Assimilation of Pseudo-GLM
Data Using the Ensemble Kalman Filter

BLAKE J. ALLEN

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

EDWARD R. MANSELL

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

DAVID C. DOWELL

NOAA/OAR/Earth System Research Laboratory, Boulder, Colorado

WIEBKE DEIERLING

National Center for Atmospheric Research, Boulder, Colorado

(Manuscript received 30 March 2016, in final form 13 June 2016)

ABSTRACT

Total lightning observations that will be available from the GOES-R Geostationary Lightning Mapper (GLM) have the potential to be useful in the initialization of convection-resolving numerical weather models, particularly in areas where other types of convective-scale observations are sparse or nonexistent. This study used the ensemble Kalman filter (EnKF) to assimilate real-data pseudo-GLM flash extent density (FED) observations at convection-resolving scale for a nonsevere multicell storm case (6 June 2000) and a tornadic supercell case (8 May 2003).

For each case, pseudo-GLM FED observations were generated from ground-based lightning mapping array data with a spacing approximately equal to the nadir pixel width of the GLM, and tests were done to examine different FED observation operators and the utility of temporally averaging observations to smooth rapid variations in flash rates.

The best results were obtained when assimilating 1-min temporal resolution data using any of three observation operators that utilized graupel mass or graupel volume. Each of these three observation operators performed well for both the weak, disorganized convection of the multicell case and the much more intense convection of the supercell case.

An observation operator using the noninductive charging rate performed poorly compared to the graupel mass and graupel volume operators, a result that appears likely to be due to the inability of the noninductive charging rate to account for advection of space charge after charge separation occurs.

1. Introduction

Lightning location and frequency observations can pinpoint deep convective storms, and the availability of lightning data in areas with poor radar coverage has the potential to supplement current radar data assimilation techniques in numerical weather prediction models. The upcoming launch of the Geostationary Lightning Mapper (GLM; Goodman et al. 2013) on board the Geostationary Operational Environmental Satellite R-series (GOES-R) will greatly extend the availability of total lightning data [intracloud and cloud-to-ground (IC + CG)] by providing continuous coverage over much of the Western Hemisphere. The GLM will provide information on convective activity in mountainous terrain, where radar coverage is poor, and over much of the Atlantic and eastern Pacific Oceans, where both radar

coverage and the detection efficiency of ground-based lightning detection networks are poor.

To test the feasibility of using GLM observations to help initialize and modulate storms in convection-resolving ensemble model frameworks, this study assimilated Lightning Mapping Array (LMA)-derived pseudo-GLM flash extent density (FED) observations using the ensemble Kalman filter (EnKF) to create analyses in a storm-scale model. In addition to demonstrating the value of assimilating such data, another goal of this study was to further develop observation operators to use with GLM data. To these ends, experiments were performed with two cases: a nonsevere multicell storm that occurred on 6 June 2000 and shifted through various morphologies during its lifetime, and a tornadic supercell that occurred on 8 May 2003. The analyses produced by assimilating the lightning data were compared both to observed characteristics of the storms and to analyses of the storms produced using the more established technique of radar radial velocity (V_r) data assimilation.

Previous studies have shown that total lightning activity is related to various measures of thunderstorm intensity (e.g., graupel and ice content, rainfall rates, and updraft characteristics), and that total lightning activity exhibits much stronger correlations and correspondences with these thunderstorm characteristics than CG lightning activity alone. This has been shown in individual storms (MacGorman et al. 1989; Wiens et al. 2005; Goodman et al. 1988), in data aggregated over multiple storms in a region (Deierling and Petersen 2008), and in globally aggregated data (Petersen et al. 2005). It is also supported by modeling studies of individual storms (Kuhlman et al. 2006), aggregated storms (Cohen 2008), and hurricane eyewalls (Fierro et al. 2015). A large body of evidence suggests these correlations arise from the primary role rebounding collisions between graupel and cloud ice play in thunderstorm charge separation [the noninductive graupel–ice mechanism; Reynolds et al. (1957); Takahashi (1978); Jayaratne et al. (1983)].

Several earlier studies have demonstrated benefits from using nudging schemes to assimilate lightning data into deterministic models, both in models with parameterized convection and in convection resolving models. Alexander et al. (1999) and Chang et al. (2001) used CG lightning data converted to rainfall rates to adjust the latent heating rate in a convective parameterization scheme (CPS), Papadopoulos et al. (2005) used CG lightning data to adjust humidity profiles and thereby indirectly trigger a CPS, and Mansell et al. (2007) used CG and total lightning data in the form of LMA source point densities to directly activate a CPS. In the latter

study, the addition of total lightning data from LMA measurements was found to improve the results over using CG-only data. Pessi and Businger (2009) again showed improvements to a forecast when using CG lightning data, converted to rainfall rates, to adjust the latent heating rate in a CPS.

In deterministic convection-resolving models, total lightning data in the form of Earth Networks Total Lightning Network (ENTLN) flash origin densities, converted to 9-km pixel-width data to mimic flash densities that will be available from the GLM, have been used to initialize observed convection using two different methods. The first method utilizes water vapor nudging in the mixed-phase region between the 0° and −20°C levels (Fierro and Reisner 2011; Fierro et al. 2012), and the second method acts by increasing the low-level temperature in the model to the convective temperature (Marchand and Fuelberg 2014).

More limited research has been done using ensemble techniques. Using primarily CG lightning data converted to rainfall rates and a synoptic-scale model, Hakim et al. (2008) was the first to assimilate lightning data using in an ensemble framework. As a direct precursor to this study, Mansell (2014) performed an observing system simulation experiment (OSSE) in which simulated total lightning data in the form of FEDs were assimilated, using the EnKF, into the same storm-scale model used in this research. In that study, the data assimilation technique showed the ability to modulate the strength of convection and suppress spurious convection. This study continues the research begun in Mansell (2014) by working with real-data cases.

2. Data assimilation framework

a. Numerical model

The model used for this work was the National Severe Storms Laboratory (NSSL) Collaborative Model for Multiscale Atmospheric Simulation (NCOMMAS; Wicker and Wilhelmson 1995; Coniglio et al. 2006), which is based on the equation set of Klemp and Wilhelmson (1978). The NSSL two-moment, four-ice category bulk microphysics scheme of Mansell et al. (2010), based on the scheme of Ziegler (1985), was used. The microphysics scheme predicts mass mixing ratio and number concentration for six hydrometeor types: cloud droplets, rain, cloud ice, graupel, snow, and hail. Additionally, the scheme predicts graupel and hail particle densities. The electrification scheme of Mansell et al. (2005) was activated during the development of lightning observation operators (section 5); the electrification scheme includes graupel-ice noninductive charge separation based on Saunders and Peck (1998) and

Brooks et al. (1997), as implemented in Mansell et al. (2010), and the scheme also includes the branched lightning simulation of Mansell et al. (2002). For the lightning data assimilation experiments (section 6), electrification was turned off except when using the noninductive charging rate observation operator. In experiments using that observation operator the non-inductive charging parameterization was activated but lightning was not simulated.

Horizontal grid spacing was set to 1 km for both cases, while vertical grid spacing varied from 200 m at low levels to a maximum of 700 m in the multicell case and was set to a constant value of 500 m in the supercell case (the cases are described in detail in section 3). The domain in each case had 50 vertical levels. In the multicell case, the domain was 112 km × 112 km in the horizontal and 18 km in the vertical, while for the supercell case the domain was 100 km × 100 km × 25 km. In the supercell case, the domain moved along with the convection ($u = 14\,\mathrm{m\,s}^{-1}$, $v = 8\,\mathrm{m\,s}^{-1}$) to keep the simulated storm near the center of the domain, while in the multicell case the domain was stationary.

b. Ensemble Kalman filter setup

The EnKF implementation used in this study is based on the ensemble square root filter (EnSRF) of Whitaker and Hamill (2002). The set of equations used to update the model state are as shown in Dowell and Wicker (2009).

Localization of observations was achieved using a factor calculated from the monotonically decreasing function of Gaspari and Cohn (1999). When assimilating lightning data, the factor decreased from unity at the location of the observation to zero 15 km away from the observation in the horizontal, while in the vertical the factor decreased from unity at the assumed altitude of the observation (6.5 km) to zero 36 km away. The localization is thus effectively two dimensional, allowing the 2D pseudo-GLM FED observations to exert influence throughout the vertical extent of the domain.

Development and testing of observation operators for use with pseudo-GLM FED data were objectives of this research, and the observation operators used are discussed in section 5.

c. Initiation of convection and maintenance of ensemble spread

When assimilating lightning data in this study, all ensemble members were first initialized with the same horizontally homogeneous sounding and, therefore, the kinematic and thermodynamic fields were identical across the ensemble at the beginning of each experiment. Additionally, convection was absent at the beginning of the simulations and, therefore, the values of

all liquid water and ice water fields are initially zero in all ensemble members. To introduce convection into the model and to develop ensemble spread, a combination of two methods was used. The first method adds randomly placed thermal bubbles to the ensemble members in regions where radar reflectivity is absent in the ensemble mean but observations show nonzero FEDs. This scheme is similar to the one used with radar data in Tanamachi et al. (2013), and the use of the scheme with lightning data is facilitated by assuming an observed reflectivity of 40 dBZ wherever the observed FED is nonzero. The bubble addition ended after the first 50 min of the simulations since continued forcing after convection was established tended to degrade the results. The second method introduces smooth, random noise to the u, v, θ, and T_d variables of each ensemble member using the method of Caya et al. (2005) as implemented in Dowell and Wicker (2009). For each of these methods, the model adjustments were made at 6-min intervals.

3. Observational data and storm cases

Each storm case in this study was observed both by radar and lightning mapping array. Choosing both weak multicell convection and strong supercellular convection allowed the performance of the data assimilation technique to be tested on opposite ends of the storm severity spectrum. An overview of the two cases follows.

a. Case 1: 6 June 2000 multicell storm

This storm occurred in northeast Colorado during the Severe Thunderstorm Electrification and Precipitation Study (STEPS; Lang et al. 2004). As shown in Fig. 1a, the environment had little instability and limited wind shear, leading to weak, disorganized convection. Although the storm was small and weak, it persisted long enough to evolve through a variety of formations. Initially, somewhat before 2200 UTC, the storm developed as a single cell, with a new cell later developing on its northern flank around 0000 UTC. The initial cell then weakened somewhat as the new cell strengthened and the overall area of precipitation became elongated along a north-northeast–south-southwest (NNE–SSW) axis by 0100 UTC. No severe weather was reported with the storm. Radar data were collected on the storm by the CSU–CHILL radar, the NCAR S-Pol radar, and the KGLD (Goodland, Kansas) WSR-88D, and lightning data were collected by the STEPS LMA. The maximum FED (after conversion from LMA data to pseudo-GLM data, see section 4) in the storm was 21 flashes min^{-1} pixel^{-1}, where a pixel has a size of approximately 8 km × 8 km.
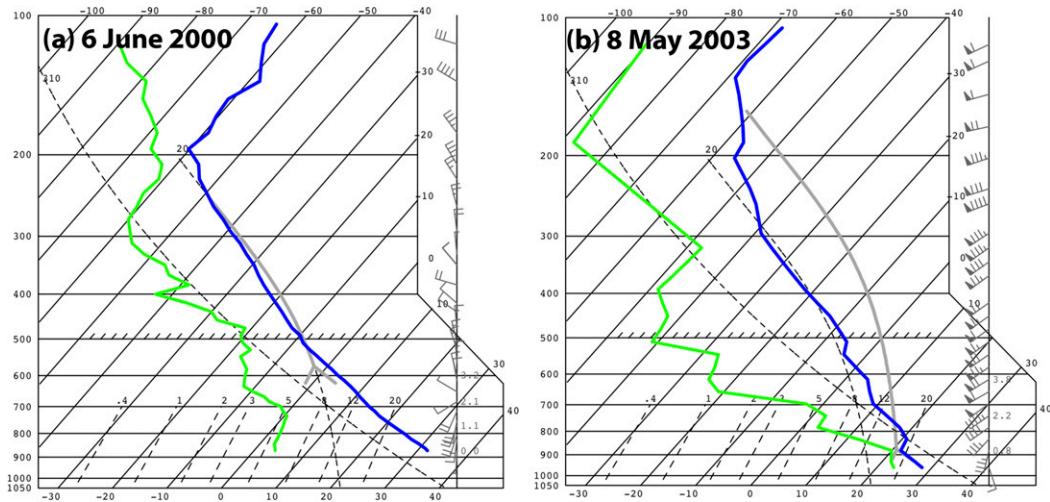
FIG. 1. Sounding used for (a) the multicell case and (b) the supercell case. For the multicell case, the thermo-dynamic profile has 315 J kg^{-1} MLCAPE and 100 J kg^{-1} MLCIN. The sounding consists of data from a mobile sounding taken during STEPS for levels below 15.3 km, with higher altitudes using data from the 0000 UTC 7 Jun 2000 North Platte, NE, sounding. For the supercell case, the thermodynamic profile has 4093 J kg^{-1} MLCAPE and 15 J kg^{-1} MLCIN, and the sounding comes entirely from data collected during the 0000 UTC 9 May 2003 Norman, OK, radiosonde launch. In both cases the blue line corresponds to temperature, the green line corresponds to dewpoint, and the gray line corresponds to the moist adiabat followed by a parcel averaged over the lowest 90 hPa of the sounding.

b. Case 2: 8 May 2003 tornadic supercell

This storm occurred in central Oklahoma in an environment strongly supportive of severe convection and tornadoes (Fig. 1b). The storm's initial development began around 2040 UTC and over the next hour the storm strengthened and began to take on supercell characteristics. By 2150 UTC the storm contained a strong mesocyclone and it first produced a short-lived and weak tornado around 2200 UTC. After a second weak tornado, a long-tracked F4 tornado formed at 2210 UTC and passed through parts of Moore and southeast Oklahoma City, Oklahoma. The mesocyclone eventually occluded and the tornado dissipated around 2250 UTC. The storm later developed additional mesocyclones but did not produce any further tornadoes before dissipating (Burgess 2004; Romine et al. 2008). The storm was observed by the KTLX (Oklahoma City) WSR-88D, the polarimetric KOUN (Norman, Oklahoma) radar, and a nearby Terminal Doppler Weather Radar (TDWR), and the storm has been the subject of many storm-scale data assimilation experiments in which either V_r data or both V_r and radar reflectivity data were assimilated (see, e.g., Dowell et al. 2004a, 2011; Dowell and Wicker 2009; Yussouf et al. 2013). Electrically, the storm was observed by the Oklahoma LMA, and conversion of the data to pseudo-GLM observations yielded FEDs up to 128 flashes min^{-1} pixel^{-1}.

4. Lightning observations and their processing

All lightning observations used in this study originated as LMA data. LMA data consist of "source points," which are geolocated based on VHF emissions that occur during the lightning breakdown and leader propagation processes. The GLM, on the other hand, will detect near-infrared pulses produced by lightning on a grid with pixel widths varying from approximately 8 km directly beneath the satellite to around 14 km at the edge of the instrument's domain. To make use of LMA observations as a proxy for the data that will be available from the GLM, the LMA observations must be converted to pseudo-GLM observations. The process by which this was done in this study is as follows. First, each minute of LMA data were processed with a flash sorting algorithm that groups source points into individual lightning flashes based on spatial and temporal thresholds as in MacGorman et al. (2008). Next, the sorted points were mapped onto a grid with an approximately 8-km pixel width, with the FED in each pixel being increased by one for each flash that passes over the pixel during the minute of data (Fig. 2).

All observations were assigned a height of 6.5 km, which is located in the mixed-phase region for typical warm-season midlatitude convection and corresponds approximately to the mean maximum in the vertical distribution of lightning sources as indicated by mapping arrays (MacGorman and Rust 1998). Although significant storm-to-storm variations can occur in the
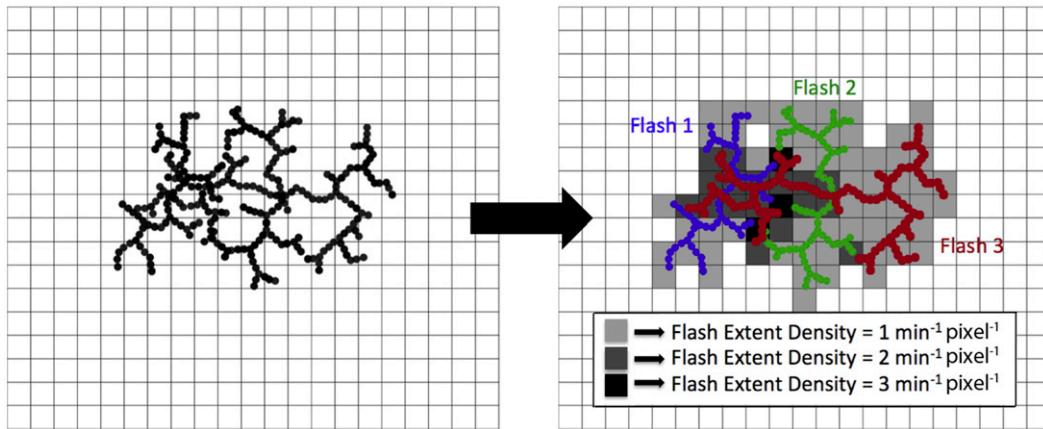
FIG. 2. Illustration of raw source points (left) separated into individual flashes and (right) mapped to the $8 \times 8 \, km^2$ pseudo-GLM grid as FED data.

height of the lightning source maximum (Bruning et al. 2007; Tessendorf et al. 2007), the impact of any vertical misplacement of the observations is limited by the large vertical localization radius used in this study. The final product is pseudo-GLM FED data with 1-min temporal resolution. Figure 3 shows an example of this for 1 min of data in the supercell case. Time series of the maximum FED for each case are shown in Fig. 4.

Experiments were performed in which the data were assimilated into the model every 60 s both without and with temporal smoothing performed using a running mean with a 3-min window. To see if the computational

cost of the assimilation process could be lowered, tests were also performed with longer assimilation intervals (3 and 6 min), but the resulting analyses were significantly degraded (not shown).

The true observation error variance in the flash extent densities generated in this manner is unknown, and additional error arises from the observation operator. Setting the assumed observation error variance too low results in overfitting and can cause excessive adjustments to model state variables when the model-derived flash rates differ significantly from the observations during assimilation, which in turn can lead to numerical
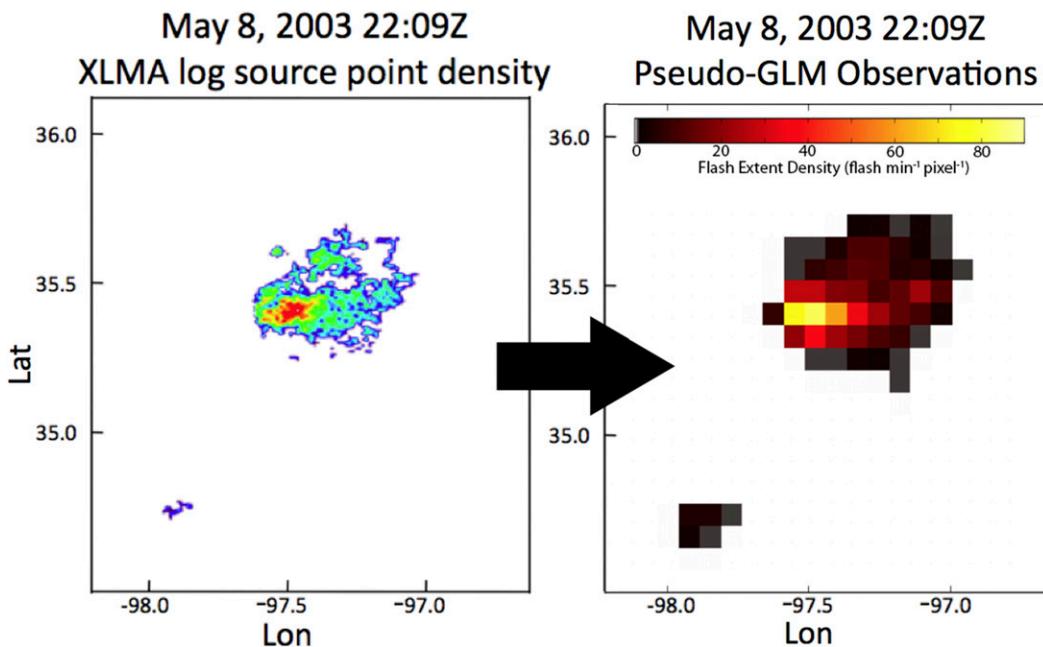

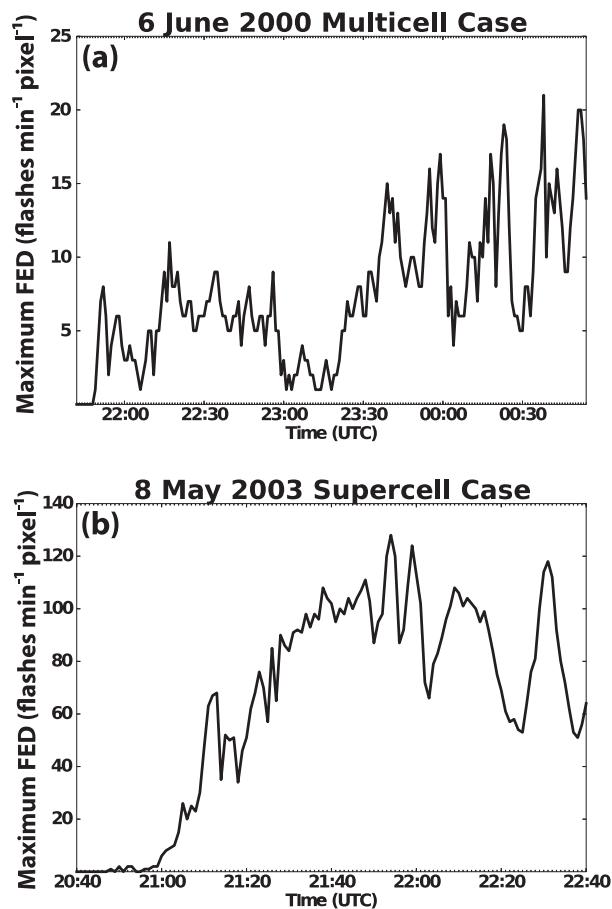FIG. 3. Example output of the Lightning Mapping Array to pseudo-GLM conversion process for actual storm data.

6 June 2000 Multicell Case

8 May 2003 Supercell Case

Fig. 4. Time series of the maximum flash extent density in the (a) multicell case and (b) supercell case.

Total Flash Rate vs. Graupel Mass

(a) $r^2 = 0.646$

Flash Rate = $2.33*10^{-8}(GM)$

Total Flash Rate vs. Graupel Volume

(b) $r^2 = 0.591$

Flash Rate = 0.026(GV)

Total Flash Rate vs. Noninductive Charging Rate

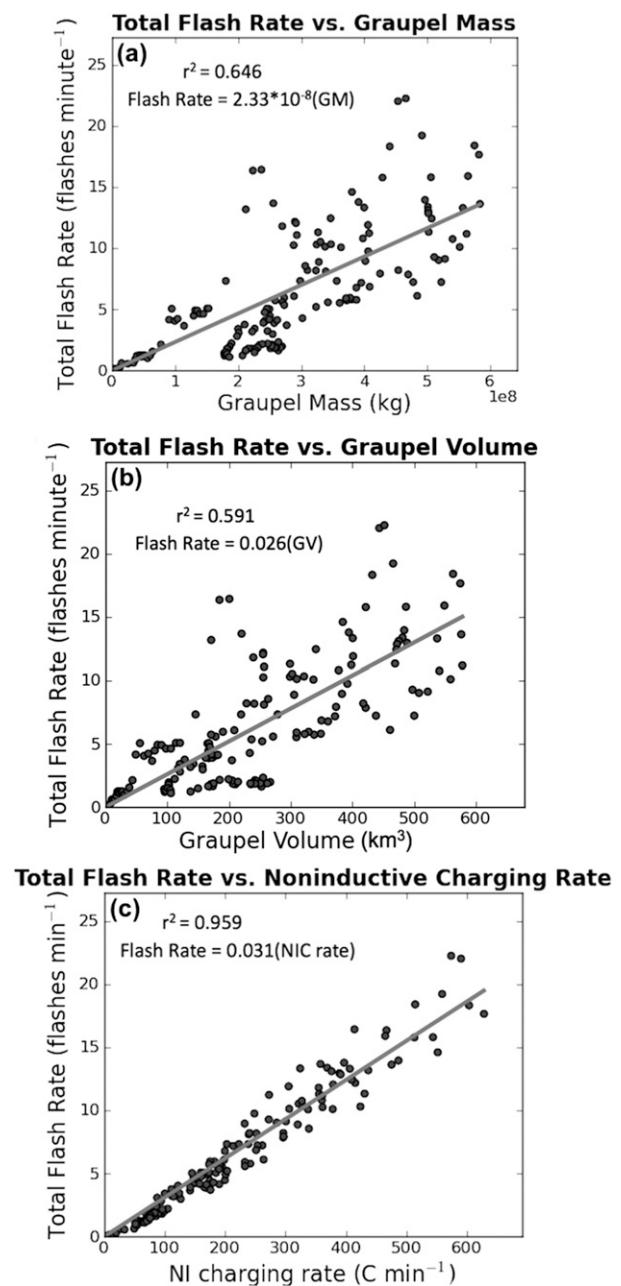(c) $r^2 = 0.959$

Flash Rate = 0.031(NIC rate)

Fig. 5. Scatterplots of flash rate and (a) graupel volume, (b) graupel mass, and (c) noninductive charging rate data from the multicell case, with best-fit lines overlaid.

instability in outlying members of the ensemble due to updraft velocities exceeding the CFL limit. On the other hand, setting the observation error variance too high results in underfitting and limits the ability of the observations to effectively influence the simulations.

Optimal results were achieved with the observation error variance set to 10%–15% of the maximum FED in each storm. This corresponded to an observation error variance setting of 2.1 flashes min^{-1} pixel^{-1} for the multicell case (maximum FED of 21 flashes min^{-1} pixel^{-1}) and a setting of 19.2 flashes min^{-1} pixel^{-1} for the supercell case (maximum FED of 128 flashes min^{-1} pixel^{-1}). The need for different observation error variance settings between cases was not entirely unexpected as the error in pseudo-GLM FEDs generated using the above method likely increases as flash rates increase due to occasional flash miscounting during the automated process of separating LMA sources into individual flashes (K. Calhoun 2016, personal communication). Given the differences in how FEDs were derived for this study compared to how the same quantity will be measured by the GLM,

observation error variance settings will likely need to be adjusted for the assimilation of actual GLM data.

5. Observation operator development

As noted earlier, a major objective of this research was to develop and test new FED observation operators for use with GLM data. A model-based linear best-fit observation operator between flash rate and graupel
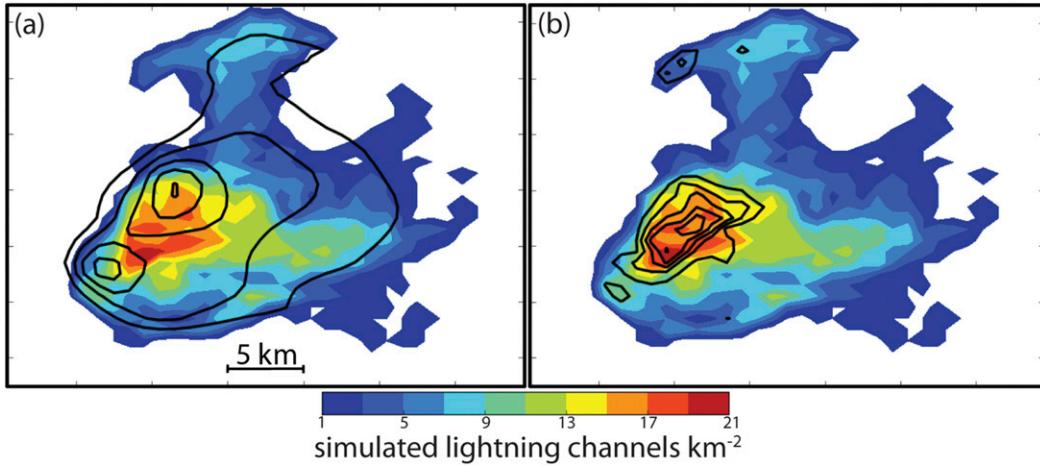
FIG. 6. (a) Vertically integrated simulated lightning channel density (filled contours) and graupel mass mixing ratio [black contours, intervals $0.2 \times$ (maximum graupel mass mixing ratio)]. (b) Vertically integrated simulated lightning channel density (filled contours) and noninductive charging rate [black contours, intervals $0.2 \times$ (maximum NI charging rate)]. Plots from V_r assimilation tests using lightning simulation for the supercell case at 2200 UTC.

volume found by Cohen (2008) was tested in Mansell (2014) along with a second linear best-fit operator found using the flash rates and graupel volume of the storm simulated in that study. Additionally, Mansell (2014) tested an observation operator that used lightning generated explicitly by each ensemble member. The simulated lightning scheme (Mansell et al. 2002) is very computationally expensive and the observation operator using it had comparable results to the linear best-fit operator created from the storm simulated in that study. Therefore, explicitly simulated FED was not tested as an observation operator here, and the current study instead focused on investigating additional operationally efficient linear best-fit observation operators.

Explicitly simulated lightning was used, however, in conjunction with radar radial velocity assimilation, to create first guesses for new linear best-fit observation operators based on relationships between model microphysical quantities and model flash rates. To accomplish this, V_r data from each case (CSU–CHILL data for the multicell case, KOUN data for the supercell case) were assimilated into NCOMMAS using the EnSRF. A similar method was used by Calhoun et al. (2014) to study supercell lightning. Electrification and lightning were simulated during these V_r data assimilation tests, and then linear best-fit relationships between flash rates and graupel volume, graupel mass, and noninductive charging rate were derived. All relationships
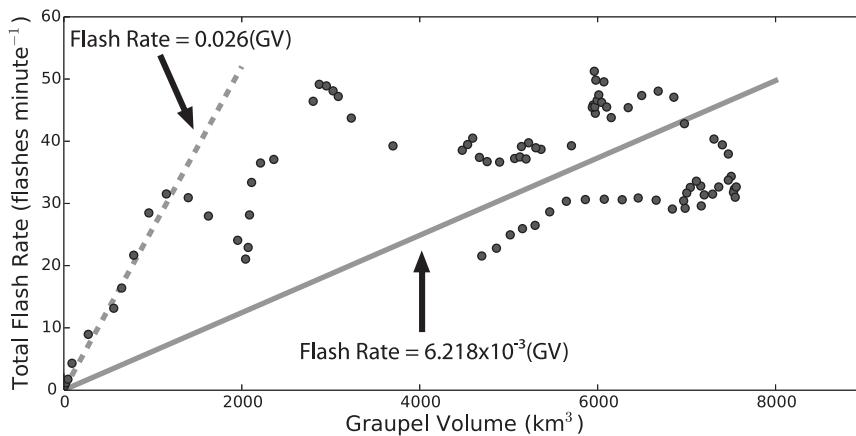

FIG. 7. Scatterplot of flash rate and graupel volume data from the supercell case, with the best-fit lines from this case (solid line) and the the multicell case (dashed line) overlaid.

TABLE 1. Data assimilation experiments performed. Observation operators use either graupel volume (GV), graupel mass (GM), or noninductive charging rate (NIC). Experiment GVa3 differs from experiment GVa1 in that it used observations smoothed with a 3-min running average.

Lightning data assimilation expt	
Observation operator	Expt name
FED = 0.017 (GV)	GVa1
	GVa3
FED = 1.5 × 0.017 (GV)	GVb1
FED = 2.088 × 10⁻⁸ (GM)	GM
FED = 0.039 (NIC)	NIC

were calculated with flash rates and microphysical variables integrated over the entire domain. Since the unconstrained best-fit lines tended to have physically unrealistic y intercepts slightly above zero (indicating the presence of lightning without graupel), the y intercepts were forced to zero for the linear fits used as first guesses.

When the observation operators were used during the lightning data assimilation experiments, each microphysical quantity was summed over a volume spanning the vertical extent of the domain and covering a $16 \times 16 \, \text{km}^2$ area in the horizontal dimensions centered on the observation being assimilated. Graupel mass and noninductive charging rate were calculated by directly summing the values of those variables at each grid point in the $16 \times 16 \, \text{km}^2$ area, while graupel volume was calculated by summing the volume of each grid box within the $16 \times 16 \, \text{km}^2$ area that contained a graupel mass mixing ratio exceeding $0.5 \, \text{g kg}^{-1}$. As in Mansell (2014), the summing of microphysical quantities over a width larger than the 8-km GLM pixel width allows for influence from lightning occurring somewhat outside or on the edge of a GLM pixel. While Mansell (2014) found optimal results using a $12 \times 12 \, \text{km}^2$ area with synthetic observations, the larger area was found to be superior with the LMA-derived FEDs assimilated in this study.

In the multicell case, the correlations between flash rates and graupel volume or graupel mass were similar, while stronger correlations were found between flash rates and noninductive charging rate (Fig. 5). Compared to graupel volume or graupel mass fields alone, noninductive charging rate depends on additional factors such as updraft/downdraft characteristics [via the relative fall velocities of colliding particles as in Mansell et al. (2005)]. In that way, it bears some similarity to the radar-derived flux product of Deierling et al. (2008), which was found in that study to have a slightly higher correlation with flash rates than individual components of the flux product such as graupel mass. Even though flash rates may be better correlated with measures of noninductive charging rate when whole-storm quantities

are considered, the same may not be true in subsections of storms on the scale of a GLM pixel. For example, Fig. 6 shows that graupel content better corresponds spatially with lightning activity in the strongly sheared storm of 8 May 2003. Unlike charging rates, graupel mass (or graupel volume) accounts for horizontal advection of charged particles after charge separation occurs, thus producing better results for FED estimation.

Relationships found between flash rates and microphysical variables when assimilating V_r data for the supercell case were more complicated than in the multicell case. Beyond a certain level of graupel mass or graupel volume, the flash rates no longer exhibited a linear relationship with the graupel content (Fig. 7 shows this for graupel volume). This significant deviation from a linear relationship was unexpected—the simulated flash rate only slightly exceeds 50 flashes min^{-1} for this case, while the maximum observed flash rate was above 250 flashes min^{-1}. It is unclear why this occurred, as strong storms simulated with a similar lightning model setup in other studies have achieved more realistic flash rates (e.g., Calhoun et al. 2014). Nevertheless, the linear relationship found when fitting data from the multicell case was also found to fit the data from the supercell case quite well at lower ($\lesssim 1000 \, \text{km}^3$) graupel volumes (Fig. 7, dashed line).

Attempts were made to assimilate the lightning data from the supercell case using a piecewise linear observation operator fitted to different portions of the graupel volume scatterplot, but results were poor compared to operators composed of a single linear relationship that only corresponded closely to the lower graupel content portions of the graupel mass or graupel volume data. It should be noted that the observation operator calculates a flash extent density from the model state by summing the values of a microphysical variable over what is, depending on storm height, roughly a $2500 \, \text{km}^3$ subsection of the domain, such that a given observation operator calculation will only cover a fraction of a large storm such as the 8 May 2003 supercell. Therefore, accuracy of the observation operator is mainly important at lower microphysical values than those found in an entire large storm. For these reasons, the first guesses for the novel observation operators used in the lightning data assimilation experiments were all based on the multicell case. This method also has the advantage of using the same observation operators for both the weak and strong convection, although with the assumed observation error variance differing between cases.

6. Lightning data assimilation results

Once first-guess observation operators had been created, experiments were performed in which analyses
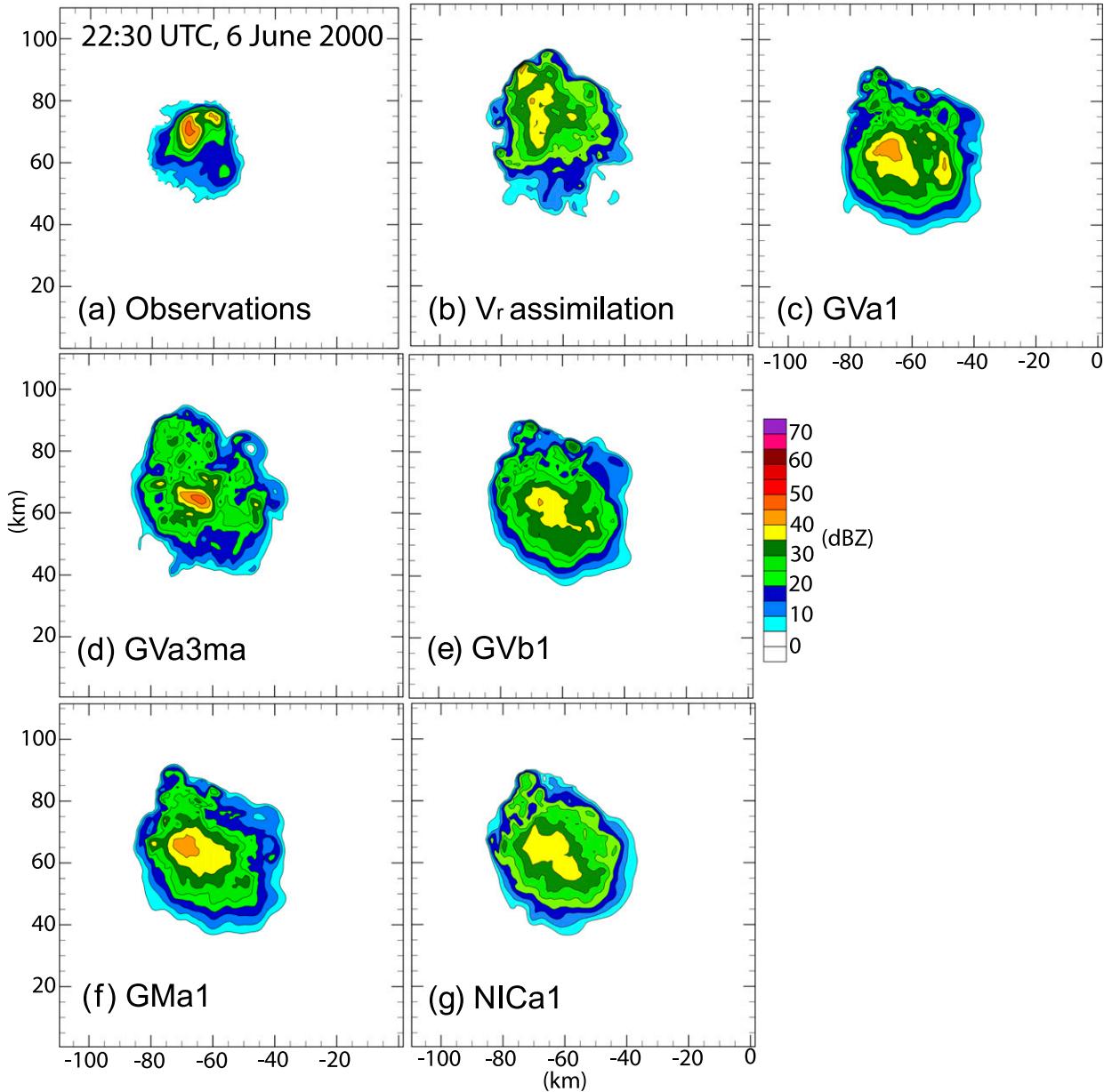
FIG. 8. Radar reflectivity near 6 km AGL at 2230 UTC 6 Jun 2000. (a) CSU–CHILL observations, (b) EnKF analysis means for CSU–CHILL V_r assimilation, and (c)–(g) the lightning data assimilation experiments.

were produced by assimilating only pseudo-GLM FED observations. For each case, experiments were performed with the three different types of observation operators noted in section 5, with the most testing done using graupel volume-based operators.

Observations of a given FED will result in the introduction of larger amounts of graupel into the model when using observation operators with shallower slopes, and the larger graupel amounts will result in higher reflectivity values. The slopes of the first-guess observation

operators were thus varied based on model over- or underestimation of reflectivity values, and the experiments presented below use the observation operators that performed best. When possible, the results of the lightning data assimilation experiments were compared to observations, and for further reference the results were compared to analyses created by assimilating V_r data alone. A list of the observation operators used with the lightning data assimilation experiments performed is shown in Table 1, along with abbreviated
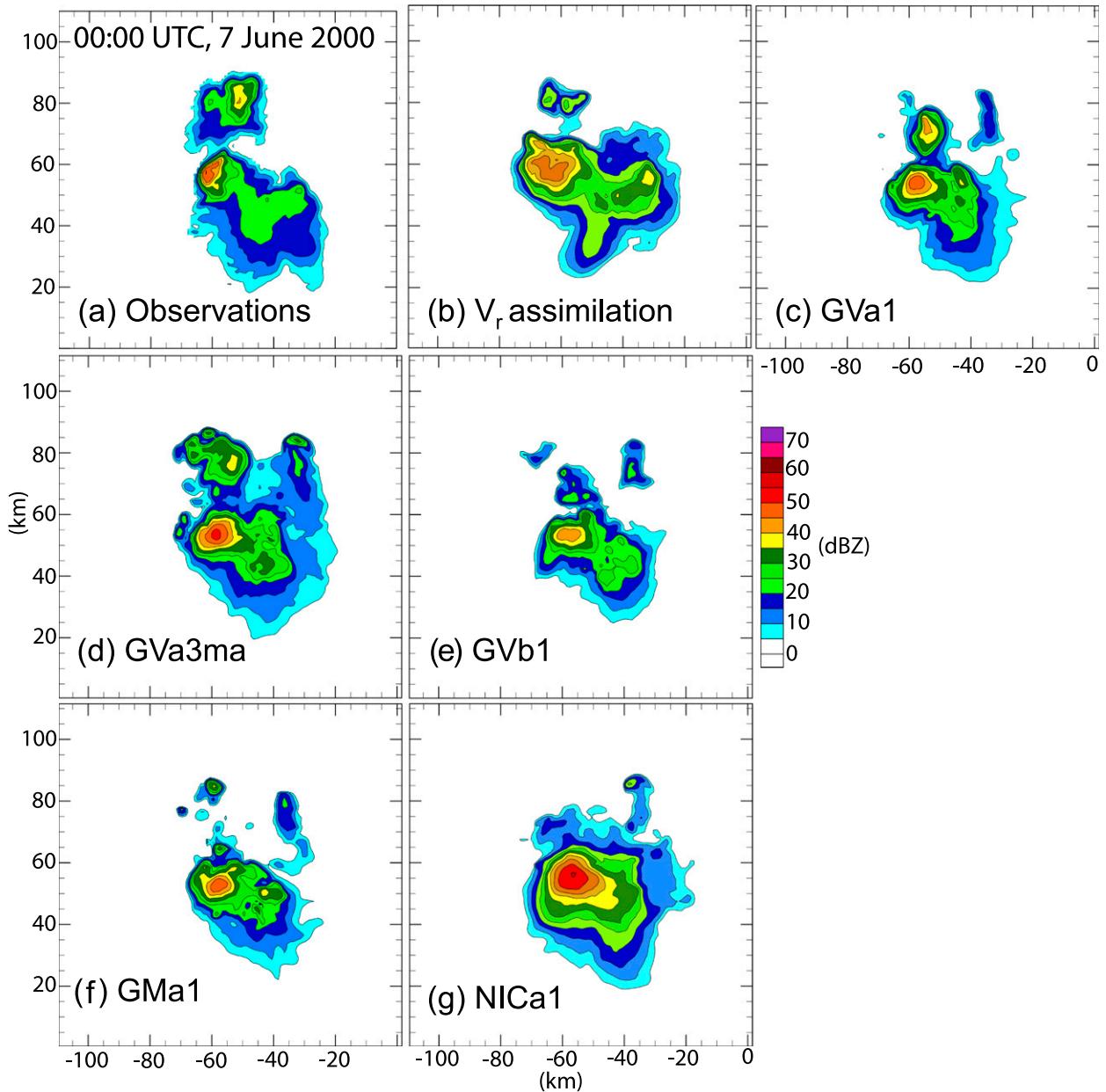
FIG. 9. As in Fig. 8, but at 0000 UTC 7 Jun 2000.

names for each experiment that will be used when discussing the results.

a. Case 1: 6 June 2000

In the multicell case, the development of the storm began shortly after 2130 UTC, and the assimilation experiments were initiated at 2142 UTC.

At 2230 UTC, the observed storm consisted of a small cell with a maximum in reflectivity slightly over 45 dBZ at 6 km above ground level (AGL), as shown in Fig. 8a. The various lightning data assimilation experiments had

by this time all produced a small storm with reflectivity of at least 35 dBZ (Figs. 8c–g). In all experiments, the areal extent of the analyzed storms was somewhat larger than that of the observed storm. Some spatial broadening of convection is expected due to the low spatial resolution of the pseudo-GLM data, but given that the convection in the V_r assimilation experiment is also wider than the observed storm, a possible factor here is the thermal bubble addition used to initially create convection, which continued for the first 50 min of the simulations and was just ending at the time of Fig. 8.
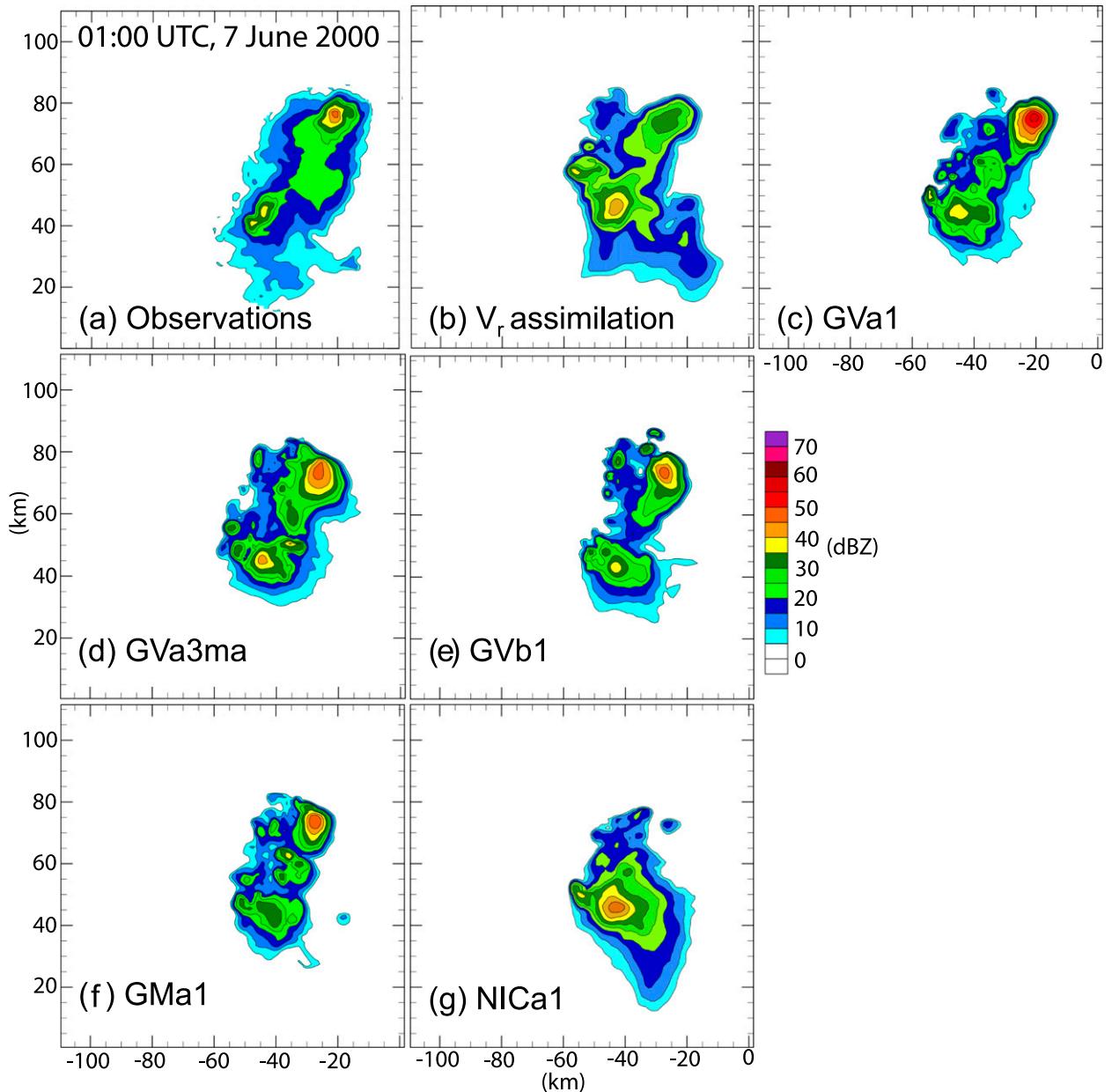
FIG. 10. As in Fig. 8, but at 0100 UTC 7 Jun 2000.

After another 90 min, at 0000 UTC 7 June, the initial cell of the observed storm had drifted south-southeasterly, and a new, weak cell had developed on its northern flank (Fig. 9a). In some of the experiments, the simulated storm tracked this evolution well, while in others, the development of the northern cell had not yet occurred. Specifically, the experiments using observation operator GVa (Figs. 9c,d) had produced the northern cell, with its maximum reflectivity at 6 km most closely matching observations in experiment GVa1. Among the experiments using other

observation operators, GVb1 and GM (Figs. 9e,f) showed the beginnings of the northern cell, while it was almost entirely absent from experiment NIC (Fig. 9g). In most experiments the extension of lower reflectivity values to the southeast of the southern cell was captured as well. Experiment NIC produced the most intense southern cell at this time, with reflectivity values well above those observed. When assimilating V_r data (Fig. 9b), the northern cell was only in the very early stages of development at this time, and the areal extent of the region with reflectivity above
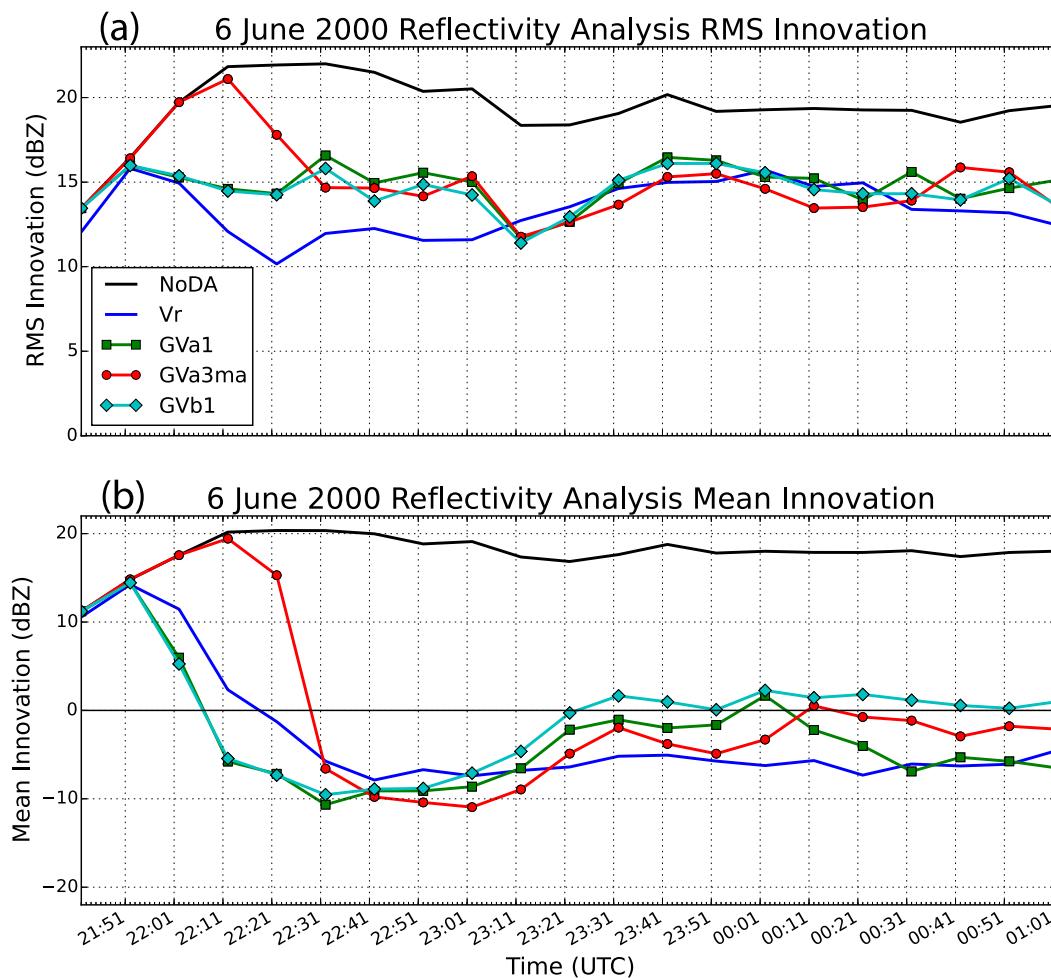
FIG. 11. (a) Radar reflectivity RMS innovation and (b) mean innovation for case-1 graupel volume observation operator experiments verified against CSU–CHILL reflectivity data. For comparison, the RMS innovation and mean innovation is also shown for the CSU–CHILL V_r assimilation experiment and for an experiment in which no data were assimilated.

35 dBZ in the southern cell had become somewhat larger than observed.

An hour later, at 0100 UTC, the southern cell had moved southeastward and weakened somewhat, while the northern cell had strengthened and moved eastward and slightly southward (Fig. 10a). With the exception of experiment NIC, in which the new cell was essentially absent, the lightning assimilation experiments tracked this evolution well (Figs. 10c–g). The V_r assimilation experiment had produced a northern cell by this time (Fig. 10b), but its maximum reflectivity remained well below that of the observed northern cell, while the southern cell was somewhat stronger than in the observed storm.

Interestingly, when comparing the results of the data assimilation experiments to the observed reflectivity near 6 km AGL, the best performing lightning assimilation experiments produced results superior to the V_r

assimilation experiment throughout the time period simulated. Given the higher resolution and larger number and depth of observations available with V_r data compared to pseudo-GLM data, this was an unexpected result. This may be due in part to the weak wind fields present in the storm and the possibility that V_r data may be less correlated with the radar reflectivity (i.e., precipitation hydrometeors) than pseudo-GLM data. Although the microphysical variables involved in calculating radar reflectivity can be updated by the EnKF when V_r observations are assimilated, the only explicit physical connection between the V_r observation operator and the microphysics is via a precipitation fall speed adjustment to the vertical wind component. The GLM FED observation operators, on the other hand, are directly tied to the microphysics, which may produce an advantage for lightning data assimilation when it
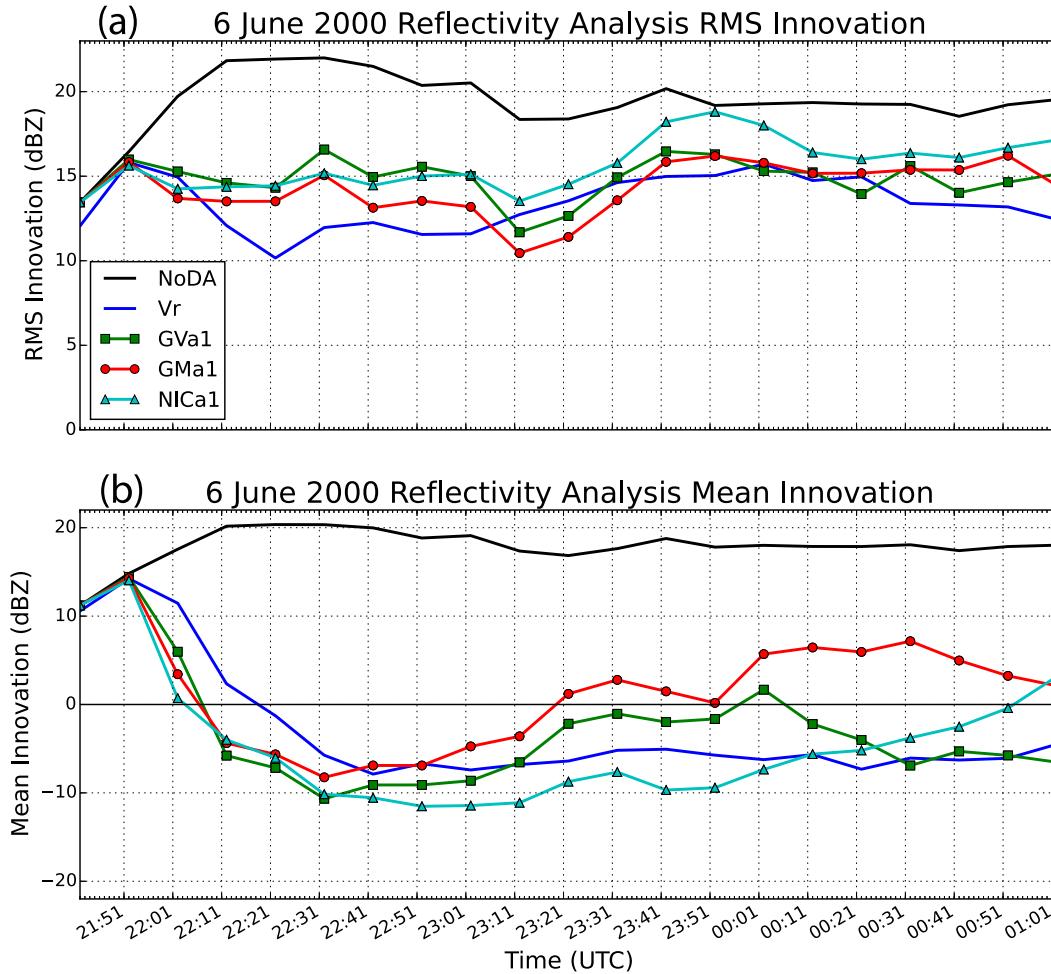
FIG. 12. As in Fig. 11, but for experiments with observation operators using different microphysical variables.

comes to reproducing the microphysical structure of the storm, especially with respect to graupel. Assimilation of reflectivity data along with V_r data has been shown to improve the model representation of convective microphysical structures at the cost of some error increase in the model representation of the velocity field (Dowell et al. 2011), but even in a test assimilating both V_r and reflectivity data for this case (not shown), the GVa lightning data assimilation experiments were faster to capture the development of the northern cell at 6 km.

To evaluate the results of the experiments throughout the domain rather than only at a specific height as in the radar reflectivity plots, time series of the analysis RMS innovation and mean innovation as verified against CSU–CHILL reflectivity observations are shown in Figs. 11 and 12. Here, the innovation is defined as the difference between the observation and ensemble mean model state mapped to observation space, $y^o - \overline{H(\mathbf{x}^a)}$, where y^o is the observed value, H is the observation operator, and $\mathbf{x}^a$ is the analysis model state. Since the

innovation magnitudes are sensitive to the areal coverage and strength of any convection present in the model domain, innovation values from an experiment in which no data assimilation or warm bubble addition was performed (and no convection developed in the model) are included as a reference (the "NoDA" experiment). In each experiment, the innovation is only calculated in areas where either the mean model reflectivity or the observed reflectivity is greater than 10 dBZ.

Figure 11a compares the RMS reflectivity innovations from experiments using the various graupel volume operators. By this measure, the V_r experiment tended to perform the best during the first 90 min of the simulation as it lowered the RMS innovation relative to the NoDA experiment almost twice as much as the lightning assimilation runs. However, during the second half of the simulation the RMS innovations of the V_r and lightning assimilation experiments are similar, and all of the lightning assimilation experiments provided significant improvement compared to the NoDA experiment throughout the time simulated.
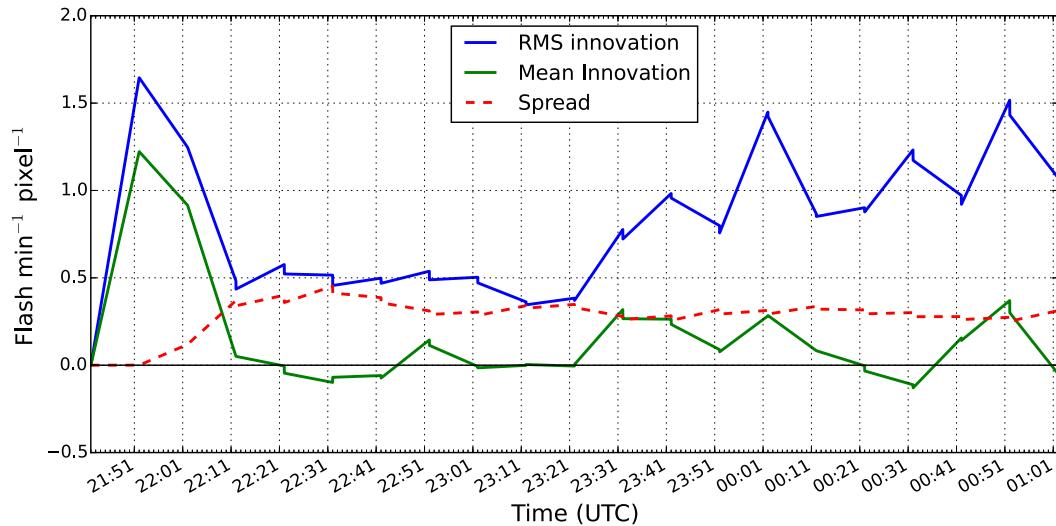
FIG. 13. Flash extent density RMS innovation, mean innovation, and ensemble spread for the case-1 GVa1 experiment.

All experiments have a substantially negative mean innovation (i.e., a high reflectivity bias, Fig. 11b) after convection initially develops, which could be due to the warm bubbles initiating convection that is stronger or has a larger areal extent than observed, as in the qualitative evaluation at 6 km AGL. Such a bias could also arise from various other factors, including biases that may exist in the model and/or in the observations. Another possible cause is that although the EnKF assumes the forecast error distribution estimated from the ensemble state at each analysis time is Gaussian, that assumption may at times be violated in a significant way in convective storms due to the positive-definite nature of hydrometeor fields. For example, if only a portion of the ensemble members contain graupel at a given grid point, the distribution of graupel values in the ensemble at that grid point used to estimate the forecast error will be non-Gaussian due to the absence of any ensemble members with graupel mixing ratios below zero.

Later in the simulations the V_r experiment maintains the bias toward higher reflectivity values, while the GVb1 experiment tends to produce the smallest mean innovations. After the initial period of warm bubble introduction the experiment using an observation operator with a shallower slope (GVa1) tended to produce convection that is biased more toward high reflectivity values than the experiment using the observation operator with a steeper slope (GVb1). That difference is likely due to the introduction of larger amounts of graupel into the simulation for a given FED rate when using the observation operator with the shallower slope.

Figure 12 compares the lightning assimilation experiments using the graupel mass (GM) and noninductive charging (NIC) observation operators to the GVa1 experiment. Experiment GM tends to outperform experiment GVa1 in both the RMS and mean innovation early in the experiments, but the RMS innovation of the two experiments is similar at later times, and the mean innovations become similar in magnitude although opposite in sign. The NIC experiment, meanwhile, has the worst RMS innovations after about 2300 UTC, and it has a large bias toward high reflectivity values throughout much of the simulation.

Figure 13 shows the RMS and mean innovation along with the ensemble spread (Dowell et al. 2004b) evaluated against the assimilated pseudo-GLM FED observations for the GVa1 experiment in regions where either the observed or model FED is nonzero. After an initial spike, the RMS innovation improves as convection first develops, but degrades somewhat later in the run. Throughout the period simulated, the RMS innovation stays within the assumed observation error variance of 2.1 flashes min^{-1} pixel^{-1}. Meanwhile, the mean innovation tends to stay slightly above zero for most of the rest of the run, suggesting a small underestimate of the FED by the model. The spread increases as convection develops and then undergoes a slow, small decrease later in the run.

b. Case 2: 8 May 2003

In the supercell case, the assimilation experiments were initiated at 2040 UTC, about the time of the first observed radar echo. Figure 14 shows that all of the assimilation experiments are able to capture the basic location and orientation of the observed reflectivity pattern, with some experiments showing more obvious
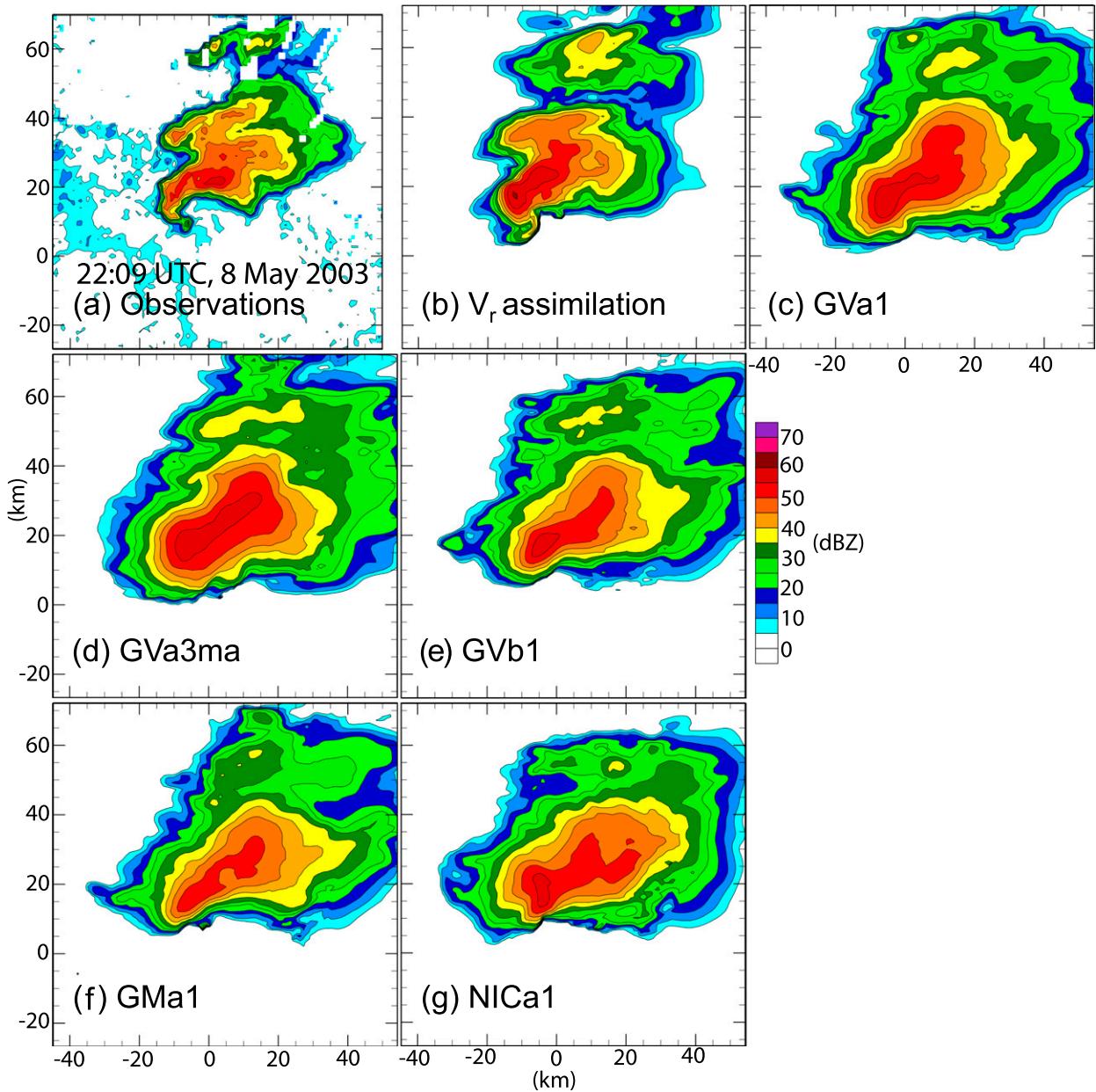
FIG. 14. (a) KOUN-observed reflectivity at 500 m at 2209 UTC, (b) EnKF analysis mean reflectivity at 250 m at 2209 UTC when assimilating KOUN V_r, and (c)–(g) EnKF analysis mean reflectivity at 250 m at 2209 UTC for each of the lightning data assimilation experiments.

supercell characteristics. Subjectively, experiments GVa1 and GVb1 appeared to produce the best results here, with experiment GM also producing good results. As with the multicell case, lower reflectivity values are present in the experiment using an observation operator with a steeper slope (GVb1) compared to the experiment using an observation operator with a shallower slope (GVa1). Averaging the pseudo-GLM observations over a longer interval but assimilating them every minute (experiment GVa3) produced somewhat poorer results than

experiment GVa1. The results for experiment NIC were also somewhat inferior to those obtained using the graupel content operators, but overall the reflectivity structures in all experiments had reasonably realistic characteristics.

Figure 15 shows the cold pools produced by each of the experiments at 2209 UTC and 0.25 km AGL in the form of analysis mean potential temperature perturbations. Differences due to the slope of the GV observation operator is once again apparent here, with
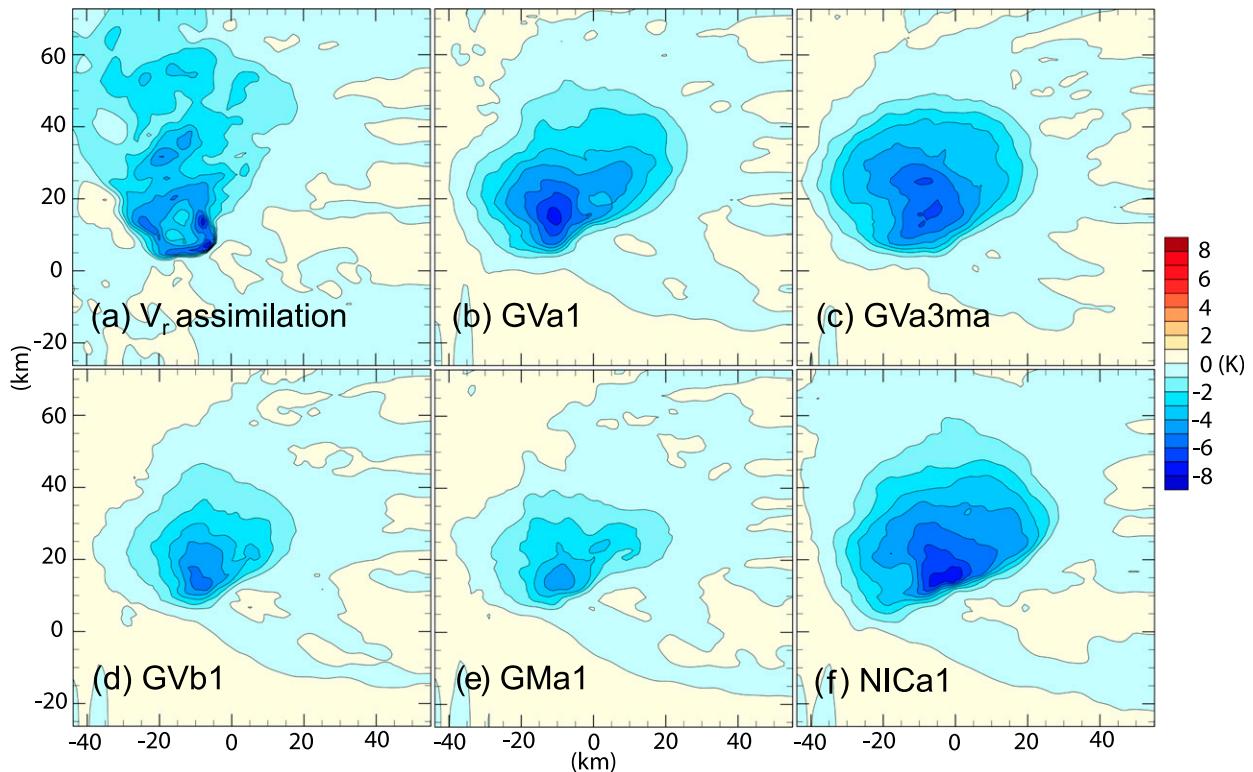
FIG. 15. (a) EnKF analysis mean potential temperature perturbation values at 250 m at 2209 UTC when assimilating KOUN V_r, and (b)–(f) EnKF analysis mean theta perturbation values at 250 m at 2209 UTC for each of the lightning data assimilation experiments.

experiment GVa1 producing a stronger cold pool than experiment GVb1. The V_r assimilation experiment produces a more detailed cold pool structure than any of the lightning assimilation runs, and although the minimum potential temperature values in the V_r assimilation experiment are similar to the minimum values in the lightning assimilation experiments with the strongest cold pools (e.g., GVa1, GVa3, NIC), the total area of perturbations in excess of −4 K is larger in the lightning assimilation experiments. Finally, compared to the V_r assimilation experiment, the cold pools are centered farther east in many of the lightning assimilation experiments. Observations of surface temperature for this case are quite limited, but the magnitude of the temperature reductions are similar to the roughly 3–6-K reductions that occurred as portions of the cold pool transited three Oklahoma Mesonet sites as shown in Yussouf et al. (2013). Although the maximum temperature perturbations found in the V_r, GVa1, and NIC experiments are roughly −8 K, those values are limited to very small areas. Given the spacing of the mesonet sites, such small features of the cold pool could have existed in the actual storm without being sampled.

To evaluate how well the lightning data assimilation experiments captured the mesocyclone of this storm, Fig. 16 shows the ensemble probability of vorticity above a $0.01 \, s^{-1}$ threshold at 1.25 km AGL for each of the lightning data assimilation experiments between 2100 and 2250 UTC (Figs. 16c–g). For reference, Fig. 16a shows the 0–2-km Warning Decision Support System-Integrated Information (WDSS-II) NEXRAD observed rotation track (Smith and Elmore 2004) for the storm, as originally shown in Yussouf et al. (2013), and Fig. 16b shows the same field as the lightning experiment panels, but for the V_r experiment which ended at 2210 UTC when KOUN data collection ceased. In all experiments except GVa3 and NIC the probability of vorticity above the threshold exceeds approximately 80% along a significant portion of the tornado's track, although in all of the experiments the vorticity probabilities drop off considerably before reaching the end of the track.

Turning to the reflectivity analysis innovation statistics from the beginning of the experiments through 2210 UTC, Fig. 17 shows that all of the GVa experiments significantly reduced the RMS innovation compared to the NoDA experiment. The RMS innovation decreased more quickly in the lightning assimilation experiments
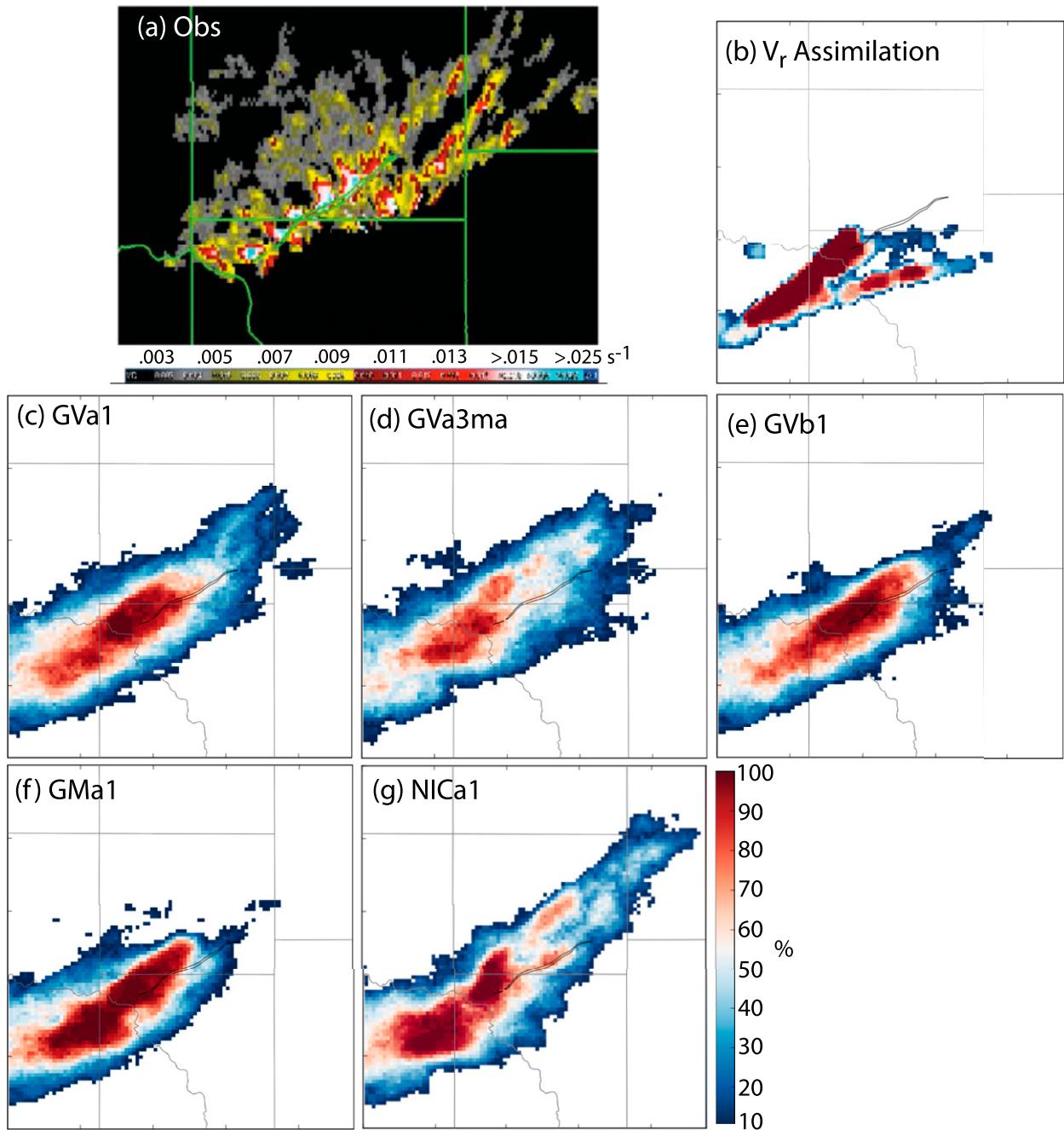
FIG. 16. (a) NEXRAD rotation track in the vicinity of the tornado track (area enclosed by thin green contour near center of plot), from Yussouf et al. (2013). (b) Probability of vorticity greater than $0.01\ s^{-1}$ at 1.25 km near the tornado track (area enclosed by thin black contour near center of plot) for the V_r assimilation experiment. (c)–(g) As in (b), but for the lightning data assimilation experiments. In (c)–(g), swaths are plotted for times between 2100 and 2250 UTC for the lightning data assimilation experiments, while the V_r assimilation experiment plot in (b) ends at 2210 UTC due to the cessation of KOUN data collection at that time.

than the V_r assimilation experiment, although later in the period the V_r assimilation produced lower RMS innovations. Little variation in the RMS innovation occurred in the GVa experiments, especially late in the period. The mean innovation suggests that the V_r experiment produced reflectivity values somewhat higher than observed for much of the period until roughly 2200 UTC.
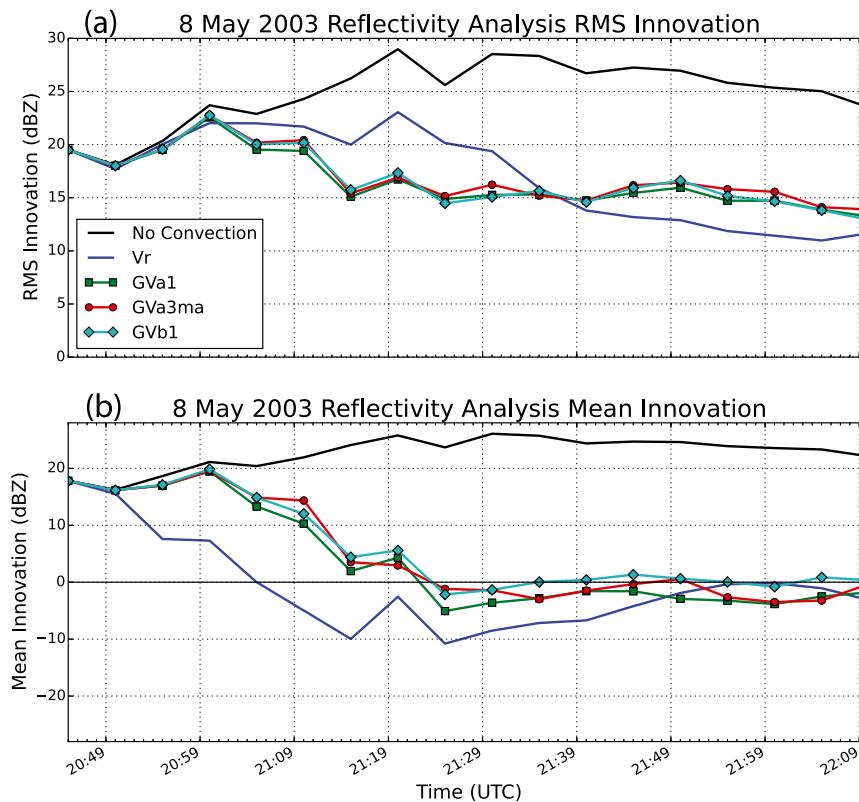
FIG. 17. Radar reflectivity (a) RMS innovation and (b) mean innovation for case-2 graupel volume observation operator experiments verified against KOUN reflectivity data. For comparison, the RMS innovation and mean innovation are also shown for the KOUN V_r assimilation experiment and for an experiment in which no data were assimilated.

Comparing the RMS innovation for experiments GVa1, GM, and NIC in Fig. 18, the GVa1 and GM experiments were very similar, with the NIC experiment having a slightly higher RMS innovation throughout much of the latter half of the period. For these experiments, the mean innovation again suggests lower and less biased reflectivity values through much of the period for the lightning assimilation experiments compared to the V_r experiment.

The pseudo-GLM RMS and mean innovation, along with the ensemble spread, are shown in Fig. 19 for the GVa1 experiment. Note that the time scale is longer than in the previous three figures to encompass the full duration of the lightning assimilation experiments. The RMS innovation in this case increases as the storm develops and then remains near 5 flashes min^{-1} pixel^{-1} for most of the remainder of the experiment. As with the multicell case, the RMS innovation remains well within the assumed observation error variance used throughout the experiment. Meanwhile, the mean innovation only contains small variations throughout the experiment and generally stays near to slightly above zero. The spread increases as convection develops and then stays

nearly constant with only a slight decrease as the run continues.

7. Conclusions

This study builds on the previous OSSE study of Mansell (2014) to demonstrate the utility of assimilating real-data, LMA-derived pseudo-GLM observations using the ensemble Kalman filter at the storm scale. Various linear best-fit observation operators were developed that related pseudo-GLM flash event densities to graupel volume, graupel mass, or noninductive charging rate. These observation operators were then used to assimilate pseudo-GLM data for two cases, and the results of these experiments were compared to observations and results produced by assimilating radar radial velocity data.

For both a nonsevere multicell storm that occurred on 6 June 2000 and a tornadic supercell that occurred on 8 May 2003, two graupel volume-based observation operators as well as a graupel mass-based observation operator showed potential to be of significant value for ensemble data assimilation in areas that lack radar data. In the multicell case, assimilating the pseudo-GLM
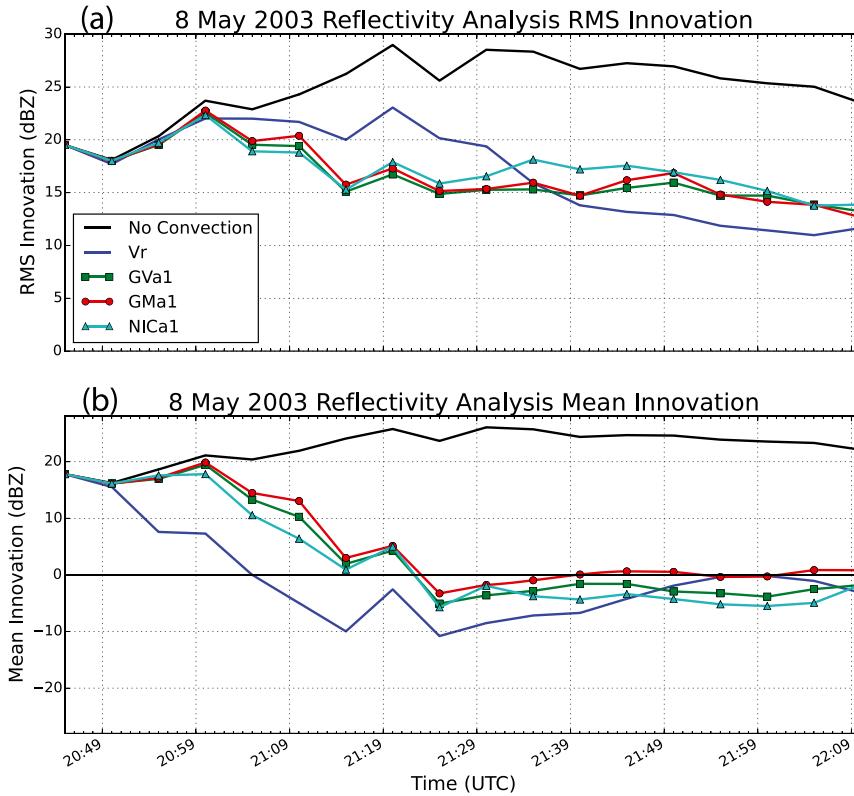
FIG. 18. As in Fig. 17, but for experiments with observation operators using different microphysical variables.

observations using the best-performing observation operators allowed for the evolution of the storm's morphology to be analyzed throughout a large portion of its lifetime. The lightning data assimilation experiments using the GVa observation operator even outperformed the radial velocity assimilation experiment in terms of how well the northern cell of the multicell

storm was represented at 6 km AGL shortly after it first appeared in the observations. In the supercell case, both the basic supercellular reflectivity structure of the storm and the development of the low-level mesocyclone were captured when assimilating pseudo-GLM observations with the best-performing observation operators.
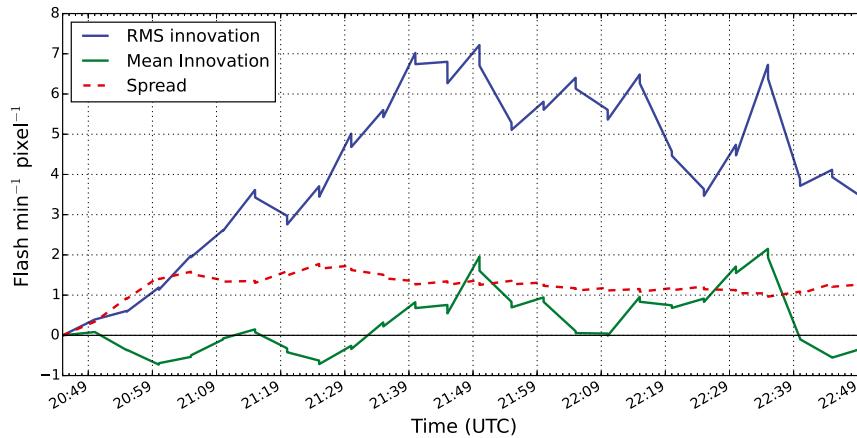

FIG. 19. Flash extent density RMS innovation, mean innovation, and ensemble spread for the case-2 GVa1 experiment.

Two findings should be emphasized:

- First, and most importantly, three observation operators (GVa, GVb, and GM) performed well when using 1-min LMA-derived pseudo-GLM data across all measures for both the strong convection of the supercell case and the much weaker convection of the multicell case.
- Second, the noninductive charging rate–based observation operator was found to have poorer results despite a significantly higher whole-storm correlation between noninductive charging rate and simulated flash rates compared to the graupel content-based observation operators. Model results suggest this may be caused by the correspondence between lightning activity and noninductive charging rate at the sub-storm scale being considerably poorer than the correspondence between those quantities when integrated across an entire storm, as shown in Fig. 6. The superior spatial correspondence between the graupel content and lightning channels is likely due to the noninductive charging rate having no way of accounting for the advection of space charge after charge separation occurs.

The results leading to the first finding are limited to analyses from only two cases, and evaluation of the forecast skill of ensembles initialized via EnKF assimilation of pseudo-GLM data will need to be carried out in the future. Nevertheless, the finding indicates that the GLM could become an important source of data for the initialization of convection-resolving ensemble numerical weather prediction models. In the future it may be useful to investigate additional observation operators, such as ice water path, which can account for the presence and advection of space charge carried on cloud ice. It will also be necessary to determine the optimal way to combine the lightning data with other convective scale observations such as radar data—it is currently unclear whether it would be beneficial to assimilate the lightning observations together with radar data or if the lightning assimilation should only be performed in areas where radar data are absent. Finally, additional future work could explore the second finding through modeling studies or by using LMA or GLM data in combination with the radar-derived flux product of Deierling et al. (2008).

REFERENCES

Alexander, G. D., J. A. Weinman, V. M. Karyampudi, W. S. Olson, and A. C. Lee, 1999: The effect of assimilating rain rates derived from satellites and lightning on the forecasts of the 1993 superstorm. *Mon. Wea. Rev.*, **127**, 1433–1457, doi:10.1175/1520-0493(1999)127<1433:TEOARR>2.0.CO;2.

Brooks, I. M., C. P. R. Saunders, R. P. Mitzeva, and S. L. Peck, 1997: The effect on thunderstorm charging of the rate of rime accretion by graupel. *Atmos. Res.*, **43**, 277–295, doi:10.1016/S0169-8095(96)00043-9.

Bruning, E. C., W. D. Rust, T. J. Schuur, D. R. MacGorman, P. R. Krehbiel, and W. Rison, 2007: Electrical and polarimetric radar observations of a multicell storm in TELEX. *Mon. Wea. Rev.*, **135**, 2525–2544, doi:10.1175/MWR3421.1.

Burgess, D. W., 2004: High resolution analyses of the 8 May 2003 Oklahoma City storm. Part I: Storm structure and evolution from radar data. *22nd Conf. on Severe Local Storms*, Hyannis, MA, Amer. Meteor. Soc., 12.4. [Available online at https://ams.confex.com/ams/11aram22sls/techprogram/paper_81939.htm.]

Calhoun, K. M., E. R. Mansell, D. R. MacGorman, and D. C. Dowell, 2014: Numerical simulations of lightning and storm charge of the 29–30 May 2004 Geary, Oklahoma, supercell thunderstorm using EnKF mobile radar data assimilation. *Mon. Wea. Rev.*, **142**, 3977–3997, doi:10.1175/MWR-D-13-00403.1.

Caya, A., J. Sun, and C. Snyder, 2005: A comparison between the 4DVAR and the ensemble Kalman filter techniques for radar data assimilation. *Mon. Wea. Rev.*, **133**, 3081–3094, doi:10.1175/MWR3021.1.

Chang, D.-E., J. A. Weinman, C. A. Morales, and W. S. Olson, 2001: The effect of spaceborne microwave and ground-based continous lightning measurements on forecasts of the 1998 Groundhog Day storm. *Mon. Wea. Rev.*, **129**, 1809–1833, doi:10.1175/1520-0493(2001)129<1809:TEOSMA>2.0.CO;2.

Cohen, A. E., 2008: Flash rate, electrical, microphysical, and kinematic relationships across a simulated storm spectrum. M.S. thesis, School of Meteorology, University of Oklahoma, Norman, OK, 71 pp.

Coniglio, M. C., D. J. Stensrud, and L. J. Wicker, 2006: Effects of upper-level shear on the structure and maintenance of strong quasi-linear mesoscale convective systems. *J. Atmos. Sci.*, **63**, 1231–1252, doi:10.1175/JAS3681.1.

Deierling, W., and W. A. Petersen, 2008: Total lightning activity as an indicator of updraft characteristics. *J. Geophys. Res.*, **113**, D16210, doi:10.1029/2007JD009598.

——, ——, J. Latham, S. Ellis, and H. J. Christian, 2008: The relationship between lightning activity and ice fluxes in thunderstorms. *J. Geophys. Res.*, **113**, D15210, doi:10.1029/2007JD009700.

Dowell, D. C., and L. J. Wicker, 2009: Additive noise for storm-scale ensemble data assimilation. *J. Atmos. Oceanic Technol.*, **26**, 911–927, doi:10.1175/2008JTECHA1156.1.

——, ——, and D. J. Stensrud, 2004a: High resolution analyses of the 8 May 2003 Oklahoma City storm. Part II: EnKF data assimilation and forecast experiments. *22nd Conf. on Severe Local Storms*, Hyannis, MA, Amer. Meteor. Soc., 12.5. [Available online at https://ams.confex.com/ams/11aram22sls/techprogram/paper_81393.htm.]

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

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

Fierro, A. O., and J. M. Reisner, 2011: High-resolution simulation of the electrification and lightning of Hurricane Rita during the period of rapid intensification. *J. Atmos. Sci.*, **68**, 477–494, doi:10.1175/2010JAS3659.1.

——, 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, doi:10.1175/MWR-D-11-00299.1.

——, ——, ——, and ——, 2015: Explicitly simulated electrification and lightning within a tropical cyclone based on the environment of Hurricane Isaac (2012). *J. Atmos. Sci.*, **72**, 4167–4193, doi:10.1175/JAS-D-14-0374.1.

Gaspari, G., and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.*, **125**, 723–757, doi: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, doi:10.1029/GL015i011p01185.

——, and Coauthors, 2013: The GOES-R Geostationary Lightning Mapper (GLM). *Atmos. Res.*, **125–126**, 34–49, doi:10.1016/j.atmosres.2013.01.006.

Hakim, G. J., P. Regulski, C. Mass, and R. Torn, 2008: Lightning data assimilation using an ensemble Kalman filter. *20th Int. Lightning Detection Conf.*, Tucson, AZ, Vaisala, 6 pp. [Available online at http://www.vaisala.com/en/events/ildcilmc/Documents/Lightning%20Data%20Assimilation%20using%20an%20Ensemble%20Kalman%20Filter.PDF.]

Jayaratne, E. R., C. P. R. Saunders, and J. Hallett, 1983: Laboratory studies of the charging of soft hail during ice crystal interactions. *Quart. J. Roy. Meteor. Soc.*, **109**, 609–630, doi:10.1002/qj.49710946111.

Klemp, J. B., and R. B. Wilhelmson, 1978: The simulation of three-dimensional convective storm dynamics. *J. Atmos. Sci.*, **35**, 1070–1096, doi:10.1175/1520-0469(1978)035<1070:TSOTDC>2.0.CO;2.

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, doi:10.1175/MWR3217.1.

Lang, T. J., and Coauthors, 2004: The Severe Thunderstorm Electrification and Precipitation Study. *Bull. Amer. Meteor. Soc.*, **85**, 1107–1125, doi:10.1175/BAMS-85-8-1107.

MacGorman, D. R., and W. D. Rust, 1998: *The Electrical Nature of Storms.* Oxford University Press, 422 pp.

——, 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, doi:10.1175/1520-0469(1989)046<0221:LRRTTS>2.0.CO;2.

——, and Coauthors, 2008: TELEX: The Thunderstorm Electrification and Lightning Experiment. *Bull. Amer. Meteor. Soc.*, **89**, 997–1013, doi:10.1175/2007BAMS2352.1.

Mansell, E. R., 2014: Ensemble assimilation of simulated total lightning flash rates. *Mon. Wea. Rev.*, **142**, 3683–3695, doi: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**, doi:10.1029/2000JD000244.

——, ——, ——, and ——, 2005: Charge structure and lightning sensitivity in a simulated multicell thunderstorm. *J. Geophys. Res.*, **110**, D12101, doi:10.1029/2004JD005287.

——, C. L. Ziegler, and D. R. MacGorman, 2007: A lightning data assimilation technique for mesoscale forecast models. *Mon. Wea. Rev.*, **135**, 1732–1748, doi:10.1175/MWR3387.1.

——, ——, and E. C. Bruning, 2010: Simulated electrification of a small thunderstorm with two-moment bulk microphysics. *J. Atmos. Sci.*, **67**, 171–194, doi:10.1175/2009JAS2965.1.

Marchand, M. R., and H. E. Fuelberg, 2014: Assimilation of lightning data using a nudging method involving low-level warming. *Mon. Wea. Rev.*, **142**, 4850–4870, doi:10.1175/MWR-D-14-00076.1.

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, doi: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, doi:10.1175/2009MWR2765.1.

Petersen, W. A., H. J. Christian, and S. A. Rutledge, 2005: TRMM observations of the global relationship between ice water content and lightning. *Geophys. Res. Lett.*, **32**, L14819, doi:10.1029/2005GL023236.

Reynolds, S. E., M. Brook, and M. F. Gourley, 1957: Thunderstorm charge separation. *J. Meteor.*, **14**, 426–436, doi:10.1175/1520-0469(1957)014<0426:TCS>2.0.CO;2.

Romine, G. S., D. W. Burgess, and R. B. Wilhelmson, 2008: A dual-polarization-radar-based assessment of the 8 May 2003 Oklahoma City area tornadic supercell. *Mon. Wea. Rev.*, **136**, 2849–2870, doi:10.1175/2008MWR2330.1.

Saunders, C. P. R., and S. L. Peck, 1998: Laboratory studies of the influence of the rime accretion rate on charge transfer during crystal/graupel collisions. *J. Geophys. Res.*, **103**, 13 949–13 956, doi:10.1029/97JD02644.

Smith, T. M., and K. L. Elmore, 2004: The use of radial velocity derivatives to diagnose rotation and divergence. *11th Conf. on Aviation, Range, and Aerospace*, Hyannis, MA, Amer. Meteor. Soc., P5.6. [Available online at https://ams.confex.com/ams/11aram22sls/techprogram/paper_81827.htm.]

Takahashi, T., 1978: Riming electrification as a charge generation mechanism in thunderstorms. *J. Atmos. Sci.*, **35**, 1536–1548, doi:10.1175/1520-0469(1978)035<1536:REAACG>2.0.CO;2.

Tanamachi, R. L., L. J. Wicker, D. C. Dowell, H. B. Bluestein, D. T. Dawson II, and M. Xue, 2013: EnKF assimilation of high-resolution, mobile Doppler radar data of the 4 May 2007 Greensburg, Kansas, supercell into a numerical cloud model. *Mon. Wea. Rev.*, **141**, 625–648, doi:10.1175/MWR-D-12-00099.1.

Tessendorf, S. A., K. C. Wiens, and S. A. Rutledge, 2007: Radar and lightning observations of normal and inverted polarity multicellular storms from STEPS. *Mon. Wea. Rev.*, **135**, 3682–3706, doi:10.1175/2007MWR1954.1.

Whitaker, J. S., and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130**, 1913–1924, doi:10.1175/1520-0493(2002)130<1913:EDAWPO>2.0.CO;2.

Wicker, L. J., and R. B. Wilhelmson, 1995: Simulation and analysis of tornado development and decay within a three-dimensional supercell thunderstorm. *J. Atmos. Sci.*, **52**, 2675–2703, doi:10.1175/1520-0469(1995)052<2675:SAAOTD>2.0.CO;2.

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, doi:10.1175/JAS3615.1.

Yussouf, N., E. R. Mansell, L. J. Wicker, D. M. Wheatley, and D. J. Stensrud, 2013: The ensemble Kalman filter analyses and forecasts of the 8 May 2003 Oklahoma City tornadic supercell storm using single and double moment microphysics schemes. *Mon. Wea. Rev.*, **141**, 3388–3412, doi:10.1175/MWR-D-12-00237.1.

Ziegler, C. L., 1985: Retrieval of thermal and microphysical variables in observed convective storms. Part I: Model development and preliminary testing. *J. Atmos. Sci.*, **42**, 1487–1509, doi:10.1175/1520-0469(1985)042<1487:ROTAMV>2.0.CO;2.