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

Jidong Gao[1] and Sijie Pan[2]

[1]NOAA/National Severe Storms Laboratory
[2]Cooperative Institute for Severe and High-Impact Weather Research and Operations, The University of Oklahoma

November 24, 2022

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 (NSSL) Warn-on-Forecast System (WoFS). The new LDA scheme first retrieves the pseudo relative humidity between the cloud base and a specific layer based on the FED value. Then on each model layer, the pseudo relative humidity is converted to dewpoint temperature according to the corresponding air temperature. Some sensitivity experiments are performed to investigate how to derive and use GLM/FED in a best possible way. The impact of assimilating this derived pseudo dewpoint temperature on short-term severe weather forecast is preliminarily assessed in this proof-of-concept study. A high-impact weather event over 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 derived dewpoint temperature observations along with radar, satellite radiance and cloud water can improve short-term (3-h) forecast skill. 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-derived dewpoint temperature, in addition to the other observations currently used in WoFS, tends to enhance the ingredients required for a thunderstorm to occur, namely moisture, instability and lifting mechanism.

Hosted file

`essoar.10511231.1.docx` available at https://authorea.com/users/548762/articles/603233-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

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*]

[1]Cooperative Institute for Severe and High-Impact Weather Research and Operations, The University of Oklahoma,

[2]NOAA/National Severe Storms Laboratory, Norman, Oklahoma

[3]School of Meteorology, University of Oklahoma, Norman, OK 73072

To be submitted to *AGU - ESS*

Key points:

- GOES-16 derived GLM 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.

*Corresponding author: Jidong Gao, NOAA/National Severe Storms Laboratory, 120 David L Boren Blvd, Norman, Oklahoma, 73072; jidong.gao@noaa.gov.

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 (NSSL) Warn-on-Forecast System (WoFS). The new LDA scheme first retrieves the pseudo relative humidity between the cloud base and a specific layer based on the FED value. Then on each model layer, the pseudo relative humidity is converted to dewpoint temperature according to the corresponding air temperature. Some sensitivity experiments are performed to investigate how to derive and use GLM/FED in a best possible way. The impact of assimilating this derived pseudo dewpoint temperature on short-term severe weather forecast is preliminarily assessed in this proof-of-concept study. A high-impact weather event over 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 derived dewpoint temperature observations along with radar, satellite radiance and cloud water can improve short-term (3-h) forecast skill. 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-derived dewpoint temperature, in addition to the other observations currently used in WoFS, tends to enhance the ingredients

required for a thunderstorm to occur, namely moisture, instability and lifting mechanism.

Plain Language Summary

Although the high temporospatial resolution Geostationary Lightning Mapper (GLM) from the newly launched NOAA GOES-16/17 could be beneficial for convective-scale numerical weather prediction, 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 derived dewpoint temperature into the Weather Research and Forecasting model for improving convective scale numerical weather prediction. First, some sensitivity experiments are performed to investigate how to derive and use GLM flash extent density in a best possible way. Then, the potential impacts of GLM on convective scale analysis and short-term severe weather prediction are examined with 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 (Stensrud and Gao 2010; Skinner et al. 2018; Zhang et al. 2019a, b; Clark et al. 2021; Hu et al. 2021). 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 reasons, such as the fidelity of microphysics schemes, that may limit the predictability of convective scale weather events still exist (Sun and Zhang 2016; Yano et al. 2018).

To provide better initial conditions for convection-allowing NWP models (CAMs), data assimilation (DA) methods require observations with higher spatial and temporal resolution. Usually, conventional observations (e.g., surface observing stations) may only provide information about the pre-storm environment. They have little impact on constructing ongoing storm structure, and thus, their impact on CAMs is significantly limited by the sparser temporospatial density of these data. Therefore, the assimilation of radar and geostationary satellite observations with high temporospatial 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 wind and hydrometeors at a resolution of approximately 1 km and a time frequency of 4 – 5 minutes. The

Advanced Baseline Imager (ABI) on board the recently launched 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 frequency (maximum a few minutes).

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

The NSSL's Warn-on-Forecast System (WoFS, Jones et al. 2020), one of the experimental systems that focus 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 Forecast 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 for 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 GLM lightning product is widely recognized as a proxy for systematically identifying 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 m s^{-1} updraft velocity and graupel (MacGorman et al. 1989; Carey and Rutledge 1998; MacGorman et al. 2005; Wiens et al. 2005; Deierling and Petersen 2008; MacGorman et al. 2011). Therefore, assimilating lightning observation into the WoFS is expected to help improve the forecast of severe thunderstorms, especially for the convection initiation phase.

Most previous studies focused on LDA, utilized lightning observations from ground-based platforms to promote convection at the observed lightning locations within convective parameterization schemes by employing empirical relationships. However, this approach is not appropriate for CAMs that can par-

tially 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 (i) the added value of LDA to forecast remains minor in areas with good coverage of radar observations and (ii) LDA typically produces overestimation of >30-dBZ reflectivities or over-prediction biases. This is because the treatment of areal coverage of non-zero flash densities in their LDA scheme does not change significantly with time lapse, resulting in progressive enhancement of convection.

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), attempted to assimilate lightning observations from spaceborne sensors for real data application using an EnKF framework. K20 adopted an empirical relationship between flash extent density (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 (GSI, Kleist et al. 2009) EnKF 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 Gan et al. (2021, G21 hereafter), 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 because of increased convergence and divergence of wind fields in the low and upper levels, respectively.

None of 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: (i) using FED instead of "flash origin density", which does not consider the areal extent of the flashes; (ii) extending the LDA algorithm to use flash rate information to create pseudo- dewpoint temperature observations for the purpose of suppressing overestimation in forecast; (iii) 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 convection with minimal overestimation in NWP models.

Although the new LDA scheme developed here is designed for real-time application for future SFEs, as a proof-of-concept, this study will first focus on a single case study: the 24 May 2021 severe weather event over northwest Kansas, where a large, multi-vortex tornado was not well predicted by the WoFS. This type of severe weather event is different from that in K20 and G21, which are usually considered as MCSs rather than supercells. Our initial focus here is on how to derive and use of GLM/FED in a best possible way through sensitivity experiments. The preliminary results show that the assimilation of FED, using the GSI-EnKF system, improves the short-term forecast of 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 observation. Section 5 describes qualitative and quantitative comparisons of forecast quality for each assimilation experiment under a quasi-real time situation, followed by the conclusions in section 6.

2. Warn-on-Forecast System (WoFS) and LDA scheme

a. Overview

The WoFS system is an on-demand ensemble DA and forecasting system designed to provide guidance of hazardous weather events, such as tornadoes, damage 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 et al. 2019, 2020). The GSI-EnKF system assimilates radar radial velocity and reflectivity data (Johnson et al. 2015; Wang and 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 DA cycles start from 1500 UTC, and end at 0300 UTC with assimilation of available conventional, radar and satellite observations every 15 minutes. All observations are assimilated into the WRF model with a regional domain using 3-km horizontal resolution and 51 vertical levels. The regional domain covers an area of 300 $\times$ 300 grid points (900 km $\times$ 900 km), and its center is 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-hour forecasts from 1400 UTC HRRRE analyses, and the boundary conditions are generated by using forecasts of the analyses of first 9 HRRRE members at 1200 UTC. A 6-h and a 3-h forecast are launched at the top and the bottom of each hour, respectively. For more details about parameterization schemes, the authors suggest audiences to see Skinner et al. (2018); Wheatley et al. (2015); Mansell et al. (2010). A comprehensive description of the inflation methods to maintain diversity between members in the WoFS can be found in Dowell and Wicker (2009); Anderson (2009) and Hu et al. (2019).

b. Description of observations used in the WoFS for SFE 2021

This study evaluates added values of a LDA scheme to the WoFS real-time performance. Therefore, observations used in the WoFS real-time experiments are also adopted in this study. The WoFS system for SFE 2021 assimilates conventional observations hourly, radial velocity and radar reflectivity, CWP and all sky radiance at 15-min intervals. The conventional observations contained in hourly prepbufr files include temperature, dewpoint, winds and pressure available at the surface and upper levels. For real-time application, the prepbufr conventional observations are assimilated using a 15-min lag, e.g., using the conventional data available at 1500 UTC into 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 prepbufr observations.

Radial velocity observations are used within a range of 150 km from the radar and are thinned to a 5-km resolution Cartesian space using Cressman objective analysis (Cressman 1959) prior to DA procedure. For the reflectivity observations, the 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 m and 7.3 m infrared bands, respectively measuring upper- and low-level water vapor content of atmosphere, and CWP that is one of cloud property products from GOES-16/17 representing vertically aggregated information of hydrometeors. These observations are analyzed to the same MRMS 5-km space for cloudy areas and a 15-km space 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, 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 is ranging from 0.8 to 0.85 in natural log pressure $[-\ln(P/P_0)]$. Here P_0 is the reference pressure 1000 hPa.

c. 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 eastern Pacific Oceans. It has a variable pixel pitch that prevents the horizontal resolution of ground samples from exceeding 14 km over the CONUS, and makes most of ground samples having horizontal resolution less than 10 km (Bruning et al. 2019). The detection efficiency of the GLM is highly dependent on the time of day. An expected flash detection efficiency exceeds 70% for daytime and exceeds 90% for nighttime. A detection, also called an event in GLM datasets, is recorded at individual pixels exceeding 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, groups that are located within 16.5 km and occurred within 330 ms from a GLM flash that may contain multiple strokes. 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. Considering unevenly distributed observed events, it finally makes super-resolution remapping of GLM FED onto a finer resolution space (e.g., 3 km resolution), which is superior to regular image remapping technique such as linear interpolation. By applying this method to raw 20-s GLM data, it is possible to derive FED observations with any spatial resolution or accumulated over any time length. 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 a FED is greater than 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) Eq. 8:

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

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

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

The philosophy behind the equation is very similar to Reisner and Jeffery (2009). The Sigmoid function is monotonic, continuous and differentiable and widely used in machine-learning related research (Gagne et al. 2019). This function is used to eliminate possible discontinuities associated with step function to prevent 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 Fig. 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 to 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. 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 approximate to 0 and is approaching maximum for FED greater than 4 flashes per unit area per minute (hereafter fl area^{-1} min^{-1}). The derived pseudo RH value is 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 limited spurious convection will be generated. The observation error of derived dewpoint temperature was described by Lin and Hubbard (2004).

3. Experimental Design and Verification

a. Experiment configuration

The experiment configurations are the same as the 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. Little 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 could 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 different spatial resolution, accumulation window, and horizontal and vertical localization radius are considered by this study (Table 1). These sensitivity experiments only assimilate GLM FED observations using the WoFS configuration described above.

The first set of experiments are named RES_12km_5min, RES_6km_5min, RES_3km_5min, RES_3km_1min, RES_3km_3min, and RES_3km_10min, respectively, based on spatial resolution and accumulation window used for FED derivation. The frequency of the FED observations valid at 2030 UTC is also shown in Fig. 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 by the time. Additionally, the number of FED > 3 fl area^{-1} min^{-1} in RES_3km_1min is much less than that in RES_3km_3min and RES_3km_5min due to information loss using such a short accumulation window. The optimal resolution and accumulation window is determined by the object-based verification method (Skinner et al. 2018; Flora et al. 2019) that is briefly described in the following subsection.

The second set of sensitivity experiments examine the impact of the localization radius for GLM FED assimilation. The horizontal radius ranges from 6 km 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_3km_5min that has the best performance in the first set of sensitivity experiments.

The final set of three experiments (Table 2) are performed to evaluate any value added 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 prepbufr file and Oklahoma Mesonet, radar radial velocity and reflectivity observations, and GOES-16 CWP, 6.2 m and 7.3 m infrared radiances. This experiment acts as a retrospective run to match the performance of the WoFS system in its real-time runs. The final experiment is called "RetroFED". It assimilates pseudo dewpoint temperature observations derived from FED observations, in addition to satellite, radar and conventional observations. In section 4, we further describe the analyses and 3-h forecasts initialized at 2030 UTC, as well as object-based verification.

b. 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 rotations that are highly related to severe weather were identified as objects and were used for evaluation. Similar to S18, single thresholds based on the 99.95$^{\text{th}}$ percentile value in both the forecasts of SFE 2021 and MRMS verification dataset 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 m^2 s^{-2}. Objects are created every 5 minutes starting from 15 minutes into the forecast by collecting ± 15 minutes 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 Eq. (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, which the maximum searching distance is 40 km and the maximum searching time is 15 minutes. The matching procedure allows predicted objects to be classified as four performance metrics "hits", "misses", "false alarms", and "correct nulls", allowing calculation of contingency table statistics for comparing forecast skills between experiments.

4. Sensitivity Experiments

a. 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 higher resolution should not significantly change the FED values but it will change the observation density (as shown in Fig. 2 a – c). 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 FED observations. This conclusion is proved by comparing the FED observations accumulated over different time lengths (Fig. 2a, d – f). Compared with RES_3km_5min, RES_3km_10min provides a wider coverage and a smaller peak value (blue circle in Fig. 2d vs 2a). On the other hand, RES_3km_1min provides a limited coverage and a smaller peak value (blue circle in Fig. 2f vs 2a). The FED observations of RES_3km_5min and RES_3km_3min (Fig. 2a vs 2e) are comparable.

While the qualitative comparison of FED observations using different derivation strategies shows significant differences, it is important to quantify the different impact of assimilation of these observations and to determine the skill of subsequent forecasts. The FED observations are assimilated into the WoFS with 15-min assimilation cycles since 1800 UTC, followed by 3-h 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 val-

ues of POD and CSI are to 1, the better the forecast. The perfect prediction score is located in the upper right corner of the chart. Figure 3 shows 30 – 180-min aggregated forecast performance for each experiment that assimilates FED with various spatial resolutions. The overall quality of 1-h rotation forecasts from experiments assimilating observations with various resolutions is fairly reasonable, as evidenced by CSI values exceeding 0.3, except RES_12km_5min. RES_3km_5min performs better than RES_6km_5min and RES_12km_5min in terms of CSI or frequency bias for the first 2h forecast (Fig. 3 a – d). On average, RES_3km_5min produces CSI less than 0.2 and frequency bias approximately 1.2 comparable to RES_6km_5min and RES_12km_5min for the forecast later than 2.5h (Fig. 3 e – f), although RES_12km_5min produces some individual forecasts associated with CSI greater than 0.3. The performance resulting from the assimilation of observations with different accumulation windows is mixed. The differences among these experiments are relatively small and the skills are gradually decreasing as a function of forecast time (Fig. 4). RES_3km_5min performs best at the 1.5h (Fig. 4c) and 2.5h (Fig. 4e), while RES_3km_3min outperforms RES_3km_5min at 0.5h (Fig. 4a) and 2h (Fig. 4d) and 3h (Fig. 4f). RES_3km_1min generally performs the worst among all experiments due to the narrow accumulation window and the limited coverage. Overall, RES_3km_3min or RES_3km_5min performs better against the other two experiments. Thus, 3-km spatial resolution and 5-min accumulation window will be selected as the optimal parameters for generating FED observations in the following experiments.

b. 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 km, 9 km and 12 km is combined with each vertical radius selected from 0.7, 0.85 and 1.0 in the log pressure unit. So a total of 9 experiments can be performed. But for simplicity, only 5 experiments are discussed here, named as RES_3km_5min (that uses 9 km and 0.85), RADII_h12v1.0, RADII_h12v0.7, RADII_h6v1.0 and RADII_h6v0.7, so that the impact of changing horizontal and vertical localization on the forecast can be easily separated.

By comparing experiments using the same vertical localization (Fig. 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 (Fig. 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-h forecast launched from 2000 UTC (Fig. 6) provides an example of how different localization combinations affect short-term forecast of convection. Among the five experiments, improvements of POD and FAR are observed when a smaller horizontal and a larger vertical localization are applied. Objects associated with

south bias are partially corrected and an additional object is generated (red box in Fig. 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

a. The event overview

Four long-lived supercells developed in west Kansas on 24 May 2021. Composite radar reflectivity at 2200 UTC shows all these supercells, along with warnings issued and severe weather reports collected by National Weather Service (Fig. 7a). The first supercell appeared in central west 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 to 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 started to develop between the Storm 1 and the Storm 2 since 2030 UTC, split around 2215 UTC and was continually developing and propagating to the south until 0500 UTC. At 2200 UTC, Storm 3 already produced several tornadoes and was associated with reflectivity greater than 50 dBZ. Corresponding FED observations at 2200 UTC (Fig. 7b) show that Storm 3 generated the highest FED values (~ 8 fl area^{-1} min^{-1}) compared to the other thunderstorms (~ 6 fl area^{-1} min^{-1}). It is usually true that more frequent flash events can be regarded as an indicator of strong convection.

b. Comparisons of the experiment results

By 2030 UTC 24 May, Storm 3 was well-organized and produced multiple tornadoes in the past half hour. It continued to produce tornadoes until 2330 UTC, followed by its dissipating stage. However, the realtime WoFS run initiated from 2000 and 2030 UTC failed to predict this long-lived supercell. The forecasts launched after 2100 UTC also claimed a rather low probability less than 40% of generating strong rotation. 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 3-h 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 three experiments. Since Storm 2 entered its dissipating stage after 2200 UTC, it is not discussed in the following paragraph.

Obvious differences between Control and Retro can be found in the forecast of Storm 1, Storm 3 and Storm 4. For the first hour prediction, the Control experiment predicts widely distributed rotation objects in west Kansas (Fig. 8a). Some ensemble members place the supercells between observed Storm 1 and Storm 3/4, and some ensemble members place the supercells further south to

observed Storm 4 (Fig. 8a). Although Control successfully predicts convection near Storm 3 by 2130 UTC in 1/3 of the ensemble members, the locations are dispersedly distributed and deviate from the observed object to a certain degree. Retro generates a more concentrated distribution of Storm 1 and Storm 4 around their actual locations, but appears to miss the Storm 3 in most of the ensemble members (Fig. 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 (Fig. 8c). Compared to the observed object, the coverage of Storm 3 in this forecast is further extended eastward. Transitioning to the second and the third hour forecasts, all experiments exhibit similar biases as the 1-h forecast. Control still generates chaos ensemble forecasts (Fig. 8 d, g), while Retro generates discrete convections surrounding Storm 3 for 3 – 6 ensemble members (Fig. 8 e, h). The members of RetroFED that successfully predict Storm 3 place the convection to the south of the observed object at 2230 UTC (Fig. 8f), and a little east to the object at 2330 UTC (Fig. 8i). Compared to the other experiments, RetroFED also produces fewer instances of Storm 1 at 2330 UTC, which should start to dissipate from 2300 UTC. It is worth noting that all experiments predict an eastward propagation of Storm 4 which is not agreed with the observation.

By focusing on 0 – 3-h forecasts of UH greater than 60 m^2 s^{-2}, we find that all experiments generate a long swath of greater 50% probabilities associated with Storm 4 (Fig. 9). The high probability values in all experiments initially correspond well with the observed mid-level azimuthal shear track, but have a misplacement at a later time 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 (Fig. 9 a, d, g). However, there is no probability swath near Storm 1, and a rather low-probability (10% - 20%) swath existing south to the 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 (Fig. 9b). Retro also generates two separated UH probability swaths associated with Storm 3 for the forecasts initiated at 2100 and 2130 UTC. However, the maximum probability values for these two swaths do not exceed 50% (Fig. 9 e, h). With respect to RetroFED, all three UH probability swaths are generated (Fig. 9 c, f, i). 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, fast movement of Storm 4 still exists in RetroFED. The swath of Storm 3 has a maximum probability greater than 60% in the forecast from 2030 UTC, and around 80% in the forecasts from 2100 and 2130 UTC. The only defect of RetroFED is that the swath is slightly south biased, 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 Retro FED is needed. The analyzed ensemble mean precipitable water valid at 2030 UTC shows several significant differences in the vertically accumulated moisture (Fig. 10). Both Retro and RetroFED successfully create moist air with precipitable water greater than 1.5 inches at the locations of the ongoing convection at the analysis time. Conversely, Control only creates such an amount of moisture for storm 4, and its coverage is relatively small. A water vapor band with more than 1.3 inches precipitable water in southwest-northeast direction is shown in all experiments. However, the water vapor content values between Storm 1 and Storm 3 are somewhat decreased in Retro. Since water vapor is mostly concentrated at lower levels, the boundary layer inflow from south and southeast (see hodograph in Fig. 11) transports less water vapor and less latent heat into the area that Storm 3 should be initiated, leading to very low probability of generating strong convection.

To further evaluate the environment in which Storm 3 initiated, the skew-T diagrams at the location at 2030 UTC just south to the Storm 3 are analyzed for all three experiments (Fig. 11). Compared with Control, the dewpoint temperature in the troposphere in Retro is overall lower as expected, especially at the layer between 500 hPa and 700 hPa and the layer around 800 hPa (Fig. 11b vs 11a). At the same time, the assimilation of additional FED alleviates this problem to some degree, and even adds more water vapor to these layers. In addition to changes in moisture, the air temperature in Retro has also changed. In Retro, the air temperature between the surface and 800 hPa is clearly increased compared with Control. Increased temperature above surface raises the level of free convection from 850 hPa to 800 hPa in Control, increases the convection inhibition value from -24.3 J kg^{-1} to -79.1 J kg^{-1}, decreases the convective available potential energy from 1996.7 J kg^{-1} to 1116.4 J $kg^{-}1$, and eventually leads to a more stable boundary layer. Assimilation of additional FED observations is able to restore the atmospheric instability to the level in Control (Fig. 11c). As a result, RetroFED likely sustains Storm 3 in its corresponding region owing to an environment more favorable to thunderstorm development compared to the other two experiments.

Beside thermodynamic conditions, dynamic mechanisms are also the key factors for triggering strong convection. An examination of the hodographs in Fig. 11 show that all experiments produce strong wind speed shear and strong wind directional shear, which are conditions favorable to producing a rotating updraft. The strong speed shear can tilt a storm and vertically displace updraft from downdraft, allowing the storm to sustain 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 dispersedly places the lifting process in different locations. This is the reason why Control produces unorganized tracks among ensemble members, as seen in

Fig. 8.

To examine the lifting process of Storm 3, vertical cross sections of wind, specific humidity, temperature and vertical vorticity for the DA analysis at 2030 UTC are compared among all experiments. The cross sections in Fig. 12 are made along the propagation direction of Storm 3. Obviously, Control fails to generate any lifting mechanism, including both vertical motion and vertical vorticity at 2030 UTC (Fig. 12a), and within the ensuing forecasts. The specific humidity and temperature are rarely changed in horizontal direction for Control. Retro somehow produces weak vertical motion associated with weak vertical vorticity and large gradient of specific humidity (Fig. 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, unfortunately, the surrounding environment does not support the development of thunderstorms. The analyzed reflectivity will dissipate 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 (Fig. 12c). RetroFED also generates more abundant water vapor accompanied with the bulge of isotherms near the center of Storm 3, indicating that warm moist air is being transported from surface to upper air by the updraft. As a result of abundant moisture, moderate instability condition, strong vertical speed and directional wind shear, and appropriate lifting mechanism, RetroFED eventually gives the best forecast of rotation object among all experiments (Fig.8, 9).

c. 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 by using the object-based verification method. Figure 13 shows 0.5 – 3-h 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 significantly better than that from Control. The CSI score exceeds 0.55 for 0.5-h forecast (Fig. 13a) and exceeds 0.5 for 1-h forecast (Fig. 13b), for both Retro and RetroFED. Control generates the lowest skill throughout the whole 3-h 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 0.5 and 1-h forecasts. After the 1-h forecast, the POD values for both Retro and RetroFED also start to decrease as a function of forecast time (Fig. 13c - 13e). The decreasing rate of POD for RetroFED is approximately 0.05 per 30 minutes and is faster than that for Retro which is approximately 0.035 per 30 minutes. However, the increasing rate of FAR for RetroFED is slower than that for Retro. This evolution of POD and FAR is consistent with the instance 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 the possible inaccuracy of the derived FED observations, sensitivity experiments are conducted by using various accumulation windows and spatial resolution during the data deriving process. The optimal combination of parameters for deriving FED observations is obtained by evaluating the 0 – 3-h forecast of rotation objects. In general, the experiment assimilating dense FED observations (3 km) clearly benefits the rotation forecast for most of the forecast times, compared to those assimilating sparse observations (6 km and 12 km). The combination of a higher resolution of 3 km and a moderate accumulation window of 5 minutes (i.e., RES_3km_5min) gives the overall best performance of forecast 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 assimilation of FED partially alleviates the misplacement error of rotation object in the forecast. Additionally, it can generate individual predicted objects that match well with the observed objects in ensemble members. These improvements may 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 radius respectively and produces a slightly better performance for 0 – 3-h forecast.

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 compared to 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 abundant water

vapor content that accompanies the boundary layer inflow. It allows the 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 better development of the tornadic supercell in northwest Kansas and the best forecast skill of both tornadic and nontornadic supercells.

Overall, this study demonstrates a new method for the potential for the assimilation of FED observations, via pseudo dewpoint temperature, to improve the prediction of the tornadic supercells on 24 May 2021. However, because our initial focus is on how to derive and use GLM/FED in a best possible way through sensitivity experiments, only one case was examined. So the findings in this study are preliminary. Many 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.

Acknowledgements

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 advices 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 (OSCER).

Data Availability Statement

The source codes of WRF model version 3.9 could be downloaded after filling in the E-mail address (https://www2.mmm.ucar.edu/wrf/users/download/get_source.html). The GLM data can be downloaded from Amazon Web Services (https://registry.opendata.aws/noaa-goes/).

REFERENCES

https://doi.org/10.1175/WAF-D-19-0071.1

https://doi.org/10.1002/qj.4168

https://doi.org/10.1175/2009MWR3027.1

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. Weather Rev.*, **137**, 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**, 72–83, https://doi.org/10.1111/j.1600-0870.2008.00361.x.Benjamin, S. G., and Coauthors, (2016). A North American Hourly Assimilation and Model Forecast Cycle: The Rapid Refresh. *Mon. Weather Rev.*, **144**, 1669–1694, https://doi.org/10.1175/MWR-D-15-0242.1.Brock, F. V., K. C. Crawford, R. L. Elliott, G. W. Cuperus, S. J. Stadler, H. L. Johnson, and M. D. Eilts, (1995). The Oklahoma Mesonet: A technical overview. J. Atmos. Oceanic Technol., 12, 5-19.Bruning, E. C., and Coauthors, (2019). Meteorological Imagery for the Geostationary Lightning Mapper. *J. Geophys. Res. Atmospheres*, **124**, 14285–14309, https://doi.org/10.1029/2019JD030874.Carey, L. D., and S. A. Rutledge, (1998). Electrical and multiparameter radar observations of a severe hailstorm. *J. Geophys. Res. Atmospheres*, **103**, 13979–14000, https://doi.org/10.1029/97JD02626.Clark, A. J., and Coauthors, (2021). A Real-Time, Virtual Spring Forecasting Experiment to Advance Severe Weather Prediction. *Bull. Am. Meteorol. Soc.*, **102**, E814–E816, https://doi.org/10.1175/BAMS-D-20-0268.1.Cressman, G. P., (1959). AN OPERATIONAL OBJECTIVE ANALYSIS SYSTEM. *Mon. Weather Rev.*, **87**, 367–374, https://doi.org/10.1175/1520-0493(1959)087<0367:AOOAS>2.0.CO;2.Crum, T. D., R. L. Alberty, and D. W. Burgess, (1993). Recording, Archiving, and Using WSR-88D Data. *Bull. Am. Meteorol. Soc.*, **74**, 645–654, https://doi.org/10.1175/1520-0477(1993)074<0645:RAAUWD>2.0.CO;2.Davis, C., B. Brown, and R. Bullock, (2006). Object-Based Verification of Precipitation Forecasts. Part I: Methodology and Application to Mesoscale Rain Areas. *Mon. Weather Rev.*, **134**, 1772–1784, https://doi.org/10.1175/MWR3145.1.Deierling, W., and W. A. Petersen, (2008). Total lightning activity as an indicator of updraft characteristics. *J. Geophys. Res. Atmospheres*, **113**, https://doi.org/10.1029/2007JD009598.Dowell, D. C., and L. J. Wicker, (2009). Additive Noise for Storm-Scale Ensemble Data Assimilation. *J. Atmospheric Ocean. Technol.*, **26**, 911–927, https://doi.org/10.1175/2008JTECHA1156.1.——, ——, and C. Snyder, (2011). Ensemble Kalman Filter Assimilation of Radar Observations of the 8 May 2003 Oklahoma City Supercell: Influences of Reflectivity Observations on Storm-Scale Analyses. *Mon. Weather Rev.*, **139**, 272–294, https://doi.org/10.1175/2010MWR3438.1.Fierro, A. O., J. Gao, C. L. Ziegler, K. M. Calhoun, E. R. Mansell, and D. R. MacGorman, (2016). Assimilation of Flash Extent Data in the Variational Framework at Convection-Allowing Scales: Proof-of-Concept and Evaluation for the Short-Term Forecast of the 24 May 2011 Tornado Outbreak. *Mon. Weather Rev.*, **144**, 4373–4393, https://doi.org/10.1175/MWR-D-16-0053.1.——, Y. Wang, J. Gao, and E. R. Mansell, (2019). Variational Assimilation of Radar Data and GLM Lightning-Derived Water Vapor for the Short-Term Forecasts of High-Impact Convective Events. *Mon. Weather Rev.*, **147**, 4045–4069,

https://doi.org/10.1175/MWR-D-18-0421.1.Flora, M. L., P. S. Skinner, C. K. Potvin, A. E. Reinhart, T. A. Jones, N. Yussouf, and K. H. Knopfmeier, (2019). Object-Based Verification of Short-Term, Storm-Scale Probabilistic Mesocyclone Guidance from an Experimental Warn-on-Forecast System. *Weather Forecast.*, **34**, 1721–1739, https://doi.org/10.1175/WAF-D-19-0094.1.Gagne, D. J., S. E. Haupt, D. W. Nychka, and G. Thompson, (2019). Interpretable Deep Learning for Spatial Analysis of Severe Hailstorms, *Mon. Weather Rev.*, **147**, 2827–2845, https://doi.org/10.1175/MWR-D-18-0316.1Gan, R., Y. Yang, X. Qiu, R. Wang, X. Qiu, and L. Zhu, (2021). Assimilation of the Maximum Vertical Velocity Converted From Total Lightning Data Through the EnSRF Method. *J. Geophys. Res. Atmospheres*, **126**, e2020JD034300, https://doi.org/10.1029/2020JD034300.Gao, J., and D. J. Stensrud, (2014). Some Observing System Simulation Experiments with a Hybrid 3DEnVAR System for Storm-Scale Radar Data Assimilation. *Mon. Weather Rev.*, **142**, 3326–3346, https://doi.org/10.1175/MWR-D-14-00025.1.——, M. Xue, K. Brewster, and K. K. Droegemeier, (2004). A Three-Dimensional Variational Data Analysis Method with Recursive Filter for Doppler Radars. *J. Atmospheric Ocean. Technol.*, **21**, 457–469, https://doi.org/10.1175/1520-0426(2004)021<0457:ATVDAM>2.0.CO;2.Honda, T., and Coauthors, 2018: Assimilating All-Sky Himawari-8 Satellite Infrared Radiances: A Case of Typhoon Soudelor (2015). *Mon. Weather Rev.*, **146**, 213–229, https://doi.org/10.1175/MWR-D-16-0357.1.Hu, J., N. Yussouf, D. D. Turner, T. A. Jones, and X. Wang, (2019). Impact of Ground-Based Remote Sensing Boundary Layer Observations on Short-Term Probabilistic Forecasts of a Tornadic Supercell Event. *Weather Forecast.*, **34**, 1453–1476, https://doi.org/10.1175/WAF-D-18-0200.1.Hu, J.,_Fierro, A., Y. Wang, J. Gao, and E. Mansell, (2020). An Evaluation of the Impact of Assimilating GLM-observed total lightning data on short-term forecasts of high-impact convective events. *Mon. Wea. Rev.* **148**, 1005-1028. Hu, J.,_J. Gao**,** Y. Wang, S. Pan, A. Fierro, P. S. Skinner, K. Knopfmeier, E. Mansell, and P. Heiselman, (2021). Evaluation of a Warn-on-Forecast 3DVAR analysis and forecast system on quasi-real-time short-term forecasts of high impact weather events. *Quart. J. Royal Meteorol. Soc.* Johnson, A., X. Wang, J. R. Carley, L. J. Wicker, and C. Karstens, (2015). A Comparison of Multiscale GSI-Based EnKF and 3DVar Data Assimilation Using Radar and Conventional Observations for Midlatitude Convective-Scale Precipitation Forecasts. *Mon. Weather Rev.*, **143**, 3087–3108, https://doi.org/10.1175/MWR-D-14-00345.1.Jones, T. A., J. A. Otkin, D. J. Stensrud, and K. Knopfmeier, (2013). Assimilation of Satellite Infrared Radiances and Doppler Radar Observations during a Cool Season Observing System Simulation Experiment. *Mon. Weather Rev.*, **141**, 3273–3299, https://doi.org/10.1175/MWR-D-12-00267.1.——, D. Stensrud, L. Wicker, P. Minnis, and R. Palikonda, (2015). Simultaneous Radar and Satellite Data Storm-Scale Assimilation Using an Ensemble Kalman Filter Approach for 24 May 2011. *Mon. Weather Rev.*, **143**, 165–194, https://doi.org/10.1175/MWR-D-14-00180.1.——, K. Knopfmeier, D. Wheatley, G. Creager, P. Minnis, and R. Palikonda, (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 Forecast.*, **31**, 297–327, https://doi.org/10.1175/WAF-D-15-0107.1.——, and Coauthors, (2020). Assimilation of GOES-16 Radiances and Retrievals into the Warn-on-Forecast System. *Mon. Weather Rev.*, **148**, 1829–1859, https://doi.org/10.1175/MWR-D-19-0379.1.Kleist, D. T., D. F. Parrish, J. C. Derber, R. Treadon, W.-S. Wu, and S. Lord, (2009). Introduction of the GSI into the NCEP Global Data Assimilation System. *Weather Forecast.*, **24**, 1691–1705, https://doi.org/10.1175/2009WAF2222201.1.Kong, R., M. Xue, A. O. Fierro, Y. Jung, C. Liu, E. R. Mansell, and D. R. MacGorman, (2020). Assimilation of GOES-R Geostationary Lightning Mapper Flash Extent Density Data in GSI EnKF for the Analysis and Short-Term Forecast of a Mesoscale Convective System. *Mon. Weather Rev.*, **148**, 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. *Bull. Am. Meteorol. Soc.*, **86**, 225–234, https://doi.org/10.1175/BAMS-86-2-225.Lin, X., and K. G. Hubbard, (2004). Uncertainties of Derived Dewpoint Temperature and Relative Humidity. *J. Appl. Meteorol. Climatol.*, **43**, 821–825, https://doi.org/10.1175/2100.1.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. Atmospheric Sci.*, **46**, 221–251, https://doi.org/10.1175/1520-0469(1989)046<0221:LRRTTS>2.0.CO;2.——, W. D. Rust, P. Krehbiel, W. Rison, E. Bruning, and K. Wiens, (2005). The Electrical Structure of Two Supercell Storms during STEPS. *Mon. Weather Rev.*, **133**, 2583–2607, https://doi.org/10.1175/MWR2994.1.——, 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. Weather Rev.*, **139**, 3871–3886, https://doi.org/10.1175/MWR-D-11-00047.1.Mach, D. M., (2020). Geostationary Lightning Mapper Clustering Algorithm Stability. *J. Geophys. Res. Atmospheres*, **125**, e2019JD031900, https://doi.org/10.1029/2019JD031900.Mansell, E. R., C. L. Ziegler, and E. C. Bruning, (2010). Simulated Electrification of a Small Thunderstorm with Two-Moment Bulk Microphysics. *J. Atmospheric Sci.*, **67**, 171–194, https://doi.org/10.1175/2009JAS2965.1.Minamide, M., and F. Zhang, (2019). An adaptive background error inflation method for assimilating all-sky radiances. *Q. J. R. Meteorol. Soc.*, **145**, 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/10.25921/stkw-7w73.Polkinghorne, R., and T. Vukicevic, (2011). Data Assimilation of Cloud-Affected Radiances in a Cloud-Resolving Model. *Mon. Weather Rev.*, **139**, 755–773, https://doi.org/10.1175/2010MWR3360.1.Reisner, J. M., and C. A. Jeffery, (2009). A Smooth Cloud Model. *Mon. Weather Rev.*, **137**, 1825–1843, https://doi.org/10.1175/2008MWR2576.1.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.Schmit, T. J., P. Griffith, M. M. Gunshor, J. M. Daniels, S. J. Goodman, and W. J. Lebair, (2017). A Closer Look at the ABI on the GOES-R Series. *Bull. Am. Meteorol. Soc.*, **98**, 681–698, https://doi.org/10.1175/BAMS-D-15-00230.1.Schultz, C. J., W. A. Petersen, and L. D. Carey, (2011). Lightning and Severe Weather: A Comparison between Total and Cloud-to-Ground Lightning Trends. *Weather Forecast.*, **26**, 744–755, https://doi.org/10.1175/WAF-D-10-05026.1.Skamarock, W., J. Klemp, J. Dudhia, D. Gill, D. Barker, W. Wang, X.-Y. Huang, and M. Duda, (2008). A Description of the Advanced Research WRF Version 3 (No. NCAR/TN-475+STR). *Univ. Corp. Atmospheric Res.*, http://dx.doi.org/10.5065/D68S4MVH.Skinner, P. S., and Coauthors, (2018). Object-Based Verification of a Prototype Warn-on-Forecast System. *Weather Forecast.*, **33**, 1225–1250, https://doi.org/10.1175/WAF-D-18-0020.1.Smith, T. M., and Coauthors, (2016). Multi-Radar Multi-Sensor (MRMS) Severe Weather and Aviation Products: Initial Operating Capabilities. *Bull. Am. Meteorol. 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. DOI: Sun, Y. Q., and F. Zhang, (2016). Intrinsic versus Practical Limits of Atmospheric Predictability and the Significance of the Butterfly Effect. *J. Atmospheric Sci.*, **73**, 1419–1438, https://doi.org/10.1175/JAS-D-15-0142.1.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. Weather Rev.*, **145**, 1447–1471, https://doi.org/10.1175/MWR-D-16-0231.1.Wheatley, D. M., K. H. Knopfmeier, T. A. Jones, and G. J. Creager, (2015). Storm-Scale Data Assimilation and Ensemble Forecasting with the NSSL Experimental Warn-on-Forecast System. Part I: Radar Data Experiments. *Weather Forecast.*, **30**, 1795–1817, https://doi.org/10.1175/WAF-D-15-0043.1.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. Atmospheric Sci.*, **62**, 4151–4177, https://doi.org/10.1175/JAS3615.1.Yano, J.-I., and Coauthors, (2018). Scientific Challenges of Convective-Scale Numerical Weather Prediction. *Bull. Am. Meteorol. Soc.*, **99**, 699–710, https://doi.org/10.1175/BAMS-D-17-0125.1.Yussouf, Nusrat, and Kent H. Knopfmeier, (2019). Application of the Warn-on-Forecast system for flash-flood-producing heavy convective rainfall events. *Q. J. R. Meteorol. Soc.,* **145**, no. 723 (May): 2385-2403. https://doi.org/10.1002/qj.3568.Yussouf, Nusrat, Thomas A. Jones, and Patrick S. Skinner, (2020). Probabilistic high-impact rainfall forecasts from landfalling tropical cyclones using Warn-on-Forecast system. *Q. J. R. Meteorol. Soc.,* **146**, no. 730 (March): 2050-2065. https://doi.org/10.1002/qj.3779.Zhang, F., M. Minamide, and E. E. Clothiaux, (2016). Potential impacts of assimilating all-sky infrared satellite radiances from GOES-R on convection-permitting

analysis and prediction of tropical cyclones. *Geophys. Res. Lett.*, **43**, 2954–2963, https://doi.org/10.1002/2016GL068468.——, ——, R. G. Nystrom, X. Chen, S.-J. Lin, and L. M. Harris, (2019). Improving Harvey Forecasts with Next-Generation Weather Satellites: Advanced Hurricane Analysis and Prediction with Assimilation of GOES-R All-Sky Radiances. *Bull. Am. Meteorol. Soc.*, **100**, 1217–1222, https://doi.org/10.1175/BAMS-D-18-0149.1.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. Weather Rev.*, **146**, 3363–3381, https://doi.org/10.1175/MWR-D-18-0062.1.——, D. J. Stensrud, and F. Zhang, (2019b). Simultaneous Assimilation of Radar and All-Sky Satellite Infrared Radiance Observations for Convection-Allowing Ensemble Analysis and Prediction of Severe Thunderstorms. *Mon. Weather Rev.*, **147**, 4389–4409, https://doi.org/10.1175/MWR-D-19-0163.1.

Experiment	Spatial Resolution of GLM FED	Accumulation Window of GLM FED	Horizontal Lo
RES_12km_5min	12 km	5 min	9 km
RES_6km_5min	6 km	5 min	9 km
RES_3km_5min	3 km	5 min	9 km
RES_3km_1min	3 km	1 min	9 km
RES_3km_3min	3 km	3 min	9 km
RES_3km_10min	3 km	10 min	9 km
RADII_h12v1.0	3 km	5 min	12 km
RADII_h12v0.7	3 km	5 min	12 km
RADII_h6_1.0	3 km	5 min	6 km
RADII_h6v0.7	3 km	5 min	6 km

Table 1 Experiment configurations for different sensitivity experiments.

Experiment	Conventional Observations (PrepBUFR+Mesonet)	Radial Velocity + Reflectivity	GOES-16
Control	Y	N	N
Retro	Y	Y	Y
RetroFED	Y	Y	Y

Table 2 Assimilated observation types for three different experiments Control, Retro, RetroFED.

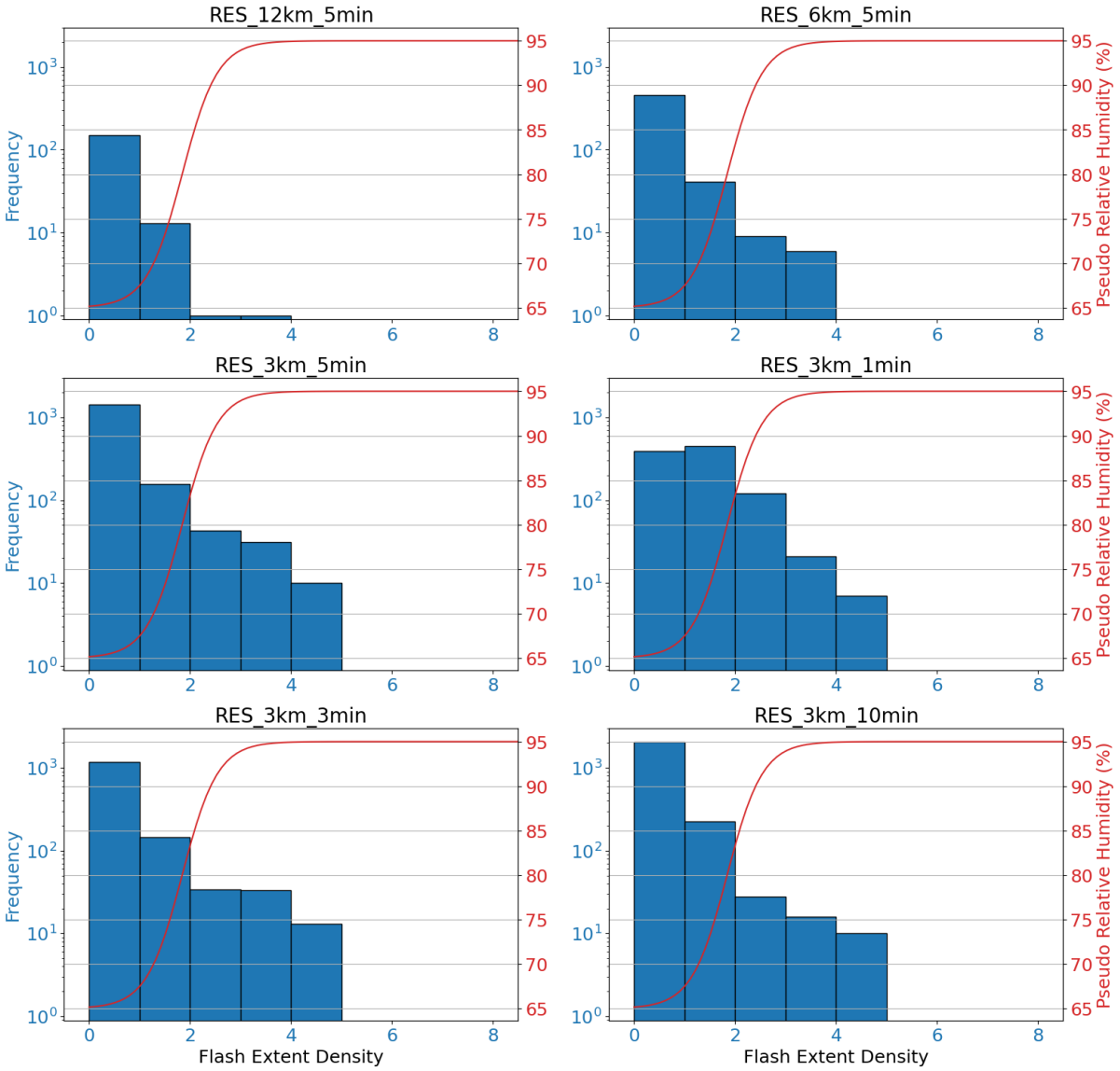

Figure 1 Frequency of the FED (fl area^{-1} min^{-1}) observations using different accumulation strategies (blue bars) at 2030 UTC, 24 May 2021. The red line represents pseudo relative humidity as a function of FED value.

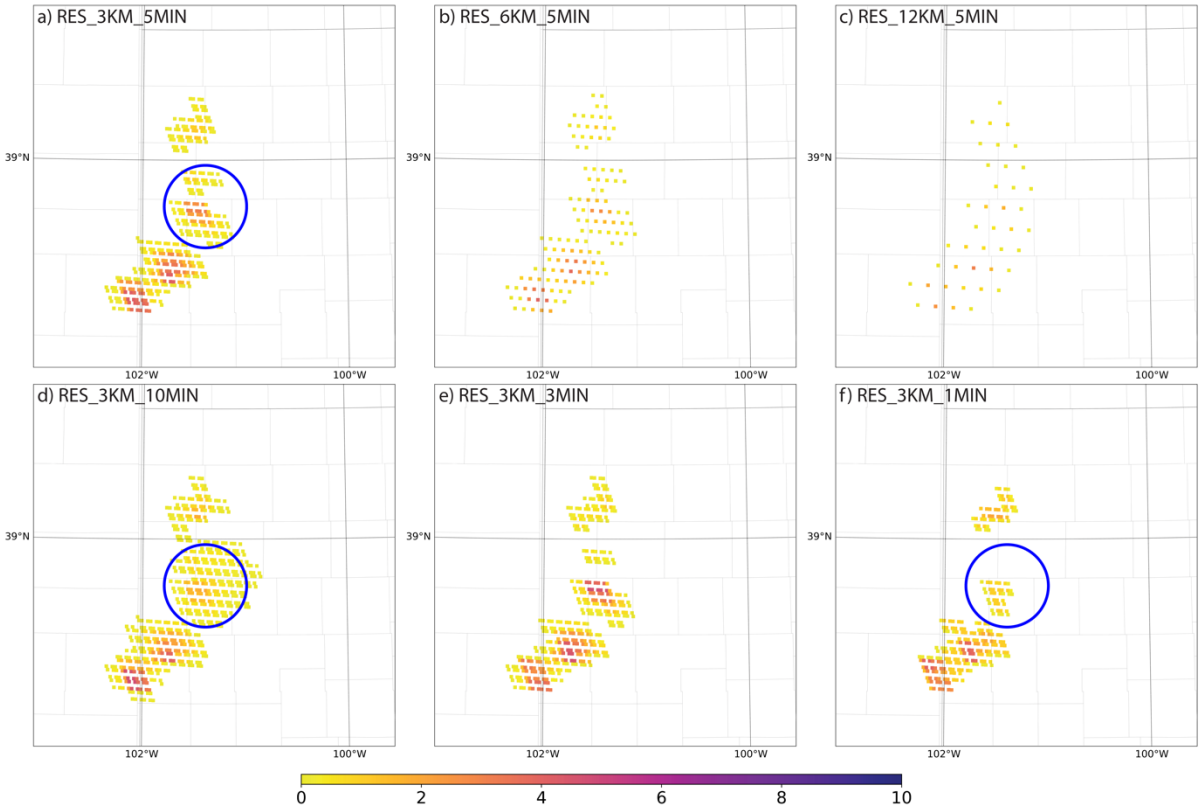

Figure 2 Derived FED observations at 2030 UTC, 24 May 2021 from the raw FED data for (a) 3-km resolution and 5-min window, (b) 6-km resolution and 5-min window, (c) 12-km resolution and 5-min window, (d) 3-km resolution and 10-min window, (e) 3-km resolution and 3-min window, and (f) 3-km resolution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

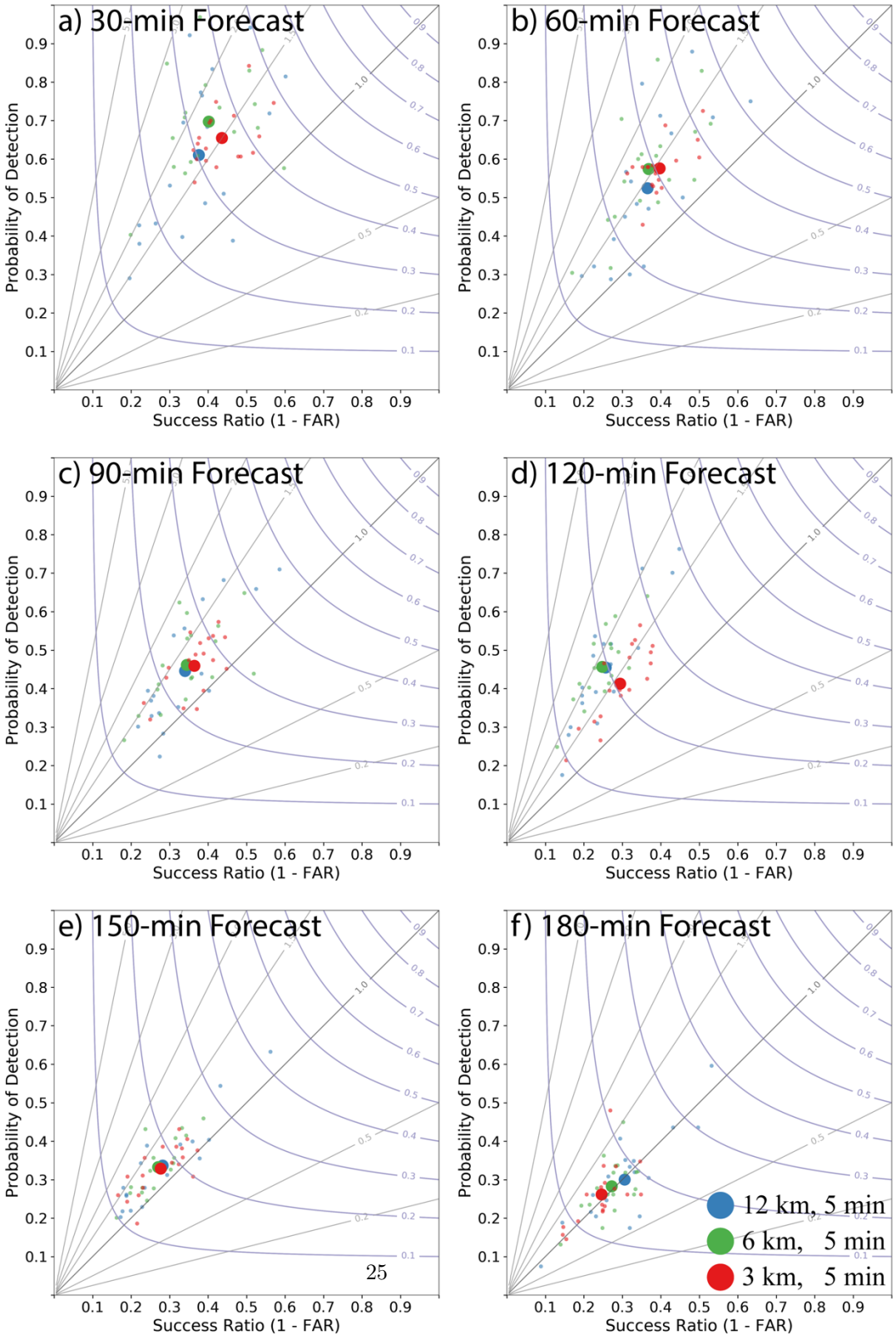
a) 30-min Forecast
b) 60-min Forecast
c) 90-min Forecast
d) 120-min Forecast
e) 150-min Forecast
f) 180-min Forecast
Probability of Detection
Success Ratio (1 - FAR)
12 km, 5 min
6 km, 5 min
3 km, 5 min

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_12km_5min (blue), RES_6km_5min (green), and RES_3km_5min (red) aggregated over 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.

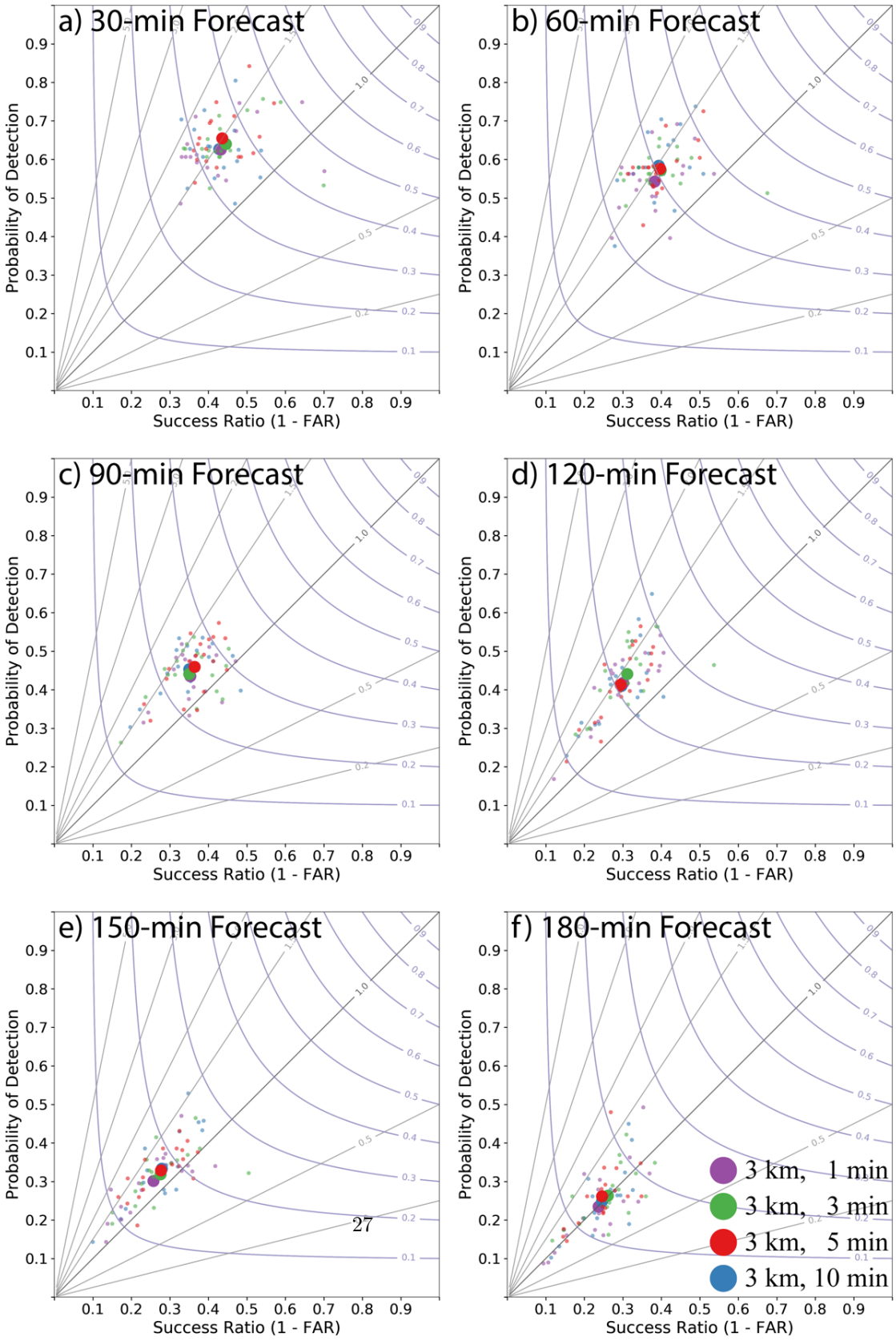
a) 30-min Forecast
b) 60-min Forecast
c) 90-min Forecast
d) 120-min Forecast
e) 150-min Forecast
f) 180-min Forecast
Probability of Detection
Success Ratio (1 - FAR)
3 km, 1 min
3 km, 3 min
3 km, 5 min
3 km, 10 min

Figure 4 Same as Fig. 3, but for the sensitivity experiments RES_3km_1min (purple), RES_3km_3min (green), RES_3km_5min (red), and RES_3km_10min (blue).

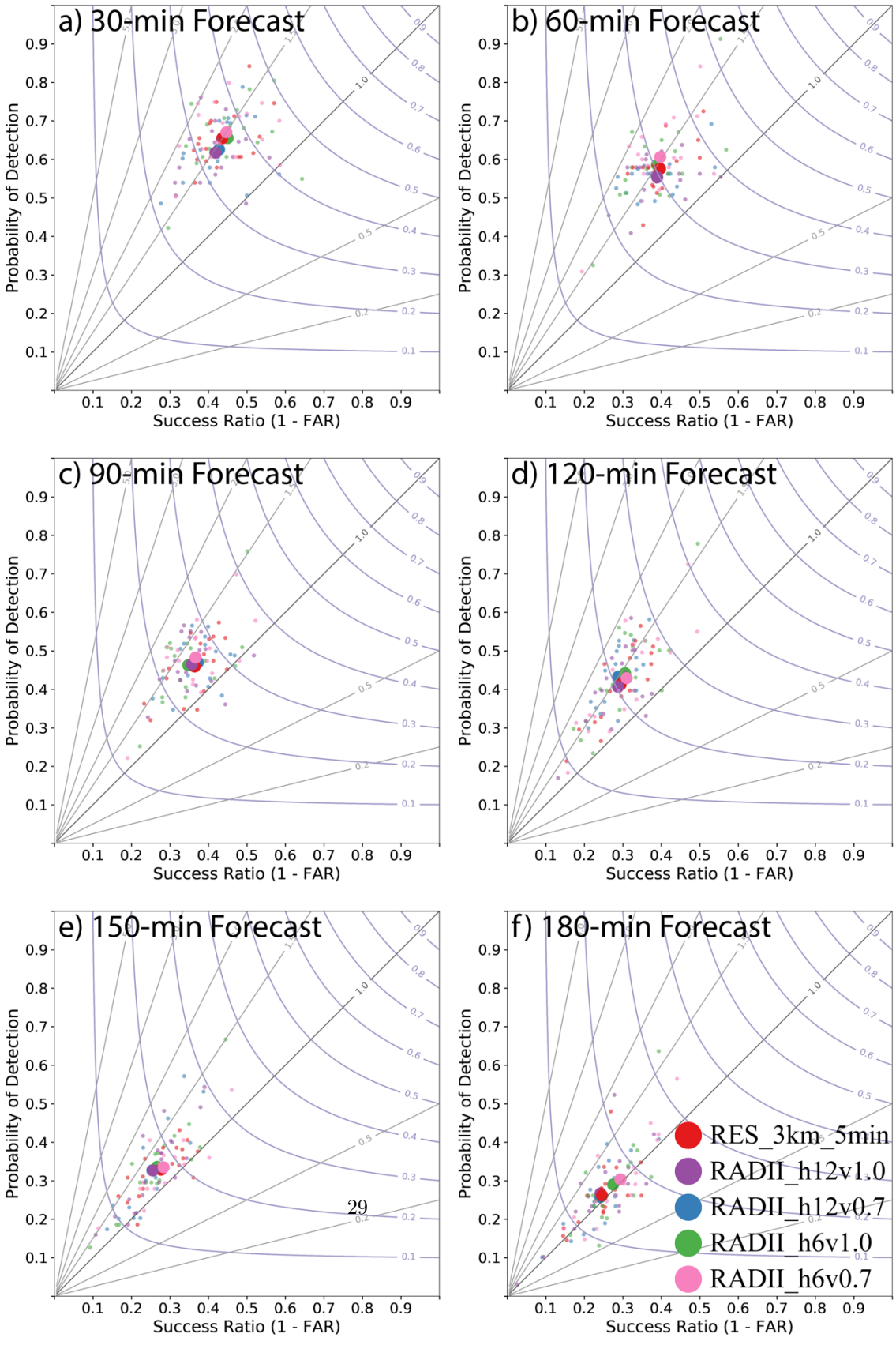
a) 30-min Forecast
b) 60-min Forecast
c) 90-min Forecast
d) 120-min Forecast
e) 150-min Forecast
f) 180-min Forecast
Probability of Detection
Success Ratio (1 - FAR)
RES_3km_5min
RADII_h12v1.0
RADII_h12v0.7
RADII_h6v1.0
RADII_h6v0.7

Figure 5 Same as Fig. 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_3km_5min (red).

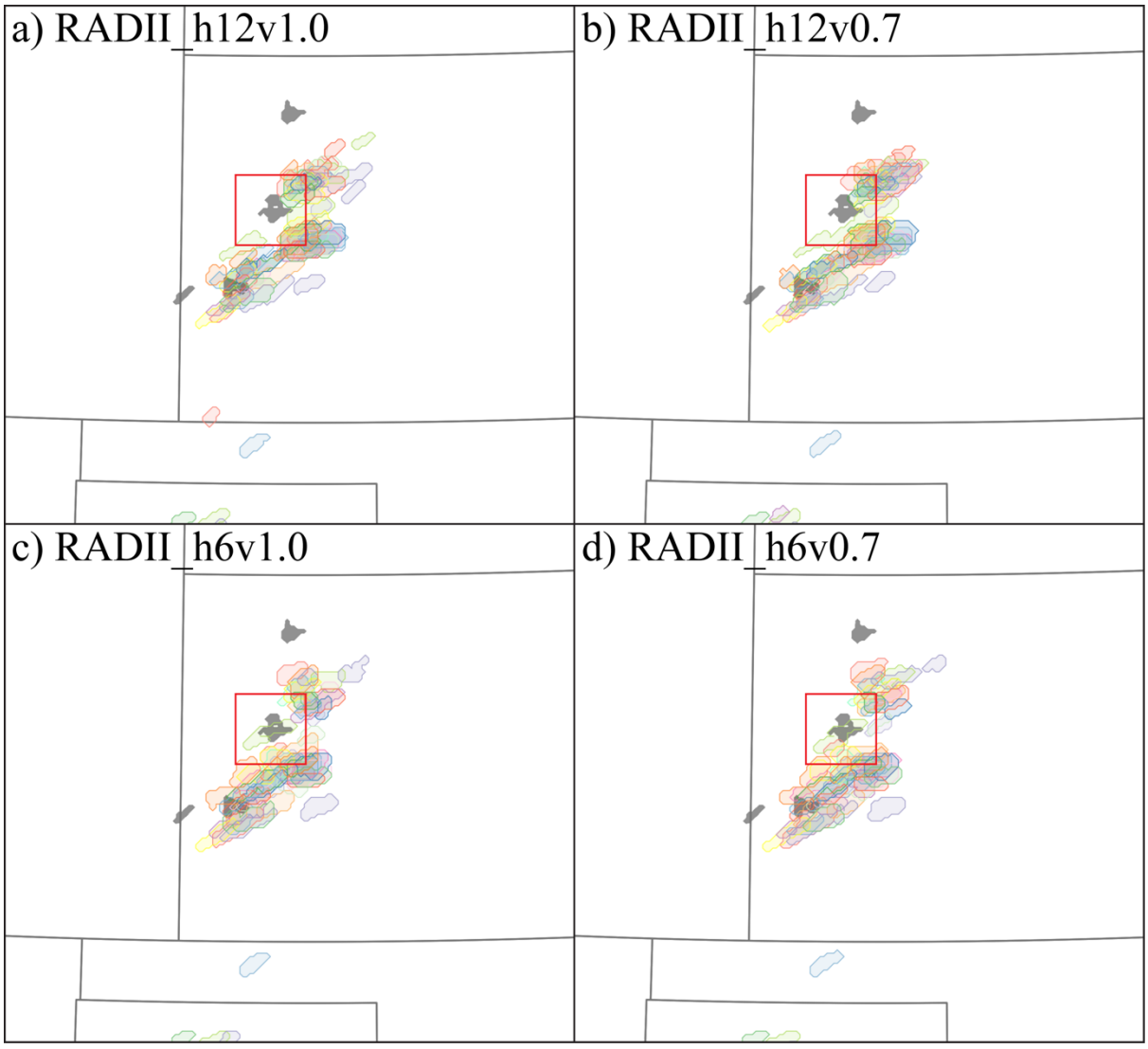

Figure 6 Predicted rotation objects derived from the ensemble forecasts valid at 2200 UTC for 2-h forecasts initiated at 2000 UTC 24 May 2021 for different localization radius. Each ensemble member is plotted as a different color, and the gray shades represent observed MRMS rotation objects at the same time. Red boxes highlight the improvement of experiments using less horizontal localization radius.

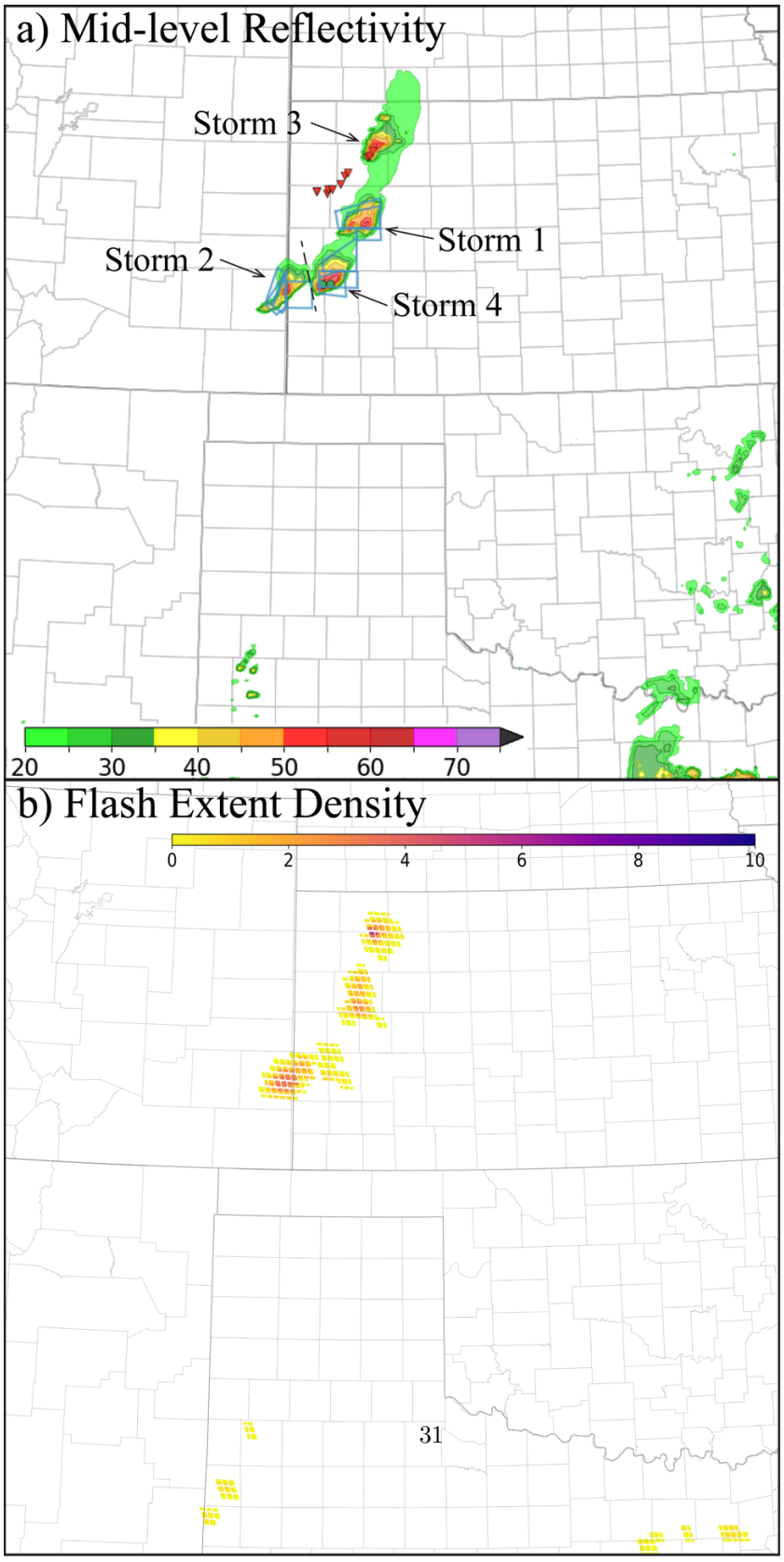
a) Mid-level Reflectivity
Storm 3
Storm 1
Storm 2
Storm 4
20
30
40
50
60
70
b) Flash Extent Density
0
2
4
6
8
10

Figure 7 MRMS 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.

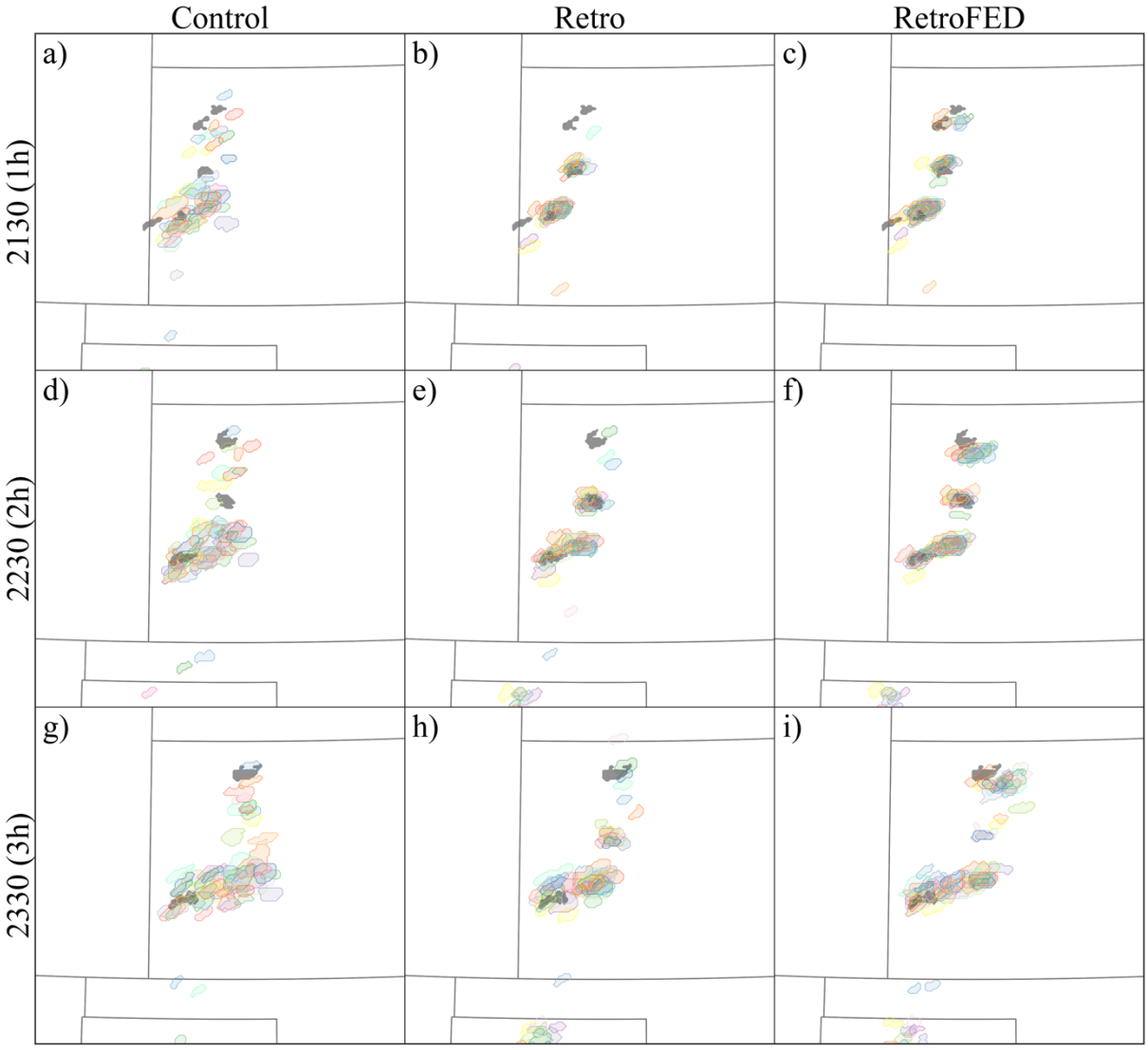

Figure 8 Similar to Fig. 6, but for predicted rotation objects for 1 – 3-h forecasts initiated at 2030 UTC for three experiments. The left, middle and right columns are corresponding to Control, Retro, and RetroFED, respectively.

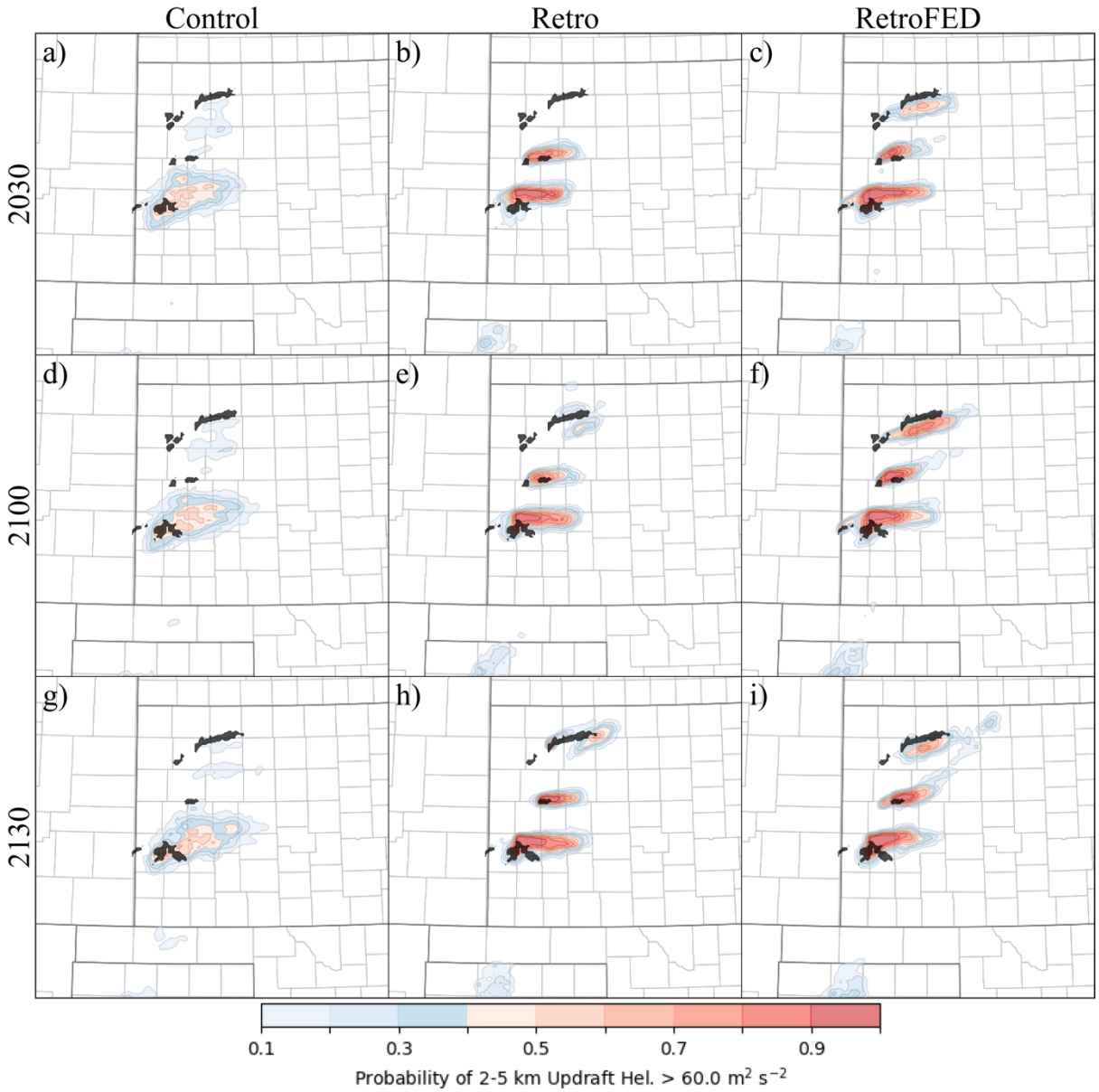

Figure 9 Probability of 2 – 5 km UH greater than 60 m^2 s^{-2} over 3-h forecasts initiated at 2030 UTC (a - c), 2100 UTC (d - f) and 2130 UTC (g - i), 24 May 2021 for each experiment (each column). MRMS 2 – 5 km azimuthal shear tracks during this period (black shades) are overlaid over the probability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

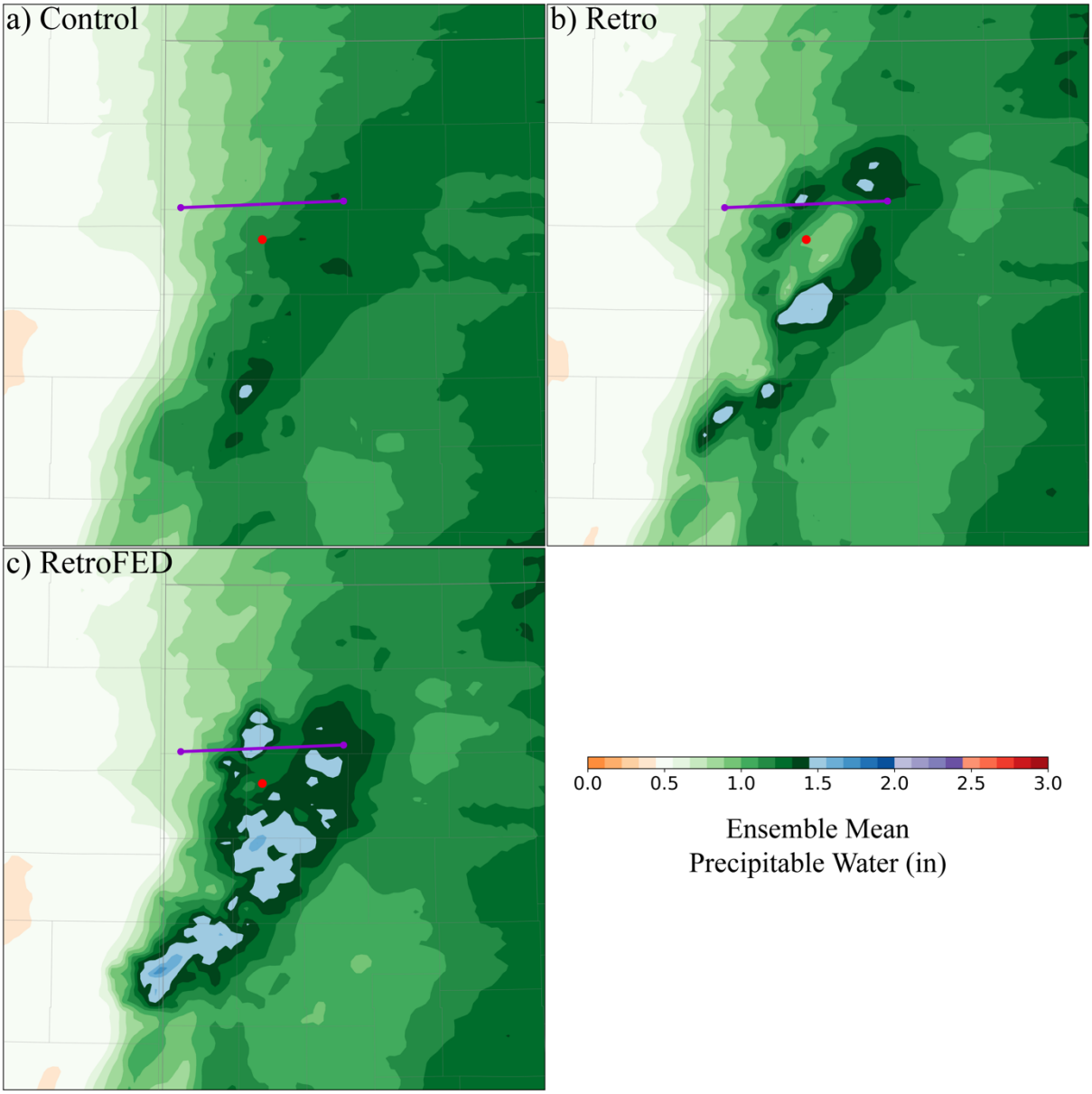

Figure 10 Analyzed ensemble mean precipitable water for (a) Control, (b) Retro and (c) RedtroFED valid at 2030 UTC, 24 May 2021. The red dot represents the location for the sounding plot in Fig. 11. The red dot is the location of the skew-T profiles shown in Fig. 11. The purple line represents the location for the cross section plot in Fig. 12.

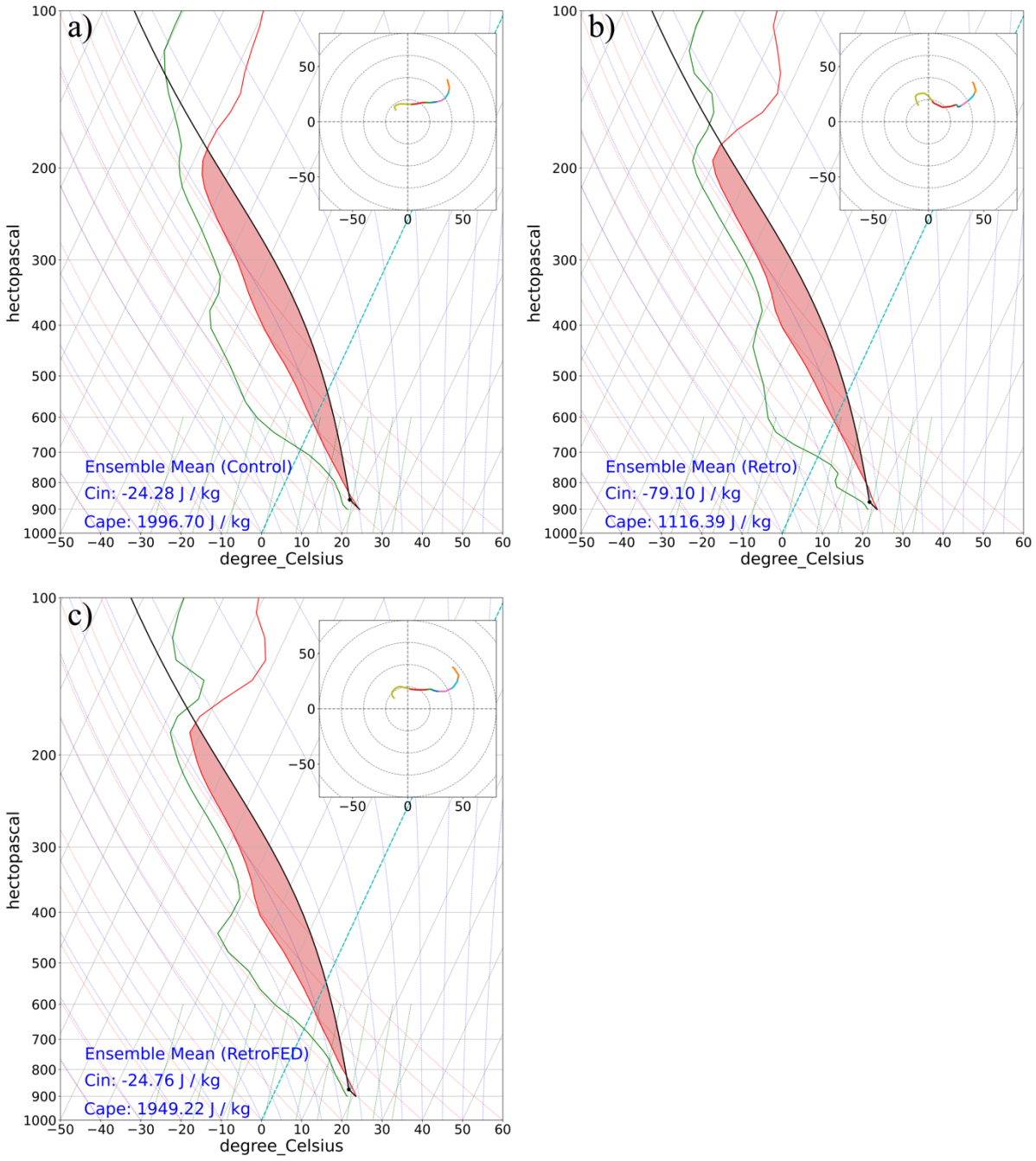

Figure 11 Skew-T profile from analyzed ensemble mean for (a) Control, (b) Retro, and (c) RetroFED valid at 2030 UTC, 24 May 2021. The location of the

skew-T profile is marked in Fig. 10 by red dot.

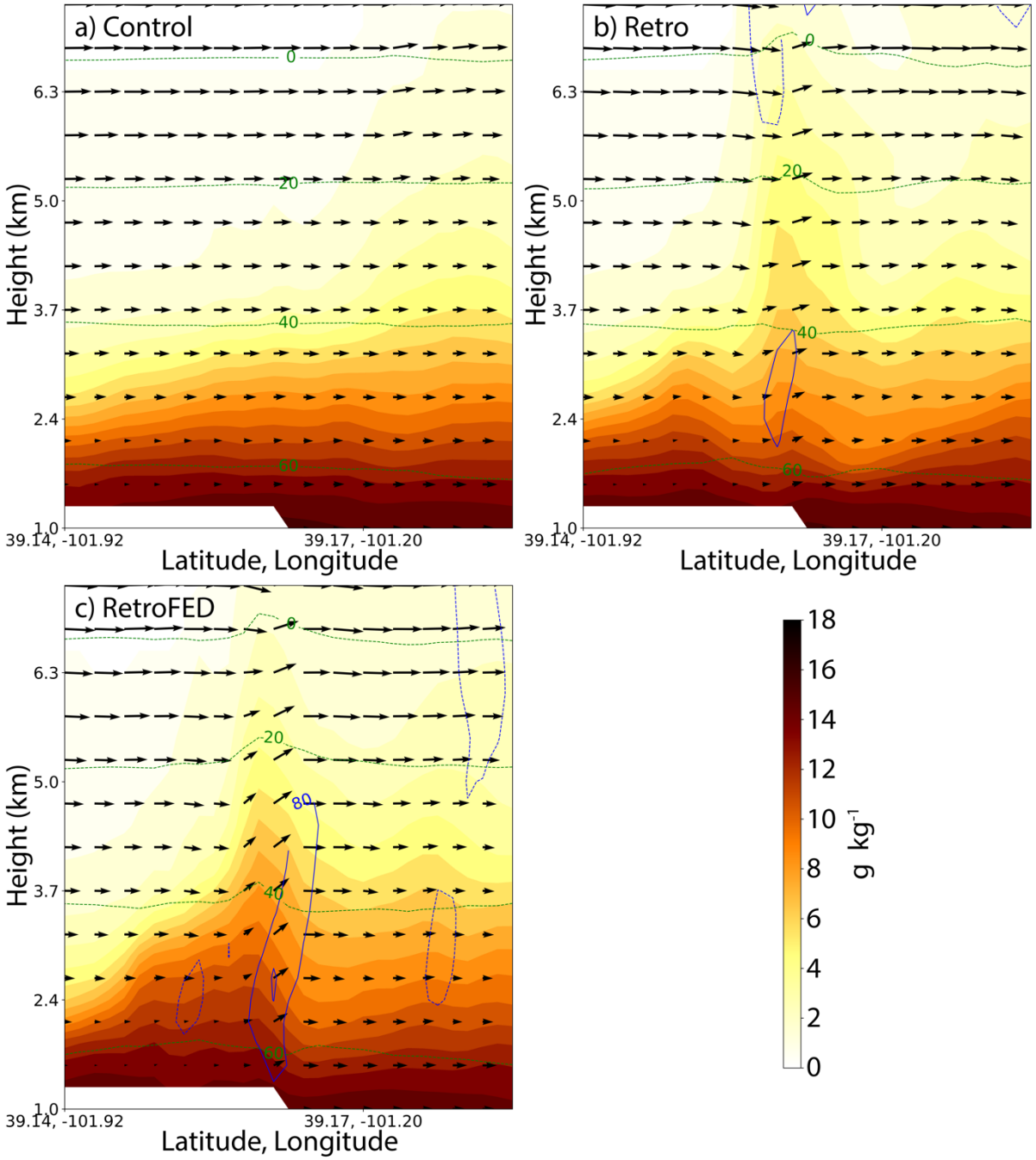

Figure 12 Vertical cross section from analyzed ensemble mean of wind (vectors),

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.

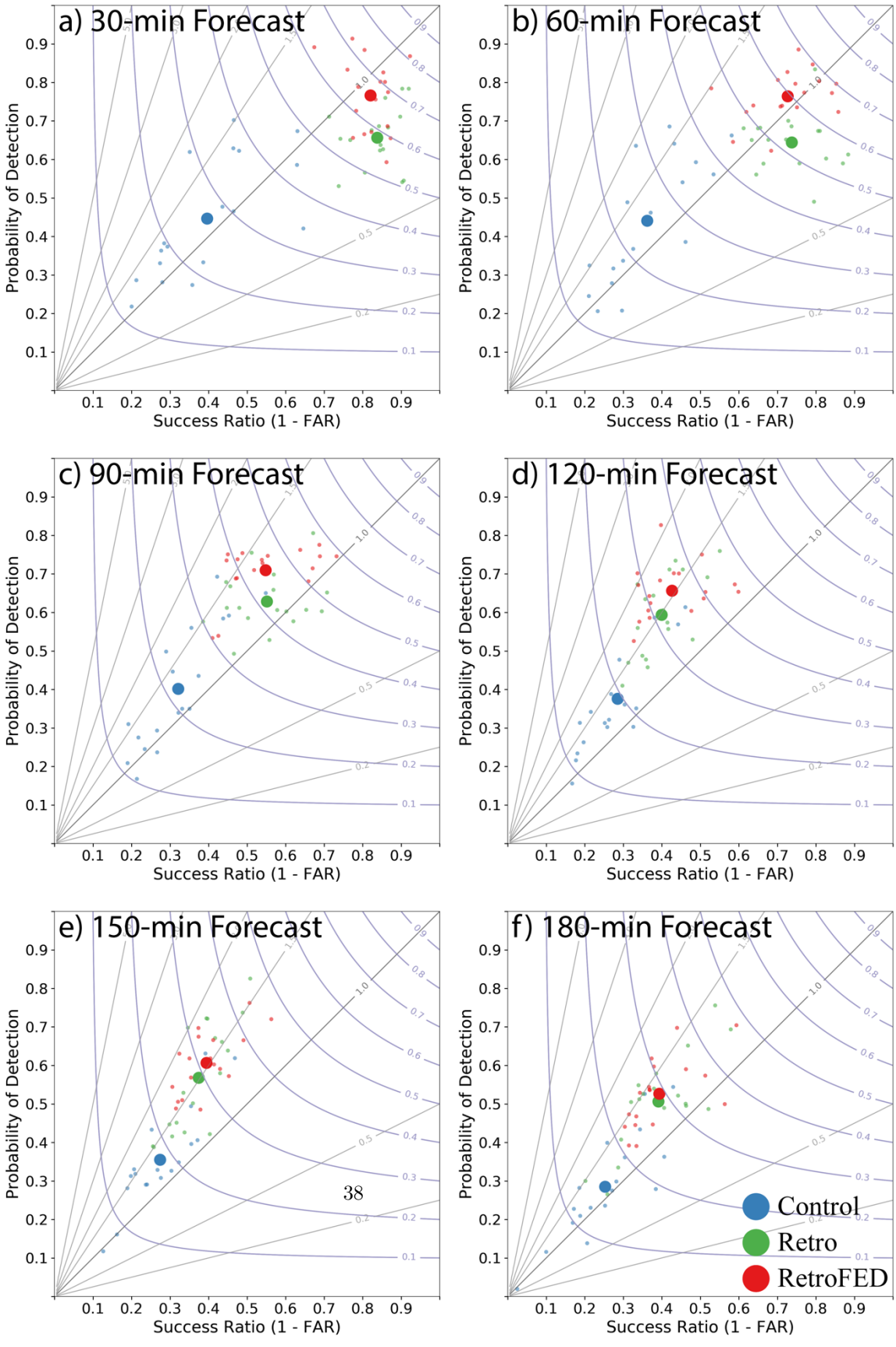
a) 30-min Forecast
b) 60-min Forecast
c) 90-min Forecast
d) 120-min Forecast
e) 150-min Forecast
f) 180-min Forecast
Probability of Detection
Success Ratio (1 - FAR)
Control
Retro
RetroFED

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