Impact of Radar Reflectivity Data Assimilation Frequency on Convection-Allowing Forecasts of Diverse Cases over the Continental United States

Yue Yang[a] and Xuguang Wang[a]

[a] School of Meteorology, University of Oklahoma, Norman, Oklahoma

ABSTRACT: The sensitivity of convection-allowing forecasts over the continental United States to radar reflectivity data assimilation (DA) frequency is explored within the Gridpoint Statistical Interpolation (GSI)-based ensemble–variational (EnVar) system. Experiments with reflectivity DA intervals of 60, 20, and 5 min (RAIN60, RAIN20, and RAIN5, respectively) are conducted using 10 diverse cases. Quantitative verification indicates that the degree of sensitivity depends on storm features during the radar DA period. Five developing storms show high sensitivity, whereas five mature or decaying storms do not. The 20-min interval is the most reliable given its best overall performance compared to the 5- and 60-min intervals. Diagnostics suggest that the differences in analyzed cold pools (ACPs) among RAIN60, RAIN20, and RAIN5 vary by storm features during the radar DA period. Such ACP differences result in different forecast skills. In the case where RAIN20 outperforms RAIN60 and the case where RAIN5 outperforms RAIN20, assimilation of reflectivity with a higher frequency commonly produces enhanced and widespread ACPs, promoting broader storms that match better with reality than a lower frequency. In the case where RAIN5 performs worse than RAIN20, the model imbalance of RAIN5 overwhelms information gain associated with frequent assimilation, producing overestimated and spuriously fast-moving ACPs. In the cases where little sensitivity to the reflectivity DA frequency is found, similar ACPs are produced.

KEYWORDS: Cold pools; Convective storms; Radars/Radar observations; Data assimilation

---

1. Introduction

Past studies have demonstrated that the performance of numerical weather prediction (NWP) can be sensitive to the data assimilation (DA) frequency (e.g., Hu and Xue 2007; Johnson et al. 2015; Johnson and Wang 2017; Pan and Wang 2019; Stratman et al. 2020; Labriola et al. 2021). The DA frequency is proportional to the inverse of the cycling interval. A long DA interval means a low frequency, and a short DA interval means a high frequency. In this study, only observations around the center of the DA window (i.e., analysis time) are assimilated for each cycle. Hence the total number of assimilated observations grows as the DA frequency increases. The use of a long DA interval can save computational cost and allow the model to adjust toward the analysis increments. However, such a long adjustment time may excessively smooth the highly localized increments and therefore may not be appropriate for small-scale (e.g., convective scale) DA (Houtekamer and Zhang 2016). Furthermore, Fertig et al. (2007) attributed the loss of accuracy associated with a long DA interval to the increased model error and the failure of DA in capturing model nonlinearity. Given the deficiencies of using a long DA interval, previous studies have illustrated potential benefits from using more frequent assimilation of conventional data (Gauntlett and Seaman 1974), mesonet and other surface datasets (Knopfmeier and Stensrud 2013), and radar data (Yussouf and Stensrud 2010; Supinie et al. 2017; Davis et al. 2021). Increasing DA frequency is found to promote higher-quality analyses, especially for rapidly changing convective-scale weather (Xue et al. 2006; Houtekamer and

Zhang 2016; Miyoshi et al. 2016). However, DA with a too-short interval not only increases computational cost but also likely degrades the quality of analyses and forecasts (Lange and Craig 2014; Xue and Martin 2006; Liu and Xue 2008). Too frequent interruption of the model may potentially accumulate noises through successive cycles and therefore result in imbalance (e.g., Lynch and Huang 1992). Such substantial model imbalance may offset the benefits from more frequent DA (Houtekamer and Zhang 2016; Brousseau et al. 2008). Therefore, uncertainty remains to determine the optimal DA frequency for NWP.

The uncertainty in the optimal DA frequency applies to the convective-scale radar DA. With high temporal and spatial resolutions, radar data can depict internal structures of convective storms in detail. However, it is challenging to obtain a balance between fully taking advantage of such high-frequency observations and accurately initializing the model (e.g., S. Wang et al. 2013; Johnson and Wang 2017). Earlier studies have provided some general guidance or suggestions on choosing the optimal radar DA frequency. As proposed in Peña and Kalnay (2004), the DA interval should be determined by scales of targeted phenomena. For example, the convective-scale radar observations were assimilated with a shorter interval (5 min) and the synoptic-mesoscale observations with a longer interval (3 h), as in Johnson et al. (2015). Houtekamer and Zhang (2016) preferred to shorten the DA interval when assimilating small-scale features compared to large-scale features. Pan and Wang (2019) suggested that the optimal radar DA interval needs to be set close to the model adjustment time following DA. To examine the sensitivity to radar DA frequency, previous studies usually focused on a specific phenomenon, such as mesoscale convective systems (MCSs; Lange and Craig 2014; Johnson and Wang 2017; Labriola et al. 2021), supercell storms

---

Corresponding author: Xuguang Wang, xuguang.wang@ou.edu

(Zhang et al. 2004; Xue et al. 2006; S. Wang et al. 2013; Lei et al. 2007), tornadic supercells (Hu and Xue 2007; Stratman et al. 2020), hurricanes (Dong and Xue 2013; Davis et al. 2021), and squall lines (Pan and Wang 2019). Configurations of radar DA and the optimal radar DA interval from these studies are summarized in Table 1. It is shown that the optimal radar DA frequency varies depending on the phenomena, types of radar, DA methods, and domain sizes.

It is widely known that the operational convection-allowing prediction system, such as the High-Resolution Rapid Refresh (HRRR; Smith et al. 2008; Benjamin et al. 2016) model and the North American Mesoscale (NAM; Carley et al. 2015; Rogers et al. 2017; Wu et al. 2017; Wang and Wang 2021) forecast system, is hourly updated over the continental United States (CONUS). For example, for the radar reflectivity assimilation, version 4 of HRRR and NAM as of early 2018 use hourly frequency and cloud analysis method (Alexander et al. 2020; Duda et al. 2019; Dowell et al. 2022). As shown in Table 1, however, little attention has been paid to the sensitivity to radar DA frequency over the CONUS that simultaneously houses multiple convection modes and scales. Additionally, earlier work was limited to a single convective event feature with high sensitivity to the radar DA frequency. This study further explores the sensitivity of convective-scale forecasts to the radar reflectivity DA frequency (hereafter RDF) by using 10 diverse cases and by evaluating the CONUS domain. The optimal radar DA frequency could be dependent on DA methods. Different from the earlier work, we adopt the Gridpoint Statistical Interpolation (GSI)-based ensemble–variational (EnVar) DA system (X. Wang et al. 2013) with the capability of directly assimilating reflectivity (Wang and Wang 2017). Besides the quantitative verification, diagnostics are performed to physically understand the impact of RDF on the analyses and subsequent forecasts.

The remaining of this paper is organized as follows. The GSI-based EnVar system, experiment setup, 10 convective cases, and verification methods are introduced in section 2. Section 3 examines the sensitivity of convection-allowing forecasts over the CONUS to RDF using the 10 cases. Based on the classification from section 3, sections 4 and 5 physically understand the various degrees of sensitivity to RDF. Finally, general conclusions together with discussion are given in section 6.

2. Experiment setup

a. GSI-based EnVar system

1) SYSTEM DESCRIPTION

The GSI-based EnVar system and its variants (Wang 2010; X. Wang et al. 2013; Wang and Lei 2014) have been widely implemented in operations and research for NWP. In this system, the forecast ensemble provides flow-dependent background error covariances for the EnVar to update the control forecast, and the ensemble perturbations are updated using the ensemble Kalman filter (EnKF). The hybrid EnVar system also includes static background error covariances.

TABLE 1. Summary of previous studies in the literature on the radar data assimilation (DA) frequency. Abbreviations of DA systems are as follows: EnKF is ensemble Kalman filter, 4DEnSRF/EnSRF is four-/three-dimensional ensemble square root filter, LETKF is local ensemble transform Kalman filter, and 3DVar is three-dimensional variational DA. Abbreviations of radar data are as follows: OSSE is observing system simulation experiment, PAR is phased-array radar, WSR-88D is Weather Surveillance Radar-1988 Doppler, MRMS is the Multi-Radar Multi-Sensor system, and CINRAD-98D is China New Generation 1998 Doppler.

| Radar DA system	Radar data	Radar DA interval	Phenomenon	Radar DA domain	Optimal interval	Literature
EnKF	OSSE	2 and 5 min	A supercell storm	70 km × 70 km	2 min	Zhang et al. (2004)
EnSRF	OSSE	1, 2.5, and 5 min	A supercell storm	78 km × 150 km	1 min	Xue et al. (2006)
EnKF	OSSE	1.25 and 5 min	A supercell storm	64 km × 64 km	1.25 min	Lei et al. (2007)
4DEnSRF vs EnSRF	OSSE	1, 3, 5, 10, and 20 min	A supercell storm	120 km × 120 km	5–10 min, 20 min for real-time application	S. Wang et al. (2013)
LETKF	OSSE	5 and 20 min	A case with convective storms	396 km × 396 km	20 min	Lange and Craig (2014)
4DEnSRF	PAR	1, 3, 5, 15 min, and (1 + 15) min	A tornadic supercell event	400 km × 400 km	More frequent, more accurate	Stratman et al. (2020)
3DVar and cloud analysis system	WSR-88D	5, 10, and 15 min	A tornadic thunderstorm case	The general Oklahoma City area	10 min	Hu and Xue (2007)
EnSRF	WSR-88D	10, 30, and 60 min	A Hurricane Ike (2008) case	Southern United States	10 and 30 min	Dong and Xue (2013)
EnKF	WSR-88D	5, 10, and 15 min	A nocturnal MCS	The Great Plains	10 min	Johnson and Wang (2017)
EnKF	MRMS	5, 10, and 15 min	An MCS	433 × 241 grid points with 3-km grid spacing	10–15 min	Labriola et al. (2021)
3DVar and cloud analysis system	CINRAD-98D	10, 20, 30, and 60 min	A squall-line system	Southern China	20 min	Pan and Wang (2019)
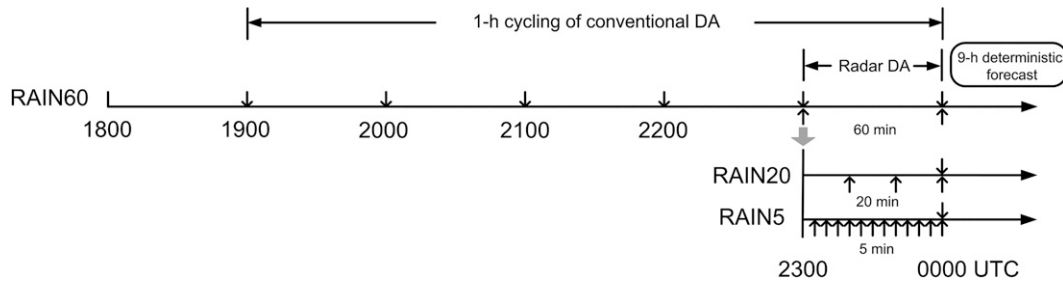
FIG. 1. Flowchart of data assimilation (DA) cycling setup for experiments with various reflectivity DA frequencies. The downward black arrows denote the hourly assimilation of conventional observations from 1900 to 0000 UTC, and the upward black arrows represent the reflectivity DA for the last hour from 2300 to 0000 UTC. Experiments using 60-, 20-, and 5-min reflectivity DA intervals are named RAIN60, RAIN20, and RAIN5, respectively. The thick gray arrow indicates that these experiments share the same analysis of the first reflectivity DA cycle at 2300 UTC. Note that analyses at 2300 and 0000 UTC include assimilation of both conventional and reflectivity observations. The 0–9-h deterministic forecasts are then launched from 0000 UTC.

It is proven that the GSI-based EnVar system is beneficial for global and tropical cyclone forecasts using the National Centers for Environmental Prediction (NCEP) Global Forecast System (GFS) model (Hamill et al. 2011; X. Wang et al. 2013; Wang and Lei 2014). Moreover, many efforts have been devoted to integrating the GSI-based EnVar system with regional models, such as the Advanced Research version of the Weather Research and Forecasting (WRF-ARW) Model (Wang and Wang 2017; Gasperoni et al. 2020; Johnson and Wang 2020; Johnson et al. 2020), the Nonhydrostatic Multiscale Model on the B grid (NMMB) (Carley 2012; Duda et al. 2019; Gasperoni et al. 2020; Johnson et al. 2020; Wang and Wang 2021), and the Hurricane WRF (HWRF) model (Lu et al. 2017), to extend the capability of convective-scale DA.

The DA cycling in this study broadly follows the two-way coupled GSI-based EnVar system developed by X. Wang et al. (2013, their Fig. 1b). The initial conditions and lateral boundary conditions (ICs and LBCs) for an ensemble of 40 members are generated by adding 20 perturbations each from the Global Ensemble Forecast System (GEFS; Zhou et al. 2017) and the Short-Range Ensemble Forecast (SREF) system (Du et al. 2015) to the GFS control analysis. The LBCs for the control member are driven by the GFS control forecasts during the DA period. Upon cycling, the mean of the 40-member prior ensemble is adopted as the first-guess field of the control member, which is then updated using the GSI-based EnVar to generate the control analysis; while the 40-member prior ensemble is analyzed via the GSI-based EnKF and further re-centered using the control analysis as its mean to obtain the final analysis ensemble. Subsequently, the final analysis ensemble is integrated forward to generate the prior ensemble for the next cycle. In addition to the conventional observations (i.e., routine surface observations from METARs, mesonet observations, radiosondes, flight-level aircraft, ships, and ocean buoys), both the EnKF and EnVar are extended to directly assimilate reflectivity observations (Johnson et al. 2015; the third approach in Wang and Wang 2017). The Multi-Radar Multi-Sensor (MRMS) system (https://www.nssl.noaa.gov/projects/mrms/; Smith et al. 2016) provides the quality-controlled fields

of three-dimensional reflectivity, which are created at a 0.01° latitude–longitude grid spacing with irregular 33 vertical levels and are available every 2 min over the CONUS and southern Canada. As shown in Fig. 1, starting from 1800 UTC, conventional data are assimilated hourly for 6 h from 1900 to 0000 UTC, with hourly or subhourly reflectivity DA from 2300 to 0000 UTC. The 1-h reflectivity DA period is designed following the 2017–19 Spring Forecasting Experiments in the Hazardous Weather Testbed by the University of Oklahoma Multiscale data Assimilation and Predictability (MAP) laboratory (e.g., Gasperoni et al. 2020; Johnson et al. 2020).

2) MODEL CONFIGURATION

The WRF-ARW version 3.9 is employed in this study. The CONUS-covering domain centered at (38.5°N, 97.5°W) consists of 1621×1121 grid points with a 3-km convection-allowing grid spacing. A total of 51 full-η levels extend vertically to the model top of 50 hPa. The physics suite mirrors the HRRR configuration (Alexander et al. 2020; Dowell et al. 2022), including the Mellor–Yamada–Nakanishi–Niino (MYNN) level 2.5 planetary boundary layer scheme (Nakanishi and Niino 2006, 2009; Olson et al. 2019), the MYNN surface layer scheme, the aerosol-aware Thompson microphysics scheme (Thompson and Eidhammer 2014), the Rapid Radiative Transfer Model for general circulation models (RRTMG) shortwave and longwave radiation schemes (Iacono et al. 2008), and the RUC land surface model (Benjamin et al. 2004). No cumulus scheme is adopted due to the 3-km grid spacing.

b. Experimental design

Three experiments with reflectivity DA intervals of 60, 20, and 5 min, respectively, during 2300–0000 UTC (Fig. 1) are conducted on each case to test the sensitivity of convection-allowing forecasts to RDF. The reflectivity data below 10 km above ground level (AGL) are assimilated. The choice of a 1-h reflectivity DA interval emulates the HRRR (Alexander et al. 2020; Dowell et al. 2022) and NAM (Rogers et al. 2017) systems. These experiments are named RAIN60, RAIN20, and

TABLE 2. Selected 10 retrospective cases used to test the sensitivity to reflectivity DA frequency (RDF). The case name is set to the date of the forecast. Characteristics of synoptic forcing determined subjectively (weak: mainly triggered by local thermodynamic effects in the absence of strong vertical wind shear and organized synoptic-scale ascent; strong: mainly triggered by synoptic-scale systems with strong vertical wind shear and organized synoptic-scale ascent; moderate: triggered by the mixed effects of local thermodynamics and weak synoptic-scale systems with weak vertical wind shear and/or disorganized synoptic-scale ascent), morphology, and storm stage during the radar DA period for the dominant storm are included. The last column shows the sensitivity of each case to RDF discussed in section 3.

Case	Synoptic forcing	Dominant storm morphology	Dominant storm stage	Sensitivity to RDF
2 May 2018	Moderate	Widely spaced discrete supercells across IA, NE, and KS	Developing	Sensitive case
3 May 2018	Strong	Squall line with trailing stratiform area across NE and KS, numerous discrete supercells across OK and TX, upscale growth into a linear MCS	Developing	Sensitive case
14 May 2018	Weak	A broken band of dryline convection across the eastern TX Panhandle, OK, and southern KS	Developing to mature	Insensitive case
16 May 2018	Weak	Widely scattered storms across parts of the Southern Plains	Mature	Insensitive case
2 May 2019	Weak	Locally linear dryline convection across northwest TX and south-central OK	Developing	Sensitive case
3 May 2019	Weak	An MCS over central TX	Decaying	Insensitive case
6 May 2019	Moderate	Isolated and widely scattered supercells with modest upscale growth across a broad of north–south expanse of High Plains	Developing	Sensitive case
9 May 2019	Strong	A well-developed linear MCS with the bowing apex across southwest AR, northwest LA, and east TX	Mature	Insensitive case
21 May 2019	Strong	Severe thunderstorms across northwest TX and western OK	Mature	Insensitive case
27 May 2019	Moderate	Along the dryline, an intensified squall line over northeast NM into southeast CO, a cluster of severe thunderstorms over northwest KS and southwest NE	Developing	Sensitive case

RAIN5, respectively. Here "RAIN" is short for the radar interval. Specifically, RAIN60 assimilates reflectivity data at 2300 and 0000 UTC; RAIN20 at 2300, 2320, 2340, and 0000 UTC; RAIN5 at 2300, 2305, 2310, …, 2355, and 0000 UTC. At each cycle, the 2-min reflectivity data around the analysis time are ingested. Therefore, increasing the RDF results in more observations being assimilated. The capability of assimilating reflectivity data every 60 min for a 1-h period to capture the primary characteristics of storm systems has been confirmed by Wang and Wang (2021) using 10 cases. After the final reflectivity DA at 0000 UTC, free deterministic forecasts are launched from the final control analyses for 9 h. The hourly outputs are used for evaluation, together with 1-min outputs within the first forecast hour for diagnostics of model balance.

c. Case selection

Ten retrospective cases in May 2018 and 2019 are selected to recognize the sensitivity of convection-allowing forecasts over the CONUS to RDF. To enhance the robustness of systematic evaluation, 10 chosen cases differ in a wide range of features, namely, synoptic forcing, convective morphologies, geographical locations, and storm stages during the 1-h radar DA period for dominant storms (Table 2). Figure 2 shows the observed composite reflectivity from the MRMS system valid at 0000 UTC for each case. Specifically, both well-organized MCSs (e.g., on 3 May 2018 and 21 May 2019) and discrete cellular convection (e.g., on 16 May 2018 and 6 May 2019) are included. Various initiating mechanisms and synoptic forcing are involved as well, such as synoptic stationary fronts with weak large-scale ascent, synoptic cold fronts enhanced by upper-level shortwave troughs, and dryline–stationary-front intersections.

d. Verification methods

The verification focuses on nearly the eastern two-thirds of the CONUS, which covers 25°–50°N latitude and 110°–80°W longitude, due to the low coverage of observations over mountains (e.g., Fierro et al. 2015; Duda et al. 2019). The locations of the convective events within the verification domain for each case are shown in Fig. 2. Products of composite reflectivity (Smith et al. 2016) and 1-h accumulated quantitative precipitation estimation (1-h QPE; Zhang et al. 2016) from the MRMS system are used for assessment. Thresholds of 20, 30, 40, and 50 dBZ for composite reflectivity and 0.254, 2.54, 6.35, and 12.7 mm (0.01, 0.1, 0.25, and 0.5 in., respectively) for 1-h QPE that have been widely used in earlier work (e.g., Duda et al. 2019; Gasperoni et al. 2020; Dunkerley 2021) are evaluated. Using the grid-to-grid approach, traditional metrics (e.g., frequency bias (FBIAS) and equitable threat score) may misrepresent the useful forecast skill owing to much weight given to the smallest scales (Stratman et al. 2013). Gilleland et al. (2009) also treated the traditional verification scores as inappropriate skill measures for forecasts of discontinuous fields from a high-resolution model. Therefore, the neighborhood method, the fractions skill score (FSS; Roberts and Lean 2008) that directly compares the fraction fields between forecasts and observations, is applicable. As the extension of
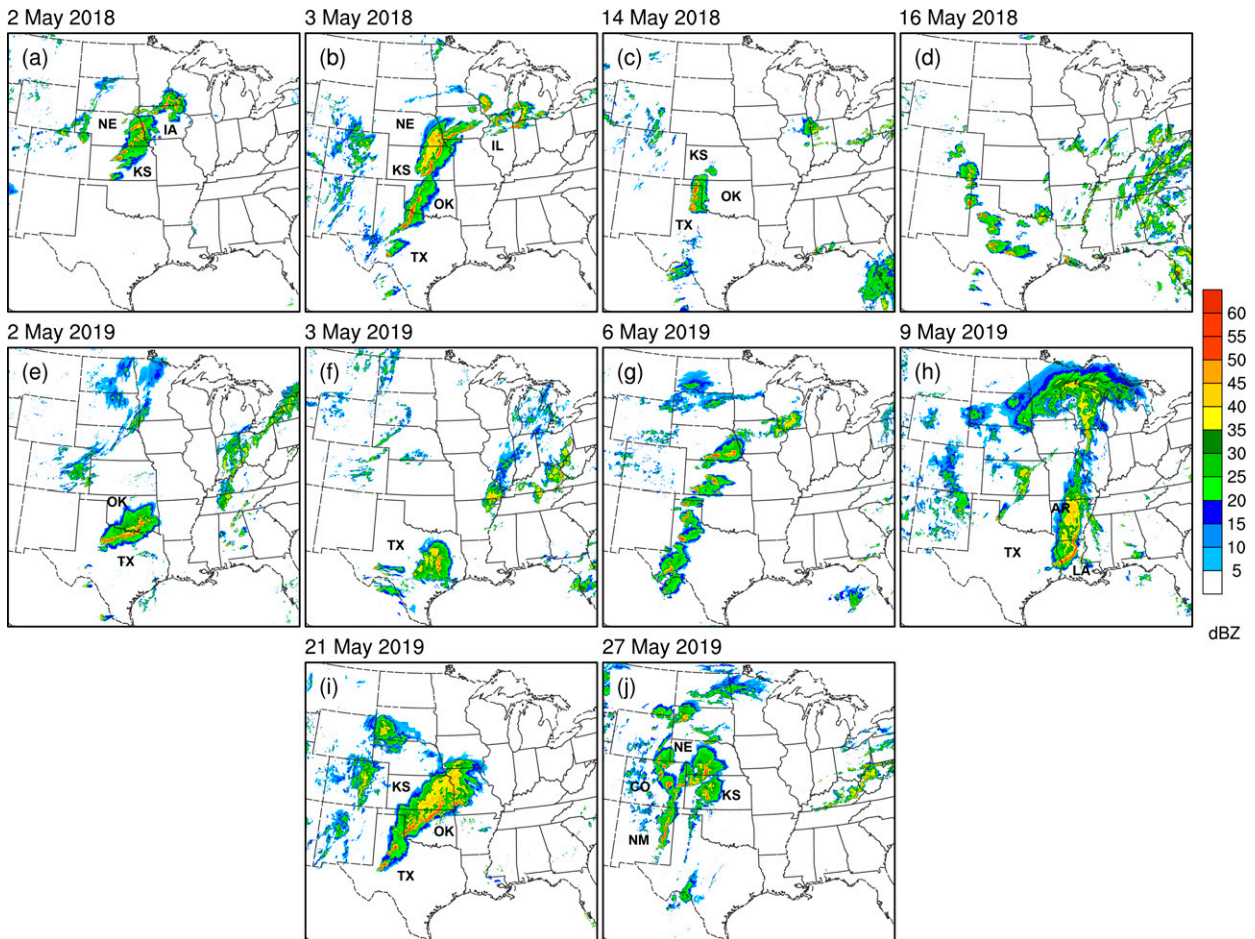
FIG. 2. Observed composite reflectivity from the MRMS system valid at 0000 UTC of the forecast date for 10 cases, which are arranged in chronological order as in Table 2. The corresponding date for each case is indicated in each panel.

the neighborhood-based fractions Brier score (FBS), FSS is given by

$$
\text{FSS} = 1 - \frac{\text{FBS}}{\text{FBS}_{\text{worst}}} = 1 - \frac{\frac{1}{N_v} \sum_{i=1}^{N_v} [\text{NP}_{F(i)} - \text{NP}_{O(i)}]^2}{\frac{1}{N_v} \left[\sum_{i=1}^{N_v} \text{NP}_{F(i)}^2 + \sum_{i=1}^{N_v} \text{NP}_{O(i)}^2 \right]}, \quad (1)
$$

where $\text{NP}_{F(i)}$ and $\text{NP}_{O(i)}$ are the neighborhood forecast probability and observed neighborhood probability, respectively, at the ith grid point in the fraction fields. The total number of grid points is N_v within the verification domain. $\text{FBS}_{\text{worst}}$ is obtained in the situation of no overlap between nonzero fractions. The FSS with a 48-km neighborhood radius (16 grid points; e.g., Johnson and Wang 2012; Duda et al. 2014; Gasperoni et al. 2020) is applied in the present study to verify the closeness between forecasts and observations. The neighborhood radius of 48 km is chosen as a trade-off between alleviating the double-penalty problem (Roberts 2008; Roberts and Lean 2008; Mittermaier and Roberts 2010; Mittermaier et al. 2013) and losing detail at increasing radii (Johnson and

Wang 2012; Gasperoni et al. 2020). The FSS between 0 (no skill) and 1 (a perfect forecast) is positively orientated. Obtaining a high FSS requires a similar agreement between forecasts and observations in the event coverage (Schwartz et al. 2010; Schwartz and Sobash 2017). No bias correction is performed prior to the calculation of FSSs. Therefore, the differences in FSS reflect both the systematic and random error differences.

To further evaluate the impact of RDF on the individual spatial scale, the intensity-scale (IS) verification is performed for biased forecasts of composite reflectivity (Stratman et al. 2013) and 1-h QPE (Casati et al. 2004; Casati 2010). For a square domain of $2^L \times 2^L$ grid points ($L = 10$) that fully covers the verification domain in Fig. 2, the forecast and observation fields are first transformed into binary fields according to a certain threshold of variables. Then the binary difference between the forecast and observation fields is decomposed using the two-dimensional Haar wavelet decomposition at resolutions of $2^{l-1} = 1, 2, 4, \ldots, 2^L$ grid points ($2^{l-1} \times 3$ km) for scale components $l = 1, 2, 3, \ldots, L + 1$, respectively. The resolution of the wavelet-decomposed scale can be treated as half of the wavelength (Johnson et al. 2014).
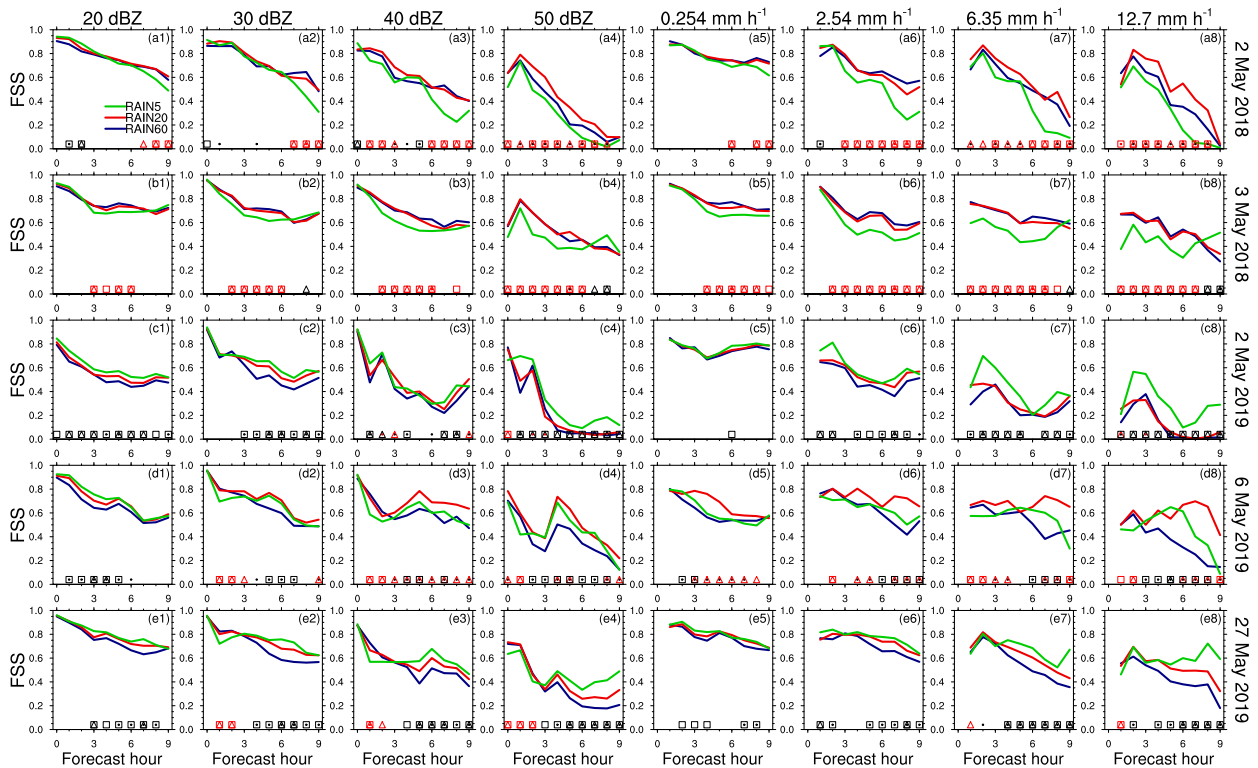
FIG. 3. FSSs of composite reflectivity at 20-, 30-, 40-, and 50-dBZ thresholds and 1-h QPE at 0.254-, 2.54-, 6.35-, and 12.7-mm thresholds for experiments using 60- (blue contours), 20- (red contours), and 5-min (green contours) reflectivity DA intervals from five sensitive cases. A 48-km radius is employed to calculate neighborhood-based fields. FSS differences exceeding 0.04 for RAIN60 vs RAIN20, RAIN20 vs RAIN5, and RAIN60 vs RAIN5 are indicated by the dot, triangle, and square symbols along the x axes, respectively. The black symbols represent that the experiment with a higher DA frequency performs better, and the red symbols represent that the experiment with a lower DA frequency performs better.

Next, the mean square error (MSE) is estimated from the binary field difference for each threshold and each scale component. Finally, based on a reference skill defined by MSE$_{random}$, the IS skill score [SS; Eq. (5) in Casati 2010] is calculated following:

$$SS = 1 - \frac{MSE}{MSE_{random}/(L+1)}, \qquad (2)$$

$$MSE_{random} = FBIAS \times BR \times (1 - BR) + BR$$
$$\times (1 - FBIAS \times BR), \qquad (3)$$

where BR is the base rate. Formed by the spatial scale of the error and the variable intensity, the IS skill score with a positive value delineates a skillful forecast, whereas a negative value indicates no skill.

3. Sensitivity to reflectivity DA frequency

For three RDF experiments, the FSSs of composite reflectivity and 1-h QPE are calculated for 10 cases (Figs. 3 and 4). It is found that the FSSs of both fields share similar trends for each case. However, the performance of the three RDF

experiments is case-dependent. In some cases, three RDF experiments remarkably differ from each other in the forecast skills, while in other cases the RDF has little impact on the FSSs.

To objectively evaluate the sensitivity to RDF for the 10 cases, a criterion is defined to distinguish between sensitive and insensitive cases. For each case, the performance of three RDF experiments is examined using the pairwise comparison. When the FSS difference between one pair of experiments is no less than 0.04 for at least five consecutive forecast hours at two or more thresholds of composite reflectivity and 1-h QPE, this case is classified as a sensitive case. For instance, the 3 May 2018 case (Figs. 3b1–b8) is a sensitive case due to the comparison between RAIN20 and RAIN5 satisfying the predefined criterion, and the 6 May 2019 case (Figs. 3d1–d8) shows the sensitivity as the FSS differences between RAIN60 and RAIN20 match the criterion. Otherwise, it is categorized as an insensitive case, for example, the 3 May 2019 case in Figs. 4c1–c8. Based on the proposed criterion, 10 retrospective cases are classified into two groups. Five sensitive cases in Fig. 3 occurred on 2 and 3 May 2018, and 2, 6, and 27 May 2019, respectively. The remaining five cases in Fig. 4 on 14 and 16 May 2018, and 3, 9, and 21 May 2019, respectively, are insensitive. It is found that shrinking the verification
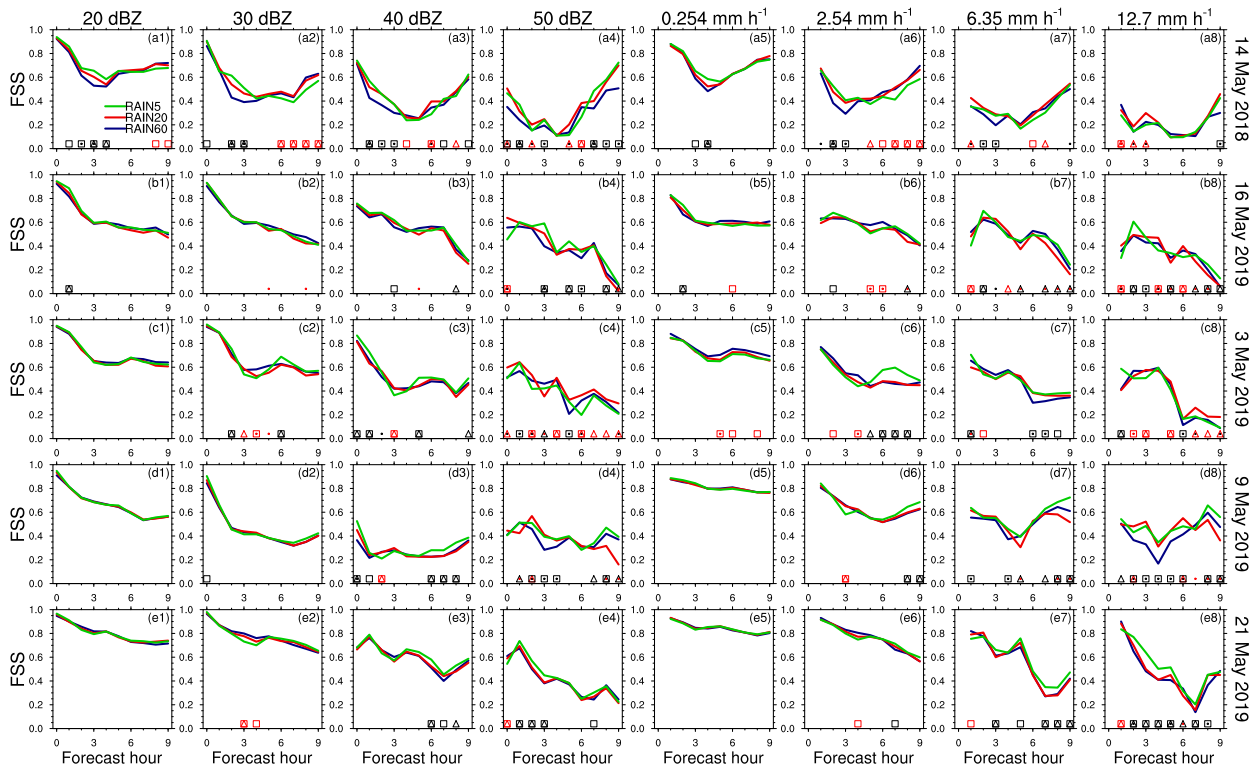
FIG. 4. As in Fig. 3, but for five insensitive cases.

domain to cover the dominant storm has little change on the FSS results in Figs. 3 and 4 (not shown). The sensitivity to RDF over the storm-concentrated domain is also consistent with that over the CONUS domain (not shown). Together with the dominant storm features in Table 2, subjective analysis shows that the cases in each group share a common storm stage during the radar DA period. The sensitive cases have developing storms during the radar DA period, such as linear MCSs with intensifying reflectivity in the 2 May 2019 case and discrete storms with upscale growth in the 6 May 2019 case. In contrast, the insensitive cases feature mature or decaying convections, like a decaying isolated MCS in the 3 May 2019 case and mature linear MCSs in the 21 May 2019 case. Therefore, the sensitivity of CONUS convection-allowing forecasts to RDF heavily relies on the dominant storm life cycle.

Among the five sensitive cases, the performance along with the increase of RDF varies by case based on the predefined criterion. In terms of the pairwise comparison, the general evaluation results associated with the storm features are presented in Table 3, where the frequency gradually increases from a 60-min interval, which is commonly used in operations (Alexander et al. 2020; Rogers et al. 2017; Dowell et al. 2022). When the frequency first increases to a 20-min interval, RAIN20 substantially outperforms RAIN60 (at least five consecutive black dots at two thresholds) for three cases (i.e., on 2 May 2018, 6, and 27 May 2019). The mean FSS differences between RAIN20 and RAIN60 are above 0.052 and 0.088 for composite reflectivity and 1-h QPE, respectively. These cases during

the radar DA period belong to rapidly developing storms, which are usually characterized by fast-moving, upscale-growing, and intensifying features. For the other two cases (i.e., 3 May 2018 and 2 May 2019), RAIN20 achieves comparable performance with RAIN60 at most thresholds. By contrast, these cases during the radar DA period generally develop more slowly in terms of movement and growth in size and intensity. In other words, RAIN20 is always better than or comparable to RAIN60, depending on storm evolutions. When further shortening the interval from 20 to 5 min for the same sensitive cases, RAIN5 has forecast skills superior to RAIN20 (at least five consecutive black triangles at two thresholds) for the two dryline convection (i.e., on 2 and 27 May 2019). The mean FSS differences between RAIN5 and RAIN20 are greater than 0.046 and 0.071 for composite reflectivity and 1-h QPE, respectively. In the remaining three non-dryline cases (i.e., on 2 and 3 May 2018, and 6 May 2019), RAIN20 performs better than RAIN5 (at least five consecutive red triangles at two thresholds) with at least 0.062 and 0.110 higher mean FSSs for composite reflectivity and 1-h QPE, respectively. Therefore, the relative performance between RAIN20 and RAIN5 varies depending on the convection mode.

Overall, the sensitivity of convection-allowing forecasts to the RDF among the three designed frequencies for the 10 cases over the CONUS depends on the dominant storm features during the radar DA period, including life cycles, development characteristics, and convection modes of the storms. It is

TABLE 3. Performance of five sensitive cases in Fig. 3 when gradually increasing the reflectivity DA frequency for pairwise comparisons. The 60-min interval is first shortened to the 20-min interval (RAIN60 vs RAIN20). Compared to the 20-min interval, the DA interval is further shortened to 5 min (RAIN20 vs RAIN5). Here the mean FSS difference over the forecast period between the better experiment and the inferior experiment represents the maximum of the mean among all thresholds for each of composite reflectivity (cref) and 1-h QPE. The cases in bold will be discussed in detail later in sections 4 and 5.

Comparison	Better	Case	Mean FSS difference for cref (1-h QPE)	Storm features during radar DA period
RAIN60 vs RAIN20	RAIN20	2 May 2018	0.063 (0.089)	Fast-moving, upscale-growing, and intensifying storms
		6 May 2019	0.114 (0.221)	
		27 May 2019	0.052 (0.088)	
	Comparable	**3 May 2018**	—	Slow-moving, slow-growing, and mild-intensifying storms
		2 May 2019	—	
RAIN20 vs RAIN5	RAIN20	2 May 2018	0.128 (0.200)	Non-dryline convection
		3 May 2018	0.062 (0.110)	
		6 May 2019	0.079 (0.123)	
	RAIN5	**2 May 2019**	0.086 (0.171)	Dryline convection
		27 May 2019	0.046 (0.071)	

noted, however, that RAIN20 has the least number of instances that perform the worst among the three experiments with varied frequencies for the 10 cases. From a practical forecast perspective, the 20-min reflectivity DA interval, therefore, is a more reliable option than the others for the convective-scale DA over the CONUS, which simultaneously houses diverse convection life cycles, scales, and modes.

4. Physical understanding of high sensitivity to reflectivity DA frequency

Based on the results of pairwise comparisons in Table 3, three representative, sensitive cases are selected as examples to physically understand the three situations with high sensitivity to RDF, respectively, by comparing two RDF experiments. The 6 May 2019 case characterized by fast-moving, upscale-growing, and intensifying storms, and the 3 May 2018 squall-line case that are not triggered by the dryline are used to figure out why RAIN20 outperforms RAIN60 and RAIN5, respectively. The 2 May 2019 dryline convection case is used to explain the better performance of RAIN5 than RAIN20. The impact of RDF on the model balance is first investigated, followed by the diagnostics about the influences of RDF on the analyses and subsequent forecasts.

a. Better performance of RAIN20 than RAIN60

1) INITIAL ADJUSTMENT

In each DA cycle, the model balance is interrupted by the DA procedure and then gradually rebuilt during the model integration (Lynch and Huang 1992; Huang and Lynch 1993; Hu and Xue 2007; Pan and Wang 2019). To measure the model balance, past studies on the convective-scale DA have applied the maximum vertical velocity (Wmax) to reflect the dynamic variable adjustment (Hu and Xue 2007; Pan and Wang 2019). In the present study, the Wmax is calculated over the CONUS every minute within the first forecast hour (Fig. 5). For RAIN60 and RAIN20 from the 6 May 2019 case, the curves of Wmax share similar shapes (Fig. 5a). The Wmax

decreases to reach the minimum below 13 m s^{-1} after initialization, and then it grows rapidly and stabilizes at 24–25 m s^{-1} in ~15 min. Compared to RAIN60, RAIN20 has ~4 m s^{-1} stronger Wmax at the initialization time, and the larger Wmax sustains for 10 min. Both RAIN60 and RAIN20 stabilize after a 15–20-min adjustment, indicating they both reach model balance before the next DA cycle.

2) IMPACT OF DA FREQUENCY ON ANALYSIS

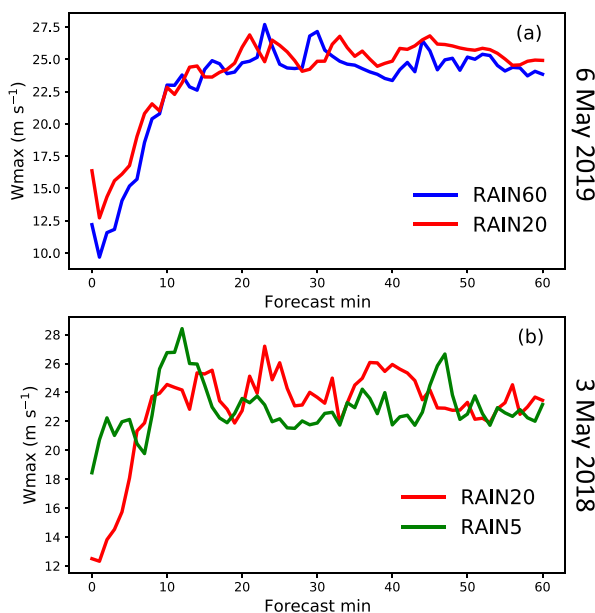Although RAIN20 and RAIN60 share the same first reflectivity DA analysis at 2300 UTC 5 May 2019 (Fig. 1), they

FIG. 5. The 1-h evolution of the maximum vertical velocity (Wmax; m s^{-1}) per min over the CONUS in forecasts starting at (a) 0000 UTC 6 May 2019 from RAIN60 (blue contour) and RAIN20 (red contour) and at (b) 0000 UTC 3 May 2018 from RAIN20 (red contour) and RAIN5 (green contour).
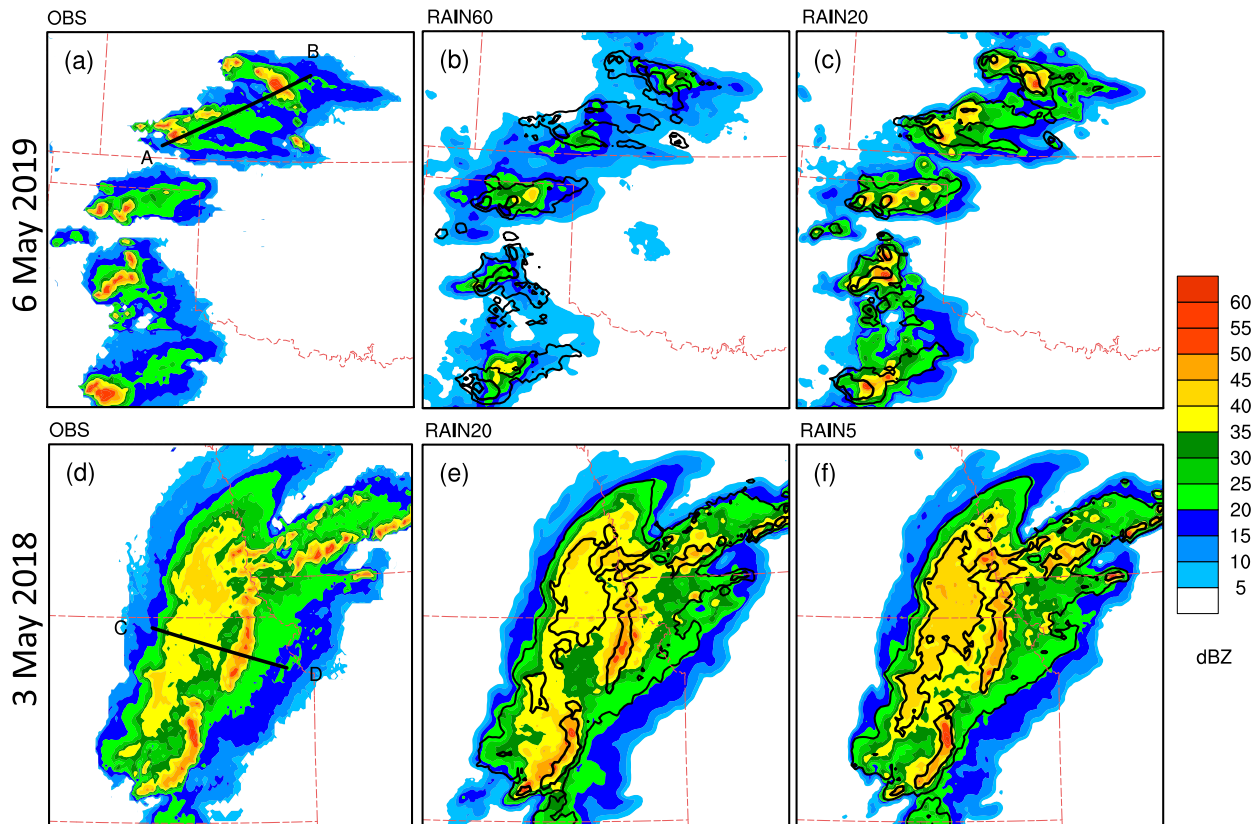
FIG. 6. (a) Observed composite reflectivity (shaded; dBZ) and composite reflectivity first guess (shaded; dBZ) for (b) RAIN60 and (c) RAIN20 valid at 0000 UTC 6 May 2019. In (b) and (c), composite reflectivity analyses of 20 and 40 dBZ from RAIN60 and RAIN20 are overlaid on the corresponding first-guess fields with thin and thick black contours, respectively. (d)–(f) As in the 6 May 2019 case, but for RAIN20 and RAIN5 valid at 0000 UTC 3 May 2018.

have remarkable differences in the composite reflectivity first guess fields at 0000 UTC 6 May (Figs. 6b,c). Due to the 1-h integration in RAIN60, the increments from the 2300 UTC cycle are excessively smoothed, and the ensemble spread becomes large. As a result, the peak intensity of the storm in the control member of RAIN60 is less than 45 dBZ, though some individual ensemble members predict intense storms (not shown). The frequent updates in RAIN20, in contrast, maintain the reflectivity strength with the storm morphology approaching the observations (Fig. 6a). After the final reflectivity DA at 0000 UTC, the differences in composite reflectivity between RAIN60 and RAIN20 become much smaller than that of their first-guess fields (Figs. 6b,c).

Cross sections below 15 km AGL along the line A–B in Fig. 6a are plotted to compare the differences between RAIN60 and RAIN20 in storm vertical structures at 0000 UTC. Figure 7a shows that three reflectivity patches labeled by P1, P2, and P3 are observed. The reflectivity observations are collected above 500 m AGL for the MRMS system, and no radar coverage exists below 2 km AGL for many areas over the CONUS (Maddox et al. 2002). Compared to RAIN60, the first guess of RAIN20 better captures the observed reflectivity pattern (Figs. 7b,c). The longer first-guess forecast in RAIN60 leads to at least 15 dBZ smaller reflectivity than RAIN20. Additionally, the updraft first guess from

RAIN20 is larger than RAIN60, and the maximum difference is 4 m s^{-1} in P3. Such stronger updrafts in RAIN20 coincide well with the larger reflectivity. Likewise, the cross correlation between reflectivity and strong updraft is better estimated by the prior ensemble of RAIN20 compared to RAIN60 (not shown). Therefore, while the differences in the final reflectivity analyses between RAIN60 and RAIN20 are smaller compared to that of their first guesses, the large differences in the updrafts remain after the final DA. RAIN60 still has 3–4 m s^{-1} weaker upper-level updrafts than RAIN20 in P1 and P3 (Figs. 7d,e). The lower top heights of the analyzed storms in RAIN60 and RAIN20 than the observed storm can be largely attributed to the model (i.e., dynamic core, physics, etc.) because the first guess storms extend below 12 km AGL lower than the observations.

The different vertical motions in the analyses of RAIN60 and RAIN20 cause analyzed cold pools (ACPs) to vary in intensity and coverage. Compared to RAIN60, RAIN20 produces 1–2 K colder and broader ACPs in the vicinities of P1 and P3 (Figs. 8a,b). Verified against the observed 2-m temperature (T2) from METARs, the cold pool indicated by the green rectangle from RAIN20 is more consistent with reality, and the cold pool from RAIN60 is weaker and narrower. The trend that the cold pool gets broader and intensified with the increase in RDF is similar to Pan and Wang (2019) and is
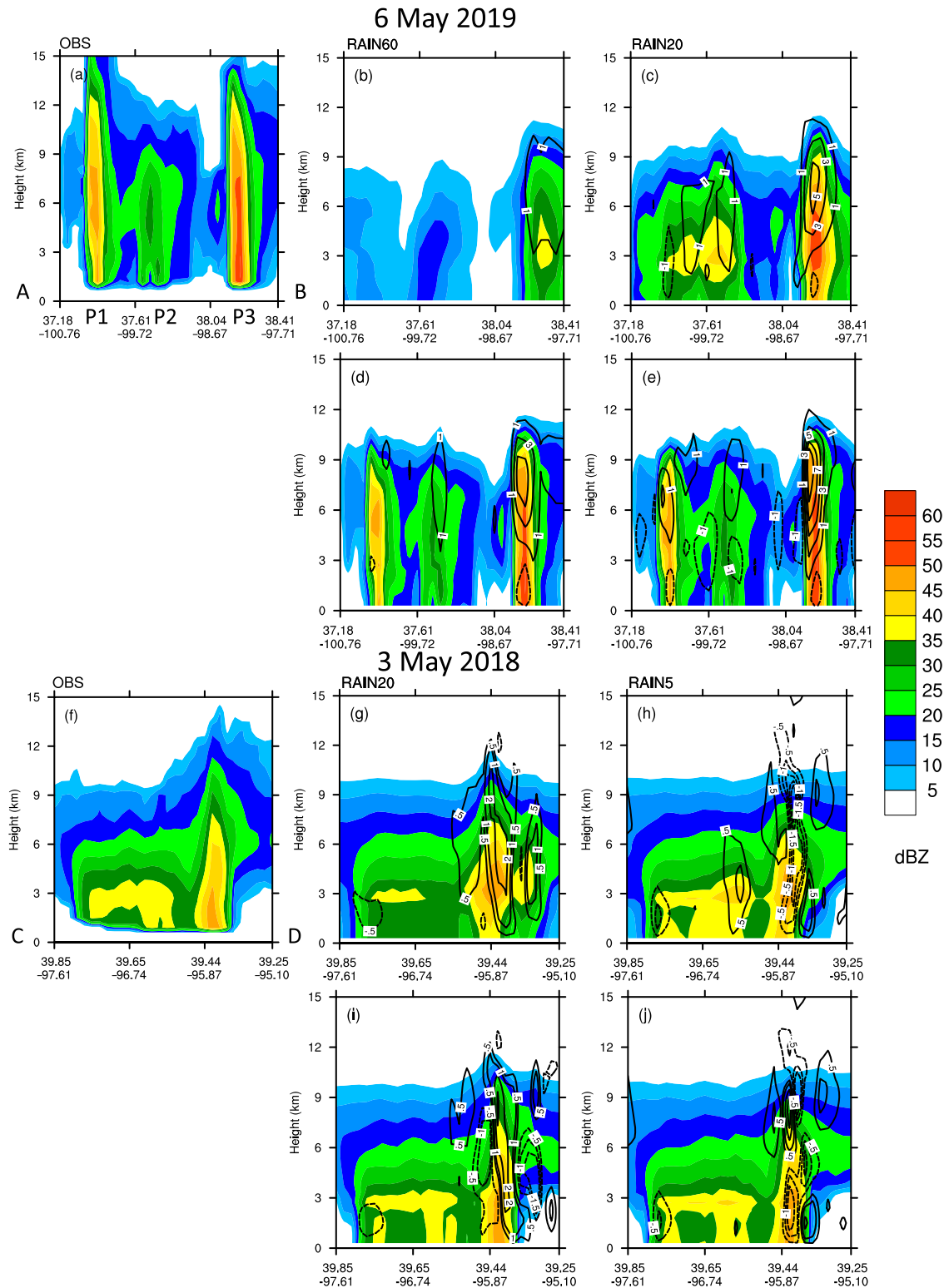
FIG. 7. Vertical cross sections of (a) observed reflectivity (shaded; dBZ) and reflectivity (b),(c) first guess and (d),(e) analysis (shaded; dBZ) for RAIN60 in (b) and (d) and RAIN20 in (c) and (e) along the line A–B in Fig. 6a valid at 0000 UTC 6 May 2019. Black contours indicate vertical velocity (dashed line: downdraft; solid line: updraft) in a 2 m s^{-1} interval. (f)–(j) As in (a)–(e), but for RAIN20 and RAIN5 along the line C–D in Fig. 6d valid at 0000 UTC 3 May 2018. The vertical velocity is in a 0.5 m s^{-1} interval below 1 m s^{-1} and a 1 m s^{-1} interval above 1 m s^{-1}.
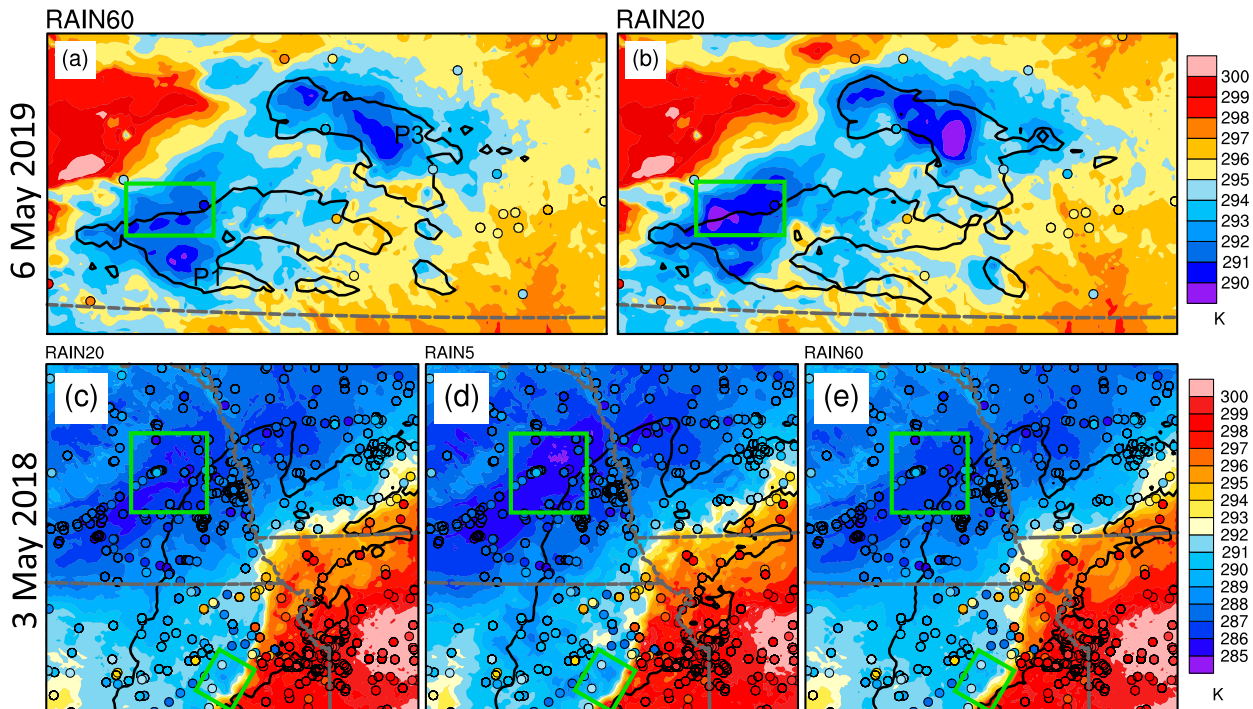
FIG. 8. The 2-m temperature analysis (shaded; K) for (a) RAIN60 and (b) RAIN20 valid at 0000 UTC 6 May 2019 and for (c) RAIN20, (d) RAIN5, and (e) RAIN60 valid at 0000 UTC 3 May 2018. Observations of 2-m temperature are overlaid using shaded circles. The rectangles outlined by the solid green lines denote the regions with remarkable differences between experiments that can be evaluated using surface observations. The black contours delineate analyzed composite reflectivity at 20 dBZ. P1 and P3 in (a) roughly mark the locations of P1 and P3 in Fig. 7a.

mainly attributed to two factors. First, temperature shows a negative correlation with reflectivity at low levels (Dowell et al. 2011; Stratman et al. 2020). Relative to RAIN60, the larger reflectivity via more frequent reflectivity DA in RAIN20 facilitates the lower low-level temperature, such as in P1 (Figs. 7d,e). Second, Dowell et al. (2011) emphasized that low-level cooling is related to the evaporation of rain. When more moisture is elevated through stronger updrafts in RAIN20, more precipitation loading and evaporation cooling are generated than in RAIN60, such as in P1 and P3 (Figs. 7d,e).

3) IMPACT OF DA FREQUENCY ON FORECASTS

Dowell et al. (2004) summarized that low-level mesocyclogenesis in supercells, convective initiations, and the development of long-lived convective systems heavily depend on cold pool features. Thus, the differently analyzed cold pools in RAIN60 and RAIN20 lead to various storm predictions. During the forecast phase, the 2-m potential temperature perturbation (θ') is used to define the coverage and intensity of the cold pool (e.g., Tompkins 2001). Taking an MCS that grew upscale from isolated storms as an example, Figs. 9a–d depict the predicted θ' and 20-dBZ composite reflectivity overlaid with observed 20-dBZ composite reflectivity. Throughout the forecast period, the storms from RAIN60 and RAIN20 propagate slowly in a west–east direction as in the observations. At 0300 UTC, the ~4 K colder θ' in RAIN20 corresponds to

~4 m s^{-1} stronger cold pool outflow than RAIN60 across Oklahoma (OK) and Kansas (KS) (Figs. 9a,b; see state locations in Fig. 2). In RAIN20, the enhanced low-level divergence maintains and provokes the convection owing to the convergence resulting from the cold pool outflow and the ambient southerly flow (Fig. 9b). Compared to RAIN60, RAIN20 predicts discrete storms with larger coverage, especially for the storm across OK and KS. Subsequently, the developing storms in RAIN20 further enlarge and enhance the cold pools. The strengthening cold pools serve as continued initiation zones for additional convection, promoting the MCS formation at 0500 UTC (Fig. 9d). However, weaker cold pools from RAIN60 gradually scatter and shrink the storms (Fig. 9c). As a result, the larger predicted MCS in RAIN20 agrees better with the observations than RAIN60, which is aligned with the higher FSSs for nearly all lead times and thresholds considered (Figs. 3d1–d4).

The IS skill score of 1-h QPE is implemented to examine the spatial scales of the error where RAIN20 outperforms RAIN60 during the forecast period. The 5-h forecasts of RAIN60 and RAIN20 are decomposed as an example. For RAIN20, the scales above 384 km have positive scores nearly approaching 1.0, and the little to no skill exists on scales less than 48 km for intense precipitation events (i.e., 12.7 mm h^{-1}) (Fig. 10a). Scales larger than 96 km exhibit higher positive scores due to well-predicted large-scale events (e.g., fronts), and scales below 48 km are less skillful because of the lower predictability for small-scale events (e.g., convective showers), especially
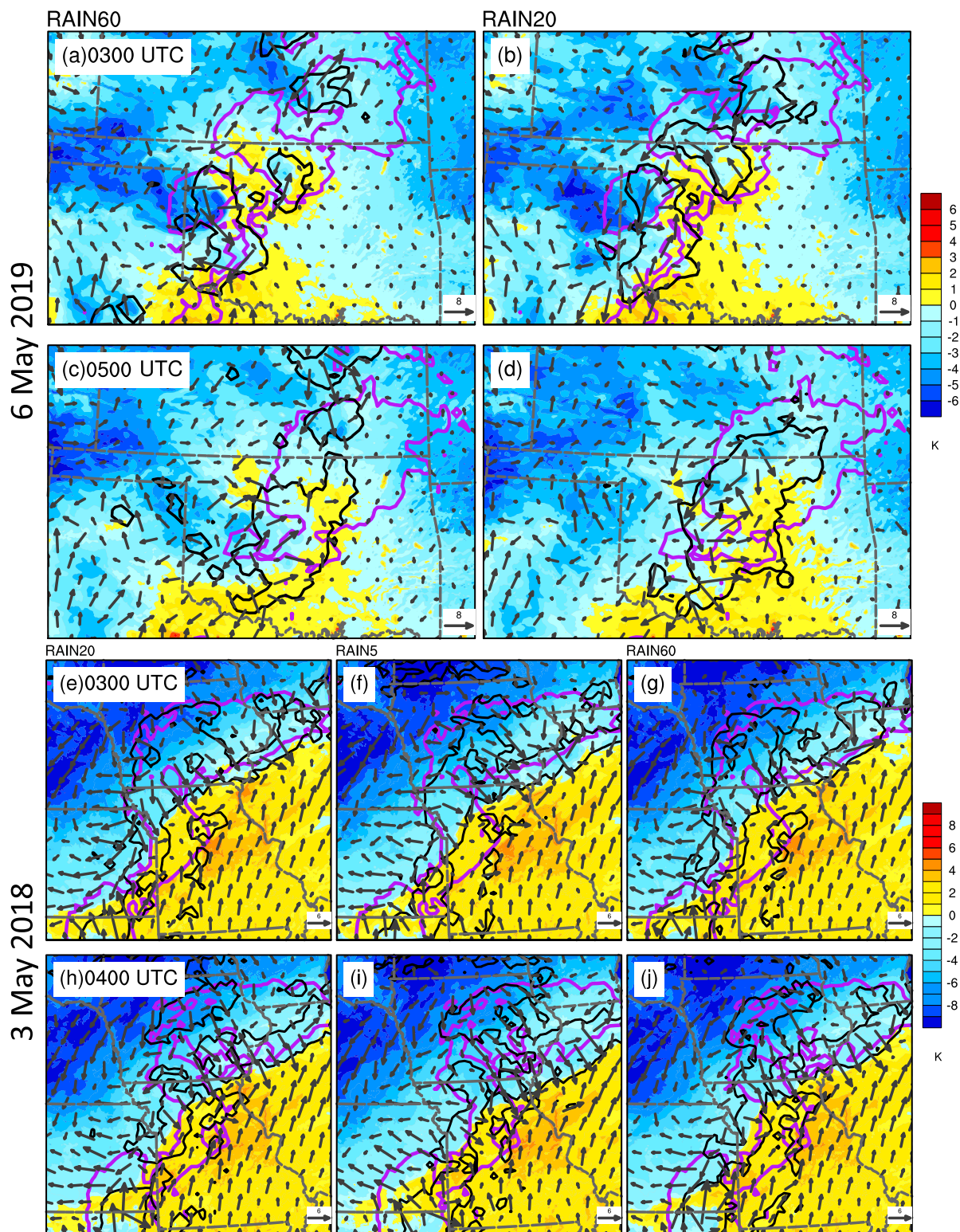
FIG. 9. Forecasts of 2-m potential temperature perturbation (shaded; K), 20-dBZ composite reflectivity (black contours), and 10-m wind vector perturbations (gray arrows; m s^{-1}) overlaid with observed composite reflectivity at 20 dBZ (purple contours) for (a),(c) RAIN60 and (b),(d) RAIN20 valid at 0300 UTC 6 May 2019 in (a) and (b) and 0500 UTC 6 May 2019 in (c) and (d). (e)–(j) As in (a)–(d), but for RAIN20 in (e) and (h), RAIN5 in (f) and (i), and RAIN60 in (g) and (j), valid at 0300 UTC 3 May 2018 in (e)–(g) and 0400 UTC 3 May 2018 in (h)–(j).
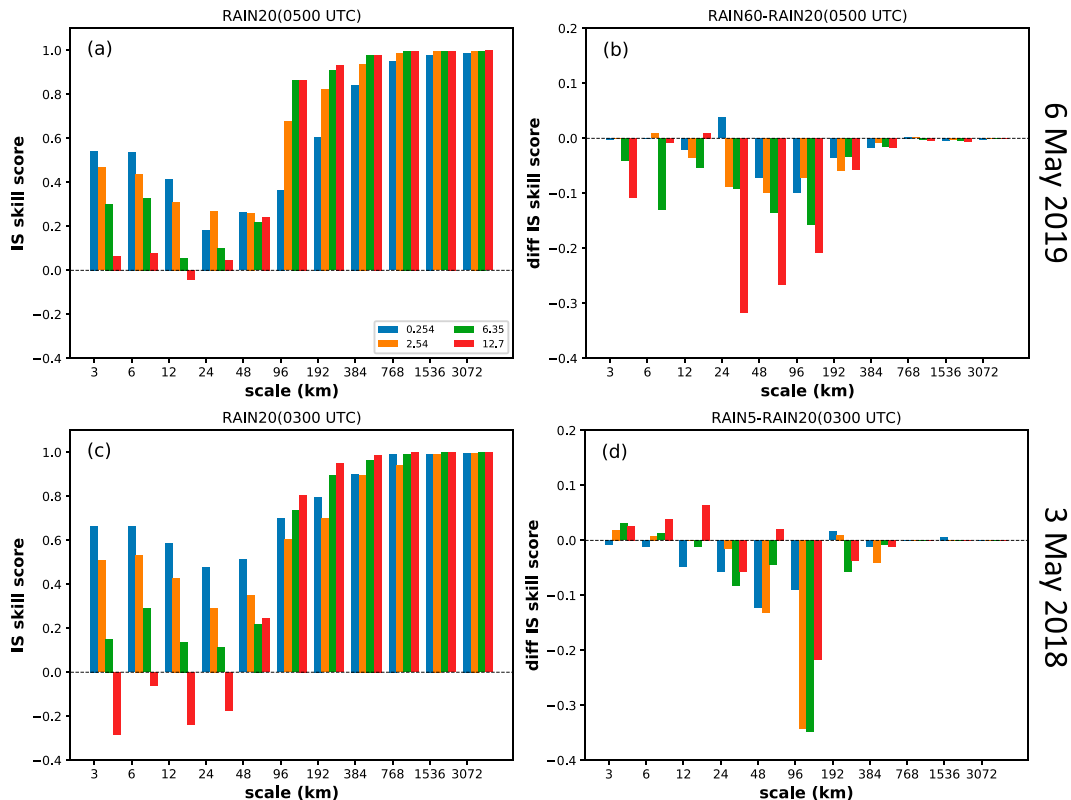
FIG. 10. (a) Intensity-scale (IS) skill score from RAIN20 and (b) differences of IS skill score between RAIN60 and
RAIN20 for 1-h accumulated quantitative precipitation estimation (1-h QPE) at the thresholds of 0.254, 2.54, 6.35,
and 12.7 mm valid at 0500 UTC 6 May 2019. (c),(d) As in (a) and (b), but for RAIN20 and differences between
RAIN5 and RAIN20, respectively, valid at 0300 UTC 3 May 2018.

with heavy rain (>12.7 mm h^{-1}). The slightly large positive
scores for small scales and low thresholds can be attributed to
the fact that few such events cause smaller errors compared to
the random forecast equipartitioned across all scales [Eq. (2);
Casati 2010; Stratman et al. 2013]. Relative to RAIN20, the
inferior skills of RAIN60 concentrate on the spatial scales of
the error between 24 and 96 km (Fig. 10b), which are the
dominant scales of predicted storms. In addition to examining
one forecast lead time, Fig. 11 averages the IS skill score of 1-
h QPE over 9 forecast hours. On average, RAIN20 has
higher forecast skills than RAIN60 across almost all thresh-
olds and scales except for a few small-scale low-intensity
events, i.e., the 2.54 mm h^{-1} threshold at 6 km (Fig. 11b).
For thresholds less than 12.7 mm h^{-1}, the minimum skill
usually occurs at medium scales of 12–48 km (Figs. 11a–c).
At the 12.7-mm h^{-1} threshold, the little to no skill shifts to-
ward the smallest scale (Fig. 11d). These results are consis-
tent with the expectation that intense small-scale events
have low predictability.

b. Better performance of RAIN20 than RAIN5

1) INITIAL ADJUSTMENT

For the 3 May 2018 squall-line case, RAIN5 has a distinctly
different Wmax curve from RAIN20 (Fig. 5b). During the first
15 min, RAIN5 has a dramatic oscillation, whereas RAIN20
experiences rapid growth and then stabilizes. In RAIN5, the
more frequent assimilation results in 6 m s^{-1} larger initialized
Wmax than RAIN20. RAIN5 needs at least 15 min to build a
stable Wmax around 23 m s^{-1}, and RAIN20 reaches a stable
Wmax around 10 min. Therefore, 20 min is sufficient for
RAIN20 to attain a new balance before the next reflectivity
DA. In contrast, 5 min is too short for RAIN5 to reach its bal-
ance before the next DA cycle. As a result, RAIN5 may accu-
mulate model imbalance through successive cycles.

2) IMPACT OF DA FREQUENCY ON ANALYSIS

Initialized from the same analysis valid at 2300 UTC 2 May
2018, both composite reflectivity first guess fields at 0000 UTC
3 May from RAIN20 and RAIN5 capture the structure of the
observed squall line, containing convective lines and strati-
form zone (Figs. 6d–f). However, the reflectivity intensity
within the stratiform region differs more remarkably between
RAIN20 and RAIN5. Compared to RAIN20, RAIN5 has up
to 10 dBZ larger composite reflectivity within the stratiform
region and a larger area covered by 40 dBZ. After the final re-
flectivity DA at 0000 UTC, the 40-dBZ coverage in RAIN5 is
overestimated, and the stratiform region in RAIN20 fits the
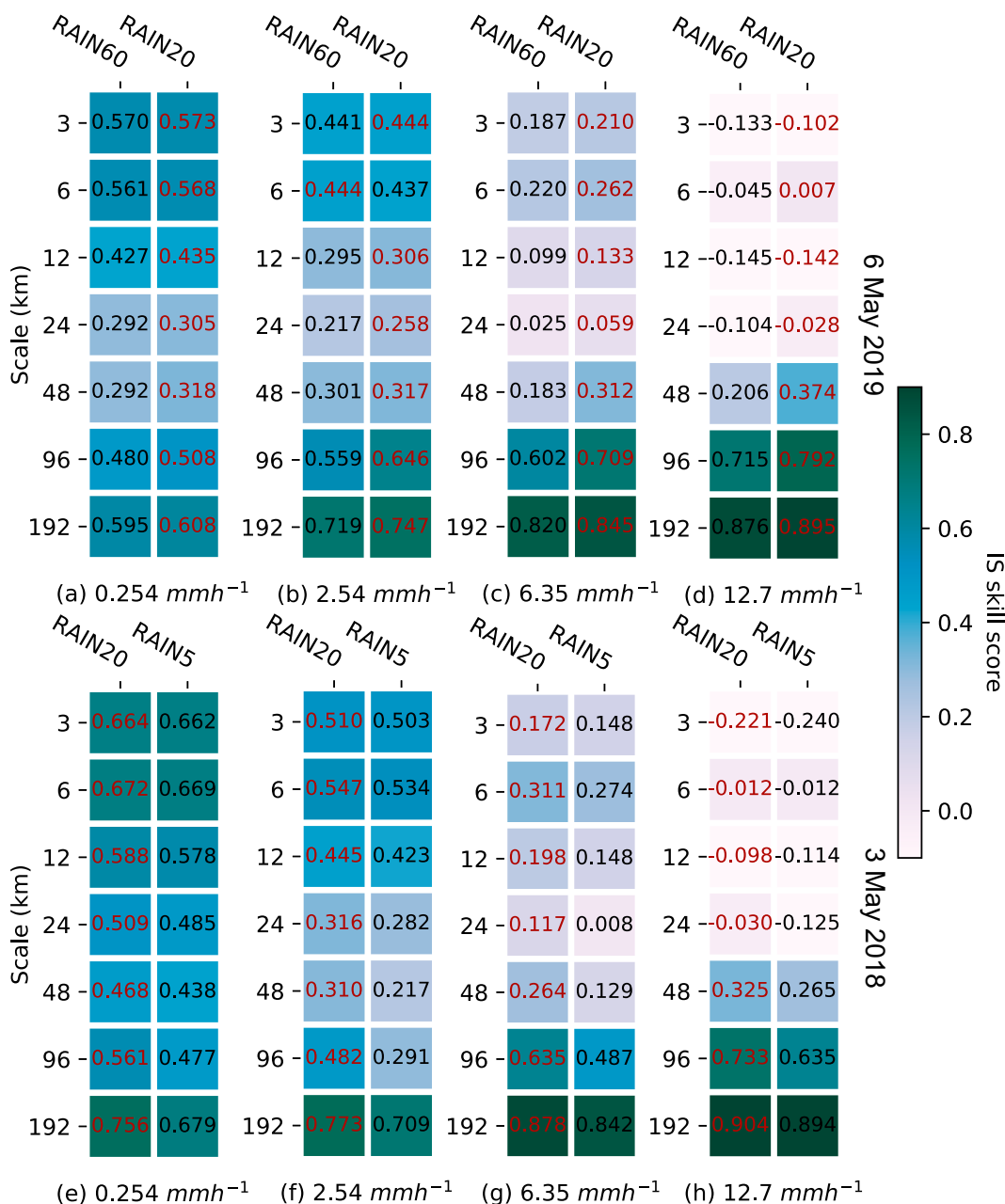observations more closely (Figs. 6d–f).

FIG. 11. IS skill score for precipitation at thresholds of (a) 0.254, (b) 2.54, (c) 6.35, and (d) 12.7 mm h^{-1} averaged over forecast hour 1–9 below the 192-km spatial scale of the error for RAIN60 and RAIN20 from the 6 May 2019 case. Under the same threshold and scale, the number in red (black) means the larger (smaller) score between RAIN60 and RAIN20. (e)–(h) As in (a)–(d), but for RAIN20 and RAIN5 from the 3 May 2018 case. Under the same threshold and scale, the number in red (black) means the larger (smaller) score between RAIN20 and RAIN5.

As shown in the observed cross section along the line C–D in Fig. 6d, the leading convective tower with intense radar echoes (convective region) is followed by a trailing stratiform region, separated by the transition region with lesser reflectivity (Fig. 7f). Such a reflectivity pattern is captured by the first guesses of RAIN20 and RAIN5 at 0000 UTC, but RAIN5 has a more intense stratiform region than RAIN20 by 5–10 dBZ

(Figs. 7g,h). Within the stratiform region, the reflectivity values greater than 40 dBZ in RAIN5 mostly concentrate near the melting level (~3 km AGL). In addition, the first guesses of RAIN20 and RAIN5 have quite different vertical motions along the cross section. The strong convective updraft core in RAIN20 is nearly collocated with the reflectivity peak, while the leading reflectivity peak in RAIN5 corresponds to the
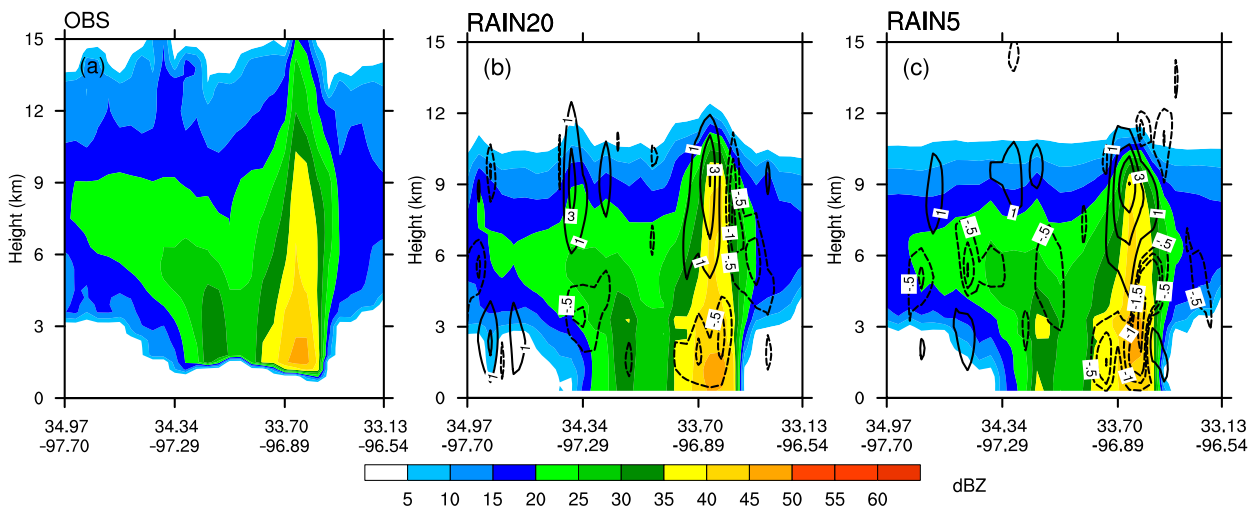
FIG. 12. Vertical cross sections of (a) observed reflectivity (shaded; dBZ) and (b),(c) reflectivity analysis (shaded; dBZ) for RAIN20 in (b) and RAIN5 in (c) valid at 0000 UTC 2 May 2019. Black contours indicate vertical velocity (m s^{-1}). Dashed contours (−0.5, −1, −1.5, and −3 m s^{-1}) represent downdrafts, and solid contours from 1 to 9 m s^{-1} at a 2 m s^{-1} interval represent updrafts.

downdrafts throughout all vertical levels under 13 km. Such an inconsistent storm structure in the first guess of RAIN5 could be attributed to the model imbalance introduced through too frequent reflectivity DA (Fig. 5b). Specifically, the updraft analysis at 2355 UTC may not be strong enough to support the increased hydrometeors from the preceding reflectivity DA cycles due to the imbalance. Then the excessive hydrometeors begin to descend in the subsequent 5-min forecast, leading to downdrafts within the first-guess convective region at 0000 UTC in RAIN5 (Fig. 7h). At the final analysis time, the analyzed reflectivity pattern of RAIN20 is comparable to RAIN5 except for the higher reflectivity near 3 km AGL in RAIN5 within the stratiform region (Figs. 7i,j). Moreover, large differences in vertical motions remain between RAIN20 and RAIN5. For the reflectivity core in the convective region below 4.5 km, RAIN20 has updrafts exceeding 2 m s^{-1}, while RAIN5 maintains downdrafts larger than 1.5 m s^{-1}. As a result, RAIN5 generates broader cold pools than RAIN20 with 1–2 K colder T2 at the final analysis time (Figs. 8c,d). The 3 May 2018 case also suggests that using a higher RDF produces stronger and broader cold pools than using a lower frequency. Based on the verification against the surface observations from METARs and mesonet, the lower T2 in RAIN5 spuriously extends toward the south to southeast, especially for the northern region indicated by the green rectangle. The T2 in RAIN20 is closer to the verifying observations, which indicates the more realistic analyzed vertical motion.

3) IMPACT OF DA FREQUENCY ON FORECASTS

At forecast hour 3, the broader cold pools with 2–3 K colder θ′ in RAIN5 correspond to the ~2 m s^{-1} larger outflow relative to RAIN20, resulting in more enhanced convergence at the boundaries between cold pools outflow and ambient southerly flow (Figs. 9e,f). In addition, the

southeastward surging cold pools from RAIN5 propagate faster than RAIN20, as do the predicted storms. Compared to the observations, RAIN20 performs better than RAIN5 in capturing the storm location. The poorer skills of RAIN5 for composite reflectivity (Figs. 3b1–b4) can therefore be attributed to larger displacement errors, e.g., in northwestern Illinois (IL) and northeastern OK (Fig. 9f; see state locations in Fig. 2). The developing storms in RAIN5 continue to enlarge the coverage of the cold pools, leading to increased storm displacements at 0400 UTC (Fig. 9i). In contrast, the predicted storms in RAIN20 match the observed storm locations better (Fig. 9h). Similar effects of the stronger cold pool on the storm evolution are discussed by earlier studies (e.g., Stratman et al. 2020).

Finally, the IS skill score of 1-h QPE for the 3-h forecast is leveraged to demonstrate the spatial scales of the error where RAIN20 outperforms RAIN5 (Figs. 10c,d). The IS skill score of RAIN20 from the 3 May 2018 case exhibits a similar pattern to that from the 6 May 2019 case (Figs. 10a,c). Likewise, the model can well predict large-scale events but has low predictability for intense small-scale events. As a result, the predictability of precipitation events at the 12.7 mm h^{-1} threshold with errors on scales less than 48 km is usually low, and the scales larger than 384 km have scores over 0.9 across all thresholds (Fig. 10c). The dominant spatial scale where RAIN20 outperforms RAIN5 is ~96 km (Fig. 10d), matching the scale of this squall-line case. The time-averaged IS skill score confirms that RAIN20 has consistently higher skills than RAIN5 across all thresholds and scales (Figs. 11e–h).

c. Better performance of RAIN5 than RAIN20

Following the procedure in section 4b, the better performance of RAIN5 than RAIN20 in the 2 May 2019 dryline convection case is briefly investigated. The model balance of RAIN20 and RAIN5 measured by the time series of Wmax is similar to that from the 3 May 2018 squall-line case (Fig. 5b).
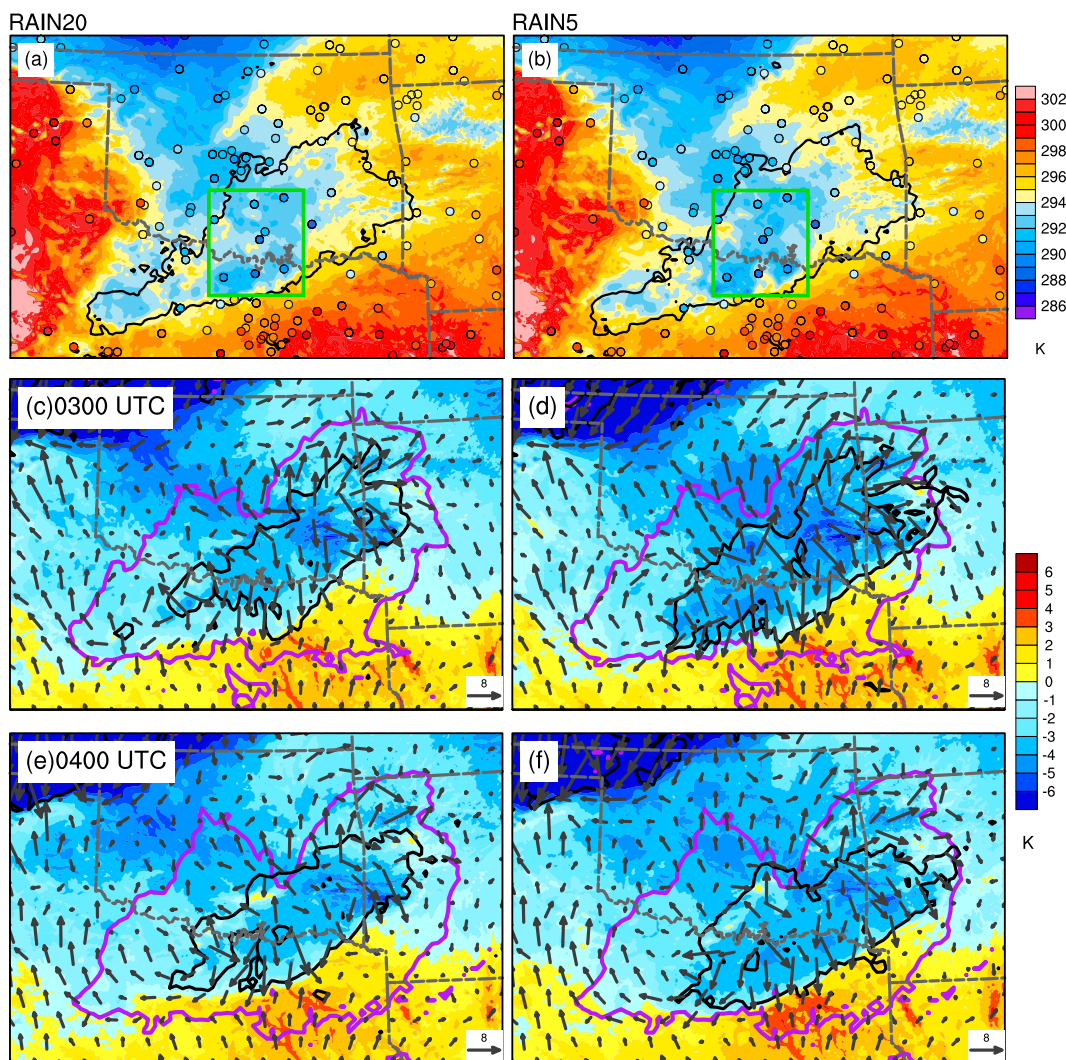
FIG. 13. The 2-m temperature analysis (shaded; K) for (a) RAIN20 and (b) RAIN5 valid at 0000 UTC 2 May 2019. Observations of 2-m temperature are overlaid using shaded circles. The rectangles outlined by the solid green lines denote the region with remarkable differences between RAIN20 and RAIN5 that can be evaluated using surface observations. The black contours delineate analyzed composite reflectivity at 20 dBZ. Forecasts of 2-m potential temperature perturbation (shaded; K), 20-dBZ composite reflectivity (black contours), and 10-m wind vector perturbations (gray arrows; m s^{-1}) overlaid with observed composite reflectivity at 20 dBZ (purple contours) for (c),(e) RAIN20 and (d),(f) RAIN5 valid at 0300 UTC 2 May 2019 in (c) and (d) and 0400 UTC 2 May 2019 in (e) and (f).

Thus, RAIN5 in the 2 May 2019 case is unable to recover its dynamically consistent state within 5 min. At the final analysis time, RAIN20 and RAIN5 in the 2 May 2019 case also produce similar reflectivity patterns along the vertical cross section that match well with the observations (Fig. 12). However, unlike the 3 May 2018 case (Figs. 7i,j), the analyses of RAIN20 and RAIN5 in the 2 May 2019 case have more similar vertical motion patterns along the vertical cross section. In RAIN5, the updrafts between 6 and 12 km AGL and the downdrafts below 6 km AGL are 2 m s^{-1} larger than in RAIN20. In other words, the model imbalance of RAIN5 in this dryline convection case does not markedly change the coherent relationship between reflectivity and updraft.

Compared to RAIN20, the stronger vertical motions and more moisture aloft in RAIN5 (Fig. 12) are beneficial for the creation of stronger cool pools at the final analysis time (green rectangles in Figs. 13a,b). The trend that the ACPs get stronger and broader as the RDF increases agrees with the facts found in sections 4a and 4b. The T2 analysis within the storm from RAIN5 fits better with the surface observations from METARs than RAIN20. In RAIN5, the cold pool outflow is ~3 m s^{-1} larger than RAIN20 at forecast hour 3 (Figs. 13c,d). Compared to RAIN20, the enhanced convergence in RAIN5 maintains and triggers a storm with a larger areal coverage, which is more consistent with the observed storm. As the forecast goes on, the developing storm in RAIN5 rapidly

expands the cold pool, which, in turn, further enlarges the predicted storm northwestward and southeastward, resulting in the larger coverage overlapping the observations better than RAIN20 at 0400 UTC (Figs. 13e,f). Therefore, RAIN5 has a consistently better performance than RAIN20 and offers consistency in higher FSSs at almost all lead times except the 40-dBZ threshold (Figs. 3c1–c4).

It is noted that the negative influence of model imbalance in RAIN5 does not offset the benefits from frequent reflectivity DA for this dryline convection case, while the model imbalance degrades the performance of RAIN5 for the squall-line case in section 4b. In the dryline convection cases (Table 3), the drier air behind the dryline lifts the moist air ahead of it, triggering the storm formation. In other words, the buoyancy affected by moisture gradients dominantly produces the updraft (i.e., buoyancy forcing), and the hydrometeors are directly related to the updraft. In comparison, the updraft arising from the vertical wind shear (i.e., dynamic forcing) is mainly responsible for the organization of storms that are not initiated by the dryline (Table 3). Therefore, it is hypothesized that the various forcings of updrafts (Markowski and Richardson 2010; Marion and Trapp 2018) may determine the different degrees of negative effects from model imbalance. The model imbalance probably has fewer (more) negative effects on the relationship between reflectivity and vertical motion when the updraft is induced by the buoyancy (dynamic) forcing. Finally, the benefits from frequent reflectivity DA overcome (succumb to) the disadvantages of the model imbalance for dryline (non-dryline) convection cases, resulting in RAIN5 being superior (inferior) to RAIN20. Further investigation associated with the various impacts of model imbalance depending on convective-scale features is warranted using more cases. In addition, the imbalance between thermal and dynamic fields inside the non-dryline storms may be attributed to the insufficiently adjusted kinematic fields. Further assimilation of radial velocity may help to alleviate the issue and will be conducted in the future.

5. Reasons for low sensitivity to reflectivity DA frequency

While section 4 reveals the physical mechanism for the high sensitivity of convection-allowing DA and forecasts over the CONUS to RDF, such sensitivity may not apply to all comparisons for all cases as discussed in section 3, e.g., the comparable situation between RAIN60 and RAIN20 in two sensitive cases (Table 3) and three RDF experiments in five insensitive cases (Fig. 4). The 3 May 2018 case featuring slow-moving, slow-growing, and mild-intensifying storms has comparable performance between RAIN60 and RAIN20 (Figs. 3b1–b8). The 3 May 2019 decaying MCS case gets mixed performance among RAIN60, RAIN20, and RAIN5 (Figs. 4c1–c8). These two cases are investigated to detail the reasons for the low sensitivity to RDF.

a. Comparable performance between RAIN60 and RAIN20

For the 3 May 2018 case, the comparison between RAIN20 and RAIN60 confirms that the ACPs get intensified and

extended as the RDF increases (Figs. 8c,e). For example, the T2 analysis in RAIN20 is 1–2 K colder than RAIN60 in southeastern Nebraska (NE) and northeastern KS (green rectangles). However, such differences between RAIN20 and RAIN60 exist only in local and scattered areas. It is difficult to verify the better T2 analysis field using the surface observations from METARs and mesonet. Compared to the 6 May 2019 case (Figs. 8a,b), the coverage with different T2 between RAIN60 and RAIN20 in the 3 May 2018 case is much smaller relative to its storm size. Such a discrepancy between the two cases may be related to the different storm evolutions during the radar DA period. Unlike the upscale growth and location change in the 6 May 2019 case, the storms in the 3 May 2018 case mildly grew in intensity without substantially changing the size and location. Therefore, localized and small differences in ACPs between RAIN60 and RAIN20 are obtained for the 3 May 2018 case. During the forecast period, the slightly different cold pools have similar effects on the storm evolution. As a result, RAIN60 and RAIN20 predict similar storm coverage throughout the 9-h lead times (Figs. 9e,g,h,j), leading to comparable FSSs at all thresholds (Figs. 3b1–b8).

b. Mixed performance among RAIN60, RAIN20, and RAIN5

For the 3 May 2019 MCS case, the reflectivity is assimilated during its early stage of decay. The size and intensity of the reflectivity core gradually reduced, and the location and coverage of the MCS changed slightly. Due to the decaying feature during the radar DA period, the differences in the T2 analyses among RAIN60, RAIN20, and RAIN5 are localized and are not as pronounced as in the comparisons with high sensitivity (Figs. 14a–c versus Figs. 8a–d). The cold pool at the final analysis time becomes slightly intensified and extensive (green arrows) as the RDF increases, although ACPs from all three experiments are larger than that indicated by the surface observations from METARs. Analogous to the 3 May 2018 case in section 5a, the slightly different cold pool analyses are consistent with the similar coverage of predicted storms for RAIN60, RAIN20, and RAIN5 (Figs. 14d–i), which coincides with their similar FSSs (Figs. 4c1–c8).

6. Conclusions and discussion

Using a single severe weather case over a limited-area domain, previous studies have demonstrated the high sensitivity of forecast performance to the radar reflectivity DA frequency (RDF). It remains unknown whether such high sensitivity can be applied to all convection events. The impact of RDF on convection-allowing forecasts over the CONUS that simultaneously contains multiscale features and multiple convection modes, such as in the operational HRRR, is still unclear. Therefore, 10 CONUS diverse cases in May of 2018 and 2019 are selected to systematically explore the degrees of sensitivity of convection-allowing forecasts to RDF. For each case, experiments with three reflectivity DA frequencies (60-, 20-, and 5-min intervals; named RAIN60, RAIN20, and RAIN5, respectively) are compared within the GSI-based EnVar system.
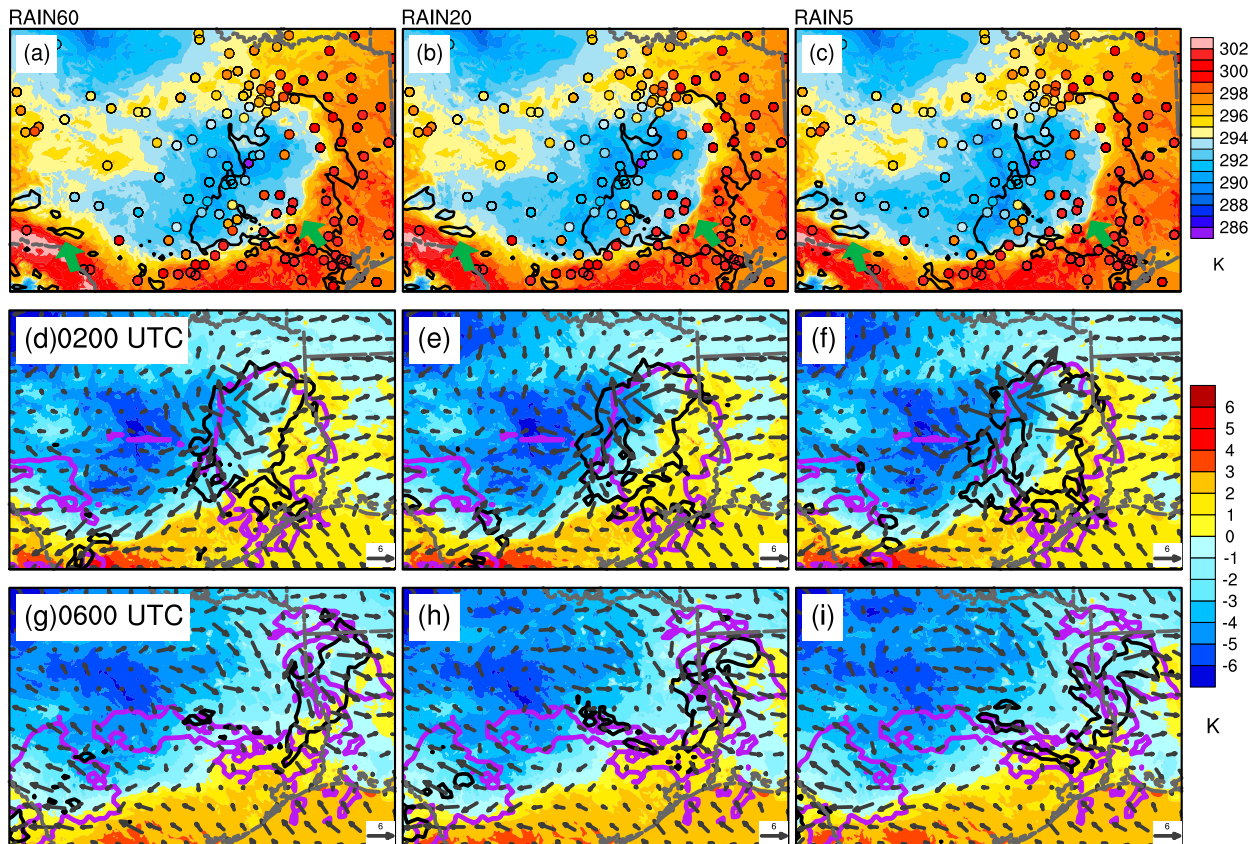
RAIN60

RAIN20

RAIN5

FIG. 14. The 2-m temperature analysis (shaded; K) for (a) RAIN60, (b) RAIN20, and (c) RAIN5 valid at 0000 UTC 3 May 2019. Observations of 2-m temperature are overlaid using shaded circles. The green arrows denote the regions with remarkable differences among RAIN60, RAIN20, and RAIN5. The black contours delineate analyzed composite reflectivity at 20 dBZ. Forecasts of 2-m potential temperature perturbation (shaded; K), 20-dBZ composite reflectivity (black contours), and 10-m wind vector perturbations (gray arrows; m s^{-1}) overlaid with observed composite reflectivity at 20 dBZ (purple contours) for (d),(g) RAIN60; (e),(h) RAIN20; and (f),(i) RAIN5 valid at 0200 UTC 3 May 2019 in (d)–(f) and 0600 UTC 3 May 2019 in (g)–(i).

For the 10 retrospective cases, the 0–9-h forecasts from RAIN60, RAIN20, and RAIN5 are evaluated using the FSS in terms of composite reflectivity and 1-h accumulated quantitative precipitation estimation. Based on the quantitative evaluation, the 10 diverse cases are categorized as 5 sensitive cases and 5 insensitive cases. Sensitive cases are the ones that have at least one pair of experiments distinguished by large and lasting FSS differences. It turns out that the classification is closely related to the storm life cycle during the radar DA period. The developing storms show strong sensitivity to RDF, whereas the mature or decaying storms do not. The pairwise comparisons are performed when RDF gradually increases in the sensitive cases. For the five sensitive cases, RAIN20 outperforms RAIN60 for three cases, which are associated with fast-moving, upscale-growing, and intensifying storms. RAIN20 is comparable to RAIN60 in the other two cases, where storms mildly enhance strength with relatively slight size growth and slow location changes. Further increasing RDF for the five sensitive cases, RAIN5 performs better than RAIN20 in two dryline convection cases, and RAIN5 is inferior to RAIN20 in three non-dryline convection cases.

Due to the least number of occurrences with the poorest performance, the 20-min reflectivity DA interval is recommended over the CONUS that simultaneously houses multiple storm life cycles, development features, and convection modes.

Detailed comparisons and diagnostics reveal that the analyzed cold pools (ACPs) become intensified and widespread as the RDF increases. The differences in ACPs among RAIN60, RAIN20, and RAIN5 result in different forecast skills of storms. The various degrees of sensitivity to RDF can be attributed to the magnitudes of ACP differences that vary with storm features during the radar DA period. The 6 May 2019 case associated with fast-moving, upscale-growing, and intensifying storms, and the 3 May 2018 squall-line case that is not triggered by the dryline are first performed to physically understand the better performance of RAIN20 compared to RAIN60 and RAIN5, respectively. For the 6 May 2019 case, the longer integration between two cycles in RAIN60 results in weaker and narrower ACPs than RAIN20. As a result, RAIN60 underpredicts the observed storms compared to RAIN20. For the 3 May 2018 case, the accumulated model imbalance introduced by the 5-min reflectivity DA causes the

inconsistency between reflectivity and vertical motions. The low-level downdrafts in RAIN5 facilitate spuriously stronger and wider ACPs than RAIN20. Thus, RAIN5 predicts faster-surging storms with larger displacement errors than RAIN20. Overall, RAIN20 in the two cases corrects the dynamic and thermodynamic fields in a relatively more coherent fashion and produces the ACPs that match better with the observations than RAIN60 and RAIN5, respectively.

In terms of the outperformance of RAIN5 compared to RAIN20, a dryline convection case on 2 May 2019 is investigated. Different from the squall-line case above, the negative influence of model imbalance in RAIN5 does not offset the benefits from the frequent DA. It is hypothesized that the limited negative influence of model imbalance is likely due to its dominant buoyancy forcing rather than dynamic forcing for the dryline case. Compared to RAIN20, the stronger ACPs in RAIN5 fit better with the observations, resulting in a larger storm that is closer to reality.

The 3 May 2018 case featuring slow-moving, slow-growing, and mild-intensifying storms with the comparable performance of RAIN60 and RAIN20, and the 3 May 2019 decaying MCS case with mixed forecasts among RAIN60, RAIN20, and RAIN5 are studied to investigate the reasons for the low sensitivity to RDF. In the two cases, the differences in ACPs among these experiments are not remarkable because of slow or slight changes of the storms during the radar DA period.

Due to the limited computational resources available, this study performs the sensitivity test using 10 cases. Future work should be warranted to extend the study with more cases, such as the full convectively active season. The model adjustment time following radar DA can rely heavily on the model dynamics (e.g., Pan and Wang 2019). Additional experiments are planned to investigate the RDF impact using the next-generation Limited Area Model based on the nonhydrostatic Finite Volume Cubed-Sphere Dynamical Core (FV3-LAM). The impact of RDF may be dependent on DA methods. Further work is also planned to perform the investigation of the RDF associated with different levels of nonlinearity using 4DEnVar (Wang and Lei 2014; Davis et al. 2021) and 3DEnVar.

Data availability statement. The Multi-Radar Multi-Sensor (MRMS) data assimilated and used for composite reflectivity and 1-h accumulated precipitation verifications can be accessed from the NCEP FTP service (https://mrms.ncep.noaa.gov/data). The model input and output data for the assimilation experiments are archived locally and available upon request to the corresponding author.

REFERENCES

Alexander, C., and Coauthors, 2020: Rapid Refresh (RAP) and High-Resolution Rapid Refresh (HRRR) model development. *30th Conf. on Weather Analysis and Forecasting (WAF)/26th Conf. on Numerical Weather Prediction (NWP)*, Boston, MA, Amer. Meteor. Soc., 8A.1, https://ams.confex.com/ams/2020Annual/webprogram/Paper370205.html.

Benjamin, S. G., G. A. Grell, J. M. Brown, T. G. Smirnova, and R. Bleck, 2004: Mesoscale weather prediction with the RUC hybrid isentropic–terrain-following coordinate model. *Mon. Wea. Rev.*, **132**, 473–494, https://doi.org/10.1175/1520-0493(2004)132<0473:MWPWTR>2.0.CO;2.

——, and Coauthors, 2016: A North American hourly assimilation and model forecast cycle: The Rapid Refresh. *Mon. Wea. Rev.*, **144**, 1669–1694, https://doi.org/10.1175/MWR-D-15-0242.1.

Brousseau, P., and Coauthors, 2008: A prototype convective-scale data assimilation system for operation: The AROME-RUC. HIRLAM Tech. Rep. 68, 8 pp.

Carley, J. R., 2012: Hybrid ensemble-3DVar radar data assimilation for the short-term prediction of convective storms. Ph.D. dissertation, Purdue University, 205 pp.

——, E. Rogers, S. Liu, B. Ferrier, E. Aligo, M. Pyle, X. Zhang, and G. DiMego, 2015: A status update for the NAMRR, an hourly-updated version of NAM Forecast System. *19th Conf. on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface (IOAS-AOLS)*, Phoenix, AZ, Amer. Meteor. Soc., 4.4, https://ams.confex.com/ams/95Annual/webprogram/Paper263759.html.

Casati, B., 2010: New developments of the intensity-scale technique within the Spatial Verification Methods Intercomparison Project. *Wea. Forecasting*, **25**, 113–143, https://doi.org/10.1175/2009WAF2222257.1.

——, G. Ross, and D. B. Stephenson, 2004: A new intensity-scale approach for the verification of spatial precipitation forecasts. *Meteor. Appl.*, **11**, 141–154, https://doi.org/10.1017/S1350482704001239.

Davis, B., X. Wang, and X. Lu, 2021: A comparison of HWRF six-hourly 4DEnVar and hourly 3DEnVar assimilation of inner core tail Doppler radar observations for the prediction of Hurricane Edouard (2014). *Atmosphere*, **12**, 942, https://doi.org/10.3390/atmos12080942.

Dong, J. L., and M. Xue, 2013: Assimilation of radial velocity and reflectivity data from coastal WSR-88D radars using an ensemble Kalman filter for the analysis and forecast of landfalling Hurricane Ike (2008). *Quart. J. Roy. Meteor. Soc.*, **139**, 467–487, https://doi.org/10.1002/qj.1970.

Dowell, D. C., F. Zhang, L. J. Wicker, C. Snyder, and N. A. Crook, 2004: Wind and temperature retrievals in the 17 May 1981 Arcadia, Oklahoma, supercell: Ensemble Kalman filter experiments. *Mon. Wea. Rev.*, **132**, 1982–2005, https://doi.org/10.1175/1520-0493(2004)132<1982:WATRIT>2.0.CO;2.

——, L. J. Wicker, and C. Snyder, 2011: Ensemble Kalman filter assimilation of radar observations of the 8 May 2003 Oklahoma City supercell: Influences of reflectivity observations on storm-scale analyses. *Mon. Wea. Rev.*, **139**, 272–294, https://doi.org/10.1175/2010MWR3438.1.

——, and Coauthors, 2022: The High-Resolution Rapid Refresh (HRRR): An hourly updating convection-allowing forecast

model. Part I: Motivation and system description. *Wea. Forecasting*, **37**, 1371–1395, https://doi.org/10.1175/WAF-D-21-0151.1.

Du, J., G. DiMego, D. Jovic, B. Ferrier, B. Yang, and B. Zhou, 2015: Short Range Ensemble Forecast (SREF) system at NCEP: Recent development and future transition. *27th Conf. on Weather Analysis and Forecasting/23rd Conf. on Numerical Weather Prediction*, Chicago, IL, Amer. Meteor. Soc., 2A.5, https://ams.confex.com/ams/27WAF23NWP/webprogram/Paper273421.html.

Duda, J. D., X. Wang, F. Kong, and M. Xue, 2014: Using varied microphysics to account for uncertainty in warm-season QPF in a convection-allowing ensemble. *Mon. Wea. Rev.*, **142**, 2198–2219, https://doi.org/10.1175/MWR-D-13-00297.1.

——, ——, Y. Wang, and J. R. Carley, 2019: Comparing the assimilation of radar reflectivity using the direct GSI-based Ensemble-Variational (EnVar) and indirect cloud analysis methods in convection-allowing forecasts over the continental United States. *Mon. Wea. Rev.*, **147**, 1655–1678, https://doi.org/10.1175/MWR-D-18-0171.1.

Dunkerley, D. L., 2021: Light and low-intensity rainfalls: A review of their classification, occurrence, and importance in land surface, ecological and environmental processes. *Earth-Sci. Rev.*, **214**, 103529, https://doi.org/10.1016/j.earscirev.2021.103529.

Fertig, E. J., J. Harlim, and B. R. Hunt, 2007: A comparative study of 4D-VAR and a 4D ensemble Kalman filter: Perfect model simulations with Lorenz-96. *Tellus*, **59A**, 96–100, https://doi.org/10.1111/j.1600-0870.2006.00205.x.

Fierro, A. O., A. J. Clark, E. R. Mansell, D. R. MacGorman, S. R. Dembek, and C. L. Ziegler, 2015: Impact of storm-scale lightning data assimilation on WRF-ARW precipitation forecasts during the 2013 warm season over the contiguous United States. *Mon. Wea. Rev.*, **143**, 757–777, https://doi.org/10.1175/MWR-D-14-00183.1.

Gasperoni, N. A., X. Wang, and Y. Wang, 2020: A comparison of methods to sample model errors for convection-allowing ensemble forecasts in the setting of multiscale initial conditions produced by the GSI-Based EnVar assimilation system. *Mon. Wea. Rev.*, **148**, 1177–1203, https://doi.org/10.1175/MWR-D-19-0124.1.

Gauntlett, D. J., and R. S. Seaman, 1974: Four-dimensional data assimilation experiments in the Southern Hemisphere. *J. Appl. Meteor.*, **13**, 845–853, https://doi.org/10.1175/1520-0450(1974)013<0845:FDDAEI>2.0.CO;2.

Gilleland, E., D. Ahijevych, B. G. Brown, B. Casati, and E. E. Ebert, 2009: Intercomparison of spatial forecast verification methods. *Wea. Forecasting*, **24**, 1416–1430, https://doi.org/10.1175/2009WAF2222269.1.

Hamill, T., J. S. Whitaker, M. Fiorino, and S. J. Benjamin, 2011: Global ensemble predictions of 2009's tropical cyclones initialized with an ensemble Kalman filter. *Mon. Wea. Rev.*, **139**, 668–688, https://doi.org/10.1175/2010MWR3456.1.

Houtekamer, P. L., and F. Zhang, 2016: Review of the ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.*, **144**, 4489–4532, https://doi.org/10.1175/MWR-D-15-0440.1.

Hu, M., and M. Xue, 2007: Impact of configurations of rapid intermittent assimilation of WSR-88D radar data for the 8 May 2003 Oklahoma City tornadic thunderstorm case. *Mon. Wea. Rev.*, **135**, 507–525, https://doi.org/10.1175/MWR3313.1.

Huang, X.-Y., and P. Lynch, 1993: Diabatic digital-filtering initialization: Application to the HIRLAM model. *Mon. Wea. Rev.*, **121**, 589–603, https://doi.org/10.1175/1520-0493(1993)121<0589:DDFIAT>2.0.CO;2.

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

Johnson, A., and X. Wang, 2012: Verification and calibration of neighborhood and object-based probabilistic precipitation forecasts from a multimodel convection-allowing ensemble. *Mon. Wea. Rev.*, **140**, 3054–3077, https://doi.org/10.1175/MWR-D-11-00356.1.

——, and ——, 2017: Design and implementation of a GSI-based convection-allowing ensemble data assimilation and forecast system for the PECAN field experiment. Part I: Optimal configurations for nocturnal convection prediction using retrospective cases. *Wea. Forecasting*, **32**, 289–315, https://doi.org/10.1175/WAF-D-16-0102.1.

——, and ——, 2020: Interactions between physics diversity and multiscale initial condition perturbations for storm-scale ensemble forecasting. *Mon. Wea. Rev.*, **148**, 3549–3565, https://doi.org/10.1175/MWR-D-20-0112.1.

——, and Coauthors, 2014: Multiscale characteristics and evolution of perturbations for warm season convection-allowing precipitation forecasts: Dependence on background flow and method of perturbation. *Mon. Wea. Rev.*, **142**, 1053–1073, https://doi.org/10.1175/MWR-D-13-00204.1.

——, X. Wang, J. R. Carley, L. J. Wicker, and C. Karstens, 2015: A comparison of multiscale GSI-based EnKF and 3DVar data assimilation using radar and conventional observations for midlatitude convective-scale precipitation forecasts. *Mon. Wea. Rev.*, **143**, 3087–3108, https://doi.org/10.1175/MWR-D-14-00345.1.

——, ——, Y. Wang, A. Reinhart, A. J. Clark, and I. L. Jirak, 2020: Neighborhood- and object-based probabilistic verification of the OU MAP ensemble forecasts during 2017 and 2018 Hazardous Weather Testbeds. *Wea. Forecasting*, **35**, 169–191, https://doi.org/10.1175/WAF-D-19-0060.1.

Knopfmeier, K. H., and D. J. Stensrud, 2013: Influence of mesonet observations on the accuracy of surface analyses generated by an ensemble Kalman filter. *Wea. Forecasting*, **28**, 815–841, https://doi.org/10.1175/WAF-D-12-00078.1.

Labriola, J., Y. Jung, C. Liu, and M. Xue, 2021: Evaluating forecast performance and sensitivity to the GSI EnKF data assimilation configuration for the 28–29 May 2017 mesoscale convective system case. *Wea. Forecasting*, **36**, 127–146, https://doi.org/10.1175/WAF-D-20-0071.1.

Lange, H., and G. C. Craig, 2014: The impact of data assimilation length scales on analysis and prediction of convective storms. *Mon. Wea. Rev.*, **142**, 3781–3808, https://doi.org/10.1175/MWR-D-13-00304.1.

Lei, T., M. Xue, T. Y. Yu, and M. Teshiba, 2007: Study on the optimal scanning strategies of phase-array radar through ensemble Kalman filter assimilation of simulated data. *33rd Int. Conf. on Radar Meteorology*, Cairns, Australia, Amer. Meteor. Soc., P7.1, http://ams.confex.com/ams/pdfpapers/124022.pdf.

Liu, H., and M. Xue, 2008: Prediction of convective initiation and storm evolution on 12 June 2002 during IHOP_2002. Part I: Control simulation and sensitivity experiments. *Mon. Wea. Rev.*, **136**, 2261–2282, https://doi.org/10.1175/2007MWR2161.1.

Lu, X., X. Wang, Y. Li, M. Tong, and X. Ma, 2017: GSI-based ensemble-variational hybrid data assimilation for HWRF for hurricane initialization and prediction: Impact of various error covariances for airborne radar observation assimilation.

Quart. J. Roy. Meteor. Soc., **143**, 223–239, https://doi.org/10.1002/qj.2914.

Lynch, P., and X. Huang, 1992: Initialization of the HIRLAM model using a digital filter. *Mon. Wea. Rev.*, **120**, 1019–1034, https://doi.org/10.1175/1520-0493(1992)120<1019:IOTHMU>2.0.CO;2.

Maddox, R. A., J. Zhang, J. J. Gourley, and K. W. Howard, 2002: Weather radar coverage over the contiguous United States. *Wea. Forecasting*, **17**, 927–934, https://doi.org/10.1175/1520-0434(2002)017<0927:WRCOTC>2.0.CO;2.

Marion, G. R., and R. J. Trapp, 2018: The dynamical coupling of convective updrafts, downdrafts, and cold pools in simulated supercell thunderstorms. *J. Geophys. Res. Atmos.*, **124**, 664–683, https://doi.org/10.1029/2018JD029055.

Markowski, P., and Y. Richardson, 2010: *Mesoscale Meteorology in Midlatitudes.* Wiley-Blackwell, 430 pp.

Mittermaier, M., and N. Roberts, 2010: Intercomparison of spatial forecast verification methods: Identifying skillful spatial scales using the fractions skill score. *Wea. Forecasting*, **25**, 343–354, https://doi.org/10.1175/2009WAF2222260.1.

——, ——, and S. A. Thompson, 2013: A long-term assessment of precipitation forecast skill using the fractions skill score. *Meteor. Appl.*, **20**, 176–186, https://doi.org/10.1002/met.296.

Miyoshi, T., and Coauthors, 2016: "Big Data Assimilation" revolutionizing severe weather prediction. *Bull. Amer. Meteor. Soc.*, **97**, 1347–1354, https://doi.org/10.1175/BAMS-D-15-00144.1.

Nakanishi, M., and H. Niino, 2006: An improved Mellor–Yamada level 3 model: Its numerical stability and application to a regional prediction of advecting fog. *Bound.-Layer Meteor.*, **119**, 397–407, https://doi.org/10.1007/s10546-005-9030-8.

——, and ——, 2009: Development of an improved turbulent closure model for the atmospheric boundary layer. *J. Meteor. Soc. Japan*, **87**, 895–912, https://doi.org/10.2151/jmsj.87.895.

Olson, J. B., J. S. Kenyon, W. M. Angevine, J. M. Brown, M. Pagowski, and K. Sušelj, 2019: A description of the MYNN-EDMF scheme and the coupling to other components in WRF-ARW. NOAA Tech. Memo. OAR GSD 61, 42 pp., https://doi.org/10.25923/n9wm-be49.

Pan, Y., and M. Wang, 2019: Impact of the assimilation frequency of radar data with the ARPS 3DVar and cloud analysis system on forecasts of a squall line in southern China. *Adv. Atmos. Sci.*, **36**, 160–172, https://doi.org/10.1007/s00376-018-8087-5.

Peña, M., and E. Kalnay, 2004: Separating fast and slow modes in coupled chaotic systems. *Nonlinear Processes Geophys.*, **11**, 319–327, https://doi.org/10.5194/npg-11-319-2004.

Roberts, N., 2008: Assessing the spatial and temporal variation in the skill of precipitation forecasts from an NWP model. *Meteor. Appl.*, **15**, 163–169, https://doi.org/10.1002/met.57.

——, 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.

Rogers, E., and Coauthors, 2017: Upgrades to the NCEP North American Mesoscale (NAM) system. *Working Group on Numerical Experimentation Blue Book*, 2 pp., http://wmc.meteoinfo.ru/bluebook/uploads/2017/docs/05_Rogers_Eric_mesoscale_modeling.pdf.

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

——, and Coauthors, 2010: Toward improved convection-allowing ensembles: Model physics sensitivities and optimizing probabilistic guidance with small ensemble membership. *Wea. Forecasting*, **25**, 263–280, https://doi.org/10.1175/2009WAF2222267.1.

Smith, T. L., S. G. Benjamin, J. M. Brown, S. Weygandt, T. Smirnova, and B. Schwartz, 2008: Convection forecasts from the hourly updated, 3-km High Resolution Rapid Refresh (HRRR) model. *24th Conf. on Severe Local Storms*, Savannah, GA, Amer. Meteor. Soc., 11.1, https://ams.confex.com/ams/pdfpapers/142055.pdf.

Smith, T. M., and Coauthors, 2016: Multi-Radar Multi-Sensor (MRMS) severe weather and aviation products: Initial operating capabilities. *Bull. Amer. Meteor. Soc.*, **97**, 1617–1630, https://doi.org/10.1175/BAMS-D-14-00173.1.

Stratman, D. R., M. C. Coniglio, S. E. Koch, and M. Xue, 2013: Use of multiple verification methods to evaluate forecasts of convection from hot- and cold-start convection-allowing models. *Wea. Forecasting*, **28**, 119–138, https://doi.org/10.1175/WAF-D-12-00022.1.

——, N. Yussouf, Y. Jung, T. A. Supinie, M. Xue, P. S. Skinner, and B. J. Putnam, 2020: Optimal temporal frequency of NSSL phased array radar observations for an experimental Warn-on-Forecast system. *Wea. Forecasting*, **35**, 193–214, https://doi.org/10.1175/WAF-D-19-0165.1.

Supinie, T. A., N. Yussouf, Y. Jung, M. Xue, J. Cheng, and S. Wang, 2017: Comparison of the analyses and forecasts of a tornadic supercell storm from assimilating phased-array radar and WSR-88D observations. *Wea. Forecasting*, **32**, 1379–1401, https://doi.org/10.1175/WAF-D-16-0159.1.

Thompson, G., and T. Eidhammer, 2014: A study of aerosol impacts on clouds and precipitation development in a large winter cyclone. *J. Atmos. Sci.*, **71**, 3636–3658, https://doi.org/10.1175/JAS-D-13-0305.1.

Tompkins, A. M., 2001: Organization of tropical convection in low vertical wind shears: The role of cold pools. *J. Atmos. Sci.*, **58**, 1650–1672, https://doi.org/10.1175/1520-0469(2001)058<1650:OOTCIL>2.0.CO;2.

Wang, S., M. Xue, and J. Min, 2013: A four-dimensional asynchronous ensemble square-root filter (4DEnSRF) algorithm and tests with simulated radar data. *Quart. J. Roy. Meteor. Soc.*, **139**, 805–819, https://doi.org/10.1002/qj.1987.

Wang, X., 2010: Incorporating ensemble covariance in the grid point statistical interpolation variational minimization: A mathematical framework. *Mon. Wea. Rev.*, **138**, 2990–2995, https://doi.org/10.1175/2010MWR3245.1.

——, and T. Lei, 2014: GSI-based four-dimensional ensemble–variational (4DEnsVar) data assimilation: Formulation and single-resolution experiments with real data for NCEP Global Forecast System. *Mon. Wea. Rev.*, **142**, 3303–3325, https://doi.org/10.1175/MWR-D-13-00303.1.

——, D. Parrish, D. Kleist, and J. Whitaker, 2013: GSI 3DVar based ensemble–variational hybrid data assimilation for NCEP Global Forecast System: Single-resolution experiments. *Mon. Wea. Rev.*, **141**, 4098–4117, https://doi.org/10.1175/MWR-D-12-00141.1.

Wang, Y., and X. Wang, 2017: Direct assimilation of radar reflectivity without tangent linear and adjoint of the nonlinear observation operator in the GSI-based EnVar system: Methodology and experiment with the 8 May 2003 Oklahoma City tornadic supercell. *Mon. Wea. Rev.*, **145**, 1447–1471, https://doi.org/10.1175/MWR-D-16-0231.1.

——, and ——, 2021: Rapid update with EnVar direct radar reflectivity data assimilation for the NOAA regional convection-

allowing NMMB model over the CONUS: System description and initial experiment results. *Atmosphere*, **12**, 1286, https://doi.org/10.3390/atmos12101286.

Wu, W.-S., D. F. Parrish, E. Rogers, and Y. Lin, 2017: Regional ensemble-variational data assimilation using global ensemble forecasts. *Wea. Forecasting*, **32**, 83–96, https://doi.org/10.1175/WAF-D-16-0045.1.

Xue, M., and W. J. Martin, 2006: A high-resolution modeling study of the 24 May 2002 dryline case during IHOP. Part I: Numerical simulation and general evolution of the dryline and convection. *Mon. Wea. Rev.*, **134**, 149–171, https://doi.org/10.1175/MWR3071.1.

——, M. Tong, and K. K. Droegemeier, 2006: An OSSE framework based on the ensemble square root Kalman filter for evaluating the impact of data from radar networks on thunderstorm analysis and forecasting. *J. Atmos. Oceanic Technol.*, **23**, 46–66, https://doi.org/10.1175/JTECH1835.1.

Yussouf, N., and D. J. Stensrud, 2010: Impact of phased-array radar observations over a short assimilation period: Observing system simulation experiments using an ensemble Kalman filter. *Mon. Wea. Rev.*, **138**, 517–538, https://doi.org/10.1175/2009MWR2925.1.

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.

Zhang, J., and Coauthors, 2016: Multi-Radar Multi-Sensor (MRMS) quantitative precipitation estimation: Initial operating capabilities. *Bull. Amer. Meteor. Soc.*, **97**, 621–638, https://doi.org/10.1175/BAMS-D-14-00174.1.

Zhou, X., Y. Zhu, D. Hou, Y. Luo, J. Peng, and R. Wobus, 2017: Performance of the new NCEP global ensemble forecast system in a parallel experiment. *Wea. Forecasting*, **32**, 1989–2004, https://doi.org/10.1175/WAF-D-17-0023.1.