Analysis of Cloud Variability and Sampling Errors in Surface and Satellite Measurements

Z. Li, M. C. Cribb, and F.-L. Chang
Earth System Science Interdisciplinary Center
University of Maryland
College Park, Maryland

A. P. Trishchenko and Y. Luo
Canada Centre for Remote Sensing
Ottawa, Ontario, Canada

Introduction

Radiation measurements have been widely employed for evaluating cloud parameterization schemes and model simulation results. As the most comprehensive program aiming to improve cloud parameterization schemes, the Atmospheric Radiation Measurement (ARM) Program has an essential goal to make observations on the scale of a general circulation model gridbox, so as to define the physics underlying some of the important parameterizations in the general circulation models used in climate change studies. While ARM has deployed a network of radiation stations over a domain of about 400 km^2, extensive radiation and cloud measurements are taken at a single location (the central facility) within a domain of 100 km^2. An important question is thus raised as to whether these measurements are adequate to represent grid mean values given the high variability of cloud and the surface. In contrast to the coarse-resolution general circulation models, fine-resolution cloud system models (CSMs) are playing an increasing role in revealing the fundamentals of cloud-radiation interactions. As computing power improves, CSMs will have a very high resolution (a few to a few tenths of km), which raises another question: on what scale does modeling need to occur in order to capture the physical properties that drive the system. Answers to both questions hinge on cloud variability and observation density. It is not clear that we can understand cloud behavior well enough from the ARM-deployed instruments nor are we sure which types of measurements and at what spatial density are required for closure of the boundary fluxes in single-column models. To address these questions, we need more sound guidance toward adding additional observation sites and deploying mobile facilities. While it is a formidable task to resolve these questions, we attempt here to analyze related issues by using satellite and ground observations of radiative quantities. Taking advantage of the high spatial and temporal resolution of geostationary operational environmental satellite (GOES) data, we mimic ground-based measurements of varying density and temporal frequency and characterize their observation uncertainties caused by cloud variability at different scales in different seasons. Such scale-dependent statistics of observation uncertainties provide critical constraints on model-observation comparisons, and are thus valuable for improving and validating cloud parameterization schemes.

Models and Input Data

The model of Li et al. (1993) was used to calculate the surface solar radiation budget (SSRB). Correction factors accounting for the effects of cloud-top altitude and aerosol loading on the net surface radiation, developed by Masuda et al. (1995), were applied to the SSRB. The top of the atmosphere (TOA) albedo data are derived from high-resolution (approximately 4-km pixel-level data) GOES data developed by the National Aeronautics and Space Administration's Langley Cloud and Radiation Research Group. It encompasses a domain of approximately 400 km^2 centered on the Southern Great Plains (SGP) central facility (CF) site in north-central Oklahoma and produces output for about every half-hour on a daily basis. Datasets covering the periods of March, May, July, August, September, and October of the year 2000 were used in this study. In addition to TOA broadband radiative fluxes, the data set provides cloud macrophysical properties such as cloud-top height and pixel cloudiness and cloud microphysical properties such as effective radius.

Another input parameter to the models is the precipitable water content and this quantity was interpolated from the $2.5^o x 2.5^o$ global National Centers for Environmental Prediction reanalysis (Kalnay et al. 1986). This data set covers the period from 1948 to the present and each file contains data for one year with values of precipitable water at four times during each day (0, 6, 12, 18 UTC) in the year. For the simulations performed here, the subset of data encompassing the SGP region for the year 2000 was extracted first. During subsequent calculations when this input is required at a given latitude and longitude and time point, the precipitable water amounts of the four closest points in space before and after the time point were isolated. Bilinear interpolation was performed on each set of four points corresponding to the times before and after the time point required then a linear interpolation in time was done.

Aerosol optical depths at 550 nm were derived from Aerosol Robotic Network measurements taken at the CF (Holben et al. 2001). Observations of aerosol optical depth were taken by the sunphotometer at wavelengths of 340, 380, 440, 500, 670, 870, and 1020 nm and the final data undergo cloud screening and data quality assurance procedures. Data from 1994 until September 2001 was used to derive the Ångstrom exponent, a, from which the aerosol optical depths at 550 nm were then calculated. These optical depths generated from data at the CF were assumed to be representative of the whole SGP domain, given that aerosol loading is not expected to vary greatly over this generally rural area. Exceptions can occur during the spring season when localized increases in aerosol load can occur within the SGP domain due to the agricultural practice of burning fields in preparation for reseeding.

Methodology

Areal means of SSRB were calculated over domains of various spatial size and temporal intervals. The model of Li et al. (1993) was first applied to each individual pixel and the correction factors accounting for aerosol and cloud-top height on atmospheric absorption were added, depending on whether the pixel was identified as clear or cloudy. Areal mean net surface radiation was calculated from all pixels within a region. Different domain sizes were selected centered on the CF. They were computed for every half-hour during the daytime for all the days when satellite data are available. Domain sizes were chosen to be representative of typical scales used in various SCM and general circulation modeling schemes and range from 10 km to 400 km (10, 20, 50, 100, 200, 300, 400 km). With the daytime SSRB calculated for each half-hour in a day, temporal averaging on various scales can be done. Averaging intervals chosen in this study were 1, 2, 4, 8 hours, together with daily means as well as 5- and 10-day means, and monthly means.

Results

To ensure that the satellite-inferred SSRB can reasonably reproduce surface measurements, satellite-retrieved values are first validated against ground-based observations. From 1-km satellite retrievals at half-hour intervals, mean surface net radiation was computed over a gridbox of 4 km centered on the CF and averaged over one hour. The gridbox mean fluxes were compared to ground measurements averaged over an hour. Observed surface net radiation was derived from measurements of downwelling surfaces fluxes (corrected for the thermal offset) from the Solar Infrared Radiation Station and measurements of broadband surface albedo from upward and downward pointing radiometers deployed at the CF. In Figure 1, comparisons of surface net radiation from satellite and surface observations are presented. In general, the two sets of data agree fairly well, especially in terms of relative differences. Biases range from 2.2 Wm^{-2} to 27.0 Wm^{-2} and the root-mean-square error (RMSE) from 30 to 48 Wm^{-2} with the smallest biases occurring in March and the smallest RMSE in September. Seasonable changes in surface albedo may play an important role in the biases, while the RMSE is dictated primarily by cloud variability. To a large extent, the scatter in the plots is caused by the mismatch of satellite estimates and ground measurements in time and space. To gain further insight into the discrepancies related to data sampling, two comparisons are made over a spatial domain of 400 km. Figure 2 is a comparison between the mean of the satellite retrievals over the domain as a function of the mean of the surface observations made at all radiation stations within this domain. While the biases are generally compatible to those shown in Figure 1, the RMSEs are reduced substantially. The reduced RMSEs, albeit still significant, is attributed almost exclusively to the sampling uncertainties. This is clearly seen from Figure 3, which shows comparisons between one retrieval (the mean of all pixels falling inside the 400-km domain) and another retrieval (the mean of the retrievals over 4-km gridboxes surrounding each extended facility located within the domain). Since the two sets of data were all retrieved from satellite data using the same algorithm, their differences attest to the sampling uncertainties. The fact that the RMSE values are similar indicates that the bulk of scattering in the satellite-surface comparisons originated from a combination of sampling and cloud variability. This finding suggests a useful tool to investigate observation sampling errors by simulating ground observations and comparing them with areal mean fluxes using satellite retrieval data alone.

Spatial and Temporal Averaging

We first investigate sampling uncertainties incurred by using single-point data to represent a grid of varying size. Such single-point measurements have been widely employed in validating general circulation models with gridbox sizes of ~200 km. Point measurements of surface net radiation at the CF were simulated and compared with areal means in order to evaluate their representativeness over regions of varying scales typical of model grid cells. Figure 4 shows an example of the March 2000 comparisons of satellite-estimated net SSRB averaged over a 4-km gridbox surrounding the CF (a proxy of ground observations at the CF) against those retrieved over areas of varying size (10, 100, 200, and

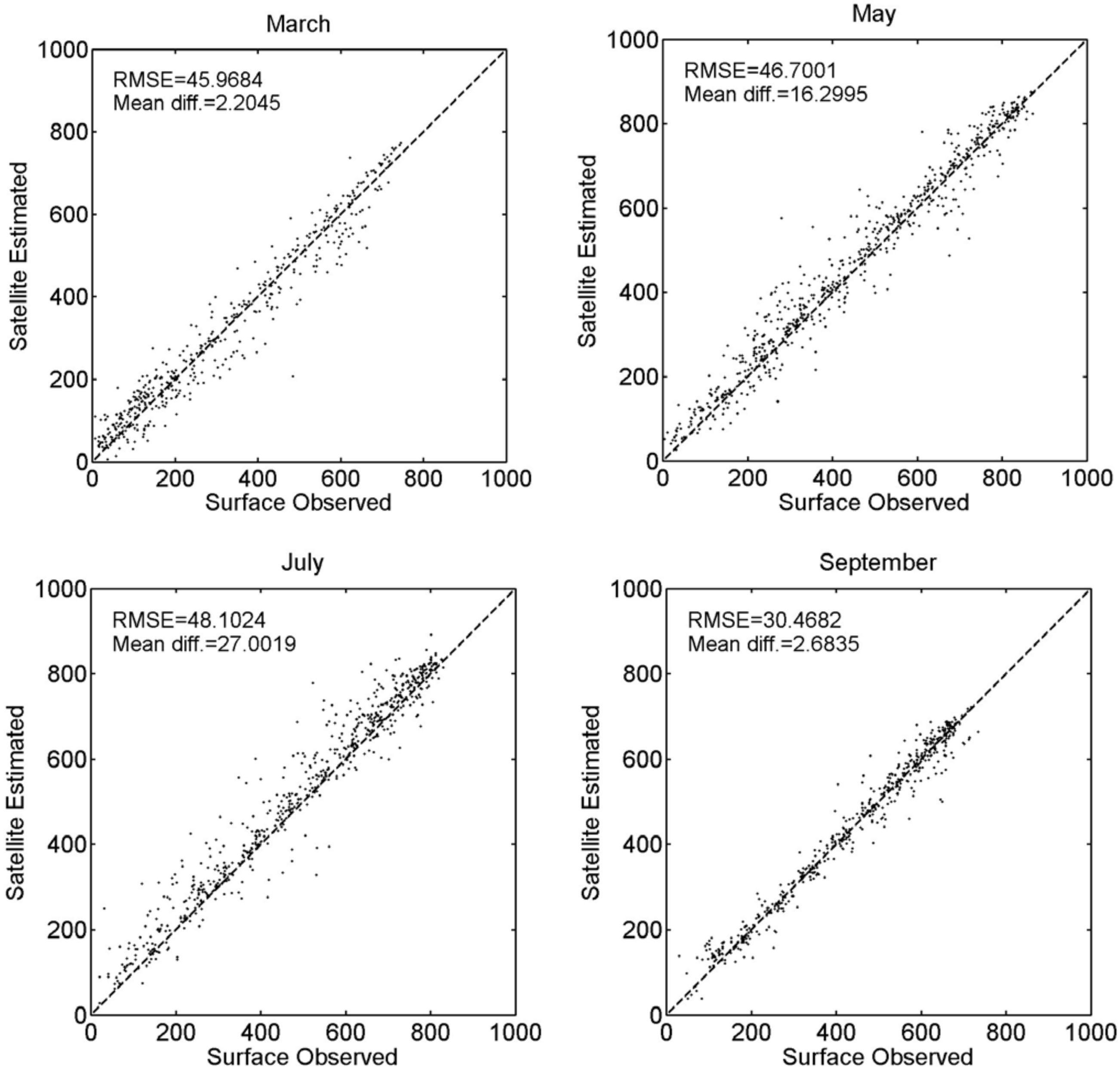

Figure 1. Comparisons of surface net radiation observed at the CF and estimated from satellite (averaged over a 4-km gridbox centered on the CF) for March, May, July, and September.

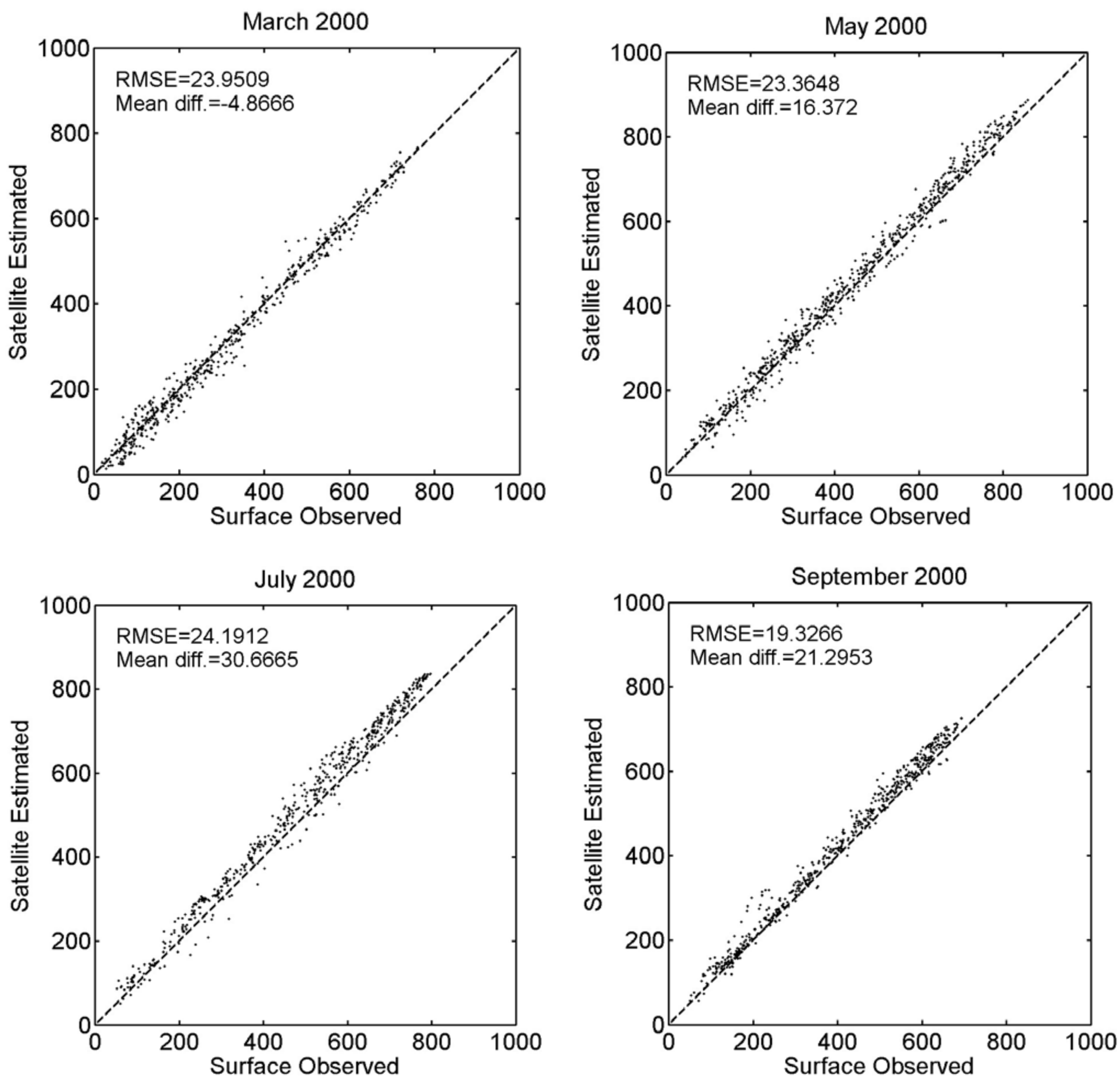

Figure 2. Same as Figure 1 but for a 400-km domain. The satellite estimates are mean values of all pixels over the entire domain and the ground observations are from 21 radiation stations.

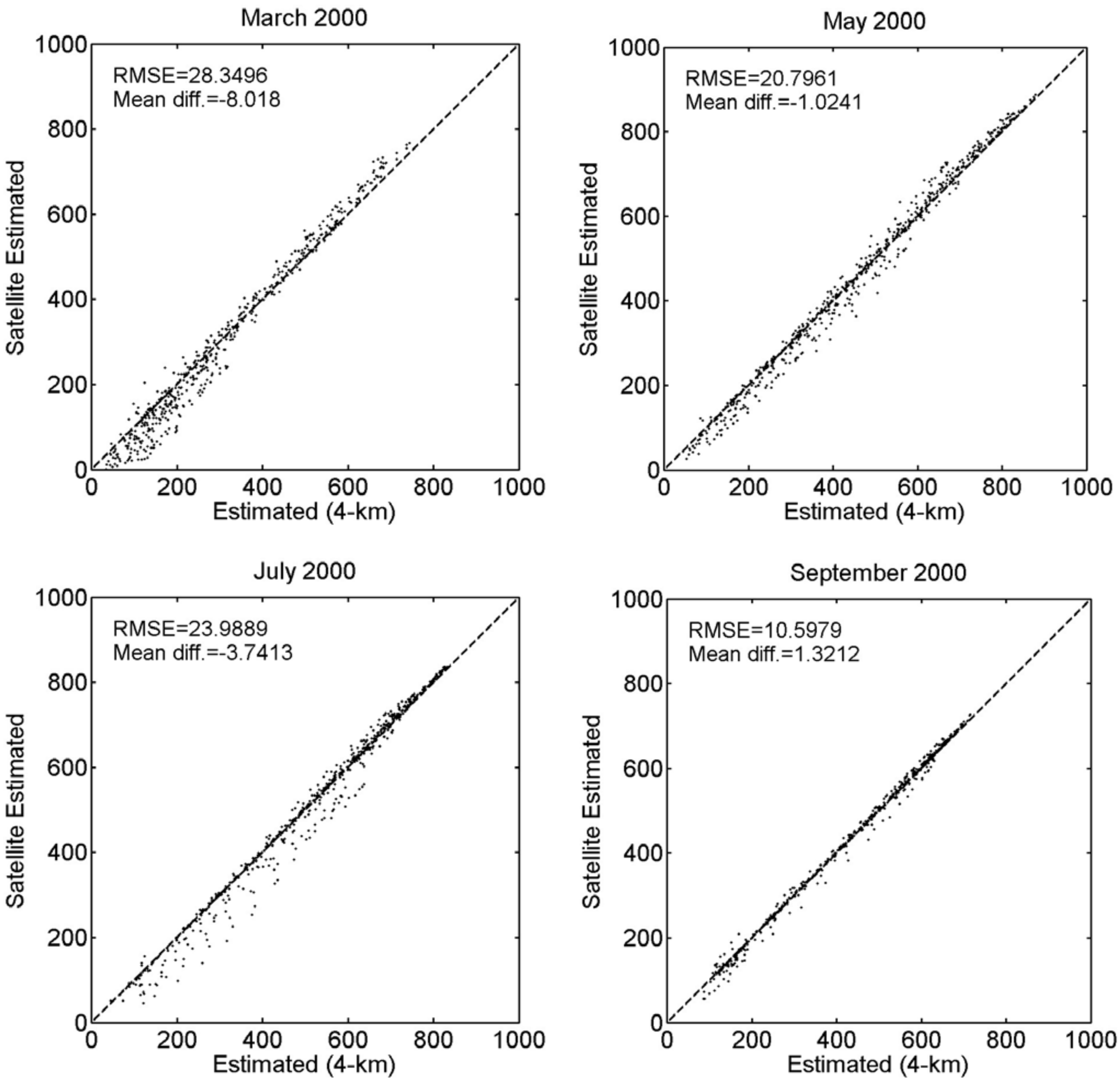

Figure 3. Similar to Figure 2 except that the ground data are simulated by satellite estimates over each station (4-km).

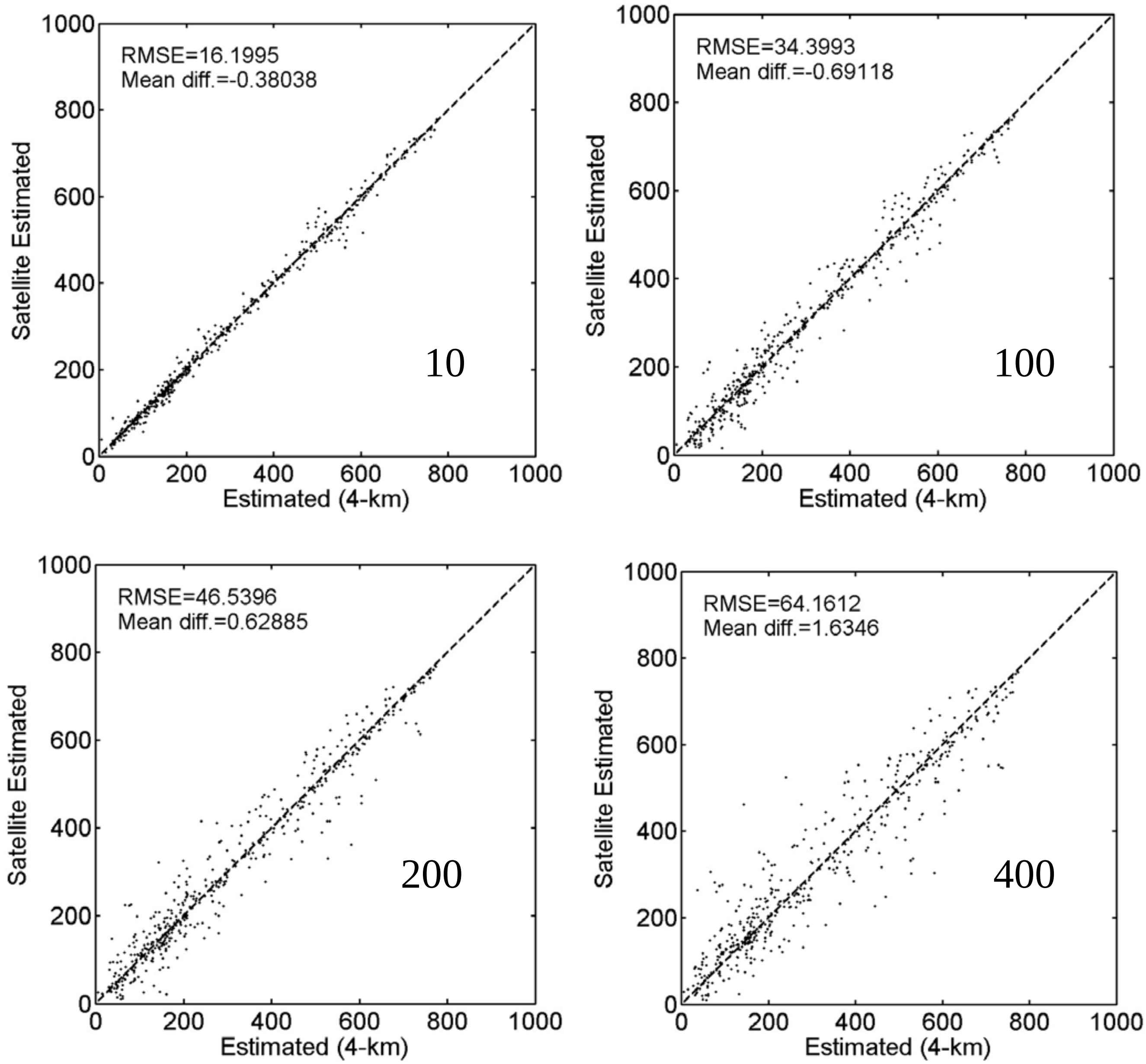

Figure 4. Modeled surface net radiation (averaged over 1 hour and for 10,100,200,400 km) as a function of modeled surface net radiation (averaged over 1 hour and over a 4-km domain). The green line is the "best-fit" line through the data points. Data shown is for the month of March 2000. Units are in W m^{-2}.

400 km), typical of grid cells for cloud resolving models up to general circulation models. As the gridbox size increases, the RMSE increases dramatically; more variability is incurred by changes in the cloud conditions. Note that the RMSE values denote the inherent uncertainties of ground-based observations due to spatial sampling. As a result, one should not expect, nor attempt to reduce, their model-observation differences to within this range. For example, for a cloud-resolving model with a 10-km gridbox, the best accuracy one may achieve would be 16 Wm^{-2} to within the hourly ground-based observations, which increases to 46 and 64 Wm^{-2} for a general circulation model of 200 and 400 km, respectively. Similar trends, but with different RMSE magnitudes, exist for the other months. Besides the gridbox size, the RMSE changes with temporal averaging intervals and cloud regimes. The latter is echoed partially in the comparisons for different months. Figure 5 illustrates the variations of RMSE with all variables by plotting the RMSE as a function of spatial domain size for varying averaging intervals (instantaneous, 1-hour, 4-hour, 1-day, 5-day, and 10-day) for March, May, July, and September. The RMSE varies considerably with month; September shows the least variability. For a 400-km gridbox, the RMSE is less than those for other months by more than 30 Wm^{-2}. July is the most variable month, presumably caused by the prevalence of small convective clouds. However, as the interval increases, the difference among different months diminishes. For a 1-day interval, the RMSE decreases substantially relative to those averaged over hours but the differences among various months are still significant. For typical general circulation model grids of 200 km, the daily sampling error ranges from 16 Wm^{-2} to 28 Wm^{-2} in September and July, respectively; they decrease to 10 Wm^{-2} or less if the model grid is reduced to 10 km. For the 10-day average, however, the differences among the months almost vanish.

Physical processes in the atmosphere occur at various scales in both space and time domains. Over time, the physical state of the atmospheric system can change substantially over a region so it is important to determine appropriate time scales upon which radiative quantities are averaged so that matching errors between model estimates and validating measurements are minimized. Decreases in the matching errors in the SSRB with averaging intervals are shown in Figure 6 for different domain sizes and months. A sharp decrease in error occurs as the averaging interval increases to a day; beyond that, there is a tendency to level off to a stable value of small magnitude.

Figure 7 shows the difference between the monthly mean SSRB over different domain sizes and the monthly mean surface net radiation measured at the CF for the months of March, May, July, and September. The magnitude of the inherent errors due to sampling in surface net radiation is less than 10 Wm^{-2} for all months and domain sizes and the error diminishes to less than 3 Wm^{-2} for typical general circulation model grids of 200 km or less. While this number bodes well with the general requirement of 5 Wm^{-2} for climate studies (Suttles and Ohring 1986), it could be an unrealistic goal for regions/seasons, such as September 2000 over the larger 400-km SGP domain. In this case, the inherent sampling error in monthly mean surface observations is 8 Wm^{-2}, exceeding substantially the required 5 Wm^{-2} unless multiple stations are deployed.

Multiple Stations

Since single-point measurements do not represent well areal means over a large domain, multiple radiation stations distributed over a large domain around the CF may improve the spatial representation of the surface net radiation. General improvement is expected as more observations capture more of the

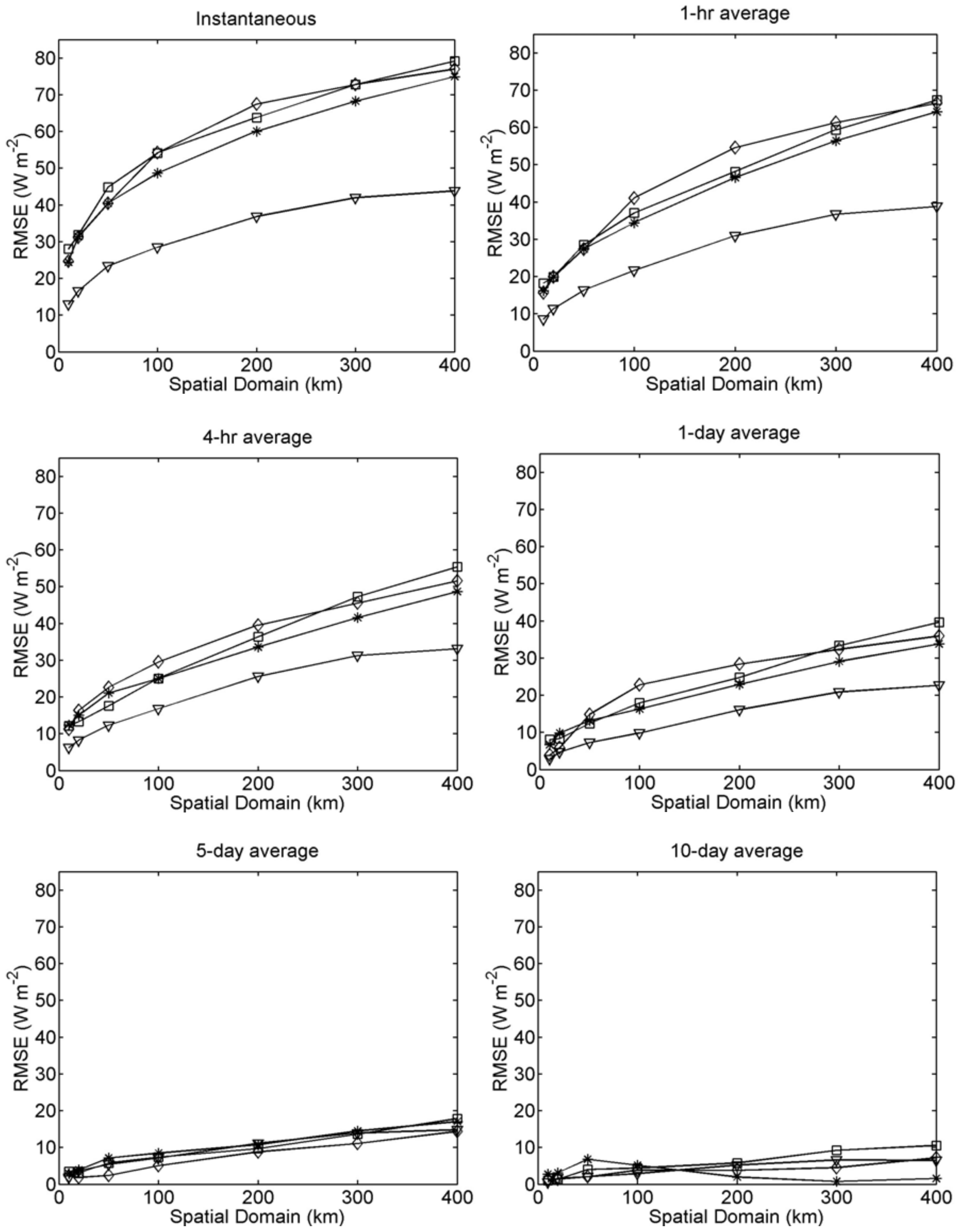

Figure 5. Standard deviation as a function of domain size for the SSRB differences between satellite-simulated "point measurements" and areal mean values over domains of varying size up to 400 km. The comparisons are for instantaneous values and temporally averaged values over different intervals in March (stars), May (squares), July (diamonds), and September (triangles).

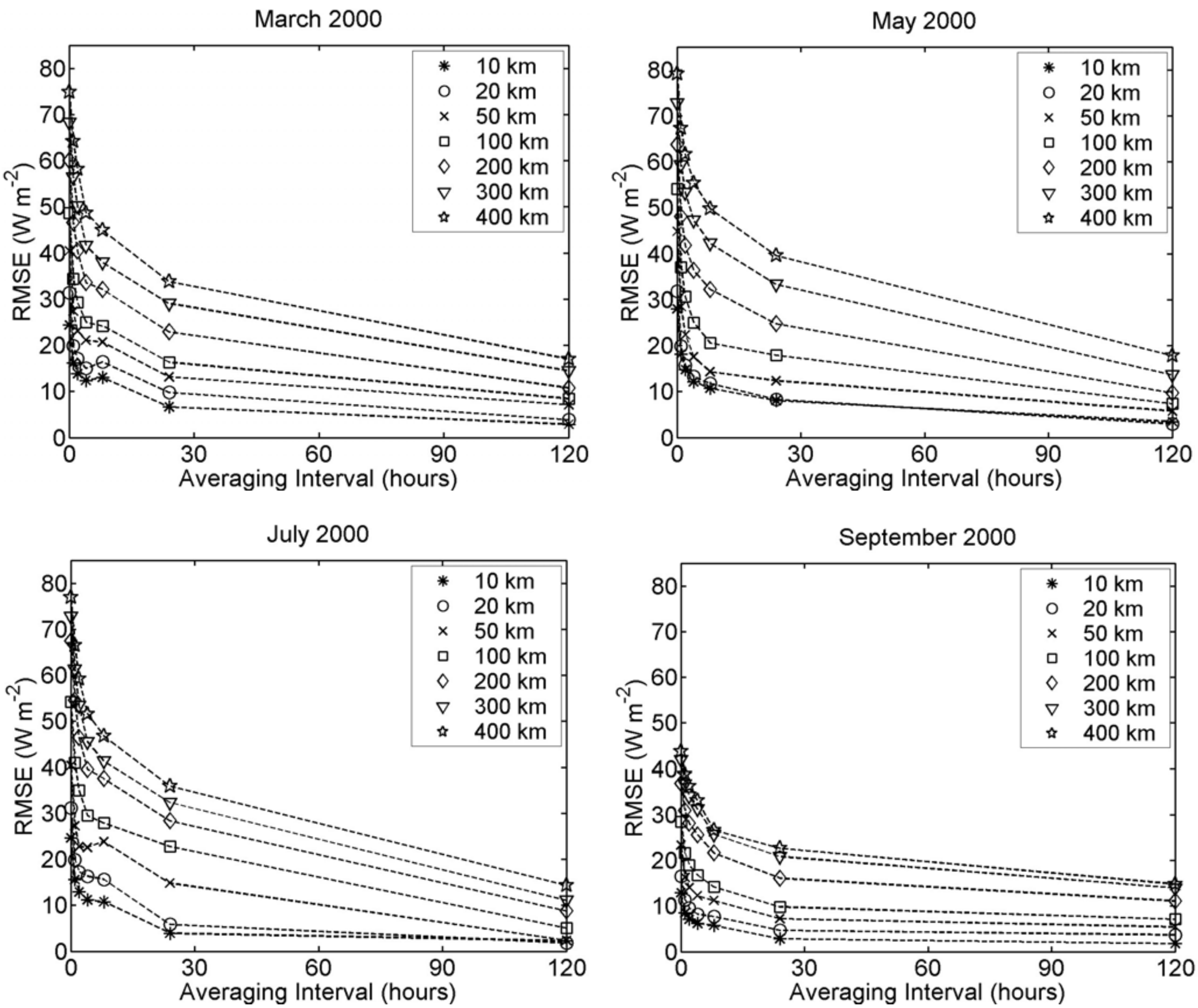

Figure 6. Similar to Figure 5 but plotted as a function of the averaging period (instantaneous, 1-hr, 2-hr, 4-hr, 8-hr, day, 5-day) for different domains and months.

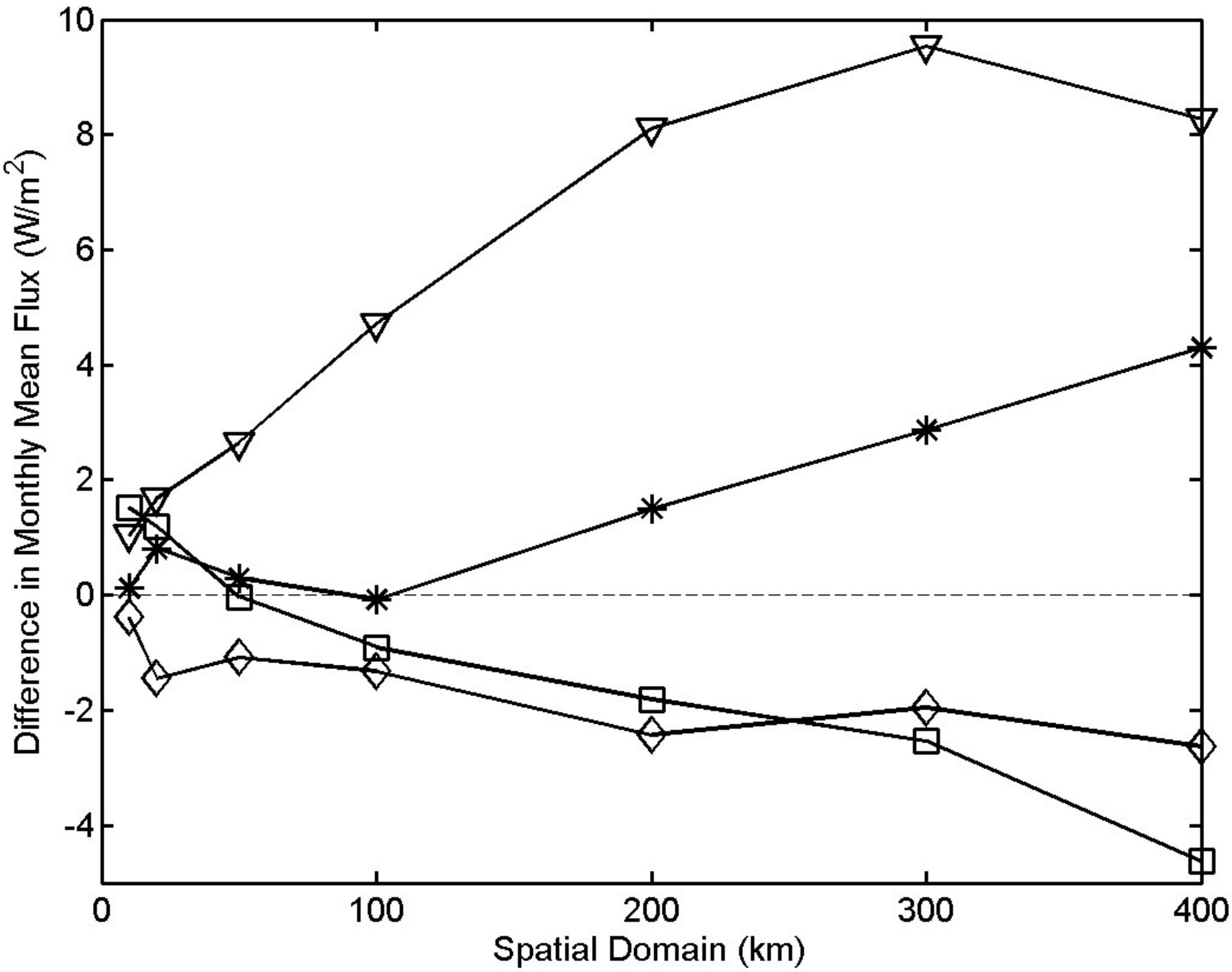

Figure 7. Difference between monthly mean SSRB simulated for "point measurements" and areal mean values as a function of domain size varying from 10 km to 400 km for March (stars), May (squares), July (diamonds), and September (triangles).

variability in the SSRB, but the question remains as to how much improvement is gained as the number of stations increases, and how many stations are really needed to meet certain accuracy requirements.

To address these questions, domains of different sizes centered on the CF were selected: 100, 200, 300, and 400 km, containing 1, 7, 12, and 21 observation sites, respectively. A diagram of the domains used and the locations of the observation sites are given in Figure 8. For each particular domain size, satellite-estimated surface net radiation was calculated within a 4-km gridbox centered on each site as a proxy for actual surface measurements. The mean values averaged over all sites inside a particular domain are compared to the satellite-estimated areal means over the entire domain. It is seen that the sampling error is the largest for the 100-km grid in which there is only one station (the central facility). The magnitudes of the RMSE for other gridbox sizes vary significantly with month but in general, there

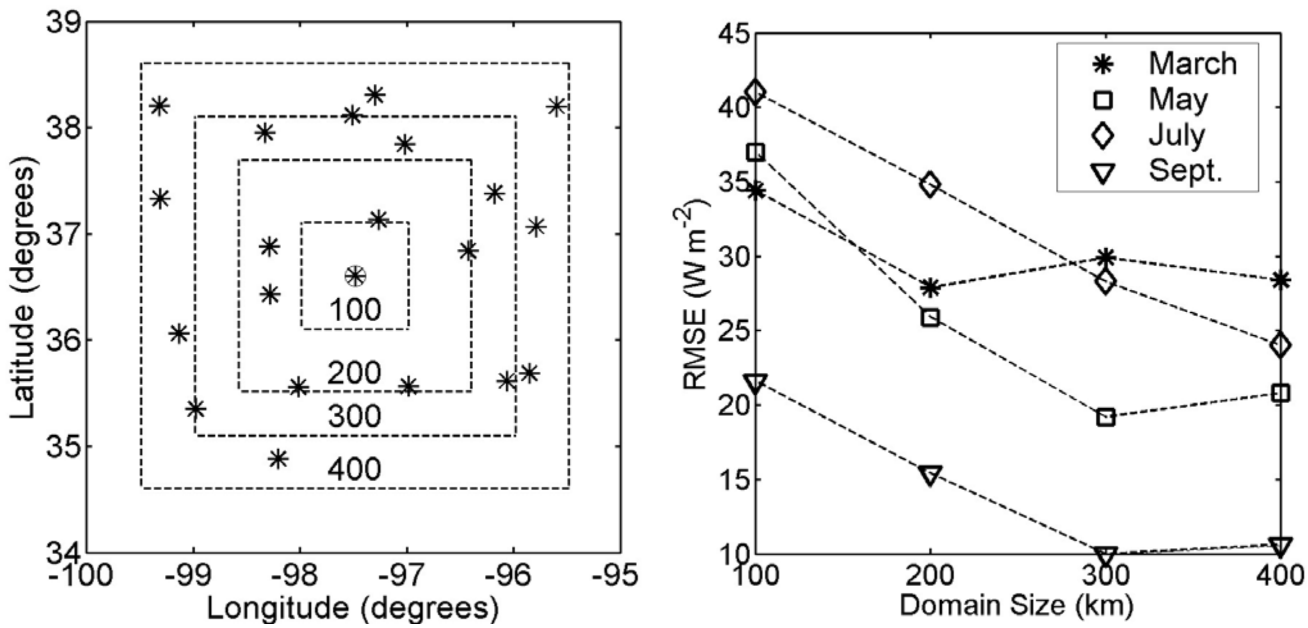

Figure 8. Left panel: Location of extended facilities within the GOES domain with the CF in red. Dashed boxes from small to large represent grid sizes of 100 km, 200 km, 300 km, and 400 km, respectively. Right panel: Error in surface net radiation (between satellite-estimated areal mean and observed) as a function of domain size where the number of observation sites increases with increasing domain size.

is a decrease in the RMSE as more observational sites are included in the calculations over the increasingly larger domain sizes.

Summary

Radiation measurements have been widely employed for evaluating cloud parameterization schemes and model simulation results. In this study, we take advantage of the high spatial and temporal resolution of the Geostationary Operational Environmental Satellite dataset of cloud properties to mimic ground-based measurements of surface net radiation of varying density and temporal frequency and characterize their observation uncertainties caused by cloud variability at different scales in different seasons. Such scale-dependent statistics of observation uncertainties provide critical constraints on model-observation comparisons, and are thus valuable for improving and validating cloud parameterization schemes. In terms of spatial averaging, a single observation site (the central facility) does an increasingly poor job of representing areal means of surface net radiation as the domain size increases. Averaging the surface net radiation at more observation sites results in a decrease in error as the domain size (and number of observation sites) increases. As for temporal averaging, increasing the time interval over which means of surface net radiation are taken leads to a general decrease in error for all domain sizes and all seasons. Instantaneous measurements incur the greatest error in surface net radiation. Some ramifications for climate and cloud modeling studies are that for fine-scale models with short integration intervals, such as cloud-resolving models, point measurements at a single site like the CF do not provide the best representation of radiative quantities. For large-scale models such as general circulation models, use of multiple observation sites does a reasonable job at capturing the surface net radiation field over a large domain. When modeled radiation quantities are compared against ground observations, the inherent uncertainties due to sampling errors must be taken into consideration. Such inherent uncertainties are simulated from our satellite retrievals, which are given as functions of model domain size, averaging periods, number of observation stations, and month. If the difference between modeled and observed radiation quantities is comparable or less than the corresponding inherent uncertainty, no insight may be gained with regard to the model's performance. Such statistics are thus valuable for validating models when testing their parameterization schemes.

Acknowledgements

This research is supported by the U.S. Department of Energy under the Atmospheric Radiation Measurement Program, Grant No. DE-FG02-97ER62361.

Corresponding Author

Z. Li, zli@atmos.umd.edu, (301) 405-6699

References

Holben, B. N., D. Tanré, A. Smirnov, T. F. Eck, I. Slutsker, N. Abuhassan, W. W. Newcomb, J. S. Schafer, B. Chatenet, F. Lavenu, Y. J. Kaufman, J. Vande Castle, A. Setzer, B. Markham, D. Clark, R. Frouin, R. Halthore, A. Kareli, N. T. O'Neill, C. Pietras, R. T. Pinker, K. Voss, and G. Zibordi, 2001: An emerging ground-based aerosol climatology: Aerosol optical depth from AERONET. *J. Geophys. Res.*, **106**, 12,067–12,097.

Kalnay, E., et al., 1996: The NCEP/NCAR reanalysis 40-year project. *Bull. Amer. Meteor. Soc.*, **77**, 437–471.

Li, Z., H. G. Leighton, K. Masuda, and T. Takashima, 1993: Estimation of SW flux absorbed at the surface from TOA reflected flux. *J. Clim.*, **6**, 317–330.

Masuda, K., H. G. Leighton, and Z. Li, 1995: A new parameterization for the determination of solar flux absorbed at the surface from satellite measurements. *J. Clim.*, **8**, 1615–1629.

Suttles, J. T., and G. Ohring, 1986: *Report of the Workshop on Surface Radiation Budget for Climate Applications.* Technical Report WCP-115, World Meteorological Organization.