Latest Tools for Viewing and Quality Checking ARM Data

S. Moore and G. Hughes
Missing Research Corporation
Santa Barbara, California

Introduction

The DOE Atmospheric Radiation Measurement (ARM) program has acquired an incredibly large quantity of data over its period of operation, all of which must be reviewed in some manner in order to ensure that the data is of "known and reasonable" quality (ARM Science Plan). To accomplish this, Mission Research Corporation (MRC) coordinates with the ARM Data Quality Office to develop software tools that quality-check data products in a timely and continuous fashion. These tools work with the Data Quality Health and Status (DQ HandS) Explorer (Peppler et al. 2004) by analyzing ARM data streams, providing assessments of data quality in the form of color-coded tables, and generating custom visualizations.

Tools for interactively visualizing ARM data sets have also been developed by us for inspecting each of the 600+ data streams that pass through the ARM Data Management Facility (DMF) each day. NCVweb is a Web based application developed for this task, and is designed to allow easy access to ARM data at the DMF and at the ARM archive.

Here we will review some of NCVweb's key features as well as introduce new tools for viewing the quality of Atmospheric Emitted Radiance Interferometer (AERI) data, its corresponding Value Added Products (VAPs) and Quality Measurement Experiments (QMEs).

NCVweb Visualization Tool

With NCVweb, the user eliminates the step of transferring files to their own computer for visualization thereby saving time and effort. The tool works with all ARM data streams that incorporate the NetCDF file format. Since it operates by probing each file for the details necessary for visualization, little or no work goes into preparing the tool for new data streams. Any problem with the data stream is easily identified with NCVweb and accessible by anyone with a Web browser. Best of all, all quantities in the data stream are easily viewed and plotted to the user's liking.

The NCVweb code has been significantly improved this past year to allow for different user interface modes specific to the type of data selected. For example, the user interface presented when three-dimensional data is available in the selected file can now be quite different than when only two-dimensional data is available. The user interface adapts to not only the type of files selected, but also to the desired visualization. This major change will allow for easy inclusion of advanced analysis modes in the future, without cluttering the simpler visualization screens used for basic plotting.

NCVweb can be used at the ARM Data Quality web site or from DQ HandS via the interactive plot link. NCVweb is also integrated into the ARM archive to allow easy viewing of standing-order data or of data ordered by an archive user (http://arm.mrcsb.com/). Users are able to view data sets before deciding on whether to download files.

NCVweb can be used in a multitude of ways. Figure 1 demonstrates its use visualizing shortwave and longwave down-welling irradiance at the Southern Great Plains central facility. In this example, the tool seamlessly joins multiple files together for plotting purposes; indicates the diurnal cycle with an alternating day/night colored background; interprets and converts ARM time to day of year; detects and does not plot data points designated as "missing" and generates plot labels based on the data viewed. The user can zoom in on any region, change the plot symbol, change the plot size, and request detailed information concerning file contents, the file header, statistics of each quantity, or convert the data to ASCII format. NCVweb also detects and displays suspicious data with a red marker, based on quality control flags embedded within ARM "b1" level files.
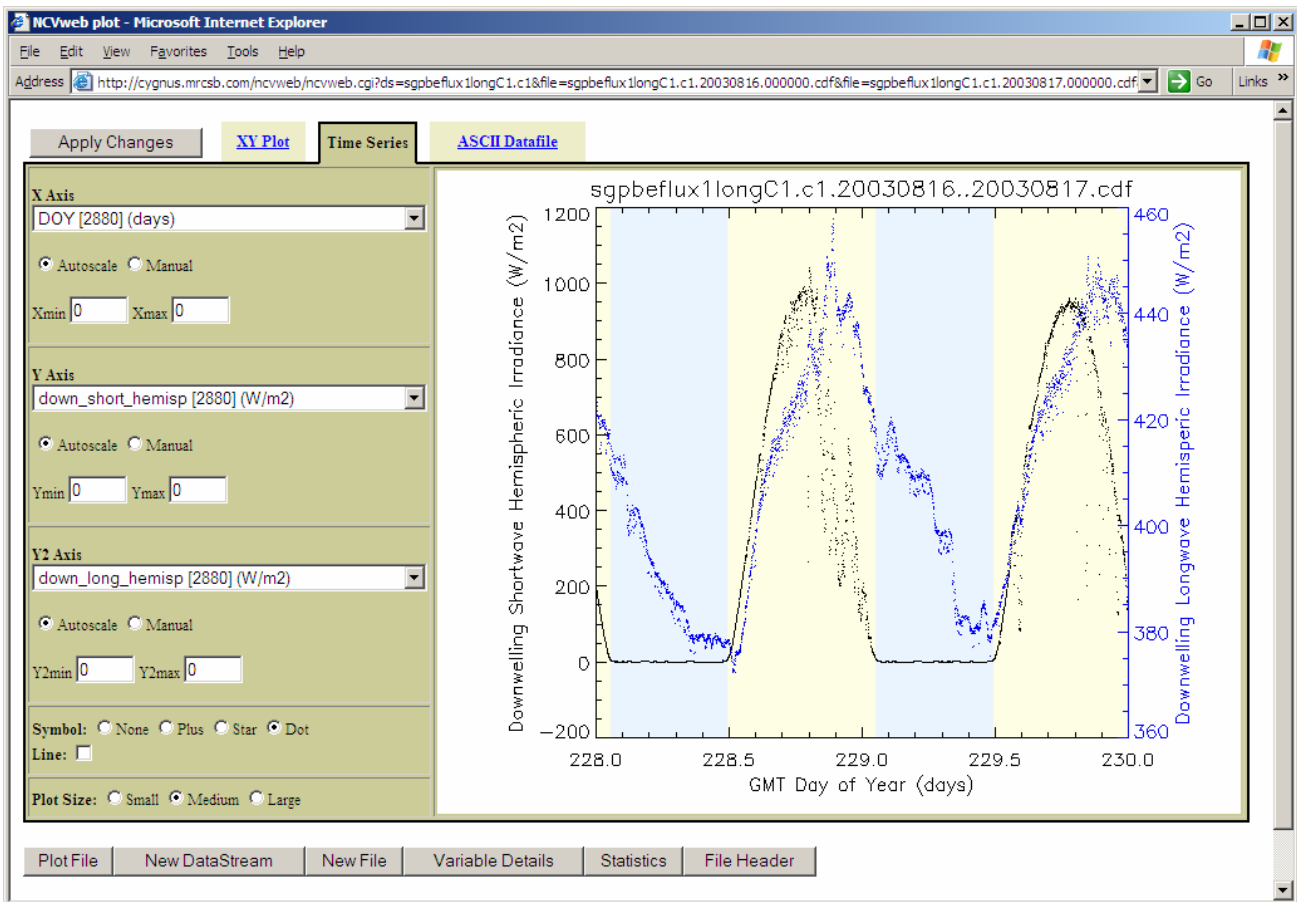

Figure 1. Example of NCVweb, showing a time series of two variables (shortwave and longwave irradiance) concurrently.

Three dimensional features are presented in Figure 2. Cloud structure is visible in this color image plot. Reds and purples indicate high backscatter detected by the micropulse LIDAR instrument at the Southern Great Plains central facility. Aerosols are visible in the first few kilometers above ground level, and clouds between 8 and 14 km can be seen throughout the day. NCVweb makes it easy to adjust the range of colors used to represent the primary measurement. Subtle features can be detected and enhanced by judicious choice of the Z-axis plot range and color table—something that cannot be done with "canned" plots available elsewhere. Once interesting features are observed in this manner, more detail can be viewed by plotting cross-sections through the regions of interest. NCVweb allows both vertical and horizontal cross-sections with any compatible variables used for either axis.
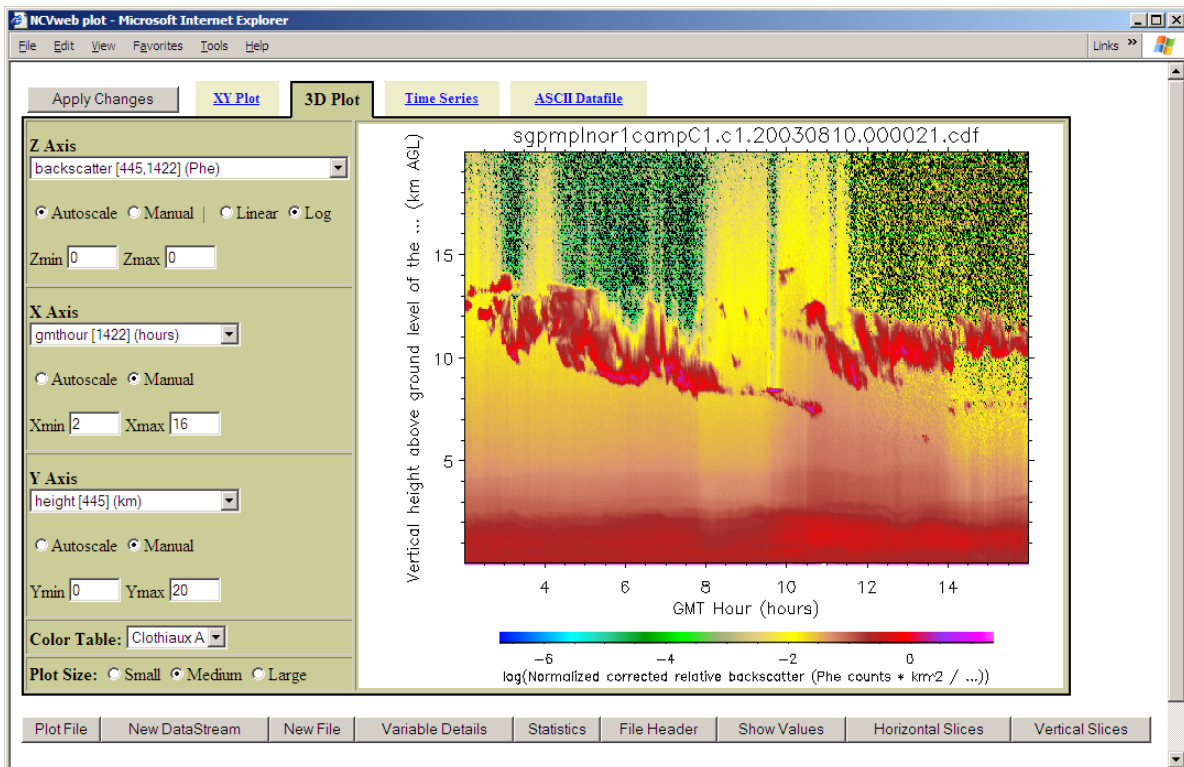

Figure 2. Example of NCVweb three dimensional features. Micropulse LIDAR data is displayed.

AERI Automated Quality Control Software

The approach for quality analysis of AERI Radiance, Engineering and Summary data streams has been to perform min/max/delta checks based on parameters supplied by the University of Wisconsin AERI Status web site http://webmads.ssec.wisc.edu/ or http://cimss.ssec.wisc.edu/aeri/. We generate lookup tables based on limits that are available at the AERI web site for warning and failing levels for the essential AERI engineering quantities. As the data streams are processed, the lookup tables are consulted to determine whether the values should be flagged as suspicious. We then generate appropriate metric files that are readable by DQ HandS as well as a suite of diagnostic plots for each key quantity in each data stream (Figure 3). Future effort will be directed toward quality analysis of radiance data and engineering quantities not listed at the University of Wisconsin site.
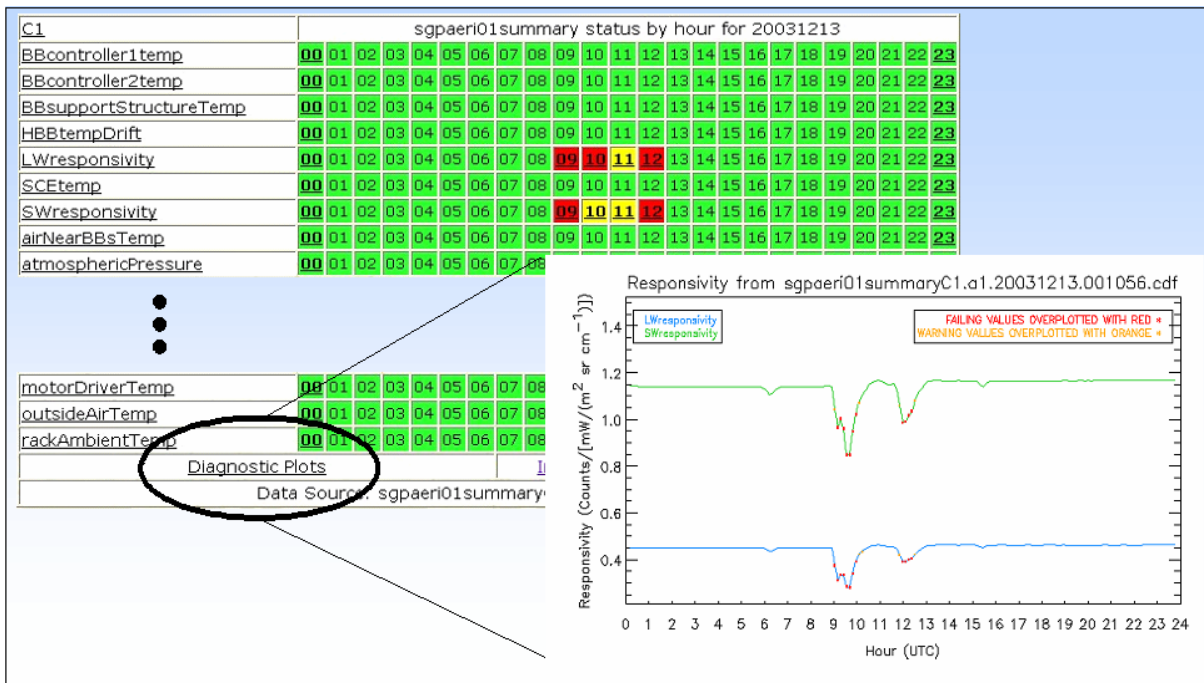
Figure 3. AERI Data Quality display in DQ HandS and sample diagnostic plot for instrument responsivities.

Our approach for quality analysis of AERI PROF VAP has been to establish nominal ranges and nominal lapse rates for absolute temperature, dewpoint temperature and water vapor mixing ratio from historical sonde profiles (Figure 4a). The AERI profiles, which are based on models, are tested against the ranges derived from sonde data, which are based on direct measurements (Figure 4b). If a retrieved AERI profile predicts a value that is outside the sonde range, then there is a defensible reason to raise suspicions about the AERI value.

The sonde ranges for each parameter and their lapse rates are established by calculating the monthly means and standard deviations within 50 meter altitude bins. For a sonde temperature that is normally distributed within the altitude bin, our acceptable range contains 99.4% of observed sonde temperature values. If the AERI PROF retrieved temperature falls outside this range, there is only a 0.6% chance that the retrieved temperature could have come from the observed temperature distribution.

For automated quality analysis of AERI PROF, each data value in the retrieved AERI profile is compared to the historical sonde range for the month of observation and corresponding altitude bin. If the AERI PROF value falls outside of the acceptable sonde range (currently 2.75 standard deviations from historical values) then the AERI PROF value is flagged as failing max/min check. If the calculated AERI PROF lapse rate falls outside a similar range for average sonde lapse rates, then the AERI PROF value is flagged as failing delta check. Hourly statistics for missing, failing and not available values are calculated for display in DQ HandS. 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. Statistics are also calculated for missing and not available values and reported to DQ HandS via the usual metrics file.
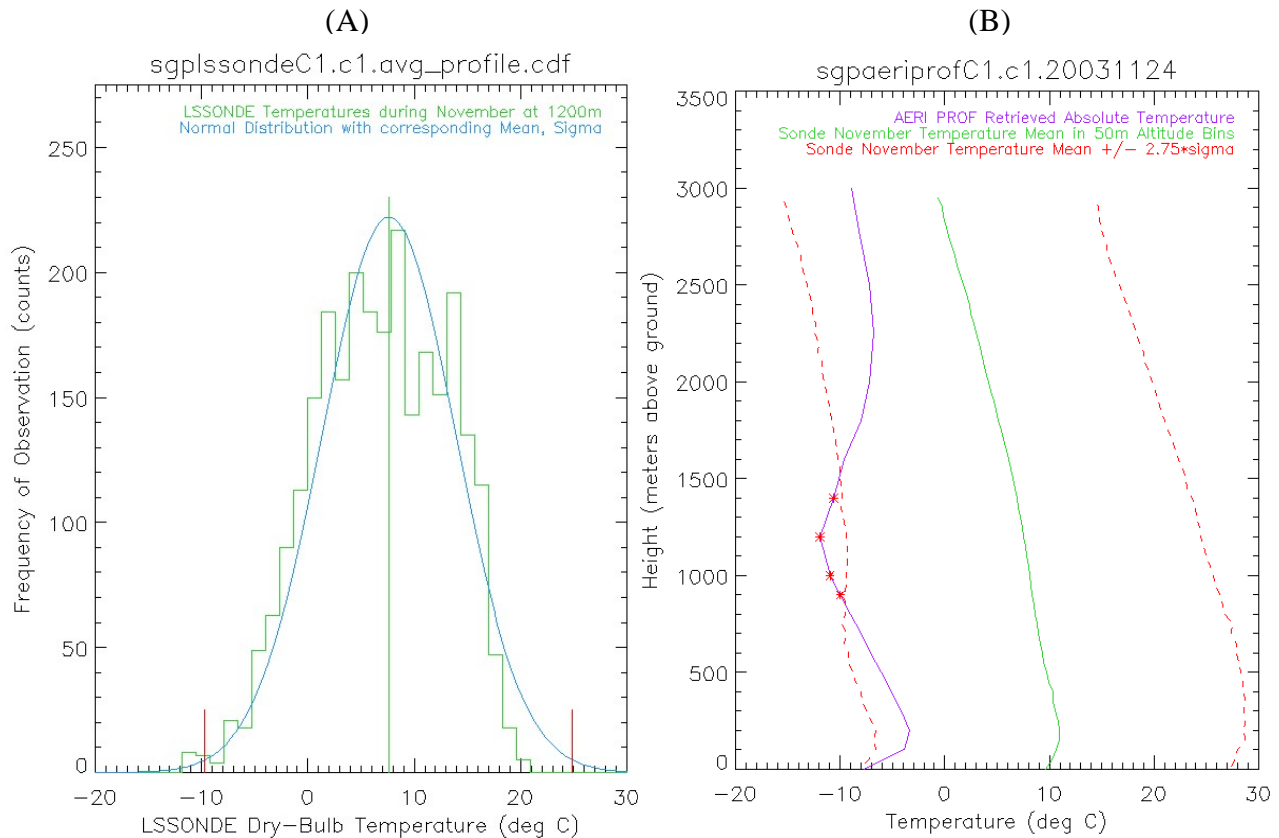
(A)   (B)

Figure 4. (a) Histogram of SGP LSSONDE profile temperatures for November, calculated in 50m altitude bins. (b) Application of outlier test to retrieved profiles that fall outside of height dependent bounds.

Analysis of AERI QME's (sgpqmeaerilbl, sgpqmeaeriprof, and others) has been to establish similar empirical distributions for each variable from historical data and then check new values against this historical database. The algorithm we use will flag values that are extreme with respect to the historical distribution of the data - which is very useful to detect when the QME product is either not being computed correctly, or when there are serious instrument problems.

Corresponding Author

S. Moore, moore@mrcsb.com, (805) 963-8761

References

ARM Science Plan, Ch. 11: Data Management and Documentation Plan, accessible at http://www.arm.gov/docs/documents/technical/sciplan/sp-11.pdf

Peppler, R., Sonntag, K., Dean, A., Shafer C., 2004: Quality of Atmospheric Radiation Measurement (ARM) Program Data, 14th Conference on Applied Climatology, 84th Annual AMS meeting, http://ams.confex.com/ams/pdfpapers/68127.pdf. Also see http://dq.arm.gov/cgi-bin/dqmenu.pl

For a complete list of NCVweb installations, visit http://arm.mrcsb.com/

The web site http://webmads.ssec.wisc.edu/ lists the acceptable ranges from most AERI related quantities. Also see http://cimss.ssec.wisc.edu/aeri/ for detailed information concerning the AERI instrument.