# Sensitivity of Supercell Simulations to Initial-Condition Resolution

COREY K. POTVIN

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

ELISA M. MURILLO

*School of Meteorology, University of Oklahoma, and National Weather Center Research Experiences for Undergraduates, Norman, Oklahoma, and University of Louisiana at Monroe, Monroe, Louisiana*

MONTGOMERY L. FLORA

*School of Meteorology, University of Oklahoma, Norman, Oklahoma*

DUSTAN M. WHEATLEY

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

(Manuscript received 28 March 2016, in final form 2 September 2016)

ABSTRACT

Observational and model resolution limitations currently preclude analysis of the smallest scales important to numerical prediction of convective storms. These missing scales can be recovered if the forecast model is integrated on a sufficiently fine grid, but not before errors are introduced that subsequently grow in scale and magnitude. This study is the first to systematically evaluate the impact of these initial-condition (IC) resolution errors on high-resolution forecasts of organized convection. This is done by comparing high-resolution supercell simulations generated using identical model settings but successively coarsened ICs. Consistent with the Warn-on-Forecast paradigm, the simulations are initialized with ongoing storms and integrated for 2 h. Both idealized and full-physics experiments are performed in order to examine how more realistic model settings modulate the error evolution.

In all experiments, scales removed from the IC (wavelengths $<$ 2, 4, 8, or 16 km) regenerate within 10–20 min of model integration. While the forecast errors arising from the initial absence of these scales become quantitatively large in many instances, the qualitative storm evolution is relatively insensitive to the IC resolution. It therefore appears that adopting much finer forecast (e.g., 250 m) than analysis (e.g., 3 km) grids for data assimilation and prediction would improve supercell forecasts given limited computational resources. This motivates continued development of mixed-resolution systems. The relative insensitivity to IC resolution further suggests that convective forecasting can be more readily advanced by improving model physics and numerics and expanding extrastorm observational coverage than by increasing intrastorm observational density.

## 1. Introduction

There have been numerous simulation studies of the predictability of organized convective storms. Focuses of these investigations include sensitivity to model grid spacing (e.g., Weisman et al. 1997; Adlerman and Droegemeier 2002; Bryan et al. 2003; Fiori et al. 2010; Bryan and Morrison 2012; Verrelle et al. 2015; Potvin and Flora 2015), prescribed small-scale (e.g., Hohenegger and Schär 2007; Zhang et al. 2016) and large-scale errors in the initial condition (Wandishin et al. 2008, 2010; Cintineo and Stensrud 2013; Durran and Weyn 2016), and data assimilation analysis errors arising from lack of low-level observations (e.g., Dong et al. 2011; Potvin and Wicker 2013a). Little has been done, however, to

*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

DOI: 10.1175/JAS-D-16-0098.1

examine the sensitivity of storm forecasts to initial-condition (IC) resolution. Impacts of limited IC resolution are of critical concern to the envisioned Warn-on-Forecast (WoF) paradigm (Stensrud et al. 2009, 2013), in which kilometer-scale ensemble forecasts will be initialized from analyses of storms and used to generate unprecedentedly high-resolution probabilistic convective hazard guidance.

If storm evolution is relatively insensitive to IC resolution, then sacrificing analysis grid resolution for forecast grid resolution could yield more accurate forecasts given limited computational resources. The simplest way to do this is to perform data assimilation on a relatively coarse grid, then interpolate to a finer grid on which the model is subsequently integrated over the desired forecast period. Another is to use mixed-resolution data assimilation methods to obtain finer analyses than can be achieved with single-resolution methods. Mixed-resolution approaches have long been used in variational data assimilation schemes (e.g., Courtier et al. 1994) and have more recently been developed for ensemble (e.g., Gao and Xue 2008; Rainwater and Hunt 2013) and hybrid variational–ensemble methods (e.g., Buehner et al. 2010; Kleist and Ide 2015). Combining the downscaling and mixed-resolution data assimilation approaches could be particularly powerful.

If, on the other hand, IC resolution errors substantially impact storm evolution, then the accuracy of convective-scale forecasts is strongly constrained by the resolution of intrastorm observations and of real-time data assimilation systems. Even in the limit of arbitrarily fine data assimilation grids, unobserved scales must still be generated by the model, either during the data assimilation (for cycled methods) or during the subsequent free model integration, and so the evolution of IC resolution error remains a critical consideration. Therefore, better understanding of the sensitivity of convective forecasts to IC resolution is needed to estimate practical predictability limits; inform design of next-generation observation, data assimilation, and prediction systems; and interpret numerical forecast output. In addition, improved understanding of the impact of small-scale errors helps illuminate the importance to storm evolution of the initial storm state versus the larger-scale environment.

Toward these ends, we conduct both idealized and full-physics supercell simulations to investigate the forecast impacts of IC resolution errors. Our experiments are designed such that limited IC resolution is essentially the only source of forecast error; the IC is otherwise perfect (apart from random perturbations, the necessity of which is explained later), as is the forecast model. To our knowledge, this is the first study to isolate the impact of IC resolution on the evolution of convection, apart from the set of experiments in Potvin and Flora (2015) that inspired the present investigation. Our results strongly support the viability of downscaling data assimilation analyses and/or using mixed-resolution data assimilation methods to initialize finer-scale forecasts, at least in the limit of a perfect model. The (perhaps surprising) lack of sensitivity to IC resolution exhibited in our experiments indicates that the evolution of many physically and societally important aspects of supercells, including low-level rotation and heavy rainfall, are primarily determined by larger scales.

It should be borne in mind that in addition to IC resolution errors, other operational forecast error sources (e.g., physics parameterization, finite grid resolution, large-scale analysis errors) will remain significant for decades to come. It is also important to note that the error growth modeled herein presumably underestimates the error growth within real convection due to the inability of the prescribed grid spacing (250 or 333 m) to resolve the finest scales important to convective evolution [$O$(10–100) m; Bryan et al. 2003]. For these two reasons, this study provides an upper-limit estimate of the predictability of storms initialized with missing scales. Thus, IC resolution requirements inferred from our experiments should be regarded as necessary, but not sufficient, conditions for accurate storm prediction.

The rest of the paper is organized as follows. Section 2 motivates our use of an ensemble simulation framework to investigate IC resolution errors and describes the model configuration and verification methods used for our simulations. Section 3 presents and discusses quantitative and qualitative analyses of the evolution of IC resolution errors in our experiments. Section 4 summarizes the major implications of the results and outlines avenues for future work.

## 2. Methodology

### a. *Motivation for an ensemble framework*

Convective storm evolution is known to be particularly sensitive to small-scale IC errors owing to the rapid upscale growth of perturbations within moist convection (e.g., Zhang et al. 2006; Hohenegger and Schär 2007). The rapid divergence of solutions initialized with very similar initial conditions strongly motivates adoption of an ensemble framework for our experiments. Figure 1 illustrates why this is so. The green curve in each schematic represents a forecast trajectory (simplified to two dimensions) for a control simulation that we label "TRUTH." The black (dashed) curve represents a

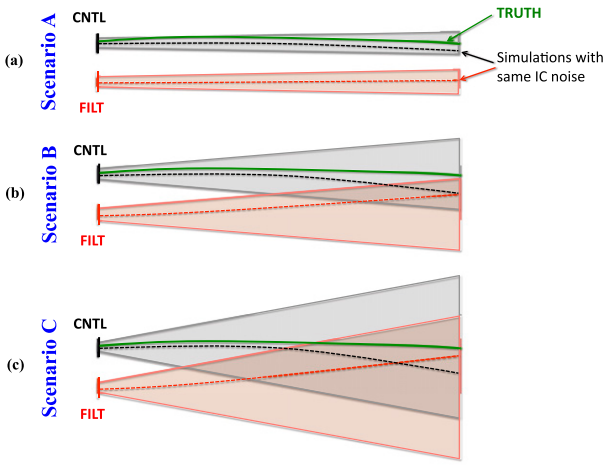


FIG. 1. Schematic illustrating the need to consider potential flow dependence of evolution of initial-condition resolution errors. (a) Scenario A: Filtered-IC (FILT) simulations (red) initialized from noisy realizations of TRUTH (green) exhibit qualitatively similar differences from their unfiltered (CNTL) counterparts (gray), rendering the deterministic framework adequate for characterizing the impact of initial-condition resolution errors. (b) Scenario B: Differences between FILT and CNTL pairs are in some instances small relative to the ranges of the CNTL and FILT solutions, requiring an ensemble approach to generalize the errors. (c) Scenario C: The solution is so sensitive to tiny initial-condition perturbations that the error correlation with initial-condition resolution is largely lost before the end of the forecast period.

forecast trajectory using the same IC as for TRUTH but with small-scale, low-amplitude noise added. The gray shaded area represents the solution space spanned by an ensemble of such forecasts that we call "CNTL." The red dashed curve represents a forecast trajectory initialized from the same IC (including noise) as the indicated CNTL forecast trajectory (black dashed curve), but with the IC resolution degraded prior to adding the perturbations. We call the ensemble of filtered-IC simulations "FILT." Now, consider three potential scenarios as the two ensembles evolve over the forecast period of interest. If, at all forecast times, the bias arising from the degraded IC resolution is large relative to the ensemble forecast spread arising from the random IC perturbations (scenario A; Fig. 1a), then any pair of corresponding simulations sampled from CNTL and FILT would qualitatively represent the differences between the two ensembles in general. A deterministic framework would then suffice for assessing the general impact of degraded IC resolution on forecasts of storms approximating those in TRUTH. Now, consider the scenario where the forecast bias is commensurate with the forecast spread (scenario B; Fig. 1b). The differences between a given pair of simulations in this case are highly dependent on the particular realization of noise added to the initial conditions. That is, the sensitivity to initial-condition resolution is so flow dependent in this scenario that comparing only a single pair of simulations, rather than the two ensembles in their entirety, could give a false impression of the systemic impact of IC resolution errors (i.e., forecast bias). Put another way: examining the impact of IC resolution errors on a single simulation is impractical in this case since the error evolution cannot be generalized beyond that one storm realization. Thus, this second scenario demonstrates that the sensitivity of supercells to random perturbations may partly obscure the IC resolution dependence in the deterministic framework. In the third scenario (scenario C; Fig. 1c), the evolution of small-scale errors is so rapid that the forecast and initial-condition errors become decoupled within the period of interest; that is, beyond some lead time, the impact of IC resolution is lost because of the limited intrinsic predictability (Lorenz 1969). Under this scenario, the deterministic framework would lead to false attribution to IC resolution errors of forecast differences arising largely or entirely from the sensitivity of convection to random IC perturbations. The eventual loss of the effect of IC resolution would only be detectable using the ensemble framework. Since the relative frequencies of scenario A (deterministic approach suitable) and scenarios B and C (deterministic approach not suitable) is uncertain a priori, we exercise caution and adopt the ensemble framework for this study. It will be shown at the beginning of section 3 that the impact of IC resolution errors in our simulations does indeed exhibit strong sensitivity to minor IC perturbations, making the ensemble approach necessary for investigating the evolution of IC resolution errors in forecasts of convective storms.

We initialize each CNTL (FILT) ensemble member by adding a different realization of random gridpoint perturbations to the TRUTH (filtered TRUTH) IC. Only the horizontal wind components and potential temperature are perturbed, and only below 10 km AGL; the perturbations are sampled from uniform distributions of $[-1, 1]\,\mathrm{m\,s^{-1}}$ and $[-0.5, 0.5]\,\mathrm{K}$, respectively. Since the role of the IC perturbations is solely to create the ensembles necessary to investigate IC resolution impacts in scenarios B and C, we chose the above noise characteristics rather arbitrarily. The only two requirements of the perturbations, which were confirmed to be met in our experiments a posteriori, were that they produce ensemble solution spaces that are 1) large enough for the mean error evolution to be reasonably general (critical for scenarios B and C) but 2) not so large that the impact of the IC noise dominate the impact of IC resolution error. Satisfying the latter requirement ensures that the analyzed

sensitivity to IC resolution is not unduly dependent upon the prescribed noise characteristics, since the errors arising from the latter largely cancel in the ensemble-mean metrics examined herein. It is confirmed in section 3a that the required error cancellation takes place. Moreover, repeating one of the sets of full-physics experiments described later (6MAY) with the IC perturbations inserted before rather than after filtering does not substantially impact the results, further demonstrating that the IC resolution sensitivity analyzed herein is not strongly tied to the initial noise characteristics. All CNTL and FILT ensembles comprise nine members; this ensemble size was found to produce sufficient cancellation of IC perturbation effects for the impact of IC resolution to be identified. The process for generating CNTL and FILT ensembles from TRUTH is depicted in Fig. 2.

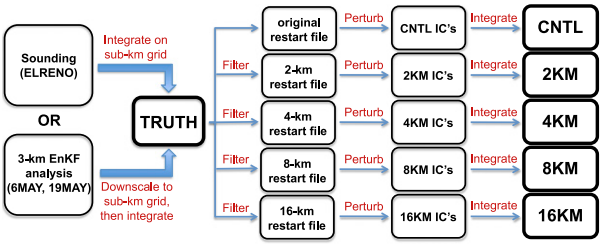


FIG. 2. Schematic summarizing the generation of the unfiltered (CNTL) and filtered-IC (2KM, 4KM, 8KM, 16KM) ensembles. A truth simulation (TRUTH) is initialized either by a warm bubble (idealized model framework) or by interpolating a real-world data assimilation analysis to a subkilometer grid (full-physics model framework). The WRF restart file generated at the desired ensemble initialization time is then filtered to remove IC scales below prescribed cutoff wavelengths. Next, random perturbations are added to the original and filtered restart files to generate ICs for each ensemble. Finally, the ensembles are integrated over the desired forecast period.

### *b. Motivation for a simulation framework*

Our experiments are designed to isolate the impact of IC resolution errors on forecasts. In practice, these errors would be modulated by errors in model physics and resolution, observations (including limited coverage), and data assimilation schemes. Our lack of knowledge of the characteristics of these various error sources and their interactions (both of which are, presumably, quite flow dependent) makes it difficult to distinguish their individual contributions to the total error in forecasts of real-world events. Consequently, while real-world experiments (in which forecasts are verified using observations of the real atmosphere) are better suited to exploring the total sensitivity of (contemporary) prediction systems, simulation experiments (in which forecasts are verified using the truth simulation from which they are generated) enable exploration of impacts of individual error sources. In addition, the lack of reliable atmospheric-state information in real-world analyses strongly limits the experimental parameter space that can be explored. For example, the impact on real-world forecasts of accurately initializing scales down to, say, 2 km cannot currently be examined since most of the atmospheric-state space cannot be well estimated at such fine scales. On the other hand, simulated atmospheric states are completely known, vastly increasing the range of hypotheses that can be tested.

The combination of the simulation and real-world approaches can provide much more insight into storm predictability than either alone. For example, if it is

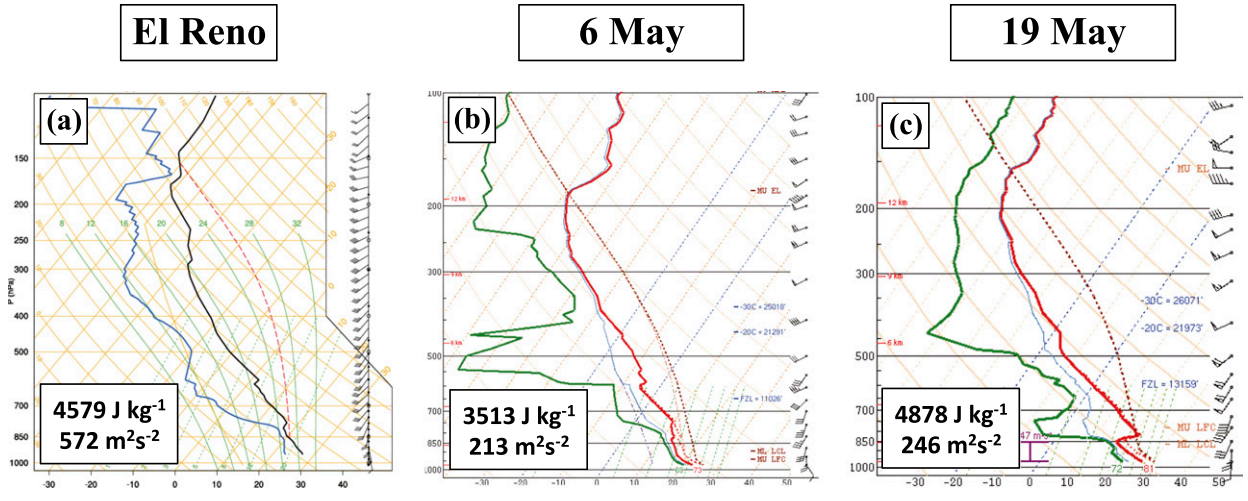


FIG. 3. Skew *T* plots for (a) El Reno, (b) 6 May 2015, and (c) 19 May 2013 simulations. Surface-based convective available potential energy (J kg$^{-1}$) and 0–3-km-AGL storm-relative helicity (m$^2$ s$^{-2}$) are indicated in the lower-left corner of each plot.

TABLE 1. ARW settings used in simulations on innermost grids.

| | ELRENO | 6MAY | 19MAY |
|---|---|---|---|
| Horizontal grid spacing (m) | 333 | 250 | 333 |
| Time step (s) | 1 | 1/2 | 2/3 |
| Time-integration scheme | Third-order Runge–Kutta | Third-order Runge–Kutta | Third-order Runge–Kutta |
| Horizontal/vertical advection | Fifth order/third order | Fifth order/third order | Fifth order/third order |
| Lateral/top boundary conditions | Open/Rayleigh damping layer | One-way nested/Rayleigh damping layer | One-way nested/Rayleigh damping layer |
| Turbulence parameterization | 1.5-order TKE closure | 1.5-order TKE closure | 1.5-order TKE closure |
| Microphysics parameterization | Thompson | Thompson | Thompson |
| Radiation parameterization | — | Rapid Radiative Transfer Model | Rapid Radiative Transfer Model |
| Surface-layer physics | — | Yonsei University | Mellor–Yamada–Janjić |
| Planetary boundary layer parameterization | — | — | — |
| Explicit numerical diffusion | — | — | — |

found that total errors in real-world forecasts scarcely decrease as IC resolution increases, this would suggest that contemporary model errors and/or large-scale analysis errors are dominating the effects of initially missing scales—something that could not have been known without performing both simulation and real-world experiments. While neither approach is fundamentally superior to the other, we adopt the simulation framework for our investigation since it is better suited to our goals of isolating and precisely evaluating the evolution of IC resolution errors.

### c. *Idealized simulation configuration*

Idealized simulations are generated using version 3.6.1 of the Advanced Research Weather Research and Forecasting (ARW) Model (Skamarock et al. 2008) with typical cloud model settings. The horizontal grid spacing and model time step are set to 333 m and 1 s, respectively. We use 50 vertical model levels, with the grid spacing increasing from ~100 m near the surface to ~700 m between 15 and 20 km AGL (model top). The model base state is provided by a Rapid Update Cycle (RUC; Benjamin et al. 2004) sounding analyzed near the 24 May 2011 El Reno, Oklahoma, tornadic supercell (Fig. 3a). Radiation and surface physics are neglected. The Thompson microphysics scheme (Thompson et al. 2004, 2008), which uses five hydrometeor categories and predicts two moments of the rain and cloud ice particle size distributions, is used. Turbulence is parameterized using the 1.5-order TKE closure. Additional model details are given in Table 1.

Convection is initiated in the truth simulation (TRUTH_ELRENO) by a thermal bubble with maximum potential temperature perturbation of 5 K. The control ensemble (CNTL_ELRENO) is initialized from TRUTH_ELRENO (as described in section 2a and Fig. 2) after 90 min of integration, by which time an intense supercell has developed, then integrated until time $t = 120$ min (all times in this paper are relative to the beginning of the control ensemble). The supercell remains strong during the entire forecast period (e.g., Figs. 4a,d) and produces several tornado-like vortices (TLVs; described later).[1] Filtered-IC ensembles are generated in the same way except that the WRF restart file is horizontally filtered prior to adding the random perturbations. As in Potvin and Flora (2015), we use the Raymond (1988) implicit tangent filter, which both is very efficient and has a sharp response function. The filter is empirically tuned to effect cutoff wavelengths (defined herein as producing a filter response of 0.5) of 2, 4, 8, and 16 km. The corresponding ensembles are labeled 2KM_ELRENO, 4KM_ELRENO, 8KM_ELRENO, and 16KM_ELRENO, respectively. We use the same nomenclature for the full-physics ensembles described in the next subsection as for the ELRENO ensembles. When not referring to one particular case, we omit the case identifier; for example, "2KM" refers to an ensemble (or ensembles) initialized with a 2-km cutoff wavelength (e.g., Fig. 2).

### d. *Full-physics simulation configuration*

To complement the idealized simulations, we additionally generate two sets of observationally constrained, full-physics simulations. The two full-physics TRUTH simulations are initialized by downscaling 3-km ensemble Kalman filter (EnKF; Evensen 1994; Snyder and Zhang 2003) member analyses of the 19 May 2013 and 6 May 2015 Oklahoma tornadic supercell events. Observed proximity soundings for the two cases are shown in Figs. 3b and 3c. The initial and lateral boundary conditions for the 3-km EnKF analyses are provided by

[1] In this study, a TLV is defined as a vortex that occurs within a mesocyclone and has vorticity $> 0.1\ s^{-1}$ at most model levels below $z = 500$ m.

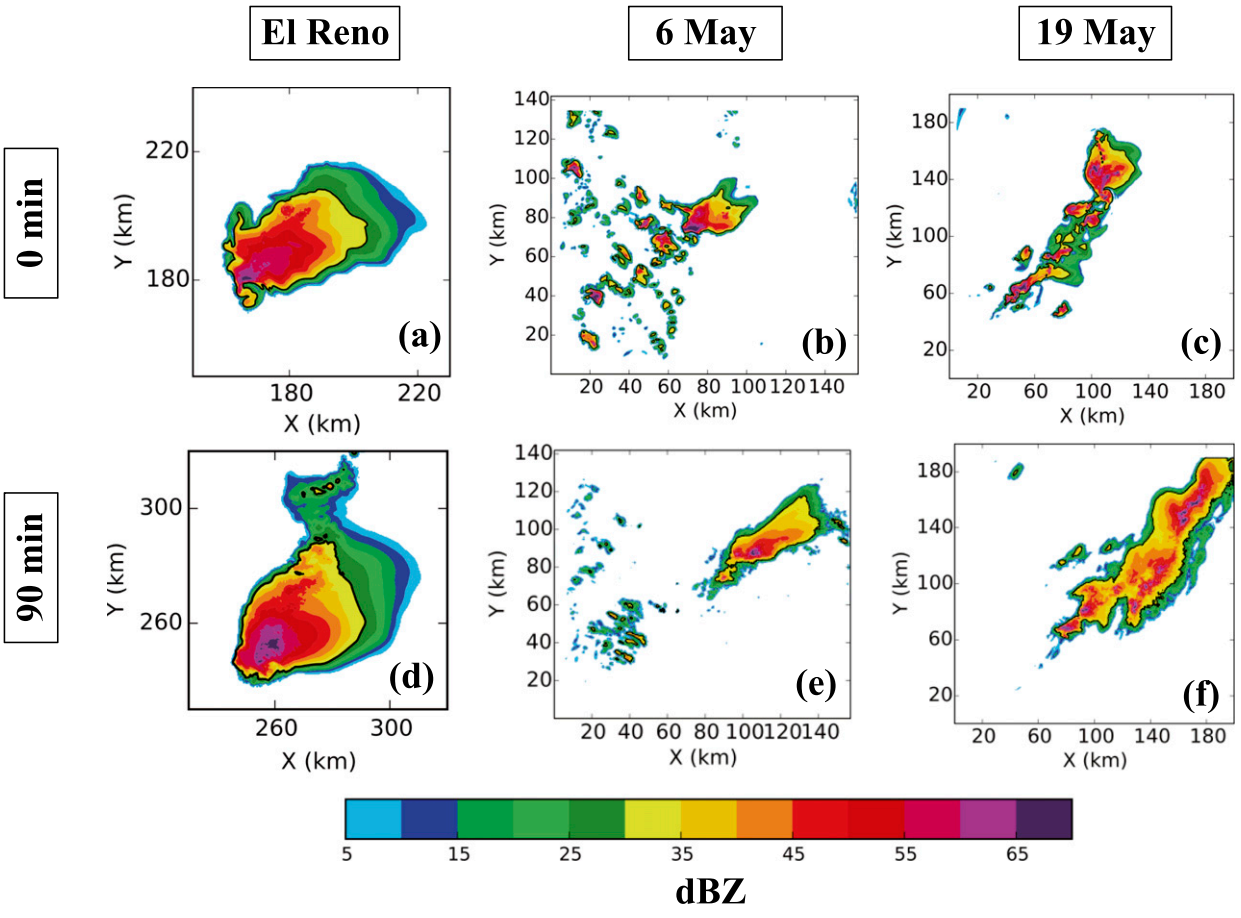


FIG. 4. Reflectivity (dBZ) at ~2 km AGL at (a)–(c) $t = 0$ and (d)–(f) $t = 90$ min in the (a),(d) El Reno, (b),(e) 6 May 2015, and (c),(f) 19 May 2013 TRUTH simulations.

the NSSL Experimental WoF System for Ensembles (NEWS-e), which assimilated conventional observations from the Meteorological Assimilation Data Ingest System (MADIS; Miller et al. 2007) onto a 15-km parent grid, within which a 3-km grid was one-way nested to spin up smaller-scale ensemble covariances in preparation for radar data assimilation [full details of the NEWS-e configuration appear in Wheatley et al. (2015) and Jones et al. (2016)]. We assimilate reflectivity and radial velocity from three nearby WSR-88Ds as well as Oklahoma Mesonet observations onto the 3-km grid using a WRF-coupled local ensemble transform Kalman filter (LETKF; Hunt et al. 2007) system configured similarly to the NEWS-e. The WRF-LETKF was adapted from the NSSL Collaborative Model for Multiscale Atmospheric Simulation (NCOMMAS)—LETKF developed and used by Thompson et al. (2015). Assimilation onto the 3-km grid begins at 1900 UTC in both cases, and the 2145 and 2045 UTC analyses are used for 19 May 2013 and 6 May 2015, respectively, to generate the TRUTH simulations.

The IC for the TRUTH simulation for each of the two cases is generated by restarting one of the 3-km ensemble analysis members in a one-way, double-nested configuration with 1-km and either 250- or 333-m grids,[2] then integrating for 20 min to allow initially unresolved scales to generate on the subkilometer grid. Key model settings for each case are shown in Table 1. For the 6 May 2015 case, the choice of ensemble member analysis to downscale and the model settings for the innermost grid were empirically tuned to obtain a simulated storm that qualitatively replicates many of the important aspects of the observed storm evolution, including the storm path, retention of supercellular characteristics through the forecast period (Figs. 4b,e), and production of a TLV (discussed later) at about the time

[2] There is nothing significant about our use of 250-m (rather than 333-m) grid spacing for one of the experiments. The 250-m simulation was adopted from a separate research project, and we saw no reason to rerun it on a 333-m grid.

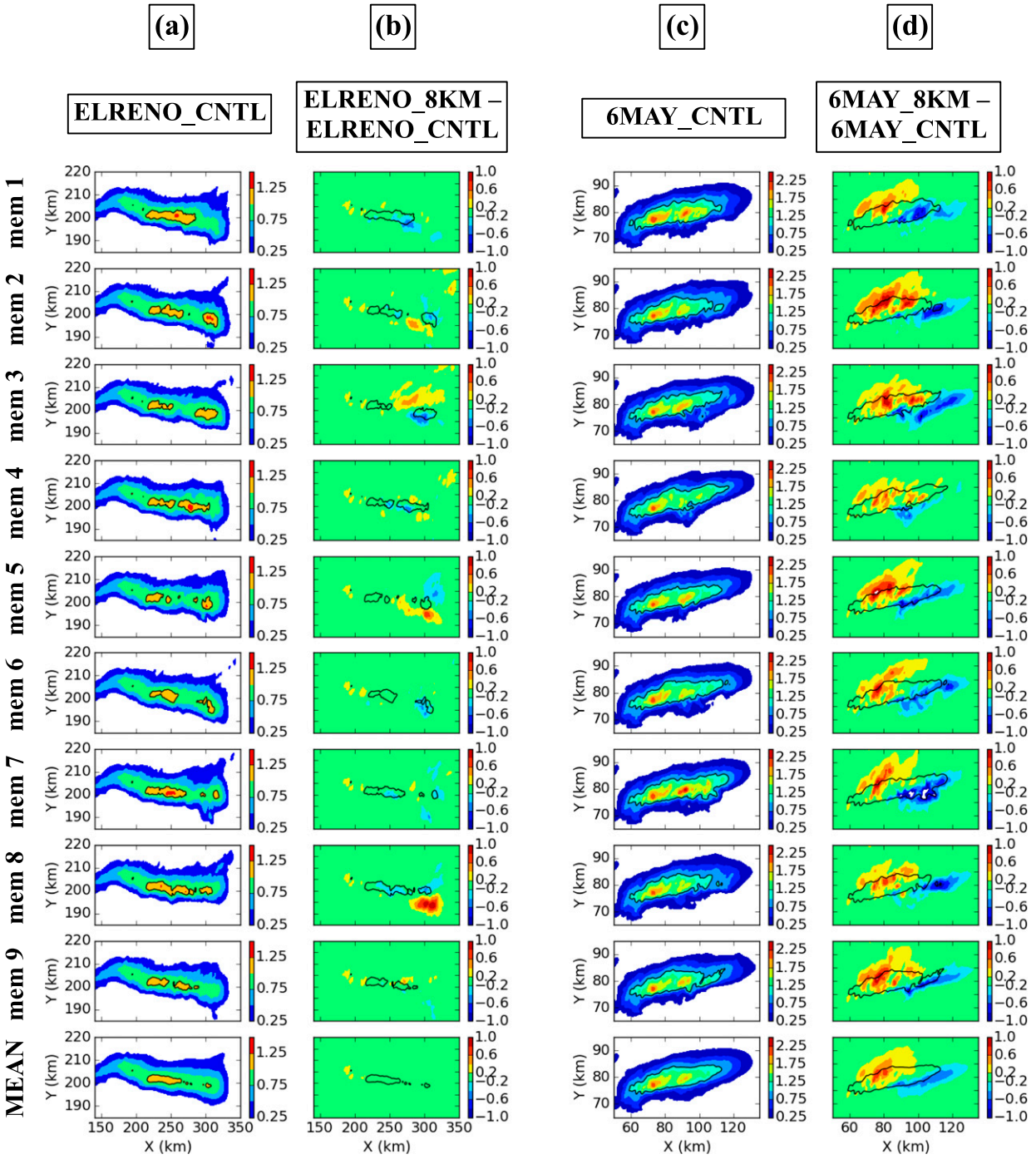


FIG. 5. Accumulated rainfall (in.) for each of the nine (a) CNTL_ELRENO and (c) CNTL_6MAY members, and the errors arising in each of the (b) 8KM_ELRENO and (d) 8KM_6MAY members. (bottom) The ensemble means are also shown. The ELRENO fields are rotated 40° clockwise to reduce the plotting domain sizes. The CNTL 1-in. contour is overlaid in each panel to aid comparisons.

that an EF3 tornado developed near Bridgecreek, Oklahoma. The TLV in TRUTH_6MAY, however, is much shorter lived than the actual tornado (~5 vs ~20 min), which calls into question how well the model replicates the processes contributing to the actual tornado evolution (which is not surprising, especially given the relatively coarse resolution).

Extensive experimentation failed to generate a downscaled 19 May 2013 simulation that qualitatively reproduces the actual storm evolution. In all of the

downscaled simulations, including the one arbitrarily adopted for our sensitivity experiments (TRUTH_19MAY), the modeled storms grow upscale into a mesoscale convective system (MCS; Figs. 4c,f), while the observed storms remain discrete through the period of interest (not shown). A potentially related shortcoming is that while a TLV does occur in the simulation (discussed later), it is much briefer than the EF4 Norman–Shawnee, Oklahoma, tornado that was produced by the corresponding observed storm (~10 vs ~60 min). Despite the limited fidelity of TRUTH_19MAY, it adds value to our investigation in two ways. First, our use of observationally constrained initial and boundary conditions and "full" model physics presumably renders this simulation more physically realistic than the idealized simulation described in the previous section (TRUTH_ELRENO). In other words, while TRUTH_19MAY fails to reproduce the real-world event from which it was generated, it likely better represents the physical evolution of organized convection than does TRUTH_ELRENO. Second, the 19MAY experiments enable exploration of IC resolution impacts for a somewhat different storm mode (supercells evolving into an MCS) than the other two cases (persistently discrete supercells).

### e. *Verification of ensemble forecasts*

Like most previous predictability studies, we make use of a point-to-point error metric in assessing the evolution of errors in our experiments—namely, the root-mean-square error (RMSE). In some instances, the RMSE is normalized by the mean square of the true field to obtain the relative RMSE (RRMSE). Since we seek to characterize errors arising from degraded IC resolution, not from random IC perturbations (as discussed in section 2a), all RMSE and RRMSE herein measure differences from the CNTL ensemble, not the TRUTH simulation. It is critical to note that the double penalty produced by phase errors (Wilks 2006) can reduce the utility of RMSE (and RRMSE) and similar metrics for evaluating forecasts of highly localized phenomena like convective storms. For example, a larger RMSE could be produced for a simulation in which the storm is structurally identical to, but substantially displaced from, the storm in TRUTH than for a simulation that fails to produce the storm altogether. Another major limitation of verification techniques involving simple metrics like RMSE is that unless they are carefully tailored to the application at hand, their relation to subjective assessments of forecast quality can be unreliable (even in the absence of phase errors). This is because forecasts having very different types and degrees of error can yield very similar

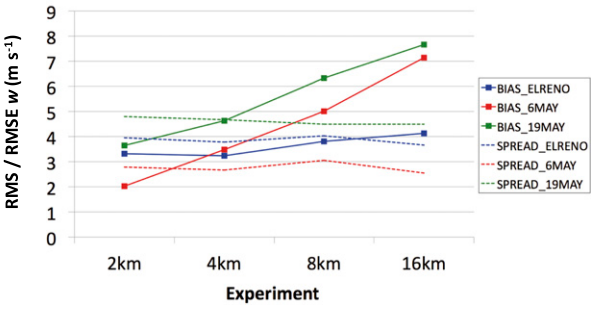


FIG. 6. Two different RMSE statistics valid for horizontal $w$ (m s$^{-1}$) composites from the ELRENO (blue), 6MAY (red), and 19MAY (green) experiments. Marked curves indicate RMS differences between the CNTL and 2-, 4-, 8-, and 16-km probability-matched ensemble means and, therefore, measure ensemble forecast bias arising from IC resolution errors. Dashed curves indicate RMS differences between the probability-matched ensemble means and individual ensemble members and, therefore, measure the ensemble spread arising from IC random errors.

RMSEs. These limitations make point-to-point-based verification methods insufficient for assessing the quality of convective storm forecasts.

In light of these considerations, our verification of the ensemble forecasts includes a subjective, storm-feature-based approach, similar to some previous predictability studies of organized convection (e.g., Wandishin et al. 2008, 2010; Cintineo and Stensrud 2013; Zhang et al. 2016). In addition, to avoid the smearing that commonly arises in ensemble-mean fields due to small displacements of features between ensemble members, we compute probability-matched means (PMMs; Ebert 2001) of the forecasts. Finally, the PMMs are computed not for the full model fields, but for horizontal ($x$–$y$) and time–height composites thereof. Herein, "composite" refers to a field of maxima computed over the cross section of each plotted coordinate; for example, the composite value at a given time–height coordinate is the maximum of the field on the intersecting $x$–$y$ plane. Judiciously chosen composites are an excellent way to distill the most dynamically or operationally important information from model output. Object-based verification methods present another promising approach to evaluating forecasts of storms (e.g., Skinner et al. 2016), but such techniques have not yet advanced to the point where they should replace subjective verification, except in cases where the latter would be unduly time consuming.

In preliminary work, we found that computing the PMM within a prescribed neighborhood centered on each coordinate generally produced more representative PMMs than the traditional, global approach. This is because ensemble biases often caused the global method to incorporate spatiotemporally distant

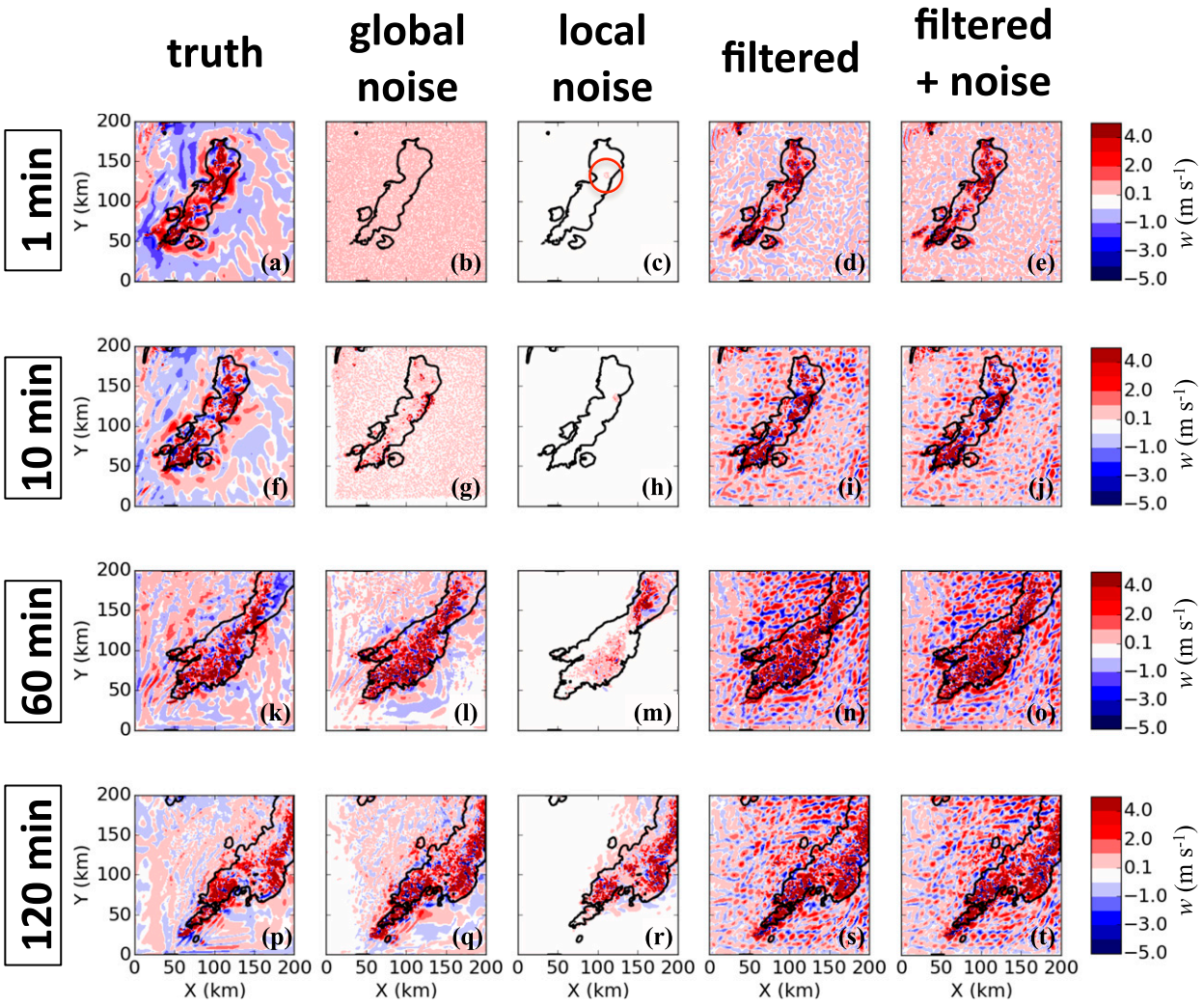


FIG. 7. Vertical velocity ($m\,s^{-1}$) (first column) values and (remaining columns) errors ~5 km AGL in 19MAY simulations, valid at (a)–(e) $t = 0$, (f)–(j) $t = 10$, (k)–(o) $t = 60$, and (p)–(t) $t = 120$ min. (a),(f),(k),(p) TRUTH; (b),(g),(l),(q) CNTL member 1; (c),(h),(m),(r) CNTL member 1, but with noise localized to region indicated by a circle; (d),(i),(n),(s) 16-km member 1, except without noise; and (e),(j),(o),(t) 16-km member 1. The TRUTH 40-dB$Z$ contour is overlaid in each panel to aid comparisons.

ensemble output into the probability matching in certain regions, leading to PMMs that were locally unrepresentative of the ensemble. The neighborhood PMM was therefore adopted for all of the analyses presented herein. Horizontal composite PMMs use neighborhoods of 60 × 60 grid points, and time–height composite PMMs use neighborhoods of 20 min × 20 model levels.

## 3. Results

### a. *Quantitative error analysis*

We begin with a demonstration of the necessity of the ensemble framework for examining the evolution of IC resolution errors in forecasts of convective storms. Rainfall differences between CNTL_ELRENO (Fig. 5a) and 8KM_ELRENO (Fig. 5b), and between CNTL_6MAY (Fig. 5c) and 8KM_6MAY (Fig. 5d), vary substantially with the IC noise realization. This indicates that the ensemble approach is required in both cases to accurately characterize the forecast biases arising from the IC resolution errors. The forecast biases themselves fundamentally differ in significance between the two cases, however. While a northwestward displacement of heaviest rainfall in 8KM_6MAY dominates the inter-member variability (arising from the IC noise), there are no visually evident biases in the 8KM_ELRENO ensemble (cf. bottom panels of Figs. 5b and 5d). Instead, the errors appear to arise largely from the sensitivity of the flow to random perturbations (i.e., limited intrinsic

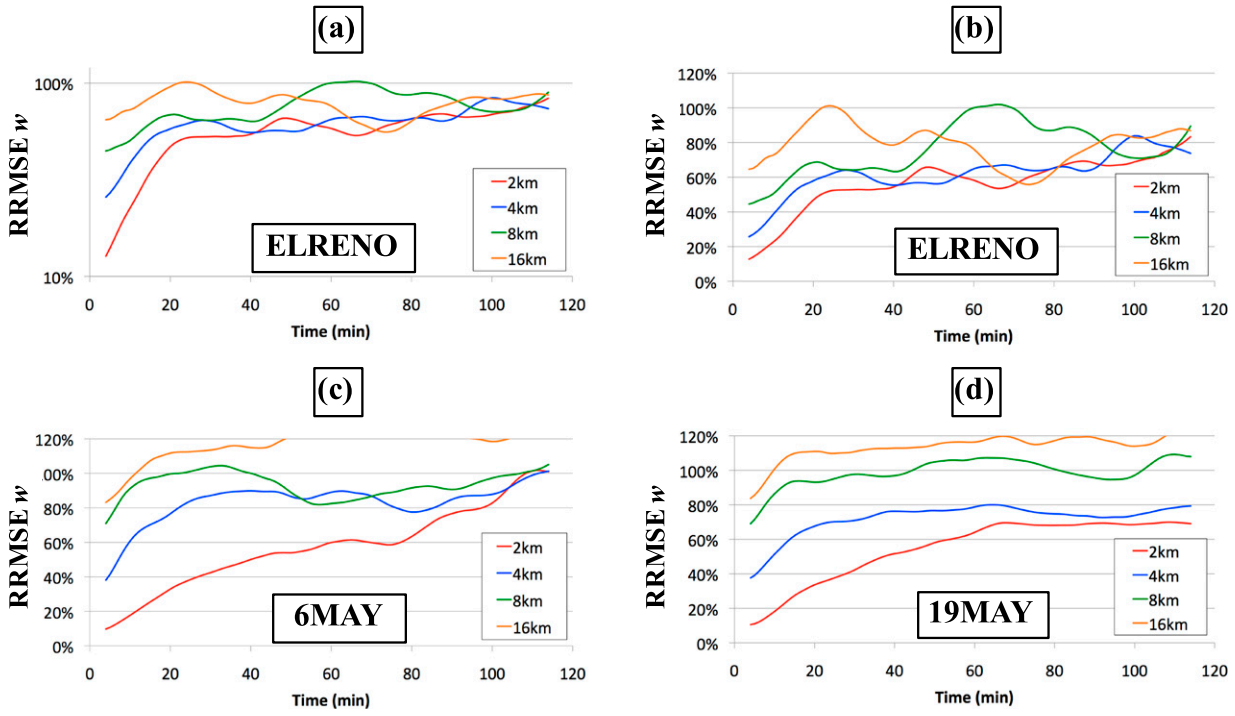


FIG. 8. RRMSE time series, averaged using a 9-min centered window, for ensemble-mean $w$ (m s$^{-1}$) in (a),(b) ELRENO, (c) 6MAY, and (d) 19MAY experiments, colored by IC cutoff wavelength (red = 2 km; blue = 4 km; green = 8 km; orange = 16 km). Here (a) differs from (b) only in its use of a logarithmic abscissa to highlight the more rapid growth of smaller IC errors. RRMSE calculations were restricted to regions where ensemble-mean reflectivity > 5 dB$Z$.

predictability), rendering the ensemble approach particularly critical in that case, since the differences between any pair of corresponding ensemble members in CNTL_ELRENO and 8KM_ELRENO would be both potentially large and unrepresentative of the forecast bias arising from IC resolution error. The errors in 8KM_6MAY and 8KM_ELRENO therefore fall into scenarios B and C from section 2a, respectively (Fig. 1). In all of our experiments, the errors arising from the IC perturbations cancel in the ensemble means, as required to distinguish the impact of limited IC resolution.

Figure 6 provides a more general view of the relative contributions of the random IC perturbations and the IC resolution errors in our simulations. For horizontal composites of $w$ (shown), potential temperature $\theta$ at the lowest model level, and rainfall for all three cases,

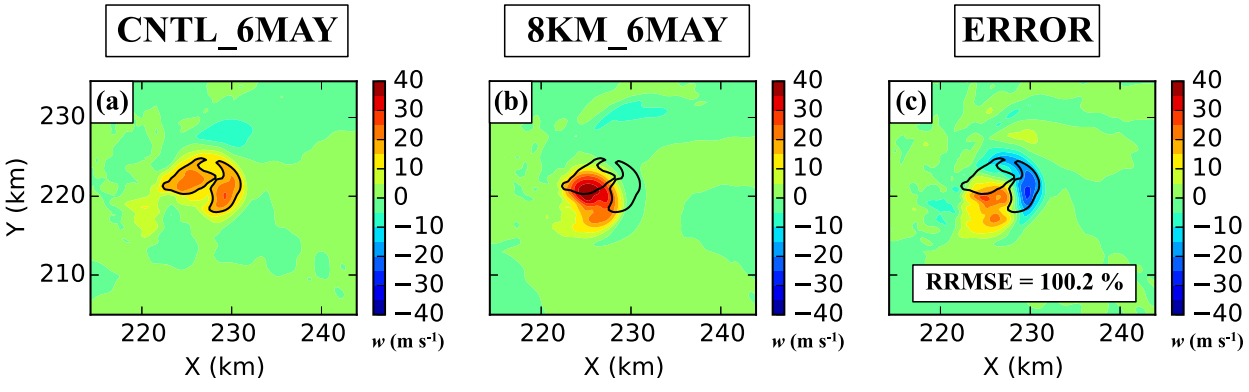


FIG. 9. An illustration of the limited utility of point-to-point verification statistics for evaluating forecasts of convective storms. Cross sections are shown at $z \approx 1.7$ km of (simple) ensemble means of $w$ (m s$^{-1}$), valid $t = 60$ min, for (a) CNTL_6MAY, (b) 8KM_6MAY, and (c) CNTL_6MAY minus 8KM_6MAY. The CNTL 15 m s$^{-1}$ contour is overlaid in each panel to aid comparison. The RRMSE (~100%) is annotated in (c).

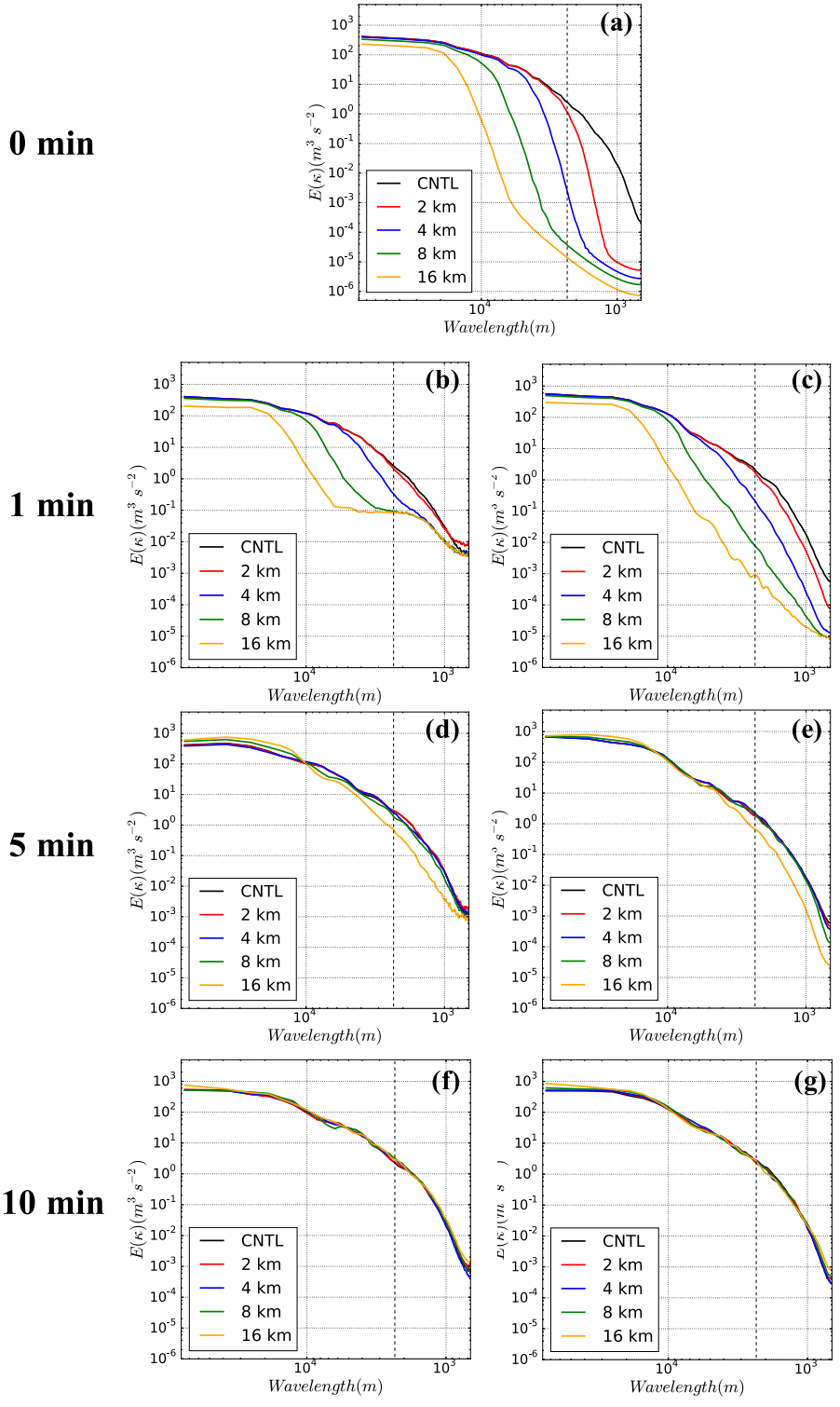


FIG. 10. Two-dimensional vertical velocity energy spectra averaged over lowest 5 km of ELRENO simulations at (a) $t = 0$, (b),(c) $t = 1$, (d),(e) $t = 5$, and (f),(g) $t = 10$ min. Results are shown for an arbitrarily selected ensemble member in (b),(d), and (f) and for simulations repeated without IC noise in (c),(e), and (g). Black, red, blue, green, and orange curves correspond to the CNTL and 2-, 4-, 8-, and 16-km experiments, respectively. The dashed gray line indicates the ARW effective model resolution ($\approx 7\Delta x$) derived by Skamarock (2004).

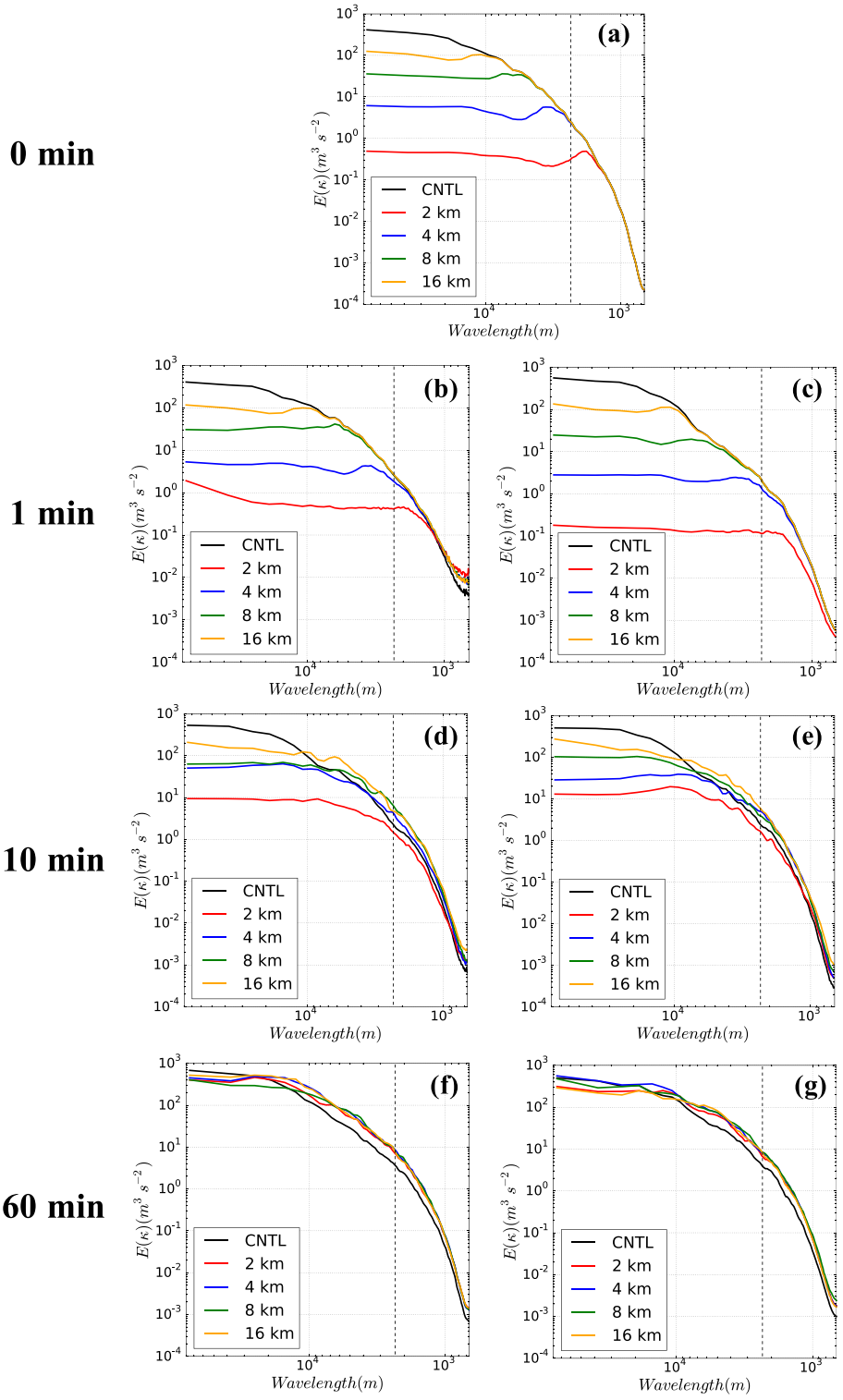


FIG. 11. As in Fig. 10, but for CNTL energy spectra (black) and filtered-IC ELRENO error energy spectra (colors) at (a) $t = 0$, (b),(c) $t = 1$, (d),(e) $t = 10$, and (f),(g) $t = 60$ min.

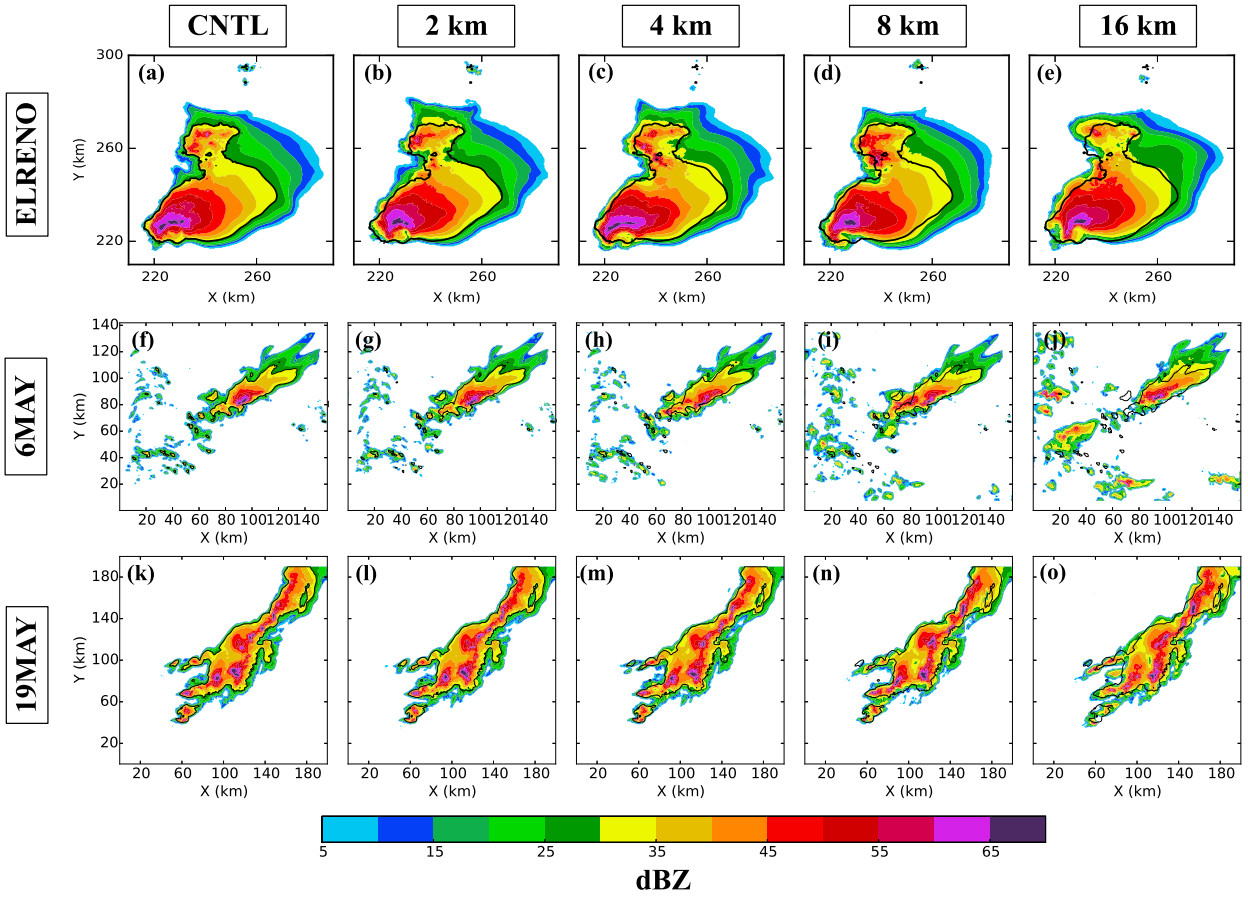


FIG. 12. Reflectivity (dB$Z$; shading) ~2 km AGL at $t$ = 60 min in (a)–(e) ELRENO, (f)–(j) 6MAY, and (k)–(o) 19MAY probability-matched ensemble means. (left)–(right) CNTL and 2-, 4-, 8-, and 16-km ensembles. The CNTL 30-dB$Z$ contour is overlaid in each panel to facilitate comparison.

the RMS differences between the probability-matched ensemble means of the control and of the 2- and 4-km ensembles in each case (ELRENO, 6MAY, and 19MAY) are similar to the RMS differences between the probability-matched ensemble means and individual ensemble members. In other words, ensemble forecast biases arising from IC resolution errors are comparable to the ensemble spread arising from the IC random errors. The forecast biases become dominant in the 8- and 16-km ensembles for 6MAY and 19MAY, but not for ELRENO. These results support the visual inferences from Fig. 5.

The much smaller sensitivity of the ELRENO simulations to IC resolution relative to both full-physics simulations raises an important question: is this a consequence of using idealized model settings or of some aspect of the ELRENO storm or larger-scale environment (apart from its initial homogeneity)? Definitively answering this question is beyond the scope of this investigation, but the error evolution analysis that follows suggests that intrastorm error growth may be slower on the examined time scales if IC resolution errors do not exist in the storm environment—that is, if the environment is homogeneous, as in the ELRENO simulations. Figure 7 depicts the evolution of four types of IC error in the 19MAY experiments; results for 6MAY are very similar (not shown). The first column of Fig. 7 shows the true $w$ field ~5 km AGL at several times in the simulation. Comparing $w$ errors arising solely from the introduction of the random IC perturbations (second column) with $w$ errors arising solely from degraded IC resolution (fourth column) reveals that, within the storm environment, errors associated with degraded IC resolution grow much larger in situ than do the random perturbations. Errors that arise anywhere in the domain, either within (second column) or outside of the storms (not shown), are ultimately propagated throughout by gravity- and soundlike waves, initiating rapid error growth in convectively unstable regions as in Hohenegger and Schär (2007). This is evidenced by the rapid spread of errors from a localized source to areas far upstream (cf. Figs. 7m and 7h). The waves presumed responsible for the

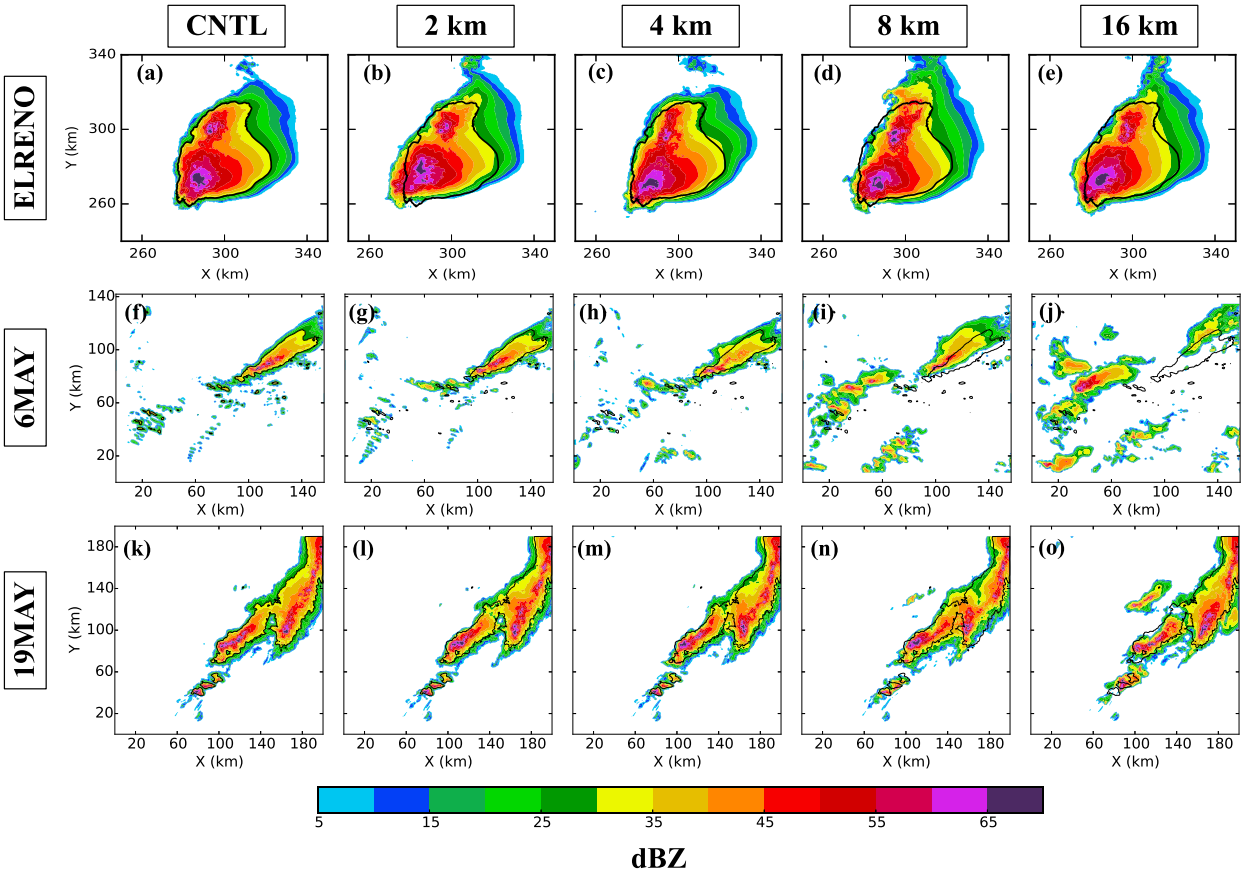


FIG. 13. As in Fig. 12, but at $t = 120$ min.

error propagation are visually evident in horizontal cross sections of $w$ and pressure (not shown) and are an expected consequence of the model response to both rapid storm evolution and to the imbalances created by the introduction of IC errors (e.g., Potvin and Wicker 2013b). Since the storm environment in ELRENO is homogeneous and therefore not subject to IC resolution errors, the error evolution therein is much slower (not shown), and therefore intrastorm error growth is not so exacerbated by error propagation from outside the storm (or from in situ storm interactions with environmental errors). We therefore hypothesize that the use of a homogeneous environment in simulations artificially slows IC resolution error growth.

At the suggestion of a reviewer, we tested this hypothesis by repeating one of the sets of full-physics simulations—we chose 6MAY—using idealized settings (6MAY_IDEAL). The TRUTH simulation for this new case was initialized using a 3-K thermal bubble and a representative sounding computed at a model gridpoint within the TRUTH_6MAY storm inflow. Plotting RMSE $w$ as a function of IC cutoff wavelength as in Fig. 6 (not shown) revealed that the sensitivity of 6MAY_IDEAL to IC resolution is substantially reduced relative to that of 6MAY. This supports our hypothesis that homogeneous initial conditions reduce IC resolution sensitivity and that using idealized simulations alone to investigate this aspect of predictability therefore limits the practical applicability of the findings.

Much of the preceding discussion has focused on methodological considerations for studying IC resolution errors. We now turn to deeper investigation of the evolution of IC resolution errors in our experiments. RRMSE time series for ensemble-mean $w$ (Fig. 8) reveal that the correlation between IC resolution and forecast bias generally weakens over time in all three cases, reflecting the limited intrinsic predictability of the flow. Similar behavior occurs in other examined variables (not shown). In the ELRENO case, the IC resolution is effectively ''forgotten'' after 60 min, but again, the weakened impact of IC resolution errors may largely be an artifact of the initially homogeneous environment. Bias growth in all the 19MAY simulations largely ceases after about an hour. This is consistent with the expected increase in influence of larger scales (relative to smaller scales) on storm evolution as the supercells organize

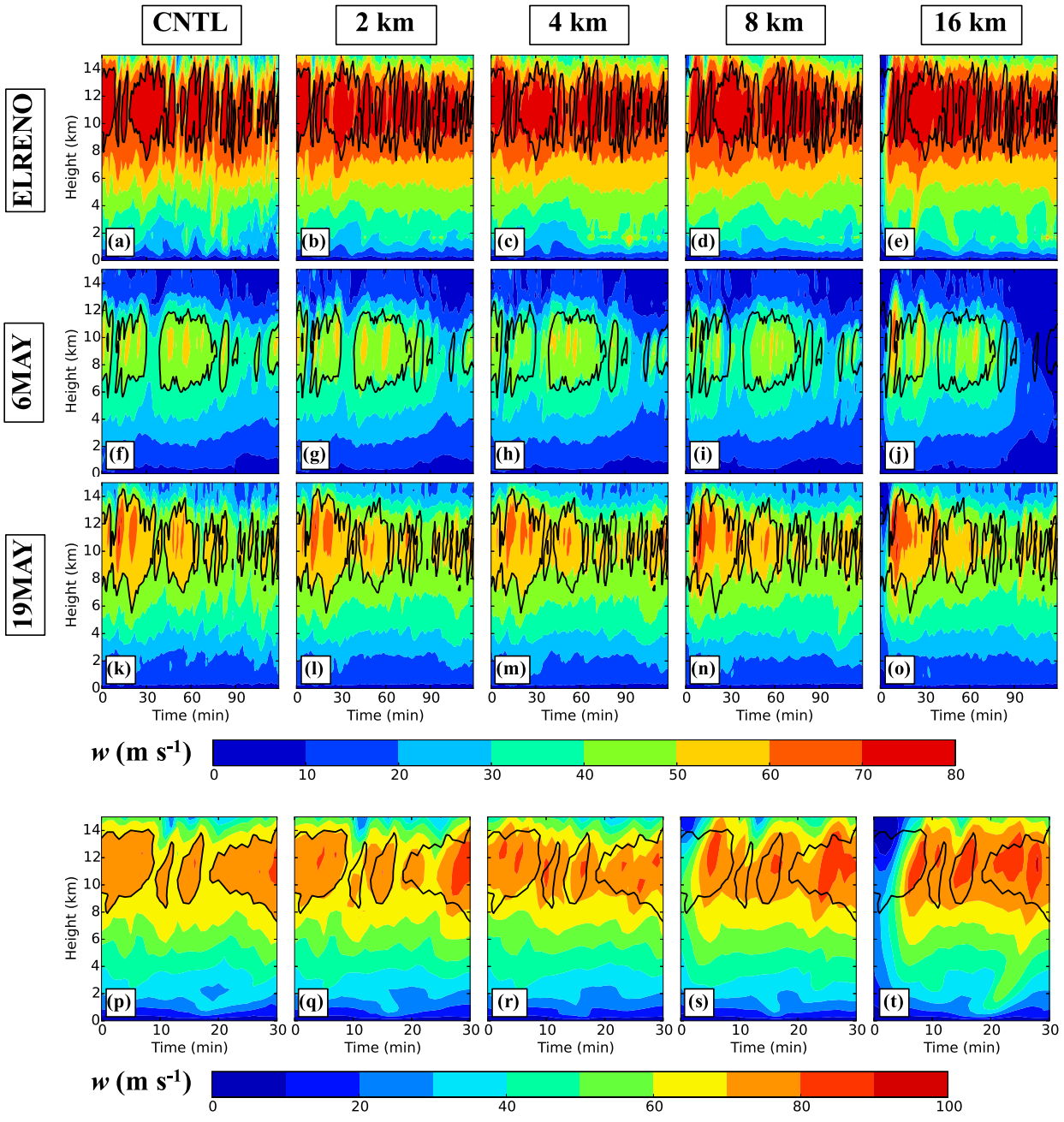


FIG. 14. Time–height vertical velocity ($m\,s^{-1}$) composites for (a)–(e) ELRENO, (f)–(j) 6MAY, and (k)–(o) 19MAY probability-matched ensemble means. (p)–(t) As in (a)–(e), except only the first 30 min are plotted to show more detail. (left)–(right) CNTL and 2-, 4-, 8-, and 16-km ensembles. Composite calculations for 6MAY were restricted to the primary storm to clarify comparisons. The CNTL 70, 40, and 50 $m\,s^{-1}$ contours are overlaid in each of the ELRENO, 6MAY, and 19MAY panels, respectively.

upscale into an MCS. This result strongly suggests that the evolution of IC resolution errors is dependent upon convective mode and that 19MAY is therefore less representative than 6MAY of the error evolution in discrete supercells.

It may be tempting to interpret RRMSE > 100% in Fig. 8 as an indication of total loss of forecast skill. This judgment would be premature, as Fig. 9 demonstrates. A 60-min forecast of the storm updraft in Fig. 9a that looked similar to Fig. 9b would be very skillful by contemporary standards, despite having substantial phase, shape, and magnitude errors that yield a RRMSE of roughly 100%. The weakened link between point-to-point error metrics and forecast skill for highly localized phenomena like

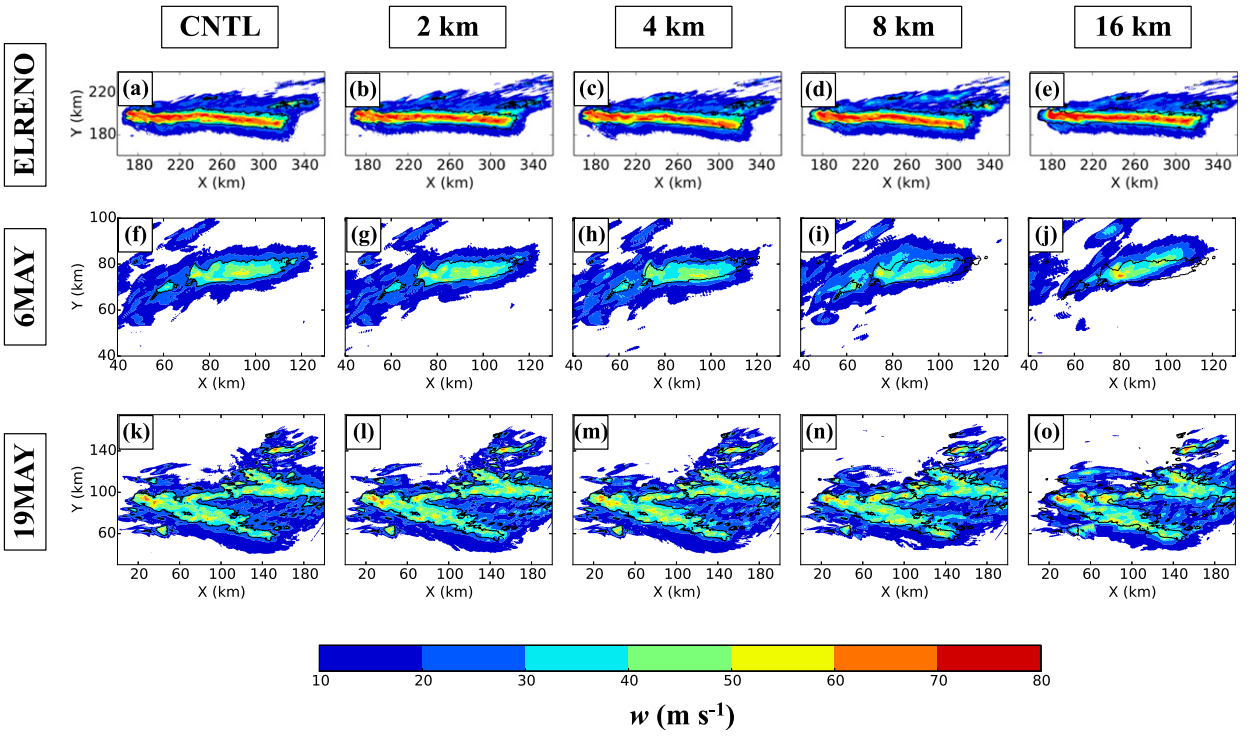


FIG. 15. Horizontal composites (i.e., maxima over all heights and times) of vertical velocity ($m\,s^{-1}$) for (a)–(e) ELRENO, (f)–(j) 6MAY, and (k)–(o) 19MAY probability-matched ensemble means. (left)–(right) CNTL and 2-, 4-, 8-, and 16-km ensembles. The ELRENO and 19MAY fields are rotated 40° clockwise to reduce the plotting domain size. The CNTL 30 $m\,s^{-1}$ contour is overlaid in each panel.

storms is further illustrated by the composite plots in section 3b, which demonstrate that forecast skill is retained throughout the forecast periods in all of our experiments.

Two-dimensional energy spectra for $w$ (shown for ELRENO in Fig. 10) and other variables, computed as in Potvin and Flora (2015), indicate that initially missing scales are rapidly regenerated as the model integration proceeds. This is true even in simulations without IC perturbations (Figs. 10c,e,g), indicating that the insertion of the IC noise and its associated energy (cf. Figs. 10b,c) is not responsible for the rapid recovery of missing scales, though it does slightly accelerate the process (cf. Figs. 10d,e). By 10 min into each simulation, most of the energy at the affected scales has been recovered (Fig. 10f). The brevity of the periods over which interscale interactions are spuriously absent in the simulations helps explain the limited qualitative impact of the IC resolution errors (shown in section 3b).

Two-dimensional error energy spectra for $w$ (shown for ELRENO in Fig. 11) illuminate the evolution of both the IC random and IC resolution errors. In simulations initialized without noise, forecast errors initially decrease (cf. Figs. 11a and 11c) as filtered scales (corresponding to errors of up to 100%) rapidly regenerate. Upscale error growth then becomes dominant, and as in the squall-line simulations of Durran and Weyn (2016), error growth at larger scales does not await error saturation at smaller scales (cf. Figs. 11c and 11e), in contrast to the upscale cascade characteristic of spatially homogeneous, statistically stationary turbulence when initial errors are predominantly small scale (e.g., Leith and Kraichnan 1972). In the original simulations (initialized with noise), the reduction in forecast error as filtered scales regenerate is offset by both upscale and downscale growth of the IC perturbations (cf. Figs. 11a–c). Whether or not IC noise is present, error saturation occurs at all scales well before the end of the simulation (Figs. 11f,g). Again, as Fig. 9 suggests, and the qualitative verification in the next subsection confirms, this does not correspond to total loss of forecast skill. The convergence of error spectra of finer-IC simulations to those of coarser-IC simulations is consistent with the IC resolution being ''forgotten'' before the end of the simulation. This result is broadly consistent with findings by Durran and Gingrich (2014) for simulations of spatially homogeneous, statistically stationary turbulence initialized with and without perturbations at wavelengths $< 400$ km. Error spectra computed from the full-physics experiments (not shown) resemble those from ELRENO, except that the largest-scale errors remain higher in the coarser-resolution-IC simulations through the end of the

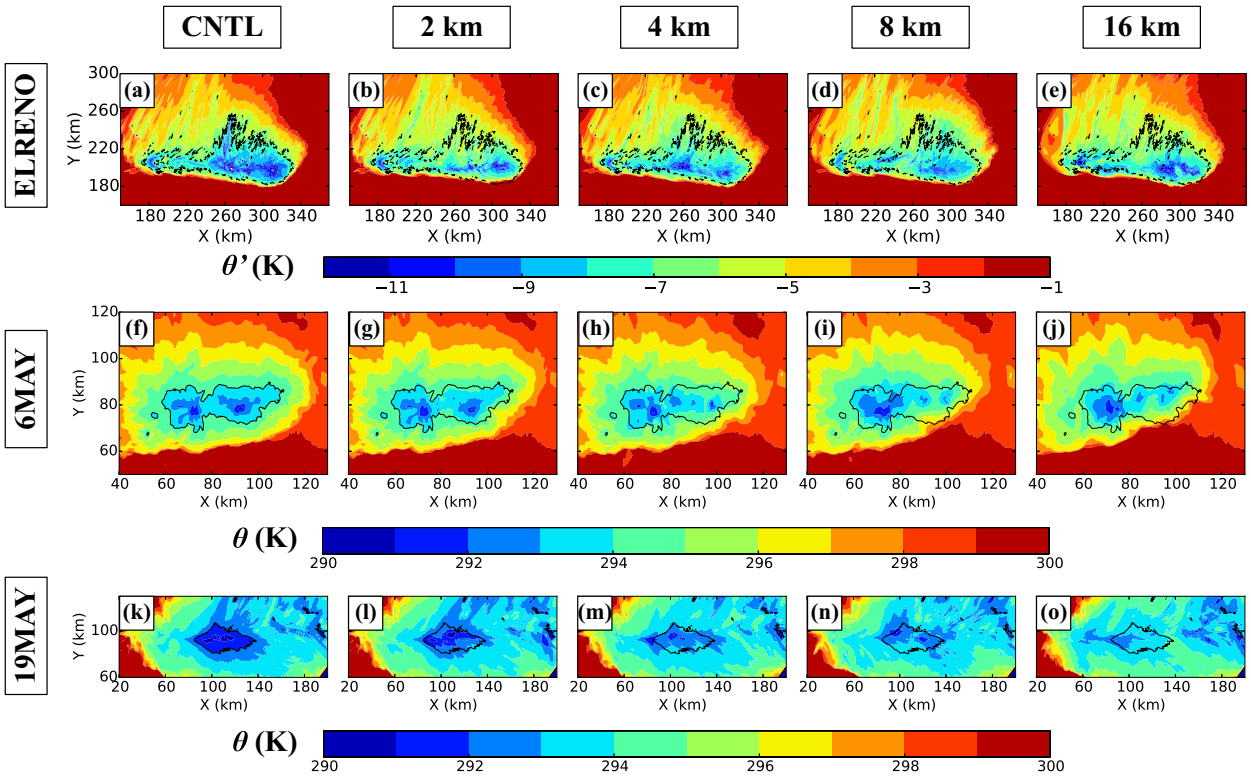


FIG. 16. Horizontal composites of (a)–(e) lowest-level perturbation potential temperature (K) for ELRENO probability-matched ensemble mean and lowest-level potential temperature (K) for (f)–(j) 6MAY and (k)–(o) 19MAY probability-matched ensemble means. (left)–(right) CNTL and 2-, 4-, 8-, and 16-km ensembles. The ELRENO and 19MAY fields are rotated 40° clockwise to reduce the plotting domain size. The CNTL −7-, 294-, and 292-K contours are overlaid in each ELRENO, 6MAY, and 19MAY panel, respectively.

forecast period, consistent with the stronger correlation between IC resolution and forecast error at longer lead times in those cases (Fig. 8).

### b. *Qualitative error analysis*

Quantitative analysis alone provides a limited view of the practical impacts of IC errors on storm simulations and forecasts. We therefore present qualitative (subjective) evaluations of the error evolution in our experiments, beginning with the reflectivity field (Figs. 12 and 13). In all three cases, even the 16-km simulations well match the TRUTH solution at $t = 60$ min (cf. Figs. 12a,e, 12f,j, and 12k,o), except that significant spurious convection has developed in 16KM_6MAY. Displacement errors in the primary storms are minimal, indicating that storm propagation has not been substantially affected by the degraded IC resolution. Even at $t = 120$ min, the 16KM_ELRENO and 16KM_19MAY simulations are remarkably similar to their corresponding TRUTH simulations (cf. Figs. 13a,e and 13k,o). The primary storm in the 16KM_6MAY simulation, however, is much too weak, and perhaps as a consequence, displaced northeastward of the TRUTH_6MAY storm (cf. Figs. 13f and 13j). In addition, significant spurious convection now exists in both the 8KM_6MAY and 16KM_6MAY simulations (Figs. 13i,j). Among the three cases, the sensitivity exhibited by the 6MAY simulations may be the most representative of supercells, given the hypothesized reductions in IC resolution error growth in the ELRENO and 19MAY experiments by the homogeneous environment in the former and the upscale growth of convection in the latter. Even in the 6MAY case, though, the practical impact of 16-km IC resolution errors on the general propagation and evolution of the reflectivity fields of the primary storm is minor through at least 1 h of forecast lead time.

A similar conclusion holds for the evolution of the updraft of the primary storm (Figs. 14 and 15). As before, the most significant errors occur in the 16KM_6MAY simulation (Figs. 14j and 15j). In all other simulations, the overall updraft evolution is captured quite well, despite the large RRMSE in $w$ (Fig. 8). In all three cases, the time required for regeneration of the initially missing scales is reflected in the updrafts being too weak early in the forecast period. This is most clearly seen in the 16KM time–height composites (cf. Figs. 14a,e, 14f,j, 14k,o, and 14p,t). Consistent with the $w$ energy spectra (Fig. 10), the time required for the filtered-IC updrafts to attain the magnitudes of the TRUTH updrafts increases with the filter

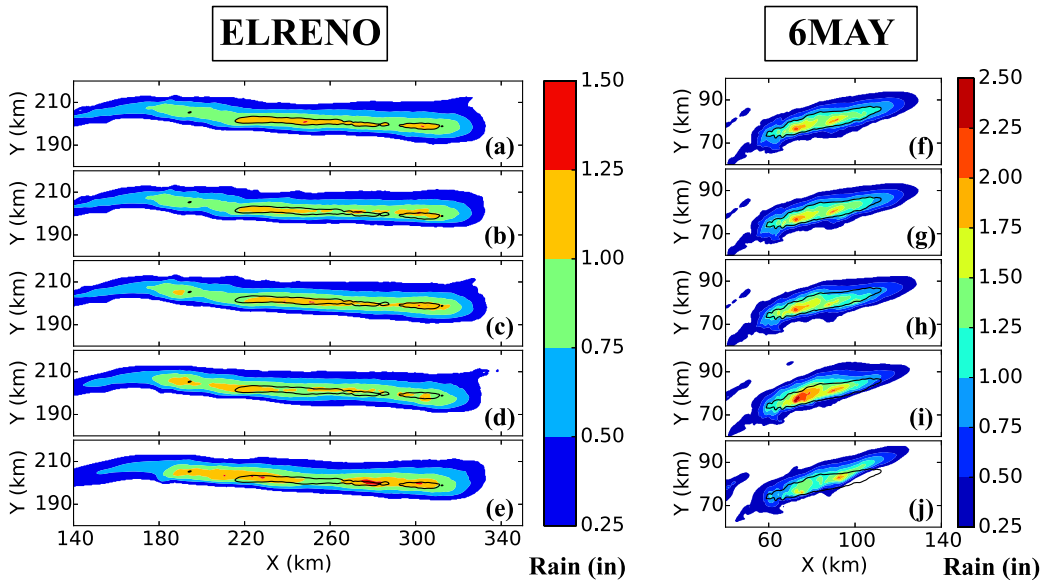


FIG. 17. Horizontal composites of total rainfall (in.) for (a)–(e) ELRENO and (f)–(j) 6MAY probability-matched ensemble means for (a),(f) CNTL and (b),(g) 2-, (c),(h) 4-, (d),(i) 8-, and (e),(j) 16-km ensembles. The ELRENO fields are rotated 40° clockwise to reduce the plotting domain size. The CNTL 1-in. contour is overlaid in each panel.

cutoff wavelength. Forecast errors distinctly increase as IC resolution is degraded in the 6MAY and 19MAY experiments, but not in the ELRENO experiments, consistent with the RRMSE time series (Fig. 8).

The relative insensitivity of storm evolution to IC resolution also extends to $\theta$ (shown for the lowest model level in Fig. 16) and total rainfall (Fig. 17). The dependence of forecast errors on IC resolution is again generally stronger in the full-physics than in the ELRENO simulations. The cold pool cores are too warm in all of the filtered-IC simulations, as can readily be deduced by comparing composite temperatures within the contour-outlined regions (e.g., cf. Figs. 16a–e), but the number of cases examined is too small to deduce whether this is a general result. The 16KM_19MAY cold pools are substantially too warm prior to the upscale growth into an MCS (cf. Figs. 16l–o and Fig. 16k within 292-K contour) but become more consistent with the TRUTH cold pools thereafter (cf. at $x > 140$ km). Rainfall errors in all the ELRENO experiments are modest (Figs. 17a–e), whereas more practically significant rainfall errors occur in the 8KM_6MAY, 16KM_6MAY (Figs. 17i,j), and 16KM_19MAY (not shown) experiments. Still, none of the rainfall errors are particularly large by contemporary forecasting standards.

Finally, we turn to the most sensitive variable of those examined—vertical vorticity (Figs. 18 and 19). Intense TLVs occur at multiple times in TRUTH_ELRENO (Figs. 18a, 19a). During the first three episodes, which occur within roughly the first hour of the simulation, vertical vorticity errors generally increase as the IC resolution is degraded (Figs. 18a–e and 19). During the last hour of the simulation, the correlation between forecast error and IC resolution error becomes weak, consistent with previous results. In both the 6MAY and 19MAY cases, a TLV that occurs within the first 15 min of TRUTH does not occur in the 16KM simulations (Figs. 18j,o), which is not surprising given the required time for missing scales to regenerate. While timing errors in the subsequent weaker low-level spinups generally increase with filter cutoff wavelength, gross similarities with the TRUTH simulations are retained even in the 8KM and 16KM experiments. Generally speaking, the evolution of low-level rotation is not unduly sensitive to IC resolution in our experiments.

## 4. Conclusions and future work

Idealized and full-physics supercell simulations were run using differing initial-condition (IC) resolutions to investigate 0–2-h forecast impacts of unanalyzed IC scales. The primary conclusions are as follow:

1) The strong sensitivity of convection to random perturbations renders the deterministic framework inappropriate for investigating impacts of IC resolution errors. More generally, predictability studies in which the errors of interest are potentially sensitive to small changes in the flow should adopt an ensemble approach to ensure representative results.

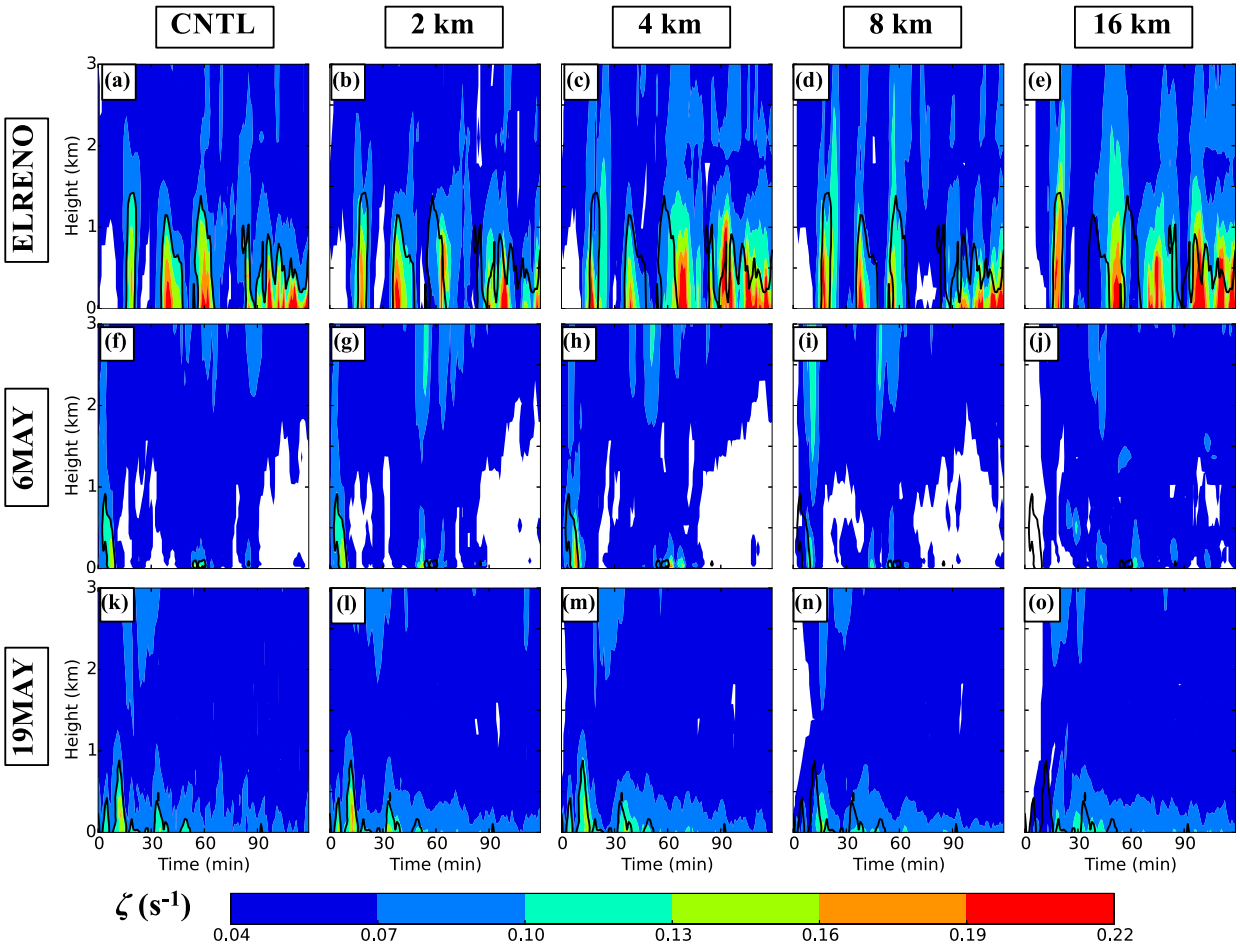


FIG. 18. Time–height vertical vorticity ($s^{-1}$) composites for (a)–(e) ELRENO, (f)–(j) 6MAY, and (k)–(o) 19MAY probability-matched ensemble means. (left)–(right) CNTL and 2-, 4-, 8-, and 16-km ensembles. The CNTL 0.1-$s^{-1}$ contour is overlaid in each panel.

2) Missing scales in the ICs of simulated supercells regenerate within the first 10–20 min of model integration.
3) Errors arising from the initially missing scales evolve similarly to previously investigated initial-condition perturbations—they rapidly grow upscale and propagate throughout the domain, initiating rapid error growth within unperturbed regions of moist convection. The resulting forecast biases generally increase as initial-condition resolution decreases.
4) Homogeneous simulations may not well represent supercell forecast sensitivities to IC resolution since the growth of errors in the storm environment and their subsequent propagation into convectively active regions is artificially suppressed.
5) The initial absence of small IC scales does not noticeably modulate ensemble forecast spread.
6) Spectral analysis and other objective methods involving point-to-point verification are insufficient to assess the operational significance of errors in forecasts of organized convection. For example, error saturation can result largely from modest phase errors in discrete storm structures and, therefore, does not necessarily indicate total loss of forecast skill. A combination of objective and subjective evaluation is critical.
7) The evolution of important storm features and attributes is relatively insensitive to IC resolution errors through 2-h forecast lead times. This indicates that smaller-scale processes within organized convection are primarily determined by larger scales, at least once storms have reached maturity.
8) Near-future analysis resolution limitations of real-time ensemble prediction systems (horizontal grid spacings of 1–3 km) will not severely diminish the practical predictability of organized storms, as long as analyses are downscaled to subkilometer forecast grids. Forecast skill will instead be primarily limited by other types of model error and by the lack of observations of the storm environment.

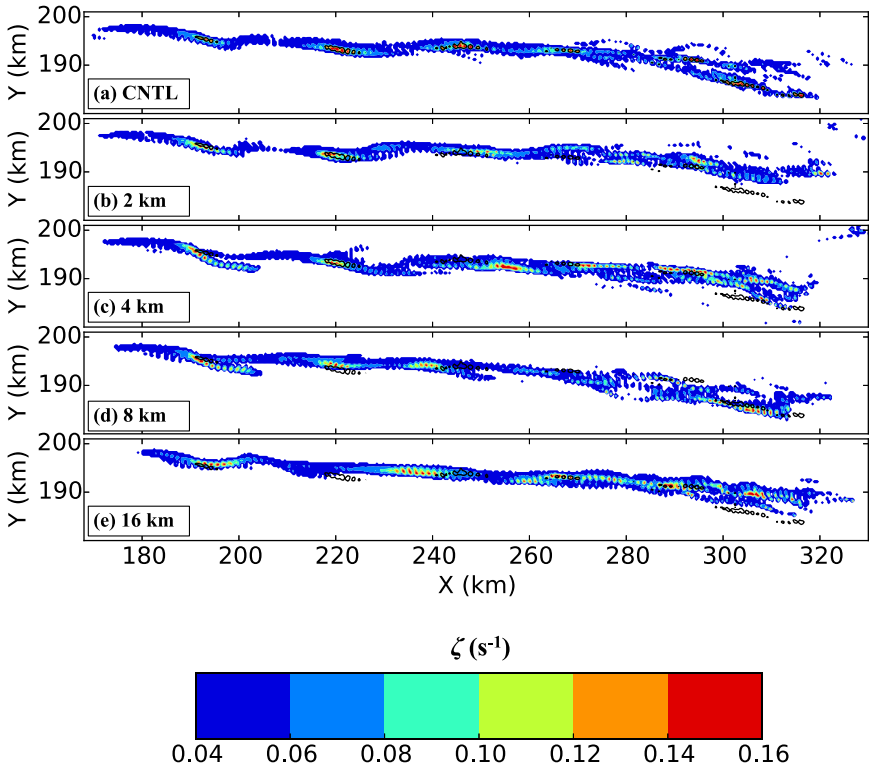


FIG. 19. Horizontal composites of vertical vorticity (maximum over lowest 2 km and all times; $s^{-1}$) for ELRENO probability-matched ensemble means: (a) CNTL and (b) 2-, (c) 4-, (d) 8-, and (e) 16-km ensembles. The ELRENO fields are rotated 40° clockwise to reduce the plotting domain size. The CNTL 0.1-$s^{-1}$ contour is overlaid in each panel.

Until real-time data assimilation systems employ convection-resolving forecast grids, limited model resolution will produce significant errors even at scales that are resolved on the model grid (e.g., Potvin and Flora 2015). Thus, downscaling coarse analyses created from traditional, single-resolution data assimilation to fine-scale forecast grids may produce larger forecast errors in practice than indicated by our experiments. Resolvable-scale IC errors arising from limited forecast model resolution can be mitigated using mixed-resolution data assimilation methods, which use one or more high-resolution ensemble members to update the analysis at each cycle. We recommend that the value added by mixed- versus single-resolution data assimilation be assessed using both simulation and real-world frameworks.

The work presented herein could be extended in other valuable ways. The relative importance of missing IC scales (and other small-scale errors) to storm prediction may vary substantially with larger-scale storm environment, as do forecast grid resolution errors (Potvin and Flora 2015). This potential sensitivity should be explored by repeating our analysis for a broad range of synoptic conditions, including the low-CAPE, high-shear scenario typified in the southeastern United States. The sensitivity to IC resolution may also vary substantially with model resolution, which is an important consideration given that operational forecast models (especially ensembles) will not explicitly model convection for years to come. It would also be useful to explore whether/how the impact of missing IC scales differs for newly developed storms versus mature storms like those examined in this study. Finally, we recommend that errors in real-world, subkilometer ensemble forecasts be compared to errors in simulated forecasts like those herein in order to assess the combined impact of contemporary parameterization and large-scale analysis errors. Better understanding of all of these aspects of storm predictability would greatly benefit the design of future data assimilation and prediction systems and interpretation of their forecast output.

*Acknowledgments.* This work was prepared by the authors with funding provided by 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. We thank Dale Durran, Chris Snyder, and an anonymous reviewer for providing valuable suggestions that substantially improved the paper. We also thank Louis Wicker and Patrick Skinner for helpful discussions and for informally reviewing an early version of the manuscript. Valuable local computing assistance was provided by Gerry Creager, Jesse Butler, and Jeff Horn. Part of the computing for this project was performed at the University of Oklahoma (OU) Supercomputing Center for Education and Research (OSCER). OSCER Director Henry Neeman, OSCER Senior System Administrator Brett Zimmerman, and Petascale Storage Administrator Patrick Calhoun provided valuable technical expertise.

## 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.

Buehner, M., P. L. Houtekamer, C. Charette, H. L. Mitchell, and B. He, 2010: Intercomparison of variational data assimilation and the ensemble Kalman filter for global deterministic NWP. Part I: Description and single-observation experiments. *Mon. Wea. Rev.*, **138**, 1550–1566, doi:10.1175/2009MWR3157.1.

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.

Courtier, P., J.-N. Thepaut, and A. Hollingsworth, 1994: A strategy for operational implementation of 4D-Var, using an incremental approach. *Quart. J. Roy. Meteor. Soc.*, **120**, 1367–1387, doi:10.1002/qj.49712051912.

Dong, J., M. Xue, and K. Droegemeier, 2011: The analysis and impact of simulated high-resolution surface observations in addition to radar data for convective storms with an ensemble Kalman filter. *Meteor. Atmos. Phys.*, **112**, 41–61, doi:10.1007/s00703-011-0130-3.

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.

——, and J. A. Weyn, 2016: Thunderstorms do not get butterflies. *Bull. Amer. Meteor. Soc.*, **97**, 237–243, doi:10.1175/BAMS-D-15-00070.1.

Ebert, E. E., 2001: Ability of a poor man's ensemble to predict the probability and distribution of precipitation. *Mon. Wea. Rev.*, **129**, 2461–2480, doi:10.1175/1520-0493(2001)129<2461:AOAPMS>2.0.CO;2.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte-Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99**, 10 143–10 162, doi:10.1029/94JC00572.

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.

Gao, J., and M. Xue, 2008: An efficient dual-resolution approach for ensemble data assimilation and tests with simulated Doppler radar data. *Mon. Wea. Rev.*, **136**, 945–963, doi:10.1175/2007MWR2120.1.

Hohenegger, C., and C. Schär, 2007: Predictability and error growth dynamics in cloud-resolving models. *J. Atmos. Sci.*, **64**, 4467–4478, doi:10.1175/2007JAS2143.1.

Hunt, B. R., E. J. Kostelich, and I. Szunyogh, 2007: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D*, **230**, 112–126, doi:10.1016/j.physd.2006.11.008.

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

Kleist, D. T., and K. Ide, 2015: An OSSE-based evaluation of hybrid variational–ensemble data assimilation for the NCEP GFS. Part I: System description and 3D-Hybrid results. *Mon. Wea. Rev.*, **143**, 433–451, doi:10.1175/MWR-D-13-00351.1.

Leith, C., and R. Kraichnan, 1972: Predictability of turbulent flows. *J. Atmos. Sci.*, **29**, 1041–1058, doi:10.1175/1520-0469(1972)029<1041:POTF>2.0.CO;2.

Lorenz, E. N., 1969: The predictability of a flow which possesses many scales of motion. *Tellus*, **21**, 289–307, doi:10.1111/j.2153-3490.1969.tb00444.x.

Miller, P. A., M. F. Barth, L. A. Benjamin, R. S. Artz, and W. R. Pendergrass, 2007: MADIS support for UrbaNet. *14th Symp. on Meteorological Observation and Instrumentation/16th Conf. on Applied Climatology*, San Antonio, TX, Amer. Meteor. Soc., JP2.5. [Available online at https://ams.confex.com/ams/87ANNUAL/techprogram/paper_119116.htm.]

Potvin, C. K., and L. J. Wicker, 2013a: Assessing ensemble forecasts of low-level supercell rotation within an OSSE framework. *Wea. Forecasting*, **28**, 940–960, doi:10.1175/WAF-D-12-00122.1.

——, and ——, 2013b: Correcting fast-mode pressure errors in storm-scale ensemble Kalman filter analyses. *Adv. Meteor.*, **2013**, 624931, doi:10.1155/2013/624931.

——, and M. L. Flora, 2015: Sensitivity of idealized supercell simulations to horizontal grid spacing: Implications for Warn-on-Forecast. *Mon. Wea. Rev.*, **143**, 2998–3024, doi:10.1175/MWR-D-14-00416.1.

Rainwater, S., and B. Hunt, 2013: Mixed-resolution ensemble data assimilation. *Mon. Wea. Rev.*, **141**, 3007–3021, doi:10.1175/MWR-D-12-00234.1.

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:HOLPIT>2.0.CO;2.

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., doi:10.5065/D68S4MVH.

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

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

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.

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.

——, 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.

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

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.

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, doi:10.1175/2008WAF2007057.1.

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

Weisman, M. L., 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.

Wheatley, D. M., K. H. Knopfmeier, T. A. Jones, and G. J. Creager, 2015: Storm-scale data assimilation and ensemble forecasting with the NSSL Experimental Warn-on-Forecast System. Part I: Radar data experiments. *Wea. Forecasting*, **30**, 1795–1817, doi:10.1175/WAF-D-15-0043.1.

Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences.* Elsevier, 627 pp.

Zhang, F., A. M. Odins, and J. W. Nielsen-Gammon, 2006: Mesoscale predictability of an extreme warm-season precipitation event. *Wea. Forecasting*, **21**, 149–166, doi:10.1175/WAF909.1.

Zhang, Y., F. Zhang, D. Stensrud, and Z. Meng, 2016: Intrinsic predictability of the 20 May 2013 tornadic thunderstorm event in Oklahoma at storm scales. *Mon. Wea. Rev.*, **144**, 1273–1298, doi:10.1175/MWR-D-15-0105.1.