Simultaneous Assimilation of Radar and All-Sky Satellite Infrared Radiance Observations for Convection-Allowing Ensemble Analysis and Prediction of Severe Thunderstorms

YUNJI ZHANG, DAVID J. STENSRUD, AND FUQING ZHANG

*Center for Advanced Data Assimilation and Predictability Techniques, and Department of Meteorology
and Atmospheric Science, The Pennsylvania State University, University Park, Pennsylvania*

ABSTRACT

This study explores the benefits of assimilating infrared (IR) brightness temperature (BT) observations from geostationary satellites jointly with radial velocity (Vr) and reflectivity (Z) observations from Doppler weather radars within an ensemble Kalman filter (EnKF) data assimilation system to the convection-allowing ensemble analysis and prediction of a tornadic supercell thunderstorm event on 12 June 2017 across Wyoming and Nebraska. While radar observations sample the three-dimensional storm structures with high fidelity, BT observations provide information about clouds prior to the formation of precipitation particles when in-storm radar observations are not yet available and also provide information on the environment outside the thunderstorms. To better understand the strengths and limitations of each observation type, the satellite and Doppler radar observations are assimilated separately and jointly, and the ensemble analyses and forecasts are compared with available observations. Results show that assimilating BT observations has the potential to increase the forecast and warning lead times of severe weather events compared with radar observations and may also potentially complement the sparse surface observations in some regions as revealed by the prob-abilistic prediction of mesocyclone tracks initialized from EnKF analyses as various times. Additionally, the assimilation of both BT and Vr observations yields the best ensemble forecasts, providing higher confidence, improved accuracy, and longer lead times on the probabilistic prediction of midlevel mesocyclones.

1. Introduction

With the recent advances in observation platforms, numerical weather prediction models, data assimilation techniques, and computational resources, the future operational warning paradigm for severe convective weather phenomena, such as tornadoes, hail, and damaging winds, has been evolving from warn-on-detection (WoD) based upon the detection or observation of certain phenomena to warn-on-forecast (WoF) based upon convection-allowing model (CAM) ensemble forecasts (Stensrud et al. 2009, 2013). Recent studies have demonstrated the capability of CAM ensemble forecasts to accurately and reliably predict rotational signatures associated with the severe weather (e.g., Wheatley et al. 2015; Yussouf et al. 2015; Sobash et al. 2016; Jones et al. 2016, 2018; Lawson et al. 2018; Skinner et al. 2018), providing confidence in the feasibility of such WoF systems for assisting severe weather preparedness in the future (Hoekstra et al. 2011; Gallo et al. 2017).

The quality of initial conditions (ICs) that are used to initialize CAM ensembles is crucial for the accuracy of the predictions, especially for short-term forecasts of severe weather events. For example, Lawson et al. (2018) found that the National Severe Storm Laboratory (NSSL) Experimental WoF System for ensembles (NEWS-e) outperformed the High-Resolution Rapid Refresh (HRRR) system for 0–3-h quantitative precipitation forecasts (QPF) in spite of similar grid spacing, model settings, and parameterization schemes, largely due to using an advanced ensemble Kalman filter (EnKF) data assimilation system that assimilated radar observations to produce more accurate NEWS-e ICs as compared with a simple latent-heating adjustment method used by HRRR to incorporate radar observations. Most current CAM data assimilation systems have horizontal grid spacing of less than 4 km and rely heavily on Doppler weather radar observations that can provide accurate three-dimensional storm structure. However, these radars are generally unable to provide information outside the storms or prior to the formation of precipitation particles. With the operational availability of the new-generation

Corresponding author: Yunji Zhang, yuz31@psu.edu

geostationary satellites *GOES-16* as GOES-East on 18 December 2017 and *GOES-17* as GOES-West on 12 February 2019, and their capability to image CONUS every 5 min with a resolution of 2 km at nadir for infrared channels, these satellite observations will cover the "blind spots" and "gaps" within the national weather radar network. The high-spatiotemporal-resolution infrared (IR) brightness temperature (BT; used interchangeably with "radiance" hereafter) observations provided by the Advanced Baseline Imager (ABI) on board the *GOES-16* and *GOES-17* satellites are expected to be an important supplement to current datasets used by data assimilation systems.

IR BT observations from geostationary satellites have been significantly underutilized for the past several decades with only the clear-sky radiances assimilated in major operational global modeling centers (Geer et al. 2018). However, the assimilation of all-sky IR BT observations using ensemble-based data assimilation techniques like the ensemble Kalman filter (EnKF) has been explored under the observing system simulation experiment (OSSE) framework in recent years (Otkin 2010, 2012; Zupanski et al. 2011; Jones et al. 2013, 2014; Cintineo et al. 2016; Zhang et al. 2016; Minamide and Zhang 2017), with some of these studies further exploring the combined assimilation of IR BT and radar observations (Jones et al. 2013, 2014; Cintineo et al. 2016). There are several studies that assimilate real BT observations from ABI and the Advanced Himawari Imager (AHI; which has similar channels and resolutions as ABI) on board Japan's *Himawari-8* satellite for various weather systems, including tropical cyclones (Minamide and Zhang 2018; Honda et al. 2018a,b), fronts (Okamoto et al. 2019), and isolate single storms (Sawada et al. 2019), and Zhang et al. (2018; hereafter Z18) presented the first successful assimilation of real ABI observations using a CAM EnKF system to improve the prediction of severe thunderstorms. Aside from the benefits resulted from assimilating all-sky compared with clear-sky observations (Okamoto et al. 2019), the flow-dependent, time-varying background error covariances and easier adaptation of complex observation operators without the requirement of adjoint and tangent linear models makes EnKF a more promising approach for the assimilation of all-sky BT observations compared with variational methods at convection-allowing scales.

Studies combining real-world rather than synthetic satellite and radar observations so far only use satellite retrievals (Jones et al. 2015, 2016; Kerr et al. 2015) or clear-sky observations instead of all-sky BT observations (Jones et al. 2018) with radar observations. As a follow up study of Z18, this current study seeks to explore the simultaneous assimilation of *GOES-16* ABI all-sky IR BT observations with radial velocity (Vr) and reflectivity (Z) observations from the Weather Surveillance Radar–1988 Doppler (WSR-88D) radars using EnKF with the Weather Research and Forecasting (WRF) Model running at a convection-allowing 1-km horizontal grid spacing. We explore how their simultaneous assimilation impacts the prediction of a severe thunderstorm event on 12 June 2017 across Wyoming, Nebraska and Colorado that has previously been presented in Z18, especially during the initiation and early development stage of the thunderstorms. Section 2 introduces the observations, data assimilation systems, numerical model, and experiment design. Results are presented in section 3, and section 4 is summary and discussion.

2. Methodology

a. Observations and preprocessing procedures

The observations assimilated in this study include reflectivity (Z) and radial velocity (Vr) observations from selected WSR-88D radars and infrared (IR) brightness temperature (BT) observations from *GOES-16* ABI, and 2-m temperature, dewpoint temperature, and 10-m wind observations from surface weather stations.

Raw level-II data of the Denver, Colorado (KFTG), and the Cheyenne, Wyoming (KCYS), WSR-88D radars are acquired through NOAA's National Centers for Environmental Information (NCEI). The Vr observations are manually dealiased and superobservations (SOs) of Z and Vr with a radial and azimuthal spacing of 2 km are generated from the raw observations following procedures similar to those in Zhang et al. (2009). The procedures include 1) all raw observations (Z and Vr observations are treated as being linked at each range and azimuth, such that either both are used or both are discarded at a given location except for nonprecipitating Z observations) with Vr magnitude smaller than $3 \, \text{m s}^{-1}$ or greater than $60 \, \text{m s}^{-1}$ or within 2 km from the radar site are discarded; 2) after dividing each elevation scan into bins with radial and azimuthal distance of 2 km, all observations within a bin are discarded if the standard deviation of the Vr observations within this bin is greater than 1.5 times the standard deviation of all Vr observations within the same elevation angle; 3) for each bin, observations are discarded if their Vr deviates more than twice the standard deviation of all surviving Vr observations (after steps 1 and 2) within the bin from the mean Vr value of the bin; 4) all raw observations within a bin are discarded if more than half the observations within the bin are discarded after previous procedures; and 5) the median value of all remaining observations within a bin is chosen as the value for Z and Vr for this SO, which is
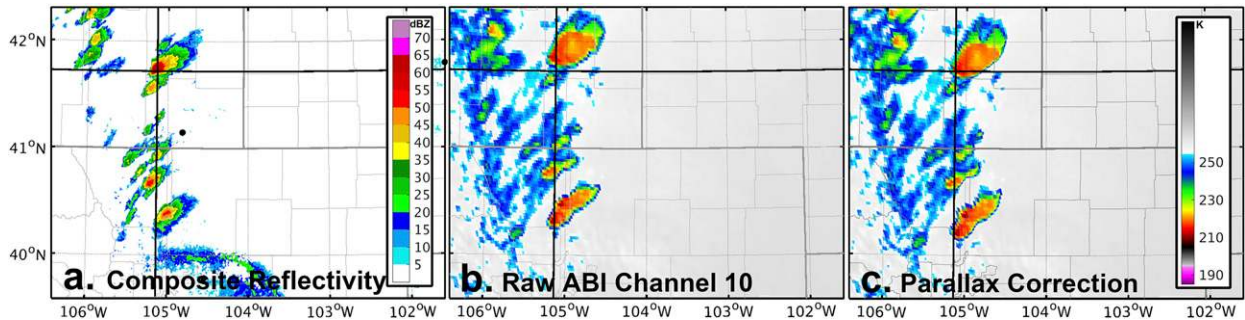
FIG. 1. Observations of (a) composite reflectivity, (b) raw ABI channel 10 brightness temperature, and (c) parallax-corrected brightness temperature at 2015 UTC 12 Jun 2017. Solid longitude and latitude lines are for easier geographical comparisons.

located at the center of the bin; if all Z observations within a bin are less than 5 dBZ, a nonprecipitating SO is assigned with 0 dBZ without a corresponding Vr observations (referred to as "0-dBZ observations"). "Z observations" only refer to nonzero precipitating Z observations unless otherwise stated. 0-dBZ observations are further thinned to a 4-km horizontal grid spacing.

BT observations from the Cloud and Moisture Imagery product (CMIP) of channel 10 (7.3 μm; the lower-tropospheric water vapor channel) of *GOES-16* ABI acquired from NOAA's Comprehensive Large Array Data Stewardship System (CLASS) are used in this study, same as in Z18. Because ABI sees the clouds slantwise, the location of the clouds from the imager are displaced from their actual locations, and this "parallax error" should be corrected to provide more accurate analysis when satellite observations are assimilated simultaneously with radar observations at storm scales. The looking angle (elevation angle of the satellite above the horizon when observed from Earth's surface) and the azimuth angle of the satellite at every grid point of the raw BT observations are determined based on Soler and Eisemann (1994), then the magnitude of parallax error of each grid point is calculated using the looking angle and the cloud top height (ACHA) product of *GOES-16* (also acquired through CLASS) and further decomposed into latitudinal and longitudinal errors using the azimuth angle to compensate the errors. Typical correction for a cloud top height of ~10 km in the targeted region of this study when *GOES-16* is situated at 89.5°W above the equator is about ~0.05° eastward and ~0.1° southward. After this correction, colder cloud top regions in BT (Fig. 1c) have a much better collocation with the storms seen in composite reflectivity (Fig. 1a) than raw observations (Fig. 1b), and this improved collocation is crucial for the simultaneous assimilation of satellite and radar observations. The parallax correction is especially important during initiation of the storms when the magnitude of the parallax error is comparable

to the scale of the clouds. It should also be pointed out that after *GOES-16* became the operational GOES-E located at 75.2°W above equator on 30 November 2017, the parallax errors associated with observations over CONUS are larger compared with when *GOES-16* was located at its test position of 89.5°W above equator (as during this event). Finally, the raw observations with parallax-corrected geographical locations are interpolated to the original geographical grids of ABI observations to maintain an even distribution of the observations with a horizontal resolution of about 2.5 km in the targeted region. Note that horizontal spacing of ACHA is 5 times the spacing of CMIP and is bilinearly interpolated when calculating parallax error.

Surface observations are acquired through the Research Data Archive (RDA) of UCAR's Computational and Information Systems Laboratory (CISL). No observations from rawinsondes or profilers are available during the EnKF cycles (see section 2c). Observation errors of surface observations are the default values assigned by the observation preprocessing program (obsproc) of the WRF data assimilation (WRFDA) system.

b. Data assimilation system

The Pennsylvania State University (PSU) WRF-EnKF cycling data assimilation system (Zhang et al. 2009; Weng and Zhang 2012) with same settings as in Z18 is used and is briefly described here. The numerical model of this system uses the fully compressible, non-hydrostatic Advanced Research WRF (ARW) dynamical core (Skamarock et al. 2008) version 3.8.1. A single model domain of 401 × 301 × 61 grids with a horizontal grid spacing of 1 km, highest model level located at 50 hPa, and 19 vertical levels in the lowest 1 km above ground level (AGL) is designed, covering the regions of northern Colorado, southeastern Wyoming, and southwest Nebraska. Physical parameterization schemes applied in the model include the six-species double-moment Thompson et al. (2008) microphysics scheme, unified

Noah land surface model (Ek et al. 2003), Monin–Obukhov–Janjić Eta scheme (Janjić 1996) for surface layer parameterization, Mellor–Yamada–Janjić TKE scheme (Janjić 1994) for PBL processes, and the Rapid Radiative Transfer Model for General Circulation Models (RRTMG) schemes (Iacono et al. 2008) for longwave and shortwave radiation. Simulated radar reflectivity is calculated using the built-in module of the Thompson et al. (2008) microphysics scheme.

The ensemble square root filter (EnSRF; Houtekamer and Mitchell 2001) variation of EnKF with 40 ensemble members acts as the data assimilation part of the system. Following Z18, the Community Radiative Transfer Model (CRTM; Han et al. 2006), version 2.1.3, is used as the observation operator to convert model variables to simulated infrared BT, with the help of the adaptive observation error inflation (AOEI; Minamide and Zhang 2017) and the adaptive background error inflation (ABEI; Minamide and Zhang 2019) techniques to better utilize BT observations (Z18). Following Z18, no additional quality control, bias correction, data thinning or SO procedures are performed on the raw BT observations. It should be pointed out that although the statistics of innovations throughout the EnKF cycles indicate that there is no significant bias in our system (figure not shown), uncertainties in CRTM and/or microphysics schemes applied in the WRF model may sometime induce considerable biases. This issue deserves continued attention as we further explore the assimilation of BT observations. The same observation height specification as in Z18 is also used, with clear-sky observations assigned at 620 hPa and cloudy-sky observations assigned at 250 hPa, respectively. The 0-dBZ observations are assimilated as $0 \, \mathrm{g \, kg^{-1}}$ mixing ratios of total precipitating hydrometeors (the sum of rain, snow and graupel) with an ad hoc observational variance equal to the background variance to avoid the nonlinearities of assimilating 0-dBZ reflectivity values as the observables. To reduce computational cost, the 0-dBZ observations are only assimilated if the maximum background value across all ensemble members at the observation location exceeds $0.1 \, \mathrm{g \, kg^{-1}}$. SOs of Z are assimilated only if the value of Z exceeds 15 dBZ, while all SOs of Vr are assimilated. Other specifications on the observational error and radius of influence (ROI; the cutoff radius of localization for EnKF) using the fifth-order compact function of Gaspari and Cohn (1999) of different observation types are listed in Table 1. ROI of BT observations are the same as in Z18. No BT observation is rejected during EnKF since AOEI is applied; Vr, Z, and surface observations for which the corresponding innovations exceed five times the observational errors are rejected. Finally, the relaxation to prior perturbation (RTPP)

TABLE 1. Observation error and radius of influence (ROI) settings of EnKF experiments.

Observation type	Error	Horizontal ROI	Vertical ROI
Surface	WRFDA	300 km	30 levels
BT	AOEI	30 km	5 × altitude
Vr	$3 \, \mathrm{m \, s^{-1}}$	20 km	20 levels
Z	5 dBZ	20 km	20 levels
0 dBZ	Equal to σ^b	40 km	20 levels

method (Zhang et al. 2004) is performed after all observations are assimilated in each EnKF analysis to maintain ensemble spread using 80% of prior perturbation and 20% of posterior perturbation.

c. Experiment design

The 40 ensemble perturbations are generated using two groups of 20-member ensemble forecasts valid at 1800 UTC. One group started from 0600 UTC and the other started from 1200 UTC using the corresponding 20-member GEFS ensemble analyses as ICs and same numerical model as in the data assimilation system is used for the ensemble forecasts. Perturbations are generated by removing the 40-member mean from each member at 1800 UTC. These perturbations are then added to the HRRR analysis at 1800 UTC to generate 40 ensemble ICs. This set of ensemble ICs at 1800 UTC is exactly the same as that in Z18. A 1-h ensemble forecast initialized from these 40 ICs is carried out to 1900 UTC, and surface observations are assimilated every 20 min (equal to report interval of METAR stations) with a time window of ±10 min till 2040 UTC (ROI showing in Table 1). The 2-m potential temperature and dewpoint temperature are used as assimilated observables for surface temperature and moisture (Fujita et al. 2007). This EnKF experiment that assimilated available conventional observations will be served as a baseline experiment and will be referred to as "CONV."

Seven data assimilation experiments are designed to compare strategies when simultaneously assimilating satellite and radar observations and isolate the influences from each of the observation types. All EnKF experiments assimilate observations from 1900 to 2040 UTC and all experiments assimilate surface observations every 20 min over this assimilation window. The "CONV" experiment only assimilates surface observations. All other experiments also assimilate either satellite and/or radar observations every 5 min over the assimilation window. Four of the experiments only assimilate one of the two remote sensing platforms in addition to surface observations: "SAT" only assimilates BT observations, "VR" only assimilates Vr observations, "REF" assimilates Z and 0-dBZ observations, and "RADAR" assimilates Z,

TABLE 2. Assimilated observations in each experiment, "Y" indicates that observations are included for each experiment.

Experiment	Surface	BT	Vr	Z and 0 dBZ
CONV	Y			
SAT	Y	Y		
VR	Y		Y	
REF	Y			Y
RADAR	Y		Y	Y
VRSAT	Y	Y	Y	
REFSAT	Y	Y		Y
RADSAT	Y	Y	Y	Y

0-dBZ, and Vr observations (i.e., all available radar observations). Three experiments simultaneously assimilate radar and satellite observations in addition to the surface observations: "VRSAT" assimilates BT and Vr observations, "REFSAT" assimilates BT, Z, and 0-dBZ observations, and "RADSAT" assimilates BT, Z, 0-dBZ and Vr observations. The EnKF cycles of all experiments end at 2040 UTC because ABI changed its scan mode right after 2040 UTC, leading to a 20 min period without any observations between 2040 and 2100 UTC. For each experiment, ensemble forecasts are initialized from 1940, 2000, 2020, and 2040 UTC ensemble analyses and end at 0000 UTC 13 June. These ensemble forecasts are referred to as "EF1940," "EF2000," "EF2020," and "EF2040," respectively.

Table 2 summarizes assimilated observations of each experiment. Figure 2 shows the number of assimilated radar observations of each experiment. Average numbers of Vr and Z are plotted since their differences between different experiments varies only by a magnitude of $O(10)$, and a constant number of 18 166 BT observations were assimilated for all cycles.

3. Results

a. Comparison of analysis mean

Before a quantitative evaluation of the experiments, we first qualitatively compare the analysis with the observations to see how the storms are simulated.

The simulated BT of ABI's channel 14 (11.2 μm, longwave window channel) of the EnKF analysis of the six experiments is compared with *GOES-16* ABI channel-14 observations as an independent verification in Figs. 3–5, showing EnKF analyses at the first EnKF cycle, when CI occurs, and the final EnKF cycle, respectively. At the first EnKF cycle at 1900 UTC, CONV and VR (Figs. 3b,c) contain extended regions of clouds and the assimilated observations are unable to remove these spurious clouds, whereas the BT observations are almost free of deep clouds (Fig. 3a). REF
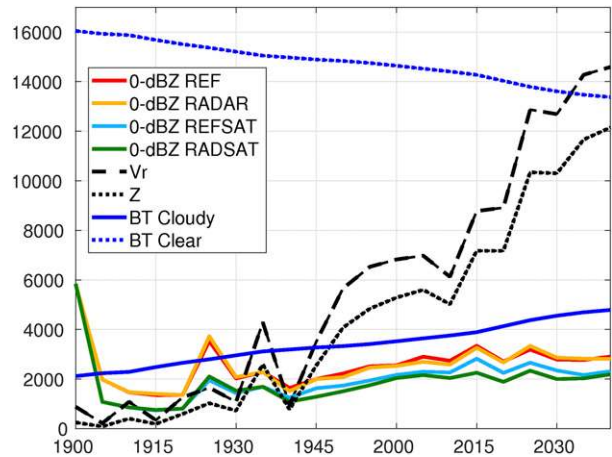

FIG. 2. Quantity of assimilated Vr, Z, 0-dBZ, and cloudy and clear-sky BT observations in different experiments. Average numbers of Vr and Z are plotted.

and RADAR (Figs. 3d,e) show a much improved analysis of clouds compared to CONV and VR due to the assimilation of 0-dBZ observations, although there remains a significant amount of cloudiness. In contrast, satellite observations, as already shown in Z18, can effectively remove almost all the spurious clouds, with the resulting analyses of SAT, VRSAT, REFSAT, and RADSAT (Figs. 3f–i) being much closer to the observations than the other experiments.

With several more cycles, 0-dBZ observations are eventually capable of eliminating spurious clouds. At 2000 UTC the EnKF analyses of REF and RADAR (Figs. 4d,e) are quite similar to the observations (Fig. 4a) as are SAT, VRSAT, REFSAT, and RADSAT that assimilated BT observations (Figs. 4f–i), while significant overestimation of cloud coverage persists in CONV and VR with additional spurious storms initiated in northeastern Colorado (Figs. 4b,c). In REF, RADAR, REFSAT, and RADSAT there are two regions of deeper clouds that are not consistent with the observations. These occur near 42°N, 105°W and 40°N, 104°W (Figs. 4d,e,h,i), with the convection near 42°N, 105°W much deeper in these four experiments than indicated in observations (Fig. 4a). The deep cloud near 40°N, 104°W is slightly weaker in REFSAT and RADSAT (Figs. 4h,i) than REF and RADAR (Figs. 4d,e), probably due to the assimilation of BT observations in these two experiments. The spurious convection near 40°N, 104°W in these four experiments is associated with an observed thin line of weak reflectivity (exceeding 20 dBZ) (Fig. 6a) associated with a dryline. In the final EnKF analyses at 2040 UTC, more than half the model domain is covered with clouds in CONV and VR (Figs. 5b,c), while all other experiments (Figs. 5d–i) generated storms that corresponds
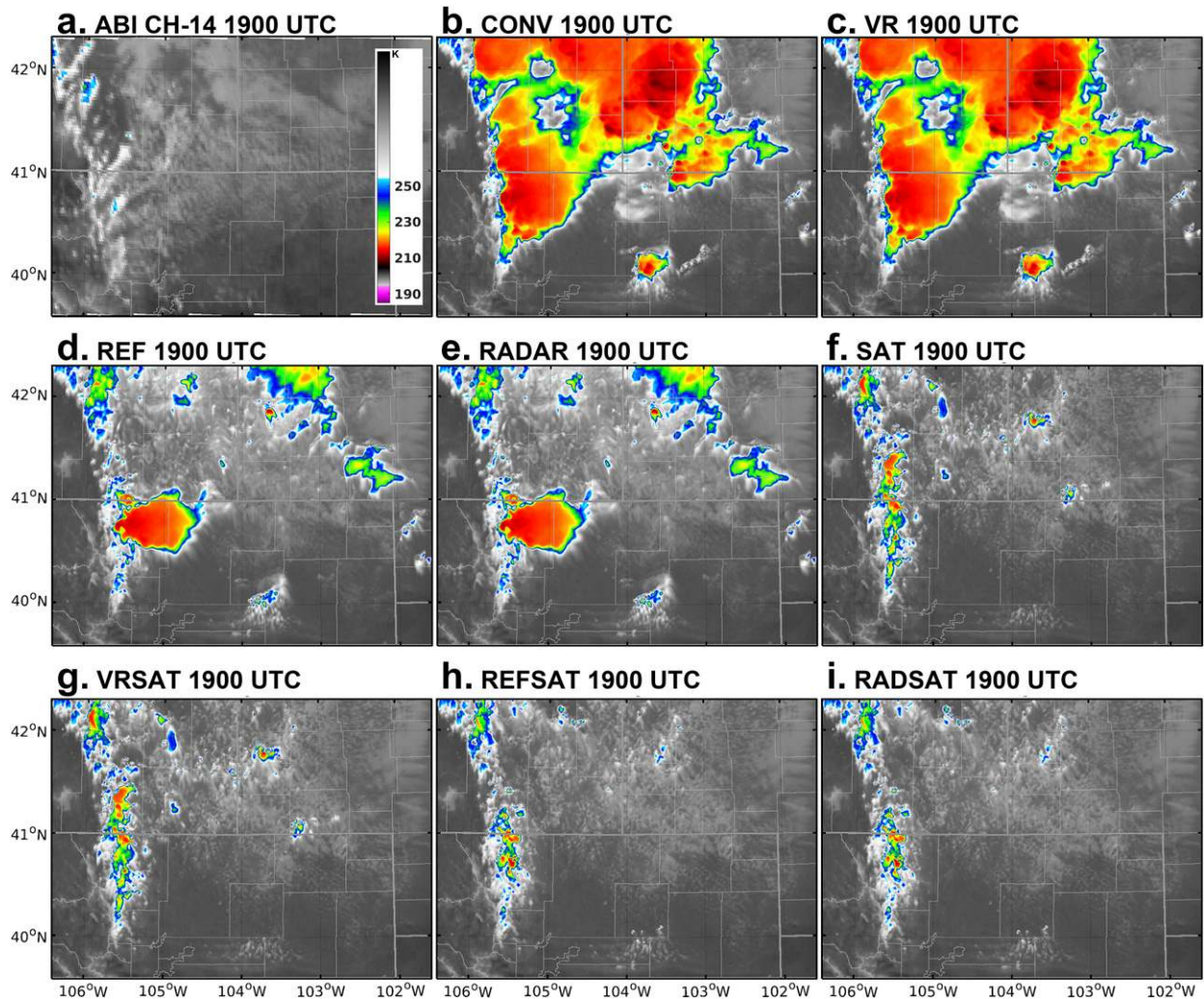
FIG. 3. ABI channel 14 brightness temperature from (a) observations and simulated from EnKF analysis mean of (b) CONV, (c) VR, (d), REF, (e) RADAR, (f) SAT, (g) VRSAT, (h) REFSAT, and (i) RADSAT experiments at 1900 UTC.

well with the observations (Fig. 5a), although the detailed extensions and strength of individual clouds are slightly different in different experiments.

Comparisons of simulated composite reflectivity with observations provide more insight on the differences in the structure of the predicted storms in each experiment (Figs. 6–8). Again, REF and RADAR (Figs. 6d,e) are not able to remove the spurious clouds as effectively as the other experiments that assimilated satellite observations (Figs. 6f–i), primarily due to the fact that even the lowest radar scan at 0.5° elevation angle can be 1–2 km above the surface in regions 100–200 km away from the radar site and there are no radar observations available to remove the excessive amount of hydrometeors near the surface. At 2000 UTC, REF, RADAR, REFSAT, and RADSAT (Figs. 7d,e,h,i) generally produce significantly stronger convection than SAT

and VRSAT (Figs. 7f,g) as well as a banded stratiform precipitation region associated with the observed dryline, consistent with their simulated BT at this time. In the final EnKF analyses at 2040 UTC, the strength and the location of the storms in REF, RADAR, SAT, VRSAT, REFSAT, and RADSAT (Figs. 8d–i) are all similar to the observed storms (Fig. 8a). Slight differences are present among the various experiments; for example, the structure of the primary storm near 41.5°N, 105°W in SAT (Fig. 8f) is less well-defined due to the inability of satellite observations to provide detailed information on storm structure underneath the cloud tops, whereas RADSAT shows a clear hook echo for the same storm (Fig. 8i) which might indicate a stronger mesocyclone compared with the observations which has a less prominent hook echo at this time (Fig. 8a). In general, comparisons of both BT and composite
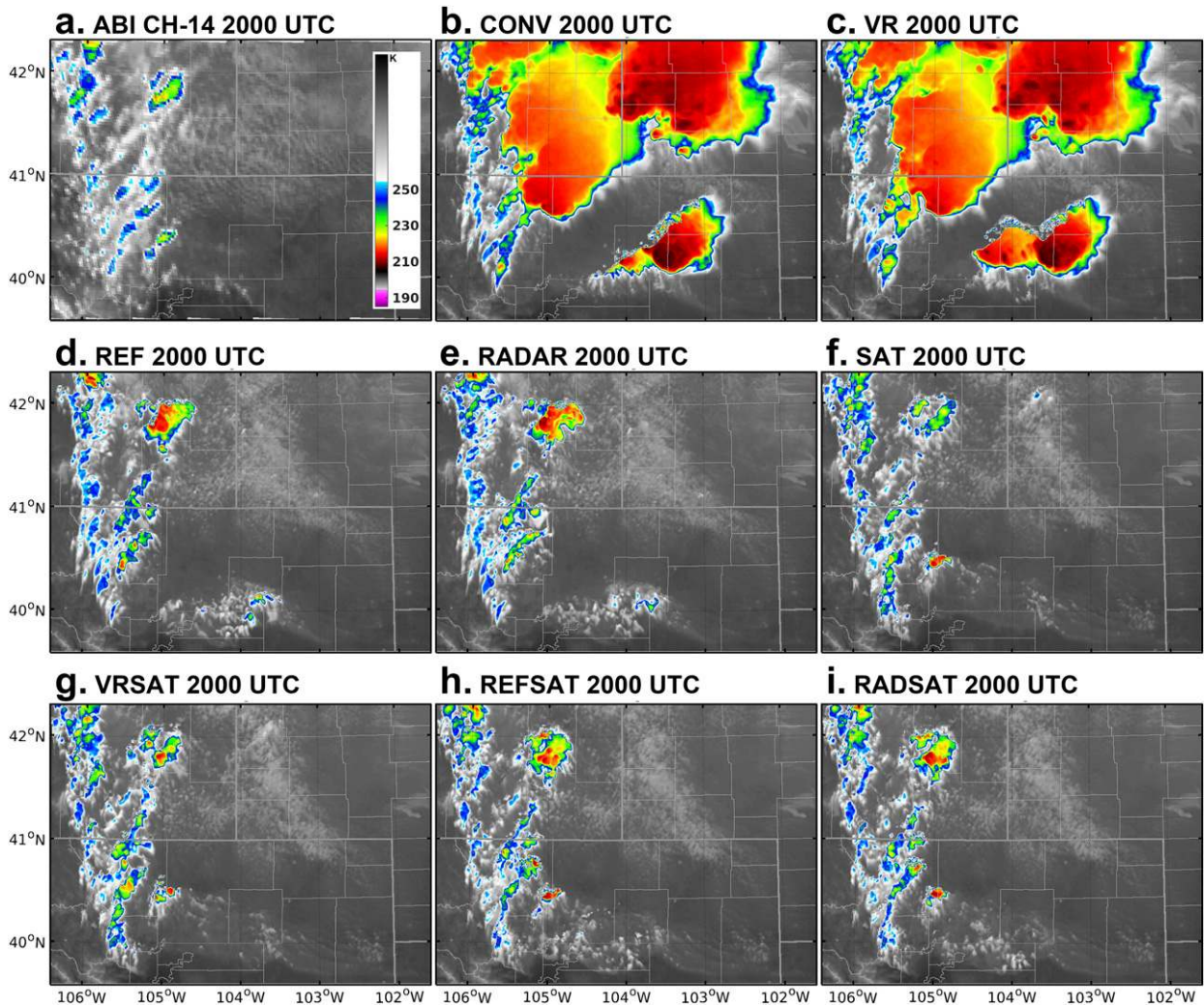
FIG. 4. As in Fig. 3, but at 2000 UTC.

reflectivity show that REF, RADAR, SAT, VRSAT, REFSAT, and RADSAT experiments are all capable of generating analyses that are close to the observations with significant improvements over CONV and VR experiments.

b. Quantitative verification of assimilated observations

Since a thorough examination on the performance of the assimilation of BT observations including RMSE, bias, and rank histogram analyses is already shown in Z18, here we only focus on observation-space root-mean-square innovation (RMSI) verifications of the experiments with different observation platforms using assimilated observations (Fig. 9). RMSI is defined as RMSI $= \sqrt{\langle (d - \langle d \rangle)^2 \rangle}$, where $d = y_o - H(\overline{x^b})$, and represents the departure of EnKF background mean from

observations. All channel 10 BT and 0-dBZ observations are used for verification, whereas Vr and Z SOs are used for verification only if observed $Z > 15$ dBZ (including those SOs rejected by the EnKF) to focus on convective regions. To be consistent with the observation preprocessing, values of simulated Z lower than 5 dBZ are set to 0 dBZ in all EnKF experiments. An additional three-point averaging is applied to the RMSI curve of each experiment to smooth out rapid changes between cycles due to large changes of observation counts in several cycles (Fig. 2); this additional averaging does not change the relative relationships between the experiments. By comparing RMSI against different observations from the different experiments depicted in Fig. 9, we can easily assess the impact of each type of observation when other observations are also assimilated, similar to the observing system experiment (OSE) framework.
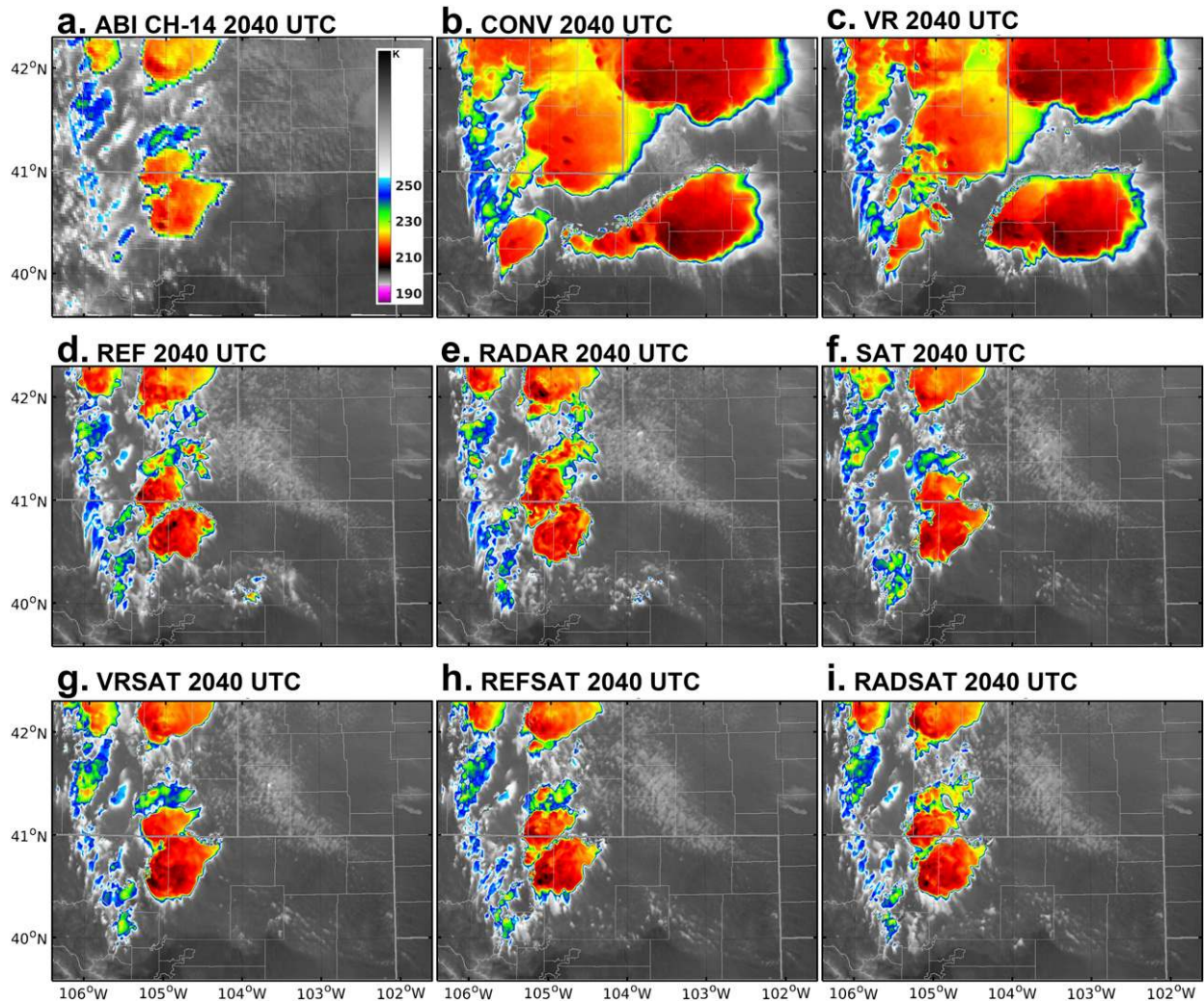
FIG. 5. As in Fig. 3, but at 2040 UTC.

The impact of assimilating BT can be inferred by comparing VRSAT and VR (purple vs yellow), REFSAT and REF (cyan vs red) and RADSAT and RADAR (green vs orange) in Fig. 9. When verified against BT observations (Fig. 9a), the assimilation of BT shows positive impact as expected. RMSI against Vr show generally neutral impact for the assimilation of BT (Fig. 9b), although some improvements occurred in VRSAT and RADSAT compared with VR and RADAR after 2000 UTC. There are also slight improvements of up to 3 dBZ RMSI reduction resulting from assimilating BT observations when verified against Z observations throughout the entire EnKF cycles (Fig. 9c), and slight improvements when verified against 0-dBZ observations (Fig. 9d). This result is reasonable since BT and Z observations both influence thermodynamic and hydrometeor fields.

The impact of assimilating Vr can be inferred by comparing RADAR and REF (orange vs red), or VRSAT and SAT (purple vs blue), and RADSAT and REFSAT (green vs cyan). Again, assimilating Vr reduces RMSI of Vr consistently throughout the entire EnKF cycles by as much as 1 m s^{-1} (Fig. 9b). Although RMSI of BT and Z show a neutral impact of Vr for most of the three above experiment pairs (Figs. 9a,c), a slightly reduced RMSI of Z between VRSAT and SAT of about 2 dBZ after 2000 UTC occurs (Fig. 9c). This result indicates that even in situations where BT and/or Z observations are assimilated, Vr observations may still have the potential to improve the analysis of hydrometeors through cross correlations.

Last, the impact of assimilating Z can be inferred by comparing RADAR and VR (orange vs yellow), REFSAT and SAT (cyan vs blue), and RADSAT and VRSAT (green vs purple). Although assimilating Z significantly reduces RMSI against Z (Fig. 9c) and there is a neutral impact on BT (Fig. 9a), it degrades analysis
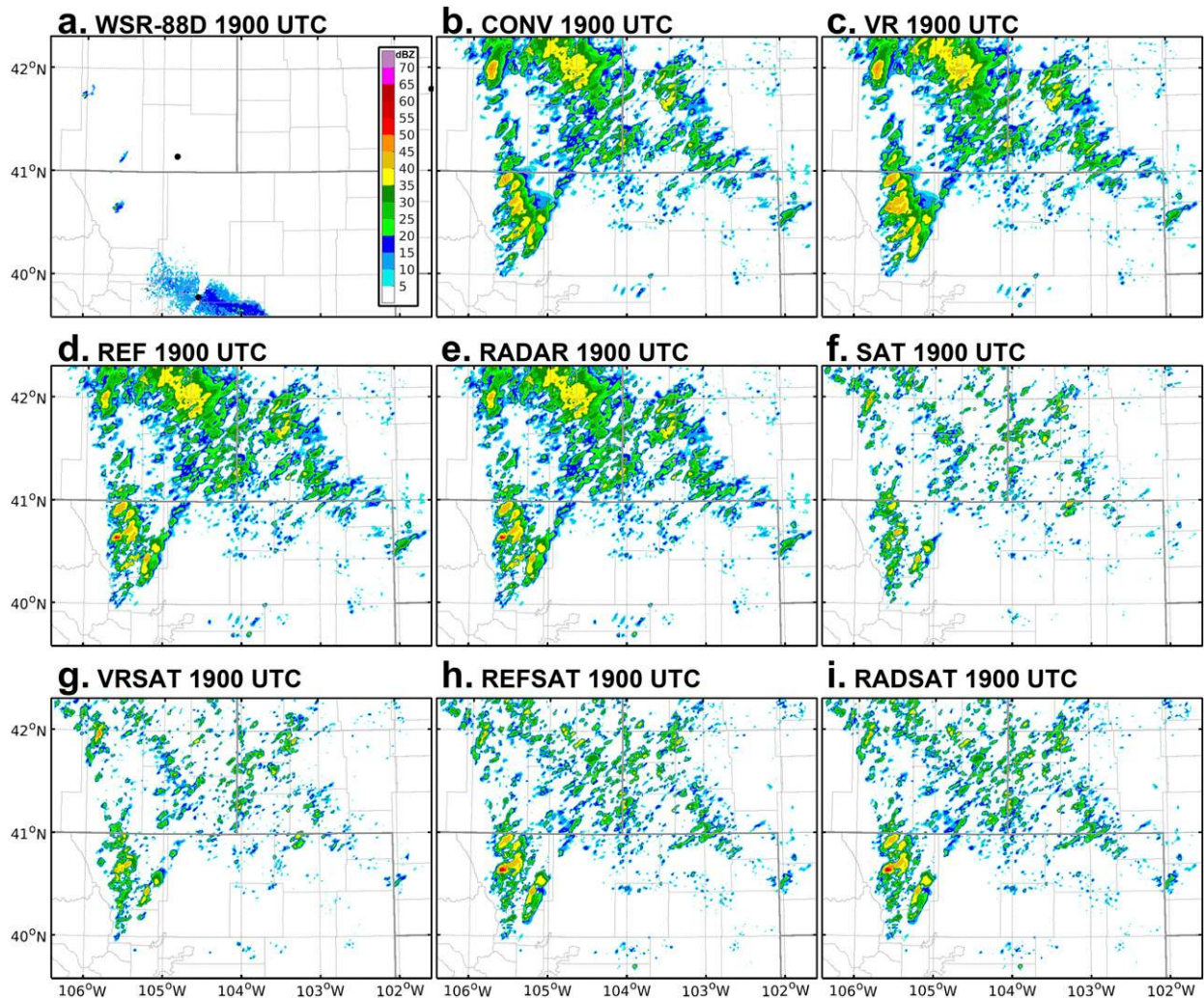
FIG. 6. Composite reflectivity from (a) observations and simulated from EnKF analysis mean of (b) CONV, (c) VR, (d), REF, (e) RADAR, (f) SAT, (g) VRSAT, (h) REFSAT, and (i) RADSAT experiments at 1900 UTC.

accuracy when verifying using Vr observations consistently by about $1\,\mathrm{m\,s^{-1}}$ throughout the entire EnKF cycles (Fig. 9b). This might be associated with the overprediction of the storm near 42°N, 105°W and the spurious stratiform precipitation region associated with the observed dryline near 40°N, 104°W that we found previously when comparing BT and composite reflectivity with the observations (Figs. 4 and 7) and might have contributed to the larger RMSI of Vr when Z observations are assimilated.

c. Evolution of surface conditions

In previous subsection, the verification of BT and 0-dBZ observations included regions outside of the convective storms, with the results showing the benefits from assimilating BT on the depiction of the environment. In this subsection, simulated surface conditions from the EnKF experiments are compared with CONV

and HRRR analysis as well as surface observations at 2000 UTC. Root-mean-square error (RMSE) and bias of EnKF background mean compared with surface observations and ensemble RMSE (''ERMSE'' hereafter; root-mean-square averages of errors of predicted value of all background members) and standard deviations of the background ensemble also are calculated.

At 2000 UTC, all experiments are several K warmer than CONV at the surface over the northwestern portion of the model domain (Figs. 10d–i) except for VR (Fig. 10c). This is due to the removal of spurious clouds in all these EnKF experiments leading to larger amounts of solar radiation reaching the surface and correspondingly less evaporative cooling associated with precipitation. SAT, VRSAT, REFSAT, and RADSAT (Figs. 10f–i) are further warmer than REF and RADAR (Figs. 10c,d), because the removal of spurious precipitation in REF and
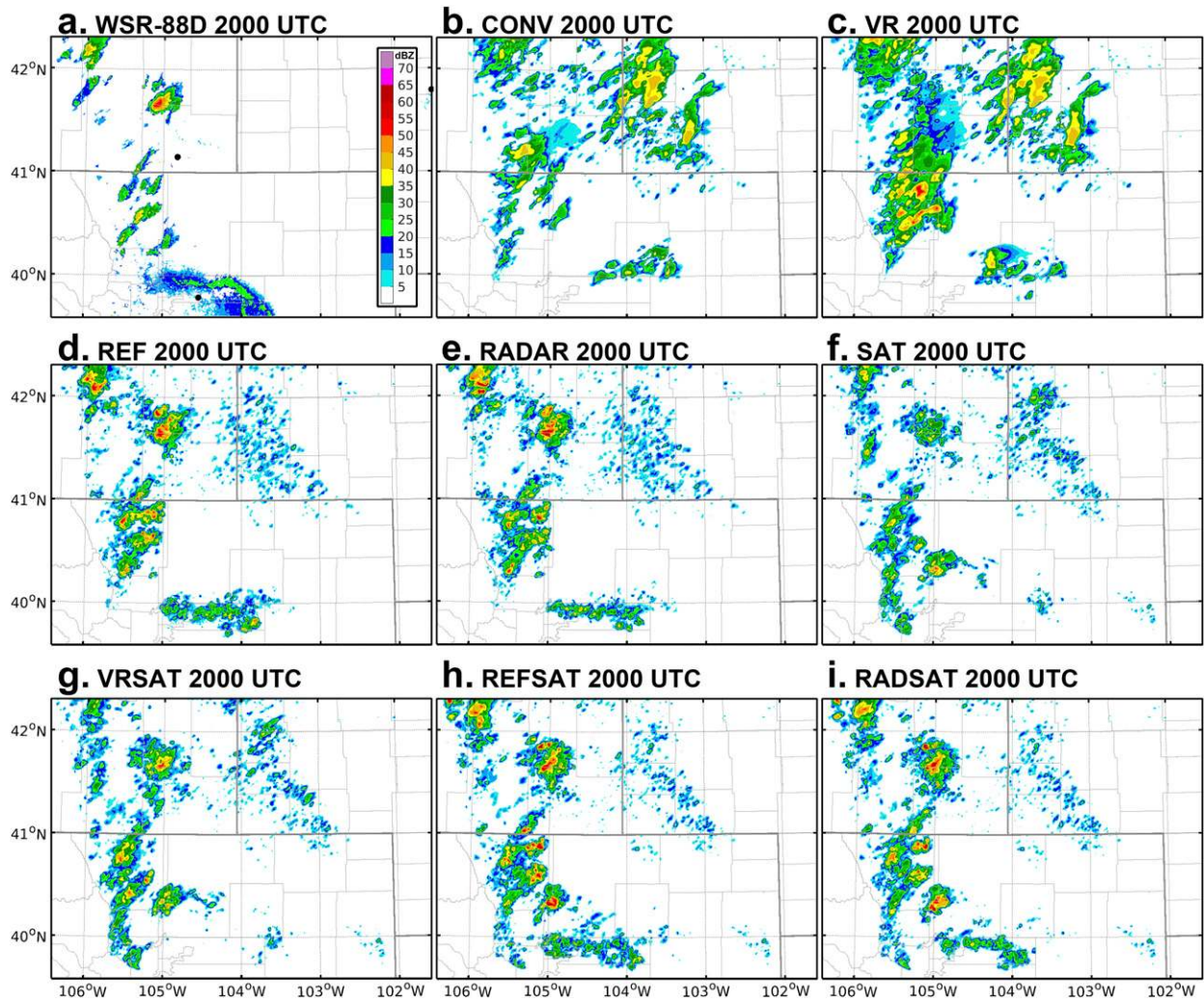
FIG. 7. As in Fig. 6, but at 2000 UTC.

RADAR at the beginning of the EnKF cycles is slower than the other four experiments as revealed in their comparison of composite reflectivity (Fig. 6). Consistent with comparisons of composite reflectivity at the same time (Fig. 3), the stronger storms at this early initiation stage in REF, RADAR, REFSAT, and RADSAT produce several localized cold pools at the surface from the evaporation of precipitation. Although the cold pool near 42°N, 105°W cannot be verified due to insufficient observation coverage, it appears that the analysis of REF, RADAR, REFSAT, and RADSAT are too cold near the southern border of the model domain between 104° and 105°W, while the analysis of SAT and VRSAT are more accurate for this location. Although on average VR has the smallest RMSE of 2-m temperature, its standard deviation is more than twice the values in the other experiments except for CONV. On the other hand, REF and RADAR have the largest RMSE and ERMSE, while

assimilating satellite observations can reduce the errors both in the ensemble mean and in the ensemble members, and VRSAT has the smallest ERMSE and the second smallest RMSE among all radar and satellite assimilation experiments.

The better analysis along the southern domain boundary for SAT and VRSAT can also be seen in surface Td (Fig. 11) which has a very large horizontal gradient in Td indicative of a surface dryline in the HRRR analysis (Fig. 11a). REF and RADAR are more than 10 K higher than the observed Td at this boundary, also leading to significantly larger RMSE and ERMSE compared with other experiments (Figs. 11d,e), while the much drier analysis of SAT and VRSAT are much closer to the observed Td with some of the smallest RMSE and ERMSE among all experiments (Figs. 11f,g). These differences result from the different locations of this dryline in SAT and VRSAT versus the other experiments
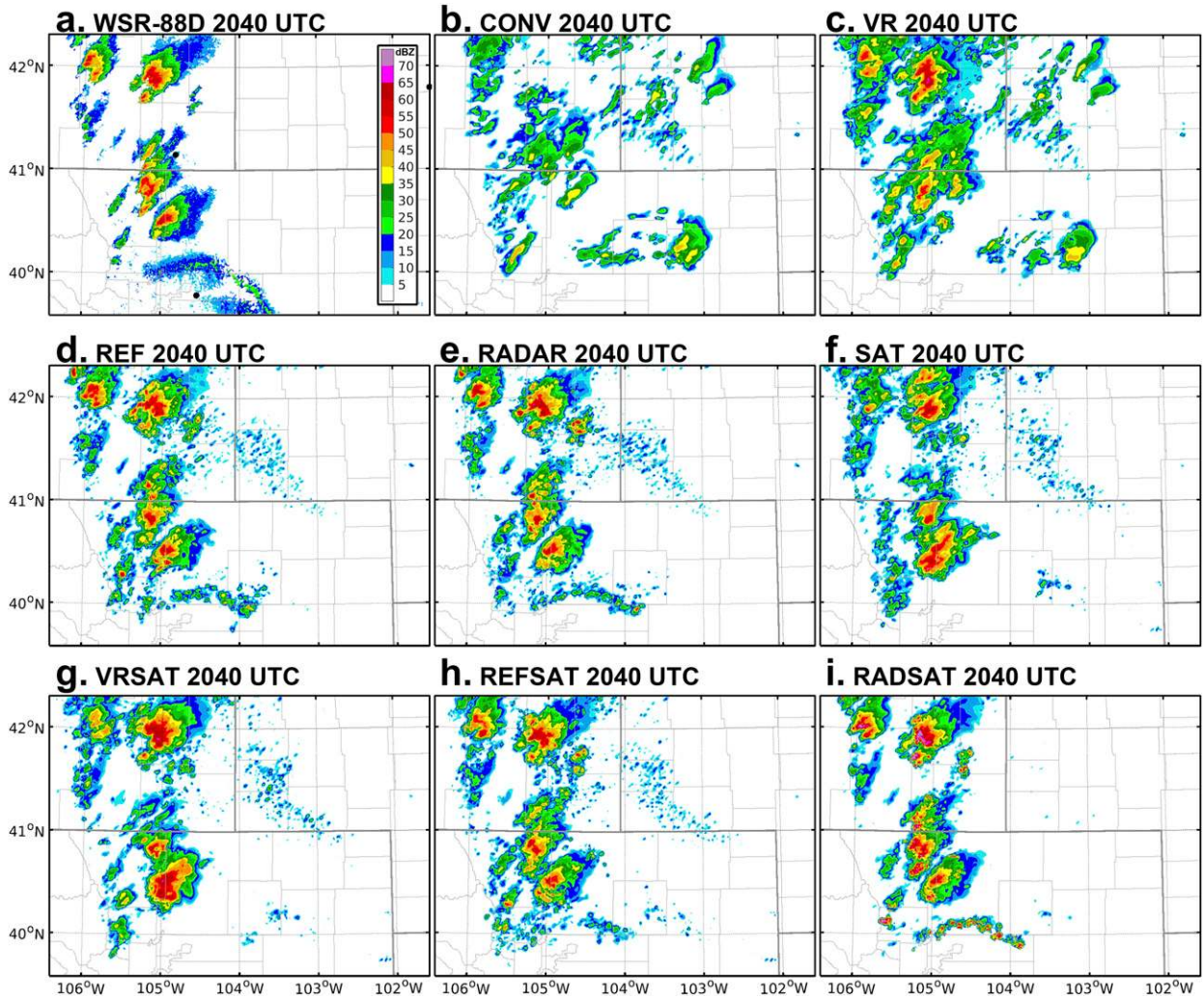
FIG. 8. As in Fig. 6, but at 2040 UTC.

(figure not shown). The northwestern portion of the model domain in SAT, VRSAT, REFSAT, and RADSAT are also drier than CONV, REF, and RADAR, consistent their warmer environment indicative of less spurious precipitation in Fig. 10.

Station-by-station time series of surface variables provide further insights, and two stations in Colorado (locations indicated in Fig. 10a) are presented here because they are the stations that are closest to the primary storms: Fort Collins (FNL) and Greeley (GXY). Time series of the ensemble background for 2-m T and Td at these two stations during EnKF are shown in Fig. 12 (omitting CONV and VR because their environmental conditions are different from the other six experiments) with median, two quartiles and two extrema, and the box-and-whisker distribution of deviations of EnKF background ensemble from surface station observations using all available times of the two stations are shown in

Figs. 13a and 13b. All experiments have a warm bias (Fig. 13a) and a dry bias (Fig. 13b) at both stations. All experiments also analyze a drop of temperature at FNL after 2000 UTC, with this drop occurring abruptly in REF and RADAR at 2015 UTC (Fig. 12a). In contrast, the drop appears gradually in SAT and VRSAT (Fig. 12b) as well as in REFSAT and RADSAT (Fig. 12c) with enhanced cooling in VRSAT compared with SAT (Fig. 12b), resulting in the smallest root-mean-square error of 2-m temperature of VRSAT at these two stations during EnKF (Fig. 13a). Furthermore, SAT and VRSAT correctly analyzed the slight increase of dewpoint (Fig. 12b) while REF, RADAR, REFSAT, and RADSAT show a flat or slight decrease trend (Figs. 12a,c), leading to larger discrepancies of these four experiments especially for the observations just before 2030 UTC compared with SAT and VRSAT. The differences among the experiments are smaller at GXY, although there is a
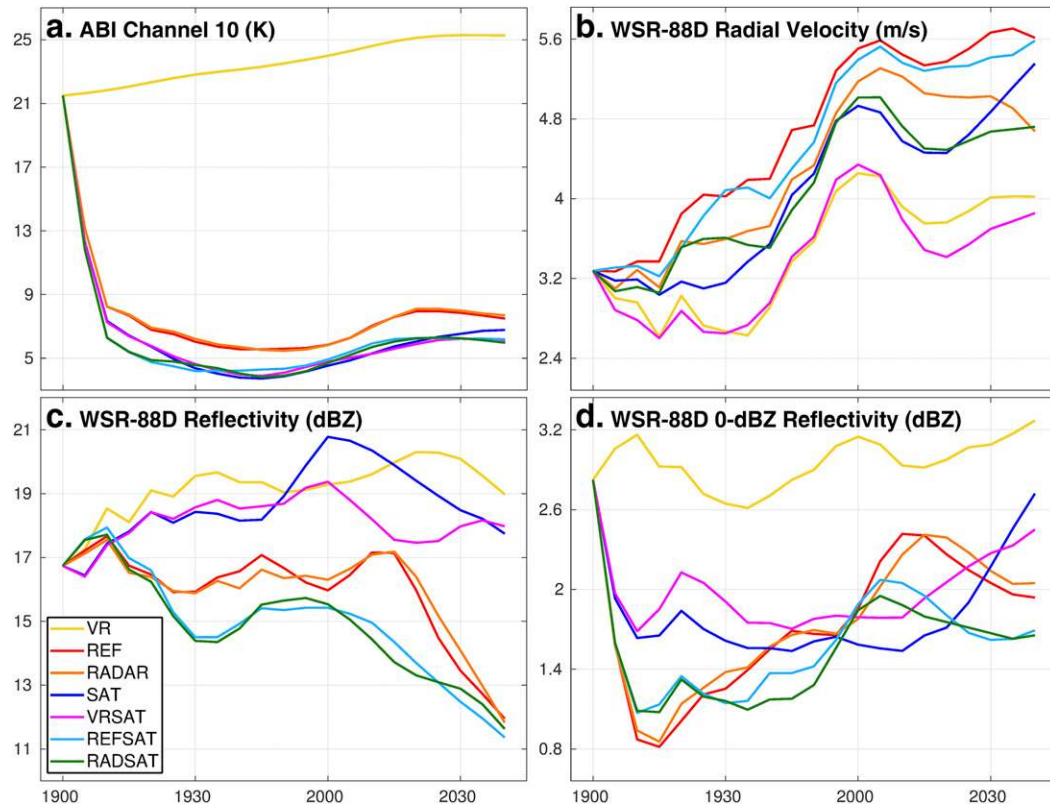
FIG. 9. Root-mean-square innovations verified against (a) ABI channel 10, (b) radial velocity, (c) nonzero reflectivity, and (d) nonprecipitating reflectivity observations. Each curve has been smoothed using a three-point averaging method.

noticeable reduction of both 2-m temperature and dewpoint toward the end of the EnKF cycles in SAT, VRSAT, REFSAT, and RADSAT (Figs. 12e,f) compared with REF and RADAR (Fig. 12d). The apparent benefit on the estimation of near-surface moisture from assimilating BT observations is clearly shown in Fig. 13b in which SAT and VRSAT have the smallest biases, smallest RMSEs and smallest spreads across ensemble members. VRSAT also has the smallest biases, RMSE and spread of 2-m temperature (Fig. 13a).

In summary, verifications with surface observations indicate that assimilating satellite BT observations has the potential to improve the accuracy of the near-surface pre-CI environment, especially for moisture which is one of the necessary ingredients of CI (Doswell et al. 1996), complementing conventional surface observations in regions where surface stations are sparse.

d. Ensemble forecasts

Ensemble forecasts out to 0000 UTC 13 June are initialized from the 1940, 2000, 2020, and 2040 UTC EnKF analyses of each experiment. Values of 2–5-km AGL updraft helicity (UH; Kain et al. 2008) exceeding

$200 \, m^2 \, s^{-2}$ is used to determine the presence of a mesocyclone, a key feature of supercell thunderstorms, in the model output. Ensemble probabilities of mesocyclone tracks over the period from 2100 UTC 12 June to 0000 UTC 13 June are used to identify the most hazardous regions of the predicted storms and the most likely locations of tornado occurrence. A neighborhood approach (Roberts and Lean 2008; Ebert 2009; Schwartz and Sobash 2017) with a neighborhood radius of 3 km is also applied (i.e., the probability value at each grid point indicates the probability of UH exceeding $200 \, m^2 \, s^{-2}$ within 3 km from the grid point). The observed mesocyclone tracks are manually identified in the radar observations when dipole structures (rotations) occur in Vr of KFTG and KCYS radars at the lowest scanning angle (typically 0.5°), and the azimuthal difference ("gate-to-gate" difference) of Vr for the dipole exceeds $10 \, m \, s^{-1}$ for at least three consecutive scans (about 9 min in precipitation scan mode) during the same time span.

For the EF1940 and EF2000 forecasts it is clear that radar data assimilation alone is not able to capture the initiation and evolution of the storms when starting from the preinitiation and initiation stages in the storm life
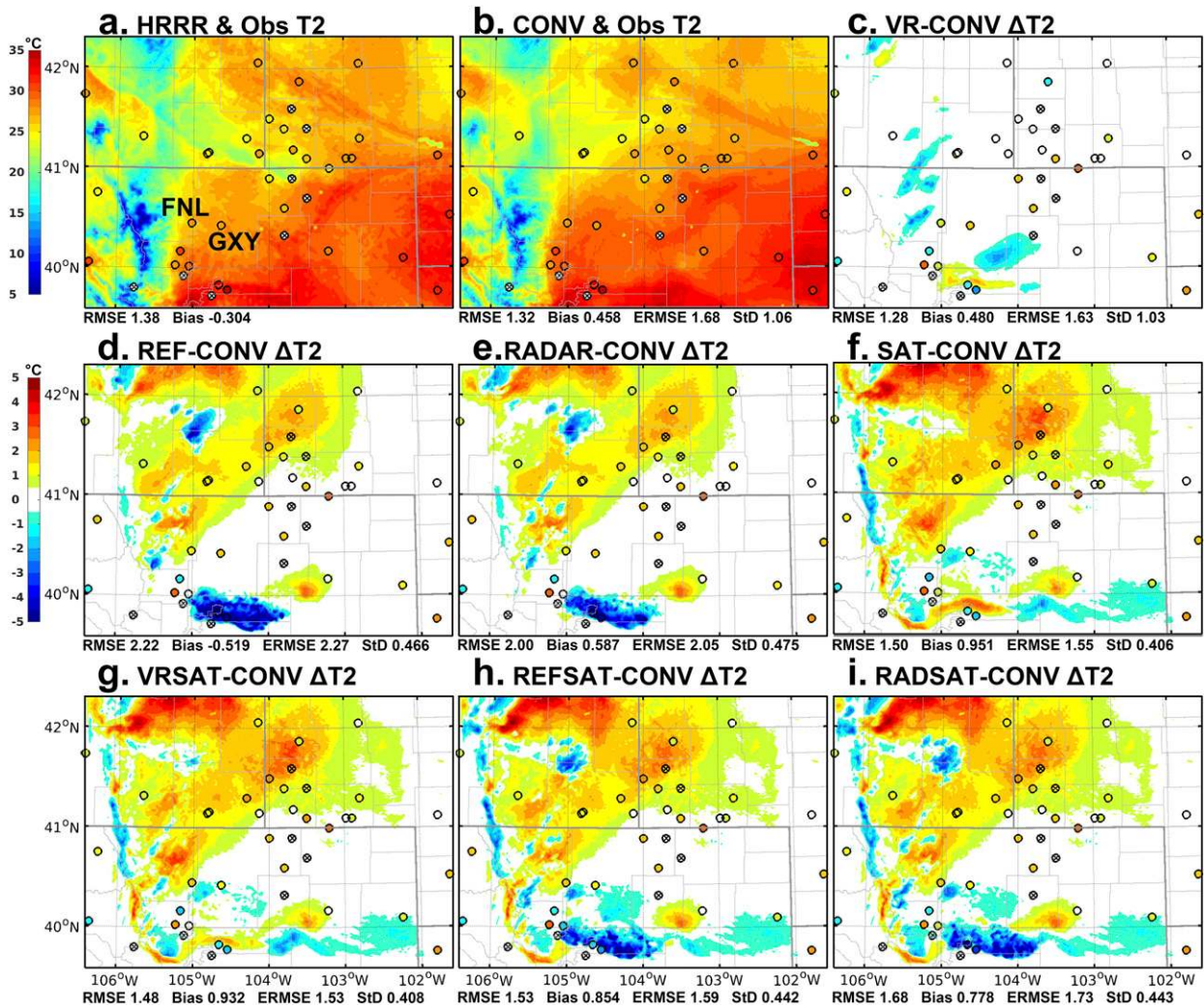
FIG. 10. Ensemble mean 2-m temperature (shaded) with observed 2-m temperature from surface stations (filled circles) from (a) HRRR analysis, and (b) EnKF background mean of CONV, and deviations of EnKF background mean from that of CONV (shaded) and observations (filled circles) of (c) VR, (d) REF, (e) RADAR, (f) SAT, (g) VRSAT, (h) REFSAT, and (i) RADSAT at 2000 UTC. Color scale of (a),(b) is indicated in the color bar to the left of (a), and color scale of (c)–(i) is indicated in the color bar to the left of (d). Locations of the two stations FNL and GXY used for verification in Figs. 12, 13, and 15 are marked in (a). Circles with a cross inside are stations that differences between model elevations and actual elevations are greater than 20 m. Statistics of deviations of EnKF background mean from the observations are listed below each panel (refer to the text for definitions).

cycle from all of the radar-only experiments (VR, REF, and RADAR). These experiments fail to predict any identifiable mesocyclone tracks near the observed mesocyclone tracks in the regions of Wyoming, Nebraska and Colorado in their EF1940 forecast (Figs. 14a1,b1,c1), and in their EF2000 forecast emphasize another storm that is several tens of kilometers to the north and in between the two observed mesocyclone tracks (Figs. 14a2,b2,c2). In contrast, all four satellite-assimilated experiments are able to produce probabilistic UH tracks that align very well with the observed tracks in their EF1940 forecasts that are initialized 20 min before the initiation of the storms (Figs. 14d1,e1,f1,g1), and their predicted tracks

became more concentrated with higher probabilities and match the observed tracks even better in their EF2000 forecasts (Figs. 14d2,e2,f2,g2) when the radar-assimilation experiments are still not able to predict the observed mesocyclone tracks. The probabilistic mesocyclone forecasts of EF1940 highlight the potential of extending forecast lead time of mesocyclones (as well as associated hazards) by as much as twenty to forty minutes with the assimilation of BT observations during preinitiation, initiation, and early development stages of severe storms when BT observations are able to provide more information about the storm and the environment than radar observations.
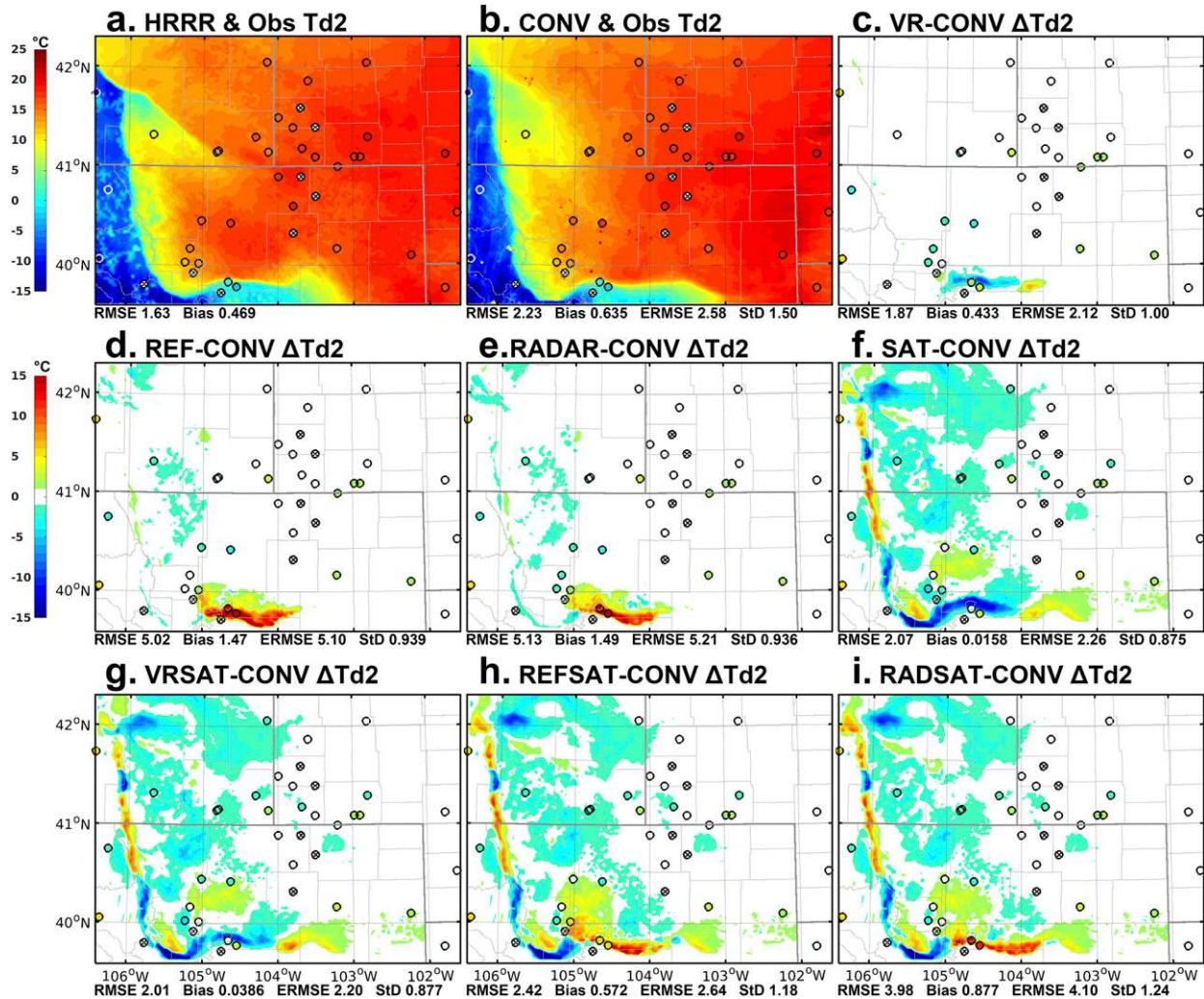
FIG. 11. As in Fig. 10, but for 2-m dewpoint temperatures.

The probabilistic forecasts of mesocyclone tracks for VR, REF, and RADAR become progressively better with more EnKF cycles, although VR failed to match the observed mesocyclone track even for the 2040 UTC ensemble forecasts from its final analyses of EnKF (Fig. 14a4). On the other hand, 2020 UTC forecasts of REF and RADAR as well as 2040 UTC forecast of RADAR all provide accurate predictions of the mesocyclone (Figs. 14b3,c3,c4). However, EF2040 forecast of REF shows a northward bias of the primary mesocyclone track toward a secondary mesocyclone track slightly to the north of the primary one (Fig. 14b4). This UH track can also be seen in some of the other ensemble forecasts and is associated with a weaker storm that occasionally develops rotation but does not consistently meet the shear criteria throughout the forecast period in observed Vr for a mesocyclone. For SAT and VRSAT, the predicted probabilistic

mesocyclone tracks consistently match the observed ones across ensembles initialized from analyses at different times (Figs. 14d,e), except for EF2040 of SAT that show a similar northward bias of the primary track toward the secondary track (Fig. 14d4) similar to EF2040 of REF (Fig. 14b4). This might be resulted from the fact that BT of infrared channels only provide information above cloud top and is unable to penetrate into the thunderstorms. Assimilating Vr observations in addition to BT observations brings improvements: the probability of the mesocyclone track at the center of the model domain in VRSAT progressively increases when more Vr observations are assimilated, such that its EF2040 forecast produces the most confident and best-defined probabilistic mesocyclone track among all forecasts of all experiments (Fig. 14e4). For REFSAT, its EF2040 forecast emphasizes the northern weaker storm, such that the observed mesocyclone track is largely missed
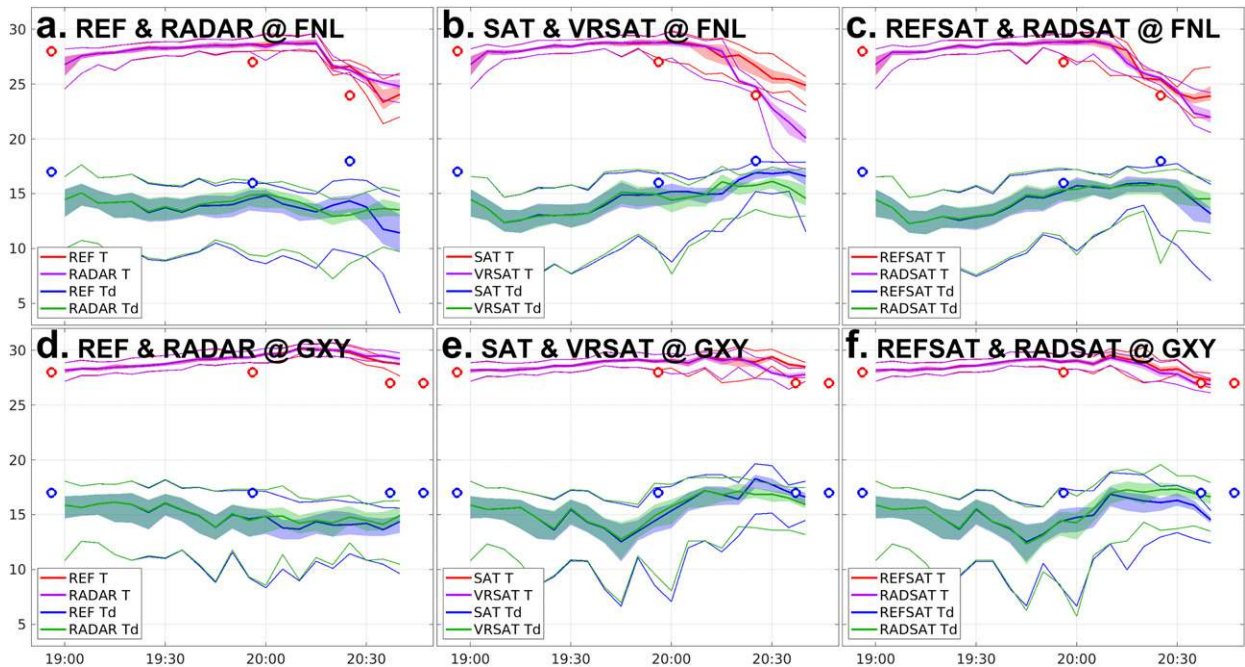
FIG. 12. EnKF background ensemble distributions of 2-m temperature and dewpoint in °C with medians (thick solid lines), regions between quartiles (shaded), and extrema (thin solid lines) and observations (circles) at (a)–(c) FNL, and (d)–(f) GXY stations for (a),(d) REF and RADAR, (b),(e) SAT and VRSAT, and (c),(f) REFSAT and RADSAT experiments.

(Fig. 14f4). RADSAT provides good forecasts from 2000 to 2020 UTC (Figs. 14g2, g3), but deviates toward the secondary track again in its EF2040 forecast (Fig. 14g4). The northward moving bias and displacement of meso-cyclones seen in several ensemble forecasts here might be produced by interactions between adjacent storms that has been reported in several storm-scale data assimilation and modeling studies (e.g., Xue et al. 2014; Wheatley et al. 2015; Carlin et al. 2017), indicating uncertainties in cur-rent storm-scale data assimilation and simulation systems.

Some variations among the ensemble forecasts also occurred for the storm that appeared near the northern boundary of the model domain and is already well or-ganized at 2000 UTC (Fig. 7a). Most ensemble forecasts have shown either a displacement or a moving bias of their predicted mesocyclone tracks compared with the observed one and only a few can provide accurate pre-dictions. The influence of the boundary conditions might have contributed to the overall degraded performance on the prediction of this storm compared with the pri-mary one in the center of the model domain.

The continuously ranked probability score (CRPS; Wilks 2011) and equitable threat score (ETS; Wilks 2011) of composite reflectivity and accumulated pre-cipitation verified against corresponding Multi-Radar Multi-Sensor (MRMS) system products and 2-m T and Td verified against surface station observations over the entire model domain reveal no significant differences among the experiments across different ensemble fore-casts, except for a worse performance of VR for some times (figure not shown). However, the time evolution of 2-m T and Td at FNL and GXY from the EF2040 ensembles of the experiments, which are initialized from the final analysis of EnKF, provide for further assess-ment (Fig. 15). At the FNL station, REF and RADAR predict a 2-m T increase immediately after initialization (Fig. 15a), while the observed temperature gradually decreases till about 2100 UTC. On the other hand, SAT and VRSAT follow the trend of observed temperature more closely and VRSAT shows the smallest warm bias among all experiments (Fig. 15b), while REFSAT and RADSAT predicted more complex evolutions of temper-ature before 2200 UTC (Fig. 15c). Many of the ensemble members of all these experiments predict a drying at FNL after about 2200 UTC (Figs. 15a–c), which is not seen in the observations, with VRSAT being one of the ensembles to have incorrectly predicted this drying (Fig. 15b).

Station GXY, although not directly influenced by the cold pools of the storms, is crossed by the dryline at 2200 UTC with a dewpoint drop of at least 8 K between 2156 and 2203 UTC. Both SAT and VRSAT capture this sharp drop of dewpoint (Fig. 15e) with all ensemble members predicting a change of at least 10 K, although there are considerable timing differences among en-semble members. The other four experiments tend to predict a more gradual decrease of dewpoint rather than
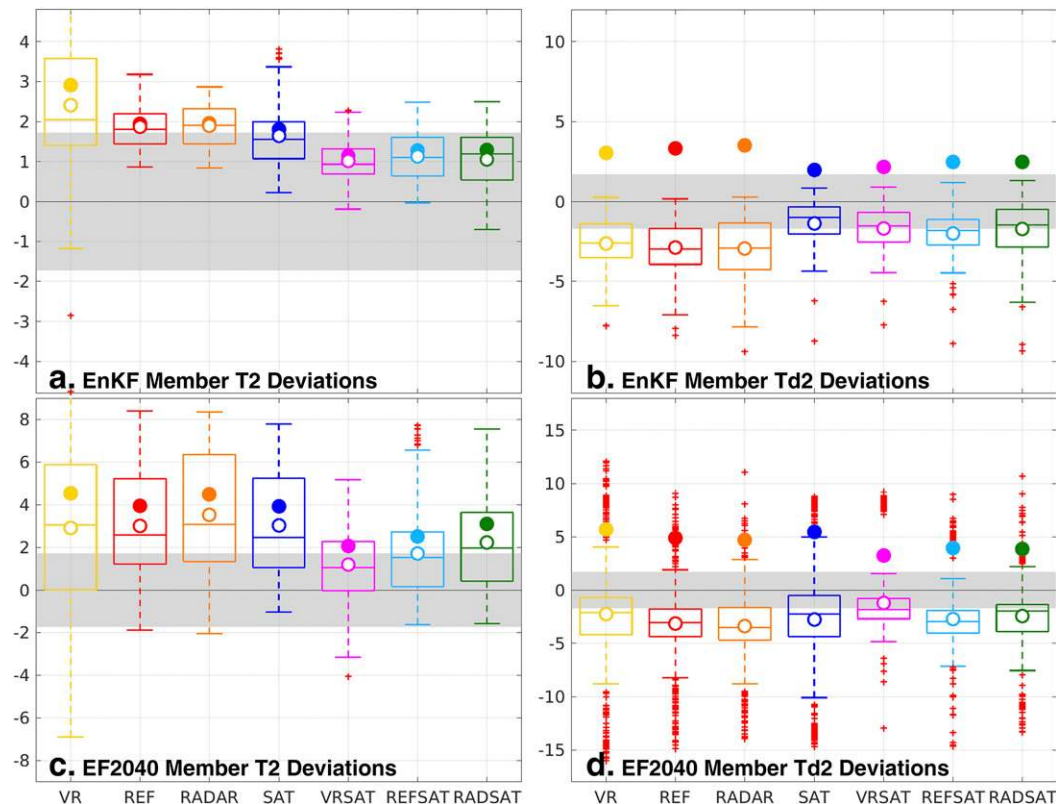
FIG. 13. Box-and-whisker distributions with medians, quartiles, extrema excluding outliers, and outliers of deviations from (a),(c) 2-m temperature and (b),(d) dewpoint temperature observations of (a),(b) EnKF background members during the entire EnKF cycles, and (c),(d) EF2040 forecasts during the entire ensemble forecasts. Filled dots are RMSEs and circles are mean biases. Shaded regions show the estimated uncertainty of surface observations.

a sharp drop (Figs. 15d,f), resulting from their different analysis of the dryline compared with SAT and VRSAT (Fig. 11), although there are a few members in these four experiments that predict a sharp decrease.

Box-and-whisker distributions of deviations of EF2040 for each experiment combining all available observations from these two stations during ensemble forecasts are presented in Figs. 13c and 13d. Consistent with previous trends seen in Fig. 15, VRSAT has the smallest magnitudes of RMSE, bias, and median for both 2-m T (Fig. 13c) and Td (Fig. 13d) and REFSAT and RADSAT have slightly larger values for these metrics. Generally, the verification results indicate that VRSAT has the overall best prediction of the surface conditions and cold pools associated with the storms, which may also have contributed to its overall best prediction of the mesocyclone tracks (Fig. 14e).

4. Conclusions and discussion

Following onto our previous study (Z18) that assimilated all-sky infrared BT observations from the ABI on board the *GOES-16* satellite using EnKF with a numerical

model running at 1-km horizontal grid spacing with results showing promising improvements in the prediction of a severe tornadic thunderstorm event on 12 June 2017 across Wyoming, Nebraska, and Colorado, the current work seeks to explore the simultaneous assimilation of both satellite and radar observations with the same data assimilation system for the same event. Aside from assimilating surface observations every 20 min, each of the experiments assimilate a different combination of radial velocity (Vr) and reflectivity (Z) observations from two WSR-88D radars and BT observations from *GOES-16* ABI every 5 min for a time span of up to 100 min. This setup allows us to isolate the influence of each observation type when assimilated together with other observations in a manner similar to the observing system experiment (OSE) approach.

The EnKF analysis of simulated BT and reflectivity indicate that all experiments (except for VR that contains no observations to remove spurious convection) are able to reproduce the initiation and early development of the storms properly. Quantitative verification using Vr, Z, and BT observations show that each of these observation types can reduce the magnitude of their own
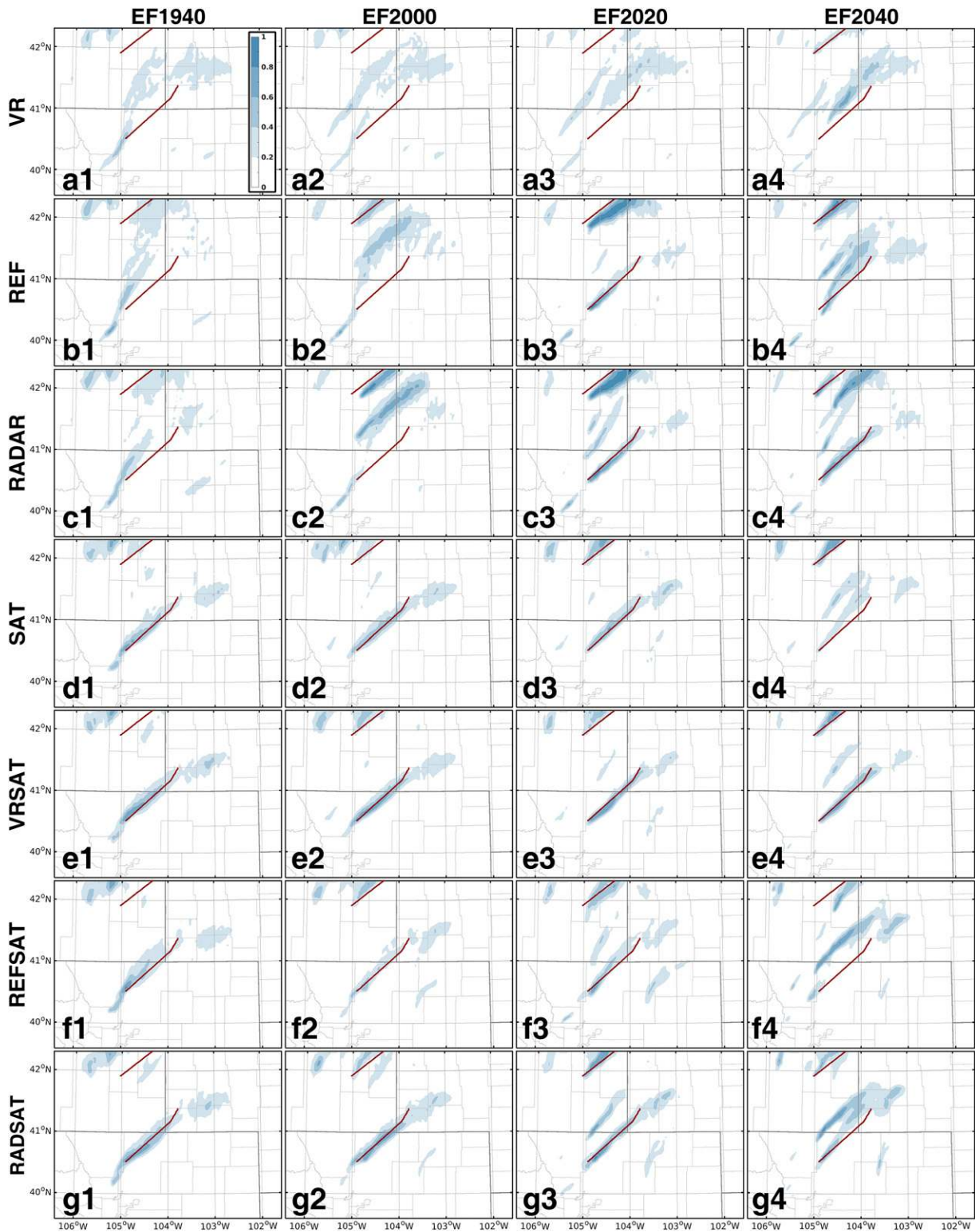
FIG. 14. Ensemble neighborhood probability of forecast UH exceeding $200\,\mathrm{m}^2\,\mathrm{s}^{-2}$ from (a) VR, (b) REF, (c) RADAR, (d) SAT, (e) VRSAT, (f) RADSAT, and (g) RADSAT experiments for 1) EF1940, 2) EF2000, 3) EF2020, and 4) EF2040 ensemble forecasts. Red solid lines are observed mesocyclone tracks identified manually from WSR-88D observations.
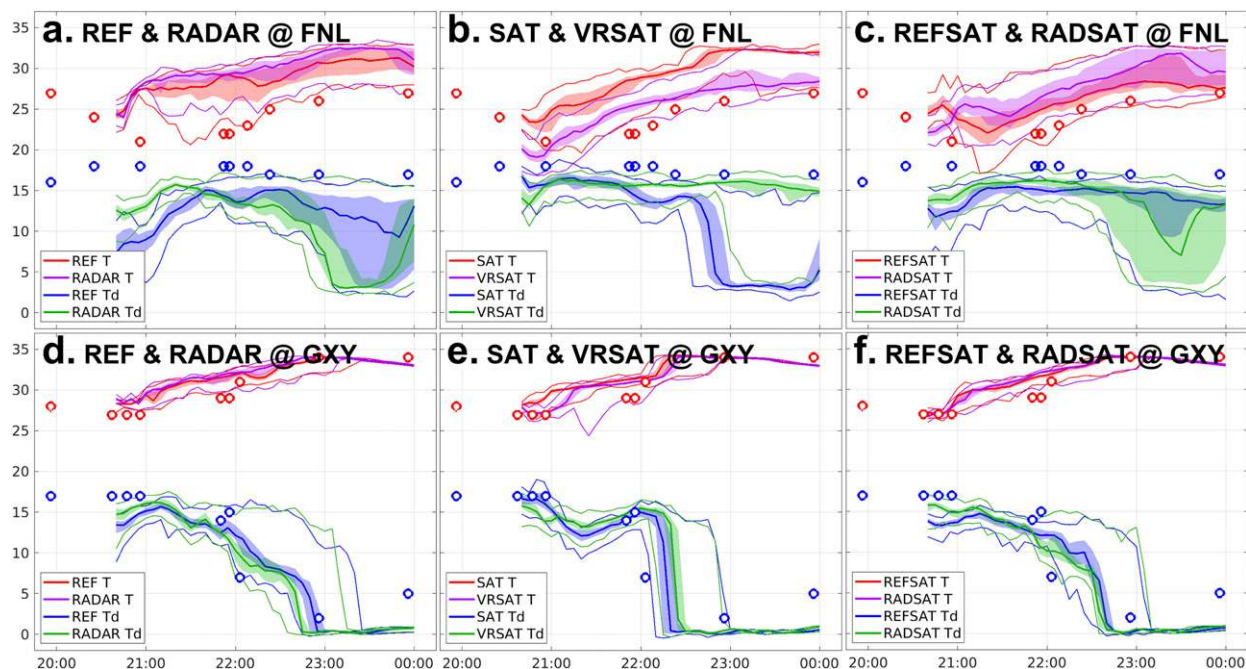
FIG. 15. As in Fig. 12, but for EF2040 ensemble forecasts.

innovations as expected when EnKF is working properly. The assimilation of Vr or BT observations also have a neutral or positive impact on the reduction of innovations of other types of observations. However, assimilating Z observations might lead to larger errors when verified against Vr observations, which might be associated with the overprediction of storm intensity and the spuriously simulated stratiform precipitation associated with an observed dryline (without precipitation) when Z observations are assimilated.

Ensemble forecasts from the 1940, 2000, 2020, and 2040 UTC EnKF analyses of each experiment also are performed, and the probabilistic midlevel mesocyclone forecasts are verified against observed mesocyclone tracks identified from radar observations. The benefits of satellite observations at the initiation and early development stages of the storms are apparent in the 1940 and 2000 UTC ensemble forecasts: experiments that assimilate BT observations are able to predict the observed mesocyclone track with high accuracy and confidence from their 1940 and 2000 UTC EnKF analysis. This occurs even before the first 5-dBZ reflectivity echo of the primary tornadic thunderstorm appears. Neither of the three experiments that solely assimilated radar observations (VR, REF, and RADAR) at this time are able to accurately predict the mesocyclone track. These two radar-assimilation experiments are able to predict the track from their 2020 and 2040 UTC EnKF analysis after more radar observations are assimilated, although predicted mesocyclone

track of EF2040 of REF show a northward moving bias that has been reported in other storm-scale simulations of severe thunderstorms. The experiment that assimilate both Vr and BT observations (VRSAT) show consistent and reliable predictions of the mesocyclone track in different ensemble forecasts and produces the most accurate predictions with the highest confidence among all experiments. On the other hand, REFSAT and RADSAT tends to emphasize an adjacent storm that has a much weaker observed rotation than the tornadic storm seen from observations with more EnKF cycles, missing the primary storm.

From the results of these data assimilation experiments, it is apparent that the availability of BT observations before storm initiation is one of its key advantages over radar observations. Satellite can not only "see" the occurrence of convective clouds several tens of minutes earlier than radar observations, but also provide valuable information on the surrounding environment that is not available from operational Doppler weather radars. The assimilation of BT observations before and during initiation as well as during the early development of the storms provide valuable information when the amount of radar observations is limited. As a consequence, up to 40 min of additional forecast lead time for the mesocyclone tracks can be gained for this event when BT observations are assimilated instead of radar observations. This conclusion is supported by the observed mesocyclone track being accurately predicted from the 1940 UTC

analysis that assimilates BT observations when the storms are barely formed, while for the experiments that assimilate radar observations this track is predicted only from analyses later than 2020 UTC. An extension of 40 min onto the forecast lead time for mesocyclone is a significant improvement and comparable to the current average warning lead times of tornadoes (Stensrud et al. 2009). This result is very promising for the future of warning operations for severe weather events, since these valuable cloud-affected IR BT observations have not been assimilated in any operational or quasi-operational weather prediction systems for any scales. On the other hand, the impact of assimilating all-sky IR BT observations from geostationary satellites when the thunderstorms are matured might be limited compared with radar observations, since cloud-affected IR BT observations mostly contain two-dimensional information at the cloud top of the thunderstorms, while the detailed three-dimensional structures of the thunderstorms that can be nicely observed by the radars are completely invisible to geostationary satellites at infrared wavelengths.

The further improvement resulting from the additional assimilation of Vr observations compared with results when only BT observations are assimilated is not surprising. IR BT of ABI channel 10 are mostly sensitive to temperature and water vapor in the lower troposphere in clear sky and to the top of clouds (in other words, characteristics of hydrometeors) in cloudy regions, and Vr observed wind fields within the storms, thus they provide complementary information on the thermodynamic and dynamic features of the storms, respectively. The benefit obtained from simultaneous assimilation of BT and Vr observations also persists into the ensemble forecasts and may extend for several hours, indicating that the environmental information provided by BT observations might help to extend the practical predictability of severe storms (e.g., Aksoy et al. 2010; Gasperoni et al. 2013). This also implies a potential benefit from assimilating BT observations in improving the analysis of environmental conditions, especially moisture and therefore instability, in regions where surface stations are sparse.

Overall, assimilating Vr and BT observations simultaneously is the best approach considering both the quality of EnKF analysis and the accuracy of ensemble forecasts for the prediction of severe storms in this event. However, this study is based on one case, and whether or not this combination can lead to a consistent improvement in the prediction of severe weather needs to be extensively evaluated using more events in the future. The benefit of the additional forecast lead times for severe thunderstorms when satellite observations are assimilated compared with radar observation assimilation only should also be evaluated with more events. Besides,

in this study BT observations are assimilated in raw spatial and temporal resolution. Although Honda et al. (2018a) indicate that assimilate BT observations over a shorter temporal interval are beneficial for the ensemble-based analysis and prediction, the optimal combination of satellite and radar observations, their respective spatial and temporal resolution, the best strategies to constrain and utilize both observations considering the balance between available computational resources, scheduling of warning processes, and efficiency and accuracy of the resulting predictions are still unclear and deserve further exploration in the future.

REFERENCES

Aksoy, A., D. C. Dowell, and C. Snyder, 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, https://doi.org/10.1175/2009MWR3086.1.

Carlin, J., J. Gao, J. C. Snyder, and A. V. Ryzhkov, 2017: Assimilation of ZDR columns for improving the spinup and forecast of convective storms in storm-scale models: Proof-of-concept experiments. *Mon. Wea. Rev.*, **145**, 5033–5057, https://doi.org/10.1175/MWR-D-17-0103.1.

Cintineo, R. M., J. A. Otkin, T. A. Jones, S. Koch, and D. J. Stensrud, 2016: Assimilation of synthetic GOES-R ABI infrared brightness temperatures and WSR-88D radar observations in a high-resolution OSSE. *Mon. Wea. Rev.*, **144**, 3159–3180, https://doi.org/10.1175/MWR-D-15-0366.1.

Doswell, C. A., H. E. Brooks, and R. A. Maddox, 1996: Flash flood forecasting: An ingredient-based methodology. *Wea. Forecasting*, **11**, 560–581, https://doi.org/10.1175/1520-0434(1996)011<0560:FFFAIB>2.0.CO;2.

Ebert, E. E., 2009: Neighborhood verification: A strategy for rewarding close forecasts. *Wea. Forecasting*, **24**, 1498–1510, https://doi.org/10.1175/2009WAF2222251.1.

Ek, M. B., 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, https://doi.org/10.1029/2002JD003296.

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 uncertainties. *Mon. Wea. Rev.*, **135**, 1846–1868, https://doi.org/10.1175/MWR3391.1.

Gallo, B. T., and Coauthors, 2017: Breaking new ground in severe weather prediction: The 2015 NOAA/Hazardous Weather Testbed Spring Forecasting Experiment. *Wea. Forecasting*, **32**, 1541–1568, https://doi.org/10.1175/WAF-D-16-0178.1.

Gaspari, G., and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.*, **125**, 723–757, https://doi.org/10.1002/qj.49712555417.

Gasperoni, N. A., M. Xue, R. D. Palmer, and J. Gao, 2013: Sensitivity of convective initiation prediction of near-surface moisture when assimilating radar reflectivity: Impact tests using OSSEs. *J. Atmos. Oceanic Technol.*, **30**, 2281–2302, https://doi.org/10.1175/JTECH-D-12-00038.1.

Geer, A. J., and Coauthors, 2018: All-sky satellite data assimilation at operational weather forecasting centres. *Quart. J. Roy. Meteor. Soc.*, **144**, 1191–1217, https://doi.org/10.1002/qj.3202.

Han, Y., P. van Delst, Q. Liu, F. Weng, B. Yan, R. Treadon, and J. Derber, 2006: JCSDA Community Radiative Transfer Model (CRTM)—version 1. NOAA Tech. Rep. NESDIS 122, 40 pp.

Hoekstra, S., K. Klockow, R. Riley, J. Brotzge, H. Brooks, and S. Erickson, 2011: A preliminary look at the social perspective of Warn-on-Forecast: Preferred tornado warning lead time and the general public's perceptions of weather risks. *Wea. Climate Soc.*, **3**, 128–140, https://doi.org/10.1175/2011WCAS1076.1.

Honda, T., and Coauthors, 2018a: Assimilating all-sky *Himawari-8* infrared radiances: A case of Typhoon Soudelor (2015). *Mon. Wea. Rev.*, **146**, 213–229, https://doi.org/10.1175/MWR-D-16-0357.1.

——, S. Kotsuki, G.-Y. Lien, Y. Maejima, K. Okamoto, and T. Miyoshi, 2018b: Assimilation of Himawari-8 all-sky radiances every 10 minutes: Impact on precipitation and flood risk prediction. *J. Geophys. Res. Atmos.*, **123**, 965–976, https://doi.org/10.1002/2017JD027096.

Houtekamer, P. L., and H. L. Mitchell, 2001: A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.*, **129**, 123–137, https://doi.org/10.1175/1520-0493(2001)129<0123:ASEKFF>2.0.CO;2.

Iacono, M., 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, https://doi.org/10.1029/2008JD009944.

Janjić, Z. I., 1994: The step-mountain eta coordinate model: Further developments of the convection, viscous sublayer, and turbulence closure schemes. *Mon. Wea. Rev.*, **122**, 927–945, https://doi.org/10.1175/1520-0493(1994)122<0927:TSMECM>2.0.CO;2.

——, 1996: The surface layer in the NCEP Eta Model. Preprints, *11th Conf. on Numerical Weather Prediction*, Norfolk, VA, Amer. Meteor. Soc., 354–355.

Jones, T. A., J. A. Otkin, D. J. Stensrud, and K. Knopfmeier, 2013: Assimilation of satellite infrared radiances and Doppler radar observations during a cool season observing system simulation experiment. *Mon. Wea. Rev.*, **141**, 3273–3299, https://doi.org/10.1175/MWR-D-12-00267.1.

——, ——, ——, and ——, 2014: Forecast evaluation of an observing system simulation experiment assimilating both radar and satellite data. *Mon. Wea. Rev.*, **142**, 107–124, https://doi.org/10.1175/MWR-D-13-00151.1.

——, D. Stensrud, L. Wicker, P. Minnis, and R. Palikonda, 2015: Simultaneous radar and satellite data storm-scale assimilation using an ensemble Kalman filter approach for 24 May 2011.

Mon. Wea. Rev., **143**, 165–194, https://doi.org/10.1175/MWR-D-14-00180.1.

——, K. Knopfmeier, D. Wheatley, G. Creager, P. Minnis, and R. Palikonda, 2016: Storm-scale data assimilation and ensemble forecasting with the NSSL experimental Warn-on-Forecast system. Part II: Combined radar and satellite data experiments. *Wea. Forecasting*, **31**, 297–327, https://doi.org/10.1175/WAF-D-15-0107.1.

——, X. Wang, P. Skinner, A. Johnson, and Y. Wang, 2018: Assimilation of *GOES-13* imager clear-sky water vapor (6.5 μm) radiances into a Warn-on-Forecast system. *Mon. Wea. Rev.*, **146**, 1077–1107, https://doi.org/10.1175/MWR-D-17-0280.1.

Kain, J. S., and Coauthors, 2008: Some practical considerations regarding horizontal resolution in the first generation of operational convection-allowing NWP. *Wea. Forecasting*, **23**, 931–952, https://doi.org/10.1175/WAF2007106.1.

Kerr, C. A., D. J. Stensrud, and X. Wang, 2015: Assimilation of cloud-top temperature and radar observations of an idealized splitting supercell using an observing system simulation experiment. *Mon. Wea. Rev.*, **143**, 1018–1034, https://doi.org/10.1175/MWR-D-14-00146.1.

Lawson, J. R., J. S. Kain, N. Yussouf, D. C. Dowell, D. M. Wheatley, K. H. Knopfmeier, and T. A. Jones, 2018: Advancing from convection-allowing NWP to Warn-on-Forecast: Evidence of progress. *Wea. Forecasting*, **33**, 599–607, https://doi.org/10.1175/WAF-D-17-0145.1.

Minamide, M., and F. Zhang, 2017: Adaptive observation error inflation for assimilating all-sky satellite radiance. *Mon. Wea. Rev.*, **145**, 1063–1081, https://doi.org/10.1175/MWR-D-16-0257.1.

——, and ——, 2018: Assimilation of all-sky infrared radiances from *Himawari-8* and impacts of moisture and hydrometer initialization on convection-permitting tropical cyclone prediction. *Mon. Wea. Rev.*, **146**, 3241–3258, https://doi.org/10.1175/MWR-D-17-0367.1.

——, and ——, 2019: Adaptive background error inflation for assimilating all-sky satellite radiance. *Quart. J. Roy. Meteor. Soc.*, **145**, 805–823, https://doi.org/10.1002/qj.3466.

Okamoto, K., Y. Sawada, and M. Kunii, 2019: Comparison of assimilating all-sky and clear-sky infrared radiances from Himawari-8 in a mesoscale system. *Quart. J. Roy. Meteor. Soc.*, **145**, 745–766, https://doi.org/10.1002/qj.3463.

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

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

Roberts, N. M., and H. W. Lean, 2008: Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Mon. Wea. Rev.*, **136**, 78–97, https://doi.org/10.1175/2007MWR2123.1.

Sawada, Y., K. Okamoto, M. Kunii, and T. Miyoshi, 2019: Assimilating every-10-minute Himawari-8 infrared radiance to improve convective predictability. *J. Geophys. Res. Atmos.*, **124**, 2546–2561, https://doi.org/10.1029/2018JD029643.

Schwartz, C. S., and R. A. Sobash, 2017: Generating probabilistic forecasts from convection-allowing ensembles using neighborhood approach: A review and recommendations. *Mon. Wea. Rev.*, **145**, 3397–3418, https://doi.org/10.1175/MWR-D-16-0400.1.

Skamarock, W. C., and Coauthors, 2008: A description of the Advanced Research WRF version 3. NCAR Tech. Note NCAR/TN-475+STR, 113 pp., https://doi.org/10.5065/D68S4MVH.

Skinner, P. S., and Coauthors, 2018: Object-based verification of a prototype Warn-on-Forecast system. *Wea. Forecasting*, **33**, 1225–1250, https://doi.org/10.1175/WAF-D-18-0020.1.

Sobash, R. A., G. S. Romine, C. S. Schwartz, D. J. Gagne II, and M. L. Weisman, 2016: Explicit forecasts of low-level rotation from convection-allowing models for next-day tornado prediction. *Wea. Forecasting*, **31**, 1591–1614, https://doi.org/10.1175/WAF-D-16-0073.1.

Soler, T., and D. W. Eisemann, 1994: Determination of look angles to geostationary communication satellites. *J. Surv. Eng.*, **120**, 115–127, https://doi.org/10.1061/(ASCE)0733-9453(1994)120:3(115).

Stensrud, D. J., and Coauthors, 2009: Convective-scale Warn-on-Forecast system: A vision for 2020. *Bull. Amer. Meteor. Soc.*, **90**, 1487–1499, https://doi.org/10.1175/2009BAMS2795.1.

——, and Coauthors, 2013: Progress and challenges with Warn-on-Forecast. *Atmos. Res.*, **123**, 2–16, https://doi.org/10.1016/j.atmosres.2012.04.004.

Thompson, G., P. R. Field, R. M. Rasmussen, and W. D. Hall, 2008: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow parameterization. *Mon. Wea. Rev.*, **136**, 5095–5115, https://doi.org/10.1175/2008MWR2387.1.

Weng, Y., and F. Zhang, 2012: Assimilating airborne Doppler radar observations with and ensemble Kalman filter for convection-permitting hurricane initialization and prediction: Katrina (2005). *Mon. Wea. Rev.*, **140**, 841–859, https://doi.org/10.1175/2011MWR3602.1.

Wheatley, D. M., K. H. Knopfmeier, T. A. Jones, and G. J. Creager, 2015: Storm-scale data assimilation and ensemble forecasting with the NSSL experimental Warn-on-Forecast system. Part I: Radar data experiments. *Wea. Forecasting*, **30**, 1795–1817, https://doi.org/10.1175/WAF-D-15-0043.1.

Wilks, D. S., 2011: *Statistical Methods in the Atmospheric Sciences*. 3rd ed. International Geophysics Series, Vol. 100, Academic Press, 704 pp.

Xue, M., M. Hu, and A. D. Schenkman, 2014: Numerical prediction of the 8 May 2003 Oklahoma City tornadic supercell and embedded tornado using ARPS with the assimilation of WSR-88D data. *Wea. Forecasting*, **29**, 39–62, https://doi.org/10.1175/WAF-D-13-00029.1.

Yussouf, N., D. C. Dowell, L. J. Wicker, K. H. Knopfmeier, and D. M. Wheatley, 2015: Storm-scale data assimilation and ensemble forecasts for the 27 April 2011 severe weather outbreak in Alabama. *Mon. Wea. Rev.*, **143**, 3044–3066, https://doi.org/10.1175/MWR-D-14-00268.1.

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, https://doi.org/10.1175/1520-0493(2004)132<1238:IOIEAO>2.0.CO;2.

——, Y. Weng, J. A. Sippel, Z. Meng, and C. H. Bishop, 2009: Cloud-resolving hurricane initialization and prediction through assimilation of Doppler radar observations with an ensemble Kalman filter. *Mon. Wea. Rev.*, **137**, 2105–2125, https://doi.org/10.1175/2009MWR2645.1.

——, M. Minamide, and E. E. Clothiaux, 2016: Potential impacts of assimilating all-sky infrared satellite radiance from GOES-R on convection-permitting analysis and prediction of tropical cyclones. *Geophys. Res. Lett.*, **43**, 2954–2963, https://doi.org/10.1002/2016GL068468.

Zhang, Y., F. Zhang, and D. J. Stensrud, 2018: Assimilating all-sky infrared radiances from *GOES-16* ABI using an ensemble Kalman filter for convection-allowing severe thunderstorms prediction. *Mon. Wea. Rev.*, **146**, 3363–3381, https://doi.org/10.1175/MWR-D-18-0062.1.

Zupanski, D., M. Zupanski, L. D. Grasso, R. Brummer, I. Jankov, D. Lindsey, M. Sengupta, and M. Demaria, 2011: Assimilating synthetic GOES-R radiances in cloudy conditions using an ensemble-based method. *Int. J. Remote Sens.*, **32**, 9637–9659, https://doi.org/10.1080/01431161.2011.572094.