

Chapter 12

The ARM Data Quality Program

RANDY A. PEPPLER, KENNETH E. KEHOE, JUSTIN W. MONROE, AND ADAM K. THEISEN

Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, Norman, Oklahoma

SEAN T. MOORE

Orbital ATK Inc., San Francisco, California

1. Introduction

As of this writing, nearly 7000 ARM Climate Research Facility data fields from 400 instruments are monitored for data quality control on a daily basis. This chapter reviews the history and evolution of ARM Program data quality assurance since the beginning of the program and describes the processes in place today. It also provides advice to those who collect field data, especially in an operational context. ARM's infrastructure was charged to produce data of "known and reasonable quality" for use by climate researchers. This is challenged by the fact that there are hundreds of different instruments of varying types in different climatic locations, translating into thousands of individual data (variable) streams. Some of these variables are geophysical variables (or will be processed by some algorithm to be such), but many are intended purely to help characterize the state of the instrument that made them (e.g., instrument temperature). The goal of the data quality program is to assess the quality of all of these variables.

To better complete this data quality mission, an ARM Data Quality Office (DQO) was formed in July 2000 to provide overall guidance and management of a program to assure that the data collected at ARM sites meet the data quality objectives and tolerances as defined by the science user community and to make estimates of that assurance publicly available. The DQO is accountable to the ARM Technical Director and works daily with the

ARM Infrastructure, Atmospheric Systems Research (ASR) science team members, and the broader ARM user community to develop an end-to-end data quality assurance system that results in continuous, consistent quantitative assessment and continual improvement of ARM data streams through improved instrument performance based on what has been learned. The DQO leads the development and implementation of data quality algorithms and visualizations, analysis of results, and the reporting of the results both to the program and to the scientific community. It is responsible for achieving efficiencies within instrument suites and across collection sites with respect to data-checking algorithms, metadata collection, and data quality reporting. It works closely with instrument mentors, site scientists, site operators, and data and engineering staff to develop the data quality tools and analyses needed.

2. History and evolution of ARM data quality inspection and assessment

a. Early programmatic efforts

The reader is referred to [Peppler et al. \(2008a\)](#) for more detail on ARM's data quality assessment history. Early programmatic efforts in data quality inspection and assessment focused on the first field site, the Southern Great Plains (SGP). These efforts included the development of self-consistency checks for individual data streams ([Blough 1992](#)) and quality measurement experiments (QME; [Miller et al. 1994](#)) for comparing multiple data streams. Self-consistency checking involved not only simple range and rate-of-change tests, but also automated statistical assessment of individual

Corresponding author address: Randy A. Peppler, CIMMS, University of Oklahoma, 120 David L. Boren Blvd., Norman, OK 73072. E-mail: rpeppler@ou.edu

data streams for internal anomalies—this was done both to detect outliers and to identify instrument failure. In each case, flags were created to notify instrument operators and data users of the issues. Some statistical assessment was accomplished using a Bayesian dynamic linear model. Early applications of these checks were made for the detection of moisture on radiometer domes, and for the detection of signal attenuation, side-lobe leakage, presence of birds, and other interference on wind profilers.

A QME concept was developed at the beginning of the program to compare multiple data streams against a set of expected outcomes of the comparison, including experimental hypothesis. The multiple data streams that served as QME input included direct observations from instruments, measurements derived from multiple instrument observations and the subsequent application of algorithms to them, and model output. The idea behind this concept was that comparisons involving multiple data streams should reveal more information about quality than single data stream self-consistency checks could allow. As such, a major function of the automated QMEs was to identify data anomalies in near-real time and to help data quality analysts identify the root cause of unusual behavior. The measurements produced by the QME were treated as official data products and were archived. An early QME example compared vertically integrated water vapor from microwave radiometers with the output of a microwave radiometer instrument performance model that used thermodynamic profiles from radiosondes to drive the model. Another early QME made hourly comparisons between infrared spectral radiances observed by a Fourier transform interferometer and the output of a line-by-line radiative transfer model (Turner et al. 2004).

Substantial effort also was made early in the program toward day-to-day data quality assurance by instrument mentors (Stokes 2016, chapter 2; Cress and Sisterson 2016, chapter 5). Instrument mentors played and continue to play a vital role by 1) independently monitoring the data produced by their assigned instruments using various analytical and interpretive techniques and 2) reporting their findings on potential problems, suggesting solutions to site operators, and actively participating in the problem-resolution process. Instrument mentors were and continue to be a first line of defense in data quality assessment and problem diagnosis and solution. During the 1990s, instrument mentor and site data quality efforts often were independent and sometimes duplicative.

b. Southern Great Plains site efforts

Site scientists for the SGP site (Sisterson et al. 2016, chapter 6) at the University of Oklahoma assisted instrument mentors by developing methods to facilitate

the graphical, automated display of data and within-file limit checks (Splitt 1996; Peppler and Splitt 1997). The idea behind these diagnostics was to make them available for viewing by instrument mentors and site operators on the web within 2 days of data ingest, regardless of the physical location of the sites or the data viewer someone used. Among the earliest diagnostic tools developed by the SGP site scientists were comparisons between hemispheric broadband solar irradiances and modeled clear-sky estimates, and respective comparisons of shortwave albedo estimates and broadband longwave observations from multiple SGP collection sites. Interpretive guidance was developed by instrument mentors to aid site scientists in evaluating plots, and the site scientists developed an e-mail reporting system for alerting site operators and instrument mentors in near-real time about possible problems. In time, the volume of data collected at SGP sites caught up with the capabilities of the SGP site scientists to review them in a timely manner, which spurred more attempts at automation that came to fruition once the DQO was formed. Scientists under contract to ARM at Mission Research Corporation also led early efforts to create automated diagnostic algorithms to evaluate some data streams.

c. Tropical Western Pacific site efforts

Efforts to display and assess data collected at Tropical Western Pacific (TWP) sites (Long et al. 2016, chapter 7) were undertaken initially by site scientists at Pennsylvania State University (PSU) with less instrument mentor involvement after initial siting and operation. This was partly because of the remote TWP locale presenting unique communication complications. Its first site installed at Manus Island in October 1996 included the core instrumentation found at the SGP Central Facility site, but unlike SGP collection sites in Oklahoma and Kansas, the extremely limited bandwidth of the network connection between Manus and ARM's Data Management Facility (DMF) led to delayed data delivery in the early years.

During this period, data examination by site scientists occurred in two stages. The first stage identified potential instrument and site maintenance issues and was directed toward on-site operations staff. The second stage involved a more detailed review of the data and was directed toward the science user community. To address the operations requirement, a compact data status message was constructed that included hourly statistics from most of the instruments along with environmental parameters such as the temperatures of instrument enclosures. These messages were sent via the Geostationary Operational Environmental Satellites (GOES) link each hour. Each day, plots of these hourly data were generated and posted on a website at PSU. Initially, this process was carried out by the site scientists but eventually

transitioned to site operators. The plots proved useful for identifying gross errors in the data, which were fed back to site operations' technicians, allowing them to plan on-site repair visits.

Once a full TWP dataset was delivered to the DMF, the site scientists produced daily data graphics and performed diagnostic tests. Such tests included the closure of the solar direct and diffuse components, net radiative flux, and comparison of integrated radiosonde water vapor with the water vapor derived from a microwave radiometer. Data gaps also were cataloged, which led to the uncovering of problems with dataloggers. After the full dataset arrived and data were examined, the site scientist produced a report describing issues with the data, and at that point the data were released to the public. This report was submitted to the DMF, though there was not a mechanism at the time to include all of its information to data users. Most of the information was subsequently converted to forms suitable for wide distribution (explanatory text, tables, or figures). The procedure for reviewing the data prior to their release ended about the time the second TWP site was installed at Nauru in November 1998. With two sites running, it became impractical to review all of the data prior to release.

As seen during this early phase of operation at the tropical sites, the TWP site scientist took the lead role in the examination of the data collected. When a question regarding a specific instrument arose, the site scientist typically contacted the appropriate instrument mentor and worked with the mentor to solve the problem. If, on the other hand, the source of a problem was already known or suspected by the site scientists, they contacted site operators directly to work on resolution. This model of putting the site scientist at the front line of data review had distinct advantages as well as disadvantages. An advantage was that site scientists have a vested interest in the instruments at the site, and looking at multiple instrument data streams provided a holistic view of site performance that was useful for problem solving. However, this system was inefficient and time consuming, limiting site scientist time for activities such as promotion of the data to the scientific community and planning and implementing TWP field campaigns. With the establishment of DQO, the role of TWP site scientist in routine data review gradually but dramatically changed.

d. North Slope of Alaska site efforts

At the North Slope of Alaska (NSA) site (Verlinde et al. 2016, chapter 8), still another model for data quality assurance was used. Site scientists and site operators jointly subjected data to a systematic program of quality checks (e.g., Delamere et al. 1999). Data streams were visually inspected on a daily basis; from these

visual inspections, metadata documenting the overall quality of the data streams were generated. Such inspections facilitated detection of instrument malfunction at the Barrow site as it was spinning up in 1998. A web-based archive of graphical images was developed and maintained to facilitate visual inspection by NSA site scientists at the Geophysical Institute of the University of Alaska, Fairbanks. These graphics were updated and made available daily. In addition to visual inspections, site scientists interpreted limits that were applied to the data. Instrument mentors had relatively little involvement in data quality assurance activities at this site after instrument installation and official data release. NSA site scientists and operators played a crucial role in discussions during 1999–2000 on how to better automate data quality checking at the site and across ARM, and how to place this information and other metadata both within data files and on the web.

e. Efforts after the World Wide Web

It should be noted that the evolution of the web, which took place during the first few years of the ARM Program, was a transformation point, especially for data quality efforts. Use of the web began as a grassroots effort and grew quickly, especially as browser technology and the Internet evolved. However, it was unevenly adopted by the three sites, which was another reason for uneven data quality treatment, especially with respect to data quality reporting and data quality coordination between different parts of the ARM infrastructure. Site scientists at the University of Oklahoma, ARM scientists at Lawrence Livermore National Laboratory, and contracted scientists at Mission Research Corporation were among the first in the program to embrace the web and develop automated algorithms that generated quick-look images that were published on the web for instrument mentors, site scientists, and others to review. There was some contention at the time about the eventual role of the web in ARM infrastructure efforts, but ultimately the use of the web for the data quality program in particular became a cornerstone of the effort.

f. Establishment of the Data Quality Office and beyond

As described above, while site-based instrument mentor, site scientist, and site operator efforts were crucial for detecting instrument malfunction and minimizing the amount of poor data collected, these efforts were unevenly developed and applied across the sites, and oftentimes were independent of each other and duplicative. This often led to varying treatments of like measurements taken at different locations, leading to uneven data quality reporting and resulting in uneven

levels of data user confidence. A key finding of the ARM Program Infrastructure Review, conducted in summer 1999, stated that “a primary mission of the ARM Infrastructure is to produce a ‘legacy data set’ that is invaluable for research on global change. We are particularly concerned about the coordination and completeness of the quality assurance information describing ARM data” ([ARM Program Infrastructure Review Committee 2001](#), p. 2). The review recommended that the program’s data quality activities should be consolidated and coordinated, and it recommended the creation of a new position of “Data Quality Manager.” This recommendation evolved in July 2000 into the establishment of the DQO at the University of Oklahoma, as described at the beginning of this chapter.

The DQO has since coordinated the program’s data quality assurance activities, in continued close consultation and participation with instrument mentors, site scientists, site operators, and others. The DQO incorporated the best practices from past site efforts to create what is seen today. From the SGP, it incorporated ideas regarding self-consistency checks for individual data streams and QMEs, data intercomparisons beginning with radiation measurements, and initial ideas for web display of diagnostic plotting and display of within-file quality control limits. These actions led to the framework for a web-based display system that began as a way to compare radiation measurements and analyze soil water and temperature sensors ([Bahrmann and Schneider 1999](#)). From the TWP, the DQO took ideas for more sophisticated data plotting, including measurement intercomparisons, and the relationships it established between the site scientists and instrument mentors. The TWP site also was the first site to interact directly with data users, something that the DQO has tried to emulate, and was the first site to coordinate data quality assurance as a whole site as opposed to individual instrument mentors distributing their own assessments. From the NSA, the DQO incorporated its systematic regimen of integrating quality control checks within a file and metadata organization structure, which was novel and elements of which are used throughout different parts of the ARM infrastructure. The NSA also was an early participant in creating a web repository for graphical products. And, as described above, individuals at the University of Oklahoma and Lawrence Livermore National Laboratory were the first to bring ARM into the web age (mid-1990s) by creating web repositories of “quicklooks” in thumbnail tabular form that allowed full graphical viewing once selected. This model, novel for its time as the powers of the web were being discovered, has been emulated countless times by those both inside and outside of the ARM Program.

Instrument mentors, site scientists, and site operators still retain strong, complementary roles in the quality assurance process. These roles are embodied through problem discovery and resolution efforts, and various weekly coordination teleconferences scheduled to discuss and resolve pressing data quality issues. Instrument mentors, as the technical authorities for their instruments, continue to provide an in-depth instrument-specific perspective on data quality, responsibility for helping resolve problems, and expert help in identifying problematic long-term data trends. They also are the final arbiters of data quality to the public as embodied in their writing of data quality reports. Site scientists, as authorities on their locale and its scientific mission, provide a broad perspective on data quality spanning the full range of site instrumentation. They also help oversee their site’s problem-resolution process and perform targeted research on topics related to site data quality issues. Site scientists interact directly with the scientific community to plan and conduct field campaigns at their sites, which have at times identified previously unknown data quality issues (see below). Site operators implement the problem-resolution process by orchestrating and conducting the corrective maintenance actions requested. They are key in ensuring the smooth, routine operation of the sites and their instruments through regular preventative maintenance and the application of periodic instrument mentor-specified calibration checks.

The next section describes the workflow of the data quality review process as it exists today, with an emphasis on what the DQO does. This process involves creation of data plots and displays of within-file data quality information as data are collected and ingested by ARM, routine data inspection, assessment, and status reporting by data quality analysts, problem reporting, problem resolution, and finally communication of data quality results to data users. Consultation of long-term data trends (to put current measurements in context), maintenance and calibration reports, and the development of data quality documentation (as interesting issues are discovered) to serve as pattern recognition are part of this process.

3. Data stream inspection and assessment, problem reporting and resolution, and reprocessing

We refer readers to [Peppler et al. \(2008b\)](#) for more details on the processes described here. Given the data volume described at the beginning of this chapter, data inspection and assessment activities must be automated and efficient, although human inspection of the results still remains a high priority.

The quality assurance model has three components. The first component is a “rapid evaluation and response” piece involving data inspection and assessment that is designed to identify gross and some more subtle issues within the data streams as fast as possible and relay that information to site operators and the instrument mentors so that the (potential) problem-resolution process can begin. The goal of this component is to minimize the amount of data that is affected by the problem. The second component involves documenting and reporting data quality issues for the scientific user; this is primarily done via text-based but machine-readable data quality reports (see below). The third component involves reprocessing of data after known problems have been identified and solved to provide end users with the best products available.

a. Inspection and assessment

The main objective of near-real-time data inspection and assessment is to quickly identify data issues and report them to instrument mentors, site scientists, and site operators so that corrective maintenance actions can be scheduled and performed, limiting the amount of unacceptable data collected. Data quality analysts at the DQO perform much of the routine data inspection, assessment, and initial problem reporting on a daily to weekly basis. This analysis is conducted not only by DQO full-time staff but also by University of Oklahoma School of Meteorology undergraduate student employees who have an interest in meteorological observations and instrumentation. These student analysts have been paramount to the DQO’s success over the years, and many have gone on to graduate school and, in some cases, became faculty at other institutions of higher education. This tasking allows full-time DQO staff to spend more time on the development of data quality checking algorithms, an activity that is done in coordination with the technical guidance of instrument mentors and site scientists. This activity has resulted in the development of a broad suite of automated tools and procedures packaged into a web-based system (<http://dq.arm.gov/dq-explorer/cgi-bin/main>), an evolution of a forerunner system described in [Peppler et al. \(2005\)](#). The ARM network configuration provides the DQO with the computing power and file services (both at the DMF) needed to facilitate data quality algorithm processing. As mentioned earlier, a system prototype was created in the late 1990s by SGP site scientists as a way to monitor solar trackers for radiometers, and was later expanded to monitor the then-new soil water and temperature system. After formation of the DQO, that system was formalized into a program-wide, web-based data quality tool that has been modernized over time.

Inspection and assessment are accomplished in a three-tier process. The first tier is the application of simple consistency checks like minimum, maximum, and delta (a comparison of consecutive values collected to detect abrupt changes) checks, and whether data exists. The second tier relates to the question “do the data make sense” for the current meteorological setting (e.g., does the temperature follow the normal diurnal cycle, do cloud base height estimates from different instruments agree to within some nominal bounds, and are there artifacts in the data that are nonphysical from a meteorological perspective?). The third tier is a more advanced statistical analysis using techniques and longer time series data that can help find more subtle problems (see [section 6](#)). This last tier can be quite powerful, but care has to be taken not to flag real signals that might be outside of a 3-sigma limit, for example.

As a process, every hour the latest available data collected by fielded instruments are ingested by the DMF, processed by the DQO algorithms, and displayed in a web-based system. The DQO processing creates a graphical summary of quality control fields (flags) within each file, a graphical summary of additional DQO-generated quality control tests, and a suite of graphical depictions of data ([Fig. 12-1](#)). Student data quality analysts primarily access the system by selecting the site, data stream, and date range of interest for each data product in their purview. Prior to final submission, a color-coded request queue alerts the analyst if the dates for the data product(s) in their submission have not been processed by the DQO. A color table then is displayed showing hourly flagging summaries for each measurement (both in-file testing and additional tests performed by the DQO). This helps an analyst to quickly screen potential problem areas. All tables and graphics are updated hourly as data arrive to the DMF.

Data quality analysts visually inspect all flagging results and data graphics. Diagnostic plots, including cross-instrument comparisons that in many ways mimic the early QMEs, display daily but are updated hourly. Inspection of these plots of primary and diagnostic variables helps identify data abnormalities not always detected through automated flagging or analysis of a data stream in isolation. A succession of daily diagnostic plots in thumbnail form, which can illustrate trends in data, is available as well (<http://plot.dmf.arm.gov/plotbrowser/>; [Fig. 12-2](#)). Analysts may select a site, one or more data streams, and a date to view thumbnails of data plots for up to 30 days at a time. The thumbnail format facilitates comparison of different instruments that measure like quantities and can provide a view of near-term trends. The example in [Fig. 12-2](#) highlights a

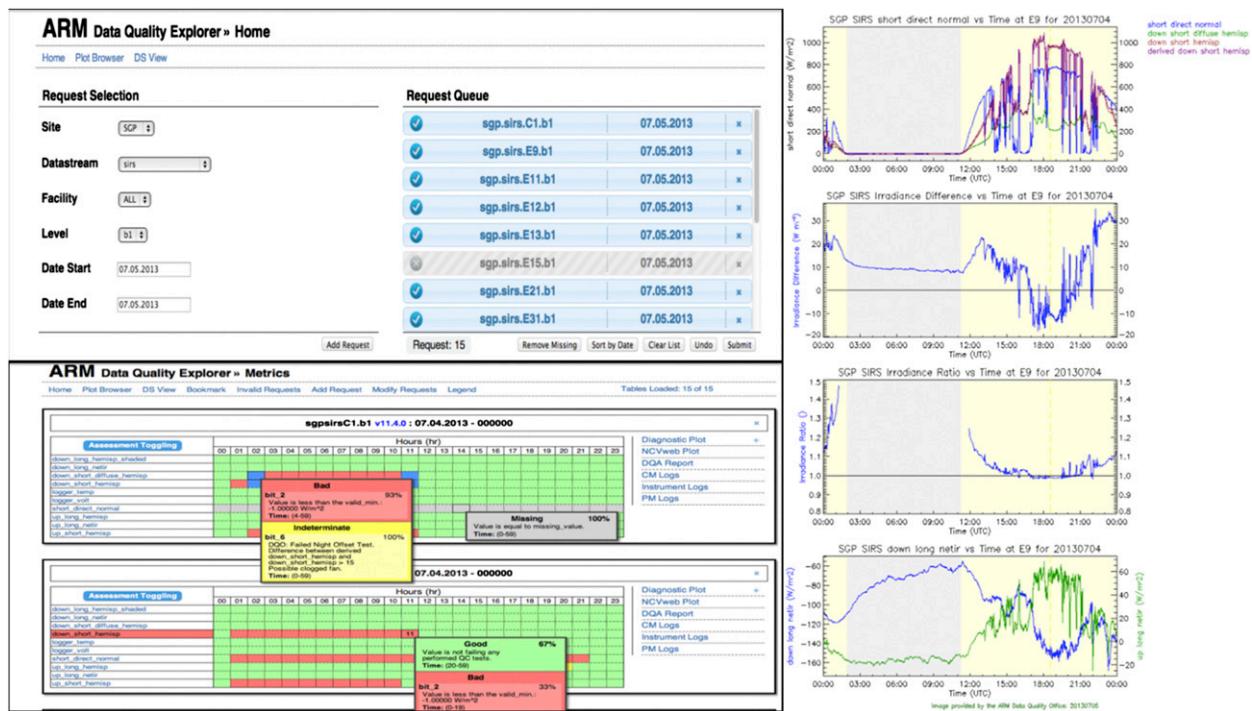


FIG. 12-1. (top left) Display showing initial data selection interface, (bottom left) sample hourly metrics table overview with associated error messages, and (right) related diagnostics plots.

case in the top panels where an unanticipated data transition takes place with 915-MHz Radar Wind Profiler precipitation consensus data, while the bottom panels as a comparison show data from another site for the same type of profiler that do not show the transition.

Another plotting capability provides the analyst with an interactive, web-based tool (Fig. 12-3). Key features include focusing on data periods of less than or more than one day, and particular data fields of interest can be specified from pulldown menus. Plots may compare data from multiple facilities, show comparisons to reference measurements, show slices of multispectral data such as atmospheric emitted radiance interferometer (AERI) radiances, or even display color-filled contours of radar reflectivity and lidar backscatter. For closer inspection, data values can be displayed in tabular form or downloaded in an ASCII comma-delimited format for easy importation into spreadsheet applications. Analysts can view file headers to obtain direct access to metadata or a summary of data field descriptions and basic data statistics. In the example shown in Fig. 12-3, temperatures are plotted for all SGP surface meteorological (MET) instrumentation sites on the same graph. In this case, a problem with the temperature sensor at the E33 facility can easily be detected.

Finally, another utility can be used to quickly create large batches of diagnostic plots that provide a detailed view of the within-file quality control describing all relevant fields in a data stream, as well as basic statistics about the data. As shown in Fig. 12-4, the top panel of the default plot for one-dimensional time series includes the data values color coded by the assessment level, while the bottom panel provides a color-coded view of the descriptions of the individual quality control (tests) applied to the data in the top panel. This example shows a plot of aerosol optical depth from the aerosol optical depth value-added product (VAP; Koontz et al. 2013). Green indicates that no tests were failed and the data are “good,” yellow indicates that the data failed a test with an assessment level of “indeterminate,” and red indicates that the data failed a test with an assessment level of “bad.” This utility also provides a number of command-line options for customizing plot generation, such as the ability to plot all fields in a file automatically, plot along a specific coordinate in a two- or three-dimensional field, plot a two-dimensional field as multiple stacked line plots, and plot short time periods covering a few hours or long time periods covering multiple years. It has proven very useful during the evaluation of ARM data products under development, and it often allows the DQO to quickly detect errors in the automated quality control algorithms.

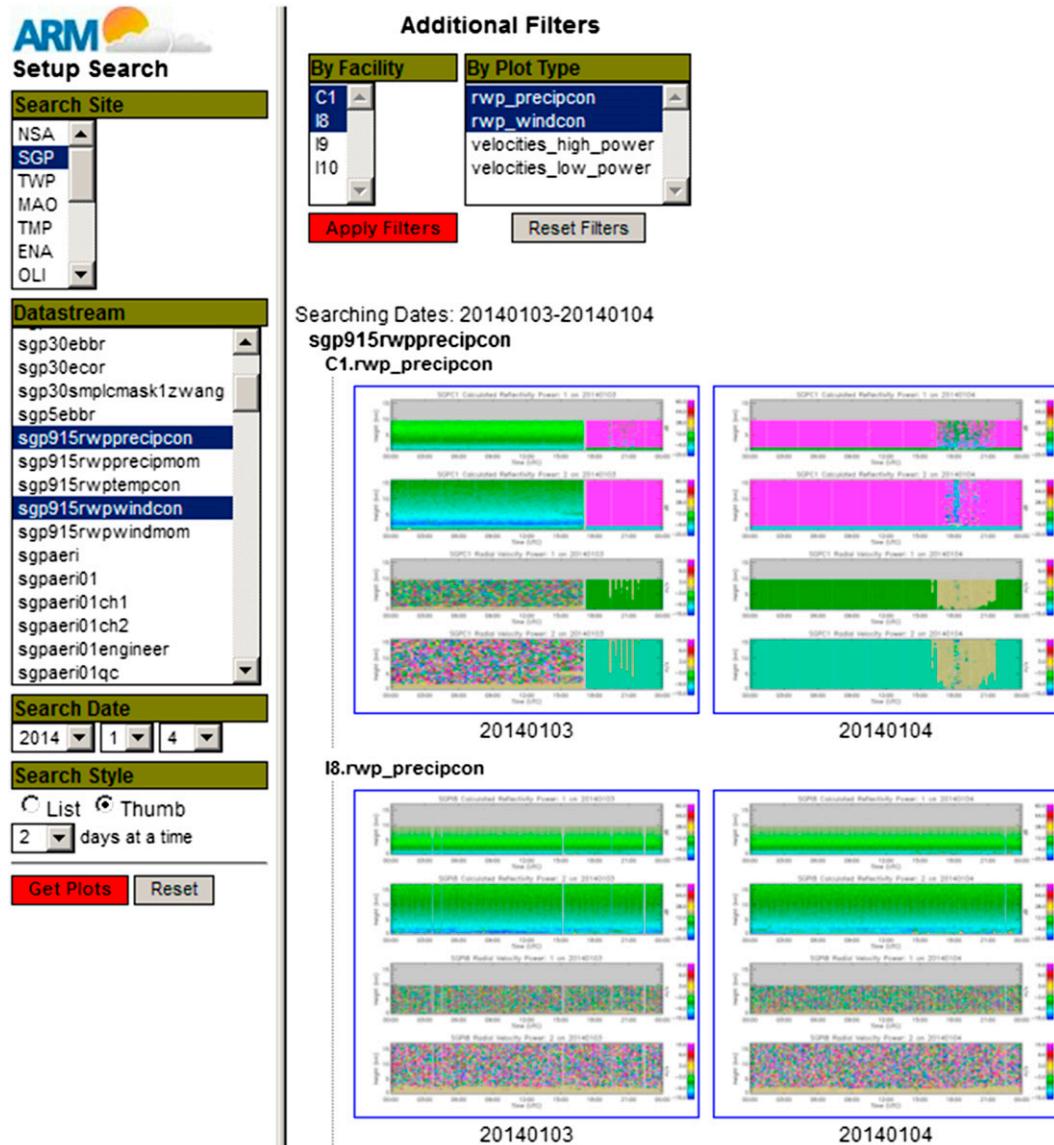


FIG. 12-2. Plot browser display showing multiday thumbnails from the SGP site. Shown are (top) two plots of 915-MHz Radar Wind Profiler (RWP) precipitation consensus data for a site showing an unanticipated data transition later in the first day that continues into the next day and (bottom) two plots from another site for the same type of profiler on the same day that does not show the transition. Clicking on a thumbnail brings up a full-resolution image.

b. Problem reporting and resolution

Once data have been inspected and assessed, a variety of reporting mechanisms allow the data quality analysts to inform instrument mentors, site operators, and site scientists of their findings. Data quality reporting mechanisms are based on web searchable and accessible databases that allow the various pieces of information produced during the quality assurance process to be neatly conveyed to problem solvers in a timely manner. The history of ARM data problem reporting

is complex and is beyond the scope of this chapter (see [Peppeler et al. 2008a,b](#)). However, early on in the program, a problem identification form system was implemented to allow ARM science team members to help document data quality issues encountered beyond those identified by site scientists and instrument mentors. This system provided scientists a mechanism to notify infrastructure members when a problem in the data was encountered. However, the system was used inconsistently and produced an “uneven plowing” of ARM data fields. The successor system described

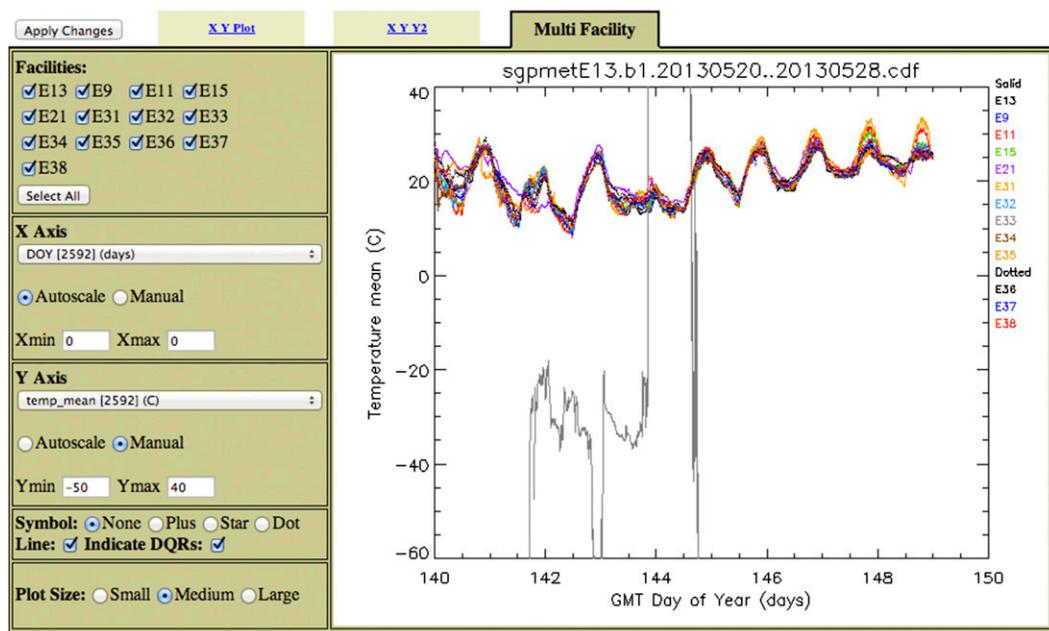


FIG. 12-3. Interactive plotting example showing 2-m temperature plotted for all SGP MET sites on the same graph. In this case, a problem temperature sensor at site E33 can be easily detected.

below has resulted in a more even and consistent treatment of ARM problem reporting and resolution.

The problem reporting system is divided into three linked processes:

- 1) Weekly reports are issued on data inspection and assessment results by DQO data quality analysts and distributed internally to instrument mentors, site operators, and site scientists.
- 2) Reports are issued describing problems discovered by data quality analysts or instrument mentors and distributed internally to instrument mentors, site operators, and site scientists so they can initiate a problem-resolution process—these online reports document the progress and status of the actions proposed and implemented.
- 3) Data quality reports documenting a known problem and its resolution as written by instrument mentors are distributed publicly to the data user community.

These data quality reports are provided to the data user along with the data they describe when an official data request is made to the ARM Data Archive. The history of a problem, including its discovery, the corrective actions taken to resolve it, and a report on its effects on data quality are typically included in these public reports and are database searchable on many criteria. The linked databases allow for the tracking of problem trends and help identify problematic instrument systems or facilities. An ARM Data Archive

service has been implemented to filter ordered data based on data quality reports. A data user may decide upfront which types of problems warrant data removal and can receive a custom product based on their particular needs. If a data user prefers more fine-grained control of data filtering, they can have their own processing codes query a data quality report web service in order to receive the details regarding any data problems in a machine-readable form.

c. Reprocessing

Last, the ARM infrastructure conducts an extensive data reprocessing program that is informed by the data quality assessment process. Reprocessing is performed to fix known data issues and has been used extensively throughout the lifetime of ARM. Reprocessing requires the modification or elimination of previous data quality reports and the subsequent reissuing of data to all who may have downloaded the data from the Data Archive (the Data Archive tracks all users of all data streams; [McCord and Voyles 2016](#), chapter 11). Reprocessing is not able to fix all problems, but it has been helpful in providing data users with the best products available. As an example of reprocessing, the MET instrumentation at the SGP Central Facility site has been known by different names since the beginning of ARM [Surface Meteorological Observation Station (SMOS) and then MET]. Each has had changing variable names over time, which has been confusing to data users. These data

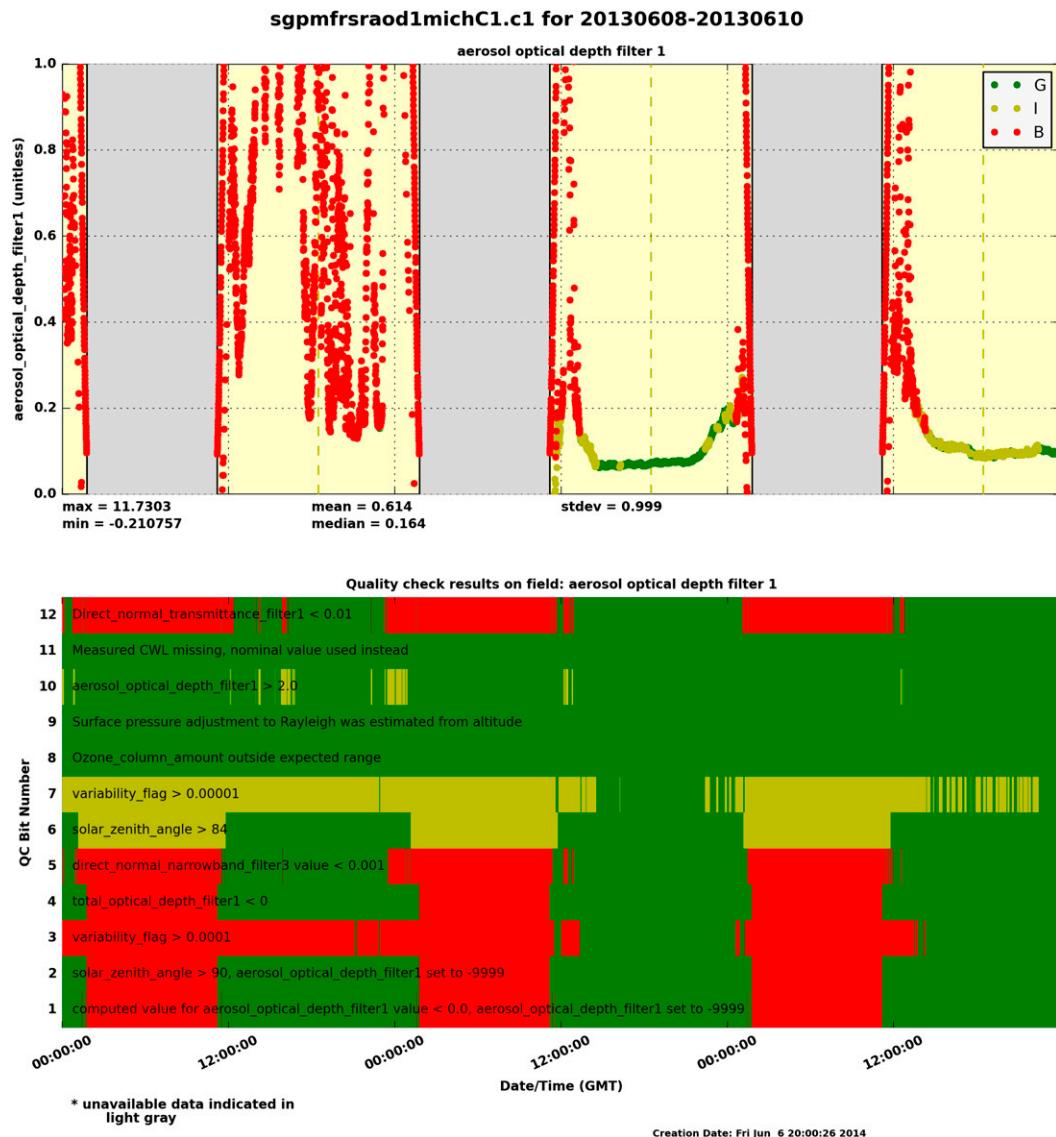


FIG. 12-4. Plot of aerosol optical depth from the SGP aerosol optical depth VAP (Koontz et al. 2013). The gray and yellow background in the top panel represents nighttime and daytime, respectively. Green indicates that no tests were failed and the data are “good,” yellow indicates that the data failed a test with an assessment level of “indeterminate,” and red indicates that the data failed a test with an assessment level of “bad.” Missing values (−9999) are automatically masked from the top-panel plot, and the y axis is scaled from 0 to 1 to show additional detail.

streams were reprocessed to provide one consistent dataset running from 1993 to the present time, greatly improving the ease of working with a long time series of the data (see Fig. 12-5). In another case, the European Centre for Medium-Range Weather Forecasts (ECMWF) model output retrieved from the ARM External Data Center was improperly ingested into the Merged Sounder (MERGESONDE) VAP (Troyan 2012). This issue adversely impacted moisture fields such as relative humidity and vapor pressure, causing the values within the boundary layer at the TWP Darwin site to essentially

drop to zero during time periods when sounding data was unavailable. A software update and subsequent reprocessing corrected this problem.

4. Role of VAPs in data quality characterization

Some of the scientific data needs of ARM data users are met through the creation of VAPs (<http://www.arm.gov/data/vaps>; Ackerman et al. 2016, chapter 3). Despite the extensive instrumentation deployed at the field sites, some measurements of interest are either impractical or

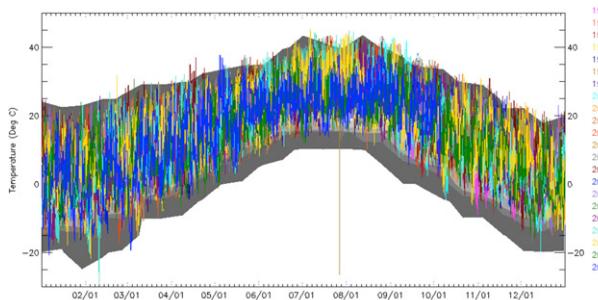


FIG. 12-5. Twenty years of temperature data from one SGP MET plotted over a PDF to indicate where varying percentages of the data normally lie (50%, 75%, 90%, 95%, 99.9%). A couple of outliers that have not been documented by data quality reports or removed by mentor supplied limits can be seen as dropouts in the data.

impossible to make directly or routinely—VAPs have filled this void. The creation and processing of VAPs have shed light not only on the usefulness of the higher-level products produced but also have provided information about the quality of the input data streams and the operation of the instruments that produced them. VAP processing and the analysis that is needed going into creating the VAPs has allowed detection of more subtle measurement inaccuracies that often defy detection through standard near-real-time data quality approaches such as limits testing or cross-measurement plotting comparisons. Two examples of VAP analysis aiding ARM data quality assurance efforts are described next.

a. Cloud radar

The Active Remotely-Sensed Cloud Locations (ARSCL) VAP (Clothiaux et al. 2000, 2001) uses millimeter cloud radar data as its primary input (Kollias et al. 2016, chapter 17). While the amount of power broadcast by the radar and returned by targets can be monitored, there are many factors involved in the operation of this complex radar that can affect data quality. ARSCL processing has revealed both radar measurement issues and radar operating characteristics. ARSCL output serves as input to the Baseline Cloud Microphysical Retrievals (MicroBase) VAP (Dunn et al. 2011), where retrievals are scrutinized both in terms of their consistency with other measurements and their relevance to the physical circumstances within which they are embedded. Consideration of consistency and situational context has been powerful for determining data quality to a degree not possible when analyzing measurements or retrievals in isolation.

b. Radiative transfer

The Broadband Heating Rate Profiles (BBHRP) VAP (Mlawer et al. 2002; McFarlane et al. 2016, chapter 20)

allowed the discovery of an unforeseen problem through its processing and subsequent user feedback. This VAP takes the output of the ARSCL and MicroBase VAPs and uses it in detailed radiative transfer model calculations to compute the radiative fluxes at both the surface and the top of the atmosphere. The BBHRP output is compared with surface and top-of-atmosphere irradiance measurements in a closure experiment framework (Mlawer and Turner 2016, chapter 14). This comparison revealed a subtle shift in model-minus-measurement flux difference statistics with direct normal shortwave measurements at the SGP Central Facility, a shift that was caused by human error when two digits of the normal incidence pyrheliometer calibration factor were inadvertently transposed while being entered into a datalogger. This error resulted in a roughly 2% error in the direct shortwave measurements, which is within the stated uncertainty of the calibrations themselves (Stoffel 2005) and as such was not detectable by standard limits and cross-comparison testing. This finding resulted in a reprocessing task to correct for this mistake, and led to a much improved direct normal solar flux dataset.

5. Role of field campaigns in data quality, measurement, and site characterization

ARM's data collection sites, including mobile facility deployments, host field campaigns to address specific scientific questions, augment routine data collections, and test and validate new instruments (<http://www.arm.gov/campaigns>). An emphasis of many campaigns has been on application of observational strategies and instrument deployments to improve the accuracy and quality of key measurements. A few are described here, which in some cases have produced climate community-wide ramifications on field measurement characterization.

a. Water vapor

Given the importance of water vapor as a greenhouse gas and its role in the life cycle of clouds and precipitation, the transfer of latent and sensible heat, and atmospheric chemistry, ARM has expended considerable observational effort, particularly at the SGP site, on the measurement of water vapor (Turner et al. 2016, chapter 13). Water vapor experiments held in 1996, 1997, and 2000, and a lidar experiment in 1999 provided key information on the quality and accuracy of onsite water vapor instrumentation (Revercomb et al. 2003). Dual-radiosonde launches revealed significant variability across and within calibration batches and showed that differences between any two radiosondes act as an altitude-independent scale factor in the lower

troposphere, such that a well-characterized reference can be used to reduce the variability. An approach subsequently was adopted by ARM to scale the radiosonde's moisture profile to agree with the precipitable water vapor observed by the microwave radiometer. This scaling significantly reduced the sonde-to-sonde variability by a factor of 2 (Turner et al. 2003). These water vapor experiments were able to verify that 1) 60-m tower-mounted in situ sensors can serve as an absolute measurement reference, 2) the SGP site's unique Raman lidar can serve as a stable transfer standard, and 3) the sensitivity of the site's microwave radiometers was excellent over a wide range of integrated water vapor. Data from the 1997 experiment figured strongly in an effort to evaluate retrievals of column water vapor and liquid water amounts from microwave radiometers (Ivanova et al. 2002).

During the first water vapor experiment in 1996, on-site humidity measurements were verified through laboratory intercomparison of in situ moisture sensors (including both capacitive chip and chilled mirror sensors) using Oklahoma Mesonet calibration facilities. Tests were made both before and after the experiment, making it possible to detect instrument problems prior to the experiment and instrument failure or drift during the experiment (Richardson et al. 2000). As a result of this work, modifications were made to humidity sensor calibration procedures and redundant humidity and temperature sensors were fielded to better detect sensor drift and calibration error.

While the aforementioned water vapor experiments were concerned with characterization of water vapor in the lower troposphere, the ARM First International Satellite Cloud Climatology Project (ISCCP) Regional Experiment (FIRE) Water Vapor Experiment (AFWEX), conducted with the National Aeronautics and Space Administration (NASA) in November–December 2000, attempted to better characterize the measurement of upper tropospheric water vapor (Ferrare et al. 2004). Results showed excellent agreement between satellite and Raman lidar observations of upper tropospheric humidity with systematic differences of about 10%; radiosondes, conversely, were found to be systematically drier by 40% relative to both satellite and lidar measurements (Soden et al. 2004). Existing strategies for correcting radiosonde dry biases were found to be inadequate in the upper troposphere, and an alternative method was suggested that considerably improved radiosonde measurement agreement with lidar observations. The alternative method was recommended as a strategy to improve the quality of the global historical record of radiosonde water vapor observations during the satellite era.

b. Atmospheric radiation

Some field campaigns have helped characterize the measurement of atmospheric radiation, especially broadband radiation (Michalsky and Long 2016, chapter 16). The second ARM Enhanced Shortwave Experiment (ARESE-II), conducted in February–April 2000 at the SGP site (Michalsky et al. 2002), focused on broadband shortwave calibration using both ground-based and aircraft-mounted radiometers and a special radiometer that could be considered a reference standard (Michalsky and Long 2016, chapter 16). A diffuse horizontal shortwave irradiance experiment held during September–October 2001 at the SGP site (Michalsky et al. 2003) characterized a nighttime offset by comparing diffuse irradiance measurements among most commercial pyranometers and a few prototypes, with the goal of reducing the uncertainty of shortwave diffuse irradiance measurements in lieu of a standard or reference for the measurement. An international pyrgeometer and absolute sky-scanning radiometer comparison held during September–October 1999 at the SGP site (Philipona et al. 2001) shed light on the reliability and consistency of atmospheric longwave radiation measurements and calculations and determined their uncertainties, also in lieu of an existing absolute standard.

c. Site characterization

Field campaigns also have contributed understanding on the representativeness of the ARM sites with respect to how well the data that are collected accomplish their scientific intent relative to desired measurement needs. The Manus Island and Nauru TWP sites had been established to make measurements representative of the surrounding oceanic area (Long et al. 2016, chapter 7). A goal of the Nauru99 field campaign was to investigate whether the small island, producing a cloud street phenomenon, was influencing the measurements made there (Post and Fairall 2000). The affirmative result led to a yearlong Nauru Island Effects Study in which a quantification of the island effect on measurements was made (Long 1998) and a method to detect the effect's ongoing occurrence and influence on collected data was developed (McFarlane et al. 2005). This study led to an explanation of the cloud street phenomenon (Matthews et al. 2007). These activities were able to quantify how well the measurements characterized the surrounding oceanic area and more generally illustrated the importance of considering spatial scales as part of the quality assurance process for siting instrumentation to measure the intended target environment.

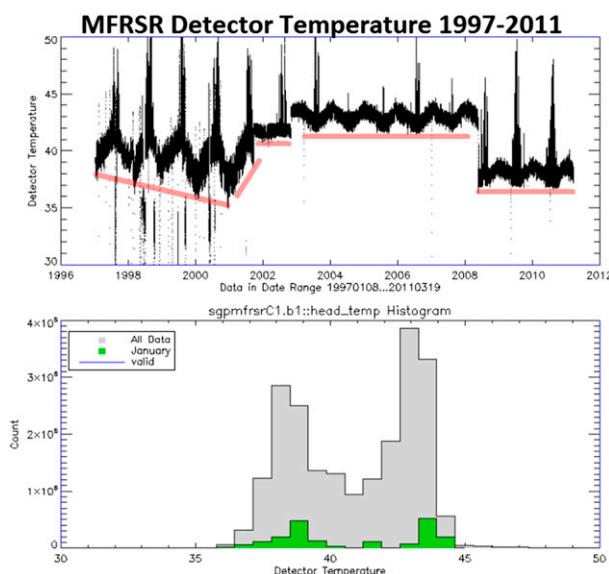


FIG. 12-6. (top) Long time series plots and (bottom) associated frequency distributions are used to detect trends and significant shifts in data that may indicate data quality problems. This example displays SGP MFRSR values over a 15-yr period at one site and indicates five distinct shifts in the data.

6. Use of ARM's historical data archive to improve current and past assessments

With over 20 years of continuous data amassed for some measurements at the ARM Data Archive (McCord and Voyles 2016, chapter 11), it is becoming possible to conduct statistical analysis on specific time scales; this work should provide valuable context for real-time measurements being made. Historical data are mined (Moore et al. 2007) to identify site-specific and time-varying (monthly or seasonal) quality control flagging limits and to facilitate better detection of subtle trends and abrupt changes in data (Fig. 12-6) that are difficult to understand when not considered in a broader context. This allows the incorporation of departures from the ARM climatology to inform the quality assurance process. Frequency distributions categorized by month and season help establish better data range limits specific to those time periods. Time series that alert analysts to outliers allow them to better distinguish bad data from unusual but valid data. Shown in Fig. 12-5 is 20 years of temperature data from one SGP MET plotted over a probability density function (PDF) to indicate where varying percentages of the data normally lie (50%, 75%, 90%, 95%, 99.9%). Figure 12-6 displays a long time series plot and associated frequency distribution of 15 years of multifilter rotating shadowband radiometer (MFRSR) values at one site—these indicate five distinct shifts in the data over that time.

In this analysis system, statistics and intelligent data range limits are used in a feedback mechanism to help data quality analysts and instrument mentors better manage validation checks. Web-based applications allow an analyst to request a particular data analysis and to view the results, allowing the dynamic creation of statistics and parametric analyses over any custom time range (e.g., day, month, year, or years).

7. Summary

The ARM data quality program necessarily has evolved from solid site-specific efforts as sites were established and matured, to a comprehensive, coordinated system that inspects, assesses, and reports on data quality incidents across all instrument systems and data collection sites. This coordinates the efforts not only of the DQO but also those of the instrument mentors, site scientists, site operators, and data system engineers. The role of scientific analysis, embodied primarily through the efforts to create VAPs and to conduct field campaigns, has allowed ARM to improve data quality by better characterizing measurements needed to improve the treatment of clouds and radiation in climate models, and better characterizing the sites where the measurements are taken so that they better fulfill the needs of the research community. The cumulative effect of these is a story of much individual and group effort that has taken place since the early 1990s.

What lessons has ARM learned over 20 years that are valuable to other field observation programs? A key lesson learned is that field measurement realities in sometimes harsh operational settings often deviate very quickly from intended instrument functioning in terms of what had been prescribed through laboratory calibration and created through inherent instrument measurement precision. Continuous calibration checking and routine scheduled maintenance therefore are essential to allow measurement systems to produce data as close as possible to their intended form. Also, intercomparison of like measurements at a particular location, to the extent possible and even for brief time periods, is key for helping establish the fidelity of the particular measurements being taken; the water vapor field campaigns, for example, were invaluable in helping ARM establish how to best measure water vapor, which had community-wide ramifications.

A comprehensive data quality assessment program is essential for documenting the quality assurance process and ultimately in producing a dataset of some prescribed known quality and usability. The program must collect and track data about the system (metadata) at every point along the assessment path, from instrument selection and

procurement to initial fielding and beta testing; to field operation, data collection, and quality inspection and assessment; to problem reporting and resolution; and to data distribution and communication of information about those data. If the process and details of the data created cannot be described, the data will have limited scientific value.

Going forward, the ARM data quality assessment process will take even greater advantage of automation opportunities as new tools and ways of thinking are developed. This automation process is necessary because the ARM data collection volume is anticipated to increase over time. This process needs to take increasing advantage of the sometimes 20-years-long ARM climatology that has been developed to better place current measurements into longer-term perspective and to create dynamic quality control flagging limits on various time scales to make them more meaningful. Another future goal is to better characterize current data in terms of inherent measurement accuracy and defined instrument precision, in concert with establishment of data spread over the long-term collection horizon.

REFERENCES

Ackerman, T. P., T. S. Cress, W. R. Ferrell, J. H. Mather, and D. D. Turner, 2016: The programmatic maturation of the ARM Program. *The Atmospheric Radiation Measurement (ARM) Program: The First 20 Years*, Meteor. Monogr., No. 57, Amer. Meteor. Soc., doi:[10.1175/AMSMONOGRAPH-D-15-0054.1](https://doi.org/10.1175/AMSMONOGRAPH-D-15-0054.1).

ARM Program Infrastructure Review Committee, 2001: The Atmospheric Radiation Measurement Program Infrastructure Review Report (AIR): Summary of Recommendations. U.S. Dept. of Energy ARM Program Doc. DOE/SC-ARM-0001, 3 pp. [Available online at <http://www.arm.gov/publications/programdocs/doe-sc-arm-0001.pdf?id=55>.]

Bahrmann, C. P., and J. M. Schneider, 1999: Near real-time assessment of SWATS data quality, resulting in an overall improvement in pre-sent-day SWATS data quality. *Proc. Ninth Atmospheric Radiation Measurement (ARM) Science Team Meeting*, San Antonio, TX, U.S. Dept. of Energy, 6 pp. [Available online at http://www.arm.gov/publications/proceedings/conf09/extended_abs/bahrmann_cp.pdf.]

Blough, D. K., 1992: Real-time statistical quality control and ARM. *Proc. 46th Annual ASQC Quality Congress*, Nashville, TN, American Society for Quality Control, 484–490.

Clothiaux, E. E., T. P. Ackerman, G. G. Mace, K. P. Moran, R. T. Marchand, M. A. Miller, and B. E. Martner, 2000: Objective determination of cloud heights and radar reflectivities using a combination of active remote sensors at the ARM CART sites. *J. Appl. Meteor.*, **39**, 645–665, doi:[10.1175/1520-0450\(2000\)039<0645:ODOCHA>2.0.CO;2](https://doi.org/10.1175/1520-0450(2000)039<0645:ODOCHA>2.0.CO;2).

—, and Coauthors, 2001: The ARM Millimeter Wave Cloud Radars (MMCRs) and the Active Remote Sensing of Clouds (ARSCL) Value Added Product (VAP). DOE Tech. Memo. ARM VAP-002.1, 56 pp. [Available online at http://www.arm.gov/publications/tech_reports/arm-vap-002-1.pdf.]

Cress, T. S., and D. L. Sisterson, 2016: Deploying the ARM sites and supporting infrastructure. *The Atmospheric Radiation Measurement (ARM) Program: The First 20 Years*, Meteor. Monogr., No. 57, Amer. Meteor. Soc., doi:[10.1175/AMSMONOGRAPH-D-15-0049.1](https://doi.org/10.1175/AMSMONOGRAPH-D-15-0049.1).

Delamere, J. S., and Coauthors, 1999: The first year of operation of the North Slope of Alaska/Adjacent Arctic Ocean ARM site: An overview of instrumentation, data streams, and data quality assurance procedures. *Proc. Ninth Atmospheric Radiation Measurement (ARM) Science Team Meeting*, San Antonio, TX, U.S. Dept. of Energy, 4 pp. [Available online at http://www.arm.gov/publications/proceedings/conf09/extended_abs/delamere1_js.pdf.]

Dunn, M., K. Johnson, and M. Jensen, 2011: The Microbase Value-Added Product: A baseline retrieval of cloud microphysical properties. Tech. Rep. DOE/SC-ARM/TR-095, 34 pp. [Available online at http://www.arm.gov/publications/tech_reports/doe-sc-arm-tr-095.pdf.]

Ferrare, R. A., and Coauthors, 2004: Characterization of upper-troposphere water vapor measurements during AFWEX using LASE. *J. Atmos. Oceanic Technol.*, **21**, 1790–1808, doi:[10.1175/1520-0450\(2002\)041<1790:ETQOGB>2.0.CO;2](https://doi.org/10.1175/1520-0450(2002)041<1790:ETQOGB>2.0.CO;2).

Ivanova, K., E. E. Clothiaux, H. N. Shirer, T. P. Ackerman, J. C. Liljegren, and M. Ausloos, 2002: Evaluating the quality of ground-based microwave radiometer measurements and retrievals using detrended fluctuations and spectral analysis methods. *J. Appl. Meteor.*, **41**, 56–68, doi:[10.1175/1520-0450\(2002\)041<0056:ETQOGB>2.0.CO;2](https://doi.org/10.1175/1520-0450(2002)041<0056:ETQOGB>2.0.CO;2).

Kollias, P., and Coauthors, 2016: Development and applications of ARM millimeter-wavelength cloud radars. *The Atmospheric Radiation Measurement (ARM) Program: The First 20 Years*, Meteor. Monogr., No. 57, Amer. Meteor. Soc., doi:[10.1175/AMSMONOGRAPH-D-15-0037.1](https://doi.org/10.1175/AMSMONOGRAPH-D-15-0037.1).

Koontz, A., C. Flynn, G. Hodges, J. Michalsky, and J. Barnard, 2013: Aerosol optical depth value-added product. Tech. Rep. DOE/SC-ARM/TR-129, 32 pp. [Available online at http://www.arm.gov/publications/tech_reports/doe-sc-arm-tr-129.pdf.]

Long, C. N., 1998: Nauru Island Effect Study (NIES) IOP Science Plan. Tech. Doc. DOE/SC-ARM-0505, 17 pp. [Available online at <http://www.arm.gov/publications/programdocs/doe-sc-arm-0505.pdf>.]

—, J. H. Mather, and T. P. Ackerman, 2016: The ARM Tropical Western Pacific (TWP) sites. *The Atmospheric Radiation Measurement (ARM) Program: The First 20 Years*, Meteor. Monogr., No. 57, Amer. Meteor. Soc., doi:[10.1175/AMSMONOGRAPH-D-15-0024.1](https://doi.org/10.1175/AMSMONOGRAPH-D-15-0024.1).

Matthews, S., J. M. Hacker, J. Cole, J. Hare, C. N. Long, and R. M. Reynolds, 2007: Modification of the atmospheric boundary layer by a small island: Observations from Nauru. *Mon. Wea. Rev.*, **135**, 891–905, doi:[10.1175/MWR3319.1](https://doi.org/10.1175/MWR3319.1).

McCord, R., and J. W. Voyles, 2016: The ARM data system and archive. *The Atmospheric Radiation Measurement (ARM) Program: The First 20 Years*, Meteor. Monogr., No. 57, Amer. Meteor. Soc., doi:[10.1175/AMSMONOGRAPH-D-15-0043.1](https://doi.org/10.1175/AMSMONOGRAPH-D-15-0043.1).

McFarlane, S. A., C. N. Long, and D. M. Flynn, 2005: Impact of island-induced clouds on surface measurements: Analysis of the ARM Nauru Island Effect Study data. *J. Appl. Meteor.*, **44**, 1045–1065, doi:[10.1175/JAM2241.1](https://doi.org/10.1175/JAM2241.1).

—, J. H. Mather, and E. J. Mlawer, 2016: ARM's progress on improving atmospheric broadband radiative fluxes and heating rates. *The Atmospheric Radiation Measurement (ARM) Program: The First 20 Years*, Meteor.

Monogr., No. 57, Amer. Meteor. Soc., doi:10.1175/AMSMONOGRAPHSD-15-0046.1.

Michalsky, J. J., and C. N. Long, 2016: ARM solar and infrared broadband and filter radiometry. *The Atmospheric Radiation Measurement (ARM) Program: The First 20 Years, Meteor. Monogr.*, No. 57, Amer. Meteor. Soc., doi:10.1175/AMSMONOGRAPHSD-15-0031.1.

—, and Coauthors, 2002: Broadband shortwave calibration results from the Atmospheric Radiation Measurement Enhanced Shortwave Experiment II. *J. Geophys. Res.*, **107**, 4287, doi:10.1029/2001JD001231.

—, and Coauthors, 2003: Results from the first ARM diffuse horizontal shortwave irradiance comparison. *J. Geophys. Res.*, **108**, 4108, doi:10.1029/2002JD002825.

Miller, N. E., J. C. Liljegren, T. R. Shippert, S. A. Clough, and P. D. Brown, 1994: Quality measurement experiments within the Atmospheric Radiation Measurement Program. *Preprints, Fifth Symp. on Global Change Studies*, Nashville, TN, Amer. Meteor. Soc., 35–39.

Mlawer, E. J., and D. D. Turner, 2016: Spectral radiation measurements and analysis in the ARM Program. *The Atmospheric Radiation Measurement (ARM) Program: The First 20 Years, Meteor. Monogr.*, No. 57, Amer. Meteor. Soc., doi:10.1175/AMSMONOGRAPHSD-15-0027.1.

—, and Coauthors, 2002: The Broadband Heating Rate Profile (BBHRP) VAP. *Proc. 12th Atmospheric Radiation Measurement (ARM) Science Team Meeting*, St. Petersburg, FL, U.S. Dept. of Energy, 12 pp. [Available online at http://www.arm.gov/publications/proceedings/conf12/extended_abs/mlawer-ej.pdf.]

Moore, S. T., K. Kehoe, R. Peppler, and K. Sonntag, 2007: Analysis of historical ARM measurements to detect trends and assess typical behavior. *16th Conf. on Applied Climatology*, San Antonio, TX, Amer. Meteor. Soc., P2.6. [Available online at <http://ams.confex.com/ams/pdpapers/119946.pdf>.]

Peppler, R. A., and M. E. Splitt, 1997: SGP Site Scientist Team data quality assessment activities. *Proc. Seventh Atmospheric Radiation Measurement (ARM) Science Team Meeting*, San Antonio, TX, U.S. Dept. of Energy, 403–406. [Available online at http://www.arm.gov/publications/proceedings/conf07/extended_abs/pepper_ra.pdf.]

—, K. E. Kehoe, K. L. Sonntag, S. T. Moore, and K. J. Doty, 2005: Improvements to and status of ARM's Data Quality Health and Status System. *15th Conf. on Applied Climatology*, Savannah, GA, Amer. Meteor. Soc., J3.13. [Available online at <http://ams.confex.com/ams/pdpapers/91618.pdf>.]

—, and Coauthors, 2008a: Quality Assurance of ARM Program Climate Research Facility data. *Tech. Rep. DOE/SC-ARM/ TR-082*, 65 pp. [Available online at http://www.arm.gov/publications/tech_reports/doe-sc-arm-tr-082.pdf.]

—, and Coauthors, 2008b: An overview of ARM Program Climate Research Facility data quality assurance. *Open Atmos. Sci. J.*, **2**, 192–216, doi:10.2174/1874282300802010192.

Philipona, R., and Coauthors, 2001: Atmospheric longwave irradiance uncertainty: Pyrgeometers compared to an absolute sky-scanning radiometer, atmospheric emitted radiance interferometer, and radiative transfer model calculations. *J. Geophys. Res.*, **106**, 28 129–28 141, doi:10.1029/2000JD000196.

Post, M. J., and C. F. Fairall, 2000: Early results from the Nauru99 campaign on NOAA ship Ronald H. Brown. *Proc. Int. Geoscience and Remote Sensing Symp.*, Honolulu, HI, IEEE, 1151–1153, doi:10.1109/IGARSS.2000.858052.

Revercomb, H. E., and Coauthors, 2003: The ARM Program's water vapor intensive observation periods. *Bull. Amer. Meteor. Soc.*, **84**, 217–236, doi:10.1175/BAMS-84-2-217.

Richardson, S. J., M. E. Splitt, and B. M. Lesht, 2000: Enhancement of ARM surface meteorological observations during the fall 1996 water vapor intensive observation period. *J. Atmos. Oceanic Technol.*, **17**, 312–322, doi:10.1175/1520-0426(2000)017<0312:EOASMO>2.0.CO;2.

Sisterson, D., R. Peppler, T. S. Cress, P. Lamb, and D. D. Turner, 2016: The ARM Southern Great Plains (SGP) site. *The Atmospheric Radiation Measurement (ARM) Program: The First 20 Years, Meteor. Monogr.*, No. 57, Amer. Meteor. Soc., doi:10.1175/AMSMONOGRAPHSD-16-0004.1.

Soden, B. J., D. D. Turner, B. M. Lesht, and L. M. Miloshevich, 2004: An analysis of satellite, radiosonde, and lidar observations of upper tropospheric water vapor from the Atmospheric Radiation Measurement Program. *J. Geophys. Res.*, **109**, D04105, doi:10.1029/2003JD003828.

Splitt, M. E., 1996: Data quality display modules—Assessment of instrument performance at the Southern Great Plains Cloud and Radiation Testbed site. *Proc. Sixth Atmospheric Radiation Measurement (ARM) Science Team Meeting*, San Antonio, TX, U.S. Dept. of Energy, 3 pp. [Available online at http://www.arm.gov/publications/proceedings/conf06/extended_abs/splitt_me.pdf.]

Stoffel, T., 2005: Solar Infrared Radiation Station (SIRS) Handbook. *Tech. Rep. ARM TR-025*, 27 pp. [Available online at http://www.arm.gov/publications/tech_reports/handbooks/sirs_handbook.doc.]

Stokes, G. M., 2016: Original ARM concept and launch. *The Atmospheric Radiation Measurement (ARM) Program: The First 20 Years, Meteor. Monogr.*, No. 57, Amer. Meteor. Soc., doi:10.1175/AMSMONOGRAPHSD-15-0021.1.

Troyan, D., 2012: Merged sounding value-added product. *Tech. Rep. DOE/SC-ARM/TR-087*, 19 pp. [Available online at http://www.arm.gov/publications/tech_reports/doe-sc-arm-tr-087.pdf.]

Turner, D. D., B. M. Lesht, S. A. Clough, J. C. Liljegren, H. E. Revercomb, and D. C. Tobin, 2003: Dry bias and variability in Vaisala RS80-H radiosondes: The ARM experience. *J. Atmos. Oceanic Technol.*, **20**, 117–132, doi:10.1175/1520-0426(2003)020<0117:DBAVIV>2.0.CO;2.

—, and Coauthors, 2004: The QME AERI LBLRTM: A closure experiment for downwelling high spectral resolution infrared radiance. *J. Atmos. Sci.*, **61**, 2657–2675, doi:10.1175/JAS3300.1.

—, E. J. Mlawer, and H. E. Revercomb, 2016: Water vapor observations in the ARM Program. *The Atmospheric Radiation Measurement (ARM) Program: The First 20 Years, Meteor. Monogr.*, No. 57, Amer. Meteor. Soc., doi:10.1175/AMSMONOGRAPHSD-15-0025.1.

Verlinde, J., B. Zak, M. D. Shupe, M. Ivey, and K. Starnes, 2016: The ARM North Slope of Alaska (NSA) sites. *The Atmospheric Radiation Measurement (ARM) Program: The First 20 Years, Meteor. Monogr.*, No. 57, Amer. Meteor. Soc., doi:10.1175/AMSMONOGRAPHSD-15-0023.1.