Development of New Observation Operators for Assimilating GOES-R Geostationary Lightning Mapper Flash Extent Density Data Using GSI EnKF: Tests with Two Convective Events over the United States

RONG KONG,[a] MING XUE,[a,b] CHENGSI LIU,[a] ALEXANDRE O. FIERRO,[c,d,e] AND EDWARD R. MANSELL[d]

[a] *Center for Analysis and Prediction of Storms, Norman, Oklahoma*
[b] *University of Oklahoma, Norman, Oklahoma*
[c] *Cooperative Institute for Mesoscale Meteorological Studies, Norman, Oklahoma*
[d] *NOAA/National Severe Storms Laboratory, Norman, Oklahoma*
[e] *Department of Forecasting Models, Zentralanstalt für Meteorologie und Geodynamik (ZAMG), Vienna, Austria*

ABSTRACT: In a prior study, GOES-R Geostationary Lightning Mapper (GLM) flash extent density (FED) data were assimilated using ensemble Kalman filter into a convection-allowing model for a mesoscale convective system (MCS) and a supercell storm. The FED observation operator based on a linear relation with column graupel mass was tuned by multiplying a factor to avoid large FED forecast bias. In this study, new observation operators are developed by fitting a third-order polynomial to GLM FED observations and the corresponding FED forecasts of graupel mass of the MCS and/or supercell cases. The new operators are used to assimilate the FED data for both cases, in three sets of experiments called MCSFit, SupercellFit, and CombinedFit, and their performances are compared with the prior results using the linear operator and with a reference simulation assimilating no FED data. The new nonlinear operators reduce the frequency biases (root-mean-square innovations) in the 0–4-h forecasts of the FED (radar reflectivity) relative to the results using the linear operator for both storm cases. The operator obtained by fitting data from the same case performs slightly better than fitting to data from the other case, while the operator obtained by fitting forecasts of both cases produce intermediate but still very similar results, and the latter is considered more general. In practice, a more general operator can be developed by fitting data from more cases.

SIGNIFICANCE STATEMENT: Prior studies found that assimilation of satellite lightning observation can benefit storm forecasts for up to 4 h. A linear lightning observation operator originally developed for assimilating pseudo-satellite lightning observations was tuned earlier through sensitivity experiments when assimilating real lightning data. However, the linear relation does not fit the model and observational data well and significant bias can exist. This study develops new lightning observation operators by fitting a high-order polynomial to satellite lightning observations and model-predicted quantities that directly relate to lightning. The new operator was found to reduce the frequency biases and root-mean-square innovations for lightning and radar reflectivity forecasts, respectively, up to several hours relative to the linear operator. The methodology can be applied to larger data samples to obtain a more general operator for use in operational data assimilation systems.

KEYWORDS: Convective storms/systems; Lightning; Satellite observations; Data assimilation

1. Introduction

Accurate initial conditions play a crucial role in improving the skill of numerical weather prediction (NWP; Stensrud and Fritsch 1994). In the absence of convective-scale data assimilation (DA), the initial conditions of storm-scale NWP often lack small-scale information that causes the commonly known spinup problem in the forecast. New observation types with high spatial and temporal resolutions, such as those of weather radars and new-generation geostationary satellites, can provide more detailed information at the convective scales than most conventional observations. In the recent decades, many studies have shown the benefit of radar DA using variational or ensemble-based methods for improving short-range weather forecasting (e.g., Xue et al. 2003; Gao et al. 2004; Tong and Xue 2005; Xue et al. 2006; Jung et al. 2008; Aksoy 2010; Aksoy et al. 2009; Kain et al. 2010; Dowell et al. 2011; Gao et al. 2013;

Liu and Xue 2016; Wang and Wang 2017; Kong et al. 2018; Liu et al. 2019; Liu et al. 2020; Kong et al. 2021).

Considering that radars do not have coverage over vast oceanic regions and suffer from beam blockage in mountainous terrain, work aimed at assimilating high-resolution satellite data in a more systematic manner has received heightened attention in recent years. Research focusing on satellite DA has primarily focused on three different satellites datasets, which are 1) satellite infrared and microwave radiances (Eyre et al. 1993; Derber and Wu 1998; Chevallier et al. 2004; Vukicevic et al. 2004; McNally et al. 2006; Vukicevic et al. 2006; Chen et al. 2008; Geer et al. 2010; Geer and Bauer 2011; Guerbette et al. 2016; Kazumori et al. 2016; Geer et al. 2018; Geer et al. 2019); 2) the cloud water path (CWP) and vertically integrated liquid (VIL) retrievals (Minnis et al. 2008a,b; Minnis et al. 2011; Jones et al. 2013; Chen et al. 2015; Jones and Stensrud 2015; Jones et al. 2018); and 3) the atmospheric motion vectors derived from cloud measurements (AMVs; Wu et al. 2014; Wu et al. 2015; Zhang and Pu 2018). Many of these studies examined impacts on tropical cyclone

predictions. The assimilation of satellite data is one of the most significant reasons for global NWP improvements in recent decades. Infrared and microwave radiances have been effectively assimilated in the operational systems of the European Centre for Medium-Range Weather Forecasts (ECMWF; Bauer et al. 2010; Geer et al. 2017; Geer et al. 2018) and NOAA/National Centers for Environmental Prediction (NCEP; Zhu et al. 2016). Certain radiances in cloudy regions have also been assimilated although the effectiveness is limited by the inability of global models in accurately predicting clouds (Zhu et al. 2016). Advances have also been made in the forecasting of storm-scale precipitation and severe weather events by directly assimilating infrared radiance (in cloudy regions) into convection-allowing NWP models (Zhang et al. 2016; Honda et al. 2018; Jones et al. 2018; Minamide and Zhang 2018; Zhang et al. 2018; Sawada et al. 2019; Jones et al. 2020). In addition to the above three satellite datasets, the GLM on board the *GOES-16/-17* R-series satellites provide total lightning data products with high spatial (roughly 8-km pixel) resolution and at 20-s time level over the Americas and adjacent ocean regions (Goodman et al. 2013). Despite their availability, the GLM lightning data are not yet assimilated into operational NWP models.

In recent years, more efforts have been made to assimilate lightning data into NWP models. Many of them use indirect assimilation method where proxy observations are created using lightning data then assimilated or the lightning data are used to adjust other model variables. For example Stefanescu et al. (2013) use convective available potential energy (CAPE) as a proxy between lightning data and model variables. Fierro et al. directly increase the water vapor mixing ratio at locations of lightning observations using a nudging method (Fierro et al. 2012, 2014) or impose water saturation within a fixed level above the lifted condensation level (Fierro et al. 2016, 2019) and Lightning data can be used to create pseudo-observations of water vapor mixing ratio (Fierro et al. 2019; Hu et al. 2020), vertical velocity (Z. X. Chen et al. 2020; Gan et al. 2021; Xiao et al. 2021), and radar reflectivity (Liu et al. 2017; Y. Chen et al. 2020) which are then assimilated. Direct assimilation of lightning data involves the development and utilization of an observation operator that links the lightning observation with model states. Apodaca et al. (2014) and Apodaca and Zupanski (2018) linked lightning data to vertical velocity of the global forecast model while Mansell (2014) and Allen et al. (2016) built the linkage with model graupel field and used an ensemble Kalman filter (EnKF) for data assimilation. Most of these studies assimilated lightning data derived from the ground-based lightning networks.

Before actual GLM data became available, efforts were made to assimilate simulated GLM FED data (Mansell 2014) or pseudo-GLM FED data derived from ground-based lightning detection networks (e.g., the ground-based lightning mapping array, Allen et al. 2016). Flash extent density (FED; with the unit of flashes per minute per pixel, hereinafter flashes min^{-1} pixel^{-1}) is the total number of flashes that occur within a grid cell over a given period (Lojou and Cummins 2005). Please refer to Bruning et al. (2019) for detailed descriptions and derivations of the GLM FED data. Mansell (2014) was the first to apply the EnKF method to the assimilation of lightning data, using simulated pseudo-GLM FED observations.

Allen et al. (2016) later applied that EnKF system to two real storm cases, assimilating pseudo-GLM FED observations derived from a ground-based lightning mapping array (LMA; Rison et al. 1999). In both studies, FED observation operators that linearly related FED to total graupel mass or total graupel volume within the FED pixel column were used. The operators were originally developed based on a linear fit to output data from cloud model simulations that contain explicit electrification physics (Mansell et al. 2002).

With the availability of operational *GOES-16/-17* GLM data over the past 5 years, studies have been carried out to assimilate actual data and examine their impacts on storm-scale predictions. For example, Fierro et al. (2019) and Hu et al. (2020) assimilate the pseudo-observation for water vapor mass retrieved from GLM lightning observations using a 3DVAR algorithm developed at National Severe Storms Laboratory (NSSL; Gao et al. 2004, 2016). These studies, while not directly assimilating lightning data themselves, but pseudo-observations derived from lightning data, were already showing the benefit of assimilating GLM data for improving shorter-term (≤6 h) storm forecasts. Compared to the 3DVar method, the ensemble Kalman filter (EnKF; Evensen 2003) can analyze all model state variables even when a very limited number of parameters are measured, because of the flow-dependent cross covariances that can be calculated between the observation prior and model state variables. For example, radar reflectivity can update all state variables within a cloud model including all hydrometeors (Tong and Xue 2005). Similarly, total lightning flash rate can update all model state variables using EnKF as long as relationship can be found between the observed flash rate and certain model state variables and the relationship is commonly known as the observation operator.

In Kong et al. (2020), the linear FED observation operators based on column graupel mass or graupel volume tested in Allen et al. (2016) were implemented within the EnKF algorithm of the GSI DA system of NCEP (Parrish and Derber 1992; Kleist et al. 2009) and tested with real GLM FED data for a mesoscale convective system (MCS) case that occurred over the central United States. In their study, 1-min FED rates were assimilated at 5-min intervals over a 1-h period to initialize the MCS on a convection-allowing 3-km grid. The GLM FED assimilation was shown to be effective in adjusting not only the model state that is directly related to FED observation in the observation operator (graupel mixing ratio), but also other model state variables (such as mixing ratios of water vapor, cloud water, cloud ice, and the temperature, pressure, and wind fields) through the cross-correlation between observation prior and model states (see Figs. 8 and 9 in Kong et al. 2020). The FED DA was found to noticeably improve the analysis and prediction of the MCS for up to several hours (~4 h). Both graupel-mass and graupel-volume-based operators were found to produce similar results. However, sensitivity experiments showed that a factor of 1/2 multiplied to the original linear FED observation operators of Allen et al. (2016) yielded best results. Otherwise, the predicted FED had significant high bias. In fact, original linear fit of operator in Allen et al. (2016) has significant scatter, and there was also apparent storm type dependency. More accurate operators could be obtained by relaxing the linear
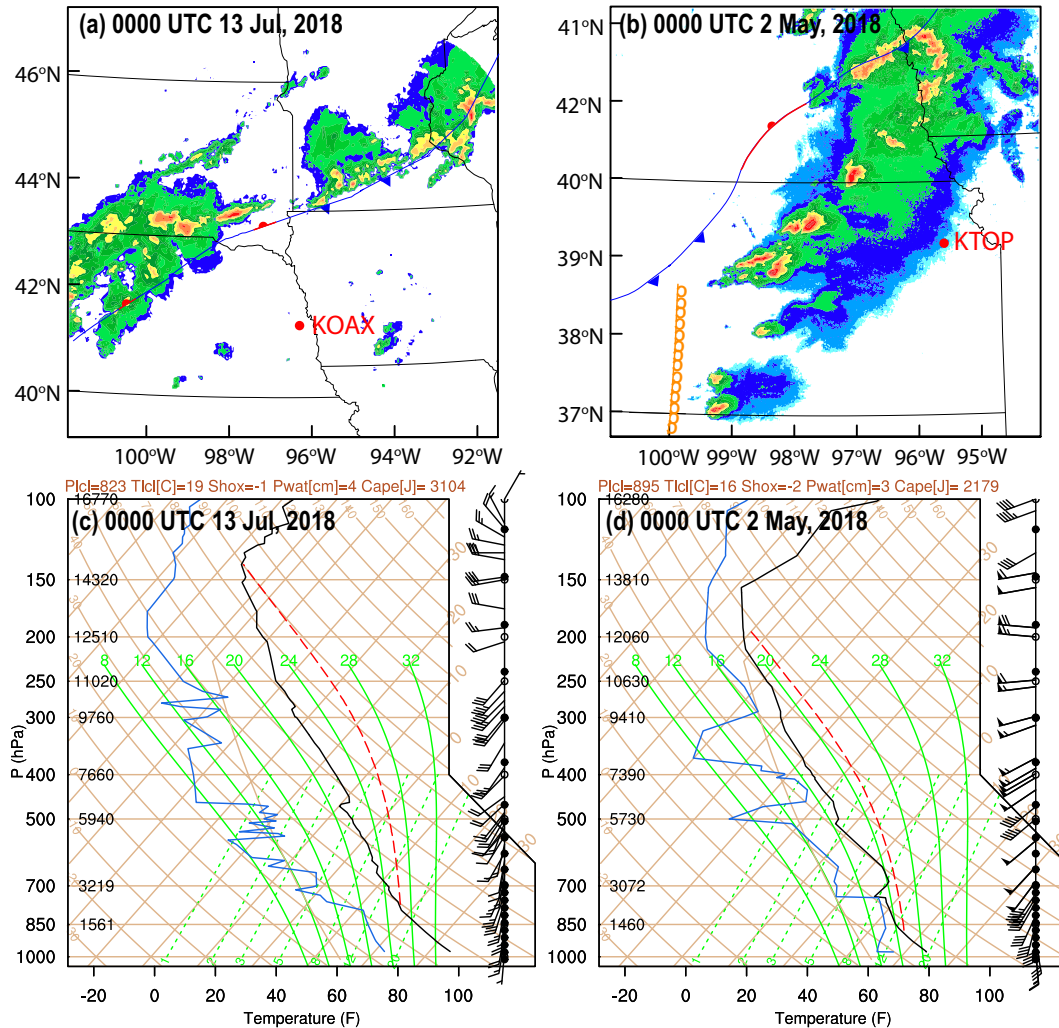
FIG. 1. Locations of the fronts (blue and red contours) and the dryline (orange contour), overlaid with the composite reflectivity observations (dBZ) remapped from WSR-88D (shaded contours) at 0000 UTC 13 Jul 2018 and 0000 UTC 2 May 2018 for the (a) MCS and (b) the supercell storm cases, respectively. (bottom) The pre-storm soundings used for (c) the MCS case and (d) the supercell case. For the MCS case, the thermodynamic profile has 3104 J kg^{-1} CAPE. The sounding comes from data collected during the 0000 UTC 13 Jul 2018 OAX (Omaha, NE), radiosonde launch [red dot in (a)]. For the supercell case, the thermodynamic profile has 4093 J kg^{-1} CAPE, and the sounding comes from data collected during the 0000 UTC 2 May 2018 Topeka, KS, radiosonde launch [red dot in (b)]. In both cases the black line corresponds to temperature, the blue line corresponds to dewpoint, and the green line corresponds to the moist adiabatic followed by a parcel averaged over the lowest level of the sounding.

relation assumption and allowing for nonlinear relations. Moreover, the amount of graupel predicted by an NWP model is often sensitive to the microphysics scheme used. To reduce systematic bias associated with the observation operator, we propose in this study to derive new operators by fitting observed FED data to model-simulated graupel mass using high-order polynomials. Doing so takes into account biases in simulated graupel mass as well as any reduced efficiency of GLM total lightning detection in optically deep storms. This approach does assume that the modeled storms are reasonably similar to real storms, and short-range forecasts with convective-scale DA can usually serve such a purpose. As a proof of concept, reasonably accurate short-range forecasts of an MCS from the control experiment of Kong et al. (2020) that assimilated GLM FED data, as well as forecasts for a supercell storm case obtained using a similar procedure are used as forecast dataset to perform the nonlinear polynomial fit. Developing and testing the new operators based on the relatively weak convection in the MCS case and the strong convection in the supercell case enable the operators to include information spanning both ends of the convective spectrum. EnKF DA experiments similar to the control experiment in Kong et al. (2020) are performed using the new operators and the results are compared to those of the control experiment of Kong et al. (2020).
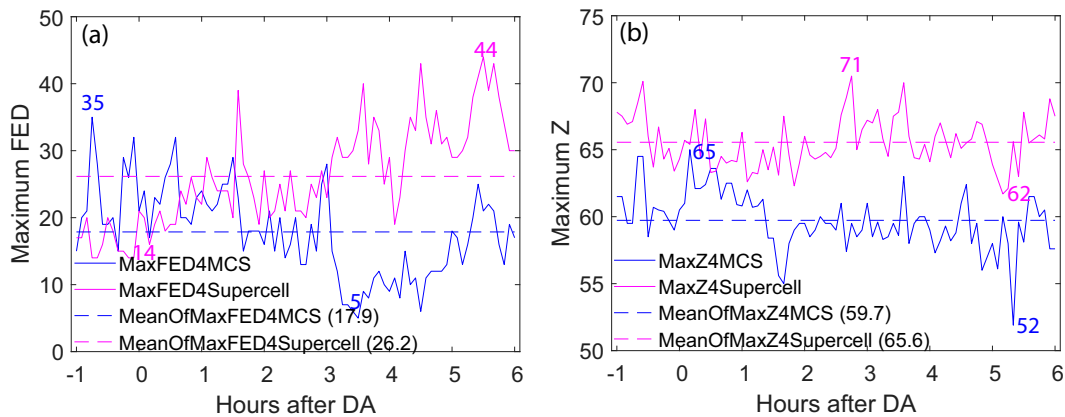
FIG. 2. Domain maximum (a) FEDs (flashes min^{-1} $pixel^{-1}$) and (b) reflectivity (dBZ) during 1-h DA window (i.e., −1 to 0 h after DA) and 0–6-h forecasts afterward (solid lines) and the corresponding time average (dashed lines) for the MCS (blue) and supercell (magenta) cases, respectively.

The rest of this paper is organized as follows. Section 2 describes the DA method, the forecast model and simulation setup, as well as the experimental design. In section 3, the FED observation data and storm cases are discussed. Section 4 describes the GLM lightning observation and the processing of FED data. In section 5, development of the new observation operators is described. The DA experiments that use the nonlinear observation operators are compared with the results using linear observation operator in section 6. A summary and conclusions are presented in section 7.

2. Case description and observational data

The two storm cases in this study were well observed by the Next Generation Weather Radar (NEXRAD) operational Weather Surveillance Radar-1988 Doppler (WSR-88D) radar network (Doviak et al. 2000) and the GOES-R GLM instrument. The two storm cases are briefly introduced below.

a. Case 1: 12 July 2018 MCS

This MCS occurred over parts of South Dakota and southern Minnesota on 13 July 2018. As is shown in Fig. 1, the storms are linked to a southwest–northeast (SW–NE)-oriented stationary front. The mixed layer convective available potential

energy (CAPE) at 1300 UTC at OAX (Omaha, Nebraska) was 3104 J kg^{-1}, providing a thermodynamic pre-storm environment favorable for deep, mixed phase moist convection. Initially, before about 2100 UTC, storms developed as a scattered field of cumulus congestus, which later developed into a well-organized west-southwest–north-northeast (WSW–NNE)-oriented MCS around 0000 UTC (Fig. 1a). The MCS persisted for more than 6 h, producing copious rainfall exceeding 103 mm in some places (not shown). The radar observations from 0000 to 0600 UTC 13 July 2018 showed radar echoes composed of stratiform precipitation alongside scattered convective areas, with the stratiform echoes covering a large area and accounting for a large portion of the rainfall (not shown). The domain maximum FED from 2300 UTC 12 July to 0600 UTC 13 July ranges from 5 to 35 flashes min^{-1} $pixel^{-1}$ (and maximum reflectivity of 52 to 65 dBZ) with a time-averaged value of 17.9 flashes min^{-1} $pixel^{-1}$ (59.7 dBZ, Fig. 2; see section 4 for processing of FED observation). No severe weather event was reported for this storm.

b. Case 2: 1 May 2018 supercell storm event

The supercell storms in this case occurred in north central Kansas and south Nebraska during the afternoon and evening hours of 1 May 2018. The supercells were associated with a

TABLE 1. Lightning data assimilation experiments, here GM is the graupel mass (in kg) summed over a 15×15 km^2 area.

| Expt name	Description
NoDA	Do not assimilate any data
K2020 (Control Expt)	Assimilate FED data using the observation operator tuned based on sensitivity experiments (Kong et al. 2020)
MCSFit	Assimilate FED data using the observation operator based on the third-order polynomial fit between the FED observations and 0–6-h free forecasts after DA in experiment K2020 for the MCS case
SupercellFit	Assimilate FED data using the observation operator based on the third-order polynomial fit between the FED observations and 0–6-h free forecasts after DA in experiment K2020 for the Supercell case
CombinedFit	Third-order polynomial fitting between the FED observations and 0–6-h free forecasts after DA in experiment K2020 for both the MCS and the supercell cases
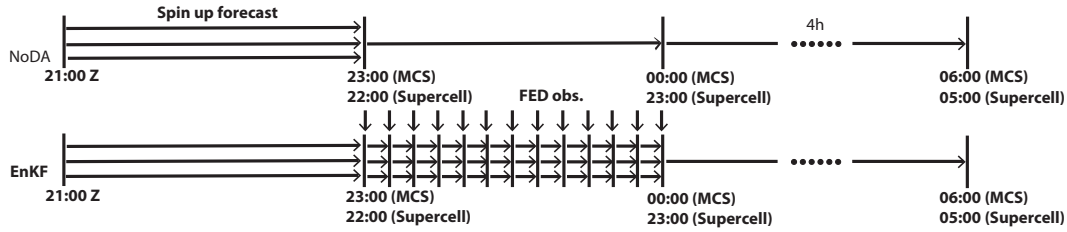
FIG. 3. Flow diagram of the (bottom) EnKF DA experiments vs (top) the no-DA run. The spinup ensemble forecasts from 2100 to 2300 UTC (2200 UTC) for the MCS (supercell) case include 40 members. FED DA occurs between 2300 (2200) and 0000 (2300) UTC with 5-min cycles for the MCS (supercell) case, and deterministic forecast is launched from the final ensemble-mean analysis at 0000 (2300) UTC and run to 0600 (0500) UTC for the MCS (supercell) case. The CTRL experiment continues the ensemble forecasts through 0000 (2300) UTC without DA for the MCS (supercell) case, when deterministic forecast continues from the ensemble-mean forecast at 0000 (2300) UTC for the MCS (supercell) case.

stationary front and a dryline in the southwest portions of the domain (Fig. 1b). The mixed layer CAPE at 0000 UTC at TOP (Topeka, Kansas) was 2179 J kg^{-1} with strong wind shear providing an environment supportive for severe storms. Scattered storms started to develop at about 2000 UTC 1 May 2018 and began to exhibit classic supercellular characteristics by about 2200–3000 UTC. By 2246 UTC, the storm produced a weak and short-lived tornado. A longer-tracked tornado formed later on at 0008 UTC 2 May 2018 and passed through parts of Ellsworth, Saline, and Ottawa county of Kansas. The tornado dissipated at around 0055 UTC, but the storms persisted for several additional hours. In addition to tornadoes,

many of the storms produced strong wind (≥ 27 m s^{-1}) and large hail (≥ 2.5 cm) reports. The domain maximum FED (reflectivity) from 2300 UTC 1 May to 0500 UTC 2 May ranges from 14 to 44 flashes min^{-1} pixel^{-1} (62 to 71 dBZ) with the time-averaged value of 26.2 flashes min^{-1} pixel^{-1} (65.6 dBZ), which are notably larger than those reported for the MCS case (Fig. 2).

c. GLM lightning observation and their processing

The GLM Level-II data are provided on a geostationary latitude–longitude grid, which include a wide variety of optically derived variables associated with lightning (Goodman et al. 2013;
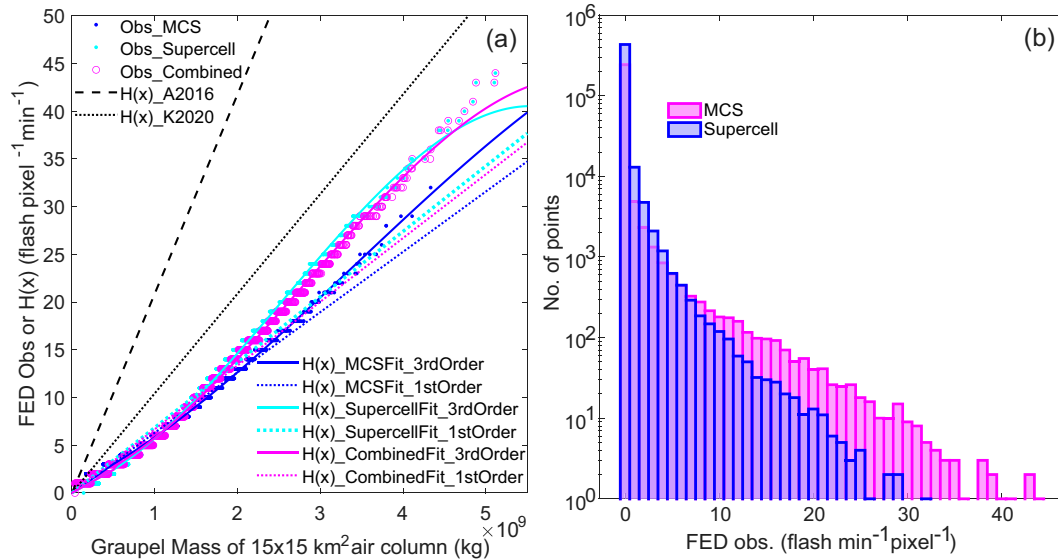

FIG. 4. (a) The observed FED and the FED simulations using the graupel-mass based observation operator from Allen et al. (2016, A2016 in our study), and the first- (MCSFit_1stOrder, SupercellFit_1stOrder, CombinedFit_1stOrder), third-order (MCSFit_3rdOrder, SupercellFit_3rdOrder, CombinedFit_3rdOrder) polynomial fittings between the observed FED and 0–6-h forecasts of the graupel mass field on a 15×15 km^2 air column for a set of control experiments (K2020) run for the MCS (MCSFit_1st/3rdOrder), supercell (SupercellFit_1st/3rdOrder), and the combination of two cases (Combined_1st/3rdOrder). In Allen et al. (2016), the simulated FED was calculated using a 16×16 km^2 air column, the right-hand side of the observation operator in A2016 is scaled by $[(16 \times 16)/(15 \times 15)]$ to be consistent with the observation operator used in Allen et al. (2016). (b) Histograms of the processed FED observations for the MCS and supercell cases, respectively.
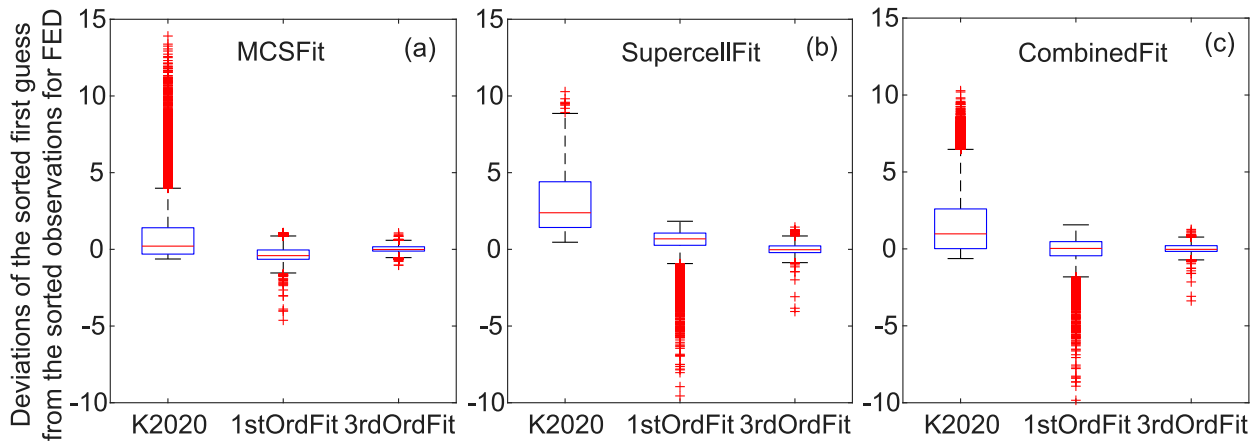
FIG. 5. The boxplots of the sorted first guess [i.e., $H(x)$] minus sorted observations (y) of the FED fields (flashes min^{-1} pixel^{-1}) for control experiment K2020, first-, and third-order polynomial fits for (a) the MCS, (b) supercell, and (c) combination of the two cases. Boxes represent the 25th–75th percentiles (bottom and top edges in blue) range or IQR with a red line indicating the median. The whiskers (dashed lines) represent the 25th percentile minus 1.5 times the IQR and the 75th percentile plus 1.5 times the IQR, extending to the most extreme data point in this data range with outliers (red crosses).

Rudlosky et al. 2019; Rudlosky and Virts 2021). Flash density is defined as the number of flashes occurring within a grid cell over a given time interval. To account for the areal extent of flashes in the flash density calculations, GLM-level optical "events" and the parent ID of each flash must be considered (Lojou and Cummins 2005; Bruning et al. 2019). Similar to Kong et al. (2020), the GLM lightning observations are remapped onto a 10-km grid before being assimilated, where the footprint is slightly larger than those used in (Mansell 2014) to alleviate gaps between adjacent pixels.

3. Data assimilation setup

a. DA method

A scalable implementation of the ensemble square root filter (EnSRF; Whitaker and Hamill 2002) that produces similar analysis as traditional EnSRF (a variant of EnKF) and updates both the state vectors and observation operators is used in GSI EnKF (Anderson and Collins 2007). Detailed information about the update equations of GSI EnKF can be found in Kong et al. (2020).

b. Forecast model setup and initialization of ensembles

The Advanced Research WRF (ARW, Skamarock et al. 2008) model (version 3.8.1) is used for the simulations, which use a 300 × 300 (500 × 500) grid with a 3-km (1-km) horizontal grid spacing for the MCS (supercell) storm case and 53 stretched vertical levels (with a model top at about 21 km). Following Kong et al. (2020), deviations of the 2100 UTC initial conditions (analyses) of the operational Short-Range Ensemble Forecast (SREF) from the North American Model (NAM) 3-h forecast valid at 2100 UTC are calculated, scaled by 75%, and then inverted to obtain a 40-member initial ensemble for the MCS (supercell) case. The scaling value was determined by sensitivity experiments. Since the resolutions of both SREF and NAM are larger than 10 km, additional smoothed, Gaussian-

distributed small-scale perturbations are added to the horizontal velocity, humidity, and potential temperature fields at the initial time to introduce convective-scale perturbations to the initial ensemble (Tong 2008; Snook et al. 2012; Johnson et al. 2014). The readers are encouraged to consult Kong et al. (2020) for more detailed information about generation of the initial ensembles. The land surface model, PBL, and microphysics (6-class Thompson bulk microphysics) schemes used are the same as those used in Kong et al. (2020). Each initial ensemble member is first integrated forward for 2 h (1 h) until 2300 UTC spin up the model fields on the 3- or 1-km grid before DA cycles start.

c. Design of data assimilation experiments

As noted earlier, the main goal of this study is to develop and test nonlinear observation operators for the assimilation of GOES-R GLM FED data. The new observation operators are developed based on the nonlinear fit between FED observations and model forecasts of FED, from control experiments using the tuned linear operator as in Kong et al. (2020). The control experiment is labeled K2020 in this study and is run for both MCS and supercell cases.

New observation operators are developed based on third-order polynomial regressions between the FED observations and 0–6-h FED forecasts of K2020. Only graupel-mass-based FED observation operators are developed and tested. This is because graupel volume is calculated only when graupel mixing ratio exceeds a fixed, specified threshold (e.g., 0.5 g kg^{-1}). As a result, the FED estimated based on graupel volume is a discontinuous function of graupel mixing ratio. This poses an issue in a variational DA framework, where discontinuous observation operator can lead to poor convergence of the cost function. Although Kong et al. (2020) showed that the two operators performed similarly for ensemble assimilation, the graupel-mass-based FED observation operator is also a better choice than the
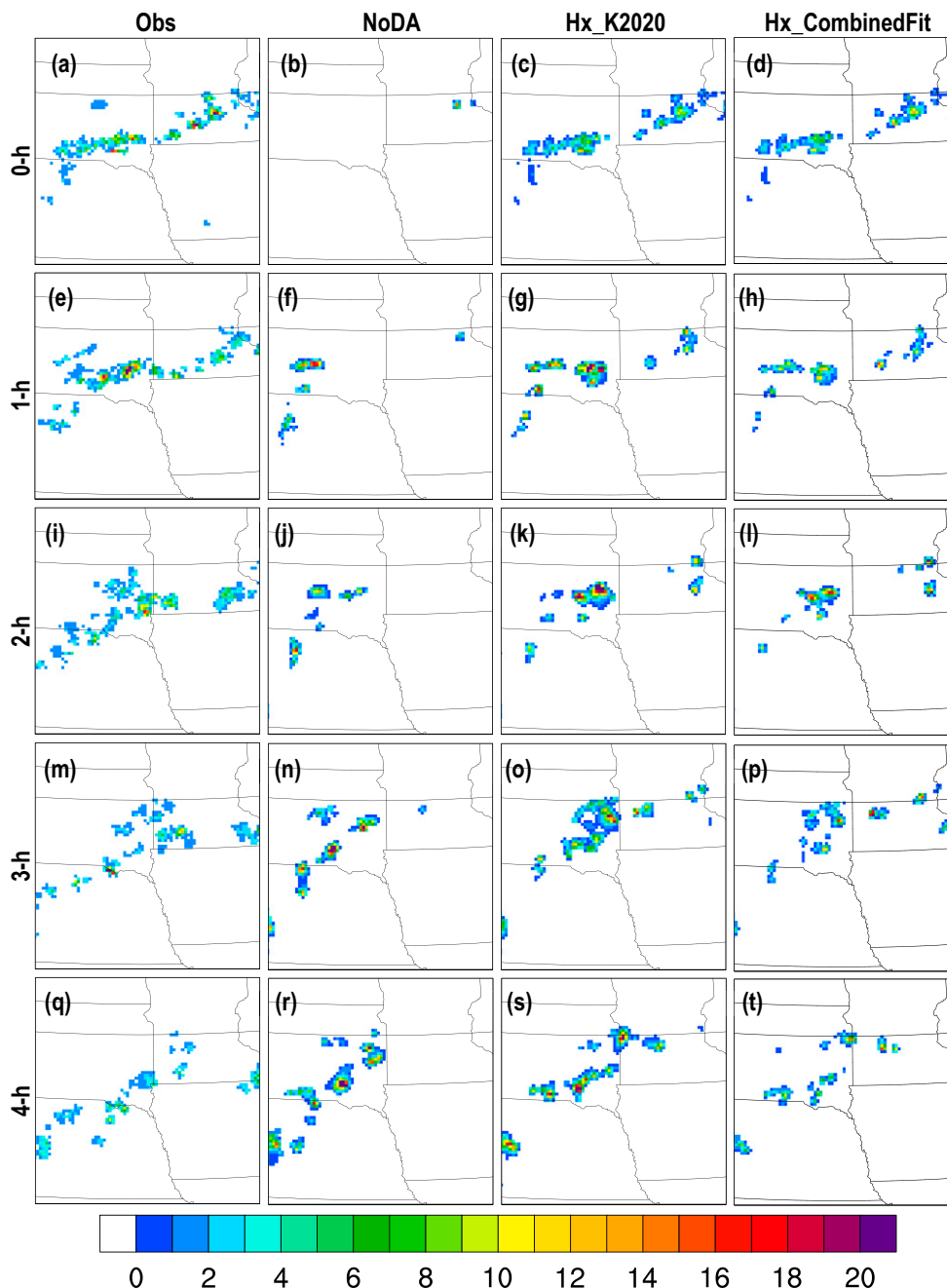
FIG. 6. Horizontal cross sections of (a),(e),(i),(m),(q) 1-min FED observations (units: flashes min^{-1} pixel^{-1}), and FED forecasts from (b),(f),(j),(n),(r) NoDA; (c),(g),(k),(o),(s) K2020; and (d),(h),(l),(p),(t) CombinedFit after (a)–(d) 0-, (e)–(h) 1-, (i)–(l) 2-, (m)–(p) 3-, and (q)–(t) 4-h free forecasts following the 1-h cycled DA period (valid from 0000 to 0400 UTC 13 Jul 2018) for the MCS case.

graupel-volume-based operator for EnKF. Since the former operator is linear function of graupel mixing ratio, the perturbations of the observation priors (needed when calculating the Kalman gain) are normally distributed if the background errors are normally distributed. While for the latter operator, the distributions of the observation priors are very non-Gaussian, even if the background errors are Gaussian. Detailed information on the

fitting procedure is given in section 5. The DA experiments that use the new operators based on data fitting for the MCS case, supercell case, and the combined data of the two cases are labeled as MCSFit, SupercellFit, and CombinedFit, respectively. These experiments are compared with control experiment K2020 and with an experiment not assimilating any data (NoDA). More details on the experiments are given in Table 1.
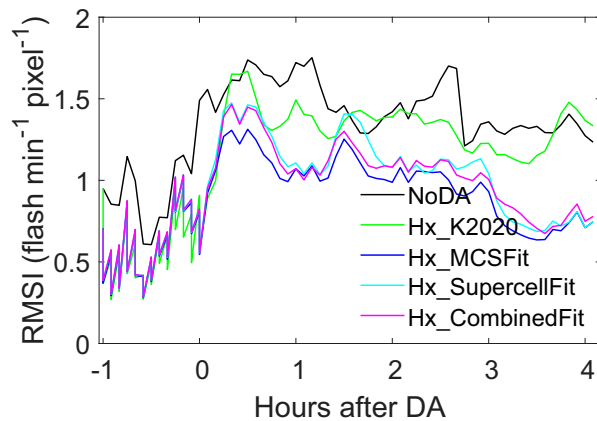
FIG. 7. RMSIs of the FED analyses and forecasts (flashes min^{-1} pixel^{-1}) within 1-h DA (corresponding to -1 to 0 h after DA) and 0–4-h FED forecasts from experiments NoDA, K2020, MCSFit, SupercellFit, and CombinedFit, for the MCS case.

Following Kong et al. (2020), 1-min FED observations are assimilated every 5 min over a 1-h period, a 6-h free forecast from the final ensemble-mean analysis is then launched. The first analysis is at 2300 UTC 12 July (2200 UTC 1 May) for the MCS (supercell) case, which is after a 2-h (1-h) integration of the initial ensemble (Fig. 3). The height of the FED observations is assumed to be 6.5 km for the vertical covariance localization purpose. The horizontal and vertical localization radii are 15 km and -4 in $\log(P/P_0)$ space (~32 km on average), respectively. The FED observation error is set to 0.5 flashes min^{-1} pixel^{-1}. To help maintain the ensemble spread after DA, a 0.95 adaptive posterior inflation that relaxes the posterior ensemble spread to 95% of the prior ensemble spread (Whitaker 2012) is applied in each DA cycle. The settings follow Kong et al. (2020) exactly.

4. Development of nonlinear FED observation operators

a. The linear observation operator used in the control experiment

The graupel-mass-based linear observation operator with the optimal tuning factor from Kong et al. (2020) is used in the control experiment of this study. This operator can be expressed as

$$\text{FED} = 1.044 \times 10^{-8}(\text{GM}), \qquad (1)$$

where GM is the graupel mass (in kg). Compared to the original formula of Allen et al. (2016), a factor of 1/2 has been multiplied in (1). Similar to Mansell (2014) and Allen et al. (2016), the graupel mass is summed over a volume spanning the vertical depth of the model domain and covering a 15×15 km^2 area in horizontal dimensions centered on the observation pixels. The FED priors that involve horizontal and vertical integrations are calculated first on the model grid then interpolated to the observation locations. In our FED DA using the EnSRF algorithm, both the model state variable that is directly related to the FED observation in the observation operator [graupel mixing ratio based on Eq. (1)] and the model fields that are not directly related to the FED observation via its operator (such as, three-dimensional

wind fields, temperature, pressure, hydrometeor mixing ratios of water vapor, rain, snow, cloud water, and cloud ice) can be updated via flow-dependent background error covariances, as discussed earlier in Introduction.

b. Development of the nonlinear observation operator

In this section, new nonlinear FED observation operators are developed based on FED observations and model forecasts of the MCS and supercell storm cases. A pair of control experiments using the tuned linear operator from Kong et al. (2020) are first performed to assimilate FED observations every 5 min for 1 h and then produce 6-h free forecasts from the ensemble-mean analyses for both the MCS and supercell cases. The model predicted FEDs every 5 min in the 6-h forecasts are used to obtain polynomial fit with observed FED data, for the MCS and supercell cases individually, and for the combined data of the two cases. The model FED is calculated by integrating the predicted graupel over the model depth and cover a 15×15 km^2 area then interpolating it to the observation locations.

With our procedure, systematic observation error or bias, and systematic discrepancies between model and observations can in principle be statistically accounted for, similarly to the practice of using observations to calibrate parameters in a computer model (Kennedy and O'Hagan 2001). Because of factors (such as spatial displacement errors) that are difficult to account for, point-by-point fitting between FED observations and graupel masses is difficult. In fact, attempt was made to do so, very large scatter was encountered that made close fitting difficult. To deal with this problem, the FED observations and calculated graupel masses are first aggregated. Data aggregation, which reduces the dimensions (degrees of freedom) of the data, is a common practice in climate model calibration to potentially reduce uncertainties of the calibrated parameters (Chang et al. 2014). The aggregated one-dimensional FED observations and graupel masses are then sorted in ascending order before being fitted to each other, under the assumption that the probability distribution function of the less accurate data (the simulated FED in our case) equals to that of the more accurate data (the observed FED, Rosenfeld et al. 1993; Anagnostou et al. 1999; Ebert 2001). New observation operators are then developed based on the third-order nonlinear fits of the sorted FED observations and model graupel masses. Since the FED observations are dominated by zero values (about 95% in both cases), the zero FED observations and the corresponding graupel masses (after sorting) are discarded to facilitate the regression process.

The new operators developed based on data of the MCS, supercell, and combination of two cases are, respectively,

$$\text{FED} = -1.173 \times 10^{-28} (\text{GM})^3 + 1.211 \times 10^{-18} (\text{GM})^2 \\ + 4.005 \times 10^{-9} (\text{GM}) + 0.733, \qquad (2)$$

$$\text{FED} = -3.762 \times 10^{-28} (\text{GM})^3 + 2.860 \times 10^{-18} (\text{GM})^2 \\ + 2.971 \times 10^{-9} (\text{GM}) + 0.269, \qquad (3)$$

$$\text{FED} = -2.988 \times 10^{-28} (\text{GM})^3 + 2.511 \times 10^{-18} (\text{GM})^2 \\ + 2.833 \times 10^{-9} (\text{GM}) + 0.720, \qquad (4)$$
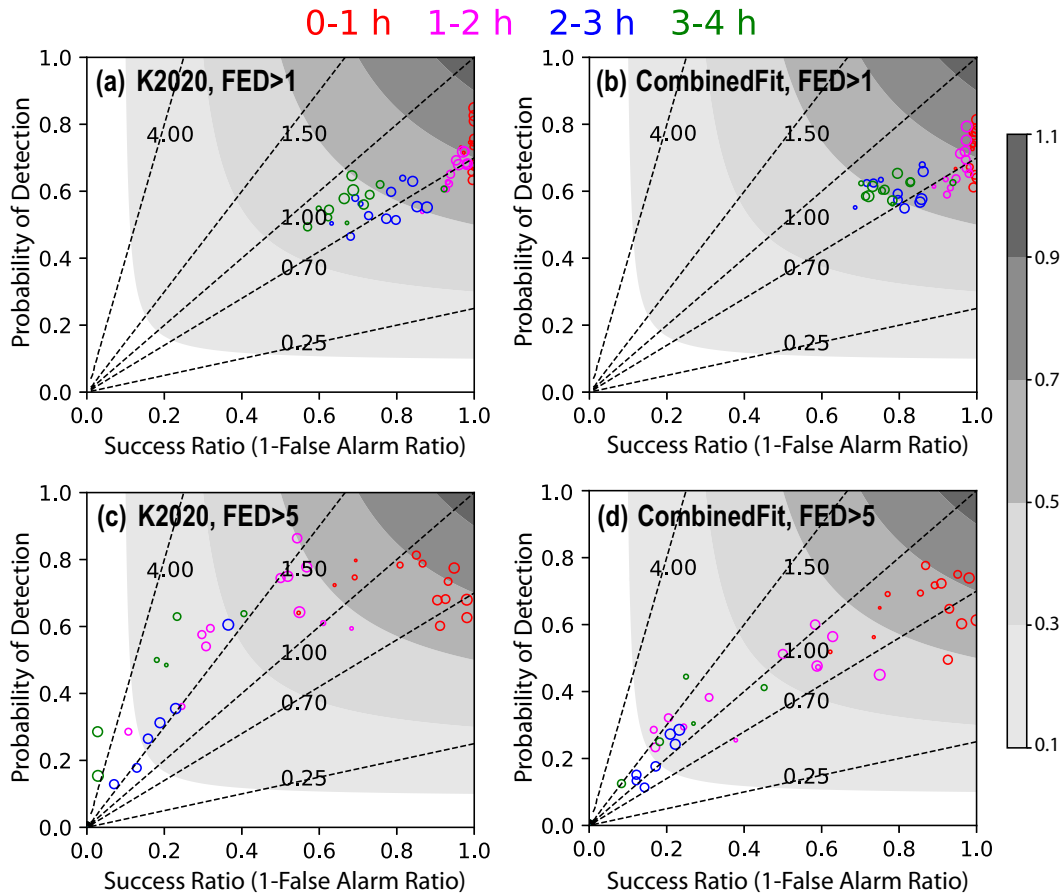
FIG. 8. Performance diagrams of the 0–4-h FED forecasts (with a 5-min interval) from (a),(c) K2020; (b),(d) Combined-Fit for thresholds of (a),(b) 1 and (c),(d) 5 flashes min^{-1} $pixel^{-1}$, respectively, for the MCS case. Performance of different forecasting times are indicated by different colors and the sizes of the circles represent the times within the specified forecast intervals (larger marker size represents longer forecast time). The shaded contour represents the critical success index.

Here GM is, again, the graupel mass (in kg) over a 15×15 km^2 area in the horizontal centered on the observation pixels. The zeroth-order intercepts in the third-order polynomial fits for MCSFit, SupercellFit, and CombinedFit are 0.733, 0.269, 0.720 flash min^{-1} $pixel^{-1}$, respectively. To avoid introducing spurious graupel analysis in clear-air region during DA, a zero-intercept linear function is applied for small graupel mass values (i.e., GM < 5×10^8 kg). The final observation operators (piecewise continuous functions) for the MCS, supercell, and combined cases are

$$FED = \begin{cases} 6.047 \times 10^{-9} \, (GM) & (GM < 5 \times 10^8 \, kg) \\ Eq. \ (2) & (GM \geq 5 \times 10^8 \, kg) \end{cases}, \quad (5)$$

$$FED = \begin{cases} 4.845 \times 10^{-9} \, (GM) & (GM < 5 \times 10^8 \, kg) \\ Eq. \ (3) & (GM \geq 5 \times 10^8 \, kg) \end{cases}, \quad (6)$$

$$FED = \begin{cases} 5.453 \times 10^{-9} \, (GM) & (GM < 5 \times 10^8 \, kg) \\ Eq. \ (4) & (GM \geq 5 \times 10^8 \, kg) \end{cases}. \quad (7)$$

In Fig. 4a, the original linear operator of Allen et al. (2016) (labeled A2016) and the tuned version of Kong et al. (2020)

(labeled K2020) are plotted together with newly fitted ones using first- and third-order polynomials. The FEDs given by K2020 are notably larger than the observations for both cases while those given by A2016 are twice as large (Fig. 4a). The third-order polynomial curves match the observations very well throughout the observation range, while the first-order curves do not fit observations well for larger values. The nonlinear fits for the MCS and supercell case are similar for lower FED values (e.g., <8 flashes min^{-1} $pixel^{-1}$, Fig. 4a) but different at larger magnitudes. The combined fit (CombinedFit_3rdOrder) is much closer to the supercell fit (SupercellFit_3rdOrder) because the supercell case contains more points with larger FED values compared to the MCS case (Fig. 4b), providing more constraints to the fitting at the larger value end. The estimated FEDs from the first-order polynomial fits for both single and combined cases are much closer to the observations relative to K2020 but are generally smaller than those from the third-order fits for higher FED values.

Figure 5 shows the box-and-whisker plots (Massart et al. 2005) of the deviations of the sorted FED simulations (i.e., the observational priors of FED) from the sorted observations for K2020, MCSFit_1stOrder (or SupercellFit_1stOrder),
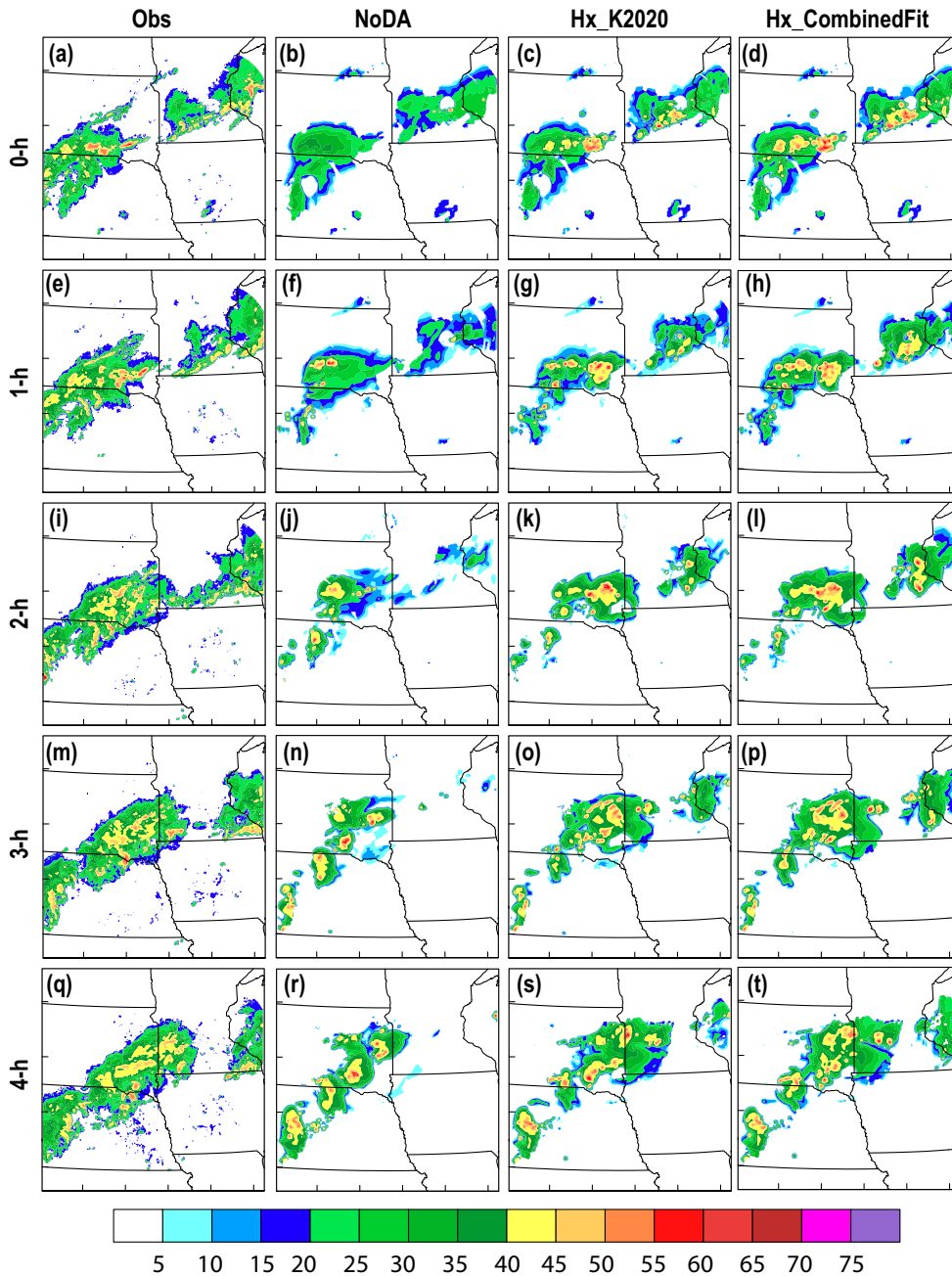
FIG. 9. Composite reflectivity fields (dBZ) for (a),(e),(i),(m),(q) observations; (b),(f),(j),(n),(r) forecasts from NoDA; (c),(g),(k),(o),(s) K2020; and (d),(h),(l),(p),(t) CombinedFit after (a)–(d) 0-, (e)–(h) 1-, (i)–(l) 2-, (m)–(p) 3-, and (q)–(t) 4-h free forecasts following the 1-h DA period for the MCS case (valid at 0000–0400 UTC).

and MCSFit_3rdOrder (or SupercellFit_3rdOrder), respectively, for the MCS (supercell) case and the combination of two storm cases. The interquartile range (IQR, distance between the 25th and 75th percentiles) and the median from the third-order fits are slightly smaller than those from the first-order fits, and both are obviously smaller than the unfitted estimation (K2020) both for the individual and combined cases, which is indicative of greater error reduction for the nonlinear fits. The number of outliers for the third-order fits are fewer than those from the first-

order fits. Based on these results, the observation operators using third-order polynomial fits are used in the new DA experiments, whose results are compared with those using the original observation operator of K2020.

5. Results using different observation operators

In this section, we examine the results of EnKF analyses and subsequent forecasts from the experiments that use the
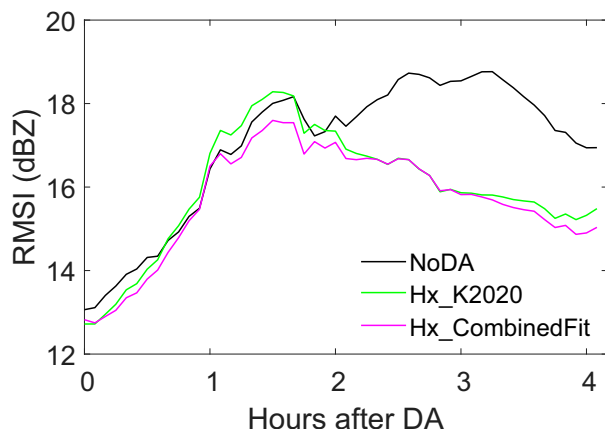
FIG. 10. RMSIs of 0–4-h composite reflectivity forecasts (dBZ) from experiments NoDA, K2020, and CombinedFit for the MCS case.

new FED observation operators (MCSFit, SupercellFit and CombinedFit) and the K2020 operator, for the two storm cases, with NoDA experiments used as reference. The results of the DA experiments are evaluated in terms of FED and radar reflectivity observations. See Table 1 for the list of experiments evaluated.

a. Case 1: 12 July 2018 MCS

To directly compare simulated and observed FED rates, observation operators are applied to the analyses and 0–4-h forecasts of the graupel mass mixing ratio fields to obtain the observation priors and posteriors. Overall, all DA experiments (shown for K2020 and CombinedFit in Fig. 6) produce similar FED analyses after 1-h DA and show obvious improvements over the NoDA case in terms of the areal coverage and intensity of the analyzed FED rates (Figs. 6a–d). MCSFit has similar FED values as CombinedFit and is therefore not shown. The FED rates in 1–4-h forecasts of CombinedFit are smaller than those of K2020 and are in closer agreement with observations (Figs. 6e–t). K2020 forecasts overestimate the FED rates in intense convective regions relative to CombinedFit. Both CombinedFit and K2020 outperform NoDA in better capturing the FEDs associated with the northeastern portion of the MCS.

The root-mean-square innovations (RMSIs; where innovation is difference from observations) of the analyzed FED during the DA period (−1 to 0 h) and the FED during the 0–4-h forecasts from experiments K2020, MCSFit, SupercellFit, and CombinedFit are plotted in Fig. 7. Relative to NoDA, all DA experiments reduce the errors of the forecast FED. Relative to K2020, MCSFit, SupercellFit, and CombinedFit reduce the RMSIs of the FED forecasts by about 25%. MCSFit performs slightly better than CombinedFit and SupercellFit in terms of the FED forecast RMSIs for this MCS case, apparently because MCSFit uses the operator fitted to data of the MCS case while the other two use operators fitted to supercell data also. The differences among the three experiments are much smaller compared to control experiment K2020. For the MCS case, most FED observations are less than 15 flashes min^{-1} pixel^{-1}

(Fig. 4b), and the fitting curves among MCSFit_3rdOrder, SupercellFit_3rdOrder, and CombinedFit_3rdOrder are not significantly different for small FEDs (<15 flashes min^{-1} pixel^{-1}, Fig. 4a), resulting in small difference in experiments MCSFit, SupercellFit, and CombinedFit. For generality, only CombinedFit will be considered in the remaining evaluations.

The performance diagrams (Roebber 2009) of FED forecasts up to 4 h at 5-min intervals against GLM FED observations are shown in Fig. 8 for K2020 and CombinedFit. For the lower FED threshold of 1 flash min^{-1} pixel^{-1} (Figs. 8a,b), CombinedFit performs similarly to K2020, with the frequency biases of 3–4-h forecasts of K2020 being slightly closer to unity. The critical success indices (CSI) of 2–4-h FED forecasts of CombinedFit are slightly larger than those of K2020. For the higher threshold of 5 flashes min^{-1} pixel^{-1}, CombinedFit clearly outperforms K2020 in terms of lower frequency biases (closer to unity), especially for >1-h FED forecasts.

To evaluate how well different DA experiments capture the structure and intensity of the storms, the 0–4-h reflectivity forecasts are compared with radar observations. Overall, all DA experiments are able to produce the observed west-southwest–east-northeast-oriented precipitation band of the MCS. All DA experiments produce reflectivity forecasts much closer to the observations relative to NoDA. NoDA fails to capture the reflectivity echoes in the northeastern part of the MCS. The areal coverage of the MCS from CombinedFit is larger than K2020 and is more consistent with the observations, especially for the northeast precipitation band in 1–4-h forecasts (Fig. 9).

Figure 10 shows the RMSIs of composite reflectivity fields from different experiments. CombinedFit produces somewhat smaller RMSIs than K2020, and both produce notably smaller RMSIs than NoDA, especially for reflectivity forecasts after 2 h, after which NoDA fails to capture the reflectivity echoes over the northeastern portions of the MCS (Fig. 9), leading to much larger RMSIs. Compared to the RMSIs of FED shown in Fig. 7, the improvement in reflectivity RMSIs of CombinedFit over K2020 is much smaller, suggesting that the large FED RMSIs of K2020 were mostly due to biases associated with the old linear operator while the forecast reflectivities are more similar.

b. Case 2: 1 May 2018 supercell storm

For the supercell storm case, the FED analyses and forecasts from NoDA, K2020, and CombinedFit are compared in Fig. 11. As can be seen, both intensity and areal coverage of FED analyses of CombinedFit (Fig. 11d) at the end of 1-h DA are similar to those of K2020 (Fig. 11c), and both are much smaller than those of NoDA (Fig. 11b) and are closer to observations (Fig. 11a). The assimilation of zero FEDs in DA experiments help suppress the overprediction of FED in NoDA. For the 0.5–4-h forecasts, the maximum FED values (shown in panels of Fig. 11) from CombinedFit are smaller or much smaller than those of K2020 and in better agreement with observations. Relative to K2020 and CombinedFit, NoDA (that also employs the same linear operator as K2020 for FED simulation) overpredicts the FED rates, and the
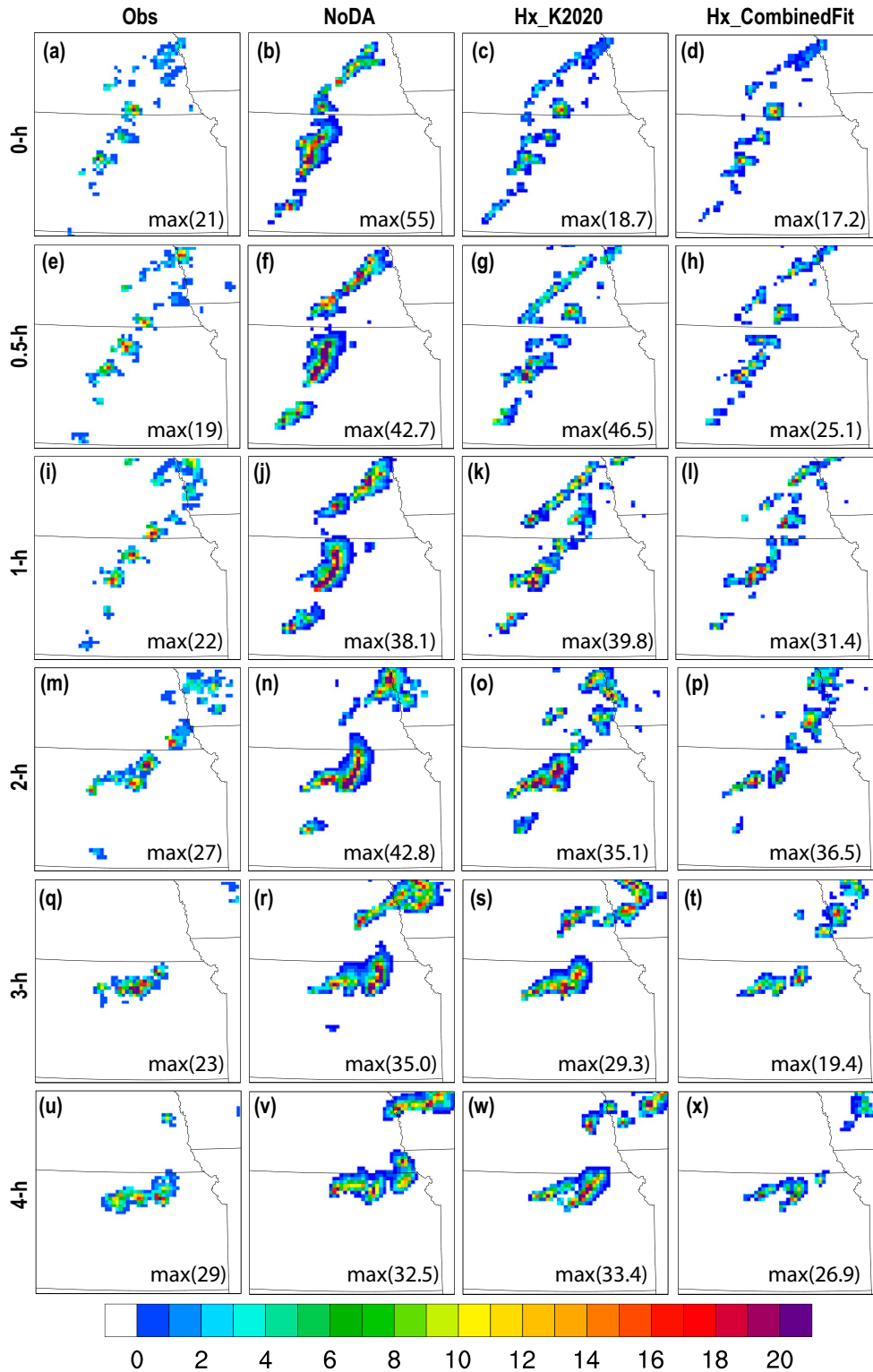
FIG. 11. Horizontal cross sections of (a),(e),(i),(m),(q),(u) 1-min FED observations (units: flashes min^{-1} $pixel^{-1}$), and FED forecasts from (b),(f),(j),(n),(r),(v) CTRL; (c),(g),(k),(o),(s),(w) K2020; and (d),(h),(l),(p),(t),(x) CombinedFit after (a)– (d) 0-, (e)–(h) 0.5-, (i)–(l) 1-, (m)–(p) 2-, (q)–(t) 3-, and (u)–(x) 4-h free forecasts following the 1-h cycled DA period (valid from 2300 1 May to 0100 UTC 2 May 2018) for the supercell case. The maximum values of each field plotted are given in the figure.
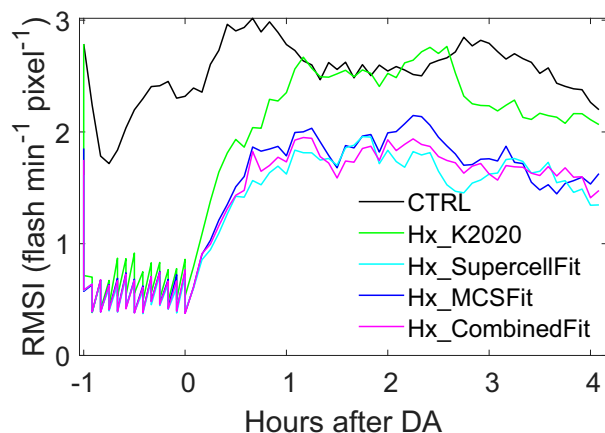
FIG. 12. RMSIs of the FED analyses and forecasts (flashes min^{-1} pixel^{-1}) within 1-h DA (corresponding to −1 to 0 h after DA) and 0–4-h FED forecasts from experiments NoDA, K2020, MCSFit, supercellFit, and CombinedFit for the supercell case.

overall structure of the storms resembles that of a bow echo more instead of the individual supercellullar storm cells seen in the observations (e.g., in Figs. 11f,j,n). The areal coverages of the FED forecasts of CombinedFit are also smaller than those of K2020 and also more consistent with observations. MCSFit significantly underestimates the FED forecasts for the simulation of high FEDs (e.g., >15 flashes min^{-1} pixel^{-1}), and the differences between SupercellFit and CombinedFit are small and, thus, only the results of CombinedFit are shown in Fig. 11. As has been discussed earlier, CombinedFit_3rdOrder is much closer to SupercellFit_3rdOrder relative to MCSFit_3rdOrder, since the supercell case contains more points with larger FED values compared to the MCS case (Fig. 4b), providing more constraints to the fitting and more similar FED estimates at the larger value end.

The RMSIs of the FED analyses during DA cycles and 0–4-h forecasts for NoDA, K2020, MCSFit, SupercellFit, and CombinedFit are shown in Fig. 12 for the supercell case. Again, all DA experiments produce notably smaller RMSIs than NoDA, underlining the positive impacts of FED DA on storm forecasts. Additionally, the DA experiments that utilize the new observation operators (MCSFit, SupercellFit and CombinedFit) produce much smaller RMSIs than K2020 using the old linear operator, especially for FED forecasts (Fig. 12). The new operators reduce the RMSIs of 0–4-h FED forecasts by about 25% relative to using the original linear operator. The RMSIs of CombinedFit lie between those of SupercellFit and MCSFit most times during the forecast, although they are the lowest between 1.5 and 2 h, and the highest for about 10 min near around 3.7 h (Fig. 12) but the differences are generally small. The results indicate a small level of sensitivity to the data used to fit the operator. Given that CombinedFit is more general, we will again focus the remaining evaluations for the supercell case on CombinedFit.

The performance diagrams of 0–4-h FED forecasts from experiments K2020 and CombinedFit verified against FED observations are plotted in Fig. 13 for the supercell case. For both

low and high thresholds (i.e., 1 and 5 flashes min^{-1} pixel^{-1}), CombinedFit produces the least frequency bias (with values closer to unity), especially for the 2–4-h FED forecasts.

The composite reflectivity forecasts of NoDA, K2020, and CombinedFit are compared against radar observations in Fig. 14. Relative to DA experiments K2020 and CombinedFit, as with the FED forecasts shown earlier, NoDA produces storms that are less isolated in nature and more like a bow echo in the central part of plotting domain that do not match observations. For the reflectivity analyses at the end of 1-h DA, CombinedFit produces more intense convection in terms of reflectivity (>45 dBZ) compared to K2020 and is in better agreement with observations. The smaller FED rates analyzed in CombinedFit relative to K2020 (Fig. 11) result in more graupel mass being added into the model during DA in K2020 and, therefore, producing overall more vigorous storms at the final analysis time (cf. Figs. 14c,d). The relatively weak storms analyzed in K2020, however, develop quickly during the first few hours of forecast, resulting in similar storm forecasts as in CombinedFit in later hours. However, CombinedFit produces fewer spurious storms relative to K2020 and NoDA (cf. magenta ellipses in Figs. 14r,s,tu,v,x). The FED estimated in CombinedFit is smaller than that from K2020, resulting in more effective suppression of spurious FED relative to K2020 over the northern portion of the domain (not shown), since less correction to the background forecast (which is more accurate) is needed when zero FEDs are assimilated in CombinedFit relative to K2020.

Figure 15 shows the RMSIs of the composite reflectivity field for different experiments for the supercell case. The RMSIs for K2020 and CombinedFit are again notably smaller than those of NoDA, indicating the ability of GLM FED DA in improving supercell storm forecasts in terms radar reflectivity also. Different from the reflectivity forecasts of the MCS case where the RMSIs of K2020 and CombinedFit are similar (Fig. 12), for this supercell case, CombinedFit further reduces the RMSIs by about 23% in the reflectivity forecasts beyond the initial 0.5 h of forecast, compared to K2020. The large impact of the new operator in this supercell case is presumably because supercell storms have large variability in intensity (see Fig. 4). There are more high flash rates in the supercell and the linear operator has larger errors at high flash rates (Fig. 4). The new nonlinear operator therefore not only improves the forecasts of FED but also reflectivity.

In summary, employing the new nonlinear operator (CombinedFit) reduces the biases in FED forecasts compared to the linear tuned operator used in K2020 for both the MCS and supercell storm cases for up to 4 h into the forecast. The errors in forecast reflectivity are also significantly smaller with the new operator for the supercell case, although the differences are much smaller for the MCS case where convection is generally weaker. The performance differences between MCSFit, SupercellFit, and CombinedFit are small for both cases and therefore the nonlinear operator obtained by fitting FED observations to aggregated forecast graupel mass data are preferred for general applications. The same procedure can be used to obtain an even more general operator by fitting data from a large number of cases in the future.
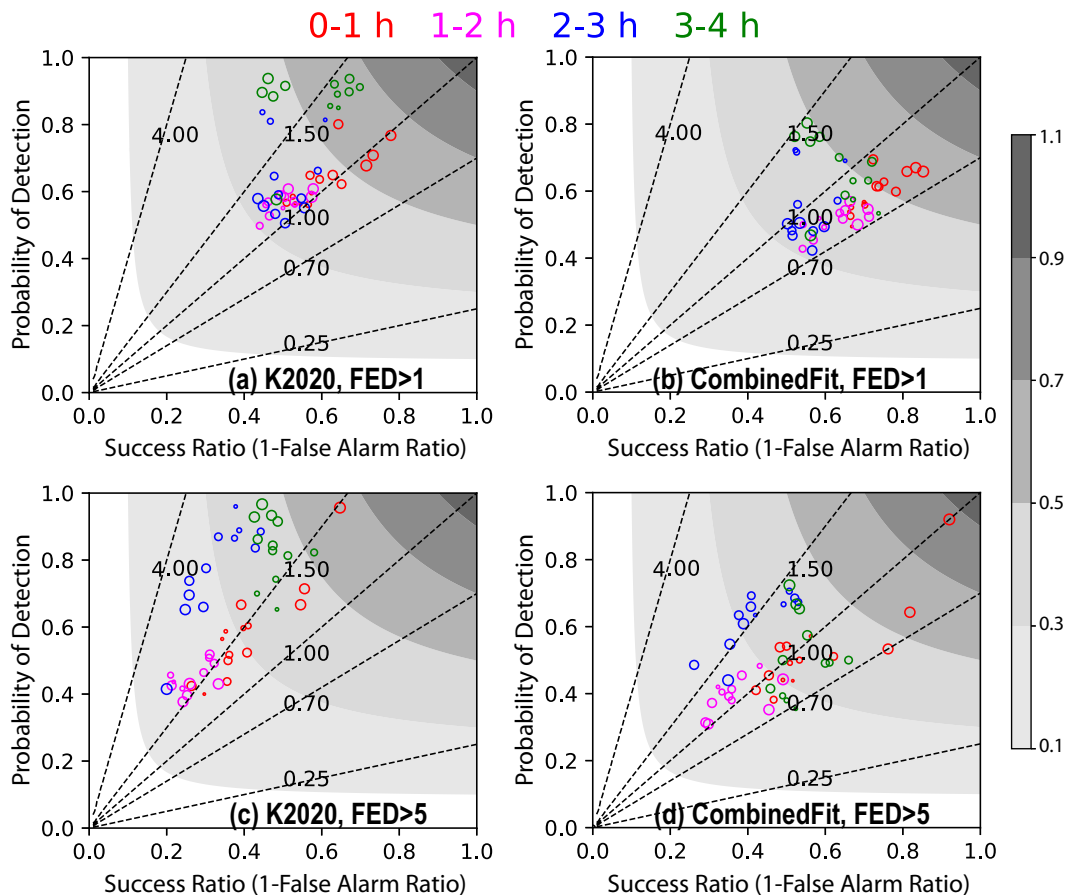
FIG. 13. Performance diagrams of the 0–6-h FED forecasts (with a 5-min interval) from (a),(c) K2020; (b),(d) Combined-Fit for thresholds of (a),(b) 1 and (c),(d) 5 flashes min^{-1} pixel^{-1}, respectively, for the supercell case. Different forecasting times are indicated by different colors and the sizes of the circles represent the times within the specified forecast intervals (larger marker size represents longer forecast time). The shaded contours indicate critical success index.

6. Summary and conclusions

In this study, new observation operators are developed to assimilate GOES-R GLM FED data into a convection-allowing NWP model. The new operators are developed based on third-order polynomial fit between FED observations and forecasts of graupel mass of an MCS and a supercell storm case initialized from cycled FED DA using linear operator. The choice of the forecasts used to do the fitting is mainly for convenience in this proof of concept. In more general practice, a large sample of forecasts starting from best initial conditions possible (e.g., those that assimilate radar data also) should be used. The same version of model (especially with the same microphysics) should be used since the fitting procedure also plays a role of bias correction. In this study, the forecasts used for the fitting are produced from initial conditions obtained after FED observations are assimilated over a 1-h period at 5-min intervals, using a tuned linear FED observation operator (denoted K2020) from Kong et al. (2020). Both observed FEDs and forecast graupel masses are sorted according to their values before fitting. The sorting alleviates the impact of mismatches in location due to spatial displacement errors, which can be

very large in convection-allowing models. The fitting is performed for the MCS and supercell data alone, and for the combined data of the two cases, leading to three operators labeled MCSFit, SupercellFit and CombinedFit. The same names are used to denote the three DA experiments using the corresponding operators, and the experiments are run for the MCS and supercell cases using the same DA configuration of Kong et al. (2020). The DA experiments with the new operators are compared with the control run K2020 that used the linear operator as well as an experiment not assimilating any FED data (NoDA). Forecasts up to 4 h are evaluated against FED and radar reflectivity observations, through both subjective comparisons of their fields and the root-mean-square innovations (RMSIs). Performance diagrams of FED forecasts are also shown to assess additional aspects of the forecasts, including frequency biases. The major findings are summarized as follows:

- All FED DA experiments produce improved (0–4 h) FED forecasts relative to the experiment that does not assimilate any data. Assimilating FED data helps to better capture the FED analyses/forecasts over the northeastern portions
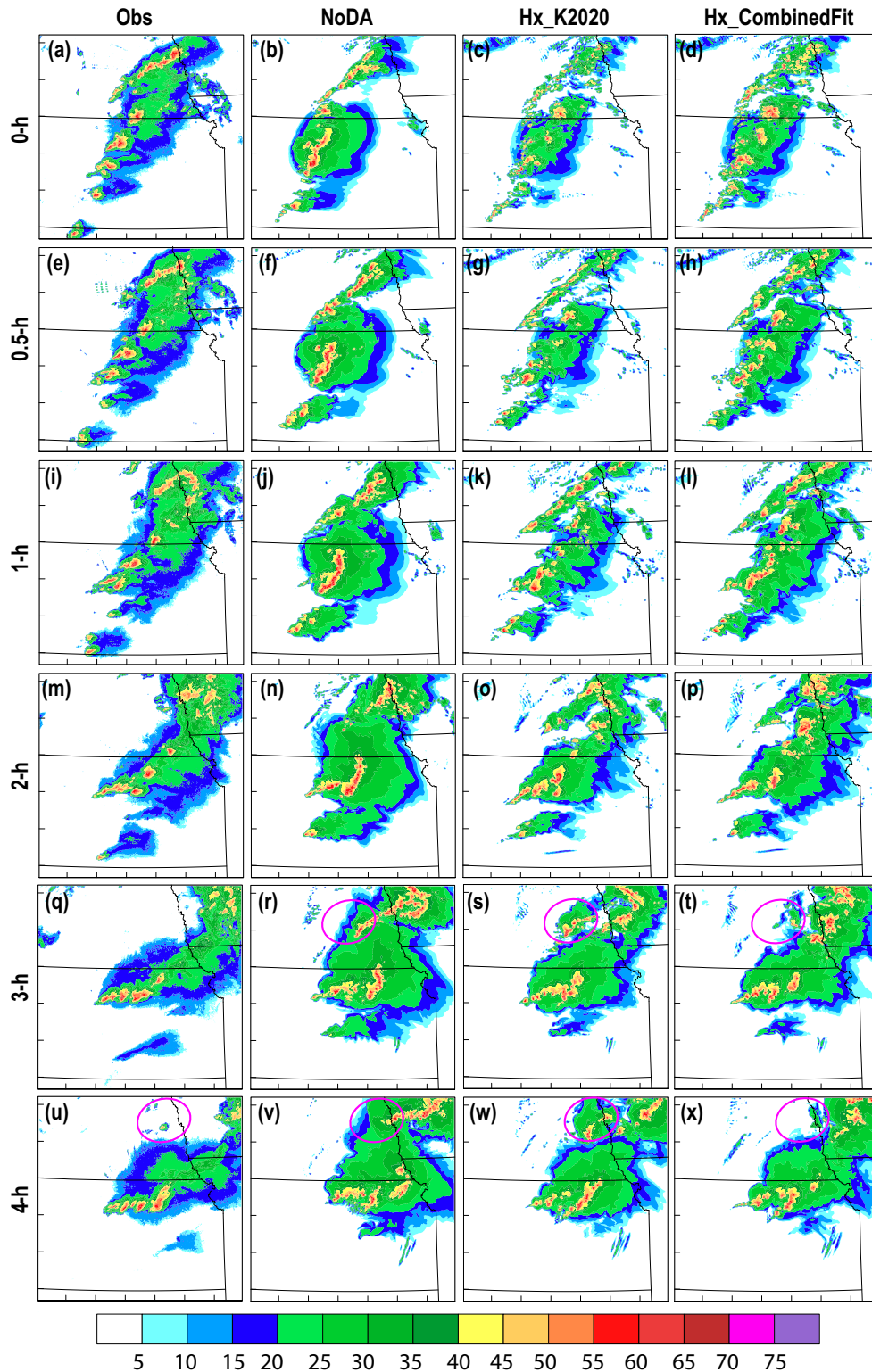
FIG. 14. Composite reflectivity fields (dB*Z*) for (a),(e),(i),(m),(q),(u) observations; (b),(f),(j),(n),(r),(v) forecasts from NoDA; (c),(g),(k),(o),(s),(w) K2020; and (d),(h),(l),(p),(t),(x) CombinedFit after (a)–(d) 0-, (e)–(h) 0.5-, (i)–(l) 1-, (m)–(p) 2-, (q)–(t) 3-, and (u)–(x) 4-h free forecasts following the 1-h cycled DA period for the supercell case (valid at 2300 UTC 1 May–0100 UTC 2 May 2018).
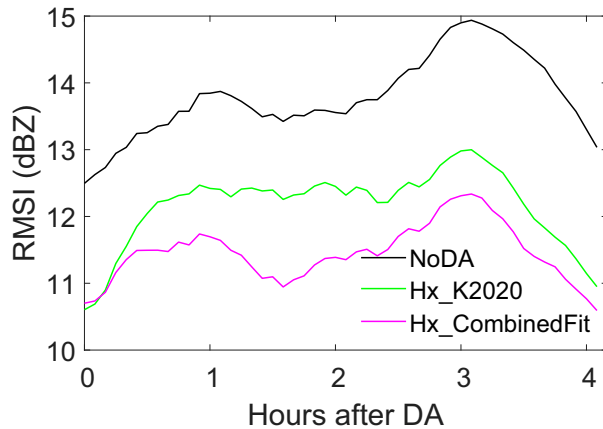
FIG. 15. RMSIs of 0–4-h composite reflectivity forecasts (dBZ) from experiments NoDA, K2020, SupercellFit, MCSFit, and CombinedFit for the supercell case.

of the MCS and the individual supercell storm cells than the experiment that does not assimilate any data. The control run using the original linear FED operator (K2020) significantly overestimates the FED rates in intense convection regions relative to using the fitted operators for both storm cases. The DA experiments that use the operators fitted based on single or combined cases do not have much difference for the MCS case; while for the supercell case, SupercellFit is more similar to the CombinedFit relative to the MCSFit, since the supercell case contains more points with larger FED values compared to the MCS case, providing more constraints to the fitting and more similar FED estimates at the larger value end.

- RMSIs of the 0–4-h FED forecasts after DA are compared in experiments K2020, MCSFit (SupercellFit), and CombinedFit for both the MCS and supercell storm case. When compared with NoDA, all the DA experiments greatly reduce the errors in FED forecasts, underscoring the positive impacts of FED DA on lightning forecasts. Relative to K2020, the nonlinear operator experiments MCSFit/SupercellFit and CombinedFit all reduce the RMSIs of the 0–4-h FED forecasts by about 25% for both storm cases. The RMSI differences between MCSFit, SupercellFit and CombinedFit remain small. For generality, only CombinedFit will be considered in the remaining evaluations.
- The performance diagrams of the 0–4-h FED forecasts verified against the FED observations were also compared between experiments K2020 and CombinedFit. For lower threshold (i.e., 1 flash min^{-1} pixel^{-1}), CombinedFit performs similarly to K2020. For higher threshold (5 flashes min^{-1} pixel^{-1}), however, CombinedFit was shown to clearly outperform K2020 in terms of frequency bias values being closer to unity (i.e., unbiased).
- For the MCS case, all experiments are able to produce the west-southwest–east-northeast-oriented precipitation band. Both DA experiments produce reflectivity forecasts that are in notably better agreement with the observations relative to NoDA. The areal coverage of the MCS from CombinedFit is

larger than K2020 and is more consistent with the observations, especially for the northeastern precipitation band at 1–4-h forecasts. For the reflectivity analyses, CombinedFit produces more intense convection in terms of higher reflectivity values (>45 dBZ) relative to K2020 and is noticeably closer to the observations. Smaller FED estimation in CombinedFit relative to K2020 results in more graupel mass being added into the domain during the DA leading to overall more vigorous storms. Meanwhile, CombinedFit produces fewer spurious storms relative to K2020 as well as NoDA.

- Assimilating FED observations using the linear operator greatly reduces the RMSIs of the composite reflectivity field relative to NoDA. Using the new fitted operator further reduces the RMSIs in the reflectivity forecasts. Especially for the supercell storm case, the RMSIs of the 0.5–4-h reflectivity forecasts are reduced by 23%. More accurate FED simulation (corresponding to less amount of FED to be corrected by zero FED) results in more effective suppression of the spurious storms in the analyses and the forecasts relative to using the old linear operator.

In summary, this proof-of-concept study provides a new way to develop the nonlinear lightning observation operator. Employing the new nonlinear operator (CombinedFit) reduces the biases and RMSIs in the 0–4-h forecasts of the FED and reflectivity fields relative to using the tuned linear operator (K2020) for both the MCS and supercell storm cases. The performance differences among the DA experiments that use the operators fitted for the MCS case, supercell case, and the combination of two cases are not significant. The operator fitted for the supercell case is much closer to the combined fit, since the supercell case contains more points with larger FED values compared to the MCS case, providing more constraints to the fitting and more similar FED estimates at the larger value end. Thus, the observation operator based on the combined fitting curve is recommended for more general applications.

Despite some of the encouraging results that have been documented in this work with the new methodology to develop nonlinear FED operators, there still exist some nonnegligible degrees of variability in the regression coefficients across different types of convective modes. More research is thus needed to develop the operator across a wider sample of cases covering all storm types to obtain a more general operator. It is hoped that through additional complementary studies, the rich storm-scale information provided by the GOES-R sensors can be better leveraged toward a more systematic usage of total lightning data in operational NWP undertakings.

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 (OSCER) hosted at the University of Oklahoma.

Data availability statement. The GLM data used in this manuscript can be ordered and downloaded at https://www.ncdc.noaa.gov/airs-web/search. The SREF and NAM data used in this study are available at https://doi.org/10.7910/DVN/PXL4JN.

REFERENCES

Aksoy, A., 2010: A multicase comparative assessment of the ensemble Kalman filter for assimilation of radar observations. Part II: Short-range ensemble forecasts. *Mon. Wea. Rev.*, **138**, 1273–1292, https://doi.org/10.1175/2009MWR3086.1.

——, D. C. Dowell, and C. Snyder, 2009: A multicase comparative assessment of the ensemble Kalman filter for assimilation of radar observations. Part I: Storm-scale analyses. *Mon. Wea. Rev.*, **137**, 1805–1824, https://doi.org/10.1175/2008MWR2691.1.

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.

Anagnostou, E. N., A. J. Negri, and R. F. Adler, 1999: A satellite infrared technique for diurnal rainfall variability studies. *J. Geophys. Res.*, **104**, 31 477–31 488, https://doi.org/10.1029/1999JD900157.

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.

Apodaca, K., and M. Zupanski, 2018: Variational and hybrid (EnVar) methodologies to add the capability to assimilate *GOES-16*/GLM observations into GDAS. *JCSDA Quart.*, **58**, 12–20.

——, ——, M. DeMaria, J. A. Knaff, and L. D. Grasso, 2014: Development of a hybrid variational-ensemble data assimilation technique for observed lightning tested in a mesoscale model. *Nonlinear Processes Geophys.*, **21**, 1027–1041, https://doi.org/10.5194/npg-21-1027-2014.

Bauer, P., A. J. Geer, P. Lopez, and D. Salmond, 2010: Direct 4D-Var assimilation of all-sky radiances. Part I: Implementation. *Quart. J. Roy. Meteor. Soc.*, **136**, 1868–1885, https://doi.org/10.1002/qj.659.

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.

Chang, W., M. Haran, R. Olson, and K. Keller, 2014: Fast dimension-reduced climate model calibration and the effect of data aggregation. *Ann. Appl. Stat.*, **8**, 649–673, https://doi.org/10.1214/14-AOAS733.

Chen, Y., F. Weng, Y. Han, and Q. Liu, 2008: Validation of the community radiative transfer model by using CloudSat data. *J. Geophys. Res.*, **113**, D00A03, https://doi.org/10.1029/2007JD009561.

——, H. Wang, J. Min, X.-Y. Huang, P. Minnis, R. Zhang, J. Haggerty, and R. Palikonda, 2015: Variational assimilation of cloud liquid/ice water path and its impact on NWP. *J. Appl.*

Meteor. Climatol., **54**, 1809–1825, https://doi.org/10.1175/JAMC-D-14-0243.1.

——, Z. Yu, W. Han, J. He, and M. Chen, 2020: Case study of a retrieval method of 3D proxy reflectivity from FY-4A lightning data and its impact on the assimilation and forecasting for severe rainfall storms. *Remote Sens.*, **12**, 1165, https://doi.org/10.3390/rs12071165.

Chen, Z. X., J. Z. Sun, X. S. Qie, Y. Zhang, Z. M. Ying, X. Xiao, and D. J. Cao, 2020: A method to update model kinematic states by assimilating satellite-observed total lightning data to improve convective analysis and forecasting. *J. Geophys. Res. Atmos.*, **125**, e2020JD033330, https://doi.org/10.1029/2020JD033330.

Chevallier, F., P. Lopez, A. M. Tompkins, M. Janiskova, and E. Moreau, 2004: The capability of 4D-Var systems to assimilate cloud-affected satellite infrared radiances. *Quart. J. Roy. Meteor. Soc.*, **130**, 917–932, https://doi.org/10.1256/qj.03.113.

Derber, J. C., and W. S. Wu, 1998: The use of TOVS cloud-cleared radiances in the NCEP SSI analysis system. *Mon. Wea. Rev.*, **126**, 2287–2299, https://doi.org/10.1175/1520-0493(1998)126<2287:TUOTCC>2.0.CO;2.

Doviak, R. J., V. Bringi, A. Ryzhkov, and A. Zahrai, 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.

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.

Ebert, E. E., 2001: Ability of a poor man's ensemble to predict the probability and distribution of precipitation. *Mon. Wea. Rev.*, **129**, 2461–2480, https://doi.org/10.1175/1520-0493(2001)129<2461:AOAPMS>2.0.CO;2.

Evensen, G., 2003: The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dyn.*, **53**, 343–367, https://doi.org/10.1007/s10236-003-0036-9.

Eyre, J. R., G. A. Kelly, A. P. McNally, E. Andersson, and A. Persson, 1993: Assimilation of TOVS radiance information through one-dimensional variational analysis. *Quart. J. Roy. Meteor. Soc.*, **119**, 1427–1463, https://doi.org/10.1002/qj.49711951411.

Fierro, A. O., 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.

——, ——, ——, 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.

Gan, R. H., Y. Yang, X. B. Qiu, R. C. Wang, X. X. Qiu, and L. J. Zhu, 2021: Assimilation of the maximum vertical velocity converted from total lightning data through the EnSRF method. *J. Geophys. Res. Atmos.*, **126**, e2020JD034300, https://doi.org/10.1029/2020JD034300.

Gao, J., M. Xue, K. Brewster, and K. Droegemeier, 2004: 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.

——, ——, 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.

——, C. Fu, D. J. Stensrud, and J. Kain, 2016: OSSEs for an ensemble 3DVAR data assimilation system with radar observations of convective storms. *J. Atmos. Sci.*, **73**, 2403–2426, https://doi.org/10.1175/JAS-D-15-0311.1.

Geer, A. J., and P. Bauer, 2011: Observation errors in all-sky data assimilation. *Quart. J. Roy. Meteor. Soc.*, **137**, 2024–2037, https://doi.org/10.1002/qj.830.

——, ——, and P. Lopez, 2010: Direct 4D-Var assimilation of all-sky radiances. Part II: Assessment. *Quart. J. Roy. Meteor. Soc.*, **136**, 1886–1905, https://doi.org/10.1002/qj.681.

——, and Coauthors, 2017: The growing impact of satellite observations sensitive to humidity, cloud and precipitation. *Quart. J. Roy. Meteor. Soc.*, **143**, 3189–3206, https://doi.org/10.1002/qj.3172.

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

——, S. Migliorini, and M. Matricardi, 2019: All-sky assimilation of infrared radiances sensitive to mid- and upper-tropospheric moisture and cloud. *Atmos. Meas. Tech.*, **12**, 4903–4929, https://doi.org/10.5194/amt-12-4903-2019.

Goodman, S. J., and Coauthors, 2013: The GOES-R Geostationary Lightning Mapper (GLM). *Atmos. Res.*, **125**, 34–49, https://doi.org/10.1016/j.atmosres.2013.01.006.

Guerbette, J., J. F. Mahfouf, and M. Plu, 2016: Towards the assimilation of all-sky microwave radiances from the SAPHIR humidity sounder in a limited area NWP model over tropical regions. *Tellus*, **68A**, 28620, https://doi.org/10.3402/tellusa.v68.28620.

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

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.

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.

Jones, T. A., and D. J. Stensrud, 2015: Assimilating cloud water path as a function of model cloud microphysics in an idealized simulation. *Mon. Wea. Rev.*, **143**, 2052–2081, https://doi.org/10.1175/MWR-D-14-00266.1.

——, ——, P. Minnis, and R. Palikonda, 2013: Evaluation of a forward operator to assimilate cloud water path into WRF-DART. *Mon. Wea. Rev.*, **141**, 2272–2289, https://doi.org/10.1175/MWR-D-12-00238.1.

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

——, and Coauthors, 2020: Assimilation of *GOES-16* radiances and retrievals into the Warn-on-Forecast System. *Mon. Wea. Rev.*, **148**, 1829–1859, https://doi.org/10.1175/MWR-D-19-0379.1.

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.

Kazumori, M., A. J. Geer, and S. J. English, 2016: Effects of all-sky assimilation of GCOM-W/AMSR2 radiances in the ECMWF numerical weather prediction system. *Quart. J. Roy. Meteor. Soc.*, **142**, 721–737, https://doi.org/10.1002/qj.2669.

Kennedy, M. C., and A. O'Hagan, 2001: Bayesian calibration of computer models. *J. Roy. Stat. Soc. B*, **63**, 425–450, https://doi.org/10.1111/1467-9868.00294.

Kleist, D. T., D. F. Parrish, J. C. Derber, R. Treadon, R. M. Errico, and R. H. Yang, 2009: Improving incremental balance in the GSI 3DVAR analysis system. *Mon. Wea. Rev.*, **137**, 1046–1060, https://doi.org/10.1175/2008MWR2623.1.

Kong, R., M. Xue, and C. S. 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.

——, ——, A. O. Fierro, Y. S. Jung, C. S. Liu, E. R. Mansell, and D. R. MacGorman, 2020: 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. *Mon. Wea. Rev.*, **148**, 2111–2133, https://doi.org/10.1175/MWR-D-19-0192.1.

——, ——, C. S. Liu, and Y. Jung, 2021: Comparisons of hybrid En3DVar with 3DVar and EnKF for radar data assimilation: Tests with the 10 May 2010 Oklahoma tornado outbreak. *Mon. Wea. Rev.*, **149**, 21–40, https://doi.org/10.1175/MWR-D-20-0053.1.

Liu, C. S., 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.

——, ——, and ——, 2020: Direct variational assimilation of radar reflectivity and radial velocity data: Issues with nonlinear reflectivity operator and solutions. *Mon. Wea. Rev.*, **148**, 1483–1502, https://doi.org/10.1175/MWR-D-19-0149.1.

Liu, S., and Coauthors, 2017: WSR-88D radar data and NLDN lightning data assimilation at NCEP. *21st Conf. on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface*, Seattle, WA, Amer. Meteor. Soc., 9.5, https://ams.confex.com/ams/97Annual/webprogram/Paper312466.html.

Lojou, J.-Y., and K. L. Cummins, 2005: On the representation of two- and three-dimensional total lightning information. *Conf. on Meteorological Applications of Lightning Data*, San Diego, CA, Amer. Meteor. Soc., 2.4, https://ams.confex.com/ams/Annual2005/techprogram/paper_86442.htm.

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.

Massart, D. L., J. Smeyers-Verbeke, X. Capron, and K. Schlesier, 2005: Visual presentation of data by means of box plots. *LC GC Eur.*, **18**, 215–218.

McNally, A. P., P. D. Watts, J. A. Smith, R. Engelen, G. A. Kelly, J. N. Thépaut, and M. Matricardi, 2006: The assimilation of AIRS radiance data at ECMWF. *Quart. J. Roy. Meteor. Soc.*, **132**, 935–957, https://doi.org/10.1256/qj.04.171.

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

Minnis, P., and Coauthors, 2008a: Near-real time cloud retrievals from operational and research meteorological satellites. *Proc. SPIE*, **7107**, 710703, https://doi.org/10.1117/12.800344.

——, and Coauthors, 2008b: Cloud detection in non-polar regions for CERES using TRMM VIRS and *Terra* and *Aqua* MODIS data. *IEEE Trans. Geosci. Remote Sens.*, **46**, 3857–3884, https://doi.org/10.1109/TGRS.2008.2001351.

——, and Coauthors, 2011: CERES Edition-2 cloud property retrievals using TRMM VIRS and *Terra* and *Aqua* MODIS data. Part I: Algorithms. *IEEE Trans. Geosci. Remote Sens.*, **49**, 4374–4400, https://doi.org/10.1109/TGRS.2011.2144601.

Parrish, D. F., and J. C. Derber, 1992: The National Meteorological Center's spectral statistical-interpolation analysis system. *Mon. Wea. Rev.*, **120**, 1747–1763, https://doi.org/10.1175/1520-0493(1992)120<1747:TNMCSS>2.0.CO;2.

Rison, W., R. J. Thomas, P. R. Krehbiel, T. Hamlin, and J. A. Harlin, 1999: GPS-based three-dimensional lightning mapping system: Initial observations. *Geophys. Res. Lett.*, **26**, 3573–3576, https://doi.org/10.1029/1999GL010856.

Roebber, P. J., 2009: Visualizing multiple measures of forecast quality. *Wea. Forecasting*, **24**, 601–608, https://doi.org/10.1175/2008WAF2222159.1.

Rosenfeld, D., D. B. Wolff, and D. Atlas, 1993: General probability-matched relations between radar reflectivity and rain rate. *J. Appl. Meteor.*, **32**, 50–72, https://doi.org/10.1175/1520-0450(1993)032<0050:GPMRBR>2.0.CO;2.

Rudlosky, S. D., and K. S. Virts, 2021: Dual Geostationary Lightning Mapper observations. *Mon. Wea. Rev.*, **149**, 979–998, https://doi.org/10.1175/MWR-D-20-0242.1.

——, S. J. Goodman, K. S. Virts, and E. C. Bruning, 2019: Initial Geostationary Lightning Mapper observations. *Geophys. Res. Lett.*, **46**, 1097–1104, https://doi.org/10.1029/2018GL081052.

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

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, 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.

Stefanescu, R., I. M. Navon, H. E. Fuelberg, and M. Marchand, 2013: 1D+4D-VAR data assimilation of lightning with WRFDA system using nonlinear observation operators. arXiv, 1306.1884, https://arxiv.org/abs/1306.1884.

Stensrud, D. J., and J. M. Fritsch, 1994: Mesoscale convective systems in weakly forced large-scale environments. Part II: Generation of a mesoscale initial condition. *Mon. Wea. Rev.*, **122**, 2068–2083, https://doi.org/10.1175/1520-0493(1994)122<2068:MCSIWF>2.0.CO;2.

Tong, M., 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.

——, 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.

Vukicevic, T., T. G. Ald, M. Z. Anski, D. Z. Anski, T. V. Haar, and A. S. Jones, 2004: Mesoscale cloud state estimation from visible and infrared satellite radiances. *Mon. Wea. Rev.*, **132**, 3066–3077, https://doi.org/10.1175/MWR2837.1.

——, M. Sengupta, A. Jones, and T. V. Haar, 2006: Cloud-resolving satellite data assimilation: Information content of IR window observations and uncertainties in estimation. *J. Atmos. Sci.*, **63**, 901–919, https://doi.org/10.1175/JAS3639.1.

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.

Whitaker, J. S., 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.

——, 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.

Wu, T.-C., H. Liu, S. J. Majumdar, C. S. Velden, and J. L. Anderson, 2014: Influence of assimilating satellite-derived atmospheric motion vector observations on numerical analyses and forecasts of tropical cyclone track and intensity. *Mon. Wea. Rev.*, **142**, 49–71, https://doi.org/10.1175/MWR-D-13-00023.1.

——, C. S. Velden, S. J. Majumdar, H. Liu, and J. L. Anderson, 2015: Understanding the influence of assimilating subsets of enhanced atmospheric motion vectors on numerical analyses and forecasts of tropical cyclone track and intensity with an ensemble Kalman filter. *Mon. Wea. Rev.*, **143**, 2506–2531, https://doi.org/10.1175/MWR-D-14-00220.1.

Xiao, X., J. Sun, X. Qie, Z. Ying, L. Ji, M. Chen, and L. Zhang, 2021: Lightning data assimilation scheme in a 4DVAR system and its impact on very short-term convective forecasting. *Mon. Wea. Rev.*, **149**, 353–373, https://doi.org/10.1175/MWR-D-19-0396.1.

Xue, M., D.-H. Wang, J.-D. Gao, K. Brewster, and K. K. Droegemeier, 2003: The Advanced Regional Prediction System (ARPS), storm-scale numerical weather prediction and data assimilation. *Meteor. Atmos. Phys.*, **82**, 139–170, https://doi.org/10.1007/s00703-001-0595-6.

——, 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.

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

Zhang, S., and Z. Pu, 2018: Impact of enhanced atmospheric motion vectors on HWRF hurricane analyses and forecasts with different data assimilation configurations. *Mon. Wea. Rev.*, **146**, 1549–1569, https://doi.org/10.1175/MWR-D-17-0136.1.

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

Zhu, Y. Q., and Coauthors, 2016: All-sky microwave radiance assimilation in NCEP's GSI analysis system. *Mon. Wea. Rev.*, **144**, 4709–4735, https://doi.org/10.1175/MWR-D-15-0445.1.