Variational Assimilation of Radar Data and GLM Lightning-Derived Water Vapor for the Short-Term Forecasts of High-Impact Convective Events

ALEXANDRE O. FIERRO AND YUNHENG WANG

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

JIDONG GAO AND EDWARD R. MANSELL

NOAA/National Severe Storms Laboratory, Norman, Oklahoma

(Manuscript received 30 November 2018, in final form 9 August 2019)

ABSTRACT

The assimilation of water vapor mass mixing ratio derived from total lightning data from the Geostationary Lightning Mapper (GLM) within a three-dimensional variational (3DVAR) system is evaluated for the analysis and short-term forecast (≤ 6 h) of a high-impact convective event over the northern Great Plains in the United States. Building on recent work, the lightning data assimilation (LDA) method adjusts water vapor mass mixing ratio within a fixed layer depth above the lifted condensation level by assuming nearly water-saturated conditions at observed lightning locations. In this algorithm, the total water vapor mass added by the LDA is balanced by an equal removal outside observed lightning locations. Additional refinements were also devised to partially alleviate the seasonal and geographical dependence of the original scheme. To gauge the added value of lightning, radar data (radial velocity and reflectivity) were also assimilated with or without lightning. Although the method was evaluated in quasi–real time for several high-impact weather events throughout 2018, this work will focus on one specific, illustrative severe weather case wherein the control simulation—which did not assimilate any data—was eventually able to initiate and forecast the majority of the observed storms. Given a relatively reasonable forecast in the control experiment, the GLM and radar assimilation experiments were still able to improve the short-term forecast of accumulated rainfall and composite radar reflectivity further, as measured by neighborhood-based metrics. These results held whether the simulations made use of one single 3DVAR analysis or high-frequency (10 min) successive cycling over a 1-h period.

1. Introduction

The timeliness and accuracy of severe weather warnings still remain a paramount challenge, given the need to resolve complex nonlinear processes over a continuum spanning a large range of scales and involving several different physical processes (e.g., Stensrud et al. 2009). In addition to model errors, the accuracy of numerical weather predictions (NWP) is strongly tied to biases and errors contained in the initial conditions often derived or downscaled from larger-scale model data. To partially alleviate this drawback, several data assimilation (DA) methods aimed at ingesting information from multiple observing platforms into NWP models have been developed (e.g., Park and Xu 2013).

For convective-scale (≤ 3 km) NWP, considerable advances have been made in assimilation methods for radar reflectivity and Doppler radial velocity data (e.g., Evensen 1994, 2003; Houtekamer and Mitchell 1998; Zhang 2005; Zhang et al. 2009; Aksoy et al. 2009; Godinez et al. 2012; Wang and Wang 2017). Recent DA works have demonstrated significant improvements in short-term forecasts (≤ 6 h) of high-impact weather events ranging from individual severe thunderstorms (e.g., Stensrud and Gao 2010; Schenkman et al. 2011; Yussouf et al. 2016; Jones et al. 2016) to tropical cyclones (e.g., Zhao and Xue 2009) with the help of radar data. Many areas of the United States, however, suffer from poor coverage by the U.S. National Weather Service (NWS) operational Weather Surveillance Radar–1988 Doppler (WSR-88D) network (e.g., Zhang et al. 2011). Over the United States, regions particularly vulnerable to poor radar coverage are mountainous terrain

Corresponding author: Alexandre O. Fierro, alex.fierro@noaa.gov

of the west (beam blockage) often prone to flash flooding or oceanic regions beyond the range of coastal radars where tropical systems develop and eventually intensify. In these circumstances, atmospheric and oceanic measurements from spaceborne, remote sensing platforms play a pivotal role in filling this gap (e.g., Jones et al. 2018; Minamide and Zhang 2018; Zhang et al. 2018). To complement the coverage of surface lightning detection networks over most of the western hemisphere, a joint collaboration between NOAA and NASA fostered the successful deployments of the *Geostationary Operational Environmental Satellite-16 and -17* (*GOES-16 and -17*) (Gurka et al. 2006; Goodman et al. 2013). For convection-allowing models, lightning data mapped from space can systematically identify areas of deep, mixed-phase convection, which can be readily assimilated into NWPs.

Earlier lightning data assimilation (LDA) studies used cloud-to-ground (CG) flash data (e.g., Jones and Macpherson 1997a,b; Alexander et al. 1999; Chang et al. 2001; Papadopoulos et al. 2005; Pessi and Businger 2009) or a combination of CG and limited-area total lightning data (Mansell et al. 2007). These studies devised empirical relationships to modify or replace latent heating rates from a convection parameterization scheme, which are not suitable for convection-allowing scales. Observational and cloud-scale modeling studies have shown much higher correlations of storm properties with total lightning rates than with CG flash rates alone, as intracloud (IC) flashes generally vastly outnumber CG flashes (e.g., MacGorman et al. 1989; Carey and Rutledge 1998; Wiens et al. 2005; Kuhlman et al. 2006; Fierro et al. 2006; Deierling and Petersen 2008; MacGorman et al. 2011; Boccippio et al. 2001; Medici et al. 2017; Weiss et al. 2012).

Thus, *total* lightning data are considered more appropriate for convection-allowing applications (Marchand and Fuelberg 2014; Fierro et al. 2012a, 2014, 2015a, hereafter F15; H. Wang et al. 2017, 2018; Mansell 2014; Fierro et al. 2016, hereafter F16). Some of these works highlighted that the assimilation of pseudo-GLM total lightning [i.e., ground-based detections scaled to assumed Geostationary Lightning Mapper (GLM) parameters] alone using either simple nudging or more sophisticated variational or ensemble Kalman filter techniques was able to notably improve analyses (Mansell 2014; Allen et al. 2016) and short-term forecasts of convective events (F16). F16 underscored, however, that the added value of lightning data assimilation (LDA) to the forecasts remains relatively minor in areas that already have good spatiotemporal coverage by the WSR-88D network. This is expected, given that the level of convective-scale information contained within successive volumetric radar scans far surpasses that of the (accumulated) two-dimensional flash densities derived from either ground-based or spaceborne platforms.

The present variational LDA method builds upon an approach first presented and evaluated in F16. At observed lightning locations, F16's method imposed incremental adjustments (i.e., increases) of water vapor mass mixing ratio (q_v) toward or near saturation within a confined layer, which promoted the initiation of convection through local enhancements of thermal buoyancy (Houze 1993; Braun 2002; Lopez and Bauer 2007; Caumont et al. 2010; Fierro et al. 2012b). Despite notable improvements in the short-term forecasts of storm structure, rainfall and reflectivity, F16's method had some shortcomings, which include: (i) the inability to effectively suppress spurious convection, (ii) the tendency to notably exacerbate any wet biases present in the simulation assimilating no data, and (iii) the q_v mass being added was not balanced by an equal removal in the domain prior to performing the 3DVAR analysis. As will be described later, the proposed LDA scheme in this study offers an attempt to address (ii) and (iii).

The present study builds upon F16 by: (i) testing the variational LDA approach using the total lightning densities derived from the GLM in lieu of ground-based systems, (ii) combining the assimilation of GLM with radar data (radial velocity and reflectivity) in a variational framework (without the use of a cloud analysis scheme), and (iii) by extending the LDA algorithm such that any water vapor mass added in the pseudo q_v observations is offset by an equal removal outside lightning locations. Additionally, the proposed LDA method is independent of flash rate. The chief motivation for this research is to leverage the systematic availability of GLM-observed total lightning data over most of the western hemisphere, which provide continuous monitoring of electrified convection over large areas devoid of ground-based observations.

Although the present LDA scheme, including its parallel experiments described later (Table 1), was tested in quasi–real time for about a dozen of cases throughout 2018, this study will place emphasis on one illustrative case study: the 7 June 2018 severe weather outbreak over the northern plains of the United States, which consisted of several mesoscale convective systems (MCSs) in addition to some isolated severe storm cells. This case was deemed broadly representative because the improvements and limitations seen did not stand out as particularly better or worse. Additionally, this case was the only one with a variety of storm modes entirely captured by the domain throughout the 6-h forecast period. Given the small number of cases tested thus far

TABLE 1. The left column lists the nomenclature/abbreviations used for all the simulations/experiments analyzed in this study. The second column from the left briefly describes the type of experiments. The second column from the right indicates the type of data that were assimilated with "dBZ" standing for radar reflectivity and "Vr" for radial velocity. The right column shows which model variable(s) is (are) impacted by the respective assimilation experiments with the symbols used to identify those variables bearing their usual meaning. For convenience, the experiments are listed in the order they appear in the text, which is the same as Fig. 3.

Experiments	Description	Data assimilated	Model variables impacted
CTRL	Control run	None	None
GLM	Lightning assimilation run	GLM flash density rates	q_v (LCL–LCL + 3 km)
RAD	Radar data assimilation run	Vr and dBZ	$q_r, q_g, q_s, q_h, u, v, w, \theta$
RAD + GLM	Lightning + radar assimilation run	GLM flash density rates, Vr, dBZ	q_v (LCL–LCL + 3 km), $q_r, q_g, q_s,$ q_h, u, v, w, θ
GLM CONS	Lightning assimilation run with balanced q_v field prior to the 3DVAR analysis	GLM flash density rates	q_v

and the relative embryonic stage of data assimilation efforts involving spaceborne total lightning data in the variational framework–let alone in combination with radar data – we believe that focusing one case study herein is reasonable. Additional rationales for selecting this case are: (i) the MCSs and isolated cells produced sufficient amount of lightning to perform the DA, (ii) the varying evolutions of individual systems with one of the major MCSs dissipating quickly, in contrast to others growing upscale, and (iii) the control forecast (without data assimilation) already produced a relatively reasonable forecast of the main storms. It will be shown that, despite a relatively good control forecast, the variational assimilation of either lightning and/or radar data (radial winds and reflectivity fields) is still able to improve the short-term evolution of some of the main storms.

2. Brief description of the synoptic setup

At about 0000 UTC 7 June 2018, large-scale environmental conditions were favorable for the development of organized convection over most of Nebraska and Iowa in the United States. The 0000 UTC analysis from the 13-km Rapid Refresh [(RAP), formerly the Rapid Update Cycle (RUC); Benjamin et al. (2004)] (Fig. 1) over the simulation domain (Fig. 2) indicated the presence of widespread, strong inertial instability with surface-based convective available potential energy (CAPE) values ranging between 2000 and 4500+ J kg^{-1} over most of these two states coupled with overall weak (> -50 J kg^{-1}) surface-based convective inhibition (CIN, Figs. 1a,b). The presence of a quasi-stationary surface boundary (Fig. 1) in central Iowa and Nebraska aided convective development further in these areas, which were characterized by near-surface (2 m) temperatures greater than 300 K (27°–33°C, Fig. 1d) with dewpoints nearing 290 K (15°–20°C, Fig. 2c). Note that areas with CIN < -200 J kg^{-1} near or along this boundary indicate where deep moist convection was ongoing in the analysis

(Figs. 1b and 3b). Over central Iowa and far eastern Nebraska, deep layer wind shear (not shown) was marginal at best [20–30 kt (1 kt $\approx$ 0.5144 m s^{-1})] to support organized convection. Over western Nebraska, however, a strengthening (30–40+ kt) low-level jet later in the evening hours favored upscale growth (Fig. 3b) into an MCS. Both of these systems produced several dozen severe wind and hail reports (not shown). Because of these widespread favorable conditions, additional severe-warned storms developed over this area (Fig. 3b), which will be discussed as needed during the analysis.

3. Data used for assimilation and validation

As indicated in the introduction, this work makes use of total lightning data from the GLM instrument on board the *GOES-16* satellite. The camera pixels of the GLM detect lightning activity day and night with a horizontal resolution ranging between about 8 km near the center of the field of view to about 12 km near its edges with an expected daily averaged detection efficiency exceeding 70% (Goodman et al. 2013; Rudlosky et al. 2019). Based on evaluation studies of its predecessor instrument–the Lightning Imaging Sensor (LIS, Christian et al. 1999; Albrecht et al. 2016)–the detection efficiency of the GLM will likely vary depending on the time of day, geographical location (e.g., Fuchs et al. 2016) and optical cloud depth (Yoshida et al. 2009). The basic lightning detections from the GLM are pixel-level lightning optical energy "events," which are algorithmically combined first into groups and then flashes. In this study, only the flash product will be considered, which is defined as one or more lightning groups that satisfy temporal and spatial coincidence thresholds [cf. Fig. 5 in Goodman et al. (2013) for an illustration]. As indicated in Mach et al. (2007), a group could be viewed as a single lightning pulse (e.g., return stroke) spanning multiple pixel events, and a flash as an amalgamation of one or
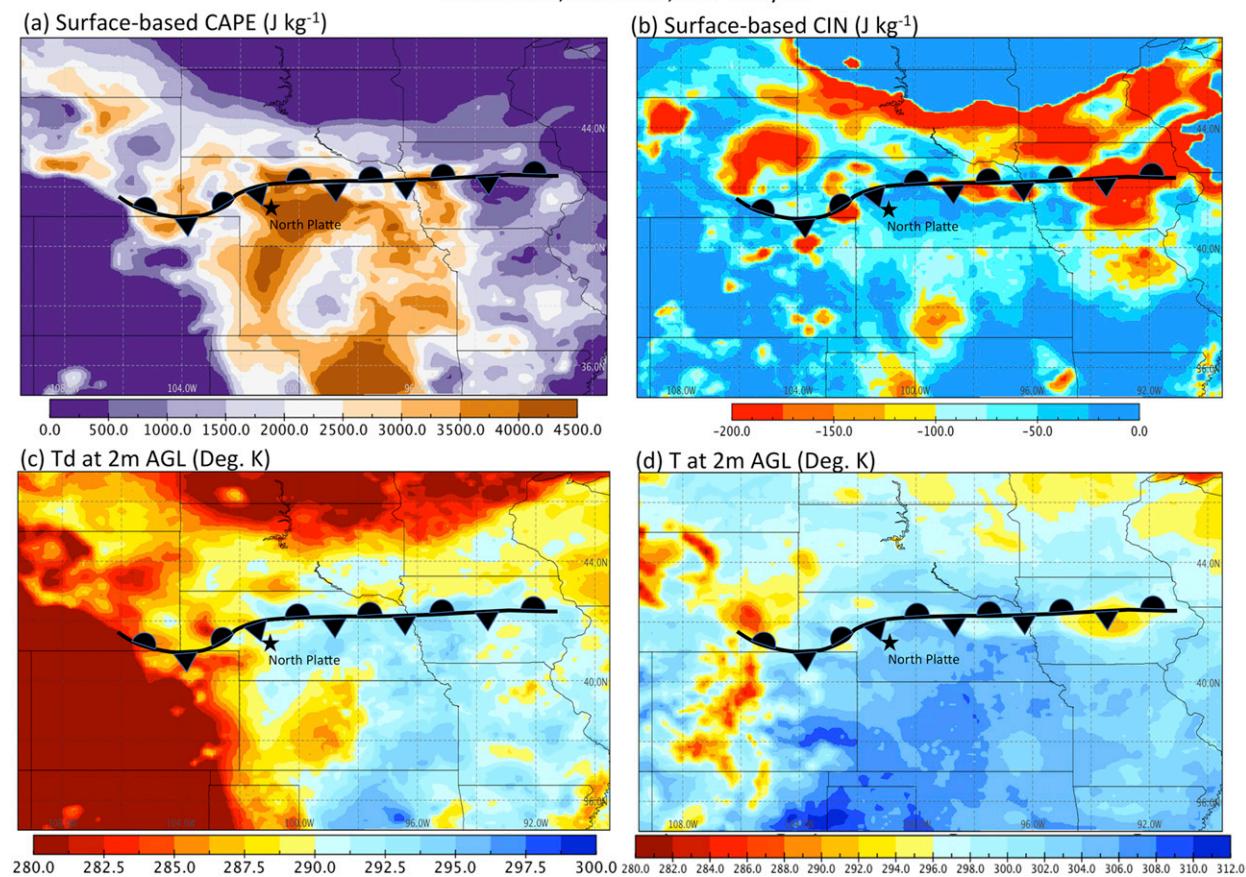
7 June 2018, 0000 UTC, RAP analysis

FIG. 1. RAP analysis output at 0000 UTC 7 Jun 2018 of (a) surface-based CAPE (J kg^{-1}), (b) surface-based CIN (J kg^{-1}), and 2 m AGL (c) dewpoint temperature (K) and (d) temperature (K). The position of the surface boundary is shown as a black scalloped–triangled curve in each panel for reference. The NWS operational sounding site at North Platte, NE (LBF), is marked by a black star.

more lightning pulses (e.g., negative CG flash producing multiple return strokes). Each flash is associated with a latitude–longitude coordinate of its pixel centroid. For the modeling experiments, the lightning flash data centroids were accumulated for the 1-h period prior to the initialization time of the free forecast (viz., 2300 UTC 6 June–0000 UTC 7 June), and then projected onto the domain's uniform 3-km Mercator grid (Fig. 2).

Following F16, the modeled reflectivity fields are evaluated against observed composite radar reflectivity fields from the National Severe Storms Laboratory (NSSL) Multi-Radar Multi-Sensor (MRMS) product (Zhang et al. 2011; Smith et al. 2016), available in 5-min increments with a horizontal grid spacing of 0.01°. For the assimilation, WSR-88D Level-II data (reflectivity factor, mean radial velocity, and spectrum width) were obtained from the National Centers for Environmental Information.[1]

The simulated accumulated precipitation fields were evaluated against the National Centers for Environmental Prediction's stage IV[2] multisensor hourly rainfall accumulation estimates (Baldwin and Mitchell 1997). To compare model and observations, the hourly stage-IV data were remapped from the native 4-km polar stereographic grid onto the 3-km Mercator grid of the simulation domain (Fig. 2).

4. Simulations setup

a. Model grid and physics configuration

The forecast model used in this study is the three-dimensional compressible nonhydrostatic WRF Model (version 3.6.1) with Advanced Research WRF (ARW) dynamic solver (WRF-ARW, Skamarock and Klemp 2008). The DA experiments are performed on one domain with a uniform horizontal grid spacing of 3 km

---

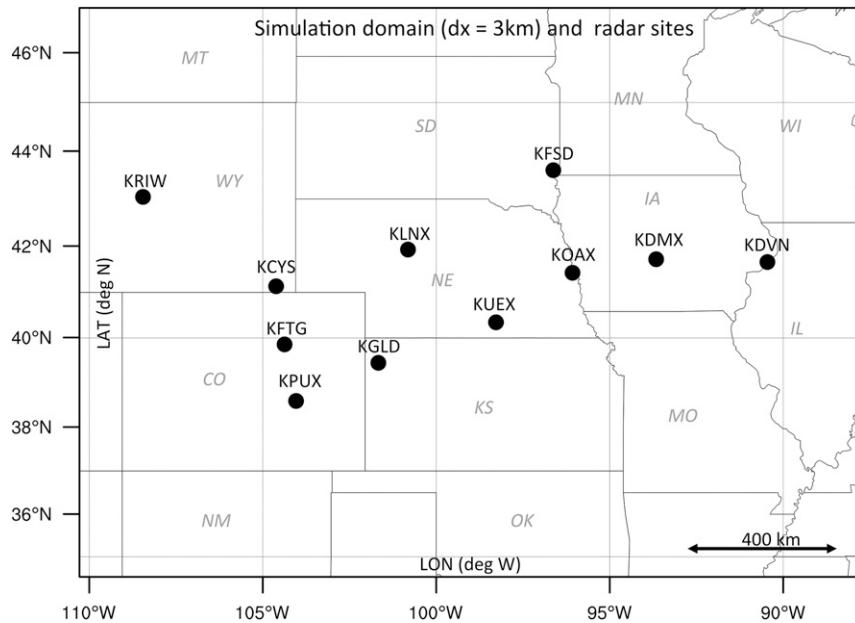[1] http://www.ncdc.noaa.gov/.

[2] https://www.eol.ucar.edu.

FIG. 2. Sketch of the simulation domain (3-km horizontal grid spacing) spanning the north-central Great Plains of the United States with the black dots denoting the locations of the WSR-88D sites used and tested in the NEW3DVAR code. The U.S. states are indicated by their usual abbreviations in gray italicized fonts.

and horizontal dimensions in grid points of 661 × 481 (Fig. 2). The stretched vertical grid has 35 levels with its top at 50 hPa (~20 km) and the computational time step is 10 s. The initial and time-dependent lateral boundary conditions were derived from the 3-hourly, 12-km North American Mesoscale Forecast System (NAM) initialized at 0000 UTC 7 June 2018 and subsequent forecast data for a 6-h period. Because this work solely focuses on short-term forecasts (≤6 h), the simulations are ended at 0600 UTC 7 June.

Convection was ongoing in the simulation domain at the analysis time (i.e., 0000 UTC 7 June, Figs. 1 and 3). As in F16, no model spinup was performed prior to the 3DVAR analysis, as conducting the DA when observed convection is active in the target area facilitates the qualitative and quantitative analysis of the impacts of the respective DA procedures. More specifically, the study will illustrate how the assimilation of GLM lightning and/or radar data will help hasten the development of most of the observed storms and, in turn, improve the placement and intensity of some of the storm-scale objects of interest.

The simulations employed the NSSL two-moment, four-ice category bulk microphysics scheme (Ziegler 1985; Mansell et al. 2010; Mansell and Ziegler 2013). No cumulus parameterization is used. The boundary layer is parameterized following the 1.5-order closure Mellor–Yamada scheme (Mellor and Yamada 1982) turbulence kinetic energy scheme adapted by Janjić (1994) with

Monin–Obukhov–Janjić similarity theory for the subgrid-scale turbulence processes (Chen et al. 1997). Lower boundary conditions for the boundary layer fluxes are provided by the Unified Noah land surface model (Ek et al. 2003). Atmospheric radiation is parameterized following Dudhia (1989) for the shortwaves and the Rapid Radiative Transfer Model (RRTM) for the longwaves (Mlawer et al. 1997) and called by the model every 10 min.

b. Data assimilation procedures

This work makes use of an upgraded version of the three-dimensional variational data assimilation system initially developed for the Advanced Regional Prediction System (ARPS) (Gao et al. 1999; Xue et al. 2001, 2003; Gao et al. 2004; Hu et al. 2006a,b; Stensrud and Gao 2010; Ge et al. 2010, 2012; Gao et al. 2013) referred to as the National Severe Storms Laboratory Experimental Warn-on-Forecast System for 3DVAR (NEWS3DVAR; Gao et al. 2016; Y. Wang et al. 2018). Though there are many advanced DA methods which could be alternatively employed to assimilate lightning and/or radar data (e.g., Park and Xu 2013; Mansell 2014; Allen et al. 2016; H. Wang et al. 2017, 2018), NEWS3DVAR is chosen for its efficiency and the need for rapid delivery of high-resolution convective-scale NWP products to the end users for future real-time applications.

The performance testing of the LDA method follows a similar set of experiments as in F16 (Table 1). A control
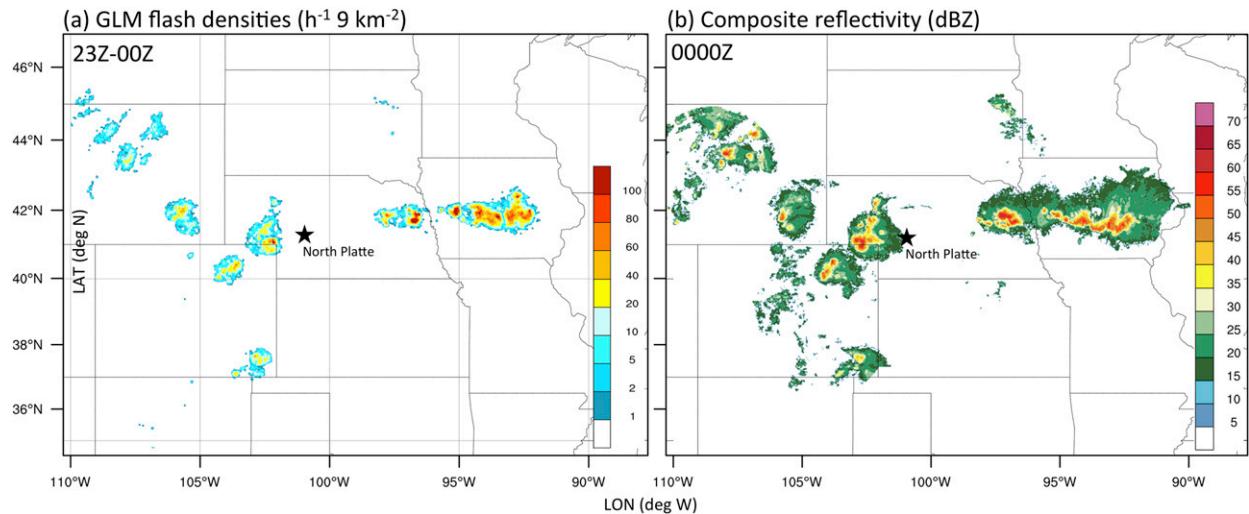
FIG. 3. (a) GLM-derived 1-h flash densities (per 9 km^2) accumulated between 2300 UTC 6 Jun until the analysis time at 0000 UTC 7 Jun 2018. (b) Composite radar reflectivity fields (dBZ) constructed using Level II data within a 5-min interval centered at 0000 UTC 7 Jun 2018. For reference, a black star denotes the location of the soundings discussed in the analysis.

run with no data assimilation (labeled as CTRL) is evaluated against the following DA experiments: (i) 3DVAR assimilation of flash density data from the GLM in Fig. 3a (denoted as GLM), (ii) 3DVAR assimilation of Level II radial velocity and radar reflectivity data (RAD), (iii) the combination of both (RAD + GLM), and (iv) the same as GLM but with q_v pseudo-observations having equal amounts of q_v mass added and removed throughout the domain (GLM CONS). Table 1 summarizes which variable(s) were adjusted in each experiment with the two following subsections describing their respective 3DVAR setup.

To provide a preliminary, broader picture of the effectiveness of the assimilation procedures, a first set of experiments employing only one 3DVAR analysis (viz., at 0000 UTC) was performed followed by a second set using successive 3DVAR cycling with 10-min frequency for both radar and lightning over a 1 h period prior to 0000 UTC. A chief motivation for high-frequency (≤15 min) cycling in cloud-scale DA applications is its ability to better depict the rapid evolution and movement of convective-scale events leading to more accurate forecasts (e.g., Gao and Stensrud 2012). Additionally, the sequential assimilation of observations has been shown to notably help alleviate model spin up issues often arising from the downscaling of coarser-resolution data (e.g., reanalysis datasets) onto the finer-scale model grid. This is because high-frequency cycling takes greater advantage of the observations by introducing a larger amount of information from these observations with more frequent smaller adjustments at finer scales. For convective-scale applications, however,

cycles generally exceeding 30 min to 1 h will be more prone to generate larger analysis increments and, thus, to overfit the observational data. Additionally, such longer cycles can cause temporal representation errors (e.g., Gao et al. 2004; Gao and Stensrud 2012). When assimilating radar reflectivity variationally as done herein without the usage of cycling, the initialization problem will more likely become underdetermined because reflectivity factor is a function of three variables (as described in the subsections below). When assimilating both lightning and radar data without cycling, the initialization will thus be more likely to be underdetermined.

The cycling frequency for the GLM and/or radar data used was 10 min to allow the variational analyses to better depict/capture the movement and evolution of individual storm-scale objects (e.g., Gao et al. 2013). Because this work chiefly focuses on the GLM lightning, the 3DVAR cycling experiments herein were conducted over the same accumulation interval selected in the original GLM experiment, namely from 2300 UTC 6 June to 0000 UTC 7 June 2018. While the cycled GLM-based experiments assimilate the same lightning information with these settings, two salient differences must be underscored for the cycled RAD-based experiments with respect to their original counterpart, which are: (i) the usage of notably more storm-scale information (radial winds and mass mixing ratios) and (ii) the usage of the 1200 UTC NAM forecast data to derive the initial conditions for the first cycle at 2300 UTC. Thus, consideration must be taken when comparing their respective forecast performances.

Multipass recursive filters from larger to finer length scales were shown to generally yield superior forecast performance over single-pass 3DVAR methods (Xie et al. 2011; Li et al. 2010). This study makes use of a two-pass recursive filter, with each pass employing a prescribed, fixed horizontal and vertical decorrelation length scale for its control variables–defined in the subsections below.

1) RADAR

In this work, volumetric WSR-88D scans of Level II radial velocity and reflectivity factor were assimilated. Prior to performing the DA, the Level II data from each radar sites (Fig. 2) are quality-controlled (e.g., de-aliasing radial velocity, removing nonmeteorological scatters) and interpolated onto the WRF grid for thinning purpose. When reflectivity data from multiple radars overlap at a given grid point, the largest value is chosen. Akin to F16, only full radar volumes contained within a 5 min interval centered on the 3DVAR analysis time were employed. Radar observations are processed and assimilated by volumes and not by individual tilts. Thus, the time lag induced by each sequential elevation sweep is not accounted for in this DA setup.

In the upgraded version of the 3DVAR program used herein (Gao et al. 2004; Gao and Stensrud 2012; Gao et al. 2016; Y. Wang et al. 2018), the capability for assimilating radar reflectivity in the 3DVAR framework was examined, which considers adjustments to the mass mixing ratios for rainwater, snow and hail/graupel. The adjustments made to these hydrometeor species are dependent on the local ambient temperature such that, for instance, snow mass is not systematically introduced at temperatures below freezing. Their background error value was set to $0.1 \, \text{g} \, \text{kg}^{-1}$. When these three species are adjusted by the 3DVAR analysis, consideration is taken to also slightly adjust θ owing to latent heat exchange during phase changes (Albers et al. 1996; Zhang et al. 1998; Brewster 2002; Hu et al. 2006a,b). While it is also possible to adjust q_v (or relative humidity) when assimilating radar reflectivity, q_v is solely adjusted by the lightning observations to avoid overlap and to facilitate the interpretation of the results (F16). To better extract convective-scale information (Gao et al. 2013), each 3DVAR analysis for the reflectivity factor employs two successive passes with respective horizontal decorrelation length scales set to 24 and 12 km. In the vertical, the decorrelation length for each of the two passes was set to 4 and 2 grid points, respectively. These decorrelation lengths for the reflectivity factor are applied to all its associated control variables, namely rain, snow and graupel/hail mass mixing ratio. The background error variances for the u, v component of the wind, θ, and

pressure are derived from statistics of the 13-km RUC 3-h forecasts in spring seasons over several years and the background error correlations are modeled by the recursive filter from Purser et al. (2003a,b).

In the 3DVAR analysis, the radial velocity data are used to adjust the three Cartesian components of the wind field. For the first and second 3DVAR pass, the control variables for the wind field use horizontal decorrelation length scales of 12 and 6 km, respectively. In the vertical, a decorrelation length of two grid points was used for both 3DVAR passes. Although not considered in this work, both θ, and the pressure fields could also be adjusted if other types of observations such as mesonet/sounding data were also assimilated. There are no assumed cross correlations between the wind field, θ and pressure.

Reflectivity and radial velocity data were used from 11 radar sites (Fig. 2). These sites were selected to provide a reasonable coverage of all the main storms during the target period of analysis, particularly above $\sim 3 \, \text{km}$ AGL where the scan volumes overlap (Zhang et al. 2011; Gao et al. 2013). Observations for reflectivity values $\leq 20 \, \text{d}BZ$ are discarded for altitudes below $1.5 \, \text{km}$ AGL and $\leq 15 \, \text{d}BZ$ at higher altitude in order to reduce unwanted influence of weak radar returns or nonmeteorological scatters.

2) LIGHTNING

Using a similar procedure as Fierro et al. (2018a), the raw, 20-s GLM total flash data first are accumulated into a 1-h interval prior to 0000 UTC 7 June and converted into gridded flash density rates (units of $9 \, \text{km}^{-2} \, \text{h}^{-1}$). Given the latitude and longitude points of the WRF grid, the projection of the GLM flash longitude and latitude centroids onto the model grid is relatively straightforward (Fierro et al. 2018a). For simplicity, this GLM-derived lightning metric is referred to as "flash origin density," because it purposively does not consider the areal extent of the flashes (Fierro et al. 2018a) to restrict the impact of the LDA near convective cores (Fierro et al. 2018b).

As mentioned in the introduction, the general LDA practice herein follows a philosophy similar to F16 wherein the intended use of total lightning observations is to derive, and then assimilate pseudo observations for q_v. At each grid point (i, j) where the GLM flash density rate exceed zero, the LDA creates pseudo observations for q_v by adjusting at that horizontal location (i, j) the background q_v value within an assumed fixed depth of 3 km above the lifted condensation level (LCL, a surrogate for cloud base). At each grid points within this 3-km deep layer, the background q_v value is adjusted to its near-saturation value (set to 95%) with respect to

water substance ($q_{v\text{sat}}$) provided that the background q_v does not already exceed that value. In other words, if the background q_v value at a grid point within one of these 3-km-deep, nonzero GLM columns already equates, for instance, $0.98 \times q_{v\text{sat}}$ this value will be used as pseudo-observation for q_v instead of $0.95 \times q_{v\text{sat}}$. The use of a fixed, shallower depth above the LCL compared to a fixed top height of 15 km above the LCL in F16, confines the adjustments within the moisture-rich layer, which, as shown by F16, yields nearly identical forecast improvements. Because the LCL exhibits notable variations as a function of season and geographical areas, the current implementation limits the maximum height of the LCL computed at lightning locations to 2 km above mean sea level (MSL). Tests using larger maximum LCL heights up to 3.5 km yielded overall similar results (not shown). The primary effect of this procedure is to enhance the influence of q_v adjustments in higher-LCL areas such as: (i) higher terrain, (ii) drier environment/boundary layer, and (iii) cold season. This additional refinement relative to F16 partially alleviates the seasonal and geographical dependence of the LDA scheme. Similar to F16, the q_v increases are not made proportional to the observed GLM gridded density rate. In contrast to F16, no arbitrary lower cutoff thresholds for flash density rates were used to differentiate between missing and valid observations points. Simply put, wherever the GLM-observed gridded hourly flash density rate is zero, the pseudo q_v observations are not created and the control variable for q_v is set to an arbitrary (large negative) value flagged by the 3DVAR system as a missing observation. Owing to the relatively recent availability of GLM data, no additional quality control algorithm is applied to lightning products, which already have a quality control procedure, and, thus, all pseudo observations for q_v are assimilated (i.e., none are rejected).

Once created, these pseudo q_v observations are then assimilated through the 3DVAR analysis. As in F16, the observation error for q_v was set to 3×10^{-3} kg kg^{-1} and the background error for q_v is set to 10×10^{-3} kg kg^{-1}. The larger background error value gives more weight to the observations during the 3DVAR analysis. Because the current code setup only permits separate namelist entries for the decorrelation length for the assimilation of the reflectivity factor, the vertical and horizontal decorrelation lengths used for q_v (i.e., assimilation of lightning) are the same as for the three Cartesian components of the wind field (i.e., assimilation of radial velocity). Namely, the control variable for q_v uses horizontal decorrelation length scales of 12 and 6 km for the first and second 3DVAR pass, respectively. In the vertical, a decorrelation length of two grid points was used for both 3DVAR passes. Lightning was assimilated only on the second 3DVAR pass, as F16 and Fierro et al. (2018b) found that q_v adjustments were most effective on horizontal length scales of 10 km or less.

As noted in the introduction, an additional refinement to F16's LDA method was devised in which the total amount of q_v mass added during the creation of the q_v pseudo observations is compensated by an equal removal outside lightning locations to maintain global mass conservation (experiment GLM CONS). To achieve this, the total q_v mass added by the LDA scheme is divided by the total sum of grid points outside the lightning areas. The gridded averaged q_v mass is then weighted by the air density value at each grid point and subtracted from the background q_v value at that grid point. Owing to the generally small areal coverage/footprint of the total area containing nonzero flash density rates (e.g., Fig. 3a), the (density-weighted) averaged q_v mass that is removed per grid point is generally quite small (maximum on the order of 9×10^{-5} g kg^{-1} herein) and, thus, is not expected to significantly impact any improvements seen in the original GLM experiment. Auxiliary sensitivity tests (not shown) revealed that these incremental negative adjustments in q_v outside the lightning areas have negligible effects on spurious convection. Although the q_v adjustments in the pseudo observations are balanced, this is not necessarily guaranteed for the q_v fields produced by the 3DVAR analysis. With this methodology, however, the post 3DVAR departure from balance in the analyzed q_v fields will generally be small.

5. Results

For all the simulations, focus is primarily directed on the two MCSs that developed in central Iowa and western Nebraska (section 2, Fig. 3b). Isolated non-severe convection that developed in the domain will be discussed whenever appropriate to highlight issues related—for instance—to spurious convection.

a. Initialization

A broad depiction of how each DA experiment adjusts the q_v field is shown in Fig. 4. As anticipated from the small horizontal decorrelation length scale chosen for lightning (6 km), the layer-averaged ($z = 3$–7 km MSL) q_v values at observed lightning locations (Figs. 3a and 4a) have adjustment areas (increases) in GLM-based DA experiments that closely match the lightning observations (Figs. 4b,c,e), which helps to avoid the development of unrealistically wide updrafts that could occur for larger length scales (F16; Fierro et al. 2018b). The relative increases in layer-averaged q_v range from a minimum of 0.2 to values
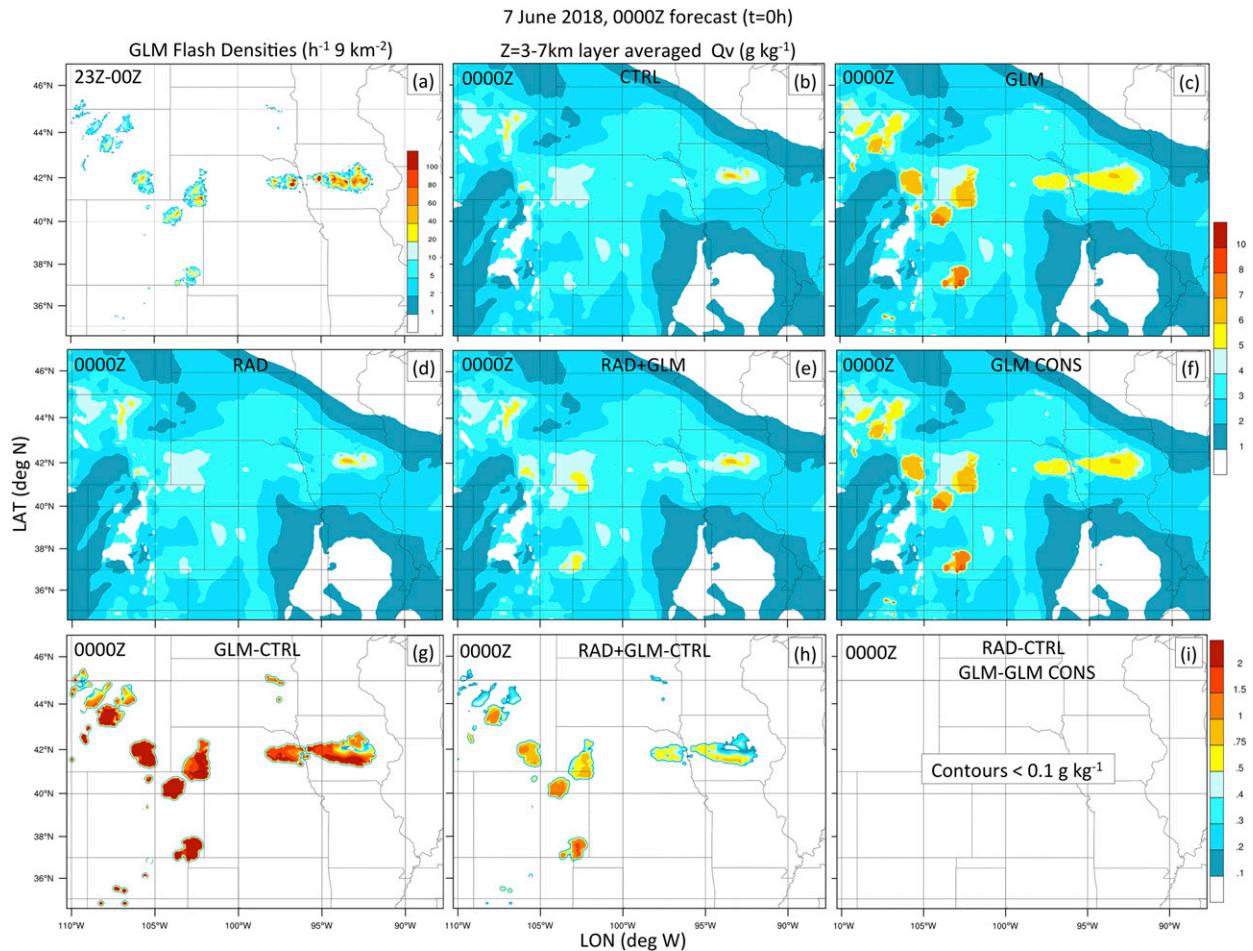
FIG. 4. (b)–(f) Horizontal cross sections at 0000 UTC 7 Jun 2018 (analysis time) of the z = 3–7 km MSL, layer-averaged q_v (g kg^{-1}) following the 3DVAR analysis for all the experiments listed and described in Table 1 with (a) showing, again, the GLM-derived flash densities of Fig. 3a, to better highlight and discern the locations where q_v adjustments were made. (g)–(i) Differences between the main DA experiments and CTRL, with (i) also showing the difference between the original GLM experiment and GLM CONS. Note that in this, and subsequent figures, simulation results for each experiment will be shown in the same order as listed in Table 1, for convenience.

exceeding $2\,\mathrm{g\,kg^{-1}}$ (Figs. 4g,h; maximum of about $4.5\,\mathrm{g\,kg^{-1}}$). The latent heating–related q_v adjustments in the radar-only DA experiment (RAD, Figs. 4d,i) are overall quite small (<0.1 g kg^{-1}, Fig. 4i). Figure 4i also highlights the nearly identical q_v fields produced by GLM and GLM CONS. Worth noting are the generally smaller differences in q_v fields between RAD + GLM and CTRL compared to between GLM and CTRL (Figs. 4g,h), which arise from the latent heating adjustments imposed by the variational assimilation scheme for radar reflectivity. Although the 3DVAR scheme adjusts the mass of particular hydrometeors within specific temperature ranges (Gao and Stensrud 2012), the condition of near saturation (RH = 95%) imposed by GLM is deemed too large by the latent-heating constraints of the radar reflectivity 3DVAR algorithm and, thus, is offset accordingly.

To provide a brief quantitative appreciation of some of the departures from observations in the initial conditions of CTRL, a skew T–log$p$ sounding analysis (Fig. 5) was performed at North Platte (LBF NWS site), western Nebraska, which was broadly representative of the prestorm environment for one of the main MCSs in this analysis (e.g., Figs. 1 and 3b). Soundings taken within a ~50-km distance from that point and away from active convection revealed overall very similar profiles and, hence, bulk kinematic/thermodynamic measures (not shown). Overall, the analyzed temperature and wind profile exhibits relatively good agreement with the observations (Figs. 5a,b). Over the area of interest (Nebraska and Iowa), however, the 2-m temperature fields from the RAP analysis and NAM analysis in CTRL exhibit spatial differences on the order of 1–2 K. Larger departures are seen where convection
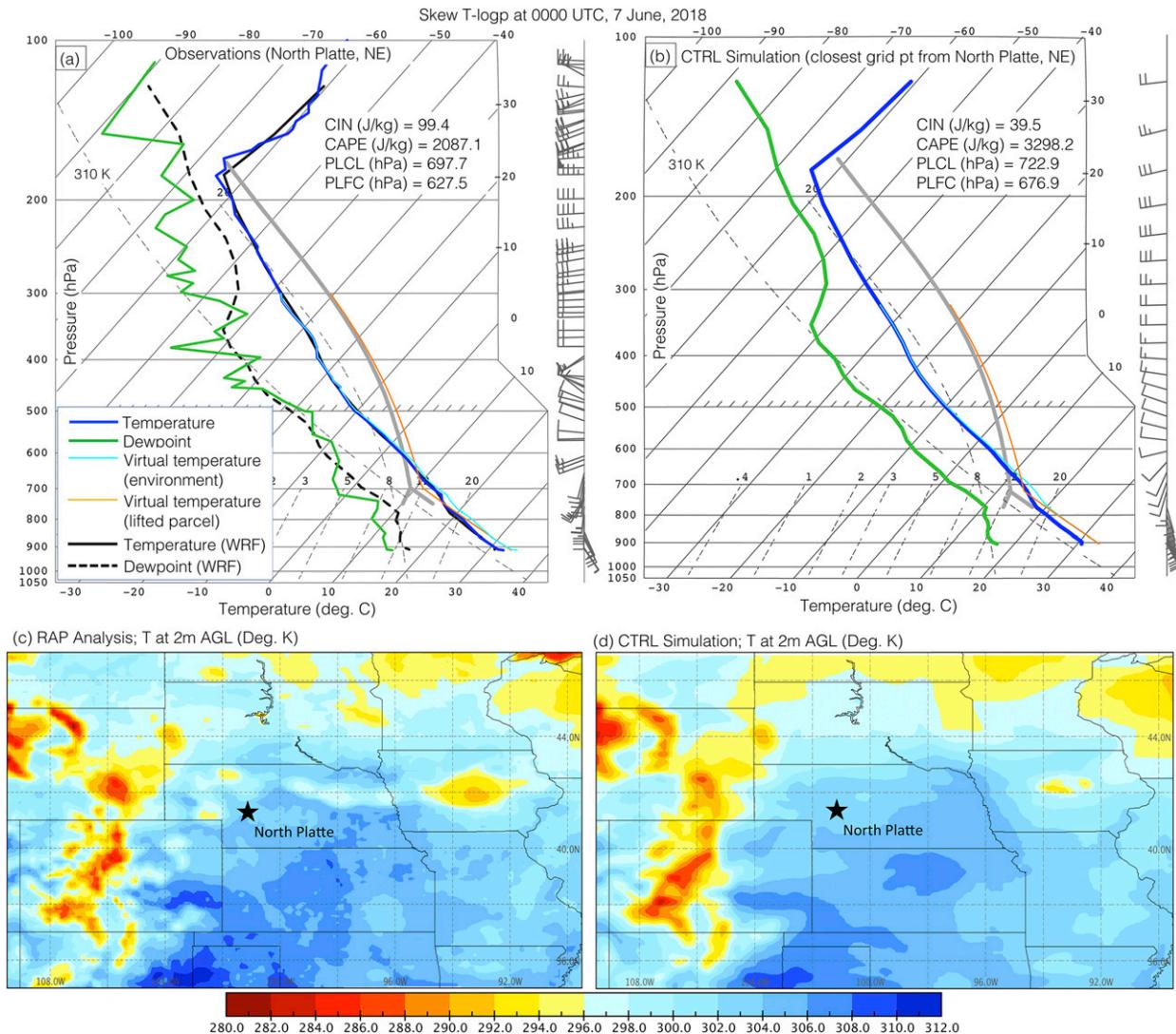
FIG. 5. Skew *T*–log*p* plots at 0000 UTC 7 Jun 2018 of (a) the North Platte, NE (LBF), NWS radiosonde site in comparison with (b) WRF forecast sounding in the closest grid column to LBF. The mixed-layer sounding parameters shown on the upper-right of each panel are computed through the lowest 90 mb, which characterizes the deepening convective boundary layer. The cyan and orange curves are the virtual temperature profiles of the environment and the lifted parcel, respectively (i.e., from which the integrated virtual buoyancy is derived to compute CAPE and CIN). Wind vectors are plotted with full barb = $5 \, \text{m s}^{-1}$, half barb = $2.5 \, \text{m s}^{-1}$, and filled triangle = $25 \, \text{m s}^{-1}$. To facilitate comparison, the WRF Model sounding temperature and dewpoint from (b) is overlaid in (a) as solid and dashed black curves, respectively. (c),(d) The 0000 UTC 2-m AGL temperature fields from the RAP analysis and the CTRL simulation, respectively. The location of the NWS operational sounding site at LBF is indicated by a black star, for reference. Sounding data for (a) are provided by the University of Wyoming public database (http://weather.uwyo.edu/upperair/sounding.html).

was analyzed (Figs. 3b and 5c) but not yet present in the model (Fig. 5d) along the nearly zonal stationary boundary across Nebraska and Iowa (Fig. 1). Figures 5a and 5b reveal, most importantly, relatively larger differences in dewpoint profile, with a generally moist bias of about $2 \, \text{g kg}^{-1}$ in the model between about LCL–700 hPa. Such departures from the observations for moisture remain particularly critical below cloud base, as highlighted by the subsequent

noteworthy differences in mixed layer CIN and CAPE between the model and the observations (Figs. 5a,b).

b. Forecast results: Single 3DVAR analysis experiments

Although CTRL is able to produce some of the observed storms already 1 h into the forecast (Figs. 6a,b), all GLM-based DA experiments help hasten the development of some of these storms further, such as the
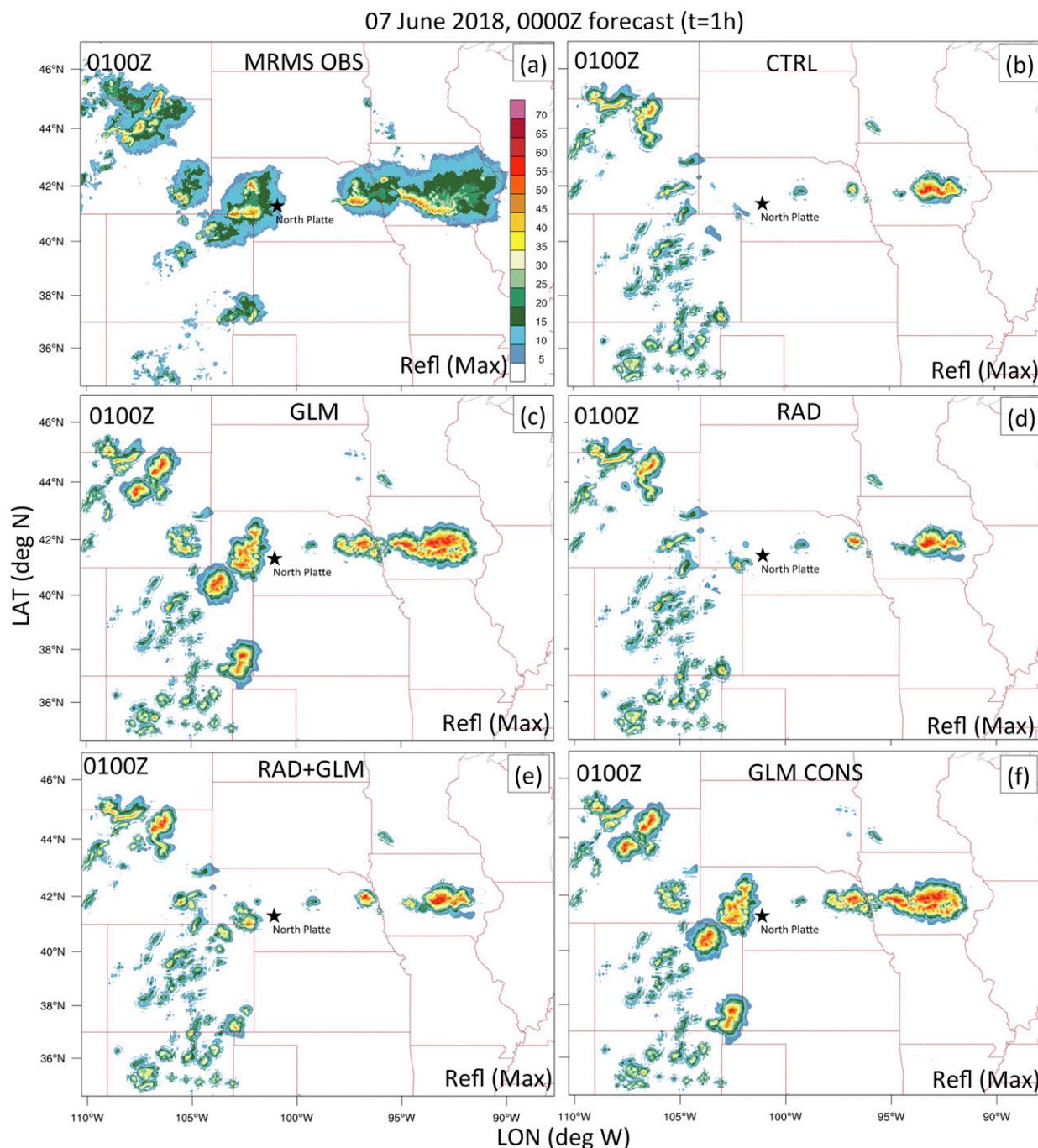
FIG. 6. Composite radar reflectivity fields (dBZ) at 0100 UTC 7 Jun 2018 (i.e., 1 h forecast) for: (a) the 1-km resolution, MRMS product (observations) interpolated onto the 3 km simulation domain, (b) CTRL, (c) GLM, (d) RAD, (e) RAD + GLM, and (f) GLM CONS experiments. Legends for color and shadings are shown in (a).

organized cluster of storms in central Iowa or some of the isolated cells in northern Wyoming (Fig. 6). In some cases, the GLM DA initiates observed storms that were still absent in CTRL at this point in time as evidenced by the presence of radar echoes exceeding 40 dBZ in eastern and western central Nebraska, north- and

southeastern Colorado (Fig. 6). This case also highlights the inability of RAD and/or GLM to effectively suppress the weaker, isolated spurious convection over Colorado and northern New Mexico (Fig. 6a).

Composite reflectivity fields six hours into the forecast (Fig. 7) underscore (i) how the more timely
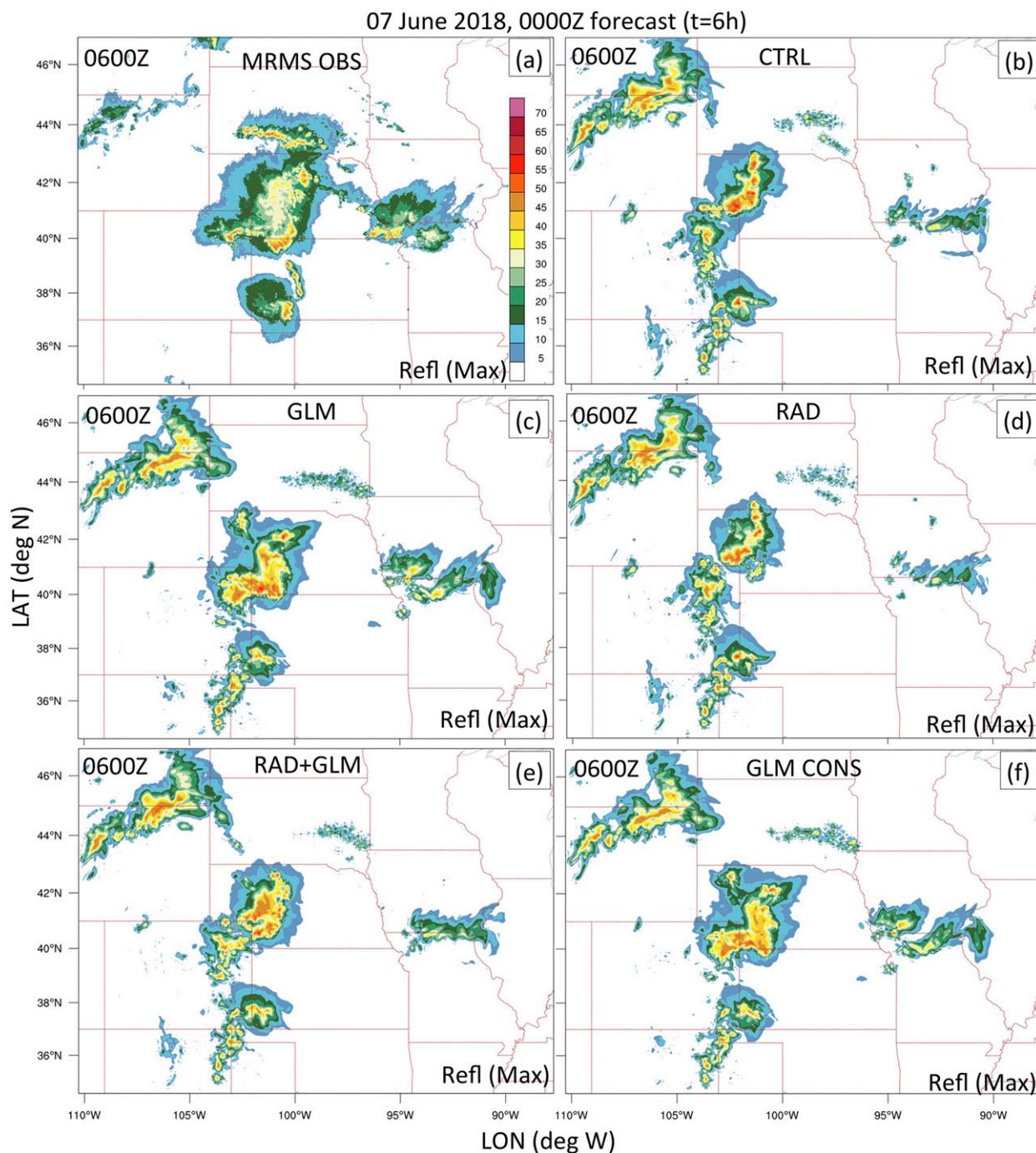
FIG. 7. As in Fig. 6, but at 0600 UTC 7 Jun 2018 (i.e., 6-h forecast).

development of the observed storms in the GLM-based experiment helps improve the placement and intensity of the main observed MCS in western Nebraska, and (ii) how some of the cells that were properly initiated in all simulations (including CTRL) in northern Wyoming (Fig. 7) are not decaying 6 h into the forecast. None of the simulations is able to reproduce the smaller, linear cluster of

storms in southern South Dakota (Fig. 7a), which developed at about 0400 UTC in the observations (not shown) and had no radar echoes or lightning at the 0000 UTC initialization time. The erroneous storms in eastern Colorado and northeastern New Mexico in CTRL (versus decay in the observations; Figs. 7a,b) are notable in all experiments (Figs. 7c–f) as they are not affected by the DA.
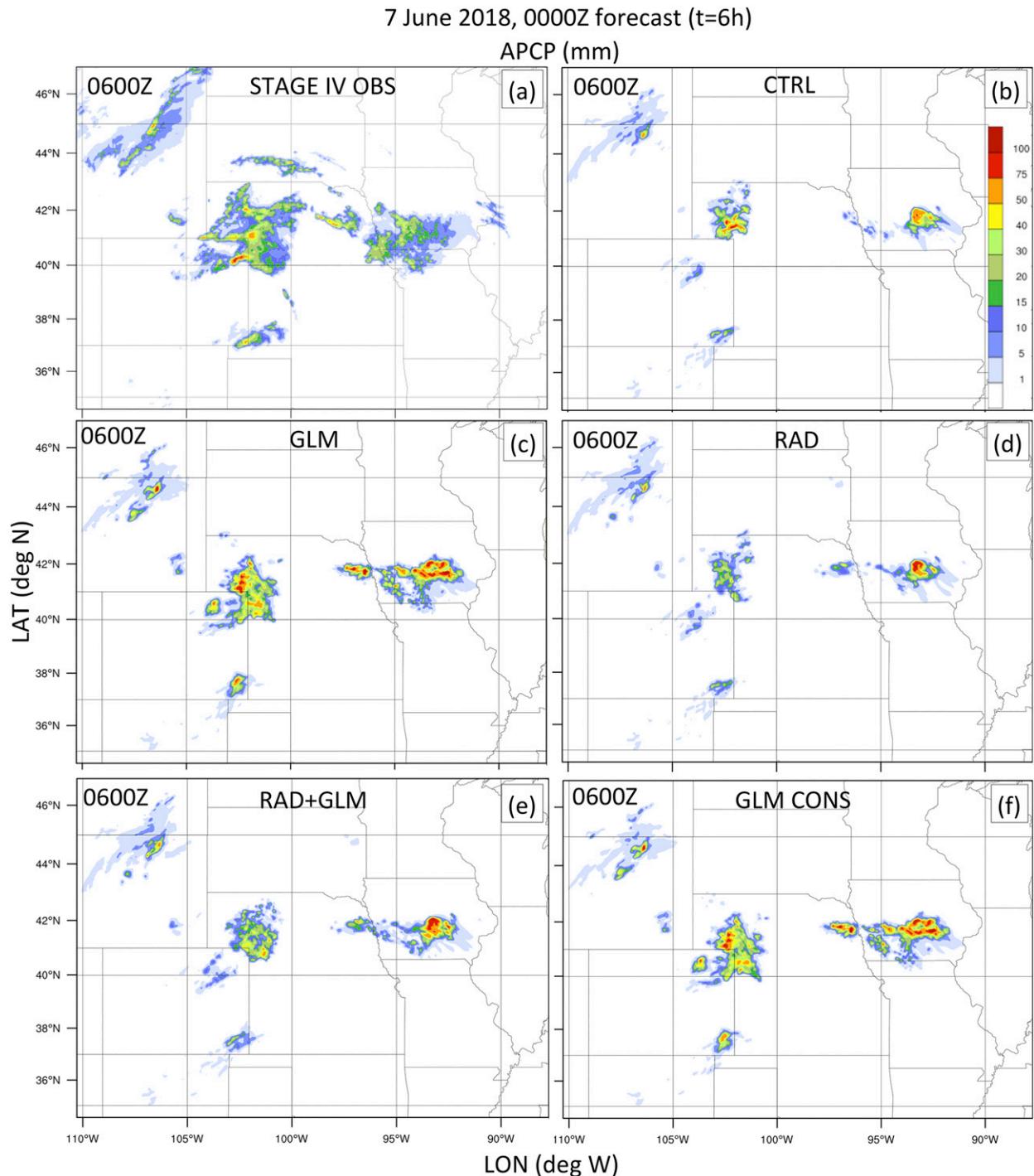
FIG. 8. As in Fig. 6, but for the 6-h accumulated precipitation fields (APCP in mm) with the observations in (a) derived from the stage IV multisensor rainfall estimates.

To provide a more complete view of the forecast performance, the accumulated precipitation (APCP) fields were evaluated in each simulation. Figure 8 generally highlights that although the DA experiments are able to improve the areal coverage of 6-hourly APCP < 40 mm,

they also generally overpredict the coverage for amounts exceeding 50 mm. This issue persisted, and was even exacerbated, when other microphysical schemes were tested such as the Hong and Lim (2006) single-moment 6-class scheme (WSM6) or the Thompson et al. (2004)
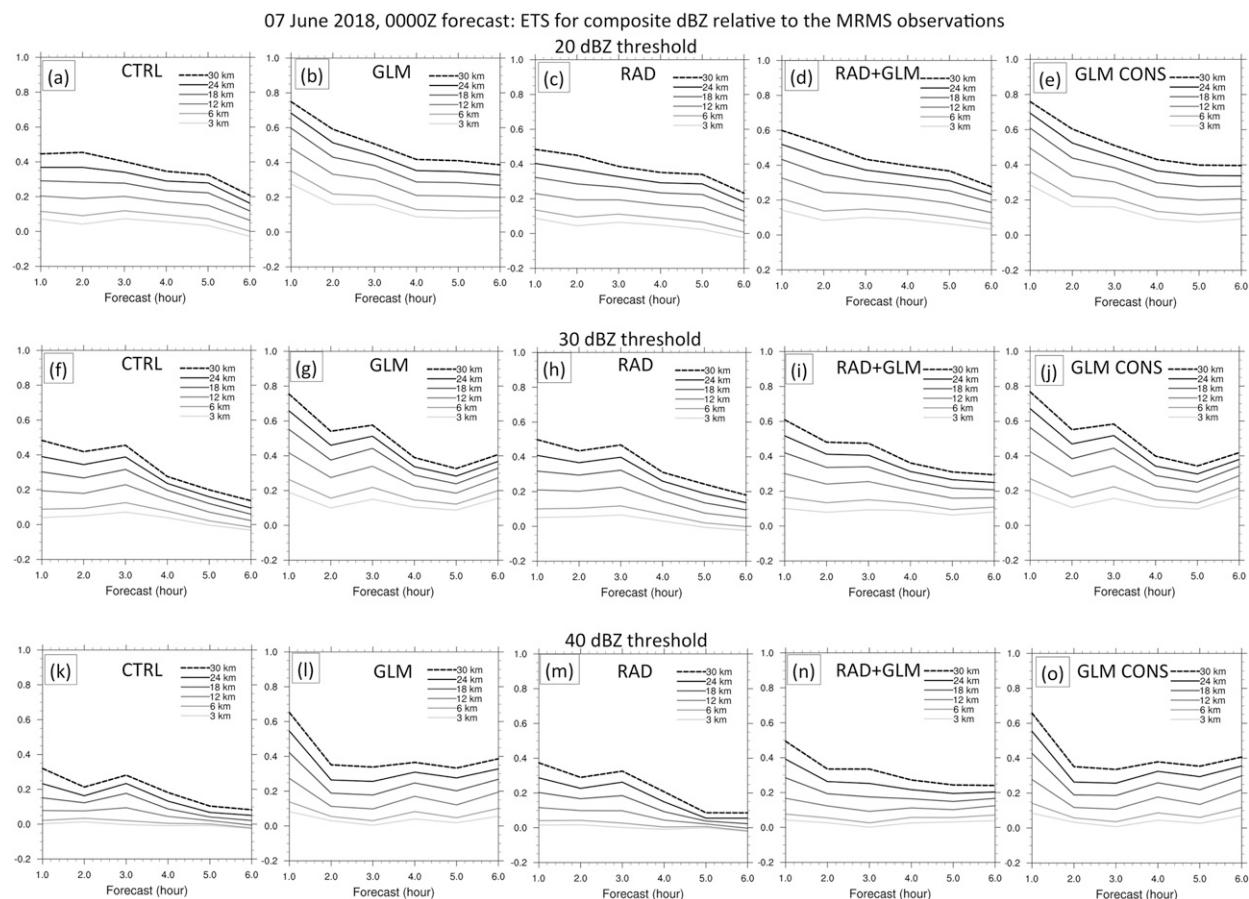
FIG. 9. Equitable threat scores (ETSs) of the simulated composite radar reflectivity fields relative to the MRMS observations over the simulation domain in Fig. 2 starting from 0100 UTC (1-h forecast) and ending at 0600 UTC 7 Jun (6-h forecast). Scores at the analysis time (0000 UTC) are not shown given that all simulations either have zero or very small reflectivity values. ETSs are shown for five different neighborhood radii (legend shown in each panel) and for three specific (composite) reflectivity thresholds indicated on each row, namely: (a)–(e) 20 dBZ, (f)–(j) 30 dBZ, and (k)–(o) 40 dBZ. For each of these thresholds, ETSs are presented from left to right for all five experiments in the order listed in Table 1, for convenience.

scheme (not shown). This is in line with F15, which underscored that any wet biases present in a simulation assimilating no data could only be exacerbated by DA schemes promoting further development of convection, especially when spurious storms are not addressed.

c. Overall forecast evaluation

To quantify the performance of composite reflectivity and APCP forecasts in each experiment, the remainder of this analysis will focus on discussing results for two neighborhood-based scores: (i) equitable threat score (ETS; Clark et al. 2010) and (ii) fractions skill score (FSS; Roberts and Lean 2008) relative to the stage IV observations and relative to the MRMS composite reflectively fields. Both score metrics were computed for neighborhood radii ranging between 3 and 30 km (i.e., grid point based up to 10 grid points) for each of the following

thresholds: 20, 30, and 40 dBZ for composite reflectivity and 1, 5, and 10 mm for APCP. Scores for APCP are computed for hourly intervals (i.e., mm h^{-1}) and scores for composite reflectivity use the instantaneous values. To complement the scores and provide an additional measure of over or underprediction, the frequency biases (referred to as bias for brevity) were also computed. No bias correction for either score metrics was applied, given their overall negligible impact (F15), especially when biases are overall low (i.e., <2.5; see later in the section). Although the ETS has been widely used as a metric to quantify the skill of a forecast for a given quantity, several studies have posited that FSS is an attractive and, arguably, more accurate measure of skill on convection-allowing grids (Mittermaier et al. 2013), because, in contrast to the ETS, FSS provides a quantitative measure of goodness and usefulness, which adds to the understanding of true forecast skill.
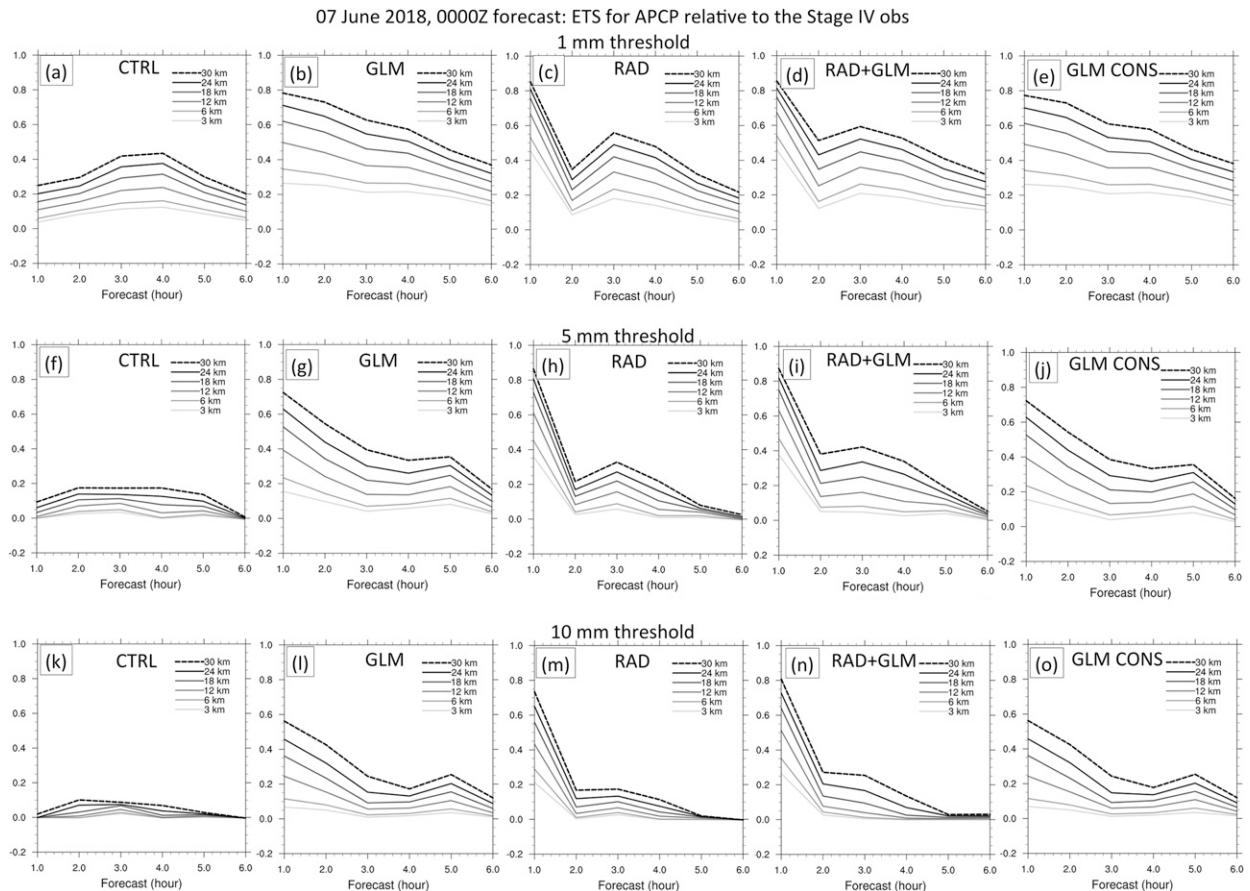
FIG. 10. As in Fig. 9, but for hourly accumulated precipitation (APCP in mm) computed for the following three thresholds: (a)–(e) 1 mm, (f)–(j) 5 mm, and (k)–(o) 10 mm.

In terms of composite radar reflectivity fields, the GLM-based experiments produce generally larger ETSs than either RAD and CTRL, especially at the 20- and 30-dBZ thresholds (Fig. 9). Given the stark similarity in the analyzed q_v fields in Fig. 4 and forecast composite reflectivity fields (Figs. 6 and 7), both GLM and GLM CONS produce nearly identical ETSs at all thresholds; which is encouraging as this confirms that imposing q_v conservation prior to applying the 3DVAR analysis does not degrade nor impact any existing gain in forecast performance. As the reflectivity threshold is increased from 30 to 40 dBZ (Figs. 9k–o) or even 50 dBZ (not shown), the ETSs decrease owing to the progressively smaller areal coverage of larger reflectivity areas. The relatively reasonable performance of CTRL in forecasting composite reflectivity fields at 1 h is evidenced by ETSs on the order of 0.4–0.5 at larger radii for the 20 and 30 dBZ thresholds, yielding to ETS differences between any of the DA experiments and CTRL generally not exceeding 0.2 (not shown).

For composite reflectivity, generally similar trends and behavior are obtained for the FSS (not shown). For the GLM-based experiments, the FSS for the lowest threshold (20 dBZ) is characterized by larger values than the ETS and, conversely, by notably lower values as the dBZ threshold is increased. Despite this, the overall differences in FSS scores relative to CTRL are very similar to those with the ETS at all thresholds (not shown).

In terms of APCP, the ETSs for the GLM-only experiments still hint at a generally similar improvement relative to CTRL (Fig. 10). A stark difference with composite reflectivity, however, is seen for the RAD-based experiments, which produce superior scores during the first hour of forecast with maximum values nearing 0.8 at all thresholds (Fig. 10). The largest improvements relative to CTRL are seen for RAD + GLM at 10 mm (Figs. 10n,k). Given notably lower ETSs in CTRL for APCP compared to composite reflectivity (cf. Figs. 9a,f,k and 10a,f,k), the ETS differences relative to CTRL are more pronounced for APCP with
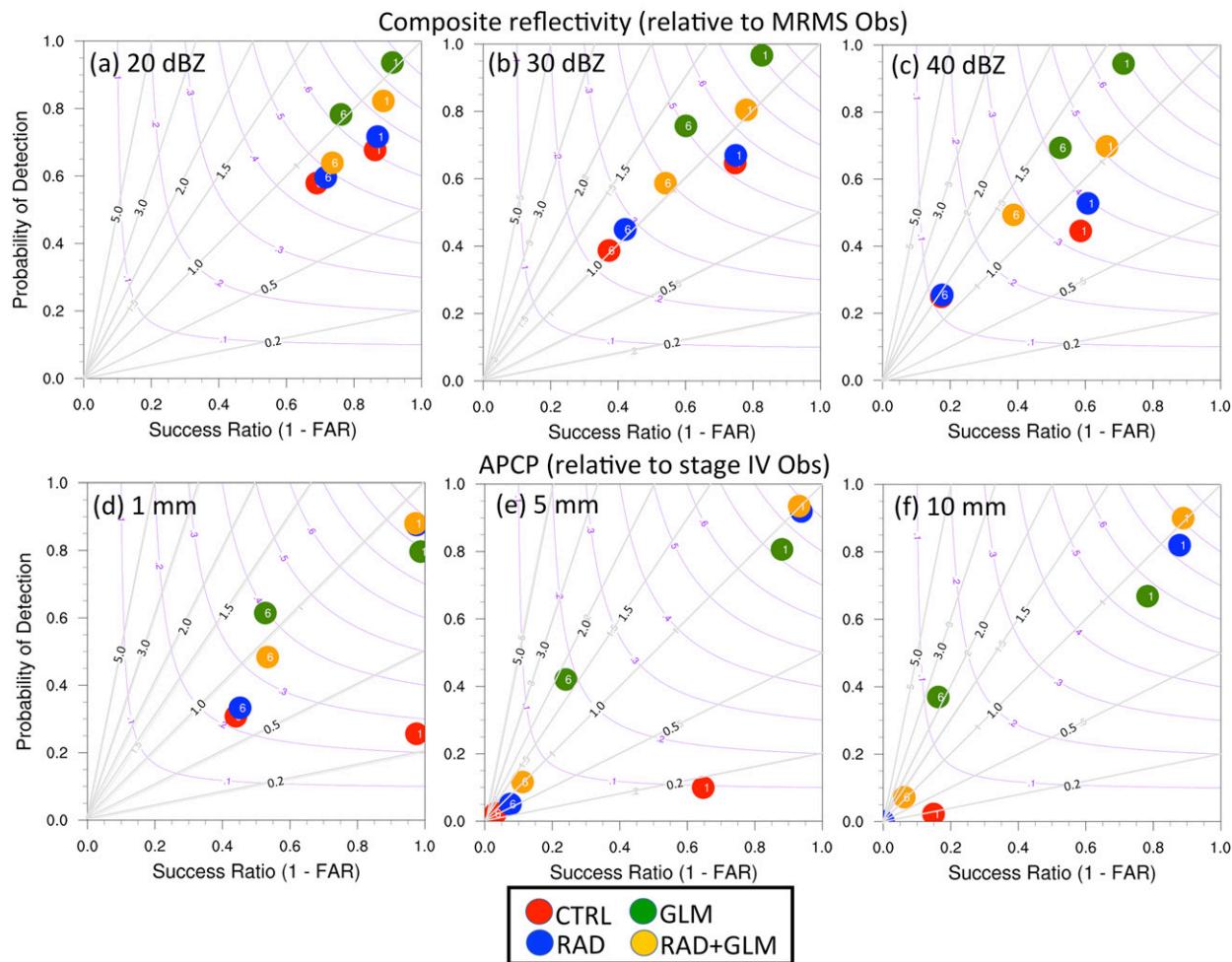
FIG. 11. Performance diagrams at 1- and 6-h forecast assuming a neighborhood radius of 30 km (10 grid points) for (a)–(c) composite reflectivity fields relative to the MRMS observations using the same thresholds as in Fig. 9, namely 20, 30, and 40 dBZ, respectively. (d)–(f) As in (a)–(c), but for accumulated precipitation, again for the same thresholds as in the previous figures: namely 1, 5, and 10 mm, respectively. In each plot, the lower-left corner stands for no forecast skill and, similarly, the upper-right corner indicates perfect skill. Purple curves represent the critical success index (CSI), and the diagonal gray lines the frequency bias. The colored dots show the results for the main experiments with legends shown at the bottom of the figure. The number inside each dot represents the forecast time in hour.

maximum values ranging between 0.6 and 0.8 for the RAD-based experiments during the first hour of forecast compared to values below 0.2 for composite reflectivity (not shown). Note again, however, how any forecast improvements in the very early portion of the forecast (≤1.5 h) are quickly lost later on (Fig. 10). Results for the FSSs gave, overall, a very similar impression (not shown).

To provide an alternative and more concise interpretation of the overall performance of the short-term forecasts for both rainfall and composite reflectivity, categorical performance diagrams (Roebber 2009) were computed, which conveniently merge the information

derived from key contingency table elements, namely: frequency bias, probability of detection (POD), critical success index, and success ratio (one minus the false alarm rate). The diagrams for composite reflectivity emphasize how, at 1-h forecast, GLM produces the best overall performance as measured by higher success ratio and POD (Figs. 11a–c), and, similarly, highlight the clearly superior performance of RAD at 1 h for rainfall (Figs. 11d–f). These diagrams also underscore concisely how any gains in forecast skill for RAD at 1-h for either rainfall or reflectivity are substantially reduced at 6 h. Note also how, in contrast to F15's nudging method, GLM is able to improve the rainfall and reflectivity

forecast without significantly affecting the bias relative to CTRL (bias remains below ~2.5, Figs. 11a–c). One potential explanation for this modest increase in bias in the GLM-based experiments could be the usage of a full hour of lightning rather than a shorter accumulation (e.g., 15–20 min). For instance, the reflectivity areas ($\geq 40\,\mathrm{dB}Z$) of the storm in northeast Colorado at 0000 UTC appears larger-than-observed (Figs. 6a,c), because the lightning area (Fig. 3a) is wider than the $\geq 40\,\mathrm{dB}Z$ reflectivity areas at 0000 UTC (Fig. 3b). At 1 h, RAD + GLM produces the best results by bringing the bias closer to 1, which is partially due to the smaller q_v adjustments relative to GLM during the 3DVAR analysis (F16; Fig. 4).

d. 3DVAR analysis experiments with successive cycling

1) FORECAST RESULTS SUMMARY

The short-term ($\leq 3\,\mathrm{h}$) forecast of reflectivity and accumulated rainfall exhibit overall similar behavior as the original DA experiments (e.g., compare Figs. 12 and 6). In terms of score, the cycled GLM experiment performs best for both reflectivity and accumulated rainfall at 1 and 6-h forecast (Fig. 13). In general, all cycled DA experiments exhibited better skill than their original counterparts (cf. Figs. 11 and 13). For the GLM cycling, one notable difference is a tendency to overestimate moderate-to-large reflectivities ($>30\,\mathrm{dB}Z$), as evidenced in Fig. 12 by their larger areal coverage and in Fig. 13 with the green dots systematically locating closer to the upper-left corner of the plot (i.e., progressively higher bias, ranging between 1.5 and 2 at $40\,\mathrm{dB}Z$). A closer inspection of the data for each (of the six) individual 3DVAR cycles revealed that the overestimation of $>30\,\mathrm{dB}Z$ reflectivities in the GLM cycling experiment can be traced back to the fact that the areal coverage of nonzero flash densities (per 10 min herein) does not differ significantly from the respective coverage of the hourly rates in Fig. 3a (not shown). Thus, at each cycle, the LDA adjusts (i.e., increases) water vapor mass over an area similar in size to that of the original GLM experiment, yielding progressively stronger convection. Future research with DA cycling is planned to address this drawback.

2) DA METRICS

To better assess the quality and performance of the DA during the cycled 3DVAR procedure, the remainder of this analysis will focus on examining key statistics including the variation of total cost function, root-mean-square (rms), and means (bias) for the innovation and analysis residuals during the DA cycles (as defined in e.g., Daley 1993; Lindskog et al. 2004).

The cost function is defined by the sum of weighted Euclidian distances between the analysis and the background field, and between the analysis and the observations. The cost function used for this study is defined by Eq. (1) in Gao et al. (2013). During the minimization process, the cost function usually decreases with iterations.

In simple terms, the rms innovation statistics (referred to as rms_innov) provide a measure of the Euclidian distance between the observations and the background fields:

$$\mathrm{rms\_innov} = \sqrt[2]{\frac{1}{N}\sum_{i=1}^{N}[y_i^o - H(x_i^b)]^2}, \quad (1)$$

where N is the number of observations, $H(x)$ is a continuous function mapping the background fields (superscript b) from model space to observation space (superscript o). Similarly, the rms analysis residual (referred to as rms_res) is a measure of the Euclidian distance between the observations and the analysis:

$$\mathrm{rms\_res} = \sqrt[2]{\frac{1}{N}\sum_{i=1}^{N}[y_i^o - H(x_i^a)]^2}, \quad (2)$$

with the subscript a standing for analysis.

In a similar manner, it is possible to determine if the DA procedure generates any biases for each control variables by analyzing the mean innovation (m_innov):

$$\mathrm{m\_innov} = \frac{1}{N}\sum_{i=1}^{N}[y_i^o - H(x_i^b)] \quad (3)$$

and mean analysis residual (m_res):

$$\mathrm{m\_res} = \frac{1}{N}\sum_{i=1}^{N}[y_i^o - H(x_i^a)]. \quad (4)$$

Because these statistics were qualitatively similar for both 3DVAR passes and because convective scales are emphasized herein, only the results for second 3DVAR pass are discussed.

Generally speaking, all DA experiments show a reduction in the total cost function with the GLM experiment producing the largest relative decrease (exceeding 80%) at the end of the 30 iterations (Fig. 14a) compared to about a 20%–30% decrease for the RAD-based experiments. Given the notably larger amount of observations being assimilated in the RAD-based experiments (i.e., volumetric scans versus 2D fields for the GLM), and nonlinearity of reflectivity observations, the cost function reduction rate is relatively small relative to the GLM experiment (Figs. 14b,c).

The rms_innov for the GLM experiment generally decrease with the DA cycles (Fig. 14d). This indicates the model gradually absorbs the q_v information from assimilating GLM data reasonably well. For the
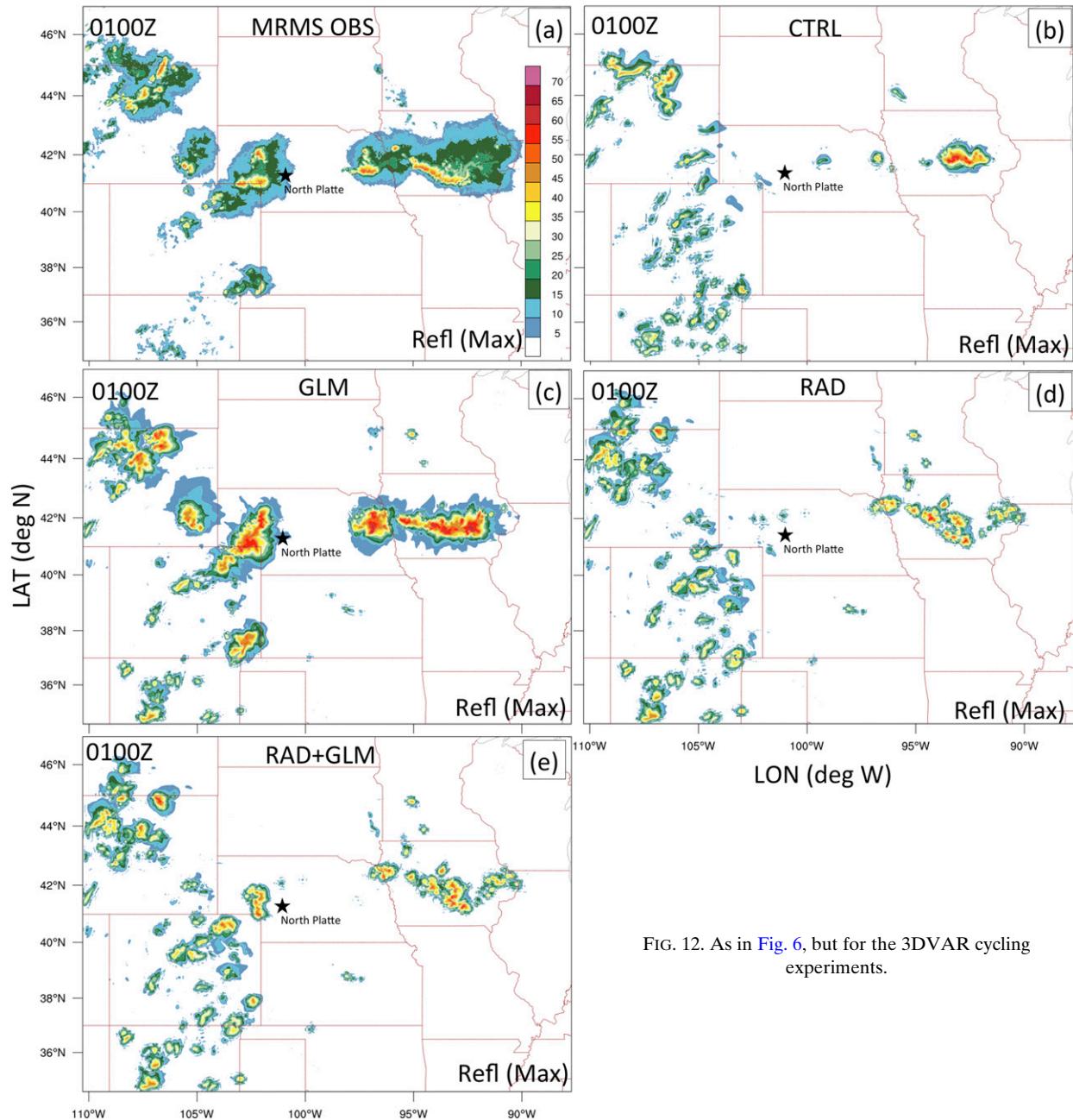
FIG. 12. As in Fig. 6, but for the 3DVAR cycling experiments.

RAD-based experiments, rms_innov for the GLM-based pseudo q_v and radial velocity data slightly decrease with the DA cycles indicating that GLM (pseudo q_v) and radial velocity information are properly assimilated (Figs. 14e,f). The gradual increases of rms_innov for reflectivity data (Figs. 14e,f), however, reflect the model's difficulty in properly absorbing precipitation-related information from reflectivity data (by adding too much reflectivity). In addition to model forecast error,

one likely factor is that reflectivity factor exhibits a complex, nonlinear relationship with the model's hydrometer variables. Additionally, the gradual increase in storm coverage with each successive 3DVAR cycle may also contribute the increase of rms_innov for reflectivity.

The rms_res statistic (Figs. 14d–f) represents how close the analyses are to the observations at the end of the 3DVAR minimization procedure. All values for rms_res are smaller than the corresponding values

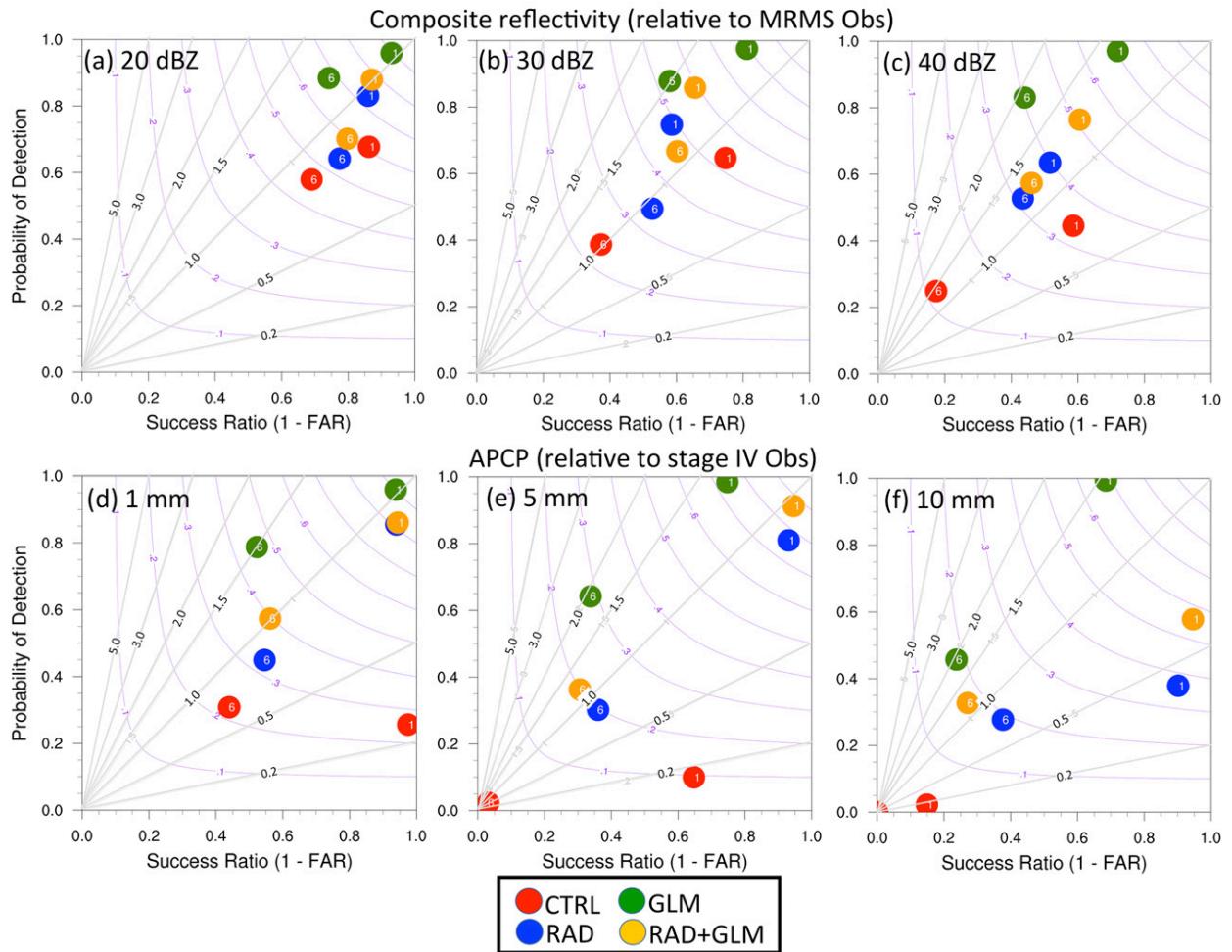CYCLING Experiments. Performance Diagrams for R=30 km: 1-h and 6-h forecast

FIG. 13. As in Fig. 11, but for the 3DVAR cycling experiments.

for rms_innov indicating that the analysis is closer to the observations for all types of observations after each DA cycles (Figs. 14d–f). Again, for GLM and radial velocity data, the values for rms_res do not change too much during the assimilation cycles indicating that the analysis generally fits these data reasonably well. While the assimilation of reflectivity is generally reasonable, the fact that rms_res increases still indicates a growing difficulty for the analysis to fit reflectivity observations. As discussed above for rms_innov, the nonlinearity of the cost function for reflectivity, in addition to phase errors and the development of spurious convection may play a role. Addressing these issues, will be deferred to future work more focused on radar data retrievals.

The innovation and analysis residual means or biases (m_innov and m_res) for q_v are generally positive (Figs. 14g,i), indicating that, overall, the background and analyzed values for q_v are lesser than the target

(near saturation) pseudo q_v observations; especially for the RAD-based experiments (in line with Fig. 4). Both m_innov and m_res for radial velocity are nearly identical and close to zero (Figs. 14h,i) consistent with the rms statistics in Figs. 14e and 14f indicating that, overall, the analyzed radial velocity field is getting closer to the observations. For reflectivity, the innovations exhibit a positive bias (~ 2 dBZ) while, in contrast, the residuals are negatively biased (~ -1 to -2 dBZ, Figs. 14h,i). This further illustrates the growing difficulty of the DA to optimally ingest, for this case, the wealth of 3D reflectivity information into the system.

6. Summary and future work

This study evaluated the adaptation of a cloud-scale LDA method in the 3DVAR variational framework for total lightning densities derived from the spaceborne
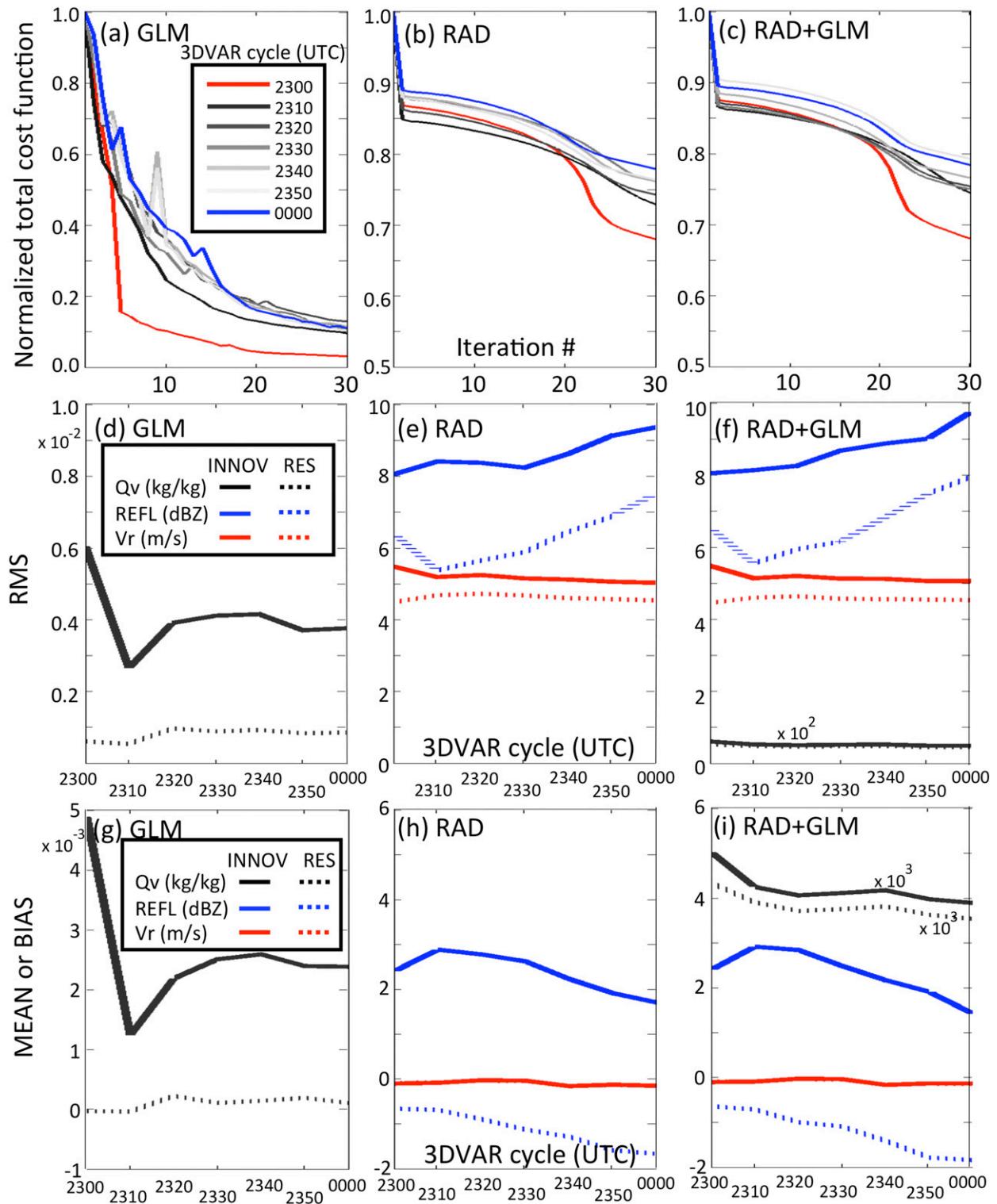
FIG. 14. (a)–(c) Normalized total cost function as a function of iteration number for the (a) GLM, (b) RAD, and (c) RAD + GLM experiments. The cost function curves are color-coded for each 3DVAR cycles following the legends shown in (a). (d)–(f) Root-mean-square statistics for the innovation (solid curve) and analysis residual (dotted curve) for water vapor mass (q_v in kg kg^{-1}, black), reflectivity factor (REFL in dBZ, blue), and radial velocity (V_r in m s^{-1}, red) as a function of 3DVAR cycle for the (d) GLM, (e) RAD, and (f) RAD + GLM experiments. (g)–(i) As in (d)–(f), but for the bias (or mean) statistics.

GLM instrument. To gauge the potential added value of the assimilation of lightning data, companion experiments assimilating no data (CTRL) or conventional NWS Level II radar data (radial velocity and radar reflectivity) with or without lightning were conducted. In addition to adapting the LDA method of F16 to GLM data, several refinements to the lightning and radar DA procedures were examined: (i) the GLM assimilation features an option to compensate the water vapor mass added at lightning locations with equal removal outside lightning areas, (ii) regional and seasonal dependence of the scheme (implicit in the conditionals for LCL) were partially alleviated, (iii) in contrast to F16, no cloud analysis scheme was used to assimilate radar reflectivity (i.e., all the observations herein were assimilated in a unified 3DVAR framework), and (iv) no arbitrary positive cut off thresholds for the flash rate are employed to identify convectively active areas.

During the course of 2018, several experiments were performed in quasi–real time over regional domains systematically positioned over the eastern two-thirds of the contiguous United States. For the sake of brevity and simplicity, this study reports the salient findings based on one illustrative case study in which the DA of either radar or GLM data improved the forecast in some areas of preexisting convection but not in other regions where convection had not yet developed. Last, this case was selected to underscore some salient limitations of the LDA scheme such as, for instance, its inability to effectively suppress spurious convection.

Overall, assimilation experiments including either radar or GLM lightning data yielded short-term forecast improvements that were similar to those reported in F15 and F16. Akin to these and other studies, any gains in composite reflectivity or rainfall forecast skill systematically seen during the first 1–1.5 h are gradually lost, and improvements over CTRL are negligible after about 3-h. This result held true for both the neighborhood-based equitable threat and fractions skill scores. As indicated in F15, this is primarily caused by the growth of both model and initial condition errors that limit predictability.

An analysis focused on examining key 3DVAR statistics that included the variation of the total cost function, root-mean-square and means (bias) for the innovation and analysis residuals during the DA cycles revealed that, overall, the radial wind and proxy variable used for the GLM lightning (i.e., q_v) were well assimilated by the system. For radar reflectivity factor, however, the nonlinearity of the cost function for reflectivity, coupled with potential phase errors and the development of spurious convection resulted in a slightly less effective assimilation for this case. Subject to ongoing

work, real-time regional forecasts conducted during spring of 2019 revealed an overall improvement of this drawback.

As mentioned above, one of the critical limitations of the LDA scheme examined here lies in its inability to suppress spurious convection. Preliminary research is underway, however, to consider this capability. Reflectivity data in both the observations and background fields will be employed to target more accurately areas where convection must be suppressed. For instance, this could be achieved by imposing, a negative adjustment in q_v (or other proxy variables for convection) at locations in the background with reflectivity areas representative of convective rainfall (e.g., $>30\,dBZ$) not supported by the observations. Where radar data are nonexistent or sparse, the suppression method could alternatively exploit spaceborne datasets able to reveal the presence of convection such as brightness temperature or specific spectral bands of all-sky radiance data [e.g., Advanced Baseline Imager instrument on board *GOES-16/17*; Goodman et al. (2013)]. Following the strategy of the NSSL Experimental Warn-on-Forecast System for ensembles (NEWS-e; Jones et al. 2016), work is also underway to implement and gauge the impact of the above DA procedure using high frequency (10–15 min) successive 3DVAR cycling. To better gauge the added value of the GLM, these tests will also focus more systematically on convection evolving over data sparse areas, such as the mountainous terrain of the western United States or the vast regions offshore such as the Gulf of Mexico.

Future work will also be devoted to examine the performance of this combined radar and lightning DA procedure in a hybrid 3DVAR and EnKF system, or ensemble of 3DVAR framework (Wang et al. 2013; Gao et al. 2016) to leverage the advantages of ensemble data assimilation and 3DVAR approaches. Another research route worth envisaging is incorporating ensemble-derived inhomogeneous background error covariance statistics that treat precipitating and nonprecipitating areas separately (Michel et al. 2011). This is critical given that the background error covariance matrix plays an important role in variational data assimilation for spreading increments along the vertical and horizontal and for achieving balance among cross-correlated fields.

The areal coverage of nonzero total lightning densities over sufficiently large regional domains is usually small, especially when compared to radar data (e.g., Fig. 3) or other spaceborne datasets (e.g., brightness temperature). The 2018 near-real-time tests revealed one noteworthy limitation for assimilating lightning data: for larger convective systems characterized by warm rain

microphysics (e.g., tropical MCSs, including hurricanes), the aggregate impact of adjusting convective-scale proxy variables for lightning (such as q_v) near or within lightning areas is generally small. For tropical cyclones [e.g., Hurricane Michael (2018)], assimilating lightning data over regions devoid of radar data was found to have negligible impacts on track and intensity prediction. On some occasions, however, the LDA alone was able to better forecast the placement and intensity of lightning-active outer rainbands (consistent with Fierro and Reisner 2011; Fierro et al. 2015b), which often are associated with hazardous weather at landfall such as quick tornado spin ups or localized flash flooding. Thus, future research should combine lightning information with auxiliary data sources, which are able to better characterize the nature of clouds and convection beyond the more confined, electrically active areas detected by the GLM.

REFERENCES

Aksoy, A., 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.

Albers, S. C., J. A. McGinley, D. L. Birkenheuer, and J. R. Smart, 1996: The Local Analysis and Prediction System (LAPS): Analyses of clouds, precipitation, and temperature. *Wea. Forecasting*, **11**, 273–287, https://doi.org/10.1175/1520-0434(1996)011<0273:TLAAPS>2.0.CO;2.

Albrecht, R., S. Goodman, D. Buechler, R. Blakeslee, and H. Christian, 2016: Where are the lightning hotspots on Earth? *Bull. Amer. Meteor. Soc.*, **97**, 2051–2068, https://doi.org/10.1175/BAMS-D-14-00193.1.

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

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.

Baldwin, M. E., and K. E. Mitchell, 1997: The NCEP hourly multisensor U.S. precipitation analysis for operations and GCIP research. Preprints, *13th Conf. on Hydrology*, Long Beach, CA, Amer. Meteor. Soc., 54–55.

Benjamin, S. G., and Coauthors, 2004: An hourly assimilation–forecast cycle: The RUC. *Mon. Wea. Rev.*, **132**, 495–518, https://doi.org/10.1175/1520-0493(2004)132<0495:AHACTR>2.0.CO;2.

Boccippio, D. J., K. L. Cummins, H. J. Christian, and S. J. Goodman, 2001: Combined satellite- and surface-based estimation of the intracloud–cloud-to-ground lightning ratio over the continental United States. *Mon. Wea. Rev.*, **129**, 108–122, https://doi.org/10.1175/1520-0493(2001)129<0108:CSASBE>2.0.CO;2.

Braun, S. A., 2002: A cloud-resolving simulation of Hurricane Bob (1991): Storm structure and eyewall buoyancy. *Mon. Wea. Rev.*, **130**, 1573–1592, https://doi.org/10.1175/1520-0493(2002)130<1573:ACRSOH>2.0.CO;2.

Brewster, K., 2002: Recent advances in the diabatic initialization of a non-hydrostatic numerical model. Preprints, *19th Conf. on Weather Analysis and Forecasting/15th Conf. on Numerical Weather Prediction/21st Conf. on Severe Local Storms*, San Antonio, TX, Amer. Meteor. Soc., J6.3, https://ams.confex.com/ams/SLS_WAF_NWP/techprogram/paper_47414.htm.

Carey, L. D., and S. A. Rutledge, 1998: Electrical and multiparameter radar observations of a severe hailstorm. *J. Geophys. Res.*, **103**, 13 979–14 000, https://doi.org/10.1029/97JD02626.

Caumont, O., V. Ducrocq, É. Wattrelot, G. Jaubert, and S. Pradier-Vabre, 2010: 1D+3DVar assimilation of radar reflectivity data: A proof of concept. *Tellus*, **62A**, 173–187, https://doi.org/10.1111/j.1600-0870.2009.00430.x.

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

Chen, F., Z. Janjić, and K. Mitchell, 1997: Impact of atmospheric surface-layer parameterizations in the new land-surface scheme of the NCEP mesoscale Eta model. *Bound.-Layer Meteor.*, **85**, 391–421, https://doi.org/10.1023/A:1000531001463.

Christian, H. J., and Coauthors, 1999: The Lightning Imaging Sensor. *Proc. 11th Int. Conf. on Atmospheric Electricity*, Guntersville, AL, International Commission on Atmospheric Electricity, 746–749.

Clark, A. J., M. Xue, and M. L. Weisman, 2010: Neighborhood-based verification of precipitation forecasts from convection-allowing NCAR WRF Model simulations and the operational NAM. *Wea. Forecasting*, **25**, 1495–1509, https://doi.org/10.1175/2010WAF2222404.1.

Daley, R., 1993: *Atmosphere Data Analysis.* Cambridge University Press, 471 pp.

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

Dudhia, J., 1989: Numerical study of convection observed during the winter monsoon experiment using a mesoscale two-dimensional

model. *J. Atmos. Sci.*, **46**, 3077–3107, https://doi.org/10.1175/1520-0469(1989)046<3077:NSOCOD>2.0.CO;2.

Ek, M. B., K. E. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayno, and J. D. Tarpley, 2003: Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model. *J. Geophys. Res.*, **108**, 8851, https://doi.org/10.1029/2002JD003296.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99**, 10 143–10 162, https://doi.org/10.1029/94JC00572.

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

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, https://doi.org/10.1175/2010JAS3659.1.

——, M. S. Gilmore, E. R. Mansell, L. J. Wicker, and J. M. Straka, 2006: Electrification and lightning in an idealized boundary-crossing supercell simulation of 2 June 1995. *Mon. Wea. Rev.*, **134**, 3149–3172, https://doi.org/10.1175/MWR3231.1.

——, E. Mansell, C. Ziegler, and D. MacGorman, 2012a: 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.

——, E. J. Zipser, M. A. LeMone, J. M. Straka, and J. Simpson, 2012b: Tropical oceanic hot towers: Need they be undilute to transport energy from the boundary layer to the upper troposphere effectively? An answer based on trajectory analysis of a simulation of a TOGA COARE convective system. *J. Atmos. Sci.*, **69**, 195–213, https://doi.org/10.1175/JAS-D-11-0147.1.

——, J. Gao, C. Ziegler, E. R. Mansell, D. R. MacGorman, and S. Dembek, 2014: Evaluation of a cloud scale lightning data assimilation technique and a 3DVAR method for the analysis and short term forecast of the 29 June 2012 derecho event. *Mon. Wea. Rev.*, **142**, 183–202, https://doi.org/10.1175/MWR-D-13-00142.1.

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

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

——, J. Gao, C. Ziegler, K. 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.

——, S. Stevenson, and R. Rabin, 2018a: Evolution of GLM-observed total lightning in Hurricane Maria (2017) during the period of maximum intensity. *Mon. Wea. Rev.*, **146**, 1641–1666, https://doi.org/10.1175/MWR-D-18-0066.1.

——, and Coauthors, 2018b: Assimilation of total lightning with GSI and NEWS3DVAR to improve short-term forecasts of

high impact weather events at cloud resolving scales. *JCSDA Quart. Newsl.*, No. 58 (Winter 2018), 5–12, https://doi.org/10.7289/V5CJ8BR2.

Fuchs, B. R., E. C. Bruning, S. A. Rutledge, L. D. Carey, P. R. Krehbiel, and W. Rison, 2016: Climatological analyses of LMA data with an open-source lightning flash-clustering algorithm. *J. Geophys. Res. Atmos.*, **121**, 8625–8648, https://doi.org/10.1002/2015JD024663.

Gao, J., and D. J. Stensrud, 2012: Assimilation of reflectivity data in a convective-scale, cycled 3DVAR framework with hydrometeor classification. *J. Atmos. Sci.*, **69**, 1054–1065, https://doi.org/10.1175/JAS-D-11-0162.1.

——, M. Xue, A. Shapiro, and K. K. Droegemeier, 1999: A variational method for the retrieval of three-dimensional wind fields from dual-Doppler radars. *Mon. Wea. Rev.*, **127**, 2128–2142, https://doi.org/10.1175/1520-0493(1999)127<2128:AVMFTA>2.0.CO;2.

——, ——, K. Brewster, and K. K. Droegemeier, 2004: A three-dimensional variational data assimilation method with recursive filter for single-Doppler radar. *J. Atmos. Oceanic Technol.*, **21**, 457–469, https://doi.org/10.1175/1520-0426(2004)021<0457:ATVDAM>2.0.CO;2.

——, and Coauthors, 2013: A real-time weather-adaptive 3DVAR analysis system for severe weather detections and warnings. *Wea. Forecasting*, **28**, 727–745, https://doi.org/10.1175/WAF-D-12-00093.1.

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

Ge, G., J. Gao, K. A. Brewster, and M. Xue, 2010: Effects of beam broadening and Earth curvature in radar data assimilation. *J. Atmos. Oceanic Technol.*, **27**, 617–636, https://doi.org/10.1175/2009JTECHA1359.1.

——, ——, and M. Xue, 2012: Diagnostic pressure equation as a weak constraint in a storm-scale three dimensional variational radar data assimilation system. *J. Atmos. Oceanic Technol.*, **29**, 1075–1092, https://doi.org/10.1175/JTECH-D-11-00201.1.

Godinez, H. C., J. M. Reisner, A. O. Fierro, S. R. Guimond, and J. Kao, 2012: Determining key model parameters of rapidly intensifying Hurricane Guillermo (1997) using the ensemble Kalman filter. *J. Atmos. Sci.*, **69**, 3147–3171, https://doi.org/10.1175/JAS-D-12-022.1.

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

Gurka, J. J., T. A. Schmit, T. M. Renkevens, M. M. Gunshor, and J. Li, 2006: 2006 update on baseline instruments for the GOES-R series. *Proc. SPIE*, **6301**, 63010H, https://doi.org/10.1117/12.683701.

Hong, S.-Y., and J.-O. J. Lim, 2006: The WRF single-moment 6-class microphysics scheme (WSM6). *J. Korean Meteor. Soc.*, **42**, 129–151.

Houtekamer, P. L., and H. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.*, **126**, 796–811, https://doi.org/10.1175/1520-0493(1998)126<0796:DAUAEK>2.0.CO;2.

Houze, R. A., 1993: *Cloud Dynamics*. Academic Press, 573 pp.

Hu, M., M. Xue, and K. Brewster, 2006a: 3DVAR and cloud analysis with WSR-88D Level-II data for the prediction of the Fort Worth, Texas, tornadic thunderstorms. Part I: Cloud analysis and its impact. *Mon. Wea. Rev.*, **134**, 675–698, https://doi.org/10.1175/MWR3092.1.

——, ——, J. Gao, and K. Brewster, 2006b: 3DVAR and cloud analysis with WSR-88D Level-II data for the prediction of the Fort Worth, Texas, tornadic thunderstorms. Part II: Impact of radial velocity analysis via 3DVAR. *Mon. Wea. Rev.*, **134**, 699–721, https://doi.org/10.1175/MWR3093.1.

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

Jones, C. D., and B. Macpherson, 1997a: A latent heat nudging scheme for the assimilation of precipitation data into an operational mesoscale model. *Meteor. Appl.*, **4**, 269–277, https://doi.org/10.1017/S1350482797000522.

——, and ——, 1997b: Sensitivity of the limited area model to the assimilation of precipitation estimates derived from lightning data. UKMO Forecasting Research Tech. Rep. 212, 11 pp.

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

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

Kuhlman, K. M., C. L. Zielger, E. R. Mansell, D. R. MacGorman, and J. M. Straka, 2006: Numerically simulated electrification and lightning of the 29 June 2000 STEPS supercell storm. *Mon. Wea. Rev.*, **134**, 2734–2757, https://doi.org/10.1175/MWR3217.1.

Li, W., Y. F. Xie, S. M. Deng, and Q. Wang, 2010: Application of the multigrid method to the two-dimensional Doppler radar radial velocity data assimilation. *J. Atmos. Oceanic Technol.*, **27**, 319–332, https://doi.org/10.1175/2009JTECHA1271.1.

Lindskog, M., K. Salonen, H. Järvinen, and D. B. Michelson, 2004: Doppler radar wind data assimilation with HIRLAM 3DVAR. *Mon. Wea. Rev.*, **132**, 1081–1092, https://doi.org/10.1175/1520-0493(2004)132<1081:DRWDAW>2.0.CO;2.

Lopez, P., and P. Bauer, 2007: "1D+4DVAR" assimilation of NCEP Stage-IV radar and gauge hourly precipitation data at ECMWF. *Mon. Wea. Rev.*, **135**, 2506–2524, https://doi.org/10.1175/MWR3409.1.

MacGorman, D. R., D. W. Burgess, V. Mazur, W. D. Rust, W. L. Taylor, and B. C. Johnson, 1989: Lightning rates relative to tornadic storm evolution on 22 May 1981. *J. Atmos. Sci.*, **46**, 221–251, https://doi.org/10.1175/1520-0469(1989)046<0221:LRRTTS>2.0.CO;2.

——, I. R. Apostolakopoulos, N. R. Lund, N. W. S. Demetriades, M. J. Murphy, and P. R. Krehbiel, 2011: The timing of cloud-to-ground lightning relative to total lightning activity. *Mon. Wea. Rev.*, **139**, 3871–3886, https://doi.org/10.1175/MWR-D-11-00047.1.

Mach, D. M., H. J. Christian, R. J. Blakeslee, D. J. Boccipio, S. J. Goodman, and W. L. Boeck, 2007: Performance assessment of the Optical Transient Detector and Lightning Imaging Sensor. *J. Geophys. Res.*, **112**, D09210, https://doi.org/10.1029/2006JD007787.

Mansell, E. R., 2014: Storm-scale ensemble Kalman filter assimilation of total lightning flash data. *Mon. Wea. Rev.*, **142**, 3683–3695, https://doi.org/10.1175/MWR-D-14-00061.1.

——, and C. L. Ziegler, 2013: Aerosol effects on simulated storm electrification and precipitation in a two-moment bulk

microphysics model. *J. Atmos. Sci.*, **70**, 2032–2050, https://doi.org/10.1175/JAS-D-12-0264.1.

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

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

Marchand, M., and H. Fuelberg, 2014: Assimilation of lightning data using a nudging method involving low-level warming. *Mon. Wea. Rev.*, **142**, 4850–4871, https://doi.org/10.1175/MWR-D-14-00076.1.

Medici, G., K. L. Cummins, W. J. Koshak, S. D. Rudlosky, R. J. Blakeslee, S. L. Goodman, and D. J. Cecil, 2017: The intra-cloud lightning fraction in the contiguous United States. *Mon. Wea. Rev.*, **145**, 4481–4499, https://doi.org/10.1175/MWR-D-16-0426.1.

Mellor, G. L., and T. Yamada, 1982: Development of turbulence closure model for geophysical fluid problems. *Rev. Geophys. Space Phys.*, **20**, 851–875, https://doi.org/10.1029/RG020i004p00851.

Michel, Y., T. Auligneì, and T. Montmerle, 2011: Heterogeneous convective-scale background error covariances with the inclusion of hydrometeor variables. *Mon. Wea. Rev.*, **139**, 2994–3015, https://doi.org/10.1175/2011MWR3632.1.

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

Mittermaier, M., N. Roberts, and S. A. Thompson, 2013: A long-term assessment of precipitation forecast skill using the Fractions Skill Score. *Meteor. Appl.*, **20**, 176–186, https://doi.org/10.1002/met.296.

Mlawer, E., S. Taubman, P. Brown, M. Iacono, and S. Clough, 1997: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *J. Geophys. Res.*, **102**, 16 663–16 682, https://doi.org/10.1029/97JD00237.

Papadopoulos, A., T. G. Chronis, and E. N. Anagnostou, 2005: Improving convective precipitation forecasting through assimilation of regional lightning measurements in a mesoscale model. *Mon. Wea. Rev.*, **133**, 1961–1977, https://doi.org/10.1175/MWR2957.1.

Park, S. K., and L. Xu, Eds., 2013: *Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications*. Vol. II, Springer-Verlag, 730 pp., https://doi.org/10.1007/978-3-642-35088-7.

Pessi, A. T., and S. Businger, 2009: The impact of lightning data assimilation on a winter storm simulation over the North Pacific Ocean. *Mon. Wea. Rev.*, **137**, 3177–3195, https://doi.org/10.1175/2009MWR2765.1.

Purser, R. J., W.-S. Wu, D. F. Parrish, and N. M. Roberts, 2003a: Numerical aspects of the application of recursive filters to variational statistical analysis. Part I: Spatially homogeneous and isotropic Gaussian covariances. *Mon. Wea. Rev.*, **131**, 1524–1535, https://doi.org/10.1175//1520-0493(2003)131<1524:NAOTAO>2.0.CO;2.

——, ——, ——, and ——, 2003b: Numerical aspects of the application of recursive filters to variational statistical analysis. Part II: Spatially inhomogeneous and anisotropic general covariances. *Mon. Wea. Rev.*, **131**, 1536–1548, https://doi.org/10.1175//2543.1.

Roberts, N. M., and H. W. Lean, 2008: Scale-selective verification of rainfall accumulations from high-resolution forecasts of

convective events. *Mon. Wea. Rev.*, **136**, 78–97, https://doi.org/10.1175/2007MWR2123.1.

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

Rudlosky, S. D., 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.

Schenkman, A. D., M. Xue, A. Shapiro, K. Brewster, and J. Gao, 2011: The analysis and prediction of the 8–9 May 2007 Oklahoma tornadic mesoscale convective system by assimilating WSR-88D and CASA radar data using 3DVAR. *Mon. Wea. Rev.*, **139**, 224–246, https://doi.org/10.1175/2010MWR3336.1.

Skamarock, W. C., and J. B. Klemp, 2008: A time-split non- hydrostatic atmospheric model for weather research and forecasting applications. *J. Comput. Phys.*, **227**, 3465–3485, https://doi.org/10.1016/j.jcp.2007.01.037.

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

Stensrud, D. J., and J. Gao, 2010: Importance of horizontally inhomogeneous environmental initial conditions to ensemble storm-scale radar data assimilation and very short-range forecasts. *Mon. Wea. Rev.*, **138**, 1250–1272, https://doi.org/10.1175/2009MWR3027.1.

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

Thompson, G., R. M. Rasmussen, and K. Manning, 2004: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part I: Description and sensitivity analysis. *Mon. Wea. Rev.*, **132**, 519–542, https://doi.org/10.1175/1520-0493(2004)132<0519:EFOWPU>2.0.CO;2.

Wang, H., and Coauthors, 2017: Improving lightning and precipitation prediction of severe convection using lightning data assimilation with NCAR WRF-RTFDDA. *J. Geophys. Res. Atmos.*, **122**, 12 296–12 316, https://doi.org/10.1002/2017JD027340.

——, and Coauthors, 2018: Continuous assimilation of lightning data using time-lagged ensembles for a convection-allowing numerical weather prediction model. *J. Geophys. Res. Atmos.*, **123**, 9652–9673, https://doi.org/10.1029/2018JD028494.

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

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

——, J. Gao, P. S. Skinner, K. Knopfmeier, T. Jones, P. L. Heiselman, and L. J. Wicker, 2018: Test of a weather-adaptive dual-resolution Hybrid 3DEnVAR and WRF-DART analysis and forecast system for severe weather events. *22nd Conf. on Integrated Observing and Assimilation Systems for the Atmosphere, Ocean, and Land Surface (IOAS-AOLS)*, Austin, TX, Amer. Meteor. Soc., 168, https://ams.confex.com/ams/98Annual/webprogram/Paper330920.html.

Weiss, S. A., D. R. MacGorman, and K. M. Calhoun, 2012: Lightning in the anvils of supercell thunderstorms. *Mon. Wea. Rev.*, **140**, 2064–2079, https://doi.org/10.1175/MWR-D-11-00312.1.

Wiens, K. C., S. A. Rutledge, and S. A. Tessendorf, 2005: The 29 June 2000 supercell observed during STEPS. Part II: Lightning and charge structure. *J. Atmos. Sci.*, **62**, 4151–4177, https://doi.org/10.1175/JAS3615.1.

Xie, Y., S. E. Koch, J. A. McGinley, S. Albers, P. Bieringer, M. Wolfson, and M. Chan, 2011: A space and time multiscale analysis system: A sequential variational analysis approach. *Mon. Wea. Rev.*, **139**, 1224–1240, https://doi.org/10.1175/2010MWR3338.1.

Xue, M., and Coauthors, 2001: The Advanced Regional Prediction System (ARPS)—A multiscale nonhydrostatic atmospheric simulation and prediction tool. Part II: Model physics and applications. *Meteor. Atmos. Phys.*, **76**, 143–165, https://doi.org/10.1007/s007030170027.

——, D. Wang, J. 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 .

Yoshida, S., T. Morimoto, T. Ushio, and Z. Kawasaki, 2009: A fifth-power relationship for lightning activity from Tropical Rainfall Measuring Mission satellite observations. *J. Geophys. Res.*, **114**, D09104, https://doi.org/10.1029/2008JD010370.

Yussouf, N., J. S. Kain, and A. J. Clark, 2016: Short-term probabilistic forecasts of the 31 May 2013 Oklahoma tornado and flash flood event using a continuous-update-cycle storm-scale ensemble system. *Wea. Forecasting*, **31**, 957–983, https://doi.org/10.1175/WAF-D-15-0160.1.

Zhang, F., 2005: Dynamics and structure of mesoscale error covariance of a winter cyclone estimated through short-range ensemble forecasts. *Mon. Wea. Rev.*, **133**, 2876–2893, https://doi.org/10.1175/MWR3009.1.

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

Zhang, J., F. Carr, and K. Brewster, 1998: ADAS cloud analysis. Preprints, *12th Conf. on Numerical Weather Prediction*, Phoenix, AZ, Amer. Meteor. Soc., 185–188.

——, and Coauthors, 2011: National Mosaic and Multi-Sensor QPE (NMQ) System: Description, results, and future plans. *Bull. Amer. Meteor. Soc.*, **92**, 1321–1338, https://doi.org/10.1175/2011BAMS-D-11-00047.1.

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

Zhao, K., and M. Xue, 2009: Assimilation of coastal Doppler radar data with the ARPS 3DVAR and cloud analysis for the prediction of Hurricane Ike (2008). *Geophys. Res. Lett.*, **36**, L12803, https://doi.org/10.1029/2009GL038658.

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, https://doi.org/10.1175/1520-0469(1985)042<1487:ROTAMV>2.0.CO;2.