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

Session Title: Next Generation ARM Data Analysis Tools
Session Date: Wednesday, March 18, 2015
Session Time: 1:30–3:30pm
Summary Authors: Jim Mather, Giri Palanisamy, and Chitra Sivaraman

Description

ARM is striving to more closely integrate observations and model simulations through reconfiguration of instruments and the development of routine modeling, beginning at the SGP megasite. To support this integration of observations and models, it will be necessary to provide data tools to support the analysis of complex, multi-source, data sets. During this session, there will be a review of existing data processing support capabilities, illustrations of potential development activities, and discussion on needs and priorities in this area.

Main Discussion

The goals of the session were:

- to describe where ARM is heading with the expansion of complex instruments and the introduction of an internal modeling element and the implications that has for data processing
- to give some examples of current capabilities
- to discuss priorities among the current ideas we have regarding future develop and solicit input from the group regarding whether there are key elements that we have missed.

To accomplish these goals, we planned a combination of talks and discussion. The session began with a query to the audience about situations where data access limited their ability to effectively use data. This was followed by a series of talks. The talks included two use cases that were intended to provide illustrations of situations where data analysis tools are important. Ed Luke provided an overview of a visualization tool used to probe radar spectra data sets and Andrew Gettelman described the CESM model diagnostics tool kit. Then Giri and Chitra provided a current view of the ARM infrastructure and where the infrastructure could be improved/expanded to support use cases such as the two discussed earlier. Giri talked about the computing infrastructure while Chitra talked about tools for visualizing and processing data.

Key Findings

There are many parallels between the CESM diagnostics use case and the LES modeling framework that ARM is setting out to implement. We can learn a lot about requirements for the LES diagnostics framework by studying the CESM diagnostics workflow and others such as the FASTER, CAPT, and AMT modeling testbeds.

Issues

Issues raised by the session attendees prior to the talks in response to a question about impediments to using data included:

- For modeling work – have to convert ARM time to CF

- Immediately convert data to ascii
- Would be useful to have synchronized aircraft data and associated quicklooks
- Would be useful to be able to merge datastreams to aid instrument to instrument diagnostics (e.g. MFRSR/SKYRAD)
- Have to load radar data at the cluster in stages for processing.

However, in thinking about developing tools to aid data access, we were also cautioned about making the data too easy to get (e.g. best estimates) or people will not think before using.

Following the talks, it was noted that we should not over-engineer tools solution (eg. with complex user interfaces) up front – but focus on designing key components that can be configured as needed later.

Needs

It was suggested that we consider supporting users moving data to other major data centers (in addition to the ORNL data center) for specific applications.

Future Plans

We (the session organizers) received positive feedback on developing remote/centralized capabilities. The use cases clearly illustrated the need for these tools. Some cautionary comments also provided food for thought in terms of being careful about doing too much too fast. In terms of next steps – we plan to do a more detailed assessment of the use cases examined in this session along with several others (e.g. development of a "VAP" by a user). We propose to develop implementation plans for these use cases and determine what tools and facilities need to be developed to implement these processes and solicit reviews of these plans.

Action Items

- Identification and description of 3-5 use cases
- Development of implementation plans for each use case – including identification of required tools and resources.