EVALUATION OF THUNDERSTORM ROTATION IN 1- AND 3-KM VERSIONS OF THE WARN-ON-FORECAST SYSTEM

Lillian E. Frey[1], Derek R. Stratman[2], and Christopher A. Kerr[3]

[1]National Weather Center Research Experiences for Undergraduates Program
Norman, Oklahoma
[2]Cooperative Institute of Severe and High-Impact Weather Research and Operations
Norman, Oklahoma
[3]NOAA/OAR National Severe Storms Laboratory
Norman, Oklahoma
[4]Wofford College
Spartanburg, South Carolina

ABSTRACT

NSSL's experimental Warn-on-Forecast System (WoFS) provides probabilistic guidance to NWS forecasters of individual thunderstorms and their associated hazards. WoFS uses ensemble data assimilation that consists of 36 analysis members and 18 forecast members. The rapid data assimilation produces a daily, on-demand forecast every 30 minutes with a regional domain of 900 x 900 km with a 3-km horizontal grid spacing (WoFS-3km). Recently, a 1-km horizontal grid spacing version of WoFS (WoFS-1km) has been under development at NSSL. WoFS-1km has a domain of 402 x 402 km and is nested within the WoFS-3km domain. Recent work has shown that WoFS-1km predicts smaller thunderstorms better than WoFS-3km. This study extends that work by focusing on the ability of WoFS-1km to predict rotation in thunderstorms compared to WoFS-3km. To assess this ability, 23 cases from 2022 and 2023 with various storm modes and environments are verified using an object-based method. Using this method, over 200,000 composite reflectivity (CREF) object hits are determined to exist in both WoFS-1km and WoFS-3km. For each forecast and observed CREF object, mid-level updraft helicity and MRMS rotation objects, respectively, are identified. Rotation object centroid displacement errors and contingency table statistics are computed. Results show that WoFS-1km predicts rotation significantly better than WoFS-3km for reflectivity object areas less than 800 km^2 and minor axes less than 24 km. Also, WoFS-1km has smaller rotation centroid displacement errors than WoFS-3km.

1. INTRODUCTION

The National Severe Storms Laboratory (NSSL) has developed a probabilistic forecasting tool, the Warn-on-Forecast System (WoFS), which is used for short-term forecasting of individual convective storms. The objective of WoFS is to help fill the forecast information gap between watch [*O*(hours)]) and warning [*O*(minutes)] time scales (Heinselman et al. 2024). Most convection-allowing models (CAMs) provide deterministic guidance of convective weather for later-day and next-day time periods, but WoFS is an ensemble data assimilation (DA) and forecast system that provides frequently updated, short-term probabilistic guidance of severe and hazardous weather, such as large hail, tornadoes, damaging winds, flash flooding, and landfalling tropical systems, for a regional domain (Wheatley et al. 2015). WoFS therefore directly supports NOAA's Forecasting a Continuum of Environmental Threats initiative (FACETs; Rothfusz et al. 2018).

Over the past several years, WoFS has been successful at predicting smaller thunderstorms, but the skill of depiction of rotation in WoFS is still being evaluated. While the traditional WoFS has a horizontal model grid spacing of 3 km (hereafter WoFS-3km), a prototype with a 1-km horizontal model grid spacing has been under development the past few years (hereafter WoFS-1km). Recent work has shown that WoFS-1km produces more skillful forecasts of individual convective storms through a reduction of false alarm storms and improved detection of smaller storms (< 400 km^2) while also better depicting storm motions (Wang et al. 2022; Kerr et al. 2023, 2024). Kerr et al. (2023) note limited differences in storm rotation forecast accuracy between WoFS-1km and

[1] *Corresponding author address:* Lillian E. Frey, Wofford College, 120 David L Boren Blvd Room 2500 Norman OK 73072, lillifrey123@gmail.com.

WoFS-3km in a dataset of 2021 cases. However, Miller et al. (2022) find that downscaling forecasts from a 3-km horizontal grid spacing to a 1.5-km horizontal grid spacing is beneficial in detecting mid-level mesocyclones. This result emphasizes the need for additional research with WoFS-1km to further explore the impact of different horizontal grid spacings on the ability of forecast systems to accurately and skillfully predict rotation in thunderstorms.

This project focuses on the continued evaluation of WoFS-1km by using the object-based method described in Kerr et al. (2023) to define rotation objects within observed and forecast thunderstorm objects and to compute contingency table statistics. This rotation information is then used to assess and compare the abilities of WoFS-1km and WoFS-3km to accurately and skillfully predict rotation in thunderstorms. The insights gained from these results will help provide guidance on the development of future versions of WoFS. allow for the decision to see if the increase in performance of WoFS-1km is worth the computational cost.

2. DATA AND METHODS

2.1 WoFS-1km and WoFS-3km

WoFS-1km and WoFS-3km are similar systems that run in parallel with each other. Both systems produce 36 analyses every 15 minutes by assimilating various in situ and remote sensing observations, such as radar reflectivity and radial velocity, satellite cloud total liquid water path, and conventional observations, using an ensemble Kalman filter (Jones et al. 2020). Ensemble forecasts with 18 members are initiated at the top of every hour and are integrated out to 3 hours from 1700 UTC until 0300 UTC with 5-minute output frequency. WoFS-3km has a spatial domain of 900 km x 900 km, while the domain for the WoFS-1km is 402 km x 402 km and is nested within the WoFS-3km domain. Both systems are initiated from initial conditions provided by the High-Resolution Rapid Refresh (HRRR) Data Assimilation System (Dowell et al. 2022). The lateral boundary conditions for WoFS-3km are generated from perturbing HRRR forecasts with perturbations from 18 members of the Global Ensemble Forecast System (Zhou et al. 2022), while WoFS-3km provides the lateral boundary conditions for WoFS-1km. To create additional ensemble member diversity, different combinations of planetary boundary layer (PBL) and shortwave/longwave radiation parameterization schemes are used during the cycling and free forecasts. Both systems use the NSSL two-moment microphysical scheme (Mansell et al. 2010), which has shown to help with the representation of storm-scale microphysical processes in supercells (Skinner et al. 2018). Additional details and differences among WoFS-1km and WoFS-3km, such as the radar observation spacing, can be found in Kerr et al. (2023). As in Kerr et al. (2024), 23 different cases are included in the dataset, which span various geographical regions and dominant storm modes (Table 1).

Table 1. Adapted from table 2 in Kerr et al 2024. There are two geographical locations for 31 March 2023.

2022 Cases	2023 Cases
22 April	26 February
29 April	24 March
2 May	31 March
4 May	31 March
6 May	19 April
11 May	9 May
12 May	11 May
16 May	12 May
19 May	15 June
30 May	10 July
31 May	20 August
	30 August

2.2 Object-Based Verification Method

To assess the abilities of WoFS-1km and WoFS-3km to accurately predict the rotation in thunderstorms, the object-based method used in previous WoFS studies (e.g., Skinner et al. 2018, Guerra et al. 2022, and Kerr et al. 2023) for reflectivity objects is adapted to include observed and forecast rotation fields. For forecast rotation fields, 5-minute maximum 2-5-km updraft helicity (UH; Kain et al. 2008) is used to represent mid-level mesocyclones. UH objects from both WoFS-1km and WoFS-3km are verified using Multi-Radar Multi-Sensor (MRMS) mid-level rotation data (Smith et al. 2016). Forecast and observed rotation objects are identified using intensity thresholds of 100 m^2s^{-2} for WoFS-3km UH, 326 m^2s^{-2} for WoFS-1km UH, and 0.0047 s^{-1} for MRMS rotation. The minimum size threshold is 9 km^2. Unlike past studies, rotation objects are identified relative to their parent storms rather than

as independent objects. This allows forecast and observed storms to be deemed rotating or non-rotating. The forecast and observed parent storms are matched using a total interest score (Davis et al. 2009; Skinner et al. 2018; Guerra et al. 2022). The reflectivity and rotation objects being considered for this project are confined to the areas where the WoFS-1km and WoFS-3km domains overlap. Further parent storm identification and object matching details are described in Kerr et al. (2024).

For this study, only forecast parent storms matched to observed storms in *both* WoFS-1km and WoFS-3km are considered, resulting in 216,742 storm objects across all cases and forecasts. After adding the rotation object information to the matched storms, rotation object centroid displacement errors are computed to evaluate the ability of WoFS-1km and WoFS-3km to accurately predict the location of rotation within thunderstorms. In addition, contingency tables are determined to assess the ability of both forecast systems to predict whether a thunderstorm is rotating or not. For these contingency tables, WoFS rotating storms matched to MRMS rotating storms are classified as hits (A; Fig. 1), WoFS rotating storms matched to MRMS non-rotating storms are false alarms (B; Fig. 1), WoFS non-rotating storms matched to MRMS rotating storms are misses (C; Fig. 1), and WoFS non-rotating storms matched to MRMS non-rotating storms are correct negatives (D; Fig. 1). From these counts, multiple contingency table metrics are computed: probability of detection (POD; Fig. 1), false alarm ratio (FAR; Fig. 1), frequency bias, and critical success index (CSI; Fig. 1).

3. RESULTS

To evaluate thunderstorm storm rotation spatial errors, rotation object centroid displacement errors are computed and displayed as a spatial heatmap with bin counts. WoFS-3km has a larger westward mean mesocyclone displacement error than WoFS-1km (i.e., ~10 km vs ~5 km, respectively). Even though WoFS-1km has a mean westward displacement error, the KDE contours reveal the peak in UH centroid location is centered over the origin. This result indicates the WoFS-1km centroids are generally located closer to the MRMS rotation object centroids than WoFS-3km and suggests WoFS-1km can predict rotation location within thunderstorms more accurately than WoFS-3km. Also, WoFS-1km has a narrower displacement error distribution than WoFS-3km, indicating WoFS-1km has a smaller spatial displacement error variance. (Fig. 2)

As in Kerr et al 2024, contingency table metrics are computed as a function of composite reflectivity object area. The POD for rotation in thunderstorms is significantly higher in WoFS-1km than WoFS-3km with differences up to ~800 km^2 due to more hits and fewer misses by WoFS-1km (not shown; Fig. 3). Another way to assess this difference, WoFS-1km has a higher POD for MRMS reflectivity object areas around 200 km^2 than WoFS-3km does for larger

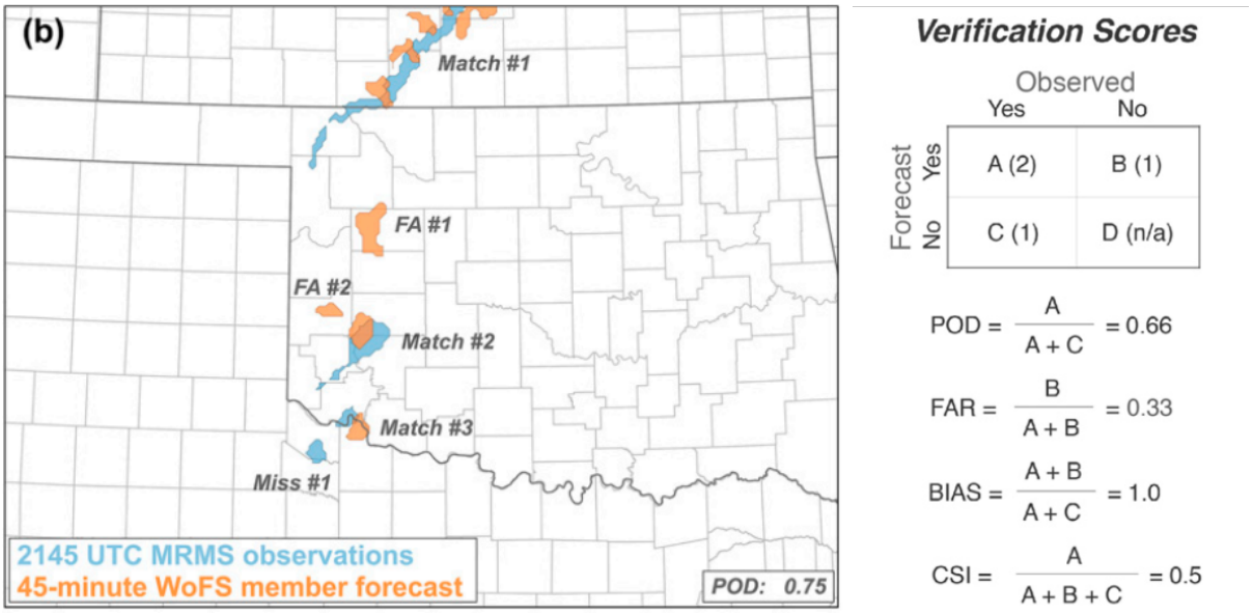

Figure 1. Adapted from figure 2b in Guerra et al. 2022 (left) and from figure 4f in Skinner et al. 2018 (right).

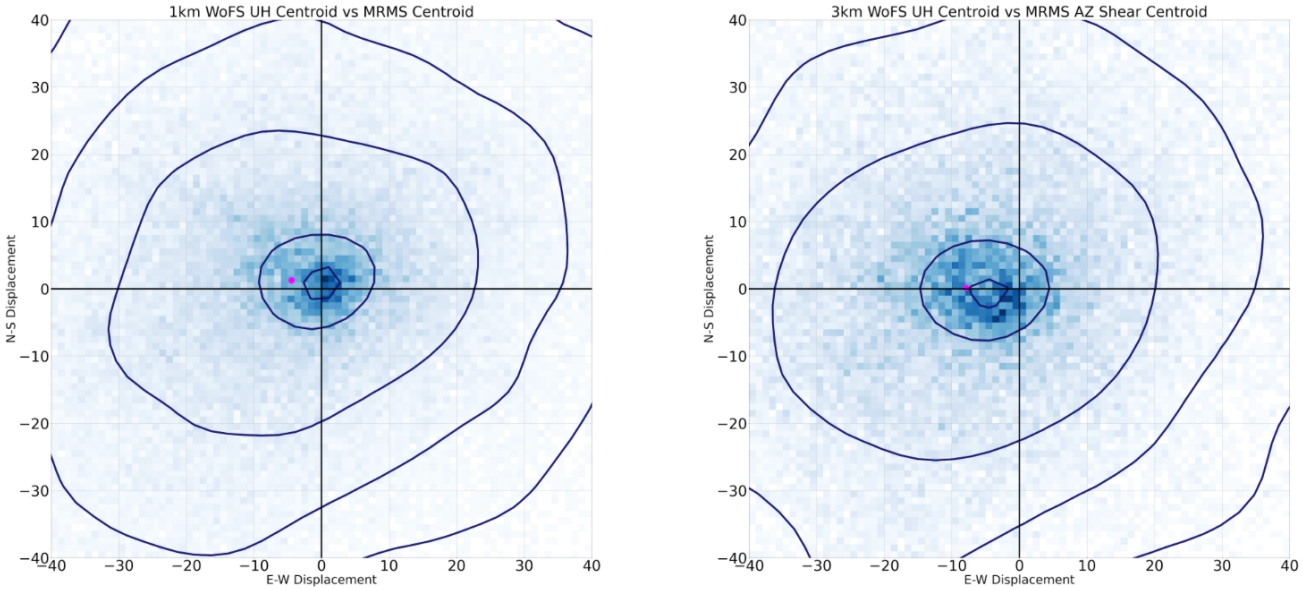

Figure 2. Heatmaps of the bin counts of UH object centroid displacement errors for WoFS-1km (left) and WoFS-3km (right). Bin sizes are 1 km x 1 km. Kernel density estimation (KDE) contours are overlaid and range from 95 percent confidence to 99.99 percent confidence. The mean centroid displacements are represented by a magenta circle.

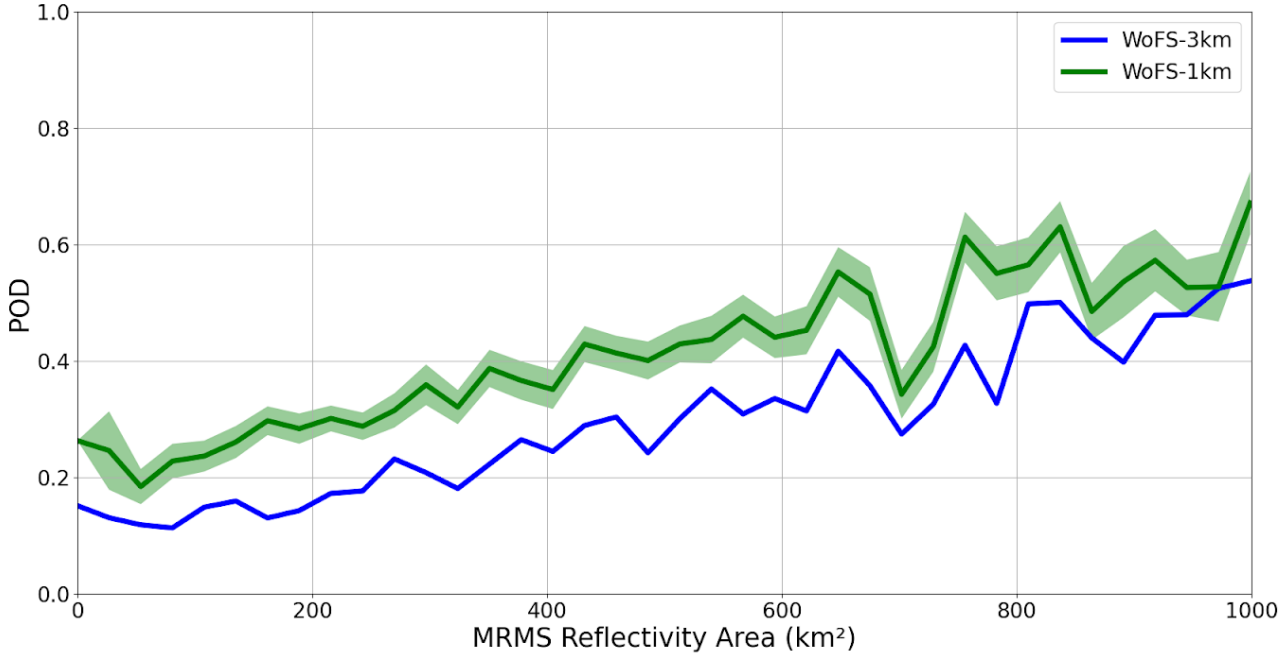

Figure 3. POD of rotation in thunderstorms as a function of MRMS reflectivity object area for WoFS-1km (green line) and WoFS-3km (blue line). Green shading surrounding WoFS-1km's line is the 95th percent confidence interval using a bootstrap resampling method (Hamill et al. 1999) on the differences between WoFS-1km and WoFS-3km. Wherever the green shading does not overlap WoFS-3km's line, the difference is statistically significant. POD values closer to 1 are better.

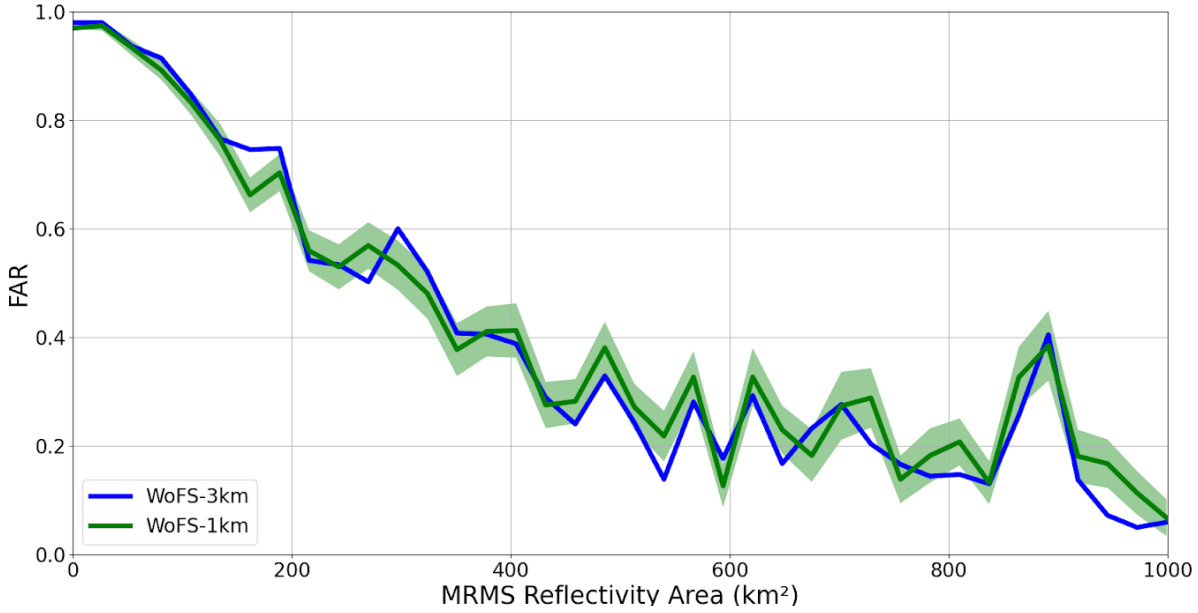

Figure 4. Same as for Figure 3, but for FAR. FAR values closer to 0 are better.

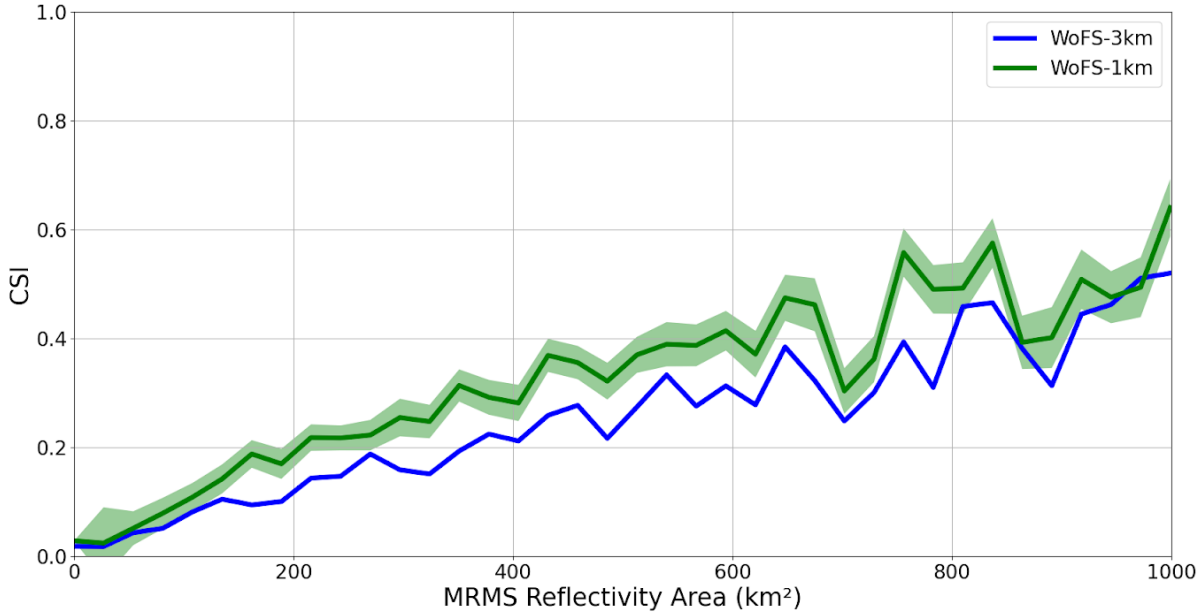

Figure 5. Same as for Figure 3, but for CSI. CSI values closer to 1 are better.

storms with object areas around 400 km^2. While WoFS-1km has more false alarms for rotation in thunderstorms than WoFS-3km (not shown), the increase is offset by more rotation hits since WoFS-1km and WoFS-3km both have statistically similar FAR values for most area bins (Fig. 4). POD and MRMS reflectivity object areas have a positive correlation since POD increases with increasing area. Conversely, FAR and MRMS reflectivity object areas have a negative correlation since FAR decreases with increasing area. This result is consistent with Kerr et al. (2024) whereas object area increases, POD increases and FAR decreases. With CSI, there is less of a disparity between the WoFS-1km and WoFS-3km (Fig. 5) As with POD, WoFS-1km is more skillful with higher CSI for MRMS reflectivity object areas up to ~800 km^2. The increase in hits and

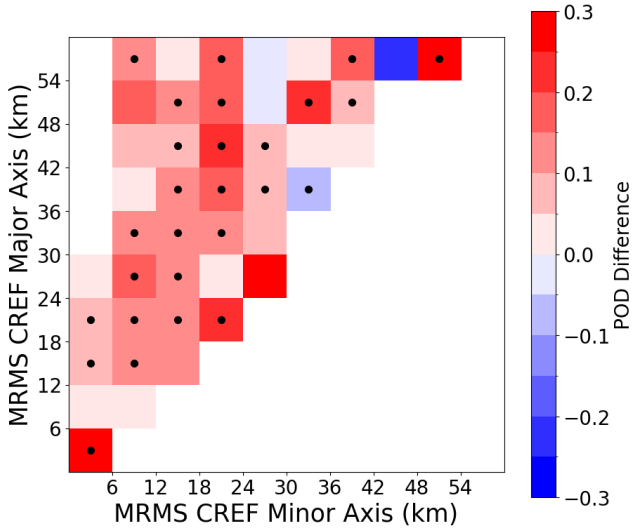

Figure 6. Heatmap of POD differences between WoFS-1km and WoFS-3km as a function of MRMS reflectivity object major and minor axis lengths where the difference equals WoFS-1km POD minus WoFS-3km POD. For each bin, statistical significance at the 95th percent confidence interval is represented by a black dot and was determined by using a bootstrap resampling method (Hamill et al. 1999) on the POD differences. Positive differences favor WoFS-1km.

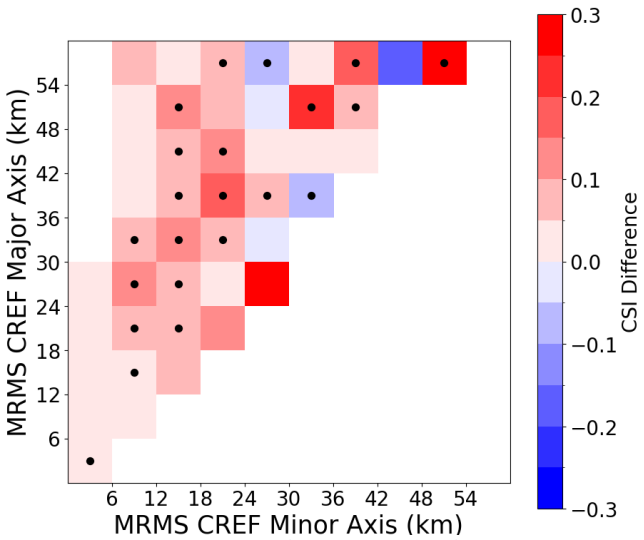

Figure 8. Same as for Figure 6, but for CSI.

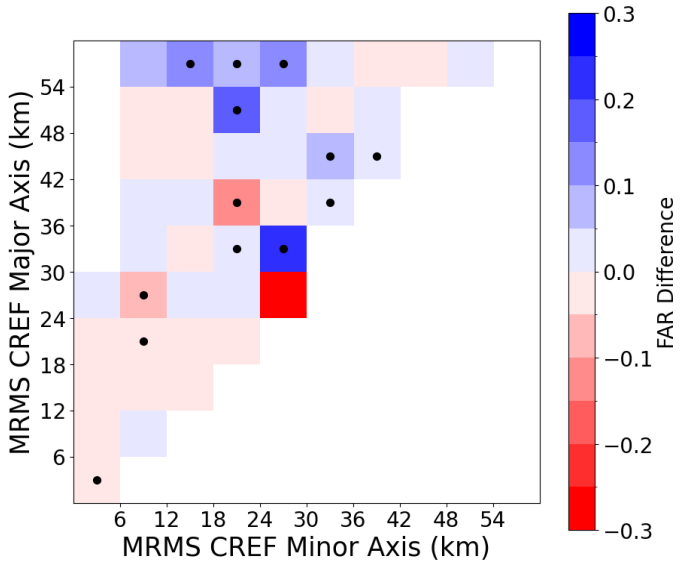

Figure 7. Same as for Figure 6, but for FAR, where negative differences favor WoFS-1km.

false alarms with a decrease in misses in WoFS-1km is driven by a higher frequency bias for all area bins (not shown).

Contingency table metric differences between WoFS-1km and WoFS-3km for rotation in thunderstorms as a function of MRMS reflectivity object major and minor axis lengths yield similar results as Kerr et al. (2024). For POD, WoFS-1km has higher values than WoFS-3km for minor axis lengths up to ~24 km with differences being statistically significant for most bins (Fig. 6). This finding indicates that WoFS-1km outperforms WoFS-3km in detecting rotation in small circular thunderstorms as well as narrower elongated thunderstorms. The FAR differences between WoFS-1km and WoFS-3km as a function of minor and major axes are not statistically

significant for most data bins (Fig. 8). This indicates that FAR difference is not dependent on the shape or size of the reflectivity objects. Similar to the POD differences, the CSI differences between WoFS-1km and WoFS-3km as a function of major and minor axes generally favor WoFS-1km for most data bins, especially those with minor axis lengths less than 24 km^2 (Fig. 8). Most of these CSI differences are also statistically significant.

4. SUMMARY AND CONCLUSIONS

NSSL's experimental WoFS provides probabilistic guidance of individual thunderstorms and their associated hazards to end users, such as NWS meteorologists. While the current system has proven to be successful, work has been underway the past few years at NSSL to develop the next-generation WoFS with 1-km horizontal grid spacing. From this work, Kerr et al. (2024) found WoFS-1km is better at predicting smaller thunderstorms than WoFS-3km. However, that work focused on the thunderstorm itself and not on the associated hazards. In this study, we begin to address that topic by focusing on the prediction of mid-level rotation within individual thunderstorms for both WoFS-1km and WoFS-3km.

In general, WoFS-1km is better able to predict mid-level rotation in thunderstorms than WoFS-3km. More specifically, WoFS-1km has smaller rotation object centroid displacement errors than WoFS-3km. Also, WoFS-1km has significantly higher POD and CSI values than WoFS-3km for rotation within thunderstorms less than ~800 km^2 in size while having similar FAR values. With respect to major and minor axis lengths, WoFS-1km is more skillful at predicting rotation in smaller circular thunderstorms and elongated thunderstorms with minor axes less than ~24 km. The improved performance of WoFS-1km over WoFS-3km further justifies the increased computational costs of running a higher resolution WoFS version for the prediction of individual convective storms.

5. ACKNOWLEDGEMENTS

The corresponding author would like to thank the co-directors of the Research Experiences for Undergraduates at the National Weather Center, Alex Marmo and Daphne LaDue, for their remarkable work. The corresponding author would also like to thank their mentors for their facilitation and guidance with this project, as well as Travis Enzensperger for the direction he provided the author.

This work was prepared by the authors with funding provided by National Science Foundation Grant No. AGS-2050267, and NOAA/Office of Oceanic and Atmospheric Research under NOAA-University of Oklahoma Cooperative Agreement #NA11OAR4320072, U.S. Department of Commerce. The statements, findings, conclusions, and recommendations are those of the author(s) and do not necessarily reflect the views of the National Science Foundation, NOAA, or the U.S. Department of Commerce.

7. REFERENCES

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.

Dowell, D. C., 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.

Gilleland, E., D. Ahijevych, B. G. Brown, B. Casati, and E. E. Ebert, 2009: Intercomparison of Spatial Forecast Verification Methods. *Wea. Forecasting*, **24**, 1416–1430, https://doi.org/10.1175/2009WAF2222269.1.

Gilleland, E., D. A. Ahijevych, B. G. Brown, and E. E. Ebert, 2010: Verifying Forecasts Spatially. *Bull. Amer. Meteor. Soc.*, **91**, 1365–1376, https://doi.org/10.1175/2010BAMS2819.1.

Guerra, J. E., P. S. Skinner, B. Matilla, M. Flora, A. Clark, K. Knopfmeier, and A. 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.

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.

Heinselman, P. L., and Coauthors, 2023: Warn-on-Forecast System: From Vision to Reality. *Wea. Forecasting*, **39**, 75–95, https://doi.org/10.1175/WAF-D-23-0147.1.

Kerr, C. A., B. C. Matilla, Y. Wang, D. R. Stratman, T. A. Jones, and N. Yussouf, 2023: Results from a Pseudo-Real-Time Next-Generation 1-km Warn-

on-Forecast System Prototype. *Wea. Forecasting*, **38**, 307–319, https://doi.org/10.1175/WAF-D-22-0080.1.

Kerr, C. A., P. S. Skinner, D. R. Stratman, B. C. Matilla, Y. Wang, N. Yussouf, 2024: Limitations of short-term thunderstorm forecasts from convection-allowing models with a 3-km horizontal grid spacing. *Wea. Forecasting*, **39**, submitted.

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, 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.

Rothfusz, L. P., R. Schneider, D. Novak, K. Klockow-McClain, A. E. Gerard, C. Karstens, G. J. Stumpf, and T. M. Smith, 2018: FACETs: A Proposed Next-Generation Paradigm for High-Impact Weather Forecasting. *Bull. Amer. Meteor. Soc.*, **99**, 2025–2043, https://doi.org/10.1175/BAMS-D-16-0100.1.

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. J. 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.

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.

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.

VandenBerg, M. A., M. C. Coniglio, and A. J. Clark, 2014: Comparison of Next-Day Convection-Allowing Forecasts of Storm motion on 1- and 4-km Grids. *Wea. Forecasting*, **29**, 878–893, https://doi.org/10.1175/WAF-D-14-00011.1.

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.

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.

Wilson, K. A., P. L. Heinselman, P. S. Skinner, J. J. Choate, and K. E. Klockow-McClain, 2019: Meteorologists' Interpretations of Storm-Scale Ensemble-Based Forecast Guidance. *Wea. Climate Soc.*, **11**, 337–354, https://doi.org/10.1175/WCAS-D-18-0084.1.

Zhou, X., and Coauthors, 2022: The development of the NCEP global ensemble forecast system version 12. *Wea. Forecasting*, **37**, 1069–1084, https://doi.org/10.1175/WAF-D-21-0112.1.