

## Sensitivity of Idealized Supercell Simulations to Horizontal Grid Spacing: Implications for Warn-on-Forecast

COREY K. POTVIN

*Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, and NOAA/OAR/National Severe Storms Laboratory, Norman, Oklahoma*

MONTGOMERY L. FLORA

*National Weather Center Research Experience for Undergraduates, Norman, Oklahoma, and Ball State University, Muncie, Indiana*

(Manuscript received 19 December 2014, in final form 19 March 2015)

### ABSTRACT

The Warn-on-Forecast (WoF) program aims to deploy real-time, convection-allowing, ensemble data assimilation and prediction systems to improve short-term forecasts of tornadoes, flooding, lightning, damaging wind, and large hail. Until convection-resolving (horizontal grid spacing  $\Delta x < 100$  m) systems become available, however, resolution errors will limit the accuracy of ensemble model output. Improved understanding of grid spacing dependence of simulated convection is therefore needed to properly calibrate and interpret ensemble output, and to optimize trade-offs between model resolution and other computationally constrained parameters like ensemble size and forecast lead time.

Toward this end, the authors examine grid spacing sensitivities of simulated supercells over  $\Delta x$  of 333 m–4 km. Storm environment and physics parameterization are varied among the simulations. The results suggest that 4-km grid spacing is too coarse to reliably simulate supercells, occasionally leading to premature storm demise, whereas 3-km simulations more often capture operationally important features, including low-level rotation tracks. Further decreasing  $\Delta x$  to 1 km enables useful forecasts of rapid changes in low-level rotation intensity, though significant errors remain (e.g., in timing).

Grid spacing dependencies vary substantially among the experiments, suggesting that accurate calibration of ensemble output requires better understanding of how storm characteristics, environment, and parameterization schemes modulate grid spacing sensitivity. Much of the sensitivity arises from poorly resolving small-scale processes that impact larger (well resolved) scales. Repeating some of the 333-m simulations with coarsened initial conditions reveals that supercell forecasts can substantially benefit from reduced grid spacing even when limited observational density precludes finescale initialization.

### 1. Introduction

Real-time ensemble data assimilation and prediction of storms on convection-allowing (e.g., grid spacing  $\Delta x \leq 4$  km) grids is close to fruition. This is expected to enable the transition to a Warn-on-Forecast (WoF; [Stensrud et al. 2009, 2013](#)) paradigm, in which National Weather Service forecasters use novel model guidance to improve short-term (e.g., 0–2 h) forecasts of tornadoes, flooding, lightning, and damaging wind and hail.

Until convection-resolving (e.g.,  $\Delta x < 100$  m; [Bryan et al. 2003](#)) systems become available, however, resolution errors will impose practical predictability limits on storm hazards. Improved understanding of grid spacing sensitivities of simulated convection is therefore needed to determine model resolution requirements for particular storm features and scales, and to optimize trade-offs between model resolution and other computationally constrained parameters such as ensemble size, parameterization scheme complexity, and forecast lead time. Calibrating and interpreting ensemble output likewise require knowledge of grid spacing dependencies.

Grid spacing sensitivity of simulated deep moist convection has been systematically examined in several previous studies. [Weisman et al. \(1997\)](#) found that the

*Corresponding author address:* Dr. Corey K. Potvin, National Severe Storms Laboratory, National Weather Center, 120 David L. Boren Blvd., Norman, OK 73072.  
E-mail: corey.potvin@noaa.gov

evolution of squall lines simulated at  $\Delta x = 1$  km was qualitatively reproduced on a 4-km grid, a result they attributed to the primary squall-line circulation being largely hydrostatic. [Bryan et al. \(2003\)](#), however, demonstrated major differences between the convective organization and system-wide vertical fluxes of squall lines simulated with 1-km versus 125-m grid spacing. They ascribed the differences largely to the increasing turbulence and associated explicit entrainment and overturning within clouds at finer resolutions. Moreover, trends of mean vertical velocity and rainwater mixing ratio magnitude were reversed between strong-shear and weak-shear simulations, indicating strong environmental sensitivity of grid spacing dependence. In squall-line simulations by [Bryan and Morrison \(2012\)](#), cloud water evaporation decreased as grid spacing increased from 250 m to 1 km, increasing total surface precipitation. This was attributed to reduced entrainment at coarser resolution due to correspondingly decreased cloud surface area and less-resolved turbulence. When grid spacing was further increased from 1 to 4 km, reduction in cloud water evaporation was largely compensated by decreased total condensation (i.e., the precipitation efficiency did not substantially change), resulting in little difference in precipitation. Other instances of nonlinear grid spacing dependence occurred; for example, maximum updraft speed increased from  $\Delta x = 250$  m to  $\Delta x = 1$  km and then substantially decreased from  $\Delta x = 1$  km to  $\Delta x = 4$  km.

Supercell simulation experiments have in many instances exhibited very different grid spacing dependence than the [Bryan and Morrison \(2012\)](#) squall-line simulations. For example, in the supercell–multicell simulations of [Verrelle et al. \(2015\)](#), total precipitation decreased monotonically as grid spacing increased. [Verrelle et al. \(2015\)](#) attributed this difference to their use of an initialization sounding with much moister midtropospheric air, and consequently reduced grid spacing sensitivity of cloud water evaporation, than that used in [Bryan and Morrison \(2012\)](#). In contrast with both [Bryan and Morrison \(2012\)](#) and [Verrelle et al. \(2015\)](#), updraft intensity decreased as grid spacing was increased from 500 m to 1 km in the supercell simulations of [Fiori et al. \(2010\)](#), [Adler et al. \(2002\)](#), and [Noda and Niino \(2003\)](#), and further decreased when grid spacing was increased to 2 km in the latter two studies.

The differences among the aforementioned studies are perhaps unsurprising given the many sources of grid spacing sensitivity, examples of which follow. As already mentioned, as grid spacing increases, entrainment and cloud evaporation tend to decrease, thus, increasing precipitation efficiency and favoring increased rainfall.

Depending on how the updraft changes with grid spacing, however, this effect can be compensated or compounded with decreased or increased (respectively) condensation. Marginally resolved updrafts tend to be weaker due not only to spatial filtering but also to reduction of nonhydrostatic effects such as linear (vertical shear induced) dynamic forcing (e.g., [Weisman et al. 1997](#); [Bryan and Morrison 2012](#)). These tendencies, however, can be offset by underestimated entrainment. Coarse resolution can delay storm evolution due to flow becoming increasingly hydrostatic (e.g., [Bryan and Morrison 2012](#)), or accelerate storm demise due to insufficient dynamic updraft forcing (e.g., [Noda and Niino 2003](#)). The sum of all these effects in any given simulation is likely strongly linked to storm characteristics (e.g., mode, updraft width), environment and, as will now be discussed, model physics.

As [Bryan and Morrison \(2012\)](#) point out, subgrid schemes may be tuned (deliberately or otherwise) to compensate for errors due to marginal grid resolution. New errors may therefore arise when such schemes are applied at finer resolution. More generally, parameterization schemes can have undesirable effects when applied outside the (typically relatively narrow) range of resolutions for which they were designed. For example, in the simulations of [Adler et al. \(2002\)](#) and [Verrelle et al. \(2015\)](#), the ratio between resolved and total (resolved plus subgrid) turbulent kinetic energy (TKE) *decreased* with increasing horizontal resolution. Both studies ascribed this effect to deficient subgrid dynamical TKE production on coarser grids owing to poorly resolved wind gradients. Underproduction of subgrid TKE and consequent underestimation of entrainment may result in overly strong updrafts and downdrafts at coarser resolution. [Verrelle et al. \(2015\)](#) attributed the unexpected increase in updraft intensity with grid spacing (over the range  $\Delta x = 0.5$ –2.0 km) in their simulations to this effect. These results demonstrate that complete understanding of convective grid spacing sensitivity requires knowledge of the grid spacing dependencies of physics.

Improved understanding of the impacts of individual forecast error sources, and therefore of their relative contributions to the total forecast error, is needed to judiciously prioritize forecast improvement efforts. In addition, knowledge of the resolution requirements for particular convective processes or features of interest to be realistically simulated is vital for numerical investigations of storms. The present work, as in the aforementioned studies, therefore adopts a simulation framework in order to isolate grid spacing effects from other error sources that are unavoidable in real-world

TABLE 1. WRF-ARW 3.4.1 settings used in all experiments unless otherwise noted.

Horizontal grid size	WK82: 661 × 661; DelCity: 451 × 451; ElReno: 901 × 901
Large/small time step	Large: 1 s; small: 1/6 s
Time-integration scheme	Third-order Runge–Kutta
Horizontal/vertical advection	Fifth-order/third-order
Lateral/top boundary conditions	Open/Rayleigh damping layer
Turbulence parameterization	1.5-order TKE closure
Microphysics parameterization	Thompson scheme
Coriolis effect	Off
Surface drag	Off
Radiation parameterization	Off
Cumulus parameterization	Off
Surface layer physics	Off
Planetary boundary layer parameterization	Off
Explicit numerical diffusion	Off
Horizontal/vertical bubble radii (control experiments only)	10 km/1.5 km

forecasts. Since the present work is primarily intended to guide the design of WoF ensemble systems, our approach differs from previous supercell grid spacing-dependence studies in several important ways. First, we initialize our coarser-resolution simulations not from a thermal bubble as in the control (finest scale) simulations, but from appropriately upscaled (to remove features unresolvable on the coarse grid) versions of the control simulations after 30 min of integration. This approach acknowledges that ensemble forecasts in the WoF paradigm will be initiated only after enough radar data (e.g., 5–10 volumes) have been assimilated for forecasts to be skillful. An incidental benefit of this approach is that some of the unrealistic thermodynamic and kinematic features associated with the thermal bubble have dissipated by the time the coarse simulations are initialized. Second, we place greater focus than previous studies on supercell features most relevant to  $O(1)$  h convective forecast and warning operations, such as low-level rotation tracks. Third, we perform experiments with a range of control simulations designed to assess the environmental sensitivity of grid spacing dependencies as well as the impact of microphysical and turbulence parameterization scheme choices. The simulations use homogeneous initial conditions in order to better identify relationships between storm environment and grid spacing dependence; future investigation is required to examine how the dependencies are modulated by horizontal heterogeneities known to impact supercell evolution (e.g., Richardson et al. 2007; Ziegler et al. 2010). Finally, we compare our coarser-resolution simulations to both the control simulations and filtered versions thereof. This allows us to assess how much of the total error arises from poor resolution of small-scale processes that interact with larger scales

(resolvable on the coarse grid) in the control simulation. Since interscale interactions can vary considerably between storms, and neglecting them can produce spatiotemporally complex effects (i.e., not just in magnitude), the errors that thereby arise cannot be reliably corrected using simple  $\Delta x$ -based scaling. Such errors therefore substantially complicate forecast calibration, and have significant methodological implications for numerical investigations of storm processes.

In summary, this work addresses three primary questions:

- 1) How accurately can important supercell features, especially those most useful to forecasting convective hazards  $O(1)$  h in advance, be simulated at different model grid spacings?
- 2) Do storm environment and parameterization scheme choices strongly modulate the grid spacing dependence of supercells?
- 3) How much does the grid spacing sensitivity of supercells arise from neglecting smaller scales that would otherwise have impacted larger (resolved) scales?

The rest of the paper is organized as follows. Section 2 details the model configuration used for our supercell simulations, including the soundings used in the three primary control experiments, along with the procedure for initializing the coarser-resolution simulations from the control simulation. Section 3 describes the grid spacing dependencies of both general and finescale storm features in the simulations, their sensitivity to storm environment and physics schemes, and the relative impact of missing scales in the initial condition. Section 4 summarizes the results, discusses their implications for WoF, and concludes with recommendations for future work.

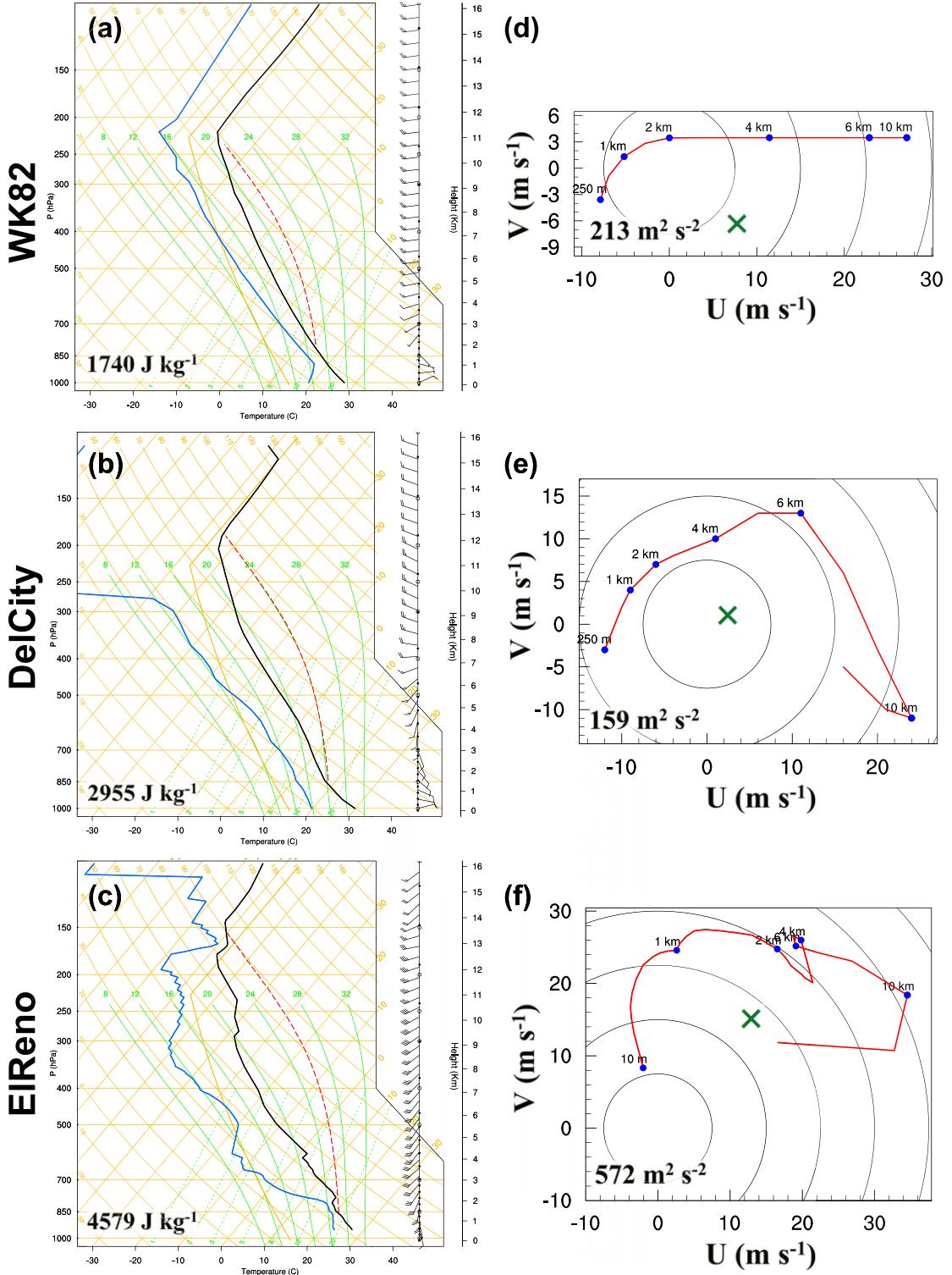


FIG. 1. Skew  $T$ -log $p$  soundings and corresponding hodographs used in default (a),(d) WK82; (b),(e) DelCity; and (c),(f) ElReno simulations. Surface-based convective available potential energy (CAPE) and 0–3 km AGL storm-relative helicity (SRH) are indicated on the skew  $T$ -log $p$  and hodograph plots, respectively. Mean storm motion vectors (used in SRH calculations and indicated by green  $\times$ s on hodographs) were computed using the locations of the hook echo signatures at  $t = 30$  and  $t = 120$  min.

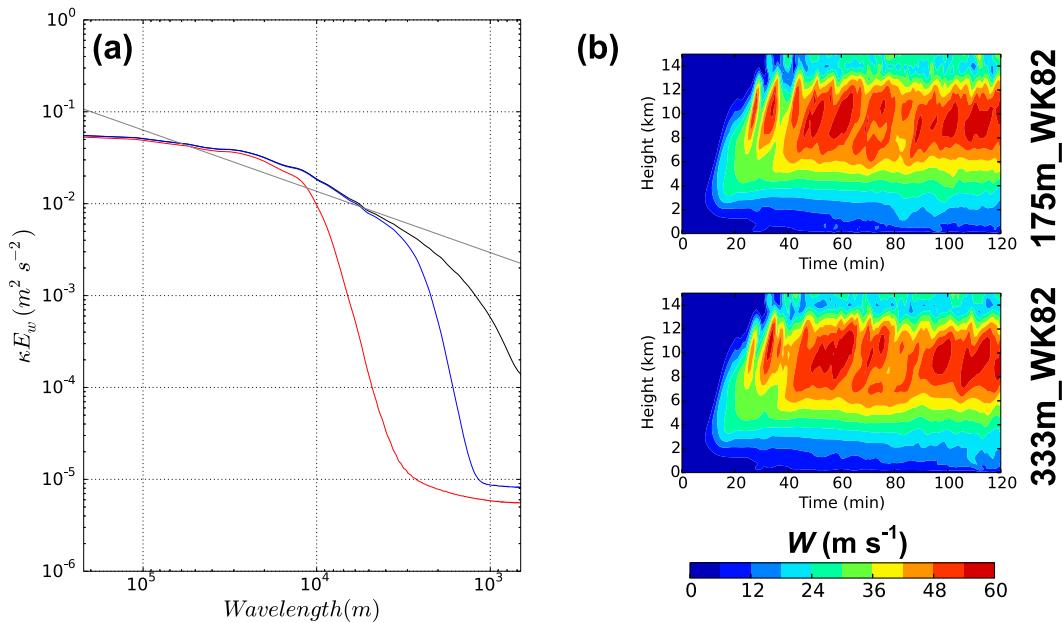


FIG. 2. (a) Two-dimensional vertical velocity spectra averaged over the lowest 5 km of WK82 simulations during  $t = 60$ –70 min:  $\Delta x = 333$  m (black),  $\Delta x = 1$  km (blue), and  $\Delta x = 4$  km (red). A gray line with  $-5/3$  slope is plotted to aid identification of the inertial subrange. (b) Time–height composites (i.e., plots of the highest-amplitude values that occurred within the horizontal domain at each model level and output time) of maximum  $w$  ( $m s^{-1}$ ) for (top) 175-m and (bottom) 333-m WK82 simulations.

## 2. Methods

### a. Model configuration

Simulations are generated using version 3.4.1 of the Advanced Research Weather Research and Forecasting Model (WRF-ARW; Skamarock et al. 2008) with typical idealized cloud model settings. The horizontal grid spacing is set to 333 m or 1, 2, 3, or 4 km. The 333-m (control) simulations serve as the “truth” against which the coarser simulations are evaluated. The model time step is fixed at 1 s for all  $\Delta x$  to isolate grid spacing sensitivities. A total of 50 vertical model levels are used, with the grid spacing increasing from  $\sim 100$  m near the surface to  $\sim 700$  m between 15 and 20 km AGL (model top). Simulations are initialized with a sounding, and in the control simulations, convection is initiated by a thermal bubble with maximum potential temperature perturbation of 3 K (first two simulations) or 5 K (third simulation). Radiation and surface physics are neglected. The Thompson microphysics scheme (Thompson et al. 2004, 2008), which uses five hydrometeor categories (including graupel) and predicts two moments of the rain and cloud particle size distributions, is used—a more sophisticated scheme than the purely single-moment parameterizations used in all (to the authors’ knowledge) previous supercell grid spacing sensitivity studies. Turbulence is parameterized using the

1.5-order TKE closure. Additional model details are given in Table 1.

### b. Control simulations

Given that the characteristic errors of physics parameterization schemes can vary substantially from case to case, as can the spatial scales of critical storm features (e.g., updraft) and thus model resolution errors, it is plausible that the grid spacing dependence of numerical supercell forecasts likewise varies considerably, even using the same model configuration. This motivated us to use multiple control simulations initialized using a range of soundings in our experiments (Fig. 1). The first, hereafter “WK82,” is the WRF idealized supercell test case sounding, which uses the temperature and mixing ratio profiles from Weisman and Klemp (1982) and a quarter-circle hodograph (Figs. 1a,b). The second is the 20 May 1977 Del City, Oklahoma, sounding with storm motion subtracted (“DelCity”; Figs. 1c,d). The third is a Rapid Update Cycle (RUC; Benjamin et al. 2004) sounding analyzed near the 24 May 2011 supercell that produced an (enhanced Fujita) EF-5 tornado near El Reno, Oklahoma (“ElReno”; Figs. 1e,f). The soundings vary considerably from each other (e.g., cf. the parameters shown in Fig. 1), providing a broad sampling of the atmospheric supercell parameter space. As will be shown, this results in substantial differences in storm evolution among these three control simulations.

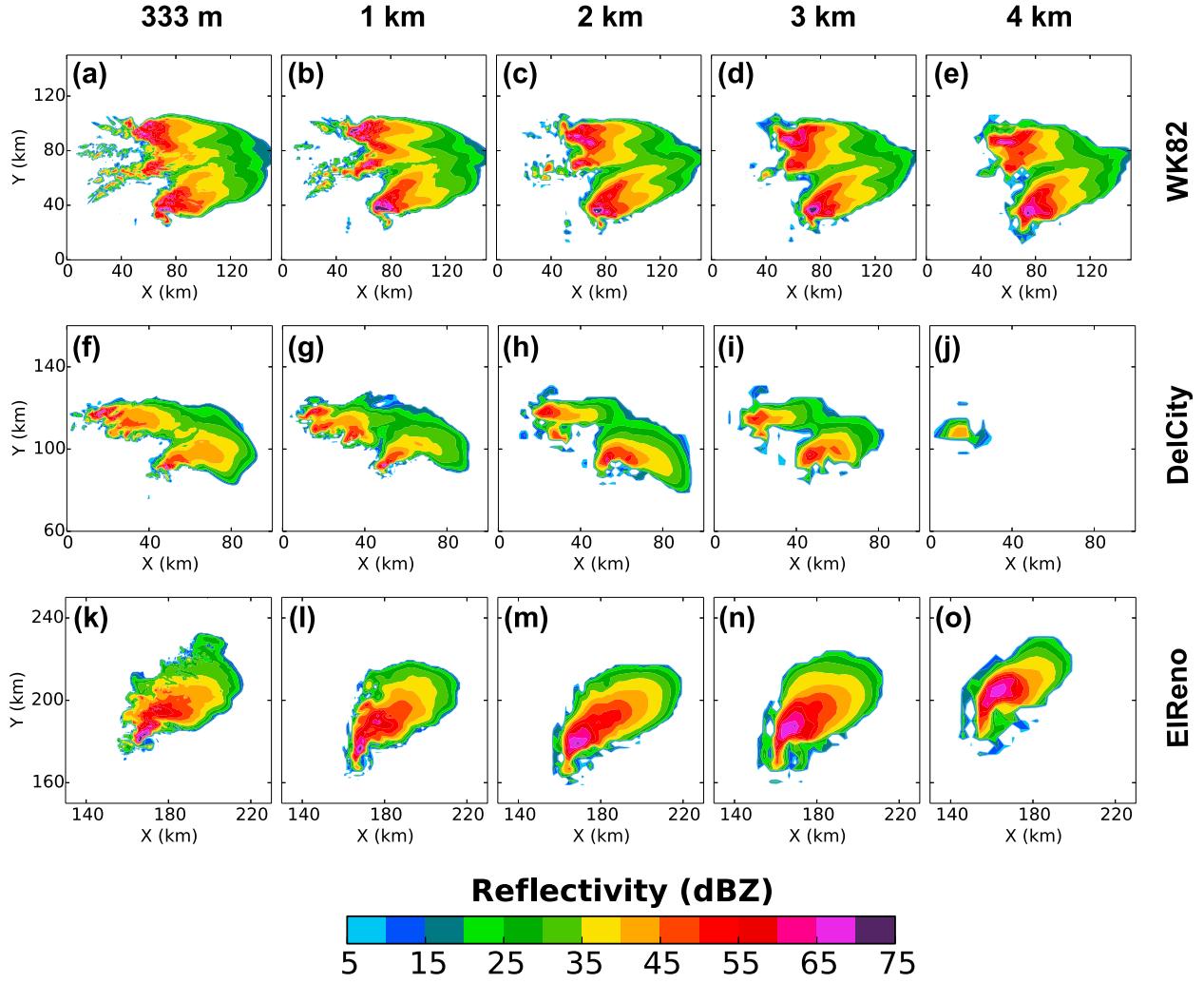


FIG. 3. Reflectivity (dBZ) at  $z \approx 2$  km,  $t = 120$  min in (top) WK82, (middle) DelCity, and (bottom) ElReno simulations with  $\Delta x =$  (a),(f),(k) 333 m; (b),(g),(l) 1 km; (c),(h),(m) 2 km; (d),(i),(n) 3 km; and (e),(j),(o) 4 km. In these and subsequent horizontal plots, the axes are relative to the southwest corner of the model domain.

Several additional control simulations, described later, were used to further explore the sensitivity of grid spacing dependence. The control simulations are run for 120 min; all times  $t$  referenced herein are relative to the beginning of the control simulations.

To help assess the realism of the control simulations, 1D vertical velocity spectra were computed (using the fast Fourier transform) and averaged over the lowest 5 km of the model domain every minute over  $t = 60$ –70. The following procedure was used to compute the spectrum at each model level and time. First, the corresponding 2D cross section of the vertical velocity field was detrended. Then, zonal (meridional) spectra were computed at each  $y$  coordinate ( $x$  coordinate). Next, the zonal (meridional) spectra were averaged in the meridional (zonal) dimensions, producing two different mean spectra that were subsequently averaged

together to create the final spectrum. For each control simulation, the existence of an inertial subrange (over which energy is transferred from larger to smaller scales with negligible dissipation; e.g., [Gage 1979](#); [Skamarock 2004](#)) is evidenced by a region of the spectra approximating the expected  $-5/3$  slope (shown for WK82 in [Fig. 2a](#)).<sup>1</sup> This suggests the 333-m grid well resolves large energy-containing eddies (i.e., the flow is nearly fully turbulent) and that our turbulence scheme therefore well parameterizes energy transfer to subgrid scales ([Bryan et al. 2003](#)). Furthermore, repeating the

<sup>1</sup> The computed inertial subranges are very similar to those obtained in [Fiori et al. \(2010\)](#). Recalculating the spectra at  $z = 5$  km within a smaller domain focused on the supercell (following that study) produces very similar results.

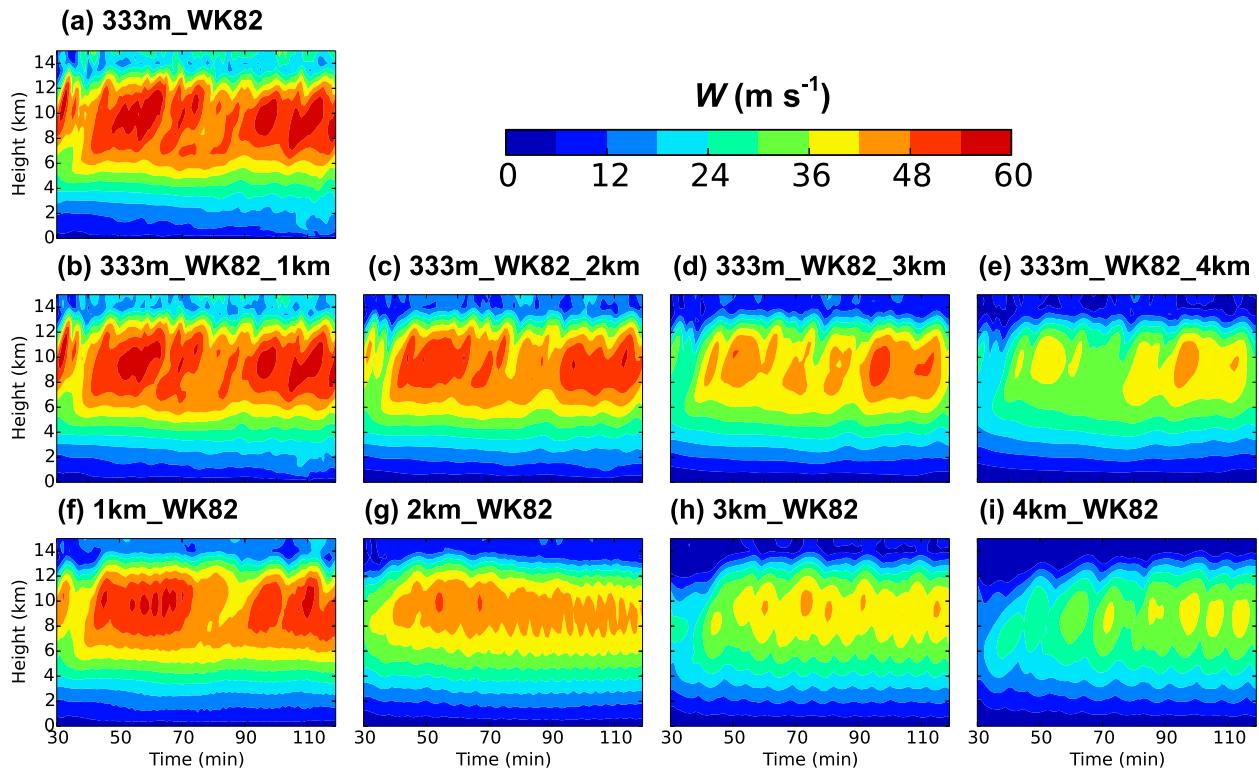


FIG. 4. Time–height maximum  $w$  ( $\text{m s}^{-1}$ ) composites for WK82 simulations: (a) 333m\_WK82, (b) 333m\_WK82\_1km, (c) 333m\_WK82\_2km, (d) 333m\_WK82\_3km, (e) 333m\_WK82\_4km, (f) 1km\_WK82, (g) 2km\_WK82, (h) 3km\_WK82, and (i) 4km\_WK82.

WK82 control simulation with 175-m grid spacing and a 0.5-s time step and comparing to the 333-m simulation using all the metrics employed in section 3 revealed the 333-m simulation to be reasonably well converged. For example, time–height composites of maximum  $w$  for the 175- and 333-m simulations are very similar (Fig. 2b). Thus, while 333-m grid spacing is too coarse to fully resolve convection, it appears to well represent the phenomena whose grid spacing dependence we seek to analyze.

LES-based turbulence closures are admittedly less realistic for the 1–4 km grids used in our coarse simulations (e.g., Bryan et al. 2003; Wyngaard 2004). However, they are still in widespread use on these scales in idealized simulations given the lack of more appropriate schemes. In most operational NWP models using 3–4-km grids, a planetary boundary layer (PBL) scheme is used to parameterize vertical mixing, and the horizontal mixing length is diagnosed from horizontal deformation or set to a constant. As will be shown in Section 3f, however, using a PBL scheme actually degraded our 3–4-km simulations. Thus, while using LES-based turbulence parameterization is certainly suboptimal on these coarser grids, it appears to be a superior alternative to PBL schemes in at least some cases in our idealized

framework. This is perhaps unsurprising given that PBL schemes are designed for still coarser [ $O(10)$  km] grids than those used here. Indeed, grid spacings of  $O(1)$  km have been termed the *terra incognita* of turbulence parameterization given the inappropriateness of both LES-based and PBL schemes on these scales (Wyngaard 2004).

### c. Coarse simulations

The 1-, 2-, 3-, and 4-km simulations, hereafter referred to collectively as the “coarse simulations,” are initialized (using WRF’s restart capability) from their respective control simulations at  $t = 30$  min. Initial conditions for the coarse simulations are obtained by filtering wavelengths in the control simulation that cannot be resolved on the coarse grid, then sampling (i.e., grid matching) the filtered simulation. More specifically, the Raymond (1988) sixth-order implicit tangent filter is used to damp wavelengths smaller than twice the coarse grid spacing by >90% (Fig. 2a). The coarse simulations are run for 90 min, thus ending at the same time as the 120-min control simulations.

The coarse simulations are compared to both the original and filtered versions of their respective control

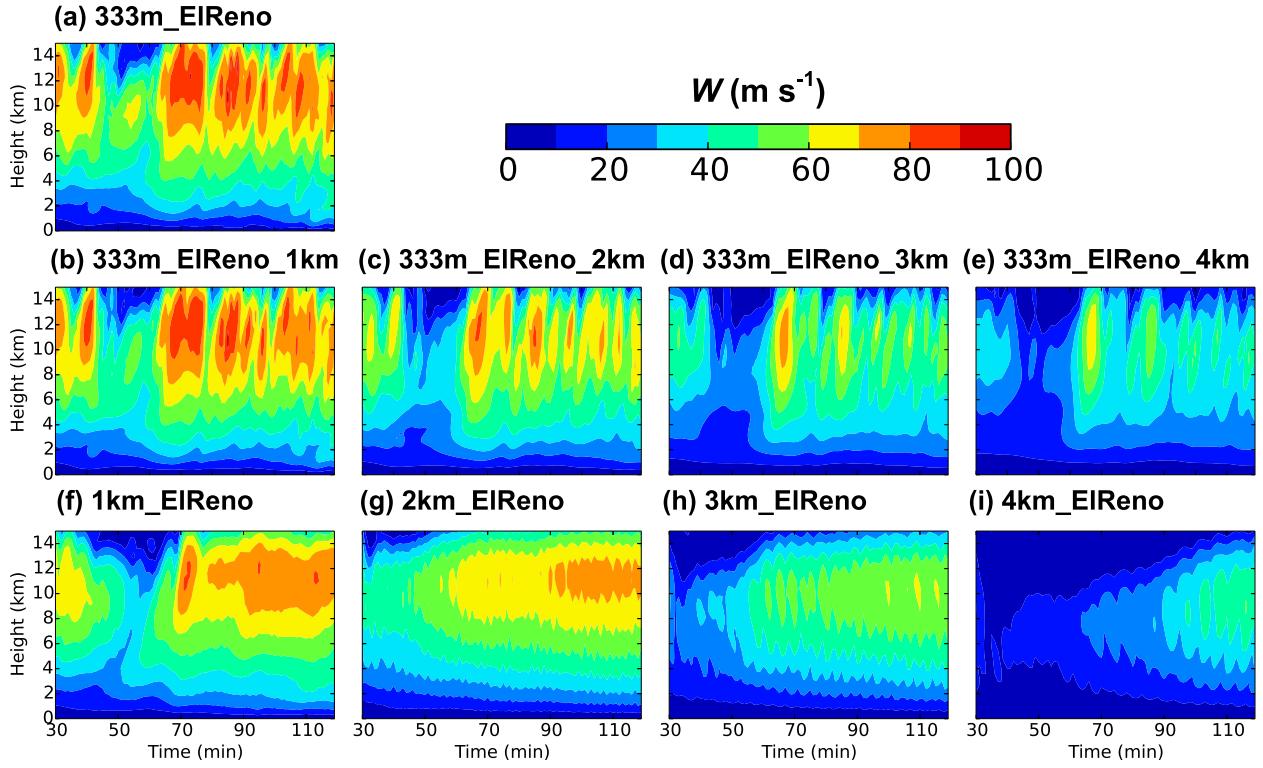


FIG. 5. As in Fig. 4, but for ElReno simulations.

simulations. The latter comparison method ignores scales in the control simulation that cannot be represented on the coarse grid, thereby isolating errors that arise due to poor resolution of upscale-interacting processes. Simulations are labeled according to their grid spacing and initialization sounding (e.g., “333m\_WK82”), and names of filtered control simulations denote the grid spacing of the corresponding coarse simulation to which they should be compared (e.g., “333m\_ElReno\_2km”).

As a final test of the suitability of the 333-m grid spacing in our control simulations, we repeated the coarse WK82 simulations but restarted from 175m\_WK82 instead of 333m\_WK82.<sup>2</sup> The grid spacing dependence (not shown) was very similar to that obtained using the 333-m control simulation. Given this result and the additional evidence presented in section 2b, we conclude that 333-m grid spacing, while not convection resolving, is sufficient to resolve the scales important to the grid spacing dependencies we seek to evaluate.

### 3. Results

#### a. General storm evolution

As expected from the range of initialization soundings used, the three control simulations vary substantially from one another. Multiple supercells are produced in 333m\_WK82 (Fig. 3a) and 333m\_DelCity (Fig. 3f) via storm splitting, whereas a single supercell is sustained in 333m\_ElReno (Fig. 3k), consistent with the strong clockwise hodograph curvature (Fig. 1f; Rotunno and Klemp 1982). Critical supercell features differ considerably among the control simulations, including the updraft (cf. Figs. 4a and 5a) and cold pool (cf. Figs. 6a and 7a).

In all the coarse WK82 and ElReno simulations, the low-level reflectivity field at  $t = 120$  min qualitatively resembles that in the control simulation, with distinctly supercellular storms being retained even in the 4-km simulations (Figs. 3e,o). By the end of the 4-km DelCity simulation, however, the left-moving storm has severely weakened, and the right-moving storm has dissipated (Fig. 3j). This result is broadly consistent with Kessler-microphysics ARPS simulations with the same sounding in which  $\Delta x < 2.6$  km was required to maintain the two storms (Noda and Niino 2003). That study showed that weakening of the linear dynamical forcing of the updraft at larger grid spacings was primarily responsible for the

<sup>2</sup> Since 175 is not a multiplicative factor of the coarse grid spacings (m) used, the initial conditions of the coarse simulations were obtained by interpolating, not sampling, the 175-m fields at  $t = 30$  min.

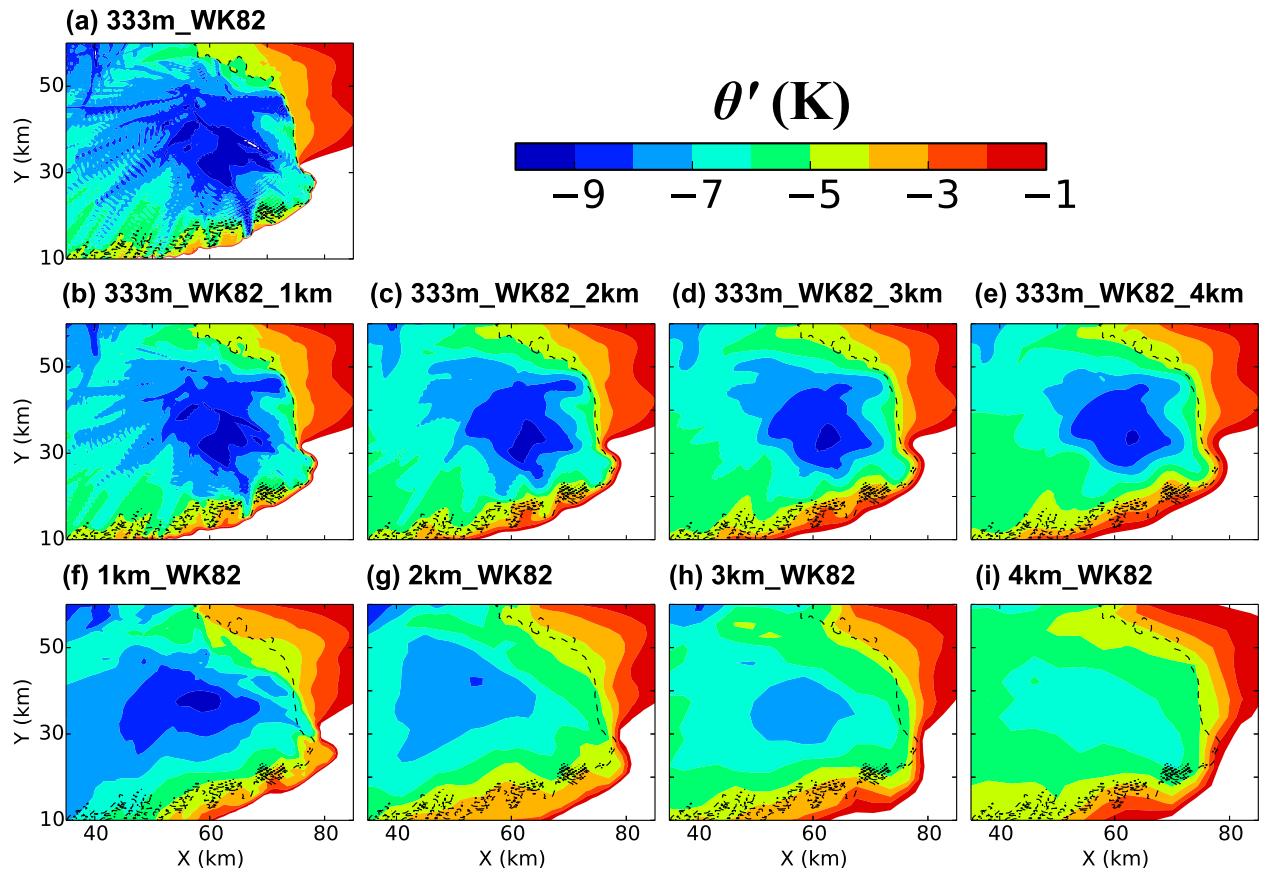


FIG. 6. Horizontal composites (i.e., plots of the highest-amplitude values that occurred at each grid point during the simulation) of minimum surface potential temperature perturbation (K) for WK82 simulations: (a) 333m\_WK82, (b) 333m\_WK82\_1km, (c) 333m\_WK82\_2km, (d) 333m\_WK82\_3km, (e) 333m\_WK82\_4km, (f) 1km\_WK82, (g) 2km\_WK82, (h) 3km\_WK82, and (i) 4km\_WK82. The 333m\_WK82 -6-K contour (black dashed curve) is plotted in all panels for reference.

storm demise, noting that for a given reference  $w$  field, larger grid spacing yields smaller horizontal  $w$  gradients and thus reduces the upward pressure gradient force that arises from the interaction of the updraft with vertical wind shear. Time-height plots<sup>3</sup> of maximum  $w$  reveal that 4km\_ElReno is also substantially too weak, though only during the first hour of the simulation (cf. Figs. 5a,i and 5e,i). The 4km\_WK82 right-moving storm<sup>4</sup> updraft is generally closer in magnitude to both the raw (cf. Figs. 4a,i) and filtered (cf. Figs. 4e,i) control simulations. That the 333m\_WK82\_4km updraft composite

matches the 333m\_WK82 updraft composite relatively well suggests that the 333m\_WK82 updraft is larger than the 333m\_DelCity and 333m\_ElReno updrafts (since larger objects are generally better resolved than smaller objects at marginal resolution). This is confirmed by both visual and quantitative comparisons of horizontal cross sections of  $w$  from the three control simulations (Fig. 8). The improved resolvability of the WK82 updraft and its horizontal gradients renders the linear dynamical forcing less sensitive to grid spacing, which may largely explain the superior performance of 4km\_WK82 relative to the other 4-km simulations. The relative insensitivity of the WK82 simulations is further investigated in section 3e.

In all three sets of experiments, maximum updraft speed increases with resolution (Figs. 4 and 5), consistent with Adlerman and Droege (2002) and Noda and Niino (2003), and in contrast to the nonmonotonic grid spacing dependence seen in Bryan and Morrison (2012) and Verrelle et al. (2015). This suggests that in

<sup>3</sup>The regular, high-frequency oscillations in the WK82 and ElReno time-height plots result from damping of field maxima when they fall between grid points. These do not arise in the DelCity simulations since those storms are nearly stationary.

<sup>4</sup>All plots other than horizontal cross sections are computed over the entire model domain in the ElReno simulations, and over a subdomain containing only the primary (initial right moving) storm in the WK82 and DelCity simulations.

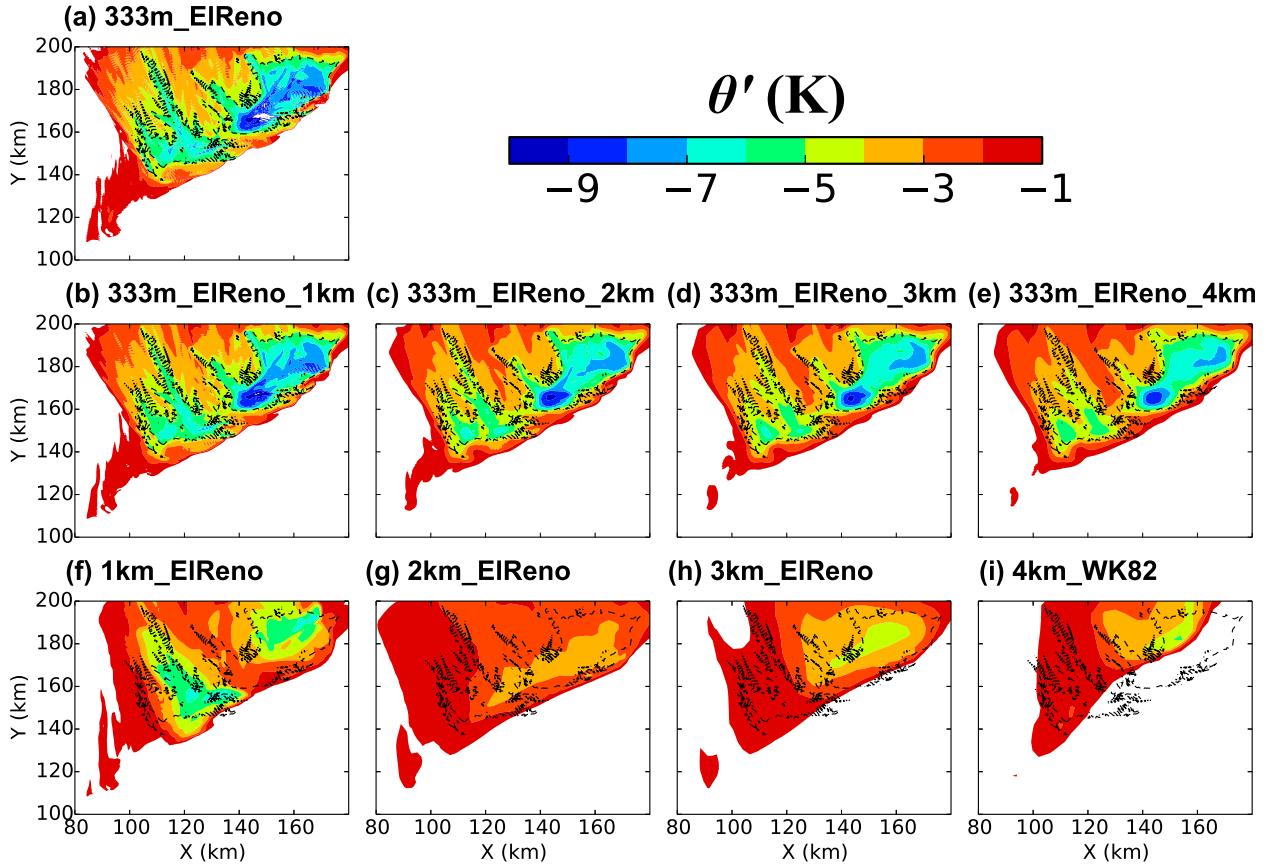


FIG. 7. As in Fig. 6, but for ElReno simulations.

our simulations, updraft underestimation due to marginal resolution of the updrafts as well as correspondingly damped nonhydrostatic effects dominates any updraft overestimation due to underrepresented subgrid turbulence (and thus entrainment) at larger grid spacings. Storm maturation as measured by maximum updraft intensity (as well as other parameters) slows as grid spacing increases, consistent with previous supercell studies (Adlerman and Droege 2002; Fiori et al. 2010; Noda and Niino 2003). This is also consistent with the demonstration by Weisman et al. (1997) that convective time scales should increase as model resolution decreases due to the increasing significance of hydrostatic effects. Delayed storm development is a critical consideration for WoF given the emphasis on longer forecast lead times.

To examine the grid spacing sensitivity of the simulated cold pools, we plotted the minimum low-level (valid at the lowest scalar level,  $\sim 50$  m above the surface) potential temperature perturbation that occurred over the duration of the simulation at each grid point. These horizontal composites reveal that 1-km grid spacing is sufficient to capture the surface cold

pool evolution quite well in the WK82 (Fig. 6) and DelCity (until storm demise; not shown) simulations. Warm biases gradually and monotonically increase with grid spacing  $> 1$  km, consistent with Verelle et al. (2015). The 1km\_ElReno cold pool (Fig. 7), on the other hand, is generally too warm by several degrees, and the dependence of the cold pool errors on grid spacing is highly nonlinear, with the warm bias increasing dramatically as  $\Delta x$  increases from 1 to 2 km, then remaining relatively steady as  $\Delta x$  increases further. Given the influence on cold pools of interactions between microphysical and turbulent processes (e.g., acceleration of evaporational cooling by entrainment), the potential sensitivity of parameterization schemes to grid spacing (section 1), and the strong dependence of storm processes on the environment, it is perhaps not surprising that the cold pool grid spacing dependence varies substantially among our experiments.

The usefulness of supercell hazard forecasts is constrained by the accuracy of predicted storm paths. Supercell motion is primarily determined by the sum of three effects: advection, forward propagation by the

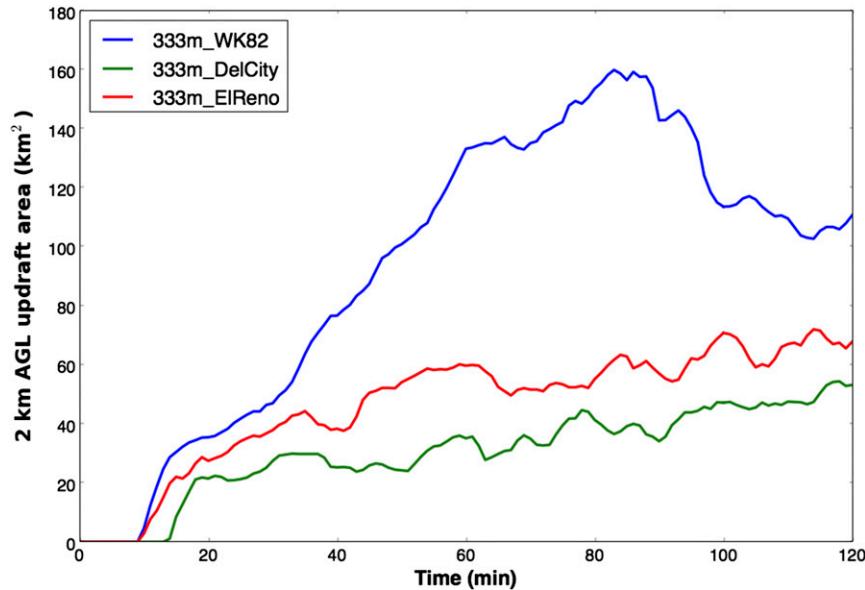


FIG. 8. Area of  $w > 5 \text{ m s}^{-1}$  at  $z \approx 2 \text{ km}$  for 333m\_WK82 (blue), 333m\_DelCity (green), and 333m\_ElReno (red). Qualitatively similar results were obtained using different heights and  $w$  thresholds.

cold pool (e.g., Corfidi 2003), and rightward propagation by linear dynamic forcing of the updraft (Rotunno and Klemp 1982). Storm advection should not generally be very sensitive to horizontal grid spacing since the background flow does not change (though changes in vertical storm structure with  $\Delta x$  may minorly impact the advection velocity). However, storm motion *can* be impacted by the grid spacing dependence of cold pool intensity (e.g., VandenBerg et al. 2014) and updraft intensity (e.g., Noda and Niino 2003).

As a metric for storm path, the location of the maximum  $w$  near 2 km AGL was computed every minute ( $t = 30\text{--}120 \text{ min}$ ) for each simulation (shown in Fig. 9a for the ElReno simulations). The displacement of the coarse simulation storm paths from the corresponding control

simulation storm paths was then computed and averaged using a five-point running mean (Fig. 9b). The 1-, 2-, and 3-km storm path errors are  $<5 \text{ km}$  throughout the 90-min simulations, suggesting that propagation errors due to grid spacing sensitivities of cold pool intensity and linear dynamic forcing are generally small (relative to current operational storm-track errors) at these resolutions. Errors in the 4km\_ElReno storm path, however, exceed 10 km after  $t = 70 \text{ min}$  and 15 km after  $t = 100 \text{ min}$  (Fig. 9b). While the 333m\_ElReno storm exhibits increasing rightward propagation during the middle third of the simulation as it strengthens, the 4km\_ElReno storm path is nearly linear until roughly halfway through the simulation, at which time the storm rapidly shifts toward the right,

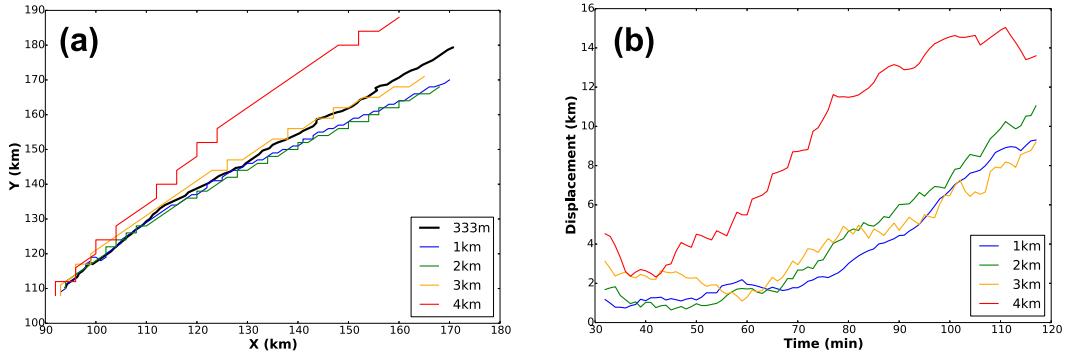


FIG. 9. (a) ElReno storm paths as defined by  $w$  maxima every minute from  $t = 30$  to  $120 \text{ min}$  at  $z \approx 2 \text{ km}$ ; (b) displacements of coarse simulation storm paths from 333m\_ElReno storm path.

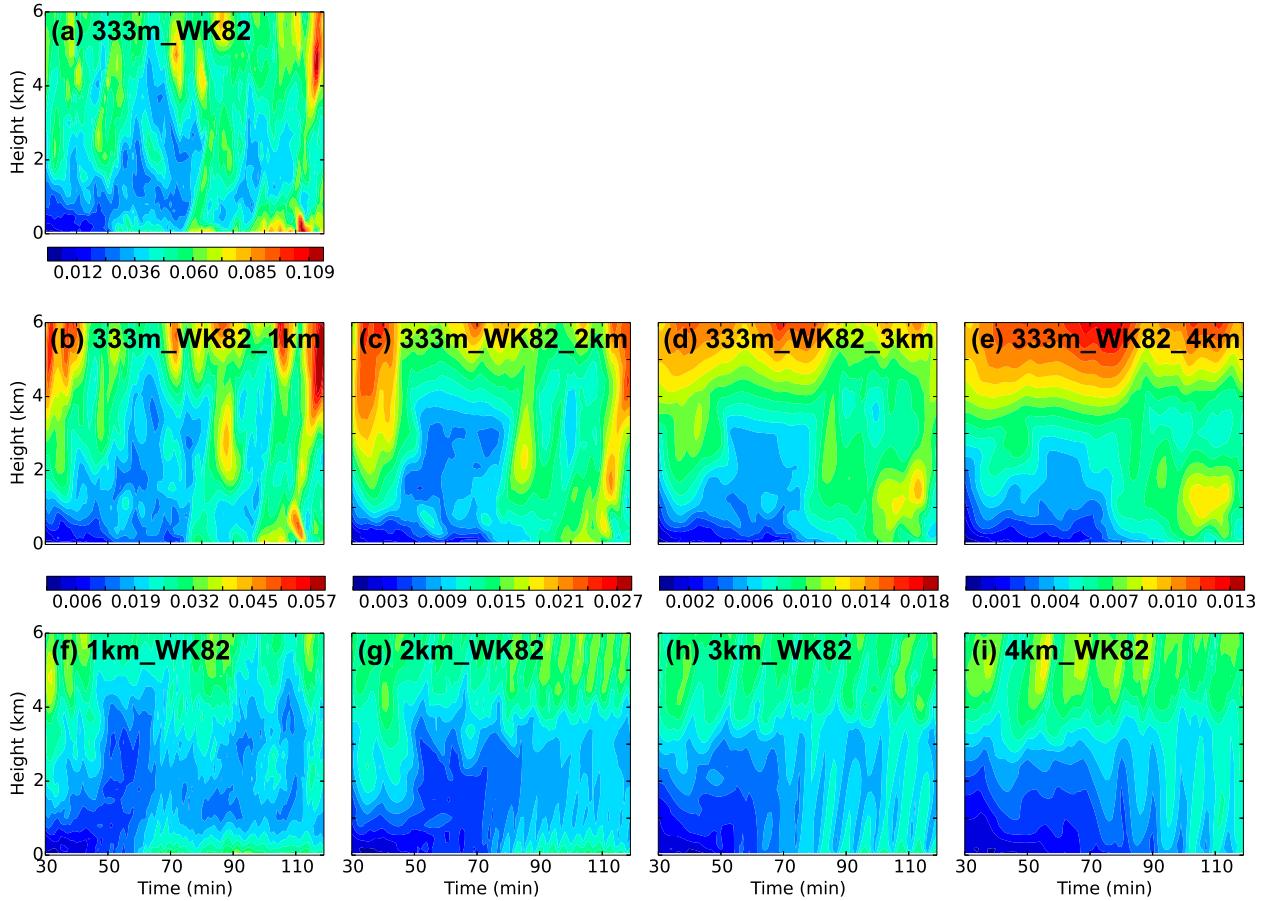


FIG. 10. Time–height  $\zeta$  ( $\text{s}^{-1}$ ) composites for WK82 simulations: (a) 333m\_WK82, (b) 333m\_WK82\_1km, (c) 333m\_WK82\_2km, (d) 333m\_WK82\_3km, (e) 333m\_WK82\_4km, (f) 1km\_WK82, (g) 2km\_WK82, (h) 3km\_WK82, and (i) 4km\_WK82. For each  $\Delta x$ , the same plotting color scale is used for the coarse simulation as for the corresponding filtered control simulation.

concurrent with rapid intensification of the main updraft (Fig. 5i). The storm motion vector remains counterclockwise to that in 333m\_ElReno through the end of the simulation, however, during which the updraft remains substantially too weak. The correlation between the biases in the 4km\_ElReno storm motion and updraft intensity strongly suggests that the storm path errors were produced primarily by inadequate linear dynamic forcing (and therefore propagation) of the updraft.

The substantial grid spacing sensitivities documented above suggest that limited horizontal resolution can significantly reduce the accuracy of model-derived guidance for convective warning operations, even apart from other initial condition and model (e.g., physics parameterization) errors. The results further suggest that grid spacing sensitivities of forecasts can strongly vary from case to case, making it difficult to correct for model resolution errors. In sections 3b and 3c, we explore the grid spacing sensitivity of model output directly tied to three important

supercell hazards: tornadoes, heavy rainfall, and damaging winds.

#### b. Low-level vorticity evolution

For model forecasts with  $\Delta x$  of 1–4 km, tornadoes and even mesocyclones are generally poorly resolved. Nevertheless, maximum low-level vertical vorticity serves as a proxy for supercell tornado potential since the development of strong low-level rotation (e.g., low-level mesocyclone) in a model may signify the onset of processes (albeit imperfectly resolved) that contribute to the genesis of tornadoes in real storms. This is the premise of using relatively coarse (e.g., 3 km) ensembles to generate probabilistic tornado forecasts in initial WoF systems (e.g., Stensrud et al. 2013; Potvin and Wicker 2013). Time–height and  $x$ – $y$  vorticity composites as well as horizontal vorticity cross sections were examined to assess the impact of limited model resolution on the evolution of low-level vorticity in our experiments (Figs. 10–15). The horizontal composites were computed over the lowest 2 km. Comparisons of the

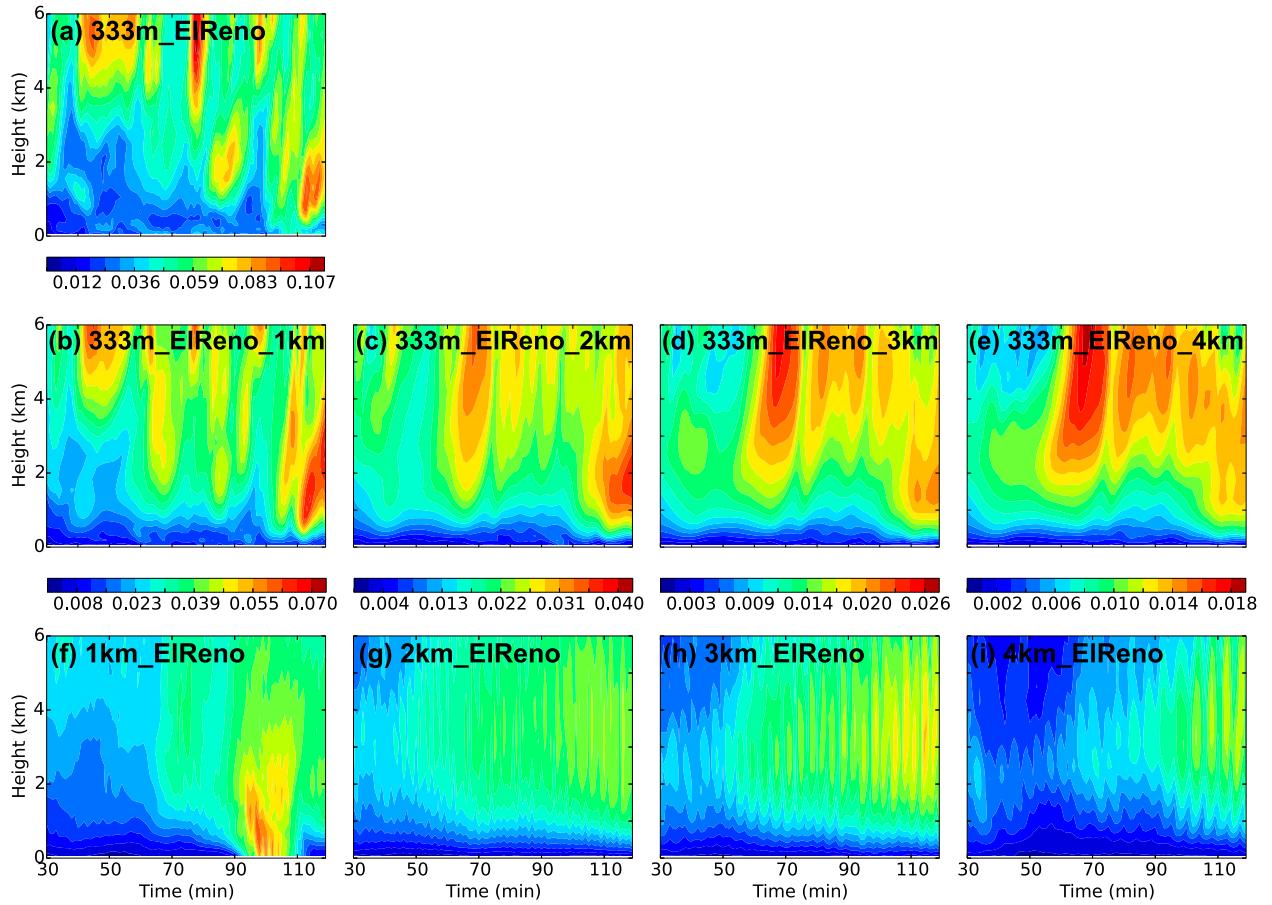
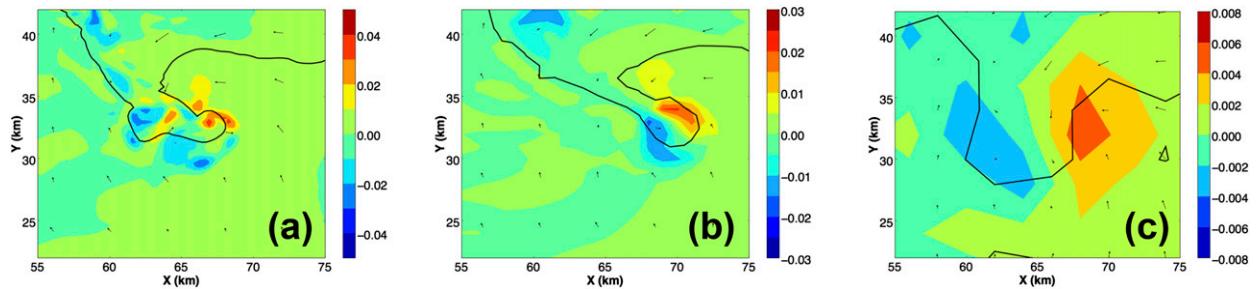


FIG. 11. As in Fig. 10, but for ElReno simulations.

vorticity composites with composites of Okubo–Weiss number ( $=D^2 - \zeta^2$ , where  $D$  is the total deformation; Okubo 1970; Weiss 1991; Markowski et al. 2011) and inspection of horizontal wind fields verified that the vorticity plots primarily capture regions of strong rotation (e.g., low-level mesocyclones) rather than of strong linear shear (not shown). Since vertical vorticity collapses to the smallest resolvable scale on grids with  $\Delta x$  greater than the Kolmogorov length scale ( $\sim 0.1$  mm),

the peak  $\zeta$ ,  $\zeta_{\max}$ , should scale inversely with  $\Delta x$  (e.g., Adlerman and Droege 2002). This effect tends to dominate  $\zeta_{\max}$  differences arising from neglecting upscale-interacting, small-scale processes, making comparisons with the filtered control simulations particularly useful for this variable. As will be shown, however, finescale spatiotemporal details of vorticity (and other variables) in supercells cannot be reliably diagnosed by downscaling from a coarser grid.

FIG. 12. Vorticity ( $s^{-1}$ ; shading), horizontal wind vectors every 4 km (arrows), and  $dBZ = 40$  (contours) valid at  $z \approx 1.5$  km,  $t = 105$  min in (a) 333m\_WK82, (b) 1km\_WK82, and (c) 4km\_WK82.

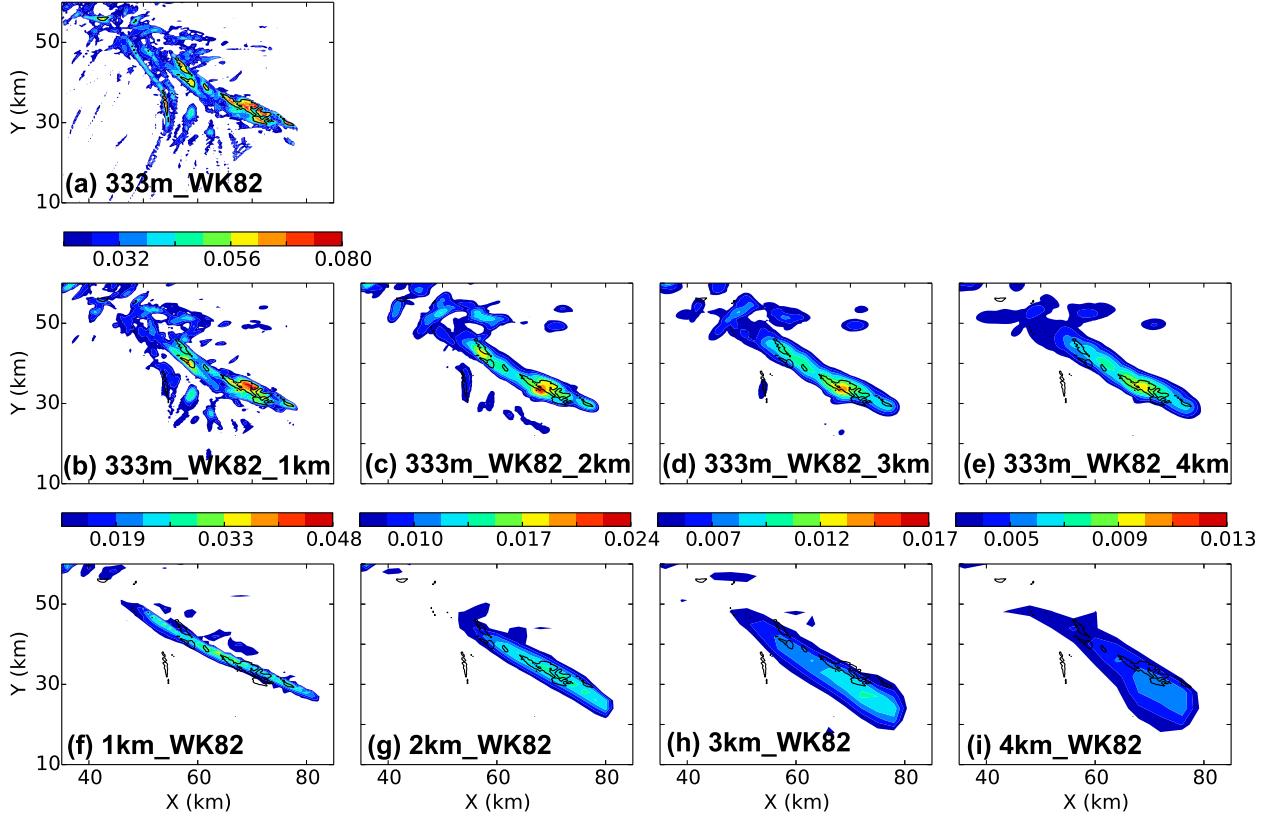


FIG. 13. Horizontal low-level (valid over the lowest 2 km of the model domain)  $\zeta$  ( $\text{s}^{-1}$ ) composites for WK82 simulations: (a) 333m\_WK82, (b) 333m\_WK82\_1km, (c) 333m\_WK82\_2km, (d) 333m\_WK82\_3km, (e) 333m\_WK82\_4km, (f) 1km\_WK82, (g) 2km\_WK82, (h) 3km\_WK82, and (i) 4km\_WK82. For each  $\Delta x$ , the same plotting color scale is used for the coarse simulation as for the corresponding filtered control simulation. The 333m\_WK82 0.05  $\text{s}^{-1}$  contour (black dashed curve) is plotted in all panels for reference.

All three sets of coarse simulations capture the basic (increasing) trend in low-level vorticity in the control simulations, even with  $\Delta x = 4 \text{ km}$  (e.g., Figs. 10 and 11). As with maximum updraft and cold pool intensity, the coarsened resolution produces a low bias in  $\zeta_{\text{max}}$  relative to the filtered control simulations. More importantly, qualitative differences in storm evolution occur in all three sets of experiments. Whereas multiple episodes of low-level vorticity intensification and weakening occur in 333m\_WK82 (Fig. 10a) and 333m\_ElReno (Fig. 11a), the maximum low-level vorticity in the 2-, 3-, and 4-km experiments increases approximately monotonically through the end of each simulation (Figs. 10g–i and 11g–i). In addition, a low-level mesocyclone cycling phase that occurs during  $t = 90\text{--}100 \text{ min}$  in 333m\_ElReno (the only cycling event in the control simulations) is absent from the coarse ElReno simulations. This result is consistent with the total cessation of mesocyclone cycling as the grid spacing was increased in Adlerman and Droege (2002), though in their case cycling was still evident at 1 km. The presence of a cyclonic–anticyclonic vortex

couplet straddling the hook echo in the WK82 (Fig. 12) and ElReno (not shown) control simulations is retained at all grid spacings. However, the particularly complex low-level vorticity field in the WK82 simulations fundamentally changes as grid spacing increases. In 333m\_WK82, multiple persistent cyclonic and anticyclonic vortices of similar intensity generally exist within the hook echo region after  $t = 80 \text{ min}$  (Fig. 12a). In the coarse WK82 experiments, however, only the cyclonic–anticyclonic vortex couplet is prominent (Figs. 12b,c). This simplification of the vorticity field is reflected in the horizontal composites (Fig. 13). The latter also show, on the other hand, that the inability of the coarse grids to resolve the finer-scale vortices in 333m\_WK82 does not produce large displacement errors in the maximum vorticity tracks.

Some of the finer-scale errors in the evolution of low-level rotation in the coarse simulations would be operationally significant, however. For example, the development of intense low-level vorticity in 1km\_ElReno

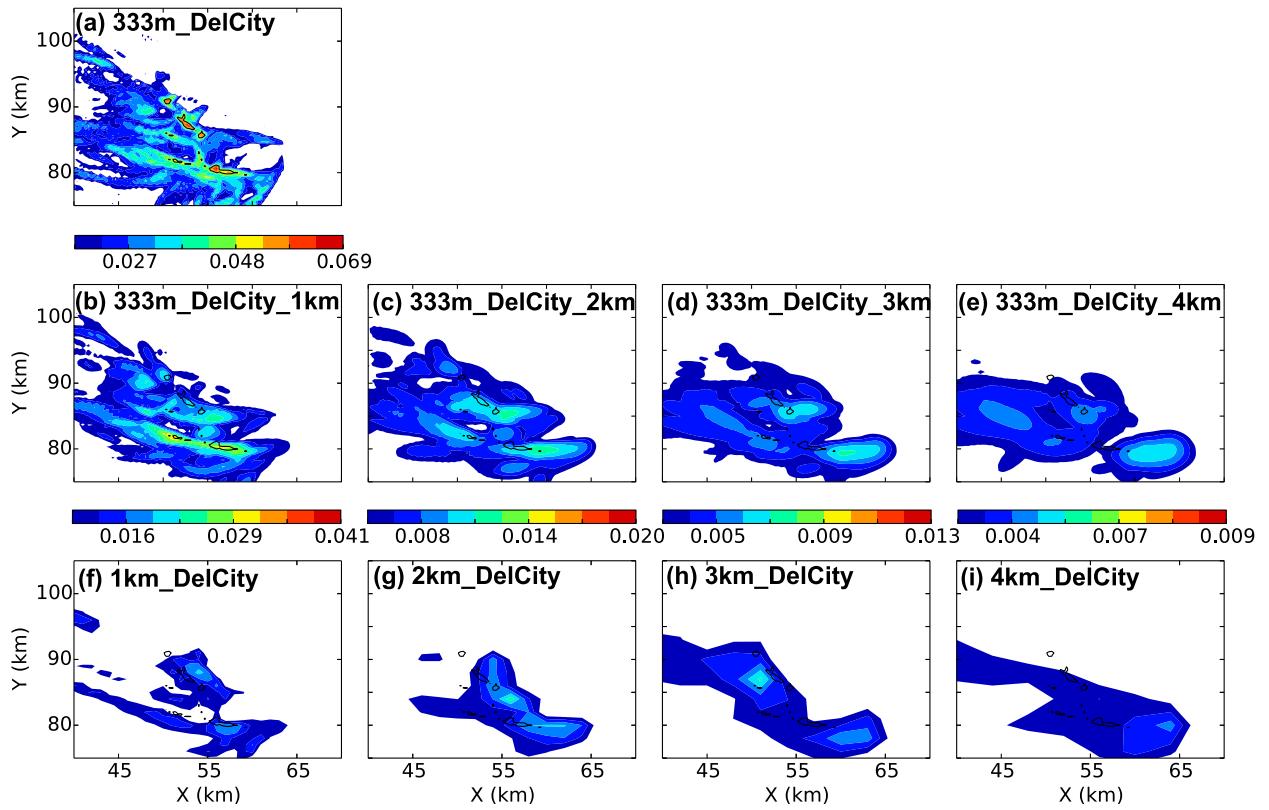


FIG. 14. As in Fig. 13, but for DelCity simulations.

is temporally offset from that in the control simulation (Figs. 11a,f vs Figs. 15a,f) and is erroneously associated with a tornado-like vortex at the surface (not shown). In addition, the intense near-ground rotation present in 333m\_WK82 around  $t = 80$  min is absent in 1km\_WK82 (cf. Figs. 10a,f). These errors suggest that increasing the precision of tornado warnings to scales of several kilometers and  $\sim 10$  min will require NWP model grid spacings  $< 1$  km, even if dramatic reductions in other types of model and initialization error occur. On the other hand, comparison of the vorticity composites from the control and coarse simulations suggests grid spacings up to 3 km are sufficient for useful prediction of the tracks (Figs. 13–15) and general intensity trends (Figs. 10 and 11) of low-level supercell rotation centers, with substantial degradation as grid spacing increases from 3 to 4 km. The former result is consistent with real-world ensemble forecast experiments being conducted by the NSSL WoF group (e.g., Wheatley et al. 2015, manuscript submitted to *Wea. Forecasting*).

### c. Rainfall and wind

Flash flooding is a sometimes overlooked but deadly supercell hazard. We therefore examined grid spacing

impacts on storm-total rainfall in our simulations. The amplitudes of rainfall maxima increase with resolution in the WK82 (Fig. 16) and DelCity (Fig. 17) simulations, but exhibit nonmonotonic grid dependence in the ElReno simulations (Fig. 18). Comparison of the filtered and raw control simulations reveals the rainfall errors arise partly from the limited ability of the coarser grids to resolve high-amplitude rainfall maxima. As is true for all other variables presented so far, however, significant errors also result from the omission of fine scales that would otherwise impact larger scales. Rainfall is severely underestimated in both 4km\_DelCity and 4km\_ElReno, which is to be expected given how underdeveloped those storms are relative to their 333-m counterparts during parts of the simulations (e.g., Figs. 3 and 5). On the other hand, 3-km grid spacing is sufficient to predict locations of local rainfall maxima with some success. The overly heavy rainfall in 2km\_ElReno and 3km\_ElReno suggests enhancement of precipitation efficiency due to increasingly underrepresented entrainment at coarser resolution may be dominating the reduction of precipitation due to underrepresentation of the updraft. This hypothesis is consistent with the ElReno environment being drier than the others at low levels of the cloud-bearing layer (Fig. 1).

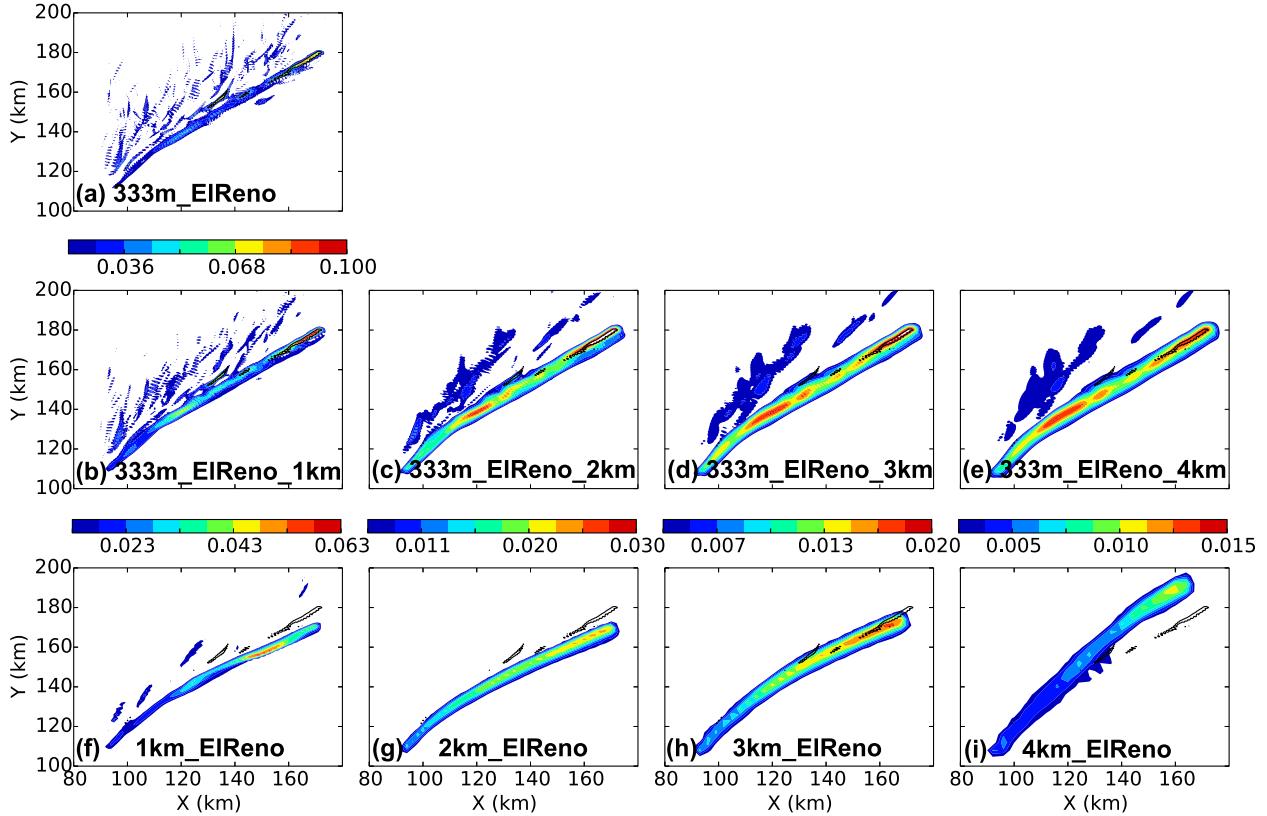


FIG. 15. As in Fig. 13, but for ElReno simulations.

Damaging straight-line winds are another common supercell hazard. Composite wind maxima at the lowest scalar level ( $\sim 50$  m above the surface) generally decrease as  $\Delta x$  increases in our experiments, and are generally too strong in the coarse WK82 simulations (not shown) and too weak in the coarse DelCity (not shown) and ElReno (in Fig. 19) simulations. As with low-level vorticity and rainfall, grid spacings up to 3 km produce low-level wind forecasts that broadly capture the locations and timing of maxima. The WK82 storm is again the least grid spacing dependent of the three.

#### *d. Grid spacing sensitivity variance by storm and variable*

To comprehensively but concisely depict the differences in grid spacing dependence between different cases and variables, we produced convergence plots using the following procedure. First, for each simulation, we computed time series of maxima of selected variables. Then, for each variable, we computed time-averaged ratios of the maxima from each coarse simulation to the maxima in the corresponding control simulation. Finally, we plotted the resulting “convergence ratios” as a

function of grid spacing (Fig. 20). This approach is traditionally used to assess whether a numerical solution is grid converged (i.e., changes little as grid spacing is further decreased); here, our purpose is more general. The grid spacing sensitivity (for  $1 \leq \Delta x \leq 4$  km) of a given variable for a particular case can be quickly assessed by examining the flatness of the associated convergence curve. The extent to which grid spacing dependencies vary by case is indicated by the distances between corresponding curves (differences in convergence ratios).

Four quantities of particular relevance to convective hazards prediction are considered: maximum vertical vorticity in the lowest 2 km, maximum vertical velocity, maximum near-surface wind speed, and maximum 5-min rainfall. The flatness of the WK82 curves relative to the ElReno and DelCity curves reflects the relative grid spacing insensitivity of the WK82 storm, documented in detail above. Maximum near-surface wind is generally less sensitive to grid spacing than are maximum rain rate, maximum updraft speed and, of course, maximum low-level vorticity. There are many differences in the grid spacing dependencies between the three storms, indicating that calibration of WoF output will be

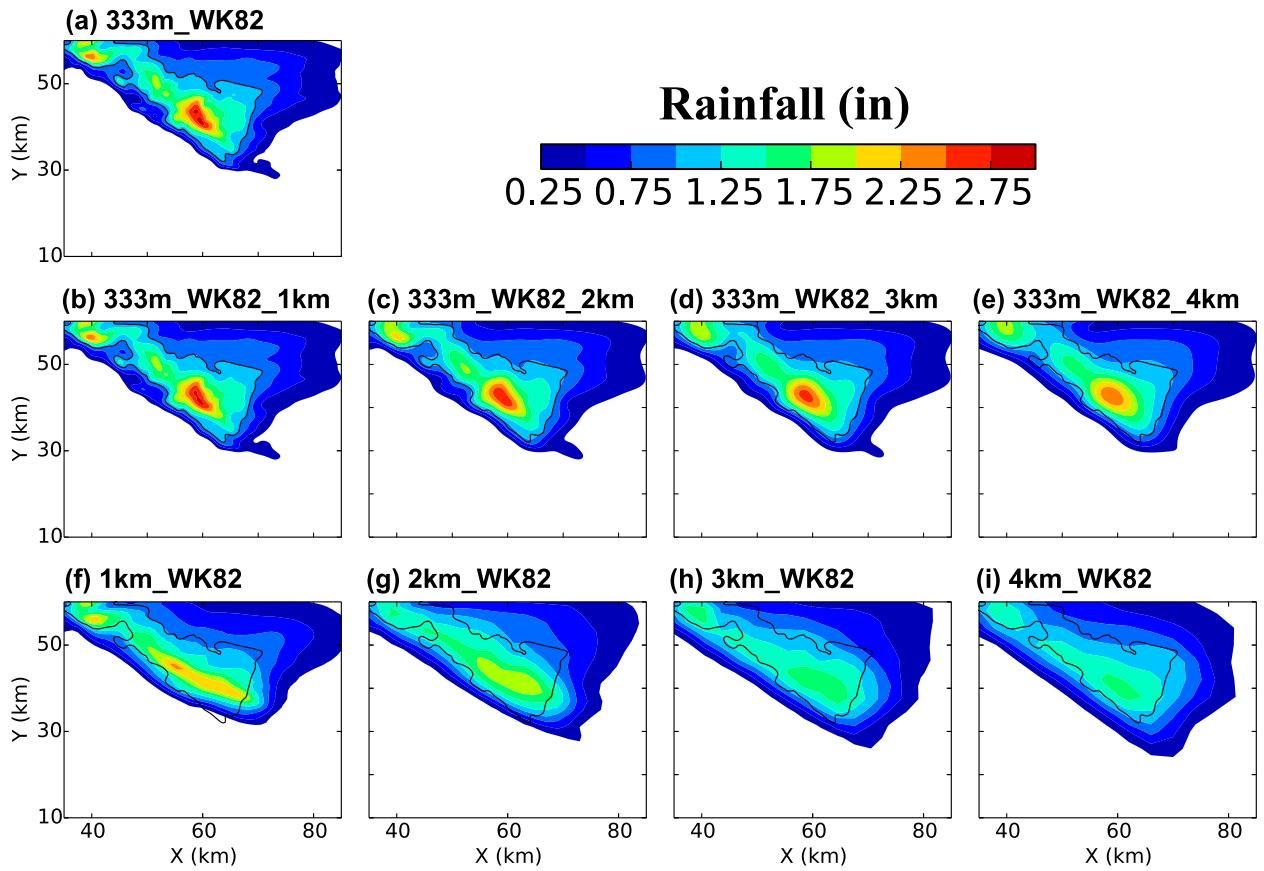


FIG. 16. Horizontal composites of total rainfall (in.) for WK82 simulations: (a) 333m\_WK82, (b) 333m\_WK82\_1km, (c) 333m\_WK82\_2km, (d) 333m\_WK82\_3km, (e) 333m\_WK82\_4km, (f) 1km\_WK82, (g) 2km\_WK82, and (h) 3km\_WK82, and (i) 4km\_WK82. The 333m\_WK82 1-in. contour (black dashed curve) is plotted in all panels for reference.

challenging unless systematic correlations between these dependencies and other variables, such as storm environment, can be identified.

#### e. Investigating the relative insensitivity of the WK82 storm to grid spacing

To explore why the WK82 storm is less sensitive to grid spacing than the other two cases, the WK82 experiments were repeated with the initialization sounding modified to better resemble the DelCity and ElReno soundings, both of which produced storms with greater grid spacing sensitivity. The four modified soundings differ from the original sounding in one of the following ways: 1) surface temperature is increased and CAPE consequently enhanced (WK82\_warm; Fig. 21b), 2) vapor mixing ratio is decreased at middle and upper levels (WK82\_dry; Fig. 21c), 3) vapor mixing ratio is decreased just above the boundary layer (WK82\_dry2; Fig. 21d), or 4) the DelCity wind profile is used. Of all four cases, only the WK82\_dry2 storm

was substantially more sensitive to grid spacing (Fig. 22). Inspection of horizontal cross sections of WK82\_dry2  $w$  revealed a much narrower updraft than that in 333m\_WK82 and the other 333m\_WK82 variants (not shown). Thus, of the seven control simulations discussed so far (DelCity, ElReno, and the five WK82 variants), the four that exhibit the least grid spacing sensitivity also contain the widest storm updrafts.

There are at least three plausible, potentially complementary, explanations for this result. First, given the primal role of the main updraft in driving supercell evolution, supercells with better-resolved (wider) updrafts should exhibit less grid spacing-dependent behavior in general. Second, the cores of wider updrafts are less influenced by entrainment, and therefore inherit less of the grid spacing sensitivity of LES-based turbulence schemes, which should again reduce the grid spacing sensitivity of other supercell processes. Third, the updrafts in the more sensitive simulations may be narrower because of increased entrainment of

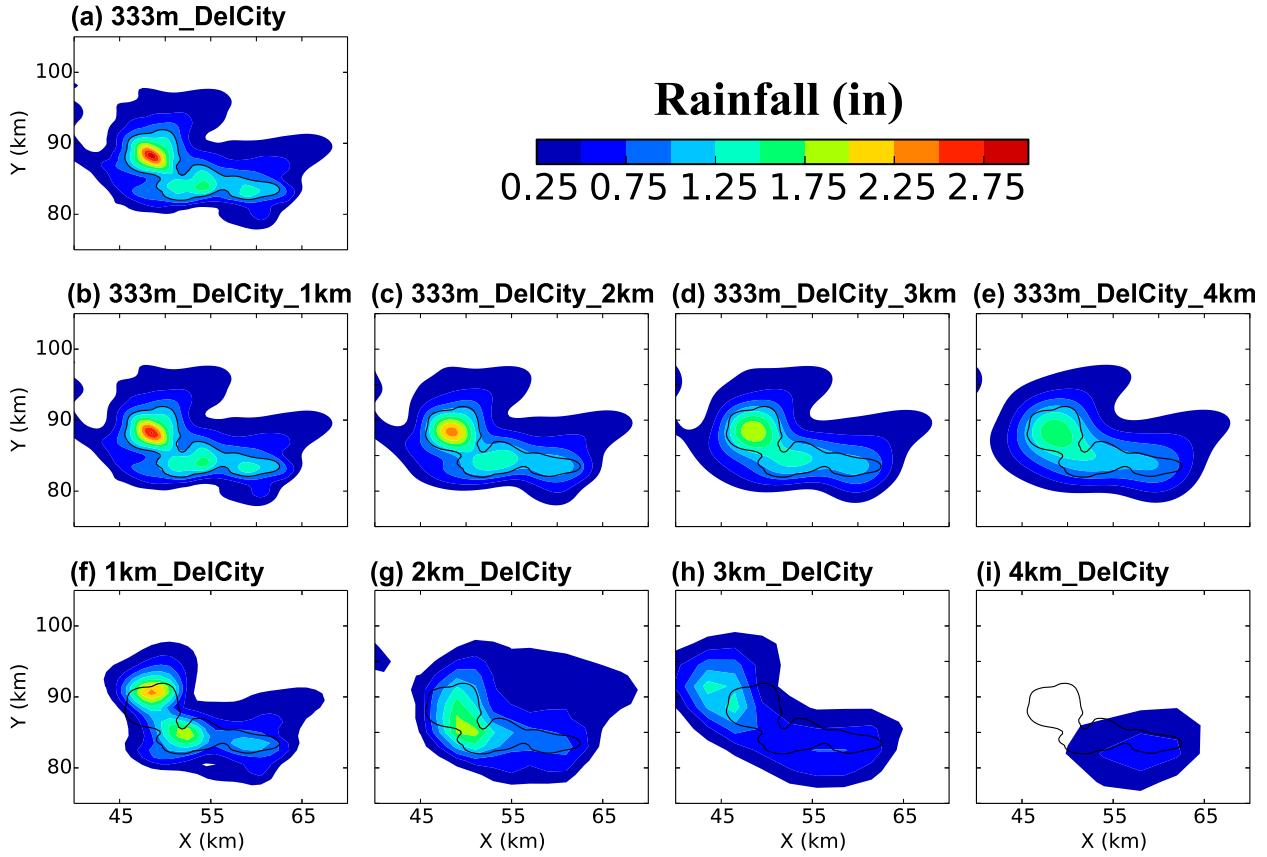


FIG. 17. As in Fig. 16, but for DelCity simulations.

dry environmental air (e.g., [Frame and Markowski 2010](#)). Increased entrainment in the simulations with drier environments could therefore be enhancing grid spacing sensitivity via the first two mechanisms involving the updraft, and/or in other ways. For example, entrainment by downdrafts and consequent sensitivities of microphysical processes, downward mass flux, and cold pool intensity may also contribute substantially to the grid spacing dependences. Regardless of their relative influence on grid spacing sensitivity, both updraft width (e.g., [Kirkpatrick et al. 2007, 2009](#)) and entrainment are strongly influenced by the storm environment. This again underscores the potential value of identifying relationships between storm environment and grid spacing dependence to anticipating errors in operational forecasts and research simulations.

*f. Impacts of microphysical and turbulence parameterization schemes on grid spacing sensitivity*

Microphysical and turbulence schemes, which are critical to modeling convective storms, are known to

exhibit substantial grid spacing dependencies ([section 1](#)). To the extent that the latter vary from scheme to scheme, so too will the grid spacing dependences of forecasts and simulations. We therefore investigated how the choices for these schemes impact the grid spacing sensitivity of supercells.

It is important to consider that in the present experimental framework, changing the physics schemes used in both the truth and coarse simulations could result in changes in the “true” storm characteristics that themselves substantially modulate the grid spacing dependence (e.g., narrower updraft). Changes in grid spacing dependence would not then be solely attributable to differences in the parameterization schemes. We therefore adopted an imperfect-model framework to assess the sensitivity of grid spacing dependence to the treatment of subgrid processes. The only difference between the new and original experiments is that either the turbulence or microphysical parameterization is changed in the coarse simulations.

Hypothesizing that the relatively poor performance of 4km\_ElReno and (especially) 4km\_DelCity resulted from the 1.5-order TKE closure underestimating

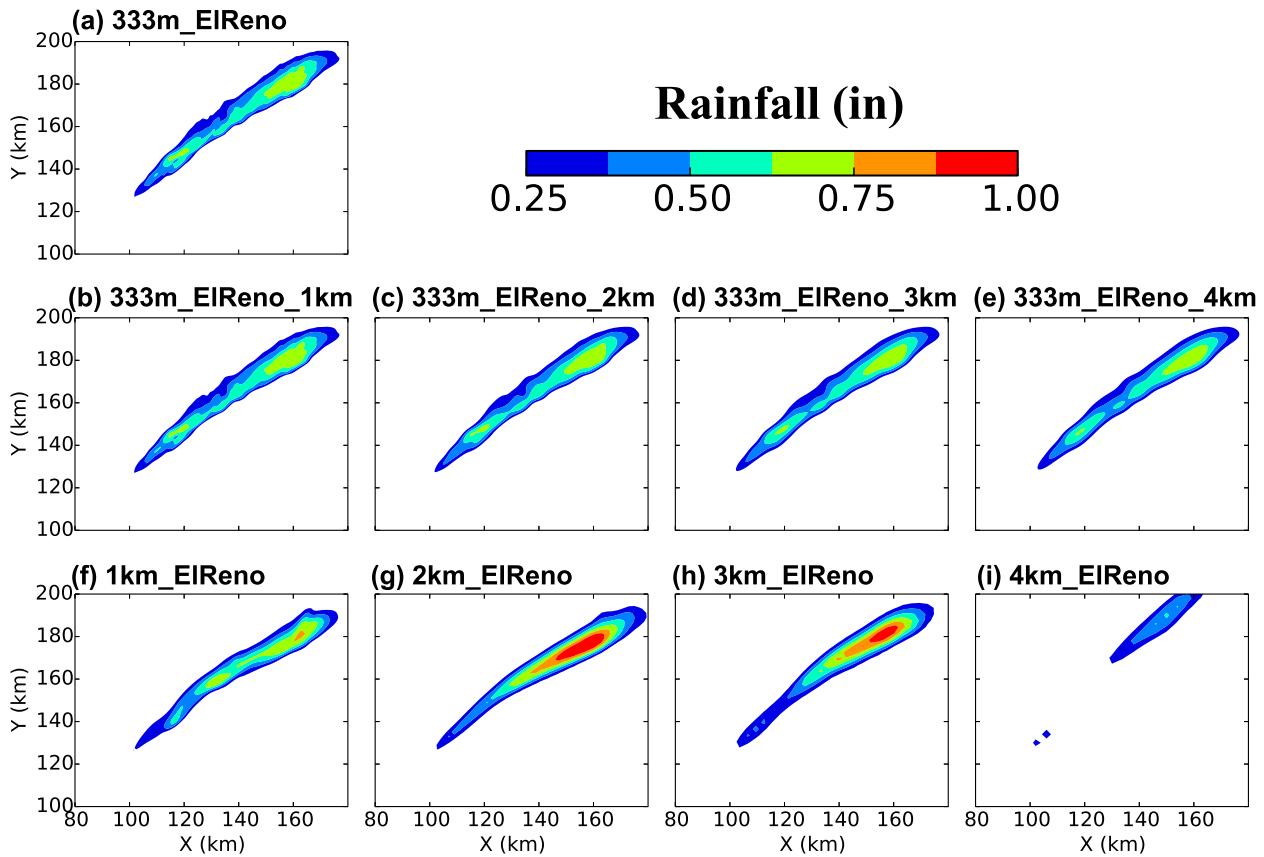


FIG. 18. As in Fig. 16, but for ElReno simulations.

subgrid vertical fluxes in the boundary layer, we repeated both simulations using either the Mellor–Yamada–Nakanishi–Niino (MYNN; Nakanishi and Niino 2004) or the Mellor–Yamada–Janjic (MYJ; Mellor and Yamada 1982; Janjić 1994) PBL scheme to represent vertical mixing, and diagnosing the horizontal mixing length from horizontal deformation (not shown). Surprisingly, storm demise was neither prevented nor delayed in the modified DelCity simulations, and was actually precipitated in the ElReno simulations. These results provide further evidence that 4-km grid spacing is too coarse to reliably resolve critical supercell processes, at least with current turbulence parameterization schemes, and may therefore be inappropriate for WoF operations. It is possible, however, that use of a land surface scheme and surface layer physics in real-world NWP mitigates some of the grid spacing sensitivity.

Repeating the coarse WK82 and ElReno simulations with the Smagorinsky scheme rather than 1.5-order TKE closure (experiment names appended with “\_SMAG”) or the fully two-moment NSSL Ziegler Variable Density scheme (Ziegler 1985; Mansell et al.

2010) instead of the Thompson microphysics (“\_ZVD”) produced substantial changes in both cases (Fig. 23). It should be noted, however, that the trends in the WK82\_SMAG and WK82\_ZVD convergence curves over the coarse (1–4 km) grid spacings do not appreciably change from WK82. This suggests much of the sensitivity of grid spacing dependence in those experiments results from resolution-independent model biases due to the use of different physics than in the truth simulations. This is not the case in the ElReno physics experiments. Most notably, the supercell rapidly dies in 3km\_ELRENO\_SMAG and 4km\_ELRENO\_SMAG. This result and the decay of the storm in the 4-km ElReno simulations that used a PBL scheme indicate very strong sensitivity to turbulence parameterization at larger grid spacings.

The above experiments reveal that parameterization scheme choices can substantially impact grid spacing dependence, and that these impacts will vary from case to case. The results, while preliminary, strongly motivate further exploration of this topic, especially the roles of other types of parameterization (e.g., surface physics schemes).

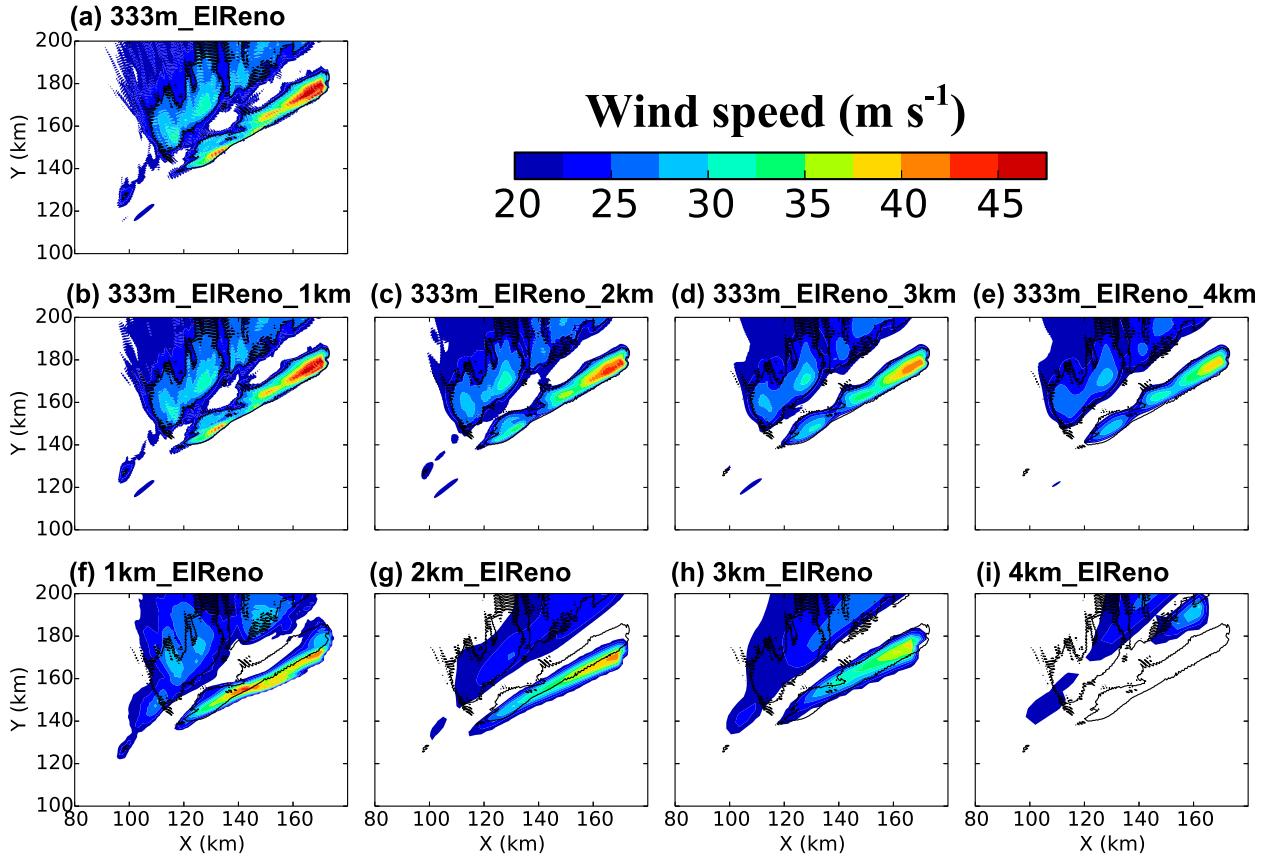


FIG. 19. Horizontal lowest-level wind speed ( $\text{m s}^{-1}$ ) composites for ElReno simulations: (a) 333m\_ElReno, (b) 333m\_ElReno\_1km, (c) 333m\_ElReno\_2km, (d) 333m\_ElReno\_3km, (e) 333m\_ElReno\_4km, (f) 1km\_ElReno, (g) 2km\_ElReno, (h) 3km\_ElReno, and (i) 4km\_ElReno. The 333m\_ElReno 30  $\text{m s}^{-1}$  contour (black dashed curves) is plotted in all panels for reference.

#### *g. Impact of resolution errors in initial condition only*

Forecast errors can arise from missing scales in both the initial condition and the model grid. To distinguish the significance of these two error sources, we filtered each control simulation at  $t = 30$  min as if preparing to restart the simulation on a 4-km grid, but instead restarted on the original 333-m grid. Thus, all errors in the subsequent simulations arise from the removal of wavelengths  $<8$  km from the initial condition. Comparing these simulations (“\_4kmIC”) to the control simulations using the same metrics as above (some of which are shown in Figs. 24 and 25) reveals that scales unresolvable in the 4-km initial condition are rapidly and reasonably accurately restored as the supercells evolve. While some significant errors occur, for example, overly strong low-level updraft and rotation (and consequently overpredicted rightward propagation) in ElReno\_4kmIC, the much higher quality of the new simulations relative to the original 4-km simulations indicates that most of the error growth in our coarse experiments results not from poor initialization of smaller

scales, but from the inability of the coarser grids to represent finescale processes that are critical to larger scales.

The fidelity with which smaller scales are restored in these experiments has potentially important implications for supercell predictability. It suggests that increasing model resolution can substantially improve short-term forecasts of supercells even when finer scales are not initialized well. That is, upscale growth of small-scale errors in the initial condition may not be a predominant source of large-scale error on  $O(1)$  h time scales. This is encouraging from a practical predictability standpoint given the current lack of intrastorm observations on scales  $<O(1)$  km. However, it is critical to consider that mesoscale errors can rapidly grow downscale, degrading representation of operationally important aspects of storm evolution (e.g., Cintineo and Stensrud 2013).<sup>5</sup>

<sup>5</sup> While upscale error growth was the focus of early atmospheric predictability studies, the importance of downscale error growth is increasingly recognized (e.g., Rotunno and Snyder 2008; Durran and Gingrich 2014).

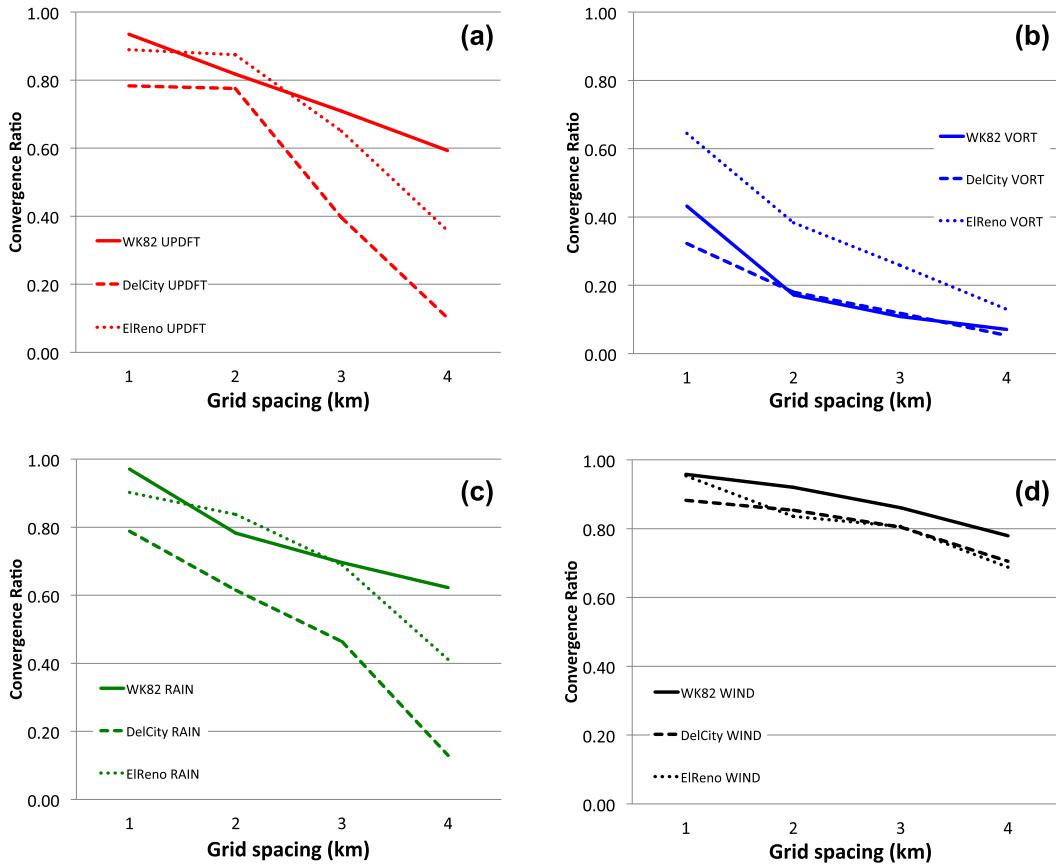


FIG. 20. Convergence plots summarizing grid spacing dependencies for WK82 (solid), DelCity (dashed), and ElReno (dotted) simulations: (a) maximum updraft speed, (b) maximum  $\zeta < 2$  km AGL, (c) maximum 5-min rainfall, and (d) maximum lowest-level wind speed.

Thus, limited observations of storm environments on the scales necessary to capture important heterogeneities (e.g., abrupt shifts in upper-level flow near jet streaks; dryline moisture gradients; and outflow boundaries from other storms) could strongly constrain the practical predictability of supercells even once substantial improvements have been made to NWP model physics and resolution. Future simulation studies are required to explore the relative contributions of different error sources (limited resolution, model physics and numerics, and small- and large-scale initialization errors) in order to help prioritize efforts to improve short-term convective forecasts.

#### 4. Conclusions

Computational and data storage constraints will force trade-offs between grid spacing and other ensemble forecast system parameters (such as ensemble size and physics scheme complexity) for the foreseeable future. Optimizing these trade-offs and properly interpreting and calibrating forecasts requires thorough understanding of

convective grid spacing sensitivity, including impacts of storm environment and parameterization schemes. The present study uses idealized simulations to explore the grid spacing dependence of supercells, with special emphasis on  $O(1)$  h predictability of thunderstorm hazards.

While 4-km horizontal grid spacing is commonly considered to be “convection allowing,” our results suggest it is only barely so in some scenarios, at least within the idealized framework adopted herein. Premature storm demise occurred in some of our 4-km simulations, and convective development was often considerably delayed. The ability of the model to resolve critical supercell processes improves considerably when grid spacing is reduced to 3 km. In particular, storm propagation (and therefore track) as well as general trends and locations of strong winds, heavy rain, and intense low-level rotation are relatively well captured. Thus, 3 km may be an approximate upper bound on the grid spacing required to realize the overarching goal of Warn-on-Forecast: ensemble forecasting systems that enable reliable prediction of convective hazards at finer

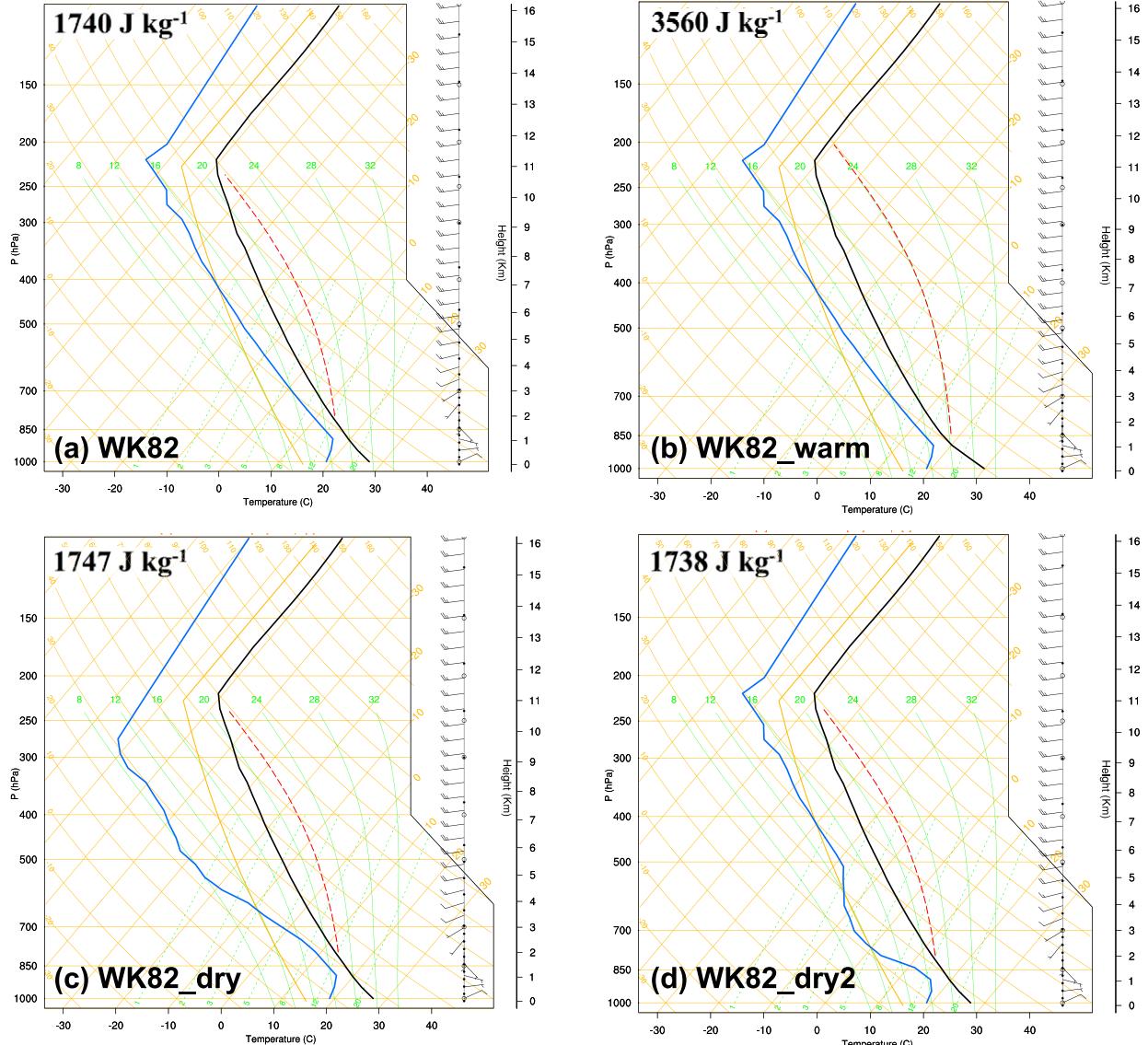


FIG. 21. Soundings used in (a) WK82, (b) WK82\_warm, (c) WK82\_dry, and (d) WK82\_dry2 experiments.

spatiotemporal scales and longer lead times than can currently be achieved. Reducing the grid spacing from 3 to 2 km yields smaller, but still substantial, improvements. Grid spacing near 1 km appears sufficient for useful prediction of episodes of rapid strengthening or weakening of low-level rotation. Operationally significant vorticity errors occurred even in our 1-km simulations, however, indicating that subkilometer grid spacing is required to maximize the spatiotemporal precision of tornado warnings. It should be noted that the above guidelines are subject to the “effective resolution” of the model (Skamarock 2004). Our experiments employed a dynamical core and numerical

schemes typical of current operational and experimental convection-allowing model systems in the United States; future improvements via the application of better numerical methods may relax grid spacing requirements to some degree.

While the considerable variance of convective grid spacing dependencies among previous (disparate) studies indicates strong sensitivity to some combination of storm mode, environment, and model configuration, the present study demonstrates that important differences can arise solely from changes in storm environment or parameterization schemes. In light of the strong impact of both of the latter on key

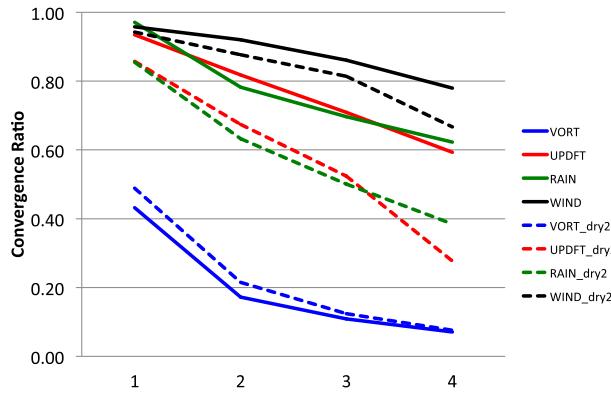


FIG. 22. Convergence plots for WK82 (solid) and WK82\_dry2 (dashed) simulations. Maximum  $\zeta < 2$  km AGL (blue), maximum updraft speed (red), maximum 5-min rainfall (green), and maximum lowest-level wind speed (black).

storm characteristics, this is not surprising. Our experiments suggest that updraft width and the degree of dry-air entrainment may strongly modulate the grid spacing dependence of storms. The substantial impact of physics scheme choice on grid spacing dependence further motivates the development and adoption of scale-aware schemes. The strong variability of grid spacing dependence indicates efforts to calibrate

forecast output will be limited unless robust relationships are found between grid spacing dependence and physics scheme characteristics as well as observable characteristics of storms and their environment. In light of the spatiotemporally complex errors that arose in our coarse simulations, calibration methods would need to be more sophisticated than simple functions of grid spacing in order to accurately recover unresolved scales that are important to Warn-on-Forecast.

Much of the error arising from limited grid spacing in our simulations resulted from poorly resolving smaller-scale processes that would otherwise have affected larger-scale storm evolution. This highlights the importance of upscale error growth in numerical supercell prediction. On the other hand, the relatively accurate recovery of smaller scales that were filtered out of the initial conditions of some of our 333-m simulations suggests that downscale interactions in supercells constrain smaller-scale errors to a greater degree than is obvious a priori. Thus, reliable prediction of sub-kilometer scales in supercells may be achievable even without dramatic advances in intrastorm observations. Additional investigation of the sensitivity of supercell evolution to initial condition resolution is required, in

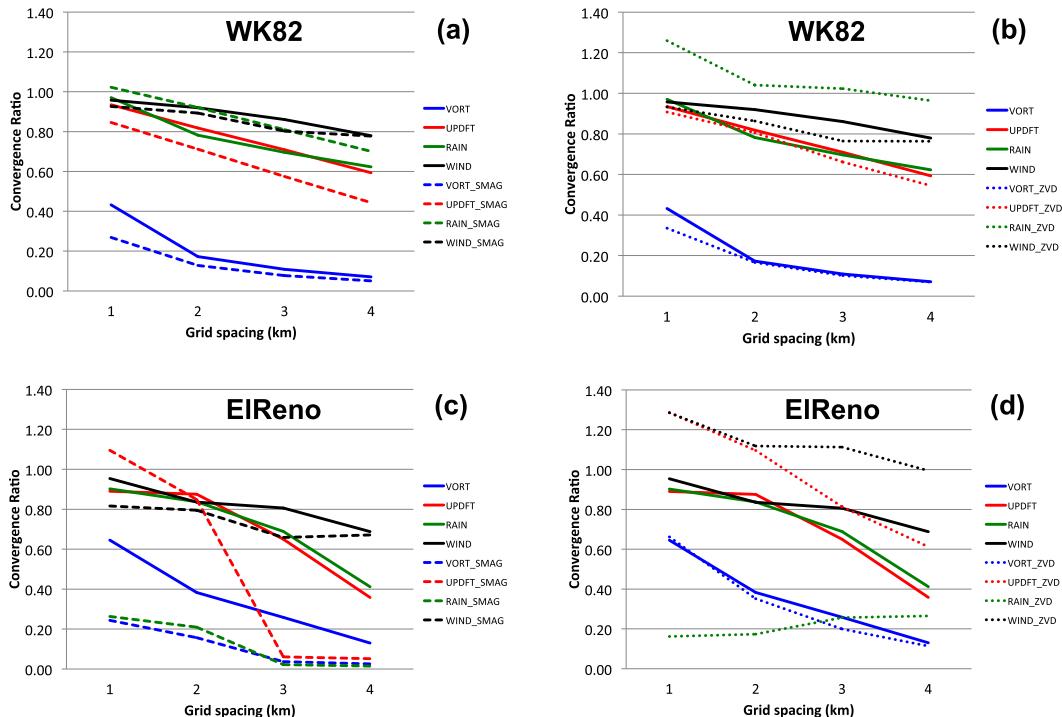


FIG. 23. Convergence plots summarizing grid spacing dependencies for (a),(b) WK82 and (c),(d) ElReno simulations using default settings (solid) vs (a),(c) Smagorinsky turbulence or (b),(d) ZVD microphysics (dashed). Maximum  $\zeta < 2$  km AGL (blue), maximum updraft speed (red), maximum 5-min rainfall (green), and maximum lowest-level wind speed (black).

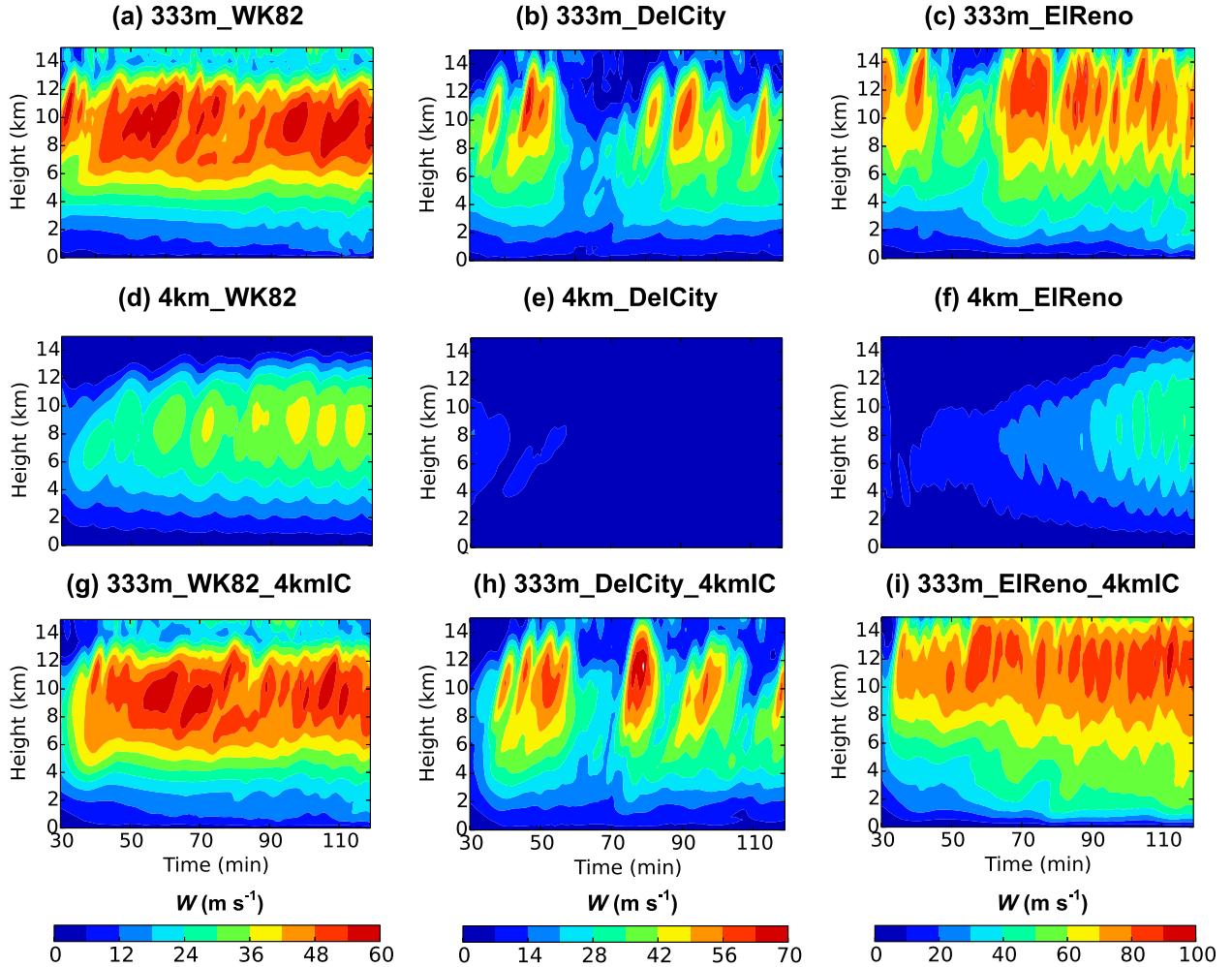
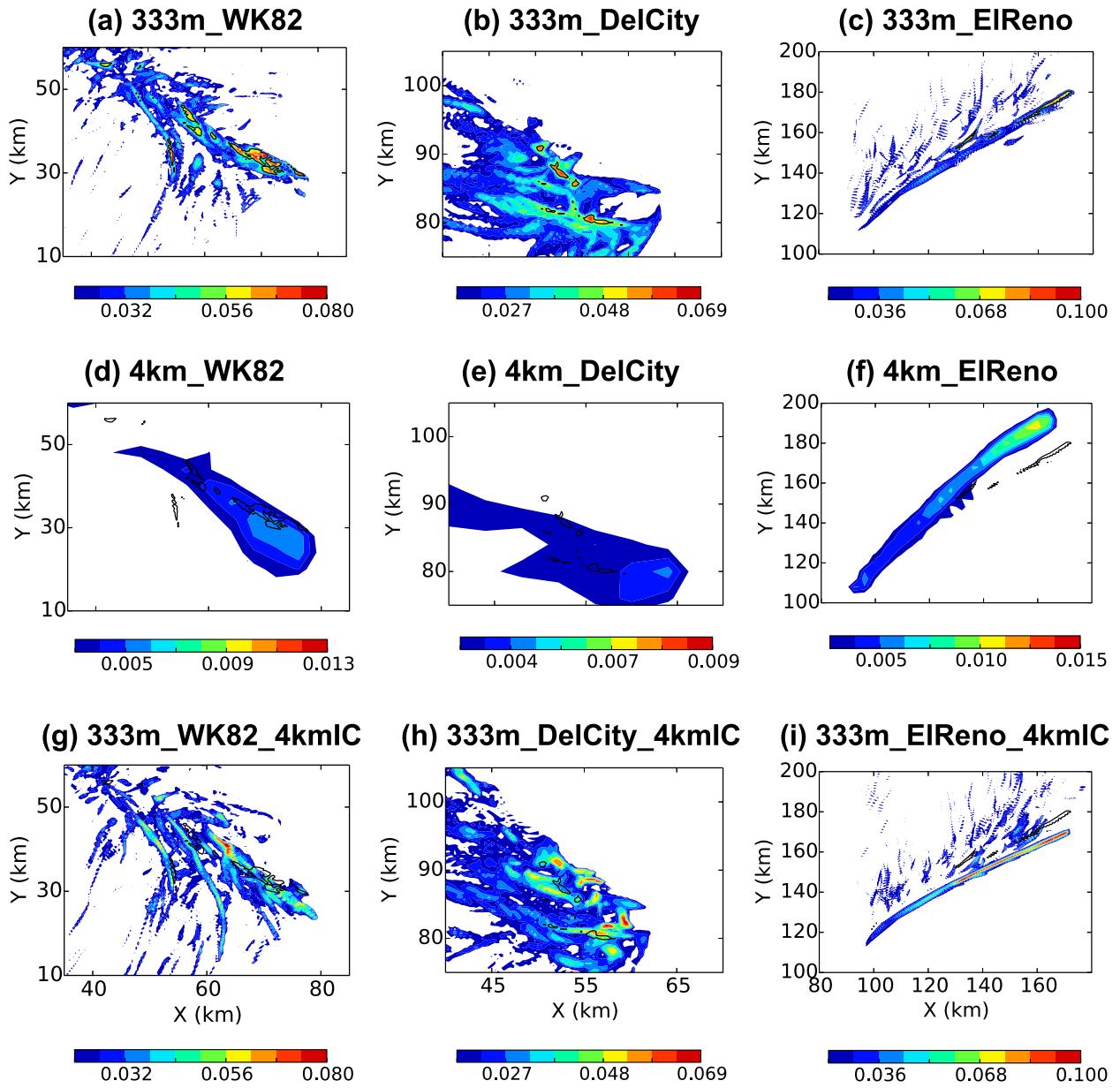


FIG. 24. Time-height  $w$  composites for (a) 333m\_WK82, (b) 333m\_DelCity, (c) 333m\_ElReno, (d) 4km\_WK82, (e) 4km\_DelCity, (f) 4km\_ElReno, (g) 333m\_WK82\_4kmIC, (h) 333m\_DelCity\_4kmIC, and (i) 333m\_ElReno\_4kmIC.

part to elucidate the importance of this error source relative to errors in the model and the analyzed storm environment. The first author plans to further explore this topic using both simulations and real case studies.

While the main conclusions from our deterministic experiments should generally extend to the ensemble framework, the grid spacing sensitivities identified here might be considerably modulated by use of a typical real-data model configuration, particularly of an inhomogeneous base state and surface physics. For example, the premature storm demise in some of our 4-km simulations may not have occurred had an external vertical motion forcing mechanism been included, such as a dryline or cloud-induced differential surface heating. Model and initial condition errors arising from sources other than finite grid resolution (e.g., physics scheme biases) may further modify these sensitivities.

Real-world grid spacing sensitivity experiments are therefore required to complement the idealized experiments presented here. Several convection-allowing model studies have already systematically examined the grid spacing dependence of real convection, but primarily at longer forecast lead times than required for convective warning operations (e.g., Kain et al. 2008; Schwartz et al. 2009; Johnson et al. 2013; Vandenberg et al. 2014). These studies have generally found inconsistent improvements from decreasing grid spacing below 4 km, but have often (reasonably) hypothesized that other model errors as well as initial/boundary condition errors dominated resolution errors at the long forecast lead times. Comparably little exploration has been done for real  $O(1)$  h forecasts, however, on which time scales the goal is to predict detailed evolution of individual storms rather than regional storm characteristics (e.g., mode, coverage, and general intensity).

FIG. 25. As in Fig. 24, but for horizontal composites of  $\zeta$  ( $\text{s}^{-1}$ )  $< 2$  km AGL.

Several efforts are under way within the NSSL WoF group to examine real-data supercell grid spacing dependence in an ensemble framework. Grid spacing sensitivity studies of other convective modes (e.g., squall lines) are also needed. Finally, perfect-model (apart from resolution) simulations like those presented herein but that use full-physics configurations (including a land surface model) and realistic initial/boundary conditions may provide insights beyond what purely idealized or real-world experiments can provide.

**Acknowledgments.** This work was prepared by the authors with funding provided by the National Science Foundation Grant AGS-1062932, 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 authors and do not necessarily reflect the views of the National Science Foundation, NOAA, or the U.S. Department of Commerce. We gratefully acknowledge the very

constructive comments of three anonymous reviewers. We thank Louis Wicker for informally reviewing an early version of this paper, and Michael Coniglio, Evgeni Fedorovich, and Louis Wicker for helpful discussions. Finally, M. Flora thanks Robert Conrick for continually lending a helpful ear during the 2014 NWC REU program.

## REFERENCES

- Adlerman, E. J., and K. K. Droegemeier, 2002: The sensitivity of numerically simulated cyclic mesocyclogenesis to variations in model physical and computational parameters. *Mon. Wea. Rev.*, **130**, 2671–2691, doi:10.1175/1520-0493(2002)130<2671:TSONSC>2.0.CO;2.
- Benjamin, S. G., and Coauthors, 2004: An hourly assimilation–forecast cycle: The RUC. *Mon. Wea. Rev.*, **132**, 495–518, doi:10.1175/1520-0493(2004)132<0495:AHACTR>2.0.CO;2.
- Bryan, G. H., and H. Morrison, 2012: Sensitivity of a simulated squall line to horizontal resolution and parameterization of microphysics. *Mon. Wea. Rev.*, **140**, 202–225, doi:10.1175/MWR-D-11-00046.1.
- , J. C. Wyngaard, and J. M. Fritsch, 2003: Resolution requirements for the simulation of deep moist convection. *Mon. Wea. Rev.*, **131**, 2394–2416, doi:10.1175/1520-0493(2003)131<2394:RRFTSO>2.0.CO;2.
- Cintineo, R. M., and D. J. Stensrud, 2013: On the predictability of supercell thunderstorm evolution. *J. Atmos. Sci.*, **70**, 1993–2011, doi:10.1175/JAS-D-12-0166.1.
- Corfidi, S. F., 2003: Cold pools and MCS propagation: Forecasting the motion of downwind-developing MCSs. *Wea. Forecasting*, **18**, 997–1017, doi:10.1175/1520-0434(2003)018<0997:CPAMPF>2.0.CO;2.
- Durran, D. R., and M. Gingrich, 2014: Atmospheric predictability: Why butterflies are not of practical importance. *J. Atmos. Sci.*, **71**, 2476–2488, doi:10.1175/JAS-D-14-0007.1.
- Fiori, E., A. Parodi, and F. Siccardi, 2010: Turbulence closure parameterization and grid spacing effects in simulated supercell storms. *J. Atmos. Sci.*, **67**, 3870–3890, doi:10.1175/2010JAS3359.1.
- Frame, J. W., and P. M. Markowski, 2010: Numerical simulations of radiative cooling beneath the anvils of supercell thunderstorms. *Mon. Wea. Rev.*, **138**, 3024–3047, doi:10.1175/2010MWR3177.1.
- Gage, K. S., 1979: Evidence for a  $k^{-5/3}$  law inertial range in mesoscale two-dimensional turbulence. *J. Atmos. Sci.*, **36**, 1950–1954, doi:10.1175/1520-0469(1979)036<1950:EFALIR>2.0.CO;2.
- Janjić, Z. I., 1994: The step-mountain Eta coordinate model: Further developments of the convection, viscous sublayer, and turbulence closure schemes. *Mon. Wea. Rev.*, **122**, 927–945, doi:10.1175/1520-0493(1994)122<0927:TSMECM>2.0.CO;2.
- Johnson, A., X. Wang, F. Kong, and M. Xue, 2013: Object-based evaluation of the impact of horizontal grid spacing on convection-allowing forecasts. *Mon. Wea. Rev.*, **141**, 3413–3425, doi:10.1175/MWR-D-13-00027.1.
- Kain, J. S., and Coauthors, 2008: Some practical considerations regarding horizontal resolution in the first generation of operational convection-allowing NWP. *Wea. Forecasting*, **23**, 931–952, doi:10.1175/WAF2007106.1.
- Kirkpatrick, J. C., E. W. McCaul, and C. Cohen, 2007: The motion of simulated convective storms as a function of basic environmental parameters. *Mon. Wea. Rev.*, **135**, 3033–3051, doi:10.1175/MWR3447.1.
- , —, and —, 2009: Variability of updraft and downdraft characteristics in a large parameter space study of convective storms. *Mon. Wea. Rev.*, **137**, 1550–1561, doi:10.1175/2008MWR2703.1.
- 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, doi:10.1175/2009JAS2965.1.
- Markowski, P. M., Y. Richardson, M. Majcen, J. Marquis, and J. Wurman, 2011: Characteristics of the wind field in three nontornadic low-level mesocyclones observed by the Doppler on Wheels radars. *Electron. J. Severe Storms Meteor.*, **6** (3). [Available online at <http://www.ejssm.org/ojs/index.php/ejssm/article/viewArticle/75>.]
- Mellor, G. L., and T. Yamada, 1982: Development of a turbulence closure model for geophysical fluid problems. *Rev. Geophys.*, **20**, 851–875, doi:10.1029/RG020i004p00851.
- Nakanishi, M., and H. Niino, 2004: An improved Mellor–Yamada level 3 model with condensation physics: Its design and verification. *Bound.-Layer Meteor.*, **112**, 1–31, doi:10.1023/B:BOUN.0000020164.04146.98.
- Noda, A., and H. Niino, 2003: Critical grid size for simulating convective storms: A case study of the Del City supercell storm. *Geophys. Res. Lett.*, **30**, 1844, doi:10.1029/2003GL017498.
- Okubo, K., 1970: Horizontal dispersion of floatable particles in the vicinity of velocity singularities such as convergences. *Deep-Sea Res.*, **17**, 445–454.
- Potvin, C. K., and L. J. Wicker, 2013: Assessing ensemble forecasts of low-level supercell rotation within an OSSE framework. *Wea. Forecasting*, **28**, 940–960, doi:10.1175/WAF-D-12-0012.1.
- Raymond, W. H., 1988: High-order low-pass implicit tangent filters for use in finite area calculations. *Mon. Wea. Rev.*, **116**, 2132–2141, doi:10.1175/1520-0493(1988)116<2132:HTLITF>2.0.CO;2.
- Richardson, Y. P., K. K. Droegemeier, and R. P. Davies-Jones, 2007: The influence of horizontal environmental variability on numerically simulated convective storms. Part I: Variations in vertical shear. *Mon. Wea. Rev.*, **135**, 3429–3455, doi:10.1175/MWR3463.1.
- Rotunno, R., and J. B. Klemp, 1982: The influence of the shear-induced pressure gradient on thunderstorm motion. *Mon. Wea. Rev.*, **110**, 136–150, doi:10.1175/1520-0493(1982)110<0136:TIOTSI>2.0.CO;2.
- , and C. Snyder, 2008: A generalization of Lorenz's model for the predictability of flows with many scales of motion. *J. Atmos. Sci.*, **65**, 1063–1076, doi:10.1175/2007JAS2449.1.
- 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, doi:10.1175/2009MWR2924.1.
- Skamarock, W. C., 2004: Evaluating mesoscale NWP models using kinetic energy spectra. *Mon. Wea. Rev.*, **132**, 3019–3032, doi:10.1175/MWR2830.1.
- , and Coauthors, 2008: A description of the Advanced Research WRF version 3. NCAR Tech. Note NCAR/TN-475+STR, 113 pp. [Available online at [http://www.mmm.ucar.edu/wrf/users/docs/arw\\_v3\\_bw.pdf](http://www.mmm.ucar.edu/wrf/users/docs/arw_v3_bw.pdf).]
- Stensrud, D. J., and Coauthors, 2009: Convective-scale warn-on-forecast system: A vision for 2020. *Bull. Amer. Meteor. Soc.*, **90**, 1487–1499, doi:10.1175/2009BAMS2795.1.

- , and Coauthors, 2013: Progress and challenges with warn-on-forecast. *Atmos. Res.*, **123**, 2–16, doi:[10.1016/j.atmosres.2012.04.004](https://doi.org/10.1016/j.atmosres.2012.04.004).
- Thompson, G., R. M. Rasmussen, and K. Manning, 2004: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part I: Description and sensitivity analysis. *Mon. Wea. Rev.*, **132**, 519–542, doi:[10.1175/1520-0493\(2004\)132<0519:EFOWPU>2.0.CO;2](https://doi.org/10.1175/1520-0493(2004)132<0519:EFOWPU>2.0.CO;2).
- , P. R. Field, R. M. Rasmussen, and W. R. Hall, 2008: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow parameterization. *Mon. Wea. Rev.*, **136**, 5095–5115, doi:[10.1175/2008MWR2387.1](https://doi.org/10.1175/2008MWR2387.1).
- 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, doi:[10.1175/WAF-D-14-00011.1](https://doi.org/10.1175/WAF-D-14-00011.1).
- Verrelle, A., D. Ricard, and C. Lac, 2015: Sensitivity of high-resolution idealized simulations of thunderstorms to horizontal resolution and turbulence parameterization. *Quart. J. Roy. Meteor. Soc.*, **141**, 433–448, doi:[10.1002/qj.2363](https://doi.org/10.1002/qj.2363).
- Weisman, M. L., and J. B. Klemp, 1982: The dependence of numerically simulated convective storms on vertical wind shear and buoyancy. *Mon. Wea. Rev.*, **110**, 504–520, doi:[10.1175/1520-0493\(1982\)110<0504:TDONSC>2.0.CO;2](https://doi.org/10.1175/1520-0493(1982)110<0504:TDONSC>2.0.CO;2).
- , W. C. Skamarock, and J. B. Klemp, 1997: The resolution dependence of explicitly modeled convective systems. *Mon. Wea. Rev.*, **125**, 527–548, doi:[10.1175/1520-0493\(1997\)125<0527:TRDOEM>2.0.CO;2](https://doi.org/10.1175/1520-0493(1997)125<0527:TRDOEM>2.0.CO;2).
- Weiss, J., 1991: The dynamics of enstrophy transfer in two-dimensional hydrodynamics. *Physica D*, **48**, 273–294, doi:[10.1016/0167-2789\(91\)90088-Q](https://doi.org/10.1016/0167-2789(91)90088-Q).
- Wyngaard, J. C., 2004: Toward numerical modeling in the “terra incognita.” *J. Atmos. Sci.*, **61**, 1816–1826, doi:[10.1175/1520-0469\(2004\)061<1816:TNMITT>2.0.CO;2](https://doi.org/10.1175/1520-0469(2004)061<1816:TNMITT>2.0.CO;2).
- Ziegler, C. L., 1985: Retrieval of thermal and microphysical variables in observed convective storms. Part I: Model development and preliminary testing. *J. Atmos. Sci.*, **42**, 1487–1509, doi:[10.1175/1520-0469\(1985\)042<1487:ROTAMV>2.0.CO;2](https://doi.org/10.1175/1520-0469(1985)042<1487:ROTAMV>2.0.CO;2).
- , E. R. Mansell, J. M. Straka, D. R. MacGorman, and D. W. Burgess, 2010: The impact of spatial variations of low-level stability on the life cycle of a simulated supercell storm. *Mon. Wea. Rev.*, **138**, 1738–1766, doi:[10.1175/2009MWR3010.1](https://doi.org/10.1175/2009MWR3010.1).