

# Assimilation of GOES-R Geostationary Lightning Mapper Flash Extent Density Data in GSI EnKF for the Analysis and Short-Term Forecast of a Mesoscale Convective System

RONG KONG

*Center for Analysis and Prediction of Storms, University of Oklahoma, Norman, Oklahoma*

MING XUE

*Center for Analysis and Prediction of Storms, and School of Meteorology, University of Oklahoma, Norman, Oklahoma*

ALEXANDRE O. FIERRO

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

YOUNGSUN JUNG AND CHENGSI LIU

*Center for Analysis and Prediction of Storms, University of Oklahoma, Norman, Oklahoma*

EDWARD R. MANSELL AND DONALD R. MACGORMAN

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

(Manuscript received 11 June 2019, in final form 9 March 2020)

## ABSTRACT

The recently launched Geostationary Operational Environmental Satellite “R-series” (GOES-R) satellites carry the Geostationary Lightning Mapper (GLM) that measures from space the total lightning rate in convective storms at high spatial and temporal frequencies. This study assimilates, for the first time, real GLM total lightning data in an ensemble Kalman filter (EnKF) framework. The lightning flash extent density (FED) products at 10-km pixel resolution are assimilated. The capabilities to assimilate GLM FED data are first implemented into the GSI-based EnKF data assimilation (DA) system and tested with a mesoscale convective system (MCS). FED observation operators based on graupel mass or graupel volume are used. The operators are first tuned through sensitivity experiments to determine an optimal multiplying factor to the operator, before being used in FED DA experiments FEDM and FEDV that use the graupel-mass or graupel-volume-based operator, respectively. Their results are compared to a control experiment (CTRL) that does not assimilate any FED data. Overall, both DA experiments outperform CTRL in terms of the analyses and short-term forecasts of FED and composite/3D reflectivity. The assimilation of FED is primarily effective in regions of deep moist convection, which helps improve short-term forecasts of convective threats, including heavy precipitation and lightning. Direct adjustments to graupel mass via observation operator as well as adjustments to other model state variables through flow-dependent ensemble cross covariance within EnKF are shown to work together to generate model-consistent analyses and overall improved forecasts.

## 1. Introduction

Hazards associated with severe weather such as flash floods, hail, lightning, and tornadoes pose significant threats to life and property. Based on a survey from the U.S. National Weather Service (NWS), lightning and

tornadoes alone are responsible, on average, for 45 and 60 fatalities annually within the United States; better forecasting and warning of such hazardous weather should help reduce the number of fatalities. Improving severe weather prediction thus remains a critical component of the NWS mission. Owing to the notable advancement in numerical weather prediction (NWP) models, computational resources, as well as development

---

*Corresponding author:* Ming Xue, mxue@ou.edu

DOI: 10.1175/MWR-D-19-0192.1

© 2020 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](https://www.ametsoc.org/PUBSReuseLicenses) ([www.ametsoc.org/PUBSReuseLicenses](https://www.ametsoc.org/PUBSReuseLicenses)).

of advanced data assimilation (DA) systems, the skill and usefulness of severe weather forecasts have improved significantly in recent years.

Modern DA techniques, in particular, provide the best estimate of the physical states within NWP models by optimally combining model forecasts and observations. A number of DA methods have been developed over the past few decades to improve the initial conditions of NWP, including the three-dimensional variational (3DVar), four-dimensional variational (4DVar), ensemble Kalman filter (EnKF), and ensemble variational hybrid DA methods. For the prediction of convective storms and associated hazards, grid spacings small enough to resolve the storms are needed (Lilly 1990; Clark et al. 2010; Sun et al. 2014). To initialize convective-scale models, observations at comparative resolutions are required, and many studies have shown that weather radar observations can effectively improve short-term forecasts of convective weather (e.g., Gao et al. 2004; Tong and Xue 2005; Xue et al. 2006; Jung et al. 2008; Aksoy et al. 2009, 2010; Kain et al. 2010; Dowell et al. 2011; Gao et al. 2013; Wang et al. 2013; Liu and Xue 2016; Kong et al. 2018; Liu et al. 2019).

Akin to radar observations, lightning data are a source of high-resolution observations that is able to indicate the presence of deep, moist, mixed-phased convection. The availability of operational total lightning data [i.e., cloud-to-ground (CG) plus intracloud (IC) flashes] can potentially provide valuable information to NWP models at the convective scale. This is because total lightning flash rates were shown to correlate well with bulk measures of convective strength such as  $10+ \text{ m s}^{-1}$  updraft and graupel mass or volume in continental storms (Goodman et al. 1988; MacGorman et al. 1989; Carey and Rutledge 1998; MacGorman et al. 2005; Wiens et al. 2005; Fierro et al. 2006; Kuhlman et al. 2006; Deierling and Petersen 2008; MacGorman et al. 2011) and in hurricanes (Fierro et al. 2015b; Fierro and Mansell 2018). Thus, leveraging the total lightning measurements to inform NWP models about the evolution and intensification of the deep moist convection is expected to help augment our ability to forecast convective threats (Papadopoulos et al. 2005; Fierro et al. 2012).

The Geostationary Lightning Mapper (GLM; Goodman et al. 2013) provides total lightning data products at high spatial and temporal resolutions over the continental United States (CONUS). In November 2017, GOES-R (operationally renamed as *GOES-16*) was repositioned from its checkout location to its final, operational location centered over  $75^\circ\text{W}$  over the Western Hemisphere, allowing the GLM to deliver observation measurements between  $52^\circ\text{N}$  and  $52^\circ\text{S}$  latitude. The GLM

provides real-time, high-frequency monitoring of lightning-active convection over CONUS and neighboring, vast data-sparse regions. Thus, GLM observations will be able to complement ground-based data in data-sparse areas such as oceans and mountainous regions.

Past works focusing on lightning DA (LDA) made use of lightning data detected from various ground-based platforms such as the very-high-frequency (VHF) Lightning Mapping Arrays (LMA, Rison et al. 1999; Thomas et al. 2004; Rust et al. 2005; Mansell et al. 2007; Allen et al. 2016), the very-low-frequency (VLF) World Wide Lightning Location Network (WWLLN; Dixon et al. 2016), the broadband Earth Network Total Lightning Network (Fierro et al. 2012, 2015b), the VLF U.S. National Lightning Detection Network (Nag and Rakov 2014), and the VLF Pacific Lightning Detection Network/Long-Range Lightning Detection Network (PacNet/LLDN; Pessi and Businger 2009). Anticipating the launch of GOES-R satellites carrying the GLM, some of the above-cited preliminary LDA studies used ground-based lightning data to create pseudoGLM data and assimilated the data primarily through nudging or direct insertion methods (Fierro et al. 2012, 2014; Dixon et al. 2016). These LDA techniques employed empirical, functional relationships between a proxy variable known to be well associated with lightning—such as relative humidity or rainfall rates—and lightning density fields. Recently, Fierro et al. (2016) applied and extended such moistening technique to a variational framework by assimilating pseudo- $q_w$  (water vapor mixing ratio) retrieved from the LMA flash extent density (FED) data with similar short-term forecast ( $\leq 3\text{h}$ ) improvements. Additionally, Fierro et al. (2016) assimilated lightning information in tandem with level-II radar data and found that the best forecast improvements were generally seen when both radar and lightning data were assimilated. Generally speaking, the above works have shown the potential benefit of the LDA in terms of achieving forecast improvements that were to certain extent comparable to that of radar DA.

With the availability of real GOES-R GLM data over the past 2 years, limited studies have been carried out assimilating such real data and examining their impacts on convective storm predictions. In Fierro et al. (2019) and Hu et al. (2020), the indirect approach that assimilates via 3DVar pseudo-moisture derived from GLM lightning data (Fierro et al. 2016) was applied to real GLM data, and positive impacts were found on short-term storm forecasting. Generally wet bias was also found in the forecasts, however.

Compared to nudging, direct insertion, or 3DVar methods, EnKF has clear advantages. With EnKF, background error covariances are estimated from an

ensemble of forecasts and are used to adjust model state variables that are directly or indirectly related to the observations (Evensen 1994; Tong and Xue 2005). Thus far, only two studies have attempted to assimilate lightning data in the EnKF framework: namely, the OSSE study of Mansell (2014) and that of Allen et al. (2016); the former assimilated simulated FED data while the latter assimilated pseudo-GLM observations derived from ground-based LMA. In Mansell (2014), observation operators for FED were derived as a function of graupel mass or graupel volume based on explicit simulations of electrification processes (e.g., noninductive charging rates) within convective storms. The assimilation of FED was found to effectively modulate convection by better depicting the location of reflectivity echoes and the spatial location probabilities of convective updrafts. Additionally, the assimilation of zero flash rates was found to partially address the presence of spurious convection in similar way as zero reflectivity data (Tong and Xue 2005). Allen et al. (2016) extended the work of Mansell (2014) by applying the EnKF DA system to two contrasting real cases; namely, a tornadic supercell, and a nonsevere multicellular storm. Their study assimilated pseudo-GLM FED observations derived from the ground-based LMA data (Rison et al. 1999). Tests were conducted to evaluate the performance of different FED observation operators with the best results obtained when assimilating 1-min FED rates using the observation operators based on graupel mass or graupel volume (Allen et al. 2016).

Though the LDA studies cited above showed promises for improving the analysis and forecast of convective storms, there still are some limitations that need to be addressed. First, the EnKF DA studies of Mansell (2014) and Allen et al. (2016) assumed a relatively simple framework (e.g., OSSE or horizontal homogeneity in the storm environment and no terrain or land surface processes) and Fierro et al. (Fierro et al. 2012, 2014, 2015a, 2016) used simpler DA procedures such as nudging or variational analysis of derived pseudo-moisture data. Thus, LDA using a more sophisticated DA procedure (such as EnKF) that is coupled with an NWP model with full model physics is desired. Second, past studies either assimilated simulated FED data (Mansell 2014) or pseudo-GLM FED data derived from the ground-based LMA data (Allen et al. 2016; Fierro et al. 2016) or the Earth Networks Total Lightning Network (Liu and Heckman 2010; Fierro et al. 2016). To the best of our knowledge, there exist currently very few studies in the literature that assimilate the real GLM FED data (Bruning et al. 2019), given that the GOES-R satellite was launched only recently.

The chief goal of this study is to enhance the DA capabilities for GLM lightning data by implementing them within the operational Gridpoint Statistical Interpolation (GSI; Wu et al. 2002; Kleist et al. 2009) framework using EnKF, and evaluate their impacts on convection-permitting/resolving resolution ( $\Delta x \leq 5$  km) forecasts.

The remainder of this manuscript is organized as follows. In section 2, we introduce the EnKF algorithms and the lightning observation operators. Section 3 provides a brief overview of the MCS case used in this study, describes the preprocessing of the GLM-FED observations and the experimental design. In section 4, FED observation operators based on the graupel mass and graupel volume are first tuned then tested in DA experiments, and their results are compared with those of control run that does not assimilate any FED data. The summary and conclusions are given in section 5, with some additional discussions on future studies.

## 2. Assimilation method and observation operators

### a. The EnKF algorithm

The GSI-based EnKF system is enhanced for it to be able to assimilate FED data derived from the GOES-R GLM observations. The algorithm used is the ensemble square-root filter (EnSRF) of Whitaker and Hamill (2002). A scalable implementation of EnSRF after Anderson and Collins (2007) is used in GSI, in which both state vectors and precalculated observation priors are updated by the filter. This ensemble filter is equivalent to the traditional EnSRF when the forward observation operator is linear. In EnSRF, the ensemble mean and ensemble perturbations are updated separately.

Specifically, the ensemble mean forecast of the model state and observation prior are first updated by Eqs. (1) and (2) as below:

$$\bar{\mathbf{x}}^a = \bar{\mathbf{x}}^b + \mathbf{K}_x[\mathbf{y}^o - \bar{\mathbf{y}}^b], \quad (1)$$

$$\bar{\mathbf{y}}^a = \bar{\mathbf{y}}^b + \mathbf{K}_y[\mathbf{y}^o - \bar{\mathbf{y}}^b]. \quad (2)$$

In the above,  $\mathbf{x}$  and  $\mathbf{y}$  are the state vector and observation vector, respectively. Superscripts  $a$ ,  $b$ , and  $o$  denotes analysis, background, and observation, respectively, while overbar denotes ensemble mean.  $\mathbf{y}^b = H(\mathbf{x}^b)$  is the observation prior,  $H$  is the observation operator (which can be nonlinear), and  $\mathbf{K}_x$  and  $\mathbf{K}_y$  are the Kalman gains used to update the means of the state vector and observation prior, respectively, via the background error covariances among the state (observation prior) and observed quantities to be calculated from the ensemble. Subscript  $x$  and  $y$  denote the state and observation prior, respectively.

The ensemble perturbations, denoted by prime, of state variables and observation priors are then updated by Eqs. (3) and (4), respectively, for ensemble member  $k$ :

$$x_k^a = x_k^{b'} - \tilde{\mathbf{K}}_x y_k^{b'}, \quad (3)$$

$$y_k^a = y_k^{b'} - \tilde{\mathbf{K}}_y y_k^{b'}, \quad (4)$$

The final analyses for the ensemble members are

$$x_k^a = \bar{x}^a + x_k^a. \quad (5)$$

Here,

$$\mathbf{K}_x = [\boldsymbol{\rho}_{xy} \circ (\hat{\mathbf{P}}^b \mathbf{H}^T)] (\mathbf{H} \hat{\mathbf{P}}^b \mathbf{H}^T + \mathbf{R})^{-1}, \quad (6)$$

$$\mathbf{K}_y = [\boldsymbol{\rho}_{yy} \circ (\mathbf{H} \hat{\mathbf{P}}^b \mathbf{H}^T)] (\mathbf{H} \hat{\mathbf{P}}^b \mathbf{H}^T + \mathbf{R})^{-1}, \quad (7)$$

$$\hat{\mathbf{P}}^b \mathbf{H}^T = \frac{1}{N-1} \sum_{k=1}^N \mathbf{x}_k^{b'} \mathbf{y}_k^{b'T}, \quad (8)$$

$$\mathbf{H} \hat{\mathbf{P}}^b \mathbf{H}^T = \frac{1}{N-1} \sum_{k=1}^N \mathbf{y}_k^{b'} \mathbf{y}_k^{b'T}. \quad (9)$$

In the above equations,  $\mathbf{H}$  is the tangent linear observation operator,  $\hat{\mathbf{P}}^b$  is the ensemble covariance estimated from an ensemble of background states  $\mathbf{x}_k^b$  with  $N$  being the ensemble size, and  $\hat{\mathbf{P}}^b \mathbf{H}^T$  is the background error covariance between the model state  $\mathbf{x}$  and observation prior  $\mathbf{y}$ , while  $\mathbf{H} \hat{\mathbf{P}}^b \mathbf{H}^T$  is that between observation prior  $\mathbf{y}$  itself. It is these ensemble-derived statistics that allow for the updating of all model state variables with the FED observations, including state variables that do not appear in the observation operator of FED.

A distance-dependent localization function (Gaspari and Cohn 1999) is used for the localization of the ensemble covariance;  $\boldsymbol{\rho}_{xy} \circ$  and  $\boldsymbol{\rho}_{yy} \circ$  are the Schur product of the correlation matrix  $\boldsymbol{\rho}_{xy}$  and  $\boldsymbol{\rho}_{yy}$ , respectively. Subscripts  $xy$  and  $yy$  are introduced to illustrate the fact that the correlation matrix  $\boldsymbol{\rho}_{xy}$  is calculated between model and observation space, while  $\boldsymbol{\rho}_{yy}$  is calculated purely in the observation space;  $\tilde{\mathbf{K}}_x$  and  $\tilde{\mathbf{K}}_y$  are the “reduced” Kalman gain used to update the ensemble perturbations of state and observation priors, respectively. When serially assimilating independent observations,  $\tilde{\mathbf{K}}_x$  and  $\tilde{\mathbf{K}}_y$  are used so that the correct analysis error covariance can be estimated without perturbing the observations (Whitaker and Hamill 2002):

$$\tilde{\mathbf{K}}_x = \left( 1 + \sqrt{\frac{\mathbf{R}}{\mathbf{H} \hat{\mathbf{P}}^b \mathbf{H}^T + \mathbf{R}}} \right) \mathbf{K}_x, \quad (10)$$

$$\tilde{\mathbf{K}}_y = \left( 1 + \sqrt{\frac{\mathbf{R}}{\mathbf{H} \hat{\mathbf{P}}^b \mathbf{H}^T + \mathbf{R}}} \right) \mathbf{K}_y. \quad (11)$$

### b. The FED observation operators

Two out of the three FED observation operators from Allen et al. (2016) are evaluated in this study. These operators relate lightning to either the simulated graupel mass or graupel volume. In Allen et al. (2016), a first guess is created by radar radial velocity DA, with explicit electrification and lightning being simulated during the radial velocity DA tests. Then linear best-fit observation operators are derived based on relationships between model microphysical quantities (graupel volume, graupel mass, and noninductive charging rates) and model flash rates in the analysis.

The lightning observation operators based on graupel mass (FEDM) and graupel volume (FEDV) from Allen et al. (2016) are given by

$$\text{FEDM} = 2.088 \times 10^{-8} (\text{GM}). \quad (12)$$

$$\text{FEDV} = 1.5 \times 0.017 (\text{GV}). \quad (13)$$

Before directly applying the above observation operators to assimilate real GLM FED observations, two possible issues need be considered. First, the observation operators from Allen et al. (2016) were derived based on idealized model simulation data, which can have significant biases, given inevitable errors in the simulation of both microphysical and electrification processes. Second, recent preliminary ground-truth studies have revealed notably lower flash rates from the GLM compared to ground-based LMA for high-flash-rate storms (Carey et al. 2019). Akin to its predecessor instrument, the Lightning Imaging Sensor (LIS; Christian et al. 2000), the detection efficiency of the GLM is likely affected by the cloud optical depth and/or the time of day (Fierro et al. 2019). To consider the above possible issues, sensitivity experiments on the observation operators are conducted by adding a tuning factor ( $C_M$  or  $C_V$ ) to the right-hand side of Eqs. (12) and (13), which are shown as below:

$$\text{FEDM} = C_M \times 2.088 \times 10^{-8} (\text{GM}), \quad (14)$$

$$\text{FEDV} = C_V \times 1.5 \times 0.017 (\text{GV}). \quad (15)$$

In this paper, we will test values of  $C_M$  and  $C_V$  equaling to 1, 1/2, 1/4, and 1/6. GM (in kg) is the graupel mass and GV (in  $\text{m}^3$ ) the graupel volume. Graupel volume is the sum of all grid cells characterized by a graupel mixing ratio exceeding  $0.5 \text{ g kg}^{-1}$  within an assumed 12-km pixel from the GLM centroid. Similar to Mansell (2014) and Allen et al. (2016), each microphysical quantity is summed

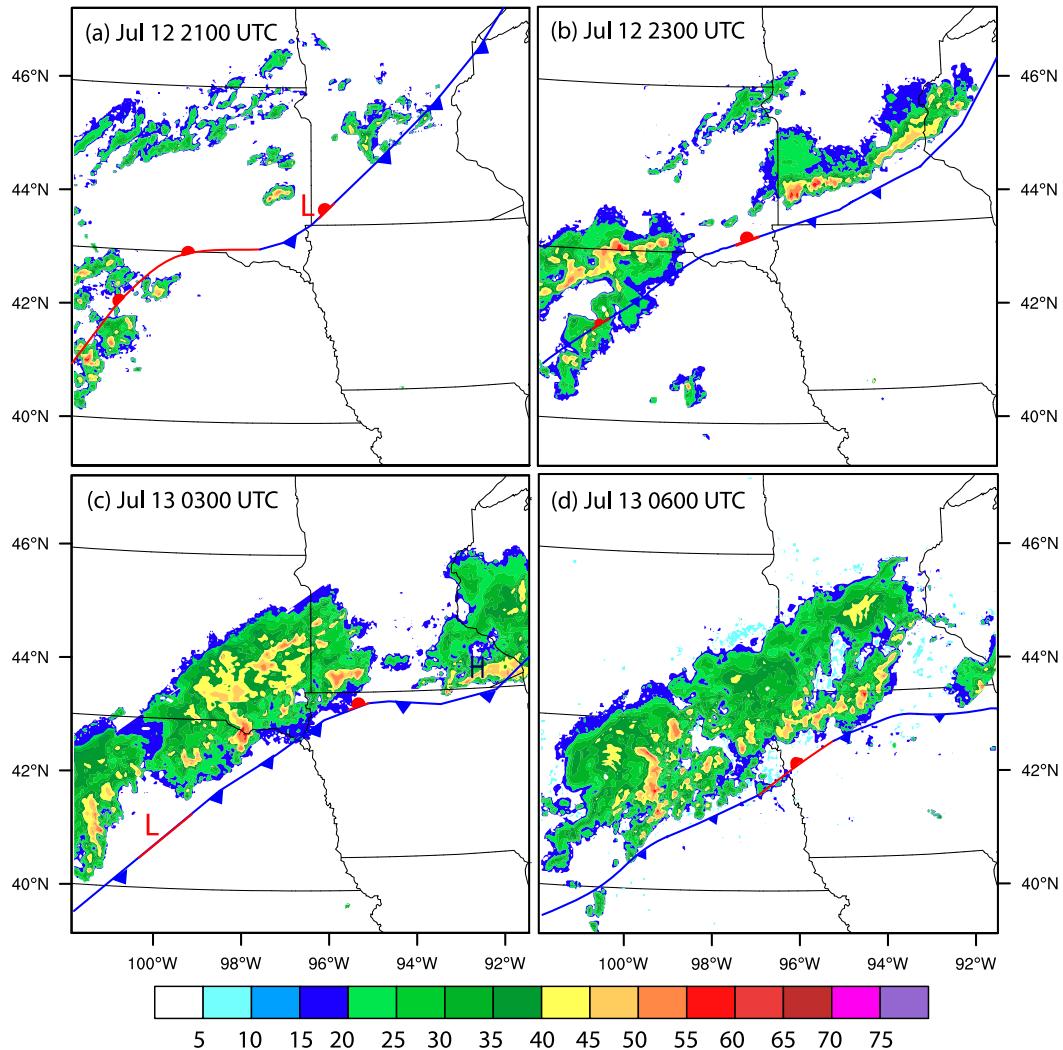


FIG. 1. Locations of the fronts (reproduced based on the surface analysis from the SPC), overlaid with the composite reflectivity observations (dBZ) remapped from the WSR-88D radars (shaded contour) at 2100 and 2300 UTC 12 Jul 0300, and 0600 UTC 13 Jul 2018, respectively.

over a volume spanning the vertical extent of the domain and covering a  $12 \times 12 \text{ km}^2$  area in horizontal dimensions centered on the model grid columns before being interpolated to the observation pixel locations; this makes the horizontal coverage of the graupel mass or graupel volume operator calculations more or less matches the  $8 \times 8 \text{ km}^2$ – $12 \times 12 \text{ km}^2$  pixel resolution of the GLM data.

### 3. The data assimilation experiments

#### a. Case overview

Scattered showers and thunderstorms developed and tracked over portions of South Dakota, and southern Minnesota on 13 July 2018. The tail-end of a cold front was the dominant forcing mechanism causing precipitation in these area (Fig. 1) with storms producing wind

gusts in excess of  $23 \text{ m s}^{-1}$  and localized flooding. At 2100 UTC 12 July, some scattered storms started to develop along the front (Fig. 1a). Two hours later at 2300 UTC, storms became more organized and developed into an MCS (Fig. 1b, the DA experiments will start at this time). The MCS spreads across southeast South Dakota, northeast Nebraska, south of Iowa, and northwest Missouri, and lasted more than 6 h (Figs. 1c,d), producing heavy rain and lightning.

#### b. GOES-R GLM FED data and their processing

The FED data used in this study are derived from the raw, 20-s Level-2 GLM data, which provide three different lightning metrics at a pixel resolution of approximately 8 km over CONUS (Goodman et al. 2013) (i.e., the flashes, groups, and events) (Fig. 2). The definition of

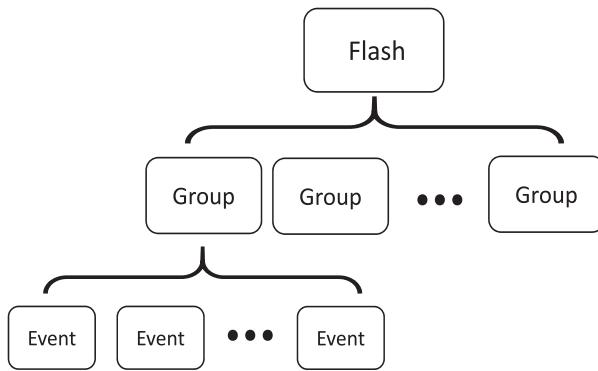


FIG. 2. GLM lightning detection product data parent-to-child relationships.

and relationship between flashes, groups, and events are governed by parent to child selection process based on fixed spatial and temporal thresholds. An event is the most elementary product and is the optical signal (photons) emitted by a lightning discharge over an individual sensor pixel for a 2-ms integration period. A group consists of an amalgamation of simultaneous events detected within adjacent pixels, and a flash combines groups that occur within about 300 ms and 16 km of each other. The readers are invited to consult Goodman et al. (2013) and Mach (2020) for more detailed information about the GLM products.

To derive the FED, the GLM data are accumulated within an assumed  $10 \times 10 \text{ km}^2$  grid box. In the GLM

data, locations of flash centroids, which are the mean constituent event latitude weighted by their energies, are also stored. If only the flash centroids are employed to derive the flash density, the resulting FED will be underestimated (Fig. 3a). Considering the parent to child relationships information (i.e., number identification or “id”) that are stored for each flash, group, and event, the FED algorithm herein tracked the event ids for different groups and the group ids for different flashes, accounting for the possibility of a single flash passing through multiple pixels. The locations of the events belonging to different flashes are used directly to count the number of flashes passing through each pixel, so that the location and rates of FED observation can be more accurate (Fig. 3b). The GLM data processing approach used in this study is consistent with the one used by Bruning et al. (2019), and GLM FED product is already part of the National Weather Service realtime display system. This fact also supports our choice of FED as the quantity to assimilate and facilitates the potential operational implementation of the DA capabilities.

### c. Forecast model setup and initialization of ensemble

The forecast model used in this study is the three-dimensional compressible non-hydrostatic WRF Model with Advanced Research dynamic core (WRF-ARW) (Skamarock et al. 2008). The data assimilation and forecast experiments use a single domain with a

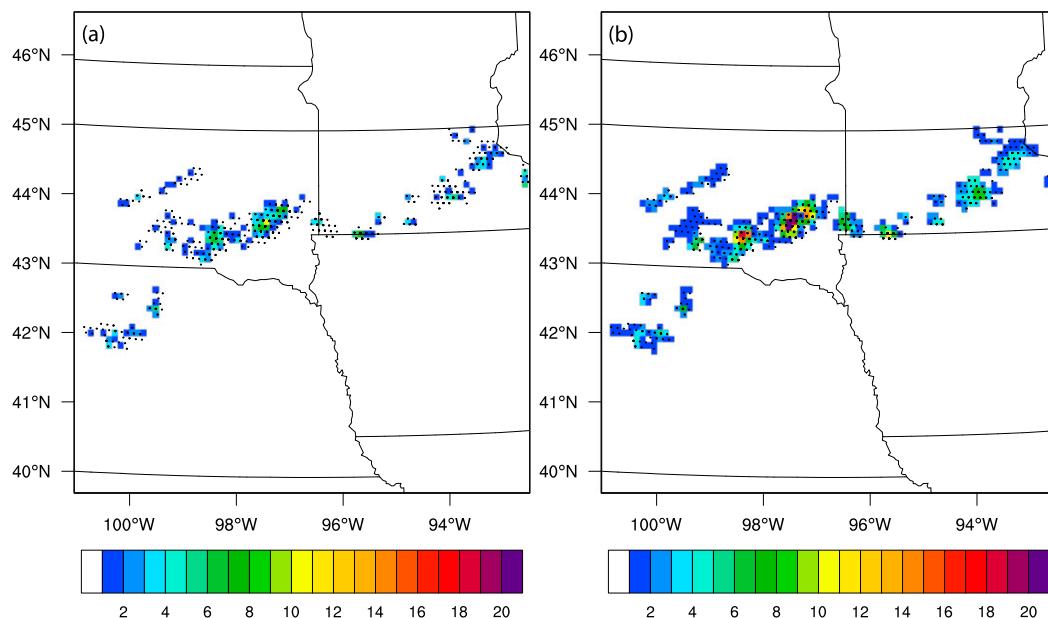


FIG. 3. The 1-min FED (units: flash  $\text{min}^{-1} \text{ pixel}^{-1}$ , shading) calculated based on (a) locations of GLM flash centroids, and (b) the locations of the events with the flash and group IDs being tracked, overlaid with the locations of the events (black dots), in the full model grid of a 10-km horizontal grid spacing.

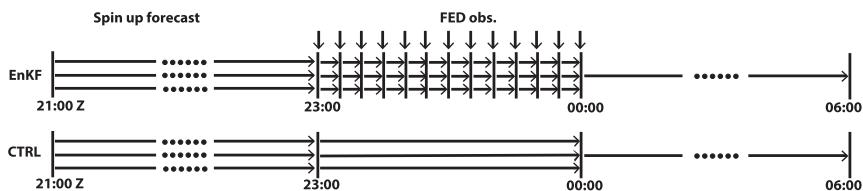


FIG. 4. Flow diagram of the (top) EnKF DA experiments vs (bottom) the control run. The spinup ensemble forecasts from 2100 to 2300 UTC include 40 members. FED DA occurs between 2300 and 0000 UTC with 5-min cycles, and deterministic forecast is launched from the final ensemble mean analysis at 0000 UTC and run to 0600 UTC. The CNTL experiment continues the ensemble forecasts through 0000 UTC without DA, when deterministic forecast continues from the ensemble mean forecast at 0000 UTC.

horizontal grid spacing of 3 km and  $300 \times 300$  horizontal grid points (Fig. 3a). The stretched vertical grid consists of 53 levels with its top set at 100 hPa ( $\sim 20.8$  km). A large time step of 15 s is used.

To perform EnKF assimilation of FED data, 2-h-long spinup ensemble forecasts are first performed between 2100 and 2300 UTC to develop evolved ensemble perturbations on the 3-km model grid (Fig. 4). The initial conditions of these forecasts are created by adding to the 3-h forecast of the 1800 UTC cycle operational North American Mesoscale Forecast System (NAM) perturbations derived from the operational Short-Range Ensemble Forecast (SREF) 2100 UTC initial conditions. SREF perturbations and their negative version are used to form 40 perturbations and they are reduced by 25% before being added to the NAM forecast at 2100 UTC. The scaling factor was found through experiments. Furthermore, because the SREF perturbations contain only wavelengths longer than 32 km (16 km is the grid spacing of SREF), additional smaller-scale perturbations of horizontal velocity, potential temperature and humidity are added to each initial condition, and the perturbations have random Gaussian distributions, and have horizontal and vertical spatial correlation scales of 20 and 3 km, respectively. Adding spatially smoothed small-scale random perturbations to the initial conditions is a common practice for starting convective-scale ensemble forecasts and data assimilation (e.g., Tong and Xue 2008; Snook et al. 2011; Johnson et al. 2014).

Similar to the configurations of the 2018 Hazardous Weather Testbed (HWT) Spring Experiment employed at the Center for the Analysis and Prediction of Storms (CAPS), the ensembles used in our EnKF DA experiments consist of 40 WRF ARW members with different physics options. The 40 WRF runs use different surface layer and planetary boundary layer (PBL) parameterization schemes. The surface layer schemes used include the revised MM5 Monin–Obukhov scheme (Jiménez and Dudhia 2012), Monin–Obukhov (Janjić Eta) scheme (Janjić 1996, 2002), and that from the

Mellor–Yamada–Nakanishi–Niino (MYNN) PBL package (Nakanishi 2001). The PBL parameterization schemes used include the Yonsei University (YSU; Hong et al. 2006), Mellor–Yamada–Janjić (MYJ) turbulent kinetic energy (TKE), the MYNN 2.5-level TKE (Nakanishi and Niino 2006), and the Asymmetrical Convective Model version 2 (ACM2; Pleim 2007) schemes. The same partially double-moment 6-class Thompson microphysical schemes (Thompson et al. 2008), with different intercept parameter (ranged from 50 000 to 4 000 000  $\text{m}^{-3}$ ) and effective density (ranged from 421 to 673  $\text{kg m}^{-3}$ ) for graupel (for which only mixing ratio is predicted), are used in different runs. The use of different physics packages and perturbing hydrometeor particle size parameters aim to include model uncertainties, which help increase ensemble spread that tends to be too small (Fujita et al. 2007; Xue et al. 2010; Snook et al. 2012).

#### d. Design of data assimilation experiments

As is mentioned in 2b, the FED observation operators based on graupel mass and graupel volume are first tuned by multiplying the right-hand side of the observation operator by different tuning factors [1, 1/2, 1/4, and 1/6 in Eqs. (14) and (15)] in a set of sensitivity experiments that assimilate the real GLM FED observations using the cycled EnKF DA configurations to be described below. These experiments are named FEDM(V)\_1Hx, FEDM(V)\_1/2Hx, FEDM(V)\_1/4Hx, and FEDM(V)\_1/6Hx, respectively, based on the operator and multiplying factor value used (Table 1). The optimal tuning factor is determined for each observation operator formulation (graupel-mass or graupel-volume-based) in terms of the equitable threat score (ETS; Mason 2003) of 0–6-h forecasts of FED calculated using the corresponding operator. The graupel-mass-based and graupel-volume-based experiments, using the optimal tuning factor, become the respective “control” experiments (FEDM and FEDV) that will be evaluated in more details. Their analyses and shorter-term forecasts are compared with those from

TABLE 1. Descriptions of the experiments.

Expt	Tuning factor in Eqs. (14) and (15)	Data being assimilated	Updating of the state variables
CTRL	N/A	No data assimilated	N/A
FEDM_1Hx	1	Assimilate lightning data using graupel-mass based observation operator	Update all the model state variables
FEDM_1/2Hx	1/2		
FEDM_1/4Hx	1/4		
FEDM_1/6Hx	1/6		
FEDM	Optimal factor (1/2)		
FEDV_1Hx	1	Assimilate lightning data using graupel-volume based observation operator	Update all the model state variables
FEDV_1/2Hx	1/2		
FEDV_1/4Hx	1/4		
FEDV_1/6Hx	1/6		
FEDV	Optimal factor (1/2)		
FEDV_OnlyQg			Only updates graupel mixing ratio
FEDV_NoTQv			Does not update temperature and water vapor mixing ratio
FEDV_NoW			Does not update vertical velocity

control experiment (CTRL) that does not assimilate any data.

As mentioned earlier, 2-h spinup ensemble forecasts are first run from 2100 to 2300 UTC 12 July, when the first EnKF analysis is performed in the DA experiments (Fig. 4). The EnKF DA cycles every 5 min are run through 0000 UTC, when the ensemble mean analysis is used to launch 6-h deterministic forecast until 0600 UTC 13 July. This allows for the assessment of the impact of FED DA on the forecast of several key storm-scale features such as accumulated rainfall and composite reflectivity. For CTRL, the ensemble forecasts proceed through 0000 UTC without assimilating any data, and deterministic forecast continues at 0000 UTC from its ensemble mean forecast. The forecast of CTRL provides a baseline for comparing the forecasts of DA experiments.

With EnKF, the FED data cannot only adjust the graupel field, which is directly related to FED through the observation operator, but also other model state variables through flow-dependent background error covariances. To assess the impact of background error cross correlations and the updating of indirectly related variables, three more experiments are conducted; they are the same as FEDV except that different model state variables are updated. The first of these experiments updates only the graupel mixing ratio (FEDV\_OnlyQg). The second and third experiments differ from FEDV in that the updating of water vapor mixing ratio and potential temperature (FEDV\_NoTQv), and of vertical velocity (FEDV\_NoW) are turned off, respectively. FEDV is chosen around which these sensitivity experiments are performed, because FEDV is found to slightly outperform FEDM. Additional details pertaining to the experiment design are listed in Table 1.

The FED observations are assimilated every 5 min over a 1-h period using GSI EnKF. Zero FED observations are assimilated also to help suppress spurious storms (Mansell et al. 2002), similar to the effect of assimilating zero reflectivity data (Tong and Xue 2005). A 0.95 adaptive posterior inflation (Whitaker and Hamill 2012), or relaxation to prior spread, is used to help maintain the ensemble spread within the DA cycles. Akin to Allen et al. (2016), the FED observations are assigned a height of 6.5 km. Also similar to Allen et al. (2016), who used 15 and 36 km for horizontal and vertical localizations, the horizontal and vertical localization radii used are 15 km and 4 in  $-\log(P/P_{\text{ref}})$  space ( $\sim 32$  km on average), respectively. Here  $P$  and  $P_{\text{ref}}$  are the pressure on model levels and reference pressure (1000 hPa), respectively. With the large vertical localization radius, the 2D GLM-FED observations effectively are allowed to influence the entire domain depth. However, the actual range of influence is determined by the ensemble-derived vertical spatial covariance, which should only be significant within the troposphere.

## 4. Results

### a. Sensitivity experiments on the FED observation operators

The sensitivity experiments with different observation operator tuning factors (Table 1) are run through the 5-min assimilation cycles over 1 h using EnKF, followed by a 6-h forecast from the final ensemble mean analysis, as described earlier. The same observation operator used in the DA is also used to calculate simulated FED from the forecasts and verified against the observed FED. ETSS of the FED forecasts after DA are examined for sensitivity

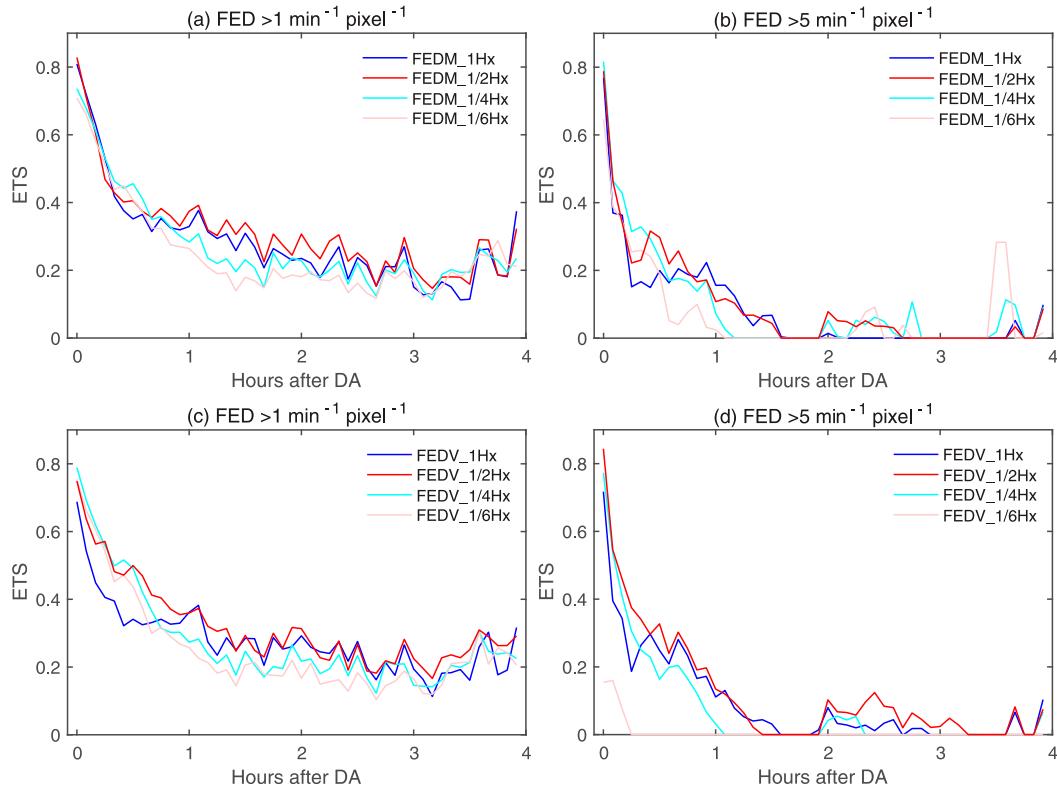


FIG. 5. ETSs of FED forecasts after DA for sensitivity experiments that multiply (a),(b) the graupel-mass-based (FEDM) and (c),(d) the graupel-volume-based (FEDV) observation operators by tuning factors of 1, 1/2, 1/4, and 1/6, for (a),(c) 1 flash  $\text{min}^{-1} \text{pixel}^{-1}$  and (b),(d) 5 flashes  $\text{min}^{-1} \text{pixel}^{-1}$  threshold, respectively.

experiments with different tuning factors in the observation operators (Fig. 5). For the threshold of 1 flash  $\text{min}^{-1} \text{pixel}^{-1}$ , FEDM\_1/2Hx performs slightly better than FEDM\_1Hx, and both are obviously better than FEDM\_1/4Hx and FEDM\_1/6Hx in terms of ETSs, especially for the 0.8–2.5-h forecasts (Fig. 5a). FEDV\_1/2Hx gives higher ETSs than FEDV\_1Hx for the first 1-h forecasts but similar (slightly better) for the rest of the forecasts (Fig. 5c); FEDV\_1/2Hx has obvious higher ETSs than FEDV\_1/4Hx and FEDV\_1/6Hx for most of the forecast time (Fig. 5c). For the threshold of 5 flashes  $\text{min}^{-1} \text{pixel}^{-1}$ , all the ETSs drop quickly within the first 1.5-h forecasts (Figs. 5b,d); FEDM\_1/2Hx performs slightly better than FEDM\_1Hx for the 0.3–0.8-h forecasts but similar for the rest of the forecasts, and both are obvious better than FEDM\_1/4Hx and FEDM\_1/6Hx (Fig. 5b). The ETSs from FEDV\_1/2Hx are higher than those of FEDV\_1Hx for FED forecasts within the first 0.4 h and 2–3.2 h but similar for the rest of the forecasts, and both are obviously higher than FEDV\_1/4Hx and FEDV\_1/6Hx (Fig. 5d). Overall FEDM(V)\_1/2Hx performs the best. Thus, 1/2 will be selected as the optimal tuning factor. Hereafter, FEDM\_1/2Hx and FEDV\_1/2Hx

are renamed as FEDM and FEDV, respectively (See Table 1).

#### b. Analyses and forecasts of FED in FEDM and FEDV

In the next a few subsections, we examine in more details the results of EnKF analyses and subsequent forecasts from experiments that use the optimal tuning parameter of 1/2 in the observation operators. During the EnKF DA cycles, the observation priors and posteriors [i.e.,  $H(\mathbf{x}^b)$  and  $H(\mathbf{x}^a)$ ] are calculated and compared to the observations. The FED forecast from CTRL and analyses from FEDM and FEDV after 1-h of DA are shown in Fig. 6. Overall, FEDM and FEDV perform similarly in capturing the intense convection regions that are absent in CTRL, although the FED rates produced by FEDM (Fig. 6c) are stronger than those of FEDV (Fig. 6d) and are more consistent with the observations for the analysis of high FED values. This is presumably because FEDV is mainly sensitive to the spatial distribution of graupel while FEDM is also sensitive to the graupel amount.

It is also noted that there exist wide areas of low magnitude FED in FEDM around the convective

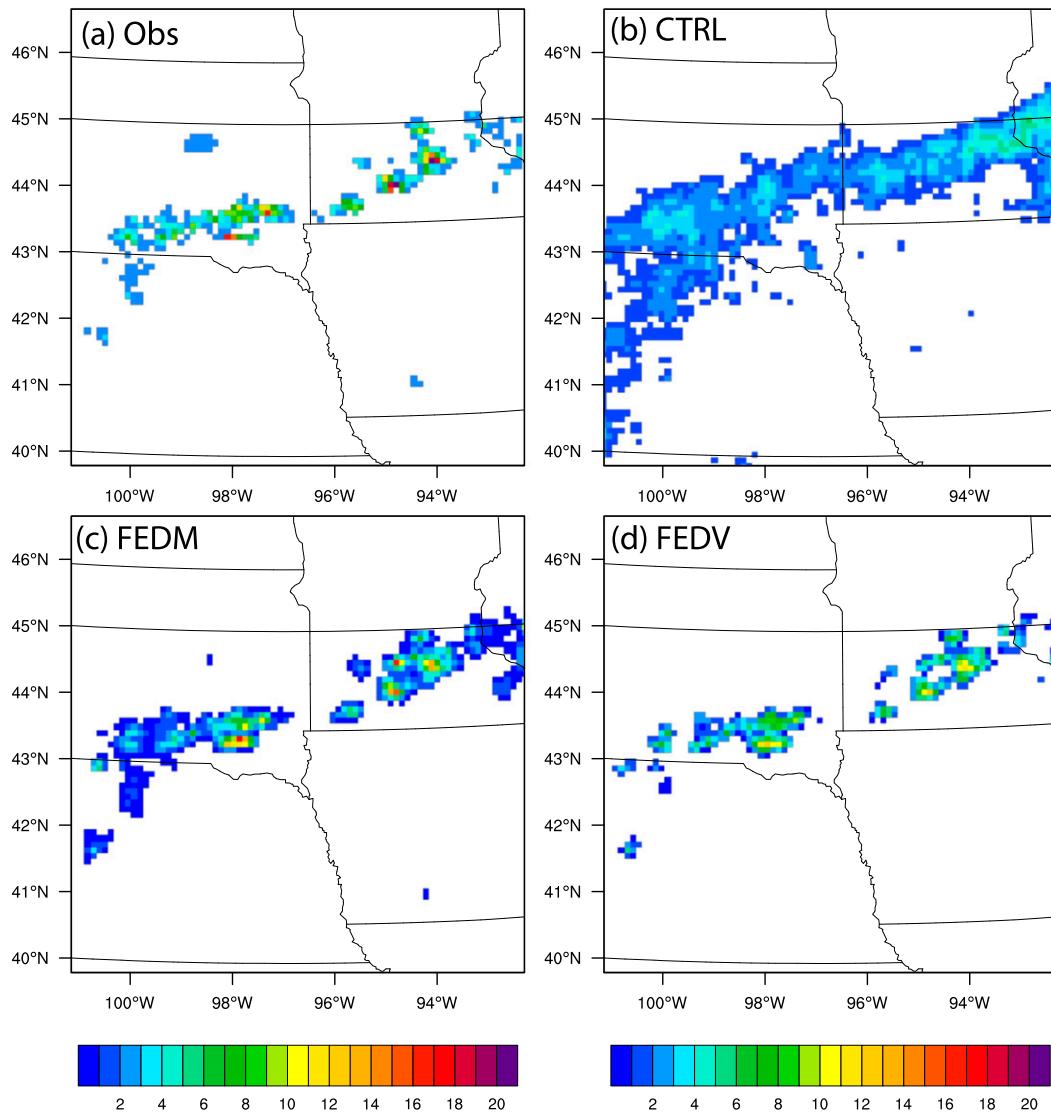


FIG. 6. Horizontal cross sections of (a) 1-min FED observations (units:  $\text{min}^{-1} \text{pixel}^{-1}$ ), (b) FED forecast from CTRL, (c) FED analyses from experiments FEDM, and (d) FEDV after a 1-h cycled DA period (valid at 0000 UTC 13 Jul).

core regions (Fig. 6c), which are not found in FEDV (Fig. 6d) or the observations (Fig. 6a). In the observation operator of FEDV, FED is no longer a function of graupel mixing ratio once the graupel mixing ratio exceeds the  $0.5 \text{ g kg}^{-1}$  threshold, and FEDV does not see graupel mixing ratios less than the threshold either. For this reason, few values of low FED are found in the analysis of FEDV (Fig. 6d). In contrast, because of the presence of the low graupel mixing ratios outside the convective cores, low values of FED are still seen in the analysis of FEDM (Fig. 6c).

A more accurate analysis in terms of the observed quantity, however, does not necessarily guarantee

forecast improvements. To examine the influence of the lightning DA on forecasts, 1–4-h forecasts of FED are compared against the observations (Fig. 7). Compared to CTRL (Figs. 7b,f,j,n), both FEDM (Figs. 7c,g,k,o) and FEDV (Figs. 7d,h,l,p) are able to better capture the intensity and distribution of lightning, although the overall FED rates are slightly underestimated and suffer from noteworthy displacement errors at larger values (i.e.,  $>10 \text{ min}^{-1} \text{ pixel}^{-1}$ ). The intensity and spatial distribution of the FED forecasts from FEDV are somewhat closer to the observations than FEDM. The FED forecasts from CTRL are notably inferior as evidenced by simulated rates that are distinctly smaller than the observations,

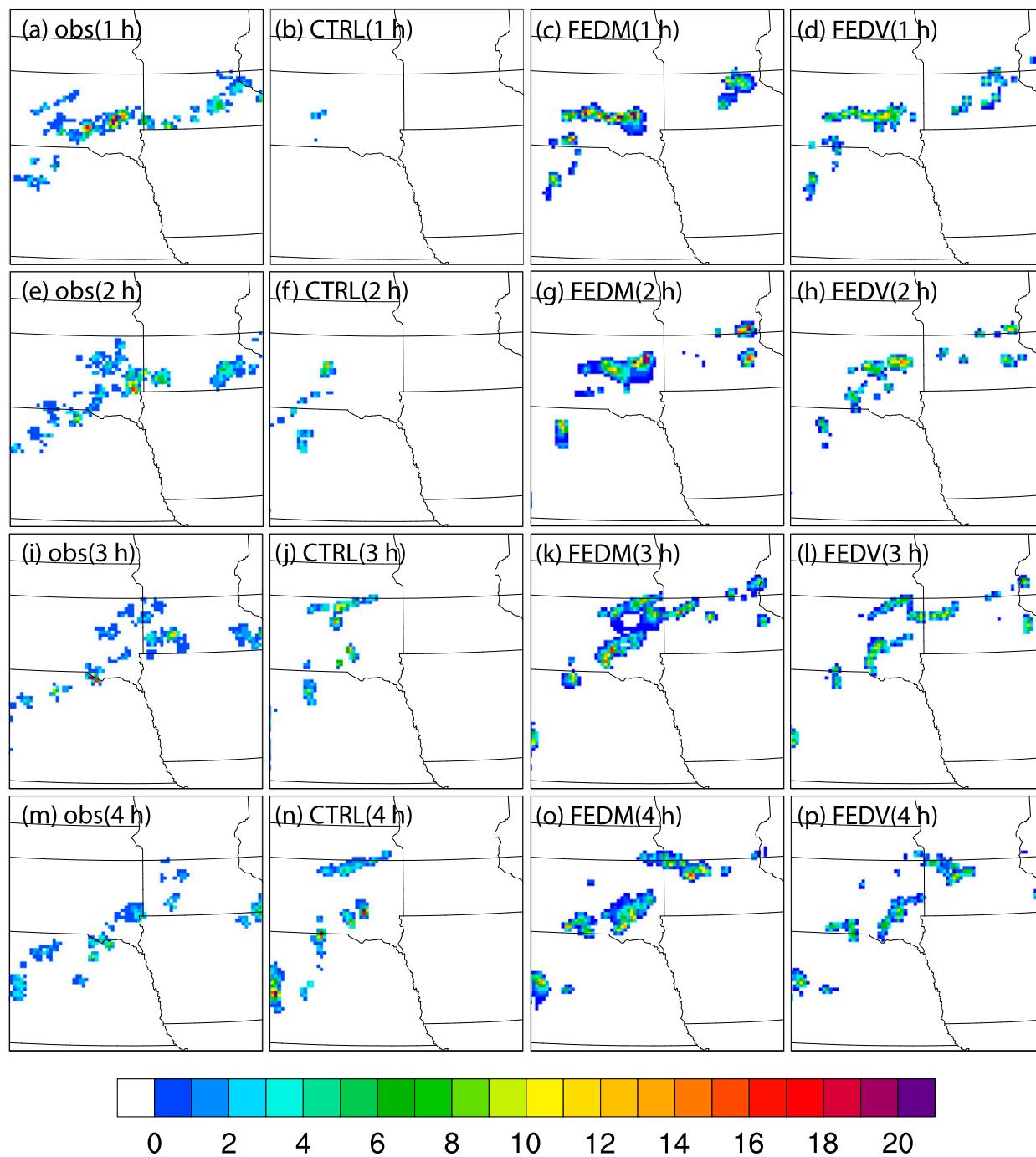


FIG. 7. Horizontal cross sections of (a),(e),(i),(m) 1-min FED observations (units: dBZ), and FED forecasts from (b),(f),(j),(n) CTRL; (c),(g),(k),(o) FEDM; and (d),(h),(l),(p) FEDV after (a)–(d) 1-, (e)–(h) 2-, (i)–(l) 3-, and (m)–(p) 4-h free forecasts following the 1-h cycled DA period (valid from 0100 to 0400 UTC).

especially within the first 2 h. The lack of convective-scale structures consistent with FED observations in CNTL should be the main reason. Additional spinup is needed when the forecast is launched from an ensemble mean (which tends to smooth out smaller-scale

structures) can be another. The 3–4-h forecasts of CNTL are better as convection in the southwestern part of the MCS is built up, but the areal coverage of the FED forecasts is much smaller than those in FEDM/FEDV and observations. Meanwhile, the FED

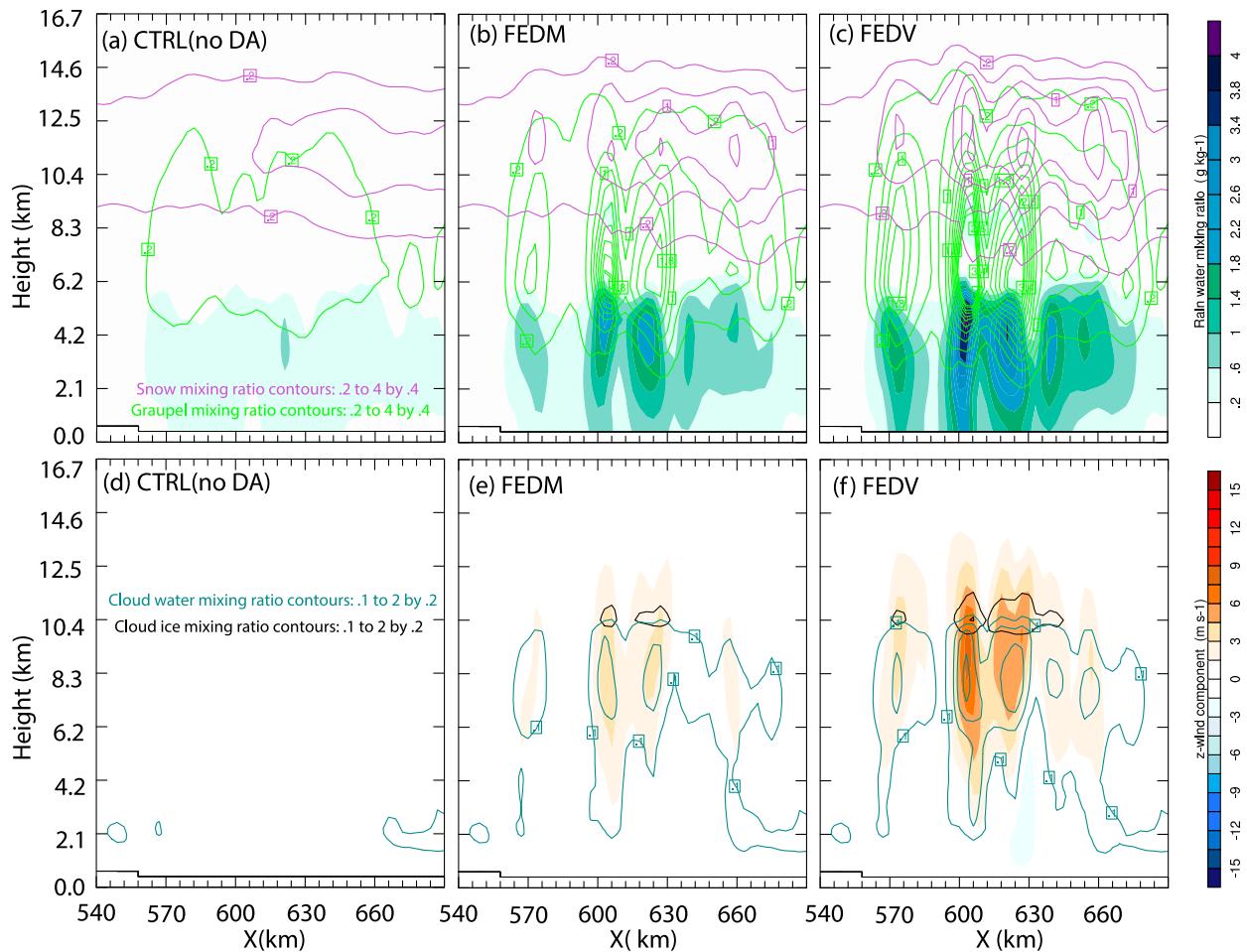


FIG. 8. Vertical cross sections of (a)–(c) the mixing ratios ( $\text{g kg}^{-1}$ ) of rain (shading), snow (purple contours), and graupel (green contours), and (d)–(f) vertical velocity ( $\text{m s}^{-1}$ , shading) and mixing ratios ( $\text{g kg}^{-1}$ ) of cloud water (blue contours) and cloud ice (black contours) after the first EnKF analysis for the three main experiments listed in Table 1 (CTRL, FEDM, and FEDV as indicated in the panels). The cross section passes through the maximum value of graupel field in the FEDM analysis at ( $x = 603 \text{ km}$ ,  $y = 558 \text{ km}$ ).

forecasts on the northeastern part of the MCS is still absent in CTRL.

### c. Analyses of state variables indirectly related to FED

In the observation operator, FED is directly related to graupel so that FED EnKF DA can directly adjust graupel mixing ratio through the observation operator connection. The FED observations also adjust other model state variables through the background error cross correlations in EnKF. To examine the direct impact of FED DA on the analyses of model state variables, vertical cross sections of the mixing ratios of rain, snow and graupel (Figs. 8a–c) and the mixing ratios of cloud water, cloud ice, and vertical velocity (Figs. 8d–f) are plotted after the first analysis at 2300 UTC for FEDM and FEDV, and compared to the fields from CTRL, which are really the background fields of the first

analysis. The vertical cross sections are chosen to pass through the maximum graupel mixing ratio value in the analysis of FEDM. By virtue of the observation operators used, a large amount of graupel is added into the model in both FEDM and FEDV (Figs. 8b,c). Coincidentally, rainwater, snow, cloud water, cloud ice, and vertical velocity (Figs. 8e,f) all increase substantially, which can only be realized within EnKF through (positive) background error correlations between FED and these state variables.

To see that positive background error cross correlations indeed exist between FED and most state variables where high FED value is found, we examine background error correlations between FED and the state variables [i.e.,  $\hat{\mathbf{P}}^b \mathbf{H}^T$  in Eq. (8)] from the background ensemble forecasts at 2300 UTC, before the first EnKF analysis is performed. The error correlations are calculated between the FED observation prior at a grid point near the

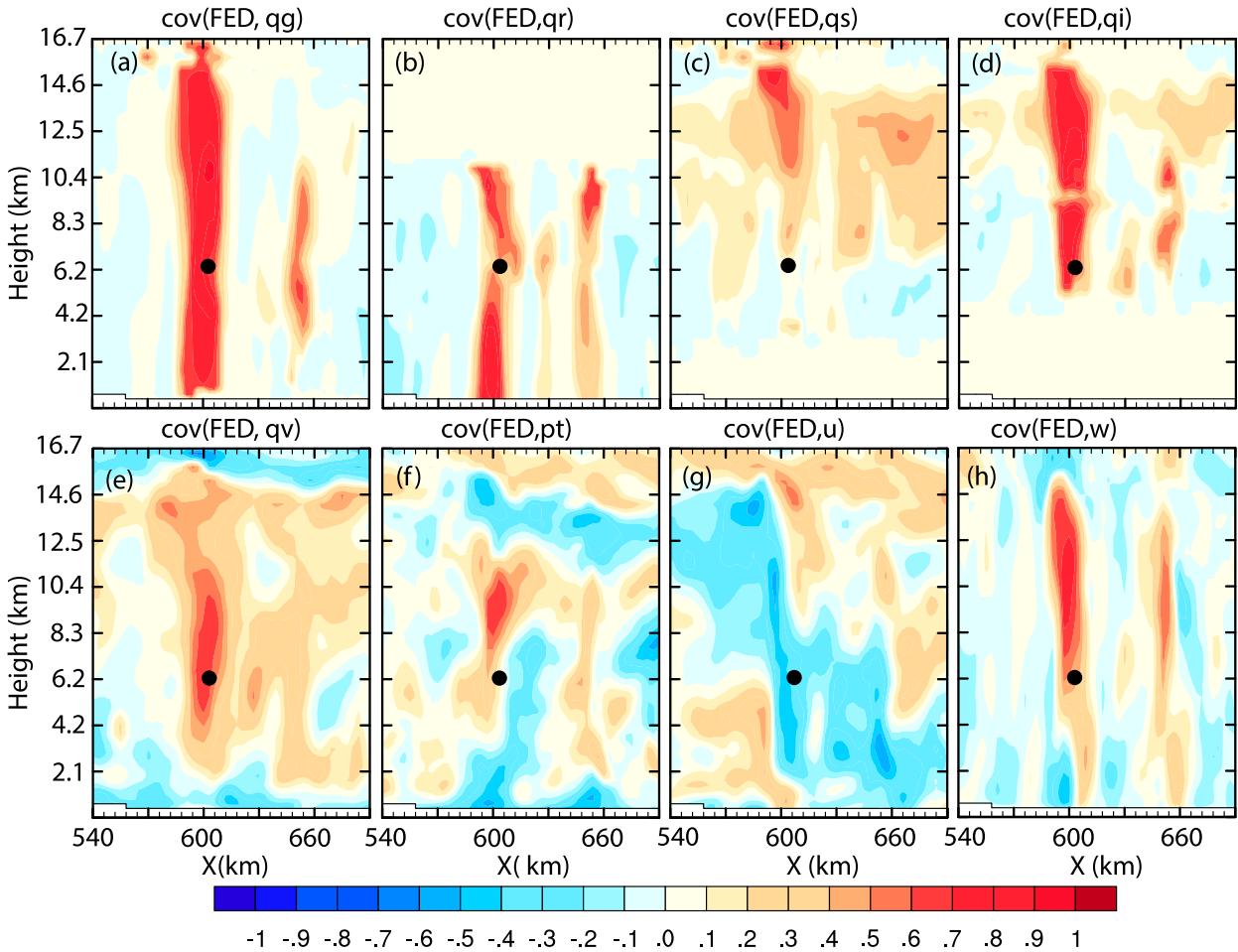


FIG. 9. Background error correlations between a simulated FED (observation prior, units:  $\text{min}^{-1} \text{pixel}^{-1}$ ) located at the black dot location and model state variables  $q_g, q_r, q_s, q_i, q_v, p, u,$  and  $w,$  calculated from ensemble forecasts valid at 2300 UTC 12 Jul for experiment FEDM in the same cross sections as those in Fig. 8, which passes through the maximum value of graupel field in the first FEDM analysis.

maximum graupel analysis increment and model state variables at each grid point. Figure 9 shows the cross-variable correlations in the same vertical cross section as the fields in Fig. 8, before covariance localization is applied. Positive correlations can be seen between the FED and almost all state variables shown in the vertical column close to the FED point, and the correlations are close to 0.95 at many locations, which will result in a positive Kalman gain according to Eq. (6). With positive Kalman gain ( $\mathbf{K}_x$ ) and positive innovation ( $\mathbf{y}^o - \bar{\mathbf{y}}^b$ ), the analysis increment of the mean state ( $\bar{\mathbf{x}}^a - \bar{\mathbf{x}}^b$ ) will also be positive based on Eq. (1). As evidenced by Fig. 8, the positive correlations shown in Fig. 9 are clearly associated with an updraft increase at the location after DA, which supports higher cloud and hydrometeor values, and positive temperature and moisture perturbations. Because of such positive correlations, the mixing ratios of rain, snow, cloud ice, water vapor as well as temperature, moisture and vertical velocity are generally adjusted

upward from the background values where graupel mixing ratio is adjusted upward. In subsequent DA cycles, the background error correlations should become more reliable when the model state estimation becomes more accurate, and the assimilation of FED data should become more effective (by producing more physical and model consistent increments; not shown). The benefits of cross correlations between radar reflectivity and model state variables within EnKF were first discussed in Tong and Xue (2005), and the cross correlations play similar roles here.

*d. Analyses and forecasts of reflectivity*

The earlier results show how the FED DA generally acts to increase the total mass of hydrometeors (and, hence reflectivity) wherever the initial background storms in the forecast domain are too weak. To evaluate how well the lightning DA captures the overall structure and intensity of the storms, the reflectivity

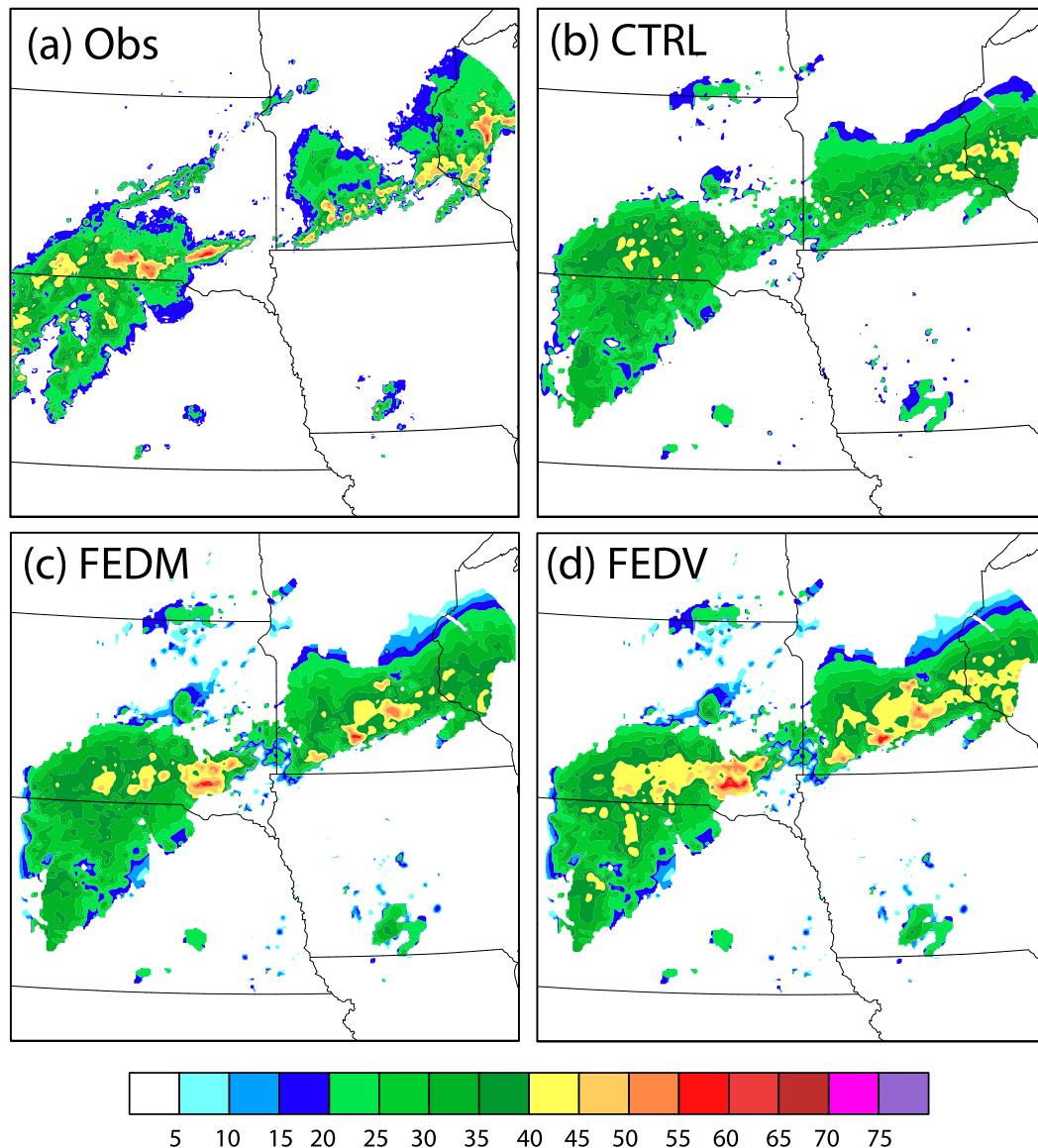


FIG. 10. Composite reflectivity fields (dBZ) from (a) the observations, (b) CTRL, (c) FEDM, and (d) FEDV after 1-h FED DA (valid at 0000 UTC).

fields after 1 h of DA and after 1–4-h forecasts are computed and compared with observations from the network of the U.S. Weather Surveillance Radar-1988 Doppler (WSR-88D), also called Next-Generation Radar (NEXRAD; Doviak et al. 2000). Figure 10 shows composite reflectivity fields at the end of DA window for the observations (Fig. 10a), CTRL (Fig. 10b), ensemble mean analyses for FEDM (Fig. 10c) and FEDV (Fig. 10d). Overall, all three experiments are able to capture the west-southwest–east-northeast-oriented precipitation band (i.e., the MCS). The areal coverage of moderate-to-large reflectivities ( $\geq 40$  dBZ) in FEDM and FEDV are similar with both being notably larger than

CTRL and, in turn, more consistent with the observations. The reflectivity is generally too weak in CTRL, lacking convective cores.

To examine the vertical distributions of key storm-scale features, vertical cross sections of reflectivity fields at the end of the DA window for all the experiments are evaluated against the observations (Fig. 11). The cross sections in Fig. 11 are made through the maximum observed reflectivity. Overall, FEDV produced stronger reflectivity cores than FEDM. The reflectivity in the FED DA experiments are notably higher than in CTRL, and closer to the observed intensity. The intense reflectivity column associated with deep convection is well

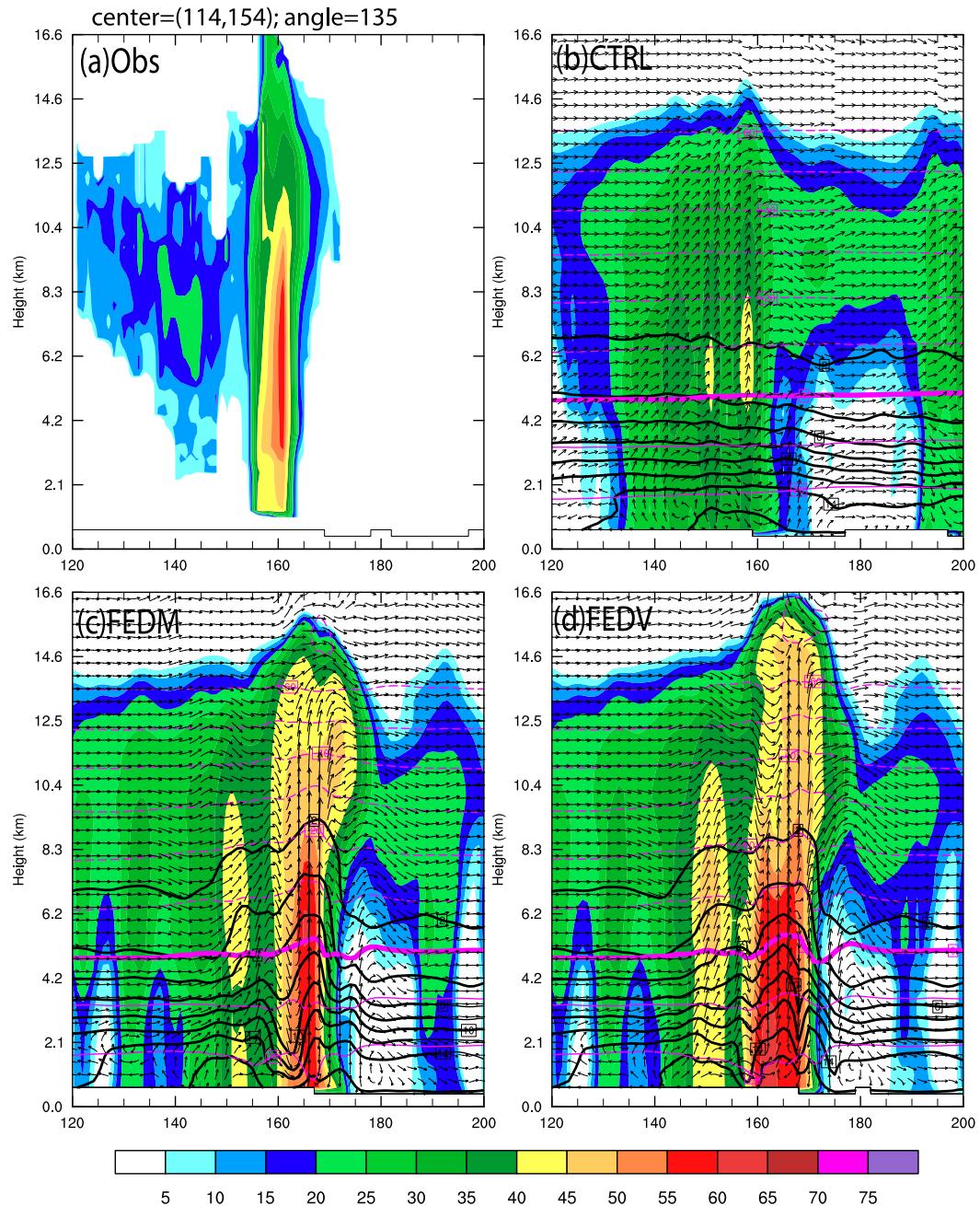


FIG. 11. Vertical cross section of reflectivity (dBZ), temperature (magenta contour, with 10°C interval), water vapor mixing ratio (black contour, with 2 g kg<sup>-1</sup> interval), and wind vector fields (m s<sup>-1</sup>) for (a) the observations, (b) CTRL, (c) FEDM, and (d) FEDV after 1-h FED DA (valid at 0000 UTC), through the maximum reflectivity value in the observations.

captured in both DA experiments. In both FED DA cases, the vertical mass flux of water vapor is larger than CTRL due to generally larger vertical velocities (stronger storms). Some displacement errors, however, must be noted between the observed and analyzed reflectivity, which, based on our examination, appears to be due to the displacement (parallax) errors of the GLM lightning

observations. Unlike the ground-based lightning mapping networks, the GLM on board the satellite must deal with parallax, which is the effect whereby the position or direction of an object appears to differ when viewed from different positions (Vicente et al. 2002). It is also noted that the analyzed storm is much broader than that of the observed one. Allen et al. (2016) also reported that the

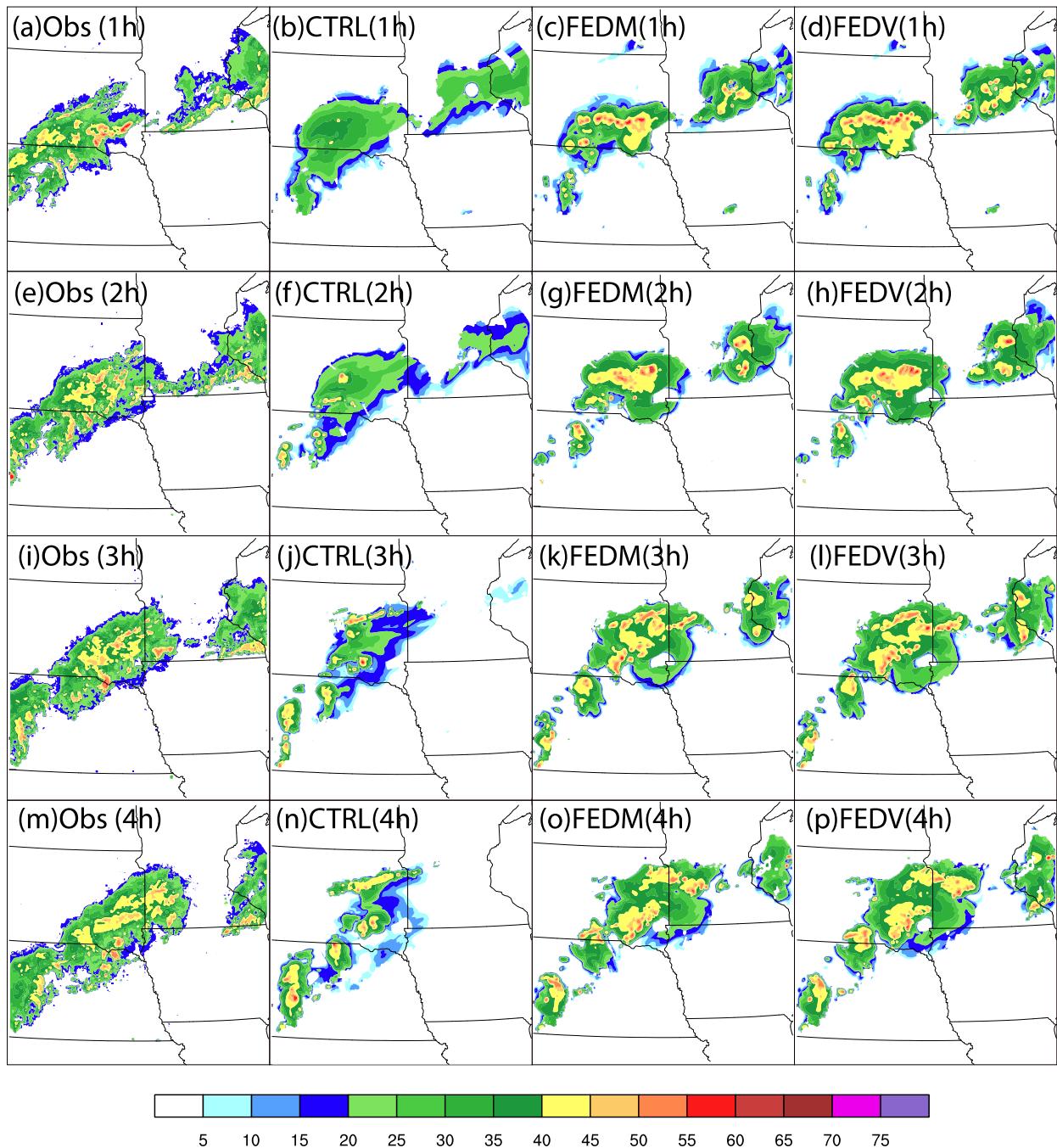


FIG. 12. Composite reflectivity fields (dBZ) for: (a),(e),(i),(m) observations, (b),(f),(j),(n) forecasts from CTRL, (c),(g),(k),(o) FEDM, and (d),(h),(l),(p) FEDV after (a)–(d) 1-, (e)–(h) 2-, (i)–(l) 3-, and (m)–(p) 4-h free forecasts following the 1-h cycled DA period (valid at 0100–0400 UTC).

areal extent of the analyzed storms was somewhat larger than that of the observed storm. Same as [Allen et al. \(2016\)](#), the spatial broadening of convection is believed to be related to low spatial resolution of the GLM data.

The 1–4-h reflectivity forecasts show that both lightning DA experiments are able to better forecast the

intensity and distribution of the precipitation band of the MCS while CTRL significantly underpredicts the intensity of the convective cores within this MCS ([Fig. 12](#)). Clearly the primary impact of the FED DA is on intense convection, where lightning occurs more systematically. It is also noted that in FEDM and

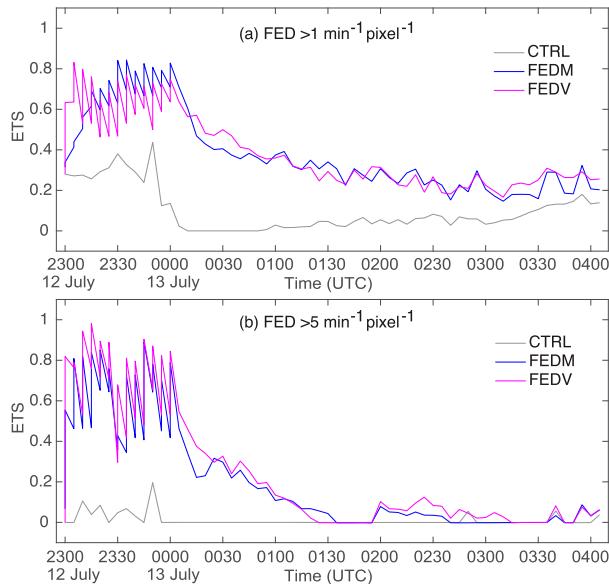


FIG. 13. ETSs of the FED forecasts for CTRL (gray curve), analyses (corresponding to higher scores) and forecasts (corresponding to lower scores) within the 1-h DA window, and 0–4-h free forecasts for FEDM (blue curve) and FEDV (magenta curve), for (a) 1 flashes  $\text{min}^{-1} \text{pixel}^{-1}$  and (b) 5 flashes  $\text{min}^{-1} \text{pixel}^{-1}$  FED threshold.

FEDV, the leading edge of the MCS in the final analysis tends to weaken in the forecasts. Although EnKF tries to produce analyses of model state variables that are consistent with the model through ensemble-derived background error covariance, the error covariance is not always reliable (due to small ensemble size, model and ensemble-mean state errors), and thus rapid adjustments tend to occur immediately in the forecast, and small-scale errors tend to grow very fast. Overall, positive impact of FED DA on the forecast is still evident.

#### e. Objective verification

To assess the performance of the experiments in a quantitative manner, the gridpoint-based ETS for the FED forecasts, and analyses during the DA cycles, are computed for CTRL, FEDM, and FEDV, for the FED thresholds of 1 and 5 flashes  $\text{min}^{-1} \text{pixel}^{-1}$ , respectively. The ETS of the FED analyses from both FEDM and FEDV noticeably increased relative to the corresponding background forecasts for both thresholds (Fig. 13). For the 1 flash  $\text{min}^{-1} \text{pixel}^{-1}$  threshold, the two DA experiments clearly outperform CTRL. FEDV slightly outperforms FEDM within the first 20-min DA window, but underperforms FEDM for the rest of the DA window. During 10–50 min of forecast, however, FEDV produces slightly larger ETSs than FEDM but becomes similar to FEDM for the rest of the FED forecasts (Fig. 13a). The ETSs for the 5 flashes  $\text{min}^{-1} \text{pixel}^{-1}$

threshold are almost 0 in CTRL because it fails to produce intense lightning convection. At this higher threshold, FEDV slightly outperforms FEDM in terms of ETS in both FED analyses and forecasts (Fig. 13b). Overall, the performances of the two forms of graupel-based observation operators for FED can be considered as equivalent.

One salient limitation of the pointwise ETS is the double penalty imposed on the scores when storm-scale features exhibit small displacement errors/biases. An alternative neighborhood-based forecast evaluation metric alleviating this drawback—the fractions skill score (FSS; Roberts and Lean 2008)—is also used to evaluate the 1–4-h forecasts of hourly accumulated precipitation against precipitation estimates from the National Severe Storms Laboratory Multi-Radar Multi-Sensor (MRMS) product (Zhang et al. 2016) (Fig. 14). For stratiform precipitation at the 5  $\text{mm h}^{-1}$  threshold, FEDV performs slightly better than FEDM and both are more skillful than CTRL for up to 4 h of precipitation forecast (Fig. 14a). For convective precipitation at the 20  $\text{mm h}^{-1}$  threshold, FEDM and FEDV are more skillful than CTRL until about 3 h, after which the skill decreases sharply for all cases (Fig. 14b). The FSS exhibits a higher degree of sensitivity to the neighborhood radius chosen for convective precipitation (20  $\text{mm h}^{-1}$ , Fig. 14b) than for stratiform precipitation (5  $\text{mm h}^{-1}$ , Fig. 14a), which is not surprising given the small sizes of convective precipitation regions and hence higher sensitivity to displacement error.

#### f. Sensitivity tests on state variable updating in EnKF

Sensitivity tests are conducted to assess the impacts of background error cross correlations and the updating of the state variables that are indirectly related to the FED observations on storm analysis and forecast. Based on the ETS of FED analyses and forecasts shown in Fig. 13, FEDV overall has higher scores than FEDM, especially in the forecasts. For this reason, we perform our sensitivity experiments based on FEDV experiment. When only updating the graupel field that is directly related to the FED observation in experiment FEDV\_OnlyQg, the ETSs for the 1–4-h free forecasts of FED are much lower than those of the original FEDV experiment that update all state variables (Fig. 15). For the 1 flash  $\text{min}^{-1} \text{pixel}^{-1}$  threshold, the difference lasts throughout the 4 h of forecast, although the difference is larger in the first 2.5 h. For the highest threshold of 5 flashes  $\text{min}^{-1} \text{pixel}^{-1}$ , the difference is mostly in the first 1.5 h. The most noticeable differences occur in the first 20 min of forecast when the ETSs of FEDV\_OnlyQg drops quickly toward zero before recovering, for both thresholds (Fig. 15), while those of FEDV drop much more slowly to levels around

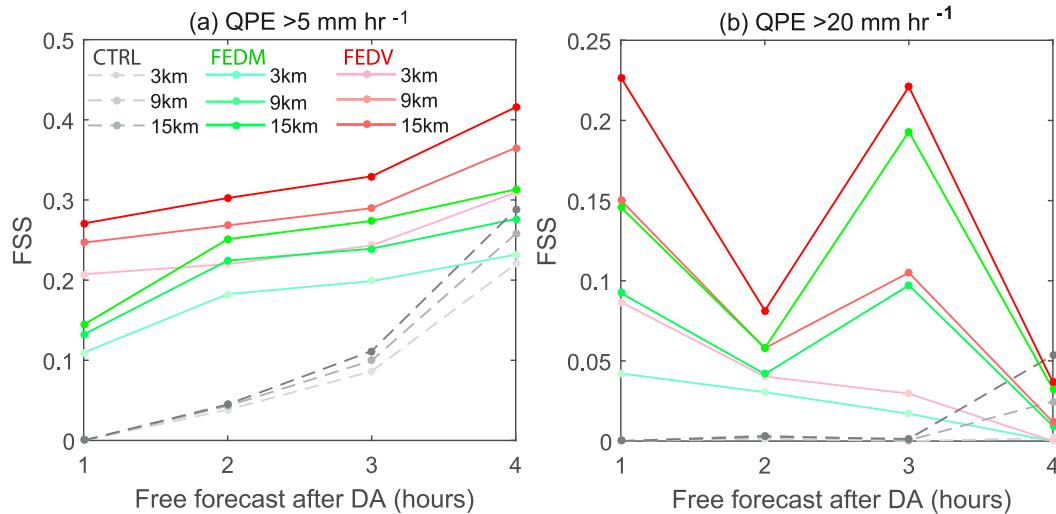


FIG. 14. FSSs for the 1–4-h forecasts of hourly accumulated precipitation for CTRL (gray), FEDM (green), and FEDV (red), relative to the MRMS hourly rainfall estimates for thresholds of (a)  $5 \text{ mm h}^{-1}$  and (b)  $20 \text{ mm h}^{-1}$  using 3-, 9-, and 15-km neighborhood radii. Legends for colors are shown in (a).

0.3. Faster drops in the scores are also seen within the DA cycles. This clearly suggests that even though FEDV\_OnlyQg is able to analyze FED fields that are close to observed FED due to direct link between graupel mixing ratio and FED, other state variables are analyzed poorly, and their errors cause the forecast to deteriorate quickly. In FEDV, other state variables are consistently updated within the DA process, resulting much slower error growth in the forecast.

Experiment FEDV\_NoTQv, which withholds the updating of water vapor and potential temperature fields, performs slightly better than FEDV\_OnlyQg but exhibits much lower scores than FEDV. This indicates the critical importance of the temperature and moisture fields since that are what support the moist convection. Not allowing the filter to update the vertical velocity in experiment FEDV\_NoW does not have much influence on the FED forecasts relative to FEDV, suggesting that the vertical velocity may be able to respond quickly within the model forecast to accurate analysis of thermodynamic conditions including temperature and moisture perturbations within the storms. In summary, the adjustments to model state variables that do not directly appear in the FED observation operator, especially to the temperature and moisture fields, play an important role sustaining the benefits of FED assimilation, and in improving the overall analysis and forecast of an MCS in this study in terms of simulated FED. Such adjustments are made possible with the flow-dependent error covariance calculated from the forecast ensemble.

We also compared the FSS scores of the sensitivity experiments for 3-, 9-, and 15-km neighborhood radii

with those of CTRL and FEDM, the general behaviors are very similar to those of ETS; they are therefore not shown.

## 5. Summary and conclusions

In this study, the GSI-based EnKF system is enhanced to include the ability to assimilate GOES-R GLM data into a convection-allowing NWP model. Real GOES-R

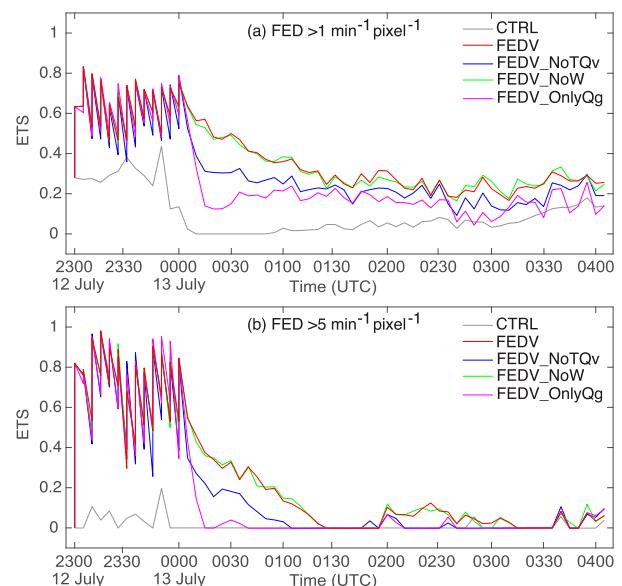


FIG. 15. As in Fig. 13, but for the original FEDV experiment (labeled FEDV), sensitivity experiments FEDV\_OnlyQg, FEDV\_NoTQv, and FEDV\_NoW, for (a) the 1 flashes  $\text{min}^{-1} \text{ pixel}^{-1}$  and (b) 5 flashes  $\text{min}^{-1} \text{ pixel}^{-1}$  FED thresholds.

GLM FED data are assimilated for an MCS case that occurred over the central U.S. plains, and the impact of assimilating such data is evaluated. Two versions of the FED observation operator are implemented and tested. They are the observation operators from Allen et al. (2016) that are based on the graupel mass and graupel volume, respectively. To account for possible biases in the original observation operators, sensitivity experiments are conducted by multiplying the right-hand sides of two versions of the observation operators by different tuning factors. The optimal tuning factors are obtained in terms of ETS in 0–6-h forecasts of the FED field after DA. The two primary sets of DA experiments using the two versions of the operator and the corresponding optimal tuning factors are labeled FEDM and FEDV, respectively. Additional sensitivity experiments are conducted, in which different model state variables are updated by the FED observations. These experiments are to examine the impact of updating state variables that are not directly connected to FED via observation operator, but through ensemble background error correlation only. To the best of our knowledge, this study represents that first time that real GOES GLM data are directly assimilated (using EnKF method).

The raw, 20-s Level-2 GLM data, including the lightning flashes, groups and events data, are processed to form FED data every minute. The 1-min FED observations are assimilated every 5 min for a 1-h period using EnKF. Experiments FEDM and FEDV are first evaluated, by comparing with the results of a control run (CTRL) not assimilating any data. The comparisons are made based on analyses and forecasts of FED, composite and 3D reflectivity fields, and vertical velocities. Objective verifications are also made using gridpoint-based ETS and neighborhood-based FSS for forecasts of hourly FED and rainfall. The main findings are summarized as follows:

- Neighborhood ETSs of 0–6-h FED forecasts are calculated for experiments FEDM(V)\_1Hx, FEDM(V)\_1/2Hx, FEDM(V)\_1/4Hx, and FEDM(V)\_1/6Hx that multiply the right-hand side of their corresponding observation operator by a different tuning factor (1, 1/2, 1/4, and 1/6), respectively. FEDM(V)\_1/2Hx outperforms experiments with other tuning factors in terms of the ETSs of FED forecasts. Thus, 1/2 is determined to be the optimal tuning factor in both FEDM and FEDV and used in the rest of experiments and evaluations.
- In terms of FED, experiments FEDM and FEDV perform similarly, and both outperform CTRL as evidenced by higher ETS scores of FED analyses and short-term forecasts (<3 h). FEDM and FEDV better capture the higher lightning density rate areas in both the analysis and forecast fields while CTRL mostly misses them.
- By virtue of the graupel-based observation operators used during the assimilation of GLM FED data, a large amount of graupel is added into the model by the analysis at the time of first analysis (given that the initial background values were too low). Positive ensemble-derived background error cross correlations between FED (i.e., graupel mixing ratio) at a main updraft location and other hydrometeor species such as rainwater, snow, cloud water, and cloud ice result in their upward adjustment in the first DA cycle when the background convection was too weak. Additionally, positive correlations between FED/graupel mass and vertical motion lead to more intense convective updrafts at observed lightning locations and overall stronger convective storms.
- The analyzed reflectivity at the end of the 1-h DA window and forecast reflectivity at up to 4 h of forecast time are compared with observed reflectivity. Both FEDM and FEDV DA experiments and CTRL are able to capture the overall west-southwest–east-northeast-oriented MCS in the analyses. The areal coverage of moderate-to-high reflectivity (i.e.,  $\geq 40$  dBZ) produced by FEDM and FEDV are, however, notably larger than in CTRL, and match observations better. The analyzed vertical velocities in the FED DA experiments (and, thus graupel mass fluxes) are overall stronger than those in CTRL, leading to more lightning-active convection. The reflectivity forecasts of FEDM and FEDV also better capture the intensity and distribution of the observed MCS relative to CTRL. Overall, the most noticeable direct impact of the FED DA is on lightning-producing convective cores.
- The 1–4-h forecast of hourly accumulated precipitation from the three main experiments are verified against precipitation estimates from the Multi-Radar Multi-Sensor (MRMS) using the standard gridpoint-based ETS formulation and the neighborhood-based fractions skill scores for light and heavy rainfall thresholds corresponding to stratiform and convection precipitation, respectively. Overall, FEDM and FEDV perform similarly and are more skillful than CTRL up to 4 h into the forecasts.
- Sensitivity experiments updating different model state variables show that the adjustments to state variables are that not directly connected to FED through observation operator plays a very important role producing consistent analysis among the model fields, enabling more sustained impact of assimilating FED data. The updating of temperature and moisture fields are found to be especially important. If they

are not updated by the filter, the forecast error grows very fast.

In summary, the FED DA experiments using graupel mass and graupel volume-based observation operators perform similarly and notably better than CTRL without DA, in terms of the accuracy of the analyses and short-term forecasts of FED, composite/3D reflectivity fields, and the expected locations of updraft cores. The FED DA using the graupel-based observation operators from Allen et al. (2016), with a proper tuning factor multiplied, is shown to be quite effective in initializing regions of intense convection for this case and, thus, helps improve short-term (up to 4 h) forecast of hazards associated with deep convection such as convective rainfall and high FED rates. EnKF adjustments to graupel and other model state variables (such as temperature, moisture and vertical velocity) through ensemble covariances generate forecast-model-consistent analyses and forecasts with higher skill scores for this case.

Given the still infancy stage of lightning DA using ensemble methods, in particular with space-borne lightning data, future studies are clearly needed, with emphasis on building statistics with additional cases with particular attention given to data-sparse areas such as oceanic and/or mountainous regions.

Given that the observation operators tested were originally derived statistically from model simulations with electrification rather than based on true observations or direct physics modeling, there still exist large uncertainties with the operators, and further calibrations are likely needed. In the case of large optical depth above the lightning flash source level, there may be significant extinction to the GLM observed flashes, hence underestimating the FED. Building light extinction into the FED observation operator is a possible solution to this problem, similar to accounting for signal attenuation in the reflectivity observation operator of short-wavelength radars (Xue et al. 2009). In addition, even though FEDM and FEDV are found to perform similarly in this study, further tests with more cases are needed to help us make the choice for operational implementation. In future studies, we also plan to refine the operation operator.

It is hoped that through further studies, the current observation operators for GLM FED can be improved and better adapted for broader use in terms of the range of convective regimes and geographical regions considered. The EnKF DA system can also be tuned further in terms of covariance localization and inflation. Akin to Fierro et al. (2016), available Doppler weather radar data should also be assimilated in tandem with lightning data to better gauge the added value of 2D lightning

data with respect to more exhaustive 3D datasets. Other available observations, such as high-resolution surface observations, should also be assimilated together, so that the storm environment can also be improved. Additionally, performance of assimilating lightning observations with more advanced DA scheme (such as the hybrid ensemble variational DA scheme) is also worth investigation in further research. In fact, we are developing hybrid DA capabilities for FED data within the GSI framework, and results will be reported in future papers.

*Acknowledgments.* This work was primarily supported by NOAA GOES-R Program funding via Grant NA16OAR4320115. The DA experiments were conducted on the NSF Xsede supercomputing facility at Texas Advanced Computing Center. Auxiliary computer resources were provided by the Oklahoma Supercomputing Center for Education and Research (OSCCER) hosted at the University of Oklahoma.

#### REFERENCES

- Aksoy, A., D. C. Dowell, and C. Snyder, 2009: A multicaser comparative assessment of the ensemble Kalman filter for assimilation of radar observations. Part I: Storm-scale analyses. *Mon. Wea. Rev.*, **137**, 1805–1824, <https://doi.org/10.1175/2008MWR2691.1>.
- , —, and —, 2010: A multicaser 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>.
- Allen, B. J., E. R. Mansell, D. C. Dowell, and W. Deierling, 2016: Assimilation of pseudo-GLM data using the ensemble Kalman filter. *Mon. Wea. Rev.*, **144**, 3465–3486, <https://doi.org/10.1175/MWR-D-16-0117.1>.
- Anderson, J. L., and N. Collins, 2007: Scalable implementations of ensemble filter algorithms for data assimilation. *J. Atmos. Oceanic Technol.*, **24**, 1452–1463, <https://doi.org/10.1175/JTECH2049.1>.
- Bruning, E. C., and Coauthors, 2019: Meteorological imagery for the Geostationary Lightning Mapper. *J. Geophys. Res. Atmos.*, **124**, 14 285–14 309, <https://doi.org/10.1029/2019JD030874>.
- Carey, L. D., and S. A. Rutledge, 1998: Electrical and multiparameter radar observations of a severe hailstorm. *J. Geophys. Res.*, **103**, 13 979–14 000, <https://doi.org/10.1029/97JD02626>.
- , N. Curtis, S. M. Stough, C. J. Schultz, and P. M. Bitzer, 2019: A radar investigation of storm morphology and precipitation processes during discrepancies between GOES-16 GLM and LMA observed lightning flash rates and jumps. *15th Annual Symp. on New Generation Operational Environmental Satellite Systems*, Phoenix, AZ, Amer. Meteor. Soc., 1018, <https://ams.confex.com/ams/2019Annual/webprogram/Paper352786.html>.
- Christian, H. J., R. J. Blakeslee, S. J. Goodman, and D. M. Mach, 2000: Algorithm Theoretical Basis Document (ATBD) for the Lightning Imaging Sensor (LIS). NASA-MSFC Tech. Rep., 53 pp.
- Clark, A. J., W. A. Gallus, M. Xue, and F. Kong, 2010: Convection-allowing and convection-parameterizing ensemble forecasts

- of a mesoscale convective vortex and associated severe weather environment. *Wea. Forecasting*, **25**, 1052–1081, <https://doi.org/10.1175/2010WAF2222390.1>.
- Deierling, W., and W. A. Petersen, 2008: Total lightning activity as an indicator of updraft characteristics. *J. Geophys. Res.*, **113**, D16210, <https://doi.org/10.1029/2007JD009598>.
- Dixon, K., C. F. Mass, G. J. Hakim, and R. H. Holzworth, 2016: The impact of lightning data assimilation on deterministic and ensemble forecasts of convective events. *J. Atmos. Oceanic Technol.*, **33**, 1801–1823, <https://doi.org/10.1175/JTECH-D-15-0188.1>.
- Doviak, R. J., V. Bringi, A. Ryzhkov, A. Zahrai, and D. Zrnic, 2000: Considerations for polarimetric upgrades to operational WSR-88D radars. *J. Atmos. Oceanic Technol.*, **17**, 257–278, [https://doi.org/10.1175/1520-0426\(2000\)017<0257:CFPUTO>2.0.CO;2](https://doi.org/10.1175/1520-0426(2000)017<0257:CFPUTO>2.0.CO;2).
- Dowell, D. C., L. J. Wicker, and C. Snyder, 2011: Ensemble Kalman filter assimilation of radar observations of the 8 May 2003 Oklahoma City supercell: Influences of reflectivity observations on storm-scale analyses. *Mon. Wea. Rev.*, **139**, 272–294, <https://doi.org/10.1175/2010MWR3438.1>.
- Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte-Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99**, 10 143–10 162, <https://doi.org/10.1029/94JC00572>.
- Fierro, A. O., and E. R. Mansell, 2018: Relationships between electrification and storm-scale properties based on idealized simulations of an intensifying hurricane-like vortex. *J. Atmos. Sci.*, **75**, 657–674, <https://doi.org/10.1175/JAS-D-17-0202.1>.
- , M. S. Gilmore, E. R. Mansell, L. J. Wicker, and J. M. Straka, 2006: Electrification and lightning in an idealized boundary-crossing supercell simulation of 2 June 1995. *Mon. Wea. Rev.*, **134**, 3149–3172, <https://doi.org/10.1175/MWR3231.1>.
- , E. R. Mansell, C. L. Ziegler, and D. R. MacGorman, 2012: Application of a lightning data assimilation technique in the WRF-ARW model at cloud-resolving scales for the tornado outbreak of 24 May 2011. *Mon. Wea. Rev.*, **140**, 2609–2627, <https://doi.org/10.1175/MWR-D-11-00299.1>.
- , J. Gao, C. L. Ziegler, E. R. Mansell, D. R. MacGorman, and S. R. Dembek, 2014: Evaluation of a cloud-scale lightning data assimilation technique and a 3DVAR method for the analysis and short-term forecast of the 29 June 2012 derecho event. *Mon. Wea. Rev.*, **142**, 183–202, <https://doi.org/10.1175/MWR-D-13-00142.1>.
- , A. J. Clark, E. R. Mansell, D. R. MacGorman, S. R. Dembek, and C. L. Ziegler, 2015a: 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>.
- , E. R. Mansell, C. L. Ziegler, and D. R. MacGorman, 2015b: Explicitly simulated electrification and lightning within a tropical cyclone based on the environment of Hurricane Isaac (2012). *J. Atmos. Sci.*, **72**, 4167–4193, <https://doi.org/10.1175/JAS-D-14-0374.1>.
- , J. Gao, C. L. Ziegler, K. M. Calhoun, E. R. Mansell, and D. R. MacGorman, 2016: Assimilation of flash extent data in the variational framework at convection-allowing scales: Proof-of-concept and evaluation for the short-term forecast of the 24 May 2011 tornado outbreak. *Mon. Wea. Rev.*, **144**, 4373–4393, <https://doi.org/10.1175/MWR-D-16-0053.1>.
- , Y. Wang, J. Gao, and E. R. Mansell, 2019: Variational assimilation of radar data and GLM lightning-derived water vapor for the short-term forecasts of high-impact convective events. *Mon. Wea. Rev.*, **147**, 4045–4069, <https://doi.org/10.1175/MWR-D-18-0421.1>.
- 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>.
- Gao, J., M. Xue, K. Brewster, and K. K. Droegemeier, 2004: A three-dimensional variational data analysis method with recursive filter for Doppler radars. *J. Atmos. Oceanic Technol.*, **21**, 457–469, [https://doi.org/10.1175/1520-0426\(2004\)021<0457:ATVDAM>2.0.CO;2](https://doi.org/10.1175/1520-0426(2004)021<0457:ATVDAM>2.0.CO;2).
- , —, and D. J. Stensrud, 2013: The development of a hybrid EnKF-3DVAR algorithm for storm-scale data assimilation. *Adv. Meteor.*, **2013**, 1–12, <https://doi.org/10.1155/2013/512656>.
- 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>.
- Goodman, S. J., D. E. Buechler, P. D. Wright, and W. D. Rust, 1988: Lightning and precipitation history of a microburst-producing storm. *Geophys. Res. Lett.*, **15**, 1185–1188, <https://doi.org/10.1029/GL015i011p01185>.
- , and Coauthors, 2013: The GOES-R Geostationary Lightning Mapper (GLM). *Atmos. Res.*, **125–126**, 34–49, <https://doi.org/10.1016/j.atmosres.2013.01.006>.
- Hong, S., Y. Noh, and J. Dudhia, 2006: A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon. Wea. Rev.*, **134**, 2318–2341, <https://doi.org/10.1175/MWR3199.1>.
- Hu, J., A. O. Fierro, Y. Wang, J. Gao, and E. R. Mansell, 2020: Exploring the assimilation of GLM derived water vapor mass in a cycled 3DVAR framework for the short-term forecasts of high impact convective events. *Mon. Wea. Rev.*, **148**, 1005–1028, <https://doi.org/10.1175/MWR-D-19-0198.1>.
- Janjić, Z. I., 1996: The surface layer parameterization in the NCEP Eta Model. *Research Activities in Atmospheric and Oceanic Modeling*, H. Ritchie, Ed., WMO, 4.16–14.17.
- , 2002: Nonsingular implementation of the Mellor–Yamada level 2.5 scheme in the NCEP Meso model. NCEP Office Note 437, 61 pp., <http://www.emc.ncep.noaa.gov/officenotes/newernotes/on437.pdf>.
- Jiménez, P. A., and J. Dudhia, 2012: Improving the representation of resolved and unresolved topographic effects on surface wind in the WRF model. *J. Appl. Meteor. Climatol.*, **51**, 300–316, <https://doi.org/10.1175/JAMC-D-11-084.1>.
- Johnson, A., 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>.
- Jung, Y., M. Xue, G. Zhang, and J. M. Straka, 2008: Assimilation of simulated polarimetric radar data for a convective storm using the ensemble Kalman filter. Part II: Impact of polarimetric data on storm analysis. *Mon. Wea. Rev.*, **136**, 2246–2260, <https://doi.org/10.1175/2007MWR2288.1>.
- Kain, J. S., and Coauthors, 2010: Assessing advances in the assimilation of radar data and other mesoscale observations within a collaborative forecasting–research environment. *Wea. Forecasting*, **25**, 1510–1521, <https://doi.org/10.1175/2010WAF2222405.1>.
- Kleist, D. T., D. F. Parrish, J. C. Derber, R. Treadon, W. Wu, and S. Lord, 2009: Introduction of the GSI into the NCEP global data assimilation system. *Wea. Forecasting*, **24**, 1691–1705, <https://doi.org/10.1175/2009WAF2222201.1>.
- Kong, R., M. Xue, and C. Liu, 2018: Development of a hybrid En3DVar data assimilation system and comparisons with 3DVar and EnKF for radar data assimilation with observing

- system simulation experiments. *Mon. Wea. Rev.*, **146**, 175–198, <https://doi.org/10.1175/MWR-D-17-0164.1>.
- Kuhlman, K. M., C. L. Ziegler, E. R. Mansell, D. R. MacGorman, and J. M. Straka, 2006: Numerically simulated electrification and lightning of the 29 June 2000 STEPS supercell storm. *Mon. Wea. Rev.*, **134**, 2734–2757, <https://doi.org/10.1175/MWR3217.1>.
- Lilly, D. K., 1990: Numerical prediction of thunderstorms—Has its time come? *Quart. J. Roy. Meteor. Soc.*, **116**, 779–798, <https://doi.org/10.1002/qj.49711649402>.
- Liu, C., and S. Heckman, 2010: The application of total lightning detection and cell tracking for severe weather prediction. *TECO-2010-WMO Tech. Conf. on Meteorological and Environmental Instruments and Methods of Observation*, Helsinki, Finland, WMO, P2(7), [https://www.wmo.int/pages/prog/www/IMOP/publications/IOM-104\\_TECO-2010/P2012\\_2017\\_Heckman\\_USA.pdf](https://www.wmo.int/pages/prog/www/IMOP/publications/IOM-104_TECO-2010/P2012_2017_Heckman_USA.pdf).
- , and M. Xue, 2016: Relationships among four-dimensional hybrid ensemble-variational data assimilation algorithms with full and approximate ensemble covariance localization. *Mon. Wea. Rev.*, **144**, 591–606, <https://doi.org/10.1175/MWR-D-15-0203.1>.
- , —, and R. Kong, 2019: Direct assimilation of radar reflectivity data using 3DVAR: Treatment of hydrometeor background errors and OSSE tests. *Mon. Wea. Rev.*, **147**, 17–29, <https://doi.org/10.1175/MWR-D-18-0033.1>.
- MacGorman, D. R., D. W. Burgess, V. Mazur, W. D. Rust, W. L. Taylor, and B. C. Johnson, 1989: Lightning rates relative to tornadic storm evolution on 22 May 1981. *J. Atmos. Sci.*, **46**, 221–250, [https://doi.org/10.1175/1520-0469\(1989\)046<0221:LRRRTS>2.0.CO;2](https://doi.org/10.1175/1520-0469(1989)046<0221:LRRRTS>2.0.CO;2).
- , W. D. Rust, P. Krehbiel, W. Rison, E. Bruning, and K. Wiens, 2005: The electrical structure of two supercell storms during STEPS. *Mon. Wea. Rev.*, **133**, 2583–2607, <https://doi.org/10.1175/MWR2994.1>.
- , I. R. Apostolopoulos, N. R. Lund, N. W. S. Demetriades, M. J. Murphy, and P. R. Krehbiel, 2011: The timing of cloud-to-ground lightning relative to total lightning activity. *Mon. Wea. Rev.*, **139**, 3871–3886, <https://doi.org/10.1175/MWR-D-11-00047.1>.
- Mach, D. M., 2020: Geostationary lightning mapper clustering algorithm stability. *J. Geophys. Res. Atmos.*, **125**, e2019JD031900, <https://doi.org/10.1029/2019JD031900>.
- Mansell, E. R., 2014: Storm-scale ensemble Kalman filter assimilation of total lightning flash-extent data. *Mon. Wea. Rev.*, **142**, 3683–3695, <https://doi.org/10.1175/MWR-D-14-00061.1>.
- , D. R. MacGorman, C. L. Ziegler, and J. M. Straka, 2002: Simulated three-dimensional branched lightning in a numerical thunderstorm model. *J. Geophys. Res.*, **107**, 4075, <https://doi.org/10.1029/2000JD000244>.
- , C. L. Ziegler, and D. R. MacGorman, 2007: A lightning data assimilation technique for mesoscale forecast models. *Mon. Wea. Rev.*, **135**, 1732–1748, <https://doi.org/10.1175/MWR3387.1>.
- Mason, I. B., 2003: Binary events. *Forecast Verification: A Practitioner's Guide in Atmospheric Science*, I. T. Jolliffe and D. B. Stephenson, Eds., John Wiley & Sons, 37–76.
- Nag, A., and V. A. Rakov, 2014: Parameters of electric field waveforms produced by positive lightning return strokes. *IEEE Trans. Electromagn. Compat.*, **56**, 932–939, <https://doi.org/10.1109/TEMC.2013.2293628>.
- Nakanishi, M., 2001: Improvement of the Mellor–Yamada turbulence closure model based on large-eddy simulation data. *Bound.-Layer Meteor.*, **99**, 349–378, <https://doi.org/10.1023/A:1018915827400>.
- , and H. Niino, 2006: An improved Mellor–Yamada level-3 model: Its numerical stability and application to a regional prediction of advection fog. *Bound.-Layer Meteor.*, **119**, 397–407, <https://doi.org/10.1007/s10546-005-9030-8>.
- Papadopoulos, A., T. G. Chronis, and E. N. Anagnostou, 2005: Improving convective precipitation forecasting through assimilation of regional lightning measurements in a mesoscale model. *Mon. Wea. Rev.*, **133**, 1961–1977, <https://doi.org/10.1175/MWR2957.1>.
- Pessi, A. T., and S. Businger, 2009: The impact of lightning data assimilation on a winter storm simulation over the North Pacific Ocean. *Mon. Wea. Rev.*, **137**, 3177–3195, <https://doi.org/10.1175/2009MWR2765.1>.
- Pleim, J. E., 2007: A combined local and nonlocal closure model for the atmospheric boundary layer. Part I: Model description and testing. *J. Appl. Meteor. Climatol.*, **46**, 1383–1395, <https://doi.org/10.1175/JAM2539.1>.
- Rison, W., R. J. Thomas, P. R. Krehbiel, T. Hamlin, and J. Harlin, 1999: A GPS-based three-dimensional lightning mapping system: Initial observations in central New Mexico. *Geophys. Res. Lett.*, **26**, 3573–3576, <https://doi.org/10.1029/1999GL010856>.
- 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>.
- Rust, W. D., and Coauthors, 2005: Inverted-polarity electrical structures in thunderstorms in the Severe Thunderstorm Electrification and Precipitation Study (STEPS). *Atmos. Res.*, **76**, 247–271, <https://doi.org/10.1016/j.atmosres.2004.11.029>.
- 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>.
- Snook, N., M. Xue, and Y. Jung, 2011: Analysis of a tornadic mesoscale convective vortex based on ensemble Kalman filter assimilation of CASA X-band and WSR-88D radar data. *Mon. Wea. Rev.*, **139**, 3446–3468, <https://doi.org/10.1175/MWR-D-10-05053.1>.
- , —, and —, 2012: Ensemble probabilistic forecasts of a tornadic mesoscale convective system from ensemble Kalman filter analyses using WSR-88D and CASA radar data. *Mon. Wea. Rev.*, **140**, 2126–2146, <https://doi.org/10.1175/MWR-D-11-00117.1>.
- Sun, J., and Coauthors, 2014: Use of NWP for nowcasting convective precipitation: Recent progress and challenges. *Bull. Amer. Meteor. Soc.*, **95**, 409–426, <https://doi.org/10.1175/BAMS-D-11-00263.1>.
- Thomas, R. J., P. R. Krehbiel, W. Rison, S. J. Hunyady, W. P. Winn, T. Hamlin, and J. Harlin, 2004: Accuracy of the lightning mapping array. *J. Geophys. Res.*, **109**, D14207, <https://doi.org/10.1029/2004JD004549>.
- 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>.
- Tong, M., and M. Xue, 2005: Ensemble Kalman filter assimilation of Doppler radar data with a compressible nonhydrostatic model: OSS experiments. *Mon. Wea. Rev.*, **133**, 1789–1807, <https://doi.org/10.1175/MWR2898.1>.
- , and —, 2008: Simultaneous estimation of microphysical parameters and atmospheric state with simulated radar data and ensemble square root Kalman filter. Part I: Sensitivity analysis and parameter identifiability. *Mon. Wea. Rev.*, **136**, 1630–1648, <https://doi.org/10.1175/2007MWR2070.1>.

- Vicente, G. A., J. C. Davenport, and R. A. Scofield, 2002: The role of orographic and parallax corrections on real time high resolution satellite rainfall rate distribution. *Int. J. Remote Sens.*, **23**, 221–230, <https://doi.org/10.1080/01431160010006935>.
- Wang, S., M. Xue, A. D. Schenkman, and J. Z. Min, 2013: An iterative ensemble square root filter and tests with simulated radar data for storm-scale data assimilation. *Quart. J. Roy. Meteor. Soc.*, **139**, 1888–1903, <https://doi.org/10.1002/qj.2077>.
- Whitaker, J. S., and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130**, 1913–1924, [https://doi.org/10.1175/1520-0493\(2002\)130<1913:EDAWPO>2.0.CO;2](https://doi.org/10.1175/1520-0493(2002)130<1913:EDAWPO>2.0.CO;2).
- , and —, 2012: Evaluating methods to account for system errors in ensemble data assimilation. *Mon. Wea. Rev.*, **140**, 3078–3089, <https://doi.org/10.1175/MWR-D-11-00276.1>.
- Wiens, K. C., S. A. Rutledge, and S. A. Tessendorf, 2005: The 29 June 2000 supercell observed during STEPS. Part II: Lightning and charge structure. *J. Atmos. Sci.*, **62**, 4151–4177, <https://doi.org/10.1175/JAS3615.1>.
- Wu, W., R. J. Purser, and D. F. Parrish, 2002: Three-dimensional variational analysis with spatially inhomogeneous covariances. *Mon. Wea. Rev.*, **130**, 2905–2916, [https://doi.org/10.1175/1520-0493\(2002\)130<2905:TDVAWS>2.0.CO;2](https://doi.org/10.1175/1520-0493(2002)130<2905:TDVAWS>2.0.CO;2).
- Xue, M., M. Tong, and K. K. Droegemeier, 2006: An OSSE framework based on the ensemble square root Kalman filter for evaluating the impact of data from radar networks on thunderstorm analysis and forecasting. *J. Atmos. Oceanic Technol.*, **23**, 46–66, <https://doi.org/10.1175/JTECH1835.1>.
- , —, and G. Zhang, 2009: Simultaneous state estimation and attenuation correction for thunderstorms with radar data using an ensemble Kalman filter: Tests with simulated data. *Quart. J. Roy. Meteor. Soc.*, **135**, 1409–1423, <https://doi.org/10.1002/qj.453>.
- , Y. Jung, and G. Zhang, 2010: State estimation of convective storms with a two-moment microphysics scheme and an ensemble Kalman filter: Experiments with simulated radar data. *Quart. J. Roy. Meteor. Soc.*, **136**, 685–700, <https://doi.org/10.1002/qj.593>.
- 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>.