Validation of Cloud-Resolving Model Background Data for Cloud Data Assimilation

ROSANNE POLKINGHORNE, TOMISLAVA VUKICEVIC, AND K. FRANKLIN EVANS
Atmospheric and Oceanic Sciences, University of Colorado, Boulder, Colorado

(Manuscript received 18 March 2009, in final form 2 September 2009)

ABSTRACT

Results from a cloud-resolving model are systematically compared with a variety of observations, both ground based and satellite, in order to better understand the mean background errors and their correlations. This is a step in the direction of developing a background error covariance matrix for use in cloud data assimilation. Observation sources include the Geostationary Operational Environmental Satellite (GOES), the Atmospheric Emitted Radiance Interferometer (AERI), a microwave radiometer (MWR), radiosonde, and cloud radar. When exploring model biases in temperature, precipitable water vapor, and liquid water path, a warm and moist bias at night and a cool and dry bias during the day are observed. Values for the background decorrelation length of water variables are determined. In addition, a dynamic cloud mask is presented to give more control in the assimilation of cloudy satellite radiances, allowing different cloud types to be excluded from the assimilation as well as establishing values for the maximum residuals to be considered.

1. Introduction

Understanding clouds and their role in the atmospheric system is important on a wide spectrum of scales from regional weather to global climate. Quantitative observations of clouds are typically obtained by indirect remote sensing methods (Kidder and Vonder Haar 1995). Although considerable progress has been made in remotely sensing and retrieving bulk cloud properties, complex 3D cloud structure and its connection to thermodynamic fields and atmospheric motions on small spatial scales is not well specified from observations alone (Stephens and Kummerow 2007). Cloud-resolving models (CRMs) are used to study the spatial and temporal variability of clouds and their environment (Khairoutdinov and Randall 2003).

The cloud fields produced by CRMs cannot be easily compared with observed cloud fields because the initial and boundary conditions are not available on cloud-resolving scales. To improve the modeled representation, observations should be combined with the model results such that the observations optimally constrain the model solution in an approach known as data assimilation (Kalnay 2003; Lewis et al. 2006; Evensen 2006). Studies by Vukicevic et al. (2004, 2006) demonstrate the potential of such an approach by the assimilation of the Geostationary Operational Environmental Satellite (GOES) Imager Infrared (IR) observations into a mesoscale atmospheric model with parameterized cloud microphysics. The results from these studies show that skilled three-dimensional cloud fields consistent with the dynamical atmospheric environment can be obtained by satellite data assimilation, but indicate that the skill is sensitive to errors in the background modeled state, among other factors.

The Vukicevic et al. (2004, 2006) studies used a four-dimensional variational data assimilation (4DVAR) least squares data assimilation technique. This technique as well as other optimal data assimilation techniques, such as the three-dimensional variational data assimilation (3DVAR) method and the ensemble Kalman filter (EnKF; Kalnay 2003; Evensen 2006) require data of background atmospheric fields and their associated error statistics. The knowledge of the background error statistics is critical for optimality of the data assimilation results. The background error statistics for the CRMs are poorly known because of the lack of observation data available at the modeled resolution for systematic comparison to CRM simulations of 4D cloud fields. In the studies by Vukicevic et al. (2004, 2006) the background error statistics were specified using error estimates from...
prior experience with 4DVAR data assimilation without cloud analysis (Zupanski et al. 2002, 2005). Since the error statistics for the cloud prognostic variables were not available from the prior 4DVAR studies the variance errors were assigned values based on an expected relative error amplitude such as 50%–100% value of the variable per vertical layer in the model. The background error covariance matrices were modeled using a spatial unimodal correlation method, described in Zupanski et al. (2005). Although the cloud data assimilation results with the GOES IR in Vukicevic et al. (2004, 2006) are encouraging, they only provide the proof-of-concept of the data assimilation approach. This approach must be improved further, especially regarding the background error statistics, in order to provide reliable cloud analysis results for a variety of cases and applications.

The background error mean and covariance are two statistics required for finding the optimal data assimilation solution with the current optimal data assimilation techniques. The error covariance matrix defines error variances and error correlations of the control variables. The background error covariance matrix in cloudy data assimilation is not easily specified because of the large heterogeneity of cloud properties in space and time, as well as the anisotropic and flow-dependent nature of their error spatial correlations. In addition, these cloud properties have unknown correlations with other variables. It is also very likely that the distribution of the background error statistics for such fields is far from being Gaussian (Lopez 2007). However, only a few studies have been carried out on the definition of the error covariances of rain and clouds (e.g., Moreau et al. 2003; Amerault and Zou 2006; Sun and Zhang 2008).

In practice, there are three widely used methods for defining the background error covariance matrix: the National Meteorological Center (NMC) method, analysis of innovations, and the EnKF. The NMC method uses the differences between pairs of forecasts valid at the same time but with different lead times to represent the background errors (Parrish and Derber 1992). For well-observed quantities, a spatial average of the observed-background departures can be used to describe the background statistics, as in Hollingsworth and Lonnberg (1986). An ensemble of forecasts is generated by randomly perturbing background states and a background error covariance matrix is estimated from the ensemble spread of forecasts valid at a specified time (i.e., Beuhrer 2005). For a more complete description of these methods and their limitations, see Bannister (2008).

In this study, an analysis of differences between the CRM simulation results and observations is performed to investigate properties of the background errors and to provide a basis for the specification of background error statistics in cloud data assimilation. This method is related to the Hollingsworth and Lonnberg (1986) method since the observed quantities are highly correlated with cloud mixing ratio. The NMC method is not used because lagged short-range forecasts of clouds would be dominated by large phase errors. The ensemble approach is not used because not only is it computationally expensive, it is not known how to perturb the ensemble for microphysical variables as their analysis errors are not known.

A 4DVAR data assimilation system, designated as the Regional Modeling and Data Assimilation System (RAMDAS), is used rather than an EnKF algorithm because it depends less on the estimate of an initial error covariance matrix when there are several temporally distributed observations. This system has been used in studies by Vukicevic et al. (2004, 2006), Zupanski et al. (2005), and Greenwald et al. (2004). In this study we only utilize the nonlinear forecast model and observational operators from RAMDAS.

The error statistics are evaluated from the comparison of the model simulations with different types of observations from the Atmospheric Radiation Measurements (ARM) Southern Great Plains (SGP) region. The ARM-SGP region is rich in ground-based remote sensing observations, characterized by near-continuous temporal coverage. In addition, other observations such as satellite data, balloon-borne soundings, upper-air data, and data collected from aircraft are available routinely and from intensive observation periods (Stokes and Schwartz 1994; more information is available online at http://www.arm.gov/instruments).

In this study, several different types of observations are systematically compared to the forecast model results in order to better understand model biases and spatial correlation of background errors. These observations include visible reflectances and infrared radiances from satellite data, ground-based infrared radiances from the Atmospheric Emitted Radiance Interferometer (AERI), ground-based microwave radiances from the Microwave Radiometer (MWR), ground-based reflectivities from a cloud radar, as well as a ground-based radiosonde. A description of the forecast model, observation operators, and data is contained in section 2. Section 3 details the experimental design and results, and a discussion of the conclusions and future work to be done is found in section 4.

2. Physics and tools
   a. Forecast model

The forecast model used in this study is the Regional Atmospheric Modeling System (RAMS), version 4.1,
which is a nonhydrostatic, cloud-resolving research model developed at Colorado State University (Cotton et al. 2003). In this model, clouds and precipitation are explicitly predicted using a microphysics parameterization that features a one-moment scheme (mixing ratio) for cloud liquid water (Walko et al. 1995), and a two-moment scheme (mixing ratio and number concentration) for six other hydrometeor types, including pristine ice, aggregates, hail, rain, snow, and graupel (Meyers et al. 1997). Since the number concentration of cloud liquid water is not computed in the model, it is set to be $1.7 \times 10^8$ number kg$^{-1}$ (depending on air density, equivalent to about 187 cm$^{-3}$) where the cloud liquid water mixing ratio is greater than 0, as in Greenwald et al. (2002). The hydrometeor size distribution is described by a gamma distribution of prescribed width, and a bulk cloud microphysics scheme is employed. Longwave and shortwave radiative fluxes are parameterized using a two-stream model that allows radiative heating to influence the growth of water droplets and ice particle vapor deposition (Harrington et al. 2000; Wu et al. 2000). The model initial state and boundary conditions are interpolated to the model grid from FNL data (more information is available online at http://www.ncdc.noaa.gov).

b. Operators

The primary observation operator used in this study is a system for computing unpolarized radiative transfer for either collimated solar and/or thermal emission sources of radiation in both clear and cloudy plane-parallel conditions, known as the Spherical Harmonic Discrete Ordinate Method Plane Parallel for Data Assimilation (SHDOMPPDA; Evans 2007). This study employs the forward model part of SHDOMPPDA. SHDOMPPDA is similar to SHDOM (Evans 1998), the main difference between the two is that SHDOM is for 3D radiative transfer, while SHDOMPPDA is 1D. This model has been well tested and has been demonstrated to be accurate (Evans 2007).

Given the model fields of temperature, pressure, humidity, mixing ratios, and number concentrations of cloud liquid, cloud ice, and snow, optical properties such as optical depth, single-scattering albedo, and the Legendre series of the phase function can be calculated for each layer. Scattering properties of these hydrometeors are stored in precomputed tables. The scattering tables are generated using two methods: one method generates scattering tables using Mie theory for a gamma distribution of spherical particles, and the other generates these tables for gamma-size distributions of mixtures of six ice crystal shapes at wavelengths from 0.2 to 100 $\mu$m using a precomputed database (Yang et al. 2005) of optical properties for individual particle lengths from 2 to 9500 $\mu$m. The ice crystal shape mixture of Baum et al. (2005) is used.

Extinction by gases is calculated by spectrally integrating over molecular absorption lines with a $k$ distribution. In this method, the spectral absorption information for the layers in a base atmosphere are computed with the Line-By-Line Radiative Transfer Model (LBLRTM; Clough et al. 2005). The $k$-distribution absorption parameters are calculated with an exponential sum fit of the transmission over paths to space from all atmosphere levels and many viewing directions. There are $k$-distribution absorption coefficients for dry air, a quadratic water vapor absorption dependence, and a linear ozone absorption dependence (other trace gases are assumed fixed). These coefficients are tabulated for each pressure level in a base atmosphere and a range of temperatures around the base atmosphere. Given an input atmospheric profile, the $k$-distribution coefficients are interpolated in pressure and temperature and used to calculate the molecular absorption optical depth for each “$k$.”

SHDOMPPDA is used in the calculation of simulated radiances for the selected GOES Imager satellite channels. It is also used in a form modified to produce bottom-of-the-atmosphere rather than top-of-the-atmosphere radiances to simulate AERI brightness temperatures using four computational streams. The operator used to simulate brightness temperatures in the microwave part of the spectrum does the radiative transfer integration for a purely absorbing atmosphere. The radiosonde profile is simulated by averaging the model temperature and relative humidity over pressure surrounding each RAMS level. This operator also follows the sonde drift, in that it compares the RAMS grid point closest to the sonde as it ascends. The optical and scattering properties for each of the above-mentioned radiative transfer calculations are stored in precomputed tables. The closest RAMS column to each instrument is used.

c. Data

The forecast model output is compared to various different datasets in order to calculate mean background errors. Because of the high spatial and temporal resolution (4-km horizontal resolution, approximately every 30 min) of the GOES Imager datasets available in the ARM archive, channels 1 (0.63 $\mu$m), 2 (3.9 $\mu$m), 4 (10.7 $\mu$m), and 5 (12.0 $\mu$m) are used (Menzel and Purdom 1994). Data from the GOES-8 satellite are used in this study. In addition, ground-based data from four different instruments at five ARM sites (Fig. 1) are included in the analysis. These locations include the Central Facility near Lamont, Oklahoma (C1), and four boundary
facilities: Hillsboro, Kansas (B1), Vici, Oklahoma (B4), Morris, Oklahoma (B5), and Purcell, Oklahoma (B6).

At each of these five locations, infrared radiances from four microwindows of the AERI (Knuteson et al. 2004) are calculated. The wavenumber ranges of the four microwindows, 830.0–834.5 cm$^{-1}$, 898.5–904.7 cm$^{-1}$, 1095.0–1098.2 cm$^{-1}$, and 1231.3–1232.2 cm$^{-1}$ are chosen from those in Turner et al. (2003). These zenith-viewing infrared radiances are mostly sensitive to cloud temperature and optical depth (for ice clouds).

Radiances from the MWR (Morris 2005) at each of the 5 ARM sites considered here are used at 23.8 and 31.4 GHz. This gives information about liquid water and water vapor along a line-of-sight path. The observational operator compares model-simulated and MWR-measured brightness temperatures (the temperature a blackbody must have in order to duplicate the observed intensity of radiation at the specified wavelength), but it is also useful to relate the brightness temperature differences to precipitable water vapor (PWV) and liquid water path (LWP) using a simple linear regression. We use the method of Liljegren (1994) in which the optical depth ($t$) is first computed at each channel using Eq. (1):

$$ t = \ln \left( \frac{T_{\text{MR}} - T_c}{T_{\text{MR}} - T_B} \right), $$

(1)
where $T_{MR}$ is the atmospheric mean radiating temperature and $T_c$ is the cosmic radiating temperature equal to 2.75 K. Here $T_{MR}$ is calculated from mean climatological values. The PWV and LWP are then calculated using Eqs. (2) and (3):

$$\text{PWV} = a_0 + a_1 t + a_2 t^2$$

$$\text{LWP} = b_0 + b_1 t + b_2 t^2$$

The retrieval coefficients for these two equations have been computed for each calendar month by E. R. Westwater and M. J. Falls using 10 yr of NWS radiosonde data from Oklahoma City, Oklahoma (Liljegren 1994).

Zenith-pointing cloud radar reflectivities at 35 GHz obtained from the Millimeter Cloud Radar (MMCR; Moran et al. 1998) are used at the central facility to determine cloud boundaries. The Actively Remotely-Sensed Cloud Locations Process (ARSCL) product (Clothiaux et al. 2000) is used. Radiosonde data, available from launches typically twice a day at the central facility, provide temperature and humidity profiles.

3. Experimental design

The RAMS forecast model is run forward in time for 12-h periods, starting at 0000 and 1200 UTC for a total of 14 days or 28 model runs. This is done for assurance that the model state has not degraded with the large accumulation of model errors seen in longer runs because of the small domain used in this study. The model runs start at 0000 UTC 21 March 2000 and end at 0000 UTC 4 April 2000. The simulation is centered on the ARM central facility in OK and is run on a 125 x 125 x 84 grid with a horizontal grid spacing of 4 km, using a 15-s time step. The vertical grid is stretched, extending up to about 16 km, with a grid spacing of around 70 m in the planetary boundary layer.

The model state is compared to the observations during the last 5 h of each 12-h run. This is to ensure that the model has spun up sufficiently to make valid comparisons. The comparisons are done approximately every half hour, when the GOES data is available, making a total of 280 comparison times. The model data is compared to the GOES satellite data at each point over the whole domain, and to the ground-based ARM data (AERI, MWR, radar, and radiosonde) at their location points when available. These ground-based data are averaged over 400 s around the comparison time to match the RAMS grid cell advection.

To see any model biases specific to cloud type, each observation point is classified into one of three groups: clear, low cloud, or high cloud. Low cloud is defined as cloud with a cloud-top height below 5 km. High cloud is defined as cloud with a cloud-top height above 5 km.

The presence of cloud and its type is determined by a simple cloud mask using the GOES satellite radiiances and reflectances. A schematic of the thresholds used in determining the presence of low and high cloud is shown in Fig. 2. The flowchart shown is used when visible reflectances are available during the day. At night, this step is skipped, meaning that if $260 < T_b < 273$ K, low cloud is identified; otherwise it is classified as clear. Thin cirrus is identified by the difference between channels 4 and 5 as in Ackerman et al. (1998). This difference is greater than zero when ice clouds are present due to the greater absorption at the longer wavelength in the infrared atmospheric window.

This cloud mask is validated against cloud radar data when such data are available using the simple definitions of high and low cloud presented previously. The thresholds presented here are tuned to give the maximum accuracy possible. Statistics are shown in Fig. 3. The validation tests show that while this cloud mask is not completely accurate, it is correct the majority of the time as seen when the data is grouped according to the mask classification (Fig. 3a) and when it is grouped according to the radar classification (Fig. 3b). This validation is also done with model data and simulated cloud radar, giving similar results that are not shown here.
Once the cloud type has been determined in the observations, several different analyses are done. The number of cloudy points in the model as determined by the cloud mask described previously is compared to the number of cloudy points in the observations at each comparison time. Average residuals are examined for each of the ground-based data according to the observed cloud type. In addition, the spatial correlation of residuals is examined. The residuals are defined as in Eq. (4): 

\[ R = \frac{(X - Y)}{\sigma}, \]  

where \( X \) is the model brightness temperature, \( Y \) is the observed brightness temperature, and \( \sigma \) is the error associated with the specified channel or instrument. The error for each channel and/or instrument is listed in Table 1.

4. Analysis

a. Model versus observed cloudiness

The cloud mask described above is used to designate each grid point as clear, low, or high cloud in the observations and then repeated using the model-produced data. The fraction of columns with each classification combination is computed and presented in Table 2. The model’s agreement with the observations is very poor, except in the case when the observations are clear. According to this data, the model has far more clear points than the observations and far fewer points with high clouds.

The fraction representing the number of points in each category out of the entire grid is then computed at each observation time for both model and observations. The difference between the model fraction and the observed fraction is plotted as a function of observation time in Fig. 4. This analysis confirms the results of the previous analysis—the model clear fraction is much higher than observed, and the high cloud fraction is much lower than observed. There are a significant number of periods in which the high cloud underestimate in Fig. 4 seems to be approximately a mirror image of the clear overestimate, indicating that the model is largely missing the high clouds altogether in those periods. The analysis of the model error relative to the AERI measurements presented in the next section suggests that for the most part, the high cloud fraction errors in the model are caused by a small amount of this cloud type in the

<table>
<thead>
<tr>
<th>Instrument and channel</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>GOES channel 1 (0.63 ( \mu )m)</td>
<td>0.1 K</td>
</tr>
<tr>
<td>GOES channel 2 (3.9 ( \mu )m)</td>
<td>1.0 K</td>
</tr>
<tr>
<td>GOES channel 4 (10.7 ( \mu )m)</td>
<td>1.0 K</td>
</tr>
<tr>
<td>GOES channel 5 (12.0 ( \mu )m)</td>
<td>1.0 K</td>
</tr>
<tr>
<td>AERI 0832 cm(^{-1})</td>
<td>1.0 K</td>
</tr>
<tr>
<td>AERI 0902 cm(^{-1})</td>
<td>1.0 K</td>
</tr>
<tr>
<td>AERI 1097 cm(^{-1})</td>
<td>1.0 K</td>
</tr>
<tr>
<td>AERI 1232 cm(^{-1})</td>
<td>1.0 K</td>
</tr>
<tr>
<td>MWR 31.4 GHz</td>
<td>0.5 K</td>
</tr>
<tr>
<td>MWR 23.8 GHz</td>
<td>0.5 K</td>
</tr>
<tr>
<td>Sonde—Temperature</td>
<td>1.0 K</td>
</tr>
<tr>
<td>Sonde—RH</td>
<td>5.0%</td>
</tr>
<tr>
<td>Radar</td>
<td>2.0 dBZ</td>
</tr>
</tbody>
</table>
model simulations. The root-mean-square differences between the model and observed fractions are presented in Table 3.

**b. Time average of ground-based observations**

The average AERI, MWR, and GOES residuals at each relevant channel are examined. The overall average residuals and root-mean-square differences for each channel are computed first. Then the information is broken down into groups including only daytime points and only nighttime points to look for biases that may exist only in the day or night. These results are shown in Figs. 5a,b. The average residuals in GOES channels 4 and 5 are quite different between the day and night, pointing to a ground temperature bias that varies with the diurnal cycle. The large average residuals seen in each AERI channel indicate a lack of cloud in the model as compared to the observations.

Each observation time and location are then classified according to cloud type using the previously described cloud mask. The average of the residuals and the root-mean-square differences are taken for each group. Each group is subdivided into daytime and nighttime groups as before. These results are shown in Figs. 6–8. The figures are grouped according to the observation classification. This paper will mainly discuss the errors where the model and observations have the same cloud type. The other cases are shown in the figures in order to demonstrate the effect of these scenarios in cloudy data assimilation.

In examining the average residuals of GOES channels 4 and 5 for the cases in which the model and observations are clear (Fig. 6a, red bars), it is apparent that the model ground temperature is about 10 K too warm during the night and about 7 K too cool during the day compared to the observed ground temperature. Additionally, the residuals of the AERI channels are much larger during the day than at night. This bias could be caused by the presence of a very weak diurnal cycle in the model. The difference between the AERI residuals in the 832 and the 902 cm$^{-1}$ channels gives some indication that the model precipitable water vapor is higher on average than is observed in these conditions.

The case where low clouds are present in both the model and observations (low-low) has small average residuals and root-mean-square difference (Figs. 7a,b, red bars). This indicates that when the locations of the model and observed low clouds coincide, the model clouds are similar to the observed clouds. The large residuals seen in the total averages (green bars) are a result of the fact that the model does not make low clouds in over 2/3 of the observed instances (see Table 2), pointing toward a lack of cloud liquid water in the model’s lower troposphere. This bias in the lower troposphere is further supported by the negative residuals of the MWR data at 31 GHz. The downwelling radiance

### Table 2. The fraction of RAMS columns with each classification combination.

<table>
<thead>
<tr>
<th></th>
<th>Clear</th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obs Clear</td>
<td>0.33</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>Obs Low</td>
<td>0.12</td>
<td>0.04</td>
<td>0.01</td>
</tr>
<tr>
<td>Obs High</td>
<td>0.31</td>
<td>0.10</td>
<td>0.07</td>
</tr>
</tbody>
</table>

**Fig. 4.** The difference between the model and observed cloud fraction for each observation time. The black line represents the clear fraction, the blue line the low-cloud fraction, and the red line the high-cloud fraction.
at this frequency is dominated by liquid water emission when liquid water clouds are present, so a negative residual signifies less liquid water in the model atmosphere than in the observed.

Analysis of the high-high case (high cloud in both the model and the observations) reveals a lack of cloud ice in the upper troposphere as well (Figs. 8a,b). The large positive GOES residuals in channels 4 and 5, as well as the large negative AERI residuals indicate that the model has a thinner high cloud than the observations. The difference between the AERI residuals in the 832 and 902 cm\(^{-1}\) channels gives a further indication that the model ice water path is lower than observed.

The MWR differences are also interpreted in terms of PWV and LWP. The brightness temperatures are used in a simple linear regression to calculate the PWV and LWP as described in section 2c. The average differences and root-mean-square differences in PWV and LWP are presented in Figs. 9a–d. According to this data, there is an average wet bias in PWV at night and dry bias during the day. In addition, when the observations are clear, the model typically has more PWV than observed. It appears that the model has too little LWP when compared to the observations in every case except when the observations are clear. The root-mean-square differences for both the PWV and LWP show that the model appears to perform closer to the observations at night than during the day.

Averaged over the absolute value of the residuals in each channel is then taken at each grid point for each cloud type. For example, if there are 10 observation times at a given grid point, and 5 of those times are classified as clear, the average of the absolute value of the residuals for the clear case is taken over those 5 times. This means that the number of values in each average is different for each grid point, although there are sufficient observation times overall to ensure a good statistical representation of each case (clear, low cloud, and high cloud) for each grid point.

The results of this test are shown in Figs. 10a–c. These figures show the average residuals for each channel and classification type. Looking at the clear case (Fig. 10a), there is a mean residual of about 10 K in the brightness temperatures of channels 4 and 5. This signifies that the ground temperature of the model is on average about 10 K warmer than is observed.

The spatial autocorrelation function of the residuals is computed for each observation time and averaged over the 280 observation times. The autocorrelation function described cloud mask criteria. The average over the absolute value of the residuals in each channel is then taken at each grid point for each cloud type. For example, if there are 10 observation times at a given grid point, and 5 of those times are classified as clear, the average of the absolute value of the residuals for the clear case is taken over those 5 times. This means that the number of values in each average is different for each grid point, although there are sufficient observation times overall to ensure a good statistical representation of each case (clear, low cloud, and high cloud) for each grid point.

The results of this test are shown in Figs. 10a–c. These figures show the average residuals for each channel and classification type. Looking at the clear case (Fig. 10a), there is a mean residual of about 10 K in the brightness temperatures of channels 4 and 5. This signifies that the ground temperature of the model is on average about 10 K warmer than is observed.

The spatial autocorrelation function of the residuals is computed for each observation time and averaged over the 280 observation times. The autocorrelation function

---

**Table 3. The root-mean-square difference between the model and observed cloud fractions.**

<table>
<thead>
<tr>
<th>Clear fraction</th>
<th>Low cloud fraction</th>
<th>High cloud fraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.4863</td>
<td>0.2121</td>
<td>0.4889</td>
</tr>
</tbody>
</table>

---

**FIG. 5.** (a) The average residuals for each instrument averaged over all cases and all times and (b) the root-mean-square differences for each instrument. The green bars represent all the cases combined, the blue bars are the nighttime cases, and the red bars are the daytime cases.
FIG. 6. The overall, clear night, and clear day (a) average residuals and (b) root-mean-square difference for each instrument. The green bars represent all the observed clear cases combined, the red bars the clear-clear case, the blue bars the clear-low case, and the yellow bars the clear-high case.
for each GOES channel is presented in Fig. 11. Since the radiative transfer model used in this study has been shown to be accurate (Evans 2007), it can be assumed that the correlations shown here are not influenced by systematic errors from this model. Assuming an exponential decay, the horizontal background decorrelation length should be about 100 km (103 km using the function for channel 4).
5. Summary and conclusions

This study is a first step toward addressing the need for a better understanding of the average background errors and their spatial correlations in cloudy data assimilation. To this end, a systematic comparison of model results to observations is made. The comparison is performed using 28 12-h CRM simulations within a 14-day period.
from 21 March 2000 through 3 April 2000 for a domain of 500 km by 500 km, centered over the ARM central facility. The error estimates obtained here are based on remote, indirect measurements of cloud properties. To locate the sources of these errors, in addition to comparing the model results to GOES Imager observations for several cases, they are also compared to a number of ground-based observations.

The comparison to the GOES imager observations from channels 1, 2, 4, and 5 (0.63, 3.9, 10.7, and 12.0 μm, respectively) in cloudy conditions is used to estimate spatially distributed errors of model cloud variables, the associated horizontal decorrelation lengths, and dependence of the mean error amplitude on the cloud type. These data are used to analyze differences in cloud fraction according to cloud type. In addition, the model deviations from the GOES imager infrared observations in clear conditions are used to estimate mean error in the model surface temperature, including diurnal variation of the error. Ground-based data from four different instruments at five ARM sites are used to further refine mean error estimates of bulk water variables including water vapor, cloud vertical boundaries, the dependence of the mean error on the cloud type, as well as the errors in the atmospheric profiles of temperature and humidity.

The comparison to AERI microwindow observations is used to look at errors in the cloud height and thickness in addition to giving some information about PWV. The model errors relative to the MWR are used to reveal the differences between the PWV and LWP in the model and observations. The radar data is used to look at errors in the cloud height and thickness. The comparison to sonde data reveals information about the profile of water vapor and temperature biases. The comparison to sonde data shows the dependence of the mean error on the cloud type, as well as the errors in the atmospheric profiles of temperature and humidity.

The error analysis is done according to the type of cloud present in the observations as determined by a simple cloud mask. This cloud mask is defined using threshold values of brightness temperature for GOES channel 4, refined by a threshold of the difference between GOES channels 4 and 5 for identification of thin cirrus conditions and a visible channel threshold during the day for identification of low clouds. A more sophisticated cloud mask could reveal more information.

There are two cloud classification combinations, the case where the observations are clear and the model has high cloud and the case where the observations have low cloud and the model has high cloud, which have less than 20 data points. Including more data points in these cases would yield a more accurate representation of the average model errors for these cases. Additionally, the horizontal decorrelation length computed in this study is based on residuals from the end of a single March over

**FIG. 9.** Differences in (a) PWV and (b) LWP for total (green bars), night (blue bars), and day (red bars). The labels "total," "clear," "low," and "high" refer to the classification of the observations. Root-mean-square differences with the same coloring coding for (c) PWV and (d) LWP.
Fig. 10. The average absolute value of the residuals for the (a) clear, (b) low cloud, and (c) high cloud cases for GOES channels 1, 2, and 4. Because of the similarity between the residuals of channels 4 and 5, the residuals for channel 5 are not shown here.
Oklahoma. It is likely that in the summer, when convective clouds are common, that this decorrelation length would be smaller.

The statistics in this study are computed for the most part in observation (2D) space, while a background error covariance matrix is defined in control variable space and as such contains the vertical dimension. In addition, this matrix must also include the statistics of the control variables and their horizontal, vertical, and intervariable correlations. The mean errors, root-mean-square differences, and horizontal decorrelation length found here in residual space are directly related to the errors in vertically averaged water variables because of the dependence of the GOES channels on cloud properties. These statistics could be used in conjunction with some method such as the NMC method (Parrish and Derber 1992) to get a full 3D matrix. Another option is to use the average satellite radiance residuals to perform cloud retrievals and form the 3D matrix of average errors and their correlations and cross correlations.

The main conclusions to be drawn from the analysis are as follows:

- The presence of large-scale features in the spatial correlation test indicate a need to have a large value of the horizontal background decorrelation length for water variables in the background error covariance matrix used in cloudy data assimilation. The spatial autocorrelation function reveals that the residuals are correlated out to approximately 100 km, which is significantly larger than the values previously used in Vukicevic et al. (2004, 2006).

- Analysis of the clear-sky data reveals a model bias in the ground temperature. The model ground temperature is about 10 K too warm at night and 7 K too cold during the day. This could be the result of a very weak diurnal temperature cycle in the model. This bias can be corrected by either improving the model’s representation of the ground temperature, or simply correcting it online during the data assimilation process.

- There are large systematic errors in the model background cloudy atmosphere caused by both errors in amplitude and location. These amplitude errors vary by cloud type and can be used as a basis of quality control. Cloud classification can be used in cloudy data assimilation to help reduce the errors in location by only allowing points with the same type of cloud in the model and observations to be considered in the data assimilation. Cases where the observations are clear and the model has some type of cloud or where the low cloud is observed and the model has high cloud can also be considered in the assimilation using this technique since it is relatively easy to remove the cloud from the model state based on the cloud-sensitive observations (Vukicevic et al. 2004). Care must be taken using this method to use the appropriate cloud-type-dependent error variance together with the appropriate error decorrelation length. This could help to prevent optimization problems caused by an undesirable dynamical response.

Overall, the analysis in this study demonstrates that there are significant systematic errors in the model background that require careful evaluation using different observation types. In a forthcoming paper, we describe how these conclusions impact the successful assimilation of cloud-affected data. Studies with other models and weather regimes would benefit from a similar analysis to that presented here.

Acknowledgments. This project was funded by NSF Project ATM 35071341. The authors also acknowledge the National Center for Atmospheric Research supercomputing facility, where the model was run. Data were obtained from the Atmospheric Radiation Measurement (ARM) Program sponsored by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research.

REFERENCES


