Relating Atmospheric Radiation Measurement Observations to General Circulation Model Scales

T. P. Barnett, J. Ritchie, J. Foat, G. Stokes
Scripps Institution of Oceanography
La Jolla, California

Introduction

The principal goal of the Atmospheric Radiation Measurement Program (ARM) is to obtain data that will be useful in improving the parameterization of clouds in General Circulation Models (GCMs). With this goal in mind, an extensive field program has been initiated in Oklahoma to obtain the necessary measurements. The field site consists of a central location where an enormous set of cloud and radiation data are being collected. Coming on line is also an array of other, much more limited measurement sites intended to set the central site observations in some perspective.

A major issue arises in trying to use the central site ARM data for the desired parameterization improvement: the data are taken at a single point location but the GCMs address average conditions in a grid box that is typically several hundred kilometers on a side. How does one relate the point measurements to the large area average? The main purpose of this paper is to indicate simple ways in which this might be accomplished.

The MESONET

The data used in this study was obtained from instruments deployed as part of the Oklahoma MESONET (Brock et al. 1995). The MESONET consists of 111 stations deployed throughout the state of Oklahoma, with at least one station in each county (Figure 1).

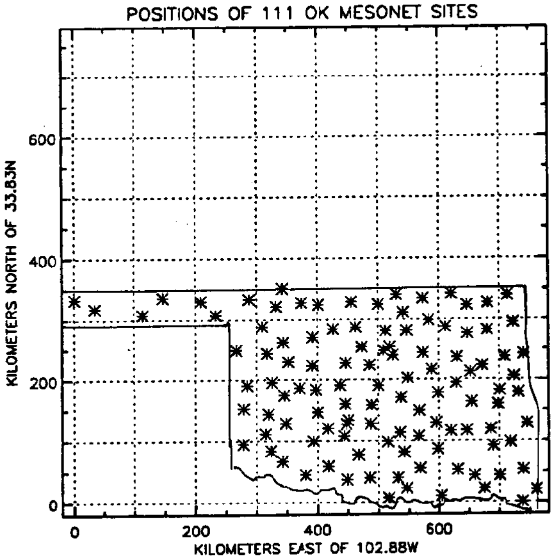

Figure 1. The '*' indicate the locations of the 111 MESONET measurement sites.

The solar radiation data used in this study was collected using a silicon photodiode detector (a Licor 200). The detector is located at a 1.75 meter height on a platform positioned south of the met tower associated with each station. The calibration of the instrument is based on a comparison with an Eppley Precision Spectral Pyranometer. Brock et al. (1995) note that the claimed absolute accuracy of such a calibration is 5%. The instrument dome is cleaned regularly but not daily. Both the temperature and radiation sensors are sampled at 3-second intervals and then averaged into 5-minute mean values.

The MESONET radiation data used in this study were for the months of June, July, and August 1994. Each month was analyzed separately as described in the next sections. Prior to analysis, each time series was visually inspected. Obviously bad data and/or data gaps were omitted in the calculations noted below. Some days had enough missing or problematical data that they were excluded from the analysis altogether. Suffice it to say, potential users of this raw data set ought to be cautious as it is rough.

Data Processing

A first major task was to remove the diurnal signal from the radiation data. This apparently simple preprocessing was required else the daily signal would dominate the analysis and lead to trivial results. It turned out that the most effective way to do this was to consider a typical or average day for a given month, and then use this average day as a basis for normalizing the data for the entire month.

The normalized data set allows for a meaningful comparison of the radiation field not only between different days and stations, but also between different times of day. Figure 2 shows a typical daily time series of incident radiation before and after normalization. As can be deduced from the sunny days in the figure, this approach was not entirely successful in removing the daily signal. This is because a daily signal independent of both geographical location in the state and time of month was used in the normalization. However, from the figure we can also see that the daily signal is dwarfed by that due to cloud activity. In this way, the normalized data set has the advantage that accentuates any variability in the data due to cloud reflection and minimizes that due to the daily cycle.

In order to obtain results relevant to the goal of improving parameterization of clouds in GCMs, we found it desirable to divide the data up into two categories: “sunny” days, characterized by high incident radiation; and “cloudy” days, characterized by low incident radiation. Clearly, there is subjectivity in such characterizations.

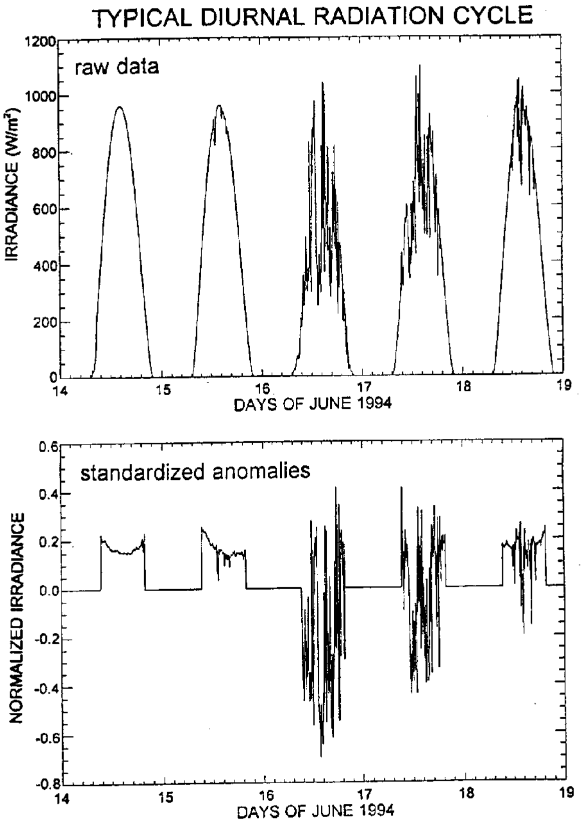

Figure 2. Typical irradiance data for five June days: (a) raw data (top, (b) standardized irradiance.

Estimating a Representative Scale: A Simple Approach

We selected a station in the middle of the MESONET (Spencer, OK) as a “central” site. We denote radiation measurements at that station by $d\ (x_o,T/t)$, with spatial location given by x_o. The data is in time series format but has been averaged in time over successive intervals T. All of the other observed time series in the MESONET array were also individually time averaged over successive intervals T and were represented by $d\ (\underline{x},T/t)$. Next, we computed the spatial average of the $d\ (\underline{x},T/t)$ that lie within a preselected radial distance, R, of x_o. This average, $\bar{d}\ (R,T/t)$, does include $(x_o,T/t)$. We denote the simple correlation between the two time series, $d\ (x_o,T/t)$ and $\bar{d}\ (R,T/t)$, by r (R,T).

A contour map of r is shown in the “two space” defined by the *radial averaging* distance R and the *temporal averaging* time T (Figure 3). As expected, the values of r are large for small R and increasing T. In the limit of R approaching 0, then r will be identically equal to one for all T. In this case, this occurs at R=30 km, the minimum bin size/station spacing that includes only the station at x_o. Likewise, for large R and small T, we expect r to be small (as observed). Note that on cloudy days, values of r first increase with T and then decrease beyond averaging times of roughly 45 minutes. This latter value turns out to be the characteristic time scale of the cloudy day data (see Temporal Considerations below). Averaging beyond this time interval effectively adds uncorrelated information to the analysis and hence reduces r.

When one selects a spatial scale comparable to a T42 GCM grid box (i.e., a radial distance of about 140 km), one needs to use averaging times of 90-120 minutes in order to have a correlation of 0.90 between the “central” site and the average of the radiation field within a “T42 GCM grid box” for sunny-day formulations. For cloudy days, the central site is representative of a grid box less than 100 km on a side at r = 0.90. Even longer averaging times will produce higher correlations for the sunny days, but the maximum for cloudy days is realized at averaging times of about 45 minutes. Thus, the expected correlation between the central site measurements and average conditions over, say, a T42 box around the central site is under 0.7 (the correlation value at R=140, T=45 min.) so they share only about 50% common variance. Investigators have to select a correlation value they are comfortable with and that selection, plus the grid size of the GCM for which they are developing a parameterization, will determine the time average, call it T', that must be applied to the central site data to make the latter representative of the GCM scale radiation field.

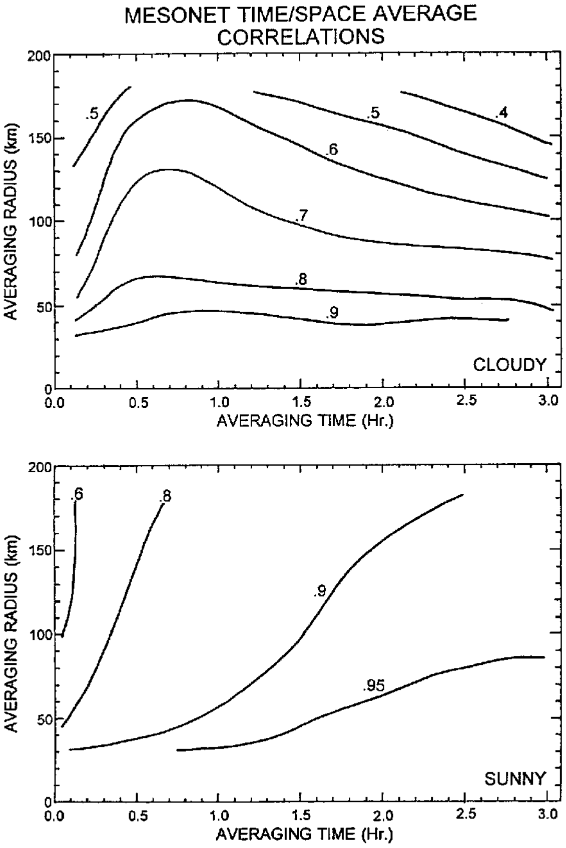

Figure 3. Correlations between a single site at the center of the MESONET array and the average of the station data around the site as functions of radius of the averaging area and interval of time averaging for cloudy conditions (top) and clear sky conditions (bottom).

The results shown in Figure 3 immediately raise several problems. For instance, if one selects a correlation level of, say, 0.90 and intends on developing a parameterization for a T42 model then T' is approximately 90-120 minutes (sunny days) or 45 minutes for cloudy days. But the integration time step for a typical T42 model is only 24 minutes. This means that the central site data can only be used for a subgrid scale parameterization routine that will be called every fourth or fifth GCM time step. One can hope that the radiation field varies slowly or linearly over the 4-5 GCM time steps and appeal to an interpolation strategy to overcome this problem. Otherwise, one simply holds the output of the parameterization fixed for 4-5 successive time steps. In either event, the parameterization is clearly non-optimal.

Another problem that is evident is that for the longest averaging time that was statistically feasible,[(a)] 3 hours, the maximum correlation is of order 0.90 for cloudy days and minimum station separation (30 km). This means that the central site is only capturing 80% of the variance represented by the area average of a very high resolution grid, say T216. A correlation of 0.9 between the central site and a grid box comparable to a T42 grid cannot be realized. In both cases, any parameterization made from the central site data alone will be less than optimal. Use of a daily average will increase this value as found by Long and Ackerman (1995), but apparently not be so useful for the parameterization problem.

The Classical Approach

The two dimensional spatial correlation functions for the radiation field were computed for clear/sunny days from data that had been averaged in time over different intervals; much the same strategy used above. A typical smoothed example of such a calculation for July is shown in Figure 4. These results are typical of June and August correlograms (not shown). Surprisingly, the spatial zero crossing for time averages of 5, 60, and 180 minutes on cloudy days are all roughly equivalent. On sunny days, a clear zero crossing is not as well defined as one might expect. The correlation functions for either sky condition are not isotropic, a result more or less independent of the averaging time. The cloudy day correlograms show longer zonal than meridional scales, while clear days are characterized by longest scale lengths in the northeast-southwest direction. These asymmetrics, which would not be revealed by an analysis such as that discussed above, will need to be accounted for in any cloud parameterizations using the central site data alone. In any event, the characteristic scale length for either sky condition is about 300-400 km if one considers only the zero crossing criteria.

The unexpected behavior of the zero crossing found above is, we believe, due to at least two factors. Inspection of Figure 2 shows that we have not been entirely successful in removing the diurnal cycle from the data (e.g., June 14; especially during clear sky conditions). This introduces low-frequency temporal information to the data. At the same time, there are situations where cloud conditions change gradually over the course of a day (e.g., June 15 or 16) and these also introduce low-frequency temporal information into the analysis. Both effects likely are coherent over large space scales and, hence, contribute to the location in lag space of the zero crossing of the correlation. We conclude that this method

---

(a) In the sense that enough realizations were available to do the statistical analyses used in this study.

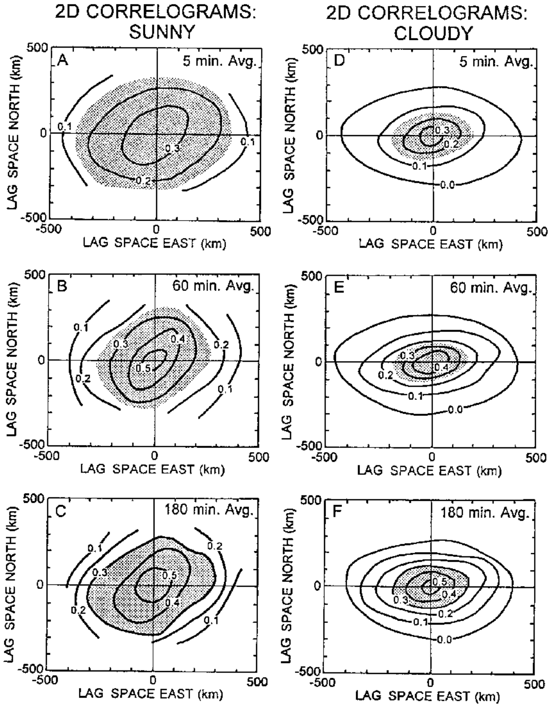

Figure 4. Two dimensional correlation function for (a) sunny and (b) cloudy days for data averaging times of 5, 60, and 180 minutes. The shaded regions are 0.05 confidence levels.

of estimating the characteristic scale length of the radiation field will be unreliable (unless a better job can be done in removing time scales associated with the diurnal cycle and longer periods). These effects are not so critical to the methods discussed above since the simple correlation is strongly weighted by the near-field pairs of data points.

Temporal Considerations

A final feature of the results presented above needs to be explained. The zero crossings of C (X,Y) were relatively insensitive to the temporal averaging interval T for values exceeding 30-45 minutes. In the process of estimating the temporal degrees of freedom we found that the temporal autocorrelation function, computed from 5-minute data of individual radiation time series, typically had first zero crossings in the range 40-50 minutes. According to Figure 3, this averaging time corresponds to a spatial radial scale for cloudy conditions of approximately 60 km if one selects an r of 0.80 or 140 km (T42 scale) if one settles for an r between 0.6-0.7. Thus, averaging times beyond the 40- to 50-minute interval were just producing new realizations of the radiation field at individual stations. Averaging these new realizations would reduce noise and cause C (X,Y) to die away more slowly. This fact, plus the weighting of the simple correlation by the more numerous near-field pairs, no doubt also contributes to the similarity in zero crossing of the 2-dimensional correlation function.

The decorrelation time of 40-50 minutes noted above was typical of all the stations in the MESONET during cloudy conditions for all three summer months studied. If this result holds in general, it implies the central site data is limited in its ability to describe spatial variations in the cloud field. As shown in Figure 3, this limitation is dependent on the level of sameness one desires between the single site and the area average it is claimed to represent. As noted above, the central site can represent an area average comparable to a T42 grid cell only if one accepts a site-area average correlation of order 0.6-0.7 (i.e., the two measures share 36-49% variance in common).

Summary

Several approaches have been taken to estimate how well the ARM central site data observations taken at a single location represent different size area averages. This is a critical estimate to know if one is going to use the central site data to parameterize conditions for GCM grid boxes.

The shortwave radiation data from the dense Oklahoma MESONET array was used to determine the relation between point measurements and area averages. It turns out that the degree of similarity between these quantities depends on averaging time of the data and the state of the cloud field. Further, investigators must decide on the degree of similarity (correlation) they are willing to accept. For instance, on cloudy days, if one requires a correlation of 0.9 between point measurements and associated area averages, then the central site will represent a region with radius 40 km (e.g., a grid cell appropriate to a spectral resolution between T106 and T216). If one relaxes the level of agreement to a correlation of 0.7, about 50% shared variance, the grid box size increases to about 250 km or T42 resolution.

If the MESONET data for the period of study are representative, they suggest a decorrelation time in the shortwave radiation field of roughly 45 minutes. This, in turn, implies a limit on the spatial area that a single-point-measurement can represent. The limit is again conditional on the level of similarity between the point measurement and area average that one is willing to accept.

References

Brock, F.V., K.C. Crawford, R.L. Elliott, G.W. Cuperus, S.J. Sadler, H.L. Johnson, and M.D. Eilts, 1995: The Oklahoma MESONET: A technical overview. *J. Atmos. and Ocean Tech*. **12**:5-19.

Long, C.N., and T.P. Ackerman, 1995: Surface measurements of solar irradiance: A study of the spatial correlation between simultaneous measurements at separated sites. *J. Appl. Met*. **34(5)**:1039-1046.