**ARM data-oriented diagnostics package for climate model evaluation**

Chengzhu Zhang

Shaocheng Xie

Lawrence Livermore National Laboratory

September 2016

Version 1.0

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. The IM release number is: LLNL-TR-704359.

**Contents**

**1. Introduction**

**2. Data Description**

**2.1 Observational Data Sets**

**2.2 CMIP Simulation Data**

**2.3 Data Limitation/Uncertainty**

**3. User’s guide**

**3.1 Package Overview / Flow Chart**

**3.2 Obtain ARM Diag**

**3.3 Set-up a Working Prototype**

**3.4 Diagnostics Examples**

**3.5 Set-up a New Case**

**1. Introduction**

A Python-based diagnostics package is currently being developed by the ARM Infrastructure Team to facilitate the use of long-term high frequency measurements from the ARM program in evaluating the regional climate simulation of clouds, radiation and precipitation. This diagnostics package computes climatological means of targeted climate model simulation and generates tables and plots for comparing the model simulation with ARM observational data. The CMIP model data sets are also included in the package to enable model inter-comparison.

Basic performance metrics are computed to measure the accuracy of mean state and variability of climate models. The evaluated physical quantities include cloud fraction, temperature, relative humidity, cloud liquid water path, total column water vapor, precipitation, sensible and latent heat fluxes and radiative fluxes, with plan to extend to more fields, such as, aerosol and microphysics properties. Process-oriented diagnostics focusing on individual cloud and precipitation-related phenomena are also being developed for the evaluation and development of specific model physical parameterizations. The version 1.0 package is designed based on data collected at ARM Southern Great Plains (SGP) Climate Research Facility, with the plan to include data from ARM sites.

The diagnostics package is currently built upon standard Python libraries and additional Python packages developed by DOE (such as CDMS and UV-CDAT). The ARM diagnostic package is available publicly with the hope that it can serve as an easy entry point for climate modelers to compare their model with ARM data.

In this report, we first present the input data, which constitutes the core content of the diagnostics package in section 2; and a user's guide documenting the workflow/structure of the version 1.0 codes, and including step-by-step instruction for running the package in section 3.

**2 Observations and Model Data Description**

**2.1 Observation Data Sets**

The observational data currently used in this package is primarily from the data collected at the DOE’s ARM Climate Research Facility SGP site with its central facility located at Oklahoma, Lamont (36.6°N, 97.5°W). In order to compare with grid-box mean variables output from climate models, majority of the observational fields are from the ARM continuous forcing and evaluation data sets [Xie et al. 2004], which represent an average over a 3 by 3 degree grid box. The domain mean quantities are derived by merging observations from various instruments within the ARM SGP surface network. Details can be seen in Zhang et al. [2001] and Xie et al. [2004]. Here, we use long-term continuous forcing data sets available from 1999 to 2011 to build representative climatology.

Other than quantities available from the continuous forcing data sets, the cloud fraction vertical profiles provided by ARSCL Value-Added product is obtained through the ARM Best Estimate-cloud data product is also included.

Table 1. Observed quantities used in the evaluation

|  |  |  |
| --- | --- | --- |
| **Quantity** | **Data Product** | **Original Data Source** |
| Surface Temperature/ Humidity | Continuous forcing | SMOS, OKM, KAS mesonet stations |
| Temperature/Humidity profile | Continuous forcing | NOAA/ NCEP Rapid Update Cycle (RUC) analysis data |
| Surface Precipitation | Continuous forcing | Arkansas-Red Basin River Forecast Center (ABRFC)  Nexrad radar precipitation estimates w/ rain gauge |
| Surface Latent/Sensible Heat | Continuous forcing | Best-Estimate Fluxes From EBBR Measurements and Bulk Aerodynamics Calculations (BAEBBR) (EBBR: Cook et al. 2007) |
| Surface Radiative Fluxes | Continuous forcing | Data Quality Assessment for ARM Radiation Data (QCRAD) (Long and Shi 2006, 2008) |
| Cloud Fraction Vertical Profile | ARMBE cloud\* | Active Remotely-Sensed Cloud Locations (ARSCL) (Clothiaux et al. 2001) |

\*: Data product for c1 site.

**2.2 CMIP5 AMIP Simulations**

Simulations of 23 models contributing to the CMIP5 [Taylor et al., 2012] multi-model experiments have been used (see Table 2 for details). We evaluate these models from the CMIP5 atmospheric only (AMIP) experiments from year 1979 to 2008. All data have been linearly interpolated to a 3° x 3° domain with center located at SGP central facility located at Oklahoma, Lamont (36.6°N, 97.5°W) to make them comparable to the continuous forcing product.

Table 2 Models used in the evaluation

|  |  |
| --- | --- |
| Modeling groups | Model name |
| Commonwealth Scientific and Industrial Research Organization and Bureau of Meteorology (BOM), Australia | ACCESS1.0  ACCESS3.0 |
| Beijing Climate Center, China Meteorological Administration | BCC-CSM1.1  BCC-CSM1.1(m) |
| College of Global Change and Earth System Science, Beijing  Normal University | BNU-ESM |
| Canadian Centre for Climate Modelling and Analysis | CanAM4 |
| National Center for Atmospheric Research | CCSM4 |
| Community Earth System Model Contributors | CESM1-CAM5 |
| Commonwealth Scientific and Industrial Research Organization in collaboration with Queensland Climate Change Centre of Excellence | CSIRO-Mk3-6-0 |
| LASG, Institute of Atmospheric Physics, Chinese Academy of  Sciences and CESS, Tsinghua University | FGOALS-g2  FGOALS-s2 |
| NOAA Geophysical Fluid Dynamics Laboratory | GFDL-HIRAM-C360  GFDL-HIRAM-C180 |
| NASA Goddard Institute for Space Studies | GISS-E2-R |
| Met Office Hadley Centre | HadGEM2-A |
| Institut Pierre-Simon Laplace | IPSL-CM5A-LR  IPSL-CM5B-LR  IPSL-CM5A-MR |
| Institute for Numerical Mathematics | Inmcm4 |
| Atmosphere and Ocean Research Institute, National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology | MIROC5 |
| Max Planck Institute for Meteorology | MPI-ESM-MR  MPI-ESM-LR |
| Norwegian Climate Centre | NorESM1-M |

\* Note that for certain quantities, especially for sub-monthly output variables; only subsets of models are available for analysis.

**2.3 Data Limitation/Uncertainty**

As an evaluation data product, the availability of the variables has to depend on the input data streams; therefore it is recommended that scientific applications be cautious for the measurement uncertainty stems from the choice of input data streams. For instance, for surface turbulence fluxes, BAEBBR is the original data stream for calculating climatology, however, it has been recognized that when comparing with QCECOR data, BAEBBR has smaller surface sensible heat flux and larger surface latent heat flux. Land type where the ECOR and EBBR stations are located is believed to cause the difference. We plan to update the data when a domain-mean QCECOR or a merged product is available. In addition, for quantities those are inquired by the modeling community (i.e., aerosol, cloud properties), we will evaluate the feasibility of adding these, with the hope to incorporate those into the next release of the package.

**3. User’s Guide**

**3.1 Package Overview / Work Flow**

Figure 1 illustrates the flowchart of creating the diagnostic results by applying the diagostics tool. The steps are straightforward. the step-by-step procedure to set-up a working prototype is presented in section 3.

Figure 1 Work flow of the diagnostics package

The project has the following structure:

arm-gcm-diagnostics/

|-- ARMDiag\_driver.py

|-- LICENSE.txt

|-- README.md

|-- config.py

|-- setup.py

|-- ARMDiag

| |-- \_\_init\_\_.py

| |-- cmip

| |-- figures

| |-- html

| |-- metrics

| |-- model

| |-- observation

| |-- samples

| |-- source

| |-- AC\_DC\_cl\_p\_plot.py

| |-- AC\_amip\_contour\_html.py

| |-- AC\_mean\_amip\_data.py

| |-- AC\_mean\_amip\_line\_taylorD\_html.py

| |-- AC\_mean\_amip\_plot.py

| |-- AC\_mean\_amip\_table.py

| |-- AC\_mean\_amip\_taylorD\_plot.py

| |-- DC\_amip\_contour\_html.py

| |-- DC\_amip\_line\_harmonicD\_plot.py

| |-- DC\_amip\_line\_html.py

| |-- DC\_mean\_amip\_line\_harmonicD\_html.py

| |-- Daily\_amip\_PDF\_plot.py

| |-- Daily\_amip\_PDF\_plot\_html.py

| |-- \_\_init\_\_.py

| |-- \_\_init\_\_.pyc

| |-- taylorD.py

| |-- write\_html.py

**3.2 Obtain ARM Diag**

ARM Diag v1 with basic sets of diagnostics is now publicly available. The data files including observation and CMIP5 model data are available through ARM archive. The analytical codes to calculate and visualize the diagnostics results are placed via repository (arm-gcm-diagnostics) at https://github.com/ARM-DOE/

For downloading data:

* Click https://www.arm.gov/data/eval/123
* Following the Data Directory link on that page, it will lead to the area that the data files are placed. A short registration is required if you do not already have an ARM account.
* DOI for the citation of the data is 10.5439/1282169

For obtaining codes:

$ git clone https://github.com/ARM-DOE/arm-gcm-diagnostics/

**3.3 Set-up a working prototype**

A working prototype has been set up for the users to run the package out-of-the-box. In this case, all the observation and CMIP data are already placed under directoris: <Your directory>/ARMDiag/observation and <Your directory>/ARMDiag/cmip, respectively, in the format of comma-separated values (csv). The demo model data are placed under <Your directory>/ARMDiag/model.

To run the package, simply type in the terminal the following:

$ python ARMDiag\_driver.py

To view the diagnostics results:

For Mac OS:

$ open <Your directory>/ARMDiag/html/ARM\_diag.html

For Linux:

$xdg-open <Your directory>/ARMDiag/html/ARM\_diag.html

**3.4 Diagnostics examples**

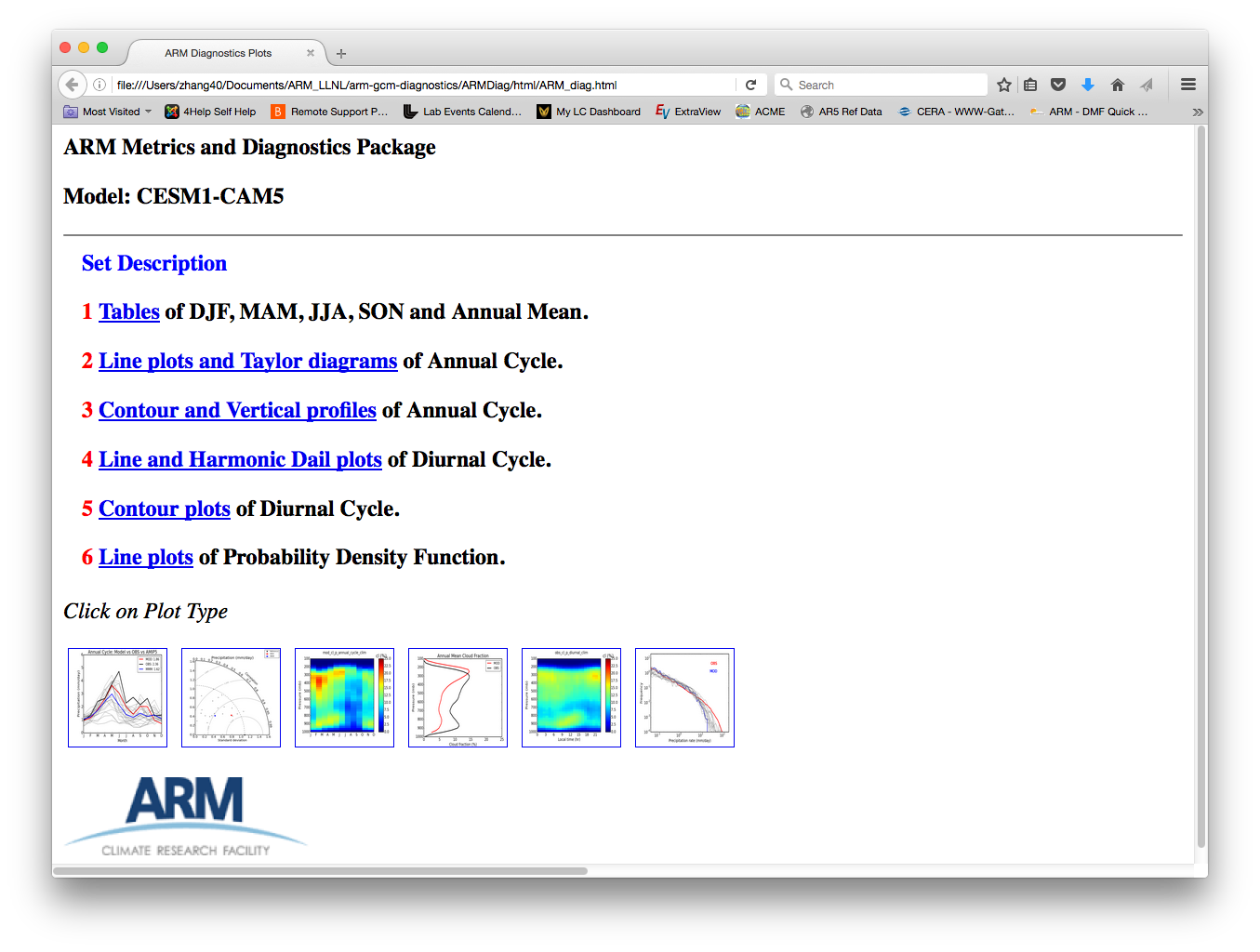
Below shows the main html page hosting the results:

Figure 2 Main html page generated to host the diagnostic results

In this release, the package provides 6 sets of diagnostics including:

* Tables summarizing DJF, MAM, JJA, SON and Annual Mean climatology using monthly output (Figure 3)
* Line plots and Taylor diagrams diagnosing annual cycle using monthly output (Figure 4)
* Contour and vertical profiles of annual cycle for quantities with vertical distribution (i.e., cloud fraction)
* Line plots of diurnal cycle for quantities without vertical distribution (i.e., precipitation)
* Contour plots of diurnal cycle for quantities with vertical distribution
* Line plots of Probability Density Functions using daily output

Among above diagnostics sets, the first two sets are most complete in the sense of the availability of models and evaluated quantities. For the other sets of diagnostics, the climatology variability is calculated based on sub-monthly model output, therefore model data availability is relatively low. In order to enable process-level study, we will put emphasis on the development of sub-monthly diagnostics in future work.

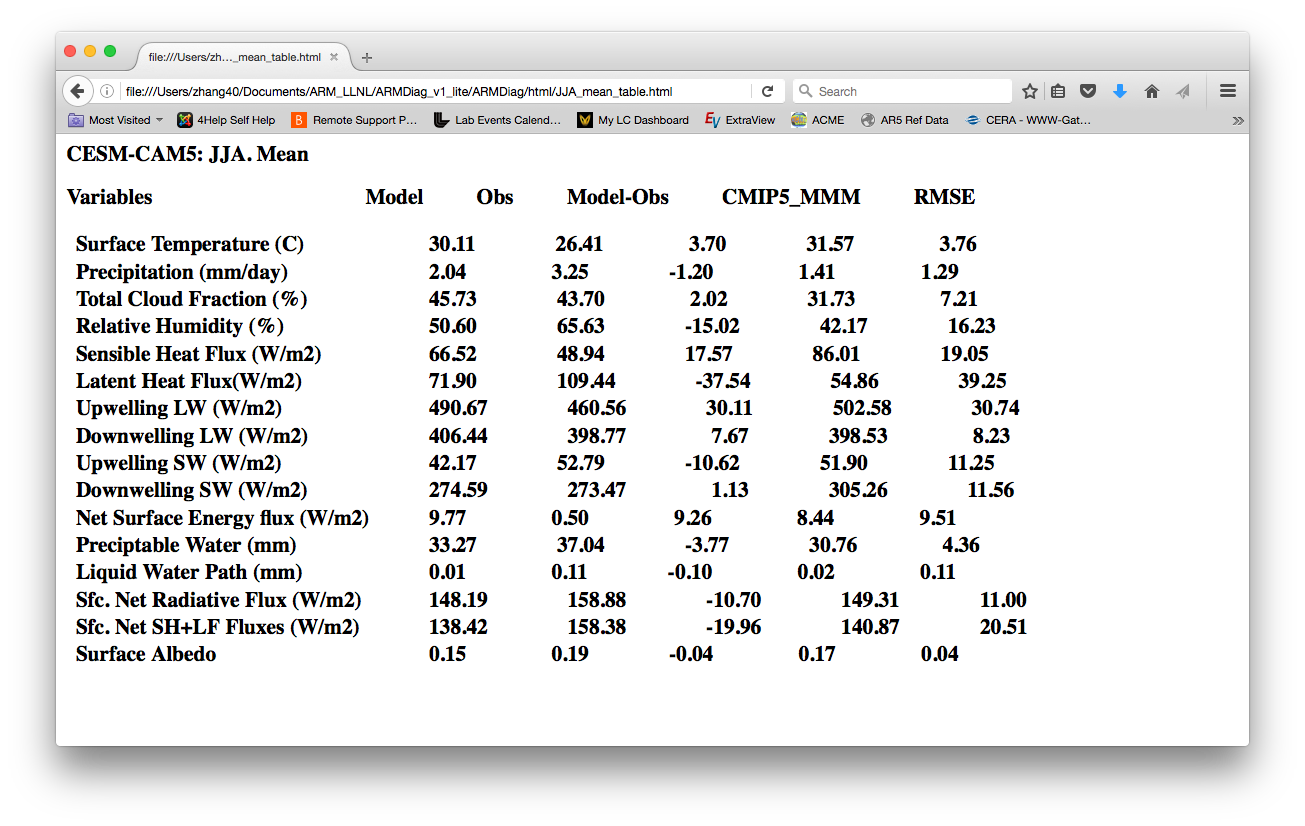


Figure 3 Tables summarizing JJA mean climatology

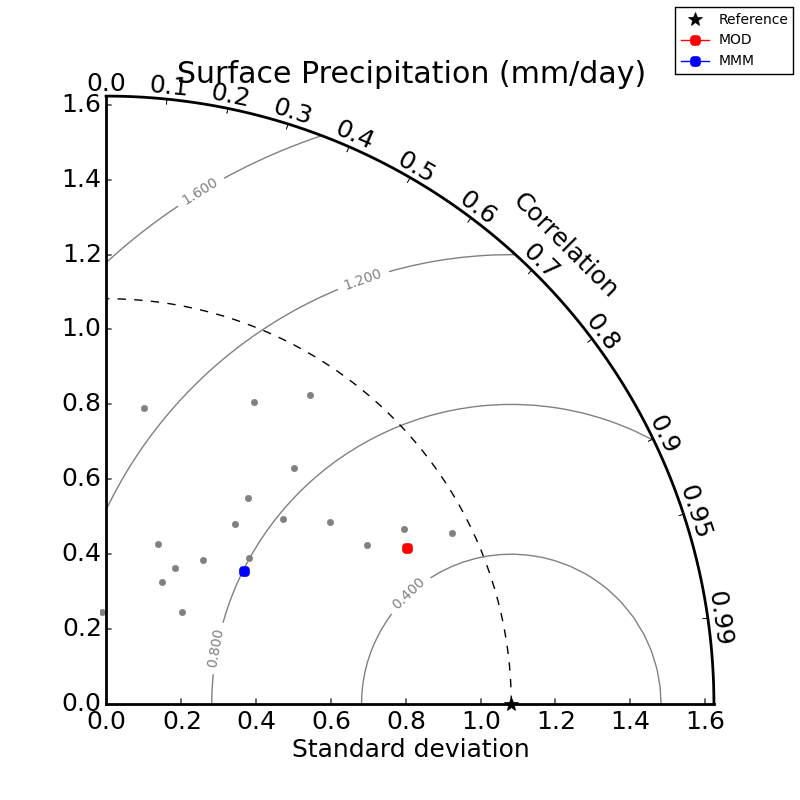
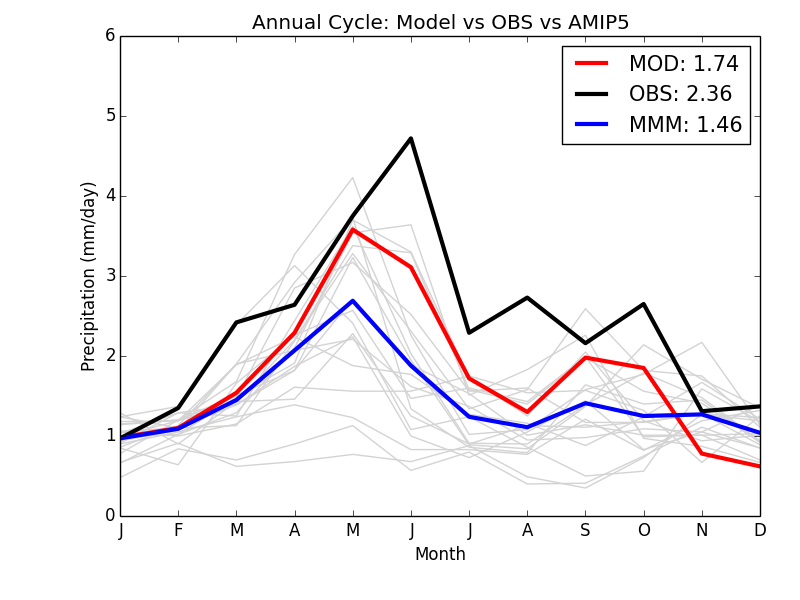


Figure 4 Line plots and Taylor diagrams diagnosing annual cycle of precipitation

**3.5 Code sample**

Within the package, we provide several code samples (in the directory ARMDiag/samples/) for the users to pre-process their model results. Below we provide example codes to process the monthly mean model data to be evaluated into form that can be read by the package, with the assumption that model results follows CMIP5 standard output regulation (see http://cmippcmdi.llnl.gov/cmip5/data\_description.html) and the Ultrascale Visualization Climate Data Analysis Tools (UVCDAT) package is installed (see https://github.com/UV-CDAT/uvcdat/wiki/install for installation guide).

import cdms2, MV2,cdutil

import numpy as np

filename= ‘input\_filename.nc’

modelname =’input\_modelname’

#Variable

#For multiple variables loop over below codes.

var=’pr’

f\_in=cdms2.open(filename)

#locate ARM SGP sites

lat0=36.6

lon0=262.5

lat=[lat0-4,lat0+4]

lon=[lon0-4,lon0+4]

dattable=f\_in(var,latitude = lat, longitude=lon, time=('1979-01-01','2008-12-31'))

ingrid=dattable.getGrid()

#Regrid to 3x3 grid centered at SGP and save the data in netcdf

outgrid=cdm.createUniformGrid(lat0,1,3,lon0,1,3,order='yx')

dat\_regrid=dattable.regrid(outgrid,regridTool='libcf',regridMethod='linear')

outfile=’output\_filename.nc’

f\_out= cdms2.open(outfile,'w')

f\_out.write(dat\_regrid)

#Convert the data into csv format

pr =f\_out(var)

pr = [x \*3600\*24 for x in pr]

pr\_yr=np.reshape(pr,(len(pr)/12,12))

pr\_ac=np.nanmean(pr\_yr,axis=0)

np.savetxt(basedir+'model/'+var+'\_model\_regrid\_3x3\_correct.csv',pr\_ac[:,:],fmt='%.3f')

**3. 5 Set-up a New Case**

* Follow sample codes and data name convention to generate model data and then place the processed data in model data directory: ARMDiag/model
* Edit config.py to change model's name accordingly
* Run the package by typing:

$ python ARMDiag\_driver.py

**References:**

Cook, D. R., 2007: Energy Balance Bowen Ratio (EBBR) handbook.U.S. Department of Energy Tech. Rep. DOE/SCARM-TR-037, 26 pp.

Clothiaux, E. E., and Coauthors, 2001: The ARM millimeter wave cloud radars (MMCRs) and the active remote sensing of clouds (ARSCL) value added product (VAP). U.S. Department of Energy Tech. Memo. ARM VAP-002.1, 56 pp.

Kato, S., N. G. Loeb, F. G. Rose, D. R. Doelling, D. A. Rutan, T. E. Caldwell, L. Yu, and R. A. Weller (2013), Surface Irradiances Consistent with CERES-Derived Top-of-Atmosphere Shortwave and Longwave Irradiances, Journal of Climate, 26(9), 2719-2740, doi: 10.1175/jcli-d-12-00436.1.

Long, C. N. and T. P. Ackerman, (2000): Identification of Clear Skies from Broadband Pyranometer Measurements and Calculation of Downwelling Shortwave Cloud Effects, JGR, 105, No. D12, 15609-15626.

Long, C. N. and Y. Shi, (2006): The QCRad Value Added Product: Surface Radiation Measurement Quality Control Testing, Including Climatologically Configurable Limits, Atmospheric Radiation Measurement Program Technical Report, ARM TR-074, 69 pp.

Long, C. N., and Y. Shi, (2008): An Automated Quality Assessment and Control Algorithm for Surface Radiation Measurements, TOASJ, 2, 23-37, doi: 10.2174/1874282300802010023.

Long, C. N. and D. D. Turner (2008): A Method for Continuous Estimation of Clear-Sky Downwelling Longwave Radiative Flux Developed Using ARM Surface Measurements, J. Geophys. Res., 113, doi:10.1029/2008JD009936.

Taylor, K. E., Stouer, R. J. & Meehl, G. A. An overvie of CMIP5 and the experiment design. Bull. Amer. Meteor. Soc. 93, 485-498 (2012).

Wang, C., L. Zhang, S.-K. Lee, L. Wu, and C. R. Mechoso (2014), A global perspective on CMIP5 climate model biases, Nature Climate Change, 4(3), 201-205, doi: 10.1038/nclimate2118.

Xie, S. C., R. T. Cederwall, and M. H. Zhang (2004), Developing long-term single-column model/cloud system-resolving model forcing data using numerical weather prediction products constrained by surface and top of the atmosphere observations, Journal of Geophysical Research-Atmospheres, 109(D1), doi: 10.1029/2003jd004045.

Xie, S., et al. (2010), ARM CLIMATE MODELING BEST ESTIMATE DATA A New Data Product for Climate Studies, Bulletin of the American Meteorological Society, 91(1), 13-+, doi: 10.1175/2009bams2891.1.

Zhang, M. H., J. L. Lin, R. T. Cederwall, J. J. Yio, and S. C. Xie (2001), Objective analysis of ARM IOP data: Method and sensitivity, Monthly Weather Review, 129(2), 295-311, doi: 10.1175/1520-0493(2001)129<0295:oaoaid>2.0.co;2.