tool
- Extraction of color attributes from automated QC now within data files
- Delivery of color and "DQR" flags to data users within data files
- Development of additional automated QC flags within data files
- Creation of Quality Assessment Record-like databases for all Cloud and Radiation Testbed (CART) sites
- Re-establishment of routine quality inspections of TWP data
- Completion of Instrument Web pages and DQ documentation
- Development or augmentation of quality-based health and status pages for all CART sites
- Development of tools to look at long-term trends and changes in DQ

Corresponding Author

Randy A. Peppler, (405) 325-6667, randy.peppler@arm.gov