Using Synthetic Brightness Temperatures to Address Uncertainties in Cloud-Top-Height Verification

JASON E. NACHAMKIN AND YI JIN

Naval Research Laboratory, Monterey, California

LEWIS D. GRASSO

Cooperative Institute for Research in the Atmosphere, Fort Collins, Colorado

KIM RICHARDSON

Naval Research Laboratory, Monterey, California

(Manuscript received 30 June 2016, in final form 18 October 2016)

ABSTRACT

Cloud-top verification is inherently difficult because of large uncertainties in the estimates of observed cloud-top height. Misplacement of cloud top associated with transmittance through optically thin cirrus is one of the most common problems. Forward radiative models permit a direct comparison of predicted and observed radiance, but uncertainties in the vertical position of clouds remain. In this work, synthetic brightness temperatures are compared with forecast cloud-top heights so as to investigate potential errors and develop filters to remove optically thin ice clouds. Results from a statistical analysis reveal that up to 50% of the clouds with brightness temperatures as high as 280 K are actually optically thin cirrus. The filters successfully removed most of the thin ice clouds, allowing for the diagnosis of very specific errors. The results indicate a strong negative bias in midtropospheric cloud cover in the model, as well as a lack of land-based convective cumuliform clouds. The model also predicted an area of persistent stratus over the North Atlantic Ocean that was not apparent in the observations. In contrast, high cloud tops associated with deep convection were well simulated, as were mesoscale areas of enhanced trade cumulus coverage in the Sargasso Sea.

1. Introduction

Technological advances in autonomous aviation and precision optical systems have increased the demand for accurate cloud forecasts. As numerical models become more sophisticated, forecasters and machine-based algorithms are increasingly turning to modeled cloud output as guidance for the horizontal and vertical location of clouds. Satellite observations, with their broad coverage and high resolution, are often useful for verifying cloud forecasts (Bikos et al. 2012). Uncertainties in cloud heights derived from satellite complicate the comparison, however. Satellite cloud-top retrievals may contain errors of up to 7 km in regions of optically thin clouds (Smith et al. 1996; Frey et al. 1999). Many recent studies have used synthetic brightness temperatures generated from model output as a means to more consistently compare model radiance with observations (Morcrette 1991; Karlsson 1996; Rikus 1997; Chevallier and Kelly 2002; Tselioudis and Jakob 2002; Lopez et al. 2003; Sun and Rikus 2004; Söhne et al. 2008; Otkin et al. 2009). These comparisons reveal general trends in model cloudiness relative to climatological satellite data (Zhang et al. 2005) and are also useful for comparing physical parameterizations—in particular, microphysical schemes (Grasso et al. 2008, 2010, 2014; Grasso and Greenwald 2004; Chaboureau and Pinty 2006; Otkin and Greenwald 2008; Jankov et al. 2011; Jin et al. 2014; Thompson et al. 2016).

Although brightness temperatures are the most consistent way to verify the forecasts, the cloud-top

---

Denotes content that is immediately available upon publication as open access.

Corresponding author e-mail: Jason Nachamkin, jason.nachamkin@nrlmry.navy.mil

DOI: 10.1175/JAMC-D-16-0240.1

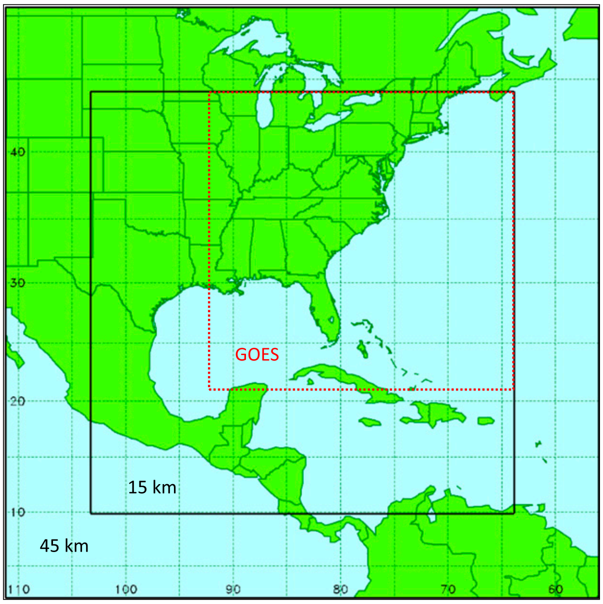

FIG. 1. The 45- and 15-km COAMPS forecast domains are depicted as labeled. The GOES verification domain is denoted by the red dotted box.

heights are still important to forecasters. Brightness temperature and cloud-top height fortunately tend to be well correlated for thicker clouds. Because cloud-top height can be determined directly from the forecast model, the synthetic brightness temperatures offer an opportunity to investigate the effects of conditional sampling on the uncertainty of the radiance-based cloud tops. Similar verification efforts by Zhang et al. (2005) and Otkin and Greenwald (2008) have successfully used cloud optical thickness in combination with brightness temperature to sort clouds into basic types. Although cloud-top heights were not directly verified, both studies were able to infer general trends such as the underrepresentation of midlevel clouds in the numerical forecasts.

This work is a continuation of the above efforts to determine the accuracy of conditionally sampled model cloud-top forecasts. A set of threshold criteria will be applied to forecast synthetic brightness temperatures generated by the Coupled Ocean–Atmosphere Mesoscale Prediction System (COAMPS[1]; Hodur 1997). Various combinations of total-condensed-water-path (TCP) and cloud-top-property thresholds will be applied in an effort to reduce cloud-top uncertainty while retaining the maximum number of viable points. These thresholds will then be applied to filter the observations from the *Geostationary Operational Environmental Satellite-13* (*GOES-13*) for comparison with the forecasts. Although the effects of the criteria on the observed retrievals are not directly known, simple errors associated with optically thin[2] ice clouds were found to dominate. The contributions of other errors associated with temperature inversions, overlapping cloud layers, or surface emissivity remain open questions. Despite these uncertainties, the resulting statistics reveal systematic model errors that stand out with greater clarity than would have been available from the unfiltered brightness temperatures, especially for low- and mid-level clouds that are often masked by misplaced upper-atmospheric thin clouds.

2. Forecast and observational data

COAMPS was used to generate a set of 48-h forecasts over eastern portions of the United States and western portions of the Atlantic Ocean for the period of 26 July–9 August 2015. The computational domain consisted of two one-way nested grids with horizontal spacings of 45 and 15 km, respectively (Fig. 1). The forecasts from the 15-km grid were used in the verification. The vertical domain consisted of 60 sigma-z levels extending from 10 m to a model top at approximately 30 km. Forecasts were initialized daily at 0000, 0600, 1200, and 1800 UTC, using the Naval Research Laboratory's Atmospheric Variational Data Assimilation System (NAVDAS; Daley and Barker 2001). The previous 6-h forecast acted as a first guess. Boundary conditions were supplied from the Navy Global Environmental Model (NAVGEM) at 3-h intervals using a Davies (1976) scheme. The explicit microphysics was parameterized using a modified version of the single-moment bulk scheme of Rutledge and Hobbs (1983, 1984) that is described by Chen et al. (2003). The mixing ratios of cloud droplets, cloud ice, rain, snow, and graupel were all predicted. Subgrid-scale convection was parameterized using the Kain–Fritsch scheme (Kain and Fritsch 1993). The Fu–Liou (Liu et al. 2009) parameterization was used for shortwave and longwave radiative transfer. Boundary layer turbulence was parameterized using a 1.5-order turbulence closure method (Mellor and Yamada 1982) in which turbulent kinetic energy is explicitly predicted. Although a shallow cloud scheme is available in COAMPS, it was not deployed for these simulations because the additional mixing removed almost all boundary layer clouds.

---

[1] COAMPS is a registered trademark of the Naval Research Laboratory.

[2] For brevity, from this point on the terms "thick/thin" will refer to "optically thick/thin" clouds.

Despite recent advances in computing power, relatively coarse grid spacing is still necessary to support the rapid production of forecast data covering large areas. Also, compromises such as the Kain–Fritsch scheme will strongly affect the cloud forecasts. In its current implementation in COAMPS, the Kain–Fritsch scheme feeds back convectively generated vapor and cloud water tendencies to the explicit microphysical scheme, which operates on the model grid. Because convection is only assumed to cover a small portion of each grid square, the tendencies are too weak to saturate the environment at most levels. As a result, convective clouds are underrepresented in the lower and middle troposphere. Near the tropopause, the tendencies become strong enough to produce extensive cirrus shields, which is reasonable considering that the scale of the anvil is often many times the size of the convection that produced it. The Kain–Fritsch scheme was primarily designed for turbulent closure and will not properly represent most convective clouds. Deep convective clouds have a horizontal length scale of approximately 5 km, whereas active/passive cumulus in the planetary boundary layer have even smaller horizontal length scales, on the order of 0.5 km. Even grids with horizontal spacings on the order of 1–2 km will not fully resolve these clouds. Because these forecasts are actively used by forecasters in the field, however, the model capabilities need to be understood.

The synthetic brightness temperatures were derived using the delta-Eddington two-stream (Deeter and Evans 1998) method described by Grasso et al. (2008) and Bikos et al. (2012). The COAMPS forecasts of temperature, pressure, relative humidity, surface temperature, and mixing ratio for the microphysical species were all used in the forward radiative transfer calculation. The microphysical size distributions and the single-scattering albedo calculation were adapted to match the assumptions in COAMPS. Synthetic brightness temperatures for the 10.7-μm infrared channel or GOES channel 4 (CH4) were derived at 6-h intervals for all forecasts on the 15-km grid through 48 h.

The observed data consisted of *GOES-13* CH4 brightness temperatures as well as retrievals of cloud water path, cloud-top height, and cloud type. Retrievals of cloud water path were obtained from the National Aeronautics and Space Administration Clouds and the Earth's Radiant Energy System (CERES), as described by Minnis et al. (2008). Daytime retrievals were primarily used for the quantitative comparisons in this work, although results from both day and night are displayed for comparison. Nachamkin et al. (2009) found that mean COAMPS TCP values were well correlated with daytime retrievals for values $\leq 200\,\mathrm{g\,m^{-2}}$. The satellite retrievals tended to underestimate TCP for thick, inhomogeneous clouds, especially at high zenith angles.[3] Cloud-top height was derived from the thermal spectra as described by Mitrescu et al. (2006). Observed brightness temperatures were matched with corresponding NAVGEM temperature fields to estimate cloud-top height. Cloud type was derived using the explicit physics algorithm outlined in Bankert et al. (2009). The observed satellite pixels were sorted into seven categories consisting of clear, partly cloudy, liquid water, supercooled water or mixed phase, glaciated (opaque ice), cirrus, and cirrus overlap. Validation of cloud-typing algorithms is difficult because of the lack of large cloud-type ground-truth datasets. Bankert et al. (2009) found that one of the most common errors was the misdiagnoses between the cirrus and overlap categories. The glaciated and mixed-phase clouds were also occasionally misdiagnosed. In this work, the high cloud types were combined into one category to aid in the identification of ice clouds (glaciated, cirrus, and cirrus overlap) that would likely contain greater cloud-top-height errors.

All of the satellite observations except cloud type were linearly averaged to the 15-km COAMPS grid from their native 4–8-km resolution. As a result of data-collection limitations, the satellite data only covered a portion of the COAMPS domain; thus, the verification domain consisted of the red dashed box in Fig. 1. For cloud type, the mode of the native pixels was used to represent the cloud type within the area centered at each COAMPS grid point. The differing interpolation methods resulted in differences in cloud coverage, especially near cloud edges or partly cloudy areas. As a result, unless otherwise noted, the cloud type was used as a diagnostic field to locate regions of ice clouds.

3. Conditional-sampling experiments

The mean coverage distributions of the brightness temperatures and cloud-top heights for all points with cloud cover are shown in Fig. 2. All of the distributions were normalized by the size of the verification domain. To emphasize better the diurnal nature of the cloud-top-height errors as well as to isolate the daytime retrievals, only those forecasts initialized at 0000 UTC were used in the calculation. The 1200 UTC forecast trends were very

---

[3] Nachamkin et al. (2009) contained three typographical errors in which the definition of the solar angle was reversed. The scientific results were not affected, but the typographical errors could lead to confusion. On page 3487, "low zenith angle" should be "high zenith angle" and "low solar zenith angles" should be "high solar zenith angles." On page 3488, "high zenith angle" should be "low zenith angle."

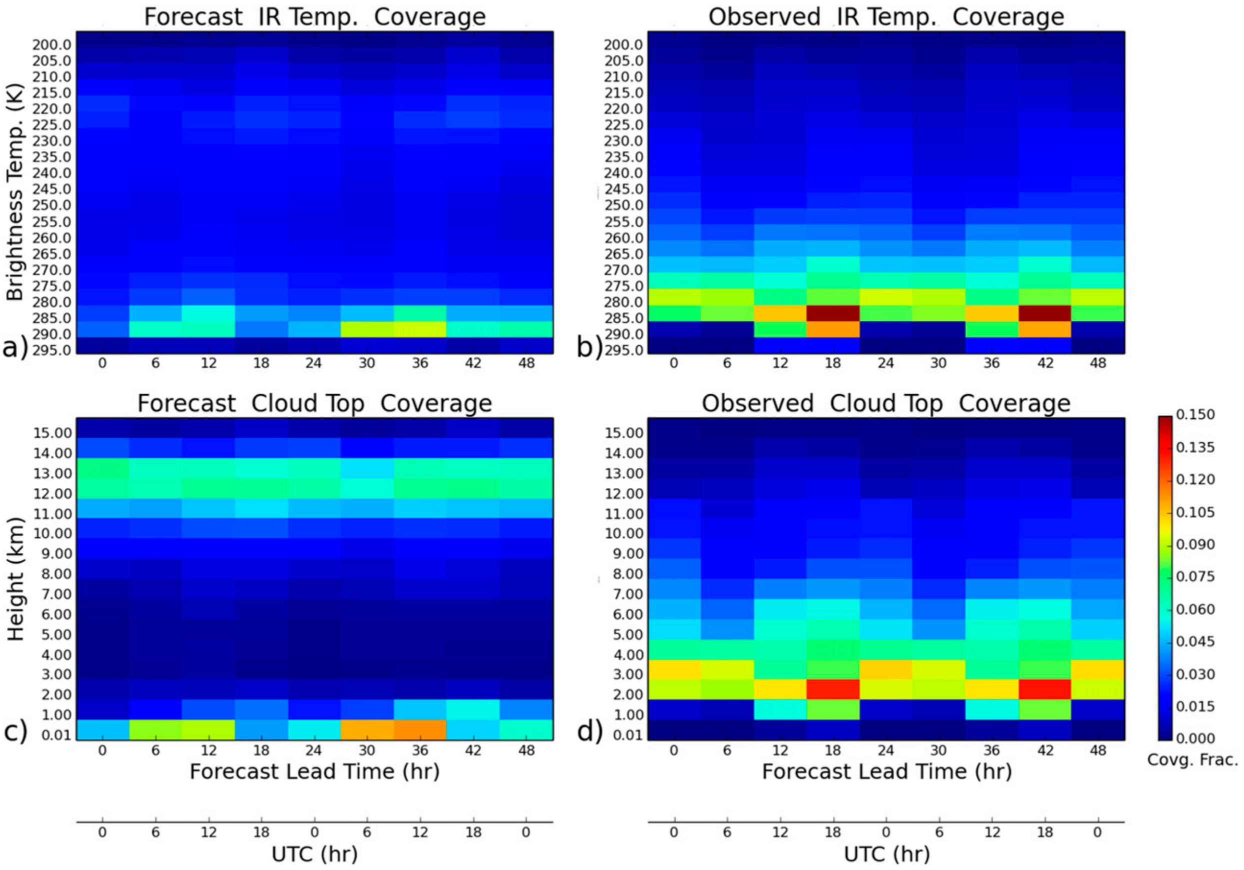

FIG. 2. The mean coverage distributions of the (top) brightness temperatures (K) and (bottom) cloud-top heights (km MSL) for all points with cloud cover within the verification domain at each forecast lead time for the (a),(c) COAMPS and (b),(d) GOES distributions. COAMPS cloud tops are defined as the highest level with total water content that exceeds 10^{-6} kg m^{-3}. The data were compiled from all model runs initialized at 0000 UTC for the 26 Jul–9 Aug verification period. The coverage fraction (shaded) in each bin is tabulated as the total number of occurrences normalized by the total number of cloudy and clear points within the verification domain.

similar to those at 0000 UTC. Because the observed cloud-top-height and TCP retrievals are generally less accurate at night, the discussion will focus on the daytime results (1200 and 1800 UTC). The strong similarity between the observed brightness temperatures and retrieved cloud-top heights reflects the dependence on the raw observations in the retrieval (Figs. 2b,d). The forecast cloud-top heights show much greater concentrations of high cloud tops between 11 and 14 km MSL[4] than are indicated by the synthetic brightness temperatures alone (Figs. 2a,c).

At this point a brief discussion about the definition of cloud-top height in the model is warranted. Cloud-top height is often used as a forecast product, but its definition is somewhat arbitrary. It can be defined by a mixing-ratio or total-water-content threshold that is exceeded from the top down at either a single level or as a top-down integrated quantity. Because the units of total water content are water mass per unit volume, that quantity would likely be superior from a user-oriented perspective. Many optical systems are sensitive to the particle number density along the line of sight. Mixing ratio also does not account for the change of air density with height. A constant mixing-ratio threshold results in a positive upper-tropospheric cloud bias because lower amounts of water per unit volume are required to exceed the threshold. For these reasons, mixing ratio probably should not be used to define cloud top or cloud base. The issue of defining the cloud top from an integrated versus a single-level value is also a matter of perspective. Satellite radiance is an integrated quantity, and setting the cloud top as the level (integrating down from the top of the model) at which the integrated total water content exceeds a threshold would be more consistent with the observations. Forecasters or pilots may be more concerned with clouds at a specific horizontal level, however. The differences between each cloud-top

---

[4] All cloud-top heights are defined to be above mean sea level.

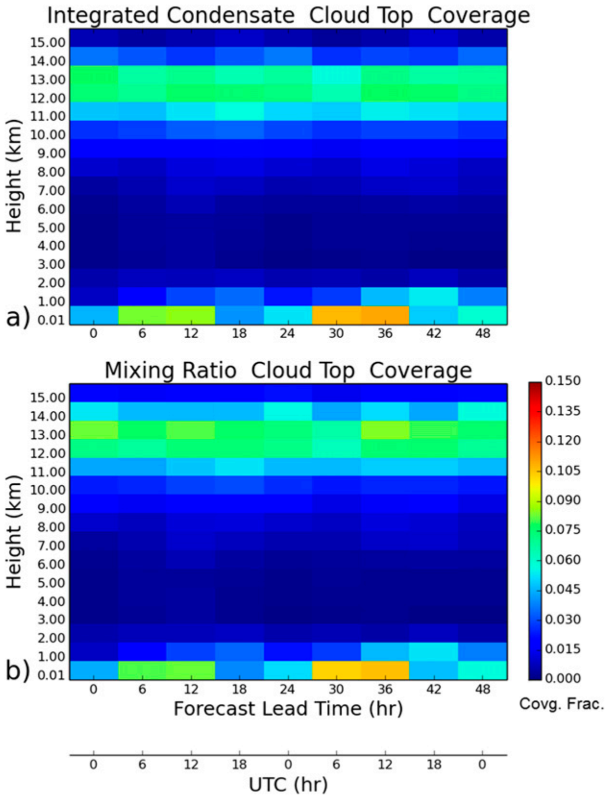

FIG. 3. Mean COAMPS coverage distributions for all clouds with tops defined by (a) integrated total water content that exceeds 10^{-6} kg m^{-2} and (b) single-level mixing ratio that exceeds 10^{-6} kg kg^{-1}. The data were compiled from all runs initialized at 0000 UTC. Coverage fractions are normalized as in Fig. 2.

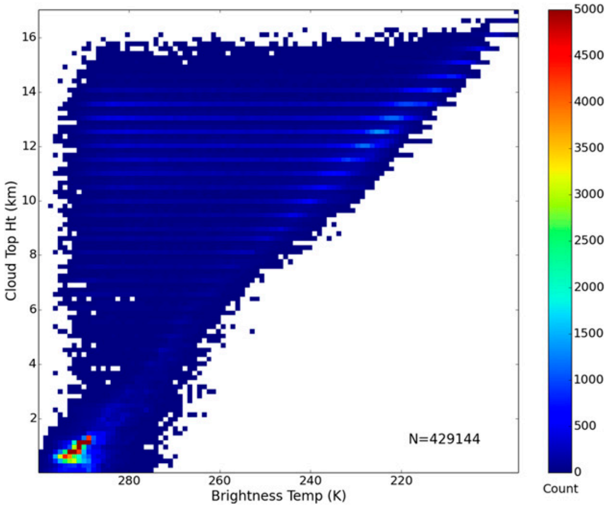

FIG. 4. Scatter/density plot of COAMPS synthetic brightness temperatures (K) and cloud-top heights (km MSL) at each cloudy point within the verification domain. The number of points in each ~0.167-km × 1-K bin is indicated by the color bar. The data were compiled from all 12-h forecasts from runs initialized at 0000 UTC. The total number of points is indicated in the bottom-right corner.

definition become apparent when comparing the cloud-top-height distributions in Fig. 3 with those in Fig. 2. In Fig. 2, forecast cloud-top height is defined as the highest level with a total water content exceeding 10^{-6} kg m^{-3}. In Fig. 3a cloud-top height is defined as the first level at which the top-down integrated total water exceeds 10^{-6} kg m^{-2}, and in Fig. 3b it is defined as the highest level that exceeds a mixing ratio of 10^{-6} kg kg^{-1}. The effect on the distributions is substantial, especially in the upper atmosphere. As expected, the mixing-ratio and integrated thresholds result in more upper-level cloudiness being diagnosed. The effect is less noticeable at lower levels, with relatively little change there. Overall, the mixing-ratio threshold results in about 5% more cloudy points than does the single-level total-water-content threshold. Situations with thin cloud layers overlapping low cloud decks may result in greater sensitivities in the distribution of low cloud tops than was noted in these results. Such cases of overlap were not readily apparent here. In the current work, the cloud definition was chosen to be the highest level with a total water content that exceeds 10^{-6} kg m^{-3} (Fig. 3) since that is most consistent with current user needs. An argument could be made for the integrated threshold (Fig. 3a), but the synthetic brightness temperatures mitigate the need for it in this study because they already account for the integrated nature of the observations. Note that the numerical value of the threshold is also arbitrary. Because 10^{-6} kg m^{-2} is widely used, that value was retained.

A snapshot scatter/density plot of the COAMPS cloud-top height versus synthetic brightness temperature for the 12-h forecasts valid at 1200 UTC is shown in Fig. 4. The horizontal lines of enhanced counts are an artifact of the discrete model sigma levels. The lines are more pronounced in the upper troposphere as a result of the greater spacing between levels as well as the increasingly horizontal nature of each sigma surface with height. The large spread at upper levels indicates poor correlations between the brightness temperatures and cloud-top height. Most of the off-diagonal values were skewed toward warm temperatures, with a nontrivial number of very large discrepancies in the upper atmosphere. At the extreme limit, brightness temperatures of 285 K were representing clouds located at 14 km, which is about 70 K above the air temperature at this level. This pattern was a continuous feature at all forecast lead times and is consistent with errors associated with optically thin cirrus. In the lower troposphere, high concentrations of warm, low-topped clouds displayed a more linear relationship between brightness temperatures and cloud-top height, indicating that these clouds

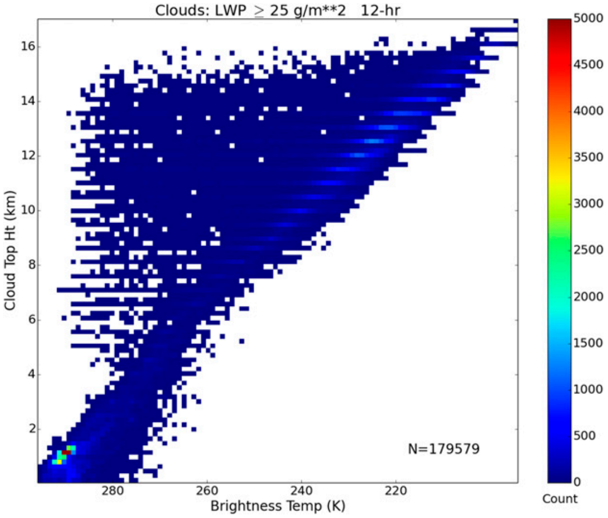

FIG. 5. As in Fig. 4, but all points with TCP values of less than 25 g m^{-2} have been removed.

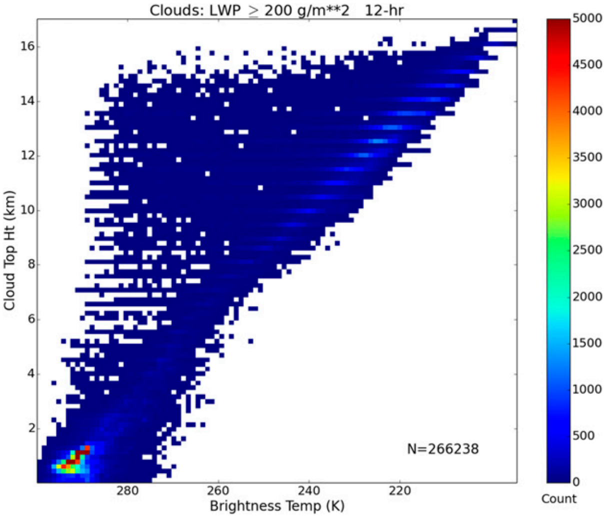

FIG. 6. As in Fig. 4, but all points with cloud tops consisting entirely of ice species and TCP values of less than 25 g m^{-2} have been removed.

were optically thick. A considerable spread was still apparent, however. Latitudinal and temporal variations in the boundary layer forecast temperature likely contributed.

In an effort to identify thicker clouds with more reliable cloud tops for verification, a TCP filter was applied to both the observed and forecast cloud data. Many of the poorly correlated forecast cloud-top heights and brightness temperatures were removed by excluding all points with corresponding TCP values of less than 25 g m^{-2} (Fig. 5). Vertical distributions of the synthetic brightness temperatures more closely matched the forecast cloud tops, especially at upper levels. A large percentage of highly correlated points below 4 km were also removed from the forecasts, however, and the remaining number of points in Fig. 5 is less than one-half that in Fig. 4. Other TCP thresholds that are lower than 25 g m^{-2} were also attempted. Those retained more low clouds but were less effective at removing the higher clouds. The observed cloud distributions (not shown) experienced a similar reduction in both upper- and lower-tropospheric cloud counts with the application of the TCP filter. Many of the lower (nonice) clouds that were removed were in regions of scattered cumulus or stratus, as based on the visual appearance of the GOES cloud field.

Microphysical information can be used to refine the effects of the TCP filter such that it acts primarily on ice clouds. Particle type is easily determined in the forecasts from the phase of the microphysical constituents at cloud top. For the observations, the necessary cloud-type (liquid or ice) information was obtained from the retrievals. Bankert et al. (2009) noted that the retrievals had more trouble sorting between ice cloud types than differentiating between ice and liquid. Mixed-phase clouds tend to be thicker, more convective, and less likely to be affected by the TCP filter if they are misidentified as ice. The discreet nature of the cloud-type information was problematic. Because the mode of the cloud type is used on the COAMPS grid, some clouds were likely misidentified. To mitigate these issues, both the TCP and cloud-type criteria were applied to the raw satellite data prior to interpolation to the COAMPS grid. Once the observed and forecast microphysical information were obtained, it was used to constrain the TCP filter to all cloud tops consisting entirely of ice. All other cloud tops were left unfiltered. Unless otherwise noted, this combined filter (removing all ice cloud tops with TCP < 25 g m^{-2}) was applied to both the observations and the forecasts for the remainder of this study.

The resulting forecast cloud-top distributions retained all of the low clouds while removing many of the uncertain high, thin ice clouds (Figs. 6–8). In terms of the total percentage of cloudiness retained, the 25 g m^{-2} criterion alone removed about 66% of the forecast cloud cover, but the cloud-type mask reduced that value to a range of 35%–50% depending on the forecast valid time (Fig. 9).[5] Many clouds were removed from the observations as well, including a large number with retrieved cloud-top heights in the lower atmosphere (Fig. 8). Examination of the visible and water vapor imagery

---

[5] Because of the limitations of the GOES nighttime observations, the diurnal variations were not investigated.

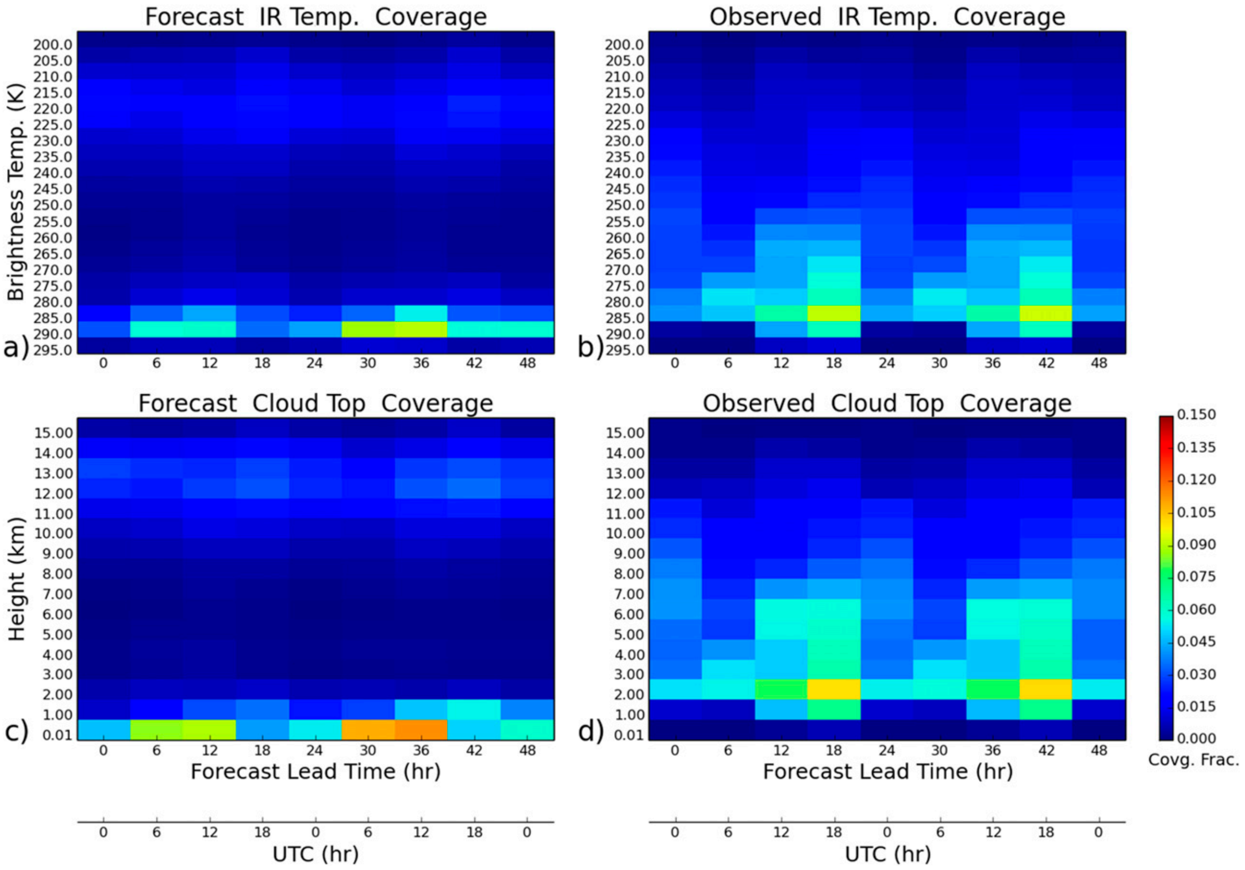

FIG. 7. As in Fig. 2, but all points with cloud tops consisting entirely of ice species and TCP values of less than 25 g m^{-2} have been removed. Coverage fraction is normalized by the same values that were used to normalize the fractions in Fig. 2.

revealed that extremely thin or scattered cirrus were being assigned cloud-top heights as low as 1.5 km. More observed clouds were retained overall, with the combined TCP and cloud-type filter removing between 10% and 27% of the total observed cloudy points (Fig. 9). The differences between the filtered and unfiltered fields suggest that thin cirrus can contribute to significant biases in the perceived cloud-top distribution. Thin cirrus accounted for about 50% of the observed cloudiness and from 30% to 60% of the forecast cloudiness in the layer with temperatures ranging from 280 to 285 K (Figs. 8 and 9). The difference in the vertical distributions was also more concentrated in the lower troposphere in the observations than in the forecasts (Fig. 8). One possible reason may be that the thickness of the model vertical levels places a lower limit on the cloud optical thickness for a given ice water content. Uncertainties in the observed TCP retrievals, as well as in the calculation of synthetic brightness temperature, also possibly contributed to these vertical variations. It is also likely that the vertical structure of the simulated cirrus clouds was very different from the observations.

The degree to which the correspondence between the forecast cloud-top height and the brightness temperature is increased by filtering is reflected in increased Pearson correlations (Fig. 10). The correlations for the original forecast fields range from 0.7 to 0.8, but they increase to 0.95 and above when the combined TCP and cloud-type filters are applied. The Pearson correlations for the TCP filter alone were slightly lower because of the removal of high-brightness-temperature clouds in the lower troposphere. Note that the Pearson coefficients for the observed fields were never less than 0.98 regardless of the threshold because of the direct dependence of the retrieved cloud-top heights on the brightness temperatures.

The impact of the conditional sampling on the observed cloud-top uncertainty is hard to gauge because of the lack of accurate cloud-top measurements. Each filter removed fewer clouds in the observations than in the forecasts, especially during the day, indicating that the observed clouds were more optically thick. Because most systematic errors in the retrievals lead to low TCP biases, the observed optical thickness was likely

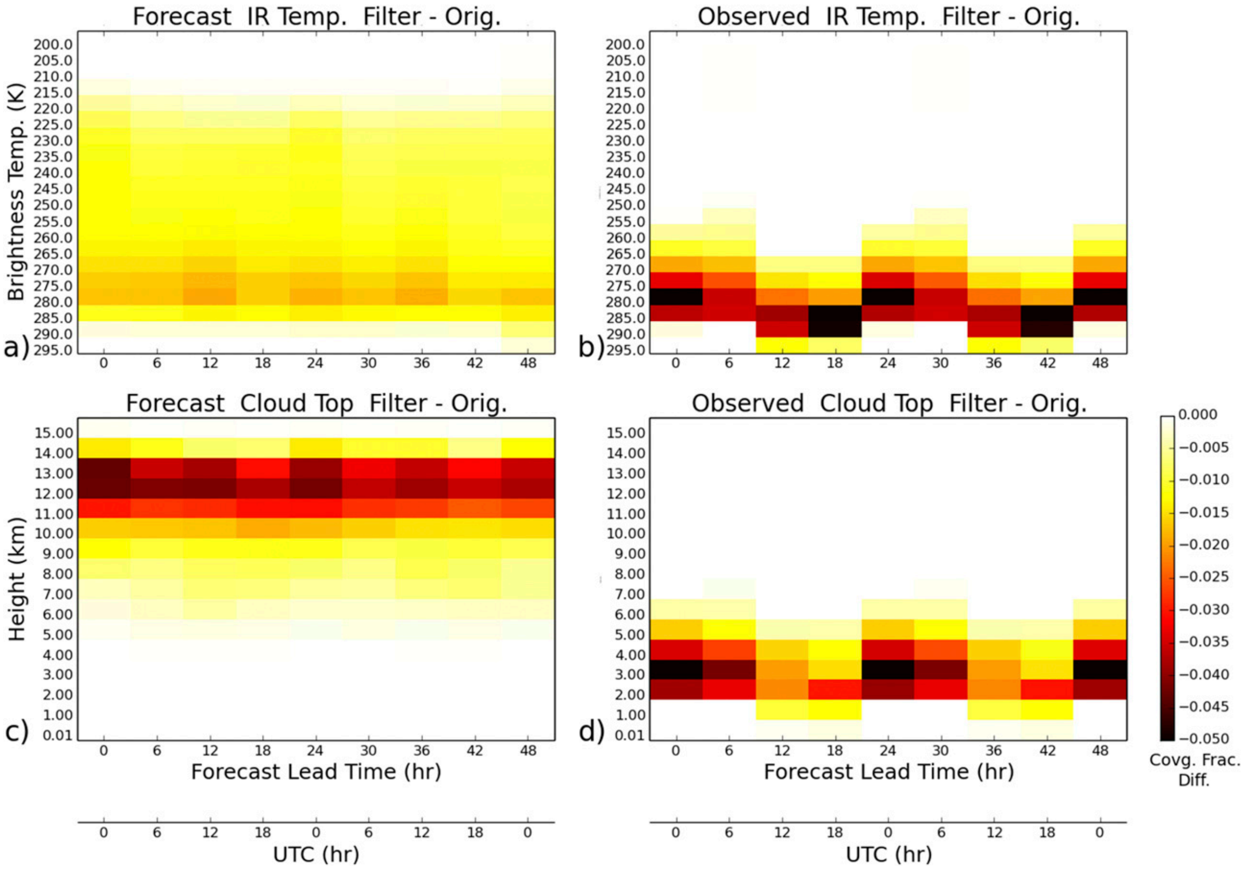

FIG. 8. Differences between the filtered mean coverage distributions shown in Fig. 7 minus the original coverage distributions in Fig. 2 for (top) brightness temperature and (bottom) cloud-top-height for all points with cloud cover at each forecast lead time for the (a),(c) forecast and (b),(d) observed values.

underestimated. Miller et al. (2000) found that the retrievals tend to underestimate TCP in areas of deep, inhomogeneous clouds. Low sun angles and variations between day and night formulations also lead to errors. These fluctuations likely contributed to the diurnal variability in the observed cloud properties in this study. Nachamkin et al. (2009) noted similar variations with sun angle. Min et al. (2004) did find that optical depth can be overestimated for inhomogeneous regions of optically thin cirrus over land. These clouds were exceedingly thin and scattered in nature, however, and would likely have been below the 25 g m^{-2} TCP threshold. Min et al. (2004) noted that optical depths for homogeneous cirrus clouds were well estimated.

4. Forecast cloud comparisons

The 2-week period covered by this study was characterized by a weak, quasi-stationary front that extended from central Florida northeastward into the Atlantic Ocean, off the east coast of the United States. The front served as a focal point for numerous convective systems and their associated high cloud tops, as indicated by the fraction of the period covered by clouds at each point (Figs. 11a,b). Episodic convective systems also moved

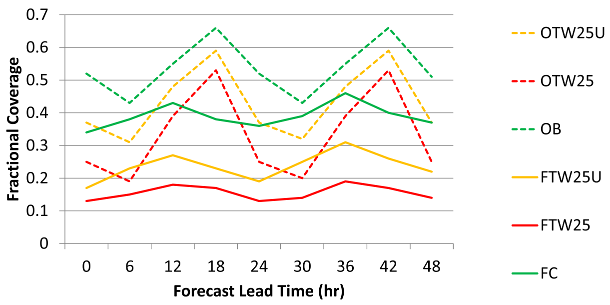

FIG. 9. Mean fraction of the verification domain covered by cloud during the 26 Jul–9 Aug period. The FC and OB lines represent the original cloud coverage from COAMPS and GOES, respectively. Coverage of clouds filtered with the 25 g m^{-2} TCP threshold is denoted by FTW25 (COAMPS) and OTW25 (GOES). Coverage of clouds filtered with the combined 25 g m^{-2} TCP/ice cloud criterion is denoted by FTW25U (COAMPS) and OTW25U (GOES).

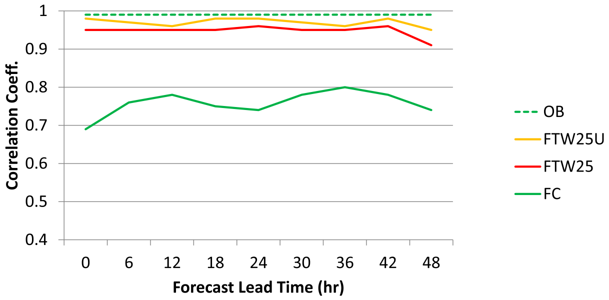

FIG. 10. Pearson correlations between cloud-top height and brightness temperature for all cloudy points over the verification domain for the 26 Jul–9 Aug period. The original GOES observations are denoted by the dashed OB line. Correlations from COAMPS clouds filtered with the 25 g m^{-2} TCP threshold are denoted by FTW25. Correlations from COAMPS clouds filtered with the combined 25 g m^{-2} TCP/ice cloud criterion are denoted by FTW25U.

into the verification domain from the Midwest, with centers tracking across Illinois, Indiana, and Ohio. Elsewhere, diurnal cumulus and scattered thunderstorms prevailed in a typical summertime regime.

Substantial differences existed between the forecast and observed filtered cloud fields. Total forecast cloud coverage within the verification domain was only 39%–74% of the observed total (Fig. 9). Although the observed cloud coverage exceeded that of the forecast, the forecast cloud-top heights were generally higher (Fig. 11). A bimodal cloud-top distribution was noted in the forecast fields, with maxima near the surface and the tropopause (Fig. 7); in contrast, the midtroposphere was characterized by a distinct lack of cloudiness. Observed cloud tops were more evenly distributed, with a single broad maximum just above the boundary layer that decreased significantly toward the surface.

A distinct maximum of low stratus clouds and/or fog existed within the forecast boundary layer during all times of the day. One of the most persistent regions of low clouds in the forecasts was located over the Atlantic, north of 40°N in the cool waters of the Labrador Current, north of the Gulf Stream (Fig. 11b). In that region, clouds with heights of 0.5 km or less were present at almost every lead time through the entire forecast. Observed cloud-top heights in that area, however, tended to be in the 4–6-km range (Fig. 11a). Multiple opportunities to sample any persistent low-level cloud layers existed in localized gaps in the observed midlevel cloud layer. Conditional samples of all cloud tops below 2 km revealed very few observed low clouds in this region (Fig. 12a). Ship and buoy data indicated no systematic bias in forecast near-surface temperatures or dewpoint temperatures, which suggested that other factors involving the turbulent mixing or radiation scheme may have been involved in producing the spurious low clouds. A general lack of midlevel clouds in the forecasts may also have had significant radiative impacts in the lower atmosphere.

Another relative maximum of occurrence of low clouds in the forecasts was apparent over Illinois eastward through Virginia (Fig. 12b). Most of these clouds occurred during two distinct periods. The first period, which occurred during 27–28 July, was characterized by boundary layer clouds and widely scattered thunderstorms. Extensive stratus on the morning of 27 July was well simulated, but the model did not forecast the conversion to nonprecipitating cumulus during the afternoon hours. Instead, the stratus decreased in coverage but no shallow cumulus clouds developed. Several regions of observed midlevel clouds were also not simulated. On 28 July, the day started out mostly clear in the observations but cumulus and scattered thunderstorms developed as the day progressed. In the model, regions of stratus persisted through the day but again cumulus failed to develop. On both days, the Kain–Fritsch parameterization produced light precipitation in the forecasts that was too broad in coverage relative to the observations. Although forecast rainfall amounts in this area were generally 5 mm or less, the cooling and moistening tended to reinforce the stratus deck. Forecast temperatures were 2°–3°C below the observed values, and dewpoints were 2°–3°C above observed values.

The second cloud event was associated with the passage of a synoptic wave on 6–8 August. Being more synoptic in nature, the overall cloud shield was better simulated but was more extensive than the observations and contained fewer breaks. In the observations, midlevel clouds obscured much of the low cloud deck. Low cloud tops that were visible tended to be 0.5 km higher than what was forecast. Vertical grid spacings of 200–300 m at cloud top may have contributed to some errors, although the systematic nature of the low bias suggests the boundary layer may have been too shallow. The model produced far less midlevel cloudiness than was observed with this system and, as a result, registered more low cloud tops than the observations did in the statistical mean.

Aside from in the Midwest, the model produced relatively few low clouds over most land locations. In contrast, extensive regions of low clouds were observed over the northeastern and southeastern United States as well as Cuba and the Bahamas. Many of these clouds were convective in nature. As with the first Midwestern cloud case above, the model failed to produce any cumulus clouds. This result is primarily due to the inability of the model to represent or resolve cumulus clouds on the 15-km horizontal grid.

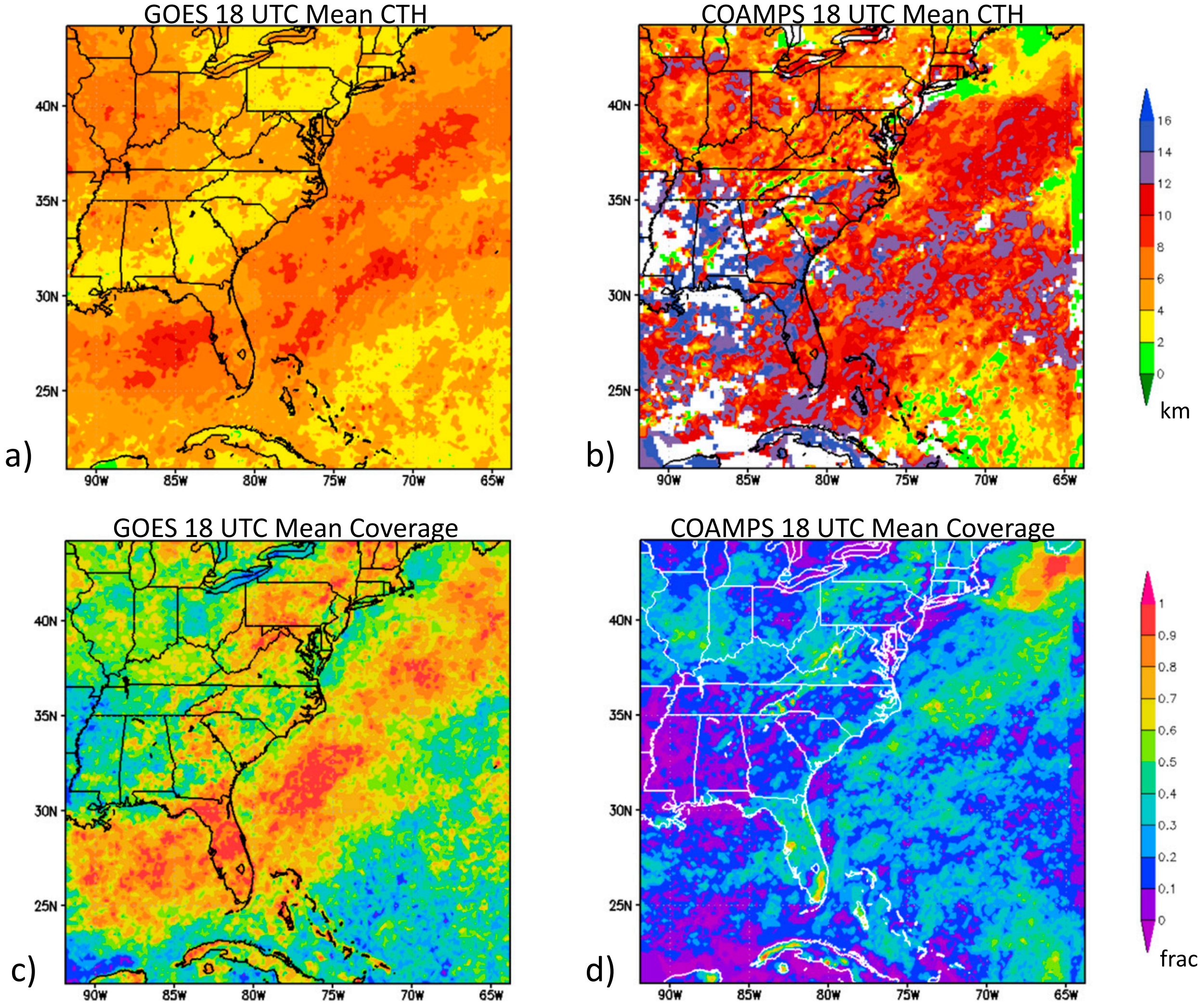

FIG. 11. Mean properties of the filtered cloud fields over the verification domain at 1800 UTC for the 26 Jul–9 Aug period. Mean cloud-top height (CTH; km MSL) is displayed for the (a) GOES observations and (b) 18-h COAMPS forecasts. In addition, the mean cloud coverage as defined by the fraction of the 2-week period with cloud cover is displayed for both (c) GOES and (d) COAMPS. White regions in (b) correspond to areas that were clear at 1800 UTC through the entire period.

The forecast performed best with the low clouds over the open waters of the Sargasso Sea (east of 70°W and south of 30°N). This region was characterized by a fairly steady state field of scattered cumulus clouds that remained little changed away from any convection. Although the simulated clouds were more widespread than was observed, their behavior was similar. Scattered areas of low cloud traversed the area in a nearly continuous fashion. The cloud tops were still too low relative to the observations. Predicted values were between 0.6 and 1.2 km, whereas observed values were from 1.6 to 1.8 km.

Higher clouds associated with precipitating convection were more consistently represented than the low or middle clouds, as indicated by the coverage of all clouds with tops of greater than or equal to 10 km (Fig. 13). The region of convection paralleling the U.S. east coast was well located, especially in its northern portions. Farther south, the 18-h forecast cloud tops were more scattered in coverage. Parameterized convection was the primary source for high cloudiness in the model, and the Kain–Fritsch parameterization tended to preferentially trigger over land along convergence zones related to differential heating. The sea breeze is especially visible over the eastern Carolinas as well as parts of Florida. Mountainous regions along the Appalachians and western Cuba also acted as focal points for excess convection. Many of these convergence zones are apparent in the total observed cloud frequencies (Fig. 11c), but the observed cloud tops

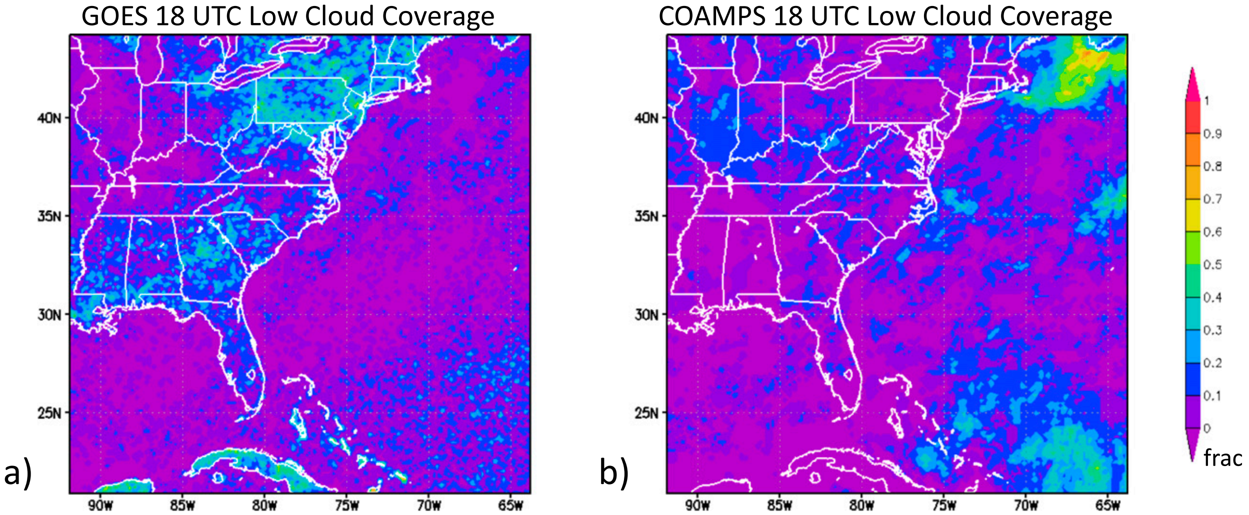

FIG. 12. Mean coverage from the filtered cloud fields for all clouds with tops of less than 2 km for the (a) GOES observations and (b) 18-h COAMPS forecasts.

were considerably lower than the forecasts, indicating that the parameterized convection was too deep. Forecast cloud-top heights were generally higher than the observations even in regions where both had deep convection. The presence of multiple lower- and middle-cloud layers contributed to reducing the observed means. The cloud-top distributions indicate a positive bias in the upper-tropospheric cloud tops (Fig. 7), however, and notable was that the forecasts also contained more than 2 times the amount of thin ice clouds that were observed (Fig. 9). These additional clouds were largely associated with overactive weak convection in the Kain–Fritsch scheme. A newer version of this scheme outlined by Kain (2004) that was tested in COAMPS was found to considerably reduce the upper-tropospheric cloud bias, but it also adversely affected the tropical-cyclone-track performance and thus was not implemented in the operational model.

A final experiment was performed to investigate the impacts of interpolating the cloud observations to the COAMPS grid. As mentioned in section 2, the cloud-type observations were interpolated to the forecast grid on the basis of the mode of the observed values within each grid box. As a result, the interpolated cloud-type fields had fewer cloudy points than the observed cloud-top fields, which were interpolated from the mean value. To generate a conservative cloud-cover estimate, the interpolated cloud-type field was used as a template for the cloud-top heights. A further reduction in cloud cover was achieved by masking out all points

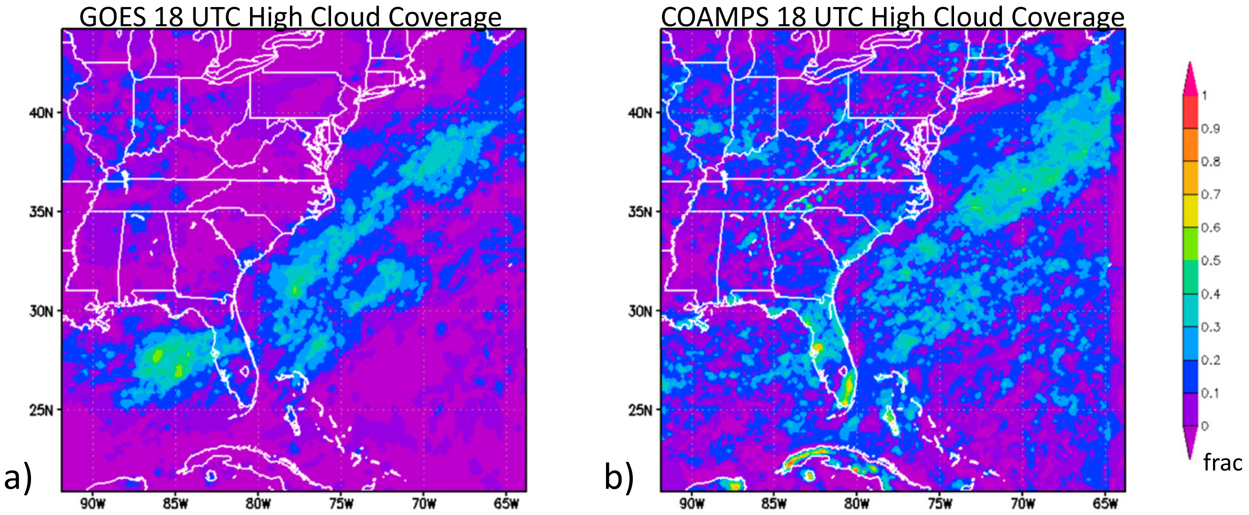

FIG. 13. As in Fig. 12, but cloud coverage for all filtered clouds with tops that are greater than or equal to 10 km is displayed.

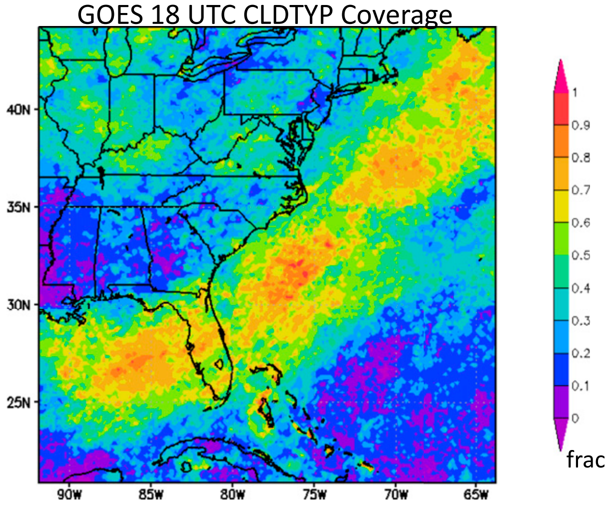

FIG. 14. Mean cloud coverage from the filtered GOES cloud field as in Fig. 11c, but templated by the interpolated cloud-type field with all clear and partly cloudy points removed.

on the verification (model) grid that were assigned the partly cloudy cloud type. The end result (Fig. 14) was an underestimate of the observed cloud cover but possibly a better comparison with the binary cloud/no-cloud modeled microphysical fields, which are unable to resolve subgrid-scale partly cloudy regimes. It also revealed regions that were sensitive to interpolation bias. The greatest reductions occurred in the same regions of enhanced geophysical convergence that were associated with increased cloud frequencies in the initial filtered cloud-cover field (Fig. 11c). Most affected were mountainous regions, as well as the narrow convergence zones directly inland from the southeastern U.S. coast. Cloud cover over the Sargasso Sea was also reduced.

When compared with the forecast cloud cover (Fig. 11d), the greatest improvements in agreement were over the Appalachians and the Ohio Valley region. Cloud-cover reductions over eastern Cuba also brought improved agreement there. These were all areas of enhanced diurnal cumulus development associated with heating and convergence. Agreement was reduced in areas where the Kain–Fritsch scheme was already active such as the southeastern U.S. coast and western Cuba. In the region of deep convective clouds along the frontal zone, the masked cloud reductions did little to improve agreement with the forecast. Many of the widespread midlevel cloud decks occurred in this area, and these were poorly simulated in the forecasts. Another area of reduced agreement was over the Sargasso Sea, where the 18-h forecasts were relatively good. The model is apparently better able to resolve the physics that lead to partly cloudy conditions in this region.

5. Summary and conclusions

Cloud-top-height verification is difficult because of the limited nature of the cloud-top-height observations. Direct comparison of the observed and synthetic radiance fields is generally considered to be the best practice because of the minimized conversion errors, but the uncertainties associated with transparent clouds are still embedded in the radiances, and errors in vertical cloud structure are difficult to extract. Results of this work indicate that up to 50% of what could be perceived as low cloud cover with high brightness temperature is, in fact, optically thin cirrus. Although satellite retrievals and model-defined cloud tops are prone to large errors, satisfactory results are possible for relatively thick clouds. When combined with targeted conditional sampling, the retrievals helped to identify localized errors in the representation of cumulus clouds over land as well as stratus in the North Atlantic. In some cases the conditional sampling may remove thin cirrus overlying a layer of water clouds. The TCP in these situations generally exceeds $25\,\mathrm{g\,m^{-2}}$, however. As next-generation satellites come online, the ability to detect cirrus will likely improve. GOES-R will observe radiances at $1.38\,\mu\mathrm{m}$, which can be used to identify thin cirrus during the day.

One significant deficiency noted in this study was the near absence of midlevel cloudiness in the forecasts. A low bias in midlevel cloudiness is not unique to COAMPS; both Cintineo et al. (2014), and Thompson et al. (2016) have also noted it in the Advanced Research version of the Weather Research and Forecasting Model. Thompson et al. (2016) attributed some of the issues to a lack of coupling between the radiation and cloud microphysics schemes. Their modifications, which allowed the effective radii of cloud ice, cloud water, and snow to evolve through time, resulted in a modest improvement in specific cloud systems. On average, however, the negative bias persisted. During the period studied for this current work, extensive altocumulus layers were frequently observed in and near regions of active convection. Although broad in extent, altocumulus layers are often only a few hundred meters thick, consist of mixed-phase clouds, and form beneath weak virtual potential temperature inversions on the order of 1–2 K (Fleishauer et al. (2002). In a combined modeling and observational study, Schmidt et al. (2014) noted that altocumulus clouds are driven by horizontal and vertical radiative heating gradients, which likely require high vertical resolution, as well as detailed coupling between the radiation and the microphysics. For this study, the midlevel vertical grid spacing in these forecasts was on the order of 200–500 m, which is too large to properly resolve these clouds. The numerical diffusion required for the leapfrog advection likely filters out any inversions spanning less than five–six vertical levels. Many of these inversions result from convective subsidence propagating outward from existing convection (Nicholls et al. 1991; Mapes 1993; Pandya and Durran 1996; Nachamkin and Cotton 2000). Whether the Kain–Fritsch scheme properly simulates the effects of the convective heating on the surrounding environment is an open question. Although heat and moisture are exchanged, convective clouds are not explicitly simulated.

Given the many possible contributing factors, the negative bias in midlevel cloudiness remains a fundamental challenge to numerical weather prediction. Ongoing work is focused on determining what measures can be taken to better simulate these clouds. Simulations at horizontal grid spacings of 1.67 km have shown some improvements in the midlevel clouds, perhaps as a result of improved depiction of deep convection. Additional simulations with midtropospheric vertical grid spacings on the order of 50 m are also planned, but these intensive simulations are not yet practical for large operational areas. Improvements in the advection and turbulent transport schemes will likely be necessary to fully capture these elusive clouds.

Aside from the midlevel clouds, biases in upper- and lower-atmospheric cloud cover attributed to parameterized convection are relatively well known (Kain 2004). Relevant questions arise about how the model should be evaluated. From a forecaster-oriented perspective the lack of cumulus clouds over land is an important problem, but most cloud forecasts are derived from models that will not fully resolve all types of clouds. Clouds and precipitation are unique in that their existence is a binary function. A staggered verification approach can be helpful whereby the smaller-scale clouds are filtered from the observations to help to separate resolvable error from unresolvable error. In regions where clouds are poorly resolved, improved parameterization or statistical postprocessing may be able to fill the gap. In other regions, such as the Sargasso Sea in this case, certain aspects of the cloud field appear to be at least partially resolvable. Results like these can help further understanding of the physical processes leading to these clouds.

Acknowledgments. This research is supported by the Office of Naval Research under Grant N0001416WX01201. Computer resources for the COAMPS simulations and data archival were supported in part by a grant of high-performance-computer time from the U.S. Department of Defense Major Shared Resource Center at Stennis Space Center in Mississippi. The work was performed on an IBM iDataPlex computer.

REFERENCES

Bankert, R. L., C. Mitrescu, S. D. Miller, and R. H. Wade, 2009: Comparison of GOES cloud classification algorithms employing explicit and implicit physics. *J. Appl. Meteor. Climatol.*, **48**, 1411–1421, doi:10.1175/2009JAMC2103.1.

Bikos, D., and Coauthors, 2012: Synthetic satellite imagery for real-time high-resolution model evaluation. *Wea. Forecasting*, **27**, 784–795, doi:10.1175/WAF-D-11-00130.1.

Chaboureau, J.-P., and J.-P. Pinty, 2006: Evaluation of a cirrus parameterization with Meteosat Second Generation. *Geophys. Res. Lett.*, **33**, L03815, doi:10.1029/2005GL024725.

Chen, S., and Coauthors, 2003: COAMPS version 3 model description—General theory and equations. Naval Research Laboratory Tech. Note NRL/PU/7500-03448, 141 pp. [Available online at http://www.nrlmry.navy.mil/coamps_docs/base/docs/COAMPS_2003.pdf.]

Chevallier, F., and G. Kelly, 2002: Model clouds as seen from space: Comparison with geostationary imagery in the 11-μm window channel. *Mon. Wea. Rev.*, **130**, 712–722, doi:10.1175/1520-0493(2002)130<0712:MCASFS>2.0.CO;2.

Cintineo, R., J. A. Otkin, M. Xue, and F. Kong, 2014: Evaluating the performance of planetary boundary layer and cloud microphysical parameterization schemes in convection permitting ensemble forecasts using synthetic *GOES-13* satellite observations. *Mon. Wea. Rev.*, **142**, 163–182, doi:10.1175/MWR-D-13-00143.1.

Daley, R., and E. Barker, 2001: NAVDAS Source Book 2001: NRL Atmospheric Variational Data Assimilation System. Naval Research Laboratory Publ. NRL/PU/7530–01-441, 163 pp. [Available online at http://www.dtic.mil/dtic/tr/fulltext/u2/a396883.pdf.]

Davies, H. C., 1976: A lateral boundary formulation for multi-level prediction models. *Quart. J. Roy. Meteor. Soc.*, **102**, 405–418, doi:10.1002/qj.49710243210.

Deeter, M., and K. F. Evans, 1998: A hybrid Eddington single-scatter radiative transfer model for computing radiances from thermally emitting atmospheres. *J. Quant. Spectrosc. Radiat. Transfer*, **60**, 635–648, doi:10.1016/S0022-4073(97)00245-8.

Fleishauer, R. P., V. E. Larson, and T. H. Vonder Haar, 2002: Observed microphysical structure of midlevel, mixed-phase clouds. *J. Atmos. Sci.*, **59**, 1779–1804, doi:10.1175/1520-0469(2002)059<1779:OMSOMM>2.0.CO;2.

Frey, R., B. Baum, W. Menzel, S. Ackerman, C. Moeller, and J. Spinhirne, 1999: A comparison of cloud top heights computed from airborne lidar and MAS radiance data using CO_2 slicing. *J. Geophys. Res.*, **104**, 24 547–24 555, doi:10.1029/1999JD900796.

Grasso, L. D., and T. Greenwald, 2004: Analysis of 10.7-μm brightness temperatures of a simulated thunderstorm with two-moment microphysics. *Mon. Wea. Rev.*, **132**, 815–825, doi:10.1175/1520-0493(2004)132<0815:AOMBTO>2.0.CO;2.

——, M. Sengupta, J. F. Dostalek, R. Brummer, and M. DeMaria, 2008: Synthetic satellite imagery for current and future environmental satellites. *Int. J. Remote Sens.*, **29**, 4373–4384, doi:10.1080/01431160801891820.

——, ——, and M. DeMaria, 2010: Comparison between observed and synthetic 6.5 and 10.7 μm *GOES-12* imagery of thunderstorms that occurred on 8 May 2003. *Int. J. Remote Sens.*, **31**, 647–663, doi:10.1080/01431160902894483.

——, D. Lindsey, K.-S. Sunny Lim, A. J. Clark, D. Bikos, and S. R. Dembek, 2014: Evaluation of and suggested improvements to the WSM6 microphysics in WRF-ARW using synthetic and observed *GOES-13* imagery. *Mon. Wea. Rev.*, **142**, 3635–3650, doi:10.1175/MWR-D-14-00005.1.

Hodur, R. M., 1997: The Naval Research Laboratory's Coupled Ocean/Atmosphere Mesoscale Prediction System (COAMPS). *Mon. Wea. Rev.*, **125**, 1414–1430, doi:10.1175/1520-0493(1997)125<1414:TNRLSC>2.0.CO;2.

Jankov, I., and Coauthors, 2011: An evaluation of five ARW-WRF microphysics schemes using synthetic GOES imagery for an atmospheric river event affecting the California coast. *J. Hydrometeor.*, **12**, 618–633, doi:10.1175/2010JHM1282.1.

Jin, Y., and Coauthors, 2014: The impact of ice phase cloud parameterizations on tropical cyclone prediction. *Mon. Wea. Rev.*, **142**, 606–625, doi:10.1175/MWR-D-13-00058.1.

Kain, J. S., 2004: The Kain–Fritsch convective parameterization: An update. *J. Appl. Meteor.*, **43**, 170–181, doi:10.1175/1520-0450(2004)043<0170:TKCPAU>2.0.CO;2.

——, and J. M. Fritsch, 1993: Convective parameterization for mesoscale models: The Kain–Fritsch scheme. *The Representation of Cumulus Convection in Numerical Models, Meteor. Monogr.*, No. 46. Amer. Meteor. Soc., 165–170.

Karlsson, K.-G., 1996: Validation of modeled cloudiness using satellite-estimated cloud climatologies. *Tellus*, **48A**, 767–785, doi:10.1034/j.1600-0870.1996.t01-1-00015.x.

Liu, M., J. E. Nachamkin, and D. L. Westphal, 2009: On the improvement of COAMPS weather forecasts using an advanced radiative transfer model. *Wea. Forecasting*, **24**, 286–306, doi:10.1175/2008WAF2222137.1.

Lopez, P., K. Finkele, P. Clark, and P. Mascart, 2003: Validation and intercomparison of three FASTEX cloud systems: Comparison with coarse-resolution simulations. *Quart. J. Roy. Meteor. Soc.*, **129**, 1841–1871, doi:10.1256/qj.01.113.

Mapes, B. E., 1993: Gregarious tropical convection. *J. Atmos. Sci.*, **50**, 2026–2037, doi:10.1175/1520-0469(1993)050<2026:GTC>2.0.CO;2.

Mellor, G. L., and T. Yamada, 1982: Development of a turbulence closure for geophysical fluid problems. *Rev. Geophys. Space Phys.*, **20**, 851–875, doi:10.1029/RG020i004p00851.

Miller, S. D., G. L. Stephens, C. K. Drummond, A. K. Heidinger, and P. T. Partain, 2000: A multisensor diagnostic satellite cloud property retrieval scheme. *J. Geophys. Res.*, **105**, 19 955–19 971, doi:10.1029/2000JD900273.

Min, Q., P. Minnis, and M. M. Khaiyer, 2004: Comparison of cirrus optical depths derived from *GOES 8* and surface measurements. *J. Geophys. Res.*, **109**, D15207, doi:10.1029/2003JD004390.

Minnis, P., and Coauthors, 2008: Near-real time cloud retrievals from operational and research meteorological satellites. *Remote Sensing of Clouds and the Atmosphere XIII*, R. H. Picard et al., Eds., International Society for Optical Engineering (SPIE Proceedings, Vol. 7107), doi:10.1117/12.800344.

Mitrescu, C., S. Miller, and R. Wade, 2006: Cloud optical and microphysical properties derived from satellite data. *14th Conf. on Satellite Meteorology and Oceanography*, Atlanta, GA, Amer. Meteor. Soc., P1.7. [Available online at https://ams.confex.com/ams/Annual2006/webprogram/Paper100515.html.]

Morcrette, J. J., 1991: Evaluation of model-generated cloudiness: Satellite-observed and model-generated diurnal variability of brightness temperatures. *Mon. Wea. Rev.*, **119**, 1205–1224, doi:10.1175/1520-0493(1991)119<1205:EOMGCS>2.0.CO;2.

Nachamkin, J. E., and W. R. Cotton, 2000: Interactions between a developing mesoscale convective system and its environment. Part II: Numerical simulation. *Mon. Wea. Rev.*, **128**, 1225–1244, doi:10.1175/1520-0493(2000)128<1225:IBADMC>2.0.CO;2.

——, J. Schmidt, and C. Mitrescu, 2009: Verification of cloud forecasts over the eastern Pacific using passive satellite retrievals. *Mon. Wea. Rev.*, **137**, 3485–3500, doi:10.1175/2009MWR2853.1.

Nicholls, M. E., R. A. Pielke, and W. R. Cotton, 1991: Thermally forced gravity waves in an atmosphere at rest. *J. Atmos. Sci.*, **48**, 1869–1884, doi:10.1175/1520-0469(1991)048<1869:TFGWIA>2.0.CO;2.

Otkin, J. A., and T. J. Greenwald, 2008: Comparison of WRF model-simulated and MODIS-derived cloud data. *Mon. Wea. Rev.*, **136**, 1957–1970, doi:10.1175/2007MWR2293.1.

——, ——, J. Sieglaff, and H.-L. Huang, 2009: Validation of a large-scale simulated brightness temperature dataset using SEVIRI satellite observations. *J. Appl. Meteor. Climatol.*, **48**, 1613–1626, doi:10.1175/2009JAMC2142.1.

Pandya, R. D., and D. Durran, 1996: The influence of convectively generated thermal forcing on the mesoscale circulation around squall lines. *J. Atmos. Sci.*, **53**, 2924–2951, doi:10.1175/1520-0469(1996)053<2924:TIOCGT>2.0.CO;2.

Rikus, L., 1997: Application of a scheme for validating clouds in an operational global NWP model. *Mon. Wea. Rev.*, **125**, 1615–1637, doi:10.1175/1520-0493(1997)125<1615:AOASFV>2.0.CO;2.

Rutledge, S. A., and P. V. Hobbs, 1983: The mesoscale and microscale structure and organization of clouds and precipitation in midlatitude cyclones. VIII: A model for the "seeder-feeder" process in warm-frontal rainbands. *J. Atmos. Sci.*, **40**, 1185–1206, doi:10.1175/1520-0469(1983)040<1185:TMAMSA>2.0.CO;2.

——, and ——, 1984: The mesoscale and microscale structure and organization of clouds and precipitation in midlatitude cyclones. XII: A diagnostic modeling study of precipitation development in narrow cold-frontal rainbands. *J. Atmos. Sci.*, **41**, 2949–2972, doi:10.1175/1520-0469(1984)041<2949:TMAMSA>2.0.CO;2.

Schmidt, J. M., P. J. Flatau, and R. D. Yates, 2014: Convective cells in altocumulus observed with a high-resolution radar. *J. Atmos. Sci.*, **71**, 2130–2154, doi:10.1175/JAS-D-13-0172.1.

Smith, W. L., Jr., and Coauthors, 1996: Comparisons of cloud heights derived from satellite, aircraft, surface lidar and LITE data. *Proc. Int. Radiation Symp.*, Fairbanks, AK, Int. Radiation Commission, 603–606.

Söhne, N., J.-P. Chaboureau, and F. Guichard, 2008: Verification of cloud cover forecast with satellite observation over West Africa. *Mon. Wea. Rev.*, **136**, 4421–4434, doi:10.1175/2008MWR2432.1.

Sun, Z., and L. Rikus, 2004: Validating model clouds and their optical properties using geostationary satellite imagery. *Mon. Wea. Rev.*, **132**, 2006–2020, doi:10.1175/1520-0493(2004)132<2006:VMCATO>2.0.CO;2.

Thompson, G., M. Tewari, K. Ikeda, S. Tessendorf, C. Weeks, J. Otkin, and F. Kong, 2016: Explicitly-coupled cloud physics and radiation parameterizations and subsequent evaluation in WRF high-resolution convective forecasts. *Atmos. Res.*, **168**, 92–104, doi:10.1016/j.atmosres.2015.09.005.

Tselioudis, G., and C. Jakob, 2002: Evaluation of midlatitude cloud properties in a weather and a climate model: Dependence on dynamic regime and spatial resolution. *J. Geophys. Res.*, **107**, 4781, doi:10.1029/2002JD002259.

Zhang, M., and Coauthors, 2005: Comparing clouds and their seasonal variations in 10 atmospheric general circulation models with satellite measurements. *J. Geophys. Res.*, **110**, D15S02, doi:10.1029/2004JD005021.