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

Session Title: ARM Data Integrator (ADI)
Session Date: Tuesday, March 17, 2015
Session Time: 12:30–1:30pm
Summary Authors: Chitra Sivaraman

Description

Learn to consolidate, transform, and merge ARM NetCDF data —without downloading data or software. The ADI is available at the ARM Data Visualization Cluster and allows you to create NetCDF data that meets your specific needs using a custom grid (10 sec., 1 min. or 1 hr.) and organized (daily, monthly or yearly files) to best support your analysis.

Main Discussion

Demonstration of ADI.

ADI framework introduction, uses cases, transformation methods, consolidation, merging, and algorithm development. ADI is a framework to analyze time-series data. It is a suite of tools, libraries, data structures, and interfaces. ADI automates repetitive data preparation and production tasks. It decreases the time and costs to implement algorithm and enforces ARM standards. It documents dependencies and automates reprocessing.

i.arm.gov -> PCM. (Processing Configuration Manager)

Five steps to transform data.

1. Define location and processing interval.

2. Identify the inputs and variables.

3. Define coordinates.

4. Define output (or can be created at run time).

5. Run the script.

One can download the demo and follow instructions to run the demo on ARM research account.

https://github.com/ARM-DOE/adi-demo

C, IDL, Python are supported by ADI. There is consideration for R in the future.

Completed Efforts: Variable History Database, Missing Value/QC bit analysis. Ongoing high priority tasks: Support ascii input and output formats, re-implement PCM, design the implementation of variable history.

Needs

1. The building of the process in menu driven, then switched to Unix mode. Which are users supposed to use? Can you use it outside of Unix?

   This could be a need that can be built within ADI where the data is automatically created with a push of the button from the UI. An email could be generated as soon as the data is processed.

Concern: We need to make sure metrics is captured. This could increase support activities for the infrastructure team.

2. Can you access DQ flags? When you transform data ADI creates it is own QC and they were demonstrated with the plots.

3. Database to map different variables (temp, temperature, T, t). Can users have access to this database? Eventually, yes. We have information about the history of the variables in JSON files. Our next step is to design the process of implementing this history.

4. How do you know which met DOD applies to the data you want to process? It is complicated but the developer can usually determine the DOD based on the site. I am not sure if the data consolidator dynamically changes the DOD based on the site.

5. Can this tool be used to make one "met" datastreams (smos + smet + mettwr)? Yes. Met has been reprocessed at most sites and they are now under one data stream.

6. What is the standard for averaging? Beginning, middle, end? Beginning of 2013 we are now using time bounds. CF convention. Time bounds is required if the data is averaged.

Future Plans

Next step: ADI@ archive; ADI on OSX

Action Items

None. Provided support for users who downloaded the demo and ran it on their desktop.