Verification of Quasi-Linear Convective Systems Predicted by the Warn-on-Forecast System (WoFS)

KELSEY C. BRITT,[a,b,c] PATRICK S. SKINNER,[a,b,c] PAMELA L. HEINSELMAN,[b,c] COREY K. POTVIN,[b,c] MONTGOMERY L. FLORA,[a,b] BRIAN MATILLA,[a,b] KENT H. KNOPFMEIER,[a,b] AND ANTHONY E. REINHART[b]

[a] *Cooperative Institute for Severe and High-Impact Weather Research and Operations, Norman, Oklahoma*
[b] *National Severe Storms Laboratory, Norman, Oklahoma*
[c] *School of Meteorology, University of Oklahoma, Norman, Oklahoma*

(Manuscript received 22 June 2023, in final form 29 September 2023, accepted 21 November 2023)

ABSTRACT: Quasi-linear convective systems (QLCSs) can produce multiple hazards (e.g., straight-line winds, flash flooding, and mesovortex tornadoes) that pose a significant threat to life and property, and are often difficult to accurately forecast. The NSSL Warn-on-Forecast System (WoFS) is a convection-allowing ensemble system developed to provide short-term, probabilistic forecasting guidance for severe convective events. Examination of WoFS's capability to predict QLCSs has yet to be systematically assessed across a large number of cases for 0–6-h forecast times. In this study, the quality of WoFS QLCS forecasts for 50 QLCS days occurring between 2017 and 2020 is evaluated using object-based verification techniques. First, a storm mode identification and classification algorithm is tuned to identify high-reflectivity, linear convective structures. The algorithm is used to identify convective line objects in WoFS forecasts and Multi-Radar Multi-Sensor system (MRMS) gridded observations. WoFS QLCS objects are matched with MRMS observed objects to generate bulk verification statistics. Results suggest WoFS's QLCS forecasts are skillful with the 3- and 6-h forecasts having similar probability of detection and false alarm ratio values near 0.59 and 0.34, respectively. The WoFS objects are larger, more intense, and less eccentric than those in MRMS. A novel centerline analysis is performed to evaluate orientation, length, and tortuosity (i.e., curvature) differences, and spatial displacements between observed and predicted convective lines. While no systematic propagation biases are found, WoFS typically has centerlines that are more tortuous and displaced to the northwest of MRMS centerlines, suggesting WoFS may be overforecasting the intensity of the QLCS's rear-inflow jet and northern bookend vortex.

SIGNIFICANCE STATEMENT: Quasi-linear convective systems (QLCSs), also known as squall lines, can be very destructive to life and property as they produce multiple hazards such as hail, severe straight-line winds, flash flooding, and tornadoes that typically form quickly and may be difficult to observe on radar. These storms can occur year-round and have the propensity to develop overnight or into the early morning hours, potentially catching the public off-guard. An ensemble prediction system called the Warn-on-Forecast System (WoFS), created by the National Severe Storms Laboratory, has shown promise in accurately forecasting a variety of severe weather events. This research evaluates the quality of the WoFS's QLCS forecasts. Results show WoFS can accurately predict these systems for forecast times out to 6 h.

KEYWORDS: Convective storms/systems; Squall lines; Forecast verification/skill; Numerical weather prediction/forecasting

1. Introduction

Quasi-linear convective systems (QLCSs) are linearly oriented mesoscale convective systems (MCSs) that typically have a longer length than width. These systems are commonly referred to as squall lines, or bow echoes when they contain one or more bowing segments along the leading convective line. QLCSs are capable of producing a variety of dangerous and destructive hazards such as straight-line winds, hail, flash flooding, and tornadoes produced from shallow, embedded mesovortices (Atkins et al. 2004; Trapp et al. 2005; Smith et al. 2012; Thompson et al. 2012; Brotzge et al. 2013; Ashley et al. 2019). In fact, a QLCS climatology study analyzing composite reflectivity data, spanning from 1996 to 2017 in the continental United States, found QLCSs account for around 21% and 28% of tornado and severe wind reports, respectively (Ashley et al. 2019). These systems can form year-round, including during the cold season (October–February; Trapp et al. 2005; Smith et al. 2012; Ashley et al. 2019), when most people may not expect severe weather to occur. QLCSs are also prone to forming during the night or in the early morning hours, potentially catching the sleeping public off-guard (Trapp et al. 2005; Ashley et al. 2019). These factors compound to make the population particularly vulnerable to QLCS hazards. Additionally, QLCS hazards (especially severe wind and tornadoes) tend to be more difficult to warn and have shorter warning lead times, which presents substantial forecasting and operational challenges (Brotzge et al. 2013).

The National Oceanic and Atmospheric Administration's (NOAA) National Severe Storms Laboratory (NSSL) is actively developing a quasi-operational, rapidly updating, convection-allowing ensemble data assimilation and prediction system as part of the Warn-on-Forecast project (WoF; Stensrud et al. 2009, 2013). This prediction system, known as the Warn-on-Forecast System (WoFS), was created to provide probabilistic numerical

Corresponding author: Kelsey C. Britt, kelsbritt@ou.edu

DOI: 10.1175/WAF-D-23-0106.1

weather prediction between the space and time scales of typical convective watches and warnings. WoFS has demonstrated an ability to provide skillful forecasts for a variety of hazardous weather events such as tornadic thunderstorms (Yussouf et al. 2013; Wheatley et al. 2015; Yussouf et al. 2015; Jones et al. 2016; Skinner et al. 2018; Flora et al. 2019; Britt et al. 2020), flash flooding events (Yussouf et al. 2016; Lawson et al. 2018; Yussouf and Knopfmeier 2019; Yussouf et al. 2020b), and tropical cyclones (Jones et al. 2019; Yussouf et al. 2020a). The capability of WoFS to accurately forecast QLCSs, however, has not been systematically analyzed. Past WoFS studies may include a QLCS event as one of their case studies (e.g., Wheatley et al. 2015; Kerr and Alsheimer 2022; Wang et al. 2022), but none have assessed WoFS's QLCS forecast performance by analyzing a large number of events. Therefore, the purpose of this study is to evaluate WoFS's capability to predict QLCS spatial extent, timing, and location at 0–6-h lead times.

To assess WoFS's skill in the prediction of QLCS events, object-based verification techniques are applied (e.g., Davis et al. 2006a,b). These verification methods have been used extensively within the WoFS framework (e.g., Skinner et al. 2018; Flora et al. 2019; Britt et al. 2020; Potvin et al. 2020; Laser et al. 2022; Miller et al. 2022; Guerra et al. 2022) to match thunderstorm objects forecasted by WoFS with observed objects in the gridded NEXRAD data from the Multi-Radar Multi-Sensor system (MRMS; Smith et al. 2016). In the current study, QLCS objects are identified using an object identification and classification algorithm (Potvin et al. 2022) and then are matched and verified using methods adapted from Skinner et al. (2018). Object-based verification metrics are then used to evaluate the overall performance of WoFS's 6-h QLCS forecasts. Furthermore, a novel centerline analysis is used to explore more specific errors between forecasted and observed QLCS objects (i.e., orientation, spatial displacement, and tortuosity/curvature), that are not considered in simple object matching. Those errors can then be connected to physical processes occurring within the QLCSs, which may help determine how WoFS may be representing these processes. Moreover, the centerline analysis may provide insight into the advantages and disadvantages of using WoFS to forecast QLCSs, and determine components of the forecast system that may need to be changed or improved in the future.

2. Data and methods

a. *Forecast and verification datasets*

The WoFS is a 36-member ensemble data assimilation and prediction system. WoFS was initialized using the High-Resolution Rapid Refresh Ensemble (HRRRE; Dowell et al. 2016) from 2017 to 2019, then the High-Resolution Rapid Refresh Data Assimilation System (HRRRDAS; Dowell et al. 2022) for 2020–present. The HRRRE/HRRRDAS provides initial and lateral boundary conditions for the 750 km × 750 km (2017/18) or the 900 km × 900 km (2019/20) WoFS domain. WoFS uses 3-km horizontal grid spacing and 51 vertical levels that extend from the surface to 10 hPa. WoFS is cycled every 15 min and uses an ensemble Kalman filter (EnKF; Houtekamer and Zhang 2016) to assimilate MRMS reflectivity (Smith et al. 2016), Weather Surveillance Radar-1988 Doppler (WSR-88D) radial velocity, Geostationary Operational Environmental Satellite (GOES) cloud water path retrievals (Jones et al. 2015, 2016, 2020), GOES clear sky radiances (Jones et al. 2018), and any available conventional observations from a variety of observing systems (e.g., Automated Surface Observing Systems, mesonets, aircraft, radiosondes). A Data Assimilation Research Testbed (DART)-based (Anderson and Collins 2007; Anderson et al. 2009) EnKF was used for all 2017/18 cases, and the Community Gridpoint Statistical Interpolation (GSI; Kleist et al. 2009; Hu et al. 2016) EnKF for the 2019/20 cases. Each WoFS member utilizes the NSSL two-moment microphysics scheme (Mansell et al. 2010) and the Rapid Update Cycle (RUC) land surface model (Smirnova et al. 2016). However, the planetary boundary layer (PBL) and radiation physics parameterizations are varied among forecast members to account for uncertainties in the model physics [see Table 1 in Skinner et al. (2018)]. Following analysis, WoFS issues 18-member forecasts at hourly intervals with durations up to 6 h.

TABLE 1. List of 2017–20 QLCS events. An asterisk indicates inclusion in the subset of events used for initial tuning of the QLCS identification algorithm (see the appendix).

2017	2018	2019	2020
4 May*	1 May*	30 Apr*	5 May
11 May*	2 May*	1 May	7 May
17 May	3 May	2 May	8 May
18 May*	7 May	3 May	13 May
19 May	9 May	6 May	15 May
24 May	12 May	7 May	21 May
27 May	14 May	8 May	22 May
	15 May	16 May	29 May
	19 May	17 May	19 Jun
	21 May	18 May	21 Jun
	28 May	20 May	26 Jun
	29 May*	21 May	17 Jul
	1 Jun	24 May	14 Aug
		28 May*	25 Aug
		29 May*	
		19 Jul	

The WoFS forecast dataset for this study consists of 50 days with QLCS events from 2017 to 2020 (Table 1). Since WoFS undergoes yearly testing and evaluation during the Hazardous Weather Testbed Spring Forecasting Experiment (e.g., Clark et al. 2020, 2021), most of the archived events are from the spring or early summer months. In this study, the majority of events occur in May and are concentrated in the southern Great Plains (Fig. 1). This study analyzes hourly 18-member forecasts issued between 2000 and 0200 UTC for each of the 50 days. The 15[1]-, 60-, 120-, 180-, 240-, 300-, and 360-min forecast times of each hourly initialized forecast are analyzed to examine forecast progression and fluctuations in verification statistics. Only the 15–180-min forecast times will be analyzed for the 2017/18 events, as WoFS forecast duration was only 3 h during those years.

[1] WoFS analyses are not used because they are generated through assimilation of MRMS reflectivity observations and closely resemble the observations (see the appendix).

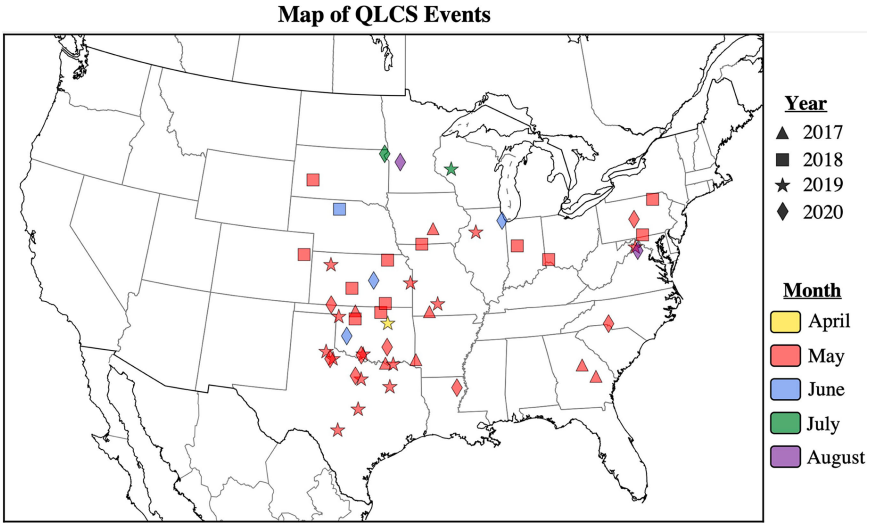

FIG. 1. Geographic and monthly distribution of the selected QLCS events. The symbols and colors represent the year and month of QLCS occurrence, respectively. Each point represents the center of the WoFS domain for each event.

Gridded composite reflectivity observations from the MRMS system serve as the verification dataset for the WoFS QLCS forecasts. MRMS provides real-time composite reflectivity from WSR-88D observations sampled across the continental United States. The MRMS data are interpolated to the 3-km WoFS grid using a Cressman filter with a 3-km radius of influence (Skinner et al. 2018).

b. QLCS object identification and classification

To identify QLCS objects across all events, a unique storm mode identification and classification algorithm is used (Potvin et al. 2022). The automated algorithm uses composite reflectivity to identify and classify storms according to a 7-class scheme (Potvin et al. 2022), but for this study only the QLCS objects are of interest. The algorithm was adapted in two stages of tuning to improve its skill in identifying intense linear structures characteristic of QLCS convective lines. The complete tuning process and how the defining thresholds for QLCS objects (Table 2) were created is outlined in the appendix.

The algorithm's identification process is two-pass. The first pass identifies regions of intense composite reflectivity using a double threshold method (Table 2). Initial storm objects in MRMS are identified as regions where composite and maximum reflectivity exceeds 40 and 45 dBZ, respectively, whereas the thresholds for composite and maximum reflectivity for WoFS storm objects must exceed 45 and 50 dBZ, respectively. Storm objects that have areas greater than 6 grid cells and exist within 12 km of each other will be merged into a single object. If the final object meets the criteria in Table 2 for minimum area, length, and eccentricity then it will be categorized as a QLCS. A second pass is applied to merge QLCS objects that are likely part of the same system. These QLCS objects have to be within 40 km of each other and have similar orientations (within 30°).

c. QLCS object matching and verification

The forecasted and observed QLCS objects, and their diagnostic properties (e.g., intensity, area, eccentricity, etc.), are extracted using the Scikit-image Python library and the regionprops function (van der Walt et al. 2014). The forecasted and observed objects are then matched according to a total interest score (TIS; Davis et al. 2006a,b), adapted from Skinner et al. (2018). The TIS is given as

$$\mathrm{TIS} = \frac{(\mathrm{md}_{\max} - \mathrm{md})}{\mathrm{md}_{\max}}, \tag{1}$$

where md is the minimum distance between the object pair and $\mathrm{md}_{\max}$ is the maximum threshold for object matching, which is set to 40 km [as in Skinner et al. (2018)]. The TIS is calculated for each of the object pairs and must be greater than 0.2 to be considered a possible match (Skinner et al.

TABLE 2. Tuned parameter thresholds for the storm ID and classification algorithm.

Parameter	Threshold value
Minimum reflectivity	40 dBZ (MRMS); 45 dBZ (WoFS)
Maximum reflectivity	45 dBZ (MRMS); 50 dBZ (WoFS)
Minimum area	54 km^2 (6 grid cells)
Merge distance	12 km (4 grid cells)
Length and eccentricity	If length >100 km, then eccentricity > 0.85; if length > 150 km, then eccentricity > 0.74

2018; Guerra et al. 2022). Unlike Skinner et al. (2018), the centroid distance is not included in the TIS calculation as the QLCS objects are often large and irregular in size and shape, which causes large variation in centroid placement. One limitation of solely using a minimum distance threshold for matching is that sometimes MRMS objects can be matched to multiple WoFS objects, and vice versa. To mitigate this occurrence, only the object pair with the highest TIS is retained, and that object pair is known as the best match. If there is a tie between pairs with the highest TIS, then the object pair with the smallest area difference is retained. The tie-breaker ensures the best possible match is made between two objects of similar size.

Object matching allows for the classification of object pairs according to 2 × 2 contingency table metrics (Brooks 2004). In this study, hits are defined as the best match (if one exists) for each MRMS object, unmatched MRMS objects are classified as misses, and false alarms are unmatched WoFS objects (Fig. 2). Extra matches may also occur, which is when a MRMS or WoFS object is matched to multiple WoFS or MRMS objects, respectively (Fig. 2). In this case, only the best match is retained and considered a hit, whereas the extra matches are ignored so as to not unfairly reward or punish WoFS (Fig. 2). Correct negatives are not produced through object matching, so verification statistics are limited to those that only use hits, misses, and false alarms in their calculations. Therefore, the verification metrics (Brooks 2004) used herein are probability of detection [POD; Eq. (2)], false alarm ratio [FAR; Eq. (3)], critical success index [CSI; Eq. (4)], and frequency bias [Eq. (5)]. They are given by the following equations:

$$\mathrm{POD} = \frac{H}{H + M}, \tag{2}$$

$$\mathrm{FAR} = \frac{F}{H + F}, \tag{3}$$

$$\mathrm{CSI} = \frac{H}{H + M + F}, \tag{4}$$

$$\mathrm{bias} = \frac{H + F}{H + M}, \tag{5}$$

where H, M, and F represent hits, misses, and false alarms, respectively. Additionally, matching the MRMS and WoFS QLCS objects enables the comparisons of their diagnostic properties, such as the area, eccentricity, aspect ratio, and the maximum intensity. One limitation of this study is that QLCS cases were selected based on the presence of a QLCS in MRMS. This selection process neglects cases where WoFS members predict a QLCS but one is not observed; therefore, the FAR values presented herein are best thought of as a lower bound of the true FAR.

d. QLCS centerlines

A centerline analysis is a technique that enables the examination of specific spatial errors between predicted and observed QLCS objects that are not available through comparison of whole QLCS objects. A centerline is defined as the line that runs lengthwise down the center of the QLCS object and is a proxy for the QLCS's leading convective line. Centerlines are generated for each of the MRMS-WoFS best-matched pairs using a customized python algorithm (Ungar 2022). First, the object is smoothed using a Gaussian filter with a radius of 5 grid points and converted to a binary object (Fig. 3b). The binary object is created by giving all points contained within the object the label 1, while all points outside the object are assigned the label 0. The binary object is then converted to a polygon using the Shapely python package (Gillies et al. 2023; Fig. 3c). The object polygon is loaded into the centerline algorithm, which identifies the polygon's perimeter and uses a Voronoi diagram technique (Aurenhammer 1991) to extract the polygon's skeleton (Fig. 3c). The object's centerline is isolated from the larger skeleton by trimming the shorter branches to find the longest shortest path between any pair of perimeter nodes (Fig. 3d). In other words, the shortest paths between all perimeter nodes are calculated and the longest of those paths is then smoothed to become the object's centerline (Fig. 4). A limitation of the centerline analysis function is that it typically has higher variability in line placement at either end of the object. To lessen this variability, only the centerline points between the 10th and 90th percentiles[2] of its length are analyzed (Fig. 4). However, there remains some uncertainty at the centerline endpoints that are, unfortunately, collocated with particularly interesting portions of the QLCS, such as bookend vortices.

Centerline characteristics examined are spatial displacements and differences in length, orientation, and tortuosity. The orientation of the centerline is defined as the angle between east (0°) and the endpoint line. For example, if the centerline orientation is 90° then the line would be purely north–south-oriented, with the maximum allowable angle being 179°. Tortuosity is a dimensionless parameter commonly used in health (Ciurică et al. 2019) and material sciences (Fu et al. 2021) to characterize the porosity of a substance, such as arteries or nerves. In this study, tortuosity is used to measure centerline curvature and is calculated using the following equation:

$$\mathrm{tortuosity} = \frac{L_c}{L_e}, \tag{6}$$

where L_c is the centerline length and L_e is the length of the centerline's endpoint line (Fig. 4). Thus, tortuosity increases as the centerline's curvature increases.

3. Results

a. WoFS overall performance

The overall forecast performance of WoFS is assessed by calculating POD, FAR, CSI, and frequency bias for all QLCS

[2] The centerline's coordinates are ordered so that they start at the QLCS's southern tip and end in the north. The percentiles are calculated over the length of the coordinate array so that there is one coordinate pair corresponding to every 5th percentile. Percentiles are used to ensure we are sampling similar parts of the centerline, regardless of its shape or length (e.g., the 50th percentile will always be at the center of the line).

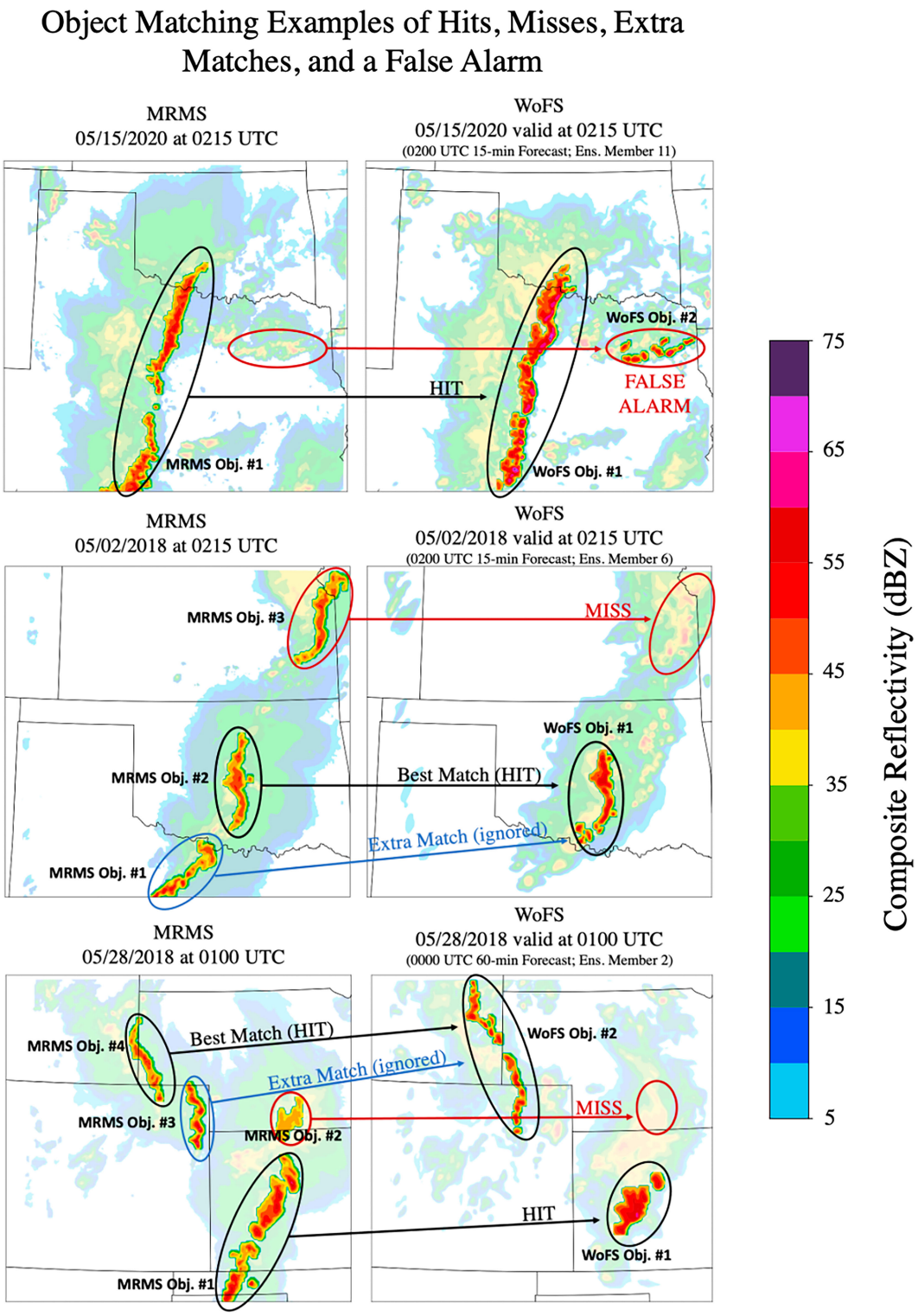

FIG. 2. (left) MRMS composite reflectivity and (right) WoFS composite reflectivity forecasts displaying examples of hits, misses, extra matches, and a false alarm in the object matching process. Best matches in MRMS are considered hits, unmatched MRMS objects are classified as misses, and false alarms are unmatched WoFS objects. Extra matches occur when a single MRMS or WoFS object is matched to multiple WoFS or MRMS objects, respectively. Only the best match is retained and the extra matches are ignored.

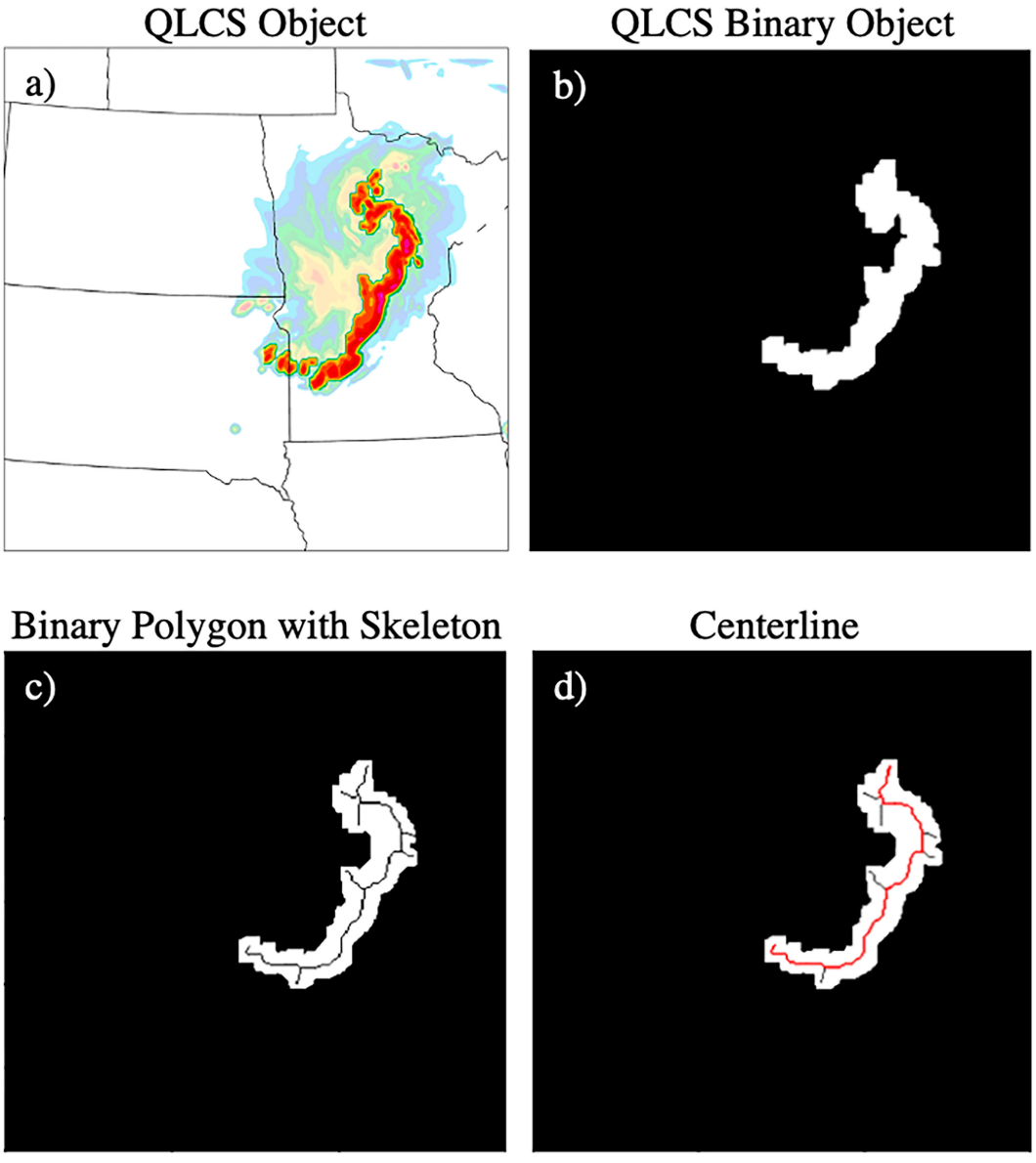

FIG. 3. Illustration of the centerline creation process. Starting with (a) the QLCS object of interest, (b) QLCS object converted to a binary object, (c) binary object converted to a polygon with object skeleton overlaid (black line), and (d) the resulting centerline (red) from the full skeleton. The centerline is then smoothed to create the final centerline product as shown in Fig. 4.

events (Fig. 5). The statistics are calculated over all events by aggregating the total number of hits, misses, and false alarms for all composite reflectivity ensemble forecasts initialized between 2000 and 0200 UTC (i.e., 36 540 total forecasts). The total number of QLCS objects identified across all events in MRMS is 3207, and the total across all events, initialization times, lead times, and ensemble members in WoFS is 50 810 (an average of approx. 2823 objects per member).

WoFS QLCS forecasts have a POD near 0.64 during the first hour of the forecast period (Fig. 5a). After 60 min, POD decreases to approximately 0.59 at the end of the 3- and 6-h forecasts (Fig. 5a). FAR values for the entire 6-h forecast period stay below 0.4, with a sharp increase in values evident within the first 60 min (Fig. 5b). This sharp increase in FAR also coincides with the increase in POD (Fig. 5a) and frequency bias (Fig. 5c), and may be attributable to WoFS's convective spinup process. During this spinup process, more spurious reflectivity echoes are present in the forecast, which increases POD, FAR, and bias. Skinner et al. (2018) found a similar trend in frequency bias when examining all WoFS reflectivity objects, in which the bias was the highest in the first

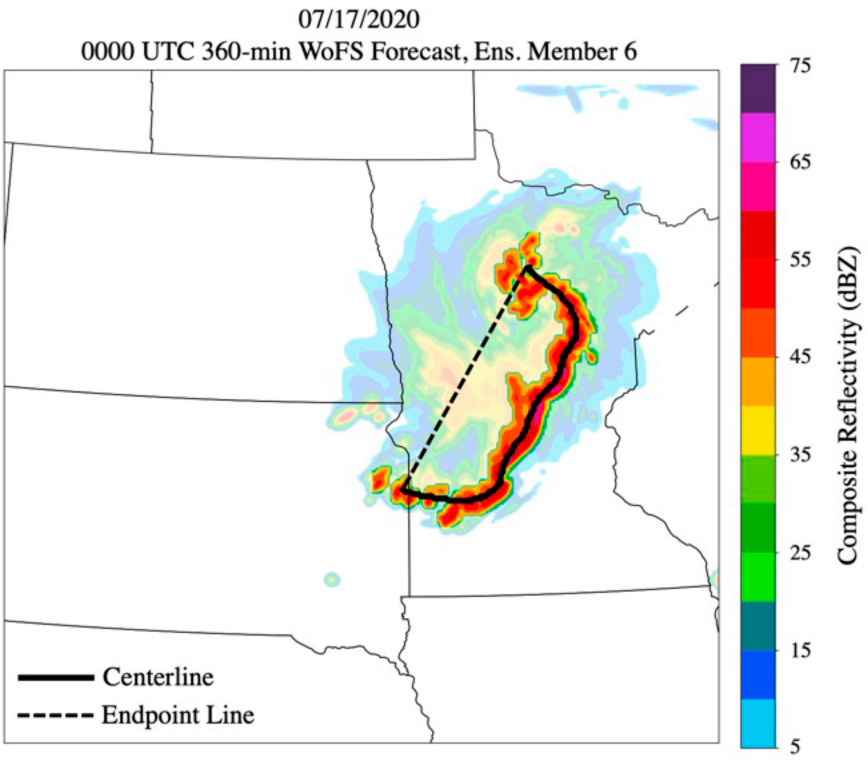

FIG. 4. Example of a centerline overlaid on the WoFS output for the 17 Jul 2020 QLCS during the 0000 UTC forecast for ensemble member 6 at the end of the 6-h forecast. The centerline (solid black line) extends from the 10th–90th percentiles of the total QLCS object's length. The endpoint line (dotted black line) is the straight path connecting the centerline's endpoints. The QLCS object is represented as the outlined area of reflectivities, with the faded area of reflectivities being the full WoFS output at this time.

60 min of the forecast. After 60 min, the bias[3] begins to decrease with time indicating these weak, spurious storms are short-lived and dissipate by the end of the forecast (Fig. 5c). CSI for QLCS forecasts is maximized at the beginning of the forecast with values around 0.52, then subsequently decreases with time with values near 0.46 and 0.44 at the end of the 3- and 6-h forecast, respectively (Fig. 5d). Examining the individual members by PBL scheme shows no substantial differences (Fig. 5), consistent with results from previous research by Potvin et al. (2020).

Verification statistics were also generated for all Spring Forecasting Experiment cases that occurred in April and May from 2017 to 2020 (total of 85 cases; Fig. 6). This dataset contains storms of all types, not just QLCSs. Therefore, comparing the two sets of verification statistics (Figs. 5 and 6) will highlight differences between forecasts for just QLCSs and for all storm types.[4] POD for all storms is maximized within the first 30 min of the forecast, but decreases consistently afterward (Fig. 6a), whereas QLCS forecasts maintain POD values around 0.6 for most of the 6-h forecast (Fig. 5a). QLCS forecasts stay below 0.4 for FAR throughout the 6-h forecast (Fig. 5b), but FAR for all storms passes 0.4 in the first 15 min of the forecast and increases to approximately 0.65 at the end of the period (Fig. 6b). Frequency bias for QLCS forecasts lingers just below 1 for the entire forecast (Fig. 5c) and is mostly above 1 for forecasts for all storms (Fig. 6c). Finally, CSI for QLCS forecasts lingers around 0.5 (Fig. 5d), while the CSI for all storms reaches 0.5 at the beginning of the forecast but decreases consistently throughout the rest of the forecast (Fig. 6d). Therefore, WoFS is exhibiting higher forecast skill for QLCS objects compared to all storm objects, and that skill is retained

[3] Frequency bias was also calculated using the total number of MRMS and WoFS objects, including the objects from the extra matches. The bias still remained between 0.8 and 0.9 as in Fig. 5c. This small change in bias when including the extra objects indicates the extra objects are nearly evenly distributed between WoFS and MRMS.

[4] Results from the full set of reflectivity objects do not use an identical methodology to the one used herein. Object identification and matching follow the methodology outlined in Guerra et al. (2022), but no object merging is performed (the extra objects are not included in the calculation of contingency table statistics). While this is not a direct comparison, it does provide a general comparison of the accuracy between QLCS and all convective storm forecasts.

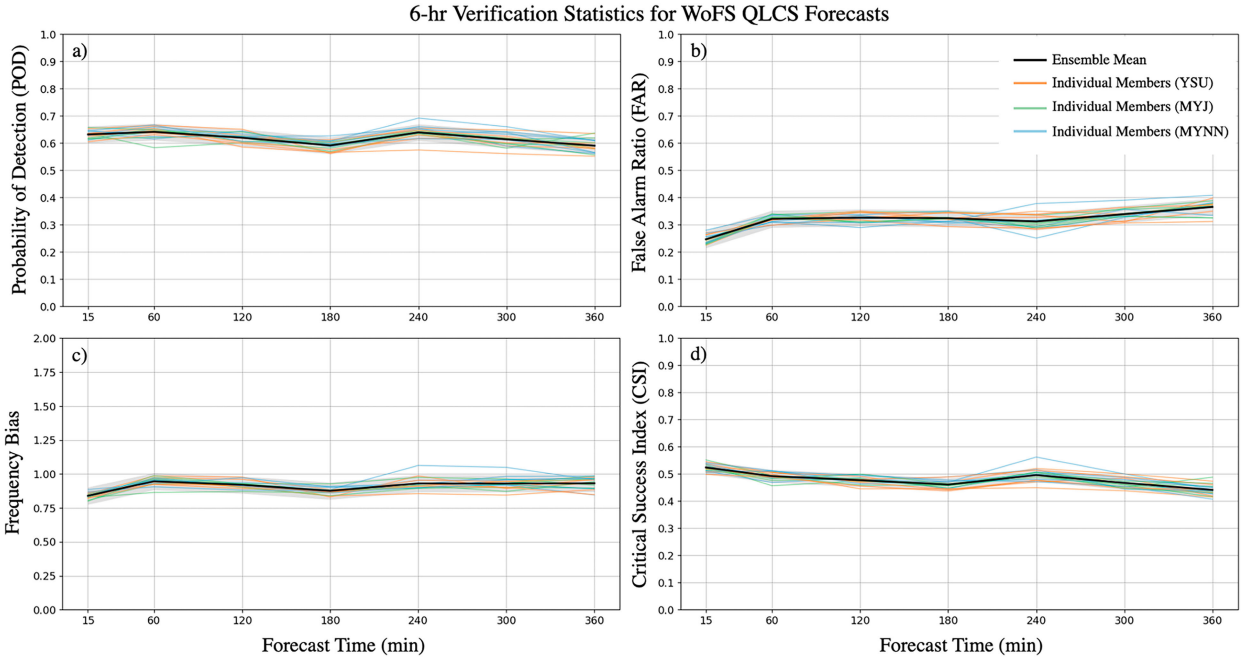

FIG. 5. Time series of verification statistics for WoFS QLCS forecasts: (a) POD, (b) FAR, (c) frequency bias, and (d) CSI. The ensemble mean is in black and the individual ensemble members (1–18) are represented by thin orange, green, and blue lines that represent the ensemble members with YSU, MYJ, and MYNN PBL schemes, respectively. Gray shading denotes the 95% confidence interval using a bootstrapping technique with replacement (N = 10 000). The ensemble mean is found by aggregating all hits, misses, and false alarms over all composite reflectivity forecasts initialized between 2000 and 0200 UTC for all years and ensemble members.

throughout the 6-h forecast. This finding suggests WoFS may be better at forecasting QLCSs over other storm modes because QLCSs can be strongly forced (e.g., by a cold front; Wheatley and Trapp 2008) and fairly large, which would make them more easily resolved on the 3-km WoFS grid, and have a longer intrinsic predictability limit (Weyn and Durran 2017, 2019).

WoFS forecast performance can also be assessed by aggregating all events on performance diagrams to examine POD and success ratio (SR) differences between events, and how performance changes with forecast time. However, instead of traditional performance diagrams, Figs. 7 and 8 are formatted like heat maps so one can easily ascertain where the highest concentration of events exist. Each point on the performance heat diagrams represents a single event's ensemble mean POD and SR for each of the hourly initialized forecasts at that specific forecast time (Fig. 7). Therefore, each event may have up to seven points on each diagram in Fig. 7, one for each of the 2000–0200 UTC forecasts. The points are then sorted into 0.1 POD × 0.1 SR bins. Finally, all points from each panel in Fig. 7 are aggregated onto a single performance heat diagram covering the 6-h forecast period (Fig. 8).

The heat diagrams are separated by forecast time to understand how WoFS performance changes with forecast progression (Fig. 7). Inspection of the performance heat maps by forecast time reveals three main areas where the majority of events are concentrated: the top-right corner, the right side where POD = 0.5, and the bottom-left corner (Figs. 7 and 8). The key reason why these locations are favored is because QLCSs are often sparse (i.e., only one or two objects) within the WoFS domain for any given forecast. For example, if there are only 2 observed QLCS objects then WoFS will either correctly identify all objects, identify one but miss the other, or miss both objects. This results in the three favored locations seen on the performance heat diagrams (Figs. 7 and 8). The top-right corner is where POD and SR equal 1, indicating a perfect forecast. Having events clustered in this location suggests WoFS is able to forecast many QLCS events very well, without missing any objects. The second location is along the right-side of the diagram where POD = 0.5 and SR = 0.8–1.0. Having events in this area means WoFS is predicting half of the objects correctly, but missing the other half. Finally, the last favored location is in the bottom-left corner near zero, indicating WoFS missed all objects in the forecast.

While the favored locations do not change with forecast time, the percentage of events in those locations does (Fig. 7). In the top-right corner, the percentage of perfect forecasts decreases with forecast time, whereas the percentage of missed forecasts (i.e., those with POD = 0 and SR = 0) increases with time. Note the decrease in sample size from 180 min (Fig. 7d) to 240 min (Fig. 7e) because the 2017 and 2018 events have a maximum forecast period of 3 h.

There are 260 missed (i.e., no hits in WoFS) QLCS events across all forecasts, accounting for 14% of all events (Fig. 8). Approximately 50% of these missed forecasts occur during the 2000–2200 UTC initializations (not shown). Past studies

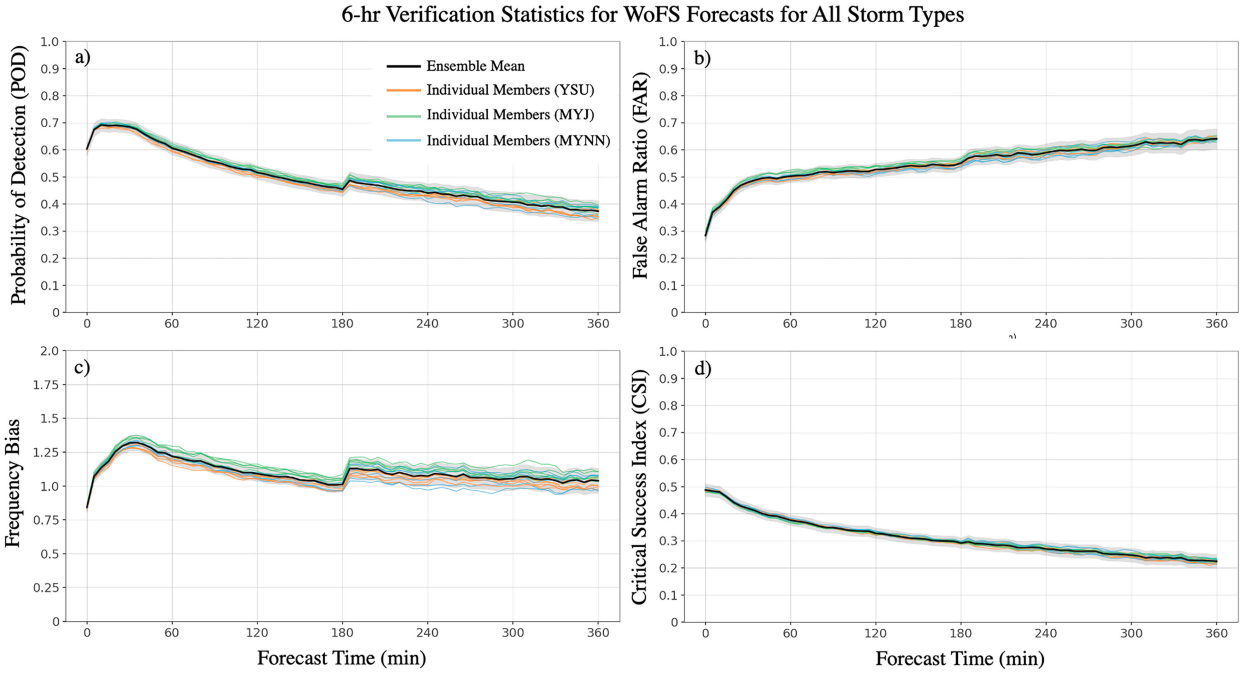

FIG. 6. As in Fig. 5, but for all Spring Forecasting Experiment (SFE) cases in April and May from 2017 to 2020 (85 total events), and data are output every 5 min. The slight increase in POD, FAR, and bias at the 180-min forecast time occurs because of the shift from DART to GSI DA software. (Figure courtesy of Patrick Skinner.)

have found the accuracy of WoFS forecasts is primarily dependent on the maturity of the storm within WoFS's initial conditions (Flora et al. 2019; Guerra et al. 2022). The higher proportion of missed events in early forecast initializations is consistent with the fact that QLCSs often haven't matured or even initialized by 2000–2200 UTC, in which case they are not accurately represented in the WoFS initial conditions.

The performance diagrams show a consistent low-bias in WoFS (Figs. 7 and 8) that was also evident in Fig. 5c where the bias was less than one for the whole 6-h forecast. There is a high concentration of events that exist below the bias = 1 line, meaning WoFS is underforecasting the number of QLCS objects in the majority of events (Figs. 7 and 8). This low-bias may be attributable to a few factors. Subjective analysis of QLCSs within MRMS and WoFS found that it is common for MRMS QLCSs to be broken into multiple objects, whereas the full convective line may be only one object in WoFS. MRMS also contains more stratiform region objects that often go unmatched. Both situations generate more objects in MRMS than WoFS, which would lead to the overall low bias. The bias may also be manipulated by changing the thresholds in the object identification algorithm.[5] The algorithm for this study was tuned to match subjective interpretation and to optimize performance (see the appendix).

[5] Changing the thresholds may change the numerical values, but the overall trends remain the same.

b. *QLCS object characteristics*

Comparing WoFS and MRMS QLCS object characteristics from each best match identified differences in area (Fig. 9a), maximum intensity (i.e., reflectivity; Fig. 9b), aspect ratio (Fig. 9c), and eccentricity (Fig. 9d). All differences are calculated by subtracting the MRMS value from the WoFS value. Therefore, positive values indicate the WoFS object value is larger than the MRMS value, and vice versa.

Analysis of object area differences shows WoFS objects are on average larger than MRMS objects at later forecast times (Fig. 9a). The positive area bias may be connected to the substantial positive bias observed in object maximum intensity (Fig. 9b), indicating WoFS objects are more intense than MRMS objects. Part of this positive intensity bias is attributable to the Cressman filter that was applied to MRMS reflectivity, which typically reduces the maximum intensity by about 5 dBZ. Additionally, WoFS uses the NSSL two-moment microphysics scheme, which was found to overpredict reflectivity values (Skinner et al. 2018). Closer visual examination of the WoFS QLCS events found the WoFS QLCS objects typically include more of the stratiform region than do the MRMS objects (Fig. 10). While WoFS QLCS objects are larger than those in MRMS, the spatial coverage of the full storm in WoFS is generally much smaller (Fig. 10). The storms in WoFS also appear to have little to no transition zone (Biggerstaff and Houze 1993), which is an area of low reflectivity that exists between the leading convective line and the trailing stratiform region (Fig. 10). The lack of a transition zone removes the boundary separating the stratiform precipitation from the convective line, making it easier for more of the stratiform region

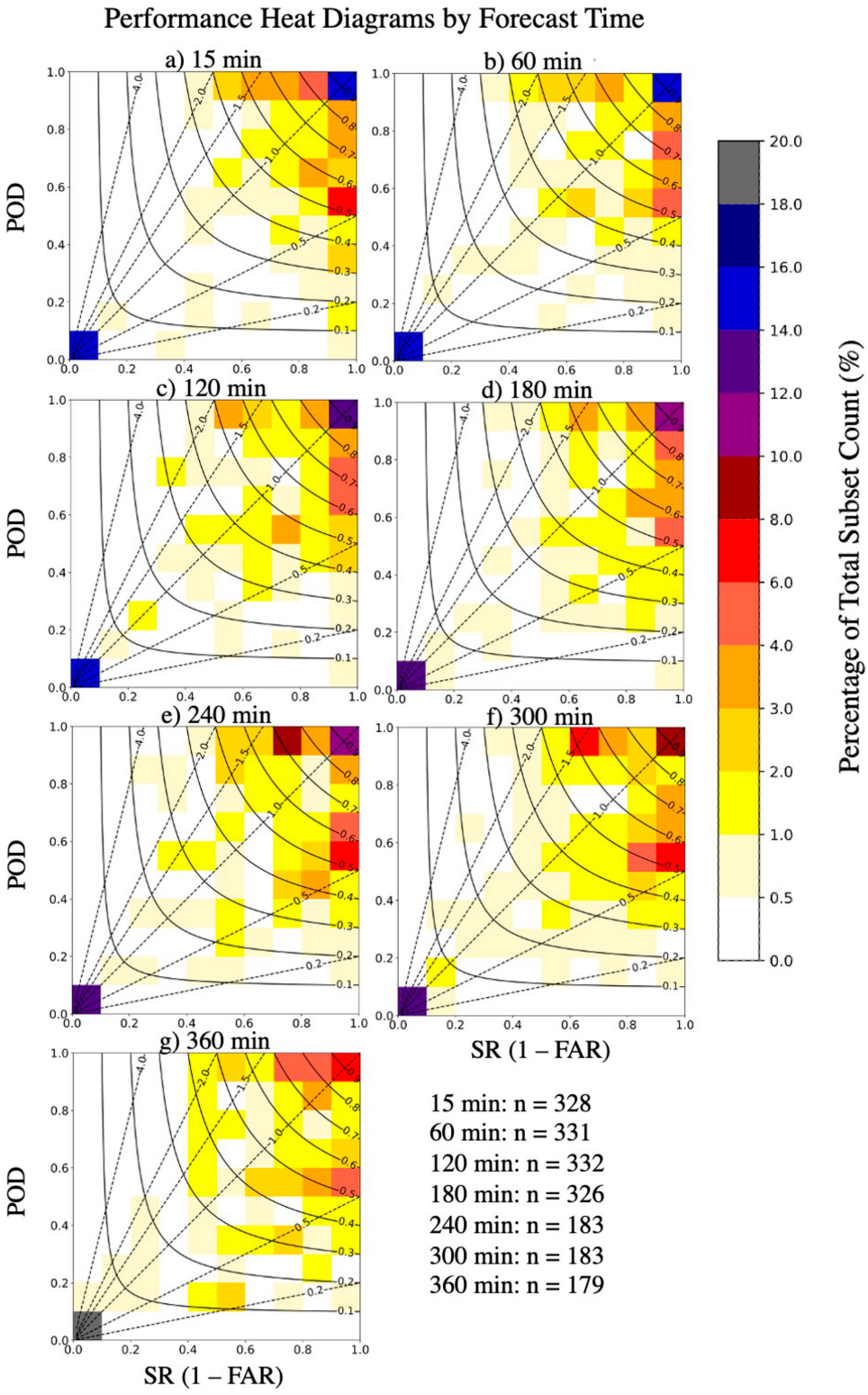

FIG. 7. Performance diagrams displayed as heat maps to easily show the locations that contain the highest percentage of points. Each point represents the event's ensemble mean POD and SR for a single forecast at that specific forecast time (up to 7 points total from the 2000–0200 UTC forecasts for each event). Each panel represents one forecast time: (a) 15, (b) 60, (c) 120, (d) 180, (e) 240, (f) 300, and (g) 360 min. The black dotted lines and curved lines represent bias and CSI, respectively. The total number of forecasts (n) used in each panel is shown in the bottom-right corner.

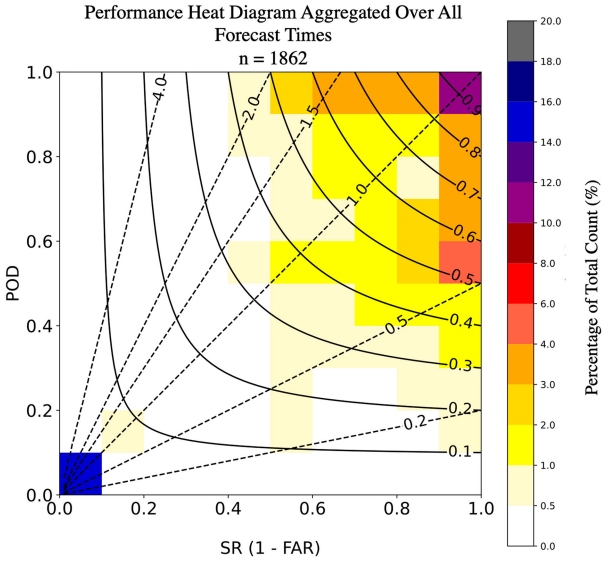

FIG. 8. All panels in Fig. 7 aggregated onto one plot for all forecast times. The total number of forecasts used is given by *n*.

to be included as part of the convective line object. This pattern is consistent with more of the stratiform region being identified with the convective line objects in WoFS and the size difference between WoFS and MRMS objects (Fig. 9a).

Differences in object aspect ratio and eccentricity also suggest WoFS is retaining too much of the stratiform region compared to MRMS (Figs. 9c,d). Aspect ratio differences show a consistent positive bias across forecast time (Fig. 9c). Recall, aspect ratio is the ratio of minor axis length to major axis length. Hence, positive aspect ratio differences indicate MRMS objects have shorter minor axes compared to WoFS objects (Fig. 9c). Additionally, if MRMS objects have shorter minor axes then the objects will also appear slimmer and more eccentric, which is evident from the negative eccentricity bias (Fig. 9d). Conversely, these biases indicate WoFS objects have wider minor axes and lower eccentricities (Figs. 9c,d), which is consistent with the WoFS objects retaining more of the stratiform region (Fig. 10).

c. *QLCS centerline analysis*

Centerline analyses allow for the quantification of specific errors in QLCS convective line characteristics that are not available in typical object-based diagnostics. Differences in centerline tortuosity (Fig. 11a), length (Fig. 11b), and orientation (Fig. 11c) between best-matched pairs are evaluated to understand how the simulated WoFS QLCSs differ from those in MRMS.

The 60–180-min forecasts show a slight negative bias in tortuosity indicating MRMS centerlines are typically more curved within this time period (Fig. 11a). Following the 180-min forecast, the bias becomes positive meaning the WoFS centerlines have slightly more curvature than MRMS at longer forecast times (Fig. 11a). The large separation between the mean and median tortuosity values (Fig. 11a) may be attributable to the high-end bow echo events. Since not all QLCSs contain bowing segments, or even transition to a bow echo, those that do will skew the mean tortuosity because they have more curvature than standard

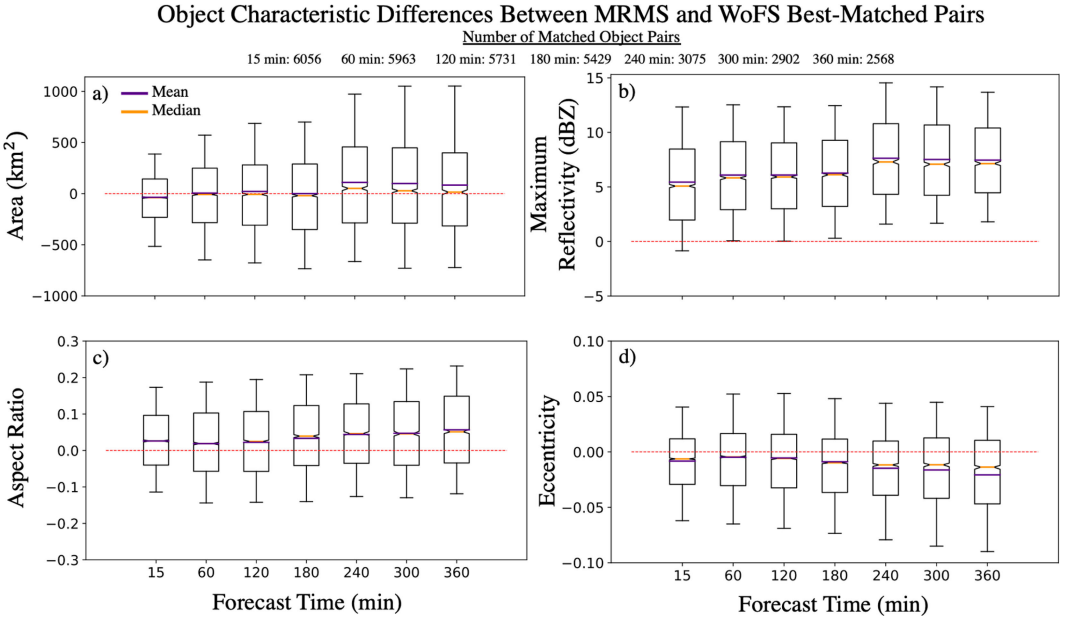

FIG. 9. Notched boxplots illustrating the differences in object (a) area, (b) maximum intensity, (c) aspect ratio, and (d) eccentricity between WoFS and MRMS best-matched pairs with increasing forecast time. Differences are calculated by subtracting the MRMS value from the WoFS value. The dotted red, solid purple, and solid orange lines represent the zero line, mean, and median, respectively. Whiskers extend to the 10th and 90th percentiles. Notches represent the 95% confidence interval of the median using a bootstrap method with replacement (N = 10 000).

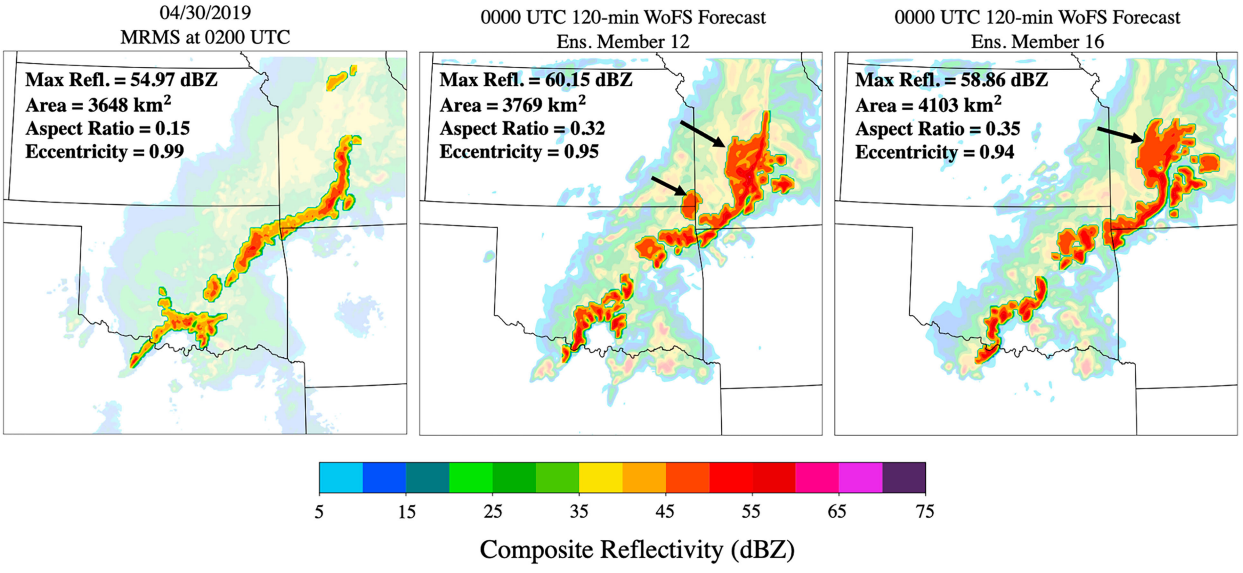

FIG. 10. Illustration of object characteristic differences in (left) MRMS, and in the 0000 UTC 120-min WoFS forecasts for ensemble member (center) 12 and (right) 16. Object properties (maximum reflectivity, area, aspect ratio, and eccentricity) are annotated in the top-left corner of each panel. Black arrows denote the areas of stratiform being included as part of the convective line object. WoFS objects tend to be more intense and include more of the stratiform region, which leads to the objects having higher aspect ratios and lower eccentricities compared to MRMS objects.

QLCSs. Centerline length differences show a small negative bias throughout the entire 6-h forecast, indicating MRMS centerlines are typically longer than those in WoFS (Fig. 11b). Finally, differences in orientation angle show a positive bias that increases with forecast time (Fig. 11c). On average, MRMS centerlines are oriented between 60° and 70° (not shown), while WoFS centerlines start off oriented between 60° and 80° then rotate cyclonically with increasing forecast time (not shown). Therefore, WoFS centerlines tend to be oriented in the same direction as MRMS centerlines (southwest–northeast direction) at the beginning of the forecast, but become more aligned in the south–north direction at the end of the forecast (see the centerline panels in Fig. 12; further discussion later).

The cyclonic rotation of WoFS centerlines and the spatial displacement between MRMS and WoFS centerlines is further explored in Figs. 13 and 14. Spatial displacement calculations between WoFS and MRMS centerlines are simplified by comparing every 5th percentile in the centerlines. The result is one coordinate pair for WoFS and MRMS corresponding to each percentile that can then be easily used to find the displacement between lines. Displacement calculations use the WoFS centerline position as the reference point and subtract the MRMS position. For example, a northern bias indicates WoFS centerlines are located to the north of MRMS centerlines. Kernel density estimation (KDE) is used to highlight areas where the majority of the full distribution's displacement occurs (Fig. 13).

Separating the displacement by forecast time allows for the examination of potential biases that may exist at certain times in the forecast or that grow with time (Fig. 13). During the first 3 h of the forecast, the displacement is evenly centered around the origin (Figs. 13a,b). However, the displacement begins to shift toward the northwest following the 3-h forecast (Figs. 13c,d), indicating WoFS centerlines are typically located to the northwest of MRMS centerlines. The full displacement distribution has the shape of a rectangle with a major axis oriented from roughly southwest to northeast (Fig. 13). When the displacement is in the northeast quadrant then the WoFS centerline is traveling ahead of MRMS, but if the displacement is located in the southwest quadrant then the MRMS centerline is leading WoFS. Therefore, more spread in the northeast and southwest directions suggests the storm propagation speed is most likely the main factor contributing to the displacement between centerlines given that the spread occurs along climatologically favorable QLCS propagation vectors. While WoFS QLCS propagation errors are often large, they are relatively unbiased overall.

Centerline displacement is dissected further by dividing the QLCS centerlines into four sections based on percentiles ranges: the southern end (0th–25th percentiles; Figs. 14a,b), south-central portion (25th–50th percentiles; Figs. 14c,d), north-central portion (50th–75th percentiles; Figs. 14e,f), and the northern end (75th–100th percentiles; Figs. 14g,h).

The displacement in the 15-min forecasts (Figs. 14a,c,e,g) is similar to Fig. 13a in that it is mostly centered around the origin. However, at the end of the 360-min forecasts there is an evident shift in the highest density of displacements away from the origin by approximately 25 km (Figs. 14b,d,f,h). The southernmost end has a mostly southern displacement bias, indicating the southern portions of WoFS QLCSs tend to be displaced to the north of those in MRMS (Fig. 14b). The south- and north-central portions of the QLCS show similar displacement biases to the north

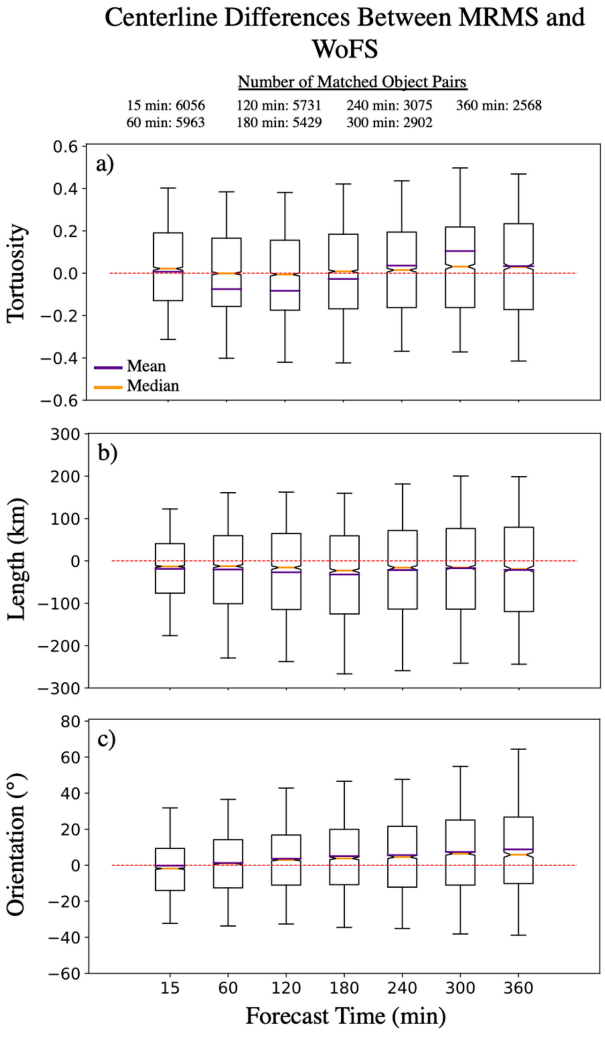

FIG. 11. As in Fig. 9, but showing the differences in (a) tortuosity, (b) length, and (c) orientation between the WoFS and MRMS best-matched centerline pairs. Differences are calculated by subtracting the MRMS value from the WoFS value.

(Figs. 14d,f), with the south-central portion (Fig. 14d) having a larger western component to its bias than seen in the north-central portion (Fig. 14f). Last, the northern ends of WoFS QLCSs are located to the northwest of that observed with MRMS (Fig. 14h). The shift in WoFS centerlines to the northwest with increasing forecast time corroborates the cyclonic change in orientation angle that was observed in WoFS centerline properties (Fig. 11c).

4. Discussion

A QLCS event that occurred on 17 July 2020 is chosen as a representative example to synthesize and discuss the results presented within this study (Fig. 12). This QLCS developed in North Dakota and progressed southeast into Minnesota and parts of South Dakota. This convective system was associated with over 100 wind, 29 hail, and 7 tornado reports.

This QLCS was predicted accurately by WoFS, with a total POD, FAR, CSI, and frequency bias of 0.94, 0.05, 0.90, and 0.99, respectively. However, there remain many differences between the simulated WoFS QLCS and the one present in MRMS (Fig. 12). First, the maximum composite reflectivities for the WoFS QLCS objects are about 10–20 dBZ higher than those in MRMS (Fig. 12), which is consistent with the large positive bias in maximum reflectivity between WoFS and MRMS objects for all events (Fig. 9b). Additionally, the WoFS QLCS object is larger in size, which corroborates the positive area bias between WoFS and MRMS QLCSs (Fig. 9a).

One of the largest differences between the simulated QLCS in WoFS compared to MRMS is WoFS predicts a more intense northern bookend vortex[6] (Fig. 12). Bookend vortices are a pair of counterrotating vortices that typically form in the midlevels (3–7 km) on either end of the bow echo (Weisman 1993; Weisman and Davis 1998; Atkins et al. 2004; Meng et al. 2012). The northern, cyclonic bookend vortex is usually favored through the convergence of planetary vorticity and will grow stronger than the southern, anticyclonic vortex (Wheatley and Trapp 2008; Atkins and Laurent 2009). Previous work (Weisman 1992, 1993; Grim et al. 2009) has shown that the combined circulation between the bookend vortices can contribute to the development and intensification of the rear inflow jet (RIJ; Smull and Houze 1987). Therefore, WoFS's intense northern bookend vortex may be indicative of a strong RIJ that is causing convective lines in WoFS QLCSs to bow out more than those in MRMS QLCSs (Fig. 12). These bowing segments may not be system-wide, but localized near the northern bookend vortex. Further evidence of this overprediction occurring in WoFS can be found in Fig. 15 where there is an intense northern bookend vortex and RIJ present in the 180-min forecast, but is absent in the analysis.

The overprediction of the northern bookend vortex and RIJ intensity in WoFS is also consistent with the results of the centerline analysis. Differences in centerline properties found WoFS centerlines tend to be more tortuous (Fig. 11a) and their orientations turn cyclonically with increasing forecast times (Fig. 11c). An intense RIJ may cause the WoFS centerlines to bow out more, thus increasing the line tortuosity with time (Fig. 12). Similarly, a stronger northern bookend vortex is consistent with the cyclonic rotation of the QLCS's orientation, and why WoFS QLCSs tend to be displaced to the northwest of MRMS QLCSs (Figs. 13 and 14). This cyclonic rotation in orientation is evident in the last centerline panel in Fig. 12.

5. Summary and conclusions

WoFS has been shown to accurately forecast a variety of severe convective events. However, evaluating WoFS's capability to accurately forecast QLCSs using a large number of events has yet to be assessed. Therefore, this study examines 50 QLCS days between 2017 and 2020 to assess WoFS's capability to predict QLCS spatial extent, timing, and propagation.

[6] Subjective analysis found this overprediction occurs at least once in 40 out of the 50 QLCS events.

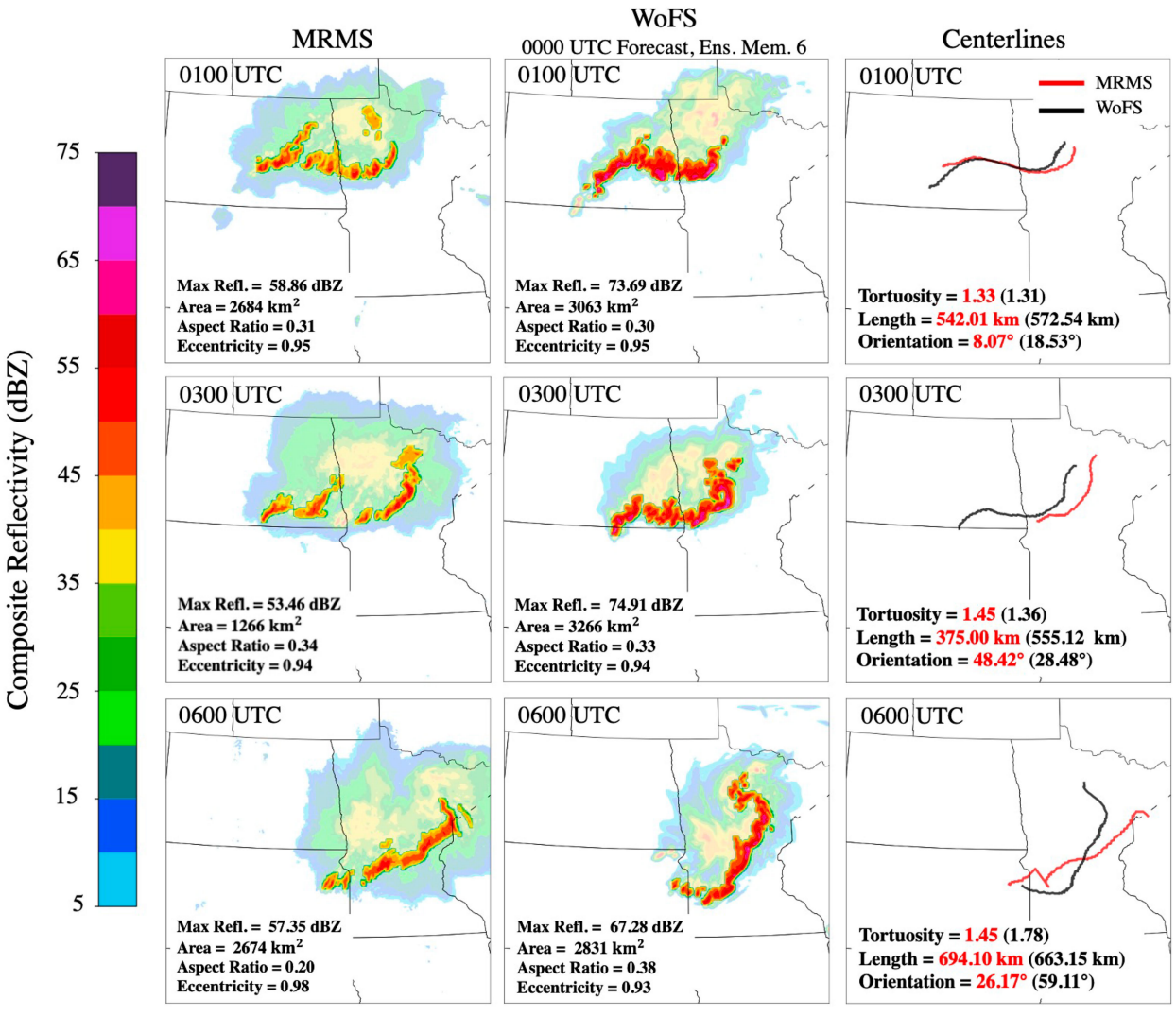

FIG. 12. (left) MRMS composite reflectivity, (center) WoFS composite reflectivity, and (right) the object centerlines from the 17 Jul 2020 QLCS event. The time that each panel is valid for is given in the top left of each panel. The WoFS 60-, 180-, and 360-min forecasts are from the 0000 UTC initialization for ensemble member 6. MRMS and WoFS centerlines are red and black, respectively, with their associated tortuosity, length, and orientation properties annotated in their respective color. Note: There are two QLCS objects present in MRMS at 0300 UTC, but the rightmost object is the best match and, therefore, used for comparison to the WoFS object.

The WoFS QLCS forecast performance is evaluated using object-based verification techniques and a novel centerline analysis used to highlight potential biases in spatial displacement, tortuosity, length, and orientation that may exist when comparing the QLCSs from WoFS and MRMS.

Results show WoFS has skill in forecasting QLCSs out to 6 h of forecast time with POD, FAR, and CSI values of the 3-h forecasts (0.57, 0.35, and 0.44) being similar to those at the end of the 6-h forecasts (0.55, 0.39, and 0.42; Fig. 5). On the other hand, verification statistics for all storm types have lower POD and CSI values, and much higher FAR values (Figs. 6a,b,d), suggesting WoFS is able to more accurately predict QLCSs than other storm modes. This is expected as QLCSs are larger and longer-lived systems and, therefore, should have longer intrinsic predictability limits than cellular convection (Weyn and Durran 2017, 2019).

We found that WoFS QLCS objects are typically larger, more intense, and less eccentric than MRMS QLCS objects, especially at longer forecast times (Fig. 9). The larger size of WoFS objects is partly attributable to them including more of the QLCS stratiform region, which in turn may be due to them having little to no transition zone that would provide a boundary between the leading convective line and the trailing stratiform region (Figs. 10 and 12). Additionally, centerline property differences found WoFS QLCSs are more tortuous and rotate cyclonically in orientation with increasing forecast time (Fig. 11). These results are consistent with WoFS overforecasting the intensity of the QLCS's northern bookend

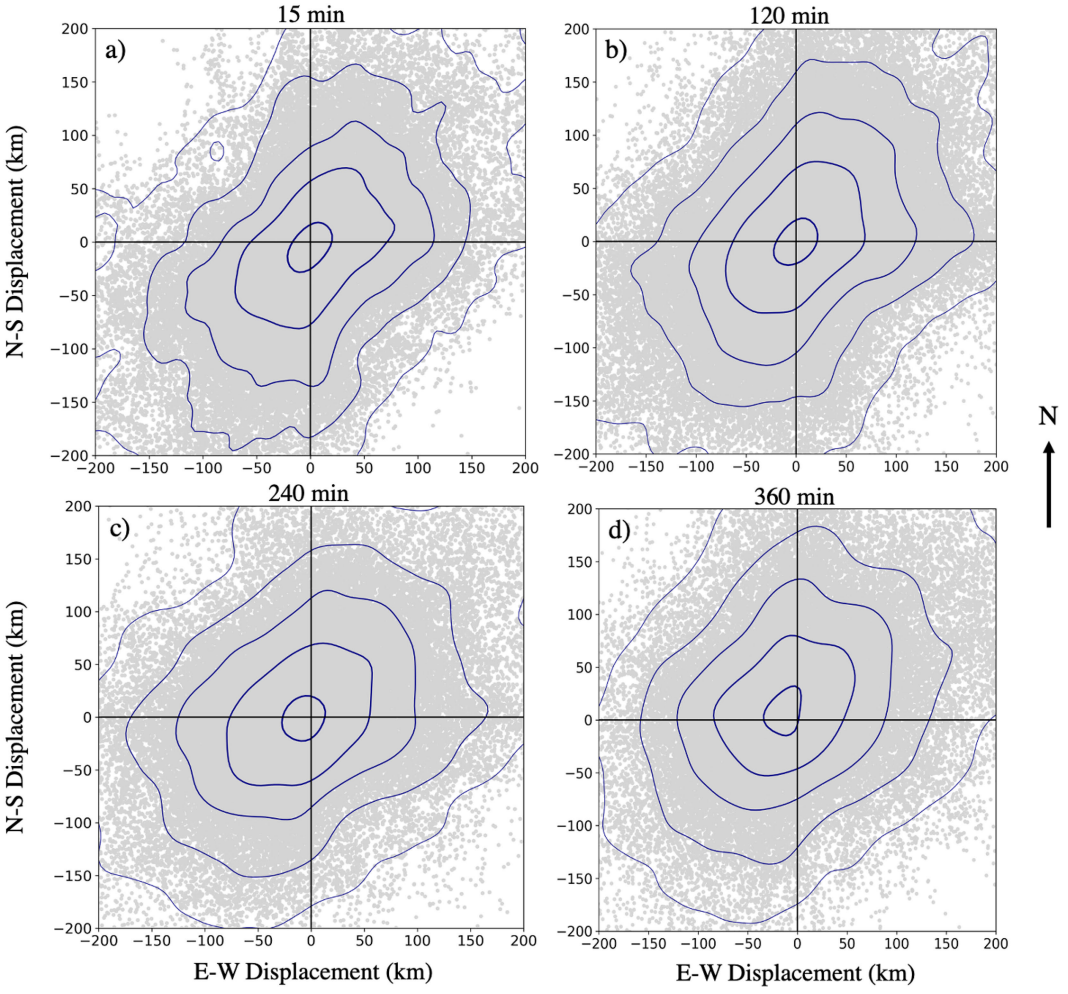

FIG. 13. North–south (N–S) and east–west (E–W) displacement between every 5th percentile along the MRMS and WoFS centerlines for the (a) 15-, (b) 120-, (c) 240-, and (d) 360-min forecasts. Blue contours represent the KDE contours for the 90th, 95th, 97.5th, 99th, and 99.9th percentiles. Displacement is calculated by subtracting the MRMS position from WoFS.

vortex and RIJ, which is evident in a representative case study from 17 July 2020 (Figs. 12 and 15).

This study focused on the systematic prediction of QLCSs. Yet to be examined is WoFS's ensemble skill in providing uncertainty information for the prediction of storm characteristics. Additionally, the veracity of the hazards associated with these systems should also be pursued. The overprediction of reflectivity and the northern bookend vortex may affect the precipitation and wind forecasts. Moreover, the change in the storm's orientation with respect to the shear field could have implications for mesovortexgenesis and tornado potential (Rotunno et al. 1988; Schaumann and Przybylinksi 2012; Gibbs 2021; Galarneau et al. 2023). Therefore, the quality of the WoFS's prediction of extreme rainfall, straight-line wind gusts, and mesovortex production should be explored. Subsequently, future work will need to be done to understand the reason for WoFS's overprediction of the intensity of the northern bookend vortex and the RIJ, as this misrepresentation may also affect the quality of the forecasted hazards. One method for diagnosing this overprediction is by analyzing potential vorticity anomalies in forecasted and observed QLCSs (Galarneau et al. 2023). Other work may include analyzing and comparing the surrounding environments around forecasted and observed QLCSs to understand how differences may impact the system's physical structure and internal dynamics. This study was unable to examine cold pool structure and depth given our dataset, but this would be worth pursuing in the future. This examination should also include in-depth case studies using radar and specialized observations from field experiments (e.g., TORUS, Houston et al. 2022; PERiLS, Kosiba et al. 2022) that may also aid in determining the overprediction of the bookend vortex and RIJ. Finally, one could also examine seasonal and geographic variations in QLCS environments by collecting events from different seasons and geographic areas to understand how those conditions influence the strength, structure, and propagation of QLCSs.

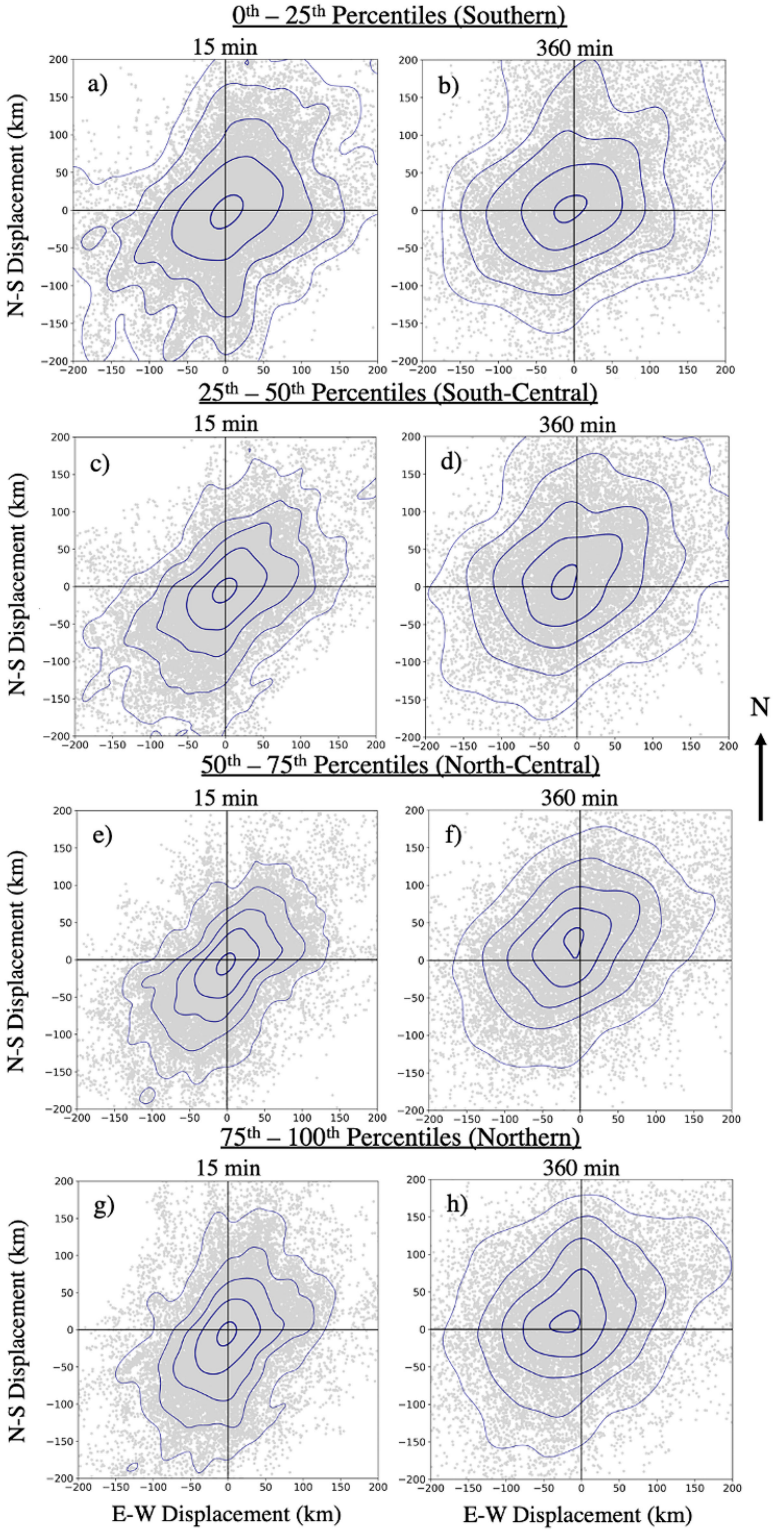

FIG. 14. As in Fig. 13, but the QLCSs are divided into four portions: (a),(b) the southern end (0th–25th percentiles); (c),(d) south-central section (25th–50th percentiles); (e),(f) north-central section (50th–75th percentiles); and (g),(h) northern end (75th–100th percentiles) for the (left) 15- and (right) 360-min forecasts.

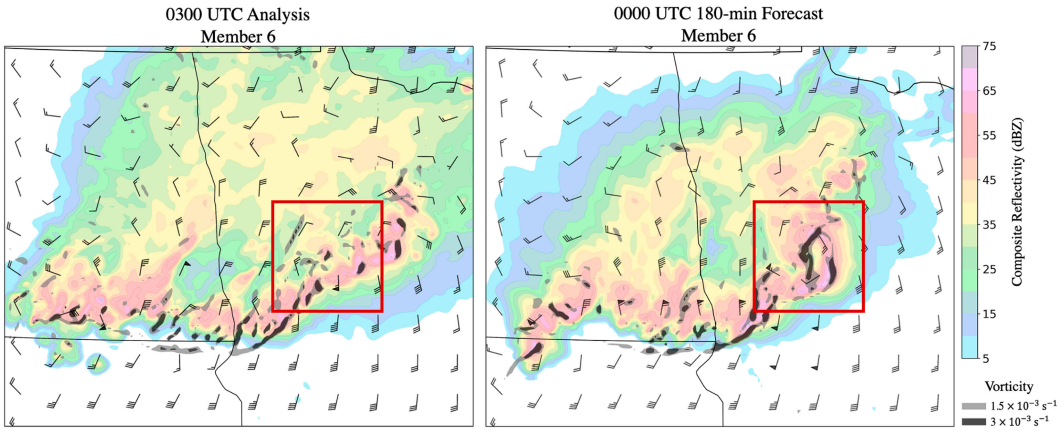

FIG. 15. Composite reflectivity, 1-km wind barbs (kt; 1 kt ≈ 0.51 m s^{-1}), and 1-km vorticity for the 17 Jul 2020 QLCS event. (left) The 0300 UTC analysis for ensemble member 6 is compared to (right) the 180-min forecast for member 6 initialized at 0000 UTC. The WoFS forecast shows an intense northern bookend vortex and rear inflow jet located within the red box that is not present in the analysis.

Acknowledgments. Funding for this research was provided by NOAA/Office of Oceanic and Atmospheric Research under NOAA–University of Oklahoma Cooperative Agreement NA11OAR4320072, U.S. Department of Commerce, and the Warn-on-Forecast project. Additional support was provided by the Joint Technology Transfer Initiative Program within the NOAA/OAR Weather Program Office under Award NA22OAR4590171 (MLF). The authors thank Harold Brooks for his feedback on how to improve the centerline analysis, as well as Tom Galarneau and the three anonymous reviewers for providing feedback to improve the manuscript. All analyses and visualizations were produced using the freely provided Anaconda Python distribution (v.2.7 and v.3.6) and the SciPy, Matplotlib, netcdf4, numpy, scikit-image, and scikit-learn libraries. Finally, a special thank you to Joachim Ungar for his creation of the centerline code, which was then adapted and expanded upon for use in this project.

Data availability statement. All datasets and software used for this study are stored on the NSSL high-performance computing server and the data are available upon request. Processing software for tracking and matching is hosted in https://github.com/WarnOnForecast/MontePython. Centerlines were generated using code created by Joachim Ungar and is available on their Github page (i.e., https://github.com/ungarj/label_centerlines).

APPENDIX

Tuning the Storm Identification and Classification Algorithm

An early version of the storm mode identification and classification algorithm from Potvin et al. (2022) underwent two tuning sessions (hereafter referred to as T1 and T2) to increase the algorithm's efficiency in identifying intense, linear QLCS convective line objects in WoFS and MRMS. T1 is broken into three stages (T1a, T1b, and T1c) and uses subjective methods to verify how well the algorithm is identifying QLCS objects. The purpose of T1 was to provide a starting point for finding the best threshold values that correctly identify and classify WoFS and MRMS objects as QLCSs, which would then be applied during T2 for the entire 2017–20 dataset (see Table 1 for the full list of case days). T2 uses object matching and verification techniques to provide an objective verification for the algorithm.

a. T1: Subjective tuning

T1 consisted of subjectively analyzing and verifying a subset of nine QLCS case days. Three cases from each year from 2017 to 2019 were randomly chosen (see the cases with an asterisk in Table 1) to examine the impacts of changing specific algorithm parameters on the identification of QLCS objects. At this time, the 2020 cases were unavailable, but because the WoFS configuration was the same as that used in 2019 those tuned parameters should still hold for the 2020 days. During the initial tuning phase, only the WoFS analyses were considered (forecasts are considered later).

T1 began by identifying QLCS objects in WoFS and MRMS using the algorithm's default thresholds (Table A1). The resulting objects were then examined to gauge how well the algorithm was identifying all possible QLCS objects in either the WoFS or MRMS output. The objects were categorized as hits, misses, or false alarms relative to subjective identification by the first author so verification statistics could be calculated to quantify the skill of the algorithm's QLCS identification. In this case, a hit was defined as an object identification that matched the subjective analysis, a miss was an object identified subjectively but not by the algorithm, and a false alarm was an object identified by the algorithm but not subjectively. Misses typically occurred when the object's reflectivity was just outside the set threshold, was too small to meet the minimum area threshold, or had a lower eccentricity. False alarms tended to be QLCS stratiform regions or a

TABLE A1. Algorithm parameter thresholds for all tuning sessions.

Parameter	Default	T1a	T1b	T1c	T2 final
WoFS minimum reflectivity	43	39	40	40	45
MRMS minimum reflectivity	36.73	38	40	40	40
WoFS maximum reflectivity	44.93	41	45	45	50
MRMS maximum reflectivity	40.73	40	45	45	45
Minimum area	135 km^2	135 km^2	135 km^2	225 km^2	54 km^2
Merge distance	9 km	12 km	12 km	12 km	12 km
Length (L) and eccentricity (E)	$L > 150$ km, $E > 0.97$	$L > 126$ km, $E > 0.74$	$L > 100$ km, $E > 0.85$; $L > 150$ km, $E > 0.74$	$L > 100$ km, $E > 0.85$; $L > 150$ km, $E > 0.74$	$L > 100$ km, $E > 0.85$; $L > 150$ km, $E > 0.74$

cluster of cells that were merged together to have the appearance of a QLCS, but were of a different storm mode type (e.g., a line of discrete supercells). The verification statistics for WoFS and MRMS using the default algorithm parameters are shown in Table A2.

The default thresholds still missed many QLCS objects in WoFS and MRMS, evident from the high number of misses in Table A2. To mitigate some of the problems that occurred while using the default parameters, T1 began by taking a close look at all the identified and unidentified QLCS objects to assess what parameters needed to be changed. This first stage of T1 is referred to as T1a. First, the default reflectivity threshold for MRMS was too low causing the algorithm to classify parts of the stratiform region of the QLCS. That threshold was then increased, as the main goal was to capture the QLCS's convective line. The reflectivity thresholds used for WoFS were too strict, which inflated the number of misses, so those were lowered. Last, the eccentricity and length thresholds were too high and had to be lowered as they were causing the algorithm to miss pieces of the QLCS's convective line. The final tuned parameters for T1a are given in Table A1 and their verification scores in Table A2.

During T1b, or the second stage of T1, the parameters were tuned separately to find the best value that maximized the POD and minimized the FAR scores. The main change between T1a and T1b was in the object length and eccentricity thresholds (Table A1). Originally, the default algorithm contained two threshold statements for length and eccentricity. However, only one statement was used in the default and T1a because the other statement was erroneously omitted (Table A1). This mistake was found during T1b resulting in the second statement being added back in (Table A1). Both length and eccentricity statements were tuned to maximize the amount of QLCS objects being identified but also minimizing the amount of more circular objects (e.g., a supercell) that may be erroneously identified (Table A1).

The final stage of T1 (referred to as T1c) addressed an issue where identified QLCS objects included too many smaller reflectivity objects as part of the larger object. To mitigate this, different minimum object area thresholds were tested and the resulting minimum area threshold was changed from 15 to 25 grid cells (Table A1). Comparing the verification statistics between T1b and T1c shows a slight decrease in POD and CSI in both WoFS and MRMS, and a small increase in FAR for WoFS. Since the values did not change drastically with the switch in minimum area thresholds, being able to accurately identify the area/extent of QLCS objects without the smaller cells attached to them took precedence. Therefore, the T1c thresholds were used as the final parameter thresholds.

b. T2: Objective tuning

The majority of the algorithm was tuned during T1, where the parameter thresholds were finalized subjectively using a subset of QLCS cases to expedite the tuning process by not having to run the algorithm over all cases. Instead, the full dataset of cases were only run through the identification algorithm during the objective tuning session (T2) using the finalized thresholds from T1c (Table A1). Once the objects were identified in WoFS and MRMS, they were then matched using a total interest score method (Skinner et al. 2018) so that the verification statistics could be calculated using an objective framework. However, once all the cases were run through the algorithm

TABLE A2. Verification statistics for the subjective tuning sessions.

Name	No. of objects	Hits	Misses	False alarms	POD (%)	FAR (%)	CSI (%)
Default WoFS	62	62	124	0	33.3	0	33.3
WoFS T1a	294	232	5	62	97.9	21.1	77.6
WoFS T1b	371	364	7	7	98.1	1.9	96.3
WoFS T1c	267	256	14	11	95.9	3.3	91.1
Default MRMS	130	111	83	19	57.2	14.6	52.1
MRMS T1a	337	247	19	90	92.9	26.7	69.4
MRMS T1b	357	345	6	12	98.3	3.4	95.0
MRMS T1c	249	245	11	4	95.7	1.6	94.2

TABLE A3. Sensitivity experiments.

Expt name	Description
Control	Thresholds from T1c (see Table A1)
Control_Smooth	As in Control, but with Gaussian smoothing with a radius of 4 grid cells
Smooth_area6	As in Control_Smooth, but with the min area threshold lowered from 25 to 6 grid cells

some new issues arose. One such problem was a time discontinuity where the algorithm would correctly identify a QLCS object at one time step but not at the subsequent time step, even though the QLCS object was still subjectively present in the forecast.

In an attempt to fix the time discontinuity problem, it was decided to add Gaussian smoothing to the QLCS objects. The smoothing was found to improve the identification of QLCS objects by making the objects more continuous and mitigating gaps between them. Additionally, the time discontinuities more often occurred with smaller QLCS objects. If the area threshold was lowered for unsmoothed objects, the issue from T1b where too many smaller objects were erroneously being included as part of the QLCS's convective line would return. However, by adding smoothing and rearranging the steps in the algorithm's quality control (QC) process, the area threshold was able to be lowered without adding in the smaller objects. Essentially, the QC was rearranged so that it would check the object area before merging all nearby objects together. Thus, small objects that did not meet the minimum area threshold were thrown out and excluded from the merging process.

Subjective and objective sensitivity tests were conducted to find the best radius of influence for the Gaussian smoothing and a new minimum area threshold. Using only the 9 May 2018 case, various smoothing radii and minimum area thresholds were subjectively analyzed. Smoothing radii of 3, 4, and 5 grid cells, and a minimum area threshold of 4, 5, 6, 7, and 8 grid cells were tested. The goal was to find the optimal radius and area that would result in objects that were more continuous, reduced any time discontinuities, and identified the majority of the object without the inclusion of smaller spurious cells. It was discovered that using a smoothing radius of 4 grid cells and a minimum area of 6 grid cells (54 km^2) accomplished this goal.

The optimal values for Gaussian smoothing and minimum area were then used in the objective sensitivity tests to understand how adding/changing each component would affect the verification statistics. These sensitivity experiments are outlined in Table A3. The experiments were conducted over a subset of four QLCS case days, which included the best case (17 July 2020), worst case (9 May 2018), and two randomly chosen days from the remaining years (18 May 2017 and 1 May 2019). Each hourly forecast from 2000 to 0200 UTC at the 0- (analysis), 30-, and 60-min lead times were examined. The resulting POD, FAR, and CSI values for the sensitivity experiments were evaluated (not shown). In all cases, the Smooth_Area6 experiment performed the best, maximizing the POD and CSI values but also minimizing the FAR. Therefore, the neighborhood was applied to all cases and the minimum area threshold was changed to 6 grid cells (54 km^2).

The final change made during the T2 stage was the decision to not include the WoFS analyses in the verification process. Fortunately, only the WoFS reflectivity thresholds would need to be changed when no longer considering the analyses. Since WoFS analyses are generated through assimilation of MRMS reflectivity observations, analysis reflectivity values will more closely resemble observations. Following the analysis, higher composite reflectivity values will be reestablished by WoFS's microphysics scheme (the NSSL two-moment scheme). Therefore, the baseline and maximum WoFS reflectivity thresholds were changed to 45 and 50 dBZ, respectively, to account for the increase in reflectivity values from the analysis to any forecast time. The final tuned parameters can be seen for T2 in Table A1 and their resulting verification statistics are evaluated throughout section 3a.

REFERENCES

Anderson, J. L., 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.

Ashley, W. S., A. M. Haberlie, and J. Strohm, 2019: A climatology of quasi-linear convective systems and their hazards in the United States. *Wea. Forecasting*, **34**, 1605–1631, https://doi.org/10.1175/WAF-D-19-0014.1.

Atkins, N. T., and M. S. Laurent, 2009: Bow echo mesovortices. Part II: Their genesis. *Mon. Wea. Rev.*, **137**, 1514–1532, https://doi.org/10.1175/2008MWR2650.1.

——, J. M. Arnott, R. W. Przybylinski, R. A. Wolf, and B. D. Ketcham, 2004: Vortex structure and evolution within bow echoes. Part I: Single-Doppler and damage analysis of the 29 June 1998 derecho. *Mon. Wea. Rev.*, **132**, 2224–2242, https://doi.org/10.1175/1520-0493(2004)132<2224:VSAEWB>2.0.CO;2.

Aurenhammer, F., 1991: Voronoi diagrams—A survey of a fundamental geometric data structure. *ACM Comput. Surv.*, **23**, 345–405, https://doi.org/10.1145/116873.116880.

Biggerstaff, M. I., and R. A. Houze Jr., 1993: Kinematics and microphysics of the transition zone of the 10–11 June 1985 squall line. *J. Atmos. Sci.*, **50**, 3091–3110, https://doi.org/10.1175/15200469(1993)050<3091:KAMOTT>2.0.CO;2.

Britt, K. C., P. S. Skinner, P. L. Heinselman, and K. H. Knopfmeier, 2020: Effects of horizontal grid spacing and inflow environment on forecasts of cyclic mesocyclogenesis in NSSL's Warn-on-Forecast System (WoFS). *Wea. Forecasting*, **35**, 2423–2444, https://doi.org/10.1175/WAF-D-20-0094.1.

Brooks, H. E., 2004: Tornado-warning performance in the past and future: A perspective from signal detection theory. *Bull.*

Amer. Meteor. Soc., **85**, 837–844, https://doi.org/10.1175/BAMS-85-6-837.

Brotzge, J. A., S. E. Nelson, R. L. Thompson, and B. T. Smith, 2013: Tornado probability of detection and lead time as a function of convective mode and environmental parameters. *Wea. Forecasting*, **28**, 1261–1276, https://doi.org/10.1175/WAF-D-12-00119.1.

Ciurică, S., and Coauthors, 2019: Arterial tortuosity. *Hypertension*, **73**, 951–960, https://doi.org/10.1161/HYPERTENSIONAHA.118.11647.

Clark, A. J., and Coauthors, 2020: A real-time simulated forecasting experiment for advancing the prediction of hazardous convective weather. *Bull. Amer. Meteor. Soc.*, **101**, E2022–E2024, https://doi.org/10.1175/BAMS-D-19-0298.1.

——, and Coauthors, 2021: A real-time, virtual spring forecasting experiment to advance severe weather prediction. *Bull. Amer. Meteor. Soc.*, **102**, E814–E816, https://doi.org/10.1175/BAMS-D-20-0268.1.

Davis, C., B. Brown, and R. 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.

Dowell, D. C., 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.

——, and Coauthors, 2022: The High-Resolution Rapid Refresh (HRRR): An hourly updating convection-allowing forecast model. Part I: Motivation and system description. *Wea. Forecasting*, **37**, 1371–1395, https://doi.org/10.1175/WAF-D-21-0151.1.

Flora, M. L., P. S. Skinner, C. K. Potvin, A. E. Reinhart, T. A. Jones, N. Yussouf, and K. H. Knopfmeier, 2019: Object-based verification of short-term, storm-scale probabilistic mesocyclone guidance from an experimental Warn-on-Forecast System. *Wea. Forecasting*, **34**, 1721–1739, https://doi.org/10.1175/WAF-D-19-0094.1.

Fu, J., H. R. Thomas, and C. Li, 2021: Tortuosity of porous media: Image analysis and physical simulation. *Earth-Sci. Rev.*, **212**, 103439, https://doi.org/10.1016/j.earscirev.2020.103439.

Galarneau, T., P. S. Skinner, M. L. Flora, and K. C. Britt, 2023: Potential vorticity diagnosis of mesoscale convective systems in the Warn-on-Forecast System. *28th Conf. on Numerical Weather Prediction*, Madison, WI, Amer. Meteor. Soc., 5.5, https://ams.confex.com/ams/WAFNWPMS/meetingapp.cgi/Paper/425560.

Gibbs, J. G., 2021: Evaluating precursor signals for QLCS tornado and higher impact straight-line wind events. *J. Oper. Meteor.*, **9**, 62–75, https://doi.org/10.15191/nwajom.2021.0905.

Gillies, S., and Coauthors, 2023: Shapely: Manipulation and analysis of geometric objects in the Cartesian plane. Zenodo, accessed 18 January 2021, https://doi.org/10.5281/zenodo.5597138.

Grim, J. A., R. M. Rauber, G. M. McFarquhar, B. F. Jewett, and D. P. Jorgensen, 2009: Development and forcing of the rear inflow jet in a rapidly developing and decaying squall line during BAMEX. *Mon. Wea. Rev.*, **137**, 1206–1229, https://doi.org/10.1175/2008MWR2503.1.

Guerra, J. E., P. S. Skinner, A. Clark, M. Flora, B. Matilla, K. Knopfmeier, and A. E. Reinhart, 2022: Quantification of NSSL Warn-on-Forecast system accuracy by storm age using object-based verification. *Wea. Forecasting*, **37**, 1973–1983, https://doi.org/10.1175/WAF-D-22-0043.1.

Houston, A. L., C. C. Weiss, E. Rasmussen, M. C. Coniglio, C. L. Ziegler, B. Argrow, and E. W. Frew, 2022: Targeted observation by radars and UAS of supercells (TORUS): Summary of the 2019 and 2022 field campaigns. *30th Conf. on Severe Local Storms*, Santa Fe, NM, Amer. Meteor. Soc., 7.1B, https://ams.confex.com/ams/30SLS/meetingapp.cgi/Paper/407665.

Houtekamer, P. L., and F. Zhang, 2016: Review of the ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.*, **144**, 4489–4532, https://doi.org/10.1175/MWR-D-15-0440.1.

Hu, M., H. Shao, D. Stark, K. Newman, C. Zhou, and X. Zhang, 2016: Gridpoint Statistical Interpolation GSI user's guide version 3.5. Developmental Testbed Center Rep., 148 pp., https://dtcenter.org/com-GSI/users/docs/users_guide/GSIUserGuide_v3.5.pdf.

Jones, T. A., D. Stensrud, L. Wicker, P. Minnis, and R. Palikonda, 2015: Simultaneous radar and satellite data storm-scale assimilation using an ensemble Kalman filter approach for 24 May 2011. *Mon. Wea. Rev.*, **143**, 165–194, https://doi.org/10.1175/MWR-D-14-00180.1.

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

——, X. Wang, P. 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.

——, P. Skinner, N. Yussouf, K. Knopfmeier, A. Reinhart, and D. Dowell, 2019: Forecasting high-impact weather in landfalling tropical cyclones using a Warn-on-Forecast System. *Bull. Amer. Meteor. Soc.*, **100**, 1405–1417, https://doi.org/10.1175/BAMS-D-18-0203.1.

——, and Coauthors, 2020: Assimilation of *GOES-16* radiances and retrievals into the Warn-on-Forecast System. *Mon. Wea. Rev.*, **148**, 1829–1859, https://doi.org/10.1175/MWR-D-19-0379.1.

Kerr, C. A., and F. Alsheimer, 2022: Storm-scale predictability and analysis of the 13 April 2020 central Savannah River area tornado outbreak. *Wea. Forecasting*, **37**, 901–913, https://doi.org/10.1175/WAF-D-21-0185.1.

Kleist, D. T., D. F. Parrish, J. C. Derber, R. Treadon, W.-S. Wu, and S. Lord, 2009: Introduction of the GSI into the NCEP Global Data Assimilation System. *Wea. Forecasting*, **24**, 1691–1705, https://doi.org/10.1175/2009WAF2222201.1.

Kosiba, K. A., J. Wurman, R. J. Trapp, M. D. Parker, and S. W. Nesbitt, 2022: The PERiLS (propagation, evolution and rotation in linear storms) southeastern tornado study. *30th Conf. on Severe Local Storms*, Santa Fe, NM, Amer. Meteor. Soc., 153, https://ams.confex.com/ams/30SLS/meetingapp.cgi/Paper/407690.

Laser, J. J., M. C. Coniglio, P. S. Skinner, and E. N. Smith, 2022: Doppler lidar and mobile radiosonde observation-based evaluation of Warn-on-Forecast System predicted near-supercell environments during TORUS 2019. *Wea. Forecasting*, **37**, 1783–1804, https://doi.org/10.1175/WAF-D-21-0190.1.

Lawson, J. R., J. S. Kain, N. Yussouf, D. C. Dowell, D. M. Wheatley, K. H. Knopfmeier, and T. A. Jones, 2018: Advancing from convection-allowing NWP to Warn-on-Forecast: Evidence of progress. *Wea. Forecasting*, **33**, 599–607, https://doi.org/10.1175/WAF-D-17-0145.1.

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.

Meng, Z., F. Zhang, P. Markowski, D. Wu, and K. Zhao, 2012: A modeling study on the development of a bowing structure and associated rear inflow within a squall line over south China. *J. Atmos. Sci.*, **69**, 1182–1207, https://doi.org/10.1175/JAS-D-11-0121.1.

Miller, W. J. S., and Coauthors, 2022: Exploring the usefulness of downscaling free forecasts from the Warn-on-Forecast System. *Wea. Forecasting*, **37**, 181–203, https://doi.org/10.1175/WAF-D-21-0079.1.

Potvin, C. K., and Coauthors, 2020: Assessing systematic impacts of PBL schemes on storm evolution in the NOAA Warn-on-Forecast System. *Mon. Wea. Rev.*, **148**, 2567–2590, https://doi.org/10.1175/MWR-D-19-0389.1.

——, and Coauthors, 2022: An iterative storm segmentation and classification algorithm for convection-allowing models and gridded radar analyses. *J. Atmos. Oceanic Technol.*, **39**, 999–1013, https://doi.org/10.1175/JTECH-D-21-0141.1.

Rotunno, R., J. B. Klemp, and M. L. Weisman, 1988: A theory for strong, long-lived squall lines. *J. Atmos. Sci.*, **45**, 463–485, https://doi.org/10.1175/1520-0469(1988)045<0463:ATFSLL>2.0.CO;2.

Schaumann, J. S., and R. W. Przybylinksi, 2012: Operational application of 0–3 km bulk shear vectors in assessing quasi linear convective system mesovortex and tornado potential. *26th Conf on Severe Local Storms*, Nashville, TN, Amer. Meteor. Soc., 142, https://ams.confex.com/ams/26SLS/webprogram/Paper212008.html.

Skinner, P. S., 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.

Smirnova, T. G., J. M. Brown, S. G. Benjamin, and J. S. Kenyon, 2016: Modifications to the Rapid Update Cycle Land Surface Model (RUC LSM) available in the Weather Research and Forecasting (WRF) Model. *Mon. Wea. Rev.*, **144**, 1851–1865, https://doi.org/10.1175/MWR-D-15-0198.1.

Smith, B. T., R. L. Thompson, J. S. Grams, C. Broyles, and H. E. Brooks, 2012: Convective modes for significant severe thunderstorms in the contiguous United States. Part I: Storm classification and climatology. *Wea. Forecasting*, **27**, 1114–1135, https://doi.org/10.1175/WAF-D-11-00115.1.

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.

Smull, B. F., and R. A. Houze Jr., 1987: Rear inflow in squall lines with trailing stratiform precipitation. *Mon. Wea. Rev.*, **115**, 2869–2889, https://doi.org/10.1175/1520-0493(1987)115<2869:RIISLW>2.0.CO;2.

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, R. L., B. T. Smith, J. S. Grams, A. R. Dean, and C. Broyles, 2012: Convective modes for significant severe thunderstorms in the contiguous United States. Part II: Supercell and QLCS tornado environments. *Wea. Forecasting*, **27**, 1136–1154, https://doi.org/10.1175/WAF-D-11-00116.%201.

Trapp, R. J., S. A. Tessendorf, E. S. Godfrey, and H. E. Brooks, 2005: Tornadoes from squall lines and bow echoes. Part I: Climatological distribution. *Wea. Forecasting*, **20**, 23–34, https://doi.org/10.1175/WAF-835.1.

Ungar, J., 2022: Label centerlines. Accessed 18 January 2021, https://github.com/ungarj/label_centerlines.

van der Walt, S., J. L. Schonberger, J. Nunez-Iglesias, F. Boulogne, J. D. Warner, N. Yager, E. Gouillart, and T. Yu, 2014: Scikit-image: Image processing in Python. *PeerJ*, **2**, e453, https://doi.org/10.7717/peerj.453.

Wang, Y., N. Yussouf, C. A. Kerr, D. R. Stratman, and B. C. Matilla, 2022: An experimental 1-km Warn-on-Forecast System for hazardous weather events. *Mon. Wea. Rev.*, **150**, 3081–3102, https://doi.org/10.1175/MWR-D-22-0094.1.

Weisman, M. L., 1992: The role of convectively generated rear-inflow jets in the evolution of long-lived mesoconvective systems. *J. Atmos. Sci.*, **49**, 1826–1847, https://doi.org/10.1175/1520-0469(1992)049<1826:TROCGR>2.0.CO;2.

——, 1993: The genesis of severe, long-lived bow echoes. *J. Atmos. Sci.*, **50**, 645–670, https://doi.org/10.1175/1520-0469(1993)050<0645:TGOSLL>2.0.CO;2.

——, and C. A. Davis, 1998: Mechanisms for the generation of mesoscale vortices within quasi-linear convective systems. *J. Atmos. Sci.*, **55**, 2603–2622, https://doi.org/10.1175/1520-0469(1998)055<2603:MFTGOM>2.0.CO;2.

Weyn, J. A., and D. R. Durran, 2017: The dependence of the predictability of mesoscale convective systems on the horizontal scale and amplitude of initial errors in idealized simulations. *J. Atmos. Sci.*, **74**, 2191–2210, https://doi.org/10.1175/JAS-D-17-0006.1.

——, and ——, 2019: The scale dependence of initial-condition sensitivities in simulations of convective systems over the southeastern United States. *Quart. J. Roy. Meteor. Soc.*, **145**, 57–74, https://doi.org/10.1002/qj.3367.

Wheatley, D. M., and R. J. Trapp, 2008: The effect of mesoscale heterogeneity on the genesis and structure of mesovortices within quasi-linear convective systems. *Mon. Wea. Rev.*, **136**, 4220–4241, https://doi.org/10.1175/2008MWR2294.1.

——, 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.

Yussouf, N., and K. H. Knopfmeier, 2019: Application of the warn-on-forecast system for flash-flood-producing heavy convective rainfall events. *Quart. J. Roy. Meteor. Soc.*, **145**, 2385–2403, https://doi.org/10.1002/qj.3568.

——, E. R. Mansell, L. J. Wicker, D. M. Wheatley, and D. J. Stensrud, 2013: The ensemble Kalman filter analyses and forecasts of the 8 May 2003 Oklahoma City tornadic supercell storm using single- and double-moment microphysics scheme. *Mon. Wea. Rev.*, **141**, 3388–3412, https://doi.org/10.1175/MWR-D-12-00237.1.

——, D. C. Dowell, L. J. Wicker, K. H. Knopfmeier, and D. M. Wheatley, 2015: Storm-scale data assimilation and ensemble forecasts for the 27 April 2011 severe weather outbreak in Alabama. *Mon. Wea. Rev.*, **143**, 3044–3066, https://doi.org/10.1175/MWR-D-14-00268.1.

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

——, T. A. Jones, and P. S. Skinner, 2020a: Probabilistic high-impact rainfall forecasts from landfalling tropical cyclones using warn-on-forecast system. *Quart. J. Roy. Meteor. Soc.*, **146**, 2050–2065, https://doi.org/10.1002/qj.3779.

——, K. A. Wilson, S. M. Martinaitis, H. Vergara, P. L. Heinselman, and J. J. Gourley, 2020b: The coupling of NSSL Warn-on-Forecast and FLASH systems for probabilistic flash flood prediction. *J. Hydrometeor.*, **21**, 123–141, https://doi.org/10.1175/JHM-D-19-0131.1.