A Comparison of MERRA and NARR Reanalysis Datasets with the DOE ARM SGP Continuous Forcing data

Aaron D. Kennedy, Xiquan Dong, and Baike Xi

Department of Atmospheric Sciences, University of North Dakota, Grand Forks, ND

Shaocheng Xie and Yunyan Zhang

Lawrence Livermore National Laboratory, Livermore, CA

Junye Chen

ESSIC, University of Maryland at College Park

To be submitted to J Climate (MERRA reanalysis special issue)

Corresponding author address: Mr. Aaron Kennedy, The Department of Atmospheric Sciences, University of North Dakota, 4149 University Ave., Box 9006, Grand Forks, ND 58202-9006. Email: aaron.kennedy@und.edu. Phone: 701-777-4478.

Abstract

In this study, the atmospheric state, precipitation, cloud fraction, and radiative fluxes from Modern Era Retrospective-analysis for Research and Applications (MERRA) and North American Regional Reanalysis (NARR) are collected and compared with the ARM SGP continuous forcing during the period 1999-2001. For the atmospheric state, the three datasets have excellent agreement for the horizontal wind components and air temperature. NARR and ARM have generally good agreement for humidity, except for several biases in the PBL and in the upper troposphere. MERRA, on the other hand, suffers from a year-round negative bias in humidity except for the month of June. For the vertical pressure velocity, significant differences exist with the largest biases occurring during the spring upwelling and summer downwelling periods. Although NARR and MERRA share many resemblances to each other, ARM outperforms these reanalyses in terms of correlation with cloud fraction. Because the ARM forcing is constrained by observed precipitation that gives the adequate mass, heat, and moisture budgets, much of the precipitation (specifically during the late spring/early summer) is caused by smaller-scale forcing that is not captured by the reanalyses. Both NARR and MERRA capture the seasonal variation of CF observed by ARM radar-lidar and GOES with high correlations (0.92-0.78), but having negative biases of 14% and 3%, respectively. Compared to the ARM observations, MERRA shows a better agreement for both SW and LW fluxes except for LW-down (due to a negative bias in water vapor), NARR has significant positive bias for SW-down and negative bias for LW-down under clear- and all-sky conditions . The NARR biases result from a combination of too few clouds and a lack of sufficient extinction by aerosols and water vapor in the atmospheric column. The results presented here represent only one location for a

limited time period, and more comparisons at different locations and longer time period are needed.

1. Introduction

In the past decade, reanalysis datasets have become increasingly common to study a variety of meteorological and climatological questions. Reanalyses blend observation and model output to create a systematic long-term description of the climate system. While it is an excellent strategy to use model output to fill holes in the observing systems and to diagnose variables unable to be measured directly, reanalyses are not error free due to the limitations of model and assimilation technology. Because the errors of reanalyses and their underlying models are relatively unknown, their benefit for answering more complex questions involving the climate is questionable. For this reason, reanalyses have been used sparingly to generate forcing which provides initial and boundary conditions for SCM/CRM studies which can help develop improvements for GCMs.

To circumnavigate these issues, extensive work has been done to derive forcing using constrained variational analysis from observations during Intensive Observation Periods (IOPs) at the Department of Energy (DOE) Atmospheric Radiation Measurement (ARM) sites (Zhang and Lin 1997, Zhang et al. 2001). More recently, Xie et al. (2003) evaluated the forcing datasets derived from ECMWF during three IOPs at the ARM SGP site. They found that although the two forcing datasets correlated well, the ECMWF derived forcing was much weaker owing to limitations in the model predicated surface radiation and precipitation fields. Unfortunately, IOPs are expensive to run from a monetary and work-load perspective. Continuously run models, however, offer long-term datasets which are valuable from a climate study perspective. To combine the benefits of long-term model results and high-quality IOP observations, Xie et al. (2004) developed a continuous forcing dataset using a combination of model (atmospheric state variables such as temperature, humidity, etc.) from Rapid Update Cycle 2 (RUC-2, Benjamin et al. 2004) and surface and TOA observations at the ARM SGP site. The end result is a forcing dataset that improves considerably on that derived from the model alone and offers itself as a good baseline to judge reanalyses.

This paper documents a comparison of the NCEP North American Regional Reanalysis (NARR, Messinger et al. 2006) and the Modern Era Retrospective Analysis for Research and Applications reanalysis (MERRA, Bosilovich et al. 2008) with the ARM continuous forcing dataset derived at the ARM SGP site during the period 1999-2001. The ARM SGP site is representative of continental climate in the mid-latitudes, and has been used in the past to evaluate a variety of model simulations including NCEP ETA (Hinkelman et al. 1999), ECMWF (Xie et al. 2004), and the NCEP GFS (Yang et al. 2006). NARR and MERRA reanalyses were chosen for this comparison for a couple of reasons. First of all, NARR includes assimilation of precipitation at a high resolution over North America and has shown improvement over the NCEP Global Reanalysis II for a variety of variables (Messinger et al. 2006). MERRA has been included because it features relatively high resolution diagnostics output during the same time period, and was released within the past year. As a result, relatively little is known about its quality.

By comparing these three datasets, this paper has the primary goal of determining the biases of the reanalyses at a location which is well observed. Such activities have been encouraged by recent studies such as Thorne and Vose (2010) which have sought to understand whether reanalyses can be used for diagnosing long-term trends. Determining biases in reanalyses will also help understand where deficiencies exist in the current underlying model parameterizations. Knowing the magnitude, when, and where reanalysis errors exist will shed light on whether developing forcing from reanalyses in the well observed mid-latitudes can be a fruitful effort and aid others who may require reanalysis information for other studies.

This paper is formatted as follows. Section 2 gives a brief summary of the various datasets used in this study. In section 3, the atmospheric state is compared between the reanalyses and the ARM continuous forcing during the period 1999-2001. Cloud fraction, total precipitation, and radiative fluxes are compared in section 4. A summary of findings and concluding remarks are provided in section 5.

2. Datasets

ARM continuous forcing, NARR, and MERRA reanalysis data sets have been collected at the ARM SGP site for the period 1999-2001. These three years were chosen because the ARM continuous forcing dataset is only available during this time period. To have cloud information at the ARM SGP site, surface observations from a vertically pointing cloud radar and micro pulse lidar pair have also been collected along with Geostationary Operational Environmental Satellites (GOES) observations. All datasets have been processed to identical temporal and spatial resolutions for comparison in sections 3 and 4. For example, the results from the two reanalyses are averaged in space to the domain of the ARM forcing, while the hourly continuous forcing is averaged in time to three hourly increments to match the reanalyses.

a. ARM Continuous Forcing

The ARM continuous forcing dataset centered on the ARM SGP Central Facility (SCF; 36.6^{o}N, 97.5^{o}W) is used for this study. Provided from January 1999 to December 2001, this forcing uses ARM surface and GOES-8 satellite observations as constraints to adjust atmospheric state variables to conserve the column integrated mass, heat, and moisture through a variational analysis approach (Zhang and Lin 1997, Zhang et al. 2001). The forcing atmospheric state is provided by hourly Rapid Update Cycle 2 (RUC-2; see Benjamin et al. 2004) analyses due to the lack of continuous sounding measurements (Xie et al. 2004). A comparison of the continuous forcing with selected IOPs by Xie et al. (2004) found root-mean-square errors within 1 m s^{-1} for horizontal wind, 0.5 K for temperature, and 0.5 g kg^{-1} for moisture for the atmospheric column. The forcing represents an average over a circular area approximately 180 km in radius centered on the ARM SCF.

b. NARR Reanalysis

The NCEP NARR is a long-term (1979-2009) climate dataset with 3-hr temporal, 32-km horizontal, and 45-layer vertical resolutions over the North American domain (Messinger et al. 2006). It contains outputs of many atmospheric variables and fluxes, and is nicely suited for diagnosis of synoptic and mesoscale conditions over the ARM SGP site. NARR uses the operational NCEP ETA model and its 3D-VAR data assimilation technique on a wide variety of observation platforms, but was principally developed to improve on NCEP reanalysis by assimilating precipitation accurately. Studies by Becker et al. (2009) and Bukovsky and Karoly (2006) found that this statement is generally true for NARR.

c. MERRA Reanalysis

NASA has recently released the Modern Era Retrospective Analysis for Research and Applications (MERRA) reanalysis dataset based on the Goddard Earth Observing System data Analysis System Version 5 (GEOS-5 DAS, Bosilovich et al. 2008). This global reanalysis covers the same time period as NARR (1979-current). MERRA takes advantage of a variety of recent satellite data streams, for example, the observations from the NASA Earth Observing System (EOS), to improve the representation of the Earth's energy and water cycles. GEOS-5 includes the GEOS-5 AGCM and the Gridpoint Statistical Interpolation (GSI) atmospheric analysis developed jointly with NOAA/NCEP/EMC. Incremental Analysis Update (IAU) technique (Bloom et al. 1996) is incorporated in the GEOS-5 to minimize the 6 hourly shock from the observation input. The model has a native spatial resolution of 72-layers in the vertical, and 2/3°×1/2° in the horizontal. In addition to the 6 hourly 3 dimensional analyses at the native spatial resolution, MERRA also provides 1 hourly 2 dimensional diagnostics at 2/3°×1/2° resolution and 3 hourly 3 dimensional diagnostics at 1.25°×1.25° resolution on 42 vertical levels.

d. Cloud observations

For several portions of the study, cloud information is used to determine its relationships with atmospheric state and to determine clear-sky radiative fluxes. Cloud information comes from two sources. Ground-based observations from the ARM 35-GHz Millimeter Wavelength Cloud Radar (MMCR, Moran et al. 1998) are combined with a Belfort laser ceilometer and Micropulse Lidar (MPL) to determine cloud bases, tops, and vertical distributions. While information is collected at 5-min intervals, it has been binned to one hour cloud fractions (CF) at the resolution of the forcing in a fashion identical to that described in Xi et al. (2010) and Kennedy et al. (2010). This cloud product is similar to The Active Remote Sensing of Clouds (ARSCL, Clothiaux et al. 2000) cloud product except the original data stream is the MACE PI product (Mace et al. 2006) which merges the original radar modes differently. Considering cloud information is only used at a 1-3 hourly resolution, the differences should between the two products is negligible.

The second source of cloud information is total cloud fractions derived from VISST (Visible Infrared Solar-Infrared Split-window Technique) retrieved satellite cloud products (Minnis et al. 2001) using algorithms developed for the NASA Clouds and Radiant Energy

System (CERES) project. Cloud properties are retrieved from half-hourly, 4-km 0.65, 3.9, 10.8 (infrared, IR), and 12.0-µm radiances taken by GOES-8. Cloudy pixels are identified using an adaptation of the method described by Minnis et al. (2008a). The areal fraction of clouds (or the amount when present, AWP) is the ratio of the number of pixels classified as cloudy to the total number of pixels within a specified area. Cloud fraction is then calculated at the resolution of the forcing by considering the quantity of 0.5°×0.5° grid boxes contained within the area of interest. Once again, this methodology is consistent with that used in the Xi et al. (2010) and Kennedy et al. (2010) studies. The reader is referred to these publications for additional details on the process.

3. Atmospheric State

NARR and MERRA reanalyses are first compared to ARM continuous forcing by evaluating the yearly and seasonal column averaged biases for atmospheric state variables including horizontal wind components, specific humidity, vertical pressure velocity (omega), and air temperature (Table 1). Considering all three datasets take into account analyzed fields from observations such as upper air soundings and surface observation networks, it is of no surprise to find that biases are quite small for many of the variables. For example, biases for horizontal wind components are less than 0.5 m s^{-1} and for temperature, reanalyses are within 0.13 K of the forcing. Although NARR shows good agreement with the ARM forcing for specific humidity (within 0.04 g kg^{-1}), MERRA has a dry bias an order of magnitude larger with values ranging from -0.17 g kg^{-1} during autumn to -0.8 g kg^{-1} during winter. The largest disagreement amongst the datasets occurs for the vertical pressure velocity with positive biases ranging from 0.07 to 0.54 mb hr^{-1} which are larger than the yearly and seasonal means.

Both specific humidity and vertical pressure velocity are crucial for developing accurate forcing required by SCM/CRM applications. For example, biases in the humidity field will directly translate to biases in cloud simulations for these models since stratiform cloud parameterizations often consider humidity to trigger cloud. For this reason and for the fact that vertical velocities are difficult to measure directly, these two variables warrant additional investigation. In doing so, it may be possible to investigate whether the reanalyses have issues within their own parameterizations.

The seasonal variations of RH and omega derived from the ARM continuous forcing and the NARR and MERRA reanalyses over the ARM SGP site during the period 1999-2001 are provided in Fig. 1. As illustrated in Figures 1a and 1b, the RH values derived from ARM and NARR are in excellent agreement and have a bimodal distribution with peaks in the boundary layer and in the upper troposphere. Although not shown, this is consistent with the seasonal variation of radar-lidar derived cloud fraction at the ARM SGP site (Kennedy et al. 2010). The decrease in RH during the late summer (August-September) is primarily due to the influence of large-scale ridging and a lack of baroclinic wave activity over Oklahoma. Some RH differences between ARM and NARR exist near the top of the troposphere during the summer and in the boundary layer throughout the year. The former of these two differences may be an issue with RUC-2 as there is no physical explanation for a peak at this level during the summer months. Despite these differences, monthly maximums are present in both datasets, especially during January and March. MERRA captures the general shape of RH at the ARM SGP site (Fig. 1c), but with a ~5% negative bias throughout the year in the upper troposphere except during the late spring and early summer when convection is most common at the ARM SGP site. During this time period, MERRA has a considerable positive bias (~10-15%) compared to ARM and NARR.

Seasonal RMSE plots (not shown) demonstrate that the largest disagreement between MERRA and ARM continuous forcing for mixing ratio occur during the spring (MAM) and summer seasons (JJA) in the boundary layer and upper troposphere. The maximum RH for MERRA occurs during June when boundary layer humidity is highest. As will be shown later, cloud fraction in MERRA also peaks in June, suggesting that this may be a byproduct of the convective parameterization used in the AGCM. Like ARM and NARR, additional peaks occur during January and March. It is concluded that the RH values from three different datasets generally agree during this 3-yr period.

Contrary to the RH comparison, significant differences exist for the omega field as shown in Figs 1d-1f. As illustrated in Fig. 1d, there are two periods of upwelling(cool colors) for the ARM dataset: one during the late spring from May-June peaking at ~1.75 mb hr^{-1} and the other in the early fall during September-October with weak upward motion. Downwelling branches occur during the late fall/early winter and the late summer in the lower troposphere. Although NARR and MERRA omega values are similar to each other, they differ considerably from ARM data. NARR is characterized by capturing the seasonal pattern of omega, however, with much different amplitudes than ARM. For upwelling motion, the largest upward motion in NARR occurs during March instead of the late spring (May-June) as shown in Fig. 1d. The upward motion during the early fall is also much weaker. Downwelling motion on the other hand, is notably stronger than ARM with maximum values around ~1 mb hr^{-1}. This is most evident during the summer months when the downwelling branch extends throughout the atmospheric column. MERRA (Fig. 1f) shares many resemblances with NARR especially with regard to the weaker spring upwelling and stronger downwelling during the summer months. Perhaps the most unique feature with MERRA is the upward motion is largest in the lower troposphere near the surface and just above the PBL.

To further investigate the RH and omega differences between the three datasets, the histograms of 3-hourly RH at 925 hPa and omega at 300 hPa for all and non-precipitating periods are presented in Fig. 2. For 925 hPa RH, there is little difference between all (Fig. 2a) and dry (Fig. 2b) conditions. ARM is characterized by having more values > 80% than NARR and MERRA, whereas MERRA has a dry bias with more values <35% than the other two. NARR RH values fall between ARM and MERRA results. For omega, histograms are given with the y-axis in a logarithmic scale. Despite having a large positive bias compared to ARM as shown in Fig. 1e, NARR occasionally produces larger upward motions although the number of events is very small (Fig. 2c). These upward motions, however, disappear under the dry period (Fig. 2d), indicating that these upward motions occur under precipitating periods. It is believed that these large upward velocities result from spurious grid scale precipitation (SGSP) as first documented by West et al. (2007). In brief, the mismatch between assimilated and ETA modeled precipitation used in NARR introduces spurious latent heating which in turn causes unreasonable upward velocities usually during times of convection. Given this only occurred several dozen times during the 3-yr period, this study agrees with the West et al. (2007) finding that *"SGSP will probably have little or no effect on long-term hydrometeorological averages prior to 2003"*. This phenomenon is a non-issue in MERRA which has a much smaller tail for upward velocities. Figures 2c and 2d demonstrate that both NARR and MERRA have more downward motion than ARM at the 300 hPa level, which is consistent with the results in Fig. 1.

Determining which dataset is closer to the atmospheric "truth" is a difficult question to answer, especially without direct measurements of vertical velocity. Therefore it is necessary to find other observed parameters that may be related to vertical velocity to evaluate the three datasets during the 3-yr period. In this study, it is hypothesized that a more accurate large-scale relative humidity and vertical motion field will have a stronger relationship with observed cloud fraction. This has the added benefit of accessing the validity of cloud parameterizations that use these variables to predict cloud fraction.

Correlations were calculated between 3-hr mean RH, omega, and cloud fraction as determined by the ARM MMCR/MPL data at the ARM SGP site during the 3-yr period. For omega, correlations are calculated at an observed CF pressure level against 300 hPa omega. Although not shown, these correlations (Fig. 3b) are higher than those calculated at each level (i.e. 925 hPa CF correlated with 925 hPa omega) because vertical motion is typically small and more turbulent at lower levels. Since these RH and omega correlations are calculated from a point observation (CF derived from ARM radar-lidar) and a forcing domain averaged mean (RH and omega), these correlations may be lower than reality because clouds might occur elsewhere in the forcing domain but were not observed by ARM radar-lidar.

As illustrated in Fig. 3a, the vertical distributions of the CF and RH correlations for the three datasets are nearly identical although values are slightly higher for ARM. Overall, RH has a moderate correlation with CF and is characterized by being bimodal, with peak values of 0.5-0.6 at the top of the boundary layer and the upper troposphere. A larger value at the lowest levels for MERRA is a result of fewer samples at the first level; unlike NARR, MERRA does not calculate variables below ground level (i.e., surface pressure less than the pressure level). Correlations for omega (Fig. 3b) are similar to the findings for RMSE in Fig. 1e where ARM has the smallest RMSE and the largest correlation (-0.45) at a level of 450 hPa. MERRA falls between ARM and NARR with a peak value of ~-0.4 and has a similar vertical distribution to those of ARM and NARR although it is slightly bimodal. In the upper troposphere, however, the rate of change in the MERRA correlation is much smaller, which results in higher correlations than those of ARM and NARR. This is most likely caused by a sampling issue because the vertical resolution of MERRA is less than those from NARR and ARM above 300 hPa (50 vs. 25 hPa).

To understand the seasonal variation of RH/omega relationship with cloud fraction, Fig. 4 is produced. The RH correlations from the three datasets have similar seasonal variations with a relatively large range, and these results are consistent with the previous findings (e.g., Figs. 1 and 3). Correlations are highest from late fall to early spring when clouds are more closely linked to baroclinic wave activity. Correlations then decrease until becoming lowest (<0.2) during the months of July and August, suggesting that cloud parameterizations that are dependant on RH to trigger clouds may need to be improved in the future.

The omega comparison basically follows that for RH except for a few important features. In particular, ARM correlations (Fig. 4d) have maxima during the months of January-February, April, and June. Although NARR and MERRA (Fig. 4e-f) capture the peaks for the winter and early spring months, they do not have a maximum during June. This warrants further investigation. Given that the ARM forcing is constrained by precipitation, this may suggest that during the late spring and early summer, precipitation is more likely caused by local forcing (i.e., isolated thunderstorms developing along weak boundaries with weak synoptic-scale support, Dong et al. 2010) that can not be captured by the reanalyses. Like the RH comparison, ARM correlations are slightly higher (0.1-0.2) than those of NARR and MERRA at any given time and height. In other words, ARM, NARR, and MERRA all agree on the hour-to-hour variation of vertical velocity and its relationship to cloud occurrence.

4. Precipitation, Cloud Fraction, and Surface Radiation

In this section, the precipitation, cloud fraction, and surface radiation derived from both NARR and MERRA are evaluated with observations at the DOE ARM SGP site during the period 1999-2001. As shown in Fig. 5, ARM and NARR precipitation have excellent agreement with each other, capturing the monthly variability in precipitation during this time period which should be expected given the design of NARR to assimilate observed precipitation. This is certainly not a new finding because it has been documented in Becker et al. (2009) and Bukovsky and Karoly (2006). The largest precipitation amounts occur during the month of June, followed by the earlier spring, and fall months. For many months, the two lines are nearly indistinguishable. MERRA on the other hand, appears to have a negative bias for most of the 3-yr period. Despite this bias, however, it does capture the monthly variability of precipitation. Figure 6 shows the scatterplots of the monthly and daily total precipitation for the three datasets. As demonstrated in Fig. 5 and Fig. 6a, NARR monthly total precipitation has excellent agreement with ARM forcing with a correlation of 0.99 and bias of -2.8 mm. MERRA monthly total precipitation (Fig. 6b), however, has a larger bias of -22.2 mm. Despite this bias, there is still a linear trend with a relatively high correlation of 0.86. Precipitation is also over simulated on occasion during low precipitation months (<50 mm), hence the intercept of 15.66 mm.

Reducing precipitation to daily totals leads towards more disagreement between ARM and reanalyses as noted by the smaller values of slope and correlation. For NARR (Fig. 6c), slope is reduced from 0.96 to 0.86 and correlation from 0.99 to 0.91. Overall, there is a ~ -0.1 mm bias per day. This panel is similar to the "Great Plains" panel in Fig. 2 from Becker et al. (2009). The more significant scattering and values at 0 for one dataset suggest that the assimilation process might introduce some uncertainty into the original observations either in time and/or location. Becker et al. (2009) found that in general, NARR has less intensity and higher frequency precipitation than the observations, so some care should be taken in analysis of individual cases. Daily precipitation correlation for MERRA (Fig. 6d) is reduced to 0.69 with a bias of -0.73 mm.

Figure 7 shows the CF comparison between ARM radar-lidar, GOES, NARR and MERRA at the ARM SGP site during the period 1999-2001. The monthly CF difference between ARM radar-lidar and GOES observations may be due to the spatial scale difference (point vs. a 2x2.5^o grid box) and remote sensing method (active vs. passive). The annual mean CF difference between ARM radar-lidar and GOES observations is within 1% (43% vs. 44%) for the entire 3-yr period. This result is consistent with the findings in the Xi et al. (2010) and Kennedy et al. (2010) studies. Cloud fraction is characterized by having maximum values during the late winter and spring (peaking in March), and then having another local maximum during June when precipitation and upward motion peaks. CF then decreases to a minimum during the summer when Oklahoma is typically under large-scale ridging. Both NARR and MERRA reanalyses capture the same seasonal variations as the ARM radar-lidar and GOES observations, but with negative biases. Of the two, however, MERRA has better agreement with a larger maximum during June and is overall, within 3-4% of observations. Correlations and RMSEs between the reanalyses and observations are also calculated based on a total of 36 monthly means and are summarized in Table 2. Although NARR has a larger RMSE against both ARM and GOES observations than MERRA, its correlations are higher, indicating that NARR captures month-to-month variability better. Note that the CF correlation between ARM and GOES is 0.91 and the RMSE is 5.8%. While the CF correlation is highest for NARR against ARM, the correlation between GOES and MERRA is nearly the same as that between GOES and NARR, and the RMSE values for MERRA are much smaller than those of NARR. This may be a matter of MERRA incorporating GOES data into its assimilation process.

Comparisons of monthly mean surface fluxes for clear-sky and all-sky conditions from the three datasets are shown in Fig. 8 and summarized in Table 3. For detailed discussion, the reader is referred to the Dong et al. (2006) study which investigated the seasonal variations of CF and surface radiative fluxes at the ARM SGP during the period 1997-2002. Despite the slightly longer time period in the Dong et al. (2006) study, the differences between this study (ARM results) and Dong et al. (2006) are within a few W m^{-2} as listed in Table 3.

Overall, the reanalyses capture the seasonal variability seen in ARM quite well, albeit with biases (Table 3). These biases are smallest for periods of clear-sky which is expected; surface fluxes in reanalyses are dependant on not only their parameterizations for surface radiation, but also clouds. Compared to the all-sky ARM results, the NARR SW-down is significantly higher (47 Wm^{-2}), and LW-down is lower (-9 Wm^{-2}), which is consistent with the negative bias of cloud fraction found in Fig. 7. Markovic et al. (2009) found similar results for NARR analyzed at six surface sites within the US and suggested that high biases in mean annual all-sky SW-down (~40 W m^{-2}) were attributed to a negative bias of CF. The clear-sky comparisons are nearly the same as their all-sky counterparts, i.e., SW-down is 25 W m^{-2} higher and LW-down is 13 W m^{-2} lower, suggesting that the impacts of water vapor and aerosols on radiative transfer in NARR also need to be improved. Given that NARR is based on the NCEP ETA model, this is consistent with Hinkelman et al. (1999) which found that ETA had an average excess of 50 W m^{-2} for SW-down with approximately half of this bias attributed to deficient extinction.

The comparisons between MERRA and ARM agree much better than those between NARR and ARM as shown in Fig. 8 and listed in Table 3. However, there are a few exceptions. MERRA has larger biases than NARR for LW-down under both clear and all sky conditions (-20 and -19 w m^{-2}). Compared to ARM and NARR, these negative biases are consistent with the drier conditions in MERRA as demonstrated in Fig. 1 because atmospheric water vapor is extremely important for LW-down fluxes (Dong et al. 2006) and is supported by the fact these biases are largest during the warm season and are nearly the same under both clear-sky and all-sky conditions.

Finally, comparisons of monthly mean TOA fluxes for clear-sky and all-sky conditions are given in Fig. 9 and are summarized in Table 4. Reanalysis fluxes under clear-sky condition have small positive biases within 5 W m^{-2} of ARM (GOES) observations. As expected, TOA SW-up fluxes for all-sky condition are highest during months with high cloud fraction, and the differences between reanalyses and ARM are related to their CF differences. For example, NARR TOA flux biases (negative for SW-up and positive for LW-up) are consistent with the year-round negative CF bias. MERRA biases vary by season depending on the amount of cloud cover produced. The peak in SW-up and minimum in LW-up during June are strongly associated the peak of CF during that month. Despite this disagreement, biases in MERRA are noticeable smaller than those of NARR as listed in Table 4.

5. Summary and Conclusions

The atmospheric state, precipitation, total cloud fraction, and surface radiative fluxes from MERRA and NARR reanalyses were collected and compared with the ARM SGP continuous forcing dataset during the period 1999-2001. Key findings are summarized below.

1. For atmospheric state, NARR and MERRA reanalyses have small column averaged biases within 0.5 m s^{-1} and 0.13 K for horizontal wind components and air temperature, respectively. Specific humidity and RH values from ARM and NARR are in excellent agreement and both have a bimodal distribution with peaks in the boundary layer and the upper troposphere. MERRA captures the general shape of RH, but with a ~5% negative bias throughout the year in the upper troposphere except during the late spring and early summer when convection is most common at the ARM SGP site.

2. Significant differences exist for the omega field. The largest differences occur for upwelling during the spring months and the magnitude of downwelling during the summer. Although NARR and MERRA share many resemblances to each other, ARM outperforms these reanalyses in terms of correlation with CF. Given that the ARM forcing is constrained by precipitation to give the adequate mass, momentum, heat, and moisture budgets, this indicates that some of the precipitation (especially during the late spring and early summer) is caused by smaller-scale forcing that is not captured by the reanalyses. This also suggests that SCMs based on the forcing derived from reanalyses would not be able to model precipitation adequately during this time period. Combined with known issues such as SGSP in NARR documented by West et al. (2007) and within this study, vertical velocity values in reanalyses should be used with caution.

3. ARM and NARR have excellent agreement for monthly precipitation amounts which are a testament to the improved precipitation assimilation into NARR. NARR has a slight (~3 mm) bias for monthly precipitation but with more variability for daily precipitation, suggesting that the assimilation of precipitation may sometimes be mistimed or misplaced. Despite this, both

monthly and daily correlations are still high. MERRA, on the other hand, only captures the monthly variation of precipitation well and contains considerable negative biases at monthly (-22.2 mm) and daily (-0.7 mm) intervals.

4. As found in Kennedy et al. (2010) and Xi et al. (2010), total CF at the ARM SGP site has good agreement between ARM and GOES satellite observations. From 1999-2001, CF peaked during the months of March and June before reaching a minimum during the summer months. Both NARR and MERRA capture this change as evidenced by high correlations (0.92-0.78), although they have negative biases (14% and 3%, respectively). MERRA correlations for CF are highest with satellite observations while NARR correlations are highest with the ARM surface observations. This is not surprising given the amount of satellite information being assimilated into MERRA.

5. Surface radiative fluxes within this study agree well with those from Dong et al. (2006). Of the two reanalyses, MERRA shows better agreement with ARM observations for all fluxes except for LW-down. NARR has significant positive biases for SW-down, SW-up, and LW-up, and these are attributed due to a combination of too few clouds and a lack of sufficient extinction by aerosols and water vapor in the atmospheric column. These results are consistent with previous studies that have investigated NARR elsewhere in the US and ETA at the ARM SGP site. MERRA biases for LW-down are attributed to the negative bias of water vapor within the atmospheric column.

The results presented here represent only one location within the well constrained continental mid-latitudes with a limited time period. However, in a companion study over the

Arctic region (Zib et al. 2010), similar results were found albeit with smaller biases. This study and Zib et al. (2010) have indicated that MERRA generally agrees better than NARR/NCEP reanalyses with ARM in both the middle latitudes and Arctic regions for CF and radiative fluxes. A potential avenue of research is expanding this analysis for a longer period using the newly developed Climate Modeling Best Estimate (CMBE) dataset by ARM (Xie et al. 2010). It is also currently planned to expand the ARM continuous forcing from 2001 to present time over the ARM SGP site, as well as other surface sites.

5. Acknowledgements

NARR reanalysis data were provided by NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their website at http://www.esrl.noaa.gov/psd. MERRA was obtained from the Goddard Earth Sciences Data and Information Services Center, Greenbelt, MD, USA, from their website at http://disc.sci.gsfc.nasa.gov/mdisc. The University of North Dakota authors were supported by NASA NEWS project under Grant NNX07AW05G, DOE ARM program under grant DE-AC52-07NA27344/B589973, and the NASA CERES project under Grant NNL04AA11G.

References

Becker, E.J., E.H. Berbery, and R.W. Higgins, 2009: Understanding the Characteristics of Daily Precipitation over the United States Using the North American Regional Reanalysis. *J. Climate*, **22**, 6268–6286

Benjamin, S. G., et al., 2004: An hourly assimilation/forecast cycle: The RUC. *Mon.Weather Rev.*, *132*, 495– 518.

Bloom S. C., L. L. Takacs, A. M. Da Silva, and D. Ledvina, 1996: Data assimilation using incremental analysis updates. *Mon. Wea. Rev*., **124**, 1256-1271

Bosilovich, M.G., J. Chen, F.R. Robertson, and R.F. Adler, 2008: Evaluation of Global Precipitation in Reanalyses. *J. Appl. Meteor. Climatol.*, **47**, 2279–2299.

Bukovsky, M.S., and D.J. Karoly, 2007: A Brief Evaluation of Precipitation from the North American Regional Reanalysis. *J. Hydrometeor.*, **8**, 837–846.

Clothiaux, E.E., T.P. Ackerman, G.G. Mace, K.P. Moran, R.T. Marchand, M.A. Miller, and B.E. Martner (2000), Objective determination of cloud heights and radar reflectivities using a combination of active remote sensors at the Atmospheric Radiation Measurement Program Cloud and Radiation Test Bed (ARM CART) sites, *J. Appl. Meteor*., **39**, 645-665.

Dong, X., B. Xi, and P. Minnis, 2006: A climatology of midlatitude continental clouds from the ARM SGP Central Facility. Part II: Cloud fraction and surface radiative forcing. *J. Climate*, **19**, 1765-1783.

Dong, X., B. Xi, A. Kennedy, Z. Feng, J. Entin, P. Houser, B. Schiffer, W. Olson, T. L'Ecuyer,

T. Liu, K-L Hsu, B. Lin, Y. Deng, andT. Jiang, 2010: Investigation the 2006 Drought and 2007 FloodExtreme Events at the SGP through an Integrative Analysis of Observations. Submitted to *J. Geophys. Res.*

Hinkelman, L. M., T. P. Ackerman, and R. T. Marchand, 1999: An evaluation of NCEP Eta model predictions of surface energy budget and cloud properties by comparison to measured ARM data. *J. Geophys. Res*., **104**, 19 535–19 549.

Kennedy, A, **X. Dong**, B. Xi, P. Minnis, A. Del Genio, A. Wolf and M. Khaiver, 2010. Evaluation of the NASA GISS Single Column Model Simulated Clouds Using Combined Surface and Satellite Observations. *J. Clim.,* doi: 10.1175/2010JCLI3353.1

Mace, G.G. and coauthors (2006), Cloud radiative forcing at the Atmospheric Radiation Measurement Program Climate Research Facility: 1. Technique, validation, and comparison To satellite-derived diagnostic quantities. *J. Geophys. Res*., **111**, D11S90, doi:10.1029/2005JD005921.

Markovic M, Jones CG, Winger K,Paquin D (2009) The surface radiation budget over North America: gridded data assessment and evaluation of regional climate models. International Journal of Climatology 29(15): 2226

Mesinger, F., and coauthors, 2006: North American Regional Reanalysis. *Bull. Amer. Meteor. Soc.*, **87**, 343–360.

Min, Q, P. Minnis, and M. M. Khaiyer, 2004: Comparison of cirrus optical depths from GOES-8 and surface measurements. *J. Geophys. Res.*, **109**, D20119, 10.1029/2003JD004390.

Minnis, P., and coauthors, 2001: A near-real time method for deriving cloud and radiation Properties from satellites for weather and climate studies. *Proc. AMS 11th Conf. Satellite Metr. and Oceano.*, Madison, WI, Oct. 15-18, 477-480.

——, and co-authors, 2008: Cloud detection in non-polar regions for CERES using TRMM VIRS and Terra and Aqua MODIS data. *IEEE Trans. Geosci. Remote Sens.*, **46**, 3857-3884.

Moran, K.P., B.E. Martner, M.J. Post, R.A. Kropfli, D.C. Welsh, and K.B. Widener, 1998: An Unattended Cloud-Profiling radar for use in climate research. *Bull. Amer. Meteor. Soc.*, **79**, 443–455.

Thorn, P.W., and R. S. Vose, 2010: Reanalyses suitable for characterizing long-term trends … are they really achievable?. *Bull. Amer. Meteor. Soc.*, **91**, 353-361.

Xi, B.,X. Dong, P. Minnis, and M. Khaiyer, 2010: A 10-year climatology of cloud cover and vertical distribution derived from both surface and GOES observations over the DOE ARM SGP Site. *J. Geophys. Res.,* 115, D12124, doi:10.1029/2009JD012800.

Xie, S., R. T. Cederwall, M. Zhang, and J. J. Yio (2003), Comparison of SCM and CSRM forcing data derived from the ECMWF model and from objective analysis at the ARM SGP site, *J. Geophys. Res.*, 108(D16), 4499, doi:10.1029/2003JD003541.

_____, R. T. Cederwall, and M. 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, *J. Geophys. Res.*, **109**, D01104, doi:10.1029/2003JD004045

_____, R.B. McCoy, S.A. Klein, R.T. Cederwall, W.J. Wiscombe, E.E. Clothiaux, K.L. Gaustad, J.C. Golaz, S.D. Hall, M.P. Jensen, K.L. Johnson, Y. Lin, C.N. Long, J.H. Mather, R.A. McCord, S.A. McFarlane, G. Palanisamy, Y. Shi, and D.D. Turner, 2010: CLOUDS AND MORE: ARM Climate Modeling Best Estimate Data. *Bull. Amer. Meteor. Soc.*, **91**, 13–20.

Yang, F., H.L. Pan, S.K. Krueger, S. Moorthi, and S.J. Lord, 2006: Evaluation of the NCEP

Global Forecast System at the ARM SGP Site. *Mon. Wea. Rev.*, **134**, 3668–3690.

Zhang, M.H., and J.L. 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**, 1503–1524.

_____,J.L. Lin, R.T. Cederwall, J.J. Yio, and S.C. Xie, 2001: Objective analysis of ARM IOP data: Method and sensitivity. *Mon. Wea. Rev.*, **129**, 295–311.

Zib, B., X. Dong, and B. Xi, 2010: Comparison of Two Extreme Minimum Arctic Sea-Ice Extents: The Record High During 1996 and Record Low During 2007. Submitted to JAMC.

Figure Captions

Figure 1. Monthly means of RH over the ARM SGP domain from 1999-2001 for (a) ARM continuous forcing, (b) NARR, and (c) MERRA. (d)-(f) are the same as (a)-(c) except for the omega field.

Figure 2. Histograms of 925 hPa RH for (a) all and (b) dry hours. (c) and (d) are the same as (a) and (b) except for 300 hPa omega Note that the y-axis for omega is logarithmic.

Figure 3. Vertical correlations of cloud fraction with (a) RH and (b) omega at a 3-hr temporal resolution.

Figure 4. Seasonal correlations of cloud fraction with RH for (a) ARM, (b) NARR, and (c) MERRA. (d)-(f) are the same as (a)-(c) except for the omega field.

Figure 5. Monthly total precipitation measured over the ARM SGP domain by ARM (black), NARR (red) and MERRA (blue) during the period 1999-2001.

Figure 6. Scatterplots of monthly total precipitation for (a) ARM vs. NARR and (b) ARM vs. MERRA. (c) and (d) are the same as (a) and (b) except for daily total precipitation.

Figure 7. Monthly mean cloud fraction for ARM (black), GOES (green), NARR (red), and MERRA (blue) during the period 1999-2001.

Figure 8. Monthly mean clear-sky (a) SW-down, (b) LW-down, (c) SW-up, and (d) LW up fluxes measured by PSPs and PIRs at the ARM SGP site. (e)-(h) are the same as (a)-(d) except for all sky conditions.

Figure 9. Monthly mean TOA clear-sky (a) SW-up and (b) LW-up fluxes measured by GOES satellite over the ARM SGP site. (c)-(d) are the same as (a)-(b) except for all sky conditions.

Table Captions

Table 1. Yearly and seasonal column averaged biases of zonal wind (m s^{-1}), meridional wind (m s^{-1}), specific humidity (g kg^{-1}), omega (mb hr^{-1}), and air temperature (K) for NARR and MERRA against ARM continuous forcing

Table 2. Correlation and RMSE of total cloud fraction from a total of 36 monthly means.

Table 3. Annual mean surface radiative fluxes and their biases compared to ARM continuous forcing.

Table 4. Annual mean TOA radiative fluxes and their biases compared to ARM continuous forcing.

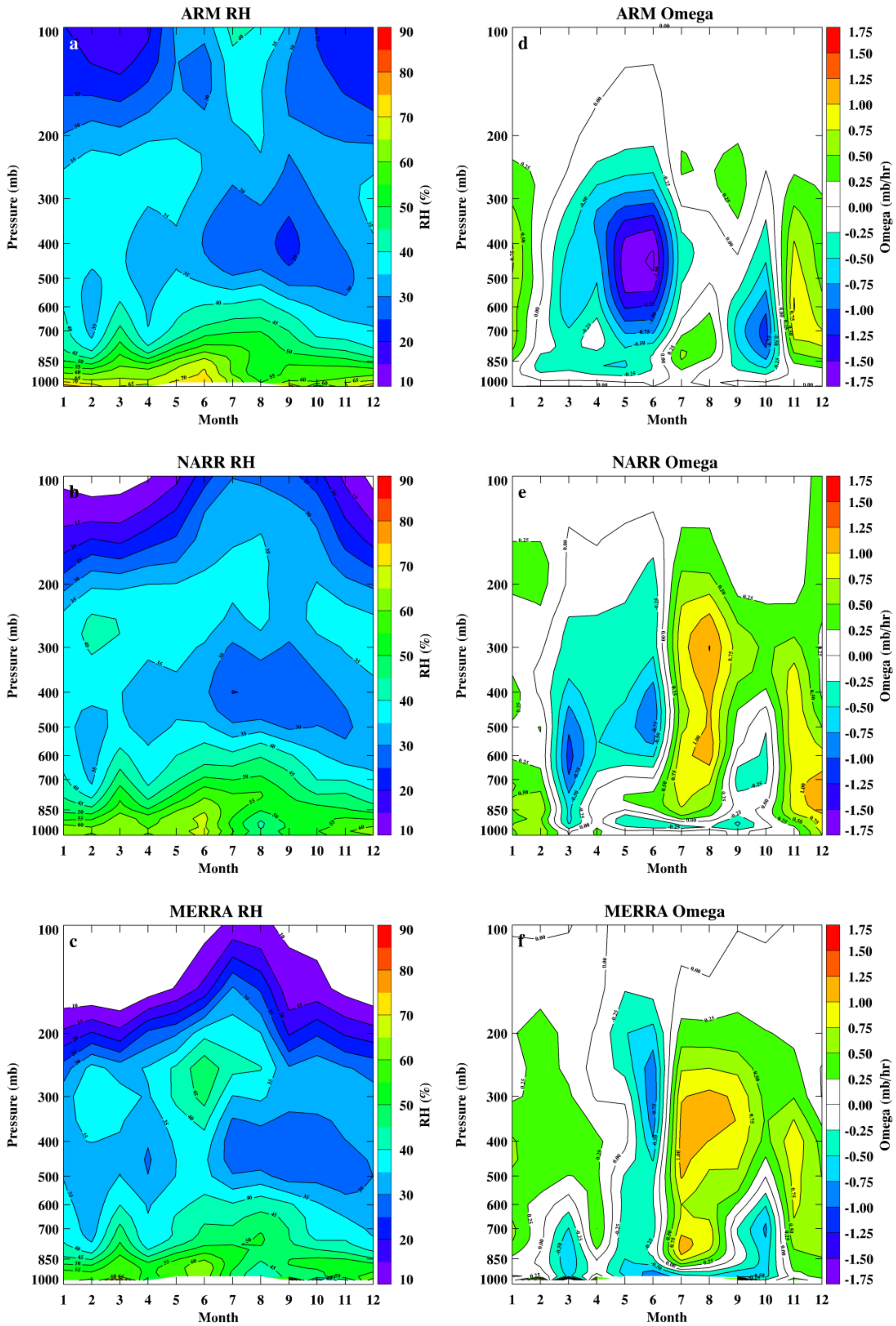

Figure 1. Monthly means of RH over the ARM SGP domain from 1999-2001 for (a) ARM continuous forcing, (b) NARR, and (c) MERRA. (d)-(f) are the same as (a)-(c) except for the omega field.

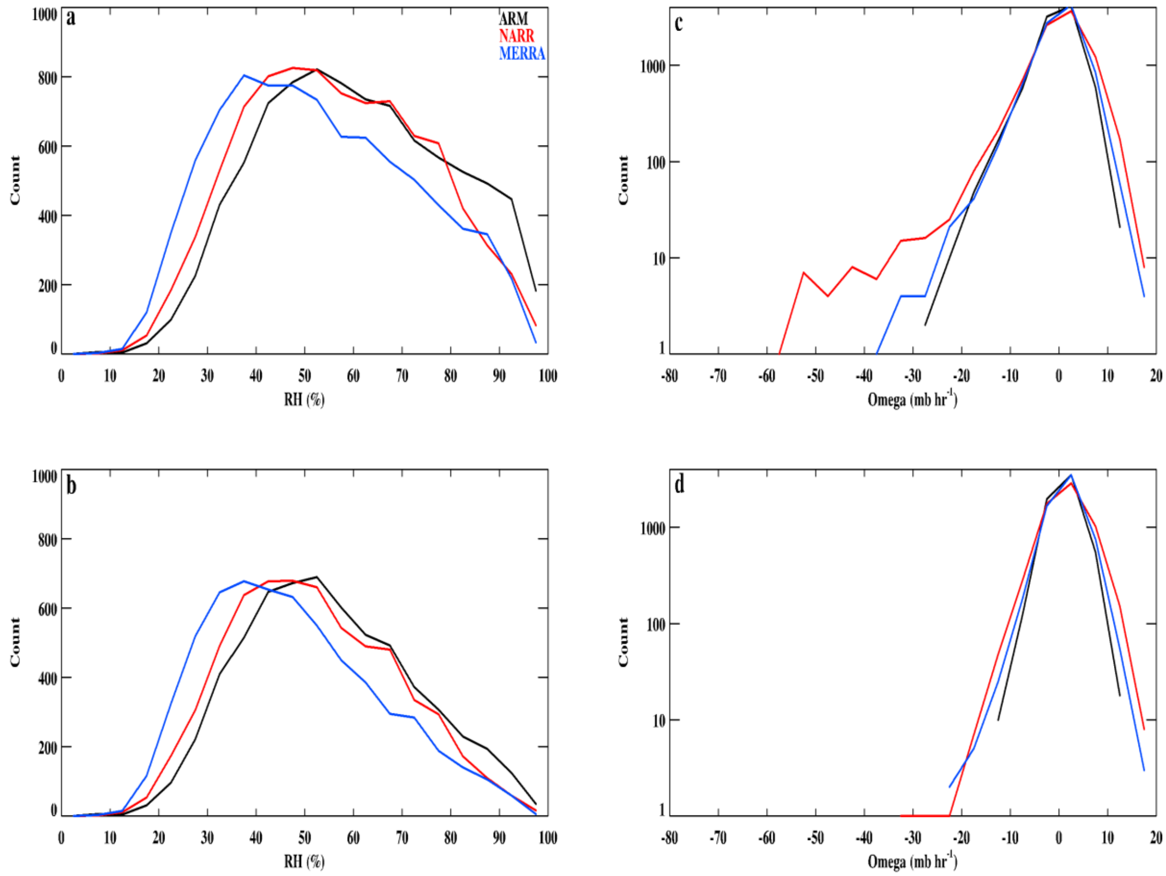

Figure 2. Histograms of 925 hPa RH for (a) all and (b) dry hours. (c) and (d) are the same as (a) and (b) except for 300 hPa omega Note that the y-axis for omega is logarithmic.

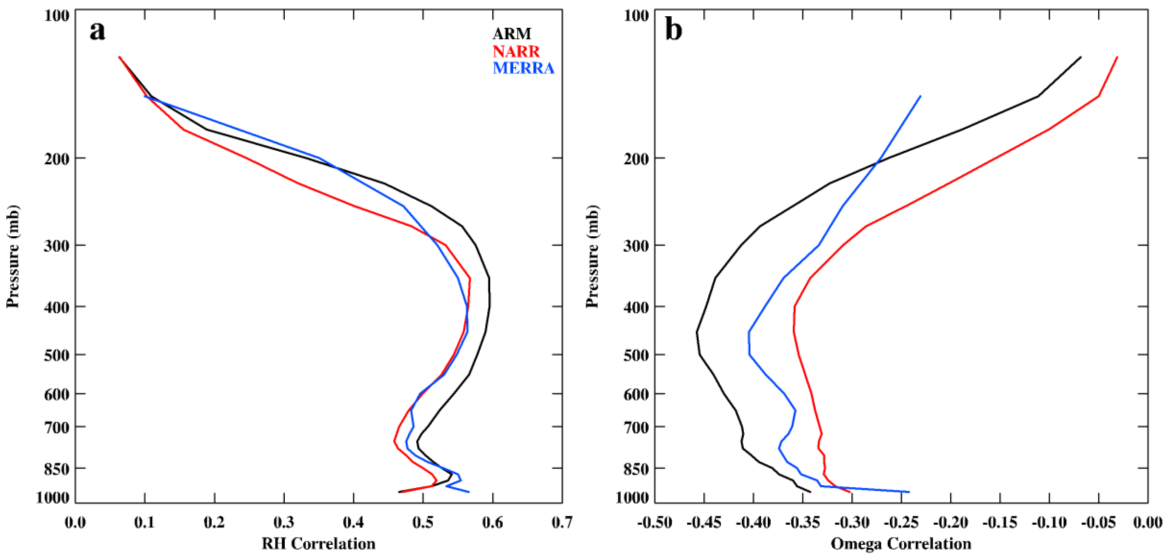

Figure 3. Vertical correlations of cloud fraction with (a) RH and (b) omega at a 3-hr temporal resolution.

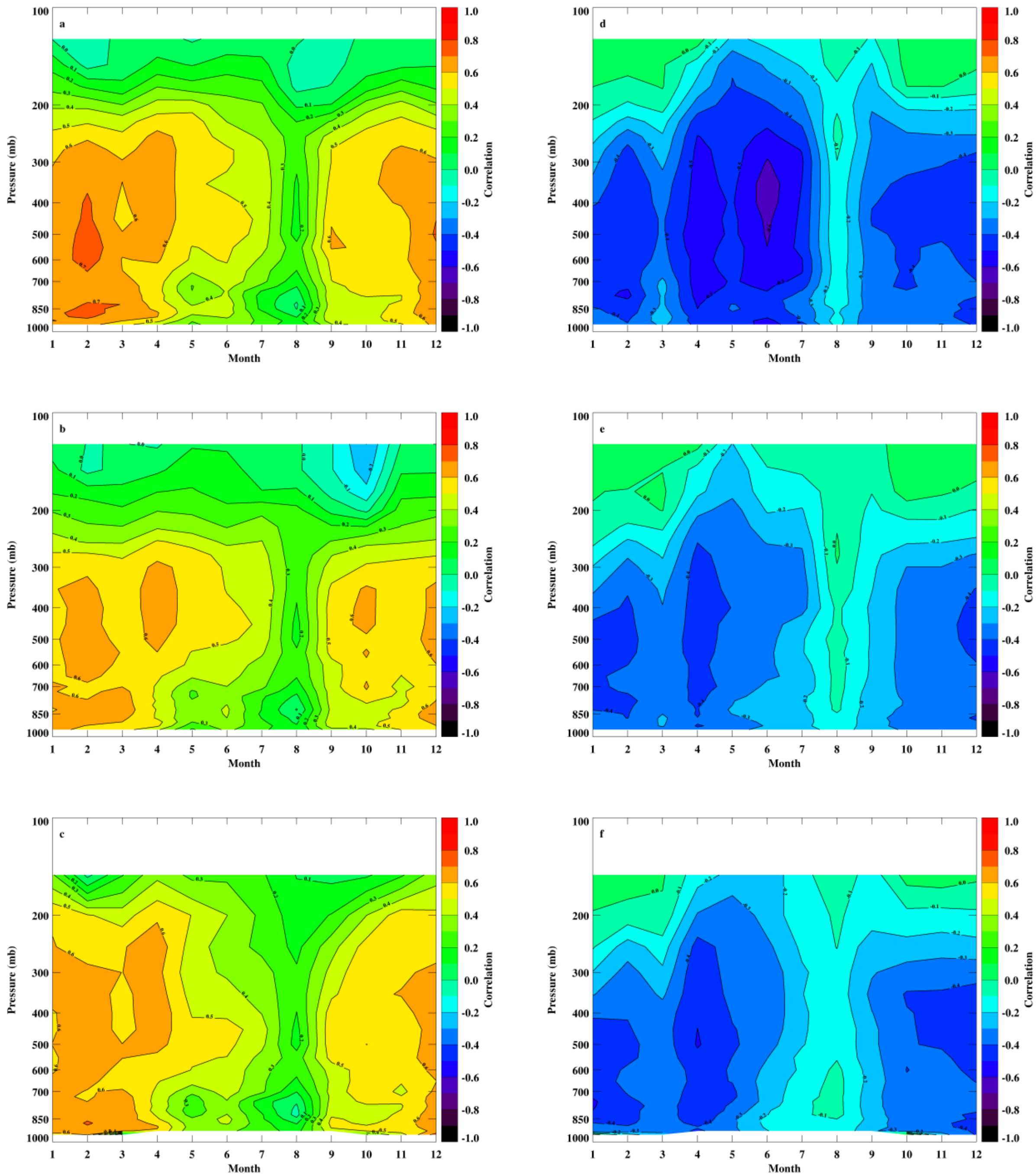

Figure 4. Seasonal correlations of cloud fraction with RH for (a) ARM, (b) NARR, and (c) MERRA. (d)-(f) are the same as (a)-(c) except for the omega field.

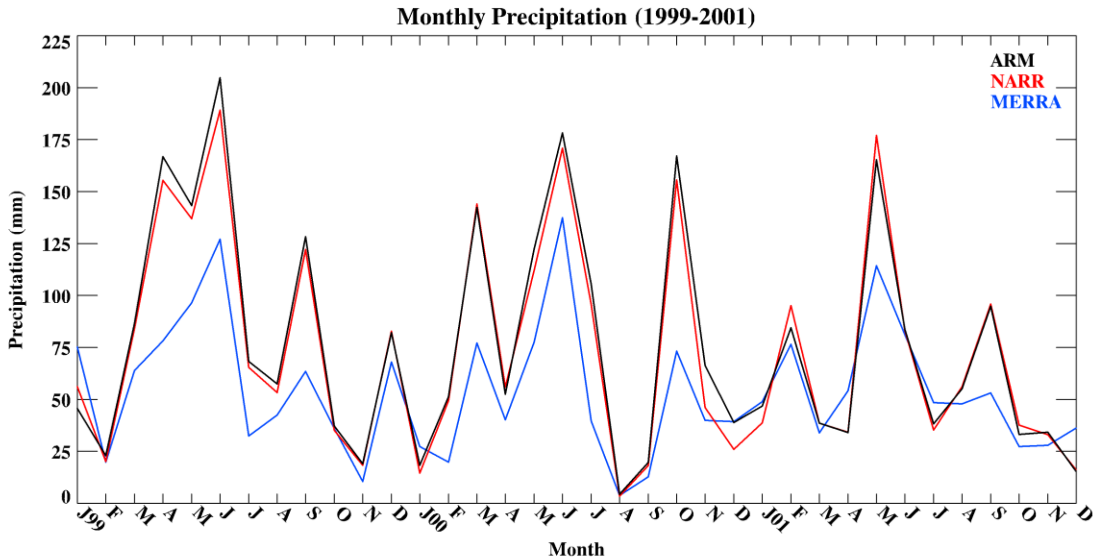

Figure 5. Monthly total precipitation measured over the ARM SGP domain by ARM (black), NARR (red) and MERRA (blue) during the period 1999-2001.

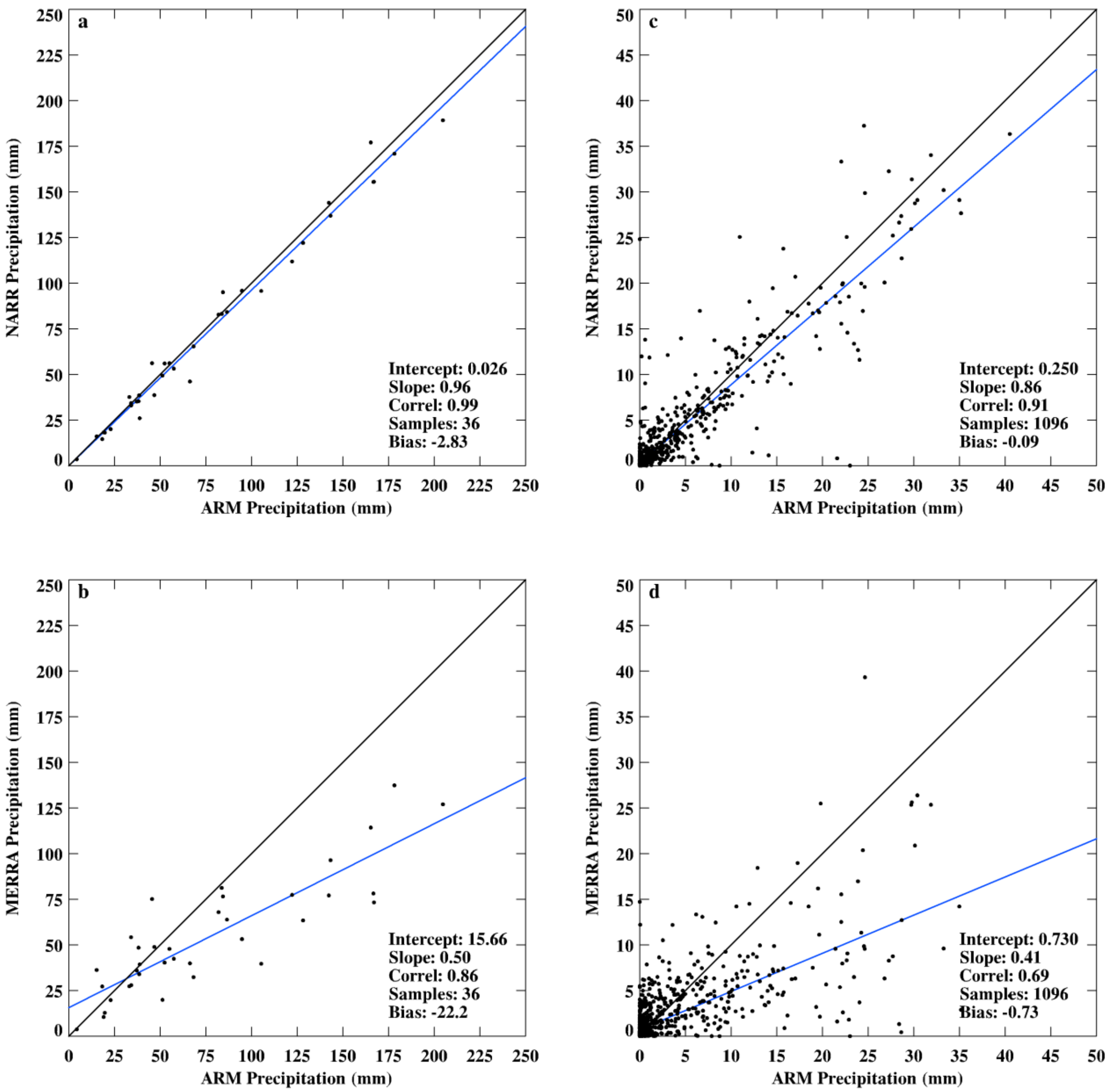

Figure 6. Scatterplots of monthly total precipitation for (a) ARM vs. NARR and (b) ARM vs. MERRA. (c) and (d) are the same as (a) and (b) except for daily total precipitation.

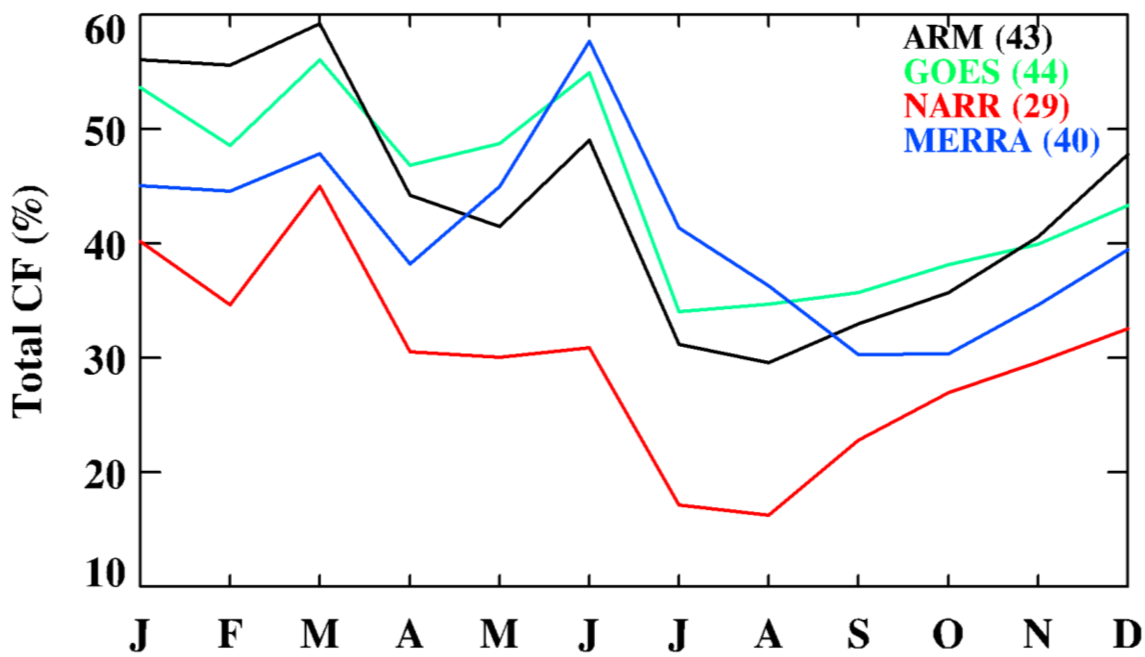

Figure 7. Monthly mean cloud fraction for ARM (black), GOES (green), NARR (red), and MERRA (blue) during the period 1999-2001.

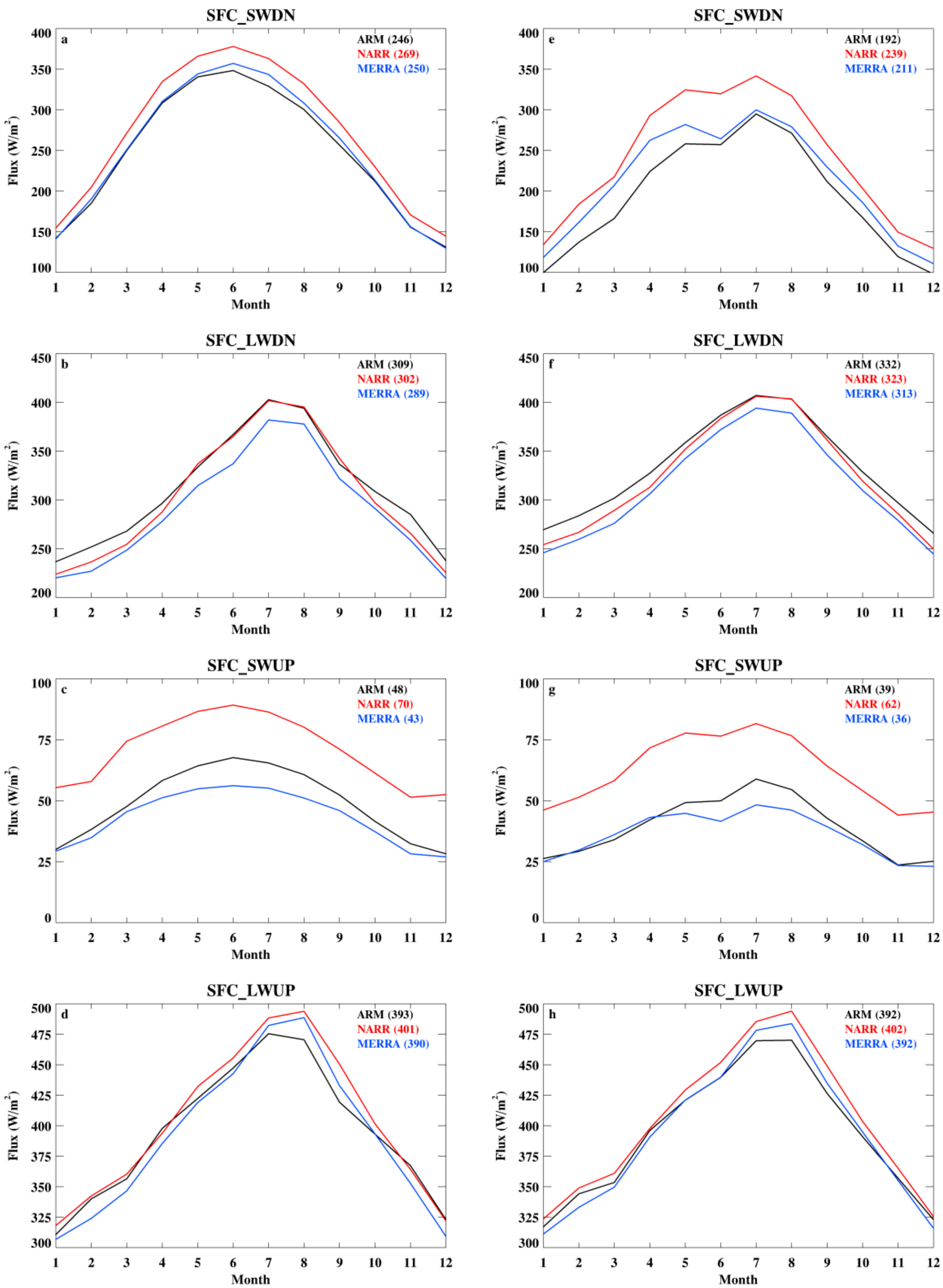

Figure 8. Monthly mean clear-sky (a) SW-down, (b) LW-down, (c) SW-up, and (d) LW up fluxes measured by PSPs and PIRs at the ARM SGP site. (e)-(h) are the same as (a)-(d) except for all sky conditions.

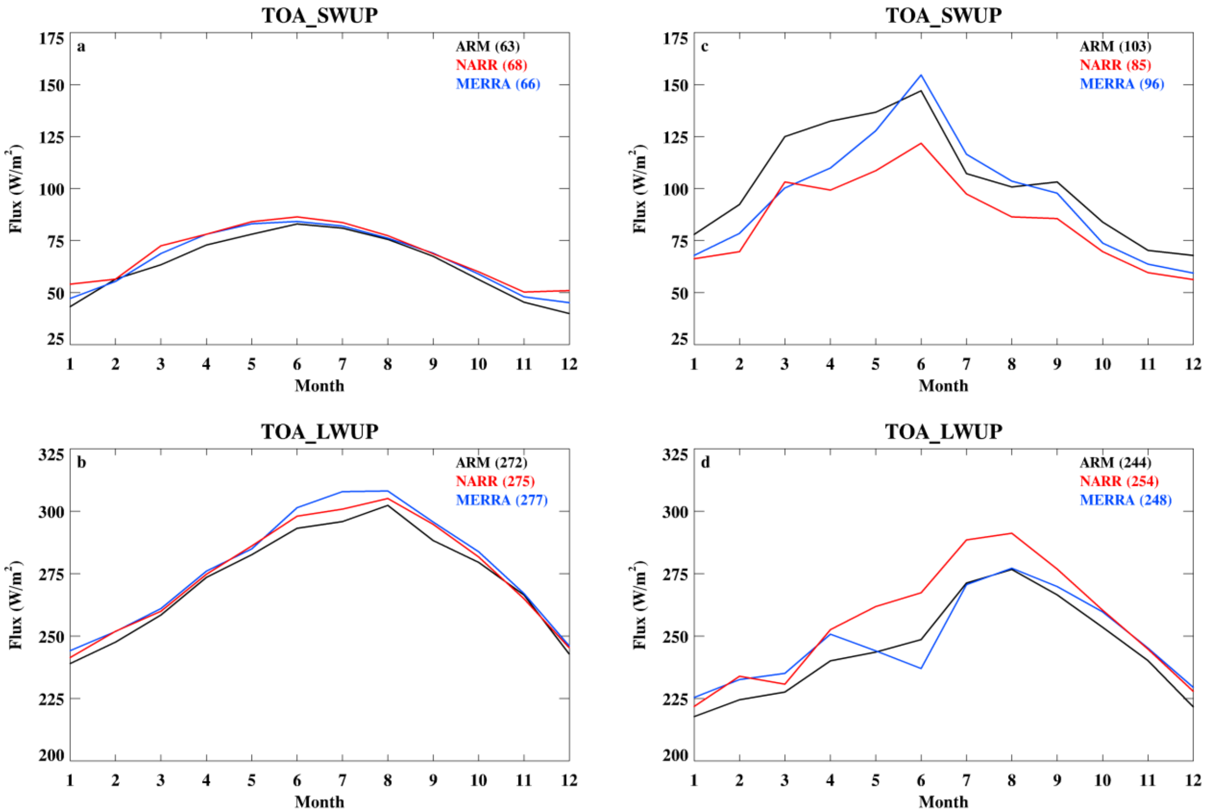

Figure 9. Monthly mean TOA clear-sky (a) SW-up and (b) LW-up fluxes measured by GOES satellite over the ARM SGP site. (c)-(d) are the same as (a)-(b) except for all sky conditions.

Table 1. Yearly and seasonal column averaged biases of zonal wind (m s^{-1}), meridional wind (m s^{-1}), specific humidity (g kg^{-1}), omega (mb hr^{-1}), and air temperature (K) for NARR and MERRA against ARM continuous forcing

NARR	YEAR	DJF	MAM	JJA	SON
U (m/s)	0.42	0.38	0.46	0.4	0.41
V (m/s)	0.04	0.13	-0.22	-0.2	0.29
Q (g/kg)	-0.01	0	0.01	-0.04	0.01
O (mb/hr)	0.34	0.26	0.22	0.54	0.33
T (K)	-0.06	0.03	-0.09	-0.1	-0.05
MERRA	**YEAR**	**DJF**	**MAM**	**JJA**	**SON**
U (m/s)	0.18	0.08	0.12	0.3	0.15
V (m/s)	0.03	-0.17	-0.3	0.25	0.36
Q (g/kg)	-0.19	-0.8	-0.16	-0.36	-0.17
O (mb/hr)	0.22	0.07	0.32	0.25	0.25
T (K)	-0.02	-0.13	-0.01	0.13	-0.07

Table 2. Correlation and RMSE of total cloud fraction from a total of 36 monthly means.

ρ	NARR	MERRA
ARM	0.92	0.78
SAT	0.9	0.86

RMSE	NARR	MERRA
ARM	14.9	9
SAT	15.6	7.1

Table 3. Annual mean surface radiative fluxes and their biases compared to ARM continuous forcing.

	Clear Sky				All Sky			
	SWDN	SWUP	LWDN	LWUP	SWDN	SWUP	LWDN	LWUP
ARM	246	48	309	393	192	39	332	392
D06	248		314		195		333	
NARR	269	70	302	401	239	62	323	402
MERRA	250	43	289	390	211	36	313	392

	Clear Sky				All Sky			
	SWDN	SWUP	LWDN	LWUP	SWDN	SWUP	LWDN	LWUP
NARR	23	22	-7	8	47	23	-9	10
MERRA	4	-5	-20	-3	19	-3	-19	0

Table 4. Annual mean TOA radiative fluxes and their biases compared to ARM continuous forcing.

	Clear Sky		All Sky	
	SWUP	LWUP	SWUP	LWUP
ARM	63	272	103	244
NARR	68	275	85	254
MERRA	66	277	96	248

	Clear Sky		All Sky	
	SWUP	LWUP	SWUP	LWUP
NARR	5	3	-18	10
MERRA	3	5	-7	4