JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 104, NO. D16, PAGES 19,535–19,549, AUGUST 27, 1999

# An evaluation of NCEP Eta model predictions of surface energy budget and cloud properties by comparison with measured ARM data

Laura M. Hinkelman, Thomas P. Ackerman, and Roger T. Marchand

Department of Meteorology, The Pennsylvania State University, University Park

**Abstract.** Time-series output from the Eta forecast model of the National Weather Service's National Centers for Environmental Prediction was evaluated by comparison with measured values for a location in Oklahoma. The measured data were drawn from the archives of the Department of Energy's Atmospheric Radiation Measurement program for the southern Great Plains site. Surface energy budget components and cloud indicators were examined for the first half of 1997. Overall, the Eta surface energy budget was found to be nearly balanced, as intended by the model physics, except for one instance when light snowfall occurred in the spring. Despite this balance, an average 50 $\mathrm{W\,m^{-2}}$ excess in incoming solar radiation was found. Approximately half of this excess was attributed to deficient extinction of shortwave radiation by water vapor and aerosols in the model, while the remainder appeared to be due to cloud treatment errors. The excess shortwave flux was offset by a smaller negative bias in the downwelling infrared flux and the use of slightly high albedos, in addition to errors of lesser magnitude in the latent and sensible heat fluxes. The upwelling infrared flux and ground heat flux were closer to measured values. Ambiguity in the definition of a cloud as well as measurement limitations hindered the analysis of cloud prediction. However, a rough cloud comparison indicated that the Eta model has more difficulty predicting convection than the movement of large storm systems. In addition, cirrus clouds were predicted too frequently in the winter, while low and midlevel clouds were often missed in the spring. This study demonstrates that single-point time-series data can be used effectively both to ascertain the quality of model output and to investigate the treatment of individual physical processes within models.

## 1. Introduction

Ten to fifteen years ago, most numerical weather prediction models incorporated relatively little of what is usually called model physics in climate models. Radiative fluxes were computed in only the crudest ways, surface processes were largely ignored, and clouds were only considered for their impact on precipitation. This has changed dramatically in recent years. The current generation of weather prediction models tends to include full-physics packages very similar to those employed in climate models. Because of the high spatial resolution of weather models and the scarcity of high-quality measurements of surface and hydrologic properties, these model fields are very attractive for use in a variety of regional budget and climatology studies. The use of data assimilation to blend observations into the model analyses makes the model fields even more attractive for such purposes. [See, for example, *Berbery and Rasmusson,* 1999.] The Global Energy and Water Cycle Experiment (GEWEX) Continental-Scale International Project (GCIP) has begun to archive routinely not only analysis products from several of these regional models, but also forecast time series at approximately 300 sites located within the Mississippi River basin [*Leese,* 1993]. Three forecast models are included in this project to date: the Eta regional model of the U.S. National Centers for Environmental Prediction, the mesoscale analysis and prediction system, or MAPS, developed at the NOAA Forecasting Systems Laboratory, and the global environmental multiscale model (GEM) of the Canadian Atmospheric Environmental Service. The study described here focuses on the Eta model.

The intent in creating this archive of weather model output is to provide research scientists with a set of high-resolution model fields that can be used to study regional hydrology and climate. However, the accuracy of model variables relating to hydrology and surface properties must be evaluated to establish the reliabil-



Paper number 1999JD900120.
0148-0227/99/1999JD900120$09.00

ity of the resulting studies. In addition to the studies conducted as part of model development [see, for example, *Chen et al.*, 1996; *Chen et al.*, 1997; and *Zhao et al.*, 1997], several investigators have recently begun to examine the operational performance of the Eta model in predicting variables of interest to the GCIP community. *Berbery et al.* [1996] evaluated Eta predictions of precipitation and water vapor flux across the United States for the months of August 1993 to March 1994. *Betts et al.* [1997] used 8 days of data from the First International Satellite Land Surface Climatology Project (ISLSCP) Field Experiment (FIFE) to assess the ability of two versions of the Eta model to predict net radiation, sensible heat flux, and latent heat flux at one $15 \times 15$ km$^2$ site. More recently, *Marshall* [1998] investigated the land-surface component of the Eta model by comparing forecasts of soil moisture and temperature for a 3-month period with measurements from the Oklahoma Mesonet. This study also included a detailed examination of the air temperature and specific humidity, latent, sensible, and ground heat fluxes, and the downwelling shortwave and net radiative fluxes at the Mesonet site in Norman over the same time period. *Yucel et al.* [1998] evaluated Eta model output of similar quantities for distributed sites in Arizona as well as in Kansas and Oklahoma over warm season periods from 1994, 1995, and 1996.

Although significant efforts to evaluate the performance of the Eta model in predicting hydrologic and surface variables have been made, further investigation is required for several reasons. First, updates to the model occur regularly, so that each version of the model must be tested separately. Second, many variables are of interest, so that it is not possible to thoroughly evaluate all of these in the course of one study. Finally, a large breadth of investigation, covering long time periods and large areas, is necessary to ascertain the general quality of the model output. While evaluations of operational model performance have often relied on comparisons between model fields over a large portion of the model domain at a given time and corresponding fields from another model or synoptic data set, such comparisons are difficult for hydrologic and surface energy budget variables because the density of observations is insufficient. For this reason, many of the studies listed above have relied on measurements over smaller areas. The long-term archiving of model time series, however, provides another approach if high-quality data are available at the time series location. In such a case, model evaluation can be based on a comparison of the two time series over a longer interval.

The purpose of this study is to provide an assessment of the quality of time series output produced by the Eta model by comparing this output with accurate observations made at a single location for an extended period of time, namely the 6 months from January to June, 1997. Measured data are obtained from the southern Great Plains (SGP) site of the Atmospheric Radiation Measurement (ARM) program, run under the auspices of the Department of Energy [*Stokes and Schwartz*, 1994]. This site is located in north central Oklahoma, within the Mississippi River basin. We examine the major components of the surface energy balance, including all four broadband radiative fluxes, latent and sensible heat fluxes, and the ground heat flux. In addition, surface air temperature, water vapor column, and cloud occurrence are evaluated. Finally, we flag areas of concern and, in some cases, suggest reasons for the discrepancies.

## 2. Data Sources

### 2.1. The Eta Model

The Eta forecast model [*Black*, 1994; *Janjic*, 1994; *Rogers et al.*, 1996] is run twice daily at the National Centers for Environmental Prediction (NCEP) for use in standard National Weather Service forecast operations. In 1997, the horizontal domain of this model covered most of North America with a grid spacing of 48 km. Up to 38 vertical levels were defined in eta pressure coordinates from the surface to about 22 km, with level separations of about 100 m near the ground increasing to about 2 km at the top of the domain. The model had a basic time step of 2 min and was run to provide hourly forecasts over 48-hour periods beginning at 0000 and 1200 UTC. A 12-hour spin-up/data assimilation period preceded each forecast run.

The physics package included radiation, cloud/precipitation, deep and shallow convection, boundary layer, and soil/vegetation schemes. The shortwave radiative transfer code followed *Lacis and Hansen* [1974], while the longwave radiative calculation employed the method of *Fels and Schwarzkopf* [1975]. These full radiative transfer routines were run every second forecast hour. Between these runs, the shortwave surface fluxes and surface albedo were adjusted every 8 min using updated values of the solar zenith angle. At the same intervals, the downward longwave flux was scaled according to the temperature of the lowest atmospheric level, while the upward infrared flux was computed from the surface skin temperature. Clouds were predicted explicitly using an algorithm developed by *Zhao and Carr* [1997] and *Zhao et al.* [1997]. This routine distinguished between frozen and liquid hydrometeors and was run every 8 min in the model. Surface sensible, latent, and ground heat fluxes were also calculated every 8 min using the land-surface scheme described by *Chen et al.* [1997]. This scheme applies a single surface energy balance equation to obtain a single effective skin temperature for the aggregate soil/vegetation/snow surface state.

Because the Eta model is in operational service, improvements and corrections are made regularly though infrequently. For example, in February of 1997, refinements to the cloud, radiation, and land-surface schemes were implemented. The most recent major update to

the Eta model occurred in June 1998. At that time, the Eta Data Assimilation System adopted fully continuous, 3-hourly, Eta-based cycling. As a result of these changes, the version of the model used currently in NCEP forecasting is not identical to that described in this paper. However, the model output from 1997 is still available for research use in the GCIP archive, and the fundamental radiative transfer, cloud prediction, and land-surface schemes remain the same. (A complete history of Eta system changes can be found at http://nic.fb4.noaa.gov:8000/research/gcip.html.)

### 2.2. Measured Energy Fluxes

The measured data used in the comparisons here were taken from the ARM SGP central facility in Lamont, Oklahoma. The goal of the ARM program is to improve the accuracy and utility of numerical models used in climate study and prediction. One thrust of the program is to obtain high-quality long-term data sets of variables that are as yet poorly understood or modeled in a number of climatologically important regions. The SGP site, located at 36.605°N, 97.485°W, at an elevation of 317 m, is the first of several locations to be outfitted with the extensive suite of instruments needed to make these measurements. Other measurement sites are under development at locations in the tropical western Pacific [*Mather et al.*, 1998] and on the North Slope of Alaska [*Stamnes et al.*, 1999].

A variety of data from the ARM southern Great Plains facility were used in this study. Downward radiative fluxes were obtained from instruments in the baseline surface radiation network (BSRN). Total (direct + diffuse) shortwave radiation was measured over the 0.285- to 2.8-$\mu$m wavelength range by an Eppley precision spectral pyranometer with a hemispheric field of view. Downwelling infrared radiation was measured using an Eppley precision infrared pyrgeometer. This instrument also had a hemispheric field of view but was sensitive to radiation in the 3.5- to 50-$\mu$m wavelength range. Both of these instruments are accurate to a few percent. Measurements were reported as 1-min averages of samples taken every second.

Upward radiative fluxes were taken from measurements by the solar and infrared observation system (SIROS) using similar instruments. These measurements were made at a height of 10 m above the surface, and values were recorded at 15-sec intervals. Expected accuracies for these instruments are also a few percent.

Surface heat fluxes were derived from data collected by the energy balance Bowen ratio (EBBR) system. Instrumentation in this system provided the air temperature and relative humidity at two heights, soil temperature and moisture, and the 5-cm soil heat flux. Latent, sensible, and ground heat fluxes computed from these values were provided as ARM products. Individual 30-min averages of the latent and sensible heat fluxes are accurate to ~10% or 10 W m$^{-2}$, whichever is larger. Ground heat flux measurements are accurate to about 15%. The latent and sensible heat fluxes determined by this system are generally incorrect when the Bowen ratio approaches -1. At these points, smoothing was applied to these heat flux time series.

The surface temperatures discussed here were measured by the surface meteorological observing system (SMOS) using a thermistor 2 m above the ground. The accuracy of this sensor is better than 0.6°C.

The water vapor content of the column of air above the central facility was determined from microwave radiometer (MWR) measurements at 23.8 and 31.4 GHz. Because the radiometers are also sensitive to liquid water, it was necessary to eliminate measurements that were made in the presence of standing water on the radiometer dome or that occurred during periods of heavy rain. This was done by discarding radiometer returns that were recorded while an electrical sensor detected water on the cover of the radiometer.

## 3. Surface Energy Budget

### 3.1. Data Description

Downward and upward solar and infrared (IR) radiative fluxes, latent, sensible, and ground heat fluxes, surface temperature, and water vapor path values predicted by the Eta model from January through June, 1997, were obtained for the model grid point nearest the ARM Southern Great Plains facility. (This grid point falls at 36.7032°N, 97.55434°W, or approximately 11 km north and 6 km west of the ARM central facility, and is assigned an elevation of 388.0 m in the Eta topography.) The water vapor path and surface temperature were provided as instantaneous hourly values, while the fluxes were all reported as hourly averaged values. Continuous time series for each variable were created from the Eta output by concatenating the predictions for hours 6 through 17 from each forecast. Of the 361 model runs needed to create these time series, only 10 were not available, and just two of these occurred consecutively (specifically, the 1200 UTC run from June 7 and the 0000 UTC run from June 8). These gaps were eliminated by substituting forecasts for the same hours that were taken from the previous or following runs. Thus the final product consisted of segments of 12 or more consecutive forecasts from 351 model runs.

Analogous time series of ARM data were made by calculating hourly averages of the measured quantities. Because the measurements were made by a number of different instruments, missing or bad values occurred at different times for each data set. Each data type was reviewed independently, and hours for which the measurements were missing or known to be unreliable were removed. The corresponding model values were also eliminated to avoid biasing the comparisons. This generally meant that about 3% fewer than the maximum 4344 hours of data were available for each vari-

able. The hourly time series were converted to series of daily mean values and 30-day running means, and these average time series were used in most comparisons. The convention that fluxes toward the surface are positive while those away from the surface are negative was used for all energy fluxes.

It must be emphasized that, although exact locations for the ARM measurement site and the Eta grid point under consideration have been given, the spatial resolutions of the two data sets are quite different. While the measurements are effectively made at one point, the Eta output corresponds to an entire grid box, with an area of about 2480 $km^2$. In making these comparisons, it is thus assumed that the time average over the ARM site is representative of the entire model grid box. However, spatial averaging does affect the values computed by the model, especially for those variables that depend on surface conditions.

### 3.2. Radiative Fluxes

The 24-hour average downward solar fluxes predicted by the Eta model and measured at the ARM site are shown in Figure 1a. A steady positive bias of about 50 $W m^{-2}$ is apparent in the predicted values. This discrepancy is similar in magnitude to those found by *Marshall* [1998] for a site in Norman, Oklahoma, between May and July of 1997 and by *Berbery et al.* [this issue] for the contiguous United States during August of 1997 and January of 1998. It appears to be somewhat smaller than the bias recorded for southern Arizona for spring and summer months in 1994 and 1995 by *Yucel et al.* [1998]. The difference between Yucel et al.'s results and those of more recent investigations can probably be largely attributed to modifications in the Eta model's treatment of clouds, ozone, and the Earth's orbit that were implemented in February 1997. Nevertheless, perhaps because these changes to the solar flux calculation were made during the period of minimal solar input for the Northern Hemisphere, no corresponding discontinuity appears in the Eta downward shortwave flux curve shown in Figure 1a.

Additional features of the curves in Figure 1a are of interest. The effect of clouds on incoming solar radiation is evident in the high-frequency oscillations of both daily mean value curves. Upon closer inspection, it appears that the Eta-predicted downward solar fluxes are closer to the measured fluxes on days when the sky is free or nearly free of clouds. This suggests that a large portion of the error is due to the prediction and treatment of clouds rather than to problems with the calculation of the downward solar flux at the top of the atmosphere or its transmission through the clear sky. A separate comparison of downward solar flux was carried out using data for only the most cloudy and cloud-free days, as determined from the ARM radiation data. Predicted downward solar fluxes were, in fact, more accurate on clear-sky days. For example, on the 39 cloudiest days, the mean difference between the ARM and Eta values was 87 $W m^{-2}$, while the difference was only 27 $W m^{-2}$ on the 32 clearest days. This indicates that approximately half of the overall excess in downward shortwave flux results from mishandling of clouds by the model, while the other half is produced by inaccuracies in the calculation of downwelling solar radiation in clear-sky conditions.

Two possible sources of the excess downward solar radiation on clear-sky days are the underestimation of shortwave extinction by water vapor and by aerosols. The *Lacis and Hansen* [1974] shortwave radiative transfer parameterization is known to produce insufficient absorption of near-infrared radiation by water vapor [*Ramaswamy and Freidenreich,* 1992]. In addition, absorption and backscattering by aerosols are not explicitly treated in the Eta model. Since February 17, 1997, the incoming solar radiation has been reduced by a constant factor of 3% in an attempt to compensate for this deficiency. However, as shown in Figure 1a, the downward solar radiation has remained too large despite the introduction of this and other corrections.

An additional test was performed to estimate the relative importance of these two sources of error. The expected downward solar flux was calculated hourly for eight clear-sky days using a radiative transfer scheme that combined the spectral model of *Kato et al.* [1999] with the two-stream algorithm devised by *Toon et al.* [1989]. (These eight days were selected because they were among the clearest for both the measurements and predictions and because sounding data for these days were immediately available.) The measured water vapor distributions were accounted for in this calculation, but aerosols were not included. The results of these computations were compared with the Eta fluxes, which only partially accounted for water vapor and aerosols, and the ARM measurements, which inherently included the effects of all the constituents of the atmospheric column over the ARM site. A plot from a representative comparison is shown in Figure 2. In this figure, the increasing difference between the measured solar flux and that calculated using the two-stream model in the afternoon has been traced to the presence of a thin layer of cloud or haze beginning at 1800 UTC. The slight time shift apparent in the Eta output relative to the observations may be due to the manner in which the solar flux values are averaged in the Eta model or to small errors in the Eta solar zenith angle computation. Looking past these differences, the results suggest that the majority of the Eta excess in clear-sky solar radiation is due to insufficient absorption by water vapor but that the contribution due to insufficient absorption and backscattering by aerosols is also significant. It is hoped that a planned upgrade in the radiative transfer routines in the Eta model will correct these deficiencies (T. Black, personal communication, 1998.)

Despite the introduction of excess solar radiation in the Eta model, the daily predicted and measured sur-

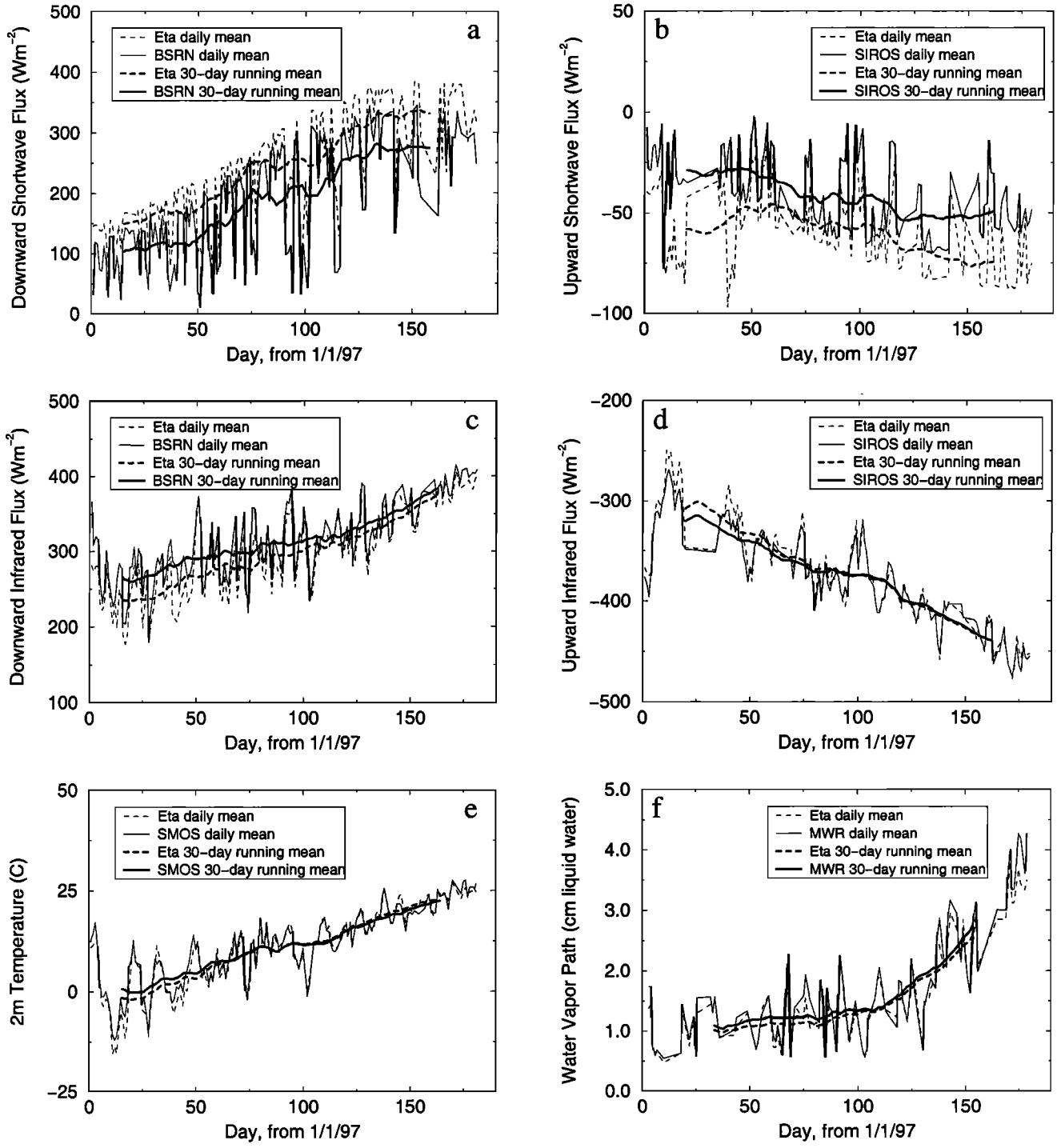


**Figure 1.** Comparison of daily average radiative fluxes, surface temperature, and column water vapor calculated by the Eta model with measured values from the Atmospheric Radiation Measurement (ARM) program southern Great Plains (SGP) site. (a) Downward shortwave flux. (b) Upward shortwave flux. (c) Downward infrared flux. (d) Upward infrared flux. (e) Temperature at 2 m. (f) Column water vapor path. Acronyms refer to instruments and are defined in section 2.2.

face (2 m) temperatures (Figure 1e) were found to agree well. The only errors evident in the Eta output were a slight negative bias in winter and a small positive bias in the summer. This implies that the excess incoming solar radiation must by compensated for by errors elsewhere in the energy budget.

The measured and predicted upward shortwave fluxes are presented in Figure 1b. As for the downward solar flux, the Eta model consistently predicts too much upward shortwave radiation, but in this case the average excess is only about 20 W m$^{-2}$. Most of this difference can be explained as the result of reflection of the excess

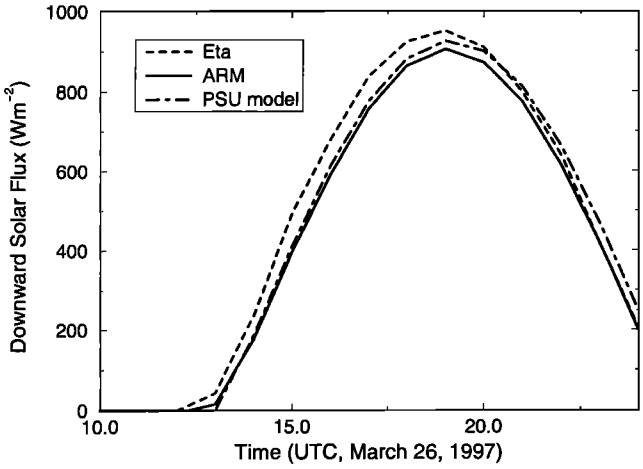


**Figure 2.** Comparison of downward solar flux predicted by the Eta model with values from our radiative transfer calculations and the ARM SGP site measurements. All data were averaged to 1-hour increments.

incoming radiation at the surface. However, the rest of the difference is likely due to the use of slightly elevated albedos for this location in the Eta model.

To test this theory, albedo was calculated for the Eta model as the ratio of the upward shortwave radiation over the downward solar radiation at each forecast hour. (Within the model, however, this process is reversed in that the upward shortwave radiation is derived from calculated albedo and incident solar radiation values.) Albedoes for the ARM site were computed using measurements of upwelling shortwave radiative flux taken over two different surfaces. The first set of measurements came from the previously described downward looking pyranometer suspended 10 m above open pasture as part of the solar and infrared observation system. The second was recorded by a multifilter radiometer suspended 25 m above a field in which winter wheat was grown. In both cases, the downwelling shortwave flux component was taken from the baseline surface radiation network pyranometer discussed above. When calculating all albedos, downward fluxes smaller than 20 W m$^{-2}$ and upward fluxes smaller than 10 W m$^{-2}$ were not considered because they are at the limits of measurement accuracy.

A comparison of the measured and predicted albedos at fixed times over the 6-month period is presented in Figure 3. Results are shown for the hours of 1300, 1600, and 1900 UTC. These curves are not smooth, as one might expect from changes in surface vegetation alone. In January and February, large albedo peaks caused by the presence of snow occur. The timing and duration of these peaks differ somewhat between the measurements and model output. Low-magnitude, higher-frequency fluctuations in all the albedo curves are due to the presence of clouds in the sky. The majority of these albedo variations in the Eta results coincide with fluctuations in the measured values, although the variations differ

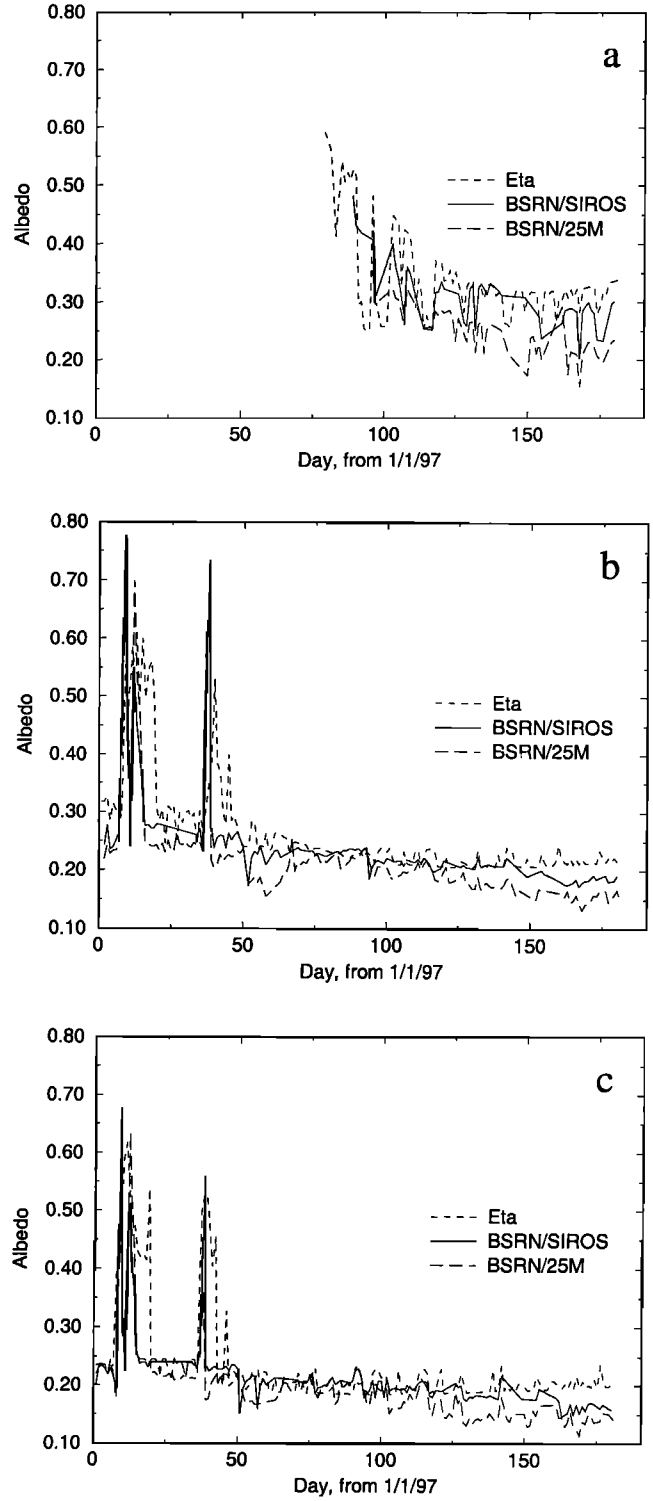


**Figure 3.** Change in albedo with season for a constant time of day: (a) 1300 UTC, (b) 1600 UTC, and (c) 1900 UTC. The "BSRN/SIROS" curve was calculated using upward solar fluxes from a pyranometer suspended 10 m above a pasture as part of the solar and infrared observation system. The "BSRN/25M" data were calculated using upward solar fluxes from a multifilter radiometer 25 m above a field in which winter wheat was grown. Measured downward solar fluxes were taken from an Eppley precision spectral pyranometer, which was part of the baseline surface radiation network.

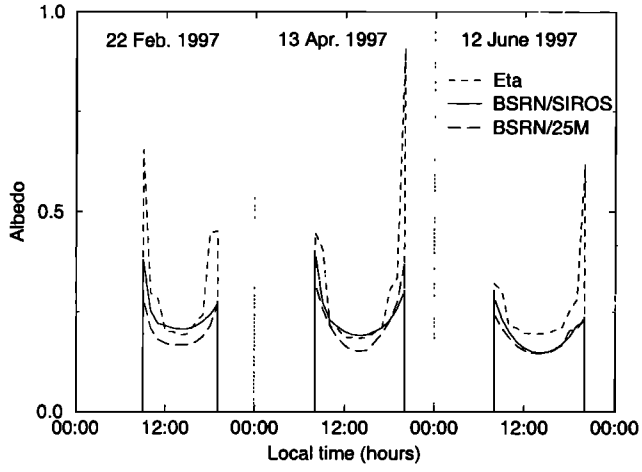


**Figure 4.** Representative daily albedo curves from Eta model output and measurements. Data descriptors are defined in Figure 3.

in magnitude. It can also be noted from Figure 3a that the magnitude of albedos and their deviations increase as the Sun approaches the horizon. (Albedos for the winter months are not shown in Figure 3a because the fluxes are below the thresholds mentioned above.) Despite all of these variations, however, the albedos from the Eta model follow the general trends of the measured values but tend to be biased slightly high.

The variation of albedo over the course of the day was also investigated. Sample albedo curves from individual clear-sky days at different times of the year are given in Figure 4. The generally positive bias in the Eta albedos is evident here. In addition, it is apparent that the model tends to create large (over 0.50) spikes in the surface albedo at low solar angles. These spikes contribute 0.8-6.1 $W m^{-2}$ to the daily-average upward solar radiation, slightly offsetting the excess in downward shortwave flux. In addition, their presence indicates that the Eta model calculation of changes in albedo with solar zenith angle is not entirely correct. Problems in accounting for the changing solar zenith angle may also account for the fact that the Eta albedos vary in discrete steps rather than smoothly like the measured values, although the sampling interval was the same for both predictions and measurements.

The predicted and measured downward infrared fluxes are presented in Figure 1c. Although the predicted downward infrared fluxes appear to track the measurement values well, the 30-day running means reveal a negative bias in the Eta predictions that decreases from about 30 $W m^{-2}$ in winter to 10 $W m^{-2}$ in spring. This bias and its evolution could be due to errors in the treatment of water vapor in the model in addition to changes in the radiation subroutines in February of 1997. In order to test this hypothesis, the predicted and measured column water vapor amounts were compared. As shown in Figure 1f, the predicted and measured water vapor paths are very similar. However, the predictions are always lower than the measured values, with an average bias of approximately 0.1 cm. Although this difference is within the measurement error for this quantity, the fact that the predicted values are consistently low could explain part of the deficiency in downward IR flux predicted by the Eta model. The reduced water vapor paths could also contribute to the excess in downward solar radiation noted above.

The upward infrared fluxes calculated by the Eta model are quite accurate, as illustrated by Figure 1d. Except during the winter, the Eta and ARM 30-day running mean curves overlap. Over the entire 6 months, the predicted upward infrared flux averages only 4.0 $W m^{-2}$ less than the measured value. This reflects the accuracy of the surface temperatures predicted by the Eta model.

### 3.3. Surface Heat Fluxes

The latent, sensible, and ground heat fluxes from the Eta model were also compared with measurements. The latent heat flux curve for the first 181 days of 1997, shown in Figure 5a, contains many of the same variations as the measured curve. However, for much of the time the predicted values are greater in magnitude than the measurements. This is especially true during the period from day 30 to day 115. The maximum difference between the two is about 30 $W m^{-2}$ at day 60.

The daily predicted sensible heat flux, given in Figure 5b, is somewhat smaller on average than the corresponding measured flux for about the first 80 days of the year. This could compensate to some extent for the excess latent heat flux computed by the model for this period. From day 80 until the end of June, the sensible heat flux obtained from the model is substantially greater in magnitude than the sensible heat flux that was measured. Again, the maximum difference in magnitude is about 30 $W m^{-2}$. Summing the sensible and latent heat fluxes, we find that the combined Eta heat flux is nearly always larger than the corresponding value from the ARM data.

The measured and predicted ground heat fluxes are in better agreement. As shown in Figure 5c, the 30-day mean curves for the two data sets match well. The average difference between the daily ARM and Eta values is just 0.6 $W m^{-2}$, which is well within the limits of measurement accuracy for this variable. However, the cross-correlation coefficient for the daily mean ground heat flux curves is less than 0.67, indicating that the measured and predicted values sometimes fluctuate differently.

### 3.4. Total Energy Budget

It would appear to be superfluous to calculate the total energy budget for a numerical weather prediction model because most models explicitly require zero net energy flux. However, the Eta energy balance equation

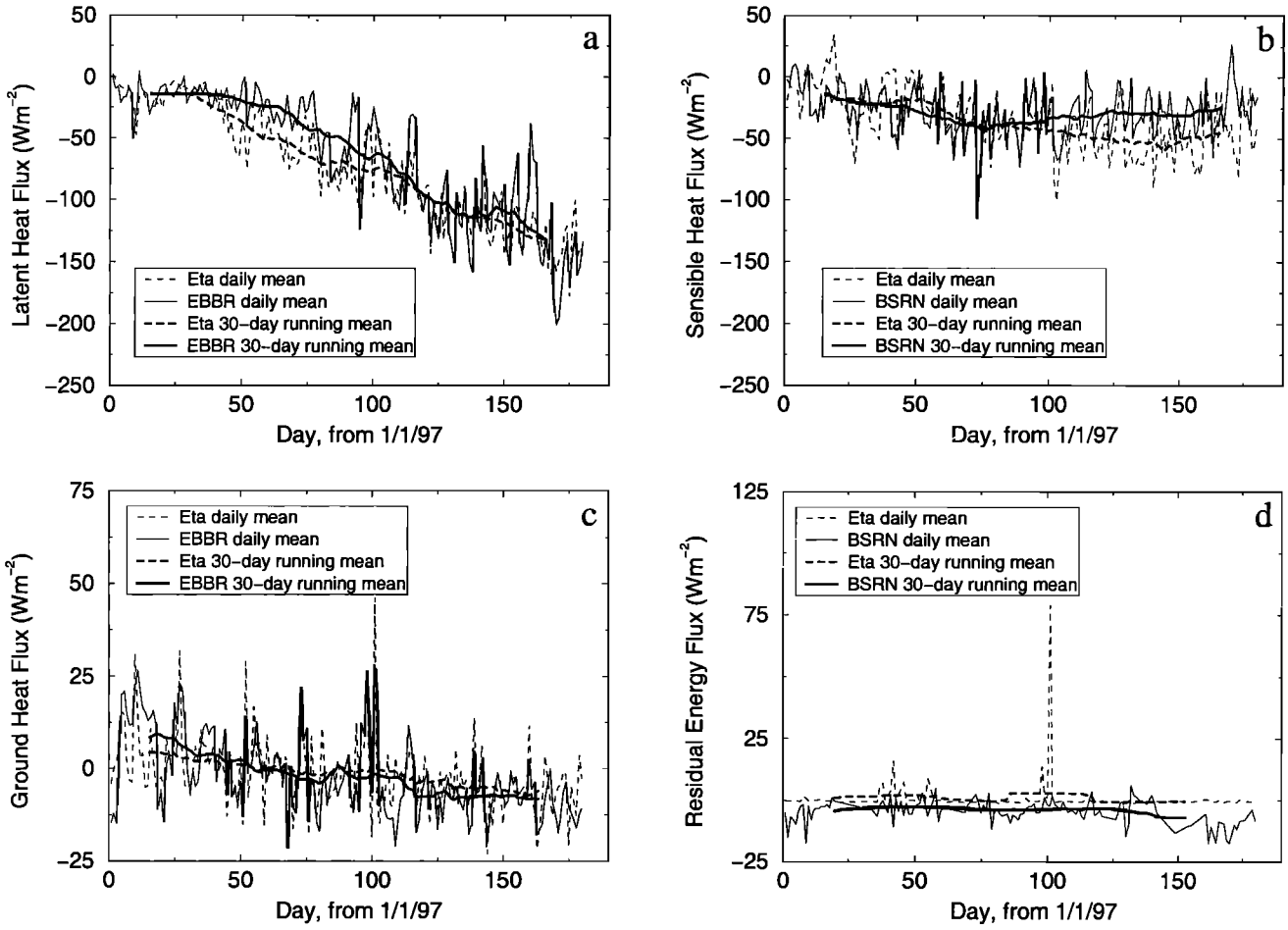


**Figure 5.** Comparison of daily average heat fluxes calculated by the Eta model with measurements from the ARM SGP site. (a) Latent heat flux. (b) Sensible heat flux. (c) Ground heat flux. (d) Residual energy flux. Acronyms refer to instruments and are defined in section 2.2.

is solved using a linearization method rather than iteration in order to reduce computation time. The linearization technique allows small (up to about 20 $\mathrm{W\,m^{-2}}$) energy budget residuals to occur. In addition, checking the net energy flux at the surface seemed a logical precaution, especially since discrepancies were found for individual energy flux terms.

The total energy budget was calculated for the measurements via

$$E_{\mathrm{tot}} = \mathrm{SW}\downarrow + \mathrm{SW}\uparrow + \mathrm{LW}\downarrow + \mathrm{LW}\uparrow + \mathrm{LHF} + \mathrm{SHF} + \mathrm{GHF}, \quad (1)$$

where $E_{\mathrm{tot}}$ is the total, or residual, energy flux, SW$\downarrow$, SW$\uparrow$, LW$\downarrow$, and LW$\uparrow$ are the downward and upward shortwave and longwave radiative fluxes, and LHF, SHF, and GHF are the latent, sensible, and ground heat fluxes, respectively. For the model output, the same equation was used with the addition of the heat flux due to snow melting. (This term was not available from the measurements.)

Daily values for the energy budget residual are shown in Figure 5d. Overall, the Eta model energy budget is within the range of accuracy of the measured budget. The measured values range between -35 and 22 $\mathrm{W\,m^{-2}}$ while all the calculated values but one fall between -2 and 18 $\mathrm{W\,m^{-2}}$. On April 11, or day 101, there is an anomalously large residual in the Eta energy balance, with a value close to 80 $\mathrm{W\,m^{-2}}$. This residual has its source in the unusually high ground heat fluxes that are calculated for this time. Since the model indicates that snow fell on days 98, 99, 101, and 102, this discrepancy may be due to mishandling of the effects of snow falling on warm ground. Given the quality of the energy budget on the remaining 180 days, this single anomaly does not seem to be a great cause for concern. However, the NCEP staff plans to investigate this problem (K. Mitchell, personal communication, 1998).

Statistics for the common energy flux terms and the energy budget residual are presented in Table 1. In addition to quantifying the comparisons made between the daily time series, this table lists the coefficient of correlation between the measured and modeled time series for each variable. These coefficients indicate that the modeled upward and downward infrared fluxes and latent heat flux track their measured counterparts very

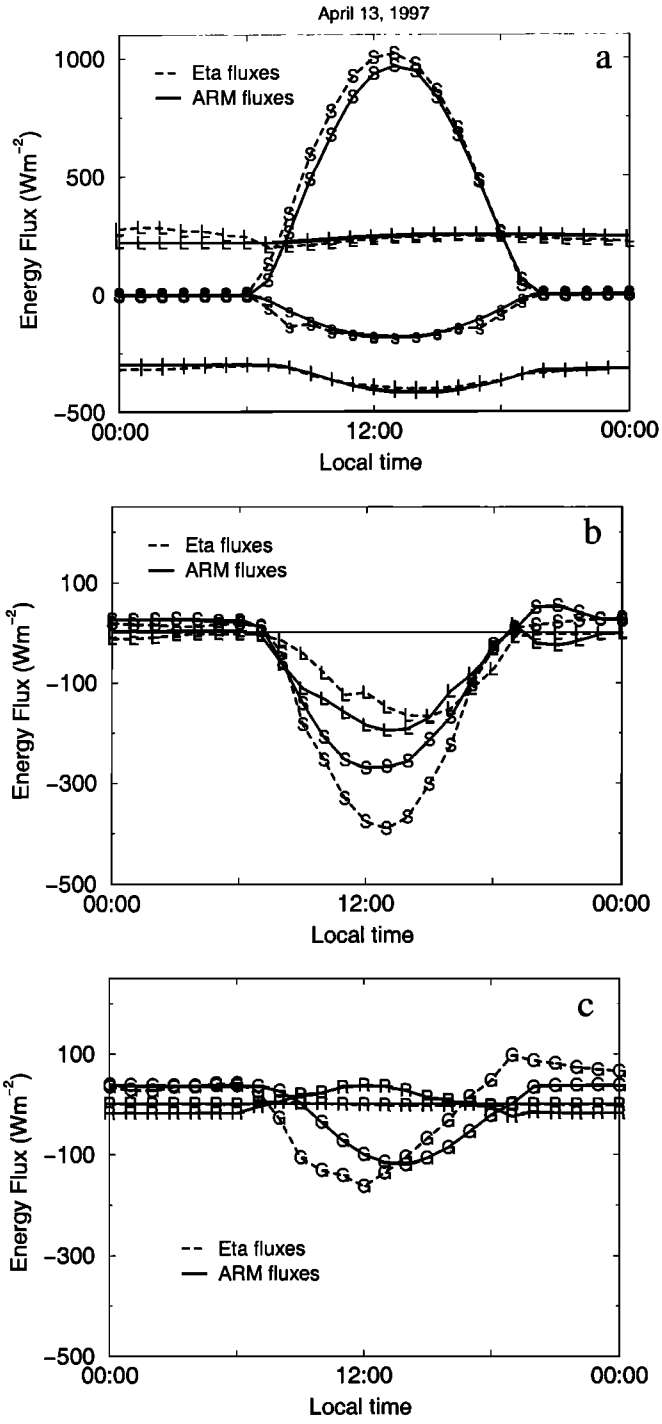


**Figure 6.** Eta and ARM energy budgets for April 13, 1997. (a) Radiative terms: S, downward shortwave; L, downward infrared; s, upward shortwave; l, upward infrared. (b) Heat fluxes: L, latent heat; S, sensible heat. (c) Heat fluxes: G, ground heat; R, energy budget residual.

well, while the ground heat and the sensible heat fluxes predicted do not track the corresponding measurements nearly as well. The residual errors in the measured and predicted energy budgets are mostly small and independent of each other. The correlation coefficient in this case is close to zero.

It should be emphasized that the energy balance and its subterms discussed to this point have been analyzed at the timescale of 1 day for ease in presentation. However, individual terms and therefore the entire energy budget can vary on much shorter timescales. An example of this is given in Figure 6, which shows the energy budget terms for one clear-sky day. From this figure, it appears that the level of agreement between the forecasted and measured energy terms varies with the time of day. To determine whether this is true for the entire 6-month period, we examined the total radiative forcing,

$$E_{\mathrm{rf}} = \mathrm{SW}\downarrow + \mathrm{SW}\uparrow + \mathrm{LW}\downarrow, \qquad (2)$$

during the daytime and nighttime hours of the first 6 months of 1997. There we found a 40 W m$^{-2}$ excess of radiative forcing in the Eta model during the day and a deficit of about 15 W m$^{-2}$ during the night. The timing and magnitude of these radiative forcing discrepancies can be expected to cause errors in other parts of the model energy budget and dynamics at specific times of the day.

## 4. Cloud Comparison

### 4.1. Data Description

The accurate prediction of cloud presence or absence, height, and liquid water content is essential to obtaining correct energy budget results. Clouds reflect solar radiation away from the Earth and reradiate absorbed infrared radiation toward the surface. Correct timing of cloud predictions is especially important for solar radiation since the intensity of incoming shortwave radiation depends strongly on time of day. With this in mind, we analyzed cloud prediction in the Eta model by comparison to measurements made by the millimeter wavelength cloud radar at the ARM SGP site.

The 35-GHz cloud radar at the SGP central facility measured the Doppler power spectrum of any hydrometeors in the column above it. From this, vertical profiles of reflectivity, Doppler velocity, and Doppler width were derived. Unfortunately, a precise determination of cloud water content is not possible from these data alone. Instead, the presence of hydrometeors was inferred from the magnitude of the radar return after correction for clutter due to noise, flying insects, or other airborne material. The received radar signal was considered to indicate the presence of a cloud when the clutter-reduced returned power was distinguishable from the noise produced by the radar receiver. However, such signals could also be caused by precipitation or unwanted returns missed by the clutter-reduction algorithm. Therefore, this criterion did not definitively indicate the presence of a cloud above the SGP site. Nevertheless, the radar's high sensitivity and the fact that radar waves can penetrate through low clouds to detect higher layers above made these some of the most reliable cloud observations available.

**Table 1.** Statistics for Daily-Averaged Values

| Parameter | Source | Mean | Standard Deviation | Number of Points | Correlation Coefficient |
|---|---|---|---|---|---|
| Downward solar flux, $W m^{-2}$ | Eta | 238.9 | 87.5 | 165 | |
| | ARM | 188.1 | 94.6 | 165 | 0.9793 |
| Upward solar flux, $W m^{-2}$ | Eta | -60.4 | 18.9 | 142 | |
| | ARM | -41.0 | 18.5 | 142 | 0.9589 |
| Downward IR flux, $W m^{-2}$ | Eta | 296.2 | 54.7 | 171 | |
| | ARM | 314.1 | 49.7 | 171 | 0.9986 |
| Upward IR flux, $W m^{-2}$ | Eta | -372.0 | 51.9 | 142 | |
| | ARM | -376.0 | 47.4 | 142 | 0.9997 |
| Latent heat flux, $W m^{-2}$ | Eta | -71.8 | 45.4 | 179 | |
| | ARM | -63.7 | 51.0 | 179 | 0.9568 |
| Sensible heat flux, $W m^{-2}$ | Eta | -35.6 | 26.0 | 179 | |
| | ARM | -27.9 | 20.1 | 179 | 0.8286 |
| Ground heat flux, $W m^{-2}$ | Eta | -1.1 | 9.8 | 175 | |
| | ARM | -1.7 | 10.9 | 175 | 0.6665 |
| Residual energy flux, $W m^{-2}$ | Eta | 0.8 | 7.3 | 132 | |
| | ARM | -4.6 | 4.5 | 132 | -0.0511 |
| Surface temperature, °C | Eta | 10.5 | 9.8 | 170 | |
| | ARM | 11.0 | 9.0 | 170 | 0.9916 |
| Water vapor path, cm | Eta | 1.60 | 0.83 | 101 | |
| | ARM | 1.69 | 0.91 | 101 | 0.9966 |

Measured cloud files were constructed by analyzing ARM SGP cloud radar reflectivity data that had been mapped onto a 5 min by 45 m time-height grid as described in the appendix. The radar height bins were mapped to the corresponding Eta model pressure levels, and the number of significant returns detected per hour at each pressure level was determined. Detection of significant return in at least 50% of the 5-min samples constituting an hour was considered to indicate that cloud had been present during that hour. To test the sensitivity of this criterion, data sets were also created for which the minimum fraction of significant returns required to indicate detection of a cloud was fixed at 17% and 83% of the 5-min samples, respectively. If fewer than six 5-min samples were available for a given hour, no cloud information was recorded.

The only cloud-related variables available from the Eta model for the entire 6-month period under consideration were the liquid and/or ice water mixing ratio, in kg $kg^{-1}$, and the cloud fraction for each pressure level and forecast hour. We chose to examine the ice/water content, calculated from the reported mixing ratios, because it has a clearer relationship to radar cloud measurements than cloud fraction does. Still, the use of this variable presented some difficulties. For instance, although liquid water contents below 0.01 or 0.02 $g m^{-3}$ are not generally detectable by current cloud measurement equipment, arbitrarily small liquid water contents can be predicted by the model. Thus it is not clear how clouds are best defined in terms of water content for comparison purposes. In this study, we have used 0.01 $g m^{-3}$ as the minimum value of liquid/ice water content required to define the occurrence of a cloud for the model output; however, other thresholds could reasonably be employed. Additionally, it may be desirable to vary this threshold with height.

### 4.2. Analysis

The characteristics of the modeled cloud climatology were assessed by calculating the frequency of cloud detection at each model level for both the model and measured data over the seasonal time periods of winter, defined as January through March, and spring, or April through June. In Figure 7a, we see that the Eta model predicts a reasonable distribution of low and middle clouds during the winter months. However, there is an overprediction of cirrus events. This trend is exaggerated if a lower liquid water content, such as 0.005 $g m^{-3}$, is used as the cloud detection threshold for the model. In the spring, the measured cloudiness increases slightly while the predicted cloudiness decreases somewhat, with the net result that cirrus clouds are predicted nearly correctly while the frequency of low and middle clouds is underpredicted by the model. Applying the lower limit to the Eta output in this case would result in an overprediction of cirrus without improving the prediction of lower clouds significantly. It is interesting to note that while each of the measured distributions decreases while passing through the boundary layer toward the surface, the Eta model shows an increase in cloud occurrence just above the surface in Figure 7b.

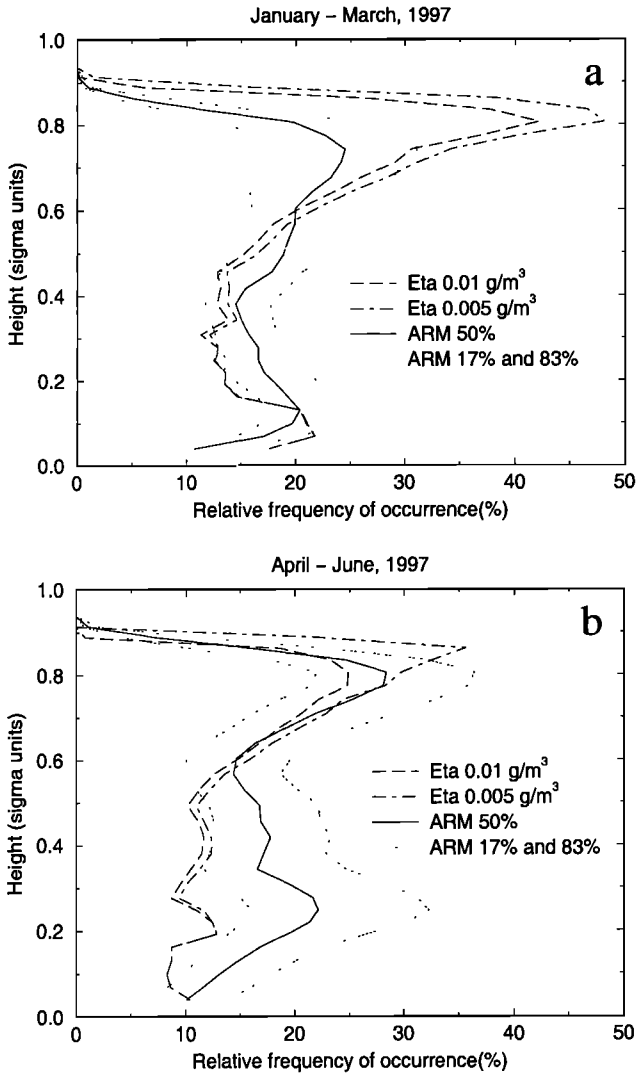


**Figure 7.** Seasonally averaged cloud profiles predicted by the Eta model and measured by the ARM SGP cloud radar.

In order to take a closer look at cloud prediction in the Eta model, the liquid water content predicted at each height for each forecast hour by the model was compared with the radar-derived cloud mask, which indicated whether a cloud was detected in each of these time-height intervals. A new cloud-prediction matrix was created by assigning each interval a code value based on the result of the comparison. Four states were possible: true positive if both the model and measurement indicated a cloud, false positive if only the model indicated a cloud, false negative if the model failed to predict cloud where it was detected by the radar, and true negative if neither the model nor the radar indicated cloud. Two examples of the resulting image maps are given in Plate 1.

Inspection of these cloud maps verifies the trends observed in the seasonal data and adds details concerning the timing of predicted clouds. In the winter months the model creates too many cirrus clouds over the SGP site. In addition, low and midlevel clouds are occasionally missed or exaggerated. During these months, however, the overall timing of cloud occurrences is correct. The model generally does not miss or create complete cloud events. This suggests that the Eta model does a good job of generating and tracking synoptic-scale cloud systems. In the spring, as convection becomes important, this trend changes. In May and June, cloud events become short in duration, and, although the remains of convective clouds in the upper troposphere are usually represented in the Eta model output, the rising portions of the clouds in the middle levels are frequently overlooked. This is due to the fact that the convective parameterization used in the Eta model does not change the cloud water mixing ratio, although adjustments are made to the radiative transfer routine and the cloud fraction when convective precipitation is found to occur. Researchers who wish to employ cloud information from the Eta model should be aware of this subtlety.

The results of this point-by-point comparison of cloud fields in the model and radar data are summarized in Table 2. Here, for brevity, clouds have been classified into three groups. Low clouds are those occurring from the surface up to 850 mbar, high clouds are those above 500 mbar, and middle clouds are those occupying the region in between. As expected, the majority of the time-height grid boxes are cloud-free and accurately predicted as such by the model. High clouds tend to be overpredicted; underprediction of upper level clouds occurs only in June. Low and middle clouds, for the most part, are underpredicted, with some exceptions occurring during the winter months. In addition, the number of accurately predicted clouds is frequently lower than the total number of prediction errors. During the period covered by this study, high clouds are predicted most accurately during April, while low and middle cloud prediction have the greatest skill during February and January, respectively.

## 5. Conclusions

Measurements from the ARM southern Great Plains site were used to evaluate the accuracy of surface radiation budget and cloud predictions from the Eta forecast model. An average of 50 $W m^{-2}$ extra downward shortwave radiation was identified in the Eta model. It was estimated that approximately one half of this error could be eliminated by correcting the impact of water vapor and aerosols on the propagation of shortwave radiation. However, such a change would also produce increased direct radiative heating of the model atmosphere. Elimination of the other half of this error appears to require improving cloud prediction and the calculation of radiative transfer through clouds. The Eta downwelling IR flux was found to be somewhat smaller than expected. This may be due in part to underprediction of the total column water vapor by the model. Smaller discrepancies were found in the other

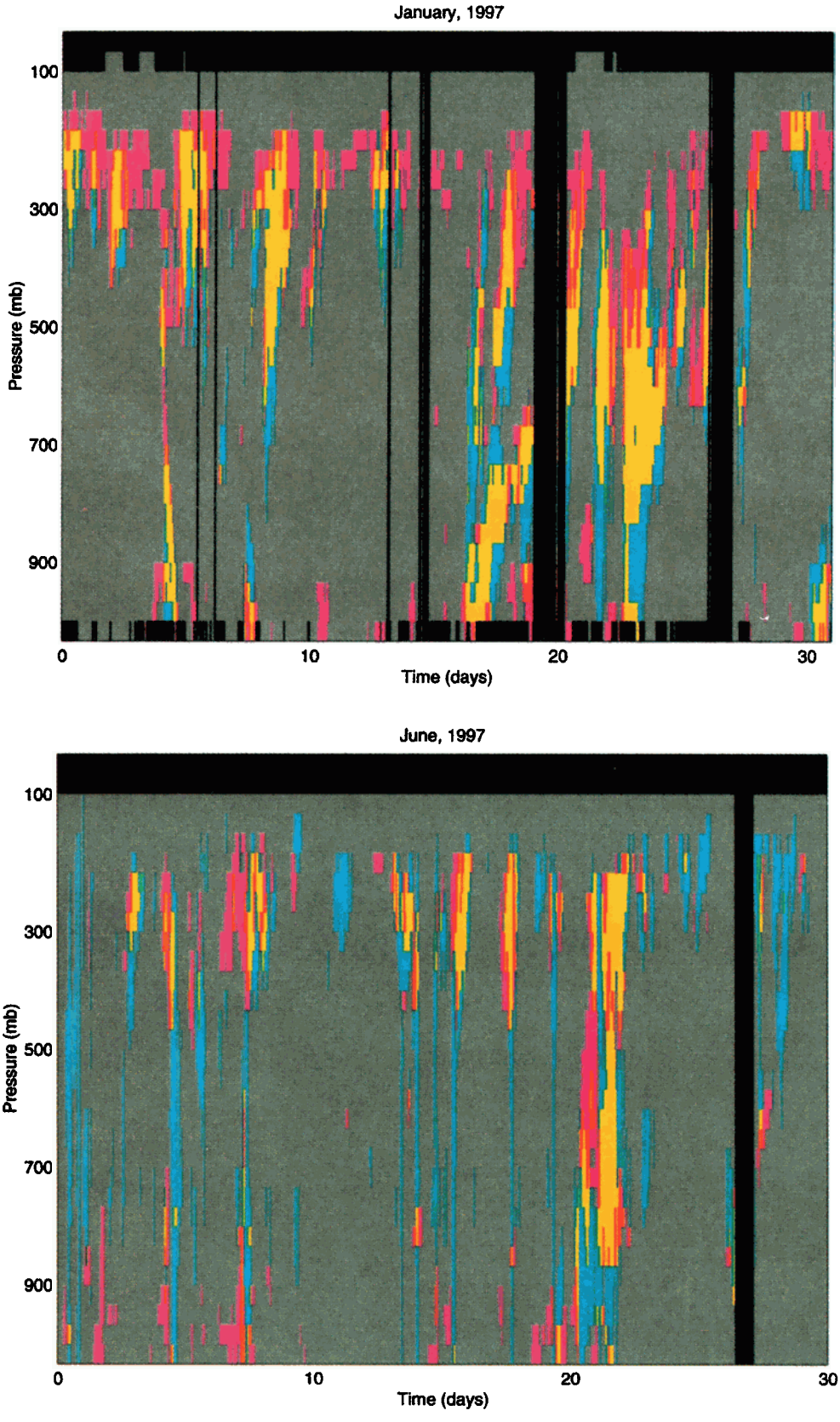


**Plate 1.** Representative comparisons of Eta model cloud predictions and ARM radar detections. Color scheme is as follows: grey, no cloud, both Eta and ARM; cyan, no cloud Eta, cloud ARM; yellow, cloud, both Eta and ARM; magenta, cloud Eta, no cloud ARM; black, no data.

**Table 2.** Statistics of Eta Cloud Prediction Relative to ARM Measurements

| | Number of Occurrences | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Surface to 850 mbar | | | | 850-500 mbar | | | | Above 500 mbar | | | |
| | True Neg. | False Neg. | False Pos. | True Pos. | True Neg. | False Neg. | False Pos. | True Pos. | True Neg. | False Neg. | False Pos. | True Pos. |
| Jan. | 4,141 | 414 | 278 | 352 | 4,756 | 474 | 229 | 697 | 5,701 | 380 | 1,281 | 846 |
| Feb. | 2,595 | 356 | 566 | 901 | 3,478 | 439 | 453 | 814 | 4,253 | 332 | 1,151 | 1,176 |
| March | 4,192 | 271 | 81 | 256 | 4,921 | 322 | 64 | 93 | 6,233 | 240 | 476 | 251 |
| April | 2,950 | 481 | 227 | 364 | 3,152 | 454 | 258 | 663 | 4,637 | 290 | 465 | 644 |
| May | 3,319 | 605 | 155 | 193 | 3,774 | 579 | 226 | 227 | 4,709 | 569 | 540 | 590 |
| June | 4,705 | 444 | 306 | 169 | 5,388 | 624 | 142 | 173 | 6,547 | 898 | 444 | 547 |
| Total | 21,902 | 2,571 | 1,613 | 2,235 | 25,469 | 2,892 | 1,372 | 2,667 | 32,080 | 2,709 | 4,357 | 4,054 |

energy flux terms. Despite this, the overall energy budget was nearly balanced and the surface temperatures were largely correct. The balance was achieved by underprediction of the downwelling infrared flux, the use of somewhat high albedos, and small offsets in the latent and sensible heat fluxes. This demonstrates that a balanced energy budget does not imply that every energy flux term is correct, so caution should be used in working with individual energy budget components from models.

Cloud prediction by the Eta model was mixed. Synoptic-scale systems were reproduced well by the model, while convective towers were frequently missed. Undertaking this analysis highlighted the difficulties inherent in working with cloud data. Since neither the measurements nor the model output explicitly specified the presence or absence of clouds, there was a dilemma regarding how clouds could be defined meaningfully and consistently in terms of the measurement and model variables. For the model, we ultimately chose a constant limit of 0.01 $g\,m^{-3}$, which is somewhat low for stratus clouds but high for cirrus. Nevertheless, excessive amounts of cirrus cloud were predicted by the model, especially in the winter. At the same time, there was always some uncertainty in exactly what was being detected by a given radar return. Another problem stemmed from the methods used by the prognostic cloud scheme. Most models that attempt to track liquid water through the atmosphere necessarily produce some residual water, so that extra condensate is present throughout the system. This not only adds to the difficulty of defining the occurrence of a cloud but may also affect the emission and transmission of infrared radiation through the atmosphere. More effort is required to improve cloud modeling. The ability to obtain routine cloud measurements that specify liquid water content and hydrometeor type would also be a dramatic improvement.

This study shows that model time series, such as those being archived as a part of the GCIP program, and measurements from a single point can be used together effectively to evaluate the accuracy of model output intended for use in climate studies. In particular, when extensive measurements, such as those from the ARM SGP site, are available, the treatment of specific processes and variables by the model can be examined in detail, as was done here for the shortwave downwelling flux. Nevertheless, this work, like that of *Mace et al.* [1998], *Yucel et al.* [1998], and *Berbery et al.* [this issue], has only begun to exploit the value of this method of inquiry. While our analysis was limited to one forecast model, one measurement point, and a 6-month time frame, measurements are available for many other dates and locations. This technique could easily be applied to other regional models and global climate models as well.

## Appendix: Radar/Lidar Cloud Mask Algorithm

The fundamental variable measured by the ARM millimeter cloud radar (MMCR) is the Doppler power spectrum, from which the vertical profile of reflectivity (or power) in dBZ, used in this study is routinely derived. The MMCR is able to sense reflectivities between −50 and +20 dBZ. This is sufficient to detect most clouds, with the possible exception of those that fail to fill the radar sample volume or that are composed of extremely small drops and are far from the radar. The signal received from a radar sample volume is considered significant when the returned power is large enough to be distinguished from the noise produced by the radar receiver. However, the source of the significant backscattered power may be precipitation or atmospheric particulates, such as dust or insects, as well as clouds.

The radar-based cloud masks used in this paper were constructed from a cloud database that is being created for analysis of clouds and radiation in climate and weather prediction models [*Marchand et al.*, 1998]. Clutter-reduced gridded cloud data with a 10-s resolu-

tion were derived from the measured radar data by applying an adaptation of the cloud detection algorithm developed by *Clothiaux et al.* [1999a] to produce a mask of significant detections for each of the radar's four operational modes. (See *Moran et al.* [1998] and *Clothiaux et al.* [1999b] for further information about these modes of operation.) Clutter reduction was performed for each radar mode using radar profiles from neighboring clear-sky periods as determined from lidar measurements. Any radar returns with reflectivities below those found in the neighboring clear-sky profiles were removed from the cloud mask, except for those that occurred above a laser-based cloud detection and were continuous in height. Boundary layer stratus clouds were most frequently affected by clutter at the ARM SGP site. During 1997, clutter contaminated approximately 70-80% of the boundary stratus cases (G. Mace, personal communication, 1998).

A set of criteria based on the signal-to-noise ratio, range resolution, target velocity, presence of pulse compression artifacts, and other factors was used to determine which mode produced the best data for each time and height. The radar reflectivity, Doppler velocity, and Doppler width data for this mode were then linearly interpolated onto a new time-height grid with 10-s by 45-m spacing. The 45-m vertical grid size was chosen because it corresponds to the spatial resolution of the radar mode with the highest range resolution. For the other three modes the range resolution is 90 m. Because of this difference in resolution and since the radar required 40 s to collect data in all four modes, the 10-s grid contains oversampled data.

Radar occurrence profiles with a temporal resolution of 5 min were obtained by examining the cloud mask on the 10-s grid. Radar height bins were flagged as cloud-occupied if, after clutter suppression, significant backscattered power was still observed during at least half of the 5-min period. A detailed analysis of cloud occurrence statistics (including cloud base distribution, cloud top distribution, cloud thickness, and cloud occurrence versus height) indicated that no significant difference exists between the full 10-s data and the data reduced to either a 5-min or 10-min grid.

**Acknowledgments.** The authors thank the following NOAA NCEP staff members for their efforts in providing guidance, technical information, model output, and samples of model algorithms to support this investigation: Kenneth Mitchell, Ying Lin, Qingun Zhao, Thomas Black, Timothy Marchok, and Eric Rogers. The advice and assistance of Gerald Mace of the University of Utah as well as Eugene Clothiaux, James Mather, George Young, Charles Pavloski, and Manajit Sengupta of The Pennsylvania State University are also appreciated. Laura Hinkelman additionally acknowledges the hospitality and support provided by the Institute for Atmospheric Measurements at the GKSS Research Center, Geesthacht, Germany, during the revision of this manuscript. Measured data were obtained from the Atmospheric Radiation Measurement (ARM) Program sponsored by the U.S. Department of Energy, Office of Energy Research, Office of Health and Environmental Research, Environmental Sciences Division. This material is based upon work supported under a National Science Foundation graduate fellowship. Additional funding was provided by the Department of Energy through grant DE-FG02-90ER-61071 and the National Atmospheric and Space Administration through prime contract NAS7-918 Jet Propulsion Laboratory subcontract 959034.

## References

Berbery, E. H., E. M. Rasmusson, and K. E. Mitchell, Studies of North American continental-scale hydrology using Eta model forecast products, *J. Geophys. Res., 101*(D3), 7305-7319, 1996.

Berbery, E. H., and E. M. Rasmusson, Mississippi moisture budgets on regional scales, *Mon. Weather Rev.*, in press, 1999.

Berbery, E. H., K. E. Mitchell, S. Benjamin, T. Smirnova, H. Ritchie, R. Hogue, and E. Radeva, Assessment of land surface energy budgets from regional and global models, *J. Geophys. Res.*, this issue.

Betts, A. K., F. Chen, K. E. Mitchell, and Z. I. Janjic, Assessment of the land surface and boundary layer models in two operational versions of the NCEP Eta model using FIFE data, *Mon. Weather Rev., 125*(11), 2896-2916, 1997.

Black, T. L., The new NMC mesoscale Eta model: Description and forecast examples, *Weather Forecasting, 9*, 265-278, 1994.

Chen, F., K. Mitchell, J. Schaake, Y. Xue, H.-L. Pan, V. Koren, Q. Y. Duan, M. Ek, and A. Betts, Modeling of land surface evaporation by four schemes and comparison with FIFE observations, *J. Geophys. Res., 101* (D3), 7251-7268, 1996.

Chen, F., Z. Janjic, and K. Mitchell, Impact of atmospheric surface-layer parameterizations in the new land-surface scheme of the NCEP mesoscale Eta model, *Boundary-Layer Meteorol., 85*, 391-421, 1997.

Clothiaux, E. E., T. P. Ackerman, G. G. Mace, K. P. Moran, R. T. Marchand, M. A. Miller, and B. E. Martner, Objective determination of cloud heights and radar reflectivities using a combination of active remote sensors at the ARM CART sites, *J. Appl. Meteor.*, in press, 1999a.

Clothiaux, E. E., K. P. Moran, B. E. Martner, T. P. Ackerman, G. G. Mace, T. Uttal, J. H. Mather, K. B. Widener, M. A. Miller, and D. J. Rodriguez, The Atmospheric Radiation Measurement program cloud radars: Operational modes, *J. Atmos. Oceanic Technol.*, in press, 1999b.

Fels, S. B., and M. D. Schwarzkopf, The simplified exchange approximation: A new method for radiative transfer calculations, *J. Atmos. Sci., 32*, 1475-1488, 1975.

Janjic, Z. I., The step-mountain eta coordinate model: Further developments of the convection, viscous sublayer, and turbulence closure schemes, *Mon. Weather Rev., 122*(5), 927-945, 1994.

Kato, S., T. P. Ackerman, J. H. Mather, and E. E. Clothiaux, The k distribution method and correlated-k approximation for a shortwave radiative transfer model, *J. Quant. Spectrosc. Radiat. Transfer, 62*, 109-121, 1999.

Lacis, A. A., and J. E. Hansen, A parameterization for the absorption of solar radiation in the earth's atmosphere, *J. Atmos. Sci., 31*, 118-133, 1974.

Leese, J. A. (Ed.), *Implementation Plan for the GEWEX Continental-Scale International Project (GCIP), Volume I, Data Collection and Operational Model Upgrade,* 83 pp., Int. GEWEX Proj. Off., Washington, D. C., 1993.

Mace, G. G., C. Jakob, and K. P. Moran, Validation of hydrometeor occurrence predicted by the ECMWF model

using millimeter wave radar data, *Geophys. Res. Lett.*, *25*(10), 1645-1648, 1998.

Marchand, R. T., E. E. Clothiaux, and T. P. Ackerman, A one-year cloud climatology for the Southern Great Plains site, paper presented at Cloud Physics Conference, Am. Meteorol. Soc., Everret, Wash., Aug. 17-21, 1998.

Marshall, C. H., Jr., Evaluation of the new land-surface and planetary boundary layer parameterization schemes in the NCEP mesoscale Eta model using Oklahoma Mesonet observations, M.S. thesis, Univ. of Okla., Norman, 1998.

Mather, J. H., T. P. Ackerman, W. E. Clements, F. J. Barnes, M. D. Ivey, L. D. Hatfield, and R. M. Reynolds, An atmospheric radiation and cloud station in the tropical western Pacific, *Bull. Am. Meteorol. Soc.*, *79*(4), 627-642, 1998.

Moran, K. P., B. E. Martner, M. J. Post, R. A. Kropfli, D. C. Welsh, and K. B. Widener, An unattended cloud-profiling radar for use in climate research, *Bull. Am. Meteorol. Soc.*, *79*(3), 443-455, 1998.

Ramaswamy, V., and S. M. Freidenreich, A study of broadband parameterizations of the solar radiative interactions with water vapor and water drops, *J. Geophys. Res.*, *97*(D11), 11,487-11,512, 1992.

Rogers, E., T. L. Black, D. G. Deaven, G. J. DiMego, Q. Zhao, M. Baldwin, N. W. Junker, and Y. Lin, Changes to the operational 'Early' Eta analysis/forecast system at the National Centers for Environmental Prediction, *Weather Forecasting, 11*, 391-413, 1996.

Stamnes, K., R. G. Ellingson, J. A. Curry, J. E. Walsh, and B. D. Zak, Review of science issues, deployment strategy, and status for the ARM North Slope of Alaska/adjacent Arctic Ocean climate research site, *J. Clim.*, *12*(1), 46-63, 1999.

Stokes, G. M., and S. E. Schwartz, The Atmospheric Radiation Measurement (ARM) program: Programmatic background and design of the cloud and radiation test bed, *Bull. Am. Meteorol. Soc.*, *75*(7), 1201-1221, 1994.

Toon, O. B., C. P. Mckay, T. P. Ackerman, and K. Santhanam, Rapid calculation of radiative heating rates and photodissociation rates in inhomogeneous multiple scattering atmospheres, *J. Geophys. Res.*, *94*(D13), 16,287-16,301, 1989.

Yucel, I., W. J. Shuttleworth, J. Washburne, and F. Chen, Evaluating NCEP Eta model derived data against observations, *Mon. Weather Rev.*, *126*(7), 1977-1991, 1998.

Zhao, Q., and F. H. Carr, A prognostic cloud scheme for operational NWP models, *Mon. Weather Rev.*, *125*, 1931-1953, 1997.

Zhao, Q., T. L. Black, and M. E. Baldwin, Implemenation of the cloud prediction scheme in the Eta model at NCEP, *Weather Forecasting, 12*, 697-712, 1997.

T. P. Ackerman, L. M. Hinkelman, and R. T. Marchand, Department of Meteorology, The Pennsylvania State University, University Park, PA 16802. (e-mail: ackerman@essc.psu.edu; laura@essc.psu.edu; roj@essc.psu.edu)

(Received September 1, 1998; revised December 22, 1998; accepted February 22, 1999.)