Instrument Cross-Comparisons and Automated Quality Control of Atmospheric Radiation Measurement Data

S. Moore and G. Hughes
ATK Mission Research
Santa Barbara, California

Introduction

Within the Atmospheric Radiation Measurement (ARM) instrument network, several different systems often measure the same quantity at the same site. For example, several ARM instruments measure time-series profiles of the atmosphere that were previously available only from balloon-borne radiosonde systems. These instruments include the Radar Wind Profilers (RWP) with Radio-Acoustic Sounding Systems (RASS), the Atmospheric Emitted Radiance Interferometer (AERI), the Microwave Radiometer Profiler (MWRP), and the Raman Lidar (RL). ARM researchers have described methods for direct cross-comparison of time-series profiles (Coulter and Lesht 1996; Turner et al. 1996) and we have extended this concept to the development of methods for automated quality control (QC) of ARM datastreams.

Automated Quality Control Using Climatologies

The current data flow architecture at the Data Management Facility (DMF) does not allow direct cross comparisons to other data streams during the generation of quality checked (b-level) instrument data files. At the moment when QC checks are performed, the availability of comparison data streams cannot be guaranteed. In order to perform automated QC at the time of data ingest, another method is required. One method is to compare key values against historical ranges.

For several ARM data sets, many years of observations are now available. By analyzing the distribution of the observations over the life of the data stream, one can get a good feel for what is typical. The historical record and statistics can be used to set appropriate limits for automated QC checks and we have implemented such checks into ARM's Data Quality Health and Status (DQ HandS) system (Peppler et al. 2005). But, when other sources of similar data exist, such as balloon-borne sondes in the case of atmospheric profiles, it is desirable to find a way to cross-check measurements against these supplementary sources.

To do this, we have built sonde-based climatologies for each of the three fixed ARM sites. Sonde data was collected from facilities at the North Slope of Alaska (NSA), the Southern Great Plains (SGP), and the Tropical Western Pacific (TWP). We have also built a climatology for the National Weather Service sonde stationed at Barrow, Alaska. These climatologies are stored using the NetCDF data format and contain monthly ranges, distributions, lapse rates, means and other statistical moments. We do this for:

wind speed and direction (NSA and SGP), temperature, dewpoint, relative humidity, and calculated water vapor mixing ratio. The statistics for each quantity are computed and stored by month and by 50 m altitude bin, up to about 25 km.

Figures 1, 2, and 3 display sample monthly profile distributions (January and July illustrated) for each of the atmospheric parameters of wind speed, temperature, and calculated water vapor mixing ratio respectively. Our web-based NetCDF data viewer, NCVweb, was used to easily visualize the distributions from our lookup table, and to create these figures. It can clearly be seen that there are large differences in these measurements at different times of the year and at various altitudes. The existence and accessibility of these monthly statistics makes it possible to refine automated QC dramatically. No longer must automated QC be restricted to checks comparing measurements to broad, all-encompassing min/max/delta limits.

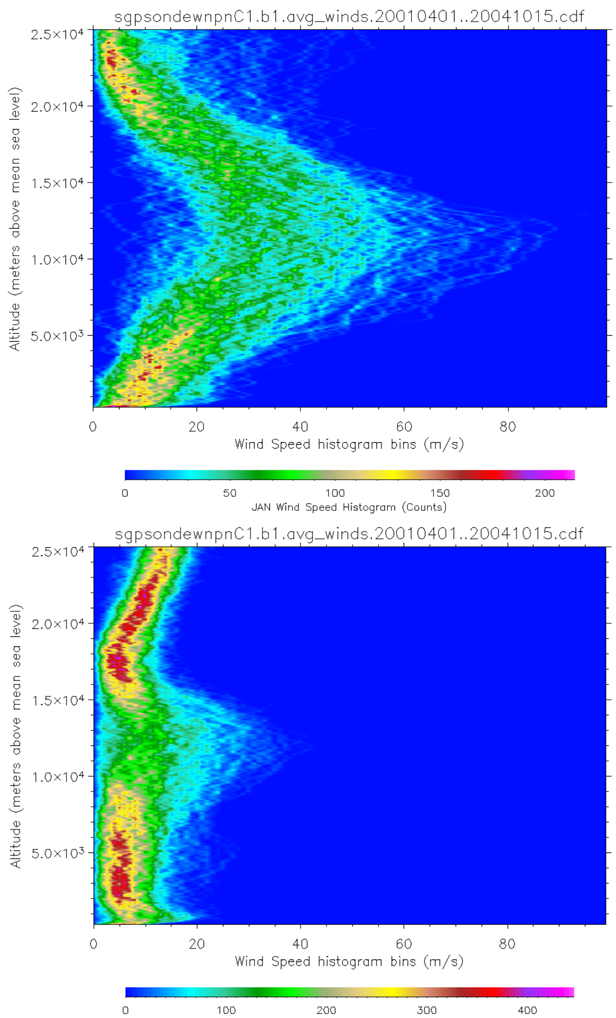

Figure 1. Monthly statistics for SGP wind speed profiles were calculated for 50 m altitude bins based on 5,068 sonde balloon launches between 01 April 2001 and 15 October 2004. Displayed are profiles for the months of January (top) and July (bottom).

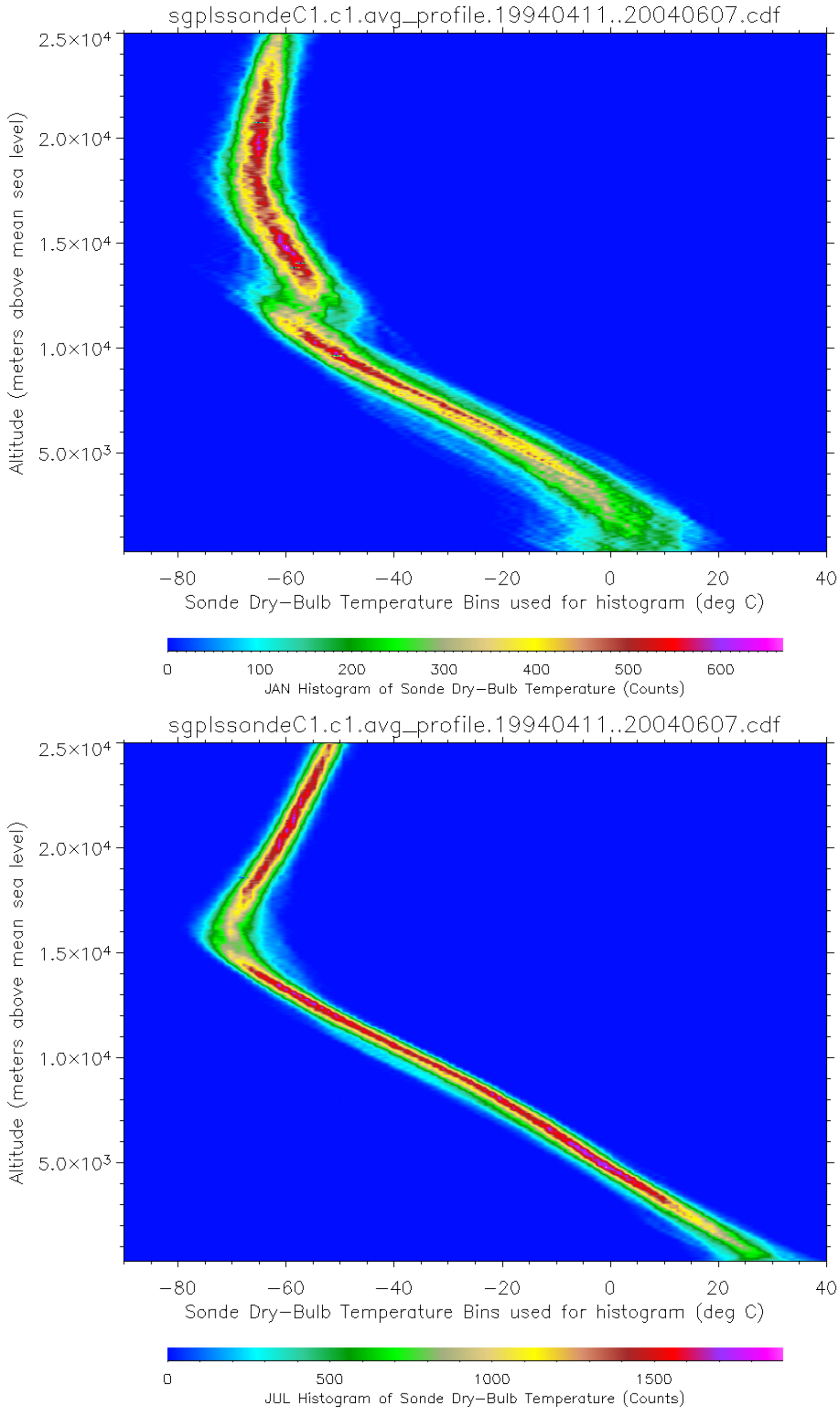

Figure 2. Monthly statistics for SGP temperature profiles were calculated for 50 m altitude bins based on 10,038 sonde balloon launches between 11 April 1994 and 07 June 2004. Displayed are profiles for the months of January (top) and July (bottom).

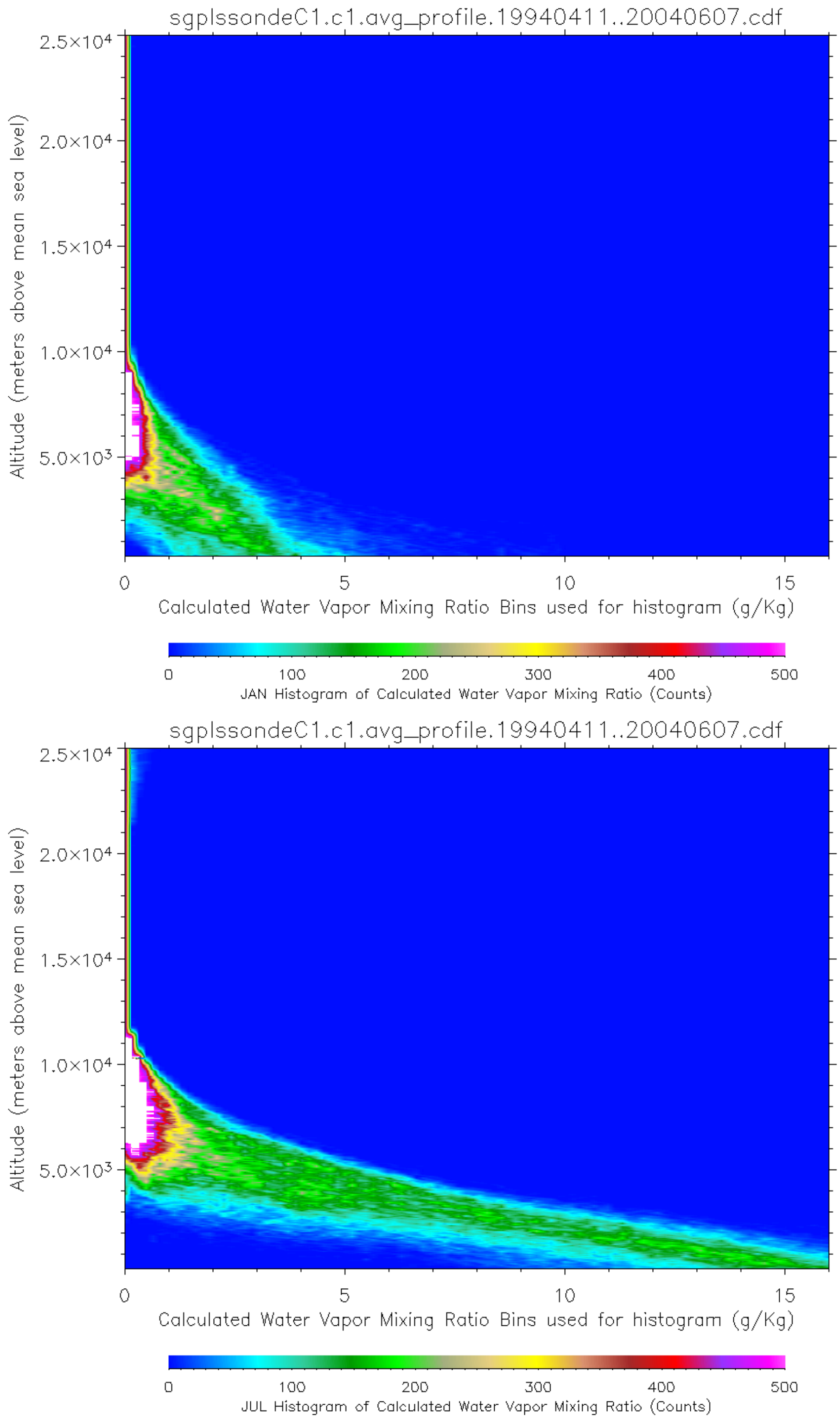

Figure 3. Monthly statistics for SGP water vapor mixing ratio profiles were calculated for 50 m altitude bins based on 10,038 sonde balloon launches between 11 April 1994 and 07 June 2004. Displayed are profiles for the months of January (top) and July (bottom).

We have implemented software to use our sonde-based climatology database and perform cross comparison quality checks with various ARM profiling instruments. To date, we have automated routines to provide these additional checks for profiles from AERI, RASS, and RWP. Figure 4a demonstrates this technique for RASS temperature data, and Figure 4b demonstrates the technique for RWP wind speed data. Instrument data is compared against average sonde profiles for any given measurement as a function of month and height. If the value falls outside an acceptable range (e.g. four standard deviations from the mean), the value is flagged as a potential min or max outlier. In addition, observation (or retrieval) lapse rates are compared against historical sonde lapse rates. If the value falls outside of an acceptable range, determined by the climatology, the value is flagged as failing delta check.

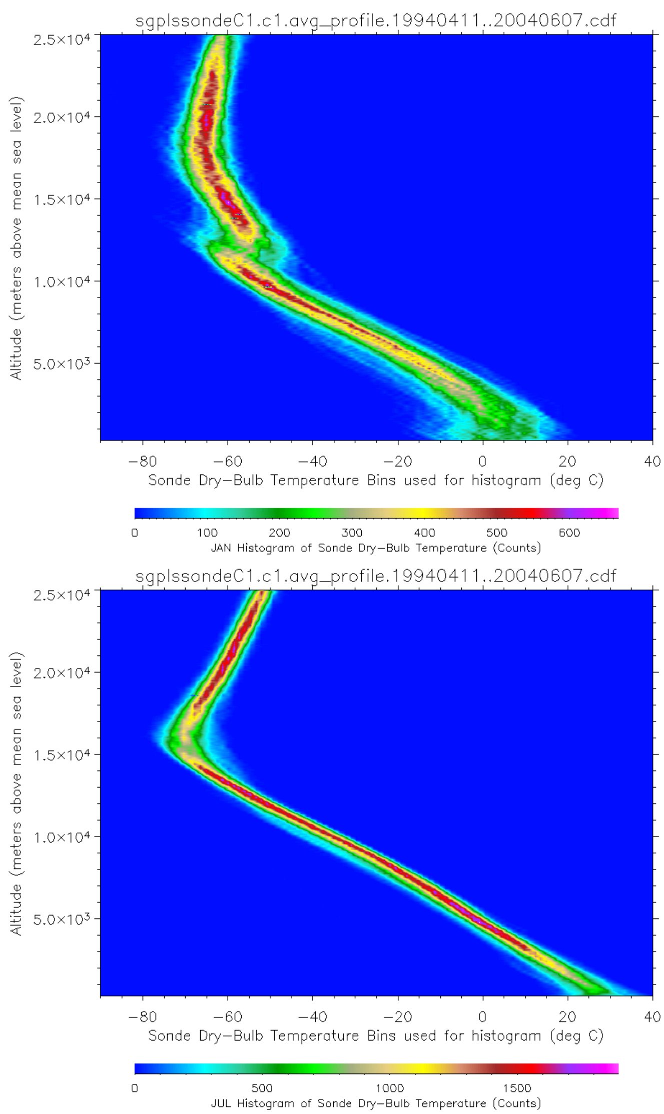

Figure 4. Monthly statistics for SGP temperature profiles were calculated for 50 m altitude bins based on 10,038 sonde balloon launches between 11 April 1994 and 07 June 2004. Displayed are profiles for the months of January (top) and July (bottom).

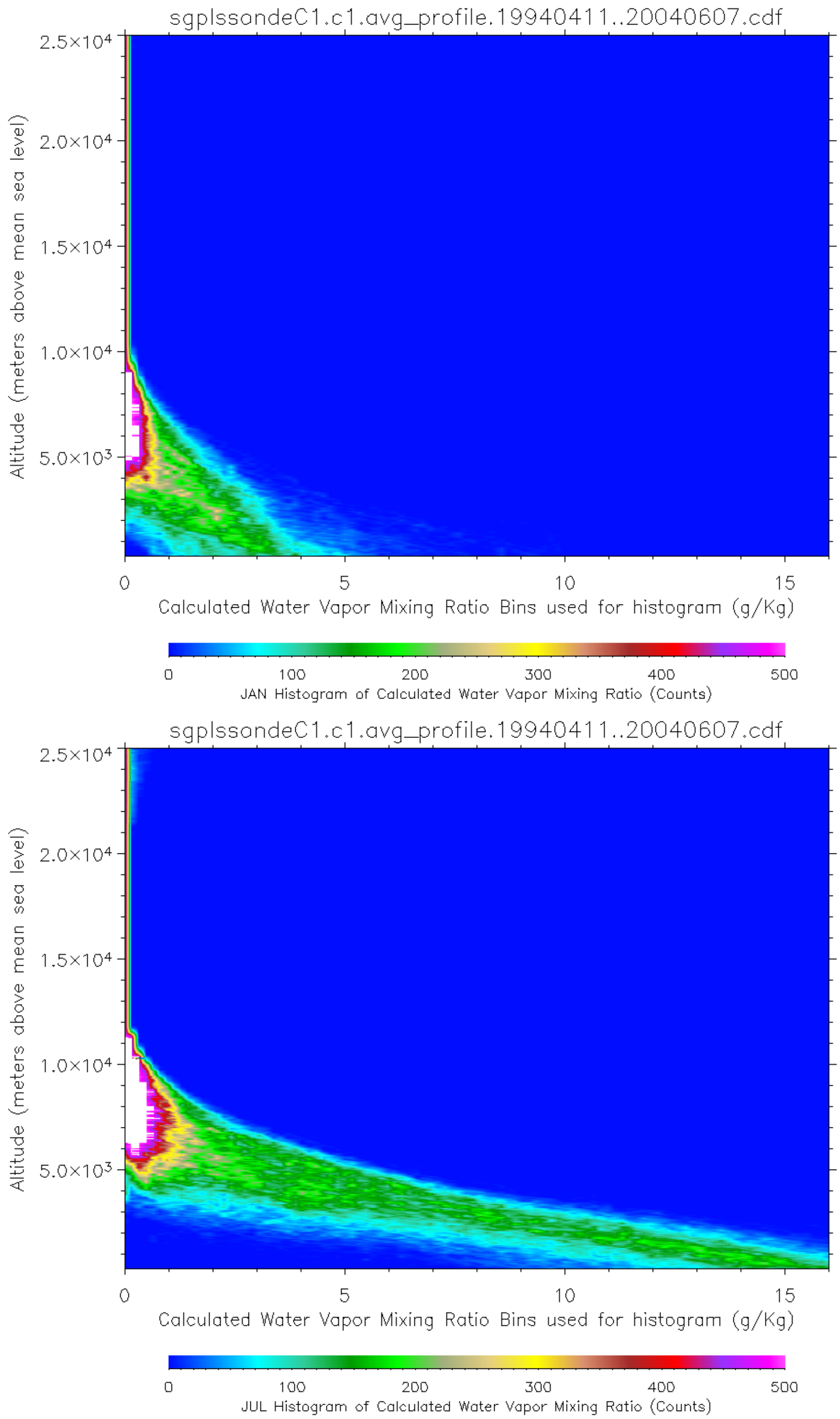

Figure 5. Monthly statistics for SGP water vapor mixing ratio profiles were calculated for 50 m altitude bins based on 10,038 sonde balloon launches between 11 April 1994 and 07 June 2004. Displayed are profiles for the months of January (top) and July (bottom).

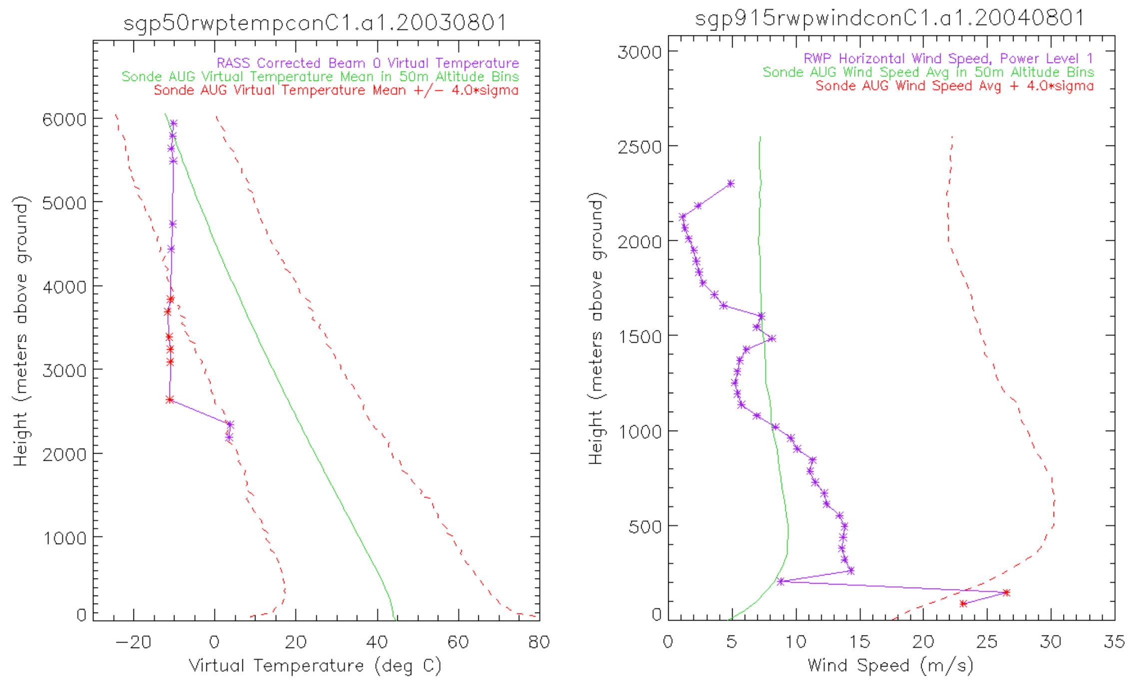

Figure 6. (a) For the month of August, the solid green line depicts the average temperature profile as reported by the sonde at the SGP site. The red dotted lines mark the spread of sonde values out to four standard deviations from the average. On 1 August 2003, RASS temperature data (purple stars) were compared against the sonde climatology, and out of bound values were flagged (red stars).
(b) For August, the solid green line depicts the average wind speed profile as reported by the sonde at the SGP site. The red dotted lines mark the spread of sonde values out to four standard deviations from the average. On 1 August 2004, RWP wind speed data (purple stars) were compared against the sonde climatology, and out of bound values were flagged (red stars).

Although ARM does not currently capture these additional flags into the NetCDF files distributed to users, the metrics based on these flags can be viewed at http://dq.arm.gov/. The total number of failing values for any given hour is divided by the total number of observations during that hour to give an hourly percentage of failing values for display with DQ HandS. Statistics are also calculated for missing and not available values.

Summary

This method complements existing quality checks for instrument data by enforcing seasonal rules for data range and rate of change, based on independent references. This method is more useful for automated QC than direct comparison methods because the procedure does not require immediate availability of other datastreams. Because the ARM data flow architecture does not allow for direct cross comparisons to other data streams, this method will be especially valuable if implemented as part of the production data ingest. The additional QC information is currently saved into files on the DQ computer at the DMF, and is made available for display by the DQ Health and Status Explorer.

Acknowledgments

This research was supported by the Office of Science (BER), U.S. Department of Energy, Grant Number DE-FG02-04ER63864.

Point of Contact

Sean Moore, sean.moore@arm.gov, (805) 963-8761 ext 243

References

Coulter, RL and BM Lesht. 1996. "Results of an Automated Comparison Between Winds and Virtual Temperatures from Radiosonde and Profilers." In *Proceedings of the Sixth Atmospheric Radiation Measurement (ARM) Science Team Meeting*, DOE CONF-9603149, March.

Peppler, RA, KE Kehoe, KL Sonntag, ST Moore, and KJ Doty. 2005. "Improvements to and Status of ARM's Data Quality Health and Status System." In *Proceedings of the Fifteenth Conference on Applied Climatology, Meeting of the American Meteorological Society*, June.

Turner, DD, TR Shippert, JC Liljegren, Y Han, and ER Westwater. 1996. "Initial analysis of water vapor and temperature profiles retrieved from integrated ground-based remote sensors." In *Proceedings of the Sixth Atmospheric Radiation Measurement (ARM) Science Team Meeting*, DOE CONF—9603149, March.