Assimilation of Satellite Infrared Radiances and Doppler Radar Observations during a Cool Season Observing System Simulation Experiment

THOMAS A. JONES

Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, Norman, Oklahoma

JASON A. OTKIN

Cooperative Institute for Meteorological Satellite Studies, Space Science and Engineering Center, University of Wisconsin–Madison, Madison, Wisconsin

DAVID J. STENSRUD

Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, and NOAA/OAR/National Severe Storms Laboratory, Norman, Oklahoma

KENT KNOPFMEIER

Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, Norman, Oklahoma

(Manuscript received 15 September 2012, in final form 15 March 2013)

ABSTRACT

An observing system simulation experiment is used to examine the impact of assimilating water vapor–sensitive satellite infrared brightness temperatures and Doppler radar reflectivity and radial velocity observations on the analysis accuracy of a cool season extratropical cyclone. Assimilation experiments are performed for four different combinations of satellite, radar, and conventional observations using an ensemble Kalman filter assimilation system. Comparison with the high-resolution ''truth'' simulation indicates that the joint assimilation of satellite and radar observations reduces errors in cloud properties compared to the case in which only conventional observations are assimilated. The satellite observations provide the most impact in the mid- to upper troposphere, whereas the radar data also improve the cloud analysis near the surface and aloft as a result of their greater vertical resolution and larger overall sample size. Errors in the wind field are also significantly reduced when radar radial velocity observations were assimilated. Overall, assimilating both satellite and radar data creates the most accurate model analysis, which indicates that both observation types provide independent and complimentary information and illustrates the potential for these datasets for improving mesoscale model analyses and ensuing forecasts.

1. Introduction

Correctly analyzing clouds and cloud cover within numerical weather prediction (NWP) models is vital to producing accurate forecasts of high-impact weather events (e.g., Polkinghorne and Vukicevic 2011). Cloud processes are inherently nonlinear with complex interactions occurring between various cloud microphysical species. Many advances in cloud microphysical parameterization schemes have been made in recent years, but the full benefit of these changes can only be realized if the initial cloud analysis is also improved through assimilation of either direct or indirect observations of cloud properties. Several potential data sources exist, but none can provide the complete answer alone.

Observations of cloud properties from surface-based disdrometers, aircraft, and surface or satellite-based lidars represent the most direct measurements, but they are not optimal for large-scale data assimilation since they provide observations for very limited spatial and temporal coverage. Satellite radiances represent the most numerous and widely available observations of the

Corresponding author address: Dr. Thomas A. Jones, Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma and NOAA/OAR/National Severe Storms Laboratory, 120 David L. Boren Blvd., Norman, OK 73072.
E-mail: thomas.jones@noaa.gov

DOI: 10.1175/MWR-D-12-00267.1

atmosphere and are a critical component of most global and regional assimilation systems. Satellite sensors sample outgoing top-of-the-atmosphere radiation in the form of visible, infrared, and microwave radiances, producing detailed information about the atmospheric state, such as moisture, temperature, wind, and cloud properties. Many prior studies have shown that forecast skill is improved when satellite-derived water vapor and temperature profile retrievals or infrared and microwave radiances are assimilated using a radiative transfer model (RTM) (e.g., Derber and Wu 1998; McNally et al. 2000; Bouttier and Kelly 2001; Chevallier et al. 2004; McNally et al. 2006; Le Marshall et al. 2006; McCarty et al. 2009; Collard and McNally 2009). Additional studies have also shown that assimilation of cloudy infrared observations in both ensemble and variational assimilation systems improves the 3D cloud structure and forecast skill in cloud-resolving and global circulation models (Vukicevic et al. 2004, 2006; Reale et al. 2008; Otkin 2010, 2012a,b; Stengel et al. 2010; Seaman et al. 2010).

Networks of surface radars provide another source of cloud information. In the contiguous United States (CONUS), a network of approximately 140 Weather Surveillance Radar-1988 Dopplers (WSR-88Ds) provide high-resolution 3D volume scans of reflectivity and radial velocity (Crum and Alberty 1993) in precipitating areas. Various cloud microphysical properties, such as phase, particle size, and number concentration can be inferred from the reflectivity data. During the past decade, many studies have examined the potential for assimilating these data into storm-scale NWP models to improve the representation of precipitation (often convective) within the model analysis (e.g., Weygandt et al. 2002; Snyder and Zhang 2003; Zhang et al. 2004; Aksoy et al. 2009, 2010; Dowell et al. 2011).

Many of these studies employ high-resolution ($<$5 km) grids that explicitly resolve the precipitation features using variational techniques (e.g., Gao et al. 1999, 2004; Weygandt et al. 2002; Gao and Stensrud 2012). Additional studies have focused on using an ensemble approach that has the advantage of providing a flow-dependent background covariance (Snyder and Zhang 2003; Zhang et al. 2004; Aksoy et al. 2009, 2010; Dowell et al. 2011; Yussouf and Stensrud 2012). Most prior research has focused on assimilating radial velocity measurements to improve the accuracy of the wind fields while also employing reflectivity data to generate an accurate microphysical representation of convection (e.g., Aksoy et al. 2009, 2010). Assimilating reflectivity has proven more difficult because of large uncertainties on the relationship between reflectivity and various cloud microphysical parameters as well as how best to handle precipitation-free areas. Using an ensemble-based approach to assimilate these data has been shown to reliably handle these uncertainties (e.g., Caya et al. 2005). Since radar reflectivity is a measure of precipitation characteristics much like cloudy satellite radiances are a measure of cloud properties, similar challenges when assimilating each into NWP models apply to both.

Satellite radiances and radar reflectivity represent two fundamentally different observation types. Infrared radiances are sensitive to temperature and moisture in clear-sky regions while also providing information about cloud microphysical properties in cloudy regions. Radar reflectivity is most sensitive to large precipitation-sized hydrometers in the cloud interior, with radial velocity observations providing information on their motion. Although radar data have higher spatial and temporal resolutions than satellite observations, satellites provide information where surface radars are not present such as over the oceans and in clear skies. In addition, nonprecipitating cloud hydrometeors are often too small to return appreciable radar reflectively, while satellite radiances are generally sensitive to these small cloud particles. Since both sensors provide complementary information, assimilating both into NWP models should prove advantageous. However, the potential gains possible from combining these datasets remain largely unexplored.

To assess this potential, this study uses an observing system simulation experiment (OSSE) to explore the use of both satellite radiances and WSR-88D reflectivity and radial velocity data for assimilation onto a mesoscale model grid for a case study that occurred on 24 December 2009. This event corresponds to a significant winter storm in the southern plains and severe convection over the Mississippi Valley. Using an OSSE allows us to examine the potential for future satellite observations such as those from the Advanced Baseline Imager (ABI) on board the Geostationary Operational Environmental Satellite-R (GOES-R) satellite (Schmit et al. 2005) to improve NWP model forecasts. To take advantage of the high temporal resolution of these observations ($\sim$5 min), an ensemble Kalman filter (EnKF) approach is used (Kalman 1960). The primary advantage of this approach is that it provides a flow-dependent and dynamically evolving estimate of the background error covariance (Evensen 1994). This is especially important for both satellite and radar data as they can vary significantly with space and time (Heemink et al. 2001).

This study represents one of the first attempts at assessing the synergistic qualities of assimilating both satellite radiances and radar reflectivity and radial velocity in a mesoscale environment. It continues recent work by Otkin (2012b) that examined the potential

TABLE 1. Observation types assimilated into each experiment.

Expt	Observation types
CONV	Conventional observations only (ASOS, ACARS, raob)
SAT	Conventional + 6.95-μm ABI clear and cloudy radiances
RAD	Conventional + radar reflectivity and radial velocity
RADSAT	Conventional + radar + satellite

impact of the three water vapor–sensitive infrared bands on the ABI to improve forecasts of a high-impact weather event. Section 2 provides a description of the event and the high-resolution truth simulation. Section 3 discusses the simulated observation types with section 4 outlining the assimilation configuration. The characteristics of the assimilated observations are provided in section 5 with section 6 describing observation diagnostics during the assimilation period.

Comparisons of truth data with experiments that assimilate conventional, satellite, and/or radar observations (Table 1) are provided in section 7. Conclusions follow in section 8.

2. Truth simulation

A truth simulation depicting the evolution of a strong midlatitude cyclone and associated areas of precipitation across the central United States on 24 December 2009 is generated using version 3.3 of the Advanced Research Weather Research and Forecasting Model (WRF-ARW) (Skamarock et al. 2008), hereafter labeled as ''truth.'' Global 0.5° FNL (final) analyses from NCEP are used to initialize and provide boundary conditions for the simulation starting at 0000 UTC 23 December 2009 on an 1100×750 gridpoint domain covering CONUS with 6-km horizontal grid spacing and 52 vertical levels (Fig. 1), which is integrated over the next 48 h and output at 5-min intervals. The vertical grid spacing decreases from <100 m in the lowest km to $\sim$700 m near the model top at 10 hPa. The truth simulation uses the WRF single-moment 6-class mixed-phase microphysics scheme (WSM6), the Yonsei University planetary boundary layer scheme (Hong et al. 2006), the Kain–Fritsch cumulus parameterization scheme (Kain and Fritsch 1993; Kain 2004), and the Rapid Radiative Transfer Model (RRTM) longwave and shortwave radiation schemes (Iacono et al. 2008) to parameterize subgrid-scale

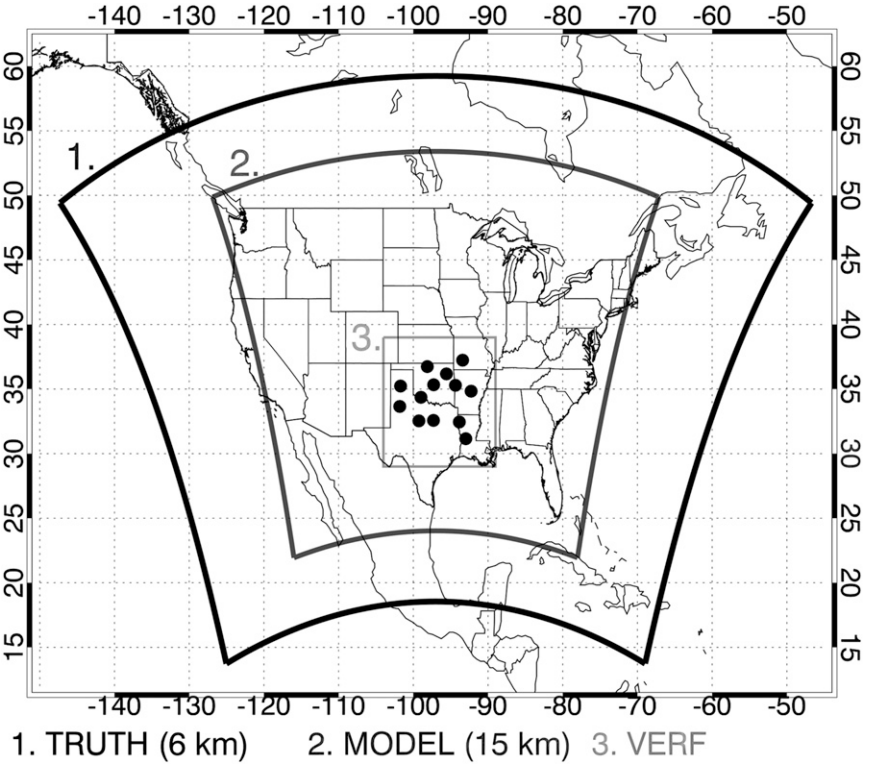

FIG. 1. 1) Truth simulation, 2) forecast model, and 3) verification domains used in this study. All domains have a horizontal resolution of 15 km with 52 vertical levels. Locations of the 13 WSR-88Ds are located within the verification domain.

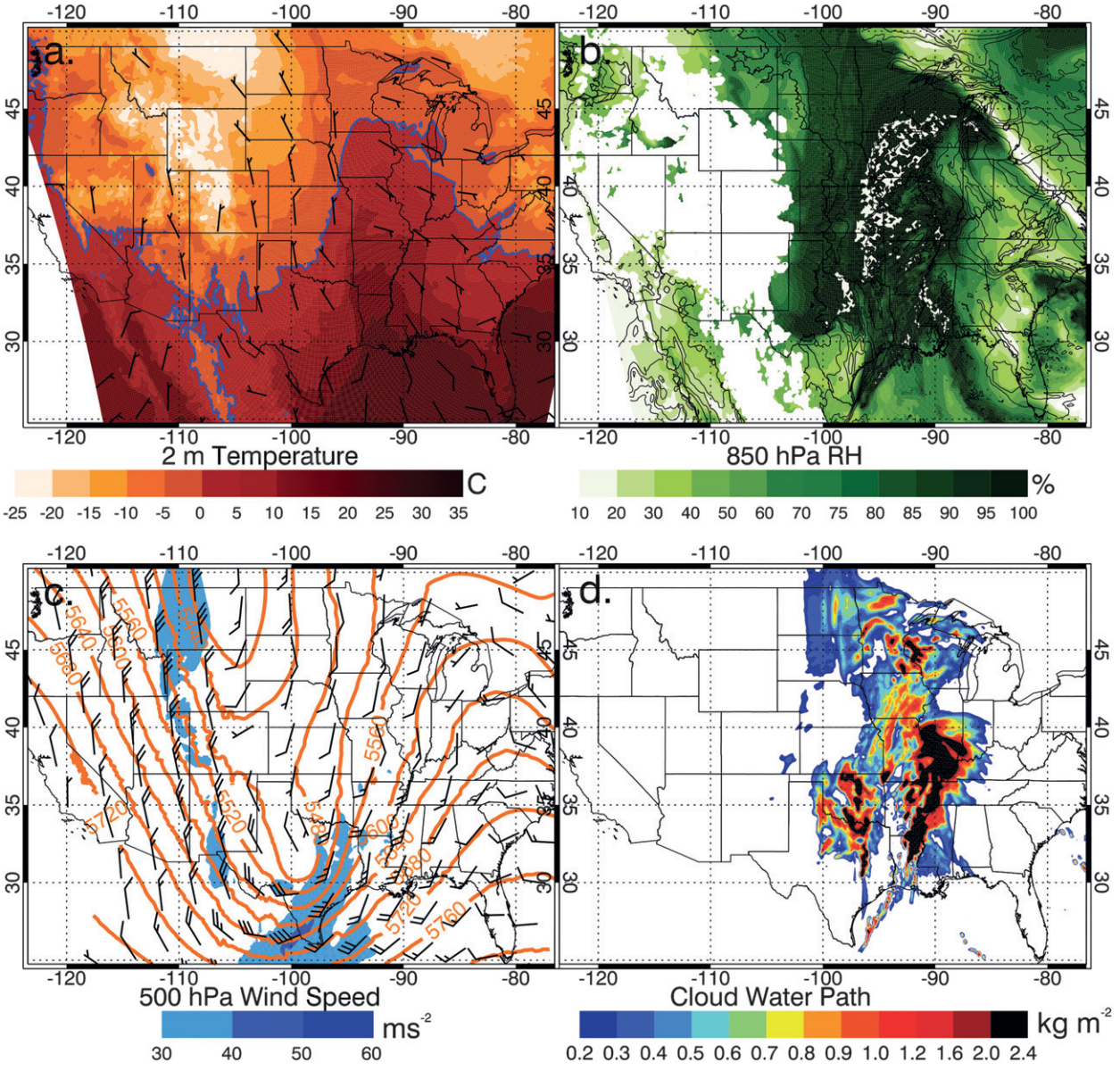

FIG. 2. Truth simulation at 1200 UTC 24 Dec 2009 for (a) 2-m temperature, (b) 850-hPa relative humidity, (c) 500-hPa heights and wind speed, and (d) cloud water path.

processes. The Noah land surface model is used to compute surface heat and moisture fluxes. The truth simulation is used as the comparison dataset to evaluate the skill of each experiment during the assimilation period.

Atmospheric conditions from the truth simulation are displayed in Fig. 2. This event is characterized by a significant winter storm in the southern plains that deposited over 15 cm of snow in many locations. At 1200 UTC, ongoing liquid precipitation transitions to snow as a result of strong cold air advection (Fig. 2a). Farther east in Louisiana and Arkansas, much warmer conditions exist with southerly winds advecting warm, moist air from the Gulf of Mexico, leading to strong convective activity. In both regions, the 850-hPa relative humidity often exceeds 90% indicating that ample moisture is present for significant precipitation to occur (Fig. 2b). The large-scale forcing for this system is provided by a deep 500-hPa trough across the central United States producing strong northwesterly flow from Canada all the way into central Texas (Fig. 2c). The high moisture content and strong large-scale forcing led to substantial cloud cover within and ahead of the trough, as shown by the high cloud water path values ($>1.0\,\mathrm{kg\,m^{-2}}$) (Fig. 2d). The wide

variety of weather conditions present during this event provides a unique opportunity to examine the relative impacts of satellite and radar data assimilation.

3. Simulated observations

a. Conventional

Simulated observations are generated from truth for a variety of observation types representing conventional, satellite, and radar data. Conventional observations, considered to be those most commonly assimilated into NWP models, include those from Automated Surface Observing System (ASOS), Aircraft Communications Addressing and Reporting System (ACARS), and radiosonde instruments. For each ASOS location, 10-m wind speed and direction, 2-m temperature and humidity, and surface pressure observations are generated from the truth simulation at 5-min intervals. Similarly, vertical profiles of temperature, humidity, and wind speed and direction are created for each radiosonde location at standard launch times of 0000 and 1200 UTC. Finally, simulated ACARS temperature and wind observations are generated along real flight tracks reported in the Meteorological Assimilation Data Ingest Files (MADIS). Observation measurement errors for each observation are drawn from uncorrelated Gaussian error distributions that are based on a given sensor's accuracy.

b. Satellite

Simulated infrared radiances are computed using the Successive Order of Interaction (SOI) forward radiative transfer model developed by Heidinger et al. (2006) and are converted to brightness temperatures (T_B) using a mathematical transformation. The SOI model uses the CompactOPTRAN code from the Community Radiative Transfer Model (CRTM; Han et al. 1995) to compute gas optical depths for each model layer. Absorption and scattering properties, such as the full scattering phase function, single scatter albedo, and extinction efficiency, for each frozen hydrometeor species (i.e., ice, snow, and graupel) are obtained from Baum et al. (2005). Lorenz–Mie calculations use a lookup table approach for the liquid species (i.e., cloud water and rainwater). Visible cloud optical depths are calculated for each microphysical species based on Han et al. (1995) and Heymsfield et al. (2003), and then converted into infrared cloud optical depths by scaling the visible optical depths by the ratio of the extinction efficiencies. Cloud-top pressure is not treated as an individual parameter. The surface emissivity for each ABI infrared band is obtained from the global emissivity database for land grid points (Seaman et al. 2010). For water points, the CRTM Infrared Sea Surface Emissivity Model computes the surface emissivity. WRF data used by the SOI model to compute simulated T_B include the water vapor mixing ratio, atmospheric temperature, surface skin temperature, 10-m wind speed, and the mixing ratios for each hydrometeor species predicted by the microphysics parameterization scheme. Previous studies have shown that the SOI model computes accurate infrared T_B for both clear- and cloudy-sky conditions (Otkin and Greenwald 2008; Otkin et al. 2009).

Simulated ABI 6.95-μm T_B are generated using the SOI forward radiative transfer model. This band is sensitive to water vapor content in the middle and upper troposphere with the peak of the vertical weighting function occurring near 6 km for clear-sky conditions (Fig. 3a). This value varies somewhat as a function of total water vapor content and cloud height. A weighting function profile specifies the relative contribution from each atmospheric layer to the radiation emitted to space and thereby determines those regions of the atmosphere that are sensed at a given wavelength. The ABI observations are computed on the 6-km truth grid, which are then averaged to a 30-km resolution prior to assimilation. Averaged observations are only used if all of the grid points within each averaging area are either clear or cloudy. Binary cloud fraction was used to avoid representativeness issues associated with a partly cloudy scene. Cloudy grid points are defined as those whose cloud optical thickness (COT) for any of the predicted microphysical species greater than zero. The observation error for the 6.95-μm T_B is set to 5 K for both clear-sky and cloudy-sky observations accounting for uncertainties in the RTM forward operator as well as those associated with model and representativeness errors. This error value was chosen based on prior results from sensitivity tests performed for this case study (Otkin 2012b). For verification purposes, simulated 6.95-μm T_B are generated from the truth and ensemble mean experiment analysis fields at 15-km resolution.

c. Radar

Synthetic WSR-88D reflectivity and radial velocity observations are generated from the truth simulation for 13 radar locations in the south-central United States (Table 2) at 5-min intervals from 1100 to 1200 UTC 24 December 2009 using the radar simulator program contained within the Data Assimilation Research Testbed (DART) software (Anderson et al. 2009). Observations are obtained up to a height of 15 km using the volume coverage pattern (VCP) 21 WSR-88D scan strategy commonly employed during convective episodes with a 6° azimuthal increment (Fig. 3b) (Crum

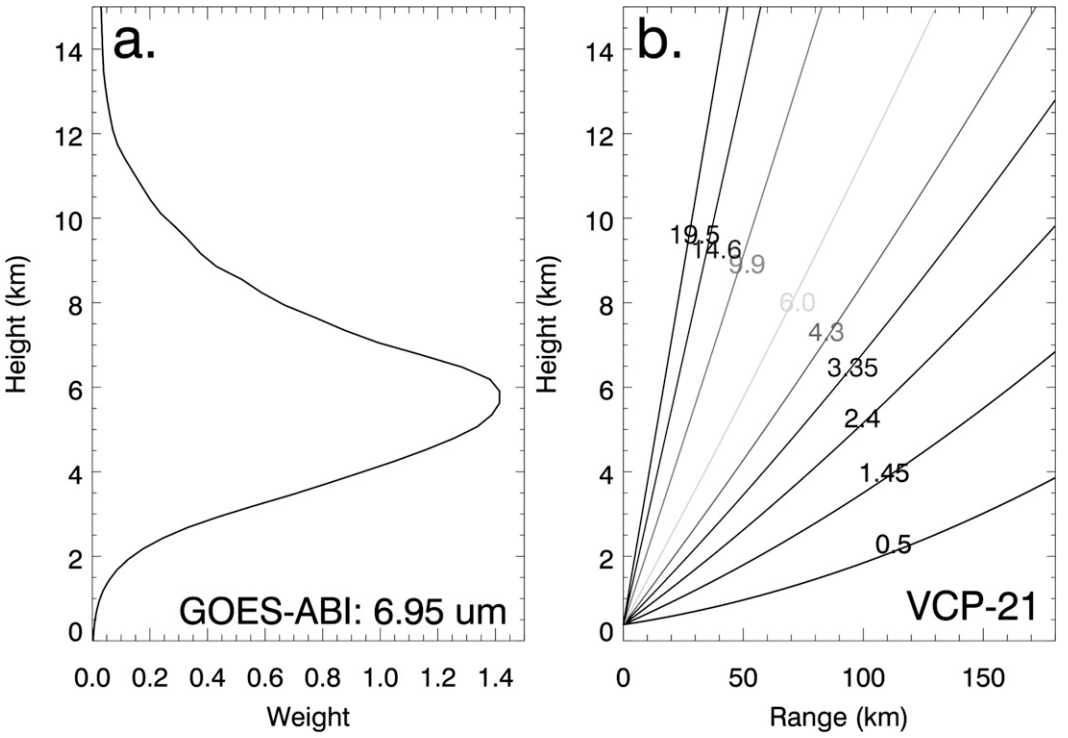

FIG. 3. (a) Simulated GOES-R ABI 6.95-μm vertical weighting function for clear-sky radiances as a function of height. (b) WSR-88D VCP-21 elevation scans as a function of height and range from a radar.

and Alberty 1993). Simulated radar gate length is set at 15 km, with the closest and farthest gates at 3 and 240 km, respectively. The entire radar volume is simulated at 5-min intervals; thus, each represents one snapshot of the entire 3D distribution of precipitation in the atmosphere. The radar simulator computes reflectivity and precipitation fall speed, which is needed to calculate the radial velocity, based on parameters consistent with the Lin (Lin et al. 1983) microphysics scheme. To account for typical observations errors, the error for reflectivity and radial velocity are set to 5 dBZ and 2 m s^{-1}, respectively.

4. Model configurations and verification

A 48-member WRF ensemble with perturbed initial and lateral boundary conditions is created using the methodology described by Torn et al. (2006). Initial and boundary conditions are taken from Global Forecast System (GFS) analysis fields with perturbations applied at 6-h intervals starting at 0000 UTC 23 December 2009. The assimilation experiments employed the same vertical resolution (52 levels) as the truth simulation, but were performed over a smaller spatial domain containing 15-km horizontal grid spacing (272 $\times$ 216 grid points) (Fig. 1). As with the truth simulation, the WRF experiments use the WSM6 mixed-phase microphysics scheme, the Yonsei University planetary boundary layer scheme (Hong et al. 2006), the Kain–Fritsch cumulus parameterization scheme (Kain and Fritsch 1993; Kain 2004), the RRTM longwave and shortwave radiation schemes (Iacono et al. 2008), and the Noah land surface model. The control variables used during each assimilation cycle include temperature, specific humidity, zonal and meridional wind, condensational heating, and cloud microphysical variables at each model level (Table 3).

TABLE 2. WSR-88D locations and elevations (AGL) where simulated radar reflectivity and radial velocity data are generated from the truth simulation.

Location	Name	Lon (°W)	Lat (°N)	Elev (m)
Fredrick, OK	KFDR	98.9764	34.3622	364
Little Rock, AR	KLZK	92.2622	34.8364	193
Amarillo, TX	KAMA	101.709	35.2333	1113
Lubbock, TX	KLBB	101.814	33.6542	1013
Dyess AFB, TX	KDYX	99.2544	32.5383	477
Fort Worth, TX	KFWS	97.3031	32.5731	228
Twin Lakes, OK	KTLX	97.2778	35.3331	384
Vance AFB, OK	KVNX	98.1278	36.7408	379
Tulsa, OK	KINX	95.5647	36.1750	224
Fort Smith, AR	KSRX	94.3617	35.2906	215
Shreveport, LA	KSHV	93.8414	32.4508	113
Springfield, MO	KSGF	93.4006	37.2353	414
Fort Polk, LA	KPOE	92.9758	31.1556	139

TABLE 3. List of control variables used during the assimilation experiments. The terms "2M" and "10M" refer to 2 m and 10 m above ground level, respectively.

	Variable name
1.	U_WIND_COMPONENT
2.	V_WIND_COMPONENT
3.	U10M_WIND_COMPONENT
4.	V10M_WIND_COMPONENT
5.	VERTICAL_VELOCITY
6.	GEOPOTENTIAL_HEIGHT
7.	POTENTIAL_TEMPERATURE
8.	POTENTIAL_TEMPERATURE_2M
9.	TEMPERATURE_2M
10.	PRESSURE
11.	SURFACE_PRESSURE
12.	SKIN_TEMPERATURE
13.	VAPOR_MIXING_RATIO
14.	VAPOR_MIXING_RATIO_2M
15.	CLOUD_LIQUID_WATER
16.	RAINWATER_MIXING_RATIO
17.	CLOUD_ICE
18.	SNOW_MIXING_RATIO
19.	GRAUPEL_MIXING_RATIO
20.	SOIL_TEMPERATURE
21.	SOIL_MOISTURE
22.	CONDENSATIONAL_HEATING

The ensemble is allowed to freely evolve starting at 0000 UTC 23 December and continues until 0900 UTC 24 December. No observations are assimilated during this time, thereby providing an initial analysis for the assimilation experiments that is sufficiently different from truth. Beginning at 0900 UTC, simulated conventional temperature, wind, and surface pressure observations are assimilated at 5-min intervals until 1100 UTC 24 December using a horizontal (vertical) covariance localization half radius of 300 (6) km with a fifth-order correlation function following Gaspari and Cohn (1999). For humidity observations, the horizontal localization radius is set to 250 km to account for the smaller spatial scale of the moisture structures. The assimilation of conventional observations during this time period ensures that the ensemble covariances are realistic at the beginning of the assimilation experiments starting at 1100 UTC.

It is necessary to ensure that the 1100 UTC ensemble mean initial conditions differ sufficiently from truth before proceeding with these experiments. Figure 4 shows the differences (ensemble mean − truth) for 500-hPa temperature (T500), water vapor mixing ratio (QV500), total cloud water path (QA500), and 850-hPa cloud water path (QA850). The total cloud water path (QALL) is defined as the summation of cloud liquid water, cloud ice, graupel, snow, and rain mixing ratio variables from the model analysis (Otkin 2010). This variable is created to provide a measure of the effects of assimilating both satellite and radar data on the bulk characteristics of cloud properties. It is clear that significant differences exist between the initial condition and truth. The 1100 UTC ensemble mean analysis has a large warm bias exceeding 2 K in the northwestern quadrant of the verification domain with a similar, but slightly smaller in magnitude, cold bias in the southwest (Fig. 4a). Farther east, another region of warm biases exists from eastern Texas northward through central Arkansas. Several differences in the water vapor analyses are also apparent with moist biases on the order of 1 g kg^{-1} being present in central Texas and Oklahoma as well as over most of Arkansas (Fig. 4b). Dry biases exist in between and in the far eastern portions of the domain. Similar differences exist in the QALL fields at 500 hPa with a large area of greater cloud water content (thicker clouds) in the analysis compared to truth located in central Arkansas (Fig. 4c). Farther east, the analysis appears to have less cloud cover indicating a displacement in the convection with the truth and analysis fields at this time. Additional differences are present farther west in Oklahoma and central Texas. These differences are also apparent at lower in the atmosphere at 850 hPa where they are present over a large portion of the domain (Fig. 4d). The differences between the truth and analysis fields indicate that the experiment design is sufficient for assessing the impact of assimilating simulated remote sensing observations.

Beginning at 1100 UTC, four assimilation experiments are initiated, with observations assimilated every 5 min until 1200 UTC. The assimilation frequency was chosen based on results from previous studies that indicated that approximately 1 h of data are required to accurately reproduce precipitation features within a model analysis (e.g., Snyder and Zhang 2003; Caya et al. 2005; Xue et al. 2006). The first experiment represents the control case (CONV) in which only conventional observations are assimilated (Table 1). The selection of this experiment as the "control" is based on several factors. For example, the need to assimilate conventional variables between 0900 and 1100 UTC suggests that continuing to assimilate these variables would be desirable. Furthermore, conventional observations are going to be part of any operational system; thus, it is important to show that the remote sensing observations show increased skill over using conventional observations alone. The following experiments assimilate simulated ABI 6.95-μm T_B across the entire domain (SAT) and WSR-88D reflectivity and radial velocity from 13 radar locations across the southern plains (RAD), in addition to conventional observations. The horizontal localization radii are set to 50 and 25 km, respectively, for the satellite and radar observations (Table 2). Vertical

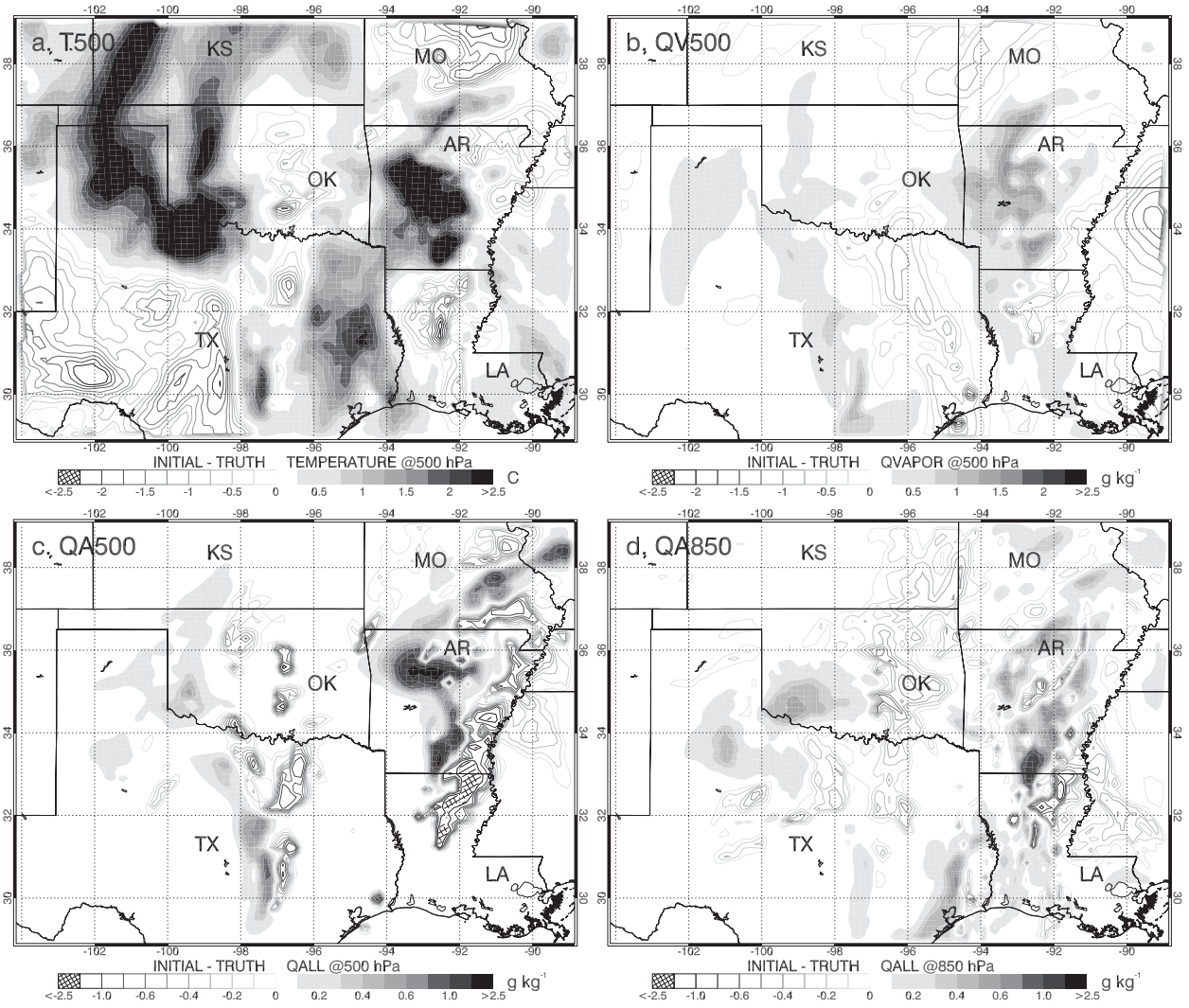

FIG. 4. Difference between 1100 UTC ensemble mean initial conditions and corresponding truth analysis for (a) 500-hPa temperature (T500), (b) water vapor mixing ratio (QV500), (c) total cloud water content (QA500), and (d) 850-hPa total cloud water content (QA850). Shaded areas indicate where the experiment is warmer (or too moist) than the truth and contour lines indicate where it is colder (or too dry).

localization does not apply to the satellite observations given that they are sensitive to broad layers of the atmosphere (Otkin 2012a). Assimilating the full CONUS-wide complement of radars proved too resource intensive, so the subset of radars corresponding to areas of precipitation are selected for assimilation. To differentiate the effects of assimilating radar reflectivity versus radial velocity, each is also assimilated separately in individual experiments. The final experiment (RADSAT) combines both the satellite and all radar data with the conventional observations to assess the combined impacts of assimilating all observation types.

The accuracy of each experiment is evaluated by comparing the ensemble mean analysis after the final assimilation cycle at 1200 UTC with the corresponding truth simulation. The truth analyses are coarsened to 15-km resolution using a simple average of the 6-km data for comparison with the assimilation results. The validation focuses on a subset of the CONUS domain that includes the region where radar data are assimilated (Fig. 1) since the radar data should have little impact on the model analysis outside of this region. Satellite data are assimilated over the full domain; thus, improvements made by these observations outside the subdomain used for validation are not fully accounted for. Although this reduces the apparent impact of the satellite observations, this is acceptable since our goal is to compare the relative impacts of assimilating radar and

satellite data where both exist. Simple statistics such as bias (experiment − truth) and root-mean-square error difference (RMSD) are computed between truth and ensemble mean analysis fields such as temperature, water vapor, and wind speed and direction. Skill scores, including probability of detection (POD), false alarm rate (FAR), and the Heidke skill score (HSS), are also computed (Wilks 2006).

5. Assimilation statistics

A total of approximately 2.05 million observations are assimilated during the RADSAT experiment over the entire domain, which decreases to 1.86 million when only the verification domain is considered (Table 4). Conventional observations represent the lowest proportion of observations assimilated, accounting for only 0.45% and 0.71% of the total for experiment and verification domains, respectively. The distribution of conventional observations assimilated between 1100 and 1200 UTC within the verification domain is shown in Fig. 5a. Radiosonde and surface observations are relatively sparse, with a slightly larger number of ACARS observations present. However, the ACARS observations are primarily limited to a few individual flight tracks or centered around major airports. Compared to conventional observations, many more satellite and radar observations are assimilated during this period (Table 4; Fig. 5b). Simulated satellite T_B (averaged to 30-km horizontal resolution) account for 1.0% of the observations being assimilated into the RADSAT experiment in the verification domain, but account for 59% of the data in the SAT experiment when no radar observations are assimilated. Approximately 80% of the satellite observations assimilated within the verification domain are classified as cloudy.

Radar reflectivity and radial velocity observations are by far the most numerous observation type assimilated, accounting for ~92% and 98% of the observations for experiment and verification domains, respectively. This is due to the high vertical and horizontal resolution of the radar observations and the inclusion of two radar variables (radial velocity and reflectivity). The much larger number of radar observations assimilated in the RAD and RADSAT experiments compared to those from satellite data strongly suggests that the radar data will have a much larger impact on the model analysis. This increases the difficultly in comparing the relative impact of the radar and satellite observations; however, thinning the radar data to a comparable resolution produces an unrealistic dataset. It should also be noted that the impact of observations is not solely a function of their number, but also the number of model grid points that they update. The larger radius of the satellite data allows it to update a greater number in the horizontal, but radar data has the potential to affect many more model levels. Since these experiments are designed to mimic observations generated under real-world conditions, the higher resolution of radar data products compared to currently available satellite products is unavoidable. Further research is under way to create and test higher-resolution satellite observations to reduce the gap between radar and satellite sample sizes.

TABLE 4. Number of assimilated observations in the RADSAT experiment for the entire model domain N_2 and within the verification domain N_3 only. Numbers correspond to the domain labels in Fig. 1.

Observation type	N_2	N_3
SIM_RAWIN_U_WIND_COMPONENT	1355	283
SIM_RAWIN_V_WIND_COMPONENT	1337	279
SIM_RAWIN_TEMPERATURE	1410	288
SIM_RAWIN_VAPOR_MIXING_RATIO	1420	289
SIM_SFC_U_WIND_COMPONENT	10 153	1820
SIM_SFC_V_WIND_COMPONENT	10 153	1820
SIM_SFC_TEMPERATURE	10 137	1820
SIM_SFC_VAPOR_MIXING_RATIO	10 153	1820
SIM_SFC_SURFACE_PRESSURE	8553	1775
SIM_ACARS_U_WIND_COMPONENT	6714	986
SIM_ACARS_V_WIND_COMPONENT	6711	985
SIM_ACARS_TEMPERATURE	6712	987
DOPPLER_RADIAL_VELOCITY	937 772	935 882
RADAR_REFLECTIVITY	897 884	895 998
ABI_BAND09_INFRARED_ T_B	148 677	18 658
TOTAL	2 059 141	1 863 690

6. Assimilation characteristics

a. Observation diagnostics

To assess the viability of assimilating simulated satellite and radar data into these model experiments, observation-space diagnostics (e.g., Dowell et al. 2011) including bias, RMSD error, and ensemble spread (SPRD) are calculated for 6.95-μm T_B, radar reflectivity, and radial velocity between 1100 and 1200 UTC at each 5-min assimilation cycle for all assimilated observations over the entire experiment domain. For reference, the observation error is not included as part of the ensemble spread variable and RMSD calculations include biases present. The latter is consistent with previous methodology by Dowell et al. (2011) and Yussouf and Stensrud (2012) among others to determine the effectiveness of ensemble assimilation process. Results are shown for the RADSAT experiment that assimilates all three observation types. Error and spread for satellite and radar observations individually within SAT and RAD are similar. Roughly

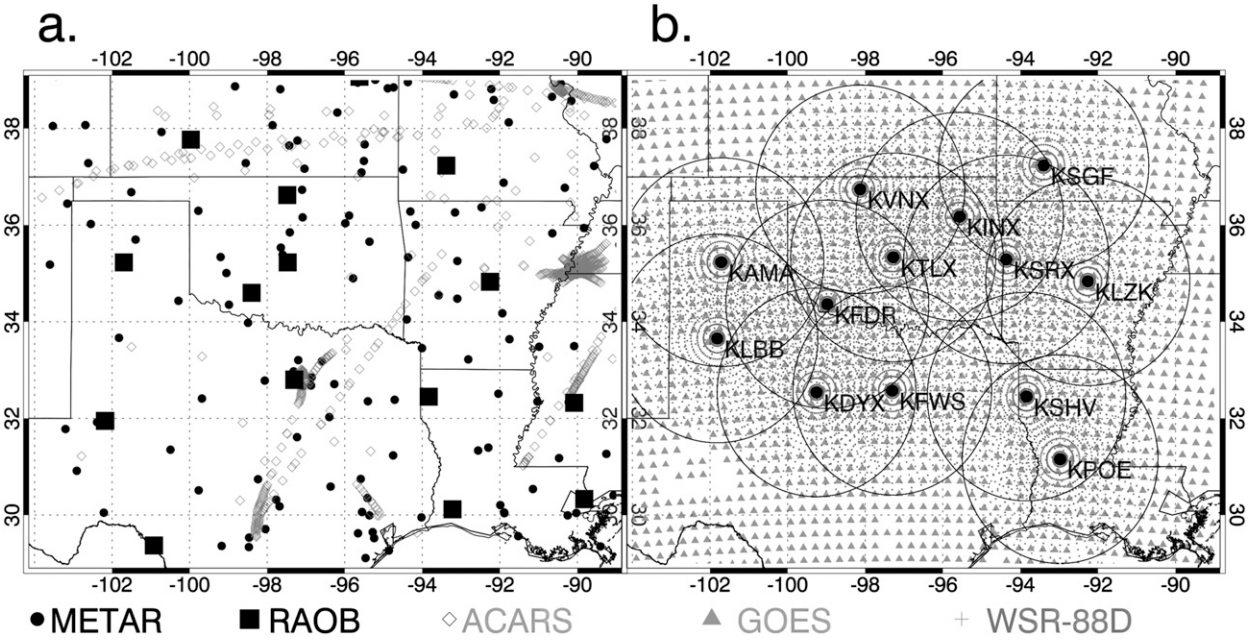

FIG. 5. Location of assimilated (a) conventional and (b) satellite and radar observations on 1200 UTC 24 May. Rings indicate 240-km radii from each WSR-88D location representing the maximum range for which radar data are simulated and assimilated.

11 000 satellite observations are assimilated during each cycle with an initial RMSD of 3.0 K, SPRD of 2.4 K, and a bias of −0.4 at 1100 UTC prior to the first assimilation (Fig. 6a). These values drop to 2.1, 1.8, and −0.4 K in the posterior analysis at the same time. The classical sawtooth pattern follows in each assimilation cycle with a general decrease in RMSD and SPRD occurring out to 1130 UTC. After this time, RMSD ranges between 1.3 and 1.5 K with SPRD ranging between 1.2 and 1.4 K. Bias remains slightly negative near −0.25 K. The ratio between the ensemble spread and RMSD stabilizes after a few assimilation cycles to a value of approximately 0.85.

The assimilation characteristics of radar reflectivity and radial velocity are similar to those present for satellite T_B. At 850 hPa, RMSD for reflectivity rapidly decreases from 7.9 dBZ in the 1100 UTC prior fields to 1.7 dBZ in the posterior fields by 1120 UTC (Fig. 6b).

The number of observations slowly increases as a function of time from 5300 to 6200 as the prior and posterior reflectivity become more similar, thereby reducing the number of rejected observations. (Note that radar reflectivity data are being assimilated throughout the atmospheric column, not just the 850-hPa level shown here for reference.) Bias ranges between 0.0 and 0.5 dBZ with SPRD being somewhat lower than RMSD throughout the assimilation period. Assimilating radial velocity has the greatest impact early in the assimilation cycle quickly reducing prior and posterior RMSD from 4.0 and 1.7 m s^{-1} to 1.7 and 1.4 m s^{-1} by 1120 UTC (Fig. 6c). SPRD is lower ranging from 0.7 m s^{-1} in prior fields to 0.4 m s^{-1} in posterior fields while bias converges to slightly below zero after the initial assimilation cycle. The diagnostics for radar data at other levels are consistent with the 850-hPa data shown here except that the number of observations being assimilated decreases as a function of height. The somewhat lower SPRD compared to RMSD for the radar reflectivity and radial velocity observations indicates that the ensemble spread is underdispersive; however, no ensemble divergence is observed during the assimilation. The effects of assimilating these data on model state variables such as temperature, humidity, and wind velocity are examined below for each assimilation cycle and experiment.

b. Bias and RMSD between 1100 and 1200 UTC

Bias and RMSD are calculated between the truth and experiment output at each assimilation cycle within the verification domain to assess the effects of assimilating radar and satellite data as a function of time for 500-hPa temperature (T500), water vapor mixing ratio (QV500), zonal and meridional wind components (U500, V500), and total cloud water content (QA500). RMSD for T500 begins at 1.2 K with a corresponding positive bias of 0.37 K for the CONV experiment with only a slight decrease out to 1155 UTC (Fig. 7a). SAT generally lowers the bias and RMSD by up to 0.2 K compared to CONV, with the improvement increasing with time. The bias and RMSD are also reduced during the RAD experiment, but

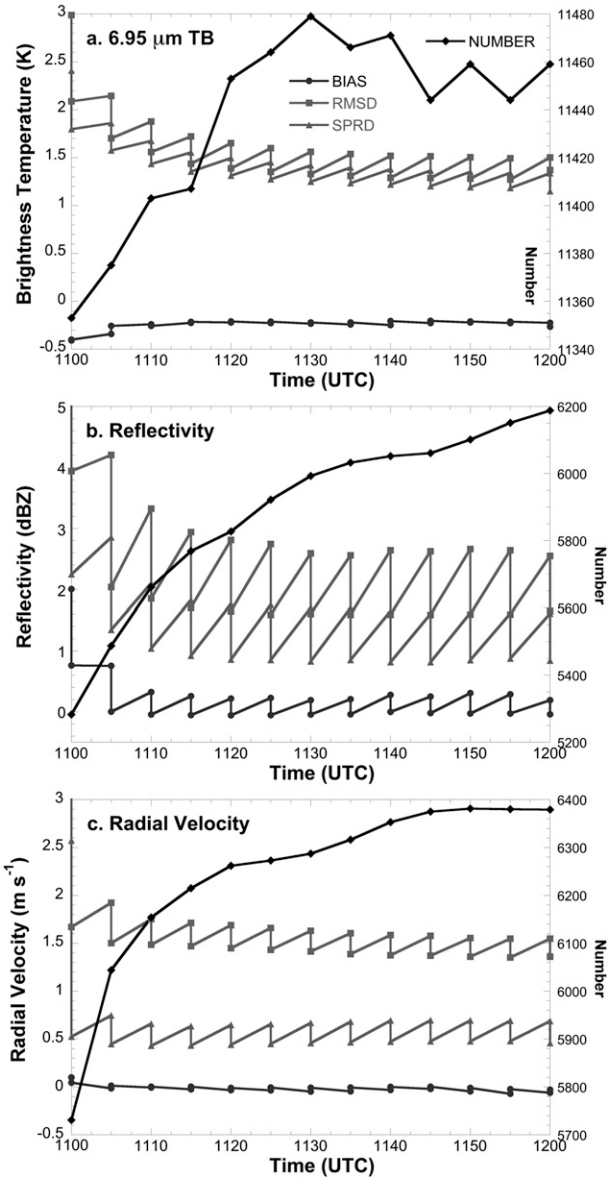

FIG. 6. Bias, RMSD, and SPRD for (a) 6.95-μm T_B, (b) 850-hPa radar reflectivity, and (c) 850-hPa radar radial velocity for prior and posterior analysis fields at 5-min intervals between 1100 and 1200 UTC. For all panels, "number" refers to the number of observations assimilated during each cycle.

the effects are generally smaller than SAT. In RADSAT, the error characteristics are similar to RAD and generally lie between SAT and RAD with respect to RMSD. A significant reduction in both bias and RMSD occurs at 1200 UTC due to the assimilation of vertical temperature profiles from simulated radiosonde observations. The assimilation of simulated radiosonde observations at 1200 UTC decreases both bias and RMSD more than assimilating either the satellite and radar observations despite the limited radiosonde observation sample size.

Larger differences between model experiments are apparent in the moisture, wind, and cloud water fields. The CONV experiment produces a bias and RMSD of -0.08 and $0.5\,\mathrm{g\,kg^{-1}}$ at 1100 UTC whose magnitude slowly decreases as a function of time (Fig. 7b). Both satellite and radar data reduce RMSD compared to CONV, with the RADSAT experiment consistently characterized by the lowest RMSD for all assimilation cycles. Bias for RAD and RADSAT closely follow CONV with SAT having a slightly larger dry bias by 1200 UTC. The effects of the radiosonde assimilation are still noticeable at 1200 UTC, but are much less significant than for temperature indicating that both the radar and satellite data are providing valuable information that augments conventional observations.

The effects of assimilating satellite data on the wind fields is comparatively small while the effect of assimilating radar data, primarily radial velocity observations, is quite large (Fig. 7c). For U500, RMSD exceeds $4.0\,\mathrm{m\,s^{-1}}$ at all cycles in CONV with SAT being almost identical. However, SAT introduces negative zonal biases of increasing magnitude that changes its sign from negative to positive with the difference exceeding $1.0\,\mathrm{m\,s^{-1}}$ after 1130 UTC. The RAD and RADSAT experiments reduce RMSD by up to $2.0\,\mathrm{m\,s^{-1}}$ compared to CONV primarily through the assimilation of radial velocity. An experiment assimilating only radar reflectivity generates wind velocity errors similar to those shown for CONV (not shown). The error characteristics for V500 are similar except that the magnitude of bias and RMSD are somewhat larger and SAT appears to perform slightly worse than CONV after 1140 UTC (Fig. 7d).

The final variable considered is QA500, which represents the summation of all liquid and frozen cloud mixing ratio variables at 500 hPa (Fig. 7e). Assimilating conventional observations has minimal effect on QA500, with a nearly constant RMSD and bias during the assimilation period. Compared to the CONV case, the SAT experiment reduces RMSD up to $0.05\,\mathrm{g\,kg^{-1}}$ with smaller improvements occurring in the bias. Much larger improvements occur in the RMSD and bias during the RAD experiment with RMSD dropping nearly $0.2\,\mathrm{g\,kg^{-1}}$ from its initial 1100 UTC value. RADSAT consistently produces the lowest RMSD for assimilation cycles with a bias similar to that of RAD, which indicates that both remote sensing datasets provide independent and complementary information.

7. Analysis verification at 1200 UTC

a. Conventional variables

Model output from each experiment at 1200 UTC is compared with the corresponding truth simulation to

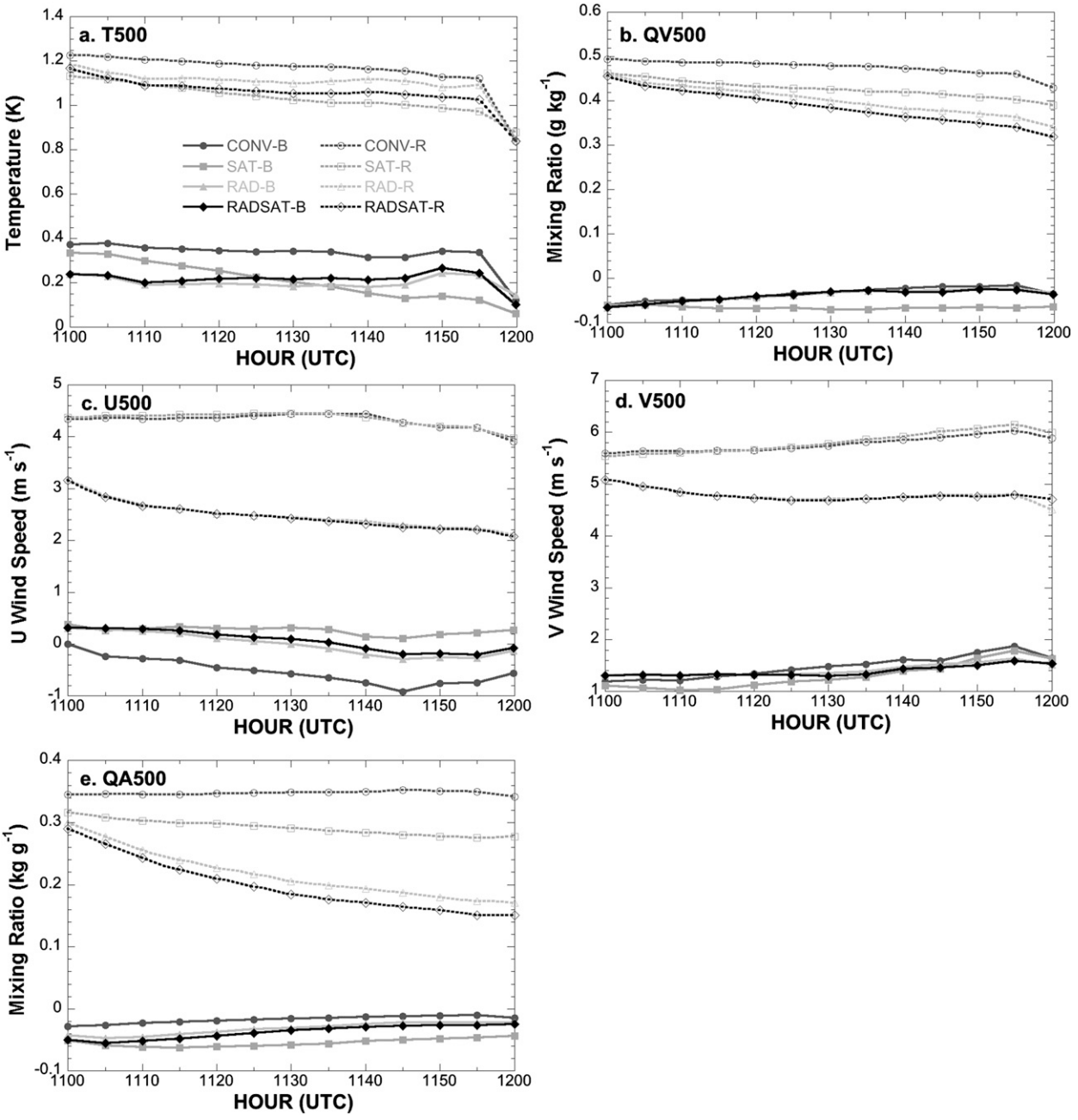

FIG. 7. Time series of bias (B) and RMSD (R) for each experiment between 1100 and 1200 UTC at 5-min intervals following the assimilation of observations at each cycle for (a) 500-hPa temperature (T500), (b) water vapor mixing ratio (QV500), (c) zonal wind speed (U500), (d) meridional wind speed (V500), and (e) total cloud water (QA500).

determine how assimilating various combinations of conventional, satellite, and radar observations affect the analysis accuracy. We compare temperature, water vapor mixing ratio, and wind speed. Figure 8 shows the difference in 500-hPa temperature between truth and each experiment. The CONV experiment contains warm biases in several regions, including Arkansas, east Texas, and the Texas Panhandle, with cold biases located elsewhere. The temperature field is not significantly improved in SAT (Fig. 8b), with the general location and magnitude of the warm and cold bias regions remaining similar to CONV. In contrast to SAT, the temperature difference field is noticeably different when radar observations are assimilated during the RAD experiment (Fig. 8c). However, these differences are small and do not result in improved bias and RMSD values compared to CONV (Fig. 8a).

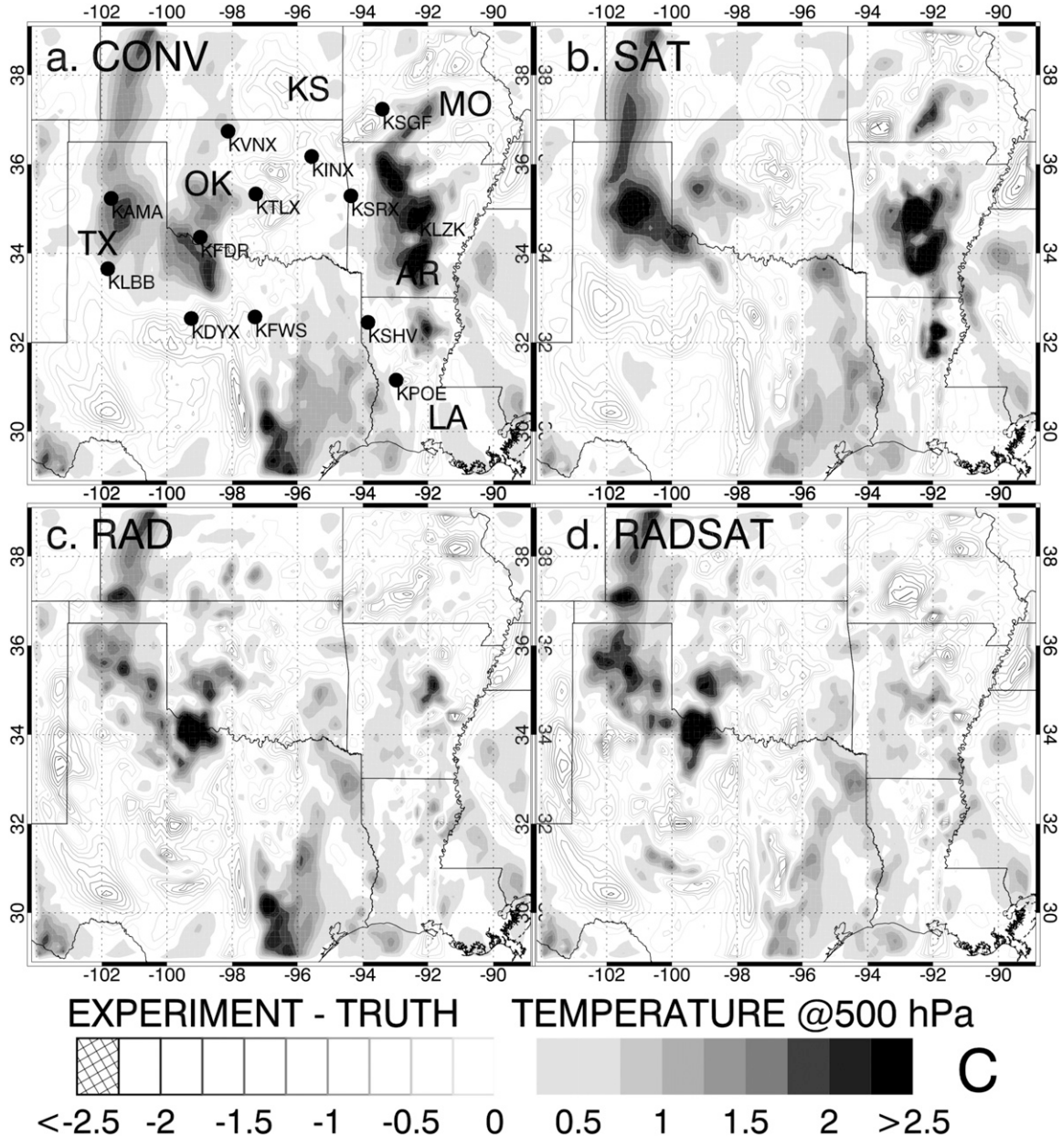

FIG. 8. (a)–(d) The 500-hPa temperature difference at 1200 UTC 24 Dec 2009 between model analysis and the truth simulation for each experiment (experiment − truth). Shaded areas indicate where the experiment is warmer than the truth and contour lines indicate where it is colder. WSR-88D locations are also shown in (a).

Differences between truth and experiment temperature fields are not confined to the 500-hPa level, but are present throughout the entire troposphere. All experiments have a warm bias below 800 hPa with SAT being slightly warmer than the other experiments (Fig. 9a). This is followed by a cold bias layer located between 750 and 550 hPa with SAT slightly cooler than the other experiments with alternating cold and warm biases in the upper troposphere. Above 800 hPa, the magnitude of the biases is less than ±0.5 K. RMSD errors decrease with height up to 500 hPa before increasing at higher levels. For most levels, SAT has the largest RMSD, with CONV, RAD, and RADSAT all being similar with each containing smaller errors at certain levels. Recall that sounding data provide much of the temperature information in CONV and the other experiments at this time; thus, only relatively small differences are apparent.

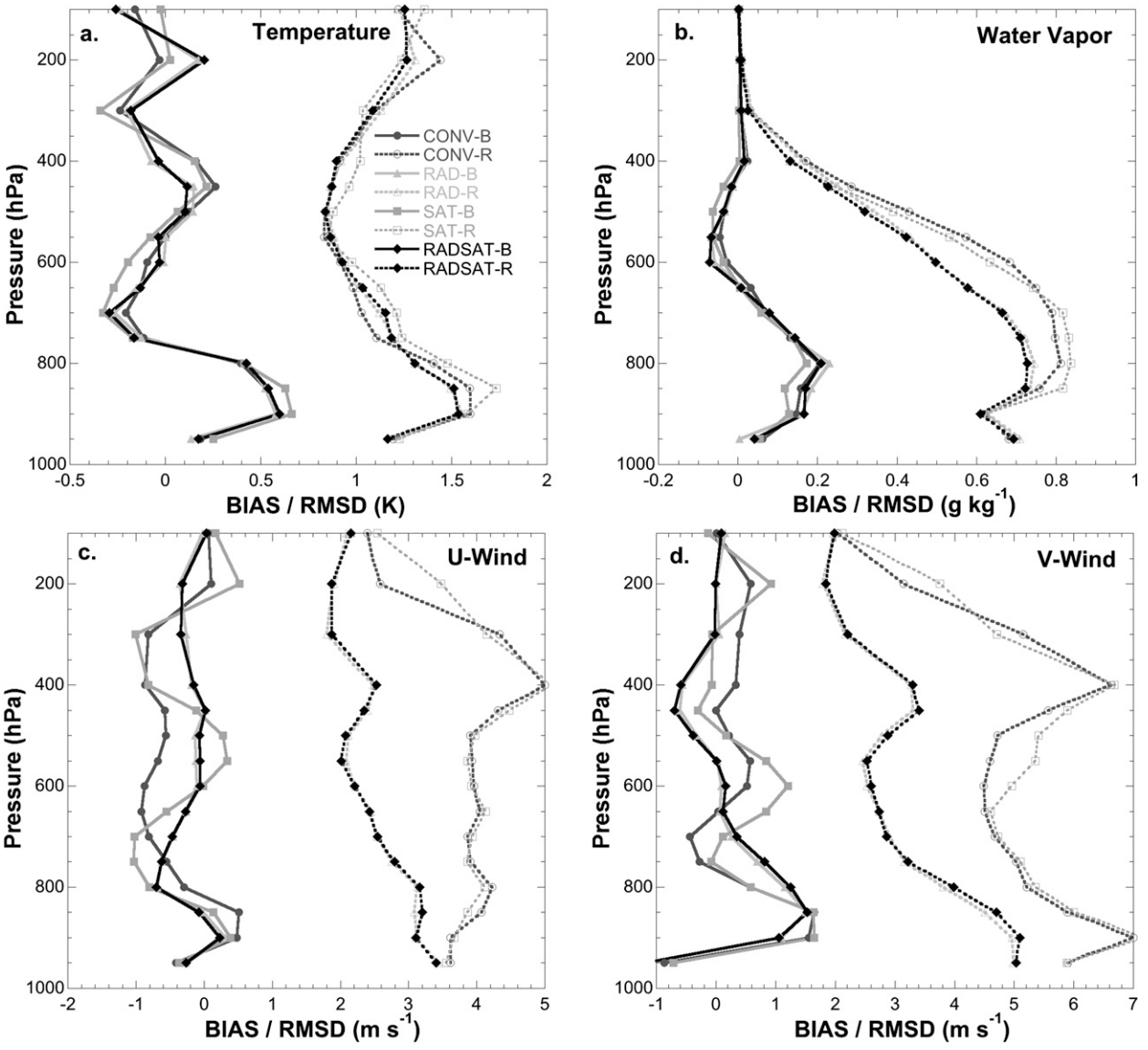

FIG. 9. Vertical profiles of bias (B) and RMSD (R) for (a) temperature, (b) water vapor mixing ratio, and (c) zonal and (d) meridional wind components computed between truth and each experiment at 1200 UTC.

Comparing the 500-hPa water vapor mixing ratio from each experiment indicates that assimilating satellite data has a greater impact on the moisture fields compared to temperature. The difference in mixing ratio between CONV and truth is generally smaller ($<0.5\,g\,kg^{-1}$) in the western half of the domain with somewhat larger differences to the east (Fig. 10a). Assimilating satellite data reduces the magnitude of the differences especially in the eastern portion of the domain (Fig. 10b). Although the bias increases in SAT, the RMSD decreases from 0.43 to $0.39\,g\,kg^{-1}$. Radar data assimilation also improves the moisture analysis with the spatial extent of the moist and dry bias regions becoming smaller with a further reduction in the RMSD to $0.34\,g\,kg^{-1}$ (Figs. 9b and 10c). The combination of radar and satellite data assimilation in RADSAT produces the most accurate analysis (Fig. 10d), with a 25% reduction in the RMSD compared to CONV (Fig. 9b). All experiments show a moist bias of up to $0.2\,g\,kg^{-1}$ between the surface and 700 hPa with a smaller dry bias above (Fig. 9b). SAT reduces the magnitude of the bias in this lower layer, but also increases RMSD slightly compared to CONV. Above 700 hPa, SAT lowers RMSD compared to CONV. The RAD experiment consistently reduces RMSD compared to CONV from 900 hPa up to at least 400 hPa above which moisture concentrations are small. RADSAT

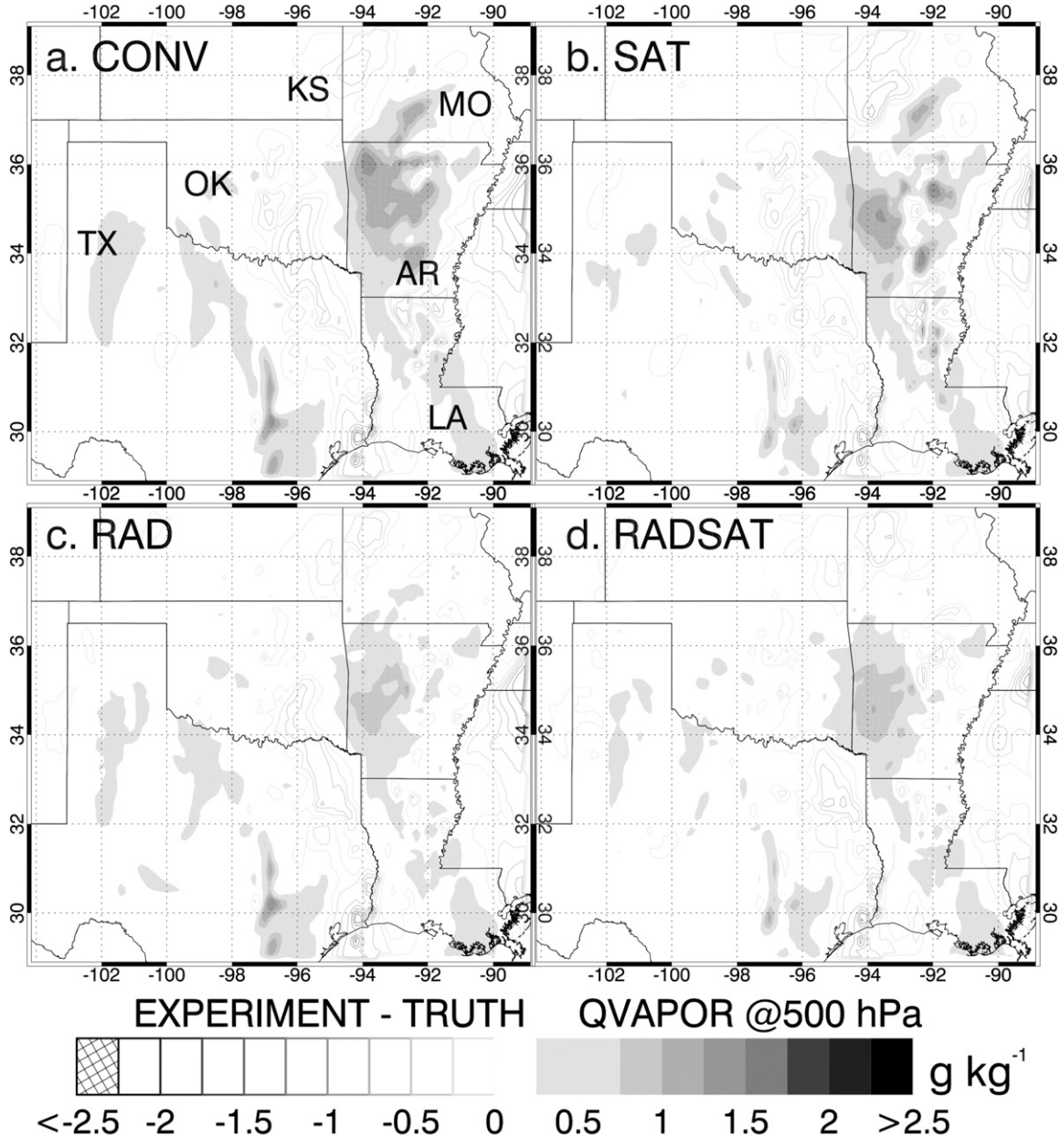

FIG. 10. As in Fig. 8, but for 500-hPa water vapor mixing ratio.

performs similarly to RAD though slightly lower RMSD values are evident at most levels. Most of the improvement in water vapor mixing ratio error can be attributed to assimilating radar reflectivity rather than radial velocity because it is more strongly correlated to atmospheric moisture content.

For wind speed and direction at 500 hPa, the CONV analysis is characterized by large areas with differences $> \pm 5\,\mathrm{m\,s^{-1}}$ (Fig. 11a). The biases in UWIND and VWIND are relatively low at -0.6 and $0.2\,\mathrm{m\,s^{-1}}$ with corresponding RMSDs of 3.9 and $4.7\,\mathrm{m\,s^{-1}}$, respectively (Figs. 9c,d). SAT appears to increase zonal wind speed in portions of central Texas and western Oklahoma compared to CONV causing a negative bias in UWIND ($-0.6\,\mathrm{m\,s^{-1}}$) to switch to a positive bias ($0.3\,\mathrm{m\,s^{-1}}$; Fig. 11b). RMSD increases slightly to $4.0\,\mathrm{m\,s^{-1}}$, indicating that assimilating satellite data alone does not improve the zonal wind field (Fig. 9c). Assimilating radar data, specifically radial velocity data, has a much larger impact on the wind field analysis as evident by visual inspection (Fig. 11c). The magnitude of the differences in wind speed greatly decreases and becomes less than $\pm 2\,\mathrm{m\,s^{-1}}$ where radar

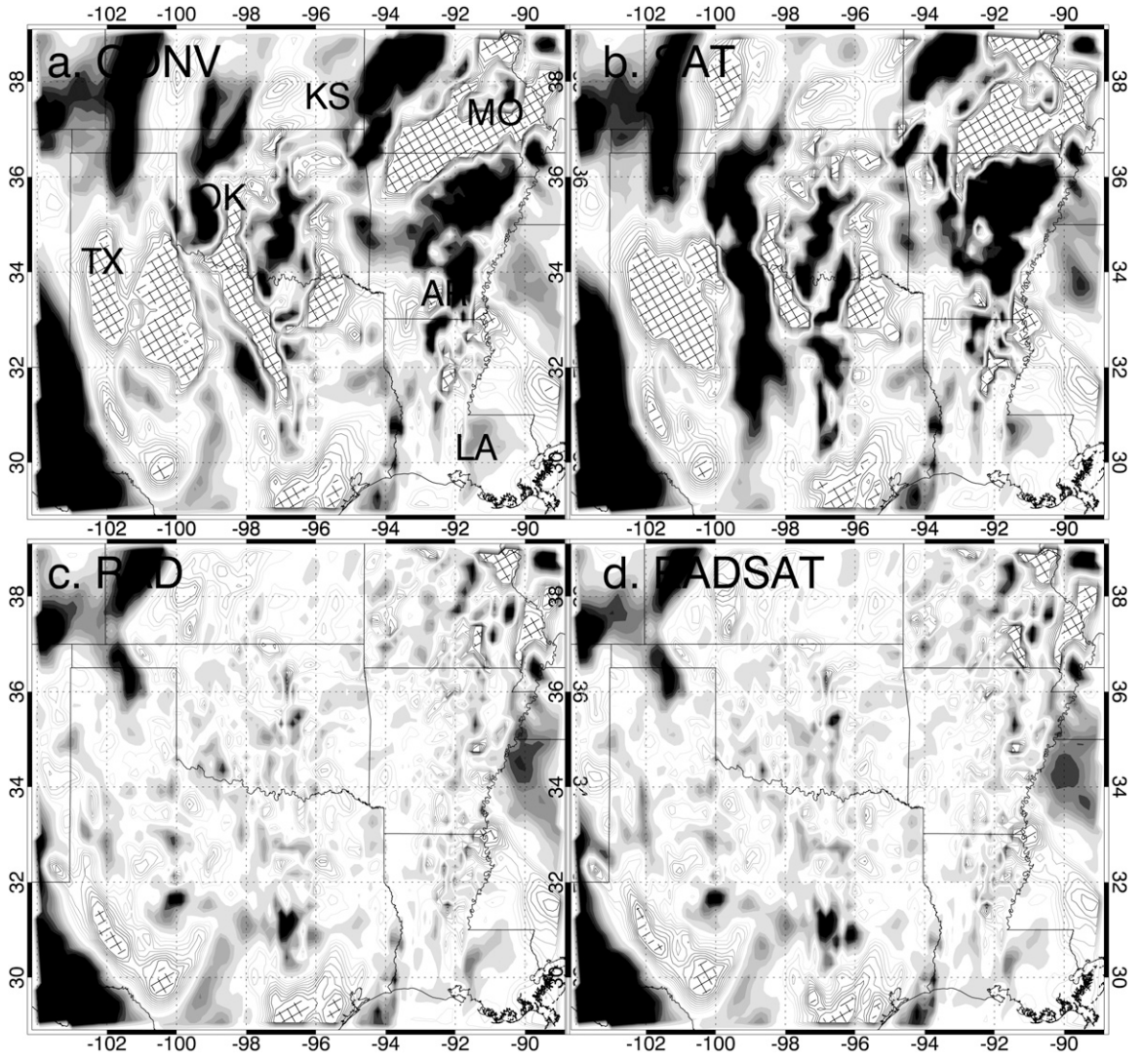

FIG. 11. As in Fig. 8, but for total wind speed between each experiment and truth.

observations are assimilated. The resulting biases are small with a large reduction in RMSD representing a >40% improvement compared to CONV. RADSAT results are quite similar to the RAD experiment, with a RMSD of 2.1 m s^{-1} (Figs. 9c,d and 11d). The differences between truth and experiment wind components differ somewhat as a function of height, but the RAD and RADSAT experiments consistently have the smallest errors (Figs. 9c,d). One interesting difference between SAT and CONV is the change in bias for both wind components. The bias in UWIND takes on an almost sinusoidal shape being -1.0 m s^{-1} at 750 hPa while increasing to 0.4 m s^{-1} at 550 hPa before being -1.0 m s^{-1} again at 300 hPa (Fig. 9c). The vertical profile of bias for the other experiments varies much less with height. Overall, these results show that the radar observations primarily in the form of radial velocity have a much larger impact than the satellite observations on the wind analysis.

b. Cloud variables

1) QALL

While assimilating both radar and/or satellite data can positively affect the analysis of temperature, humidity, and wind fields, the key metric as to the success of these experiments is how each represents the more complex cloud-related variables compared to the truth simulation. Vertical error distributions for each individual mixing ratio variables are then analyzed to assess which contributes most to improvements observed in the total cloud water content, previously defined as QALL. These errors are calculated for all grid points whether or not a cloud is present.

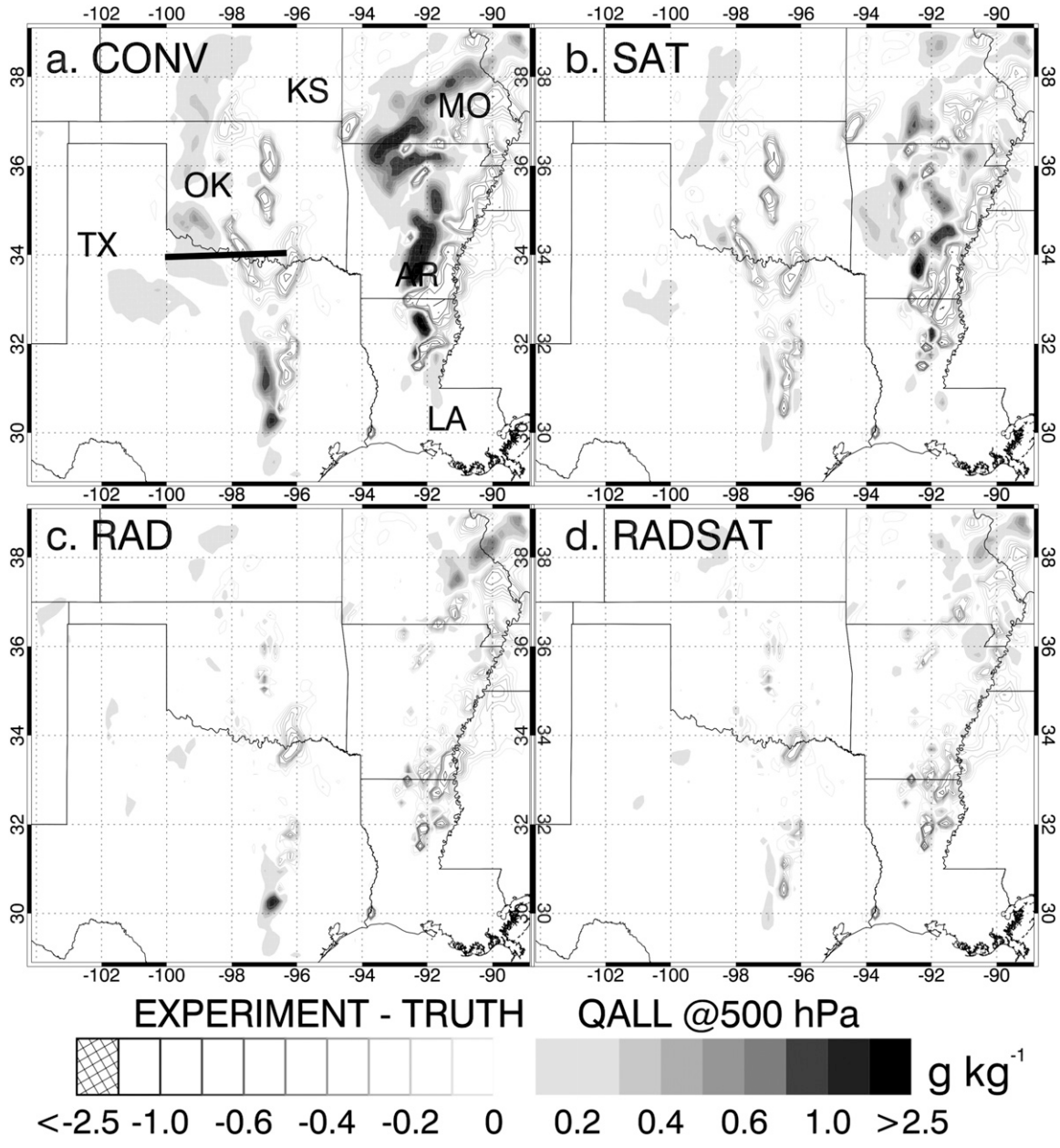

FIG. 12. (a)–(d) Difference in total cloud water (QALL) at 500 hPa at 1200 UTC 24 Dec 2009 between each experiment and the truth simulation. Gray to black indicates that the experiment has higher QALL, hence greater cloud cover, than truth. Whites and hatched indicate the opposite. Also shown in (a) is a black line representing the location of cold sector cross section shown in Fig. 14.

At 500 hPa, differences between QALL are confined to two large areas. The first lies in the central portion of the domain in Oklahoma and Texas with CONV overestimating QALL along a north–south line near 99°W while underestimating QALL just to the east (Fig. 12a). In the warm sector in the eastern portion of the domain, a larger area of overestimation exists in southern Missouri and central Arkansas, sometimes exceeding 2.5 g kg^{-1}. CONV underestimates QALL along the Mississippi River with a RMSD of 0.34 g kg^{-1} (Fig. 12a). Assimilating satellite data has a much larger impact in the QALL field at 500 hPa than at 850 hPa (Fig. 12b). The positive bias in QALL in western Oklahoma is nearly eliminated with a large reduction in positive bias also apparent in Arkansas, though the distribution and magnitude of the negative biases changes little. The RAD experiment performs even better, eliminating nearly all of the positive and negative biases in QALL except in the far

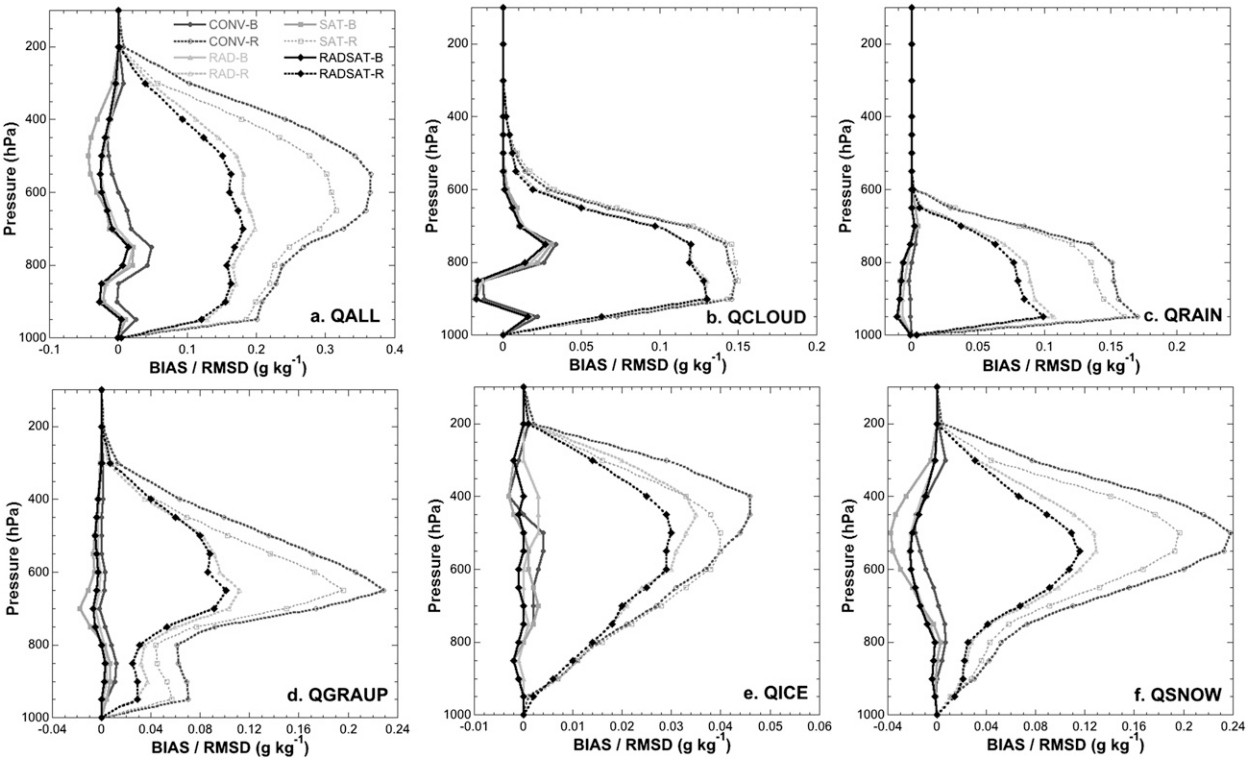

FIG. 13. Vertical profile of bias (B) and RMSD (R) for (a) QALL, (b) QCLOUD, (c) QRAIN, (d) QGRAUP, (e) QICE, and (f) QSNOW calculated over the verification domain for each experiment at 1200 UTC 24 Dec 2009.

eastern portion of the domain where there are no radar observations (Fig. 12c). The combination of radar and satellite data in RADSAT produces even better results (Fig. 12d). Similar results are present at 300 hPa, but with SAT having a larger improvement relative to CONV as compared to 500 hPa. The improved performance of the satellite observations in the middle and upper troposphere reflects their greater sensitivity to cloud top properties, with only indirect information about the lower portion of the cloud field limiting their impact at lower levels.

2) VERTICAL ERROR PROFILES

To better understand the mechanisms behind these improvements, it is important to determine which cloud hydrometeor species exerts the greatest impact on QALL errors. Figure 12a shows bias and RMSD for QALL for each experiment as a function of height. All experiments generally have a low bias, rarely exceeding $\pm 0.05\,\mathrm{g\,kg^{-1}}$ with small differences between each experiment. Much larger differences are apparent for RMSD, especially in the midtroposphere. CONV has the highest RMSD at all levels ranging from $0.2\,\mathrm{g\,kg^{-1}}$ at 950 hPa to $0.36\,\mathrm{g\,kg^{-1}}$ at 600 hPa before decreasing again at higher levels. The SAT experiment reduces RMSD primarily above 700 hPa, with RAD and RADSAT experiments significantly reducing RMSD at all levels. RADSAT also consistently has the lowest RMSD between 700 and 500 hPa.

For cloud liquid water (QCLOUD), CONV and SAT have almost identical RMSD profiles indicating that the 6.95-μm T_B have little impact (Fig. 13b). However, radar reflectivity is directly related to cloud water in the lower troposphere and has a much larger impact on the cloud mixing ratio errors. The magnitude and vertical distribution of errors for cloud rain (QRAIN) is similar to that for QCLOUD and approaches zero above 600 hPa (Fig. 13c). As with QCLOUD, the RAD and RADSAT experiments contain the smallest errors, which are about half of the corresponding CONV values.

For graupel (QGRAUP), biases are again near zero at most levels for all experiments with RMSD errors maximized at the 650-hPa level (Fig. 13d). At this level, RMSD for CONV is $0.23\,\mathrm{g\,kg^{-1}}$, which reduces to $0.20\,\mathrm{g\,kg^{-1}}$ in SAT and to near $0.10\,\mathrm{g\,kg^{-1}}$ in RAD and RADSAT. Radar data again have the greatest impact, though a consistent improvement from the satellite data is also apparent, especially above 700 hPa. Cloud ice mixing ratio (QICE) is highest in the upper troposphere near

~400 hPa (not shown) where the largest RMSD also occurs for each experiment (Fig. 13e). As for the previous cloud species, CONV consistently has the largest RMSD with RAD and RADSAT containing the smallest. The SAT experiment also reduces RMSD above 600 hPa, but has little effect below this layer. Finally, large differences are present in snow mixing ratio (QSNOW) with the maximum RMSD reduction occurring at 500 hPa (Fig. 13f). Both SAT and RAD reduce midtropospheric RMSD compared to CONV with radar data having the larger impact.

In summary, CONV has the highest RMSD values for all hydrometeor types followed by the SAT, RAD, and RADSAT cases. Assimilating radar reflectivity data accounts for much of the improvement in the RADSAT case, but the satellite T_B observations also provide useful information, as indicated by the reduction in RMSD from SAT compared to CONV in the mid- and upper troposphere. Assimilating radar data has a large impact on the liquid and frozen hydrometeors while the impact of satellite data is generally limited to the frozen species in the mid- and upper troposphere where the 6.95-μm band is most sensitive.

3) COLD-SECTOR CROSS SECTION

To visualize the effects of assimilating radar and satellite data on the vertical cloud distribution, a vertical cross section of QCLOUD, QICE, and temperature transecting the area of winter precipitation across northern Texas and southern Oklahoma is shown in Fig. 14 (refer to Fig. 13a for cross-section location). Truth QICE concentrations $>0.2\,\mathrm{g\,kg^{-1}}$ exist between 550 and 700 hPa west of 98°W, and near 400 hPa east of 98°W (Fig. 14a). QCLOUD is greatest at and below the freezing level, with local maxima near 96.7° and 98.1°W (Fig. 14a). The freezing level decreases from 800 hPa east of 98°W to near the surface farther west, which is consistent with snowfall observations at this time. The depiction of these features is much different in CONV (Fig. 14b). For instance, CONV generates a large area of cloud ice between 600 and 400 hPa west of 98°W that is not present in truth. Farther east, midtropospheric QICE concentrations are near zero compared to the much higher values generated in truth. QCLOUD also differs significantly with CONV only showing a single layer of liquid water near 800 hPa with values rarely exceeding $0.2\,\mathrm{g\,kg^{-1}}$. Temperature profiles are similar though CONV appears colder in the 900–700-hPa layer west of 98° and somewhat warmer farther east.

Assimilating remote sensing observations substantially improves the vertical representation of QICE and QCLOUD. The SAT experiment eliminates the mid- and upper-tropospheric QICE west of 98.5°W present in CONV (Fig. 14c). The distribution of QICE below 600 hPa also better matches truth, though many small-scale differences remain. The SAT experiment also resolves the dual-maximum structure in QCLOUD present in the truth simulation though it is too far west in its location. The large improvement in the QICE analysis is consistent with the improvement in QICE RMSD shown in Fig. 12e above 500 hPa whereas the improvement in QCLOUD is more limited, consistent with Fig. 12b. The temperature fields in SAT are similar to those generated by CONV.

Assimilating radar reflectivity and radial velocity data in RAD generates a much different result than either CONV or SAT. High values of QICE are present between 700 and 400 hPa from the western edge of the cross section eastward to 97°W (Fig. 14d). Several areas of localized maxima are present with those east of 98.5°W agreeing well with those in truth. However, RAD retains the mid- and upper-tropospheric QICE west of 99°W that is also present in CONV, but not SAT. Nearer the surface, the magnitude of QCLOUD increases compared to either CONV or SAT; however, RAD fails to capture the higher values east of 96.5°W below 700 hPa that are present in truth. The temperature profile is also different showing several small-scale features west of 98°W that are not present in truth, though the overall freezing level location is correct. The combined RADSAT experiment maintains improvements made during the RAD and SAT experiments, and thus generates the most accurate representation of this cross section (Fig. 14e). For instance, RADSAT does not contain the erroneous mid- and upper-tropospheric QICE present in CONV and SAT west of 99°W, but does have the higher values to the east. The location and magnitude of the maximum QICE concentrations are also more similar to truth. Overall, the greatest increase in skill along this cross section occurs when both satellite and radar data are assimilated.

c. *Satellite 6.95-μm T_B*

To assess the impact of assimilating satellite and radar data on their respective analysis fields, we compare simulated ensemble mean satellite T_B and radar reflectivity from each experiment with the corresponding truth simulation. In addition to calculating bias and RMSD statistics for each experiment, skill scores including POD, FAR, and HSS are also computed using a threshold of $T_B < 230$ K. If both the truth and experiment analysis generate a pixel where $T_B < 230$ K, then this is considered a "hit." If the experiment generates a pixel with $T_B < 230$ K and this threshold is not exceeded in truth, then it is considered a false detection. Finally, if neither truth nor the experiment exceeds this threshold, then it is considered a correct null forecast. The goal is to generate an analysis

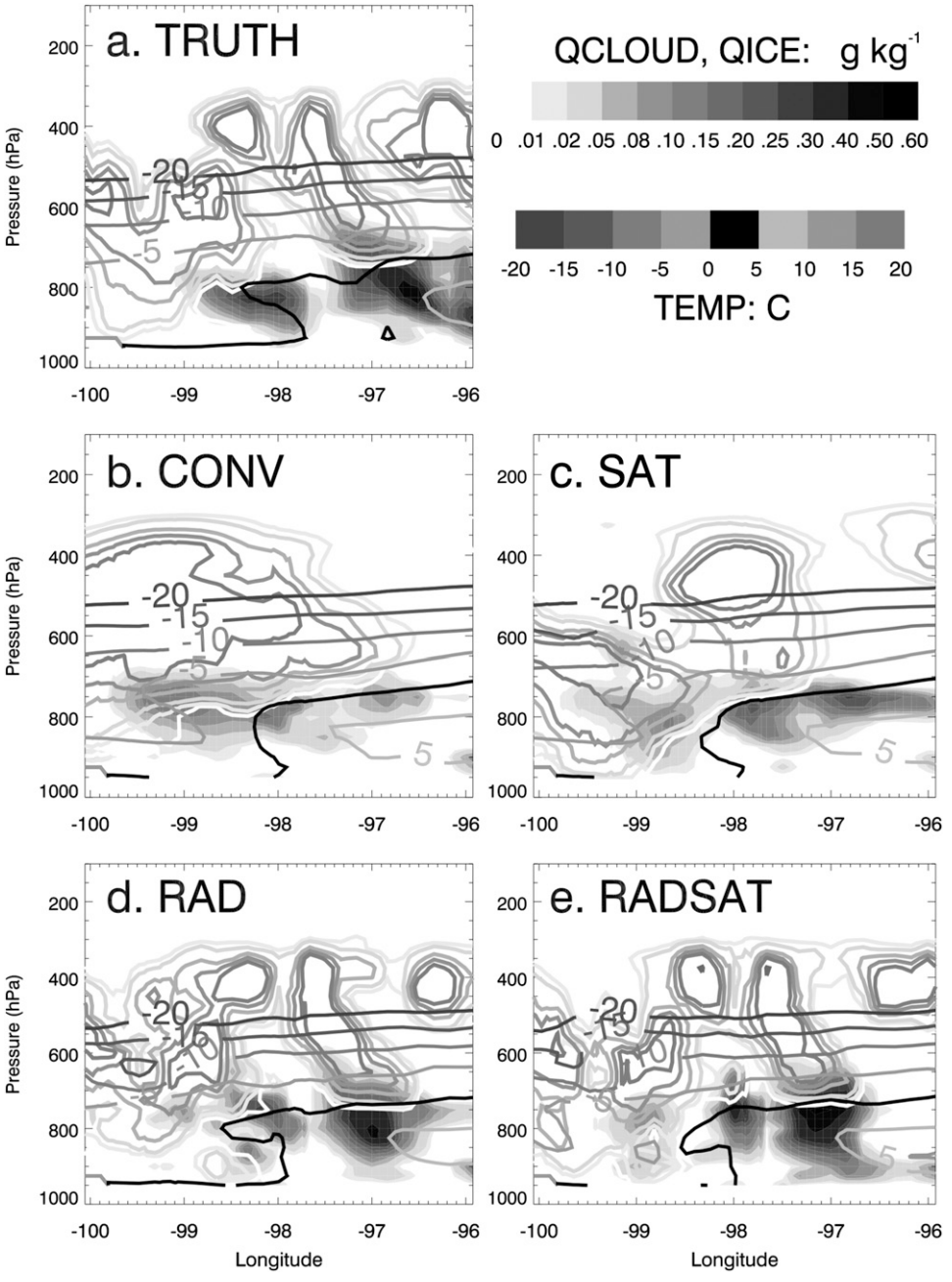

FIG. 14. Vertical profile of cloud liquid water content mixing ratio (shaded), cloud ice content (gray-shaded contours without values), and temperature (gray-shaded contours with values) from west to east between 100° and 96°W along the 34°N parallel covering an area of winter precipitation. The location is shown in Fig. 12a and labeled as COLD. Profiles from the (a) truth simulation, (b) CONV, (c) SAT, (d) RAD, and (e) RADSAT experiments are shown.

where "hits" are maximized, resulting in a high POD, but false detections are limited, thereby resulting in a low FAR. The HSS takes both into account to generate a statistic indicating the overall skill of the experiment at forecasting T_B.

Truth 6.95-μm T_B indicate the presence of high atmospheric moisture content in the eastern portion of the domain roughly along the 91°W meridian where $T_B <$ 230 K (Fig. 15a). In Oklahoma, T_B are warmer, indicating less midtropospheric moisture, which is consistent with the lower altitude of these clouds. Several finer details are also visible, such as linear banding features in Oklahoma and individual convective cells farther east. The CONV

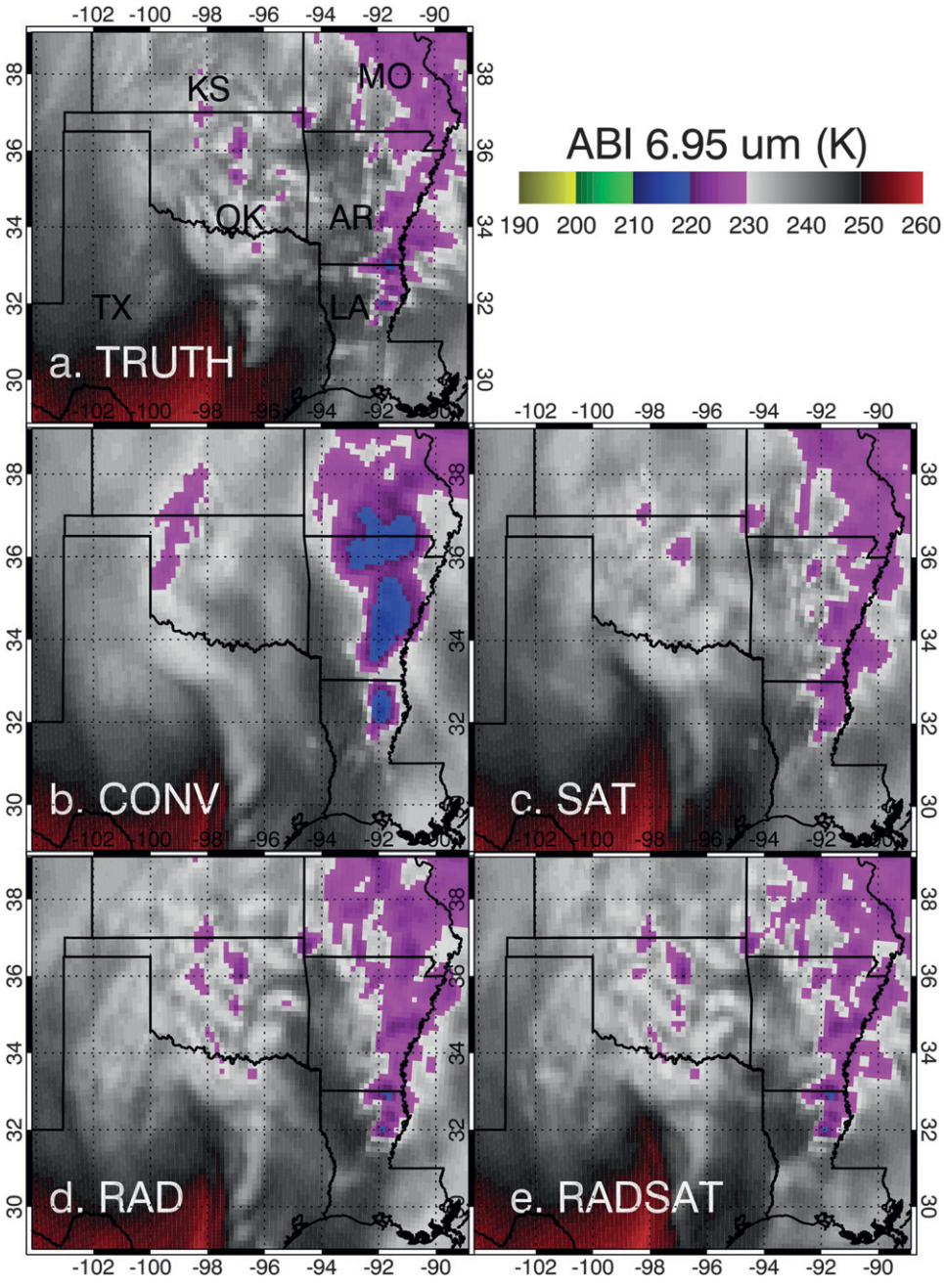

FIG. 15. Simulated GOES-R ABI 6.95-μm T_B (K) for the truth simulation and each experiment at 1200 UTC 24 Dec 2009.

experiment generates a similar T_B simulation overall, but either misses or incorrectly analyzes several important details (Fig. 15b), such as generating much colder T_B (<220 K) associated with the convection in the eastern portion of the domain (Fig. 15b). In northwest Oklahoma, CONV generates a region of very cold T_B that does not exist in the truth simulation. Finally, much of the finer-scale detail apparent in the truth simulation is lost in the CONV experiment. The excessively cold T_B in the eastern portion of the domain result in an overall cold bias of −2.3 K with a corresponding RMSD of 6 K (Table 5). The magnitude and displacement errors result in a low POD of 0.4 with a correspondingly high FAR of 0.64 resulting in a HSS of 0.37 (Table 6).

The SAT experiment generates a much better representation of 6.95-μm T_B compared to CONV (Fig. 15c). The erroneously cold (<220 K) T_B in the eastern portion

TABLE 5. Bias and RMSD between each experiment and truth at 1200 UTC for simulated ABI 6.95-μm T_B.

ABI 6.95-μm T_B	CONV	SAT	RAD	RADSAT
BIAS	2.30	−0.81	−1.97	−0.74
RMSD	6.02	2.10	4.15	1.85

TABLE 6. Probability of detection (POD), false alarm rate (FAR), and Heidke skill score (HSS) for simulated GOES-R band-9 $T_B <$ 230-K reflectivity between the truth simulation and individual experiment output at 1200 UTC. Similar statistics for 2 and 6 km AGL simulated radar reflectivity >25 dBZ are also shown. Values in italics represent the best skill for a particular variable and experiment.

	6.95-μm T_B			2-km Reflectivity			6-km Reflectivity		
Expt	POD	FAR	HSS	POD	FAR	HSS	POD	FAR	HSS
CONV	0.40	0.64	0.37	0.21	0.72	0.24	0.07	0.85	0.10
SAT	0.66	*0.22*	*0.71*	0.23	0.59	0.29	0.11	0.60	0.17
RAD	0.58	0.44	0.57	0.42	0.49	0.46	0.31	0.51	0.38
RADSAT	*0.67*	0.27	0.70	*0.43*	*0.47*	*0.47*	*0.31*	*0.50*	*0.40*

of the domain have been eliminated and the location of colder T_B shifts farther east in much better agreement to the truth simulation. To the west, the area of colder T_B in northwestern Oklahoma is no longer present while several of the smaller-scale features in this region are correctly analyzed. Since the 6.95-μm T_B from truth are being assimilated in SAT, it is expected that they would have a large impact on the corresponding analysis. The objective statistics support the visual interpretation and show a large reduction in bias and RMSD to −0.8 and 2.1 K, respectively (Table 5). Similarly, POD increases to 0.66 with an even larger decrease in FAR to 0.22 (Table 6). Corresponding skill is similarly improved with a HSS of 0.71.

Assimilating radar data also produces a significant impact to the 6.95-μm T_B analysis (Fig. 15d). Compared to CONV, the RAD experiment exhibits a similar cold bias, but does not have the large region with $T_B <$ 220 K. The overall bias is similar to CONV (−2.0 K); however, the RMSD decreases, but is still nearly twice as high as the SAT case (Table 5). Similarly, the improvement in skill scores between CONV and RAD are only about 50% of those seen for the SAT experiment (Table 6). Given that the radar data are not directly related to satellite T_B, the increase in skill is less than that observed during the SAT case. Still, the fact that assimilating radar data alone increases skill in simulated satellite T_B is very encouraging, and depicts their influence on the cloud field. The assimilation of both radar and satellite data in RADSAT slightly reduces bias and RMSD compared to the SAT experiment (Fig. 15e). Visually, RADSAT captures the details of the T_B distribution in central Oklahoma better than the other experiments and better depicts the finer-scale patterns in the T_B field throughout the entire domain. The corresponding objective skill scores are nearly the same or slightly worse than the SAT experiment (Table 6).

d. Radar reflectivity

A similar comparison is conducted using simulated radar reflectivity data as a function of height to determine the relative impacts of the observations on the radar reflectivity analysis. Skill scores are computed in the same manner as employed for the satellite date, but now using a threshold of 25 dBZ. At 2 km, the truth simulation generates a large area of >20-dBZ reflectivity in western Oklahoma and northwestern Texas associated with an area of snowfall (Fig. 16a). Higher reflectivity values are present along the eastern edge of the snowfall region with the highest values located in the deeper convection within the eastern portion of the domain. Reflectivity at 6 km is generally lower, especially in central Oklahoma and north Texas where much of the precipitation is stratiform and confined to lower altitudes (Fig. 18a). The line of convection is still apparent farther east, though reflectivity values are generally lower than at 2 km since most of the hydrometers at this level are some form of ice, which generally returns lower reflectivity values.

The CONV experiment produces too much precipitation throughout the entire domain and is too far west with the line of convection in Arkansas and Louisiana (Fig. 16b). The areas of reflectivity > 25 dBZ are much larger than truth, especially in the southern portion of the domain. These errors are evident in high bias (13.6 dBZ) and RMSD (27.2 dBZ) values (Fig. 17a, Table 6). Skill scores are also poor with a POD of 0.21, a FAR of 0.72, and a HSS value of 0.24. The SAT experiment reduces the coverage of higher reflectivity values in several locations, most notably in west central Texas and in northeastern Oklahoma (Fig. 16c). However, it also remains too aggressive with the intensity and coverage of precipitation. The bias, RMSD, and skill scores are only marginally improved, with the greatest improvement being a reduction in FAR to 0.59 (Fig. 17). Recall that 6.95-μm T_B are not very sensitive to the lower troposphere; thus, any improvement at 2 km in SAT is likely to be small. The RAD experiment, however, more accurately reproduces the finer-scale detail in the reflectivity fields than either CONV or SAT (Fig. 16d). This is evident in both the winter precipitation in central Oklahoma and the convection regions farther east. RAD also correctly analyzes the small convective feature in far northeastern Oklahoma and southwestern Missouri not

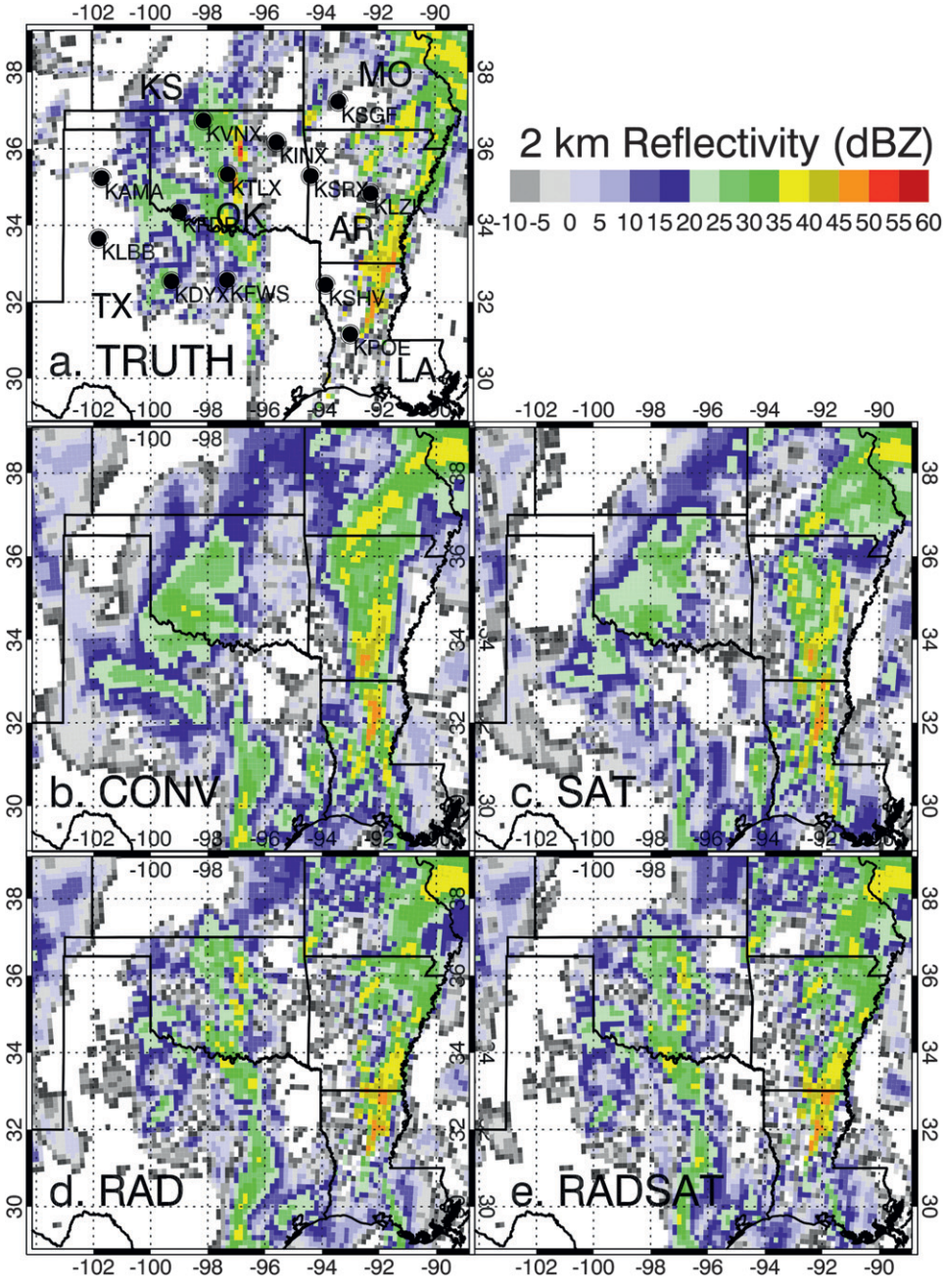

FIG. 16. Simulated WSR-88D 2-km radar reflectivity above ground level (AGL) at 1200 UTC for the truth simulation and each experiment.

evident in the CONV or SAT experiments. Still, RAD overestimates the intensity and coverage of precipitation, with a positive bias of 8.9 dBZ. Overall RMSD is reduced to 19.7 dBZ with modest improvements across all skill scores. RADSAT generates a 2-km reflectivity field that is visually similar to that from the RAD experiment (Fig. 16e). The objective statistics show that it is the most skillful experiment with the lowest bias and RMSD (8.6 and 19.1 dBZ), and the highest POD and HSS (0.43 and 0.47).

Between 2 and 6 km, POD for CONV and SAT are nearly identical while SAT continues to reduce FAR relative to CONV (Figs. 17b,c). Conversely, both RAD and RADSAT clearly outperform CONV and SAT between the surface and 6 km AGL. In both cases, POD and HSS are nearly double their CONV counterparts;

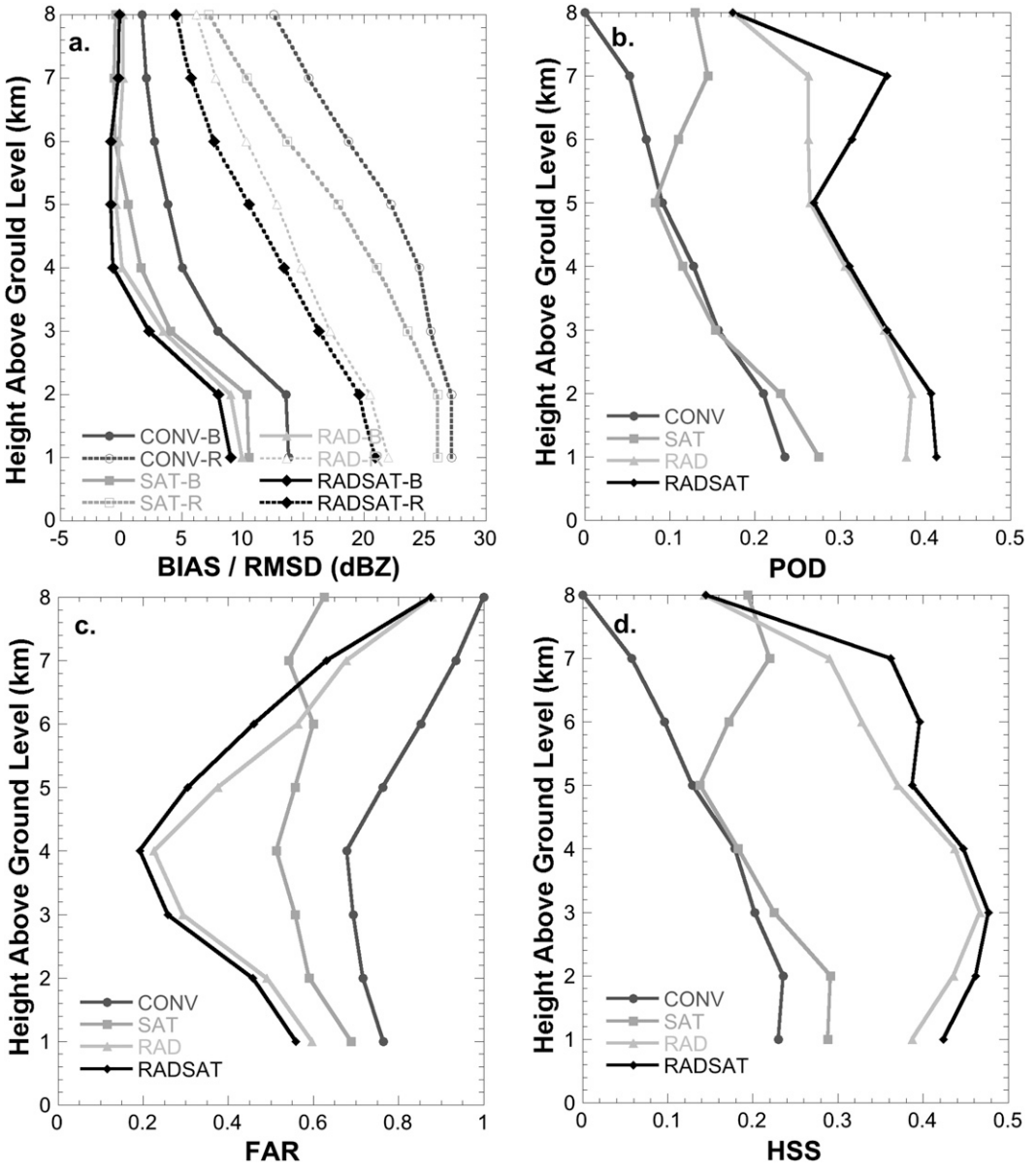

FIG. 17. Vertical profiles of (a) bias (B) and RMSD (R) for radar reflectivity relative to truth, (b) POD, (c) FAR, and (d) HSS for reflectivity >25 dBZ as a function of height from each experiment at 1200 UTC.

while FAR is reduced by half (Fig. 17). At 6 km, CONV has the largest bias and RMSD consistent with an analysis that again overestimates the precipitation at this level, especially in western Oklahoma (Fig. 18b). Overall skill is poor with a POD of 0.07 and a very high FAR of 0.85 with HSS being 0.10 (Table 6). SAT improves the 6-km reflectivity analysis much more than occurred at 2 km (Fig. 18c, Table 6), with skill scores showing a similar improvement (Table 6). Reflectivity in western Oklahoma is reduced and much finer detail in the structure of reflectivity is now apparent farther east. Above 6 km, POD and FAR calculated from SAT increase significantly compared to CONV; and begin to perform similarly to the RAD experiment (Fig. 17). This improvement reflects the much greater sensitivity of 6.95-μm T_B near this atmospheric layer compared to lower in the atmosphere.

The reflectivity structures present in the truth simulation are well represented in RAD at 6 km (Fig. 18d). Bias decreases to near zero with RMSD lowered to 9.5 dBZ. Corresponding POD and FAR are 0.31 and 0.51, respectively, much improved over the CONV experiment (Table 6). The RADSAT experiment again produces the most skillful reflectivity analysis. A small negative bias is present (-0.27 dBZ), but RMSD is lowered to 8.4 dBZ,

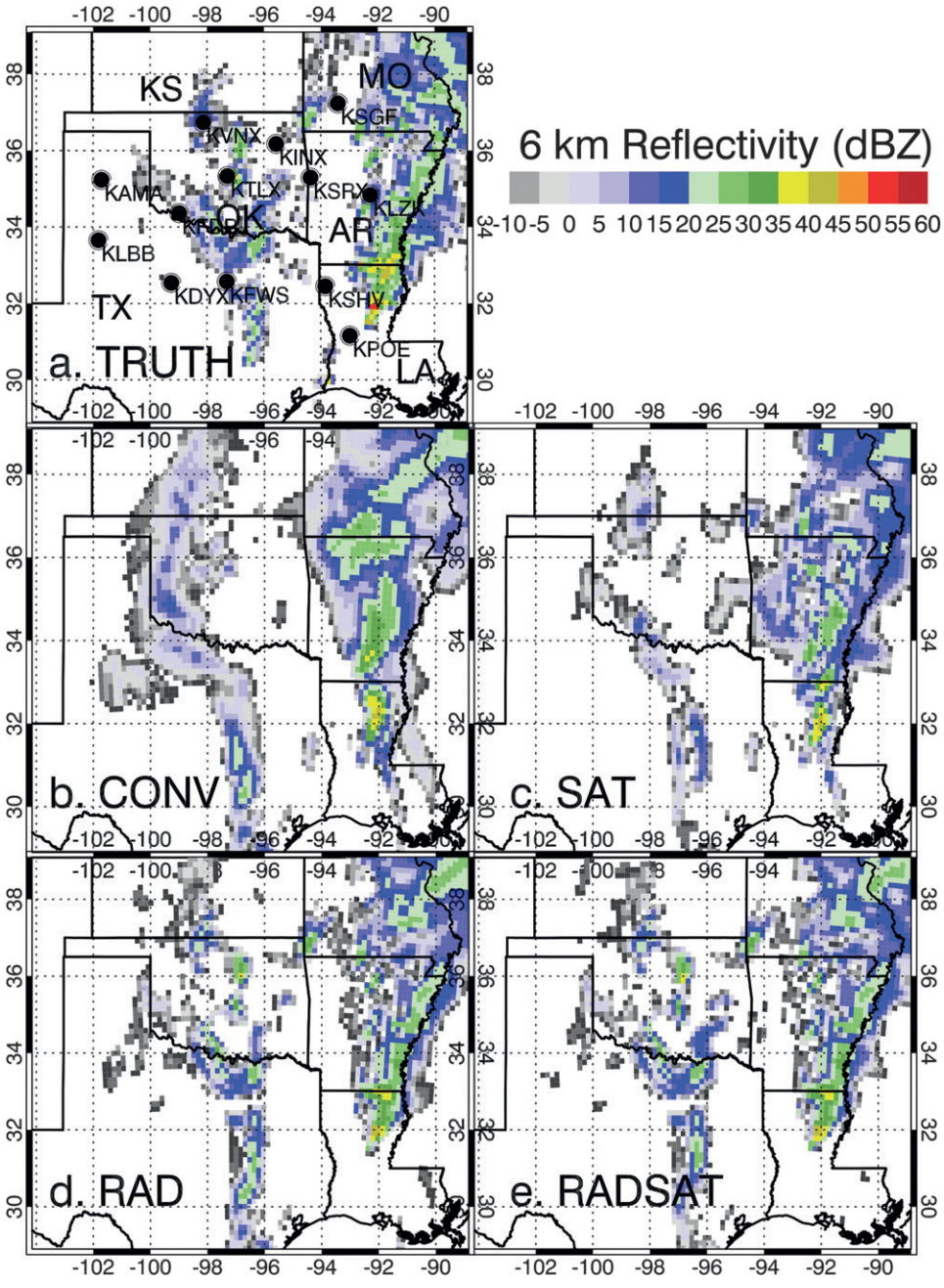

FIG. 18. As in Fig. 16, but for 6-km AGL simulated radar reflectivity.

or less than half of the original CONV value (Fig. 18e). RADSAT also produces the best skill scores with a POD of 0.31, a FAR of 0.46, and a HSS of 0.40. These scores represent over a 100% improvement compared to the CONV experiment. These results indicate that satellite data provide valuable information when analyzing radar reflectivity even when reflectivity data are also assimilated, with the greatest impact from satellite data occurring above 5 km.

8. Conclusions

Assimilating water vapor–sensitive infrared 6.95-μm T_B in combination with WSR-88D reflectivity and radial velocity often proved superior to assimilating either satellite or radar data alone. Radar data are most effective at improving low- and midtropospheric cloud liquid water as well as wind velocity throughout the troposphere, while satellite T_B generally improved the

mid- and upper-tropospheric moisture and cloud ice concentrations. In essence, radar data are most effective in the lower troposphere while the potential for satellite data lie in the upper troposphere. Both datasets provide independent, complementary information that combine to give a better representation of the atmospheric state than is possible when used individually.

For most atmospheric variables, assimilating radar data appear to have a greater positive impact than assimilating satellite data. This is partially a result of the design of this experiment focusing on the region where radar data are being assimilated. Also, the number of radar reflectivity and radial velocity observations is far greater than the number of satellite T_B observations. Recall that the radar data provided high-resolution 3D data whereas the satellite data only provide single T_Bs that represent a broad layer of the atmospheric column. Still, the importance of assimilating satellite data is apparent given that the combined RADSAT experiment typically had smaller errors than when the radar observations were assimilated separately.

Ongoing research will examine the 0–3-h forecast period to determine which observations had the greatest impact on subsequent forecasts. In addition, future research will also refine techniques to extract the greatest amount of information from a combined radar-satellite dataset to provide a path toward operational assimilation of these important observations. Several challenges became apparent during the course of this research opening up many possibilities for future research. The modest impacts of satellite observations indicates the need for additional infrared satellite channels to gain better sensitivity of the lower- and midtropospheric water vapor content. The sensitivity of model microphysics also needs to be examined as both datasets affect liquid and frozen hydrometeor species differently depending on the microphysics option selected. Finally, the lack of a vertical covariance localization could be introducing an unwanted error; thus, future research will need to determine potential vertical localization radii as a function of conditions, channel, and cloud cover to maximize the potential for satellite data in ensemble data assimilation systems.

Acknowledgments. This work was funded by the National Oceanic and Atmospheric Administration under Grant NA10NES4400013. Additional support was provided by the NOAA/National Environmental Satellite, Data, and Information Service (NESDIS). The assimilation experiments were performed using the NESDIS ''S4'' supercomputer located at the University of Wisconsin—Madison and the ''ranger'' supercomputer located at the Texas Advanced Computing Center. Ranger is part of the Extreme Science and Engineering Discovery Environment (XSEDE) network that is supported by the National Science Foundation under Grant OCI-1053575. We would also like to thank the three anonymous reviewers who provided comments that improved the overall quality of this manuscript.

REFERENCES

Aksoy, A., D. Dowell, and C. Snyder, 2009: A multicase comparative assessment of the ensemble Kalman filter for assimilation of radar observations. Part I: Storm-scale analyses. *Mon. Wea. Rev.,* **137,** 1805–1824.

——, ——, and ——, 2010: A multicase comparative assessment of the ensemble Kalman filter for assimilation of radar observations. Part II: Short-range ensemble forecasts. *Mon. Wea. Rev.,* **138,** 1273–1292.

Anderson, J. L., T. Hoar, K. Raeder, H. Liu, N. Collins, R. Torn, and A. Avellano, 2009: The Data Assimilation Research Testbed: A community data assimilation facility. *Bull. Amer. Meteor. Soc.,* **90,** 1283–1296.

Baum, B. A., P. Yang, A. J. Heymsfield, S. Platnick, M. D. King, Y.-X. Hu, and S. T. Bedka, 2005: Bulk scattering properties for the remote sensing of ice clouds. Part II: Narrowband models. *J. Appl. Meteor.,* **44,** 1896–1911.

Bouttier, F., and G. Kelly, 2001: Observing-system experiments in the ECMWF 4D-Var data assimilation system. *Quart. J. Roy. Meteor. Soc.,* **127,** 1469–1488.

Caya, A., J. Sun, and C. Snyder, 2005: A comparison between the 4DVAR and ensemble Kalman filter techniques for radar data assimilation. *Mon. Wea. Rev.,* **133,** 3081–3094.

Chevallier, F., P. Lopez, A. M. Tompkins, M. Janiskova, and E. Moreau, 2004: The capability of 4D-Var systems to assimilate cloud-affected satellite infrared radiances. *Quart. J. Roy. Meteor. Soc.,* **130,** 917–932.

Collard, A. D., and A. P. McNally, 2009: The assimilation of Infrared Atmospheric Sounding Interferometer radiances at ECMWF. *Quart. J. Roy. Meteor. Soc.,* **135,** 1044–1058.

Crum, T. D., and R. L. Alberty, 1993: The WSR-88D and the WSR-88D Operational Support Facility. *Bull. Amer. Meteor. Soc.,* **74,** 1669–1687.

Derber, J. C., and W. S. Wu, 1998: The use of TOVS cloud-cleared radiances in the NCEP SSI analysis system. *Mon. Wea. Rev.,* **126,** 2287–2299.

Dowell, D. C., L. J. Wicker, and C. Snyder, 2011: Ensemble Kalman filter assimilation of radar observations of the 8 May 2003 Oklahoma City supercell: Influences of reflectivity observations on storm-scale analyses. *Mon. Wea. Rev.,* **139,** 272–294.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.,* **99** (C5), 10 143–10 162.

Gao, J., and D. J. Stensrud, 2012: Assimilation of reflectivity data in a convective-scale, cycled 3DVAR framework with hydrometeor classification. *J. Atmos. Sci.,* **69,** 1054–1065.

——, M. Xue, A. Shapiro, and K. K. Droegemeier, 1999: A variational method for the retrieval of three-dimensional wind fields from dual-Doppler radars. *Mon. Wea. Rev.,* **127,** 2128–2142.

——, ——, K. Brewster, and K. K. Droegemeier, 2004: A three-dimensional variational data assimilation method with recursive filter for single-Doppler radar. *J. Atmos. Oceanic Technol.,* **21,** 457–469.

Gaspari, G., and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.,* **125,** 723–757.

Han, Q., W. Rossow, R. Welch, A. White, and J. Chou, 1995: Validation of satellite retrievals of cloud microphysics and liquid water path using observations from FIRE. *J. Atmos. Sci.,* **52,** 4183–4195.

Heemink, A. W., M. Verlaan, and A. J. Segers, 2001: Variance reduced ensemble Kalman filtering. *Mon. Wea. Rev.,* **129,** 1718–1728.

Heidinger, A. K., C. O'Dell, R. Bennartz, and T. Greenwald, 2006: The successive-order-of-interaction radiative transfer model. Part I: Model development. *J. Appl. Meteor. Climatol.,* **45,** 1388–1402.

Heymsfield, A. J., S. Matrosov, and B. Baum, 2003: Ice water path–optical depth relationships for cirrus and deep stratiform ice cloud layers. *J. Appl. Meteor.,* **42,** 1369–1390.

Hong, S.-Y., Y. Noh, and J. Dudhia, 2006: A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon. Wea. Rev.,* **134,** 2318–2341.

Iacono, M. J., J. S. Delamere, E. J. Mlawer, M. W. Shephard, S. A. Clough, and W. D. Collins, 2008: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *J. Geophys. Res.,* **113,** D13103, doi:10.1029/2008JD009944.

Kain, J. S., 2004: The Kain–Fritsch convective parameterization: An update. *J. Appl. Meteor.,* **43,** 170–181.

——, 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. 24, Amer. Meteor. Soc., 165–170.

Kalman, R. E., 1960: A new approach to linear filtering and prediction problems. *J. Basic Eng.,* **82,** 35–45.

Le Marshall, J., and Coauthors, 2006: Improving global analysis and forecasting with AIRS. *Bull. Amer. Meteor. Soc.,* **87,** 891–894.

Lin, Y.-L., R. D. Farley, and H. D. Orville, 1983: Bulk parameterization of the snow field in a cloud model. *J. Climate Appl. Meteor.,* **22,** 1065–1092.

McCarty, W., G. Jedloveck, and T. L. Miller, 2009: Impact of the assimilation of Atmospheric Infrared Sounder radiance measurements on short-term weather forecasts. *J. Geophys. Res.,* **114,** D18122, doi:10.1029.2008JD011626.

McNally, A. P., J. C. Derber, W. Wu, and B. B. Katz, 2000: The use of TOVS level-1b radiances in the NCEP SSI analysis system. *Quart. J. Roy. Meteor. Soc.,* **126,** 689–724.

——, P. D. Watts, J. A. Smith, R. Engelen, G. A. Kelly, J. N. Thepaut, and M. Matricardi, 2006: The assimilation of AIRS radiance data at ECMWF. *Quart. J. Roy. Meteor. Soc.,* **132,** 935–957.

Otkin, J. A., 2010: Clear and cloudy sky infrared brightness temperature assimilation using an ensemble Kalman filter. *J. Geophys. Res.,* **115,** D19207, doi:10.1029/2009JD013759.

——, 2012a: Assessing the impacts of the covariance localization radius when assimilating infrared brightness temperature observations using an ensemble Kalman filter. *Mon. Wea. Rev.,* **140,** 543–561.

——, 2012b: Assimilation of water vapor sensitive infrared brightness temperature observations during a high impact weather event. *J. Geophys. Res.,* **117,** D19203, doi:10.1029/2012JD017568.

——, and T. J. Greenwald, 2008: Comparison of WRF model-simulated and MODIS-derived cloud data. *Mon. Wea. Rev.,* **136,** 1957–1970.

——, ——, 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.

Polkinghorne, R., and T. Vukicevic, 2011: Data assimilation of cloud affected radiances in a cloud resolving model. *Mon. Wea. Rev.,* **139,** 755–773.

Reale, O., J. Susskind, R. Rosenberg, E. Brin, E. Liu, L. P. Riishojgaard, J. Terry, and J. C. Jusem, 2008: Improving forecast skill by assimilation of quality-controlled AIRS temperature retrievals under partially cloudy conditions. *Geophys. Res. Lett.,* **35,** L08809, doi:10.1029/2007GL33002.

Schmit, T. J., M. M. Gunshor, W. P. Menzel, J. J. Gurka, J. Li, and A. S. Bachmeier, 2005: Introducing the next-generation advanced baseline imager on GOES-R. *Bull. Amer. Meteor. Soc.,* **86,** 1079–1096.

Seaman, C. J., M. Sengupta, and T. H. Vonder Haar, 2010: Mesoscale satellite data assimilation: Impact of cloud-affected infrared observations on a cloud-free initial model state. *Tellus,* **62A,** 298–318.

Skamarock, W. C., and Coauthors, 2008: A description of the Advanced Research WRF version 3. NCAR Tech Note NCAR/TN-475+STR, 113 pp. [Available from UCAR Communications, P.O. Box 3000, Boulder, CO 80307.]

Snyder, C., and F. Zhang, 2003: Assimilation of simulated Doppler radar observations with an ensemble Kalman filter. *Mon. Wea. Rev.,* **131,** 1663–1677.

Stengel, M., M. Lindskog, P. Unden, N. Gustafsson, and R. Bennartz, 2010: An extended observation operator in HIRLAM 4D-VAR for the assimilation of cloud-affected satellite radiances. *Quart. J. Roy. Meteor. Soc.,* **136,** 1064–1074.

Torn, R. D., G. J. Hakim, and C. Snyder, 2006: Boundary conditions for limited area ensemble Kalman filters. *Mon. Wea. Rev.,* **134,** 2490–2502.

Vukicevic, T., T. Greenwald, M. Zupanski, D. Zupanski, T. Vonder Haar, and A. Jones, 2004: Mesoscale cloud state estimation from visible and infrared satellite radiances. *Mon. Wea. Rev.,* **132,** 3066–3077.

——, M. Sengupta, A. S. Jones, and T. Vonder Haar, 2006: Cloud-resolving satellite data assimilation: Information content of IR window observations and uncertainties in estimation. *J. Atmos. Sci.,* **63,** 901–919.

Weygandt, S. S., A. Shapiro, and K. K. Droegemeier, 2002: Retrieval of model initial fields from single-Doppler observations of a supercell thunderstorm. Part II: Thermodynamic retrieval and numerical prediction. *Mon. Wea. Rev.,* **130,** 454–476.

Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences*. 2nd ed. Academic Press, 627 pp.

Xue, M., M. Tong, and K. K. Droegemeier, 2006: An OSSE framework based on the ensemble square root Kalman filter for evaluating the impact of data from radar networks on thunderstorm analysis and forecasting. *J. Atmos. Oceanic Technol.,* **23,** 46–66.

Yussouf, N., and D. J. Stensrud, 2012: Comparison of single-parameter and multiparameter ensembles for assimilation of radar observations using the ensemble Kalman filter. *Mon. Wea. Rev.,* **140,** 562–586.

Zhang, F., C. Snyder, and J. Sun, 2004: Impacts of initial estimate and observation availability on convective scale data assimilation with an ensemble Kalman filter. *Mon. Wea. Rev.,* **132,** 1238–1253.