Assimilating AIRS Temperature and Mixing Ratio Profiles Using an Ensemble Kalman Filter Approach for Convective-Scale Forecasts

THOMAS A. JONES

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

DAVID J. STENSRUD

NOAA/OAR/National Severe Storms Laboratory, Norman, Oklahoma

(Manuscript received 22 August 2011, in final form 20 December 2011)

ABSTRACT

One satellite data product that has received great interest in the numerical weather prediction community is the temperature and mixing ratio profiles derived from the Atmospheric Infrared Sounder (AIRS) instrument on board the *Aqua* satellite. This research assesses the impact of assimilating AIRS profiles on high-resolution ensemble forecasts of southern plains severe weather events occurring on 26 May 2009 and 10 May 2010 by comparing two ensemble forecasts. In one ensemble, the 1830 and 2000 UTC level 2 AIRS temperature and dewpoint profiles are assimilated with all other routine observations into a 36-member, 15-km Weather and Research Forecast Model (WRF) ensemble using a Kalman filter approach. The other ensemble is identical, except that only routine observations are assimilated. In addition, 3-km one-way nested-grid ensemble forecasts are produced during the periods of convection. Results indicate that over the contiguous United States, the AIRS profiles do not measurably improve the ensemble mean forecasts of midtropospheric temperature and dewpoint. However, the ensemble mean dewpoint profiles in the region of severe convective development are improved by the AIRS assimilation. Comparisons of the forecast ensemble radar reflectivity probabilities between the 1- and 4-h forecast times with nearby Weather Surveillance Radar-1988 Doppler (WSR-88D) observations show that AIRS-enhanced ensembles consistently generate more skillful forecasts of the convective features at these times.

1. Introduction

Significant advancements in numerical weather prediction (NWP) have been made in the last two decades, many of which result from the assimilation of satellite data to improve the analysis of the atmospheric state where traditional observations are not available (Derber and Wu 1998; Le Marshall et al. 2007). Many satellite products have been incorporated into these models and include wind speed and direction derived from cloud motion, temperature, and humidity data from infrared (IR) and microwave radiances, rainfall rate estimates, and lightning observations (Velden et al. 1998; Hou et al. 2004; Reale et al. 2008; Mansell et al. 2007).

Corresponding author address: 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/WAF-D-11-00090.1

Operational models currently assimilate IR and microwave satellite radiances directly from multiple sensors such as the Advanced Very High Resolution Radiometer (AVHRR) and the Advanced Microwave Sounding Unit (AMSU) on polar-orbiting satellites and the Geostationary Operational Environmental Satellite (GOES) imager and sounder (Derber and Wu 1998; McNally et al. 2000, 2006). Model state variables are generally transformed into simulated satellite radiances via a radiative transfer model (RTM) built into the forward operator (McNally et al. 2006; Chen et al. 2008). Assimilating radiances directly has generally been the favored approach in operations and research studies since it avoids uncertainties and discrepancies in various retrieval algorithms that differ from satellite to satellite (Derber and Wu 1998; Errico 2000).

However, the assimilation of retrieved products remains an important research topic. Products such as temperature and humidity profiles calculated using

© 2012 American Meteorological Society

retrieval algorithms developed in concert with a particular sensor have been used by several studies to assess their impact on various forecast metrics (e.g., Reale et al. 2008; Li and Liu 2009). One advantage of this approach is that the retrieval algorithms often correct for sensor biases and calibration issues using their own RTM while also significantly reducing the data volume compared to assimilating many channels of raw radiances. By reducing data volume and eliminating the need for additional calls to the RTM in a NWP model, processing requirements are reduced, decreasing the expense of satellite data assimilation (Migliorini 2012). Another advantage of this approach is that information from partly cloudy areas is more easily incorporated into the data streams, whereas direct assimilation of cloudy radiances remains much more challenging and very resource intensive (Errico et al. 2007; Weisz et al. 2007; Pavelin et al. 2008).

Recent studies have shown the assimilation of satellite data using either retrievals or raw radiances improves predictability by 4–6 h for forecasts several days out (Pavelin et al. 2008; Chou et al. 2007, 2009). However, these improvements generally represent a global average and are not representative of the impacts possible from high-resolution satellite data assimilation on regional scales. Operational satellite data assimilation generally uses products thinned to a spatial resolution on the order of ~100 km with large uncertainties in the vertical distribution of derived atmospheric parameters (Goldberg et al. 2003). The lack of information on the vertical state of the atmosphere represents a key limitation of current NWP approaches, but recent advancements in sensors and assimilation techniques are beginning to address this issue. One such sensor is the Atmospheric Infrared Sounder (AIRS) instrument on board the *Aqua* satellite that is a hyperspectral infrared spectrometer from which temperature and mixing ratio profiles are retrieved in clear to partly cloudy conditions from near the surface into the stratosphere (Aumann et al. 2003; Susskind et al. 2003, 2006). The AIRS instrument measures over 2000 infrared channels, substantially decreasing the uncertainty of retrieved vertical profiles over other instruments, whose channels generally number two orders of magnitude less.

Several studies analyze the impact of assimilating these profiles into NWP models and find a positive impact on analysis and forecast fields (e.g., Le Marshall et al. 2006; Chou et al. 2007; Reale et al. 2008). One example focuses on the impact of assimilating AIRS profiles on tropical cyclone track and intensity forecasts as the limited amounts of in situ observations over the oceans can severely limit predictability (Wu et al. 2006; Li and Liu 2009; Liu and Li 2010; Pu and Zhang 2011). Both Li and Liu (2009) and Liu and Li (2010) find measureable improvement in 12–96-h track forecasts due to better analyses of the midtropospheric temperature and mixing ratio compared with forecasts produced without AIRS data. Smaller, but still significant, improvements are present for intensity forecasts. Pu and Zhang (2011) show similar results while also noting the possible need for bias adjustments in the assimilated profiles to further reduce forecast error. The impact of AIRS profile assimilation on a severe weather event in eastern Texas on 12–13 February 2007 is described by Chou et al. (2010) where the AIRS profiles improve the characterization of the thermodynamic environment leading to a more accurate 6-h precipitation forecast.

We use an ensemble Kalman filter (EnKF) data assimilation scheme (Kalman 1960; Evensen 2003; Liu et al. 2009; Yussouf and Stensrud 2010; Meng and Zhang 2011) rather than the more traditional variational approach (e.g., Barker at al. 2004). The primary advantage of the EnKF approach is that it provides a flow-dependent and dynamically evolving estimate of the background error covariances. Thus, there is no reliance on a priori assumptions for the background error characteristics. This is an important consideration for satellite data whose uncertainties can be a function of many factors that are time and location dependent. Both Li and Liu (2009) and Liu and Li (2010) successfully assimilate AIRS profiles using an EnKF approach for tropical cyclone applications and show improvement in forecast skill. The impact of AIRS data assimilation on severe weather forecasting examined by Chou et al. (2010) uses a variational assimilation approach, also noting some increase in forecast skill. This research builds on these previous studies by incorporating ensemble assimilation techniques with the satellite-derived atmospheric profiles to determine their combined impacts on convective forecasting applications for the first time. A key hypothesis of this research is that improvements in the mesoscale environment from assimilating AIRS profiles translate down to improvements in the near-storm environment, thereby improving forecasts on the convective scale.

Given the complexity of this undertaking, this research is presented as a ''proof of concept'' example to show that AIRS profiles can impact convective-scale forecasts in a positive fashion. Even this limited goal presents several obstacles that must be overcome. Since this research is conducted for events located over land, isolating the impacts of assimilating AIRS profiles relative to traditional observations is difficult. Previous tropical cyclone studies (Li and Liu 2009; Liu and Li 2010) assimilate AIRS profiles in an otherwise data-sparse environment over the ocean, increasing the potential for AIRS data to have a positive impact on the analysis and ensuing forecasts. The other issue is the need for adequate verification

metrics to assess the impact of the AIRS profiles. Only a limited number of upper-level observations are available that correspond to the locations and times of the satellite overpasses, resulting in few ways to assess whether or not the satellite observations being assimilated truly represent the atmospheric state. Validating its impact on convective-scale forecasts is also limited by a lack of corresponding in situ observations at the correct times and scales. To partially address these issues, cases are selected that occurred during the Verification of the Origin of Rotation in Tornadoes Experiment 2 (VORTEX2) that sampled several severe weather events in great detail (Wurman et al. 2010). These observations are used to assess the impacts of AIRS data on the overall environment while Weather Surveillance Radar-1988 Doppler (WSR-88D) radar reflectivity is used to validate the placement of forecast convection in each model run. While an analysis of multiple case studies would be ideal, the resources required for EnKF assimilation of AIRS profiles limits us to two events. The primary focus is a severe weather outbreak occurring on 10 May 2010 with a secondary case occurring on 26 May 2009.

A brief summary of the 10 May 2010 and 26 May 2009 events selected for study is provided in section 2. A description of the AIRS instruments and retrieved temperature and mixing ratio profiles is given in section 3. Section 4 provides a summary of the numerical model and data assimilation approaches. Section 5 analyzes the impact of AIRS data assimilation on a contiguous U.S. (CONUS) domain while section 6 assesses the impact of AIRS assimilation on the predicted location and intensity of convection in the southern plains compared to WSR-88D data. Section 7 summarizes the results.

2. Event summaries

a. 10 May 2010

Severe thunderstorms form in the afternoon and evening of 10 May 2010 associated with the passage of a short-wave trough across the southern Great Plains, with dozens of damaging wind, severe hail, and tornado reports between 2100 UTC 10 May and 0200 UTC 11 May (Fig. 1a). A corresponding surface low pressure center deepens as it moves eastward through southern Kansas, trailing a dryline southward into central Oklahoma, as shown at 2100 UTC (Fig. 1b). Ahead of the dryline, surface dewpoints greater than 20°C and south-southeasterly surface winds greater than 10 m s^{-1} exist over a large portion of central and southeastern Oklahoma. Surface temperatures range from ~34°C in southwestern Oklahoma to below 20°C in the northeastern part of the state where low clouds are present.

At higher levels, a strong west wind field (>40 m s^{-1} at 500 hPa) results in storm motions in excess of 25 m s^{-1} (Fig. 1c). The large change in wind speed and direction with height creates more than ample storm-relative helicity (SREH) for tornadic supercells with 0–3-km values greater than 500 m^2 s^{-2} present (Fig. 1d). Convective available potential energy (CAPE) also exceeds 2500 J kg^{-1} throughout much of central and southeastern Oklahoma (Fig. 1d). Supercell thunderstorms develop ahead of the eastward-progressing dryline initially in southern Kansas around 1800 UTC and in central and southern Oklahoma by 2000 and 2200 UTC, respectively. A second line of supercells initiates nearer the dryline itself after 2300 UTC and also moves eastward. Several of the supercells produce tornadoes in Oklahoma before evolving into a more linear structure in southeastern Oklahoma by 0100 UTC 11 May. The focus of this research lies in the Oklahoma supercells since they develop after the AIRS overpass at 2000 UTC.

b. 26 May 2009

The second event consists of a splitting supercell occurring during the afternoon and evening of 26 May 2009 in north-central Texas. Convection initiates in the vicinity of a southeasterly moving cold front that is being reenforced by outflow from morning convection. Strong instability (CAPE ~2000 J kg^{-1}) but relatively weak helicity (SREH ~100 m^{-2} s^{-2}) provide an environment only adequate for severe convection. The base of an upper-level trough is located in central Texas with weak 500-hPa winds (<10 m s^{-1}) and low vorticity, limiting the potential for a large-scale severe weather event. Convective initiation occurs ahead of the cold front, leading to the formation of a large supercell storm after 2200 UTC west of Fort Worth. As the supercell progresses eastward, it splits into left- and right-moving components around 2330 UTC with the right mover becoming outflow dominant. The left mover, with anticyclonic rotation, continues to show supercelluar characteristics until at least 0100 UTC 27 May. The isolated nature of this case makes it ideal for testing the impacts of the effects of modifying the near-storm environment from satellite data assimilation. Results from this case are used to determine whether or not assimilating AIRS profiles has consistent effects compared to those found for 10–11 May 2010.

3. AIRS data

AIRS is a hyperspectral infrared spectrometer launched on board the National Aeronautics and Space Administration's (NASA) Earth Observing System (EOS) *Aqua*

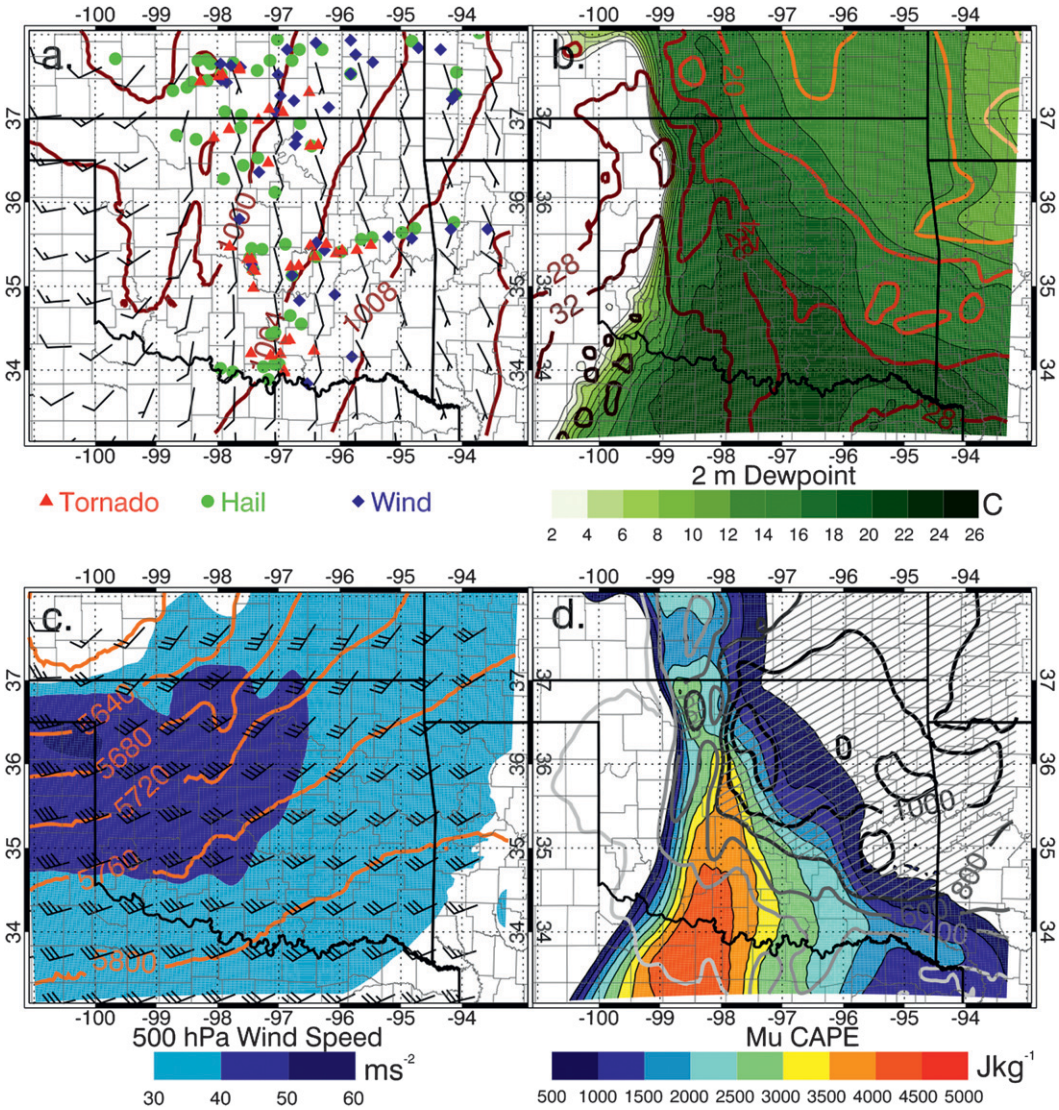

FIG. 1. (a) Tornado, hail, and wind severe weather reports over the southern Great Plains between 2100 UTC 10 May and 0200 UTC 11 May 2010 with surface wind speeds and direction and mean sea level pressure (hPa) contours at 2100 UTC overlaid; long wind barb = 10 m s^{-1} and short barb = 5 m s^{-1}. (b) Surface dewpoint and temperature (°C), (c) 500-hPa geopotential height field and wind velocity vectors (m s^{-1}), and (d) CAPE (J kg^{-1}) and 0–3-km SREH (m^2 s^{-2}); all at 2100 UTC. Hatched areas indicate locations where SREH exceeds 800 m^2 s^{-2}.

polar-orbiting satellite on 4 May 2002 with overpasses at approximately 0130 and 1330 local solar time (Aumann et al. 2003). The AIRS instrument measures infrared radiances at 2378 channels between 3.7 and 15.4 μm with a spatial resolution of 13.5 km at nadir and near 45 km at the limb with a 1650-km swathwidth, providing twice-daily coverage for most locations. AIRS retrieval algorithms resolve 1-km vertical-layer mean temperature and 2-km mean layer water vapor mixing ratio profiles from near the surface to the stratosphere as well as ozone and trace gas profiles (Susskind et al. 2003, 2006; Goldberg et al. 2003, 2004). As part of the profile retrieval, AIRS data are combined with microwave spectrum observations from the AMSU instrument (Aumann et al. 2003; Pagano et al. 2003), also located on *Aqua*. The lower resolution of the microwave instrument (45 km at nadir and up to 150 km at the limb) reduces the spatial resolution of the temperature and mixing ratio profiles created from the combined IR and microwave data. The primary advantage of AIRS over other sounding instruments (such as GOES and AMSU) is the improved vertical accuracy of the temperature and mixing ratio profiles made possible by the many more channels available from which to make retrievals (e.g., Susskind et al. 2003). The

current standard retrieval product is known as the level 2, version 5 AIRX with each file containing 6-min swath-level data (Olsen et al. 2007).

Retrieval quality is defined as the pressure level at which the quality of the retrieval becomes questionable. Below the reported layer (PBEST), it is not recommended that retrievals (if present) be used for data assimilation purposes. Temperature and mixing ratio uncertainty estimates at each pressure level for individual profiles are also provided. The cloud-clearing methods used by this algorithm allow for retrievals to be made for cloud fractions up to 80%, with somewhat decreased accuracy present in higher cloud fraction scenes. The design goal of the AIRS instrument is to provide 1-km-layer mean temperatures with a root-mean-square error (RMSE) of 1°C, and 2-km-layer mean mixing ratio RMSE of 20% for retrievals with up to 80% cloud cover (Susskind et al. 2003, 2006; Fetzer et al. 2003). Tobin et al. (2006), Divakarla et al. (2006) and Xubebao et al. (2005) have shown that observed errors are close to the design goals, with over-water retrievals having the lowest uncertainties.

One important consideration that must be addressed when assimilating AIRS profiles is their relationship to the observed atmospheric state. Whether we assume radiosonde or independent model analyses are most representative of this state, any biases in the AIRS temperature or mixing ratio profiles compared to these data could have a negative impact on the model forecasts. Calculating this bias from radiosonde data is difficult since very few nearby soundings occur close to the times of the AIRS retrievals over the CONUS. Previous research by Divakarla et al. (2006) and Pu and Zhang (2011) found that AIRS temperature retrievals between 850 and 250 hPa were 0.5°–1.0°C cooler than the nearby radiosonde measurements while a significant warm bias also exists near the tropopause height at roughly 200 hPa. The AIRS temperature retrieval averages comparatively warmer temperatures below and above (in the stratosphere), resulting in a 200-hPa temperature that is somewhat too warm. Due to this anomaly, retrievals above the tropopause height are not considered for assimilation. Moisture biases are somewhat harder to discern, but AIRS mixing ratio retrievals are somewhat drier in the 700–500-hPa layer and, also, slightly wetter above when compared to radiosondes. We must also consider that radiosonde measurements themselves are not perfect, especially for mixing ratio (McMillan et al. 2005). With the exception of filtering out retrievals above the tropopause, no additional bias correction is performed on the AIRS profiles prior to assimilation. The magnitude and uncertainties of previously documented biases make creating a robust bias correction impractical for this study.

4. Ensemble and data assimilation system

The NWP model selected for this data impact study is the Advanced Research core of the Weather Research and Forecasting Model version 3.1.1 (WRF-ARW; Skamarock et al. 2008). A 15-km domain covering most of the CONUS is created with 51 vertical levels stretching from the surface to 50 hPa. All ensemble members use a single set of physics options that include the Thompson microphysics scheme (Thompson et al. 2004); the Rapid Radiative Transfer Model (RRTM) for longwave radiation (Mlawer et al. 1997); the Dudhia scheme for shortwave radiation (Dudhia 1989); Eta Model similarity and the Noah land surface model for the surface layer and surface physics, respectively (Ek et al. 2003); the Mellor–Yamada–Janjić scheme for boundary layer physics (Janjić 2002); and the Grell–Devenyi approach for cumulus parameterization (Grell and Devenyi 2002).

WRF obtains initial and boundary conditions from the 1200 UTC run of the North American Mesoscale Model (NAM) from 10 May 2010 or 26 May 2009 at 3-h intervals. Fixed spatial and temporal covariance relationships drawn from random perturbations around a normal distribution are used to perturb the initial and boundary conditions for the 36 ensemble members (Torn et al. 2006). The data assimilation period for both cases starts at 1200 UTC and continues for 9 h until 2100 UTC and follows the process outlined by Wheatley and Stensrud (2010). Fujita et al. (2007) show that even a relatively short data assimilation period of 6 h can yield improved analyses and forecasts. New observations are assimilated at hourly intervals within a ±30-min window of each analysis time. Ensemble forecasts are then made that extend 15 h, beginning at 2100 UTC and ending 1200 UTC the following day.

The EnKF method is chosen to assimilate the observations for this study owing to past good performance in creating analyses with mesoscale structure (Zhang et al. 2004; Fujita et al. 2007; Stensrud et al. 2009). A parallel implementation of the filter within the Data Assimilation Research Test Bed (DART) software is used (Anderson and Collins 2007; Anderson et al. 2009) with a horizontal localization half-radius of 250 km and a vertical localization half-radius of 5 km. The vertical localization may seem large, but it is found to work best for the traditional observations assimilated due to the desire to have smoothly varying vertical adjustments to the analysis. Both horizontal and vertical localizations are constant for all observation types and no covariance inflation is applied. The DART package contains forward operators for all general atmospheric variables as well as for retrieved profiles from satellite data (Liu and Li 2010).

a. Experiment design

To assess the impacts of AIRS temperature and mixing ratio profile assimilation, two ensemble experiments are performed. One experiment assimilates the AIRS profiles (AIRSP) from satellite overpasses near 1830 and 2000 UTC, while the other experiment does not assimilate the AIRS profiles (NO-AIRSP). The AIRS mixing ratio is converted to dewpoint prior to assimilation for two reasons: assimilating small mixing ratio values in the upper troposphere sometimes leads to model instability, and we desire to maintain consistency with the traditional humidity variables that are also assimilated in dewpoint form. Otherwise, the experiments are identical and assimilate the same standard atmospheric observations from surface instruments, radiosondes, and aircraft. The AIRS data assimilated are the 45-km-resolution level-2 AIRS temperature and mixing ratio retrievals when the pixel-level cloud fraction is less than 50% between the surface layer and up to 250 hPa, where the best quality retrievals are found. Otherwise, no horizontal or vertical data thinning is performed. Testing (not shown) indicates that AIRS retrievals near and just above the tropopause layer significantly increase model errors. Results from both the ensemble mean and individual members are compared to determine whether or not the differences between the AIRSP and NO-AIRSP runs are consistent from member to member and case to case.

b. Nested-grid characteristics

To better resolve the convective features of these events, one-way nested-grid ensembles at 3-km horizontal resolution are generated for both experiments over a domain centered over each severe weather outbreak. The nested domain for 10 May 2010 contains southern Kansas, Oklahoma, and north Texas, and encompasses the majority of the severe weather reports that occurred (Figs. 1 and 2). The nested grid for the 26 May 2009 case is shifted southward to better capture the convection in north-central Texas (Fig. 2). No data assimilation is performed on the nested domain and the hourly mesoscale (15 km) forecasts act as boundary conditions. The only change made to the model settings is that the convective parameterization used on the mesoscale domain is turned off, as it is not needed for the higher-resolution grid. The nested-domain forecasts start at 2100 UTC and extend to 0600 UTC and are output at 15-min intervals. One limitation of this high-resolution ensemble is that the simulated storms must "spin up" before they show up in the forecast, since they are not accurately depicted in the 15-km analyses, which can take an hour or longer.

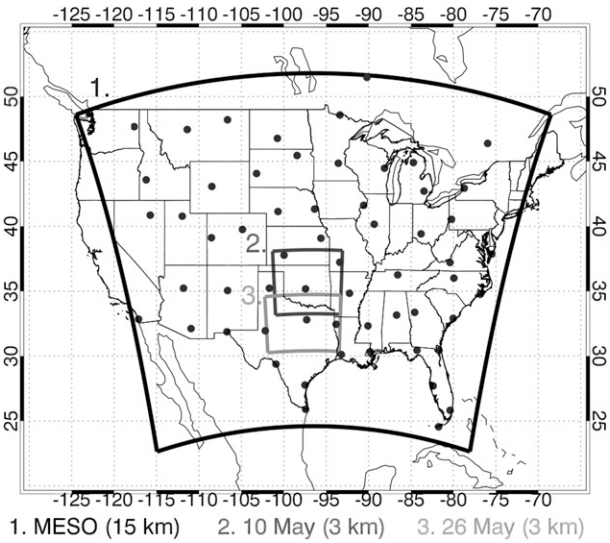

FIG. 2. Mesoscale (15 km) and convective-scale model domains (3 km) used in this study. Locations of radiosonde stations used for verification of the mesoscale temperature and dewpoint forecasts are overlaid. Note that the location of the convective-scale domain differs between the 10 May 2010 and 26 May 2009 cases (case 2 and 3, respectively). Otherwise, all other domain characteristics are identical between the two.

c. Assimilated observations

The spatial distribution of assimilated data between 1200 and 2100 UTC 10 May 2010 for both the AIRSP and NO-AIRSP model ensembles (Fig. 3) shows that observations from land surface instruments (aviation routine weather reports, METARs) make up the largest source of data (~19 000 observations each for temperature and dewpoint and ~18 000 for horizontal wind components and altimeter) and account for 63% of the total NO-AIRSP dataset (Table 1). Buoy and ship observations account for another 5% of the data. Combined, surface observations account for over two-thirds of the data being assimilated into WRF-DART after 1200 UTC. The remainder of the dataset represents upper-air observations that primarily originate from Aircraft Communications, Addressing, and Reporting System (ACARS) measurements. ACARS data include temperature, winds, and sometimes dewpoint along an aircraft flight track. To reduce redundant information in the vicinity of airports and prevent overfitting of the model to the observations, ACARS observations are superobbed by averaging all observations within a 45-km horizontal and 25-hPa vertical resolution box at hourly time intervals, following Torn and Hakim (2008). Between the surface and 200 hPa, approximately 15 000 temperature and wind observations are assimilated, though ACARS moisture observations are less frequent, only

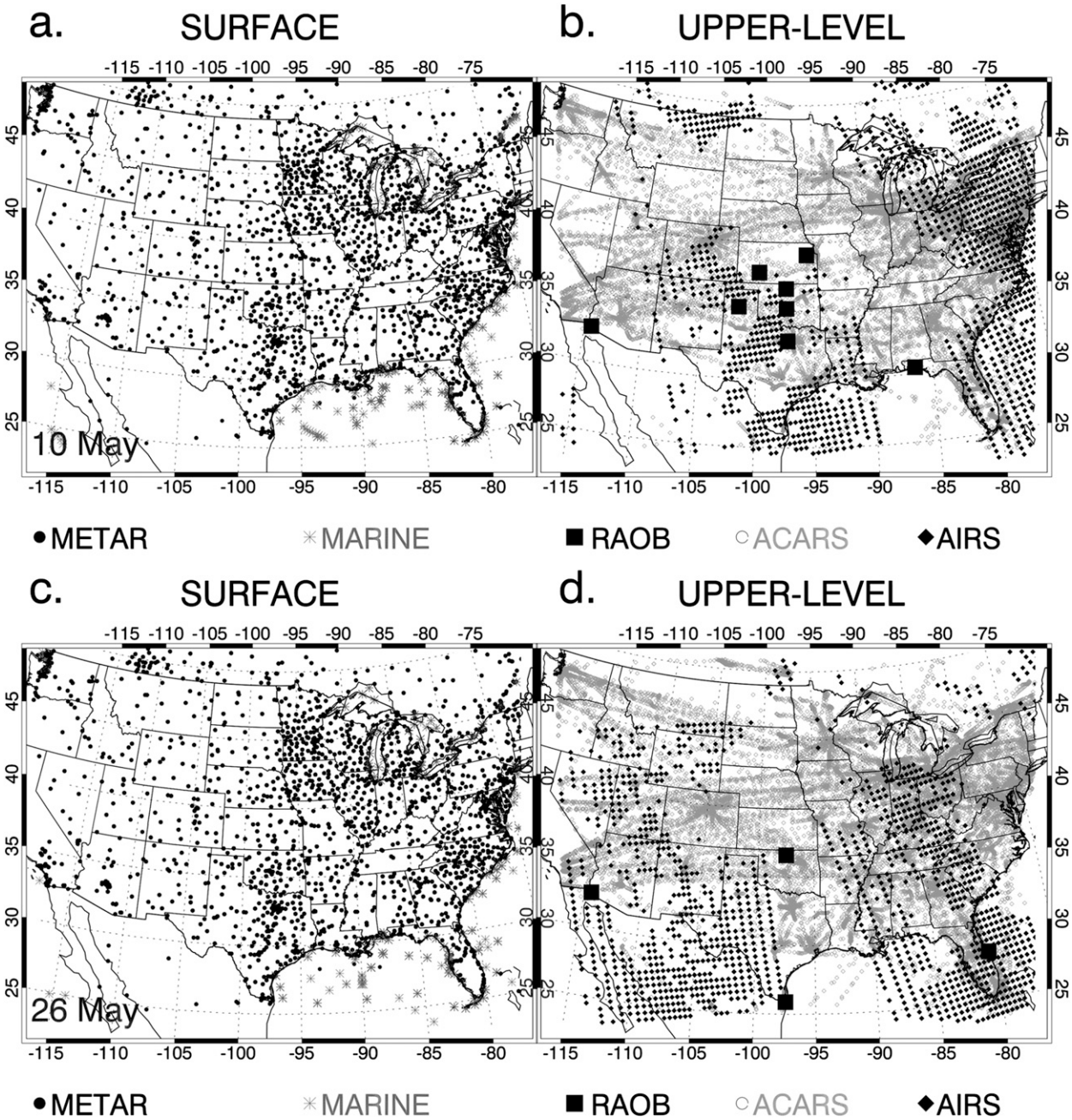

FIG. 3. (a),(c) Surface and (b),(d) upper-air observations assimilated into WRF-DART between 1200–2100 UTC 10 May 2010 and 26 May 2009. Note the large number of data points that are added by assimilating AIRS profiles. Plotted AIRS data points represent the center of each 45-km-resolution pixel being assimilated. AIRS observations over the eastern United States occur at approximately 1830 UTC, while those in the central United States come from the following orbit at ~2000 UTC. Sample sizes for each data type are listed in Table 1.

numbering approximately 3500. While ACARS data are relatively plentiful, the majority of the midtropospheric observations are located near major airports and are measured during the ascent or decent phases of a flight. Radiosonde data make up the smallest portion of data being assimilated, accounting for less than 1% of the total observations. Note that 1200 UTC radiosondes are not assimilated directly, but are assumed to be a part of the initial and boundary conditions provided by the NAM analysis.

AIRS observations occur at approximately 1830 UTC along the east coast of the United States and at

TABLE 1. Assimilated observation sample sizes for each data type used in WRF-DART for AIRSP and NO-AIRSP ensembles on 10 May 2010 and 26 May 2009. Assimilating AIRS profiles adds nearly 12 000 additional observations compared to only using traditional data types.

Observation type	10 May 2010		26 May 2009	
	N_{NOAIRS}	N_{AIRS}	N_{NOAIRS}	N_{AIRS}
RADIOSONDE_U_WIND_COMPONENT	94	94	65	65
RADIOSONDE_V_WIND_COMPONENT	96	96	65	65
RADIOSONDE_TEMPERATURE	97	98	61	61
RADIOSONDE_DEWPOINT	76	77	36	36
ACARS_U_WIND_COMPONENT	14 464	14 463	18 178	18 182
ACARS_V_WIND_COMPONENT	14 467	14 464	18 197	18 189
ACARS_TEMPERATURE	15 542	15 543	20 060	20 059
ACARS_DEWPOINT	3474	3474	4751	4746
MARINE_SFC_U_WIND_COMPONENT	2010	2010	1610	1610
MARINE_SFC_V_WIND_COMPONENT	2011	2010	1613	1613
MARINE_SFC_TEMPERATURE	1819	1819	1540	1540
MARINE_SFC_DEWPOINT	580	580	620	6260
MARINE_SFC_ALTIMETER	1461	1461	1266	1266
METAR_U_10_METER_WIND	17 963	17 963	23 457	23 457
METAR_V_10_METER_WIND	17 967	17 967	23 457	23 457
METAR_TEMPERATURE_2_METER	19 391	19 391	24 566	24 562
METAR_DEWPOINT_2_METER	18 852	18 852	24 305	24 303
METAR_ALTIMETER	18 578	18 567	23 695	23 689
AIRS_TEMPERATURE	0	5723	0	5753
AIRS_DEWPOINT	0	6272	0	6143
Total	148 942	160 924	187 542	199 416

approximately 2000 UTC in the midwestern states (Fig. 3b). The ~90-min difference from the East Coast to the Midwest represents the time required from one orbit to the next. From north to south during the same orbit, the time difference is generally less than 15 min over the CONUS. Several regions exist where no AIRS profiles are assimilated. In many cases no best quality retrievals are made, or they are removed as part of the cloud-screening process and are not included in the data assimilation stream. Approximately 30% of the best quality observations for temperature and mixing ratio are rejected during the assimilation process mainly due to being outliers when compared to the background. The remaining 5723 temperature and 6272 mixing ratio observations on 10 May are assimilated. Compared to ACARS, the number of temperature observations from AIRS is not impressive, though AIRS dewpoint observations exceed those available from ACARS (Table 1). Data from the latter swath are key, since they can influence the environments near the regions of deep convection. Upstream of the 10 May event, ACARS observations are somewhat sparse in central and western Texas whereas AIRS profiles are available over much of this region. A similar number of AIRS profiles is assimilated in the 26 May 2009 case with the swaths located somewhat west of their locations on 10 May 2010 (Table 1, Fig. 3). As with 10 May, ACARS observations upstream of the event in west Texas are rather sparse compared to what is available from AIRS (Fig. 3d).

For each temperature and mixing ratio retrieval, the retrieval algorithm also estimates the uncertainty of that observation. Since these uncertainties vary as a function of atmospheric conditions, cloud cover, and height above the surface, they are used to define the error characteristics of the AIRS profiles in lieu of an assumed constant observation error for this product. The average AIRS temperature observation error calculated from this dataset on 10 May over all levels is 1.1°C, which compares favorably to assumed values of 2.5° and 1.0°C for METAR and ACARS temperature errors, respectively. Larger differences exist for dewpoint where the corresponding AIRS and ACARS errors are 7.9° and 1.9°C, respectively. The magnitude of the AIRS dewpoint error represents a limiting factor on how much AIRS profiles can improve analysis and forecast fields.

To assess whether the assimilation of AIRS profiles is successful, the vertical profiles of the observation diagnostics are examined before and after assimilation. Diagnostics include bias, RMSE, and total spread, with each calculated by comparing the assimilated AIRS profiles with the prior or posterior model fields generated by WRF-DART. The total bias for temperature is less than ±0.5°C between the surface and 400 hPa for both prior and posterior assimilation fields for the 10 May case (Fig. 4a). Bias decreases slightly after AIRS assimilation at all levels, but the magnitude of this decrease is small (0.1°C). The decrease in RMSE is larger and present at

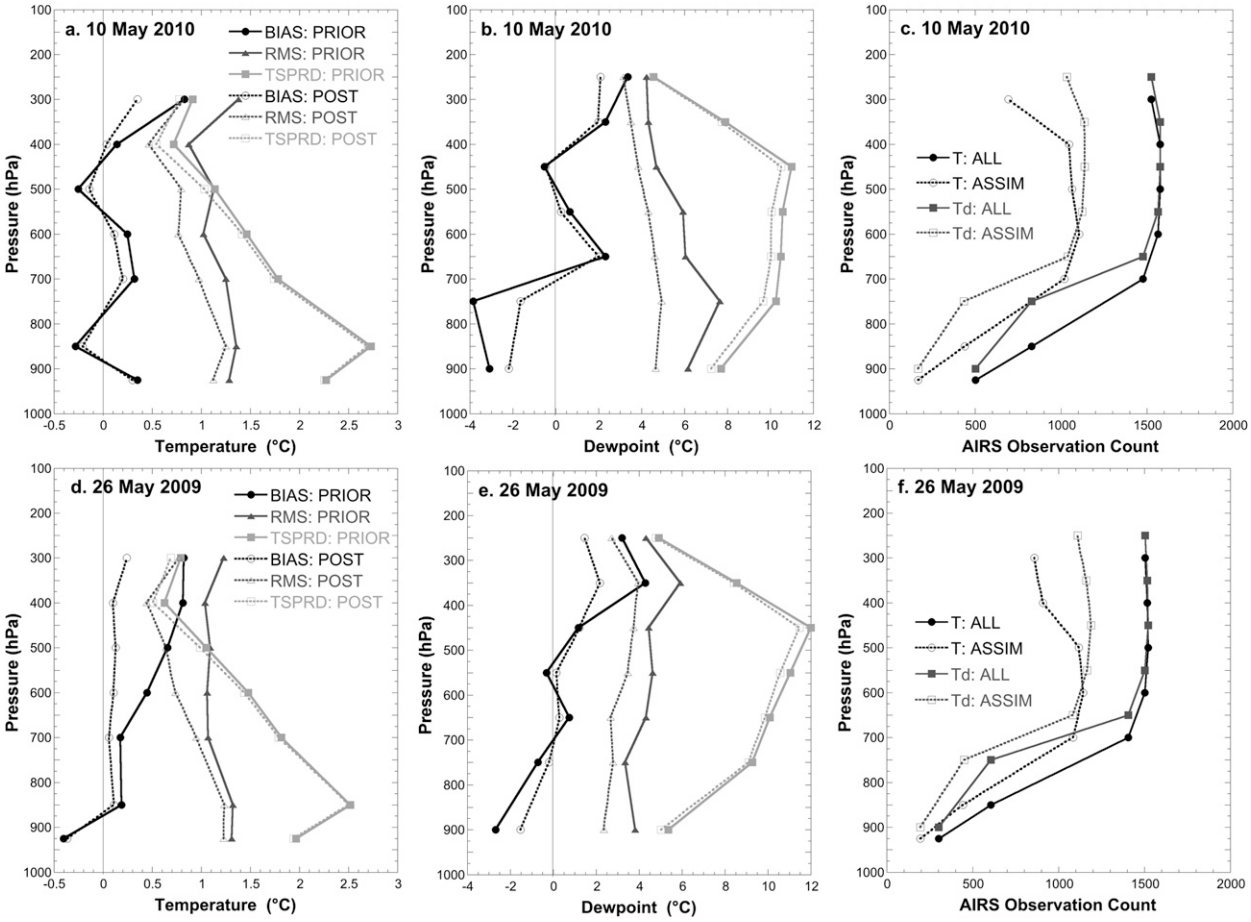

FIG. 4. Observation diagnostics comparing temperature and dewpoint fields with AIRS observations before (PRIOR) and after (POSTerior) assimilations for 10 May 2010 and 26 May 2009 over the mesoscale domain. (a) Temperature bias, RMSE (abbreviated RMS in the caption), and total spread (TSPRD) for the prior (solid lines) and posterior (dotted lines) analysis states at selected pressure levels. Error statistics decrease after assimilation at all levels. (b). Diagnostics for dewpoint. (c) Temperature and dewpoint sample sizes as a function of pressure level. Solid lines indicate all observations passing AIRS quality control thresholds while dotted lines represent the number of observations actually assimilated. (d)–(f) As in (a)–(c), but for the 26 May 2009 case.

all levels where AIRS data are assimilated (Fig. 4a). Corresponding examination of the dewpoint diagnostics shows similar results (Fig. 4b) except the magnitude of the bias, RMSE, and total spread are all much larger than for temperature. Despite the large uncertainties in AIRS mixing ratio profiles, their assimilation consistently reduces dewpoint error and spread. Both temperature and mixing ratio profiles are assimilated from near the surface to about 300 hPa, with the greatest number of observations occurring in the 700–300-hPa layer (Fig. 4c). The number of assimilated AIRS temperature observations exceeds those for the dewpoint below 550 hPa while the opposite is true above this layer. Similar results are present for the 26 May case (Figs. 4d–f) though the assimilation of temperature profiles (Fig. 4d) appears to have a greater impact compared to the 10 May case. Overall, the large number of AIRS profiles assimilated coupled with the reduction in both temperature and dewpoint analysis errors at most levels for both cases strongly indicates a successful assimilation is occurring despite the large observation errors sometimes associated with the AIRS data.

5. AIRSP versus NO-AIRSP forecast differences

a. CONUS domain (15 km)

The greatest differences between the AIRSP and NO-AIRSP temperature and dewpoint ensemble mean analyses at 500 hPa for 2100 UTC 10 May 2010 exist near the satellite data swaths where the AIRS profiles are assimilated (Figs. 3 and 5a,b). The largest differences occur in the northeastern CONUS, associated with the ~1800 UTC overpass with a second large area of differences over

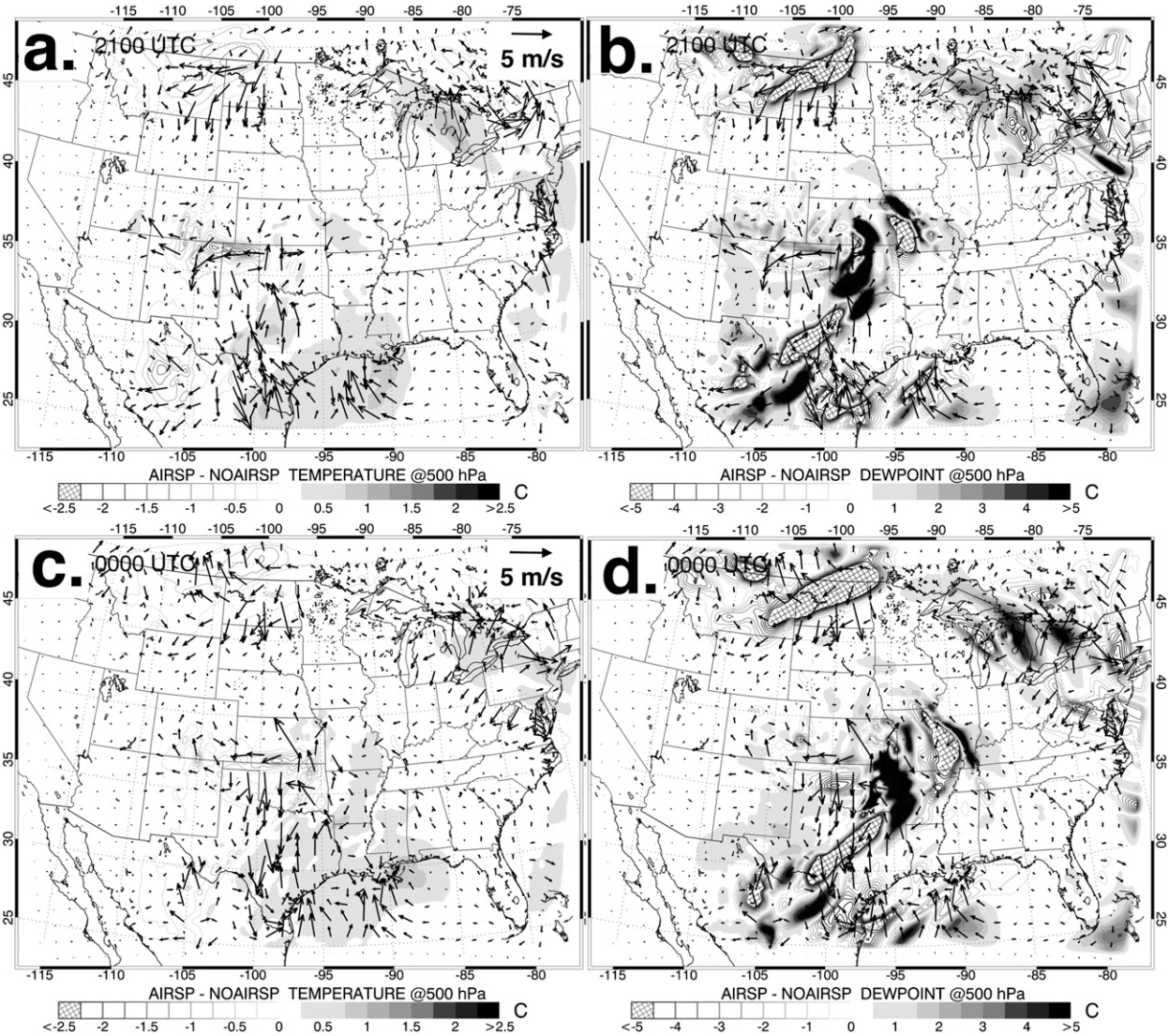

FIG. 5. Differences (AIRSP − NOAIRSP) of (a),(c) ensemble mean temperature and (b),(d) dewpoint at 500 hPa for 2100 UTC 10 May and 0000 UTC 11 May 2010. Positive (shaded) values indicate AIRSP is warmer (moister), with negative values (line contours and hatching) indicating AIRSP is cooler (drier) than the NO-AIRSP ensemble mean. Arrows display the velocity-vector difference between the ensemble means. A leftward-pointing arrow indicates that AIRSP has a slower eastward component of the wind. A downward-pointing vector indicates that AIRSP has a slower northward component.

the southern Great Plains that is associated with the ~2000 UTC overpass (Figs. 5a,b). In both cases, the magnitude of the dewpoint differences (±5°C) exceeds those for temperature (±1°C) even though the patterns remain similar. The 3-h forecast at 0000 UTC yields similar patterns, which are displaced eastward compared to the analysis at 2100 UTC (Figs. 5c,d). For both times, velocity vector differences of up to 5 m s^{-1} exist in the areas where the largest differences in dewpoint are present. Similar temperature, dewpoint, and velocity differences are also evident in the 26 May case (not shown).

Ensemble mean 3- and 15-h forecast temperature and dewpoint profiles are compared against nearby radiosonde observations at 0000 and 1200 UTC 11 May 2010 and 0000 UTC 27 May 2009 to determine whether the inclusion of AIRS data improves the forecasts within the CONUS domain. At 0000 UTC 11 May, profile data from up to 63 radiosonde stations are available for comparison with 67 stations available on 27 May 2009. Ensemble mean temperature and dewpoint forecasts are averaged within a radius 15 km around each radiosonde with drift as a function of height not considered. Bias (model − radiosonde) and RMSE are then computed

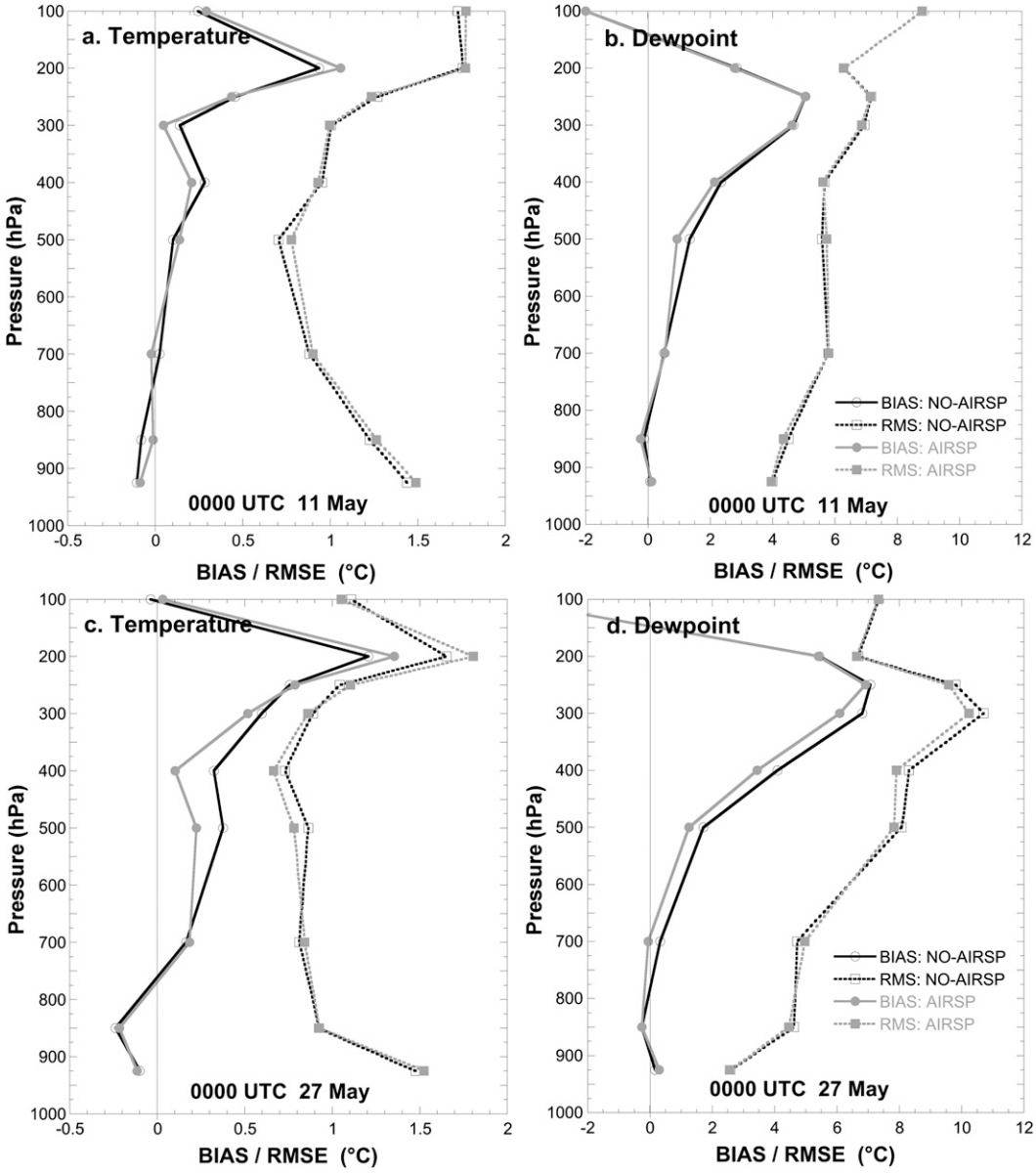

FIG. 6. The vertical profiles of (a) temperature and (b) dewpoint bias and RMSE averaged over a maximum of 63 radiosonde stations in the United States at 0000 UTC 11 May 2010 and the corresponding AIRSP and NO-AIRSP 3-h ensemble mean forecasts. Positive values indicate a warm or moist bias of the model compared to radiosonde data. Negative values indicate a cold or dry bias. (c),(d) As in (a),(b), but for the 26 May 2009 case averaged over a maximum of 67 stations.

from just above the surface (925 hPa) to above the tropopause (100 hPa).

The temperature RMSEs at 0000 UTC 11 May are less than 1.5°C between 925 and 250 hPa for both ensembles with minimum errors occurring at 500 hPa (Fig. 6a). The corresponding temperature biases are quite low, being less than ±0.3°C up to 300 hPa. Near the tropopause (200 hPa) the ensemble forecasts have a noticeable warm bias. No large differences in either bias or RMSE exist between AIRSP and NO-AIRSP at this time. Dewpoint RMSEs from both experiments increase from 4°C near the surface to 6°C between 700 and 400 hPa (Fig. 6b). Between 700 and 200 hPa, both NO-AIRSP and AIRSP are too moist compared to the radiosonde observations. No significant difference exists in the ensemble mean dewpoint forecasts when averaged over all radiosonde sites at 0000 UTC 11 May. Similar vertical profiles of bias and RMSE exist at 0000 UTC 27 May 2009 when

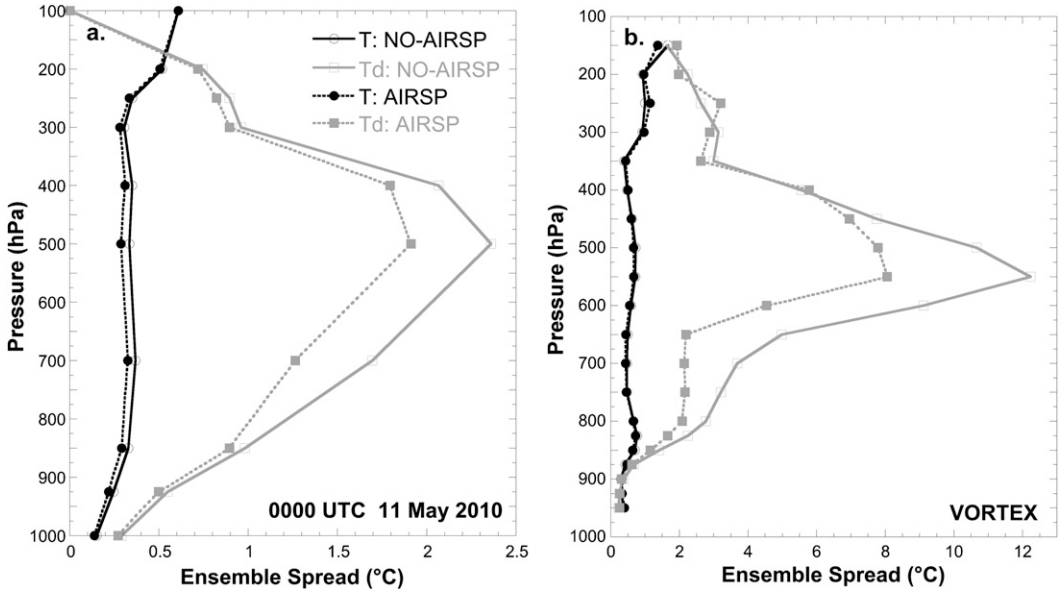

FIG. 7. (a) The vertical profile of ensemble spread at 0000 UTC 11 May 2010 of temperature (black) and dewpoint (gray) considering forecasts from all 36 members averaged over the 63 available radiosonde locations. (b) As in (a), but averaged over six VORTEX2 observations between 2145 and 2300 UTC 10 May from nested-grid forecasts.

comparing 3-h forecasts of temperature and dewpoint to radiosonde observations (Figs. 6c,d). The primary differences are that AIRSP appears to reduce the 500- and 400-hPa temperature bias and RMSE for this case compared to negligible or no reduction on 11 May (Figs. 6a,c). Slight improvements in the dewpoint bias and RMSE are also apparent at 0000 UTC 27 May. Results from the 15-h forecasts for both cases are similar, except the RMSE values are larger.

The ensemble spread of temperature or dewpoint surrounding a radiosonde defines the variability, or spread, of each forecast parameter. Temperature spread averaged across the CONUS is low on average, being less than 0.5°C below 100 hPa at 0000 UTC 11 May (Fig. 7a). For dewpoint, the spread increases to 2°C in the midtroposphere, indicating greater variability in atmospheric moisture forecasts. Differences between AIRSP and NO-AIRSP ensembles for temperature are less than 0.1°C, with AIRSP consistently having smaller ensemble spread. Greater differences exist for dewpoint with the ensemble spread reduced by 0.2°–0.5°C between 800 and 400 hPa when assimilating AIRS data. Analyzing a temperature-only AIRSP assimilation experiment verifies that the reduction in dewpoint spread is primarily a function of assimilating AIRS mixing ratio profiles as the dewpoint ensemble spread is only slightly less than NO-AIRSP at all vertical levels in that experiment (not shown). The reduction in dewpoint spread represents the only consistent difference between the AIRSP and NO-AIRSP forecasts when averaged over the CONUS domain. Similar results are found when examining the ensemble spreads at 0000 UTC 27 May 2009 (not shown).

b. *Southern plains domains (3 km)*

To examine the effects of AIRS profile assimilation on temperature and dewpoint forecasts on a smaller scale, we focus our analysis on the southern plains. The 15-km analysis is interpolated down to a 3-km grid with high-resolution forecasts generated beginning at 2100 UTC. At 2100 UTC 10 May, the 700-hPa ensemble mean temperatures from AIRSP are slightly cooler than those from NO-AIRSP over the 3-km domain, with the magnitude of these differences generally less than 0.5°C (Fig. 8a), indicating little impact on the AIRS profiles from temperature. More notably, the AIRSP ensemble is at least 5°C moister in a narrow swath extending southwest to northeast through central Oklahoma (Fig. 8b). This area corresponds to the location of a midtropospheric dry tongue in the NO-AIRSP analysis, and assimilating AIRS data removes this feature. AIRSP is also slightly drier (<3°C) to the east and west of this region compared to NO-AIRSP. These features persist at higher levels (~500 hPa), where they are displaced somewhat to the east and the magnitude of the dewpoint difference is larger (Figs. 8c,d). Maximum wind speed differences between the AIRSP and NO-AIRSP analyses are generally less than 5 m s^{-1}.

To determine whether these analysis differences improve the forecasts of temperature and dewpoint in the 3-km domain, six special radiosonde observations from

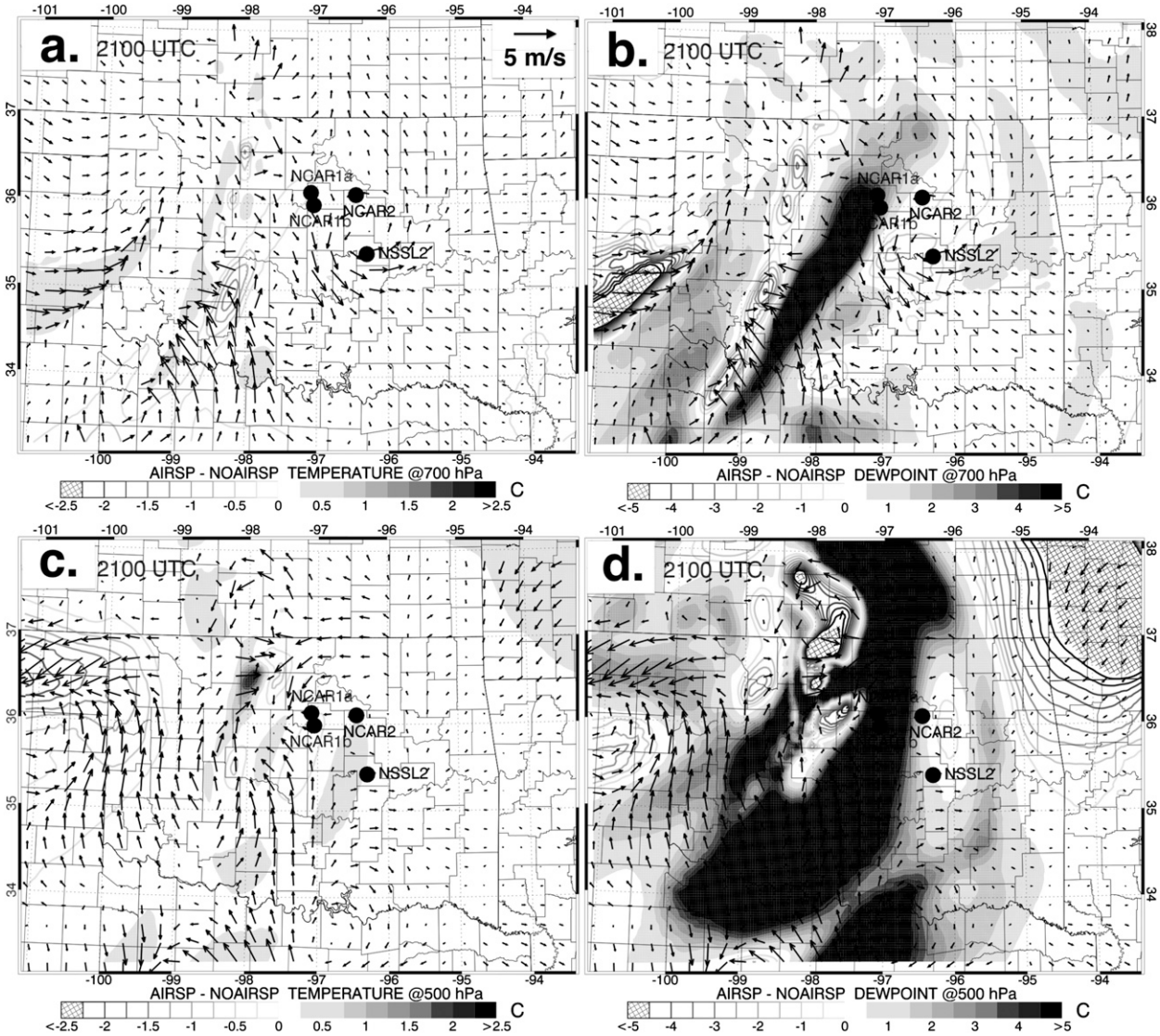

FIG. 8. Differences (AIRSP − NOAIRSP) of ensemble mean temperature and dewpoint in the nested-grid (3 km) domain at (a),(b) 700 and (c),(d) 500 hPa at 2100 UTC 10 May. Locations of special VORTEX2 radiosonde sites are denoted by NCAR1a, NCAR1b, NCAR2, and NSSL2, respectively. Otherwise, similar to Fig. 5.

the VORTEX2 experiment are compared with the AIRSP and NO-AIRSP forecasts. The locations and times of each observation are listed in Table 2 and shown in Fig. 8 and are concentrated in the preconvective environment of north-central Oklahoma between 2145 and 2258 UTC. Averaging over all six VORTEX2 sites, a warm bias exists in the model forecasts between 850 and 775 hPa and also between 350 and 250 hPa, with a cold bias present between 750 and 650 hPa (Fig. 9a). The magnitude of these biases is small, being near or below 1°C for most levels and most ensemble members. At 800 hPa, the inclusion of AIRS profiles reduces the ensemble mean warm bias from 1.5 to 1.1°C, with 31 out of 36 ensemble members showing at least some decrease in bias. At 700 hPa, the cold bias for AIRSP is larger by 0.2°C compared to NO-AIRSP. However, significant member-to-member variability also exists, with NO-AIRSP performing better for 25 members. At 300 hPa, AIRSP produces no improvement in the warm bias, but again the member-to-member variability is rather large compared to the differences between the ensemble mean biases.

Much larger differences are found between the AIRSP and NO-AIRSP ensemble mean dewpoint forecasts. A dry bias exists in the model forecasts between 850 and 750 hPa and also between 650 and 500 hPa. Above 500 hPa and up to 250 hPa, both models are too moist compared to the observations (Fig. 9b). The dewpoint

TABLE 2. Locations and times for special radiosondes launched into the prestorm environment between 2145 and 2300 UTC 10 May 2010 as part of the VORTEX2 field campaign. Locations are plotted in Fig. 8.

Station name	Lat (°N)	Lon (°W)	Time (UTC)
NCAR1a	36.131	−97.092	2145
NSSL2a	35.424	−96.298	2156
NCAR2a	36.106	−96.450	2157
NCAR2b	36.105	−96.450	2251
NSSL2b	34.423	−96.298	2254
NCAR1b	35.987	−97.051	2258

biases at 800 and 550 hPa for NO-AIRSP are −7.6° and −14.7°C, which reduce to −4.9° and −10.5°C, respectively, in AIRSP with improvements seen in 35 members. The dry bias exhibited by both ensembles is also evident from analyzing total precipitable water (TPW). When averaged over all VORTEX2 radiosondes, the dry biases are −3.9 and −2.3 mm for NO-AIRSP and AIRSP, respectively. The AIRSP ensemble is moister overall at all six locations due to the moistening of the low- to mid-tropospheric analysis from the assimilation of AIRS profiles, though even the AIRSP ensemble remains too dry.

RMSE statistics generally follow the same pattern as that observed for bias. The differences in RMSE for temperature at all levels are small (<0.2°C) (Fig. 7b). The same cannot be said for the differences in dewpoint RMSE between NO-AIRSP and AIRSP. A reduction in RMSE exists corresponding to the reduction in the dry bias between 850 and 500 hPa (Fig. 9b). At all levels, the ensemble variability is lower for AIRSP within this domain. Temperature spread decreases between 0.1° and 0.2°C below 500 hPa with dewpoint spread decreasing at least 1°C over these same levels (Fig. 7b). This represents a similar result to that observed for the larger CONUS domain, but the differences within this small region are greater. A similar comparison of ensemble mean winds with VORTEX2 radiosonde observations shows no significant differences in bias or RMSE statistics when averaging over these sites.

Examining individual radiosonde observations shows that a moist boundary layer exists between 950 and 800 hPa (Fig. 10). Above this layer, a strong capping inversion is present with a temperature increase of up to 5°C observed with a corresponding large drop in atmospheric humidity. Above the inversion, lapse rates are nearly dry adiabatic to approximately 600 hPa and approach a more moist-adiabatic rate above this level. Atmospheric moisture content generally remains low, but with moist layers being observed when the radiosonde ascends through cloud layers, some of which likely represent cirrus outflow from the severe storms west of the measurement locations. Comparing 3-km AIRSP and NO-AIRSP ensemble forecasts averaged over a radius ±5 km from the observation sites at the forecast times nearest the observations (at most 8 min) shows an overall good level of agreement between the observations and model fields. Several important differences exist with the most obvious being that the models all underestimate the height of the capping inversion to some degree. The largest difference occurs at the National Center for Atmospheric Research 2a (NCAR2a) site at 2200 UTC

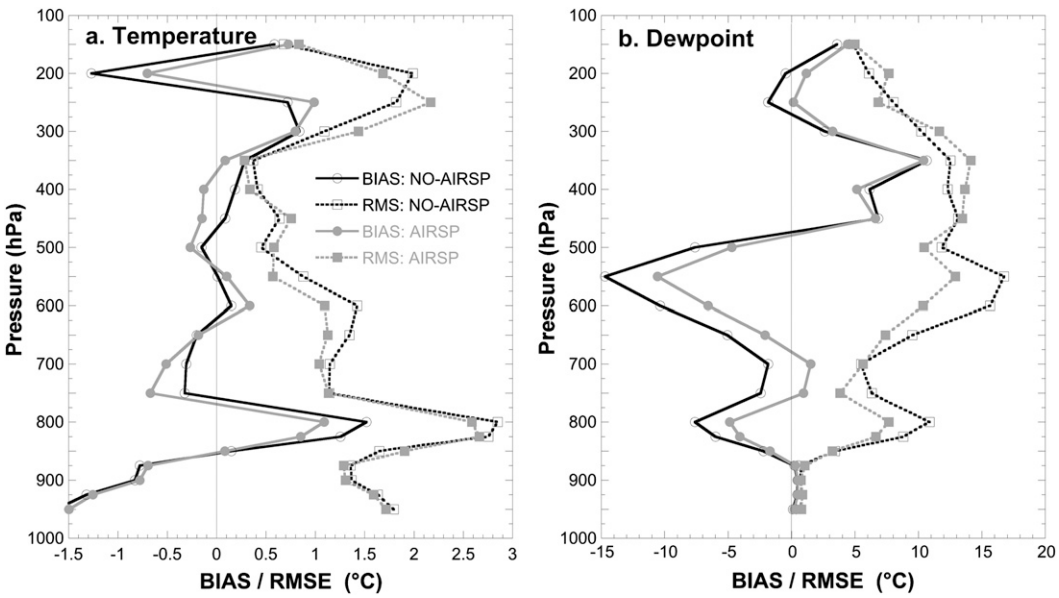

FIG. 9. Similar to Fig. 6, but computed from 3-km data in the nested grid averaged over six VORTEX2 soundings between 2145 and 2300 UTC 10 May.

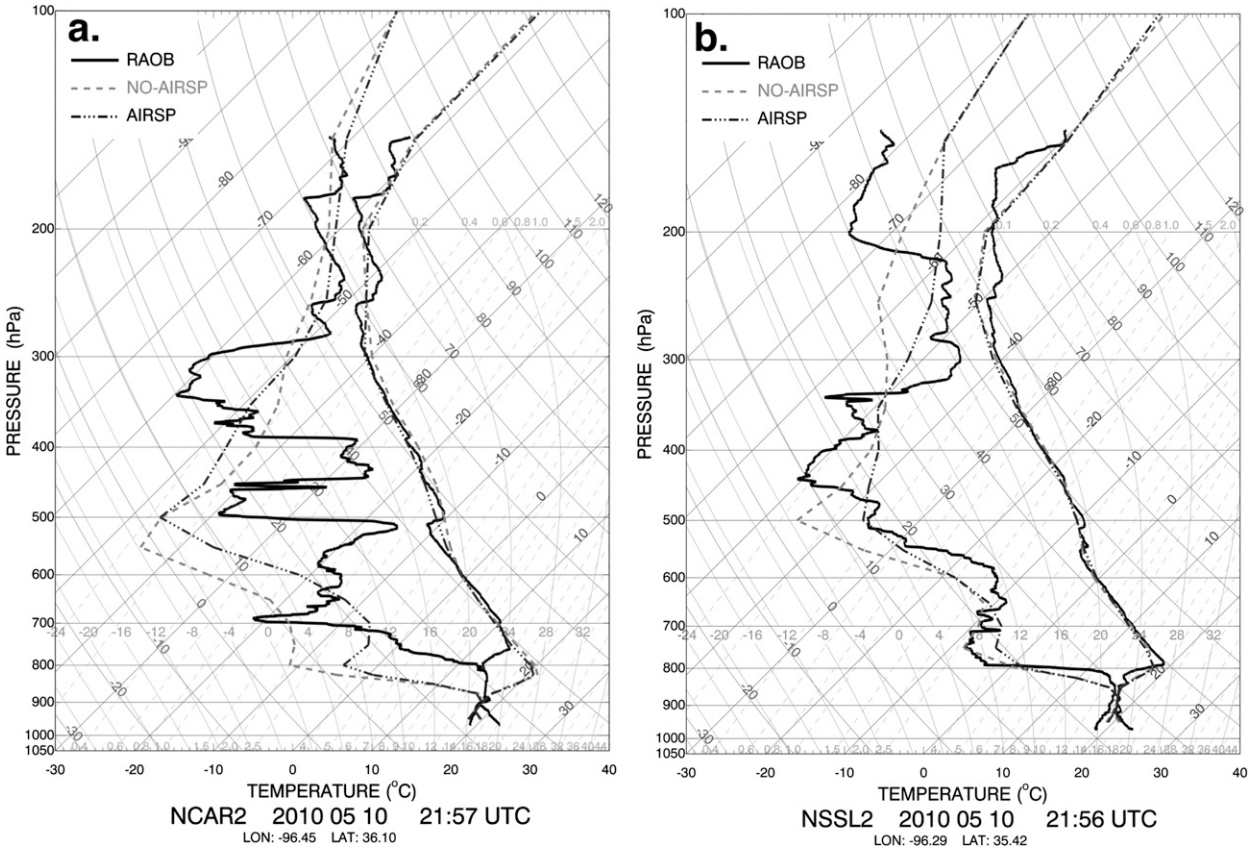

FIG. 10. Temperature and dewpoint profiles (solid lines) from radiosondes at approximately 2200 UTC 10 May at the (a) NCAR2a and (b) NSSL2a observation sites. See Fig. 8 for locations. Ensemble mean profiles from NO-AIRSP (light gray, dashed lines) and AIRSP (dark gray, dash–dot lines) at 2200 UTC are overlaid.

(Fig. 10a). Here, the modeled capping inversion is located approximately 70 hPa too low. However, the model temperature profile quickly recovers to within ±2°C of the observations above the capping inversion. No significant difference exists between AIRSP and NO-AIRSP temperature profiles, indicating that AIRS data are not having a large impact on the resulting mid- to upper-level temperature forecasts. An experiment where only AIRS temperature is assimilated yields results that are not substantially different from NO-AIRSP at these locations (not shown).

More interesting differences exist between modeled and observed dewpoint profiles. While both NO-AIRSP and AIRSP dewpoint profiles follow the same general trends, they do not capture the finer-scale changes in moisture present in the VORTEX2 measurements. For example, the ensemble mean dewpoint profiles near the NCAR2a site at 2200 UTC (which represents the nearest observation to the oncoming convection) do not appear to capture the cloud layers present aloft indicated by the nearly saturated layers in the observations (Fig. 10a). However, the inclusion of AIRS profiles tends to improve the forecast dewpoints at these two sites. The NO-AIRSP dewpoint profile at the NCAR2a site at 2200 UTC is generally too dry compared to the observations between the boundary layer and 400 hPa (Fig. 10a). Assimilating AIRS profiles moistens the environment, moving the dewpoint profile closer to the observations between 800 and 500 hPa. Overall, the forecast still remains too dry compared to the observations. Similar improvements are observed at the National Severe Storms Laboratory 2a (NSSL2a) site at 2200 UTC though the AIRSP dewpoint between 750 and 650 hPa becomes somewhat too moist compared to the observations (Fig. 10b). Since cloud-layer moisture is known to influence deep convection (Ferrier et al. 1996; Gilmore and Wicker 1998; Tompkins 2001), the AIRS data assimilation is expected to impact the subsequent ensemble forecasts of convection on the 3-km domain.

Overall, the comparisons between ensemble short-term forecasts with VORTEX2 radiosonde observations indicate that the temperature profiles are in excellent agreement with the observations, with the most significant difference being an underestimation of the capping

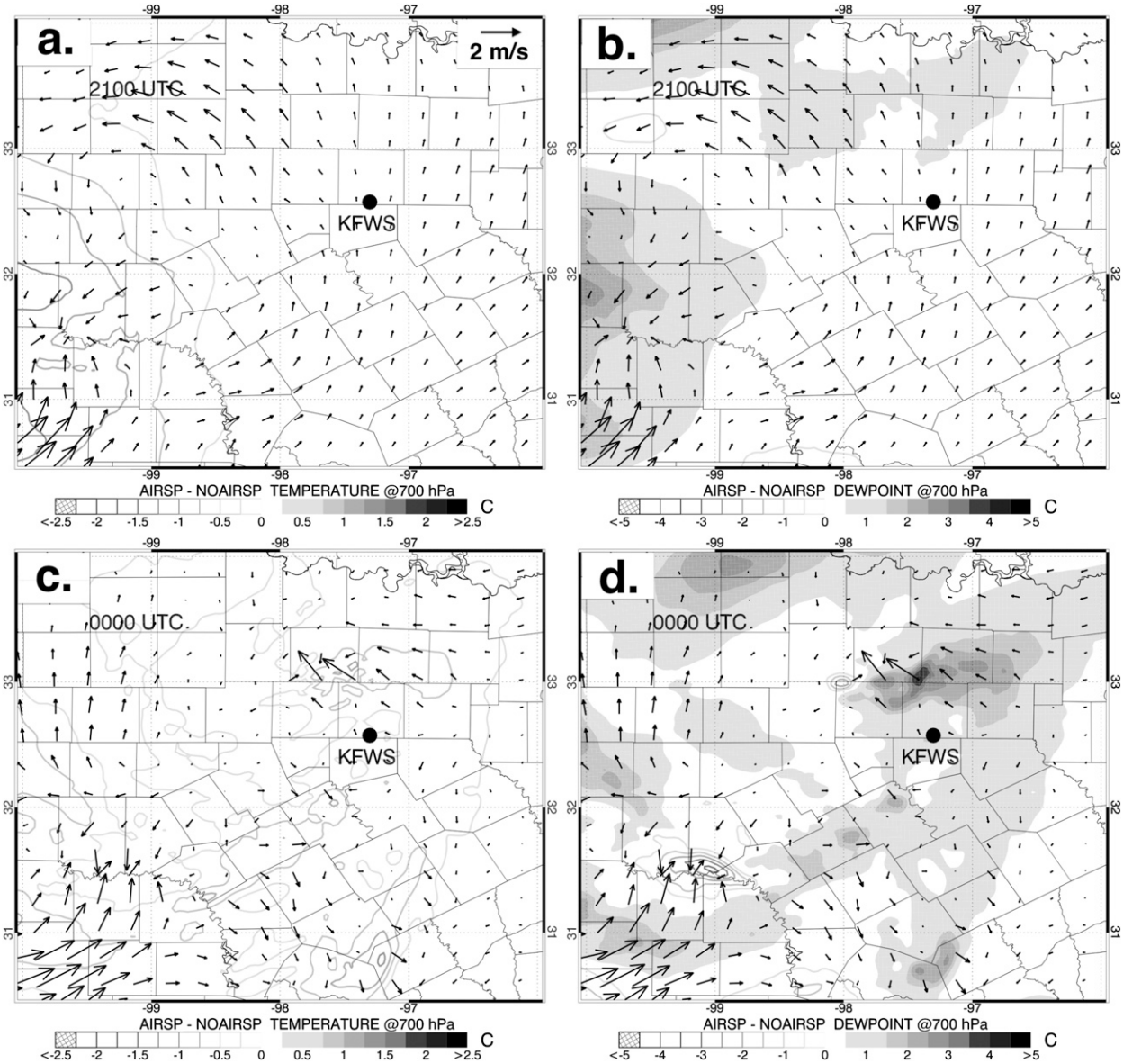

FIG. 11. Differences (AIRSP − NOAIRSP) in ensemble (a),(c) mean temperature and (b),(d) dewpoint at 700 hPa in the nested-grid (3 km) domain at (a),(b) 2100 UTC 26 May and (c),(d) 0000 UTC 27 May.

inversion height by WRF. Agreement between dewpoint profiles is not as good owing to the smoother moisture profile present in the model analyses. No significant difference between NO-AIRSP and AIRSP temperature analyses exists, though dewpoint differences up to 10°C are present with AIRSP showing better agreement with the VORTEX2 observations. This improvement in the moisture field should provide a positive impact to the forecasts of convection.

Unfortunately, the nature of the VORTEX2 radiosonde observations for the 26 May 2009 case limits the potential for accurate validation of the forecast temperature and dewpoint profiles. The few upper-level observations that are available occur very near (<5 km) or inside the ongoing convection preventing a comparison between larger-scale model forecasts and the surrounding environment. However, the impact of assimilating AIRS profiles can be seen in the differences in temperature and dewpoint analyses and forecasts between NO-AIRSP and AIRSP ensembles. At 2100 UTC, the effects of assimilating AIRS profiles are evident with AIRSP generating a cooler, moister analysis west and southwest of Fort Worth, Texas (approximated by the location of the nearby radar, KFWS) (Figs. 11a,b). The magnitude

of these differences at 700 hPa is greatest where profiles are being assimilated in the western portion of the 3-km domain with temperature decreasing up to 2°C and dewpoint increasing up to 3°C (Figs. 3c and 11a,b). Also present is an increase in the north and eastward components of the wind in the southwest and an increase in the western component to the northwest (Fig. 11a). The analysis differences have a significant impact on the forecasts as shown by the 0000 UTC 27 May temperature and dewpoint difference plots (Figs. 11c,d). The temperature forecast from AIRSP at 700 hPa is generally 0.5°–1.5°C cooler in the western half of the domain as compared to NO-AIRSP. AIRSP also continues to generate a moister environment, with dewpoints increasing 0.5°–2.0°C over large areas of the domain (Fig. 11d). The largest increase in dewpoint occurs north of KFWS and is a result of AIRSP forecasting a thunderstorm in this location whereas only a few members of NO-AIRSP generate a similar feature. The following section compares this forecast with the observations and will show that AIRSP produces the better forecast of this event.

6. Forecast evaluation of simulated radar reflectivity

a. 10–11 May 2010

Simulated reflectivity forecasts generated by AIRSP and NO-AIRSP ensembles for 10 May 2010 are compared with independent observations from the KTLX WSR-88D radar interpolated to a 1-km horizontal, 0.5-km vertical resolution grid. Reflectivity at 2 km above ground level (AGL) on 2230 UTC 10 May shows the line of supercells extending from southern Oklahoma north-northeastward into northern Oklahoma (Fig. 12a). The strongest cells are located in central Oklahoma, near KTLX with a second area of intensifying convection located in southern Oklahoma (Fig. 12a). Both areas of convection are producing severe weather by this time. The same features are present 30 min later at 2300 UTC with both areas of convection moving eastward at 20 to 25 m s^{-1} (Fig. 12c). The original convection north of the Twin Lakes, Oklahoma, radar (KTLX) weakens and is replaced by a secondary line of convection forming along the dryline that remains located in central Oklahoma at this time. The location of the dryline can be seen in the clear-air reflectivity returns west of the radar and corresponds well to the 20°C dewpoint contour from both model forecasts. By 2330 UTC, the original convection has evolved into three distinct supercells; one in central Oklahoma centered near 35.4°N and −96.4°W with two more located farther south near 34.25° ± 0.25°N and −97.25° ± 0.25°W (Fig. 13a). Advancing to 0000 UTC, the same features are present, but with a significant increase in the intensity of the westernmost line of convection in central Oklahoma (Fig. 13c). The eastward progression of the dryline is again evident from both radar observations and model forecasts. However, both models appear too slow in progressing the dryline compared to observations.

For all forecast times, the probability of reflectivity greater than 30 dBZ at 2 km AGL is calculated with greater probability values, indicating locations where more members generate convection at a particular location. At 2230 UTC (1.5-h forecast), the location of the 30% probability contours generated by NO-AIRSP in central and northern Oklahoma is quite near the WSR-88D observations at this time (Fig. 12a). However, NO-AIRSP fails to properly forecast the convection in the southern portion of the state. In the corresponding AIRSP forecast, significant convection is generated in this region, representing a forecast improvement (Fig. 12b). The convection generated by AIRSP is located slightly to the west of the observations, but is in the correct southwest–northeast orientation. The only negative impact occurs over a small area just north of KTLX where AIRSP fails to capture the southern portion of this supercell compared to the NO-AIRSP forecast (Figs. 12a,b). At 2300 UTC (2-h forecast), the locations of the 30% probability contours from the NO-AIRSP ensemble are displaced to the west of the observations for the southern convection in the 3-km domain, but show a similar spatial pattern (Fig. 12c). The overall coverage is somewhat lower compared to the observations. In contrast, simulated reflectivity forecasts from AIRSP yield much higher probabilities associated with the southern and central Oklahoma convection (Fig. 12d).

Moving to 2330 and 0000 UTC (2.5- and 3-h forecasts), we find that probabilities for simulated reflectivity greater than 30 dBZ are greater and spread over a larger area in the AIRSP forecast compared to the NO-AIRSP forecast (Figs. 13a,b). The southern Oklahoma cells are well forecast by both models with higher probabilities produced by AIRSP. Both ensembles overforecast convection between 34.0° and 34.5°N at 2330 UTC and, to a lesser extent, at 0000 UTC. Both also generate a feature that appears to correspond to the supercell located at approximately 35.5°N and −96.0°W, but the forecast is approximately 0.5°N too far north (Figs. 13c,d). The location of the NO-AIRSP 30% probability contour is slightly closer to the actual radar data, but the AIRSP forecast has a larger area of higher probabilities (>50%) associated with this feature. Neither model accurately captures the location and intensity of the central Oklahoma convection associated with the secondary line of storms. The ensembles do not generate convection

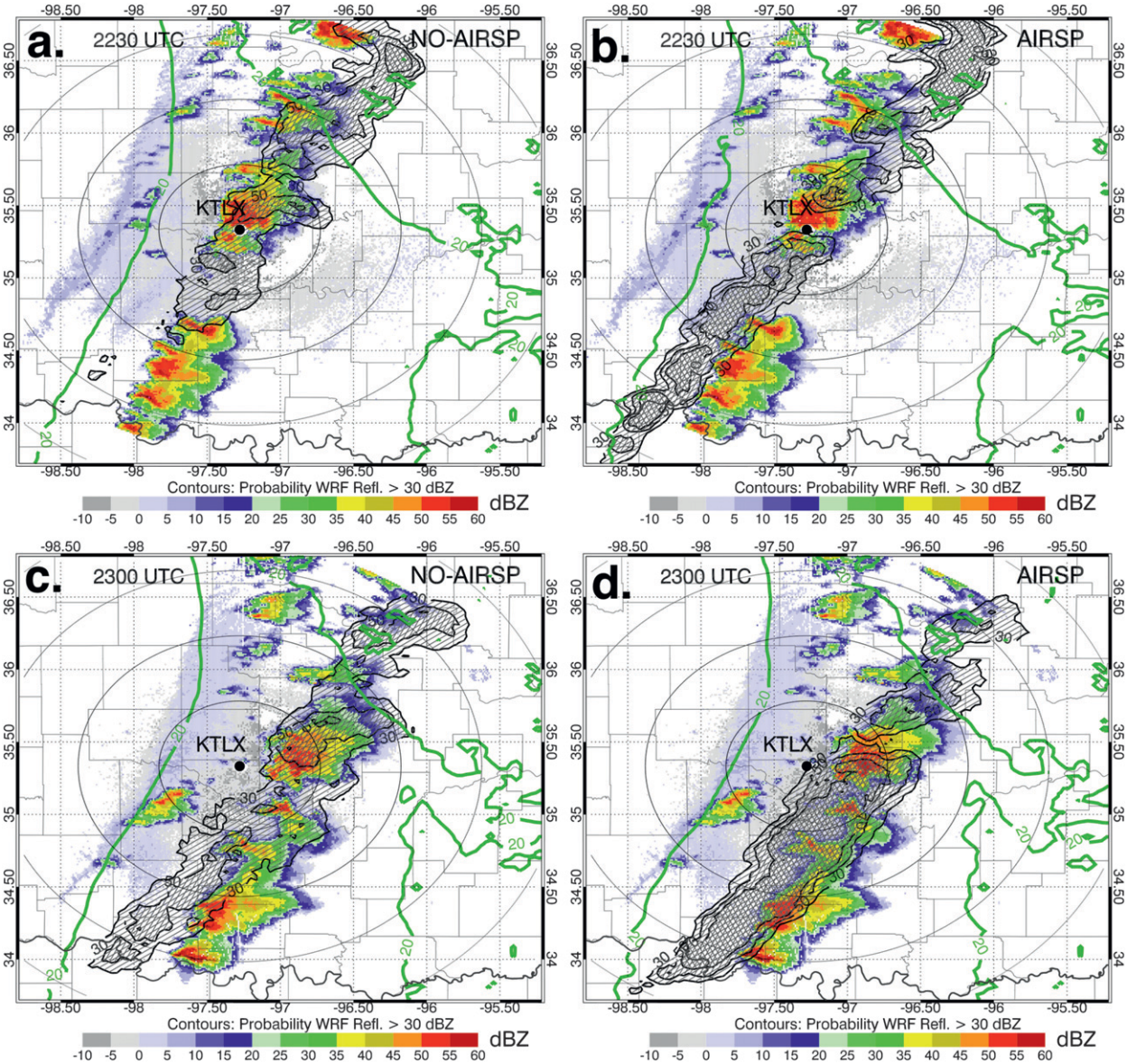

FIG. 12. KTLX WSR-88D radar reflectivity (dBZ) at approximately 2.0 km (AGL) at (a),(b) 2230 and (c),(d) 2300 UTC showing a line of supercells in central OK. Contours of simulated radar reflectivity probability greater than 30 dBZ for the (a),(c) NO-AIRSP ensemble and (b),(d) AIRSP ensemble are overlaid. The 50-km range rings from KTLX are shown for reference. The 20°C isodrosotherm generated by each model at the appropriate forecast time is overlaid on each panel to indicate the approximate location of the dryline in each forecast. Note that the domain plotted for this and following Figs. 13 and 15 is smaller than the 3-km nested grid in order to emphasize the forecast differences where active convection is occurring.

along the secondary line until approximately 0030 UTC 11 May, by which time most of the other convection has moved out of the domain (not shown).

Given the fast-moving nature of this event and the multiple convective features, providing an accurate assessment of forecast skill is difficult. The verification method selected uses a nearest-neighbor approach to relate forecast radar reflectivity and observed WSR-88D reflectivity at 2 km AGL and create a 2×2 contingency table (Wilks 2006). For each ensemble member, a forecast event occurs when the model reflectivity at 2 km exceeds 30 dBZ and a forecast nonevent occurs when the reflectivity at 2 km is less than 30 dBZ. For the observed event, WSR-88D reflectivities within a radius of 10 km from the center of each 3-km-resolution model pixel are examined to determine if any observed values exceed the 30-dBZ threshold. An observed event occurs when at least five WSR-88D radar observations within

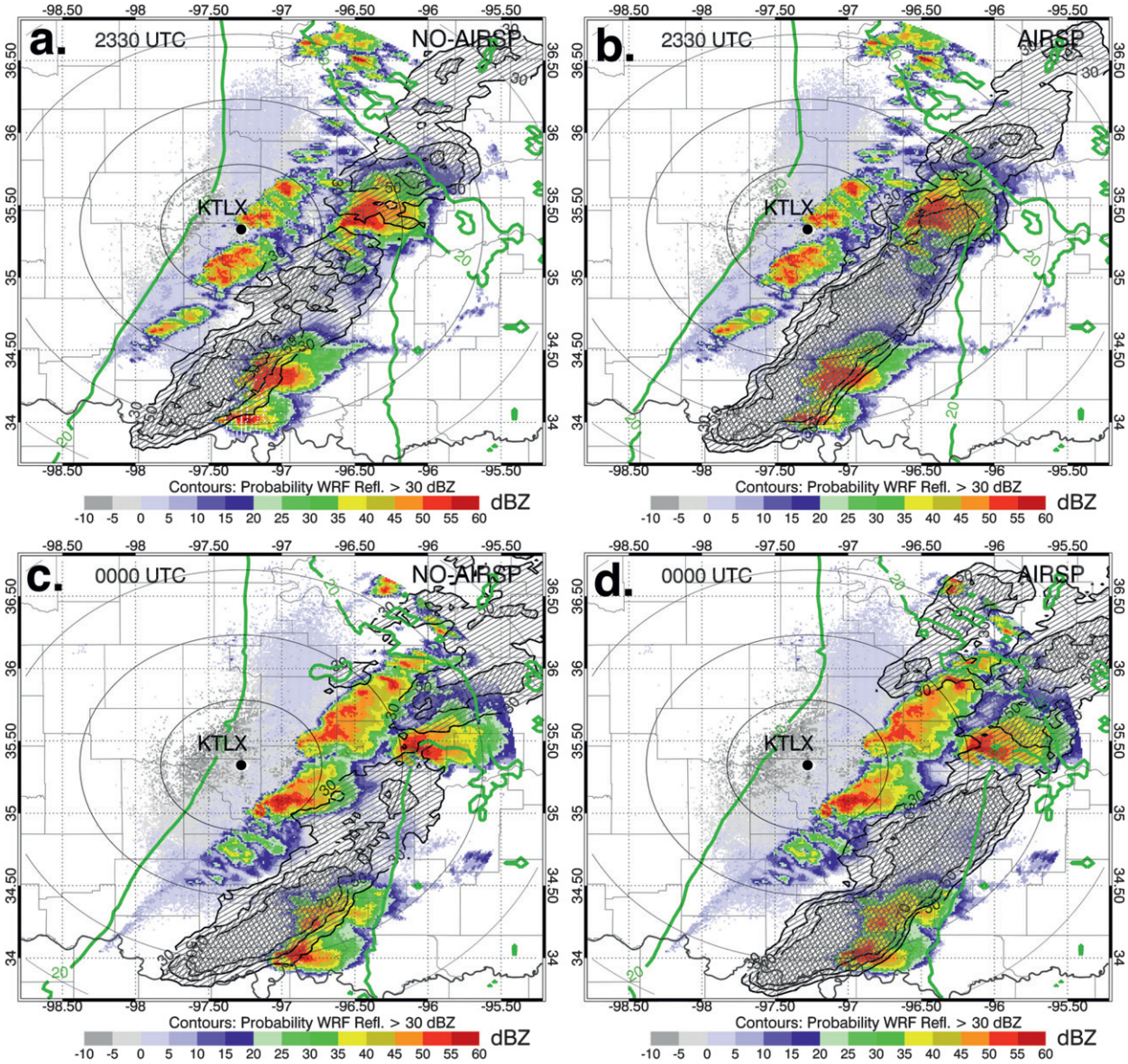

FIG. 13. As in Fig. 12, but for (a),(b) 2330 and (c),(d) 0000 UTC.

this radius exceed the 30-dBZ threshold. Otherwise, it is classified as a nonevent. This process is repeated for all grid boxes within 210 km from the center of the radar (KTLX or KFWS) at each forecast interval to build the contingency table. The probability of detection (POD), false alarm rate (FAR), and Heidke skill score (HSS) for reflectivity greater than 30 dBZ are then computed for each ensemble member and averaged over all members to determine the model skill for each forecast interval (Wilks 2006). The displacement error already noted in the simulated reflectivity forecasts cannot be completely taken into account using this method, but it still remains a useful tool for estimating the differences in model skill since both ensemble forecasts have very similar temporal evolutions.

Prior to 2230 UTC, the skill scores for both NO-AIRSP and AIRSP are similar. Afterward, both HSS and POD increase approximately 0.15 by 2300 UTC in the AIRSP forecast (Figs. 14a,b). The corresponding change in skill with the NO-AIRSP is much smaller, resulting in a large difference in model skill between NO-AIRSP and AIRSP with AIRSP being more skillful out to at least 0000 UTC. FAR increases for both forecasts until 2330 UTC, but the NO-AIRSP value is consistently higher out to 2330 UTC and becoming similar thereafter (Fig. 14c). After 0015 UTC, HSS and POD

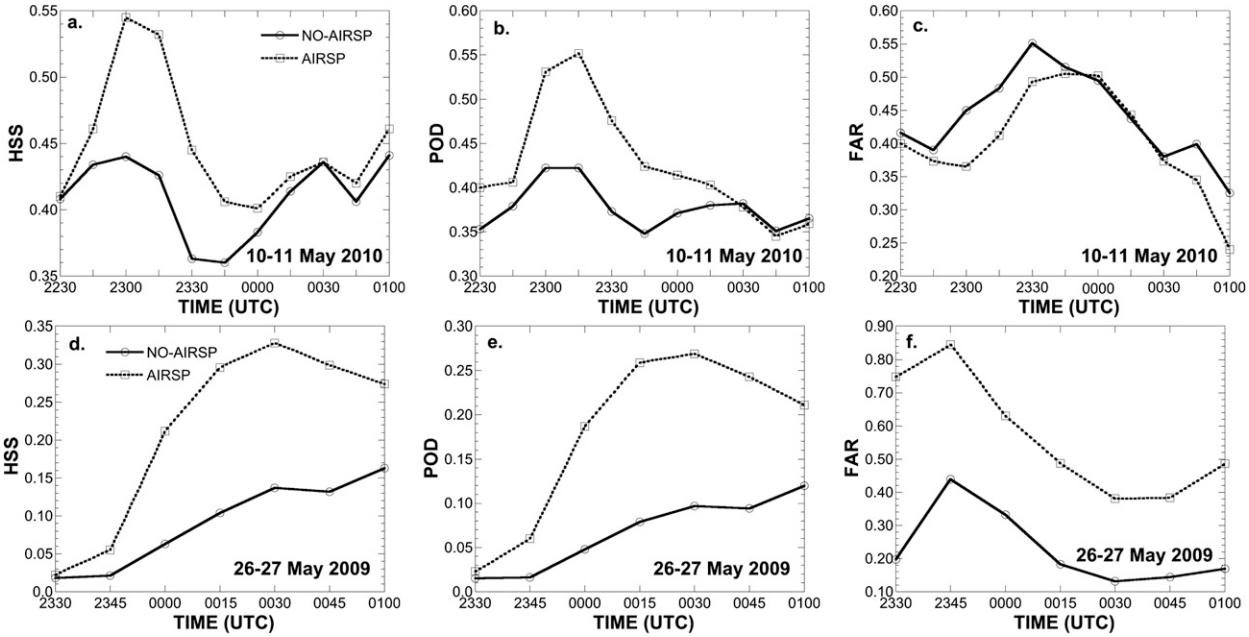

FIG. 14. (a) HSS, (b) POD, and (c) FAR between 2230 UTC 10 May and 0100 UTC 11 May of reflectivity greater than 30 dB*Z* compared to the WSR-88D reflectivity averaged over all 36 ensemble members for NO-AIRSP and AIRSP. Higher values of HSS and POD and lower values of FAR correspond to greater skill. (d)–(f) As in (a)–(c), but for the 26–27 May 2009 case between 2330 and 0100 UTC.

become similar for both NO-AIRSP and AIRSP though FAR in AIRSP decreases somewhat faster than for NO-AIRSP. The time lag discussed above limits the overall skill of both models at the later forecast times, but the differences observed between 2230 and 0000 UTC strongly indicate that assimilating AIRS profiles provides a positive impact on the convective-scale forecast.

b. 26–27 May 2009

The same verification approach is applied to the 26–27 May 2009 case with the primary focus between 2330 and 0100 UTC. As seen in the 10–11 May 2010 case, the effects of assimilating AIRS profiles on midtropospheric temperature and humidity yield significant differences in convective-scale forecasts. The differences are again apparent in a comparison of 2-km (AGL) radar reflectivity from KFWS at 0015 UTC with the simulated reflectivity forecasts from NO-AIRSP and AIRSP (Fig. 15). At this forecast time, only 13 out of 36 ensemble members for NO-AIRSP generate reflectivity greater than 30 dB*Z* within the domain shown in Fig. 15. The result is that the 30% probability contour is not generated at any one location. AIRSP performs much better with 32 members forecasting reflectivity greater than 30 dB*Z*. At 0015 UTC, 30% and 50% probability contours are present nearby the observed left split supercell around 33.1°N and −97.5°W (Fig. 15b). However, neither ensemble is able to accurately forecast the right split located just northwest of KFWS at this time.

The differences between the NO-AIRSP and AIRSP forecasts are also evident in the reflectivity verification (Figs. 14d–f). By 0000 UTC, the HSS and POD from AIRSP rapidly increase compared to NO-AIRSP, resulting in a large difference in skill that peaks at 0030 UTC (Figs. 14d,e). After this time, AIRSP skill gradually decreases, although it continues to perform better than NO-AIRSP out to 0100 UTC. The only negative associated with the AIRSP forecast is a much larger FAR compared to NO-AIRS (Fig. 14f). At 2330 and 2345 UTC, FAR exceeds 0.7 before dropping quickly after 0000 UTC. The high FAR early on is a result of the forecast convection being approximately one county away from its observed location, which is beyond the search radius (10 km) used for verification (not shown). The FAR for NO-AIRSP is much lower at all times, but only because most members fail to generate strong convection.

7. Conclusions

The ensemble WRF-DART system successfully assimilates AIRS temperature and mixing ratio profiles and yields a net positive impact on the analyses and short-term regional ensemble forecasts of convection in Oklahoma for 10–11 May 2010 and Texas on 26–27 May

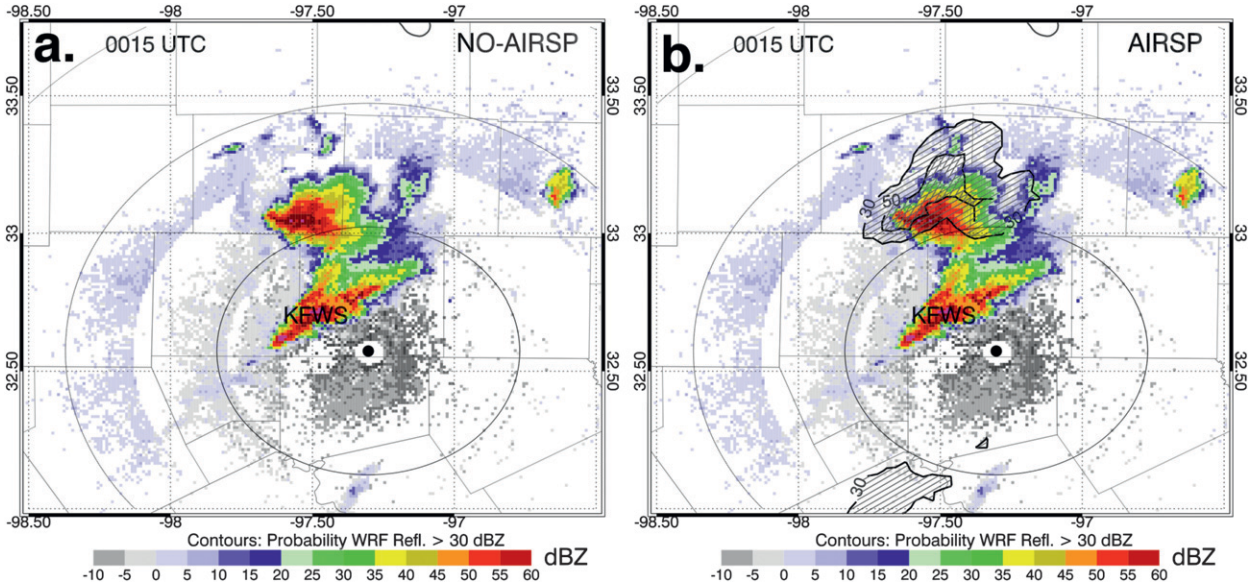

FIG. 15. WSR-88D radar reflectivity at 0015 UTC 27 May 2009 from KFWS with probability contours of simulated reflectivity greater than 30 dBZ from (a) NO-AIRSP and (b) AIRSP overlaid. Probability is less than 30% over the plotted domain of NO-AIRSP at this time.

2009. The greatest impacts on the model forecast fields on the CONUS domain occur near and downstream of the locations of assimilated AIRS profiles in the midtroposphere (700–300 hPa). Differences in AIRSP and NO-AIRSP ensemble mean temperature forecasts are generally less than ±2°C, with several factors potentially limiting the impact of assimilating AIRS temperature profiles. There is often good agreement between the AIRS temperature retrievals with observations and the background fields. In addition, the distribution and number of ACARS temperature observations outweighs the AIRS retrievals in the assimilation when both are present. Thus, differences in the temperature fields between AIRSP and NO-AIRSP are limited except in very localized regions where AIRS data exist with few other observations present (e.g., western Texas).

ACARS moisture observations are much sparser compared to their temperature counterparts, thereby permitting more potential impact from assimilating AIRS mixing ratio profiles. However, the influence of the AIRS dewpoint profiles in WRF-DART is limited by their large observation errors. Nonetheless, despite this limitation, observation diagnostics strongly indicate that assimilating AIRS mixing ratio profiles reduces the model bias and RMSE compared to analyses that do not assimilate these data. The large differences in AIRSP minus NO-AIRSP dewpoint fields at the final analysis time of 2100 UTC exceed 5°C and represent a strong indication that the AIRS profiles are providing unique information to the assimilation process.

Overall, the average improvement over the CONUS due to AIRS data assimilation is negligible when verified against radiosonde observations for 3-h ensemble mean forecasts. For midtropospheric (700–300 hPa) dewpoint, the AIRSP ensemble spread decreases compared to NO-AIRSP due to the inclusion of moisture data where none existed previously. Without the AIRS retrievals, the model relies primarily on the initial and boundary conditions with little in the way of correction for midtropospheric moisture. The reduction of dewpoint spread and forecast uncertainty is seen as one of the important results of this study. In addition to directly affecting temperature and dewpoint fields, the assimilation of AIRS profiles also indirectly impacts wind speed and direction. The overall differences are generally small (<5 m s^{-1}) with no apparent bias or RMSE improvement.

Over the high-resolution southern Great Plains domain for 10 May 2010, comparisons with special upper-air observations made available from the VORTEX2 project reveal that assimilation of AIRS data improves the characterization of the moisture fields for 0–3-h forecasts. The changes to the thermodynamic conditions engendered by assimilating AIRS profiles lead to important differences in the forecast convection. AIRSP generally produces 0–3-h forecasts that contain more favorable environments for strong convection in the warm sector. Examining simulated reflectivity forecasts shows that AIRSP generates more intense and widespread convection, which is more consistent with actual WSR-88D observations and storm reports. The greatest improvement

in skill on 10–11 May occurs for the 2- and 3-h forecast times where AIRSP captures the coverage and intensity of the convection in central and southern Oklahoma much better than does NO-AIRSP. The comparison for 26 May 2009 produces similar results, with AIRSP forecasting a thunderstorm in the approximate location as observed, whereas most members of NO-AIRSP fail to forecast any convection. The greatest forecast improvement occurs for the 3–4-h forecast times for this case. It is important to stress that this work represents only two case studies and some of the improvements observed here may not be reproduced in other cases.

The overall results of this research are positive and highlight the potential for satellite data assimilation to improve high-resolution storm-scale forecasts, which is a key goal of the Warn-on-Forecast project (Stensrud et al. 2009). However, routine assimilation of AIRS profiles in the fashion described here will be difficult due to the low temporal resolution of these data as well as questions concerning their real-time availability. The proposed solution to this problem is to include an AIRS-like instrument on board the next generation of GOES satellites. While this instrument was deleted from the suite of instruments to be placed on GOES-R, it is hoped that research such as this that show its positive impact on convective-scale forecasts may result in this sensor's reintroduction on future satellites.

Acknowledgments. This research would not have been possible without the support of the entire FRDD team. In particular, we would like to thank Nusrat Yussouf for processing the KTLX WSR-88D radar data used in Fig. 9. AIRS L2 data were acquired from the NASA data server (http://mirador.gsfc.nasa.gov/). Mobile radiosonde data collected by the VORTEX2 experiment were made available by Mike Coniglio. We also appreciate the detailed comments provided by the anonymous reviewers that provided a basis for several significant improvements in the presentation of this study. This research was supported by the NOAA/National Environmental Satellite, Data, and Information Service. Partial funding for this research was also provided by NOAA/Office of Oceanic and Atmospheric Research under NOAA–University of Oklahoma Cooperative Agreement NA17RJ1227, under the U.S. Department of Commerce.

REFERENCES

Anderson, J. L., and N. Collins, 2007: Scalable implementations of ensemble filter algorithms for data assimilation. *J. Atmos. Oceanic Technol.,* **24,** 1452–1463.

——, 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.

Aumann, H. H., and Coauthors, 2003: AIRS/AMSU/HSB on the Aqua mission: Design, science objectives, data products and processing systems. *IEEE Trans. Geosci. Remote Sens.,* **41,** 253–264.

Barker, D. M., W. Huang, Y.-R. Guo, and Q. N. Xiao, 2004: A three-dimensional data assimilation system for use with MM5: Implementation and initial results. *Mon. Wea. Rev.,* **132,** 897–914.

Chen, Y., F. Weng, Y. Han, and Q. Liu, 2008: Validation of the Community Radiative Transfer Model by using CloudSat data. *J. Geophys. Res.,* **113,** D00A03, doi:10.1029/2007JD009561.

Chou, S.-H., B. Zavodsky, G. Jedlovec, and W. Lapenta, 2007: Assimilation of Atmospheric Infrared Sounder (AIRS) profiles using WRF-VAR. *Ninth Annual WRF Users' Workshop,* Boulder, CO, NCAR, P5.14.

——, ——, and ——, 2009: Data assimilation and regional forecasts using Atmospheric Infrared Sounder (AIRS) profiles. Preprints, *16th Conf. on Satellite Meteorology and Oceanography/ Fifth Symp. on Future Operational Environmental Satellite Systems—NPOESS and GOES-R,* Phoenix, AZ, Amer. Meteor. Soc., JP6.11. [Available online at http://ams.confex.com/ams/pdfpapers/147745.pdf.]

——, ——, and ——, 2010: Regional precipitation forecast with Atmospheric Infrared Sounder (AIRS) profile assimilation. *11th Annual WRF Users' Workshop,* Boulder, CO, NCAR, 8 pp. [Available online at http://www.mmm.ucar.edu/wrf/users/workshops/WS2010/abstracts/P-88.pdf.]

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.

Divakarla, M. G., C. D. Barnet, M. D. Goldberg, L. M. McMillin, E. Maddy, W. Wolf, L. Zhou, and X. Liu, 2006: Validation of Atmospheric Infrared Sounder temperature and water vapor retrievals with matched radiosonde measurements and forecasts. *J. Geophys. Res.,* **111,** D09S15, doi:10.1029/2005JD006116.

Dudhia, J., 1989: Numerical study of convection observed during the Winter Monsoon Experiment using a mesoscale two-dimensional model. *J. Atmos. Sci.,* **46,** 3077–3107.

Ek, M., K. E. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayno, and J. D. Tarpley, 2003: Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta Model. *J. Geophys. Res.,* **108,** 8851, doi:10.1029/2002JD003296.

Errico, R. M., 2000: The dynamical balance of leading singular vectors in a primitive-equation model. *Quart. J. Roy. Meteor. Soc.,* **126,** 1601–1618.

——, G. Ohring, F. Weng, P. Bauer, B. Ferrier, J.-F. Mahfouf, and J. Turk, 2007: Assimilation of satellite cloud and precipitation observations in numerical weather prediction models. *J. Atmos. Sci.,* **64,** 3737–3741.

Evensen, G., 2003: The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dyn.,* **53,** 343–367.

Ferrier, B. S., J. Simpson, and W.-K. Tao, 1996: Factors responsible for precipitation efficiencies in midlatitude and tropical squall simulations. *Mon. Wea. Rev.,* **124,** 2100–2125.

Fetzer, E. J., and Coauthors, 2003: AIRS/AMSU/HSB validation. *IEEE Trans. Geosci. Remote Sens.,* **41,** 418–431.

Fujita, T., D. J. Stensrud, and D. C. Dowell, 2007: Surface data assimilation using an ensemble Kalman filter approach with

initial condition and model physics uncertainty. *Mon. Wea. Rev.,* **135,** 1846–1868.

Gilmore, M. S., and L. J. Wicker, 1998: The influence of midtropospheric dryness on supercell morphology and evolution. *Mon. Wea. Rev.,* **126,** 943–958.

Goldberg, M. D., Y. Qu, L. M. McMillin, W. Wolf, L. Zhou, and M. Divakarla, 2003: AIRS near-real-time products and algorithms in support of operational numerical weather prediction. *IEEE Trans. Geosci. Remote Sens.,* **41,** 379–389.

——, C. D. Barnet, W. Wolf, L. Zhou, and M. Divakarla, 2004: Distributed real-time operational products from AIRS. *Atmospheric and Environmental Remote Sensing Data Processing and Utilization: An End-to-End System Perspective,* H.-L. A. Huang and H. J. Bloom, Eds., International Society for Optical Engineering (SPIE Proceedings, Vol. 5548), doi:10.1117/12.559820.

Grell, G. A., and D. Devenyi, 2002: A generalized approach to parameterizing convection combining ensemble and data assimilation techniques. *Geophys. Res. Lett.,* **29,** 1693, doi:10.1029/2002GL015311.

Hou, A. Y., S. Q. Zhang, and O. Reale, 2004: Variational continuous assimilation of TMI and SSM/I rain rates: Impact on GEOS-3 hurricane analyses and forecasts. *Mon. Wea. Rev.,* **132,** 2094–2109.

Janjić, Z. I., 2002: Nonsingular implementation of the Mellor–Yamada level 2.5 scheme in the NCEP Meso Model. NCEP Office Note 437, 61 pp.

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.

——, and Coauthors, 2007: The Joint Center for Satellite Data Assimilation. *Bull. Amer. Meteor. Soc.,* **88,** 329–340.

Li, J., and H. Liu, 2009: Improved hurricane track and intensity forecast using single field-of-view advanced IR sounding measurements. *Geophys. Res. Lett.,* **36,** L11813, doi:10.1029/2009GL038285.

Liu, H., and J. Li, 2010: An improvement in forecasting rapid intensification of Typhoon Sinlaku, 2008: Using clear-sky full spatial resolution advanced IR soundings. *J. Appl. Meteor. Climatol.,* **49,** 821–827.

Liu, J., H. Li, E. Kalnay, E. J. Kostelich, and I. Szunyogh, 2009: Univariate and multivariate assimilation of AIRS humidity retrievals with the local ensemble transform Kalman filter. *Mon. Wea. Rev.,* **137,** 3918–3932.

Mansell, E., C. Ziegler, and D. MacGorman, 2007: A lightning data assimilation technique for mesoscale forecast models. *Mon. Wea. Rev.,* **135,** 1732–1748.

McMillan, W. W., and Coauthors, 2005: Daily global maps of carbon monoxide from NASA's Atmospheric Infrared Sounder. *Geophys. Res. Lett.,* **32,** L11801, doi:10.1029/2004GL021821.

McNally, A. P., J. C. Derber, W.-S. 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.

Meng, Z., and F. Zhang, 2011: Limited-area ensemble-based data assimilation: A review. *Mon. Wea. Rev.,* **139,** 2025–2045.

Migliorini, S., 2012: On the equivalence between radiance and retrieval assimilation. *Mon. Wea. Rev.,* **140,** 258–265.

Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997: RRTM, a validated correlated-k model for the longwave. *J. Geophys. Res.,* **102** (D14), 16 663–16 682.

Olsen, E. T., and Coauthors, cited 2007: AIRS/AMSU/HSB version 5 data release user guide. [Available online at http://disc.sci.gsfc.nasa.gov/AIRS/documentation/v5_docs/.]

Pagano, T. S., H. H. Aumann, D. E. Hagan, and K. Overoye, 2003: Pre-launch and in-flight radiometric calibration of the Atmospheric Infrared Sounder (AIRS). *IEEE Trans. Geosci. Remote Sens.,* **41,** 265–273.

Pavelin, E., S. English, and J. Eyre, 2008: The assimilation of cloud-affected infrared satellite radiances for numerical weather prediction. *Quart. J. Roy. Meteor. Soc.,* **134,** 737–749.

Pu, Z., and L. Zhang, 2011: Validation of AIRS temperature and moisture profiles over tropical oceans and their impact on numerical simulations of tropical cyclones. *J. Geophys. Res.,* **115,** D24114, doi:10.1029/2010JD014258.

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:10.1029/2007GL033002.

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.]

Stensrud, D. J., and Coauthors, 2009: Convective-scale warn-on-forecast system: A vision for 2020. *Bull. Amer. Meteor. Soc.,* **90,** 1487–1499.

Susskind, J., C. D. Barnet, and J. M. Blaisdell, 2003: Retrieval of atmospheric and surface parameters from AIRS/AMSU/HSB data in the presence of clouds. *IEEE Trans. Geosci. Remote Sens.,* **41,** 390–409.

——, ——, ——, L. Iredell, F. Keita, L. Kouvaris, G. Molnar, and M. Chahine, 2006: Accuracy of geophysical parameters derived from Atmospheric Infrared Sounder/Advanced Microwave Sounding Unit as a function of fractional cloud cover. *J. Geophys. Res.,* **111,** D09S17, doi:10.1029/2005JD006272.

Thompson, G., R. M. Rasmussen, and K. Manning, 2004: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part I: Description and sensitivity analysis. *Mon. Wea. Rev.,* **132,** 519–542.

Tobin, D. C., H. E. Revercomb, C. C. Moeller, and T. S. Pagano, 2006: Use of Atmospheric Infrared Sounder high-spectral resolution spectra to assess the calibration of Moderate Resolution Imaging Spectroradiometer on EOS Aqua. *J. Geophys. Res.,* **111,** D09S05, doi:10.1029/2005JD006095.

Tompkins, A. M., 2001: Organization of tropical convection in low vertical wind shears: The role of water vapor. *J. Atmos. Sci.,* **58,** 529–545.

Torn, R. D., and G. J. Hakim, 2008: Performance characteristics of a pseudo-operational ensemble Kalman filter. *Mon. Wea. Rev.,* **136,** 3947–3963.

——, ——, and C. Snyder, 2006: Boundary conditions for limited-area ensemble Kalman filters. *Mon. Wea. Rev.,* **134,** 2490–2502.

Velden, C. S., T. L. Olander, and S. Wanzong, 1998: The impact of multispectral *GOES-8* wind information on Atlantic tropical cyclone track forecasts in 1995. Part I: Dataset methodology, description, and case analysis. *Mon. Wea. Rev.,* **126,** 1202–1218.

Weisz, E., J. Li, J. Li, D. K. Zhou, H.-L. Huang, M. D. Goldberg, and P. Yang, 2007: Cloudy sounding and cloud-top height retrieval from AIRS alone single field-of-view radiance

measurements. *Geophys. Res. Lett.,* **34,** L12802, doi:10.1029/2007GL030219.

Wheatley, D. M., and D. J. Stensrud, 2010: The impact of assimilating surface pressure observations on severe weather events in a WRF mesoscale ensemble system. *Mon. Wea. Rev.,* **138,** 1673–1694.

Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences*. 2nd ed. Academic Press, 467 pp.

Wu, L., S. A. Braun, J. J. Qu, and X. Hao, 2006: Simulating the formation of Hurricane Isabel (2003) with AIRS data. *Geophys. Res. Lett.,* **33,** L04804, doi:10.1029/2005GL024665.

Wurman, J., L. Wicker, Y. Richardson, E. N. Rasmussen, P. Markowski, D. Dowell, D. Burgess, and H. Bluestein, 2010: An overview of the VORTEX2 field campaign. Preprints, *25th Conf. on Severe Local Storms,* Denver, CO, Amer. Meteor. Soc., 5.1. [Available online at http://ams.confex.com/ams/25SLS/techprogram/paper_176068.htm.]

Xubebao, W., L. Jun, Z. Wenjian, and W. Fang, 2005: Atmospheric profile retrieval with AIRS data and validation at the ARM CART site. *Adv. Atmos. Sci.,* **22,** 647–654, doi:10.1007/BF02918708.

Yussouf, N., and D. J. Stensrud, 2010: Impact of high temporal frequency phased array radar data to storm-scale ensemble data assimilation using observation system simulation experiments. *Mon. Wea. Rev.,* **138,** 517–538.

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.