



RESEARCH ARTICLE

10.1029/2022EA002378

Key Points:

- Geostationary Lightning Mapper (GLM)-based product has the potential to benefit convective scale data assimilation and forecasts
- A new method for assimilating pseudo dewpoint temperature as a function of GLM flash extent density is developed
- The short-range forecast of a severe weather event can be improved by assimilating GLM data with the new method

Correspondence to:

J. Gao,
jidong.gao@noaa.gov

Citation:



Pan, S., & Gao, J. (2022). A method for assimilating pseudo dewpoint temperature as a function of GLM flash extent density in GSI-based EnKF data assimilation system—A proof of concept study. *Earth and Space Science*, 9, e2022EA002378. <https://doi.org/10.1029/2022EA002378>

Received 29 APR 2022
 Accepted 9 OCT 2022

Author Contributions:

Conceptualization: Sijie Pan, Jidong Gao
Funding acquisition: Jidong Gao
Investigation: Sijie Pan
Methodology: Sijie Pan, Jidong Gao
Software: Sijie Pan
Supervision: Jidong Gao
Validation: Sijie Pan
Writing – original draft: Sijie Pan
Writing – review & editing: Jidong Gao

A Method for Assimilating Pseudo Dewpoint Temperature as a Function of GLM Flash Extent Density in GSI-Based EnKF Data Assimilation System—A Proof of Concept Study

Sijie Pan^{1,2,3}  and Jidong Gao^{2,3} 

¹Cooperative Institute for Severe and High-Impact Weather Research and Operations, The University of Oklahoma, Norman, OK, USA, ²NOAA/National Severe Storms Laboratory, Norman, OK, USA, ³School of Meteorology, University of Oklahoma, Norman, OK, USA

Abstract In this study, a new lightning data assimilation (LDA) scheme using Geostationary Lightning Mapper (GLM) flash extent density (FED) is developed and implemented in the National Severe Storms Laboratory Warn-on-Forecast System (WoFS). The new LDA scheme first assigns a pseudo relative humidity between the cloud base and a specific layer based on the FED value. Then at each model layer, the pseudo relative humidity is converted to pseudo dewpoint temperature according to the corresponding air temperature. Some sensitivity experiments are performed to investigate how to assign and use GLM/FED in an optimum way. The impact of assimilating this pseudo dewpoint temperature on a short-term severe weather forecast is preliminarily assessed in this proof-of-concept study. A high-impact weather event in Kansas on 24 May 2021 is used to evaluate the performance of the new scheme on analyses and subsequent short-term forecasts. The results show that the assimilation of additional FED-based dewpoint temperature observations along with radar, satellite radiance, and cloud water can improve short-term (3-hr) forecast skill in terms of quantitative and qualitative verifications against the observations. The improvement is primarily due to the direct and indirect adjustment of dynamic and thermodynamic conditions through the LDA process. More specifically, the assimilation of FED-based dewpoint temperature, in addition to the other observations currently used in WoFS, tends to enhance the ingredients required for thunderstorm formation, namely moisture, instability, and lifting mechanism.

Plain Language Summary Although the high temporospatial resolution Geostationary Lightning Mapper (GLM) from the NOAA GOES-16/17 could be beneficial for convective-scale numerical weather prediction (NWP), only a few studies have explored this potential. In this study, a new ensemble-based GLM data assimilation scheme is developed to assimilate the GLM-based dewpoint temperature into the Weather Research and Forecasting model for improving convective scale NWP. First, some sensitivity experiments are performed to investigate how to use GLM flash extent density in an optimal way. Then, the potential impacts of GLM on convective scale analysis and short-term severe weather prediction are examined for a severe weather event. It is demonstrated that assimilating the GLM product noticeably improves the analysis for key model variables, especially thermodynamic variables. As a result, short-term severe weather forecasts can be improved in terms of rotational tracks and storm strength.

1. Introduction

According to NOAA National Centers for Environmental Information (NCEI, 2021) statistics on the 2021 U.S. Billion-Dollar Weather Disasters, severe storms have caused, on average over the past decade, about 84 fatalities and \$16.8 billion economic losses each year. Scientists have made significant strides toward improving the accuracy of convective-scale forecasts (Clark et al., 2021; Hu et al., 2021; Skinner et al., 2018; Stensrud & Gao, 2010; Zhang, Minamide, et al., 2019; Zhang, Stensrud, & Zhang, 2019). However, the accuracy of severe weather forecasts still suffers due to the inaccurate initial conditions for numerical weather prediction (NWP) models and the complex non-linear interactions between processes of different length scales. More accurate initial conditions can significantly improve convective-scale NWP, although other issues, such as the fidelity of microphysics schemes, may still limit the predictability of convective scale weather events (Sun & Zhang, 2016; Yano et al., 2018).

To provide better initial conditions for convection-allowing models (CAMs), data assimilation (DA) methods require observations with higher spatial and temporal resolution. Usually, conventional observations (e.g., surface

observing stations) only provide information about the pre-storm environment. They have little impact on representing ongoing storm structure, and thus, their impact on CAMs is significantly limited by the sparser temporal density of these data. Therefore, the assimilation of radar and geostationary satellite observations with high temporal resolution in DA has received more attention in the past two decades. For example, the US Weather Surveillance Radar—1988 Doppler (WSR-88D, Crum et al., 1993) network can provide information about storm wind and hydrometeors at a resolution of approximately 1 km and a time frequency of 4–5 min. The Advanced Baseline Imager (ABI) on board the Geostationary Operational Environmental Satellite—R Series (GOES-R) (Schmit et al., 2017) can provide information on temperature, humidity, and wind at 5–10 km resolution and various time intervals from 5 to 60 min.

In recent years, numerous studies have been made to assess the impact on forecast of DA methods that incorporate radar and satellite observations into CAMs (e.g., Aksoy et al., 2009; Dowell et al., 2011; Gao & Stensrud, 2014; Gao et al., 2004; Honda et al., 2018; Johnson et al., 2015; Jones et al., 2015, 2016; Minamide & Zhang, 2019; Polkinghorne & Vukicevic, 2011; Wang & Wang, 2017; Wheatley et al., 2015; Zhang et al., 2016, 2018; Zhang, Minamide, et al., 2019; Zhang, Stensrud, & Zhang, 2019). These studies have demonstrated that assimilating radar and satellite observations can improve the prediction of high-impact weather events, ranging from severe thunderstorms (e.g., Jones et al., 2016; Zhang, Stensrud, & Zhang, 2019) to tropical cyclones (e.g., Honda et al., 2018; Zhang, Minamide, et al., 2019). The assimilation of radar observations, especially radial velocity and reflectivity, enables construction of the three-dimensional internal storm structure (Clark et al., 2021; Dowell & Wicker, 2009; Gao & Stensrud, 2014; Gao et al., 2004). The assimilation of satellite observations can also play an important role in improving the prediction of storms over areas where surface and radar observations are limited, such as oceans or regions where complex terrain can block radar beams (Fierro et al., 2016, 2019; Jones et al., 2013, 2015, 2016, 2020).

National Severe Storms Laboratory's (NSSL's) Warn-on-Forecast System (WoFS, Jones et al., 2020), one of the experimental systems that focuses on short-term forecasting applications, provides reasonable estimates of the initial conditions by optimally combining forecast backgrounds and observations. In the 2021 NOAA Hazardous Weather Testbed (HWT) Spring Forecasting Experiment (SFE, Clark et al., 2021), the assimilation of radar observations from WSR-88D and satellite observations from GOES-R helped to establish the foundation for good performance of the WoFS. The satellite observations used in HWT/SFE 2021 include information on nonprecipitating clouds and environmental conditions for all-sky. Thus far, however, the benefits of assimilating lightning data provided by the GOES-R series to the accuracy of WoF predictions has been limited. The Geostationary Lightning Mapper (GLM) lightning product (Goodman et al., 2013) is widely recognized as a proxy for the occurrence of deep moist convection (Schultz et al., 2011). In particular, prior studies found that total lightning flash rates are indicative of convective cores having $>10 \text{ m s}^{-1}$ updraft velocity and graupel (Carey & Rutledge, 1998; Deierling & Petersen, 2008; MacGorman et al., 1989, 2005, 2011; Wiens et al., 2005). Therefore, assimilating lightning observations into the WoFS is expected to help improve the forecast of severe thunderstorms, especially for the convection initiation phase.

Most previous studies focused on lightning data assimilation (LDA) have applied empirical relationships in convective parameterization schemes to force lightning observations from ground-based platforms to promote convection at observed lightning locations. However, this approach is not appropriate for CAMs that can partially resolve convective features without parameterization schemes. Following the operational availability of GOES-16/17 GLM products since 2018, only a few studies have been performed on the assimilation of real GLM products into convective scale NWP models. Fierro et al. (2019) and Hu et al. (2020) implemented an LDA method in a variational framework by assimilating pseudo water vapor content. Basically, their study showed the potential ability of LDA to achieve improvement of short-range forecasts for a specific set of severe weather events. However, they also underscored that (a) the added value of LDA to forecasts remains minor in areas with good coverage of radar observations and (b) LDA typically produces overestimation of $>30\text{-dBZ}$ reflectivities. This is because the treatment of areal coverage of non-zero flash densities in their LDA scheme does not change significantly with time, resulting in unrealistic progressive enhancement of convection. As shown below, our new method may have potential to improve the treatment of areal coverage of non-zero flash densities.

Compared to the variational DA method with a deterministic member forecast, use of the ensemble Kalman Filter (EnKF) has obvious advantages from two aspects. First, the EnKF technique introduces flow-dependent background error covariances, which allow more accurate spatial and cross-variable correlations between model

states and directly observed variables. Second, an ensemble of forecasts is able to generate probabilistic guidance for severe weather events (Skinner et al., 2018; Flora et al., 2019). Only a few studies, namely Kong et al. (2020, K20 hereafter) and Gan et al. (2021, G21 hereafter), attempted to assimilate lightning observations from spaceborne sensors for real data application using an EnKF framework. K20 adopted an empirical relationship between FED and graupel mass or volume, and demonstrated the potential benefit of the LDA to short-term severe weather forecasts using the Community Gridpoint Statistical Interpolation EnKF (GSI-EnKF, Whitaker and Hamill 2002) system. They highlighted that the positive correlations between FED and model state variables, especially temperature and moisture, are the most important factors leading to a better forecast. Although more intense convective updrafts were produced by positive correlations between FED and vertical velocity, those updrafts had little influence on the forecast. In G21, the LDA scheme adopted an empirical relationship between flash rate and maximum vertical velocity. They suggested a similar conclusion, that convective scale short-term forecasts have been improved primarily because of increased convergence and divergence of wind fields in the low and upper levels, respectively.

None of the aforementioned studies examined the impact of LDA on the forecast of a tornadic supercell event. The present study adopts the assumption used in Fierro et al. (2019) but improves on three aspects. These aspects include: (a) using FED instead of “flash origin density,” which does not consider the areal extent of the flashes; (b) extending the LDA algorithm to use flash rate information to create pseudo dewpoint temperature observations; (c) combining the assimilation of FED with other observations used by the WoFS. The primary motivation of this research is to leverage the potential for GLM FED to aid the development of properly placed convection with minimal overestimation in NWP models. In this new research, relative humidity is added based on the FED information which adds moisture information in a way which should help reduce over-prediction biases.

Although the new LDA scheme developed here is designed as a real-time application for future SFEs, as a proof-of-concept, this research will first focus on a single case study: the 24 May 2021 severe weather event over northwest Kansas, where a large supercell with multi-vortex tornadoes occurred, but was not well predicted by the WoFS. This type of severe weather event is different from that in K20 and G21, which are better described as MCSs rather than supercells. Our initial focus here is on how to use GLM/FED in an optimum way through sensitivity experiments. The preliminary results show that the assimilation of FED, using the GSI-EnKF system, improves the short-term forecast of both tornadic and nontornadic supercells. Section 2 describes the WoFS system, the FED observations, and the LDA scheme. Section 3 introduces the experiment configurations and verification method used in this study. In Section 4, sensitivity experiments examine the impact of assimilating GLM FED observations using different spatial resolutions, accumulation windows, and cutoff radii of FED observations. Section 5 describes qualitative and quantitative comparisons of forecast quality for each assimilation experiment under a quasi-real time situation, followed by conclusions in Section 6.

2. Warn-On-Forecast System and LDA Scheme

2.1. Overview

The WoFS system is an on-demand ensemble DA and forecasting system designed to provide guidance of hazardous weather events, such as tornadoes, damaging winds, large hail, and flash flooding. The current WoFS uses a customized Advanced Research Weather Research and Forecasting Model (WRF-ARW) based on version 3.8.1 (Skamarock et al., 2008), coupled with the GSI-based EnKF system (Jones et al., 2016, 2020; Yussouf & Knopfmeier, 2019; Yussouf et al., 2020). The GSI-EnKF system assimilates radar radial velocity and reflectivity data (Johnson et al., 2015; Wang & Wang, 2017), satellite cloud water path (CWP, Jones et al., 2013) and GOES-16 ABI radiance (Jones et al., 2020). This research extends the system to include an LDA method that assimilates FED as an indicator of strong ongoing convection.

As mentioned in the first section, one advantage of the EnKF method is that the flow dependent error covariances can be estimated so that prognostic variables, including three-dimensional wind fields, temperature, humidity, pressure, diabatic heating, and hydrometeors can be updated without using forward operators that directly link model variables with observations. The configurations used in this research are the same as those used in the real-time WoFS run during the SFE 2021 but with the additional assimilation of GLM FED observations, which are described later. The WoFS assimilates available conventional data every 1-hr; radar and satellite observations every 15-min starting from 1500 UTC and cycling until 0300 UTC. All observations are assimilated into the

WRF model with a regional domain having 3-km grid spacing and 51 vertical levels. The regional domain covers an area of 301×301 grid points (900×900 km), with its center determined by the Day 1 Convective Outlook from NOAA Storm Prediction Center. The initial and boundary conditions for the WoFS 36-member ensemble are provided by an experimental 36-member HRRR ensemble (HRRRE, Benjamin et al., 2016). The initial conditions use 1-hr forecasts from 1400 UTC HRRRE analyses, and the boundary conditions are generated by using forecasts of the analyses of the first 9 HRRRE members at 1200 UTC. Then, 6- and 3-hr forecasts are launched at the top and the bottom of each hour, respectively. For more details about parameterization schemes, see Skinner et al. (2018); Wheatley et al. (2015); Mansell et al. (2010). A comprehensive description of the additive noise and adaptive inflation methods used to maintain diversity among members in the real-time WoFS can be found in Dowell and Wicker (2009); Anderson (2009) and Hu et al. (2019). The same settings for these inflation methods are used in this study.

2.2. Description of Observations Used in the WoFS for SFE 2021

This study examines the added value of an LDA scheme to the WoFS real-time performance. Therefore, observations used in the WoFS real-time experiments are also incorporated in this study. The WoFS system for SFE 2021 assimilates conventional observations hourly and radar radial velocity and reflectivity, CWP, and all sky radiance at 15-min intervals. The conventional observations contained in hourly files include temperature, dewpoint, winds, and pressure available at the surface and upper levels. For real-time application, the conventional observations are assimilated using a 15-min lag, for example, using the conventional data available at 1500 UTC in the 1515 UTC analysis. Oklahoma and west Texas Mesonet observations (Brock et al., 1995) that have a denser spatial resolution are assimilated as a complement to traditional observations in PrepBUFR format (https://rda.ucar.edu/docs/formats/bufr/BUFR_PrepBUFR_User_Guide_v1.pdf).

Radial velocity observations are used within a range of 150 km from the radar and are thinned to a 5-km Cartesian grid using Cressman objective analysis (Cressman, 1959) prior to the DA procedure. For reflectivity observations, NSSL provides an alternative three-dimensional quality-controlled option, called Multi-Radar Multi-Sensor (MRMS) reflectivity (Smith et al., 2016), which has horizontal resolution at 0.01° (approximately 1 km) and vertical resolution ranging from 250 m to 1 km from sea level surface to 20 km mean sea level. Akin to radial velocity observations, MRMS reflectivities are thinned to 5-km horizontal resolution.

The satellite observations assimilated into the WoFS include radiances from 6.2 to 7.3 μm infrared bands, respectively measuring upper- and low-level water vapor content, and CWP, which is one of the cloud property products from GOES-16/17 representing vertically-aggregated hydrometeors. These observations are analyzed to the same MRMS 5-km grid for cloudy areas and a 15-km grid for clear sky. A parallax correction is applied to satellite data for cloudy pixels defined by the L2 cloud height product. More detailed descriptions about all assimilated observations, including their associated errors and localization radii in SFE 2021 can be found in Jones et al. (2020), except that vertical localization radii, for conventional observations ranges from 0.8 to 0.85 in natural log pressure $[-\ln(P/P_0)]$. Here $P_0 = 1,000$ hPa is the reference pressure.

2.3. GLM Flash Extent Density and the Assimilation Scheme

The GLM instruments carried by GOES-16/17 are able to measure optical signals emitted by lightning discharges over most of the Americas and central and eastern Pacific Ocean. It has a variable pixel pitch that prevents the horizontal resolution of ground samples from being finer than 14 km over the CONUS, and makes most ground samples have horizontal resolution less than 10 km (Bruning et al., 2019). The detection efficiency of the GLM is highly dependent on the time of day. The expected flash detection efficiency exceeds 70% for daytime and 90% for nighttime. A detection, also called an event in GLM data sets, is recorded at individual pixels that exceed background within the 2-ms integration period once non-lightning artifacts are recognized and eliminated. A single event and simultaneous events at adjacent pixels are clustered into a group that could be considered as a single lightning pulse. Finally, a GLM flash may consist of one or more sequential groups occurring within 16.5 km and within 330 ms. More details about non-lightning artifacts filtering and GLM products can be found in Rudlosky et al. (2019) and Mach (2020).

In this study, the method and the open-source python package called *glmtools*, developed by Bruning et al. (2019), is adopted to derive FED observations. Instead of simply accumulating flashes on each grid, this method considers

a connection between events, groups, and flashes because GLM events are not sampled with even spacing. It takes into account the actual spatial footprint of each GLM event and is able to fill the gaps between pixels. Observed events with an even distribution could be counted on a target grid matching the GLM detector configuration. However, from the geostationary perspective, evenly distributed events are unevenly distributed on a regular grid (such as latitude-longitude grid) because of longitude convergence toward the pole, leading to overlapping events in a single grid on the equator side and gaps on the pole side. The remapping technique introduced by Bruning et al. (2019) is superior to regular image remapping techniques which usually do not consider the situations described above. Basically, accumulating the GLM data has the following six steps: (a) reconstruction of the event-flash linkage from the GLM L2 data; (b) computation of fixed grid coordinates matching to geolocated GLM L2 data, including lightning ellipsoid removal and fixed grid navigation; (c) spatial aggregation of each event, and accumulation of flash and group properties; (d) partitioning of aggregated events by the target grid; (e) accumulation on the target grid using the pre-accumulated properties; (f) producing a time accumulation by simple addition. By applying this method to raw 20-s GLM data, it is possible to derive FED observations for any spatial resolution or time window. This also allows sensitivity experiments to examine the impact of assimilating derived FED data with different resolutions and accumulation windows.

The derived GLM FED observations are assimilated into the WRF simulations using the following procedure: whenever $FED > 0$ at a given latitude-longitude coordinate, pseudo dewpoints are provided for model levels (between cloud base to 650 hPa) associated with the location of the observed FED based on Lawrence (2005) Equation 8:

$$T_d = \frac{B \left[\ln \left(\frac{RH}{100} \right) + \frac{AT_{air}}{B + T_{air}} \right]}{A - \ln \left(\frac{RH}{100} \right) - \frac{AT_{air}}{B + T_{air}}} \quad (1)$$

Based on Teten's suggestions, the values of A and B are 17.269 and 237.3 respectively. T_{air} is the air temperature in Celsius, and RH is the pseudo relative humidity (RH) calculated via a sigmoid equation that is a function of the FED observation:

$$RH = \left(X + \frac{Y}{1 + e^{C-D*FED}} \right) \% \quad (2)$$

The philosophy behind the equation is very similar to Reisner and Jeffery (2009) in which a hyperbolic tangent equation was used to avoid sharp variations in cloud variables. This Sigmoid function is a function widely used in machine-learning related research for binary classification (Gagne II et al., 2019). It is monotonic, continuous, and differentiable and is limited to a range from 0 to 1. These features make it ideal as a bridge between FED and the amount of pseudo relative humidity (Figure 1). This function can also be used to eliminate possible discontinuities associated with the step function to prevent the LDA scheme from generating discrete pseudo observations at adjacent points. With the sigmoid function, it is convenient to select the asymptotic value (through X , Y , and C), and the slope (through D) of the function. The red line in Figure 1 gives a more intuitive perception of the relationship between pseudo RH and FED observations used in this study. C , D , X , and Y are specified as 5.25, 2.85, 65, and 30 respectively. These values are chosen to make the sigmoid function shift to the right and have a steeper changing rate and are determined by a trial-and-error technique to make sure they produce reasonable RH values based on information from FED. The value of pseudo RH within a layer between cloud base and 650 hPa is proportional to the value of FED. It is minimized when FED is near 0 and is maximized when FED approaches and exceeds four flashes per unit area per min (hereafter $fl \text{ area}^{-1} \text{ min}^{-1}$). The derived pseudo RH value is constrained to lie between 65% and 95%. By combining a steeper RH slope with such asymptotic values, it is expected that the possibility of isolated convection initiation and its ensuing development will be increased, while limiting the generation of spurious convection. The observation error of derived dewpoint temperature, u_{T_d} was described in Lin and Hubbard (2004) and can be written as:

$$u_{T_d} = 2 \left[\left(\frac{\partial T_d}{\partial T_{air}} \right)^2 u^2(T_{air}) + \left(\frac{\partial T_d}{\partial RH} \right)^2 u^2(RH) \right]^{1/2} \quad (3)$$

where $u(T_{air})$ is the background error for the forecast model and $u(RH)$ is set to 0.05.

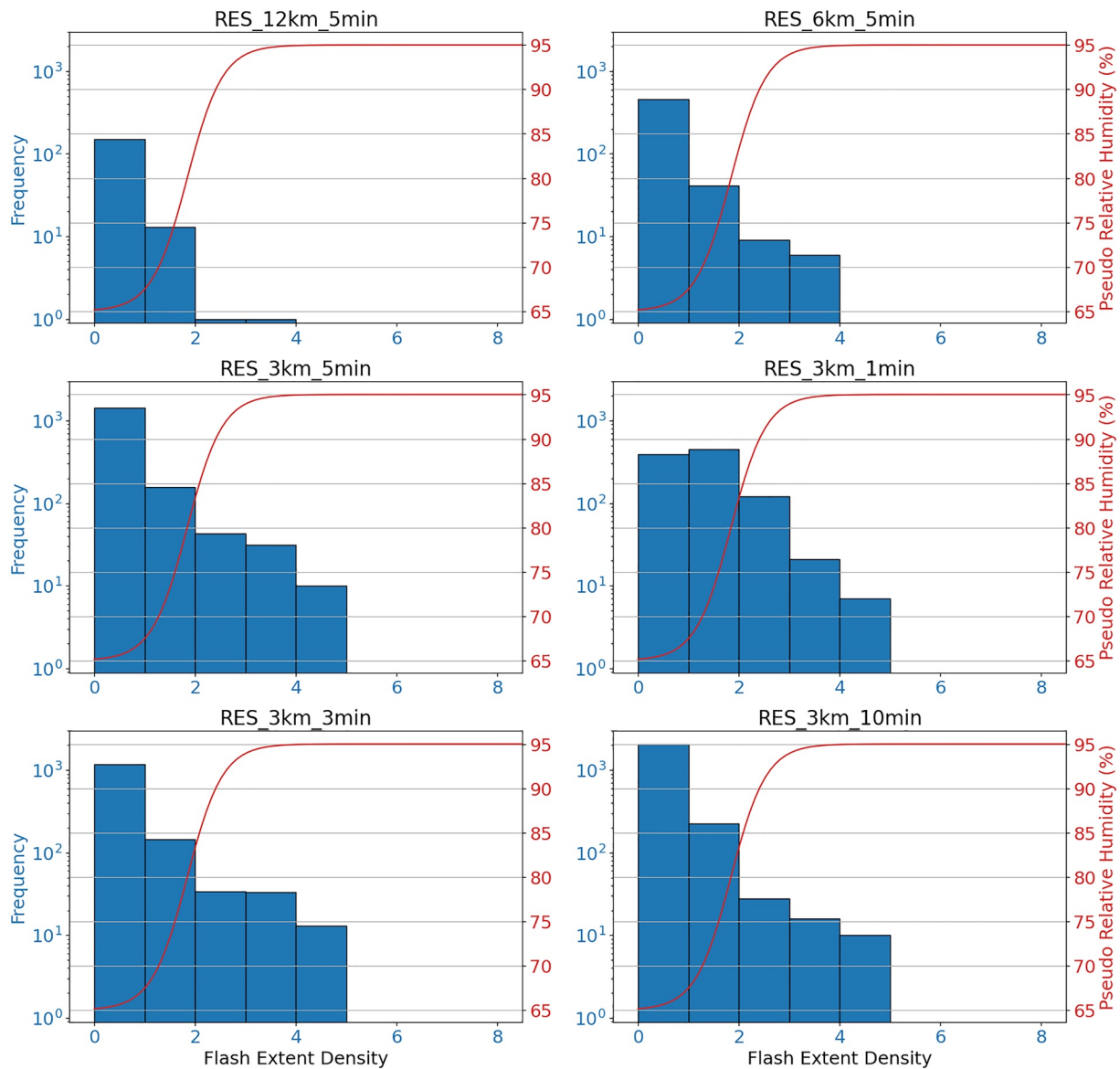


Figure 1. Frequency of the flash extent density (FED) ($\text{fl area}^{-1} \text{min}^{-1}$) observations using different accumulation strategies (blue bars) at 2030 UTC, 24 May 2021. The red line represents pseudo relative humidity as a function of FED value.

3. Experimental Design and Verification

3.1. Experiment Configuration

The experiment configurations are the same as the WoFS configurations used in SFE 2021, except that the assimilation of radar and satellite observations starts from 1800 UTC 24 May 2021 and the assimilation cycles end at 2200 UTC, the same day. The EnKF DA cycles between 1500 and 1800 UTC only assimilate conventional observations, and this is considered as a spin-up process. All experiments have exactly the same initial and boundary conditions at 1800 UTC. This allows the evaluation of the impact of LDA on the forecast of severe weather events with minimum computational resources. Few differences have been found in the subjective comparison between forecasts from this setup and those from the real-time runs for the 24 May 2021 event (not shown).

As mentioned in Section 2c, the FED observations can be derived with any spatial resolution and time period by using raw 20-s GLM data. The performance of LDA could be sensitive to the spatial resolution or accumulation strategy for FED observations. Therefore, two sets of sensitivity experiments using FED observations with

Table 1
Experiment Configurations for Different Sensitivity Experiments

Experiment	Spatial resolution of GLM FED (km)	Accumulation window of GLM FED (min)	Horizontal localization (km)	Vertical localization
RES_12 km_5 min	12	5	9	0.85
RES_6 km_5 min	6	5	9	0.85
RES_3 km_5 min	3	5	9	0.85
RES_3 km_1 min	3	1	9	0.85
RES_3 km_3 min	3	3	9	0.85
RES_3 km_10 min	3	10	9	0.85
RADII_h12v1.0	3	5	12	1.0
RADII_h12v0.7	3	5	12	0.7
RADII_h6_1.0	3	5	6	1.0
RADII_h6v0.7	3	5	6	0.7

different spatial resolutions, accumulation windows, and horizontal and vertical localization radii are considered by this study (Table 1). These sensitivity experiments only assimilate GLM FED observations using the WoFS configuration described above.

The experiments in the first set are named RES_12 km_5 min, RES_6 km_5 min, RES_3 km_5 min, RES_3 km_1 min, RES_3 km_3 min, and RES_3 km_10 min, respectively, based on the spatial resolution and accumulation window used in the FED derivation. The frequency of the FED observations valid at 2030 UTC is also shown in Figure 1. As expected, more observations are generated with higher spatial resolution using the super-resolution remapping technique. On the other hand, observations are more likely concentrated around lower values when the accumulation window is greater. This is because peak values are smoothed over the longer time window. Additionally, the number of points with $FED > 3 \text{ fl area}^{-1} \text{ min}^{-1}$ in RES_3 km_1 min is much less than that in RES_3 km_3 min and RES_3 km_5 min due to information loss using such a short accumulation window. The optimal resolution and accumulation window is determined by the object-based verification method (Flora et al., 2019; Skinner et al., 2018) that is briefly described in the following subsection.

The sensitivity experiments in the second set examine the impact of the localization radius for GLM FED assimilation. The horizontal radius ranges from 6 to 12 km and the vertical radius ranges from 1.0 to 0.7 in the log pressure unit. So four more experiments are performed and named as RADII_h12v1.0, RADII_h12v0.7, RADII_h6v1.0 and RADII_h6v0.7 (Table 1). Experiments in this set are based on the configuration of RES_3 km_5 min that has the best performance in the first set of sensitivity experiments.

Three experiments in the final set (Table 2) are performed to assess the potential benefits from the assimilation of FED observations, once the optimal parameters (i.e., spatial resolution and accumulation window for FED derivation, horizontal and vertical localization) are determined by the two previous sensitivity experiments. The first one is the “Control” experiment, which only assimilates conventional observations. The second is the “Retro” experiment, which assimilates conventional observations from hourly prepbuf file and Oklahoma Mesonet, radar radial velocity and reflectivity observations, and GOES-16 CWP, 6.2 and 7.3 μm infrared radiances. This experiment acts as a retrospective run to match the performance of the WoFS in its real-time runs. The final experiment is called “RetroFED”. It assimilates pseudo dewpoint temperature observations derived from FED observations,

Table 2
Assimilated Observation Types for Three Different Experiments: Control, Retro, and RetroFED

Experiment	Conventional observations (PrepBUFR + Mesonet)	Radial velocity + reflectivity	GOES-16 radiance + CWP	GOES-16 GLM FED
Control	Y	N	N	N
Retro	Y	Y	Y	N
RetroFED	Y	Y	Y	Y

in addition to satellite, radar, and conventional observations. In Section 4, we further describe the analyses and 3-hr forecasts initialized at 2030 UTC, as well as object-based verification.

3.2. Verification

To quantitatively evaluate the quality of ensemble forecasts in each experiment, this study adopts the object-based verification technique developed by Skinner et al. (2018, S18 hereafter). Strong mid-level rotation signatures that are highly related to severe weather were identified as objects and were used for evaluation. Similar to S18, single thresholds based on the 99.95th percentile value in both the real-time WoFS forecasts of SFE 2021 and MRMS verification data set were used to identify objects for this case. Observed mid-level rotation objects are identified by determining locations where 2–5 km MRMS azimuthal wind shear is greater than 0.0025 s^{-1} , while WoFS mid-level rotation objects are defined as the locations where predicted 2–5 km updraft helicity (UH) is greater than $40.74 \text{ m}^2 \text{ s}^{-2}$ (Skinner et al., 2018). Objects are created every 5-min starting from 15 min into the forecast by collecting ± 15 min azimuthal wind shear or UH. Once isolated objects are identified, they are merged if the distance between objects is less than 10 km.

Predicted and observed objects are matched using the total interest score (Davis et al., 2006; also see Equation 1 in S18), which takes into account spatial and timing displacement. To remove edge artifacts, the object-based verification is calculated over a 286×286 gridpoint domain. A qualified match must have a total interest score greater than 0.2, where the maximum searching distance is 40 km and the maximum searching time is 15 min. The matching procedure allows predicted objects to be classified by four performance metrics: “hits,” “misses,” “false alarms,” and “correct nulls,” allowing calculation of contingency table statistics for comparing forecast skills among experiments.

4. Sensitivity Experiments

4.1. Sensitivity Experiments for the FED Observations

The first set of sensitivity experiments only assimilates FED observations with different spatial resolutions and accumulation windows. Figure 2 gives an example of assimilated FED observations at 2030 UTC. Since the super-resolution resampling takes into account the actual spatial footprint of GLM events and fills the gaps between consecutive flash events, resampling at finer spatial resolution should not significantly change the FED values but it will change the observation density (as shown in Figures 2a–2c). Conversely, a greater accumulation window may smooth small time-scale features, and a smaller accumulation window may lose information about consecutive flash events while using the 15-min assimilation cycles. Therefore, an inappropriate accumulation window may eventually lead to lower peak values of GLM FED observations. This conclusion is proved by comparing the FED observations accumulated over different time lengths (Figures 2a, 2d, and 2f). Compared with RES_3 km_5 min, RES_3 km_10 min provides a wider coverage and a smaller peak value (blue circle in Figure 2d vs. Figure 2a). On the other hand, RES_3 km_1 min provides a limited coverage and a smaller peak value (blue circle in Figure 2f vs. Figure 2a). The FED observations of RES_3 km_5 min and RES_3 km_3 min (Figure 2a vs. Figure 2e) are comparable and preferable.

While the qualitative comparison of FED observations using different derivation strategies shows substantial differences, it is important to quantify the different impacts of assimilating these observations and to determine the skill of subsequent forecasts. The FED observations are assimilated into the WoFS with 15-min assimilation cycles starting at 1800 UTC, followed by 3-hr forecasts initiating from 2000, 2030, 2100, 2130, and 2200 UTC. The localization radius used here is initially set to 9 km horizontally and 0.85 log pressure vertically. The quality of predicted rotation objects is assessed using contingency table statistics and performance diagrams related to probability of detection (POD), false alarm rate (FAR), critical success index (CSI), and frequency bias. The closer the values of POD and CSI are to 1, the better the forecast. The perfect prediction score is located in the upper right corner of the table. Figure 3 shows 30–180-min aggregated forecast performance for each experiment that assimilates FED with various spatial resolutions. The overall quality of 60-min rotation forecasts from experiments assimilating observations with various resolutions is fairly reasonable, as evidenced by CSI values exceeding 0.25. RES_3 km_5 min performs better than RES_6 km_5 min and RES_12 km_5 min in terms of CSI or frequency bias for the first 120-min forecast (Figures 3a–3d). On average, RES_3 km_5 min produces CSIs less than 0.2 and frequency biases of approximately 1.2 for the 150- and 180-min forecasts, comparable to

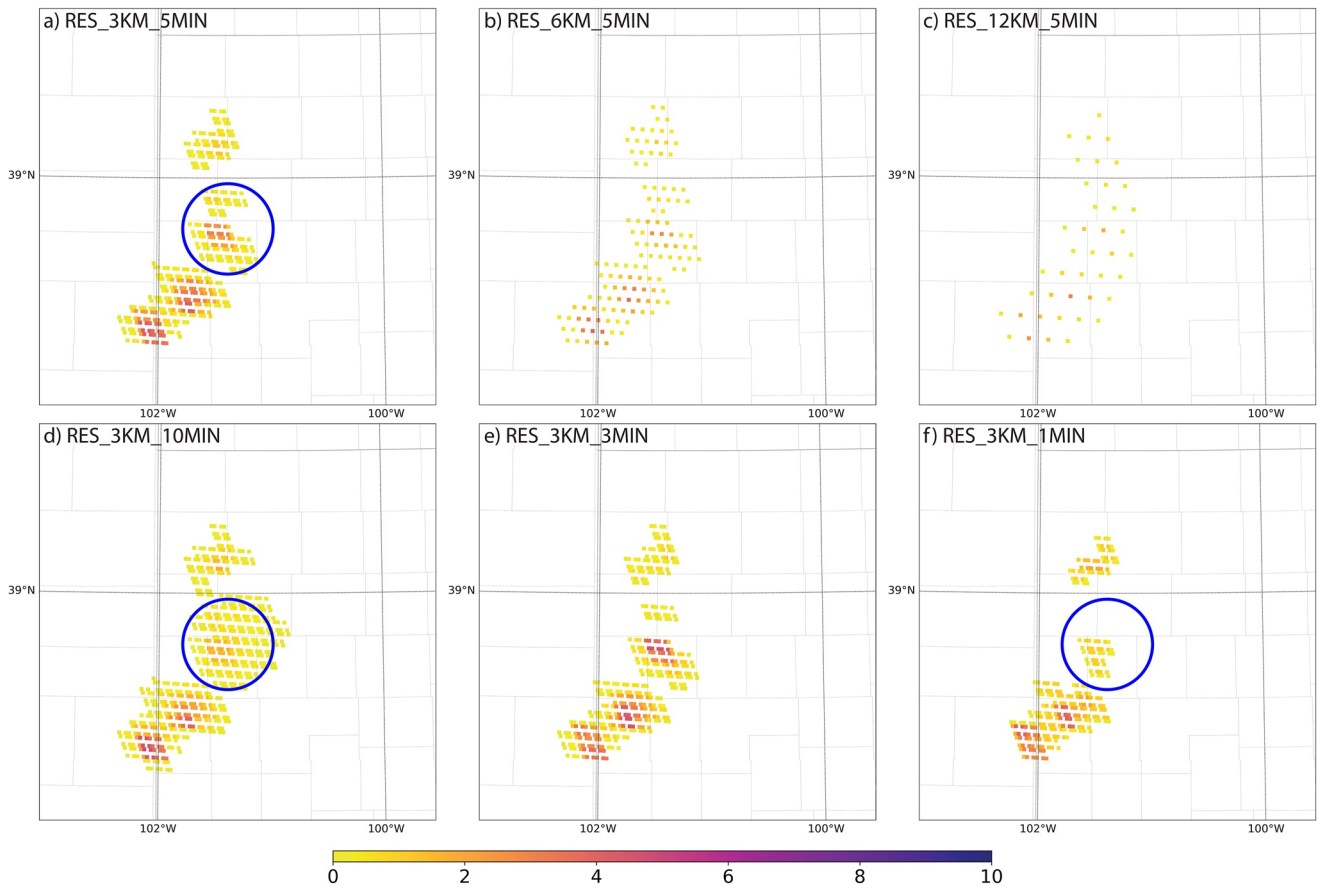


Figure 2. Derived flash extent density (FED) observations at 2030 UTC, 24 May 2021 from the raw FED data for (a) 3-km resolution and 5-min window, (b) 6-km resolution and 5-min window, (c) 12-km resolution and 5-min window, (d) 3-km resolution and 10-min window, (e) 3-km resolution and 3-min window, and (f) 3-km resolution and 1-min window. The blue circles highlight the huge difference between the experiments using different accumulation windows.

RES_6 km_5 min and RES_12 km_5 min (Figures 3e and 3f). But RES_12 km_5 min produces some individual forecasts associated with CSIs greater than 0.3.

The differences in performance among the experiments with different accumulation windows is relatively small and mixed. The skills gradually decrease as a function of forecast time (Figure 4). RES_3 km_5 min performs best at the 90-min (Figure 4c) and 150-min (Figure 4e), while RES_3 km_3 min outperforms RES_3 km_5 min at 30-min (Figure 4a), 120-min (Figure 4d), and 180-min (Figure 4f). RES_3 km_1 min generally performs the worst among all experiments due to the narrow accumulation window and the limited coverage. Overall, RES_3 km_3 min or RES_3 km_5 min performs better than the other two experiments. Thus the 3-km spatial resolution and 5-min accumulation window will be selected as the optimal parameters for generating FED observations in the following experiments.

4.2. Sensitivity Experiments to Ensemble Covariances Localization

The second set of sensitivity experiments aim to select the optimal localization radius for assimilating FED derived pseudo dewpoint temperature. In order to construct multiple assimilation configurations with different horizontal and vertical localizations, each horizontal radius selected from 6, 9, and 12 km could be combined with each vertical radius selected from 0.7, 0.85, and 1.0 in the log pressure unit. This implies a total of nine experiments could be performed. But for simplicity, only five experiments are discussed here, named as RES_3 km_5 min (that uses 9 km and 0.85), RADII_h12v1.0, RADII_h12v0.7, RADII_h6v1.0 and RADII_h6v0.7. The impact of changing horizontal and vertical localization on the forecast can be easily separately tested.

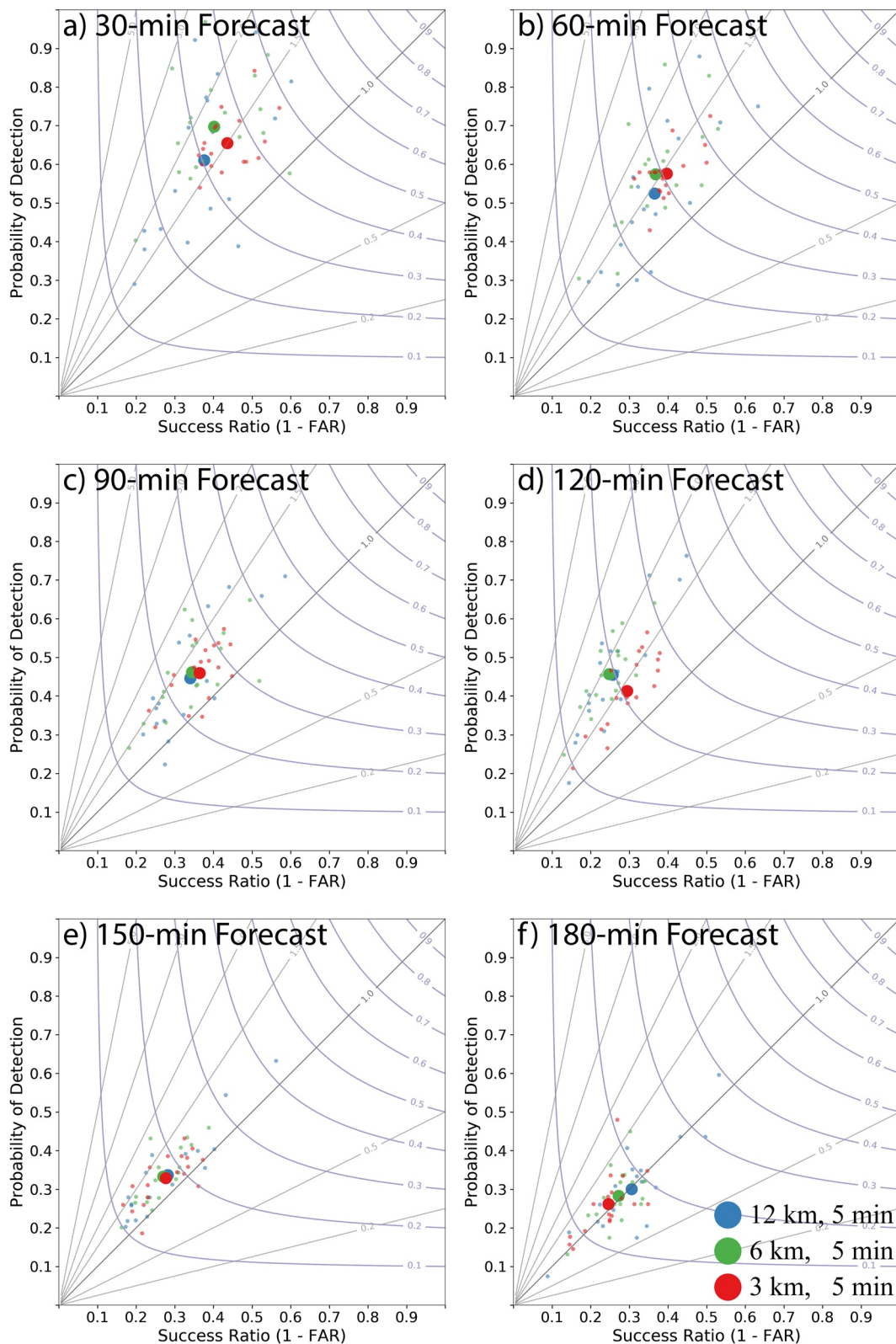


Figure 3. Performance diagram of (a) 30-min, (b) 60-min, (c) 90-min, (d) 120-min, (e) 150-min, (f) 180-min forecasts for the sensitivity experiments RES_12 km_5 min (blue), RES_6 km_5 min (green), and RES_3 km_5 min (red) aggregated over 180 min forecasts initialized at 2000 UTC, 2030 UTC, 2100 UTC, 2130 UTC, and 2200 UTC, 24 May 2021. Small dots represent scores of individual ensemble members, and large dots represent the ensemble mean from each experiment.

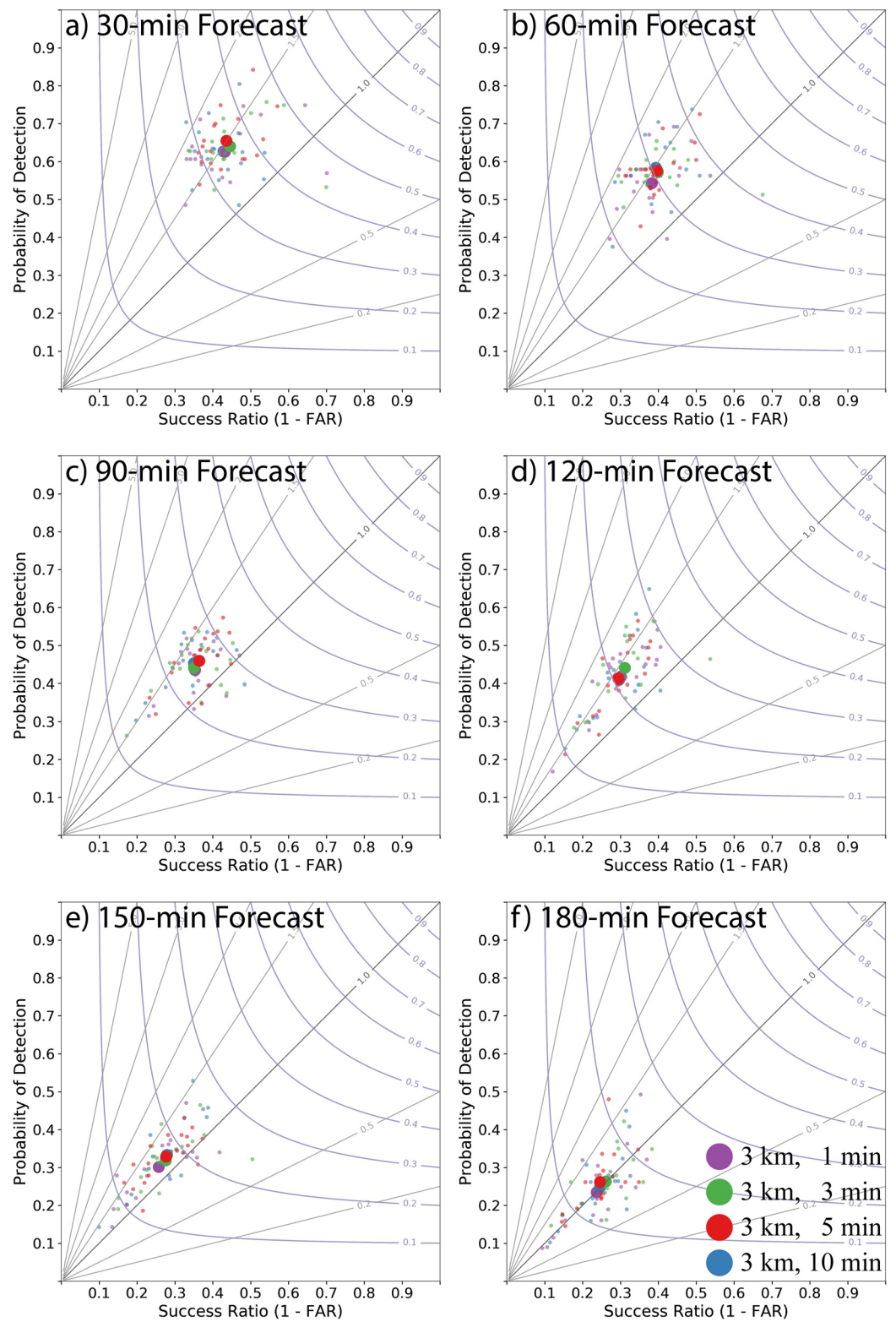


Figure 4. Same as Figure 3, but for the sensitivity experiments RES_3 km_1 min (purple), RES_3 km_3 min (green), RES_3 km_5 min (red), and RES_3 km_10 min (blue).

By comparing experiments using the same vertical localization (Figure 5, purple vs. green, or blue vs. pink), we see that larger horizontal localization leads to slightly worse rotation predictions. However, larger vertical localization (represented by a smaller value) leads to slightly better rotation predictions (Figure 5, purple vs. blue, or green vs. pink). Assimilation using smaller horizontal or larger vertical localization typically increases the POD or decreases the FAR. The 2-hr forecast launched from 2000 UTC (Figure 6) provides an example of how different localization combinations can affect short-term forecasts of convection. Among the five experiments, improvements of POD and FAR are observed when a smaller horizontal localization and a larger vertical localization are applied. Objects associated with south bias are partially corrected, and an additional object is generated (red box in Figure 6). In addition, prediction of the northernmost rotation object in Kansas is not good in all experiments, and the performance among all experiments is quite close. Generally speaking, RADII_h6v0.7 slightly outperforms other experiments. Thus, the combination of the horizontal localization of 6 km and the vertical localization of 0.7 is selected as the configuration for the following case study.

5. Results

5.1. The Event Overview

Four long-lived supercells developed in western Kansas on 24 May 2021. Composite radar reflectivity at 2200 UTC shows all of these supercells, along with warnings issued and severe weather reports collected by the National Weather Service (Figure 7a). The first supercell appeared in western central Kansas (Storm 1) between 1830 and 2300 UTC. The second supercell (Storm 2) was located southwest of Storm 1 between 1900 and 2230 UTC. The third one (Storm 3), which was located north of Storm 1 between 1930 UTC and 0000 UTC and produced several tornadoes during this period, was not forecasted in the real-time WoFS run. The last supercell (Storm 4) developed between Storm 1 and Storm 2 by 2030 UTC, split around 2215 UTC and was continually developing and propagating to the south until 0500 UTC. At 2200 UTC, Storm 3 had already produced several tornadoes and was associated with reflectivity greater than 50 dBZ. Corresponding FED observations at 2200 UTC (Figure 7b) show that Storm 3 generated the highest FED values (about $8 \text{ fl area}^{-1} \text{ min}^{-1}$) compared to the other thunderstorms (about $6 \text{ fl area}^{-1} \text{ min}^{-1}$). It is usually true that more frequent flash rates serve as an indicator of strong convection (Liu et al., 2012).

5.2. Comparisons of the Experiment Results

By 2030 UTC 24 May, Storm 3 was well-organized and produced multiple tornadoes in the prior half hour. It continued to produce tornadoes until 2330 UTC, followed by its dissipating stage. However, the realtime WoFS run initiated from 2000 to 2030 UTC failed to predict this long-lived supercell. The forecasts launched after 2100 UTC also claimed a rather low probability of less than 40% of generating strong rotation for Storm 3. To assess the impact of assimilating additional GLM FED data on the forecast of thunderstorms, the evolutions of the rotation objects for all ensemble members during the 3-hr forecast period starting from 2030 UTC as well as its initial storm environment are generated.

An examination of Figure 8 shows significant differences in predicting rotation objects among the three experiments Control, Retro, and RetroFed. Since Storm 2 entered its dissipating stage after 2200 UTC, it is not discussed. Obvious differences between Control and Retro can be found in the forecast of Storm 1, 3, and 4. For the first hour prediction, the Control experiment predicts widely distributed rotation objects in west Kansas (Figure 8a). Some ensemble members place the supercells between observed Storm 1 and Storm 3/4, while some ensemble members place the supercells further south than observed Storm 4 (Figure 8a). Although Control successfully predicts rotation near Storm 3 by 2130 UTC in 1/3 of the ensemble members, the locations are widely distributed and deviate from the observed object due to the lack of the assimilation of high-density observations. Retro generates a more concentrated distribution of Storm 1 and 4 around their actual locations, but appears to miss Storm 3 in most of the ensemble members (Figure 8b). RetroFED maintains the concentrated distribution for the forecasts of Storm 1 and Storm 4, while generating more instances of Storm 3 in northwest Kansas (Figure 8c). Compared to the observed object, the coverage of Storm 3 in this forecast has an eastward bias.

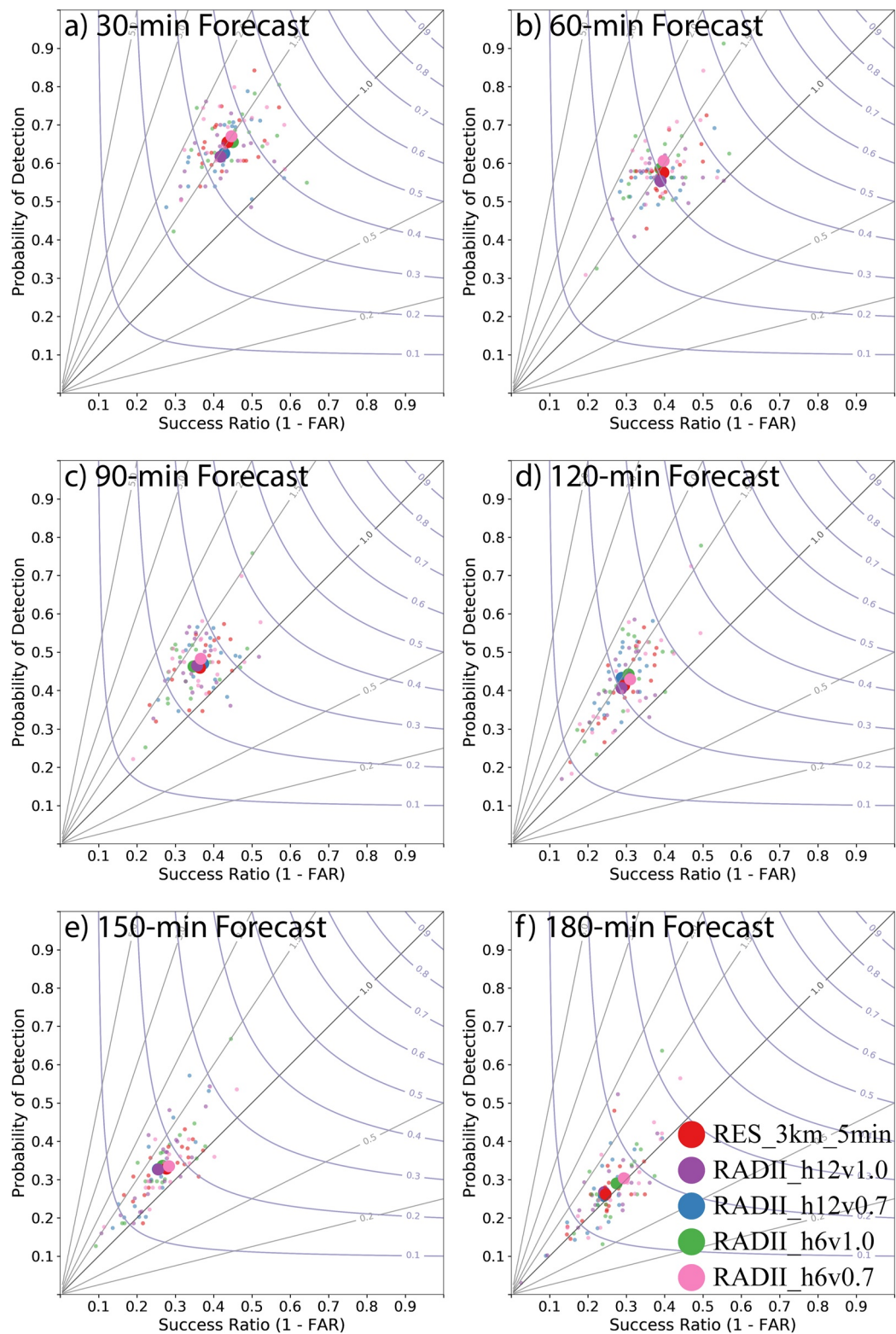


Figure 5. Same as Figure 3, but for the sensitivity experiment RADII_h12v1.0 (purple), RADII_h12v0.7 (blue), RADII_h6v1.0 (green), RADII_h6v0.5 (pink), and the reference experiment RES_3 km_5 min (red).

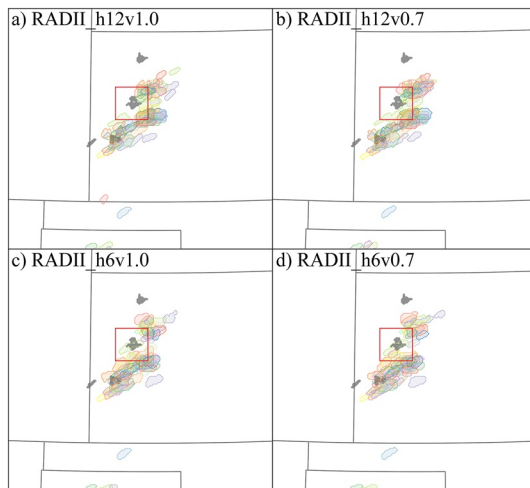


Figure 6. Predicted rotation objects derived from the ensemble forecasts valid at 2200 UTC for 2-hr forecasts initiated at 2000 UTC 24 May 2021 for different localization radii. Each ensemble member is plotted as a different color, and the gray shading represent observed Multi-Radar Multi-Sensor rotation objects at the same time. Red boxes highlight the improvement of experiments using smaller horizontal localization radii.

Transitioning to the second and the third hour forecasts, all experiments exhibit similar biases as the 1-hr forecast. Control still generates overly diverse ensemble forecasts (Figures 8d and 8g), while Retro generates rotation objects near Storm 3 for 3–6 ensemble members (Figures 8e and 8h). The members of RetroFED that successfully predict Storm 3 place the rotation objects to the south of the observed object at 2230 UTC (Figure 8f), and a little east of the object at 2330 UTC (Figure 8i). Compared to the other experiments, RetroFED also produces fewer instances of Storm 1 at 2330 UTC, which actually began dissipating around 2300 UTC. It is worth noting that all experiments predict an eastward propagation of Storm 4 which does not agree with the observations.

By focusing on 0–3-hr forecasts of UH greater than $60 \text{ m}^2 \text{ s}^{-2}$, we find that all experiments generate a long swath of probabilities $>50\%$ associated with Storm 4 (Figure 9). The high probability values in all experiments initially correspond well with the observed mid-level azimuthal shear track, but they show a displacement at later times due to the faster storm propagation in the ensemble forecasts for the southernmost storm (Storm 4). Because of the dispersive distribution of convection among the ensemble members, Control produces a wider swath of probability for Storm 4 compared to Retro and RetroFED (Figures 9a, 9d, and 9g). However, there is no probability swath near Storm 1, and a rather low-probability (10%–20%) swath further south of Storm 3 that was associated with several tornadoes. While Retro does not generate any UH probability swath close to the observed track of Storm 3 at

2030 UTC, it does forecast high UH probabilities greater than 80% associated with Storm 1. This storm did not produce any tornadoes, but did produce a few hail events at a later time (Figure 9b). Retro also generates two separate UH probability swaths associated with Storm 3 for the forecasts initiated at 2100 and 2130 UTC, but the maximum probability values for these two swaths do not exceed 70% for Storm 3 (Figures 9e and 9h). With respect to RetroFED, all three UH probability swaths are generated (Figures 9c, 9f, 9i). The southern two swaths (i.e., Storm 1 and Storm 4) match well with corresponding azimuthal shear tracks, hail and tornado reports (not shown). Similar to Control and Retro, a fast-motion bias of Storm 4 still exists in RetroFED. The swath of Storm 3 has a maximum UH probability greater than 60% in the forecast from 2030 UTC, and around 80% in the forecasts from 2100 to 2130 UTC. The only issue of the RetroFED forecast is that the UH swath is slightly south biased, as compared to the observations. Later forecasts with more assimilation cycles fail to correct this spatial displacement error.

To understand the reason for improvements to the prediction of the Kansas supercells in RetroFED, an assessment of the impact of assimilating radar and satellite observations in Retro and additional FED observations in RetroFED is needed. The analyzed ensemble mean precipitable water (PW) valid at 2030 UTC shows several significant differences in the vertically accumulated moisture (Figure 10). Both Retro and RetroFED successfully create moist air with $\text{PW} > 1.5$ inches at the locations of the ongoing convection at the analysis time, but Control only creates such an amount of PW for storm 4, and its coverage is relatively small. A band with $\text{PW} > 1.3$ inches lying along a southwest-northeast direction is seen in all experiments. However, PW values between Storm 1 and Storm 3 are somewhat decreased in Retro. Since water vapor is concentrated at lower levels, the boundary layer inflow from the south and southeast (see hodograph in Figure 11) transports less water vapor and less latent heat into the area where Storm 3 should be initiated, reducing the probability of generating strong convection.

To further evaluate the environment in which Storm 3 initiated, skew-T diagrams at 2030 UTC just south of Storm 3 are analyzed for all three experiments (Figure 11). The skew-T profiles represent the ensemble mean of the WoFS forecast and are averaged over a 30×30 -km area centered on the location indicated by the red dots in Figure 10. Compared with Control, the dewpoint temperature below 650 hPa in Retro is overall lower as expected (Figure 11b vs. Figure 11a). At the same time, the assimilation of FED alleviates this problem to some degree, and even adds more water vapor to the layer between 500 and 700 hPa. In addition to changes in moisture, the air temperature in Retro has also changed. In Retro, the lapse rate between the surface and

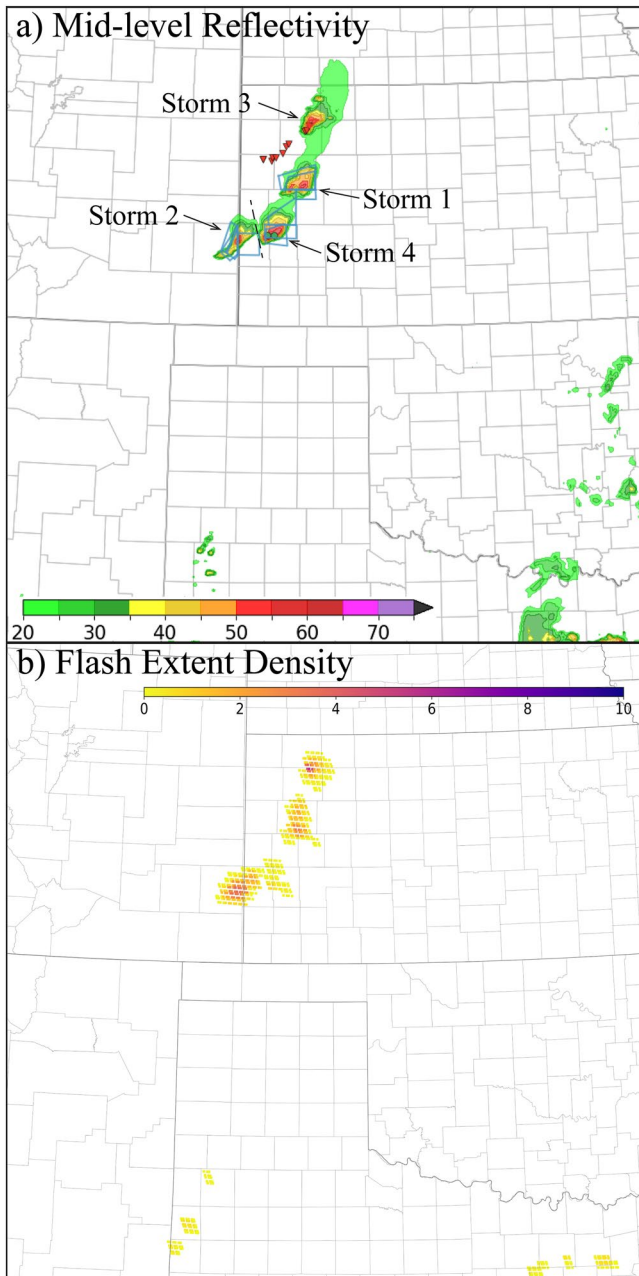


Figure 7. Multi-Radar Multi-Sensor mid-level reflectivity and flash extent density valid at 2200 UTC for 24 May 2021. Storm Prediction Center's severe weather event reports (red triangle for tornadoes, green dot for hail) and warning (red polygon for tornadoes and blue polygon for severe thunderstorms) issued by NWS at this time are shown.

800 hPa is clearly decreased compared with Control. This extra warmth aloft raises the level of free convection from 850 to 800 hPa in Retro, increases the convection inhibition value of the lowest 100-hPa mixed layer from -25.2 to -70.1 J kg^{-1} , decreases the convective available potential energy of the lowest 100-hPa mixed layer from 2006.4 to 972.3 J kg^{-1} , thereby creating a more stable boundary layer. Assimilation of FED observations is able to restore the atmospheric instability to the level seen in Control ($1,899.8 \text{ J kg}^{-1}$, Figure 11c). As a result, RetroFED likely sustains Storm 3 in its corresponding region because the environment is more favorable to thunderstorm development compared to Retro.

Aside from thermodynamic conditions, dynamic mechanisms are also key factors for triggering strong convection. An examination of the hodographs in Figure 11 show that all experiments produce strong speed and directional shear, which are conditions favorable to producing a rotating updraft. The strong speed shear can tilt a storm and vertically displace the updraft from the downdraft, allowing the storm to persist for a longer lifetime. On the other hand, strong directional shear in the lower troposphere typically generates high values of helicity and is important to tornadogenesis. Although Control has a thunderstorm-favorable environment in terms of instability and wind shear, it places the lifting process in different locations. This is likely the reason why Control produces unorganized tracks among ensemble members, as seen in Figure 8.

To examine the lifting process of Storm 3, vertical cross sections of ensemble mean wind, specific humidity, temperature and vertical vorticity for the DA analysis at 2030 UTC are compared among all experiments. The cross sections in Figure 12 are made along the propagation direction of Storm 3. Obviously, Control fails to generate any concentrated lifting mechanism, including both vertical motion and vertical vorticity both at 2030 UTC (Figure 12a), and within the ensuing forecasts. The specific humidity and temperature remain mostly unchanged in the horizontal direction for Control. Meanwhile, Retro produces weak vertical motion associated with weak vertical vorticity and the large gradient of specific humidity (Figure 12b), likely due to the assimilation of radar observations. Although the assimilation of radar and satellite observations enforces the lifting mechanism at the right location, the surrounding environment unfortunately does not support the development of thunderstorms. The analyzed reflectivity dissipates immediately once the forecast is launched (not shown). By assimilating additional FED observations, RetroFED not only improves the environment, but also strengthens internal vertical motion (Figure 12c). RetroFED also generates more abundant water vapor accompanied by the bulge of isotherms near the center of Storm 3, indicating that warm moist air is being transported upward from near the surface by the updraft. As a result of abundant moisture, moderate instability, strong vertical speed and directional wind shear, and appropriate lifting mechanism, RetroFED eventually gives the best forecast of the Storm 3's rotation object among all experiments (Figures 8 and 9).

5.3. Object-Based Verification

Similar to the sensitivity experiments, the previously discussed differences and the forecast skills among various experiments are evaluated in a quantitative manner using the object-based verification method. Figure 13 shows 30–180-min forecast performance for rotation objects accumulated over all forecasts launched from 2000, 2030, 2100, 2130, and 2200 UTC. The overall quality of rotation object forecasts from Retro and RetroFED is

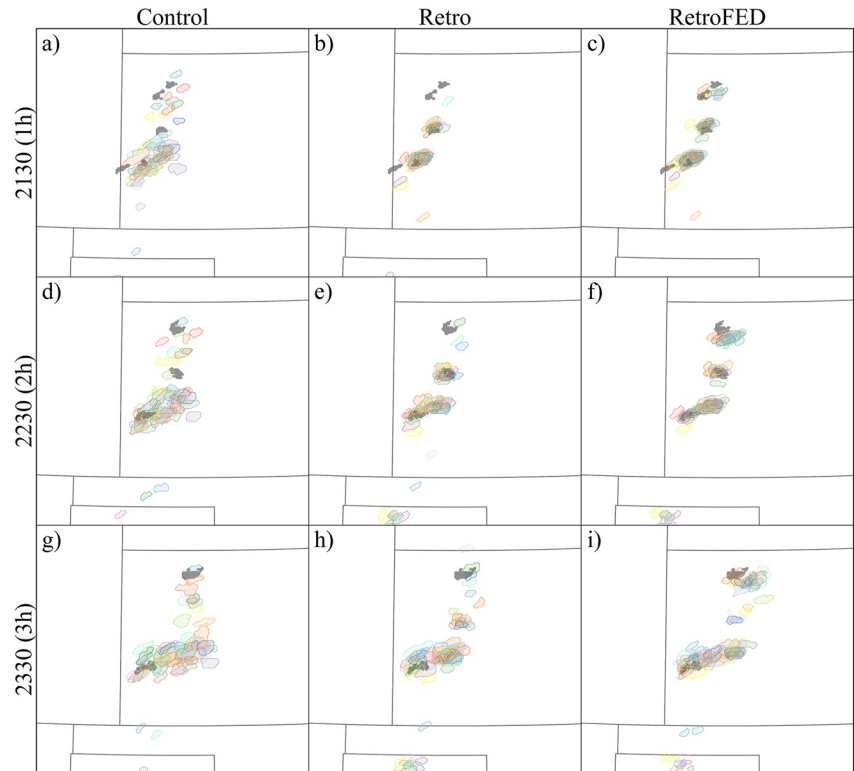


Figure 8. Similar to Figure 6, but for predicted rotation objects for 1–3-hr forecasts initiated at 2030 UTC for three experiments. The left, middle, and right columns correspond to Control, Retro, and RetroFED, respectively.

significantly better than that from Control. The CSI score exceeds 0.55 for the 30-min forecast (Figure 13a) and exceeds 0.5 for the 60-min forecast (Figure 13b), for both Retro and RetroFED. Control generates the lowest skill throughout the whole 180-min forecast period. For Retro and RetroFED, the increase of FAR, with nearly the same level of POD, is the reason for the slight decrease of CSI between the 30- and 60-min forecasts. After the 60-min forecast, the POD values for both Retro and RetroFED also start to decrease as a function of forecast time (Figures 13c–13e). The decreasing rate of POD for RetroFED is approximately 0.05 per 30 min and is faster than that for Retro which is approximately 0.035 per 30 min. However, the rate of increase of FAR for RetroFED is smaller than that for Retro. This evolution of POD and FAR is consistent with the finding shown in the previous section, where Storm 3 is predicted by RetroFED with the slight south bias, and Storm 1 is overpredicted by Retro at a later forecast time. In general, RetroFED outperforms Retro for predicting rotation objects at all forecast times, primarily through the improved forecast for Storm 1 and Storm 3. This indicates that assimilation of FED observations via pseudo dewpoint temperature has the potential to improve the skill of model forecasts.

6. Summary and Conclusions

In this study, a simple DA scheme that indirectly assimilates FED observations into a convection-allowing NWP model is developed and examined. Since the raw 20-s GOES-16 GLM data provides three different matrices “flashes,” “groups,” and “events,” the FED observation is derived by considering the relationship between flashes, groups, and events within a specified period prior to the analysis time. Then, the actual footprint of each GLM event is remapped onto a target grid with a custom resolution. To alleviate possible inaccuracies of the derived FED observations, sensitivity experiments are conducted by using various accumulation windows and spatial resolutions during the data assembly process. The optimal combination of parameters for deriving FED observations is obtained by validating 0–3-hr forecasts of rotation objects. In general, the experiment assimilating dense FED observations (3 km) clearly benefits the rotation forecasts for most of the forecast

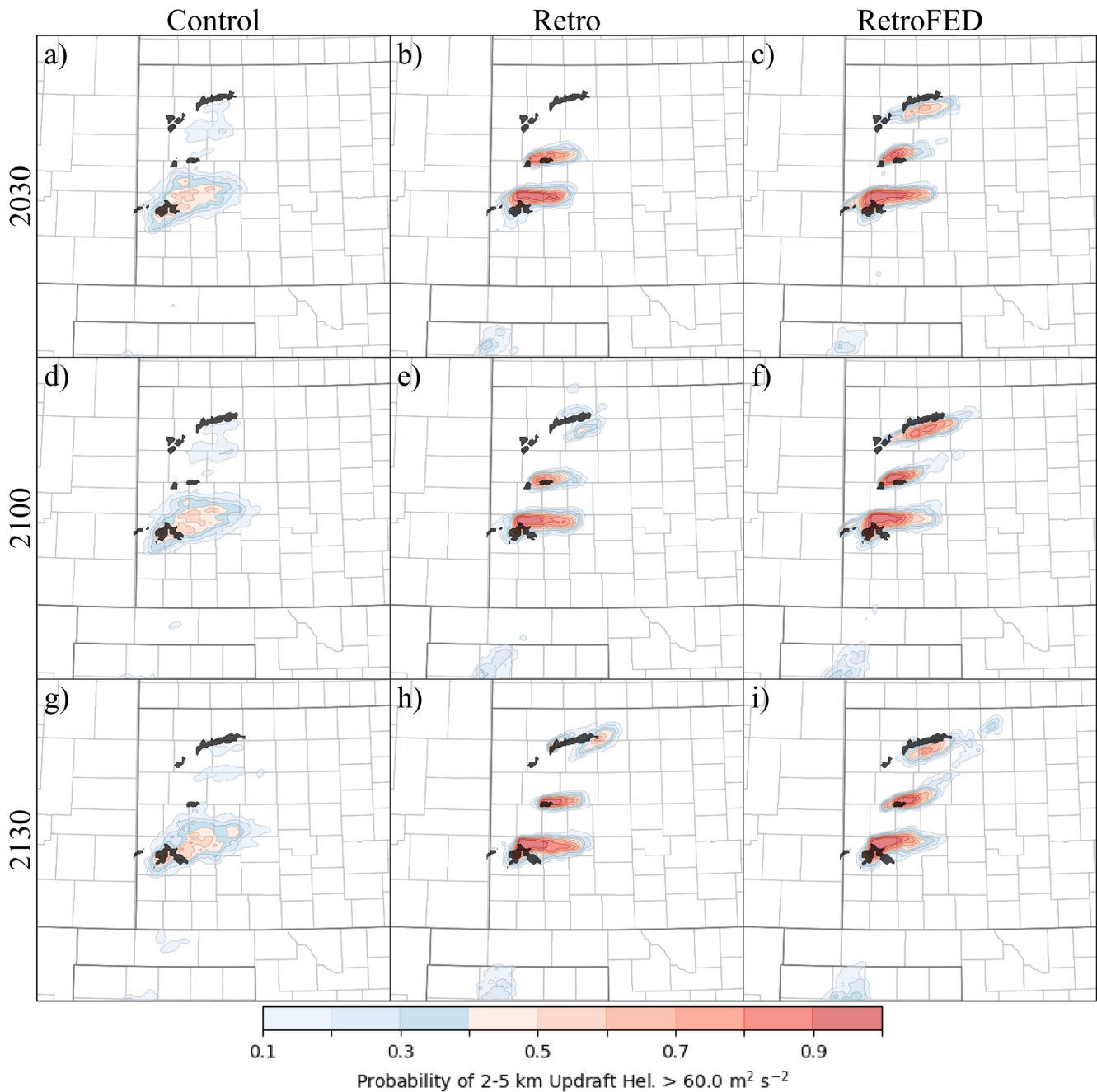


Figure 9. Probability of 2–5 km UH greater than $60 \text{ m}^2 \text{ s}^{-2}$ over 3-hr forecasts initiated at 2030 UTC (a–c), 2100 UTC (d–f), and 2130 UTC (g–i), 24 May 2021 for each experiment (each column). Multi-Radar Multi-Sensor 2–5 km azimuthal shear tracks during this period (black shades) are overlaid over the probability plots.

times, compared to those assimilating sparse observations (6 and 12 km). The combination of a finer resolution of 3 km and a moderate accumulation window of 5 min (i.e., RES_3 km_5 min) gives the best overall forecast performance among the experiments. Thus, this combination is selected for the other experiments conducted in this research.

Additional sensitivity experiments are conducted in which different horizontal and vertical localization radii are applied to FED observations. The results show that the use of narrower horizontal localization and broader vertical localization in the assimilation of FED partially alleviates the displacement errors of rotation objects in the forecast. Additionally, it can generate individual predicted objects that match well with the observed objects in

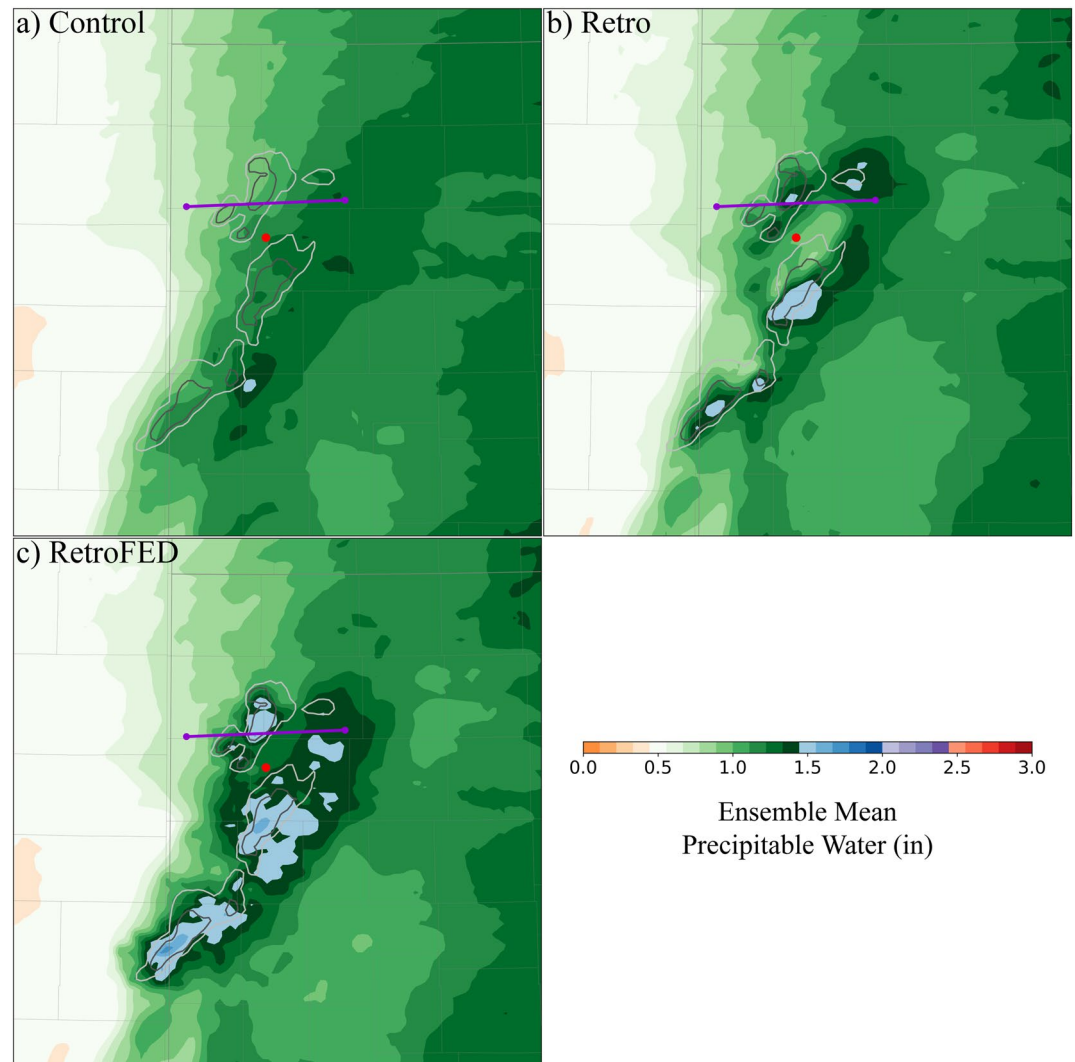


Figure 10. Analyzed ensemble mean precipitable water for (a) Control, (b) Retro, and (c) RedtroFED valid at 2030 UTC, 24 May 2021. The red dot represents the location which a 30×30 -km area is centered for the sounding plot in Figure 11. The purple line represents the location for the cross section plot in Figure 12. The light gray and dark gray contours are for the observed 25 and 40 dBZ composite reflectivity, respectively.

ensemble members. These improvements slightly increase the POD and decrease the FAR. Among the sensitivity experiments, RADII_h6v0.7 uses 6 km and 0.7 as the horizontal and vertical localization radii respectively and produces a slightly better performance for the 0–3-hr forecasts.

Once the optimal parameters for retrieving and assimilating FED observations have been determined, three additional experiments, named as Control, Retro, and RetroFED are performed to assess the value of assimilating FED observations into the WoFS. Control only assimilates conventional observations, while Retro assimilates radar and satellite observations in addition to conventional observations. The third experiment, RetroFED assimilates FED observations in addition to conventional, radar, and satellite observations. By comparing forecasts from Control and Retro, assimilation of radar and satellite data into the WoFS does improve supercell forecasts as compared to the assimilation of only conventional observations. Similar results were reported by Wheatley et al. (2015) and Jones et al. (2016, 2020). Although significant improvements are observed in Retro compared with Control, the forecast can further be improved when additional FED observations are assimilated. Qualitatively, three positive aspects are found in RetroFED. The most important one is the enhanced water vapor

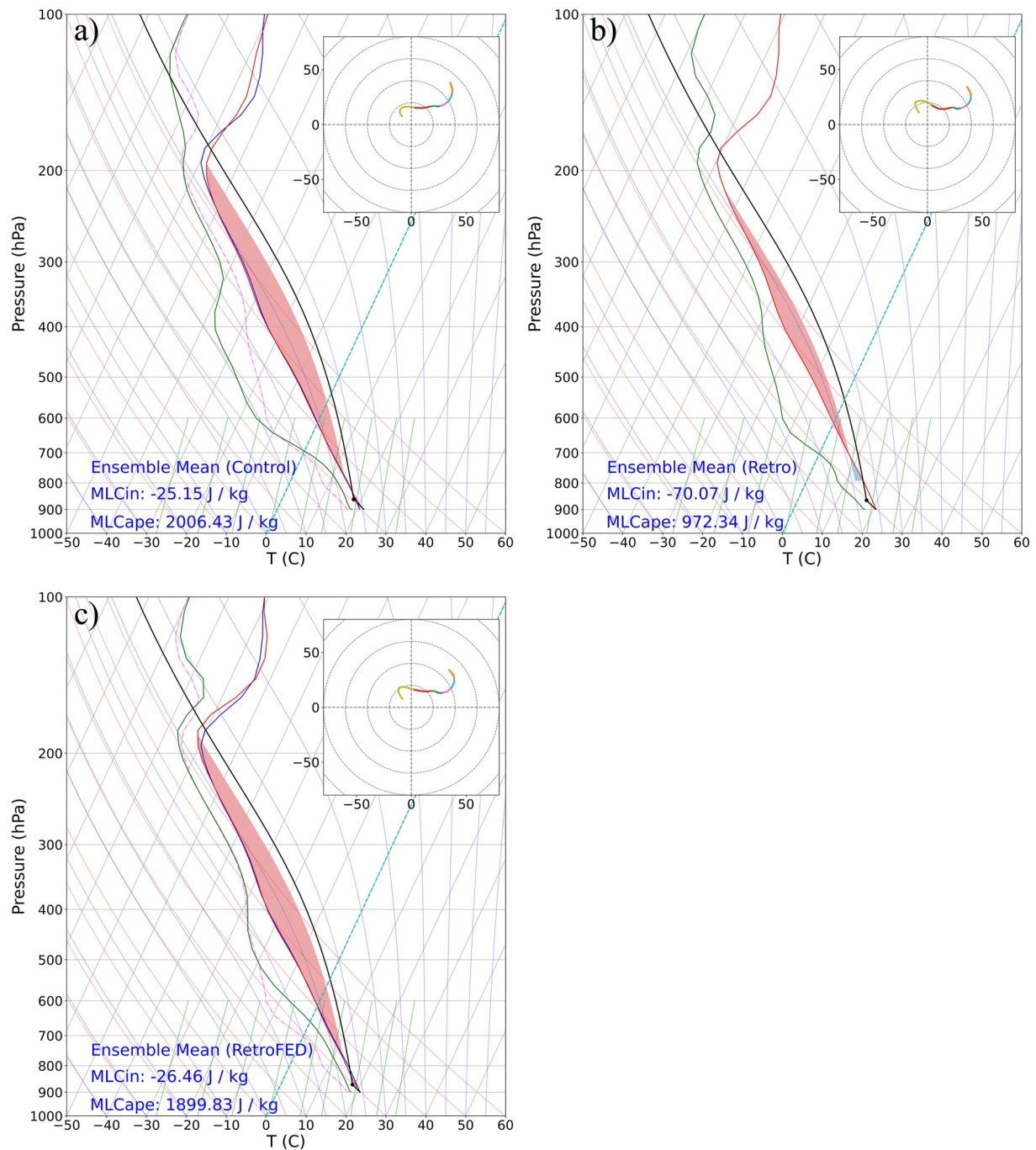


Figure 11. Skew-T profile from analyzed ensemble mean for (a) Control, (b) Retro, and (c) RetroFED valid at 2030 UTC, 24 May 2021. The skew-T profile is averaged over a 30×30 -km area centered at the location indicated by the red dot in Figure 10. For comparison, the temperature and dewpoint profile of Retro experiment are plotted as blue and magenta lines in (a) and (c), respectively.

content that accompanies the boundary layer inflow, which allows that inflow to transport more latent heat into the storm, thus helping to initiate or sustain convection in the appropriate area. The second aspect is related to the indirect adjustment of the lower troposphere temperature. Compared with Retro, RetroFED slightly cools the atmosphere above the surface, resulting in a steeper lapse rate as well as an unstable environment. Finally, the assimilation of FED observations also indirectly intensifies the updraft via flow-dependent error covariances. The rich moisture in the lower troposphere, increased instability and enhanced updraft ultimately leads to a

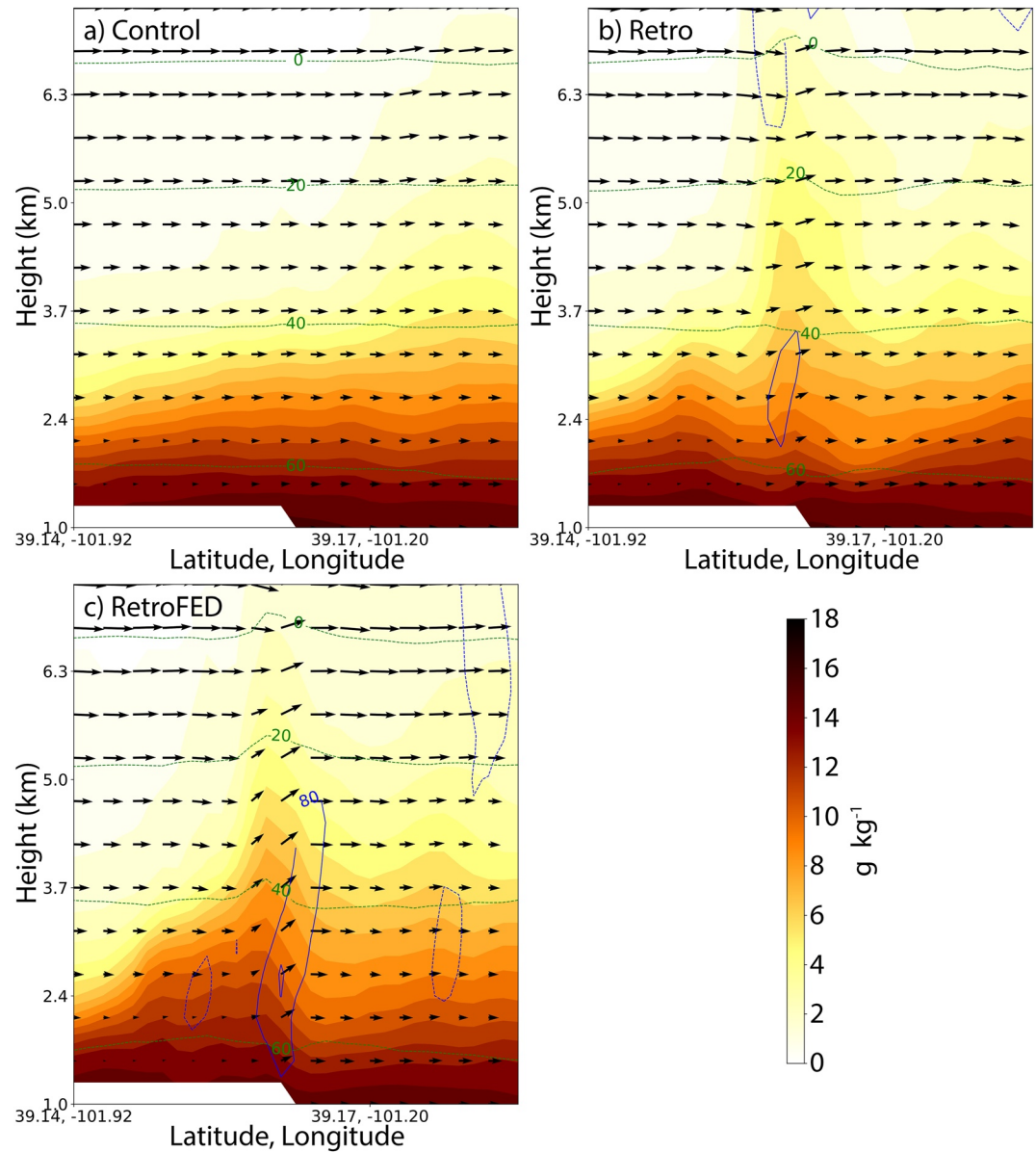


Figure 12. Vertical cross section from analyzed ensemble mean of wind (vectors), specific humidity (color shades, in g kg^{-1}), temperature (green contours, in F), and vertical vorticity (blue contours in s^{-1}) for (a) Control, (b) Retro, and (c) RetroFED valid at 2030 UTC, 24 May 2021. The blue dashed lines represent negative vertical vorticities.

stronger development of the tornadic supercell in northwest Kansas and the best forecast skill for both tornadic and nontornadic supercells.

Overall, this study demonstrates a new method for the potential for assimilation of FED observations, via pseudo dewpoint temperature, to improve the prediction of tornadic supercells on 24 May 2021. However, because our initial focus is on how to derive and use GLM/FED in an optimum way through sensitivity experiments, only one case was examined. Thus the findings in this study are only preliminary. Hazardous weather events may occur in various synoptic- and meso-scale environments. Future work will focus on the inclusion of more cases with different atmospheric conditions, such as frontal systems, dry line systems, and weak forcing situations.

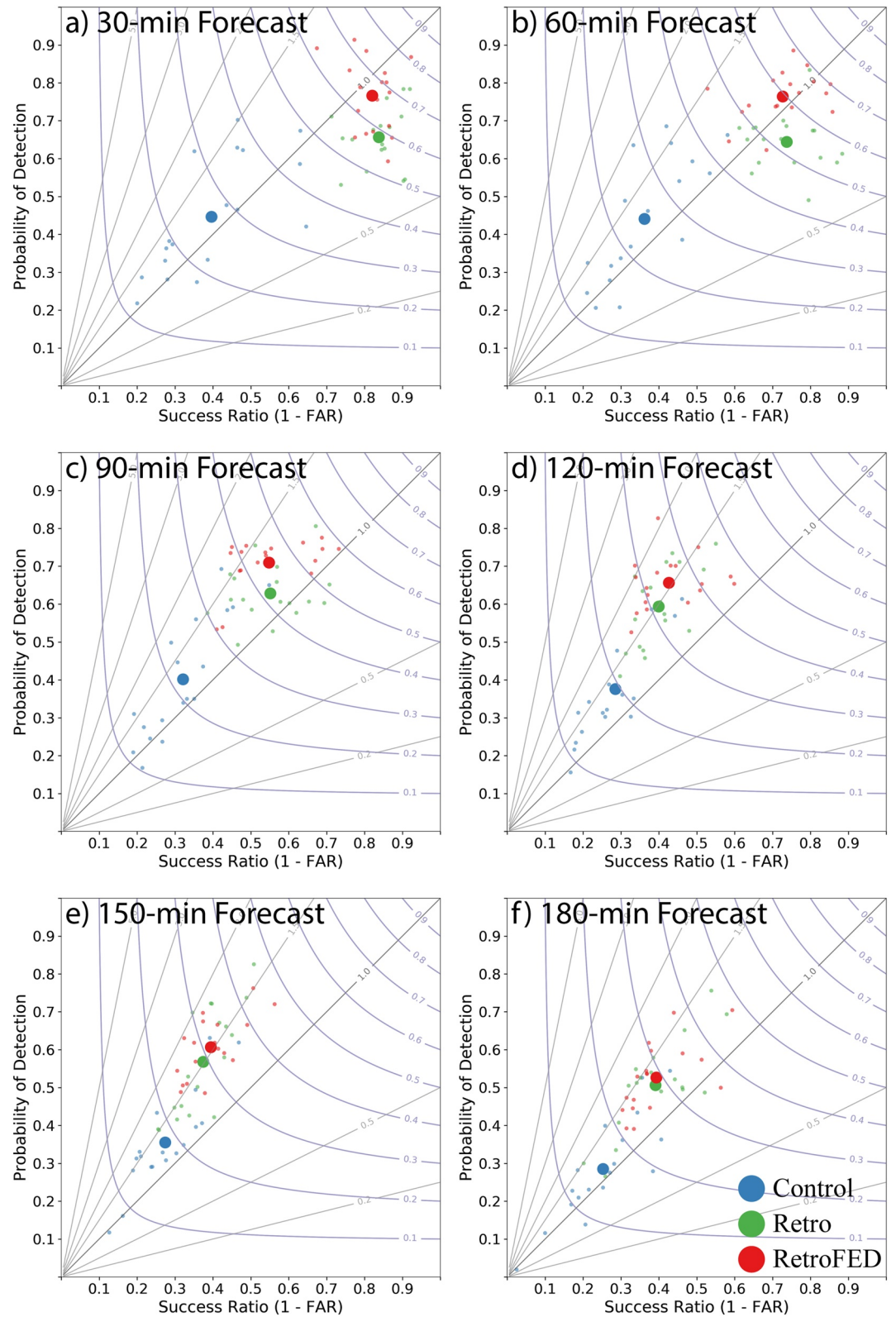


Figure 13. Same as Figure 3, but for experiments Control (blue), Retro (green), and RetroFED (red).

Data Availability Statement

For this work, the (WSR-88D Level-II) data (reflectivity factor and radial velocity) used in this research can be accessed at <http://www.ncdc.noaa.gov/> by filling in locations of radar site and date. The Geostationary Lightning Mapper Level 2 Lightning Detection product (GLM-L2-LCFA) can be accessed on date from Amazon Web Services S3 Bucket (<https://registry.opendata.aws/noaa-goes/>) via the Amazon Command Line Interface (<https://github.com/aws/aws-cli.git>) or can be browsed by selecting the date and time on <https://noaa-goes16.s3.amazonaws.com/index.html#GLM-L2-LCFA/>. The Geostationary Lightning Mapper Tools (*glmtools*) (Bruning et al., 2019) which is used for remapping raw GLM data can be downloaded from <https://github.com/deeplycloudy/glmtools.git>. The community version 1.3 of (GSI-EnKF) data assimilation software can be downloaded from <https://dtcenter.org/community-code/gridpoint-statistical-interpolation-gsi/download>. The (WRF) source code version 3.9 (Skamarock et al., 2008) is publicly available at NCAR/UCAR (<https://github.com/wrf-model/WRF>).

Acknowledgments

Funding was provided by NOAA/ Office of Oceanic and Atmospheric Research under NOAA-University of Oklahoma Cooperative Agreement #NA16OAR4320115 and #NA21OAR4320204, U.S. Department of Commerce. The second author is partially supported by NSF AGS-2136161. We thank Dr. Patrick Skinner for allowing us to use his objective verification program. Discussions with Drs. Cameron Homeyer, Junjun Hu, Thomas Jones, Brian Matilla, Greg McFarquhar, and Xuguang Wang are greatly appreciated for their many kind advice and helpful suggestions. Drs. Jacob Carlin and Pamela Heinselman kindly proofread and edited this manuscript, which led to significant improvement of the manuscript. Computing resources were partially provided by OU Supercomputing Center for Education & Research (OSCAR).

References

- Aksoy, A., Dowell, D. C., & Snyder, C. (2009). A multicase comparative assessment of the ensemble Kalman filter for assimilation of radar observations. Part I: Storm-scale analyses. *Monthly Weather Review*, *137*(6), 1805–1824. <https://doi.org/10.1175/2008MWR2691.1>
- Anderson, J. L. (2009). Spatially and temporally varying adaptive covariance inflation for ensemble filters. *Tellus A*, *61*(1), 72–83. <https://doi.org/10.1111/j.1600-0870.2008.00361.x>
- Benjamin, S. G., Weygandt, S. S., Brown, J. M., Hu, M., Alexander, C. R., Smirnova, T. G., et al. (2016). A North American hourly assimilation and model forecast cycle: The rapid refresh. *Monthly Weather Review*, *144*(4), 1669–1694. <https://doi.org/10.1175/MWR-D-15-0242.1>
- Brock, F. V., Crawford, K. C., Elliott, R. L., Cuperus, G. W., Stadler, S. J., Johnson, H. L., & Eilts, M. D. (1995). The Oklahoma Mesonet: A technical overview [Dataset]. *Journal of Atmospheric and Oceanic Technology*, *12*(1), 5–19. [https://doi.org/10.1175/1520-0426\(1995\)012<0005:TOMATO>2.0.CO;2](https://doi.org/10.1175/1520-0426(1995)012<0005:TOMATO>2.0.CO;2)
- Bruning, E. C., Tillier, C. E., Edgington, S. F., Rudlosky, S. D., Zajic, J., Gravelle, C., et al. (2019). Meteorological imagery for the geostationary lightning mapper [Software]. *Journal of Geophysical Research: Atmospheres*, *124*(24), 14285–14309. <https://doi.org/10.1029/2019JD030874>
- Carey, L. D., & Rutledge, S. A. (1998). Electrical and multiparameter radar observations of a severe hailstorm. *Journal of Geophysical Research*, *103*(D12), 13979–14000. <https://doi.org/10.1029/97JD02626>
- Clark, A. J., Jirak, I. L., Gallo, B. T., Roberts, B., Dean, A. R., Knopfmeier, K. H., et al. (2021). A real-time, virtual spring forecasting experiment to advance severe weather prediction. *Bulletin of the American Meteorological Society*, *102*(4), E814–E816. <https://doi.org/10.1175/BAMS-D-20-0268.1>
- Cressman, G. P. (1959). An operational objective analysis system. *Monthly Weather Review*, *87*(10), 367–374. [https://doi.org/10.1175/1520-0493\(1959\)087<0367:AOOAS>2.0.CO;2](https://doi.org/10.1175/1520-0493(1959)087<0367:AOOAS>2.0.CO;2)
- Crum, T. D., Alberty, R. L., & Burgess, D. W. (1993). Recording, archiving, and using WSR-88D data [Dataset]. *Bulletin of the American Meteorological Society*, *74*(4), 645–654. [https://doi.org/10.1175/1520-0477\(1993\)074<0645:RAAUWD>2.0.CO;2](https://doi.org/10.1175/1520-0477(1993)074<0645:RAAUWD>2.0.CO;2)
- Davis, C., Brown, B., & Bullock, R. (2006). Object-based verification of precipitation forecasts. Part I: Methodology and application to mesoscale rain areas. *Monthly Weather Review*, *134*(7), 1772–1784. <https://doi.org/10.1175/MWR3145.1>
- Deierling, W., & Petersen, W. A. (2008). Total lightning activity as an indicator of updraft characteristics. *Journal of Geophysical Research*, *113*(D16), D16210. <https://doi.org/10.1029/2007JD009598>
- Dowell, D. C., & Wicker, L. J. (2009). Additive noise for storm-scale ensemble data assimilation. *Journal of Atmospheric and Oceanic Technology*, *26*(5), 911–927. <https://doi.org/10.1175/2008JTECHA1156.1>
- Dowell, D. C., Wicker, L. J., & Snyder, C. (2011). Ensemble Kalman filter assimilation of radar observations of the 8 May 2003 Oklahoma City supercell: Influences of reflectivity observations on storm-scale analyses. *Monthly Weather Review*, *139*(1), 272–294. <https://doi.org/10.1175/2010MWR3438.1>
- Fierro, A. O., Gao, J., Ziegler, C. L., Calhoun, K. M., Mansell, E. R., & MacGorman, D. R. (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. *Monthly Weather Review*, *144*(11), 4373–4393. <https://doi.org/10.1175/MWR-D-16-0053.1>
- Fierro, A. O., Wang, Y., Gao, J., & Mansell, E. R. (2019). Variational assimilation of radar data and GLM lightning-derived water vapor for the short-term forecasts of high-impact convective events. *Monthly Weather Review*, *147*(11), 4045–4069. <https://doi.org/10.1175/MWR-D-18-0421.1>
- Flora, M. L., Skinner, P. S., Potvin, C. K., Reinhart, A. E., Jones, T. A., Yussouf, N., & Knopfmeier, K. H. (2019). Object-based verification of short-term, storm-scale probabilistic mesocyclone guidance from an experimental warn-on-forecast system. *Weather and Forecasting*, *34*(6), 1721–1739. <https://doi.org/10.1175/WAF-D-19-0094.1>
- Gagne, D. J., II, Haupt, S. E., Nychka, D. W., & Thompson, G. (2019). Interpretable deep learning for spatial analysis of severe Hailstorms. *Monthly Weather Review*, *147*(8), 2827–2845. <https://doi.org/10.1175/MWR-D-18-0316.1>
- Gan, R., Yang, Y., Qiu, X., Wang, R., Qiu, X., & Zhu, L. (2021). Assimilation of the maximum vertical velocity converted from total lightning data through the EnSRF method. *Journal of Geophysical Research: Atmospheres*, *126*(9), e2020JD034300. <https://doi.org/10.1029/2020JD034300>
- Gao, J., & Stensrud, D. J. (2014). Some observing system simulation experiments with a Hybrid 3D-EnVAR system for storm-scale radar data assimilation. *Monthly Weather Review*, *142*(9), 3326–3346. <https://doi.org/10.1175/MWR-D-14-00025.1>
- Gao, J., Xue, M., Brewster, K., & Droegeimer, K. K. (2004). A three-dimensional variational data analysis method with recursive filter for Doppler radars. *Journal of Atmospheric and Oceanic Technology*, *21*(3), 457–469. [https://doi.org/10.1175/1520-0426\(2004\)021<0457:ATVDAM>2.0.CO;2](https://doi.org/10.1175/1520-0426(2004)021<0457:ATVDAM>2.0.CO;2)
- Goodman, S. J., Blakeslee, R. J., Koshak, W. J., Mach, D., Bailey, J., Buechler, D., et al. (2013). The GOES-R geostationary lightning mapper (GLM) [Dataset]. *Atmospheric Research*, *125*, 34–49. <https://doi.org/10.1016/j.atmosres.2013.01.006>
- Honda, T., Miyoshi, T., Lien, G. Y., Nishizawa, S., Yoshida, R., Adachi, S. A., et al. (2018). Assimilating all-sky Himawari-8 satellite infrared radiances: A case of Typhoon Soudelor (2015). *Monthly Weather Review*, *146*(1), 213–229. <https://doi.org/10.1175/MWR-D-16-0357.1>

- Hu, J., Fierro, A. O., Wang, Y., Gao, J., & Mansell, E. R. (2020). Exploring the assimilation of GLM-derived water vapor mass in a cycled 3DVAR framework for the short-term forecasts of high-impact convective events. *Monthly Weather Review*, *148*(3), 1005–1028. <https://doi.org/10.1175/MWR-D-19-0198.1>
- Hu, J., Gao, J., Wang, Y., Pan, S., Fierro, A. O., Skinner, P. S., et al. (2021). Evaluation of an experimental Warn-on-Forecast 3DVAR analysis and forecast system on quasi-real-time short-term forecasts of high-impact weather events. *Quarterly Journal of the Royal Meteorological Society*, *147*(741), 4063–4082. <https://doi.org/10.1002/qj.4168>
- Hu, J., Yussouf, N., Turner, D. D., Jones, T. A., & Wang, X. (2019). Impact of ground-based remote sensing boundary layer observations on short-term probabilistic forecasts of a tornadic supercell event. *Weather and Forecasting*, *34*(5), 1453–1476. <https://doi.org/10.1175/WAF-D-18-0200.1>
- Johnson, A., Wang, X., Carley, J. R., Wicker, L. J., & Karstens, C. (2015). A comparison of multiscale GSI-based EnKF and 3DVar data assimilation using radar and conventional observations for midlatitude convective-scale precipitation forecasts. *Monthly Weather Review*, *143*(8), 3087–3108. <https://doi.org/10.1175/MWR-D-14-00345.1>
- Jones, T. A., Knopfmeier, K., Wheatley, D., Creager, G., Minnis, P., & Palikonda, R. (2016). Storm-scale data assimilation and ensemble forecasting with the NSSL Experimental Warn-on-Forecast System. Part II: Combined radar and satellite data experiments. *Weather and Forecasting*, *31*(1), 297–327. <https://doi.org/10.1175/WAF-D-15-0107.1>
- Jones, T. A., Otkin, J. A., Stensrud, D. J., & Knopfmeier, K. (2013). Assimilation of satellite infrared radiances and Doppler radar observations during a cool season observing system simulation experiment. *Monthly Weather Review*, *141*(10), 3273–3299. <https://doi.org/10.1175/MWR-D-12-00267.1>
- Jones, T. A., Skinner, P., Yussouf, N., Knopfmeier, K., Reinhart, A., Wang, X., et al., (2020). Assimilation of GOES-16 radiances and retrievals into the warn-on-forecast system [Software]. *Monthly Weather Review*, *148*(5), 1829–1859. <https://doi.org/10.1175/MWR-D-19-0379.1>
- Jones, T. A., Stensrud, D., Wicker, L., Minnis, P., & Palikonda, R. (2015). Simultaneous radar and satellite data storm-scale assimilation using an ensemble Kalman filter approach for 24 May 2011. *Monthly Weather Review*, *143*(1), 165–194. <https://doi.org/10.1175/MWR-D-14-00180.1>
- Kong, R., Xue, M., Fierro, A. O., Jung, Y., Liu, C., Mansell, E. R., & MacGorman, D. R. (2020). Assimilation of GOES-R geostationary lightning mapper flash extent density data in GSI EnKF for the analysis and short-term forecast of a mesoscale convective system. *Monthly Weather Review*, *148*(5), 2111–2133. <https://doi.org/10.1175/MWR-D-19-0192.1>
- Lawrence, M. G. (2005). The relationship between relative humidity and the dewpoint temperature in moist air: A simple conversion and applications. *Bulletin of the American Meteorological Society*, *86*(2), 225–234. <https://doi.org/10.1175/BAMS-86-2-225>
- Lin, X., & Hubbard, K. G. (2004). Uncertainties of derived dewpoint temperature and relative humidity. *Journal of Applied Meteorology and Climatology*, *43*(5), 821–825. <https://doi.org/10.1175/2100.1>
- Liu, C., Cecil, D. J., Zipser, E. J., Kronfeld, K., & Robertson, R. (2012). Relationships between lightning flash rates and radar reflectivity vertical structures in thunderstorms over the tropics and subtropics. *Journal of Geophysical Research*, *117*(D6), 1–19. <https://doi.org/10.1029/2011JD017123>
- MacGorman, D. R., Apostolopoulos, I. R., Lund, N. R., Demetriades, N. W. S., Murphy, M. J., & Krehbiel, P. R. (2011). The timing of cloud-to-ground lightning relative to total lightning activity. *Monthly Weather Review*, *139*(12), 3871–3886. <https://doi.org/10.1175/MWR-D-11-00047.1>
- MacGorman, D. R., Burgess, D. W., Mazur, V., Rust, W. D., Taylor, W. L., & Johnson, B. C. (1989). Lightning rates relative to tornadic storm evolution on 22 May 1981. *Journal of the Atmospheric Sciences*, *46*(2), 221–251. [https://doi.org/10.1175/1520-0469\(1989\)046<0221:LRRRTS>2.0.CO;2](https://doi.org/10.1175/1520-0469(1989)046<0221:LRRRTS>2.0.CO;2)
- MacGorman, D. R., Rust, W. D., Krehbiel, P., Rison, W., Bruning, E., & Wiens, K. (2005). The electrical structure of two supercell storms during STEPS. *Monthly Weather Review*, *133*(9), 2583–2607. <https://doi.org/10.1175/MWR2994.1>
- Mach, D. M. (2020). Geostationary lightning mapper clustering algorithm stability. *Journal of Geophysical Research: Atmospheres*, *125*(5), e2019JD031900. <https://doi.org/10.1029/2019JD031900>
- Mansell, E. R., Ziegler, C. L., & Bruning, E. C. (2010). Simulated electrification of a small thunderstorm with two-moment bulk microphysics. *Journal of the Atmospheric Sciences*, *67*(1), 171–194. <https://doi.org/10.1175/2009AS2965.1>
- Minamide, M., & Zhang, F. (2019). An adaptive background error inflation method for assimilating all-sky radiances. *Quarterly Journal of the Royal Meteorological Society*, *145*(719), 805–823. <https://doi.org/10.1002/qj.3466>
- NOAA National Centers for Environmental Information (NCEI). (2021). Billion-dollar weather and climate disasters. <https://doi.org/10.25911/stkw-7w73>
- Polkinghorne, R., & Vukicevic, T. (2011). Data assimilation of cloud-affected radiances in a cloud-resolving model. *Monthly Weather Review*, *139*(3), 755–773. <https://doi.org/10.1175/2010MWR3360.1>
- Reisner, J. M., & Jeffery, C. A. (2009). A smooth cloud model. *Monthly Weather Review*, *137*(6), 1825–1843. <https://doi.org/10.1175/2008MWR2576.1>
- Rudlosky, S. D., Goodman, S. J., Virts, K. S., & Bruning, E. C. (2019). Initial geostationary lightning mapper observations. *Geophysical Research Letters*, *46*(2), 1097–1104. <https://doi.org/10.1029/2018GL081052>
- Schmit, T. J., Griffith, P., Gunshor, M. M., Daniels, J. M., Goodman, S. J., & Lehair, W. J. (2017). A closer look at the ABI on the GOES-R series [Dataset]. *Bulletin of the American Meteorological Society*, *98*(4), 681–698. <https://doi.org/10.1175/BAMS-D-15-00230.1>
- Schultz, C. J., Petersen, W. A., & Carey, L. D. (2011). Lightning and severe weather: A comparison between total and cloud-to-ground lightning trends. *Weather and Forecasting*, *26*(5), 744–755. <https://doi.org/10.1175/WAF-D-10-05026.1>
- Skamarock, W., Klemp, J., Dudhia, J., Gill, D., Barker, D., Wang, W., et al. (2008). *A description of the advanced research WRF version 3 (No. NCAR/TN-475+STR) [WRF]*. University Corporation for Atmospheric Research. <https://doi.org/10.5065/D68S4MVH>
- Skinner, P. S., Wheatley, D. M., Knopfmeier, K. H., Reinhart, A. E., Choate, J. J., Jones, T. A., et al. (2018). Object-based verification of a prototype warn-on-forecast system. *Weather and Forecasting*, *33*(5), 1225–1250. <https://doi.org/10.1175/WAF-D-18-0020.1>
- Smith, T. M., Lakshmanan, V., Stumpf, G. J., Ortega, K. L., Hondl, K., Cooper, K., et al. (2016). Multi-radar multi-sensor (MRMS) severe weather and aviation products: Initial operating capabilities. *Bulletin of the American Meteorological Society*, *97*(9), 1617–1630. <https://doi.org/10.1175/BAMS-D-14-00173.1>
- Stensrud, D. J., & Gao, J. (2010). Importance of horizontally inhomogeneous environmental initial conditions to ensemble storm-scale radar data assimilation and very short-range forecasts. *Monthly Weather Review*, *138*(4), 1250–1272. <https://doi.org/10.1175/2009MWR3027.1>
- Sun, Y. Q., & Zhang, F. (2016). Intrinsic versus practical limits of atmospheric predictability and the significance of the butterfly effect. *Journal of the Atmospheric Sciences*, *73*(3), 1419–1438. <https://doi.org/10.1175/JAS-D-15-0142.1>
- Wang, Y., & Wang, X. (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. *Monthly Weather Review*, *145*(4), 1447–1471. <https://doi.org/10.1175/MWR-D-16-0231.1>

- Wheatley, D. M., Knopfmeier, K. H., Jones, T. A., & Creager, G. J. (2015). Storm-scale data assimilation and ensemble forecasting with the NSSL experimental warn-on-forecast system. Part I: Radar data experiments. *Weather and Forecasting*, *30*(6), 1795–1817. <https://doi.org/10.1175/WAF-D-15-0043.1>
- Whitaker, J. S., & Hamill, T. M. (2002). Ensemble data assimilation without perturbed observations [GSI-EnKF]. *Monthly Weather Review*, *130*(7), 1913–1924. [https://doi.org/10.1175/1520-0493\(2002\)130<1913:EDAWPO>2.0.CO;2](https://doi.org/10.1175/1520-0493(2002)130<1913:EDAWPO>2.0.CO;2)
- Wiens, K. C., Rutledge, S. A., & Tessendorf, S. A. (2005). The 29 June 2000 supercell observed during STEPS. Part II: Lightning and charge structure. *Journal of the Atmospheric Sciences*, *62*(12), 4151–4177. <https://doi.org/10.1175/JAS3615.1>
- Yano, J.-I., Ziemianski, M. Z., Cullen, M., Termonia, P., Onvlee, J., Bengtsson, L., et al. (2018). Scientific challenges of convective-scale numerical weather prediction. *Bulletin of the American Meteorological Society*, *99*(4), 699–710. <https://doi.org/10.1175/BAMS-D-17-0125.1>
- Yussouf, N., Jones, T. A., & Skinner, P. S. (2020). Probabilistic high-impact rainfall forecasts from landfalling tropical cyclones using Warn-on-Forecast system. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 2050–2065. <https://doi.org/10.1002/qj.3779>
- Yussouf, N., & Knopfmeier, K. H. (2019). Application of the Warn-on-Forecast system for flash-flood-producing heavy convective rainfall events. *Quarterly Journal of the Royal Meteorological Society*, *145*(723), 2385–2403. <https://doi.org/10.1002/qj.3568>
- Zhang, F., Minamide, M., & Clothiaux, E. E. (2016). Potential impacts of assimilating all-sky infrared satellite radiances from GOES-R on convection-permitting analysis and prediction of tropical cyclones. *Geophysical Research Letters*, *43*(6), 2954–2963. <https://doi.org/10.1002/2016GL068468>
- Zhang, F., Minamide, M., Nystrom, R. G., Chen, X., Lin, S.-J., & Harris, L. M. (2019). Improving Harvey forecasts with next-generation weather satellites: Advanced Hurricane analysis and prediction with assimilation of GOES-R all-sky radiances. *Bulletin of the American Meteorological Society*, *100*(7), 1217–1222. <https://doi.org/10.1175/BAMS-D-18-0149.1>
- Zhang, F., Stensrud, D. J., & Zhang, F. (2019). Simultaneous assimilation of radar and all-sky satellite infrared radiance observations for convection-allowing ensemble analysis and prediction of severe thunderstorms. *Monthly Weather Review*, *147*(12), 4389–4409. <https://doi.org/10.1175/MWR-D-19-0163.1>
- Zhang, Y., Zhang, F., & Stensrud, D. J. (2018). Assimilating all-sky infrared radiances from GOES-16 ABI using an ensemble Kalman filter for convection-allowing severe thunderstorms prediction. *Monthly Weather Review*, *146*(10), 3363–3381. <https://doi.org/10.1175/MWR-D-18-0062.1>