

ARM data-oriented metrics and diagnostics package for climate model evaluation

Chengzhu Zhang
Shaocheng Xie

Lawrence Livermore National Laboratory

July 2017
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 Test Case**
 - 3.4 Diagnostics Examples**

1. Introduction

A Python-based metrics and diagnostics package is currently being developed by the ARM Infrastructure Team at Lawrence Livermore National Laboratory 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 metrics and 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 as demonstrated in Zhang et al. (2017). The mean of the CMIP model can be served as a reference for individual models.

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 extend to other ARM sites.

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

In this report, we first present the input data, which constitutes the core content of the metrics and 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 used in this study are primarily from that collected at the DOE's ARM Climate Research Facility SGP surface network with its central facility located at Lamont, Oklahoma (36.6°N, 97.5°W). In order to compare with grid-box mean variables output from climate models, the majority of the observational fields are from the ARM continuous forcing evaluation data sets [Xie et al., 2004] which attempts to determine the spatial average for a region of approximately 3° latitude-longitude centered on the central facility. The long-term continuous forcing data sets are available from 1999 to

2011 that allow us to build representative climatologies. In this data set, the vertical profiles of the atmospheric state variables (temperature and specific humidity) are from the National Oceanic and Atmospheric Administration (NOAA) rapid update cycle (RUC) analysis, but are adjusted to conserve the column integrated mass, dry static energy, and moisture through a constrained variational analysis approach developed by Zhang and Lin [1997] and Zhang et al. [2001] using observed surface and Top-of-the-Atmosphere (TOA) fluxes as the constraints. The surface quantities include both radiation and turbulence fluxes, which are first interpolated into $0.5^\circ \times 0.5^\circ$ grids within the ARM SGP domain that covers a $3^\circ \times 3^\circ$ area (See Figure 1 from Tang et al., 2016)) before the domain mean is calculated.

Table 1 summarizes all data sets used and provides additional information on data sources and estimated uncertainties. From the continuous forcing product, the surface screen-level temperature and humidity are based on 9 Surface Meteorological Observation Stations (SMOS), 127 Oklahoma and 13 Kansas mesonet stations (OKM and KAM). Note that the number of stations varies with time. The precipitation rate is obtained from the Arkansas-Red Basin River Forecast Center (ABRFC) precipitation product, which provides hourly gridded (4 km x 4km) precipitation field by combining both WSR-88D Nexrad radar precipitation estimates and rain gauge reports, with the missing periods supplemented by the station data [Breidenbach et al., 1998, Fulton et al., 1998]. The column water vapor available in continuous forcing is derived from the microwave radiometer retrieval from the single ARM central facility station.

The derived all-sky radiative fluxes: including downwelling/upwelling shortwave and longwave radiative fluxes in the continuous forcing datasets are based on 14 radiometers in the Solar and Infrared Observation Stations (SIROS). The Data Quality Assessment for ARM Radiation Data (QCRAD) methodology is applied to use climatological analyses of the surface radiation measurements to control the quality of the data [Long and Shi, 2006].

The surface sensible heat and latent heat fluxes are measured at ARM's Energy Balance Bowen Ratio (EBBR) stations since 1993 and Quality Controlled Eddy Correlation fluxes (QCECOR) stations since 2003 [Berg and Lamb, 2016]. The vertical fluxes of sensible and latent heat produced by the EBBR systems are estimated from the vertical temperature and humidity gradients. The Bulk Aerodynamic technique is applied to the EBBR data streams (BAEBBR) to address sunrise and sunset spikes in the fluxes data [Cook, 2011a: EBBR handbook]. The ECOR technique estimates the vertical fluxes by correlating the vertical wind component with temperature (sensible heat flux) and humidity (latent heat flux) [Cook, 2011b: ECOR handbook]. The EBBR stations are often deployed at stable land, such as pasture and grassland, while QCECOR stations are usually at disturbed land such as cropland and wooded land. The multiyear monthly climatology of surface

latent and sensible heat fluxes is constructed by averaging over the measurements from available EBBR and QCECOR stations during the period from 1999 to 2011. Measurements from up to 19 EBBR and 13 QCECOR stations are used to calculate the domain mean.

Soil moisture data are from the Soil Water and Temperature Systems (SWATS) [Bond, 2005: SWATS handbook]. Two profiles of sensors are installed one meter apart from each other that perform measurement at eight different depths. To calculate the soil moisture variable equivalent to the model output variable (mrsos: soil moisture integrated over uppermost 10 cm layer), the volumetric soil moisture measured by two sensors for top 5 cm and 15 cm depths are averaged for each site. Data from a total of 22 sites are used for generating the domain mean climatology from 1999-2011.

The Aerosol Optical Depths (AODs) are from MultiFilter Rotating Shadowband Radiometer (MFRSR) deployed at ARM sites. A review of the utility of the narrowband radiometer can be found in Michalsky and Long [2016] and McComiskey and Ferrare [2016]. To compare with model output of AOD output at 550 nm (od550aer), the monthly mean AOD500 is extrapolated to AOD550 following,

$$AOD_{550} = AOD_{500} * (550 / 500)^{-\alpha}, \quad (1)$$

where the Ångström exponent, α is estimated using the AODs measured at the wavelengths 415nm and 615 nm. We also note that the AOD climatology obtained from the central facility can well represent the domain mean climatology calculated by averaging all available MFRSR stations.

Other quantities such as, the cloud fraction vertical profiles provided by ARSCL Value-Added product is also included. Retrieved properties such as liquid water path and ice water path climatology are generated from ACRED data product. The multiyear monthly climatology is constructed for all observed variables analyzed in this paper. The climatology of the observational datasets is formed for the period from 1999 to 2011, except for that of the variables from ACRED product which uses data available from 2002 to 2008.

Table 1. Observed quantities used in the evaluation

Quantities	Data Products	Data Source/ Instruments	Time resolution	Spatial info
------------	---------------	--------------------------	-----------------	--------------

Surface Screen-Level Temperature/ Humidity	Continuous forcing	Surface Meteorological Observation System (SMOS), Oklahoma and Kansas mesonet stations (OKM and KAM)[Tang et al. 2016]	Mo, da, hr	sgp domain averaged
Temperature/Humidity profile/wind speed/large scale tendencies	Continuous forcing	NOAA/ NCEP Rapid Update Cycle (RUC) analysis data [Tang et al. 2016]	Mo, da, hr	sgp domain averaged
Surface Precipitation	Continuous forcing	Arkansas-Red Basin River Forecast Center (ABRFC) Nexrad radar precipitation estimates w/ rain gauge [Breidenbach et al., 1998, Fulton et al., 1998]	mo, da, hr	sgp domain averaged
Precipitable Water	Continuous forcing	Microwave Radiometer (MWR) water liquid & vapor along line of sight (LOS) path (MWRLOS)	mo, da, hr	sgp domain averaged
Surface All Sky Radiative Fluxes	Continuous forcing	Data Quality Assessment for ARM Radiation Data (QCRAD) [Long and Shi, 2006, 2008]	mo, da, hr	sgp domain averaged
Aerosol Optical Depth 550nm	MFRSRAOD1 MICH	Multifilter Rotating Shadowband Radiometer (MFRSR) [Knootz et al.,2013]	mo	sgp Site C1 and E13 averaged
Surface Latent/Sensible Heat	BAEBBR	Best-Estimate Fluxes From EBBR Measurements and Bulk Aerodynamics Calculations (BAEBBR) [Cook, 2011a]	mo	sgp domain averaged
	QCECOR	Quality Controlled Eddy Correlation Flux Measurement [Cook, 2011b]	mo	sgp domain averaged
Surface Soil Moisture Content (10 cm)	SWATS	Soil Water and Temperature System (SWATS)) [Bond, 2005]	mo	sgp domain averaged
Cloud Fraction	ARSCL	Active Remote Sensing of Clouds [Clothiaux et al, 2001]	mo, da, hr	sgp Site C1
Ice Water Content/Liquid Water Content	ACRED	ARM Cloud Retrieval Ensemble Dataset [MACE and MICROBASE] [Zhao et al. 2012]	mo, da, hr	sgp Site C1

mo, da, hr: data are processed into monthly mean, daily mean and hourly mean.

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 $3^\circ \times 3^\circ$ 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),	ACCESS1.0 ACCESS3.0

Australia	
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

The ARM data used in the package have been gone through stringent data quality control and are representing the "best" estimate of the selected quantities. Fully addressing data uncertainty is a challenging task and ARM is making efforts to address this issue. More information will be provided once the uncertainty of these selected fields is better quantified. We recommend the user to read the references on the observational data products and contact PIs of each data product for more data quality information.

3. User's Guide

3.1 Package Overview / Work Flow

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

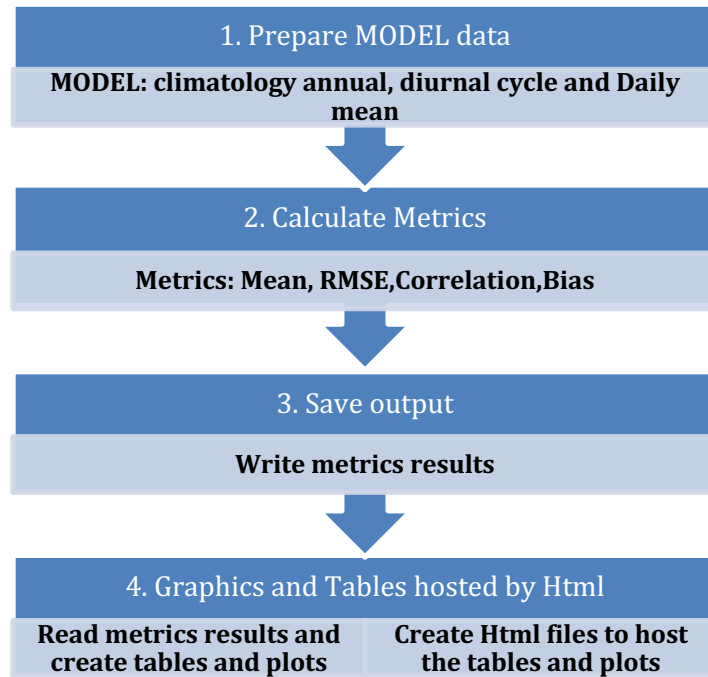


Figure 1 Work flow of the diagnostics package

The project has the following structure:

```

a | __arm_diags
  | | __.DS_Store
  | | __init__.py
  | | __arm_driver.py
  | | __arm_parameter.py
  | | __arm_parser.py
  | | __basicparameter.py
  | | __cmip
  | | __diags_all.json
  | | __examples
  | | | __diags_set1.json
  | | | __diags_set2.json
  | | | __diags_set3.json
  | | | __diags_set4.json
  
```



```

| | |__diags_set6.json
| | |__diags_sets.json
| | __misc
| | |__ARM_logo.png
| | __model
| | __observation
| | __src
| | |__init__.py
| | |__annual_cycle.py
| | |__annual_cycle_zt.py
| | |__create_htmls.py
| | |__diurnal_cycle.py
| | |__pdf_daily.py
| | |__seasonal_mean.py
| | |__taylor_diagram.py
| | |__varid_dict.py
|__ARM_gcm_diag_pkg_TechReport_v1.docx

```

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 test case

First to create a conda enviroment and then activate it:

```
$conda create -n arm_diags_env cdp cdutil genutil cdms2 numpy matplotlib
scipy -c conda-forge -c uvcdat
$source activate arm_diags_env
```

To install the package, cd <Your directory>/, type following:

```
$python setup.py install
```

A working test case has been set up for the users to run the package out-of-the-box. In this case, all the observation, CMIP data, test data should be downloaded placed under directoris:

```
<Your directory>/arm_diags/observation
```

```
<Your directory>/ arm_diags /cmip
```

```
<Your directory>/ arm_diags /model, respectively.
```

To configure basic parameter file: basicparameter.py and edit parameters such as, input and output paths, model name (used to search the file), and case name (to create a new folder for the case).

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

```
$ python arm_driver.py -p basicparameter.py
```

To view the diagnostics results:

For Mac OS:

```
$ open <Your directory>/arm_diags/case_name/html/ARM_diag.html
```

For Linux:

```
$xdg-open <Your directory>/ arm_diags/case_name/html/ARM_diag.html
```

For setting up customized runs and creating new cases, check details at:

<https://github.com/ARM-DOE/arm-gcm-diagnostics/>

3.4 Diagnostics examples

Below shows the main html page hosting the results:

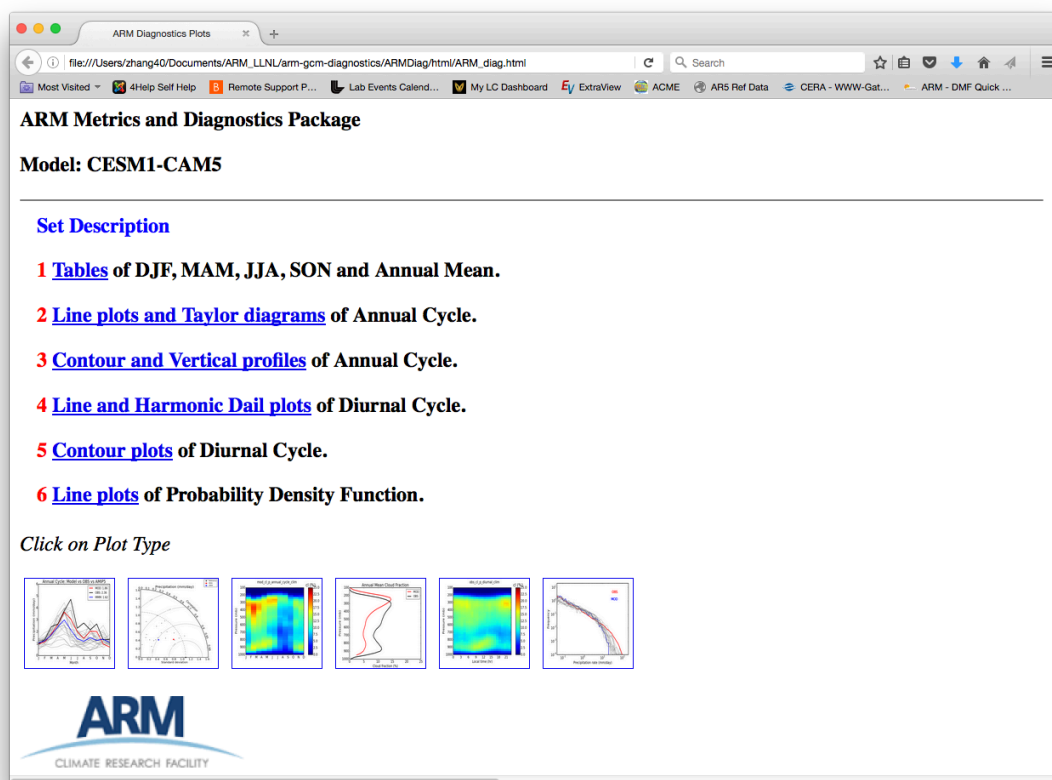


Figure 2 Main html page generated to host the diagnostic results

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.

file:///Users/zh..._mean_table.html

file:///Users/zhang40/Documents/ARM_LLNL/ARMDIAG_v1_lite/ARMDIAG/html/JJA_mean_table.html

Most Visited 4Help Self Help Remote Support P... Lab Events Calend... My LC Dashboard ExtraView ACME ARS Ref Data CERA - WWW-Gat...

CESM-CAM5: JJA. Mean

Variables	Model	Obs	Model-Obs	CMIP5_MMM	RMSE
Surface Temperature (C)	30.11	26.41	3.70	31.57	3.76
Precipitation (mm/day)	2.04	3.25	-1.20	1.41	1.29
Total Cloud Fraction (%)	45.73	43.70	2.02	31.73	7.21
Relative Humidity (%)	50.60	65.63	-15.02	42.17	16.23
Sensible Heat Flux (W/m2)	66.52	48.94	17.57	86.01	19.05
Latent Heat Flux(W/m2)	71.90	109.44	-37.54	54.86	39.25
Upwelling LW (W/m2)	490.67	460.56	30.11	502.58	30.74
Downwelling LW (W/m2)	406.44	398.77	7.67	398.53	8.23
Upwelling SW (W/m2)	42.17	52.79	-10.62	51.90	11.25
Downwelling SW (W/m2)	274.59	273.47	1.13	305.26	11.56
Net Surface Energy flux (W/m2)	9.77	0.50	9.26	8.44	9.51
Precipitable Water (mm)	33.27	37.04	-3.77	30.76	4.36
Liquid Water Path (mm)	0.01	0.11	-0.10	0.02	0.11
Sfc. Net Radiative Flux (W/m2)	148.19	158.88	-10.70	149.31	11.00
Sfc. Net SH+LF Fluxes (W/m2)	138.42	158.38	-19.96	140.87	20.51
Surface Albedo	0.15	0.19	-0.04	0.17	0.04

Figure 3 Tables summarizing JJA mean climatology

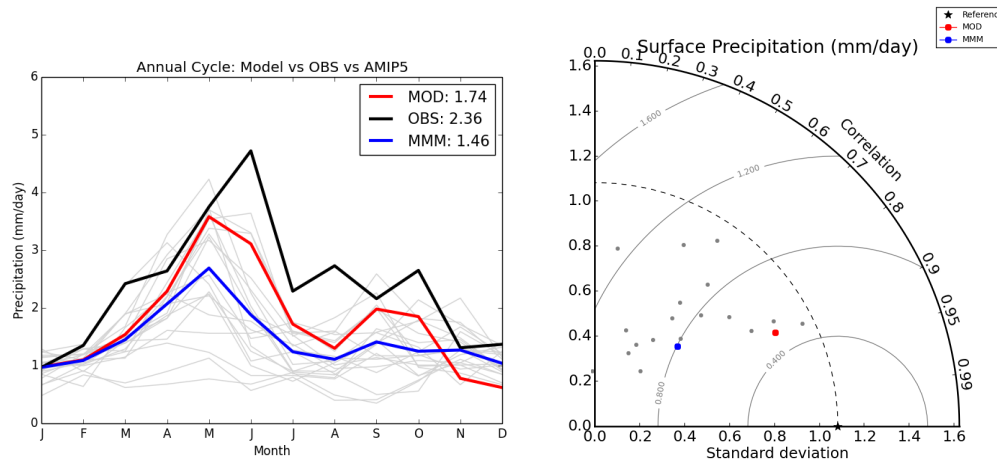


Figure 4 Line plots and Taylor diagrams for diagnosing annual cycle of precipitation

References:

- Berg, L. K., and Lamb, P. J. (2016). Surface Properties and Interactions: Coupling the Land and Atmosphere within the ARM Program. The Atmospheric Radiation Measurement (ARM) Program: The First 20 Years, Meteor. Monogr., No. 57, Amer. Meteor. Soc., doi: <http://dx.doi.org/10.1175/AMSMONOGRAPHIS-D-15-0044.1>
- Bond, D. (2005), Soil Water and Temperature System (SWATS) Handbook, ARM Technical Report TR-063, U.S. Department of Energy, Washington, D. C.
- Breidenbach, D. J. Seo, Richard Fulton. 1998. Stage II and III Post Processing of NEXRAD Precipitation Estimates in the Modernized Weather Service. AMS 78th Annual Meeting Phoenix, Arizona, January 1998.
- Cook, D. R. (2011a). Energy Balance Bowen Ratio Station (EBBR) Handbook (No. DOE/SC-ARM/TR-037). DOE Office of Science Atmospheric Radiation Measurement (ARM) Program (United States).
- Cook, D. R. (2011b). Eddy Correlation Flux Measurement System (ECOR) Handbook (No. DOE/SC-ARM/TR-052). DOE Office of Science Atmospheric Radiation Measurement (ARM) Program (United States).
- 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.
- Fulton, R.A., J.P. Breidenbach, D.J. Seo, D.A. Miller, and T. O'Bannon (1998), The WSR-88D Rainfall Algorithm. Wea. Forecasting, 13, 377–395.
- 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.
- Knoutz, et al. (2013), Aerosol Optical Depth Value-Added Product, DOE/SC-ARM/TR-129, USDOE Office of Science, Washington, D. C.
- 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.
- McComiskey, A. and Ferrare, R.A., 2016. Aerosol physical and optical properties and processes in the ARM Program. The Atmospheric Radiation Measurement (ARM) Program: The First 20 Years, Meteor. Monogr., No. 57, Amer. Meteor. Soc., <http://dx.doi.org/10.1175/AMSMONOGRAPHIS-D-15-0028.1>
- Michalsky, J. J., and C. N. Long, 2016: ARM solar and infrared broadband and filter radiometry. The Atmospheric Radiation Measurement (ARM) Program: The First 20 Years, Meteor. Monogr., No. 57, Amer. Meteor. Soc., doi:10.1175/AMSMONOGRAPHIS-D-15-0031.1.
- 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.
- Tang, S., M. Zhang, and S. Xie (2016), An ensemble constrained variational analysis of atmospheric forcing data and its application to evaluate clouds in CAM5, Journal of Geophysical Research: Atmospheres, 121(1), 33–48, doi: 10.1002/2015JD024167.
- Taylor, K. E., Stouffer, R. J. & Meehl, G. A. An overview 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, C., Xie, S., Klein, S. A., Ma, H.-Y., Tang, S., Van Weverberg, K., Morcrette, C., Petch, J. (2017). CAUSES: Diagnosis of the Summertime Warm Bias in CMIP5 Climate Models at the ARM Southern Great Plains Site. Journal of Geophysical Research, Submitted.
- Zhang, M., and J. Lin (1997), Constrained variational analysis of sounding data based on column-integrated budgets of mass, heat, moisture, and momentum: Approach and application to ARM measurements, J. Atmos. Sci., 54(11), 1503–1524.
- 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:oaoid>2.0.co;2.
- Zhao, C., Xie, S., Klein, S. A., Protat, A., Shupe, M. D., McFarlane, S. A., ... & Hogan, R. J. (2012). Toward understanding of differences in current cloud retrievals of ARM ground - based measurements. Journal of Geophysical Research: Atmospheres, 117(D10).