Updraft-Based Adaptive Assimilation of Radial Velocity Observations in a Warn-on-Forecast System

CHRISTOPHER A. KERR,[a,b] LOUIS J. WICKER,[b] AND PATRICK S. SKINNER[a,b,c]

[a] *Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, Norman, Oklahoma*
[b] *NOAA/OAR/National Severe Storms Laboratory, Norman, Oklahoma*
[c] *School of Meteorology, University of Oklahoma, Norman, Oklahoma*

(Manuscript received 1 January 2020, in final form 14 September 2020)

ABSTRACT: The Warn-on-Forecast system (WoFS) provides short-term, probabilistic forecasts of severe convective hazards including tornadoes, hail, and damaging winds. WoFS initial conditions are created through frequent assimilation of radar (reflectivity and radial velocity), satellite, and in situ observations. From 2016 to 2018, 5-km radial velocity Cressman superob analyses were created to reduce the observation counts and subsequent assimilation computational costs. The superobbing procedure smooths the radial velocity and subsequently fails to accurately depict important storm-scale features such as mesocyclones. This study retrospectively assimilates denser, 3-km radial velocity analyses in lieu of the 5-km analyses for eight case studies during the spring of 2018. Although there are forecast improvements during and shortly after convection initiation, 3-km analyses negatively impact forecasts initialized when convection is ongoing, as evidenced by model failure and initiation of spurious convection. Therefore, two additional experiments are performed using adaptive assimilation of 3-km radial velocity observations. Initially, an updraft variance mask is applied that limits radial velocity assimilation to areas where the observations are more likely to be beneficial. This experiment reduces spurious convection as well as the number of observations assimilated, in some cases even below that of the 5-km analysis experiments. The masking, however, eliminates an advantage of 3-km radial velocity assimilation for convection initiation timing. This problem is mitigated by additionally assimilating 3-km radial velocity observations in locations where large differences exist between the observed and ensemble-mean reflectivity fields, which retains the benefits of the denser radial velocity analyses while reducing the number of observations assimilated.

KEYWORDS: Convective storms; Radars/Radar observations; Short-range prediction; Cloud resolving models; Data assimilation

1. Introduction

The goal of the NOAA Warn-on-Forecast (WoF) project is to produce short-term (0–6 h), probabilistic forecasts of hazards associated with severe convection including tornadoes, hail, damaging winds, and flash flooding (Stensrud et al. 2009, 2013). Beginning in 2016, the NSSL experimental WoF System (WoFS) provided real-time, convective-scale ensemble forecasts of spring convective events (Wheatley et al. 2015; Jones et al. 2016) and a baseline of forecast quality across 64 cases from 2017 and 2018 has been established using object-based verification techniques (Skinner et al. 2018; Flora et al. 2019). Forecasts are typically initialized every 30 min between 1900 and 0300 UTC for a regional domain covering expected locations of severe thunderstorm development in the United States. This study will consider eight WoFS cases from the 2018 Hazardous Weather Testbed (HWT) Spring Forecasting Experiment (SFE; Gallo et al. 2017) conducted by the NOAA Storm Prediction Center and National Severe Storms Laboratory.

One important aspect of any data assimilation system, particularly one that utilizes dense observations, is the method to "thin" observation sets before assimilation. The process in this study is called "superobbing" and is needed when assimilating dense observations from platforms such as Doppler radars and satellites (Janjić et al. 2018). The U.S. Doppler radar network has a grid spacing of 0.25 km along each radial and an azimuthal spacing that is less than 1.5 km within 150 km of the radar (the maximum range used for the analysis) at low levels.[1] Typical convection-allowing models use grid spacings of ~3 km. Therefore, superobbing is utilized to remove features that cannot be resolved on the numerical grid (Majcen et al. 2008). Superobbing is also used to reduce random errors in the observations as well as to reduce the error correlation between observations (Lu and Xu 2009). A variety of methods have been used for thinning radar data (Albers 1995; Seko et al. 2004; Salonen et al. 2009). Lu and Xu (2009) found that superobbing the radar data to a resolution approximately twice the grid spacing improves the analysis of storm-scale features in the ensemble. Thus, for the real-time WoFS, the Doppler reflectivity and radial velocity have been superobbed to a horizontal grid that has larger grid spacing than the numerical prediction model.

A key to the success of WoFS forecasts is rapidly cycling assimilation of radar data (reflectivity and Doppler radial velocity; Snyder and Zhang 2003; Dowell et al. 2004; Tong and Xue 2005; Aksoy et al. 2009; Yussouf and Stensrud 2010),

Corresponding author: Christopher A. Kerr, christopher.kerr@noaa.gov

[1] See https://roc.noaa.gov/WSR88D/Engineering/NEXRADTechInfo.aspx for the most current specifications.

DOI: 10.1175/WAF-D-19-0251.1

satellite data (Jones et al. 2016), and conventional surface observations using an ensemble Kalman filter (EnKF). In 2018, radar observations were assimilated in the real-time system by creating superobs on two-dimensional grids using 5-km grid spacing[2] via the Cressman method (Cressman 1959). Both radar reflectivity from the Multi-Radar Multi-Sensor (MRMS; Smith et al. 2016) and radial velocity (hereinafter V_r) from individual radars are analyzed to the 5-km grid. The smoothing associated with the 5-km analysis grid spacing given the specified radius of influence may result in V_r superobs not accurately representing the most important features observed within severe convective storms since typical continental mesocyclones have horizontal diameters between 3 and 10 km (Fig. 1). Use of these smoothed V_r superobs can adversely impact or neglect to impact the analyses and forecasts of severe convection (Wicker et al. 2018) as there can be too few observations of individual convective storm features, particularly mesocyclones, to adequately produce suitable model depictions.

The focus of this paper is to explore using denser radial velocity superobs and to optimize their use to improve WoFS forecasts. This study repeats eight real-time cases from the 2018 HWT SFE to examine WoFS forecast sensitivity to variations in V_r superob method, specifically using a 3-km analysis grid spacing and a Cressman radius of influence (ROI) of 1 km. Observation ROI data assimilation sensitivity tests were performed as a precursor to this study and found a 1-km ROI to produce the most ideal results rather than a larger radius. For example, the use of 5-km superobs for radial velocity often removes information about storm rotation. The purpose of 3-km superobs is to more accurately capture the character of the storm. This observational analysis better depicts the true state of convective storms compared to the 5-km superobs utilized in the real-time experiments. In the current study, multiple sets of experiments are performed to find an optimal assimilation configuration. The initial experiment assimilates all available 3-km V_r observations. However, this method results in some model failure[3] and noisy forecasts (described below; Lange and Craig 2014; Thompson et al. 2015). A second and third experiment adaptively assimilates higher-resolution observations using a ''targeted assimilation'' approach in which observations are only assimilated if they are more likely to impact ensemble analyses and forecasts.

The next section outlines the method including ensemble design, verification techniques, and the adaptive assimilation process. Section 3 presents the various assimilation results in comparison with those from real-time experiments. Last, section 4 discusses and summarizes the importance of the results.

[2] The radius of influence for the analysis is $2^{1/2} \times \Delta x$ of the superob grid (7.07 km).

[3] The term ''model failure'' is used here to indicate that the analysis increments in one or more of the ensemble members are large enough to generate numerical instabilities and result in the numerical integration becoming unstable.

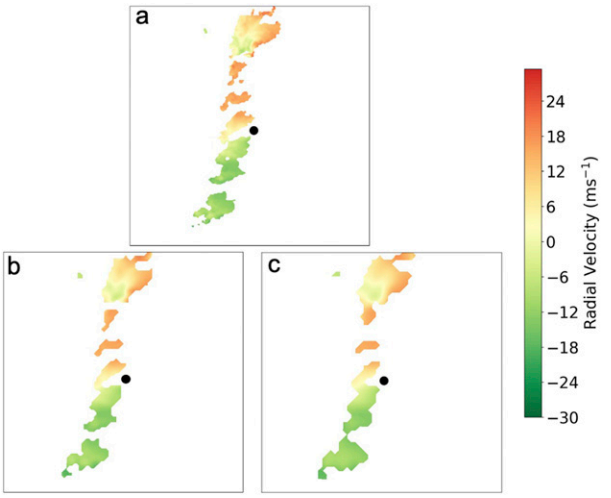

FIG. 1. Example of (a) raw, (b) 3-km superob, and (c) 5-km superob radial velocity. The radar location is denoted by the black dot.

2. Method

a. 2018 case studies

Eight convective events from the 2018 HWT SFE are examined in this study (Table 1). The events include various convective modes as well as times of convection initiation (CI). Cases include discrete supercell events (1, 2, and 16 May), mixed modes with disorganized convection (14, 28, and 29 May), and well-organized mesoscale convective systems (MCSs; 12 and 15 May). These events are also in different locations within the United States. The discrete-cell cases are confined to the central United States, and two MCS cases (12 and 15 May) are in the mid-Atlantic region. The cases are all high-impact severe weather events that resulted in many local storm reports. Considering cases involving different convective modes and locations may reveal circumstances where improved V_r assimilation is more impactful.

b. Ensemble configuration

The WoFS is an ensemble data assimilation and forecast system initialized using the High-Resolution Rapid Refresh Ensemble (HRRRE; Dowell et al. 2016). WoFS consists of 36 WRF-ARW (Skamarock et al. 2008) ensemble members with varying physical parameterization schemes (Table 2; Wheatley et al. 2015). During the 2018 SFE, WoFS was run over a 750 km × 750 km domain with 3-km horizontal grid spacing. The WoFS was initialized daily at 1800 UTC using HRRRE initial and boundary conditions before beginning frequently cycled assimilation. The 3-h, 18-member ensemble forecasts were initialized every 30 min during 1900–0300 UTC. In this study, only hourly forecasts through 2300 UTC are considered.

WoFS initial conditions were created using 1-h, 36-member HRRRE forecasts initialized at 1700 UTC (valid 1800 UTC). This is the starting point of 15-min data assimilation cycling of satellite column liquid or ice water path (Minnis et al. 2011; Jones and Stensrud 2015; Jones et al. 2016), WSR-88D

TABLE 1. The eight events presented in this study include various convective modes, locations, and SPC Categorial Outlooks, as indicated by this table.

Event date	Location	Primary convective mode	SPC Categorial Outlook
1 May 2018	Kansas and Oklahoma	Supercell	Moderate
2 May 2018	Kansas and Oklahoma	Supercell; continuous CI	Moderate
12 May 2018	Pennsylvania, New Jersey, and Delaware	Linear	Enhanced
14 May 2018	Kansas and Oklahoma	Mixed; continuous CI	Slight
15 May 2018	New York and Connecticut	Linear	Enhanced
16 May 2018	Texas	Supercell	Slight
28 May 2018	Colorado and Kansas	Mixed; continuous CI	Enhanced
29 May 2018	Kansas	Mixed; continuous CI	Enhanced

reflectivity and V_r, and surface observations using a variant of the EnKF (EAKF; Anderson 2001) within the Data Assimilation Research Testbed (Anderson et al. 2009). The boundary conditions originate from a 9-member HRRRE forecast initialized at 1500 UTC and are repeated four times to create the 36 WoFS members. The microphysical scheme used by all the WoFS ensemble members is the NSSL two-moment scheme (Mansell et al. 2010) given the desirable performance of multimoment schemes in supercell simulations (Dawson et al. 2010, 2014). Physics diversity is employed in the planetary boundary layer, shortwave, and longwave radiation schemes (Table 2). A Gaspari and Cohn (1999) localization function is applied to all observations with the horizontal and vertical cutoffs for radar observations set to 18 and 6 km, respectively. Additive noise is used to spin up convection where observed storms are not represented in the ensemble and, along with adaptive inflation (Anderson and Collins 2007), to maintain ensemble spread (Dowell and Wicker 2009; Sobash and Wicker 2015). Additional WoFS configuration details are available in Wheatley et al. (2015) and Jones et al. (2016).

c. *Verification method*

As WoFS is designed to provide short-term guidance on hazards associated with convective storms, forecasts of simulated composite reflectivity (hereinafter DZ) and updraft helicity in the 2–5-km layer (hereinafter UH25; Kain et al. 2008) are used to produce proxies for thunderstorms and mesocyclones, respectively. An object-based verification framework that permits spatial verification of convective-scale features is chosen (e.g., Davis et al. 2006a,b; Gallus 2010; Johnson et al. 2013; Clark et al. 2014; Skinner et al. 2016, 2018; Jones et al. 2018; Flora et al. 2019).

Forecast objects in simulated DZ and UH25 fields are identified for each 5-min output time available in WoFS 3-h forecasts and matched to corresponding objects identified in gridded Multi-Radar Multi-Sensor (MRMS; Smith et al. 2016) DZ fields and rotation tracks derived from 2 to 5 km azimuthal wind shear (Miller et al. 2013; Mahalik et al. 2019). Objects across the disparate fields are identified using matched percentile thresholds (e.g., Mittermaier and Roberts 2010) from a climatology of 2017 WoFS cases (Skinner et al. 2018). The resulting WoFS thresholds are 45 dBZ for DZ and 65.7 $m^2 s^{-2}$ for UH25. The MRMS thresholds are 41.241 dBZ for DZ and 0.0041 s^{-1} for azimuthal wind shear. UH25 and MRMS azimuthal wind shear values are aggregated to create 30-min mesocyclone tracks centered on each output time. Extensive quality control, particularly for MRMS azimuthal wind shear observations, is applied to limit the number of spurious mesocyclone objects. Initial quality control is performed by the MRMS system, with aliased V_r values corrected and reflectivity data associated with ground clutter removed by algorithms developed by Jing and Weiner (1993) and Lakshmanan et al. (2014), respectively. Additional quality control measures for MRMS rotation track objects include removing azimuthal wind shear values that do not occur in proximity to DZ values greater than 45 dBZ, application of an area threshold of 90 km^2

TABLE 2. Physics options applied to 36 HRRRE members, where all members have the same microphysics. The NSSL two-moment microphysics and RAP land surface parameterization are applied to all members. PBL schemes include the Yonsei University (YSU; Hong et al. 2006), Mellor–Yamada–Janjić (MYJ; Janjić 2002), and Mellor–Yamada–Nakanishi–Niino (MYNN: Nakanishi and Niino 2009) schemes. Shortwave (SW) and longwave (LW) radiation schemes include the Dudhia (1989) shortwave scheme, Rapid Radiative Transfer Model (RRTM; Mlawer et al. 1997) shortwave scheme, and the Rapid Radiative Transfer Model–Global (RRTMG; Iacono et al. 2008) shortwave and longwave schemes (adapted from Wheatley et al. 2015).

Member	PBL	SW radiation	LW radiation
1, 36	YSU	Dudhia	RRTM
2, 35	YSU	RRTMG	RRTMG
3, 34	MYJ	Dudhia	RRTM
4, 33	MYJ	RRTMG	RRTMG
5, 32	MYNN	Dudhia	RRTM
6, 31	MYNN	RRTMG	RRTMG
7, 30	YSU	Dudhia	RRTM
8, 29	YSU	RRTMG	RRTMG
9, 28	MYJ	Dudhia	RRTM
10, 27	MYJ	RRTMG	RRTMG
11, 26	MYNN	Dudhia	RRTM
12, 25	MYNN	RRTMG	RRTMG
13, 24	YSU	Dudhia	RRTM
14, 23	YSU	RRTMG	RRTMG
15, 22	MYJ	Dudhia	RRTM
16, 21	MYJ	RRTMG	RRTMG
17, 20	MYNN	Dudhia	RRTM
18, 19	MYNN	RRTMG	RRTMG

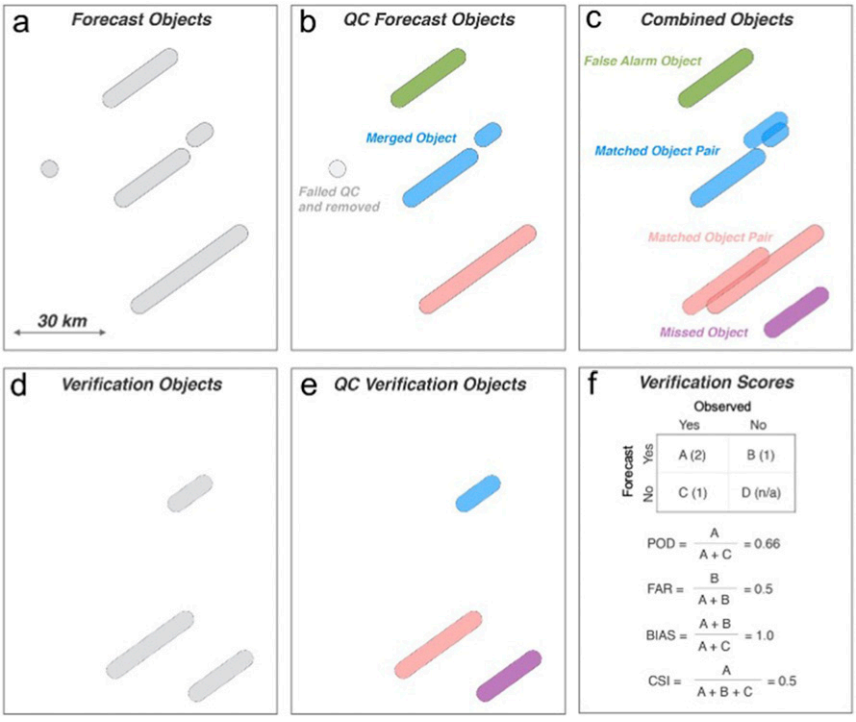

FIG. 2. Schematic illustrating the object matching and verification process. Forecast objects from (a) an ensemble member and (d) observed objects are quality controlled for (b),(e) object identification. Forecast and observed objects are then (c) matched, where matched pairs are hits, unmatched forecast objects are false alarms, and unmatched observed objects are misses. Using this information, (f) contingency-table metrics POD, FAR, bias, and CSI are calculated. This figure is adapted from Skinner et al. (2018).

(this threshold is erroneously listed as 100 km^2 in Skinner et al. 2018), and application of a continuity threshold of 10 min (i.e., a rotation track object must contain data from at least 3 different output times). Resulting objects in all fields with a boundary displacement <10 km from an adjacent object are merged.

Forecast objects are matched to verification objects in space and time using a simple total interest score (Davis et al. 2006a,b):

$$\mathrm{TI} = \left\{ \frac{\left[\frac{(\mathrm{cd}_{\max} - \mathrm{cd})}{\mathrm{cd}_{\max}}\right] + \left[\frac{(\mathrm{md}_{\max} - \mathrm{md})}{\mathrm{md}_{\max}}\right]}{2} \right\} \left[\frac{(t_{\max} - t)}{t_{\max}}\right], \quad (1)$$

where TI is the total interest score, cd is the centroid distance between an object pair, md is the minimum distance between an object pair, t is the time difference of an object pair, and the max subscript designates the maximum allowable threshold for object matching (40 km for centroid and minimum distances; 25 min for time displacement). TI values are calculated for each potential forecast and verification pair; TI > 0.2 is required to consider the objects matched. The thresholds chosen for object matching were selected to roughly match typical time and space scales of severe thunderstorm and tornado warnings issued by the U.S. National Weather Service. Object matching permits matched object pairs to be deemed "hits," unmatched forecast objects "false alarms," and unmatched verification objects "misses" (Fig. 2). This classification allows the use of the contingency-table-based metrics probability of detection (POD), false alarm ratio (FAR), frequency bias (hereinafter "bias"), and critical success index (CSI) to assess the quality of reflectivity and mesocyclone forecasts. Statistical significance of score differences is determined using a resampling technique described by Hamill (1999). Hits, misses, and false alarms are randomly sampled between two forecast sets 1000 times to create a null distribution of score differences. If the aggregated score difference exceeds the 97.5th percentile or falls below the 2.5th percentile, the score difference between two forecast sets is statistically significant at the 95% confidence level. A complete description of the object identification and matching method is available in Skinner et al. (2018).

d. V_r assimilation techniques

Whereas the real-time WoFS in 2018 assimilated 5-km V_r superobs wherever reflectivity > 20 dBZ, the goal of this study is to successfully assimilate 3-km V_r observations. First, each case is run assimilating 3-km V_r observations everywhere the reflectivity threshold described above is met. However, 3-km reflectivity superobs are used to threshold V_r, which may retain

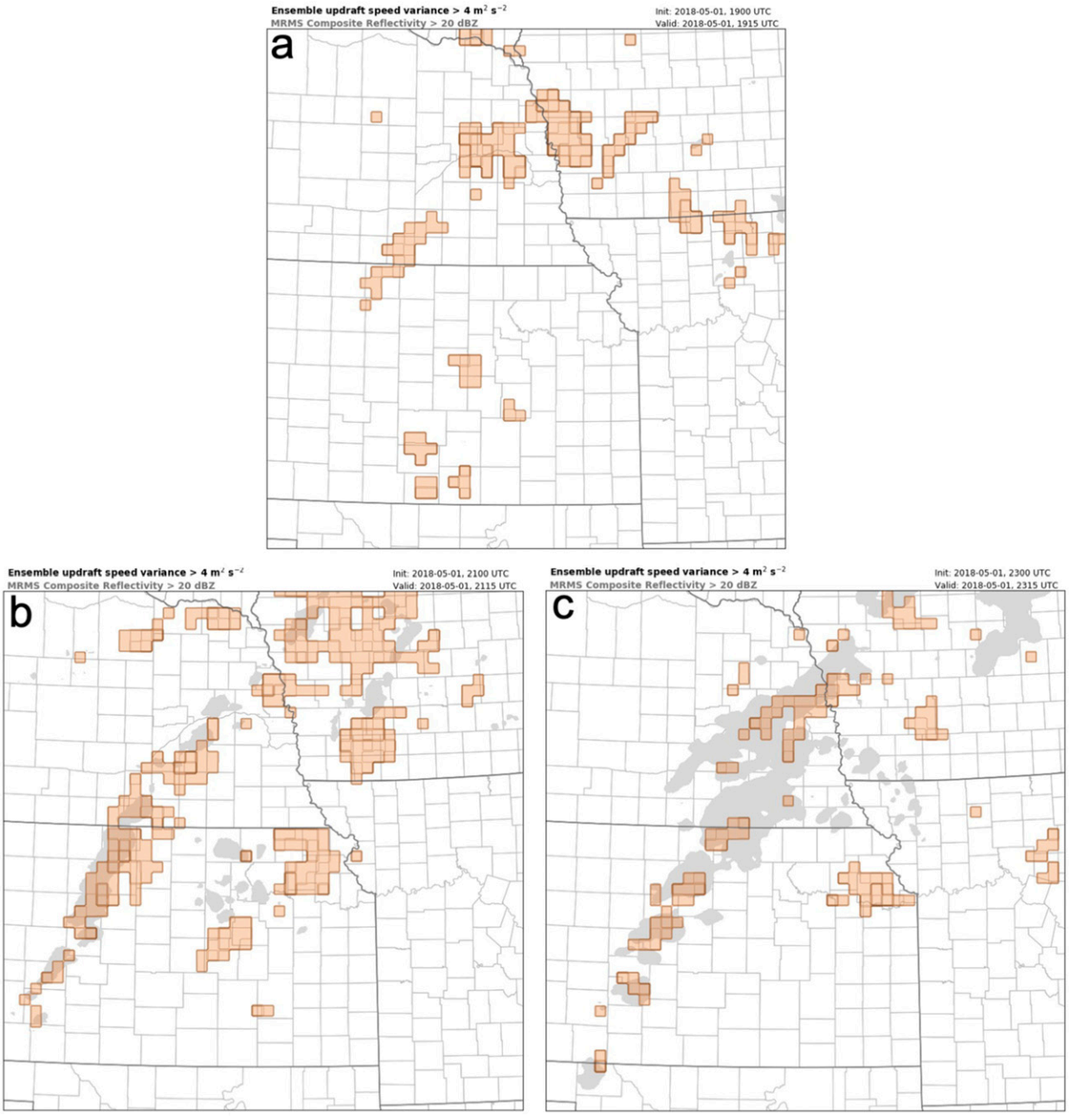

FIG. 3. Examples of areas where background column-maximum updraft ensemble variance exceeds $4\,m^2\,s^{-2}$ (orange boxes) with observed composite reflectivity (>20 dBZ; shaded gray) for 1 May 2018 at (a) 1915, (b) 2115, and (c) 2315 UTC.

V_r superobs in areas that are not included in the 5-km superob analysis. As shown in the next section, using all available 3-km V_r superobs occasionally produces forecast degradation and model imbalance. Therefore, adaptive techniques are developed whereby only those 3-km V_r superobs within the reflectivity threshold that may be most impactful on updrafts and subsequent storm dynamics in the analysis update are assimilated (Chang 2014; Kerr and Wang 2020). This technique is designed to remove observations that minimally impact analyses and forecasts except to add noise to the model state.

From the findings by Chang (2014), V_r observations are most likely to reduce vertical velocity (w) spread in areas where w ensemble variance is large. The present study considers ensemble variance of column-maximum updraft speed. To reduce the effects of noise in the variance fields, the average ensemble column-maximum updraft variance is computed at each grid point within a 5×5 grid point (15-km $\times$ 15-km) neighborhood. Examples of these averages are displayed in Figs. 3 and 4 for two cases that demonstrate how column-maximum updraft variance coverage can differ across cases

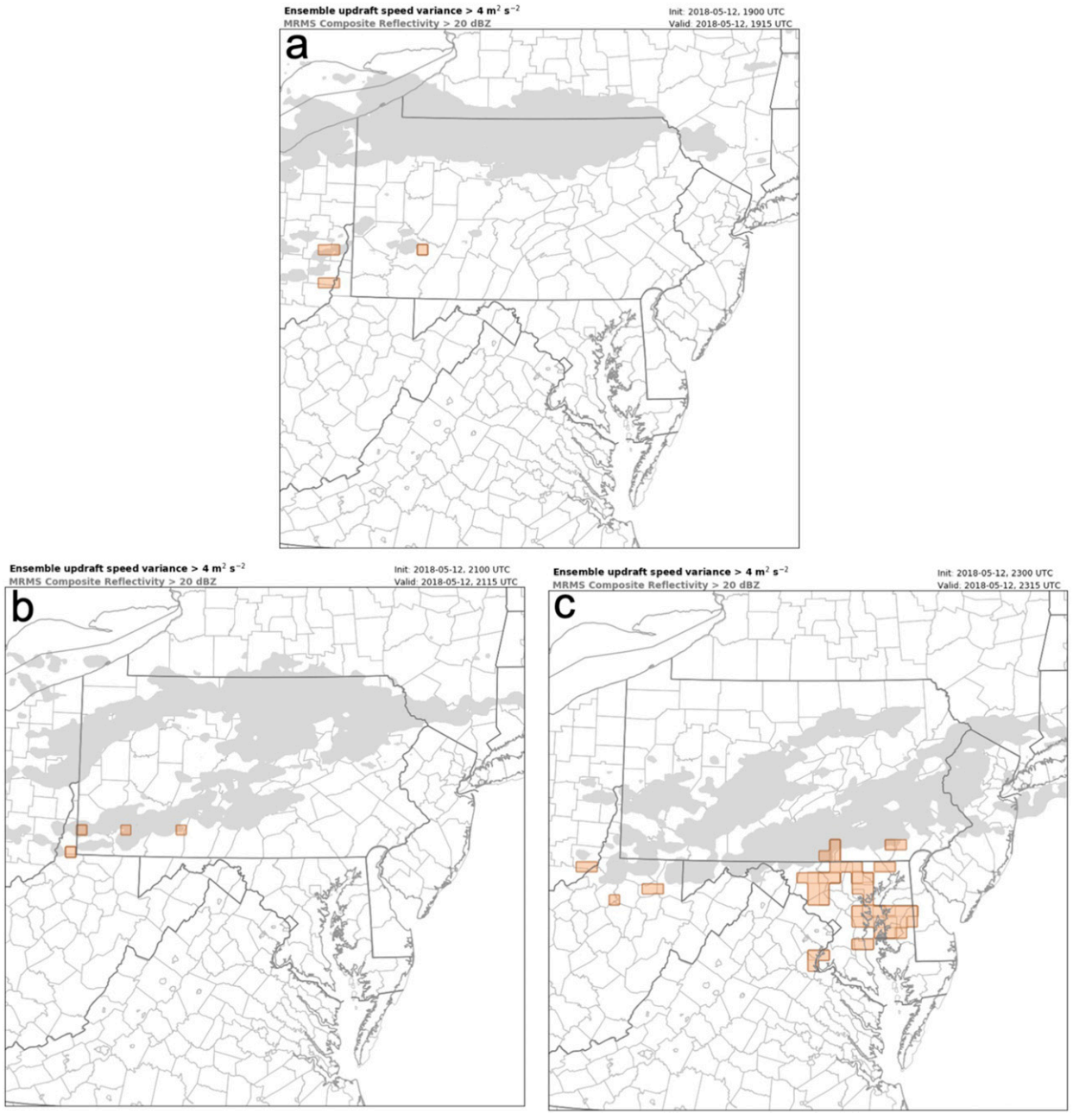

FIG. 4. As in Fig. 3, but for 12 May 2018.

and storm mode. On 1 May (Fig. 3), there is a noticeable column-maximum updraft variance coverage starting 1900–2100 UTC coincident with widespread CI before diminishing by 2300 UTC as storms mature. In other cases, such as an ongoing MCS on 12 May (Fig. 4), substantial column-maximum updraft variance is very limited throughout this timespan even though there is vast storm coverage. Based on these averaged column-maximum updraft speed variance fields, V_r observations are only assimilated if they lie within boxes exceeding a specified variance threshold of $4\,\mathrm{m^2\,s^{-2}}$ (at all vertical levels). This value is based on preliminary experiments where smaller thresholds drastically increased the number of observations assimilated without notable forecast skill improvement while thresholds over $5\,\mathrm{m^2\,s^{-2}}$ removed many useful observations and reduced forecast skill. An innovation-based threshold is also developed and only retains V_r superobs in boxes where the maximum 5-km AGL reflectivity innovation exceeds 20 dBZ (“*innovation*-mask”). The innovation threshold magnitude is based on the prerequisite that all V_r superobs must be

TABLE 3. Number of real-time-run 3-h forecast objects aggregated over all cases, lead times, and ensemble members for each odd forecast initialization time.

	1900 UTC	2100 UTC	2300 UTC
Real-time DZ	30 529	48 503	37 287
Real-time UH25	29 865	22 700	13 638

collocated with reflectivity observations exceeding 20 dBZ. The neighborhood used to compute the prior reflectivity innovations is the same as that for the updraft variances (5×5 grid points). The motivation for this additional adaptive feature is that w-masking alone will not retain V_r superobs at the onset of CI as several assimilation cycles are required to spin up model convection after observed CI.

These strategies can reduce computational costs, because 1–2 s are required to assimilate every 1000 observations, as well as reduce the potential for undesirable noise within the model state. All other observation types are assimilated as in the real-time experiments. The impacts on forecasts of adaptively assimilating observations are quantified using the object-based verification technique described above. Three experiment sets are created to assess the impact of this w-mask adaptive assimilation technique. The initial set assimilates all reflectivity-thresholded V_r observations available via 3-km superobbing (hereinafter ALL3KM set). Next, only the w-mask technique is applied throughout the entire assimilation cycling period for all

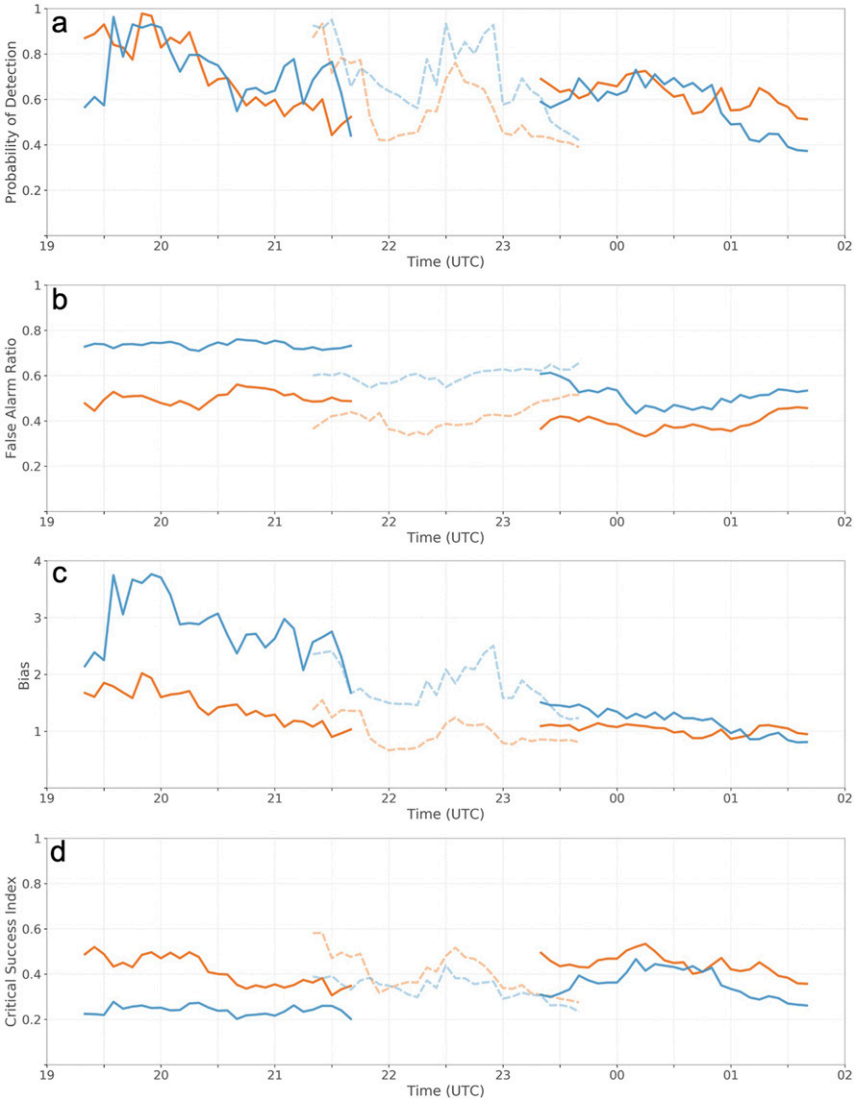

FIG. 5. Time series (3-h forecasts initialized at 1900, 2100, and 2300 UTC) of ensemble mean (a) POD, (b) FAR, (c) bias, and (d) CSI for real-time DZ (orange) and UH25 (blue). The first and last 20 min of each forecast are removed so that only objects that could be matched in time and space are considered. The 2100 UTC lines are faded and dashed to improve readability.

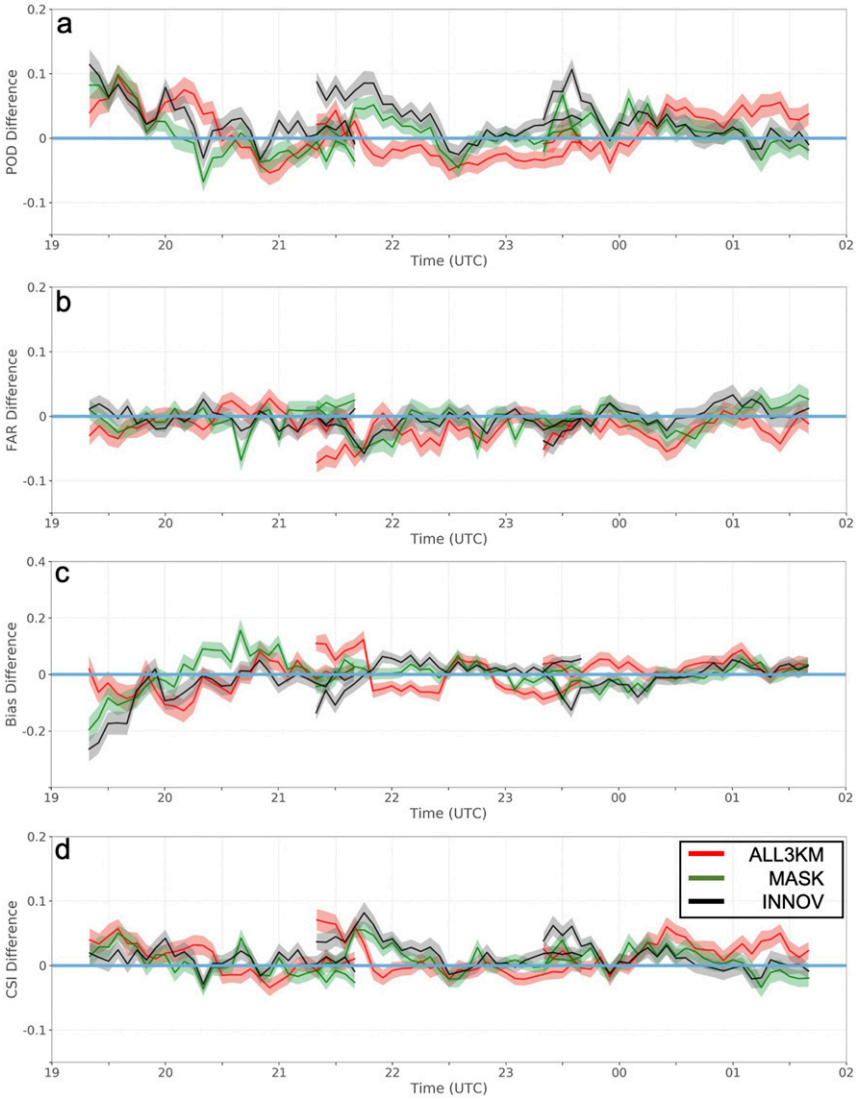

FIG. 6. Time series (3-h forecasts initialized at 1900, 2100, and 2300 UTC) of ensemble mean difference from real-time (a) POD, (b) FAR, (c) bias distance from 1 (positive and negative values signify closer to and farther from 1, respectively), and (d) CSI for DZ for ALL3KM (red), MASK (green), and INNOV (black) with 95% confidence intervals of difference from real time (shaded). A portion of the shaded area overlapping the zero line signifies a statistically insignificant difference from real time. The first and last 20 min of each forecast are removed so that only objects that could be matched in time and space are considered.

cases (hereinafter MASK set). Last, V_r observations are assimilated after applying both the *w*- and *innovation*-mask (hereinafter INNOV set).

3. Results

The number of real-time-run forecast objects for each odd forecast initialization time reveals large sample sizes for verification statistics (Table 3). Object-based verification statistics are aggregated across all eight cases for each odd hourly forecast initialization time (even hourly forecasts are similar but removed for figure readability), enabling examination of changes in forecast quality with initialization and lead time. Real-time experiment verification metrics demonstrates DZ forecasts are more skillful than UH25 forecasts as in Skinner et al. (2018; Fig. 5). FAR and bias generally decrease throughout the assimilation cycling window as spurious convection is suppressed over time as in Skinner et al. (2018).

a. *Forecasts of composite reflectivity (DZ)*

The initial set of experiments assimilates all available 3-km V_r observations (ALL3KM) across each case's domain every

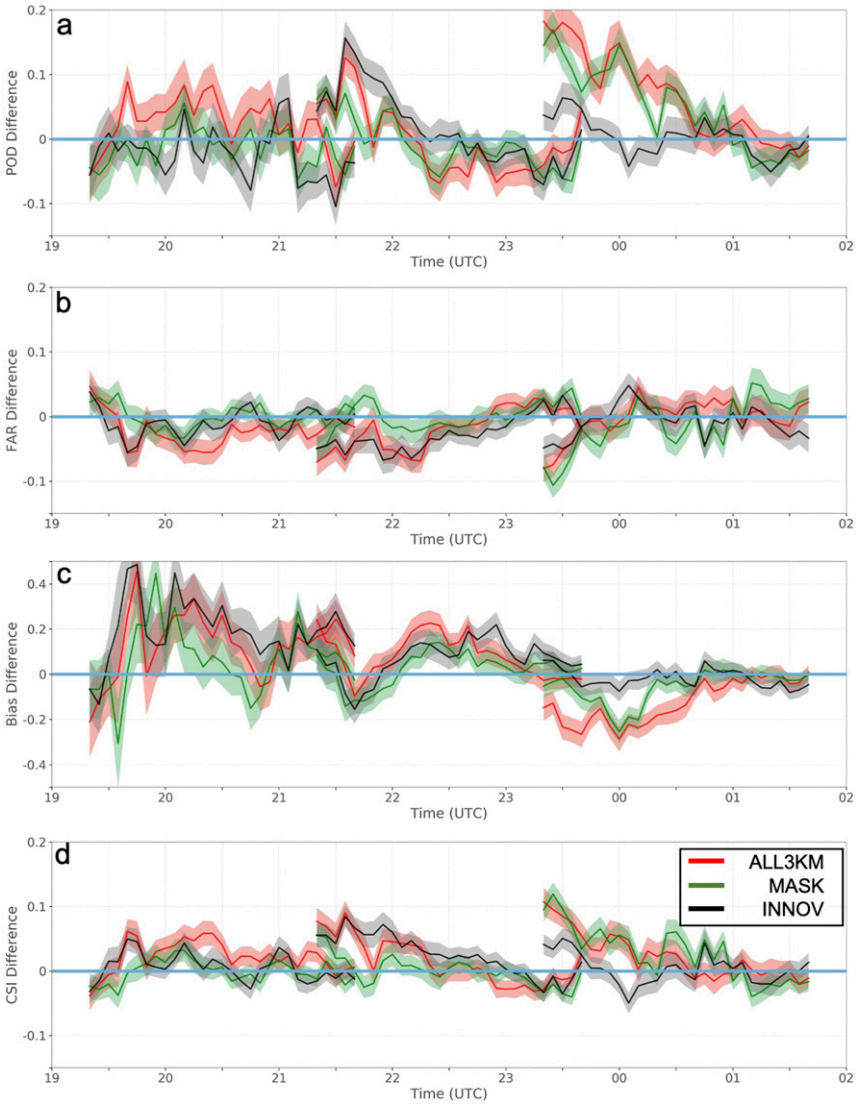

FIG. 7. As in Fig. 6, but for UH25 forecasts.

15 min. Improvements in DZ forecast skill (Fig. 6) are notable for the 1900 UTC initialization time, where POD increases during the first ~90 min. Other improvements are noted at other forecast initialization times (e.g., decreased FAR and increased CSI for 2100 UTC forecast). Comparison of the real-time and MASK DZ forecasts reveals that the *w*-masking technique has similar impacts as ALL3KM given the improvement in skill during the 1900 UTC forecasts. Unlike ALL3KM, the improvements in CSI and FAR in the 2100 UTC forecasts are not as large and POD is higher during the early lead times of the 2300 UTC forecasts. As in both ALL3KM and MASK, the INNOV 1900 UTC reflectivity forecasts have higher POD early in the forecast period. There is also a substantial increase in POD during the first half of the 2100 and 2300 UTC forecasts. There is considerable CI from 1900 to 2100 UTC across the eight cases (described below) suggesting INNOV better depicts storm spinup, increasing POD. The smaller ROI used to create the 3-km superobs is somewhat a factor in the differences relative to real time; however, the smaller superob grid spacing dominates since using a 1-km ROI for 5-km superobs does not substantially alter the forecast skill (not shown).

b. *Forecasts of mesocyclones (UH25)*

Mesocyclone forecast skill is positively impacted by 3-km V_r observations (Fig. 7). The ALL3KM 1900 UTC forecasts are substantially better, where CSI is nearly 0.1 higher for 30–90-min lead times. This improvement is due to both an increased POD and decreased FAR, signaling an improvement in storm spinup. By 2300 UTC, the prolonged assimilation cycling of 3-km V_r observations has resulted in a large increase in POD but is not suppressing spurious mesocyclones more effectively than real-time runs as reflected in comparable bias during the first 90 min of the forecast. This is not the case for reflectivity,

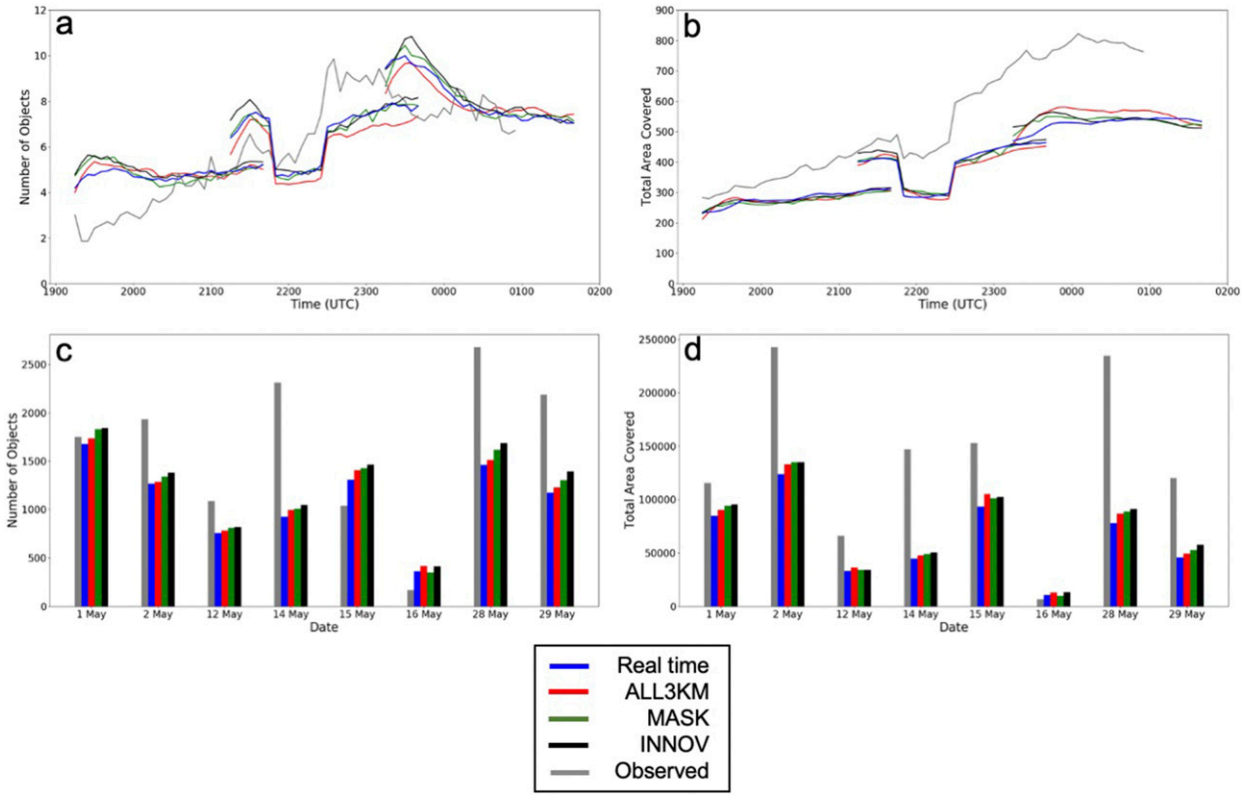

FIG. 8. Case-averaged ensemble mean reflectivity (a) object number and (b) object area covered (grid points) with time for the 1900, 2100, and 2300 UTC forecasts. Note that data are missing from 2145 to 2230 UTC 1 May; hence, there is a notable decrease in objects during this time. Also shown are (c) object number and (d) area covered for each case accumulated over all forecasts.

likely because the overall object counts are significantly higher (described in more detail in the next section). However, there are still positive impacts on ALL3KM CSI during the first 60 min of the 2300 UTC forecast.

Forecast skill is improved from real time in MASK, but improvements are modest compared to ALL3KM. This includes an increase in POD and CSI and decrease in FAR. Impacts are subtle when forecasts are initialized at 1900 UTC, likely due to the lack of storms in this timeframe and reduced ensemble variance of maximum updraft speed (Figs. 3 and 4). Ensemble variance is larger in the 2100–2300 UTC window, thus allowing V_r observations to be assimilated in critical locations. Beyond 2300 UTC, storms are typically mature in these cases resulting in reduced ensemble variance (given several hours of assimilation cycling) and do not differ substantially from the real-time runs (not shown). Last, INNOV improves 2100 UTC forecasts as CSI is significantly improved for the first ~75 min of the forecasts, given both an increase in POD and decrease in FAR. Skill in the 2300 UTC forecasts is not as improved as ALL3KM and MASK. When comparing the MASK and INNOV to the real-time set, there is definite skill-based incentive to use the INNOV observation processing method.

c. *Model failure in ALL3KM*

Many of the ALL3KM experiments develop model failure from 2300 to 0000 UTC (after >5 h of assimilation cycling) resulting in model failure, making comparisons beyond 2300 UTC difficult. Model failure in 1–4 ensemble members occurred in three cases (14, 28, 29 May) in this timeframe. Failures are associated with the vertical Courant–Friedrichs–Lewy condition (CFL) greatly exceeding O(1) near the top of storms where the vertical velocity becomes too large for the given vertical grid spacing and time step.[4] While the true cause is unknown, this is potentially a result of an excessive number of observations perturbing the model state, thus increasing model imbalances leading to large vertical velocities. Model failure is harmful for a potential real-time system, so assimilating all 3-km storm V_r observations is suboptimal given this frequent problem.

d. *V_r assimilation technique impact on forecast objects*

Changing the method of V_r assimilation can alter the forecast biases, POD, and FAR for both DZ and UH25 objects (Figs. 6 and 7). Since the same verification dataset is used for each experiment, these differences reflect changes in the counts and areal coverages of forecast objects. Reflectivity object counts and total areal coverage increase with subsequent forecasts for all experiment sets (Figs. 8a,b). There are substantially more

[4] Decreasing the model time step is a solution that can reduce model failure rates; however, this increases model integration computational costs in addition to assimilation costs.

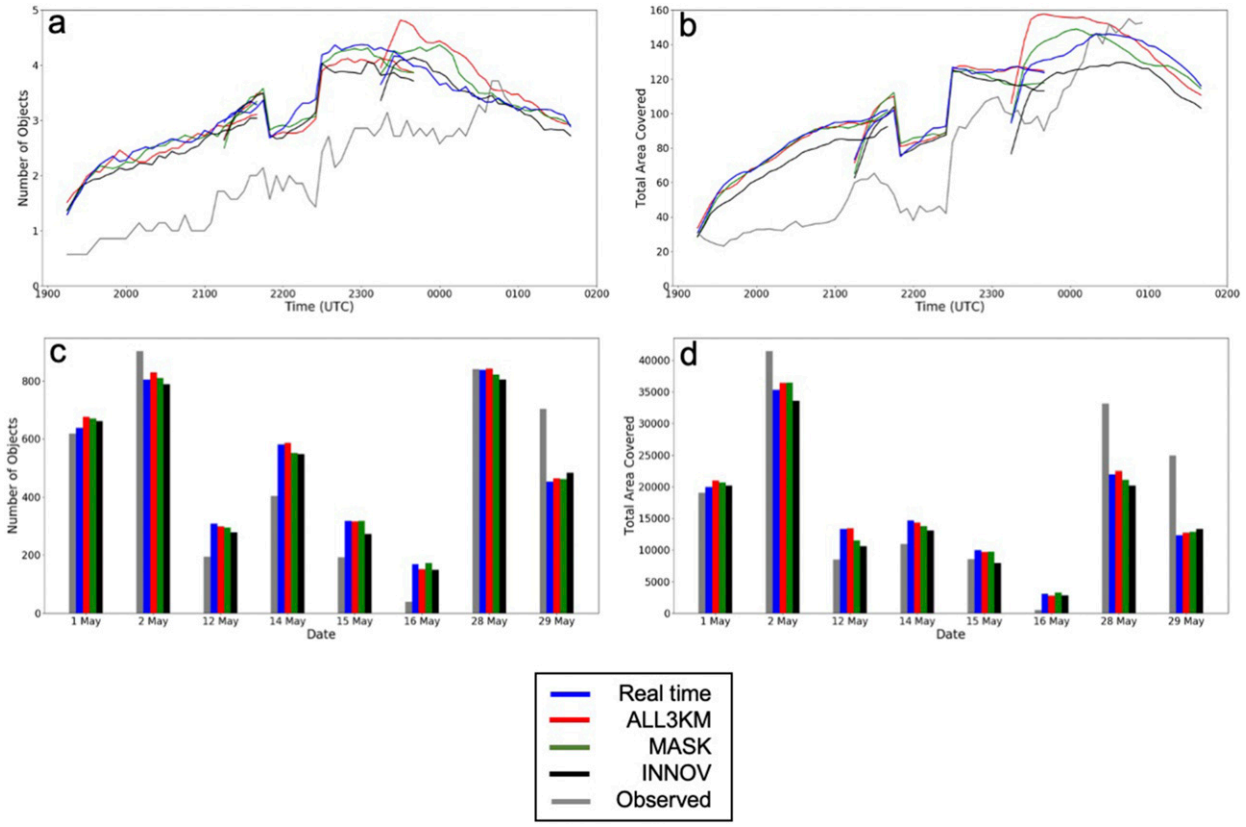

FIG. 9. As in Fig. 8, but for mesocyclone objects.

forecast objects than observed objects early in the period, but comparable numbers of forecast and observed objects beyond 2300 UTC. Objects in 2300 UTC forecasts in all experiments are generally too small given comparable observed and forecast object counts but larger observed areal coverage. ALL3KM has the fewest objects but largest areal coverage by 2300 UTC, suggesting the individual objects are large compared to MASK and INNOV, and potentially signaling more aggressive upscale growth. The four continuous CI (meaning CI occurring throughout the experiment time window) cases listed in Table 1 (2, 14, 28, 29 May) have both underpredicted reflectivity object counts and areal coverage (Figs. 8c,d). This is because the ensemble lags behind observations during storm spinup, which is occurring throughout the DA cycling window. The area–object count ratio is generally consistent across experiments for each case (i.e., areal coverage increases with increasing number of objects) suggesting that individual object sizes are mostly unaffected by adaptive assimilation. These individual cases all have negative biases and relatively low PODs (not shown). The 16 May case has late CI by comparison resulting in a low object count.

Mesocyclone object counts are overpredicted throughout the majority of the experiment timeframe (Fig. 9a). The most significant difference between experiments is the increase in ALL3KM UH25 object counts and areal coverage during the first ~90 min of the 2300 UTC forecasts (Fig. 9b). This is consistent with the increased forecast bias present in this experiment set by 2300 UTC (Fig. 8c). As in the reflectivity verification, this suggests upscale growth in ALL3KM since MCSs may produce numerous large UH objects. Generally speaking, INNOV has the least biased object counts and areal coverage throughout the experiment time window and across most individual cases (Figs. 9c,d). The consistent overprediction of mesocyclone object count and coverage could be attributed to a suboptimal threshold for object identification. The thresholds in this study were calculated from a climatology of 2017 cases in Skinner et al. (2018) and are unbiased for that dataset. However, using a suboptimal threshold should not substantially impact differences between experiments, but only verification score magnitudes (Skinner et al. 2018).

As noted previously, adaptively assimilating 3-km V_r superobs improves storm spinup. An example of this is shown in Fig. 10, where UH25 probabilities for the southernmost mesocyclone object on 2 May are enhanced in both MASK and INNOV. This storm is not well depicted in the real-time or ALL3KM forecasts. The ensemble does not have adequate column maximum updraft variance or reflectivity innovation in the vicinity of this storm during spinup (not shown). Therefore, the inclusion of either 3- or 5-km V_r observations of this southernmost storm appears to hinder development. The other convective features in this example are mature at the forecast

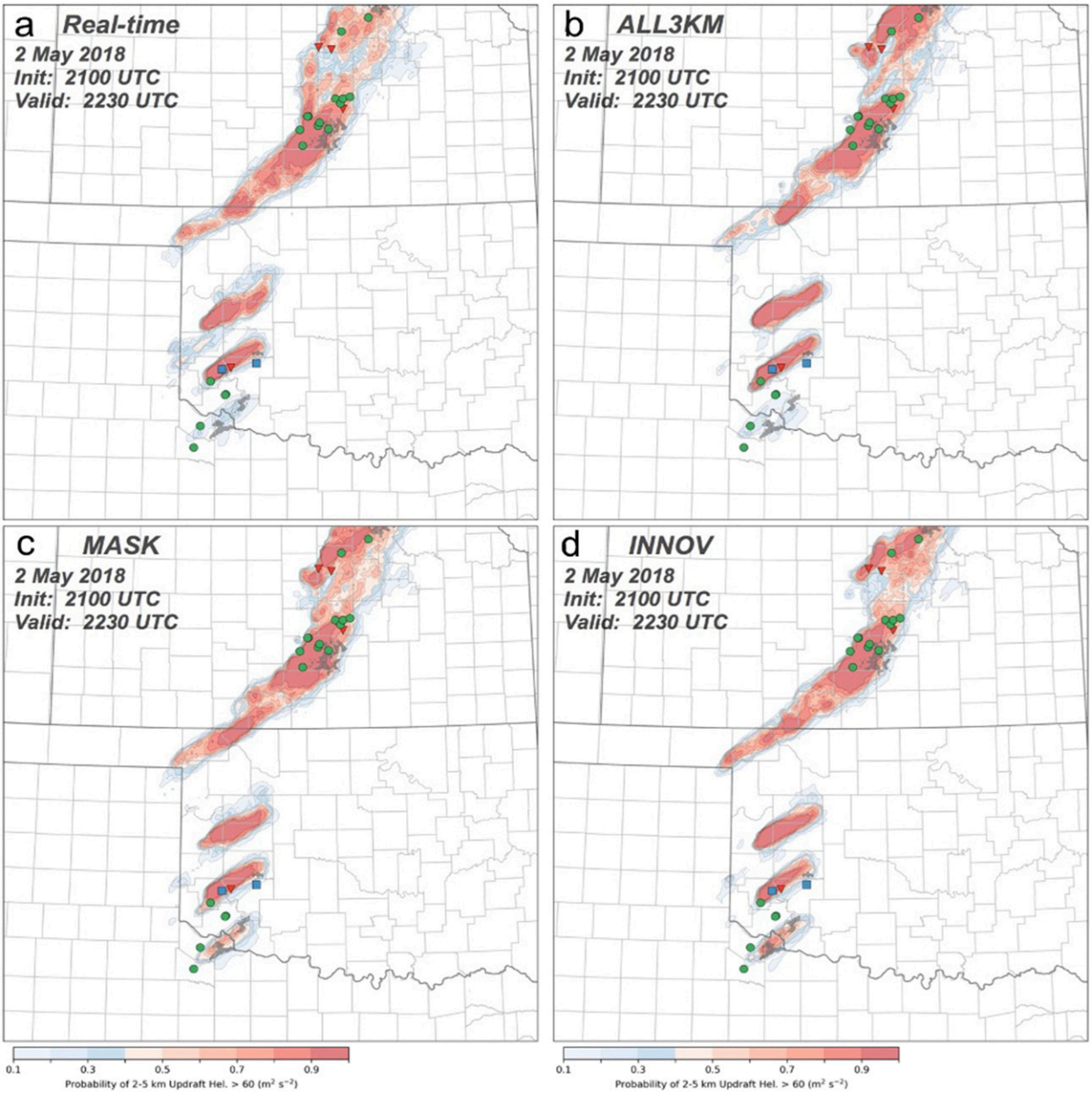

FIG. 10. The 90-min forecast (initialized 2100 UTC 2 May) probability swathes of UH25 exceeding $60\,m^2\,s^{-2}$ in (a) real time, (b) ALL3KM, (c) MASK, and (d) INNOV with observed mesocyclone objects (shaded gray). Local storm reports are tornadoes (red triangle), damaging winds (blue square), and hail (green circle).

initialization time, thus, there are only minor differences between the experiments.

e. Masking impact on number of observations

A second impact of V_r w-masking and innovation-based thresholding is the change in the number of observations assimilated compared to the real-time sets. In cases with particularly widespread regions exceeding the thresholds for w ensemble variance or reflectivity innovations, there will be significantly more observations assimilated. In cases of widespread threshold-exceeding updraft variance, there are typically more V_r observations assimilated in MASK than in the real-time runs; conversely, in cases with relatively few areas of high updraft variance, there are fewer V_r observations assimilated in MASK than in the real-time runs. Assimilating fewer observations without degrading forecast skill is also an advantage of adaptive assimilation. The total number of V_r observations assimilated in real time (5-km superobs) and ALL3KM (3-km superobs) during 1800–2300 UTC are presented for each case in Fig. 11a. ALL3KM assimilates an order of magnitude more

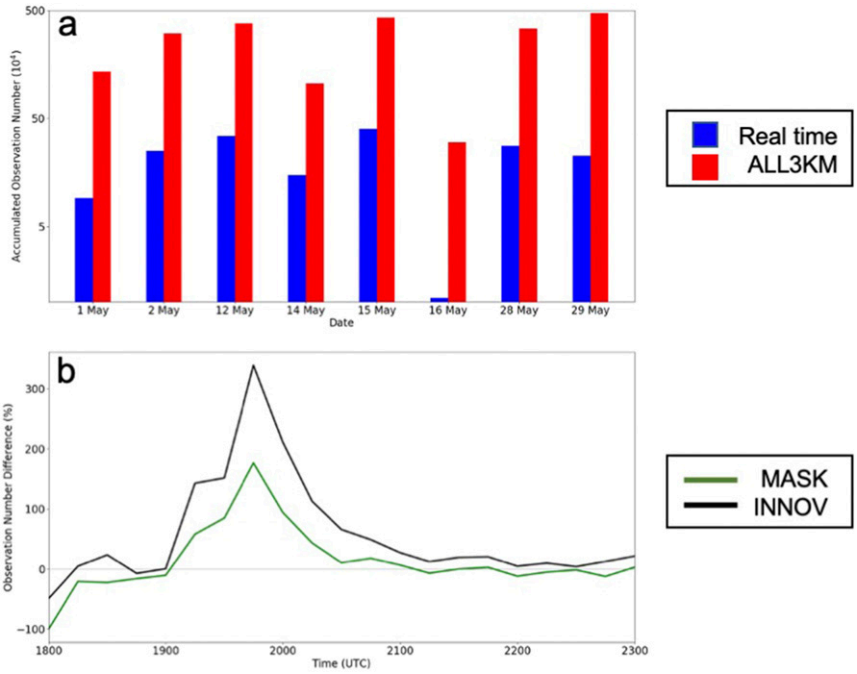

FIG. 11. (a) Accumulated V_r observations assimilated from 1800 to 2300 UTC in real time and ALL3KM for each case. (b) Average percentage change among cases in number of V_r observations assimilated in MASK experiments and INNOV relative to real time.

V_r observations than the real-time experiments. The MCS and mixed-mode convection cases have the most V_r observations, as expected. The average change in V_r observation assimilation from real-time cases for MASK and INNOV shows when the w-masking and innovation techniques are most utilized (Fig. 11b). There is an average increase in assimilated V_r counts during 1900–2100 UTC in both experiment sets since large updraft ensemble variance (and/or CI in the case of INNOV resulting in more observations) are more widespread in this timeframe. For example, at the maximum just before 2000 UTC, INNOV assimilates over 4.5×10^3 V_r observations on 1 May while only 200 were assimilated in real time. However, 12 May does not have widespread updraft variance and CI, so INNOV only assimilates 2×10^3 V_r observations at this time compared to the 1.7×10^4 assimilated in real time (not shown). Beyond 2100 UTC, updraft variance and CI diminishes, causing fewer V_r observations to be assimilated.

The percentage change in observations throughout all cycles between the MASK/INNOV and real-time experiments are designated in Fig. 12 along with overall (all forecast times) change in reflectivity POD and success ratio (SR; 1 − FAR). Some cases have a reduction in V_r observations assimilated versus the real-time experiments despite the higher-resolution observational analyses (12, 14, 15, 16, and 29 May). The 12, 15, and 16 May cases have fewer observations assimilated in both MASK and INNOV with increases in forecast skill overall. The reduction in assimilated observations is particularly important for 12 and 15 May since these two cases feature MCSs with the two highest observation counts (Fig. 12a). The cases where the assimilated observation counts in MASK and INNOV are similar to the real-time experiments have less change in forecast skill. Mesocyclone forecast skill is impacted more from case to case than reflectivity forecast skill (Fig. 13). Two supercell cases, 1 and 2 May, have increased POD and SR in 30-min forecasts (Figs. 13a,b). There is a mixture of positive and negative impacts on forecast skill across cases similar to changes in individual ALL3KM cases, but as previously noted, these statistics are valid for all forecast initialization times, and so the improved forecast skill shortly after CI (Figs. 6 and 7) is dampened in these results.

4. Conclusions and summary

Three sensitivity experiments are performed for eight case studies during the spring of 2018 to assess the impact of altering V_r superobs assimilated by WoFS. In the first set of experiments, 3-km V_r superob analyses are assimilated instead of 5-km analyses, which improves reflectivity and UH25 forecasts over the real-time experiments, particularly for early initiation times. However, assimilating entire 3-km V_r superob analyses are computationally expensive, especially in cases of substantial storm areal coverage, and introduces noise that may reduce forecast skill owing to increased spurious convection generated by large imbalances in the ensemble states. Important for real-time considerations, model failure occurred in three of eight cases when 3-km V_r superobs are assimilated.

In the second set of experiments, V_r observations are not assimilated where the ensemble variance in vertical velocity is below 4 $m^2 s^{-2}$. The impacts of this w-masking on DZ forecasts are small, but generally positive, with an increase in POD and

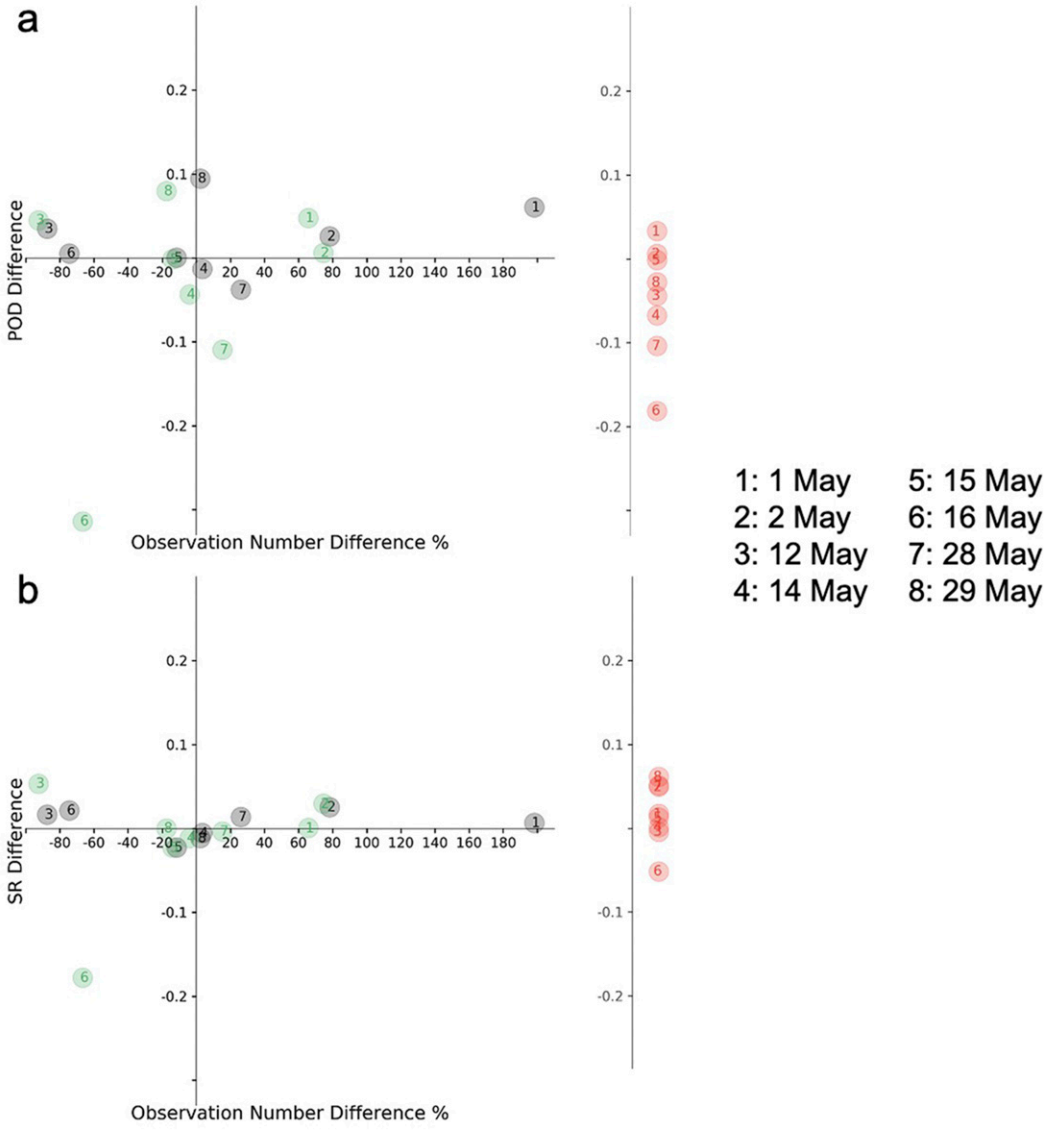

FIG. 12. Individual case percentage change in number of V_r observations assimilated (relative to real-time experiments) vs reflectivity object (a) POD change and (b) SR change multiplied by 100 for 30-min forecasts for MASK (green markers) and INNOV (gray markers). A marker in the top-left quadrant represents an increase in skill with decrease in observations, which is most optimal. Conversely, a marker in the bottom-right quadrant represents a decrease in skill with increase in observations, which is least optimal. A marker in the top-right quadrant also represents an increase in skill. The corresponding ALL3KM skill changes (red markers) from real time are displayed on the axes to the right.

associated positive bias in forecasts at various lead times (Fig. 8). However, UH25 forecasts are notably improved over real-time runs. While not every case presented has substantially improved forecast skill, many of them see improvements over the real-time experiments at various forecast lead times with only a few instances of forecast degradation (Figs. 12 and 13).

One limitation of the *w*-masking method alone is that it discards potentially beneficial observations during CI since the ensemble's updraft variance is typically small while the members spin up newly observed convection. This motivated a third set of experiments in which the denser V_r observations are assimilated wherever the updraft variance is large or the reflectivity innovation is large. This method improves both reflectivity and mesocyclone forecasts as in ALL3KM while also discarding many observations. The greatest benefits from this approach occur in forecasts initialized shortly after widespread CI. In cases where new storms are continually initiating, this technique produces improvements throughout the assimilation cycling. Given the forecast skill increases via 3-km

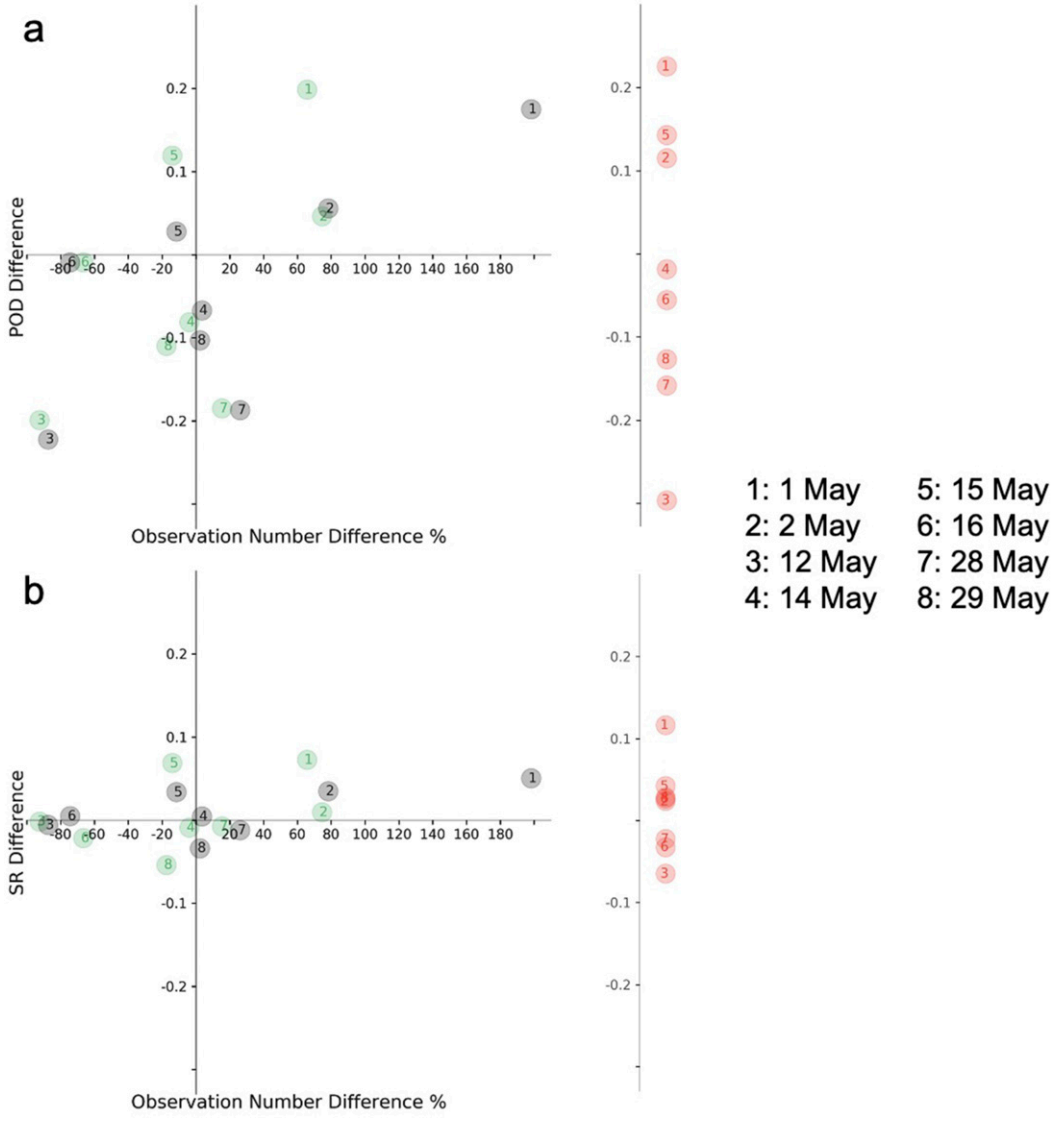

FIG. 13. As in Fig. 12, but for UH25.

superobs and the reduction in computation from the adaptive process, it is the most optimal of the three techniques tested with WoFS.

In addition to the potential forecast skill increases provided by adaptive assimilation, the number of observations assimilated after *w*-masking is reduced in some cases even though the superobbed grid spacing is 3 km instead of 5 km. This is true predominantly for the MCS and mixed mode cases that usually have more V_r observations compared to discrete supercell cases. The reduction in assimilated observations decreases the computational expense, an important consideration given the rapid assimilation cycling in WoFS. A reduction in the number of assimilated observations can also result in improved forecasts by eliminating many observations that may have spurious or low correlations with the model state fields. Two cases have low environmental CAPE (12 and 16 May) thus having weaker updrafts leading to few areas where the specified *w* variance threshold is met. However, the ALL3KM counterpart experiments have very similar forecast skill, so lowering the *w* variance threshold would only increase the computational expense via assimilation and not improve forecast skill in these instances. Future variance-based observation masking studies for other observation types may be useful, especially for data assimilation systems analyzing intermittent phenomena (e.g., convection) when radar and satellite observations are used.

Acknowledgments. The authors thank Corey Potvin, Altug Aksoy, and two anonymous reviewers for providing constructive feedback that improved this paper. We are also grateful for WoFS support provided by Kent Knopfmeier and Gerry Creager and also to Anthony Reinhart for generating the MRMS observations. This study was made possible by the CIMMS Cooperative Agreement

NA16OAR4320115. All analyses produced in the paper were generated with the freely provided Anaconda Python distribution and the basemap and scikit-image libraries.

REFERENCES

Aksoy, A., D. Dowell, and C. Snyder, 2009: A multicase comparative assessment of the ensemble Kalman filter for assimilation of radar observations. Part I: Storm-scale analyses. *Mon. Wea. Rev.*, **137**, 1805–1824, https://doi.org/10.1175/2008MWR2691.1.

Albers, S. C., 1995: The LAPS wind analysis. *Wea. Forecasting*, **10**, 342–352, https://doi.org/10.1175/1520-0434(1995)010<0342:TLWA>2.0.CO;2.

Anderson, J. L., 2001: An ensemble adjustment filter for data assimilation. *Mon. Wea. Rev.*, **129**, 2884–2903, https://doi.org/10.1175/1520-0493(2001)129<2884:AEAKFF>2.0.CO;2.

——, and N. Collins, 2007: Scalable implementations of ensemble filter algorithms for data assimilation. *J. Atmos. Oceanic Technol.*, **24**, 1452–1463, https://doi.org/10.1175/JTECH2049.1.

——, T. Hoar, K. Raeder, H. Liu, N. Collins, R. Torn, and A. Avellano, 2009: The Data Assimilation Research Testbed: A community facility. *Bull. Amer. Meteor. Soc.*, **90**, 1283–1296, https://doi.org/10.1175/2009BAMS2618.1.

Chang, W., 2014: Convective-scale radar data assimilation and adaptive radar observation with the ensemble Kalman filter. Ph.D. thesis, McGill University, 154 pp., https://escholarship.mcgill.ca/concern/theses/8s45qc87q?locale=en.

Clark, A. J., R. G. Bullock, T. L. Jensen, M. Xue, and F. Kong, 2014: Application of object-based time-domain diagnostics for tracking precipitation systems in convection-allowing models. *Wea. Forecasting*, **29**, 517–542, https://doi.org/10.1175/WAF-D-13-00098.1.

Cressman, G. P., 1959: An operational objective analysis system. *Mon. Wea. Rev.*, **87**, 367–374, https://doi.org/10.1175/1520-0493(1959)087<0367:AOOAS>2.0.CO;2.

Davis, C. A., B. G. Brown, and R. G. Bullock, 2006a: Object-based verification of precipitation forecasts. Part I: Methodology and application to mesoscale rain areas. *Mon. Wea. Rev.*, **134**, 1772–1784, https://doi.org/10.1175/MWR3145.1.

——, ——, and ——, 2006b: Object-based verification of precipitation forecasts. Part II: Application to convective rain systems. *Mon. Wea. Rev.*, **134**, 1785–1795, https://doi.org/10.1175/MWR3146.1.

Dawson, D. T., M. Xue, J. A. Milbrandt, and M. K. Yau, 2010: Comparison of evaporation and cold pool development between single-moment and multimoment bulk microphysics schemes in idealized simulations of tornadic thunderstorms. *Mon. Wea. Rev.*, **138**, 1152–1171, https://doi.org/10.1175/2009MWR2956.1.

——, E. R. Mansell, Y. Jung, L. J. Wicker, M. R. Kumjian, and M. Xue, 2014: Low-level ZDR signatures in supercell forward flanks: The role of size sorting and melting of hail. *J. Atmos. Sci.*, **71**, 276–299, https://doi.org/10.1175/JAS-D-13-0118.1.

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

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

——, and Coauthors, 2016: Development of a High- Resolution Rapid Refresh Ensemble (HRRRE) for severe weather forecasting. *28th Conf. on Severe Local Storms*, Portland, OR, Amer. Meteor. Soc., 8B.2, https://ams.confex.com/ams/28SLS/webprogram/Paper301555.html.

Dudhia, J., 1989: Numerical study of convection observed during the Winter Monsoon Experiment using a mesoscale two-dimensional model. *J. Atmos. Sci.*, **46**, 3077–3107, https://doi.org/10.1175/1520-0469(1989)046<3077:NSOCOD>2.0.CO;2.

Flora, M. L., and Coauthors, 2019: Object-based verification of short-term, storm-scale probabilistic mesocyclone guidance from an experimental Warn-on-Forecast system. *Wea. Forecasting*, **34**, 1721–1739, https://doi.org/10.1175/WAF-D-19-0094.1.

Gallo, B. T., and Coauthors, 2017: Breaking new ground in severe weather prediction: The 2015 NOAA/Hazardous Weather Testbed Spring Forecasting Experiment. *Wea. Forecasting*, **32**, 1541–1568, https://doi.org/10.1175/WAF-D-16-0178.1.

Gallus, W. A., Jr., 2010: Application of object-based verification techniques to ensemble precipitation forecasts. *Wea. Forecasting*, **25**, 144–158, https://doi.org/10.1175/2009WAF2222274.1.

Gaspari, G., and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.*, **125**, 723–757, https://doi.org/10.1002/qj.49712555417.

Hamill, T. M., 1999: Hypothesis tests for evaluating numerical precipitation forecasts. *Wea. Forecasting*, **14**, 155–167, https://doi.org/10.1175/1520-0434(1999)014<0155:HTFENP>2.0.CO;2.

Hong, S.-Y., Y. Noh, and J. Dudhia, 2006: A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon. Wea. Rev.*, **134**, 2318–2341, https://doi.org/10.1175/MWR3199.1.

Iacono, M. J., J. S. Delamere, E. J. Mlawer, M. W. Shephard, S. A. Clough, and W. D. Collins, 2008: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *J. Geophys. Res.*, **113**, D13103, https://doi.org/10.1029/2008JD009944.

Janjić, T., and Coauthors, 2018: On the representation error in data assimilation. *Quart. J. Roy. Meteor. Soc.*, **144**, 1257–1278, https://doi.org/10.1002/qj.3130.

Janjić, Z. I., 2002: Nonsingular implementation of the Mellor–Yamada level 2.5 scheme in the NCEP Meso model. NCEP Office Note 437, 61 pp., http://www.emc.ncep.noaa.gov/officenotes/newernotes/on437.pdf.

Jing, Z., and G. Weiner, 1993: Two-dimensional dealiasing of Doppler velocities. *J. Atmos. Oceanic Technol.*, **10**, 798–808, https://doi.org/10.1175/1520-0426(1993)010<0798:TDDODV>2.0.CO;2.

Johnson, A., X. Wang, F. Kong, and M. Xue, 2013: Object-based evaluation of the impact of horizontal grid spacing on convection-allowing forecasts. *Mon. Wea. Rev.*, **141**, 3413–3425, https://doi.org/10.1175/MWR-D-13-00027.1.

Jones, T. A., and D. J. Stensrud, 2015: Assimilating cloud water path as a function of model cloud microphysics in an idealized simulation. *Mon. Wea. Rev.*, **143**, 2052–2081, https://doi.org/10.1175/MWR-D-14-00266.1.

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

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

Kain, J. S., and Coauthors, 2008: Some practical considerations regarding horizontal resolution in the first generation of operational convection-allowing NWP. *Wea. Forecasting*, **23**, 931–952, https://doi.org/10.1175/WAF2007106.1.

Kerr, C. A., and X. Wang, 2020: Ensemble-based targeted observation method applied to radar radial velocity observations on idealized supercell low-level rotation forecasts: A proof of concept. *Mon. Wea. Rev.*, **148**, 877–890, https://doi.org/10.1175/MWR-D-19-0197.1.

Lakshmanan, V., C. Karstens, J. Krause, and L. Tang, 2014: Quality control of weather radar data using polarimetric variables. *J. Atmos. Oceanic Technol.*, **31**, 1234–1249, https://doi.org/10.1175/JTECH-D-13-00073.1.

Lange, H., and G. C. Craig, 2014: The impact of data assimilation length scales on analysis and prediction of convective storms. *Mon. Wea. Rev.*, **142**, 3781–3808, https://doi.org/10.1175/MWR-D-13-00304.1.

Lu, H., and Q. Xu, 2009: Trade-offs between measurement accuracy and resolutions in configuring phased-array radar velocity scans for ensemble-based storm-scale data assimilation. *J. Appl. Meteor. Climatol.*, **48**, 1230–1244, https://doi.org/10.1175/2008JAMC2009.1.

Mahalik, M. C., B. R. Smith, K. L. Elmore, D. M. Kingfield, K. L. Ortega, and T. M. Smith, 2019: Estimates of gradients in radar moments using a linear least squares derivative technique. *Wea. Forecasting*, **34**, 415–434, https://doi.org/10.1175/WAF-D-18-0095.1.

Majcen, M., P. Markowski, Y. Richardson, D. Dowell, and J. Wurman, 2008: Multipass objective analyses of Doppler radar data. *J. Atmos. Oceanic Technol.*, **25**, 1845–1858, https://doi.org/10.1175/2008JTECHA1089.1.

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

Miller, M. L., V. Lakshmanan, and T. M. Smith, 2013: An automated method for depicting mesocyclone paths and intensities. *Wea. Forecasting*, **28**, 570–585, https://doi.org/10.1175/WAF-D-12-00065.1.

Minnis, P., and Coauthors, 2011: CERES edition-2 cloud property retrievals using TRMM VIRS and Terra and Aqua MODIS data—Part I: Algorithms. *IEEE Trans. Geosci. Remote Sens.*, **49**, 4374–4400, https://doi.org/10.1109/TGRS.2011.2144601.

Mittermaier, M., and N. Roberts, 2010: Intercomparison of spatial forecast methods: Identifying skillful spatial scales using the fractions skill score. *Wea. Forecasting*, **25**, 343–354, https://doi.org/10.1175/2009WAF2222260.1.

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

Nakanishi, M., and H. Niino, 2009: Development of an improved turbulence closure model for the atmospheric boundary layer. *J. Meteor. Soc. Japan*, **87**, 895–912, https://doi.org/10.2151/jmsj.87.895.

Salonen, K., H. Järvinen, G. Haase, S. Niemelä, and R. Eresmaa, 2009: Doppler radar radial winds in HIRLAM. Part II: Optimizing the super-observation processing. *Tellus*, **61A**, 288–295, https://doi.org/10.1111/j.1600-0870.2008.00381.x.

Seko, H., T. Kawabata, T. Tsuyuki, H. Nakamura, K. Koizumi, and T. Iwabuchi, 2004: Impacts of GPS-derived water vapor and radial wind measured by Doppler radar on numerical prediction of precipitation. *J. Meteor. Soc. Japan*, **82**, 473–489, https://doi.org/10.2151/jmsj.2004.473.

Skamarock, W. C., and Coauthors, 2008: A description of the Advanced Research WRF version 3. NCAR Tech. Note NCAR/TN-475+STR, 113 pp., https://doi.org/10.5065/D68S4MVH.

Skinner, P. S., L. Wicker, D. M. Wheatley, and K. H. Knopfmeier, 2016: Application of two spatial verification methods to ensemble forecasts of low-level rotation. *Wea. Forecasting*, **31**, 713–735, https://doi.org/10.1175/WAF-D-15-0129.1.

——, and Coauthors, 2018: Object-based verification of a prototype warn-on-forecast system. *Wea. Forecasting*, **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. Amer. Meteor. Soc.*, **97**, 1617–1630, https://doi.org/10.1175/BAMS-D-14-00173.1.

Snyder, C., and F. Zhang, 2003: Assimilation of simulated Doppler radar observations with an ensemble Kalman filter. *Mon. Wea. Rev.*, **131**, 1663–1677, https://doi.org/10.1175//2555.1.

Sobash, R. A., and L. J. Wicker, 2015: On the impact of additive noise in storm-scale EnKF experiments. *Mon. Wea. Rev.*, **143**, 3067–3086, https://doi.org/10.1175/MWR-D-14-00323.1.

Stensrud, D. J., and Coauthors, 2009: Convective-scale warn-on-forecast system: A vision for 2020. *Bull. Amer. Meteor. Soc.*, **90**, 1487–1500, https://doi.org/10.1175/2009BAMS2795.1.

——, and Coauthors, 2013: Progress and challenges with warn-on-forecast. *Atmos. Res.*, **123**, 2–16, https://doi.org/10.1016/j.atmosres.2012.04.004.

Thompson, T. E., L. J. Wicker, and X. Wang, 2015: A comparison between the local ensemble transform Kalman filter and the ensemble square root filter for the assimilation of radar data in convective-scale models. *Quart. J. Roy. Meteor. Soc.*, **141**, 1163–1176, https://doi.org/10.1002/qj.2423.

Tong, M., and M. Xue, 2005: Ensemble Kalman filter assimilation of Doppler radar data with a compressible nonhydrostatic model: OSS experiments. *Mon. Wea. Rev.*, **133**, 1789–1807, https://doi.org/10.1175/MWR2898.1.

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. *Wea. Forecasting*, **30**, 1795–1817, https://doi.org/10.1175/WAF-D-15-0043.1.

Wicker, L. J., D. M. Wheatley, K. H. Knopfmeier, P. L. Heinselman, P. Skinner, J. J. Choate, T. A. Jones, and J. Poterjoy, 2018: Sensitivity experiments assessing ensemble forecast reliability for the 16 May 2017 western Oklahoma severe weather outbreak: Current status of NSSL's "Warn-on-Forecast" storm-scale prediction system. *Eighth Conf. on Transition of Research to Operations*, Austin, TX, Amer. Meteor. Soc., 14A.1, https://ams.confex.com/ams/98Annual/webprogram/Paper326888.html.

Yussouf, N., and D. J. Stensrud, 2010: Impact of phased-array radar observations over a short assimilation period: Observing system simulation experiments using an ensemble Kalman filter. *Mon. Wea. Rev.*, **138**, 517–538, https://doi.org/10.1175/2009MWR2925.1.