JAMES | Journal of Advances in Modeling Earth Systems

RESEARCH ARTICLE
10.1029/2022MS003246

This article is a companion to Zhao et al. (2021), https://doi.org/10.1029/2021MS002484 and Zhao et al. (2021), https://doi.org/10.1029/2021MS002486.

**Key Points:**

- GOES-R derived atmospheric motion vector (AMV) product has the potential to benefit convective scale data assimilation and forecasts
- The impact of assimilating GOES-R derived AMVs together with NEXRAD observations on short-range severe storm forecasts is evaluated
- The verification results for five severe weather events suggest the added forecast skill of GOES-16 AMVs over conventional radar data

**Correspondence to:**
J. Gao,
jidong.gao@noaa.gov

**Citation:**
Zhao, J., Gao, J., Jones, T., & Hu, J. (2022). Impact of assimilating high-resolution atmospheric motion vectors on convective scale short-term forecasts: 3. Experiments with radar reflectivity and radial velocity. *Journal of Advances in Modeling Earth Systems*, *14*, e2022MS003246. https://doi.org/10.1029/2022MS003246

Received 9 JUN 2022
Accepted 20 NOV 2022



# Impact of Assimilating High-Resolution Atmospheric Motion Vectors on Convective Scale Short-Term Forecasts: 3. Experiments With Radar Reflectivity and Radial Velocity

**Juan Zhao[1,2], Jidong Gao[3], Thomas Jones[2,3], and Junjun Hu[4,5]**

[1]China Meteorological Administration Training Centre, Beijing, China, [2]Cooperative Institute for Severe and High-Impact Weather Research and Operations, The University of Oklahoma, Norman, OK, USA, [3]NOAA/National Severe Storms Laboratory, Norman, OK, USA, [4]Key Laboratory of Land Surface Process and Climate Change in Cold and Arid Regions, Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences, Lanzhou, China, [5]Nagqu Station of Plateau Climate and Environment, Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences, Nagqu, China

**Abstract** Based on the idealized supercell and real case studies in Part I and II, the purpose of this subsequent study is to further investigate the impact of assimilating Geostationary Operational Environmental Satellites-16 (GOES-16) derived atmospheric motion vectors (AMVs) in addition to WSR_88D Doppler radar observations on convective scale numerical weather prediction. Five high-impact weather events that occurred in spring 2018 and 2019 are analyzed using the National Severe Storms Laboratory three-dimensional variational data assimilation (DA) system. Four types of experiments are implemented and compared: (a) the control experiment (NoDA) without assimilating any observation, (b) the radar DA experiment (RAD), (c) the GOES-16 AMV DA experiment (AMV), and (d) the experiment assimilating AMVs together with radar data (AMV_RAD). Score metrics aggregated over all cases indicate that AMV_RAD performs slightly better than RAD in 0–3 hr reflectivity and precipitation forecasts especially at higher thresholds, suggesting the added value of GOES-16 AMVs on radar data. Detailed case examinations also show that AMV_RAD generally exhibits slightly more skillful storm forecast in terms of the areal coverage, storm mode, and storm orientations, owing to improvements in the analysis of boundary locations and localized enhanced divergence signatures. In spite of encouraging objective and subjective evaluation results, AMV_RAD has difficulty in adjusting the moisture gradient associated with dryline and tends to underpredict the associated weak discrete storms.

**Plain Language Summary** With the launch of Geostationary Operational Environmental Satellites-16 (GOES-16) in November 2016, the impact of its derived high-spatiotemporal-resolution atmospheric motion vectors (AMVs) product in convective-scale numerical weather prediction has not been extensively explored. In this study, the GOES-16 AMVs together with WSR-88D Doppler radar observations are effectively assimilated by a three-dimensional variational data assimilation scheme developed at NOAA/National Severe Storms Laboratory. Both subjective and objective assessment results for five severe weather events suggest the slightly added forecast skill of GOES-16 AMVs on conventional radar data for 0–3 hr reflectivity and precipitation forecasts.

## 1. Introduction

Since the launch of a new-generation geostationary satellites, such as GOES-16/17, Himawari-8/9, Geostationary Korea Multipurpose Satellite-2A/B(GEO-KOMPSAT-2A/B), EUMETSAT-Meteosat Third Generation, and Fengyun-4A/B, etc., study on assimilation of the high-spatiotemporal-resolution satellite radiance observations, or derived products into convective scale numerical weather prediction (NWP) has been a popular research topic and of great importance for promoting the application of new satellite observations in operational severe weather forecasts (Fierro et al., 2016, 2019; Hartman et al., 2021; Jones et al., 2018, 2020; Kong et al., 2020; Minamide & Zhang, 2018; Minamide et al., 2020; Otsuka et al., 2021; Zhang et al., 2019). One of the important geostationary satellite-derived products are atmospheric motion vectors (AMVs), that can provide mesoscale or convective flow information for model initial conditions. Recent research suggested slightly positive or neutral impacts on the global and regional NWP from the assimilation of high-density or rapid-scan AMV data, especially in the tropical cyclone track and intensity forecasts (Cherubini et al., 2006; Elsberry et al., 2018; Kim & Kim, 2018; Kim et al., 2017; Lean & Bormann, 2019; Lean et al., 2016; Le Marshall et al., 2008; Li et al., 2020; Lim

et al., 2019; Mallick & Jones, 2020; Otsuka et al., 2015; Sawada et al., 2019; Velden et al., 2017; Wu et al., 2015; Yamashita, 2012, 2017). To our knowledge, there are few peer-reviewed published research investigating the impacts of assimilating high-spatiotemporal-resolution AMV data on mesoscale and convective scale weather forecasts over land (Mallick & Jones, 2020; Zhao et al., 2021a, 2021b).

An early application of satellite-derived AMVs in the convective-allowing ensemble DA system has been studied by Mallick and Jones (2020) and their results showed that assimilation of AMVs can have a positive influence on probabilistic forecasts of reflectivity objects. Zhao et al. (2021a, Part I of this series) explored the benefit of assimilating simulated GOES-R AMVs on convective scale NWP with an idealized supercell storm. In this idealized case study, assimilation of synthetic AMV observations resulted in improvements in the wind analyses within and around storms and better representations of storm top-level divergence and low-level convergence fields. Sensitivity of the analysis and forecast to observation density, DA cycling frequency and horizontal correlation scale were also tested and the conclusions served as a reference guide for real case study. Subsequently in Zhao et al. (2021b, Part II in this series), the impact of assimilating the GOES-16 AMV product was investigated with three high-impact weather events that occurred in spring 2018 and 2019 over the Great Plains of the United States. The results showed that AMV DA improves the dynamic and thermodynamic analyses associated with storm environments and the location and shape of boundaries, leading to more skillful forecasts of storm evolution and movement. It was also found that assimilation of AMVs has a remarkable positive impact on the quasi-linear or mesoscale convective system (MCS). However, AMVs have difficulty in discerning an isolated storm or a smaller cluster of cells.

Doppler radar data, which samples the three-dimensional storm structures with high frequency and density, including wind and some hydrometeor information, has been playing an important role in convective scale DA and exhibits many encouraging results (Gao & Stensrud, 2012; Gao et al., 2004, 2016; Hu et al., 2021; Stensrud & Gao, 2010; Sun, 2005; Wang & Wang, 2017). However, conventional weather radar cannot provide observations outside thunderstorms after formation of precipitation, limiting its utilization after convection initiation or in clear-air regions (Huang et al., 2022; Markowski et al., 2006). Since the Advanced Baseline Imager (ABI) aboard GOES-16/17 is capable to observe Continental United States (CONUS) area every 5 min with a resolution at nadir of 0.5 km for the visible channel and 2 km for infrared channels, high spatiotemporal-resolution AMV product can be retrieved every 15 min over CONUS with a horizontal resolution of 7.5 km (30 or 38 km) for the visible (infrared) channel derived winds (Daniels et al., 2012). AMVs are derived by tracking cloud or clear-sky water vapor, so AMVs have the potential to cover the blind spots and gaps within radar network.

The preceding two parts of this study, which represent the maximum possible amount of new information introduced by assimilating GOES-16 AMVs, indicate that AMV DA benefits short-term severe weather forecasts by improving the storm environment. It is expected that assimilating radar data together with AMV can complement each other's respective observation gaps to resolve a much wider range of scales. In this study, the following questions will be addressed. How will the assimilation of AMVs combined with radar DA impact severe storm forecast? Is there any added forecast skill of assimilating GOES-16 AMV product over radar data on short-term convective-scale NWP? To answer these questions, one control experiment without assimilating any observations and three DA experiments with different combinations of data usage will be performed with several real data cases in different weather scenarios. Comparisons between these experiments will reveal the combined impact of AMV and radar DA on short-term severe weather forecast.

The remainder of this study is organized as follows. Section 2 briefly describes the GOES-16 AMV and radar data and their associated forward operators and quality control (QC) procedures. The configurations for forecast model and assimilation settings for AMV and radar observations, as well as experimental design, are presented in Section 3. Both aggregate performance of all five real cases and the analysis/forecast results in more details from three representative cases are discussed in Section 4. Finally, the summary and conclusions are offered in Section 5.

## 2. Brief Description of Observations Used for Assimilation

### 2.1. GOES-16 AMV Data

As in Zhao et al. (2021b), this study uses the GOES-16 ABI Level 2 (L2) AMV product over the CONUS region generated by NOAA NESDIS Center for Satellite Applications and Research in NetCDF file format (Daniels

et al., 2012). By tracing a set of targets, including cloud edges or moisture gradients, the ABI AMV product is derived from six selected spectral channels: visible band 2 (0.64 μm), shortwave infrared band 7 (3.9 μm), mid-wave water vapor bands 8 (6.15 μm), 9 (7.0 μm), and 10 (7.4 μm), and longwave infrared band 14 (11.2 μm). In the AMV retrieval algorithm, band 2, 7, and 14 employ cloud tracers, while band 9 and 10 use clear-sky water vapor tracers. Both cloud and clear-sky water vapor features are tracked for band 8. It is noted that the AMVs derived from band 2 and 7 are only available during daytime and nighttime, respectively.

The GOES-16 AMVs over the CONUS scan sector are available every 15 min. Given the fact that horizontal resolution of AMVs is driven by spatial and temporal resolution of ABI imageries as well as target scene size, the horizontal resolutions of AMVs generated from visible, water vapor, and longwave infrared bands are 7.5, 30, and 38 km, respectively (Daniels et al., 2012). There are only three layers of wind observations in AMV product: 1,000–700, 700–400, and 400–100 hPa, with band-depend acceptable height ranges as indicated in Zhao et al. (2021b).

Prior to entering the assimilation system, the wind speed, and azimuth observations in raw AMV datasets are converted to zonal and meridional wind components, that will go through a seven-step QC strategy. Most of the QC procedures are performed to filter out bad data beyond the acceptable height ranges of AMVs. The criteria for pressure are selected based on the channel-dependent height ranges for AMVs specified in the GOES-R ABI Algorithm Theoretical Basis Document for AMVs (Daniels et al., 2012) as well as previous studies and the Grid-point Statistical Interpolation software (Kim et al., 2017; Lim et al., 2019; Mallick & Jones, 2020; Sawada et al., 2019; Velden et al., 2017). Moreover, poorer data quality of AMVs in middle level is also considered in the QC strategy (Lean & Bormann, 2019; Salonen et al., 2015). The QC procedures performed in this study include: (a) remove all the wind data from surface to 950 hPa and that above 100 hPa; (b) for the winds derived from both visible (0.64 μm) and shortwave infrared (3.9 μm) bands, remove data above 750 hPa; (c) for the upper-level water vapor winds (6.15 μm), reject data below 400 hPa; (d) for band 9 (7.0 μm) winds, keep the data above 450 hPa only and with a minimum wind speed threshold of 8 m/s; (e) for the longwave infrared band (11.2 μm), remove winds retrieved between 800 and 400 hPa; (f) reject all the data with a solar zenith angle larger than 68°; and (g) a relaxed gross error check, in which the threshold value between the ratio of innovation to observation error is set to 5, is performed to eliminate the observations outside of set tolerances from the interpolated model background field.

### 2.2. Radar Observations

WSR-88D level II radar reflectivity and radial velocity observations are interpolated onto model grids before they are inserted into the data assimilation scheme. The largest reflectivity value is chosen at the grid point with multiple radars overlapping. The radar observations are then passed through QC procedures, such as radial velocity dealiasing and removal of weak radar returns or non-meteorological scatters (less than 15 dBZ), prior to interpolating to model grid points using quadratic interpolation technique. The forward operator and its adjoint for reflectivity and radial velocity follow the same procedures as in Gao and Stensrud (2012) and Gao et al. (2004, 2013).

## 3. Model and Assimilation System Configurations, and Experimental Design

### 3.1. Model and Assimilation System Configurations

The forecast model is the Advanced Research version of the Weather Research and Forecasting (WRF-ARW) Model version 3.6.1 (Skamarock et al., 2008). All forecast experiments in this research follow a similar configuration as in Zhao et al. (2021b) wherein the simulation was performed on a 1.5-km spaced grid of 600 × 600 horizontal points and 51 vertical levels. Figure 2 illustrates the domain for each case, which is determined based on the "Day 1" Convective Outlook released by the Storm Prediction Center. The following physical parameterizations are selected: the National Severe Storms Laboratory (NSSL) two-moment four-ice category bulk microphysics scheme (Mansell & Ziegler, 2013; Mansell et al., 2010; Ziegler, 1985), the Rapid Radiative Transfer Model longwave radiation scheme (Mlawer et al., 1997), the Dudhia shortwave radiation scheme (Dudhia, 1989), the Rapid Update Cycle land surface scheme (Benjamin et al., 2004), and the Yonsei University (YSU) planetary boundary layer scheme (Hong et al., 2006).

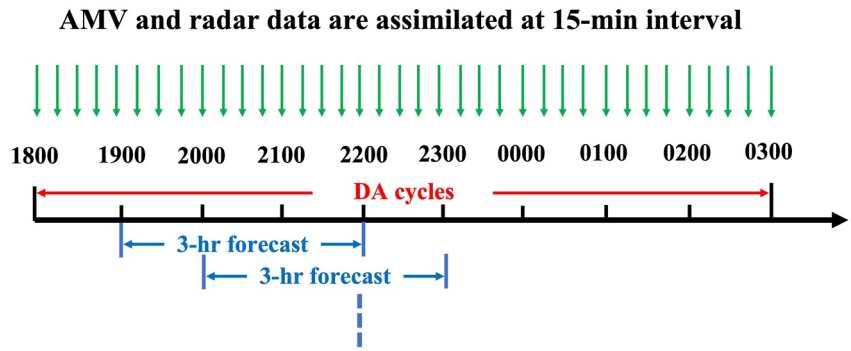


**Figure 1.** Illustration of the data assimilation and forecast cycle workflow. A 3-hr forecast is launched every hour from 19:00 to 03:00 UTC (namely, nine separate forecasts).

The AMV and radar observations are assimilated by using the NSSL Experimental Warn-on-Forecast (WoF) 3DVAR system (NSSL3DVAR). Assimilation of radar reflectivity makes adjustments to the mass mixing ratios for rainwater, snow and hail/graupel, and classification of the hydrometeor type depends on background temperature (Gao & Stensrud, 2012; Gao et al., 2016). The background errors for these hydrometeor variables were derived from a realtime ensemble Warn-on-Forecast System for short-term convective scale NWP (Jones et al., 2018, 2020; Pan, Gao, Jones, et al., 2021). The observation errors for radar reflectivity and radial velocity are 10.0 dBZ and 3.0 m $s^{-1}$, respectively. As in Zhao et al. (2021b), the observation error for the zonal and meridional wind components converted from AMV data are set to be 6 m $s^{-1}$. To account for multi-scale features represented by the two types of observations, multiple outer loops are applied in the minimization process with different correlation scale lengths, that depend on the density of assimilated observations and affect how far the observational information will spread in model space (Gao et al., 2004; Purser et al., 2003). The horizontal correlation scale lengths for assimilating radar radial velocity and reflectivity are 20 and 6 km, respectively. AMVs and radar data will be assimilated sequentially in two outer loops. Following Zhao et al. (2021b), AMVs are assimilated using the 50-km horizontal correlation scale length in the first outer loop, while radar radial velocity and reflectivity are assimilated with 20- and 6-km horizontal correlation scale lengths in the second outer loop.

### 3.2. Experimental Design

To perform a systematic evaluation of the impact of assimilating AMVs together with radar data on short-term forecasts, five high-impact weather events that occurred over the Great Plains of the United States in spring 2018 and 2019 are examined, namely, 1 May 2018, 30 April 2019, 17 May 2019, 28 May 2019, and 29 May 2019. For each case, the DA and forecast cycling occurs from 18:00 to 03:00 UTC the following day at 15-min intervals with a 3-hr free forecast launched hourly starting from 19:00 UTC (as shown in Figure 1). This is similar to WoF ensemble system real-time Spring Forecast Experiment runs (Hu et al., 2020; Zhao et al., 2021b). The 3-km High Resolution Rapid Refresh forecasts initialized at 18:00 UTC are employed to provide initial conditions and boundary conditions for the WRF model. Four types of experiments (Table 1) are performed to assess the impact of assimilation of AMV and radar data on severe storm forecast. The first type of experiments are control runs (denoted as NoDA) advancing the model forward without assimilating any observations. The second type of DA experiments assimilate WSR-88D radar reflectivity and radial velocity (referred to as RAD). The third type of DA experiments only assimilate GOES-16 AMV data (referred to as AMV), and the fourth type of DA experiments assimilate AMVs in conjunction with radar data (referred to as AMV_RAD). Additionally, horizontal and vertical correlation scale lengths for each type of observation used in each outer loop in the variational system are also listed in Table 1. The simulation domain overlaid with location of radar sites utilized for each case is presented in Figure 2, while the geographical distribution of GOES-16 AMVs assimilated into the NSSL3DVAR system at 18:00 UTC for all five cases are displayed in Figure 3. For all

**Table 1**
*List of the Data Assimilation Sensitivity Experiments*

| Experiment | AMV | Reflectivity | Radial velocity | Horizontal correlation scale (km) | Vertical correlation scale (grid points) |
|---|---|---|---|---|---|
| NoDA | | | | | |
| RAD | | ✓ | ✓ | Vr: 20 | Vr: 4 |
| | | | | Ref: 6 | Ref: 2 |
| AMV | ✓ | | | 50/20 | 5/4 |
| AMV_RAD | ✓ | ✓ | ✓ | AMV: –/50 | AMV: –/5 |
| | | | | Vr: 20/– | Vr: 4/– |
| | | | | Ref: 6/– | Ref: 2/– |

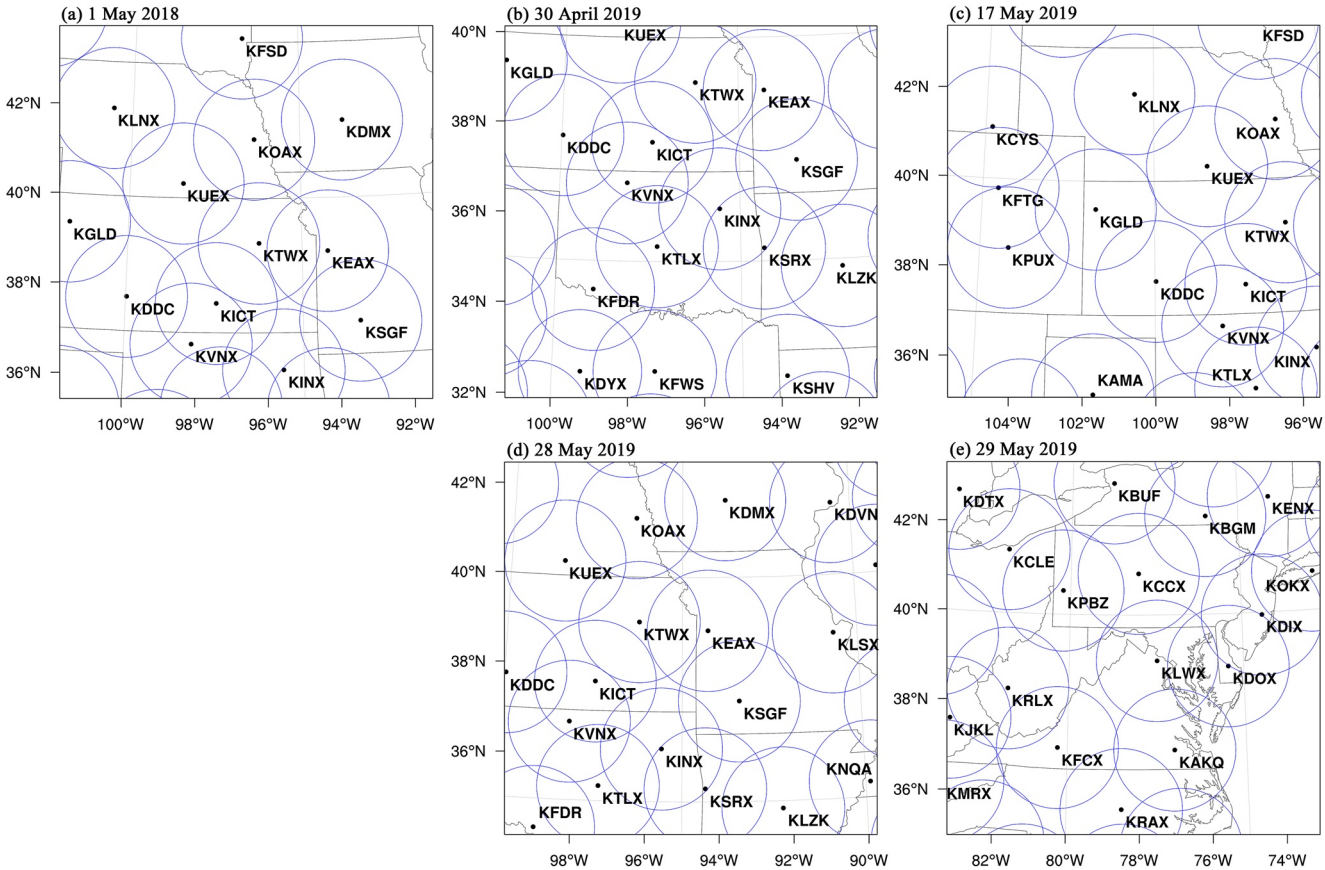


**Figure 2.** Simulation domains and locations of the radar sites for (a) 1 May 2018, (b) 30 April 2019, (c) 17 May 2019, (d) 28 May 2019, and (e) 29 May 2019. The detection range (150 km) of each radar is denoted by a blue circle.

cases, the upper level (400–100 hPa) contains most AMV data, and the middle level (700–400 hPa) has least AMV data. It is clearly seen that the number of radar observations available for assimilation is generally around three orders of magnitude larger than that of AMV data (Figure 4c). Moreover, the number of AMV observations decreases during assimilation cycles, whereas the number of radar data increases. This suggests that AMV and radar observations generally cover different areas and phases in a life cycle of convection (i.e., cloudy area and clear-sky area, prior to and after formation of precipitation) and can compensate one another's shortage in temporal and spatial coverages.

To obtain an overall insight into the impact of AMV DA on wind analyses, root-mean-square errors (RMSEs) and biases of wind components against the GOES-16 AMV observations during DA cycles are calculated for individual cases and then aggregated over all five cases. To evaluate the impact on short-term forecasts, both qualitative and quantitative assessments of reflectivity and precipitation forecasts are conducted against several available observational data sets. The neighborhood-based equitable threat score (ETS, Clark et al., 2010) is calculated using neighborhood radius of 12-km from composite reflectivity (CREF) and hourly precipitation (HPRCP) from model forecast outputs against the CREF observations from the NSSL Multi-Radar Multi-Sensor (MRMS) product (Smith et al., 2016) and the Stage IV hourly rainfall estimates from the National Centers for Environmental Prediction (Baldwin & Mitchell, 1997). Both verification data sets are interpolated onto the 1.5-km model grids prior to verification. Additionally, contingency-table based metrics including the probability of detection (POD), false alarm ratio (FAR), success ratio (SR), frequency bias (BIAS), and critical success index (CSI) are also computed to verify the CREF and HPRCP forecasts. Given that data impact may vary across different weather conditions, all score metrics are aggregated across five cases, to gain a fair and consistent measure of forecast skill. Moreover, subjective diagnostic analyses of dynamic and thermodynamic variables related to storm environments are also discussed in three cases examined in more detail in the following sections.

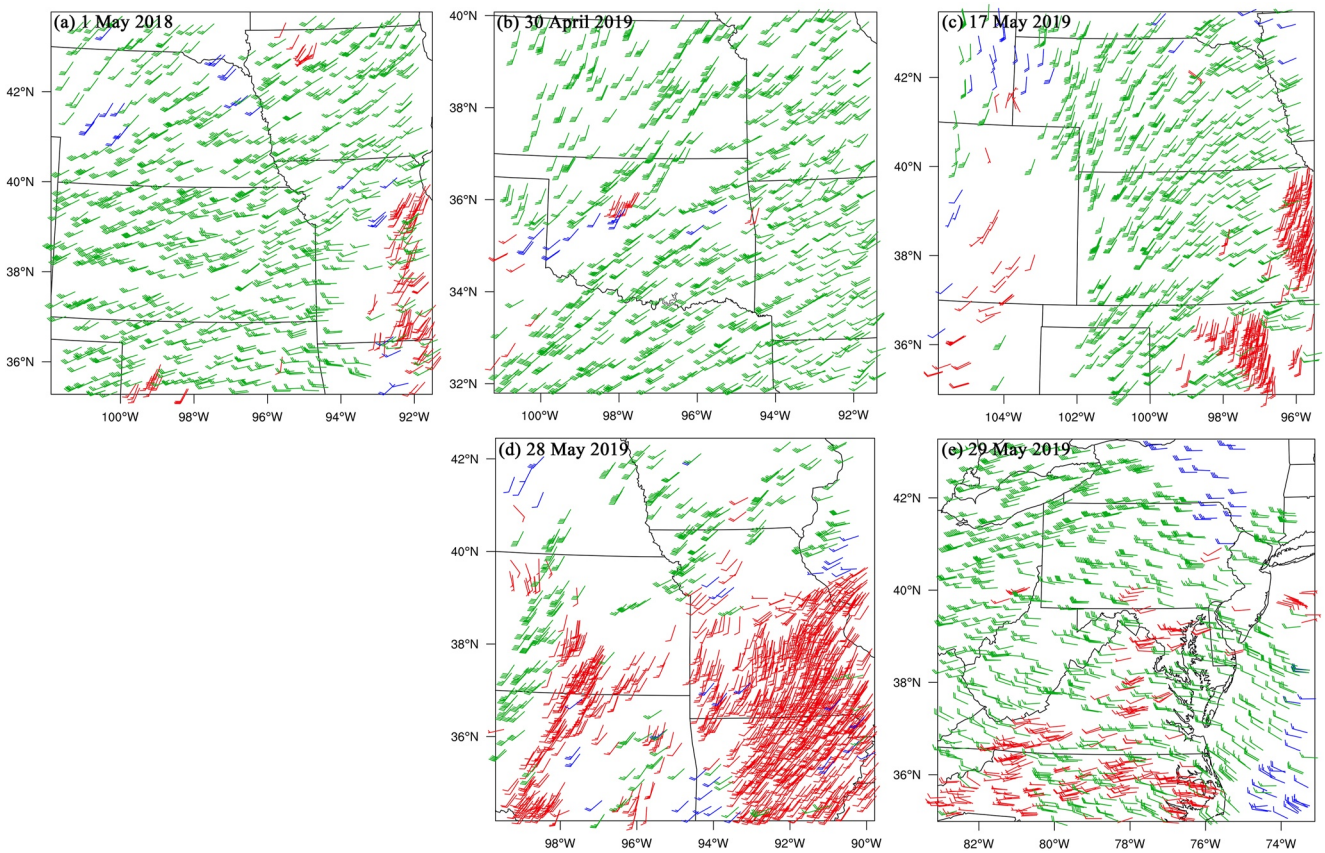


**Figure 3.** The geographical distribution of the Geostationary Operational Environmental Satellites-16 atmospheric motion vector (AMV) observations which are assimilated at 18:00 UTC (a) 1 May 2018, (b) 30 April 2019, (c) 17 May 2019, (d) 28 May 2019, and (e) 29 May 2019. Red barbs represent AMVs within the 1,000–700 hPa layer, blue AMV barbs are within the 700–400 hPa layer, and green AMV barbs are within the 400–100 hPa layer.

## 4. Results and Discussion

### 4.1. Aggregate Analysis and Forecast Performance

As mentioned above, all the quantitative verification parameters including RMSE, ETS, POD, FAR, SR, BIAS, and CSI initialized at each available forecast time from individual cases are aggregated over all five real data cases. This helps gain an overall understanding of the impact of assimilating both AMV and radar data on short-range severe storm forecasts. First of all, the influence of assimilating AMVs on wind analysis can be straightforwardly assessed by examining RMSEs and biases of wind components before and after DA during the assimilation cycles from 18:00 to 03:00 UTC. For clarity, Figures 4a and 4b compare analysis errors for $u$ component from the AMV and AMV_RAD experiments, all of which have been aggregated over five cases. Prior to assimilating AMV data, RMSEs of the zonal wind decrease remarkably within the first cycle and then increase progressively over time (Figure 4a). The RMSEs of wind analyses from both AMV and AMV_RAD experiments are significantly smaller than those in the background throughout all DA cycles, indicating that wind information in the AMV data set is absorbed by the variational system reasonably well. It is worth noting that wind analysis in AMV_RAD is slightly degraded with larger RMSEs (about 15%–30%) compared to the AMV experiment, which can be the consequence of wind analysis being fitted to both AMV and radial velocity observations in AMV_RAD. Moreover, the biases for $u$ (Figure 4b) and $v$ (figure omitted) components are reduced (closer to zero) when AMVs are assimilated alone or together with radar data.

Figure 5 shows the categorical performance diagrams (Roebber, 2009) and neighborhood-based ETS graphs of CREF forecast for various thresholds. Since the score metrics are aggregated from all available initialization times over five real cases, a comprehensive view of the performance of AMV and radar DA is highlighted by comparing all four experiments NoDA, AMV, RAD, and AMV_RAD. The error bars in the ETS graphs represent a

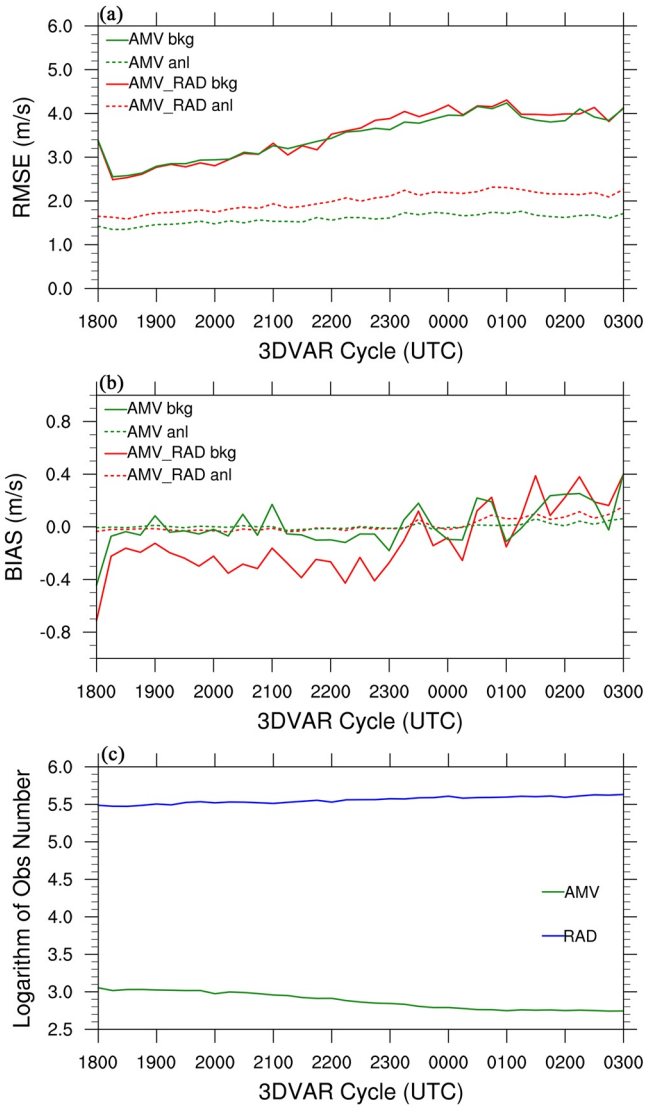


**Figure 4.** Aggregated (a) root-mean-square errors and (b) bias of *u* background (solid) and analysis (dashed) from the atmospheric motion vector (AMV) (green) and AMV_RAD (red) experiments during the assimilation cycles from 18:00 to 03:00 UTC. (c) Common logarithm of the numbers of AMV (green) and radar (blue) observations assimilated during the assimilation cycles.

95% confidence interval. As expected, the values of POD, SR, CSI, and ETS decrease with respect to both forecast lead time and reflectivity threshold values, whereas BIAS increases monotonically when the threshold increases from 30 to 50 dBZ (Figures 5a, 5c, and 5e). BIAS is close to unity at 30 dBZ (Figure 5a), while overprediction is seen at higher thresholds from all experiments, for example, 40 and 50 dBZ (Figures 5c and 5e). This indicates that the model microphysics may have difficulty in properly resolving convective clouds, which is also concluded by Hu et al. (2021). Overall, the result shows that all DA experiments outperform NoDA with larger values of POD, SR, and CSI. In contrast to NoDA, the AMV experiment produces a slightly better CREF forecast for the thresholds of 30–40 dBZ and exhibits a larger positive effect for 50 dBZ, which is also seen in Zhao et al. (2021b). Comparison among three DA experiments reveals that both RAD and AMV_RAD show an overwhelming superiority over AMV. This occurs likely because (a) radar data provides three dimensional internal structures of storms, and (b) the number of radar observations is about three orders of magnitude larger than the number of AMV data (Figure 4c). Specifically, for 30- and 40-dBZ thresholds, the CREF forecasts from AMV_RAD are comparable to those from RAD, besides that the AMV_RAD analysis has a smaller BIAS. For 50 dBZ, AMV_RAD produces the highest POD, SR, and CSI as well as the smallest BIAS, with RAD exhibiting slightly inferior performance. Aggregated ETSs of reflectivity forecast show that RAD and AMV_RAD perform remarkably better than AMV and NoDA in the first 2 hours, which is statistically significant at the 95% confidence level (Figures 5b, 5d, and 5f). AMV_RAD generally produces the highest ETS values especially for 0–1 hr forecasts at 30 dBZ and 0–3 hr forecasts at 40 and 50 dBZ. These improvements are, however, not statistically significant at the 95% confidence level.

In agreement with the CREF forecast, AMV_RAD shows slightly more skillful HPRCP forecast than RAD does (Figure 6). Both the first- and third-hour precipitation forecasts in NoDA have high POD but low SR as well as large BIAS for all thresholds (Figures 6a, 6c, and 6e), which indicates the presence of abundant false alarms and precipitation overprediction. Due to remarkable BIAS decrease in the second hour forecast, the ETSs of NoDA reach the highest in the second hour, then decrease in the third hour (Figures 6b and 6d). Compared to NoDA, the AMV experiment reduces number of false alarms (leading to a higher SR) with a lower POD as a consequence of suppressing spurious precipitation. In addition, AMV_RAD and RAD improve precipitation forecast by increasing the number of "hits" along with suppression of false alarms, especially at higher thresholds (Figures 6c and 6e). As seen in the reflectivity forecast skill, lower ETS values are obtained when a higher threshold for precipitation forecast is examined, which may result from a lower frequency of occurrence for heavy precipitation in both observation and model forecast (Figures 6b, 6d, and 6f). RAD and AMV_RAD produce higher ETSs than AMV and NoDA throughout 0–3 hr forecasts at all thresholds. The improvement in precipitation forecast is statistically significant at the 95% confidence level, except the 3-hr forecast at 15 mm threshold. Comparing AMV_RAD to RAD, the positive impact of AMVs over radar data is evident in 0–3 hr precipitation forecasts at 10 and 15 mm, though the improvement is not statistically significant at the 95% confidence level.

Since there exists variation and difference among cases in the distribution of observations available for assimilation and storm environments, more details in verification metrics from the five cases remain to be investigated. Given the above encouraging preliminary results obtained by assimilating AMV in conjunction with radar data, performance diagrams of CREF and HPRCP forecast from AMV_RAD are shown as representative of the best DA experiment (Figure 7). Except for 17 May 2019, all real cases generally exhibit similar evolution as the aggregated performance illustrated in Figures 5 and 6. In this context, the larger the threshold value, or the longer the

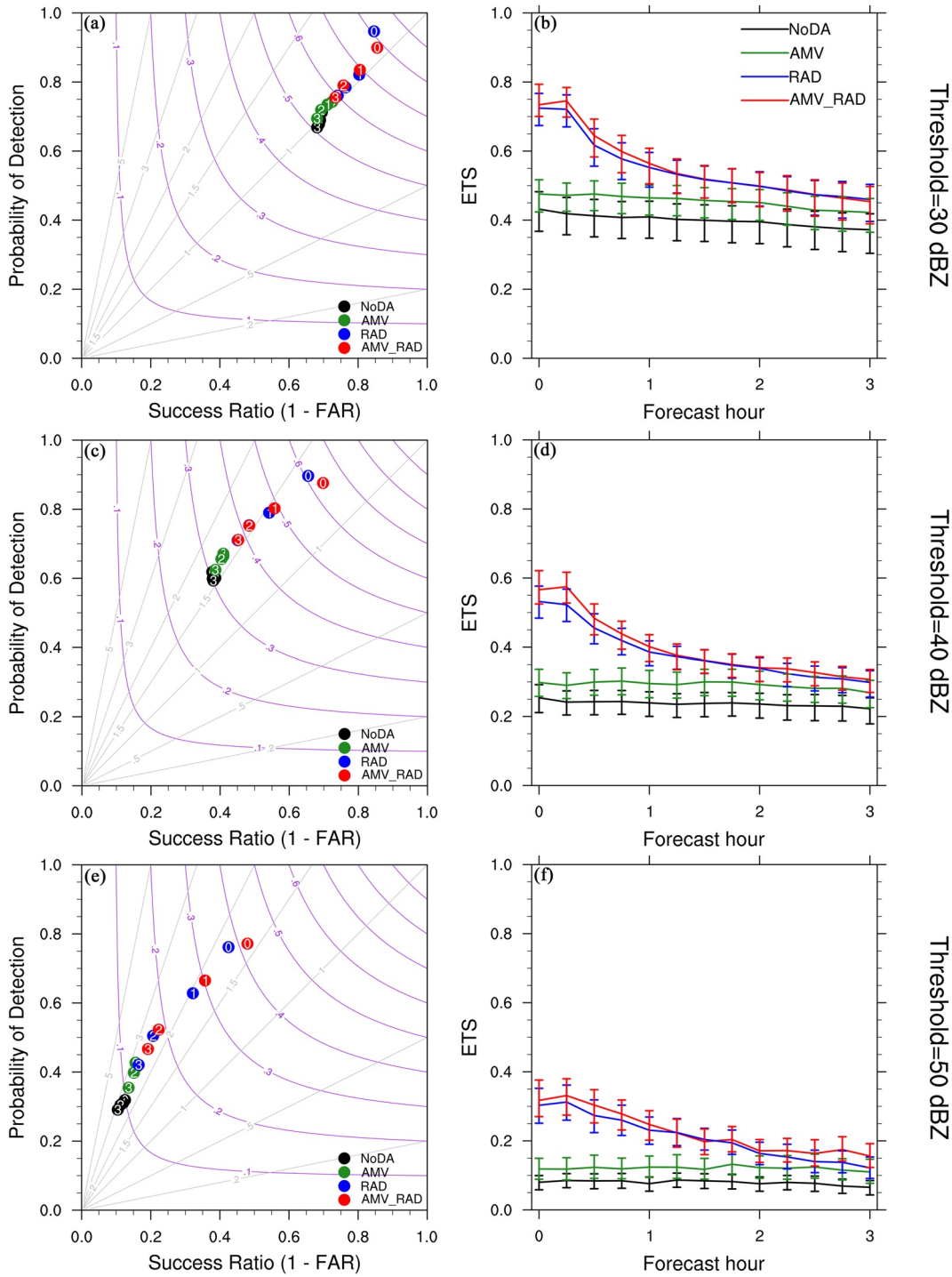


**Figure 5.** Aggregate score metrics of 0–3 hr composite reflectivity forecasts aggregated from each initialization hour from 19:00 to 03:00 UTC the next day across five cases for the NoDA (black), atmospheric motion vector (AMV) (green), RAD (blue), and AMV_RAD (red) experiments. (left) The performance diagrams, and (right) the equitable threat score for (a and b) 30, (c and d) 40, and (e and f) 50 dBZ thresholds, respectively. Results are shown for a neighborhood radius of 12-km. The numbers within the colored dots in the performance diagrams denote the forecast hour (i.e., 0-, 1-, 2-, and 3-hr forecasts). Error bars represent a 95% confidence interval.

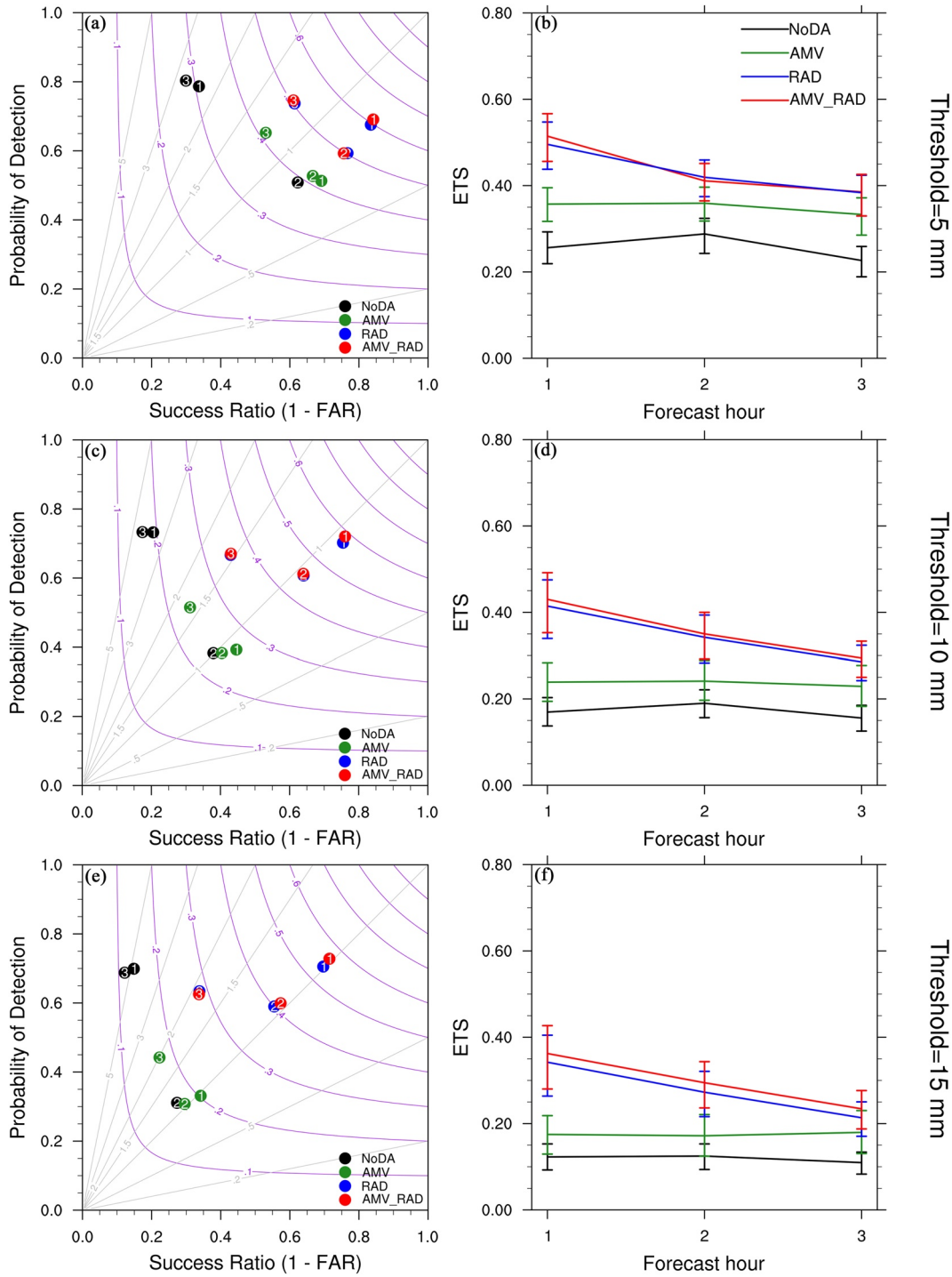


**Figure 6.** Same as in Figure 5, but for 1–3 hr hourly precipitation forecasts from each case relative to Stage IV rainfall estimates at thresholds of 5 (top), 10 (middle), and 15 mm (bottom), respectively.

forecast lead time, the lower the forecast skill in terms of score metrics. Nevertheless, persistent underprediction throughout 1–3 hr reflectivity forecasts at all thresholds from the 17 May case, which exhibits the poorest performance among five cases, can mostly be traced back to the failure to forecast scattered storms associated with the dryline across western Kansas (this detail will be discussed later). Compared to reflectivity forecast, the precipitation forecast skill exhibits more variability especially at higher thresholds. Forecasts from 30 April 2019 show the best performance in terms of high POD and CSI, although presence of spurious convection across the eastern

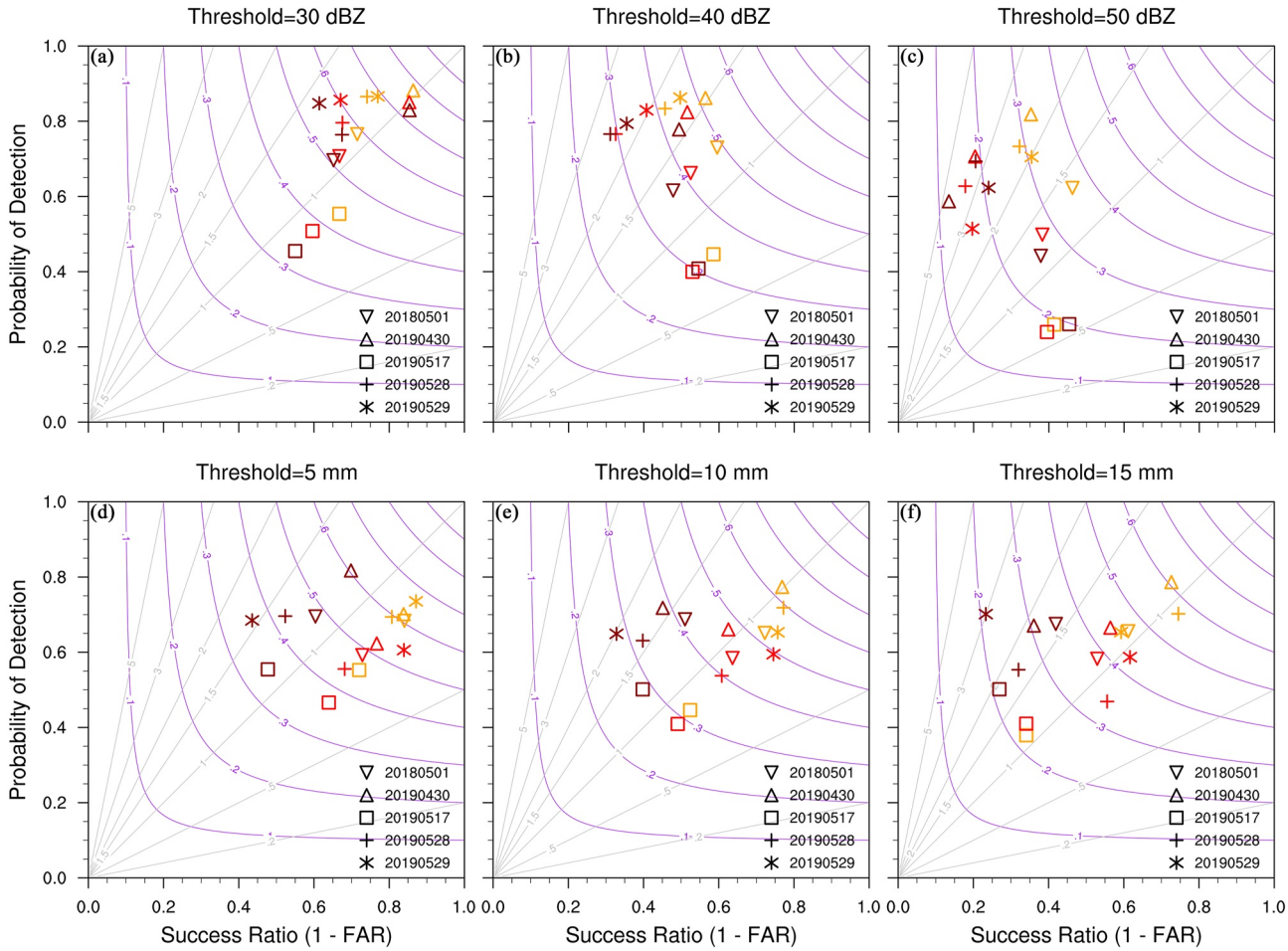


**Figure 7.** Performance diagram for 1–3 hr composite reflectivity forecasts from the AMV_RAD experiment in each case relative to Multi-Radar Multi-Sensor observations at thresholds of (a) 30, (b) 40, and (c) 50 dBZ, respectively. (d–f) Same as in (a–c), but for 1–3 hr hourly precipitation forecasts from each case relative to Stage IV rainfall estimates at thresholds of (d) 5, (e) 10, and (f) 15 mm, respectively. Orange, red, and brown colors denote the forecast hour of 1-, 2-, and 3-hr. Results are shown for a neighborhood radius of 12-km.

Oklahoma in the first 3-hr cycles leads to relatively high false alarms at 50 dBZ (Figure 7c). It is also noticed that BIAS values from 1 May 2018 generally remain within a reasonable range of 1.1–1.4 for reflectivity and 0.8–1.5 for precipitation, suggesting an overall good forecast skill. Performance of the other two events, namely 28 and 29 May 2019, lies in between the best and worst cases. Specifically, a more west-east oriented MCS occurring near the Missouri-Iowa border is predicted on 28 May 2019, while the observed MCS runs from southwest to northeast. On the other hand, the predicted storms on 29 May 2019 show a slightly westward displacement bias and move more slowly than those in the observations. Since the examined cases produced distinct storm modes under different weather situations, further detailed investigations into 1 May 2018, and 30 April 2019 are provided as representative of two best performing cases and 17 May 2019 case as the worst performing case.

### 4.2. 1 May 2018 Case

The 1 May 2018 case was characterized by the presence of a slow-moving cold front that arced across west-central Kansas into eastern Nebraska and a dryline extending southward from the front down to Texas during evening of the day. Under the favorable pre-storm environment, a quasi-linear convective system (also known as squall line) formed along the frontal boundary after 21:00 UTC and clustered storms emerged in the vicinity of surface low over Kansas as well. In addition, several discrete storms occurred along the dryline in southern Kansas. Numerous severe weather warnings were issued by the National Weather Service with subsequent storm reports including 13 tornadoes and several large hail and damaging wind events.

Changes in the model's dynamic and thermodynamic conditions due to assimilating AMV and/or radar data can be evaluated by displaying 850- and 200-hPa wind vectors and 850-hPa equivalent potential temperature fields (Figure 8). Differences between the analyses from three DA experiments and the corresponding fields from NoDA are shown in Figures 8c–8h. In NoDA, the frontal boundary is analyzed nearby the analyzed dryline (Figures 8a and 8b). AMV generates a turning point, which is the intersection between the more west-east oriented front and the dryline across the western Kansas, with a narrow corridor of low-level moisture ahead of the dryline (Figures 8c and 8d). The AMV experiment also enhances low-level cyclonic rotation for the low-pressure system (Figure 8c), the associated upper-level divergent flows (Figure 8d), as well as cold air behind the front. In contrast, with the assimilation of radar reflectivity and radial velocity, both RAD and AMV_RAD produce a much more pronounced cold pool near the center of the low (Figures 8e–8h). The temperature difference between RAD and NoDA is quite similar to that between AMV_RAD and NoDA. However, RAD does not significantly modify the analyzed wind field, while AMV_RAD strengthens both low-level cyclonic rotation and upper-level divergent airflow above convection, which is consistent with AMV observations (Figures 8g and 8h). A narrow region of increases in equivalent potential temperature (2–10 K) are evident ahead of the dryline in AMV_RAD, intensifying moisture gradient associated with the dryline.

Since wind analyses are directly related to AMV observations, the impact of assimilating AMVs on the divergence field, another critical factor affecting convective initiation and development, can be evaluated. The contours for MRMS CREF at 30 and 50 dBZ are overlapped over the divergence field, providing a source of independent verification for the analysis. Substantial discrepancies in the pattern between 200-hPa divergence and MRMS CREF (>30 dBZ) are observed in NoDA (Figure 9a). Owing to the improvement in mesoscale airflow analyses, upper-level divergence in AMV and RAD agrees better to the observed reflectivity in terms of location and magnitude, though displacement errors are still present (Figures 9b and 9c). In contrast, AMV_RAD successfully captures most of the localized enhanced divergence signatures in upper troposphere, which exhibits the best match with the area of >30 dBZ observed reflectivity (Figure 9d).

One to three hours reflectivity forecasts from all experiments are compared to the MRMS observations in Figure 10. A quasi-linear convective system (QLCS) developing along the front and severe storms initiated near the surface low are embedded in a comma shape (Figures 10a–10c). Although NoDA roughly forecasts the southwest to northeast orientation of the observed storms and northeastward movement, it generates two isolated MCSs at 1-hr forecast, rather than a QLCS over northeast Kansas and southwest Iowa (Figures 10d–10f vs. Figures 10a–10c). All three DA experiments AMV, RAD, and AMV_RAD improve storm forecast as evidenced by a more linear convective mode agreeing relatively well with the observed reflectivity. In general, AMV_RAD generates the best forecast skill in terms of areal coverage, storm mode, and storm orientation (Figures 10m–10o), which is consistent with aggregate statistics (Figure 5). The observed comma-shaped storms are particularly evident in analysis (not shown) and 1-hr forecast (Figure 10m) from AMV_RAD, while underprediction in trailing region of the QLCS over the northeastern Kansas is seen at later times (Figures 10n and 10o). Among three DA experiments, only AMV is able to fully produce the curly tail of the QLCS close to the border between Kansas and Nebraska at 2–3 hr forecasts, but with notably weaker intensity (Figures 10h and 10i). This indicates that the assimilation of high-resolution AMV data can correct the forecast of storm location and shape by capturing the mesoscale flow information embedded in the front structures. Additionally, the discrete cell near the border between Oklahoma and Kansas is considered to be resolved reasonably well by all experiments despite somewhat displacement and intensity errors. This is primarily attributed to the fact that this isolated cell is quite small and dissipates rapidly.

Concerning the precipitation, differences between 1-, 2-, and 3-hr accumulated precipitation (APCP) forecasts and the corresponding Stage IV estimates are illustrated for clarity (Figures 11d–11o). As discussed above, all DA experiments predict the QLCS structure better than NoDA, leading to improvements in APCP forecasts in terms of areal coverage, orientation, and amount. Compared to NoDA, overestimation in southeastern Nebraska and underestimation in tailing region of the QLCS are alleviated remarkably by all DA experiments (Figures 11g–11o vs. Figures 11d–11f). Meanwhile, some stronger or weaker biases associated with cells embedded in QLCS are present in the DA experiments. For example, RAD yields a smaller areal coverage of the precipitation across portions of eastern Nebraska and western Iowa, but still with notably heavier rainfall than the observed (Figures 11j–11l). As expected, the APCP forecasts from AMV_RAD are generally in better agreement with Stage IV rainfall estimates in terms of placement and amount (Figures 11m–11o).

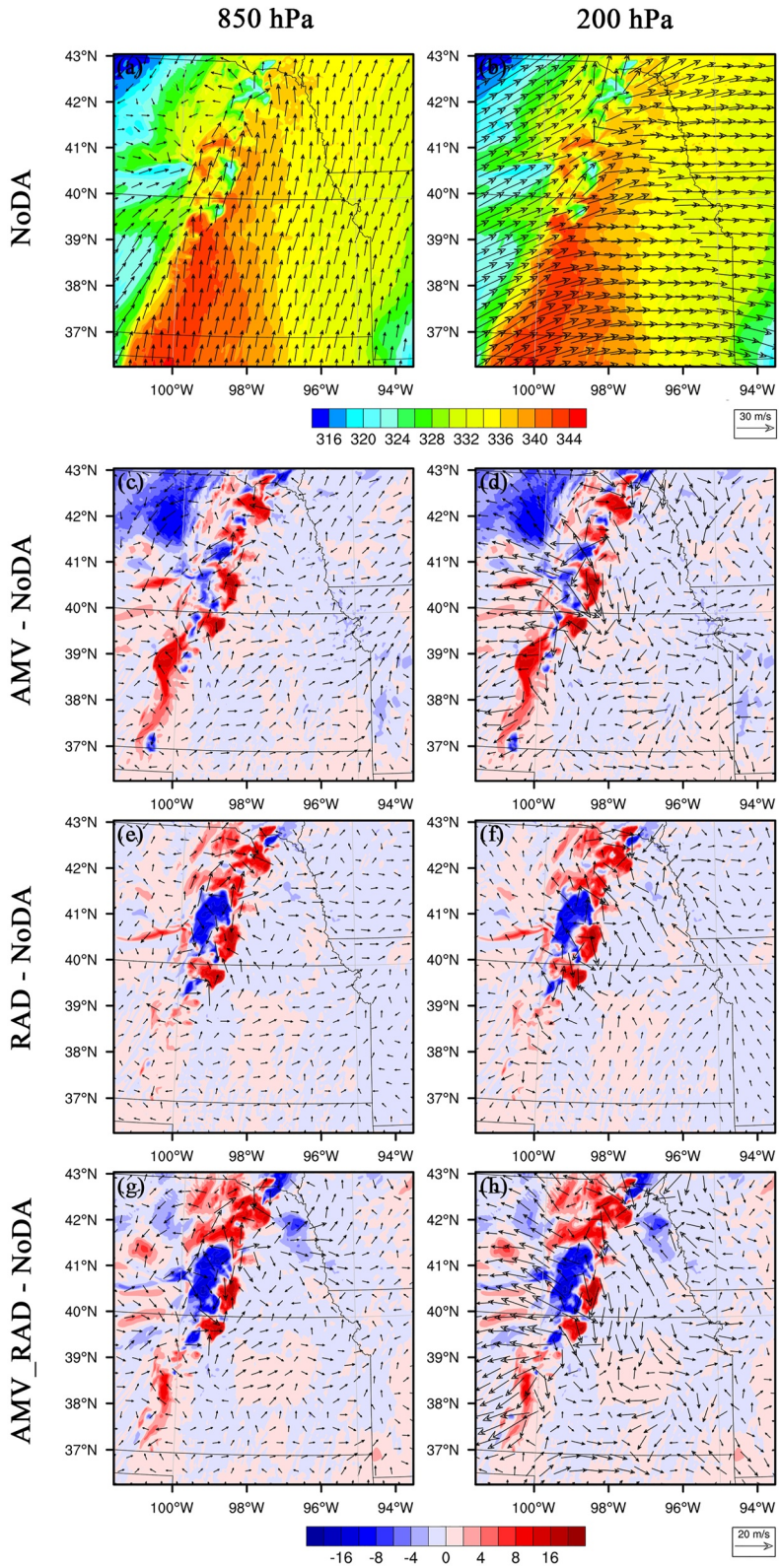


**Figure 8.** (a) 850 and (b) 200-hPa wind analyses (m $s^{-1}$, vectors) with 850-hPa equivalent potential temperature (K, shaded) indicated in colors for NoDA at 23:00 UTC 1 May 2018. The differences of analyzed 850-hPa equivalent potential temperature, 850 (left) and 200-hPa (right) wind (c and d) between atmospheric motion vector (AMV) and NoDA, (e and f) between RAD and NoDA, and (g and h) between AMV_RAD and NoDA at 23:00 UTC 1 May 2018.

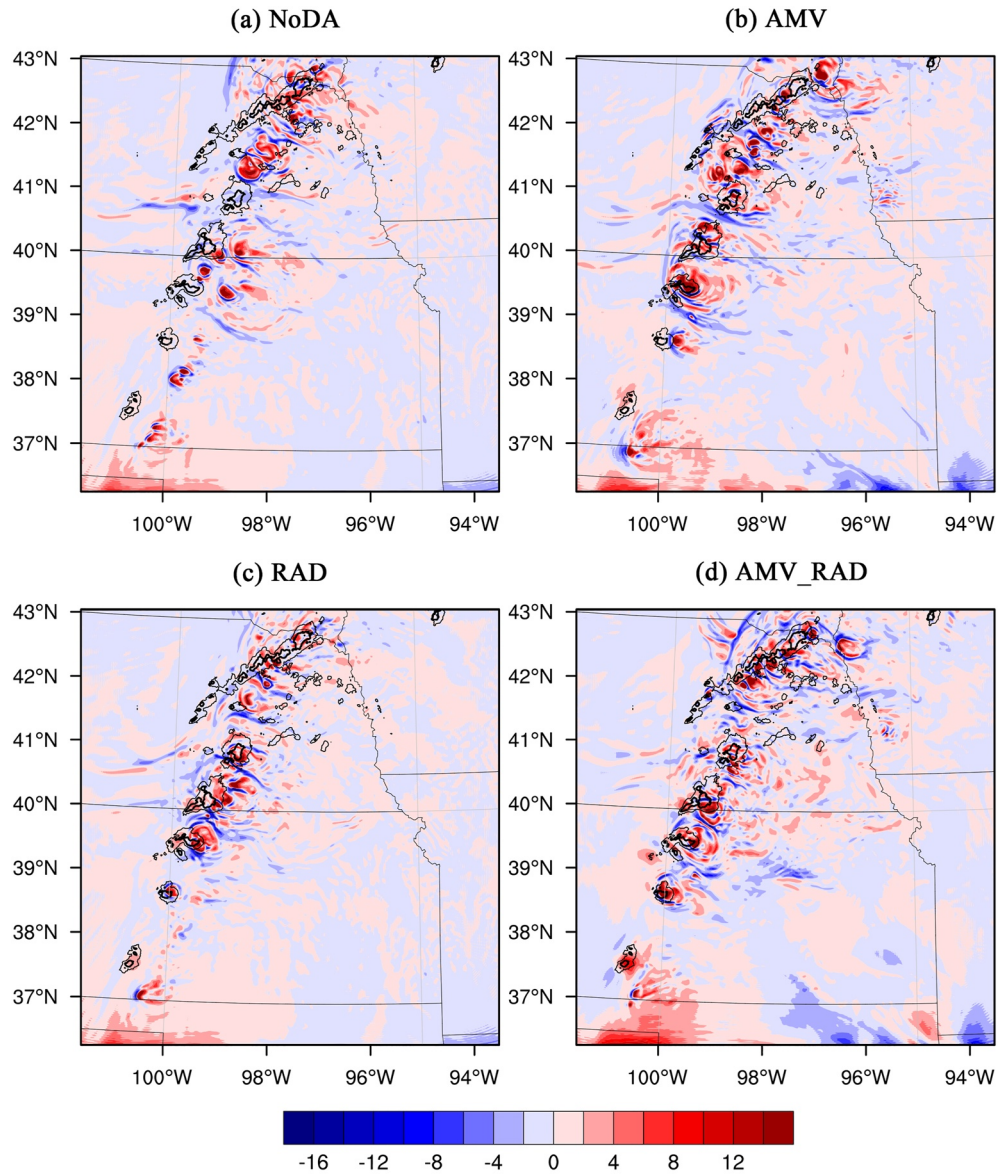


**Figure 9.** 200-hPa divergence analyses ($10^{-5}$ $s^{-1}$, shaded) for (a) NoDA, (b) atmospheric motion vector (AMV), (c) RAD, and (d) AMV_RAD experiments at 23:00 UTC 1 May 2018. The contours represent the Multi-Radar Multi-Sensor composite reflectivity at 30 and 50 dBZ.

### 4.3. 30 April 2019 Case

As a large-scale trough persisted across the western U.S. on 30 April 2019, two low-pressure systems across Colorado and the central Plains/Mid-Missouri Valley developed in response to the ejecting mid-level shortwave. A stationary front extended from Texas north-northeastward into eastern Kansas, and then transitioned to a warm front across central Illinois/Indiana (not shown). Due to the low-level southerly flow in south and east of the front maintaining a moist and buoyant low-level airmass, as well as strong lift provided by the front and low, mixed convective modes including linear and supercellular appeared in the vicinity of the above-mentioned frontal boundaries and within the warm sector (Figures 13a–13c). These thunderstorms are generally associated with surface-based instability in their inflow region, whereas other convection further to the north were initiated in a region of elevated instability. With the mean flow nearly parallel to the storm motions, a group of cells in a southwest to northeasterly direction led to significant flash flooding from central Oklahoma northeast into southwest

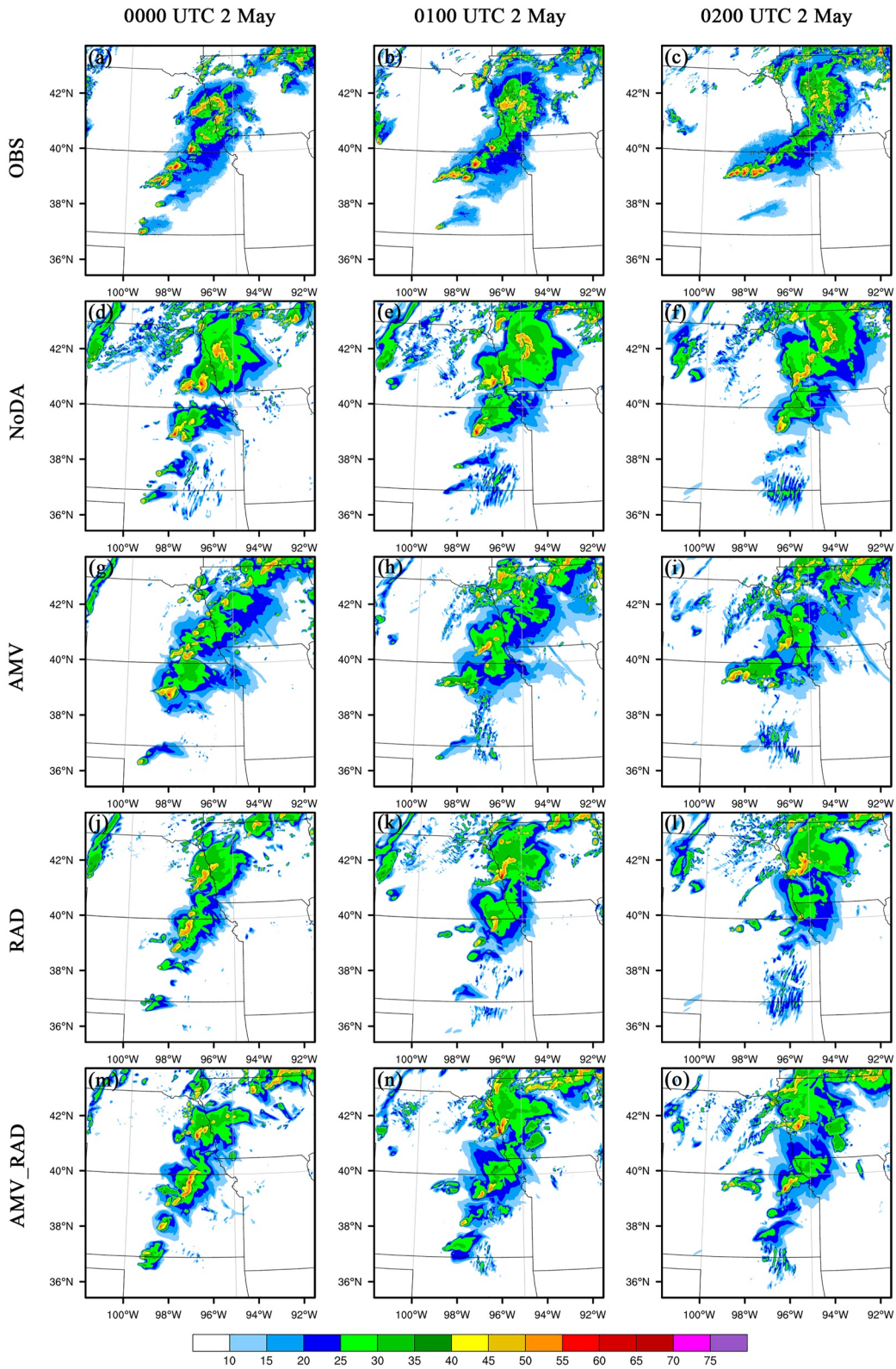


**Figure 10.** The composite reflectivity (dBZ, shaded) for (a–c) Multi-Radar Multi-Sensor observations, (d–f) NoDA, (g–i) atmospheric motion vector (AMV), (j–l) RAD, and (m–o) AMV_RAD experiments at (left, 1-hr forecast) 00:00 UTC, (middle, 2-hr forecast) 01:00 UTC, and (right, 3-hr forecast) 02:00 UTC 2 May 2018.

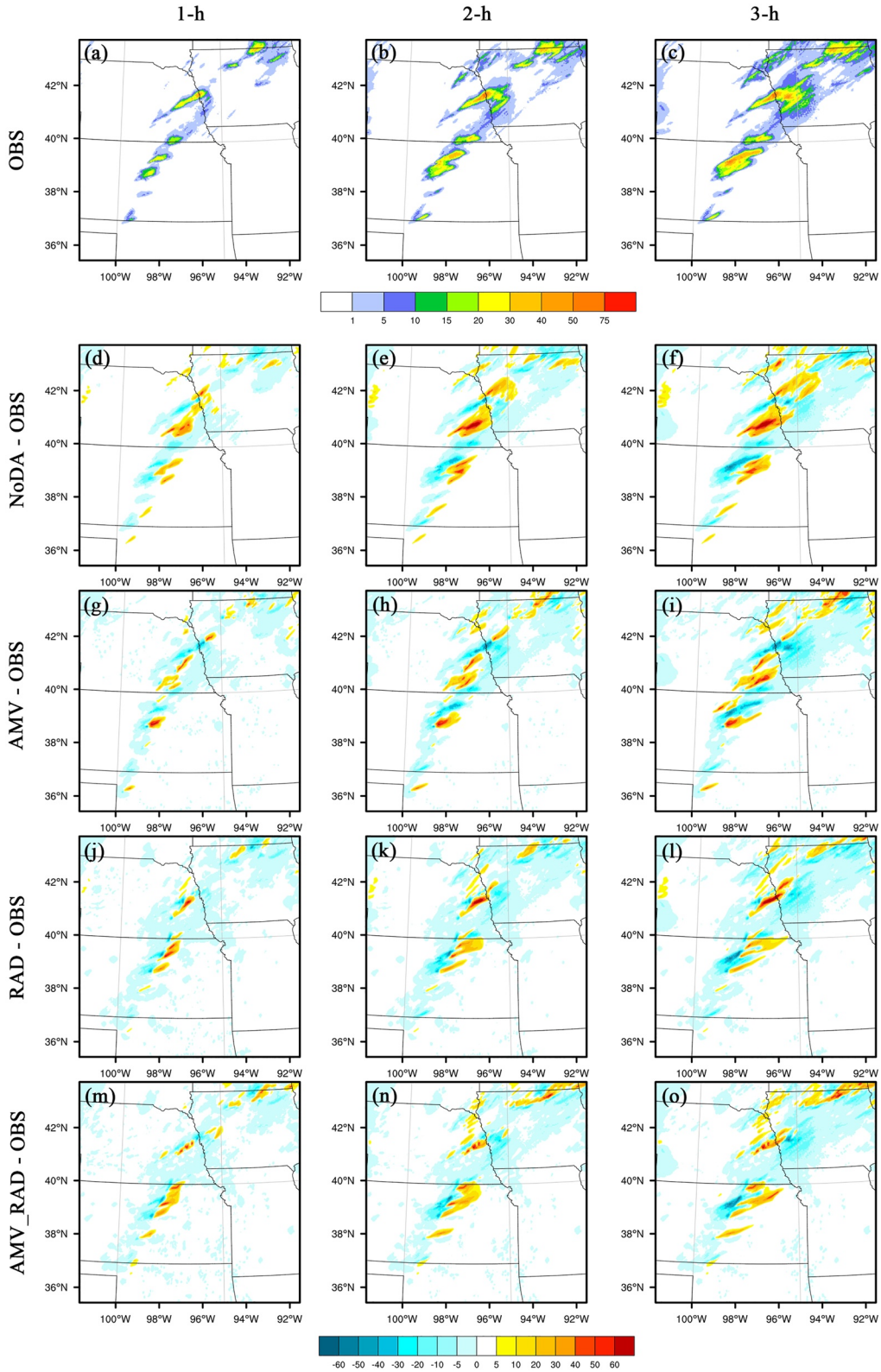


**Figure 11.** (a–c) Stage IV accumulate precipitation (APCP) estimates, and differences between the forecasted APCP from (d–f) NoDA, (g–i) atmospheric motion vector (AMV), (j–l) RAD, (m–o) AMV_RAD experiments initialized at 23:00 UTC 1 May 2018 and the Stage IV observation. The (left) 1-, (middle) 2-, and (right) 3-hr forecasts are shown.

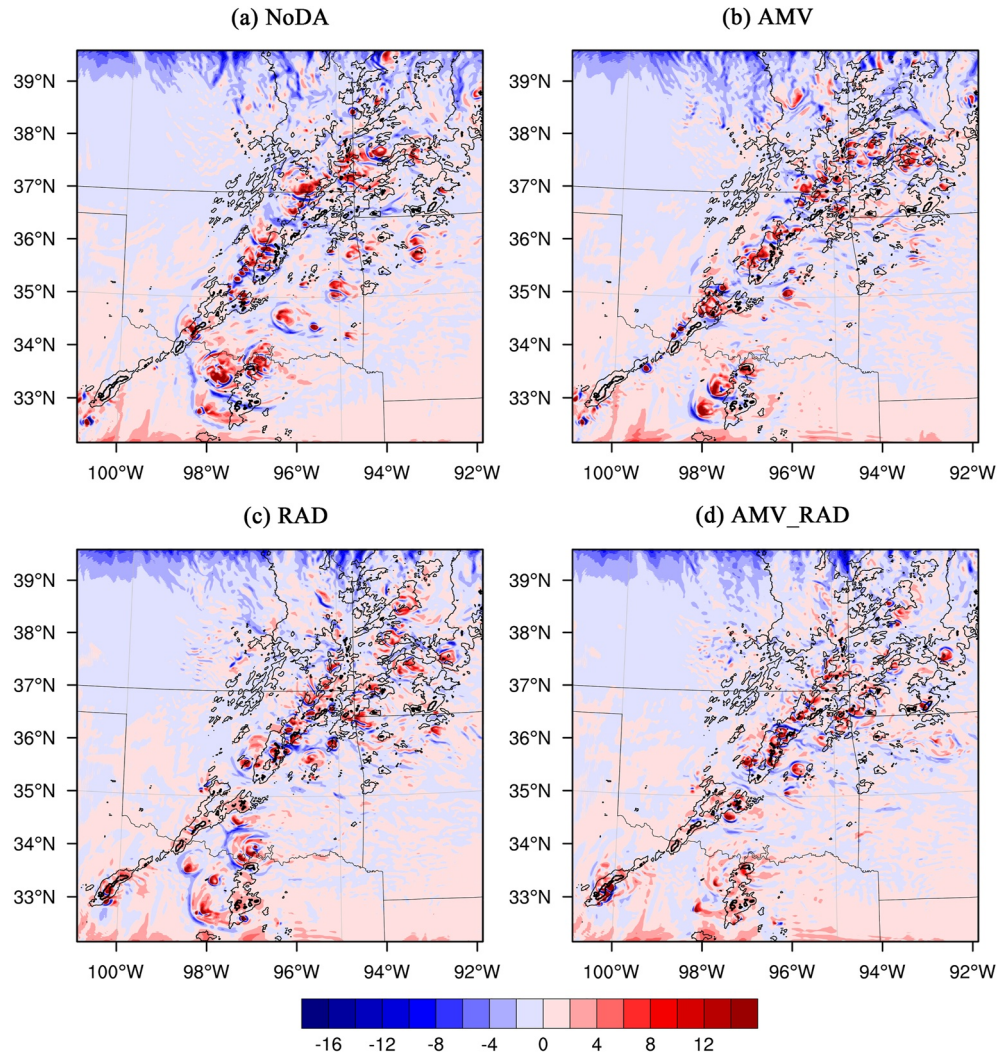


**Figure 12.** Same as in Figure 9, but for 23:00 UTC 30 April 2019.

Missouri (Figures 14a–14c). Other hazardous events including tornadoes, hail, and winds were also reported over northern Texas, central Oklahoma, northwestern Arkansas, southern Missouri, and southeastern Kansas.

Similar to the first examined case, 200-hPa divergence analyses at 23:00 UTC 30 April 2019 from different experiments are compared in Figure 12. We focus on this particular analysis and its subsequent forecasts because MRMS reflectivity shows that convection along the frontal boundaries has organized into a QLCS across central Oklahoma until 23:00 UTC (not shown). Moreover, cumulative benefit from the prior 5-hr DA cycling is also clearly perceptible. Though all experiments generally produce warm/moist and unstable conditions along south of the front (not shown), differences in the mesoscale and small-scale divergence fields are observed especially ahead of the front over Oklahoma and adjacent areas of northeastern Texas. Upper-level divergence associated with the QLCS in NoDA does not deviate substantially from the area of >30 dBZ observed reflectivity, characterized by a slightly slower-moving bias (Figure 12a). However, NoDA exhibits several spurious strong divergence cores in the warm sector across southeastern Oklahoma into northcentral Texas. It is evident that AMV successfully eliminates most spurious divergence features ahead of the QLCS, and a ribbon of divergence cores corresponds slightly better to the coverage of observed QLCS, indicating that detailed information on storm environment is introduced by assimilating AMV data (Figure 12b). However, a slightly westward displacement error is still present for the isolated storm located in the Northcentral Texas. This can be the consequence of difficulty of AMV DA in either reducing analysis errors for the variables not directly related to wind observations, such as humidity, or resolving quite fine airflow features (Zhao et al., 2021b). The assimilation of radar data also helps

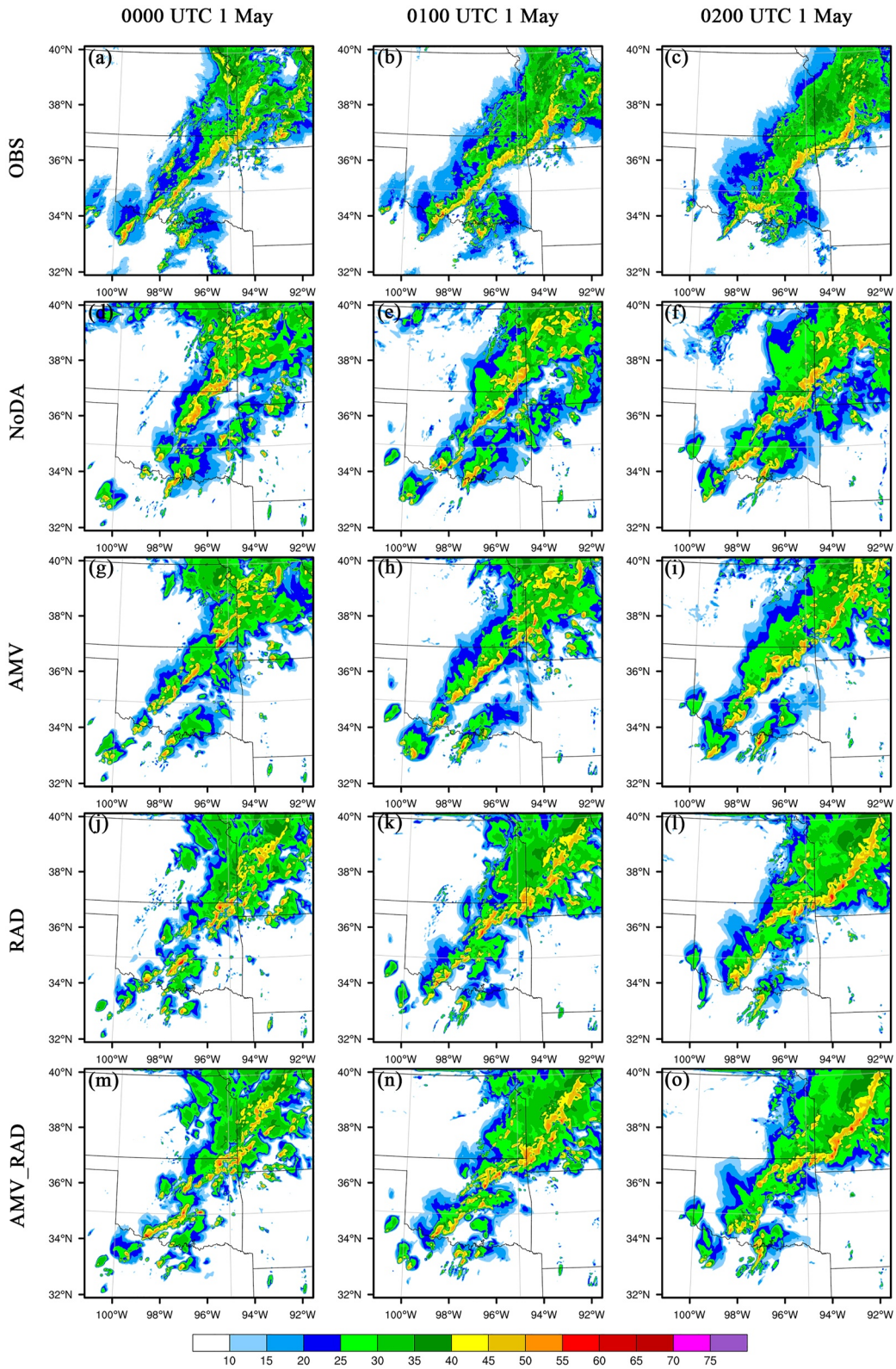


**Figure 13.** Same as in Figure 10, but for (left, 1-hr forecast) 00:00 UTC 1 May, (middle, 2-hr forecast) 01:00 UTC 1 May, and (right, 3-hr forecast) 02:00 UTC 1 May 2019.

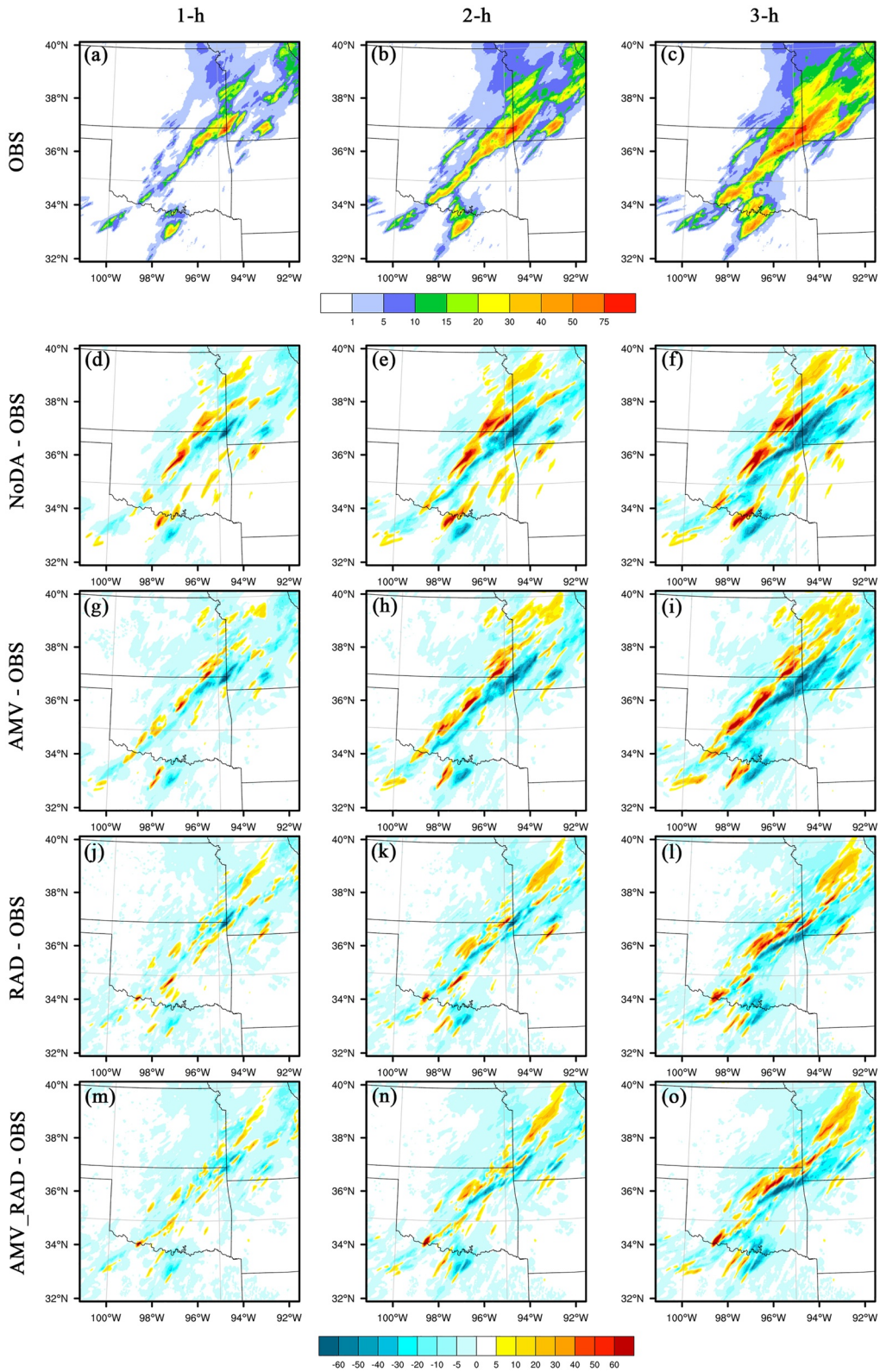


**Figure 14.** Same as in Figure 11, but for the Stage IV observations and accumulated precipitation forecast difference initialized at 23:00 UTC 30 April 2019.

suppress spurious cells and alleviate displacement errors, despite a couple of spurious divergence cores on both sides of the border between Texas and Oklahoma (Figure 12c). The pattern of divergence analysis in AMV_RAD agrees reasonably well with the coverage of reflectivity observations, as evidenced by approximately lining up with the developing arc of QLCS and coinciding with some deeper convective elements embedded within a large cluster of convection in southwestern Missouri (Figure 12d).

Between 20:00 and 23:00 UTC 30 April, a group of cells initiated along the frontal boundary of a southwest to northeasterly direction in central Oklahoma, and then organized into a QLCS moving northeastward (not shown). Further to the north, two additional large clusters of convection developed in a region of elevated instability and moved eastward, leading to a long period of heavy rain from eastern Kansas into western Missouri. At 23:00 UTC the northern flank of QLCS meets the convection clusters (Figure 13a) and they eventually merge into one linear system (Figures 13b and 13c). Meanwhile, a smaller cluster of storms developed within the warm sector close to the Texas/Oklahoma border also moves northeastward and then combines with the south flank of QLCS. NoDA produces excessive spurious convection in a quasi-linear mode that are approximately parallel to southeast of the aforementioned QLCS, while the QLCS across southern Oklahoma into northern Texas is underpredicted and less organized (Figures 13d–13f). Moreover, some deeper convective elements embedded in the elevated convection clusters far north are overestimated in NoDA. In the AMV experiment, not only is the spurious convection over eastern Oklahoma suppressed to a certain extent, but also the predicted QLCS is significantly better organized in terms of a linear mode (Figures 13g–13i). This implies that the assimilation of high-resolution AMVs can improve storm pattern and development by providing a better representation of storm environment. Although RAD and AMV_RAD correctly analyze all storms near or at their observed locations (not shown), 1-hr reflectivity forecast show that the QLCS is relatively loosely structured in RAD compared to the observed reflectivity, especially over central and southern Oklahoma (Figure 13j vs. Figure 13a). AMV_RAD alleviates this problem by providing an improved linear mode of storms over southern Oklahoma in the 1-hr forecast, though both RAD and AMV_RAD slightly overestimate the large reflectivity values in leading edge of the QLCS (Figures 13j–13o). Overall, the areal coverage, storm mode, and storm orientation in AMV_RAD agree reasonably well with the MRMS observations, especially in 3 hr forecast. All DA experiments miss the merging of QLCS and small storm cluster from northeastern Texas. This occurs because the predicted storm cluster in that area moves more slowly than the observed.

South and east of the front, southerly low-level flow maintained a moist and buoyant low-level airmass characterized by near 70°F dewpoints (not shown). Combination of abundant moisture and strong instability along and south of the front is favorable for heavy rainfall to occur. As the southwest-northeast oriented QLCS moved northeastward and merged with larger storm clusters induced by the elevated instability, a long period of heavy rain occurred across central Oklahoma into western Missouri (Figures 14a–14c). It is clearly seen that the rainband along QLCS in NoDA moves more slowly than the observed, and NoDA overpredicts the rainfall amount ahead of the rainband (Figures 14d–14f). Atmospheric Motion Vector suppresses spurious precipitation ahead of QLCS throughout 1–3 hr APCP forecast (Figures 14g–14i) and reduces the rainfall bias for isolated rain clusters near the Texas/Oklahoma border. It is evident that both RAD and AMV_RAD significantly improve the areal coverage and magnitude of 1–3 hr APCP forecasts. By comparing these two DA experiments in more detail, slightly better heavy rainfall position and cell alignment are seen in AMV_RAD at 2–3 hr forecasts (Figures 14j–14l vs. Figures 14m–14o).

### 4.4. 17 May 2019 Case

The third analyzed event occurred on 17 May 2019 in the Central Plains. An intensifying surface low over eastern Colorado/western Nebraska, along with eastward progression of an upper trough, caused strengthening low/mid-tropospheric wind fields and favorable deep-layer shear for organized and severe storms (not shown). In addition, a slow-moving dryline extended southward from the surface low, while a warm front extended eastward. Between 22:00 and 03:00 UTC, discrete storms initiated in the vicinity of intersection between the front and dryline and evolved into a linear mode spreading northeast through the destabilizing warm sector (Figures 16a–16c). Farther south along the dryline, a couple of supercell storms initiated around 23:30 UTC and progressed northeastward. As the storms moved eastward/northeastward, severe weather including tornadoes, large hail, and damaging winds occurred from southwest Kansas into western Nebraska.

Around 01:00 UTC 18 May, the surface low moved into the Colorado/Nebraska border along with a cold front stretched southward from the low system, and the slow-moving dryline persisted across the western portions of Kansas and Texas (not shown). From the 700-hPa equivalent potential temperature and water vapor mixing ratio illustrated in Figure 15, it is seen that temperature gradient associated with the cold front in NoDA is relatively weak (Figure 15a). In contrast, AMV produces remarkably colder airmass and larger northwestern wind behind the front, leading to much stronger temperature gradient near the frontal boundary (Figure 15c). Stronger cold surges are also observed in RAD and AMV_RAD, though with slightly weaker intensities (Figures 15e and 15g). Distribution of moisture is another key factor that may affect convective initiation and evolution. A significantly narrower band of large water vapor content exceeding 6 g $kg^{-1}$ is observed over western Nebraska in all three DA experiments, whereas NoDA has a broad area of high humidity in the warm sector. It is also noted that moisture gradient associated with the dryline across western Kansas and Texas is as sharp in NoDA (Figure 15b). However, the benefit from assimilating AMVs or radar data is otherwise negligible in this region. This occurs likely because the dryline is mainly attributed to sharp moisture gradient without distinct dynamic convergence, while the humidity field is not directly correlated with the targeted observations in this case study.

By adjusting storm dynamic and thermodynamic environment, AMV significantly corrects the displacement error for storm clusters over Nebraska when comparing the AMV and NoDA experiment, despite of larger areal coverage of stratiform across the eastern portions of Nebraska and Kansas especially in the 3-hr forecast (Figures 16g–16i vs. Figures 16d–16f). It is evident that reflectivity analyses in both RAD and AMV_RAD agree better with the MRMS observations than in AMV (Figures 16j and 16m vs. Figure 16g). However, the weak discrete storms in western Kansas are not maintained in the 2–3 hr forecasts for RAD and AMV_RAD, as a result of inability in producing a sharp enough moisture gradient associated with the dryline in this region (Figures 16j–16o). Comparison of AMV_RAD against RAD reveals that RAD overpredicts the coverage of convective precipitation over Nebraska, while AMV_RAD underpredicts the stratiform precipitation in that area. Additionally, the light spurious stratiform precipitation over eastern Colorado in RAD is suppressed in AMV_RAD to some degree. AMV_RAD partially alleviates overestimation of stratiform coverage across the eastern Nebraska/Kansas in the 3-hr forecast for AMV, but RAD does not produce the stratiform rain there. The above indicates that the assimilation of radar data only has difficulty in properly defining the environment for very weak convective or stratiform precipitation, which leads to rapid dissipation of weak convection or missing stratiform for this particular case. As GOES-16 AMVs can provide mesoscale flow information in the general vicinity of boundaries, AMV DA may help improve forecasts of convection location and development owing to better representation of storm environment and even enhanced divergence. However, difficulty in adjusting the humidity analysis based on AMV observations in this case limits its impact on predicting isolated cells associated with the dryline.

## 5. Summary and Conclusions

In this research, four types of experiments are carried out to assess the increase in forecast skill from assimilating new-generation geostationary satellite retrieved AMV in addition to radar observations to short-term convective-scale NWP with five high-impact weather events that occurred over the Great Plains of the United States in spring 2018 and 2019. The NoDA experiment does not assimilate any observations. The RAD and AMV experiments assimilate WSR-88D radar and GOES-16 AMV data alone, respectively. Furthermore, the AMV_RAD experiment assimilates AMVs in addition to radar observations. Comparison of analyses and forecasts from four types of experiments indicates improvements on model initial conditions and short-term severe weather forecasts by assimilating AMV and radar observations, as well as the added benefit of GOES-16 AMVs over conventional radar data.

Both subjective and objective verification against the MRMS CREF and Stage IV hourly rainfall estimates were performed on 0–3 hr forecasts from 19:00 to 03:00 UTC the next day. Statistical metrics, including RMSEs and biases of wind components, neighborhood-based ETS, POD, FAR, BIAS, and CSI of reflectivity and precipitation forecasts, were calculated for all cases. The aggregate results suggest that analyses and forecasts within the simulation domain are improved by both AMV and radar data during DA cycles. Overall, AMV_RAD performs slightly better than RAD and much better than AMV, in terms of POD, SR, CSI, and BIAS values for 0–3 hr reflectivity and precipitation forecasts.

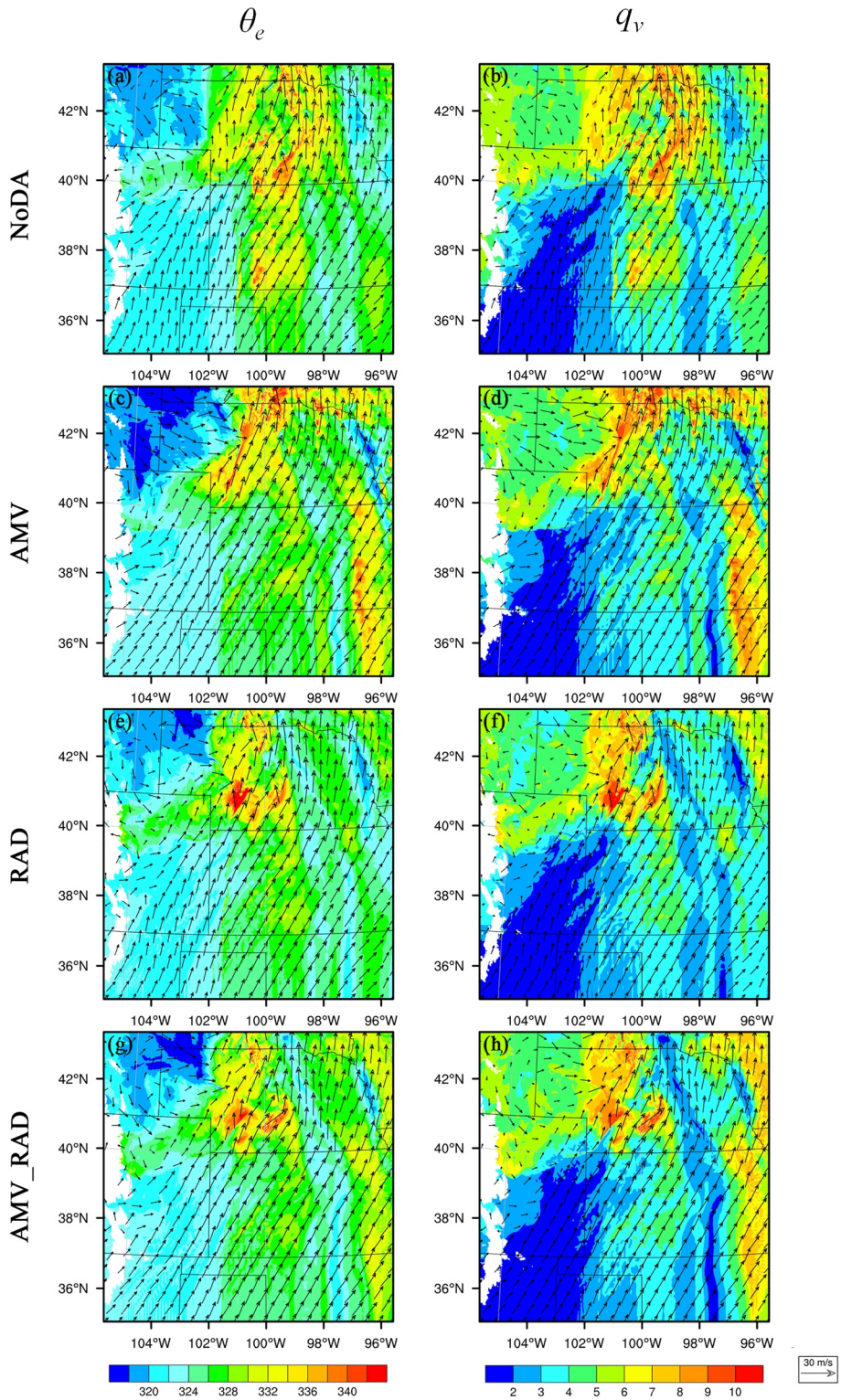


**Figure 15.** 700-hPa wind analyses (m $s^{-1}$, vectors) with (left) equivalent potential temperature (K, shaded) and (right) water vapor mixing ratio (g $kg^{-1}$, shaded) indicated in colors for (a and b) NoDA, (c and d) atmospheric motion vector (AMV), (e and f) RAD, and (g and h) AMV_RAD experiments at 01:00 UTC 18 May 2019.

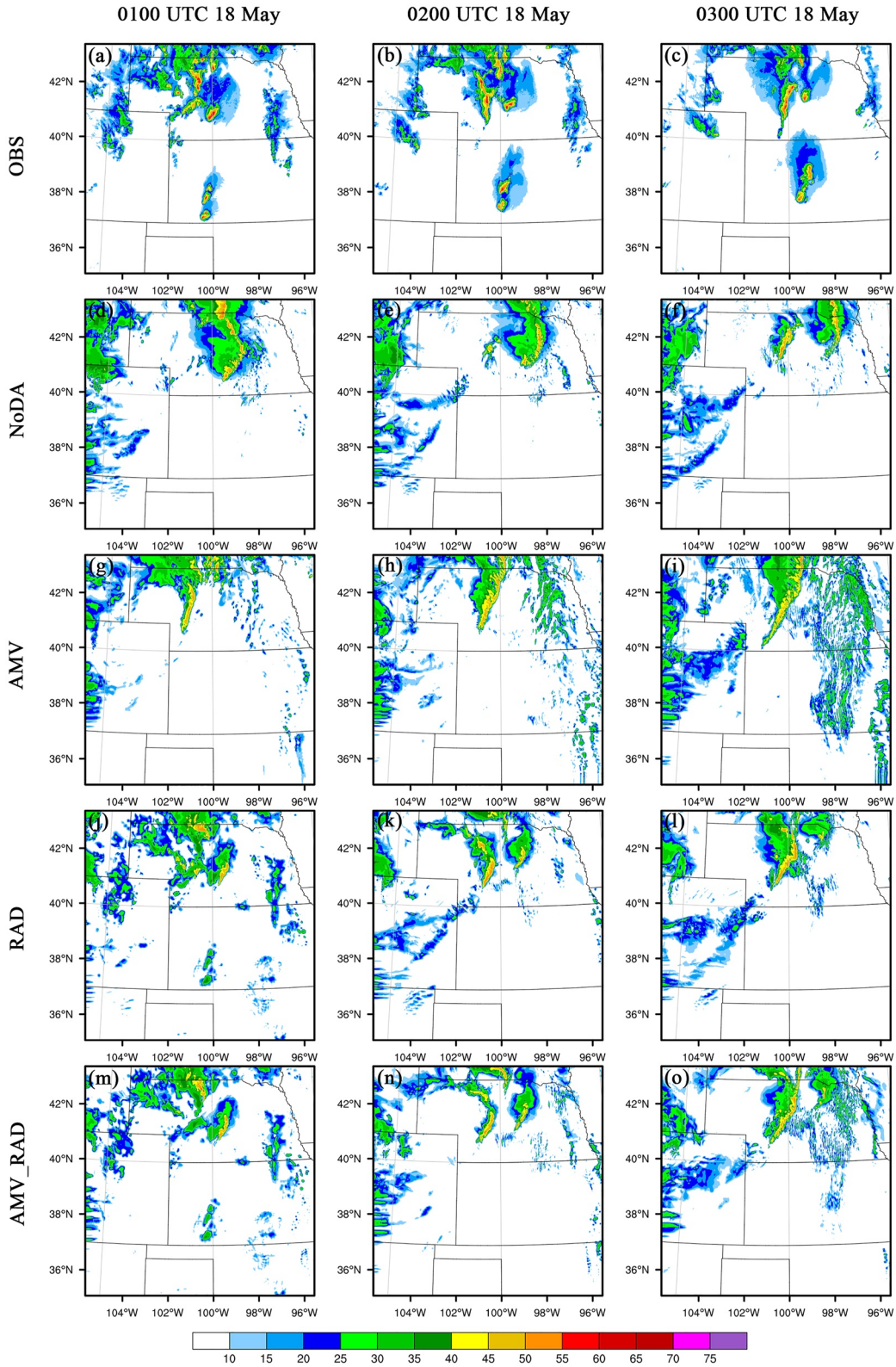


**Figure 16.** Same as in Figure 10, but for (left, analysis) 01:00 UTC, (middle, 1-hr forecast) 02:00 UTC, and (right, 2-hr forecast) 03:00 UTC 18 May 2019.

Qualitative comparison also reveals that AMV_RAD generally yields slightly better reflectivity and precipitation forecasts. The two best-performing cases (1 May 2018 and 30 April 2019) and the worst case (17 May 2019) are investigated in more detail. For the 1 May 2018 event, all three DA experiments AMV, RAD, and AMV_RAD improve the storm forecast in terms of areal coverage, storm mode, and storm orientation, especially evidenced by a more linear convective mode agreeing better with the reflectivity observations. This is mainly attributed to improvements in the analysis of boundary location and localized enhanced divergence signatures by assimilating AMVs and radar data. However, only the AMV experiment successfully captures the curly tail of QLCS at 2–3 hr forecasts, though with a weaker intensity. By adjusting the storm thermodynamic environment and mesoscale flow information embedded in the frontal boundary, high-resolution AMV data can correct forecasts of storm location and shape. For the 30 April 2019 case, assimilation of both AMV and radar observations also yields skillful forecast of QLCS as well as suppressing spurious convection in warm sector of the front. Although RAD corrects storm locations in the initial conditions, the QLCS is relatively loosely structured at 1-hr forecast compared to the observed reflectivity. Both AMV and AMV_RAD generate a better organized QLCS in a linear mode.

For 17 May 2019, the AMV experiment typically predicts severe storms that are more appropriately placed by adjusting the storm thermodynamic environment, but it fails to forecast discrete storms associated with the dryline. Reflectivity analyses from AMV_RAD and RAD are better than that from AMV, as expected. Nevertheless, weak discrete storms along the dryline analyzed by RAD and AMV_RAD could not be maintained in later forecasts, as a result of inability in producing a sharp enough moisture gradient associated with the dryline.

In general, the above results all indicate the added value of GOES-16 AMVs on conventional radar observations, even though the number of assimilated radar data is usually three orders of magnitude larger than that of AMVs. Although the five real cases are characterized by different weather situations and distinct storm modes, performance for all five cases are consistently improved with the assimilation of AMV data. Nevertheless, persistent underprediction throughout the 1–3 hr reflectivity forecasts at all thresholds is observed in the 17 May 2019 case with the poorest performance even if AMV is assimilated.

Although this study shows preliminary encouraging results of assimilating the high-resolution GOES-16 AMVs in conjunction with radar observations in convective scale NWP, some drawbacks are also observed such as the difficulty in capturing or maintaining weak discrete storms associated with moisture gradient (e.g., dryline) during the short-term forecast. More research is warranted to explore adding more dense observations to adjust the humidity field. For example, more experiments will be performed to investigate combined impact of the GOES-R derived layered precipitable water product (Pan, Gao, & Wang, 2021) together with AMVs and radar data in convective-scale analyses and forecasts. Deficiency in the static background errors of 3DVAR can be addressed by using flow-dependent background error covariances derived from ensemble forecasts in future.

## Data Availability Statement

The source codes of WRF model version 3.6.1 could be downloaded after filling in the E-mail address (https://www2.mmm.ucar.edu/wrf/users/download/get_source.html). The GOES-16 ABI Level 2 (L2) AMV data were obtained online (https://www.ncdc.noaa.gov/airs-web/search). The Multi-Radar Multi-Sensor (MRMS) and the Stage IV rainfall products, and the aggregate forecast statistics for composite reflectivity and APCP are accessible online (https://doi.org/10.5281/zenodo.4495919).

**Acknowledgments**
This work was jointly supported by Fengyun Application Pioneering Project (FY-APP-2021.0209), and the China Meteorological Administration Training Centre Key Research Program (2022CMATCZD14). The work was also supported by NOAA/Office of Oceanic and Atmospheric Research under NOAA-University of Oklahoma Cooperative Agreement NA16OAR4320115 and NA21OAR4320204, U.S. Department of Commerce. Drs. Pamela Heinselman and Qin Xu kindly provided internal review which led to improvement of the manuscript. The work was done when the first author visited CIWRO/the University of Oklahoma (OU) as a research scientist. Computing resources were provided by OU Supercomputing Center for Education & Research (OSCER).

## References

Baldwin, M. E., & Mitchell, K. E. (1997). The NCEP hourly multisensor US precipitation analysis for operations and GCIP research. In *Presented at the preprints, 13th Conference on Hydrology* (pp. 54–55). Amer. Meteor. Soc.

Benjamin, S. G., Grell, G. A., Brown, J. M., Smirnova, T. G., & Bleck, R. (2004). Mesoscale weather prediction with the RUC hybrid isentropic–terrain-following coordinate model. *Monthly Weather Review*, *132*(2), 473–494. https://doi.org/10.1175/1520-0493(2004)132<0473:MWPWTR>2.0.CO;2

Cherubini, T., Businger, S., Velden, C., & Ogasawara, R. (2006). The impact of satellite-derived atmospheric motion vectors on mesoscale forecasts over Hawaii. *Monthly Weather Review*, *134*(7), 2009–2020. https://doi.org/10.1175/MWR3163.1

Clark, A. J., Gallus, W. A., & Weisman, M. L. (2010). Neighborhood-based verification of precipitation forecasts from convection-allowing NCAR WRF model simulations and the operational NAM. *Weather and Forecasting*, *25*(5), 1495–1509. https://doi.org/10.1175/2010WAF2222404.1

Daniels, J., Bresky, W., Wanzong, S., Velden, C., & Berger, H. (2012). GOES-R advanced baseline imager (ABI) algorithm theoretical basis document for derived motion winds (No. Version 2.5). Retrieved from https://www.star.nesdis.noaa.gov/goesr/docs/ATBD/DMW.pdf

Dudhia, J. (1989). Numerical study of convection observed during the winter monsoon experiment using a mesoscale two-dimensional model. *Journal of the Atmospheric Sciences*, *46*(20), 3077–3107. https://doi.org/10.1175/1520-0469(1989)046<3077:NSOCOD>2.0.CO;2

Elsberry, R. L., Hendricks, E. A., Velden, C. S., Bell, M. M., Peng, M., Casas, E., & Zhao, Q. (2018). Demonstration with special TCI-15 datasets of potential impacts of new-generation satellite atmospheric motion vectors on navy regional and global models. *Weather and Forecasting*, *33*(6), 1617–1637. https://doi.org/10.1175/WAF-D-17-0168.1

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

Gao, J., Fu, C., Stensrud, D. J., & Kain, J. S. (2016). OSSEs for an ensemble 3DVAR data assimilation system with radar observations of convective storms. *Journal of the Atmospheric Sciences*, *73*(6), 2403–2426. https://doi.org/10.1175/JAS-D-15-0311.1

Gao, J., Smith, T. M., Stensrud, D. J., Fu, C., Calhoun, K., Manross, K. L., et al. (2013). A real-time weather-adaptive 3DVAR analysis system for severe weather detections and warnings. *Weather and Forecasting*, *28*(3), 727–745. https://doi.org/10.1175/WAF-D-12-00093.1

Gao, J., & Stensrud, D. J. (2012). Assimilation of reflectivity data in a convective-scale, cycled 3DVAR framework with hydrometeor classification. *Journal of the Atmospheric Sciences*, *69*(3), 1054–1065. https://doi.org/10.1175/JAS-D-11-0162.1

Gao, J., Xue, M., Brewster, K., & Droegemeier, 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

Hartman, C. M., Chen, X., Clothiaux, E. E., & Chan, M.-Y. (2021). Improving the analysis and forecast of Hurricane Dorian (2019) with simultaneous assimilation of GOES-16 all-sky infrared brightness temperatures and tail Doppler radar radial velocities. *Monthly Weather Review*, *149*(7), 2193–2212. https://doi.org/10.1175/MWR-D-20-0338.1

Hong, S.-Y., Noh, Y., & Dudhia, J. (2006). A new vertical diffusion package with an explicit treatment of entrainment processes. *Monthly Weather Review*, *134*(9), 2318–2341. https://doi.org/10.1175/MWR3199.1

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

Huang, Y., Wang, X., Mahre, A., Yu, T.-Y., & Bodine, D. (2022). Impacts of assimilating future clear-air radial velocity observations from phased array radar on convection initiation forecasts: An observing system simulation experiment study. *Monthly Weather Review*, *150*(7), 1563–1583. https://doi.org/10.1175/MWR-D-21-0199.1

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. *Monthly Weather Review*, *148*(5), 1829–1859. https://doi.org/10.1175/MWR-D-19-0379.1

Jones, T. A., Wang, X., Skinner, P., Johnson, A., & Wang, Y. (2018). Assimilation of *GOES-13* imager clear-sky water vapor (6.5 μ m) radiances into a Warn-on-Forecast system. *Monthly Weather Review*, *146*(4), 1077–1107. https://doi.org/10.1175/MWR-D-17-0280.1

Kim, D.-H., & Kim, H. M. (2018). Effect of assimilating Himawari-8 atmospheric motion vectors on forecast errors over East Asia. *Journal of Atmospheric and Oceanic Technology*, *35*(9), 1737–1752. https://doi.org/10.1175/JTECH-D-17-0093.1

Kim, M., Kim, H. M., Kim, J., Kim, S.-M., Velden, C., & Hoover, B. (2017). Effect of enhanced satellite-derived atmospheric motion vectors on numerical weather prediction in East Asia using an adjoint-based observation impact method. *Weather and Forecasting*, *32*(2), 579–594. https://doi.org/10.1175/WAF-D-16-0061.1

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

Lean, K., & Bormann, N. (2019). Moving to GOES-16: A new generation of GOES AMVs (EUMETSAT/ECMWF fellowship Programme research report No. No. 49). Retrieved from https://www.ecmwf.int/node/18860

Lean, K., Bormann, N., & Salonen, K. (2016). Assessment of Himawari-8 AMV data in the ECMWF system (EUMETSAT/ECMWF Fellowship Programme Research Reports No. No. 42). Retrieved from https://www.ecmwf.int/node/16972

Le Marshall, J., Seecamp, R., Dunn, M., Velden, C., Wanzong, S., Puri, K., et al. (2008). Shorter contribution the contribution of locally generated MTSat-1R atmospheric motion vectors to operational meteorology in the Australian region. *Australian Meteorological Magazine*, *57*, 359–365.

Li, J., Li, J., Velden, C., Wang, P., Schmit, T. J., & Sippel, J. (2020). Impact of rapid-scan-based dynamical information from GOES-16 on HWRF hurricane forecasts. *Journal of Geophysical Research: Atmospheres*, *125*(3), e2019JD031647. https://doi.org/10.1029/2019JD031647

Lim, A. H. N., Jung, J. A., Nebuda, S. E., Daniels, J. M., Bresky, W., Tong, M., & Tallapragada, V. (2019). Tropical cyclone forecasts impact assessment from the assimilation of hourly visible, shortwave, and clear-air water vapor atmospheric motion vectors in HWRF. *Weather and Forecasting*, *34*(1), 177–198. https://doi.org/10.1175/WAF-D-18-0072.1

Mallick, S., & Jones, T. A. (2020). Assimilation of GOES-16 satellite derived winds into the warn-on-forecast system. *Atmospheric Research*, *245*, 105131. https://doi.org/10.1016/j.atmosres.2020.105131

Mansell, E. R., & Ziegler, C. L. (2013). Aerosol effects on simulated storm electrification and precipitation in a two-moment bulk microphysics model. *Journal of the Atmospheric Sciences*, *70*(7), 2032–2050. https://doi.org/10.1175/JAS-D-12-0264.1

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/10.1175/2009JAS2965.1

Markowski, P., Hannon, C., & Rasmussen, E. (2006). Observations of convection initiation "failure" from the 12 June 2002 IHOP deployment. *Monthly Weather Review*, *134*(1), 375–405. https://doi.org/10.1175/MWR3059.1

Minamide, M., & Zhang, F. (2018). Assimilation of all-sky infrared radiances from Himawari-8 and impacts of moisture and hydrometer initialization on convection-permitting tropical cyclone prediction. *Monthly Weather Review*, *146*(10), 3241–3258. https://doi.org/10.1175/MWR-D-17-0367.1

Minamide, M., Zhang, F., & Clothiaux, E. E. (2020). Nonlinear forecast error growth of rapidly intensifying Hurricane Harvey (2017) examined through convection-permitting ensemble assimilation of GOES-16 all-sky radiances. *Journal of the Atmospheric Sciences*, *77*(12), 4277–4296. https://doi.org/10.1175/JAS-D-19-0279.1

Mlawer, E. J., Taubman, S. J., Brown, P. D., Iacono, M. J., & Clough, S. A. (1997). Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *Journal of Geophysical Research*, *102*(D14), 16663–16682. https://doi.org/10.1029/97JD00237

Otsuka, M., Kunii, M., Seko, H., Shimoji, K., Hayashi, M., & Yamashita, K. (2015). Assimilation experiments of MTSAT rapid scan atmospheric motion vectors on a heavy rainfall event. *Journal of the Meteorological Society of Japan. Ser. II*, *93*(4), 459–475. https://doi.org/10.2151/jmsj.2015-030

Otsuka, M., Seko, H., Hayashi, M., & Koizumi, K. (2021). Data validation and mesoscale assimilation of Himawari-8 optimal cloud analysis products. *Journal of Atmospheric and Oceanic Technology*, *38*(2), 223–242. https://doi.org/10.1175/JTECH-D-20-0015.1

Pan, S., Gao, J., Jones, T. A., Wang, Y., Wang, X., & Li, J. (2021). The impact of assimilating satellite-derived layered precipitable water, cloud water path, and radar data on short-range thunderstorm forecasts. *Monthly Weather Review*, *149*(5), 1359–1380. https://doi.org/10.1175/MWR-D-20-0040.1

Pan, S., Gao, J., & Wang, Y. (2021). Implementation of precipitation-dependent background error profiles in a convective scale 3DVar system for improving short-term severe weather forecasts. In *25th Conference on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface*. (IOAS-AOLS), Paper 383511.

Purser, R. J., Wu, W.-S., Parrish, D. F., & Roberts, N. M. (2003). Numerical aspects of the application of recursive filters to variational statistical analysis. Part I: Spatially homogeneous and isotropic Gaussian covariances. *Monthly Weather Review*, *131*(8), 1524–1535. https://doi.org/10.1175//1520-0493(2003)131<1524:NAOTAO>2.0.CO;2

Roebber, P. J. (2009). Visualizing multiple measures of forecast quality. *Weather and Forecasting*, *24*(2), 601–608. https://doi.org/10.1175/2008WAF2222159.1

Salonen, K., Cotton, J., Bormann, N., & Forsythe, M. (2015). Characterizing AMV height-assignment error by comparing best-fit pressure statistics from the met office and ECMWF data assimilation systems. *Journal of Applied Meteorology and Climatology*, *54*(1), 225–242. https://doi.org/10.1175/JAMC-D-14-0025.1

Sawada, M., Ma, Z., Mehra, A., Tallapragada, V., Oyama, R., & Shimoji, K. (2019). Impacts of assimilating high-resolution atmospheric motion vectors derived from Himawari-8 on tropical cyclone forecast in HWRF. *Monthly Weather Review*, *147*(10), 3721–3740. https://doi.org/10.1175/MWR-D-18-0261.1

Skamarock, W., Klemp, J., Dudhia, J., Gill, D., Barker, D., Wang, W., et al. (2008). *A description of the advanced research WRF version 3*. [Application/pdf] (p. 1002). UCAR/NCAR. https://doi.org/10.5065/D68S4MVH

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, J. (2005). Initialization and numerical forecasting of a supercell storm observed during STEPS. *Monthly Weather Review*, *133*(4), 793–813. https://doi.org/10.1175/MWR2887.1

Velden, C., Lewis, W. E., Bresky, W., Stettner, D., Daniels, J., & Wanzong, S. (2017). Assimilation of high-resolution satellite-derived atmospheric motion vectors: Impact on HWRF forecasts of tropical cyclone track and intensity. *Monthly Weather Review*, *145*(3), 1107–1125. https://doi.org/10.1175/MWR-D-16-0229.1

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

Wu, T.-C., Velden, C. S., Majumdar, S. J., Liu, H., & Anderson, J. L. (2015). Understanding the influence of assimilating subsets of enhanced atmospheric motion vectors on numerical analyses and forecasts of tropical cyclone track and intensity with an ensemble Kalman filter. *Monthly Weather Review*, *143*(7), 2506–2531. https://doi.org/10.1175/MWR-D-14-00220.1

Yamashita, K. (2012). An observing system experiment of MTSAT rapid scan AMV using JMA meso-scale operational NWP system. In *Presented at the 11th international winds workshop*. Retrieved from https://www-cdn.eumetsat.int/files/2020-04/pdf_conf_p60_s4_15_yamashit_v.pdf

Yamashita, K. (2017). Assimilation of Himawari-8 atmospheric motion vectors into the numerical weather prediction systems of Japan Meteorological Agency (JMA). In *Presented at the fifth AMS symposium on the joint center for satellite data assimilation, 97th AMS annual meeting*. Retrieved from https://ams.confex.com/ams/97Annual/webprogram/Handout/Paper304768/Assimilation_H8AMV_Koji_Yamashita.pdf

Zhang, Y., 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

Zhao, J., Gao, J., Jones, T. A., & Hu, J. (2021a). Impact of assimilating high-resolution atmospheric motion vectors on convective scale short-term forecasts: 1. Observing System Simulation Experiment (OSSE). *Journal of Advances in Modeling Earth Systems*, *13*(10), e2021MS002484. https://doi.org/10.1029/2021MS002484

Zhao, J., Gao, J., Jones, T. A., & Hu, J. (2021b). Impact of assimilating high-resolution atmospheric motion vectors on convective scale short-term forecasts: 2. Assimilation experiments of GOES-16 satellite derived winds. *Journal of Advances in Modeling Earth Systems*, *13*(10), e2021MS002486. https://doi.org/10.1029/2021MS002486

Ziegler, C. L. (1985). Retrieval of thermal and microphysical variables in observed convective storms. Part 1: Model development and preliminary testing. *Journal of the Atmospheric Sciences*, *42*(14), 1487–1509. https://doi.org/10.1175/1520-0469(1985)042<1487:ROTAMV>2.0.CO;2