On the Changes in Convection-Allowing WRF Forecasts of MCS Evolution due to Decreases in Model Horizontal and Vertical Grid Spacing. Part II: Impacts on QPFs

BRIAN J. SQUITIERI[a] AND WILLIAM A. GALLUS JR.[a]

[a] *Department of Geological and Atmospheric Sciences, Iowa State University, Ames, Iowa*

(Manuscript received 22 February 2022, in final form 29 July 2022)

ABSTRACT: Several past studies have demonstrated improvement in forecasting convective precipitation by decreasing model grid spacing to the point of explicitly resolving deep convection. Real-case convective modeling studies have attempted to identify what model grid spacing feasibly provides the most optimal forecast given computational constraints. While Part I of this manuscript investigated changes in MCS cold pool characteristics with varied vertical and horizontal grid spacing, Part II explores changes in skill for MCS spatial placement, forward speed, and QPFs among runs with decreased horizontal and vertical grid spacing by employing the same WRF-ARW runs as in Part I. QPF forecast skill significantly improved for later portions of the MCS life cycle when decreasing horizontal grid spacing from 3 to 1 km with the part double-moment Thompson microphysics scheme. Some improvements were present in QPFs with higher precipitation amounts in the early stages of MCSs simulated with the single-moment WSM6 microphysics scheme. However, significant improvements were not common with MCS placement or QPF of the entire precipitation swath with either the Thompson or WSM6 schemes, suggesting that the benefit to MCS QPFs with decreased horizontal grid spacings is limited. Furthermore, increasing vertical resolution from 50 to 100 levels worsened WSM6 scheme QPF skill in some cases, suggesting that choices of or improvement in model physics may be equally or more positively impactful to NWP forecasts than grid spacing changes.

KEYWORDS: Convective storms/systems; Numerical analysis/modeling; Forecast verification/skill; Numerical weather prediction/forecasting; Cloud resolving models; Model evaluation/performance

1. Introduction

MCSs have a considerable impact on society. In the United States, MCSs are often responsible for producing severe weather (Jirak et al. 2003; Gallus et al. 2008; Hocker and Basara 2008), but also beneficial rainfall to the agricultural industry across the central United States (Ahijevych et al. 2016; Haberlie and Ashley 2019). Given the hazards and benefits posed by MCSs, improving the forecast accuracy of these events is crucial (Stensrud and Fritsch 1993, 1994; Coniglio et al. 2007; Jirak and Cotton 2007). MCSs can be challenging to simulate in numerical weather forecasts, both in deterministic (Jankov and Gallus 2004a,b; Grams et al. 2006; Trier et al. 2011; Squitieri and Gallus 2016a,b) and ensemble modeling frameworks (Wandishin 2008, 2010; Lawson and Gallus 2016; Clarke et al. 2019; Lawson et al. 2020).

Improvements in forecasting of deep moist convection were made when 3–6-km horizontal grid spacings were implemented to explicitly resolve storm-scale processes (Done et al. 2004; Weisman et al. 2008; Clark et al. 2009, 2010b), and the models employing these grid spacings are known as convection-allowing models (CAMs). However, when decreasing horizontal grid spacing (Δx) from 3–4 to 1–2 km, the benefits of the finer Δx become less clear. Some studies have depicted negligible improvements in quantitative precipitation forecasts (QPFs) for central U.S. warm season convective events that are not worth the increase in computational costs. Kain et al. (2008) noted during the 2005 NOAA Hazardous Weather Testbed (HWT) Spring Experiment that 0000 UTC initialized 2-km WRF forecasts showed a slight improvement in the coverage of 5 and 10 mm h^{-1} diurnal convective precipitation rates. However, 2-km runs showed an even greater moist bias than Δx = 4-km runs during the late afternoon and early evening (2000–0000 UTC). Both objective evaluations and subjective evaluations made by forecasters in the 2005 HWT Spring Experiment suggested that 2-km simulations did not add appreciable value to the forecasts despite showing greater detail in deep convective structures. Schwartz et al. (2009) followed up on Kain et al. (2008) by running WRF at Δx = 2 and 4 km to compare impacts of changing Δx on hourly convective precipitation accumulations. QPFs were rigorously statistically analyzed during the 2007 HWT Spring Experiment to assess more closely the impact Δx changes have on CAM generated QPFs. QPFs from Δx = 2 km simulations demonstrated a weaker moist bias compared to 4-km simulations at higher precipitation thresholds but were not more skillful than 4-km QPFs, regardless of precipitation threshold or radius of influence applied in neighborhood statistical skill scores. Squitieri and Gallus (2020) noted a modest improvement in 1- versus 3-km deterministic MCS QPF skill scores for 3-h periods in nocturnal MCS evolution, which was likely due to better placement of the heavier MCS precipitation, but these improvements were not statistically significant.

On the other hand, some studies have noted more meaningful improvements in forecasts of warm season deep moist convection in the central United States when model grid spacing was decreased. Schwartz et al. (2017) noted that 3-km ensemble QPFs outperformed 1-km deterministic forecasts, but that

Corresponding author: Brian Squitieri, brianjs@iastate.edu

DOI: 10.1175/WAF-D-22-0042.1

1-km ensemble QPFs were better than 3-km ensembles. Aligo et al. (2009) found that, when holding Δx constant at 4 km, decreasing vertical grid spacing (Δz) closest to the surface and near the melting layer improved warm-season convective precipitation forecasts. However, simply increasing the number of vertical levels throughout the troposphere sometimes decreased the skill. Few studies have investigated changes in properties of deep convection (including accumulated precipitation) for convection-allowing forecasts of real events with simultaneously decreased horizontal and vertical grid spacings. Squitieri and Gallus (2022, hereafter Part I) focused on understanding changes in nocturnal MCS cold pool characteristics due to decreased model grid spacing, but with no comparison to observations. Part II of the present work analyzes simulated MCS placement and QPFs relative to observations to assess potential value of simultaneously decreasing Δx and Δz in same-day MCS forecasts. Kain et al. (2008) and Schwartz et al. (2009) evaluated QPFs at or beyond 24 h to focus on the performance of next-day convective forecasts, regardless of the convective mode expected, whereas the current study more closely evaluated 12–24-h deterministic QPFs specifically involving nocturnal MCSs. The authors hypothesized that a more appreciable improvement in nocturnal MCS QPFs may be achieved if Δz is decreased (better resolving melting-layer hydrometeorological processes as in Aligo et al. (2009) along with decreased Δx, where more localized swaths of higher precipitation accumulations are better resolved or placed, as in Schwartz et al. (2017) and Squitieri and Gallus (2020).

Section 2 of this paper highlights the data used for verification and model configurations employed for the experiment, and discusses the methodologies embraced for the present study, while section 3 focuses on results. Section 4 provides a brief summary and conclusion of the research findings, and section 5 facilitates a discussion of the results and potential implications for future applications.

2. Methodology

a. Observational and model data

Multi-Radar Multi-Sensor (MRMS) derived observed mosaic 1 km above ground level (AGL) radar reflectivity (Smith et al. 2016) was used to compare with simulations, which was retrieved from a database managed by the Iowa Environmental Mesonet (Iowa Environmental Mesonet 2020). The details on case selection are provided in Part I. Stage-IV 4-km gridded data (NCAR/UCAR/EOL 2015) were used to represent observed quantitative precipitation estimates (QPEs). All WRF data were regridded to 4 km for exact comparisons with Stage-IV before any postprocessing or calculations were made. The 3- and 1-km WRF runs, both with 50 and 100 vertical levels, were the same WRF-ARW version 4.2 (Skamarock et al. 2019) runs assessed in Part I of this manuscript, with the same model physics (Part I, Table 2), domains (Part I, Fig. 1), and 11 real cases (Part I, Table 1) and were simulated on the CISL-Cheyenne supercomputing system (CISL 2020). The 28 June 2020 case was excluded from the QPF analyses and only included in MCS placement and forward speed statistics since forecast skill scores were exceptionally poor for this event, serving as an outlier in aggregated statistical results involving a relatively small sample size. As in Part I of this research, the different model grid spacings are defined as such: 3 km with 50 vertical levels is 3 km–50 lev, 3 km with 100 vertical levels is 3 km–100 lev, with the same abbreviation applied to 1-km runs with 50 and 100 vertical levels (1 km–50 lev and 1 km–100 lev, respectively).

b. Calculating MCS placement, timing, and forward speeds

Observed and simulated MCS leading-line centroid locations were recorded to determine how simulated MCS spatial displacement errors varied with changing Δx or Δz. Observed and simulated MCS initiation (t0) times as well as leading-line centroid locations of the initiating system were defined in the same manner as in Squitieri and Gallus (2020), with MCS centroids subjectively placed at the leading line of deep convection (Fig. 1), and with forecast MCS initiation times and centroids recorded independently from observations. The simulated MCS initiation times used in Part I of this study were also used here in Part II. It was difficult to cleanly identify simulated MCS leading-line centroids after t5 for a subset of cases, so no spatial displacement errors were calculated past this time. Objective approaches were attempted by employing the Method for Object-Based Diagnostic Evaluation tool (MODE; Davis et al. 2006a,b).

The 1 km AGL reflectivity, composite reflectivity, and 1-h precipitation accumulations were evaluated at varying thresholds and levels of convolution to produce a consistent definition of an MCS based on object identification. All combinations of MODE criteria showed considerable differences in MCS initiation time (1–3 h) and structure of the reflectivity and precipitation objects, which would greatly impact the location of centroids and subsequent results. Specifically, merging deep convective clusters often displaced centroids in a fashion not representative of overall MCS forward motion (similarly noted in Schwartz et al. 2017), with the stratiform rain region expanding upstream of MCS motion at roughly the same pace as MCS leading-line downstream movement, causing the centroid to remain nearly stationary (a similar dilemma experienced by Clark et al. 2014). Given the complexities with identifying MCS centroids with MODE, the subjective approach was determined to be more accurate and was undertaken. Also, as in Squitieri and Gallus (2020), leading line observed and simulated MCS centroids were used to calculate MCS forward speeds. For any given hour t, the forward speed was derived by calculating the great circle distance between the MCS centroids at hours t and $t - 1$, with the distance (m) divided by 3600 s for a speed in units of meters per second. Given the struggles in getting reliable MCS centroids after t5, forward speeds were only calculated for t1–t5.

c. Derivation of QPF swaths, skill scores, and calculation of statistical significance

QPF skill scores were determined using the neighborhood equitable threat score (nETS; Clark et al. 2010a), with Stage-IV

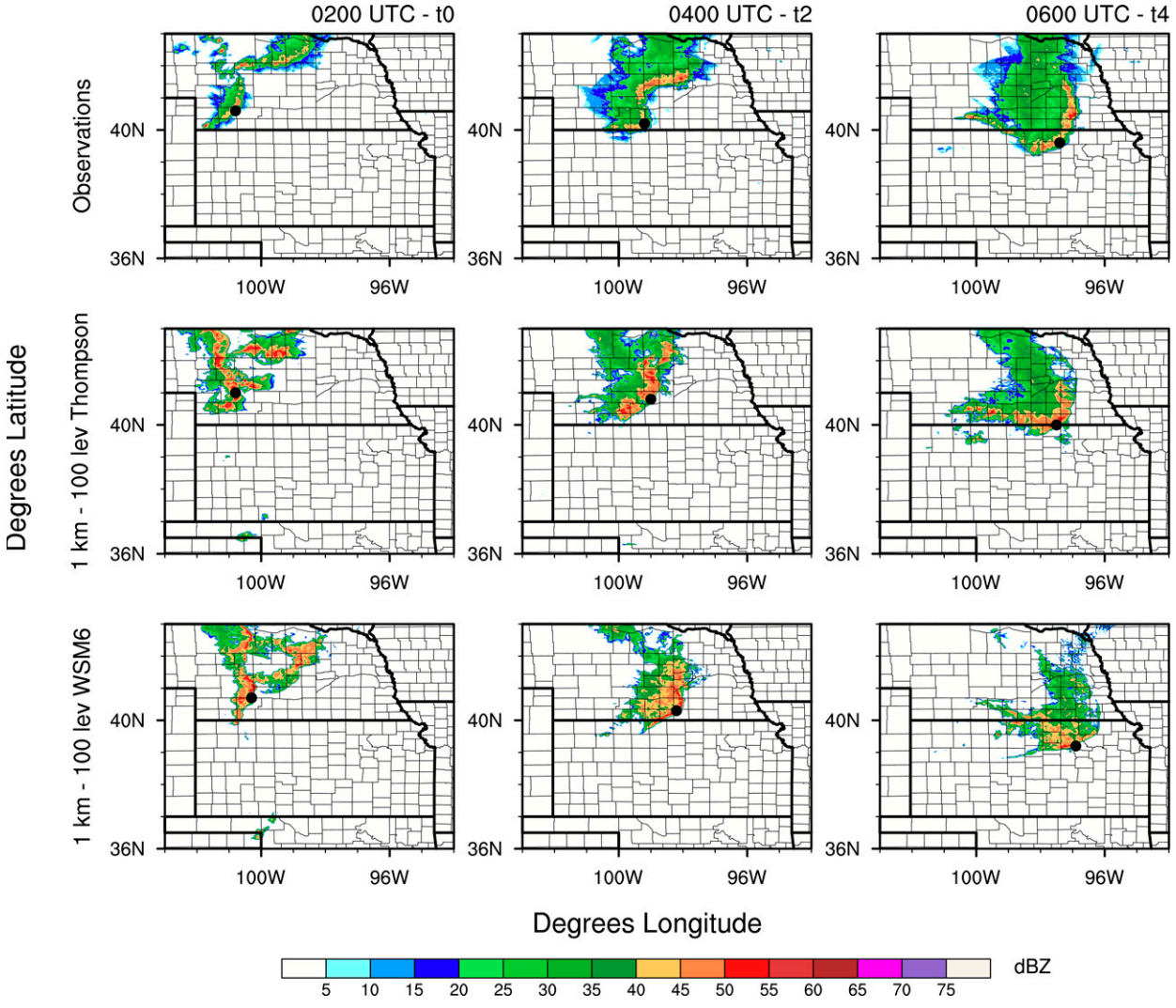


FIG. 1. Plots of 1 km AGL reflectivity (see color bar for dBZ values) for the 9 Jul 2020 nocturnal MCS across the Central Plains. Filled contours represent reflectivity for (top) observations, (middle) the 1 km–100 lev simulation with Thompson microphysics, and (bottom) WSM6 microphysics. All plots show the evolution of the observed and simulated MCS from initiation (t0) to 4 h after initiation (t4), at 2-h intervals. Black dots indicate the subjectively defined leading-line MCS centroids.

data used as observations for QPE validation. As in Squitieri and Gallus (2020), nETSs were calculated with 30-, 60-, and 120-km radii of influence, for 12.7- and 25.4-mm (0.50 and 1.0 in.) thresholds. MCS precipitation was accumulated in two ways. First, QPEs and QPFs were accumulated in 3-h overlapping bins, for every hour of MCS evolution, beginning at the observed MCS initiation time, to determine how changes in Δx and Δz may impact portions of MCS evolution. Second, QPEs and QPFs were accumulated in the 2100–1200 UTC period, encompassing the entire observed and simulated MCS precipitation swaths, to determine if decreased model grid spacing had significant impact on total QPFs for the MCS lifetime.

MCS QPE and QPF composites were also created for the 2100–1200 UTC period. For QPEs and QPFs of each case, the composite 600 km × 600 km domain was centered on the t3 observed MCS leading-line centroid, which would encompass the approximate center of the MCS accumulated QPE swath, with all cases under observations and model configurations averaged to produce the composites. The QPF composites (12.7- and 25.4-mm thresholds) were overlaid as line contours atop QPE filled contours. Some additional smoothing was performed on the lined contours to remove artifacts introduced by outlier cases. Since 25.4-mm accumulations were "patchier" compared to 12.7 mm, smoothing eliminated some of the smaller 25.4-mm accumulation areas. It is acknowledged that compositing QPEs and QPFs among multiple cases may yield a system not resembling any observed. However, the composites were created with the purpose of evaluating overlapping 12.7- and 25.4-mm QPFs and QPEs to explain nETS trends common among varying Δx and Δz configurations from multiple cases. Most of the 10 observed and simulated MCSs moved from northwest to southeast, so the

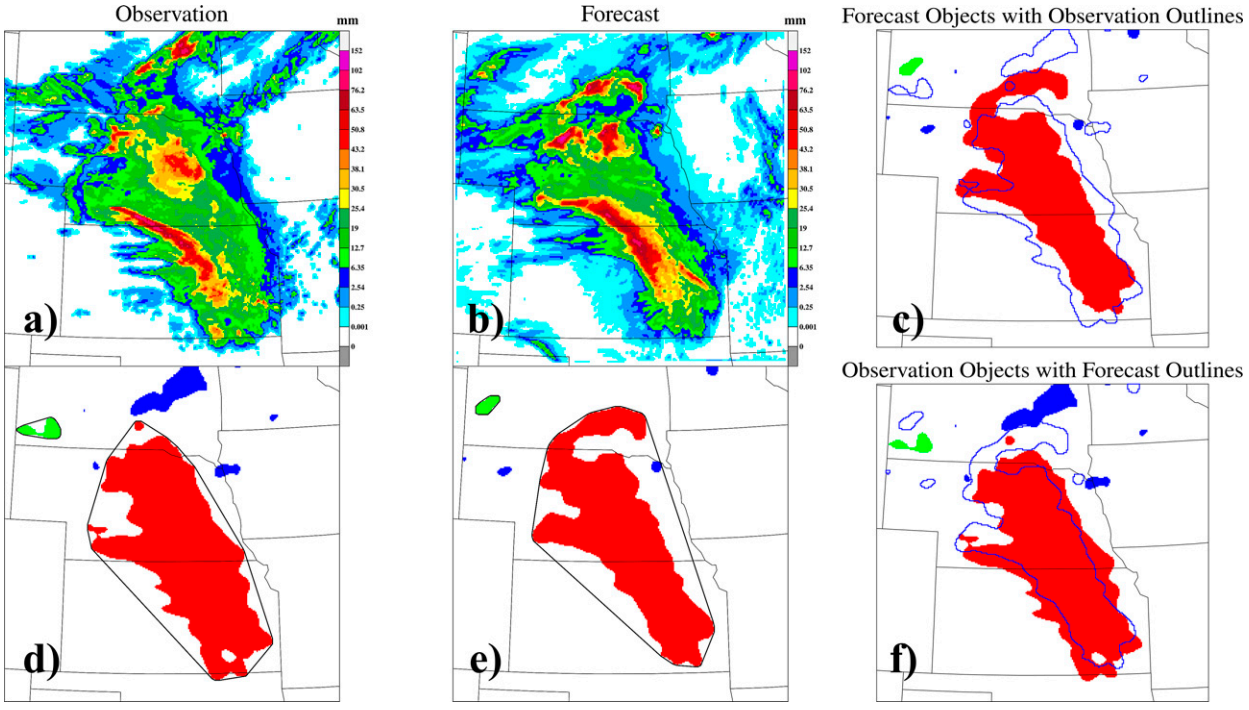


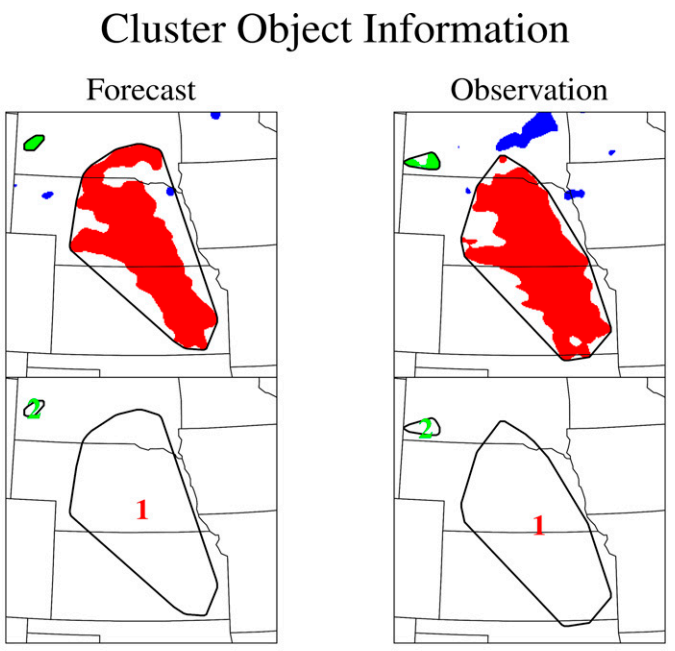
FIG. 2. MODE generated plots exemplifying the classification of 2100–1200 UTC precipitation accumulation swaths for the 9 Jul 2020 case. Total accumulations are represented by filled contours (see color scale to the right of each image) for (a) observations and (b) the 1 km–50 lev simulation. (c)–(f) All accumulation values below the 12.7-mm threshold are excluded, with a convolution radius of 20 km applied to smooth the remaining values to identify more coherent features. The resulting observed and forecast objects in (d) and (e), respectively, are delineated by filled contours, with black lined contours identifying the cluster objects that contain the precipitation objects belonging to the MCS precipitation swaths. The individual observed and forecast cluster objects are matched based on the distance between their centroids, with (c) showing lined contours of observed objects overlaid with forecast objects (filled contours), and (f) showing lined contours of forecast objects overlaid with observed objects (filled contours).

beginning of the MCS QPE and QPF swaths in the composites start from the northwest and end toward the southeast.

Though not utilized to identify MCS initiation times or placement centroids, MODE was used to compare the physical characteristics of the aforementioned 2100–1200 UTC accumulated rainfall swaths above the 12.7-mm threshold between observations and simulations for all four grid spacings. The 12.7+ mm QPE and QPF swaths (Figs. 2a,b) were categorized into observation and forecast objects (Figs. 2d,e), whose placement and spatial characteristics were compared (Figs. 2c,f). The main QPE and QPF swaths and adjacent smaller MCS precipitation objects were grouped within single cluster objects (Fig. 3). Cluster objects were chosen to be statistically analyzed as they were less sensitive to changes in grid spacing, as opposed to individual observation and forecast objects, which were more easily fragmented and thus difficult to evaluate statistically. Similarly, 25.4+ mm threshold QPE and QPF swath evaluations were attempted in MODE, but the highly localized (and fragmented) nature of the observation and forecast objects in multiple cases inhibited the generation of consistent cluster objects, even when no convolution radius was applied. A convolution radius of 5 grid squares (or 20 km) was used for the 12.7-mm QPEs and QPFs in order to smooth precipitation tracks into one coherent swath representative of an MCS QPE or QPF swath. A maximum centroid distance error threshold of 500 km was employed to ensure that the QPE and simulated MCS QPF swaths would be compared, regardless of the magnitude of spatial displacement. In previous studies, simulated MCS swaths have been displaced north or east of observations (Squitieri and Gallus 2016b; Peters et al. 2017; Vertz et al. 2021).

The QPE and QPF centroid differences, angle differences, intersection and union areas, symmetric differences (forecast area not covered by observations), differences in forecast versus observation areas and swath intensity differences above the 50th and 90th percentiles of matching cluster objects were the statistics evaluated. The goal was to determine which aspects (i.e., placement, orientation, or spatial coverage) of simulated MCS QPF swaths, per case, may have contributed most to changes in forecast skill scores with respect to changing Δx or Δz. The differences in observation and forecast areas and 50th and 90th threshold intensities were calculated by subtracting forecast values from observations.

Last, all statistical bootstrap significance calculations employed the same two-tailed hypothesis testing used in Part I of the present study, with resampling performed 10 000 times (Mendenhall and Sincich 2007). Significance testing was performed when comparing model grid spacing statistics (Part I, Table 4), with p values less than 0.05 representing statistical significance and p values between 0.05 and 0.10 suggesting



CLUS PAIR	CEN DIST	ANG DIFF	FCST AREA	OBS AREA	INTER AREA	UNION AREA	SYMM DIFF	FCST INT 50	OBS INT 50	FCST INT 90	OBS INT 90	TOT INTR
1	16.76	0.24	9919	10757	7775	12901	5126	23.70	22.27	51.25	42.50	0.9975
2	18.68	34.20	152	230	0	382	382	19.30	17.04	34.00	26.09	0.7194

FIG. 3. As in Figs. 2e and 2d, but explicitly labeling the matched observed and forecast cluster objects, with the 9 Jul 2020 swaths, delineated by black lined contours, assigned the value of "1." Provided with these plots are the MODE generated cluster object comparison statistics. The matched cluster object pairs associated with the MCS precipitation swaths were manually identified, with the comparison statistics for the appropriate matching cluster pair collected for statistical aggregation. Example output of cluster-pair statistics includes (from left to right) centroid distance displacement (km) of the simulated cluster centroid relative to the observed cluster centroid; absolute value of the difference in orientation angles (°) between the observed and forecast clusters; forecast cluster area (km^2); observed cluster area (km^2); intersection area where the observed and forecast cluster overlap (km^2), union area covered by the observed cluster, forecast cluster, or both (km^2); symmetric difference between the observed and forecast cluster, or area not covered by observations (km^2); 50th percentile of the forecast and observed cluster magnitudes (mm); 90th percentile of the forecast and observed cluster magnitudes (mm); and total interest of the matched observed and forecast cluster pair.

that a spatial displacement, MCS speed, QPF skill score, or QPF swath characteristic of a sample is noticeably different (greater or smaller) than its counterpart.

3. Results

a. *Grid spacing impacts on MCS placement and forward speeds*

MCS leading-line centroids (exemplified in Fig. 1) were evaluated for the first 5 h of MCS evolution to gauge how decreasing both Δx and Δz may impact the placement of simulated MCSs relative to observations. From t0 to t4, for both microphysics (MP) schemes tested (Thompson MP; Fig. 4 and WSM6 MP; Fig. 5), the differences in displacement errors among grid-spacing configurations in nearly all cases were smaller compared to the differences in displacement errors between any simulated MCS and observations. Similarly, Squitieri and Gallus (2016b) found that varying the selection of PBL schemes (under a given MP scheme) did not meaningfully impact simulated MCS initiation spatial displacements. The results in Squitieri and Gallus (2016b) and the present study suggest that changes in model grid spacing or selection of MP/PBL schemes have little impact on the placement of centroids for larger deep convective features such as MCSs.

For both Thompson and WSM6 MP runs, spatial displacement errors of all simulated MCSs appeared smallest at t0 and gradually increased with time to t4. Thompson MP simulated MCSs were displaced along a more south–north axis. Bootstrap significance testing trends between Δx and Δz Thompson MP configurations showed that 1 km–50 lev runs were noticeably more displaced from observations compared to all other grid spacings, especially at t2 (Fig. 4b). When evaluating MCS forward speeds (Fig. 6a), 1 km–50 lev MCSs generally moved faster than observations (as in Squitieri and Gallus 2020), especially after t3, where the 1 km–50 lev MCSs were noticeably to significantly faster than observed MCSs. The 1 km–100 lev simulations were also noticeably faster moving compared to observations. However, despite the general increase in forecast MCS speeds with time as Δx or Δz

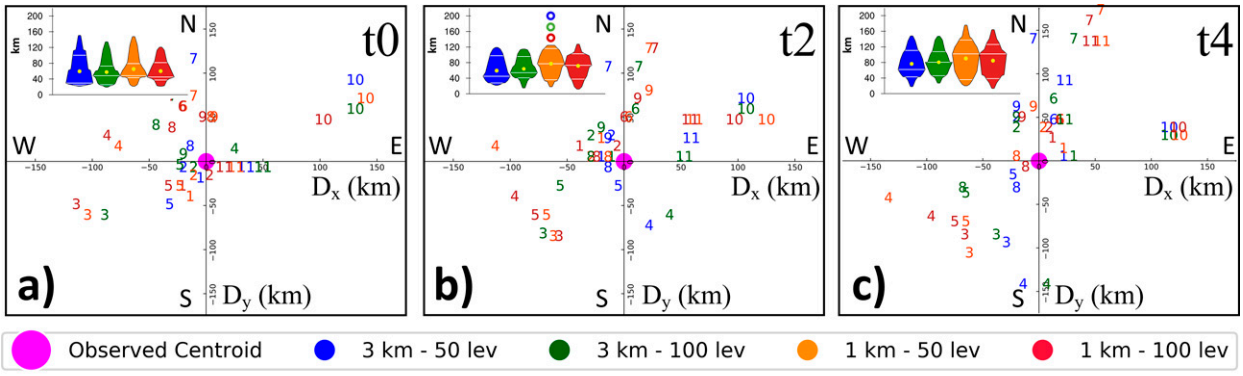


FIG. 4. Cartesian plots depicting the simulated MCS leading-line centroid spatial displacements relative to observations (delineated at the origin with a large pink dot). The *x* and *y* axes denote distance displacements (km), with positive *x* representing east, negative *x* representing west, positive *y* representing north, and negative *y* representing south. The blue, green, orange, and red numbers represent the spatial displacements of a specific case for the 3 km–50 lev, 3 km–100 lev, 1 km–50 lev, and 1 km–100 lev MCS leading-line centroids, respectively, from (a) the time of MCS initiation (t0) to (c) 4 h after MCS initiation (t4), at 2-h intervals. For each time of displacement, a violin plot aggregating the absolute magnitude of the WRF MCS leading-line centroid distance displacements (km on the *y* axis) is provided. For the violin plots, filled colors represent the distribution of absolute magnitude distances for all 11 cases at each grid-spacing configuration, with yellow dots within the filled colors representing the average distance displacement among 11 cases for a given model grid-spacing configuration. Unfilled dots associated with the violin plots show where a sample is noticeably bigger or smaller than the sample it was compared to but was not statistically significant. For example, at t2, the 1 km–50 lev centroid displacements were noticeably higher than their 3 km–100 lev and 1 km–100 lev counterparts (with a *p* value between 0.05 and 0.10).

decreased, simulated MCS spatial displacements were only slightly larger on average with runs employing finer grid spacing compared to coarser simulated MCSs. Figure 4c (t4) total spatial displacement violin plots show that the 1-km Thompson MP runs exhibited a larger spread of spatial displacements relative to 3-km simulations.

WSM6 MP runs depicted an increasingly easterly (farther downstream placement) bias of simulated leading-line MCS centroids that was evident among a majority of cases, even at MCS initiation (t0; Fig. 5a), as also noted for MCS initiation leading-line centroids for WSM6 MP runs in Squitieri and Gallus (2016b). The Fig. 5 Cartesian and total spatial displacement violin plots show (by t4; Fig. 5c) that 1-km runs exhibited larger east-southeastward spatial displacement errors compared to 3-km simulations for nearly all cases, though the total spatial displacements were not statistically significantly different between the different grid spacings. Despite the lack of statistical significance for differences in total spatial displacements in 1- versus 3-km runs, the width of the t4 WSM6 MP-based violin plots (Fig. 5c) suggests that several more 1-km simulated MCSs were displaced by 60–120 km compared to 3-km runs due to finer Δx and Δz MCSs moving faster than both observed and 3-km MCSs. Figure 6b shows that 1-km WSM6 MP-simulated MCSs (especially 50 vertical levels) moved significantly faster than observations in the t3–t4 period, with 1 km–100 lev simulated MCSs moving noticeably faster earlier in their evolution (t1–t2).

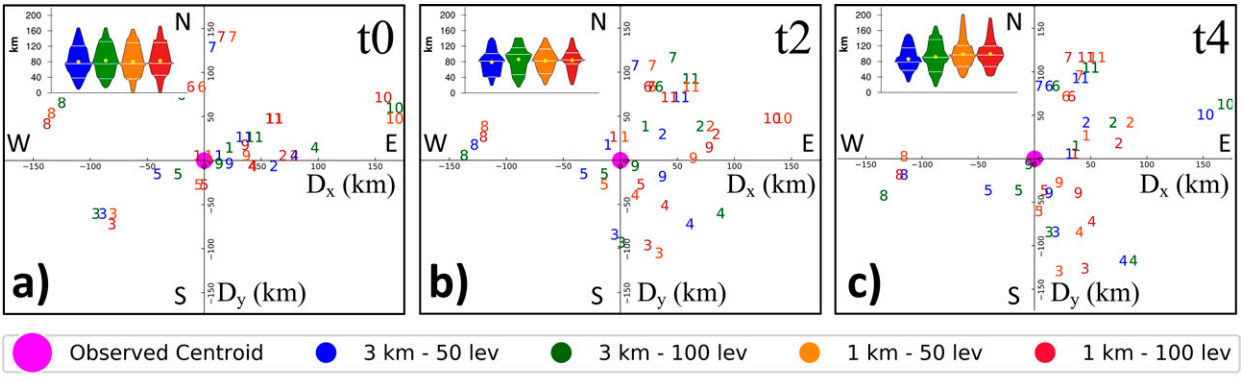


FIG. 5. As in Fig. 4, but for the WSM6 MP scheme.

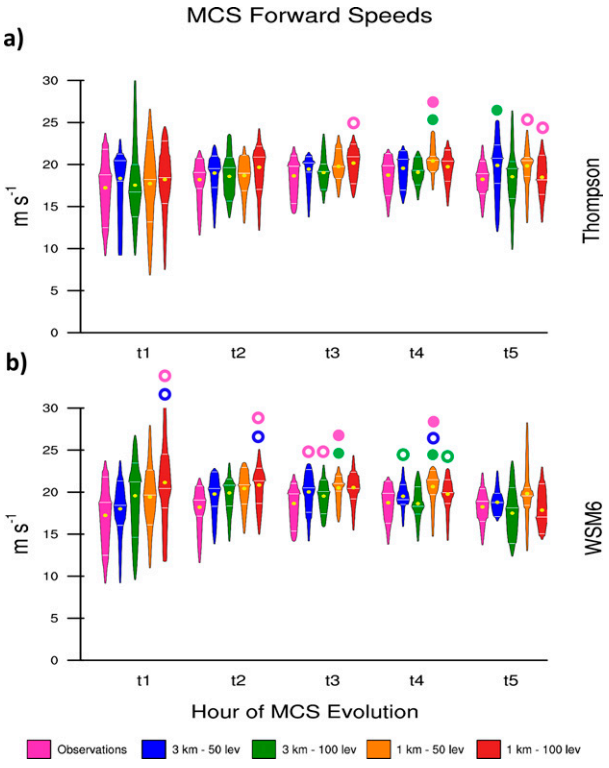


FIG. 6. Violin plots depicting MCS forward speeds (m s^{-1}) for the 11 observed MCSs (pink) and associated simulated MCSs with grid spacings of 3 km–50 lev (blue), 3 km–100 lev (green), 1 km–50 lev (orange), and 1 km–100 lev (red) using (a) the Thompson MP scheme and (b) WSM6 MP scheme, for 1–5 h after MCS initiation (t1–t5). White horizontal lines on the violins delineate the 25th, 50th, and 75th percentiles. The configuration for the violin plots and notation for statistical significance is as in Fig. 4.

b. *Changes in QPF skill scores with decreased horizontal and vertical grid spacing*

The 3-h nETSs were lowest around observed MCS initiation time (t0), with QPF skill peaking during MCS maturity (t3–t5), then decreasing toward the decaying stage (t6–t7), for all grid-spacing configurations in both Thompson MP (Fig. 7, except for the 25.4-mm threshold–120-km radii influence) and WSM6 MP-based (Fig. 8) runs. Thielen and Gallus (2019) also noted a decrease in MCS morphology accuracy toward the late stages of MCS evolution in 3- and 1-km runs, regardless of MP scheme implemented, suggesting that convection-allowing guidance still struggles to accurately depict MCS structures late in the MCS life cycle. For both Thompson and WSM6 MP simulations, greater spread was noted for nETSs at the 25.4-mm threshold (Figs. 7–8d,f) compared to 12.7-mm threshold nETSs (Figs. 7–8a,c). This was most noticeable at the 120-km radius of influence, where both Thompson and WSM6 MP nETS averages (yellow dots) across MCS evolution varied between 0.7 and 0.8 at the 12.7-mm threshold (Figs. 7–8c). At the 25.4-mm threshold, Thompson and WSM6 MP nETSs were 0.4–0.7 (Figs. 7–8f).

Thompson MP-based nETSs (for most times and for the 12.7- and 25.4-mm thresholds) increased as both Δx and Δz decreased (Fig. 7), especially as the radius of influence increased, forgiving larger spatial or temporal scale displacements. The nETS significance comparisons between model grid spacings with the Thompson MP scheme show that QPFs with the finer Δx and Δz runs were noticeably to significantly more skillful compared to QPFs in the coarser runs employing a larger radius of influence, especially later in MCS evolution. When evaluating the total distance displacement errors of Thompson MP runs around t4 (Fig. 4c), the 1-km runs showed a greater spread of displacement error among all cases compared to 3-km simulations, despite 1-km simulated MCS forward speeds being noticeably to significantly faster than observations (Fig. 6a). This suggests that refining Δx may either improve or degrade MCS placement forecast accuracy on a case-by-case basis.

However, QPE/QPF Thompson MP-based composites (Figs. 9a,d) show that 1-km runs have a larger overlap of 25.4-mm QPF atop higher QPE area compared to 3-km runs, with 12.7-mm QPFs within 120–240 km on the ordinate and 240–360 km on the abscissa better matching QPEs. The better spatial placement of MCSs or timing of higher precipitation amounts of some cases in the 1-km runs may have contributed to the significantly higher 1-km member nETSs toward the later stages of MCS evolution. Nonetheless, the statistically significantly higher 1-km nETSs were noted primarily with high (120-km) radii of influence, raising the question of how valuable these more skillful runs with finer grid spacing truly are when considering high radii of influence and associated forgiveness of large spatial displacements. For all radii of influence, 1-km simulations with 100 vertical levels did noticeably improve skill scores compared to 1 km–50 lev runs around 4–5 h after MCS initiation. However, this improvement was insignificant.

Thompson MP-based nETSs derived from 2100 to 1200 UTC accumulated QPE versus QPF swaths show that skill scores still increased with lower precipitation thresholds and increasing radii of influence (Fig. 10). However, the spread of nETSs and differences in their averages between grid spacings were much smaller compared to the 3-h bins in Fig. 7, regardless of radii of influence or precipitation threshold selected. Some slight differences in the distribution of nETSs may be noted between grid spacings when studying the shape of the violin plots in Fig. 10. For example, in Fig. 10d, the higher nETS magnitudes within the lower quartile of the 1- versus 3-km violins at the 30-km radii of influence–25.4-mm threshold nETSs suggest that slight improvements were made to 1-km forecasts of locally higher precipitation amounts embedded within simulated MCS swaths. However, when decreasing the precipitation threshold or increasing the radius of influence, the signal of modest improvement with decreased model grid spacing disappears, with one exception. Bootstrap testing nETSs of the varying Δx and Δz configurations for the 2100–1200 UTC accumulation period revealed that the 1 km–100 lev Thompson MP runs were significantly more skillful compared to 3 km–100 lev simulations for the

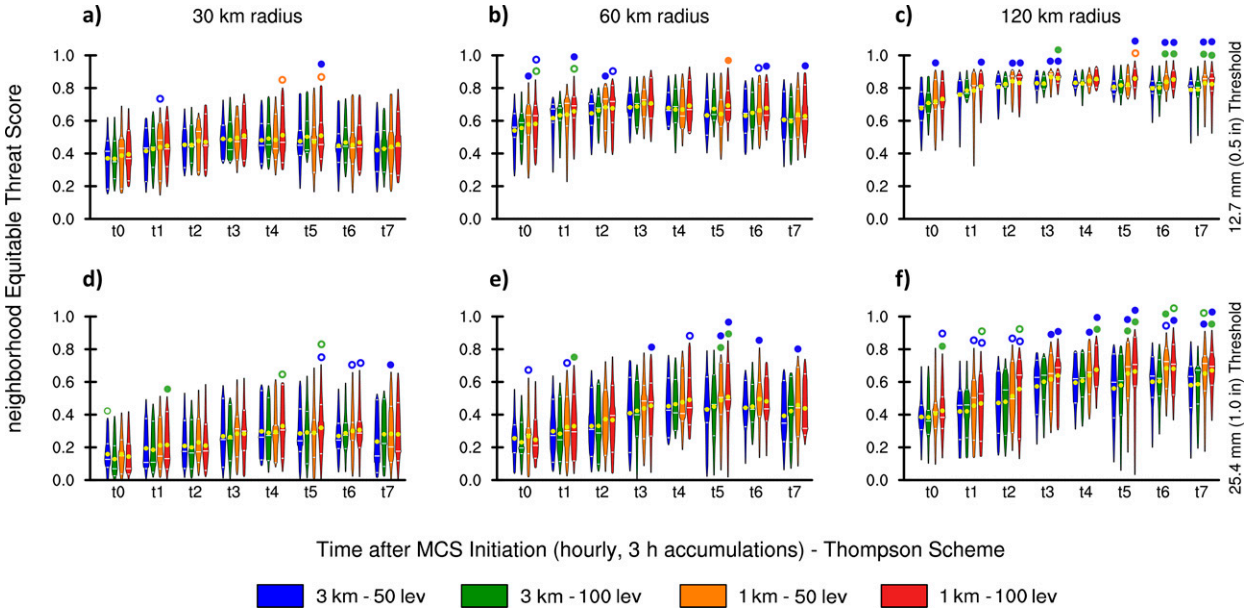


FIG. 7. Violin plots showing the distribution of nETSs for 3-h QPFs, for each model grid-spacing configuration, starting at MCS initiation time (t0), and progressing hourly to 7 h after MCS initiation (t7). The blue, green, orange, and red violins represent the 3 km–50 lev, 3 km–100 lev, 1 km–50 lev, and 1 km–100 lev nETSs, respectively. Yellow dots within each violin depict the average of the 10 cases within each set. White horizontal lines on the violins delineate the 25th, 50th, and 75th percentiles. nETSs were calculated at (a)–(c) the 12.7-mm (0.5 in.) and (d)–(f) 25.4-mm (1.0 in.) thresholds, employing a (left) 30-, (center) 60-, and (right) 120-km radius of influence. All results are for Thompson MP simulations. Between each tested pair, statistical significance was delineated on the figures, where the statistically significantly larger sample was labeled with a dot bearing the color of the significantly smaller sample. If a violin is labeled with a filled dot with the color of another member, then the violin member is statistically significantly larger than the other member(s) represented by the dot's color(s). For example, the nETS score at t7, with a 120-km radius and 12.7-mm threshold, the 1 km–100 lev QPFs were significantly more skillful than both 3 km–50 lev and 3 km–100 lev runs. The same concept applies to an unfilled dot, but for noticeable changes between samples (p value in the 0.05–0.10 range).

25.4-mm threshold at the 120-km radius of influence (Fig. 10f). Still, no other instances of statistical significance were observed between the differences in skill scores for different grid-spacing combinations for the 2100–1200 UTC accumulated QPFs. Overall, it appears that the slight improvements made to the QPF swath skill scores were driven primarily by decreasing Δx, with smaller changes noted when changing only Δz.

Results were considerably different for WSM6 MP-based nETSs with changing Δx and Δz. For the 120-km radii of influence, primarily for the 25.4-mm threshold, 1-km runs showed noticeably to significantly more skillful 3-h QPFs toward the earlier stages of MCS evolution (Fig. 8f). An analysis of the QPE/QPF composites (Figs. 9e,h) suggest that in the 3-km runs (360–480 km on the ordinate and 360–480 km on the abscissa), 25.4-mm QPF accumulations were displaced too far north and east of the QPEs. However, this signal was not as strong as the significant improvement in QPF skill toward the later part of the Thompson MP simulations in the runs with finer versus coarser grid spacing. Furthermore, for all radii of influence at the 12.7-mm threshold, the 3 km–50 lev and 1 km–50 lev WSM6 MP runs were noticeably to significantly more skillful than their 100 vertical levels counterparts toward the end of MCS evolution (t5–t7, Figs. 8a,c).

The QPE/QPF composites show that toward the end of MCS evolution (0–120 km on the ordinate and 480–600 km on the abscissa), the 12.7-mm QPFs in the 1 km–50 lev runs better overlap with 12.7-mm QPEs compared to 1 km–100 lev (Figs. 9f,h). A closer look at Stage-IV and WRF 1-h precipitation accumulation plots also suggests that WSM6 MP-simulated MCSs weakened or shrunk in size earlier in finer Δx and Δz runs compared to in coarser simulations (not shown), implying that MCSs weakened too early in their life cycle. While it was difficult to cleanly identify simulated MCS centroids after t5 for some cases (hence no spatial displacement errors calculated past this time), Fig. 5 also shows that multiple cases (3, 4, 9, 10), especially for finer Δx and Δz, became increasingly displaced from observations with time. This trend likely continued past t5, where the 1 km–100 lev simulated MCS continued to surge farther southeast compared to observations at t4–t6. Furthermore, the WSM6 MP-simulated MCS forward speeds (Fig. 6b) show that forecast MCSs with finer model grid spacing, in general, move too quickly around t4–t5, contributing to greater spatial displacements.

Slight changes were noted between coarser and finer grid spacing 2100–1200 UTC accumulated WSM6 MP-based QPF nETSs, with a few exceptions noted. At the 12.7-mm threshold, for 30–60-km radii of influence, the 3 km–100 lev swaths were noticeably to significantly more skillful than the 1 km–100 lev events, likely due to the contributing rapid

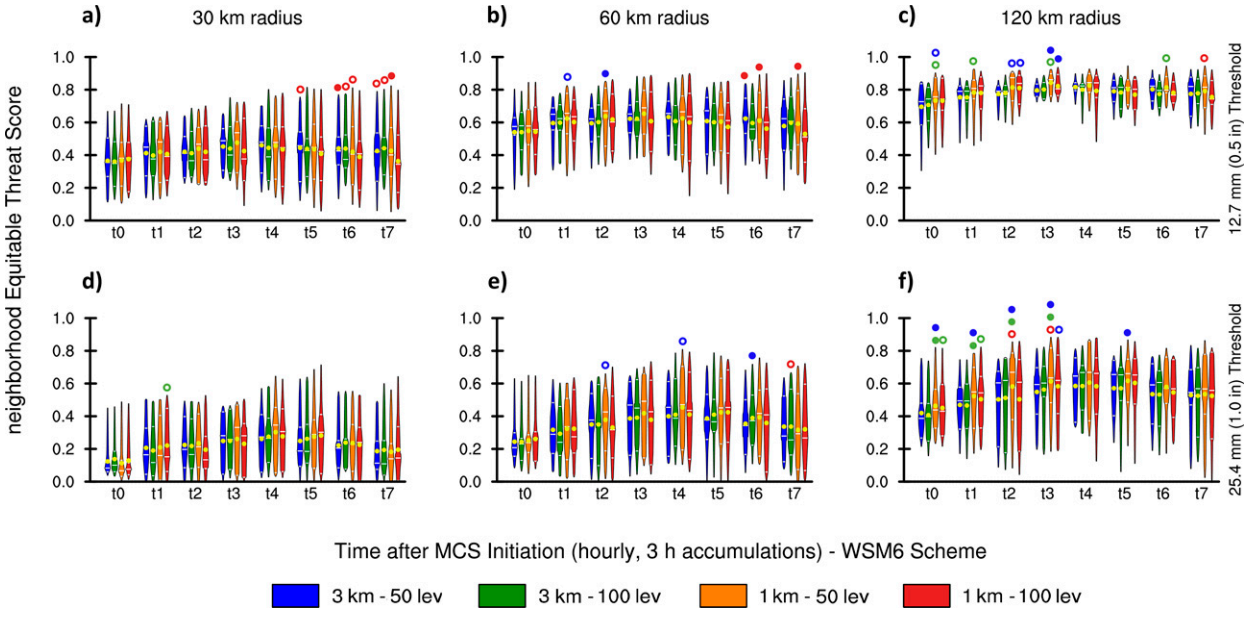


FIG. 8. As in Fig. 7, but for the WSM6 MP scheme.

decrease in overall skill of the finest Δx and Δz runs in the t6–t7 period with simulated MCSs weakening early (Figs. 11a,b). When the 2100–1200 UTC WSM6 MP accumulated QPFs were bootstrap tested, the finer 1 km–50 lev simulated swaths were significantly more skillful than the 3 km–100 lev swaths at the 12.7-mm threshold, for a 120-km radius of influence (Fig. 11c). It is interesting to note that 1 km–50 lev 3-h accumulations were only noticeably higher than 3 km–100 lev for 12.7-mm threshold swaths at 120-km radius of influence, yet statistical significance of differences was present for a larger accumulation period. Reasoning for this disparity is discussed in section 3c. On the other hand, finer Δx and Δz grid spacings were not significantly more skillful than the coarser runs at the 25.4-mm threshold–120-km radius of influence, so the significant improvements in 3-h WSM6 MP-based QPFs with decreased model grid spacing were not influential enough to significantly improve longer-term accumulated WSM6 MP-simulated MCS QPF swaths.

c. *Grid spacing impacts on MCS precipitation swath characteristics*

Individual 2100–1200 UTC QPE and QPF swath characteristics were derived in MODE, with spatial distribution statistics aggregated to determine why the skill of forecast accumulated precipitation swaths with finer grid spacing differed from that of their coarser counterparts. For Thompson MP runs, the centroids of the finer Δx and Δz 2100–1200 UTC 12.7-mm swaths were increasingly far displaced from the QPE centroids compared to the coarser simulations (Fig. 12a), as noted earlier for the later times for Thompson MP leading-line centroids. However, when comparing the differences in areal coverage between forecast and observed swaths (forecast − observed area), 1 km–100 lev forecast areal coverage was noticeably to significantly greater than all other grid spacings, with 1 km–50 lev and 1 km–100 lev QPF and QPE swaths (on average) overlapping more than 3-km QPFs and QPEs (Fig. 12c). While the 1-km simulated swaths generally have higher symmetric differences (12.7+ mm area not covered by observations, or false alarm, Fig. 12f) and greater errors in orientation angles of the swaths (Fig. 12b), 1-km simulations do have greater intersection and union areas compared to 3 km (with some noticeable to significant improvements noted), likely explaining why finer Δx and Δz runs earn higher QPF skill scores despite greater centroid spatial displacements (Figs. 12d,e).

While differences in the intensities of the forecast versus observed precipitation accumulations at the 50th percentile were similar for all model grid spacings under Thompson MP (Fig. 12g), runs with finer grid spacing more closely matched observations for some of the heaviest precipitation accumulations (90th percentile accumulations within the 12.7-mm swaths, Fig. 12h). This means that 1-km runs reduced the WRF moist bias to a degree for some of the highest accumulations, with 1 km–50 lev runs significantly improving the distribution of the heaviest precipitation over their 3-km counterparts.

Similar to the Thompson MP simulated 2100–1200 UTC swath cluster objects, WSM6 MP-simulated clusters showed a gradual increase (on average) in difference of angular orientation of the swaths, and symmetric difference in areal coverage compared to observations when decreasing Δx and Δz (Figs. 13b,f). However, while an increase in centroid spatial displacements was noted with 3 km–100 lev runs as Δx was decreased to 1 km, 3 km–50 lev runs seemed to exhibit the greatest centroid displacement from observations, with 1 km–50 lev runs significantly improving spatial displacements, and even noticeably improving spatial displacements compared to 1 km–100 lev simulations (Fig. 13a). In fact, the 1 km–50 lev cluster objects shared the most areal coverage with observations, with

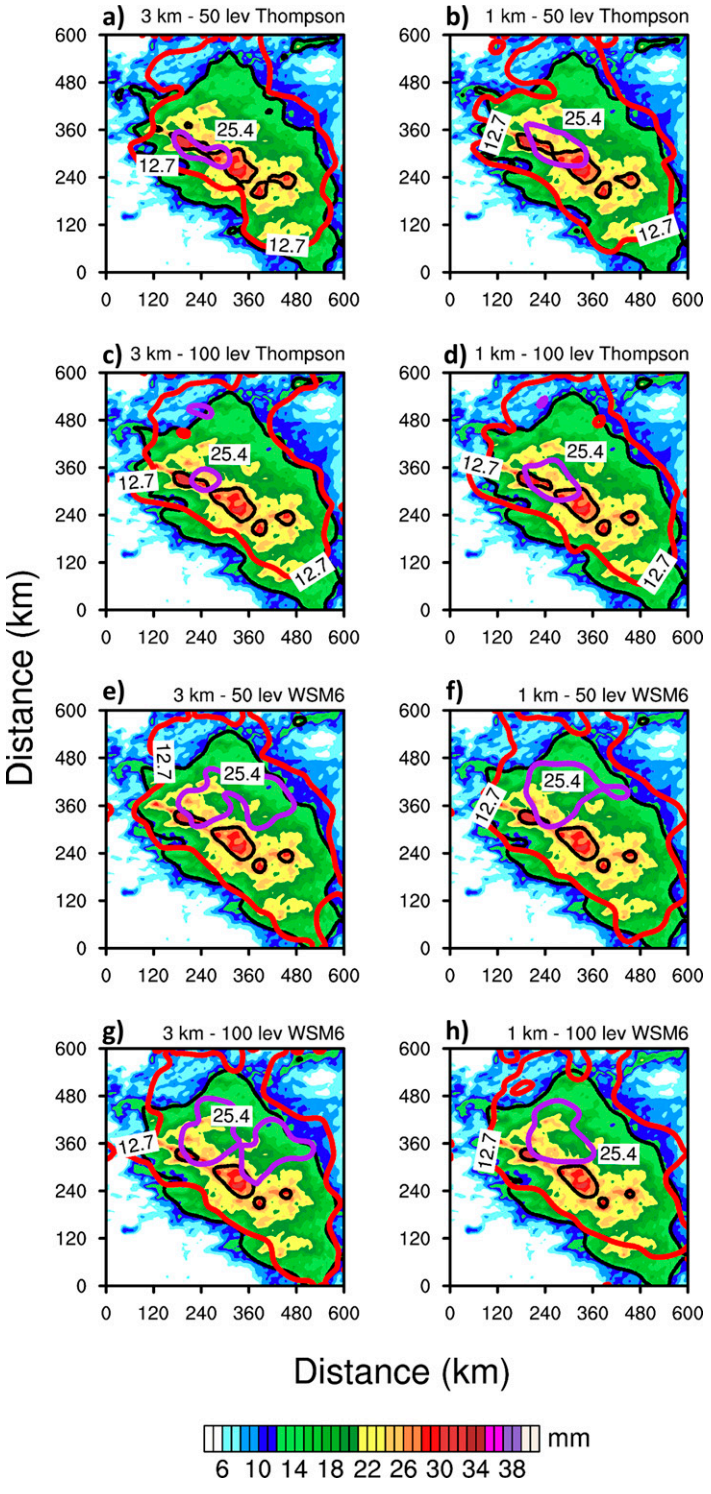


FIG. 9. MCS 2100–1200 UTC accumulated precipitation composites for the 10 real cases sampled. Filled contours in all plots (mm; see color scale at bottom) represent Stage-IV observations, with the 12.7-and 25.4-mm thresholds delineated with black contours. Red and purple contours represent 12.7- and 25.4-mm accumulations, respectively, of QPFs for (a) 3 km–50 lev Thompson MP, (b) 1 km–50 lev Thompson MP, (c) 3 km–100 lev Thompson MP, (d) 1 km–100 lev Thompson MP, (e) 3 km–50 lev WSM6 MP, (f) 1 km–50 lev WSM6 MP, (g) 3 km–100 lev WSM6 MP, and (h) 1 km–100 lev WSM6 MP simulations. The abscissa and ordinate values represent distance in km.

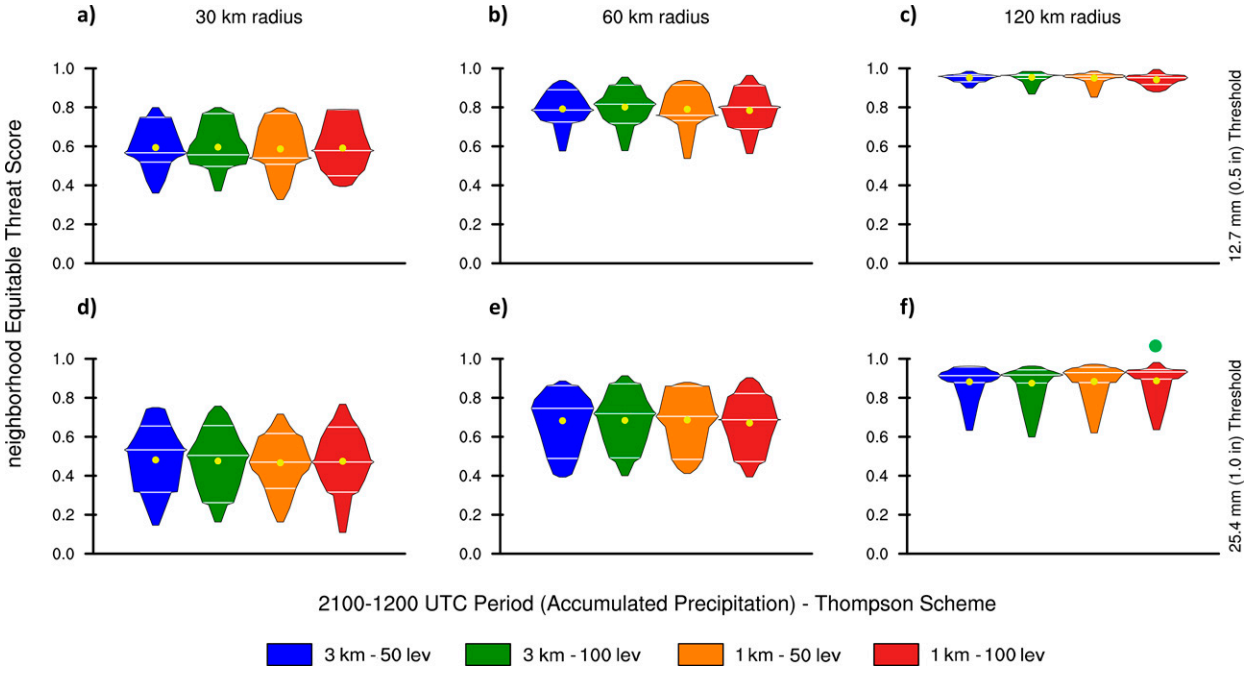


FIG. 10. As in Fig. 7, but for the 2100–1200 UTC accumulated precipitation swaths.

noticeably greater coverage compared to all 3-km runs (Fig. 13c). Also, 1 km–50 lev runs had precipitation regions that intersected or shared a union with the observed clusters noticeably to significantly more than in all 3-km runs (Figs. 13d,e). 1 km–50 lev runs also showed the greatest similarity in 50th and 90th percentile intensities (highest precipitation accumulations) to observations (Figs. 13g,h).

The 1 km–50 lev magnitudes of the highest accumulations were also significantly more similar to observations compared to 3 km–100 lev magnitudes at both the 50th and 90th percentiles, suggesting that decreasing Δx to 1 km significantly improved QPFs for the highest accumulations, even with coarser Δz. When reviewing the 3-h nETSs (Fig. 8) and nETSs for the 2100–1200 UTC accumulation period (Fig. 11), 1 km–50 lev seemed to be the most skillful forecasts when employing WSM6 MP, and the MODE statistical analyses supported these results, with 1 km–50 lev runs sharing the greatest areal coverage with observations, with the most similar magnitudes of the higher accumulations. In contrast, 3 km–100 lev simulations showed some of the greatest differences between forecast and observed accumulations (particularly those at the 90th percentile), being noticeably to significantly less skillful to not only 1-km runs, but also their 3 km–50 lev counterparts. Given earlier noted struggles with MCSs in 1 km–100 lev runs seemingly moving too fast and dissipating too early, decreases in Δz seemed to lead to degraded precipitation forecasts in WSM6 MP simulations. In Part I of this study, MCS cold pools were often most intense in magnitude and depth in 3 km–100 lev for both Thompson and WSM6 MP schemes. At least for WSM6 MP simulations, it may be deduced that the generation of rainwater (and resultant cold pool development) is too intense with 100 vertical levels.

4. Summary and conclusions

Changes in Δx or Δz did not significantly improve or degrade forecast MCS placement relative to observations at any given time of MCS evolution, regardless of the MP scheme used, though 1-km simulated MCSs generally moved noticeably to significantly faster than observed events, especially later in the MCS life cycle. However, Thompson MP simulated MCSs were typically displaced in a north to south orientation relative to observations, and this north–south spatial displacement only grew larger with time among all Δx and Δz grid spacings. Despite this increasing north–south displacement with time among all Thompson MP members, the finer Δx runs showed significant QPF skill improvement toward the later stages of MCS evolution, which likely contributed to significant QPF skill improvement between 3- and 1-km 2100–1200 UTC accumulated precipitation swaths when 100 vertical levels were employed. Based on the MODE statistical analyses, 2100–1200 UTC simulated swaths at finer grid spacing had more overlapping area of 12.7-mm accumulated precipitation with that of observations and simulated higher accumulations more representative of the observed 2100–1200 UTC accumulated swaths.

WSM6 MP simulated MCSs were displaced downstream of observations starting at initiation, with the spatial displacement errors of all Δx and Δz grid spacings increasing with time. The greater simulated MCS speeds compared to observations a few hours after MCS initiation likely contributed to the increased spatial displacements later in MCS evolution. Finer Δx simulations significantly improved 3-h QPFs at the early stages of WSM6 MP simulated MCSs; however, the 2100–1200 UTC accumulated precipitation based QPFs did

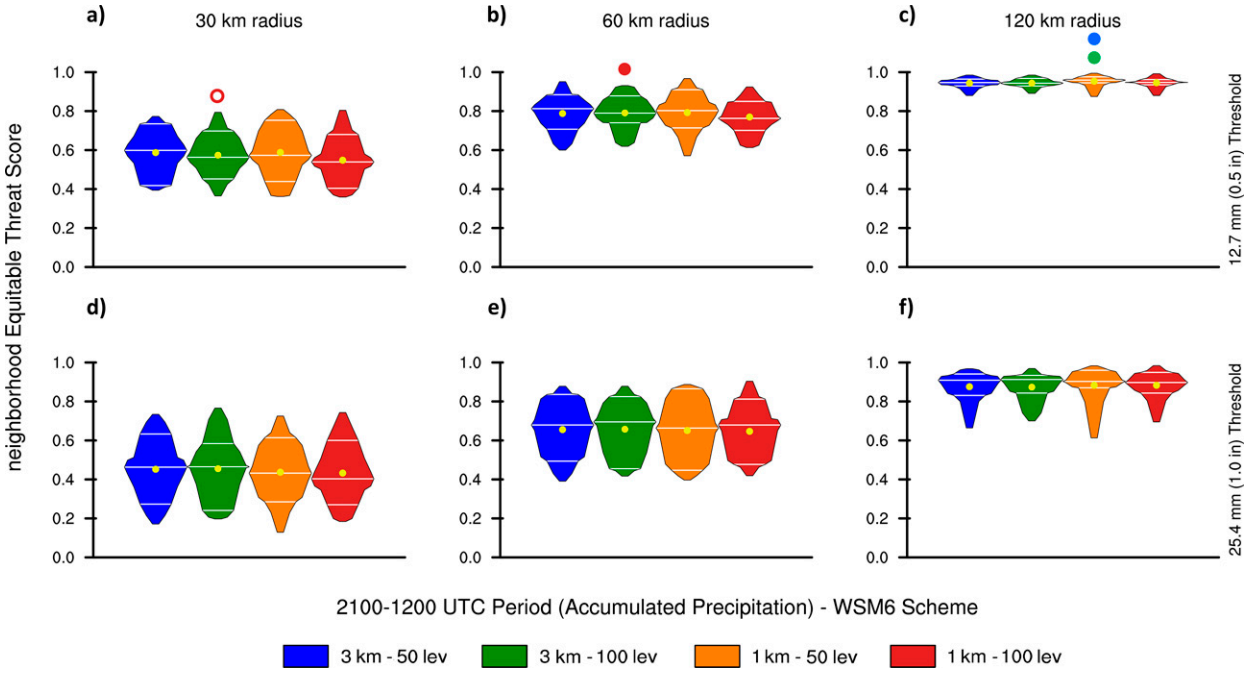


FIG. 11. As in Fig. 10, but for the WSM6 MP scheme.

not significantly improve with finer Δx. When comparing the distribution (shapes of the violin plots) of spatial displacement errors of QPFs with grid spacings holding Δx constant, noticeable changes were seldom observed between runs of varying Δz. The most noticeable difference was 50 vertical levels showing significant 12.7-mm threshold QPF improvement compared to 100 vertical levels for 1-km WSM6 MP simulations toward the end of MCS evolution. This significant improvement did not heavily contribute to the improvement of the 2100–1200 UTC accumulated precipitation swath. However, analysis of MODE generated statistics of the spatial characteristics and distribution of the 2100–1200 UTC simulated QPFs revealed that 1 km–50 lev runs shared the greatest areal overlap with observed MCS swaths, with the most accurate distribution of the accumulations at the 50th and 90th percentiles above 12.7 mm.

Even though Δz decreases seemed to slightly enhance QPF skill for the Thompson MP simulations, refining Δz in WSM6 MP simulations served more as a detriment to forecast skill, with MCSs potentially moving too fast or weakening too early compared to reality. Still, given the lack of widespread significant improvements or degradation of MCS spatial placements and QPF improvements for the entire MCS precipitation swath with changing Δx and Δz, future work may want to concentrate on changes in physics schemes to possibly have a bigger impact on numerical MCS forecasts.

5. Discussion

Similar to Squitieri and Gallus (2020), Part I and II of this research have shown that refining Δx from 3 down to 1 km supports the development of stronger and larger cold pools but does not significantly influence the spatial placement and timing of simulated MCSs. Kain et al. (2008) showed that 2- and 4-km WRF precipitation and reflectivity fields demonstrated greater similarity to each other than when either simulated field was compared to observations. Similar observations were made by forecasters during the 2010 HWT experiment (Clark et al. 2012). Part I of this research showed that 1-km simulated MCSs and their cold pools were more expansive than their 3-km counterparts mainly due to a greater coverage of updrafts along the leading line in both Thompson and WSM6 MP runs, which may also explain the greater overlap of QPEs and 1-km QPFs here in Part II of this research. Still, the results in the present study roughly agree with Kain et al. (2008) and Clark et al. (2012) that decreasing Δx to spacing finer than 4 km would provide limited added value to MCS placement forecasts and resultant QPFs while greatly increasing computational costs. Given the limited impact decreasing Δx had on simulated versus observed MCS placements at given times, and modest differences in placement of 2100–1200 UTC accumulated QPFs versus QPEs, the authors agree with Kain et al. (2008) and Clark et al. (2012), that refining Δx past 3–4 km will likely yield a limited improvement of QPFs for larger-scale deep convective events.

While 1-km simulations had at times significantly improved MCS QPF skill scores compared to 3-km runs in the present study, the improvements must be evaluated in context of the experimental setup. Case-to-case variability may have played a role in convective QPF improvements or degradations when altering Δx, as in either the convection-parameterized study of Gallus (1999), or the convection-allowing deterministic work of Squitieri and Gallus (2020). The behavior in nETS trends also varied considerably between Thompson and WSM6 MP runs. Thompson MP 1-km runs significantly

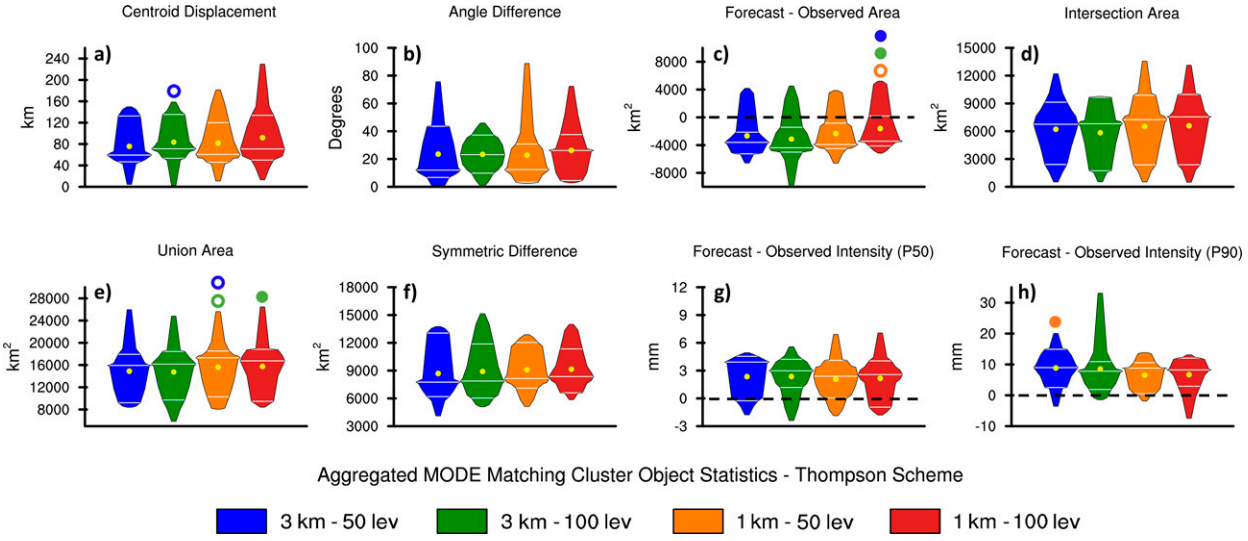


FIG. 12. Violin plots of aggregated MODE-derived statistics of matching cluster objects representative of the 2100–1200 UTC 12.7+ mm threshold precipitation accumulation swaths between observations and 3 km–50 lev (blue), 3 km–100 lev (green), 1 km–50 lev (orange), 1 km–100 lev (red) simulations, for all 10 cases studied. All simulations employed Thompson MP. Statistical significance is displayed in the same manner as in Fig. 8. The statistics displayed are (a) total distance spatial displacement (km) of the simulated precipitation swath centroid relative to the observed precipitation swath centroid; (b) absolute value of the difference in orientation angles (in degrees) between the observed and simulated precipitation swaths; (c) difference in total areal coverage of the simulated vs observed precipitation swaths (forecast minus observed; km^2); (d) total areal overlap (km^2) of the observed and simulated precipitation swaths; (e) total area (km^2) covered by either the observed or simulated precipitation swaths; (f) area of simulated precipitation (km^2) not covered by observations; (g) forecast minus observed precipitation accumulation values exceeding 50th percentile (mm); and (h) as in (g), but for the 90th percentile. The dashed line in (c), (g), and (h) delineates the "0" *y*-axis value. Yellow dots on the violins are the averages and the white horizontal lines delineate the 25th, 50th, and 75th percentiles.

improved QPFs over 3 km later in the MCS life cycle while WSM6 MP 1-km runs significantly improved QPFs over 3-km runs earlier on in MCS evolution, with some QPF degradation introduced when increasing vertical resolution of 1-km runs using 100 vertical levels. These differences in 3-h QPF trends between the single-moment WSM6 and part-two-moment Thompson MP schemes suggest that the choice of MP scheme employed may impact solutions as much as different model grid spacings (Potvin et al. 2017), including changes in Δx or Δz. Thielen and Gallus (2019) evaluated nocturnal MCS morphology in weakly forced environments and also found major differences in performance between the Thompson and WSM6 MP schemes. Just like 3-h QPF nETSs for 1 km–100 lev WSM6 MP runs degraded toward the end of MCS evolution in the present work, Thielen and Gallus (2019) showed that the morphology accuracy score for 1-km MCS simulations later in evolution was lower than 3-km when employing the WSM6 MP scheme. On the other hand, toward the end of MCS evolution, 1-km Thompson MP simulations outperformed 3-km events in Thielen and Gallus (2019), in much the same way that 1-km 3-h QPFs significantly improved over 3-km QPFs toward the end of MCS life cycles using the Thompson MP scheme.

As alluded to in Thielen and Gallus (2019), Thompson MP simulations tend to outperform single-moment MP schemes like WSM6 because of their production of greater snow mixing ratios in place of graupel at high altitudes which in turn favors more stratiform precipitation. Bao et al. (2019) also stated that the Thompson MP scheme fosters smaller frozen hydrometeors, resulting in more efficient maintenance of snow. It was found in Part I of this research that melting snow was a dominant source of rainwater production and latent cooling, especially later during MCS evolution in the finer Δz Thompson MP runs. Better representation of the upward and rearward advection of snow and subsequent larger generation of rainwater production near the melting layer with finer model grid spacing may explain the improved QPF skill scores later in MCS evolution. As noted in previous studies, MCS stratiform rain may account for up to 40% of total precipitation accumulation within a midlatitude MCS (Johnson and Hamilton 1988; Watson et al. 1988; Johnson et al. 1989; Gallus and Johnson 1995), highlighting the importance of properly resolving the MCS stratiform region. As such, the Thompson MP scheme is often preferred for simulating MCSs given the scheme's ability to simulate the stratiform region well compared to other schemes, including the full double-moment Morrison scheme (Feng et al. 2018).

On the other hand, Bao et al. (2019) mentioned that WSM6 MP frozen hydrometeor sizes and fall speeds were larger than Thompson MP, allowing for more graupel production. In Part I of this research, graupel melting was among the dominant contributors to latent cooling and cold pool production, with relatively higher amounts of graupel melting and stronger cold pools noted with 100 vertical levels runs compared to 50 vertical

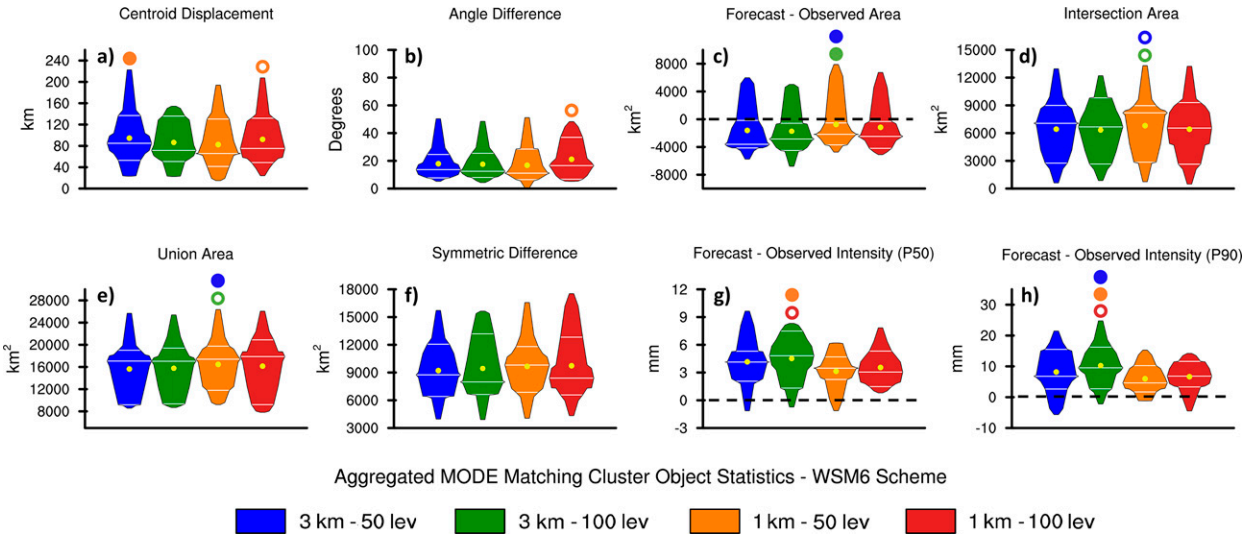


FIG. 13. As in Fig. 12, but for the WSM6 MP scheme.

levels. It is possible that WSM6 MP simulated cold pools were too strong compared to reality in runs with finer grid spacing, with excess graupel loading leading to cold pools intensifying too quickly. This may be why finer Δx and Δz WSM6 MP simulated MCSs moved too quickly early on in their life cycles and in turn, weakened too early compared to observations, which could have resulted in the poorer QPF skill scores later in the MCS life cycle.

Acknowledgments. This research was sponsored by the National Science Foundation (NSF) under Grant AGS-2022888. The authors would like to acknowledge NCAR's Computational and Information Systems Laboratory, sponsored by the National Science Foundation, for providing high-performance computing support via the Cheyenne supercomputing system (doi:10.5065/D6RX99HX). Stage-IV data were provided by NCAR/EOL under sponsorship of the National Science Foundation: http://data.eol.ucar.edu/. The authors would also like to thank Dr. Adam Clark (NSSL) for providing code to compute neighborhood-based equitable threat scores. The authors also extend gratitude to the editor and three anonymous reviewers for their careful review of this manuscript, and for providing several insightful comments that have resulted in major improvements of this research.

Data availability statement. The WRF output that supports the findings of this work are available in DataShare, Iowa State's open research data repository. All other data availability locations are as defined in-text and within the references of this manuscript.

REFERENCES

Ahijevych, D., J. O. Pinto, J. K. Williams, and M. Steiner, 2016: Probabilistic forecasts of mesoscale convective system initiation using the random forest data mining technique. *Wea. Forecasting*, **31**, 581–599, https://doi.org/10.1175/WAF-D-15-0113.1.

Aligo, E. A., W. A. Gallus Jr., and M. Segal, 2009: On the impact of WRF Model vertical grid resolution on Midwest summer rainfall forecasts. *Wea. Forecasting*, **24**, 575–594, https://doi.org/10.1175/2008WAF2007101.1.

Bao, J.-W., S. A. Michelson, and E. D. Grell, 2019: Microphysical process comparison of three microphysics parameterization schemes in the WRF Model for an idealized squall-line case study. *Mon. Wea. Rev.*, **147**, 3093–3120, https://doi.org/10.1175/MWR-D-18-0249.1.

Clark, A. J., W. A. Gallus Jr., M. Xue, and F. Kong, 2009: A comparison of precipitation forecast skill between small near-convection-permitting and large convection-parameterizing ensembles. *Wea. Forecasting*, **24**, 1121–1140, https://doi.org/10.1175/2009WAF2222222.1.

——, ——, and M. L. Weisman, 2010a: Neighborhood-based verification of precipitation forecasts from convection-allowing NCAR WRF Model simulations and the operational NAM. *Wea. Forecasting*, **25**, 1495–1509, https://doi.org/10.1175/2010WAF2222404.1.

——, ——, M. Xue, and F. Kong, 2010b: Convection-allowing and convection-parameterizing ensemble forecasts of a mesoscale convective vortex and associated severe weather environment. *Wea. Forecasting*, **25**, 1052–1081, https://doi.org/10.1175/2010WAF2222390.1.

——, and Coauthors, 2012: An overview of the 2010 Hazardous Weather Testbed Experimental Forecast Program Spring Experiment. *Bull. Amer. Meteor. Soc.*, **93**, 55–74, https://doi.org/10.1175/BAMS-D-11-00040.1.

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

Clarke, S. J., S. L. Gray, and N. M. Roberts, 2019: Downstream influence of mesoscale convective systems. Part 2: Influence on ensemble forecast skill and spread. *Quart. J. Roy. Meteor. Soc.*, **145**, 2953–2972, https://doi.org/10.1002/qj.3613.

CISL, 2020: Cheyenne: HPE/SGI ICE XA System (University Community Computing). National Center for Atmospheric Research, accessed 13 June 2022, https://doi.org/10.5065/D6RX99HX.

Coniglio, M. C., H. E. Brooks, S. J. Weiss, and S. F. Corfidi, 2007: Forecasting the maintenance of quasi-linear mesoscale convective systems. *Wea. Forecasting*, **22**, 556–570, https://doi.org/10.1175/WAF1006.1.

Davis, C., B. Brown, and R. Bullock, 2006a: Object-based verification of precipitation forecasts. Part I: Methods 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.

Done, J., C. A. Davis, and M. L. Weisman, 2004: The next generation of NWP: Explicit forecasts of convection using the Weather Research and Forecasting (WRF) Model. *Atmos. Sci. Lett.*, **5**, 110–117, https://doi.org/10.1002/asl.72.

Feng, Z., L. R. Leung, R. A. Houze Jr., S. Hagos, J. Hardin, Q. Yang, B. Han, and J. Fan, 2018: Structure and evolution of mesoscale convective systems: Sensitivity to cloud microphysics in convection-permitting simulations over the United States. *J. Adv. Model. Earth Syst.*, **10**, 1470–1494, https://doi.org/10.1029/2018MS001305.

Gallus, W. A., Jr., 1999: Eta simulations of three extreme precipitation events: Sensitivity to resolution and convective parameterization. *Wea. Forecasting*, **14**, 405–426, https://doi.org/10.1175/1520-0434(1999)014<0405:ESOTEP>2.0.CO;2.

——, and R. H. Johnson, 1995: The dynamics of circulations within the trailing stratiform region of squall lines. Part I: The 10–11 June pre-storm system. *J. Atmos. Sci.*, **52**, 2161–2187, https://doi.org/10.1175/1520-0469(1995)052<2161:TDOCWT>2.0.CO;2.

——, N. A. Snook, and E. V. Johnson, 2008: Spring and summer severe weather reports over the Midwest as a function of convective mode: A preliminary study. *Wea. Forecasting*, **23**, 101–113, https://doi.org/10.1175/2007WAF2006120.1.

Grams, J. S., W. A. Gallus Jr., S. E. Koch, L. S. Wharton, A. Loughe, and E. E. Ebert, 2006: The use of a modified Ebert–McBride technique to evaluate mesoscale model QPF as a function of convective system morphology during IHOP 2002. *Wea. Forecasting*, **21**, 288–306, https://doi.org/10.1175/WAF918.1.

Haberlie, A. M., and W. S. Ashley, 2019: A radar-based climatology of mesoscale convective systems in the United States. *J. Climate*, **32**, 1591–1606, https://doi.org/10.1175/JCLI-D-18-0559.1.

Hocker, J. E., and J. B. Basara, 2008: A 10-year spatial climatology of squall line storms across Oklahoma. *Int. J. Climatol.*, **28**, 765–775, https://doi.org/10.1002/joc.1579.

Iowa Environmental Mesonet, 2020: Documentation on IEM generated NEXRAD Mosaics. Iowa State University, accessed 22 October 2020, https://mesonet.agron.iastate.edu/docs/nexrad_composites/.

Jankov, I., and W. A. Gallus Jr., 2004a: Some contrasts between good and bad forecasts of warm season MCS rainfall. *J. Hydrol.*, **288**, 122–152, https://doi.org/10.1016/j.jhydrol.2003.11.013.

——, and ——, 2004b: MCS rainfall forecast accuracy as a function of large-scale forcing. *Wea. Forecasting*, **19**, 428–439, https://doi.org/10.1175/1520-0434(2004)019<0428:MRFAAA>2.0.CO;2.

Jirak, I. L., and W. R. Cotton, 2007: Observational analysis and predictability of mesoscale convective systems. *Wea. Forecasting*, **22**, 813–838, https://doi.org/10.1175/WAF1012.1.

——, ——, and R. L. McAnelly, 2003: Satellite and radar survey of mesoscale convective system development. *Mon. Wea. Rev.*, **131**, 2428–2449, https://doi.org/10.1175/1520-0493(2003)131<2428:SARSOM>2.0.CO;2.

Johnson, R. H., and P. J. Hamilton, 1988: The relationship of surface pressure features to the precipitation and airflow structure of an intense midlatitude squall line. *Mon. Wea. Rev.*, **116**, 1444–1473, https://doi.org/10.1175/1520-0493(1988)116<1444:TROSPF>2.0.CO;2.

——, S. Chen, and J. J. Toth, 1989: Circulations associated with a mature-to-decaying midlatitude mesoscale convective system. Part I: Surface features—Heat bursts and mesolow development. *Mon. Wea. Rev.*, **117**, 942–959, https://doi.org/10.1175/1520-0493(1989)117<0942:CAWAMT>2.0.CO;2.

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

Lawson, J. R., and W. A. Gallus Jr., 2016: On contrasting ensemble simulations of two Great Plains bow echoes. *Wea. Forecasting*, **31**, 787–810, https://doi.org/10.1175/WAF-D-15-0060.1.

——, ——, and C. K. Potvin, 2020: Sensitivity of a bowing mesoscale convective system to horizontal grid spacing in a convection-allowing ensemble. *Atmosphere*, **11**, 384, https://doi.org/10.3390/atmos11040384.

Mendenhall, W., and T. Sinchich, 2007: *Statistics for Engineering and the Sciences*. 5th ed. Pearson-Prentice Hall, 1060 pp.

NCAR/UCAR/EOL, 2015: NCEP/EMC 4KM Gridded Data (GRIB) Stage IV data. Subset used: June 2015-July 2020, Earth Observing Laboratory, accessed 6 October 2020, http://data.eol.ucar.edu/cgi-bin/codiac/fgr_form/id=21.093.

Peters, J. M., E. R. Nielsen, M. D. Parker, S. M. Hitchcock, and R. S. Schumacher, 2017: The impact of low-level moisture errors on model forecasts of an MCS observed during PECAN. *Mon. Wea. Rev.*, **145**, 3599–3624, https://doi.org/10.1175/MWR-D-16-0296.1.

Potvin, C. K., E. M. Murillo, M. L. Flora, and D. M. Wheatley, 2017: Sensitivity of supercell simulations to initial-condition resolution. *J. Atmos. Sci.*, **74**, 5–26, https://doi.org/10.1175/JAS-D-16-0098.1.

Schwartz, C. S., and Coauthors, 2009: Next-day convection-allowing WRF Model guidance: A second look at 2-km versus 4-km grid spacing. *Mon. Wea. Rev.*, **137**, 3351–3372, https://doi.org/10.1175/2009MWR2924.1.

——, G. S. Romine, K. R. Fossell, R. A. Sobash, and M. L. Weisman, 2017: Toward 1-km ensemble forecasts over large domains. *Mon. Wea. Rev.*, **145**, 2943–2969, https://doi.org/10.1175/MWR-D-16-0410.1.

Skamarock, W. C., and Coauthors, 2019: A description of the Advanced Research WRF Model version 4. NCAR Tech. Note NCAR/TN-556+STR, 145 pp., https://doi.org/10.5065/1dfh-6p97.

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.

Squitieri, B. J., and W. A. Gallus Jr., 2016a: WRF forecasts of Great Plains nocturnal low-level jet-driven MCSs. Part I: Correlation between low-level jet forecast accuracy and MCS precipitation forecast skill. *Wea. Forecasting*, **31**, 1301–1323, https://doi.org/10.1175/WAF-D-15-0151.1.

——, and ——, 2016b: WRF forecasts of Great Plains nocturnal low-level jet-driven MCSs. Part II: Differences between strongly and weakly forced low-level jet environments. *Wea. Forecasting*, **31**, 1491–1510, https://doi.org/10.1175/WAF-D-15-0150.1.

——, and ——, 2020: On the forecast sensitivity of MCS cold pools and related features to horizontal grid spacing in convection-allowing WRF simulations. *Wea. Forecasting*, **35**, 325–346, https://doi.org/10.1175/WAF-D-19-0016.1.

——, and ——, 2022: On the changes in convection-allowing WRF forecasts of MCS evolution due to decreases in model horizontal and vertical grid spacing. Part I: Changes in cold pool evolution. *Wea. Forecasting*, **37**, 1903–1923, https://doi.org/10.1175/WAF-D-22-0041.1.

Stensrud, D. J., and J. M. Fritsch, 1993: Mesoscale convective systems in weakly forced large-scale environments. Part I: Observations. *Mon. Wea. Rev.*, **121**, 3326–3344, https://doi.org/10.1175/1520-0493(1993)121<3326:MCSIWF>2.0.CO;2.

——, and ——, 1994: Mesoscale convective systems in weakly forced large-scale environments. Part III: Numerical simulations and implications for operational forecasting. *Mon. Wea. Rev.*, **122**, 2084–2104, https://doi.org/10.1175/1520-0493(1994)122<2084:MCSIWF>2.0.CO;2.

Thielen, J. E., and W. A. Gallus Jr., 2019: Influences of horizontal grid spacing and microphysics on WRF forecasts of convective morphology evolution for nocturnal MCSs in weakly forced environments. *Wea. Forecasting*, **34**, 1495–1517, https://doi.org/10.1175/WAF-D-18-0210.1.

Trier, S. B., J. H. Marsham, C. A. Davis, and D. A. Ahijevych, 2011: Numerical simulations of the postsunrise reorganization of a nocturnal mesoscale convective system during 13 June IHOP_2002. *J. Atmos. Sci.*, **68**, 2988–3011, https://doi.org/10.1175/JAS-D-11-0112.1.

Vertz, N. J., W. A. Gallus Jr., and B. J. Squitieri, 2021: Can pre-storm errors in the low-level inflow help predict spatial displacement errors in MCS initiation? *Atmosphere*, **12**, https://doi.org/10.3390/atmos12010007.

Wandishin, M. S., D. J. Stensrud, S. L. Mullen, and L. J. Wicker, 2008: On the predictability of mesoscale convective systems: Two-dimensional simulations. *Wea. Forecasting*, **23**, 773–785, https://doi.org/10.1175/2008WAF2007057.1.

——, ——, ——, and ——, 2010: On the predictability of mesoscale convective systems: Three-dimensional simulations. *Mon. Wea. Rev.*, **138**, 863–885, https://doi.org/10.1175/2009MWR2961.1.

Watson, A. I., J. Meitín, and J. B. Cunning, 1988: Evolution of the kinematic structure and precipitation characteristics of a mesoscale convective system on 20 May 1979. *Mon. Wea. Rev.*, **116**, 1555–1567, https://doi.org/10.1175/1520-0493(1988)116<1555:EOTKSA>2.0.CO;2.

Weisman, M. L., C. Davis, W. Wang, K. W. Manning, and J. B. Klemp, 2008: Experiences with 0–36-h explicit convective forecasts with the WRF-ARW Model. *Wea. Forecasting*, **23**, 407–437, https://doi.org/10.1175/2007WAF2007005.1.