UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

PREDICTABILITY OF MESOSCALE CONVECTIVE SYSTEMS IN TWO- AND THREE-DIMENSIONAL MODELS

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

MATTHEW S. WANDISHIN
Norman, Oklahoma
2008

PREDICTABILITY OF MESOSCALE CONVECTIVE SYSTEMS IN TWO- AND THREE-DIMENSIONAL MODELS

A DISSERTATION APPROVED FOR THE
SCHOOL OF METEOROLOGY

BY

_____
Dr. Lance Leslie, Chair

_____
Dr. David J. Stensrud

_____
Dr. Michael B. Richman

_____
Dr. Steven L. Mullen

_____
Dr. S. Lakshmivarahan

Acknowledgments

I must begin by expressing my indebtedness to Drs. Dave Stensrud and Steve Mullen for their investment in me through this endeavor, from putting together the funding proposal that paid my way to helping me refocus the topic of this research when my efforts bogged down a few years back to pushing me to reach, finally, the denouement. I am deeply grateful for their encouragement and for their efforts on my behalf.

Many thanks, also, to rest of my committee—Drs. Lance Leslie, Mike Richman and Varahan—for their patience through this too lengthy process and for improving the clarity and thoroughness of this dissertation.

Dr. Ted Mansell was invaluable for his assistance with the COMMAS model and software to wrangle with the model output. Many times did I darken his doorway, but Ted was always gracious and helpful. Conversations with Dr. Lou Wicker, both scientific and otherwise, have been a strong antidote to lazy thinking and easy assumptions.

Completing this project involved substantial computer resources, in terms of CPU and storage. Steve Fletcher, Brett Morrow, and the rest of the IT staff at NSSL are deserving of great praise for their easy approachability and dedication to keeping things running smoothly.

Lastly, I cannot express enough my gratitude for my family: my wife, Ann, and children, Joseph and Sara. Ann has sacrificed much for me to reach this point, especially over these last few months; I am blessed to have her. Joseph and Sara are a source of great joy and make it difficult to leave the house in the morning. To Sara's oft-stated "I

want Daddy stay home," I can only say, "Me, too." Thank you to my parents for their continued love and support. My thanks, as well, to my mother-in-law, Sally Coleman, for her prayers and for enlisting the prayers of the Holy Spirit Adoration Sisters, the "Pink Sisters," and to many others whose prayers have been a great support.

Table of Contents

List of Tables

List of Figures

Abstract

Mesoscale convective systems (MCSs) are a dominant climatological feature of the central United States and are responsible for a substantial fraction of warm season rainfall. Yet, very little is known about the predictability of MCSs. To help address this situation, a series of ensemble simulations of a MCS are performed using two- and three-dimensional versions of a storm-scale (Δx = 1 km) model and compared with an idealized control run. Ensemble member perturbations in wind speed, relative humidity and convective instability are based on current 24-h forecast errors from the North American Model (NAM). The ensemble results thus provide an upper bound on the predictability of mesoscale convective systems within realistic estimates of environmental uncertainty, assuming successful convective initiation.

The two-dimensional simulations are assessed by considering an ensemble member a success when it reproduces a convective system of at least 20 km in length (roughly the size of two convective cells) within 100 km on either side of the location of the MCS in the control run. By that standard, MCSs occur roughly 70% of the time for perturbation magnitudes consistent with 24-h forecast errors. Reducing the perturbations for all fields to one-half the 24-h error values increases the MCS forecast success rate to over 90%. The same improvement in forecast accuracy leads to a 30-40% reduction in maximum surface wind speed uncertainty, a roughly 20% reduction in the uncertainty in maximum updraft strength, and initially slower growth in the uncertainty in the size of the MCS. However, the occurrence of MCSs drops below 50% as the mid-layer mean relative humidity falls below 65%. The response of the model to reductions in forecast errors for

convective instability, moisture, and wind speed is not consistent and cannot easily be generalized, but each perturbation type is found to have a substantial impact on forecast uncertainty.

Results from the three-dimensional simulations resemble those found in two dimensions. The MCS success rate for the three-dimensional runs is around 70% for perturbation magnitudes consistent with 24-h forecast errors, while reducing the initial uncertainty improves the success rate to nearly 85%. The maximum updraft and surface wind uncertainties are of similar magnitude to their two-dimensional counterparts. However, whereas the response of the model to improvements in the initial uncertainty is inconsistent in the two-dimensional experiment, it is non-existent in three-dimensions. For most storm features, improvement in the forecast uncertainty requires the reduction of initial uncertainty for all three of the fields together. The three-dimensional runs often produce solutions that resemble bow echoes, but surface winds associated with these solutions, and the perturbation profiles that produce them, are indistinguishable from the non-bowing solutions, making any conclusions about the bow-like systems difficult.

Chapter 1: Introduction

In the central United States, mesoscale convective systems (MCSs)—a collection of thunderstorms with a contiguous precipitation area extending at least 100 km and lasting for at least 3 h (Parker and Johnson 2000)—are one of the dominant convective features in the warm season and are responsible for a substantial fraction of warm season rainfall (Fritsch et al. 1986). Schumacher and Johnson (2006) find that 74% of all warm-season extreme rain events over the eastern two-thirds of the United States during the period 1999-2003 are associated with an MCS. Consequently, variations in MCS activity can contribute to flooding (e.g., Junkers et al. 1995) or drought (Fritsch et al. 1986). At the same time, forecasts of warm season rainfall continue to possess little skill (Fritsch and Carbone 2004); the challenge is particularly acute for MCS forecasts in weakly forced environments (Stensrud and Fritsch 1994; Jankov and Gallus 2004). In addition to their importance to seasonal rainfall, MCSs are also responsible for a wide range of severe weather, including flash floods, hail, strong surface winds and tornadoes (see Weisman and Trapp 2003 and references therein).

As the increase in computational resources continues apace, the operational use of low-single-digit horizontal grid spacing numerical weather prediction models approaches. Indeed, the Met Office already produces a 4-km run over the UK (see http://www.metoffice.gov.uk) and recent Storm Prediction Center/National Severe Storms Laboratory Spring Programs have utilized daily 2- and 4-km versions of the Weather Research and Forecast (WRF) Model over a domain covering ~2/3 of the CONUS (e.g., Weiss et al. 2005). Compared to larger scales, however, very little is

known about predictability of phenomena on these scales. This is true even for phenomena as extensively studied as MCSs.

Some indirect investigations of MCS predictability consist of attempts to develop forecast aids for existing systems, whether simple extrapolations of the storm movement (Boucher and Wexler 1961) or prediction of system maintenance or decay based on environmental conditions (Coniglio et al. 2007). Early forecast experiments with high-resolution numerical models included both squall-line and supercell events (Brooks et al. 1993; Wicker et al. 1997) and achieved some qualitative predictive success. However, there is no apparent a priori information available to indicate when the model is capable of a successful forecast, rendering their value questionable.

In a more direct predictability study, Stensrud and Wicker (2004, hereafter SW04) employ a storm-scale model in two dimensions in order to make possible the production of 100-member ensembles under a variety of environmental and perturbation scenarios, for a total of 2700 simulations. Instead of using infinitesimal perturbations or initial errors consistent with analysis uncertainty, as is more commonly used, SW04 take the perspective of a forecaster looking at numerical model guidance for issuing a day 1 or day 2 prediction. If the forecaster sees the model developing an MCS in his forecast area 12 or 24 h into the run, what level of confidence can he place in the occurrence of that event or its subsequent evolution? Therefore the initial perturbations used to generate the ensemble members are based on the 12- and 24-h forecast errors of a mesoscale model compared against upper-air observations. Using perturbation magnitudes consistent with 24-h forecast errors, only 60% of the ensemble members produced an MCS located within 200 km of the control run. SW04 note that raising that success rate to 90% would

require improving the forecast error to magnitudes at or below even what is currently possible in model analysis systems, which are limited by instrument, representativeness, and model errors.

The SW04 results raise the question of the relevance of two-dimensional (2D) simulations. Modeling surface-forced thermal circulations in LES-type systems, Zeng and Pielke (1993) find the predictability estimates from 2D runs to be suitable approximations to the three-dimensional (3D) results, as long as the surface forcing is homogeneous in the third dimension. Subsequently, 2D models have been used to study the predictability of tropical squall lines (e.g., Clark et al. 2003) and isolated convection in the Southern Plains (e.g., Derbyshire et al. 2004; Petch 2004), and as a form of convective parameterization for general circulation models (e.g., Randall et al. 2003; Grabowski 2003). Khairoutdinov and Randall (2003) compare two- and three-dimensional simulations of 28-day runs of a storm-scale model. Consistent with earlier studies (see Moeng et al. 1996), they find both versions produce similar results for first-order fields: the 3D run is somewhat wetter, but the vertical profiles of temperature and water vapor errors and of vertical mass flux are very similar, the latter despite stronger 3D updrafts even. Modeling tropical convection in both two and three dimensions, Grabowski et al. (1998) also find that mean fields are well represented in 2D, and note that the biggest discrepancies are associated with the evolution of features and, thus, the temporal variability of quantities. Temporal variability is larger in 2D simulations, but this is largely a problem of sample size—the reduced domains of the 2D runs include fewer areas containing high-amplitude features, such as heavy rain, than their 3D counterparts and so the domain average of these quantities is less stable. However, if the

statistics are averaged over an ensemble of 2D simulations, then the lower variability of the 3D runs can reasonably be approximated (Grabowski 2004). More directly relevant to the present study, in a comparison of 2D and 3D simulations of squall lines, Rotunno et al. (1988) conclude that, while there exist differences in the details of the simulations, there is "no *essential* difference" between the two models as long as the low-level wind shear is not too strong. Including Coriolis effects also can also result in distinctive 3D behavior (Skamarock et al. 1994). Complete verisimilitude in the details of the modeled MCSs is not to be expected, yet confidence in the predictability results may still be reasonable, as long as the simulations are not carried forward so long as to make Coriolis effects non-negligible.

The research undertaken herein is based on, and extends, the preliminary work of SW04, including a test of the hypothesis that 2D simulations can provide realistic predictability estimates. This study is unique in the field of the predictability in that it is focused on a single weather phenomenon (MCSs) of strong interest to a broad sector of society, including agriculture, transportation, and emergency management. Furthermore, predictability is investigated using large ensembles (50-100 members), an approach not computationally practical until recently.

Chapter 2 presents an overview of predictability research from the early recognition of forecast uncertainty through current studies with storm-scale models, along with a discussion on measuring predictability. The numerical model is described in Chapter 3, along with an explanation of the perturbations used to initialize the ensemble members. Chapter 4 examines an update of the SW04 research including a description of the changes required by the addition of ice microphysics, which is not present in the SW04

simulations. The applicability of the 2D model formulation is then tested by comparing the results against ensembles that are identical in every way save the addition of the third dimension, and the results from these 3D results are presented in Chapter 5. Chapter 6 provides a summary.

Chapter 2: Review of predictability studies

The recognition of uncertainty in forecasting has been around, in some form, for over 200 years (see discussion by Murphy 1998). In the early 20th century there was a rise in the belief that these uncertainties would disappear as improvements in scientific knowledge and observations were made (see Murphy 1998; Anthes 1986). In a noteworthy exception to this optimism, H. Poincaré, in tackling the three-body problem in celestial mechanics, discovered the potential for a small initial uncertainty to result in a sufficiently large error at some later time as to render prediction impossible. Poincaré distinguished this scenario from one in which the initial cloud of uncertainty remained unchanged with time and even suggested that the challenge of weather forecasting could be attributed to the intrinsic behavior of the atmosphere (Poincaré 1908; see Ruelle 1998 and Murzi 2006). However, Poincaré's deductions lay fallow for 50 years before they were accidentally re-discovered when longer-term numerical model solutions for Rayleigh-Bernard convection were compared with solutions that were halted and restarted at some intermediate runtime (Lorenz 1963). The saved data were truncated for expediency. The small errors introduced by this truncation grew, eventually producing substantially different solutions.

In further experiments, Lorenz (1963) demonstrates that this uncertainty is not simply a result of an imperfect knowledge of the physics involved or an inability to represent those processes, but rather it is an inherent property of the system. In particular, whether a system maintains unlimited predictability depends on whether the system possesses periodic behavior; non-periodic systems have limited predictability.

Echoing Poincaré once more, three predictability scenarios are possible (Lorenz 1969): 1.) infinitely predictable systems in which errors stay the same or decay with time; 2.) systems with finite predictability, but unlimited capacity for extending the predictability—errors grow with time, but reducing initial error leads to a proportional reduction in error at time *t*; 3.) systems with intrinsic predictability limits in which errors grow with time and the predictability limit can not be improved by further reductions in the initial error. Until the last half-century, the atmosphere was believed to fall into the second category. For example, in a comprehensive examination of atmospheric predictability, Thompson (1957) views the biggest hurdle to improved forecasts to be administrative. The error at time, *t*, is assumed to be proportional to the initial error, but increases in the density of the observing network are expensive. For each, say, doubling of the network density a smaller portion of the initial uncertainty remains. Thus, successive improvements in observing network become more expensive while yielding a smaller return in terms of forecast improvement. The limit to atmospheric predictions is then a practical societal one and not one imposed by the nature of the atmospheric system itself.

A self-described metaphysical argument that the atmosphere instead falls into the third category is offered by Robinson (1967). He notes that the Navier-Stokes equation predicts the motion of a particle of fluid and proposed that, because of diffusion, each particle has a characteristic lifetime (a function of the characteristic length scale and the coefficient of self-diffusion) and so the Navier-Stokes equation is only valid for the lifetime of the particle. Robinson proceeds to make some back-of-the-envelope calculations of the predictability limits implied by his theory: 3-5 days for synoptic-scale

flow and about one hour for storm-scale motions. However, most of the early predictability studies used simplified turbulence models.

a. Turbulence-based models

Using a turbulence-based approach, the numerical experiments of Lorenz (1969) are in basic agreement with Robinson (1967). The energetics argument of Lorenz (1969) states that, regardless of grid resolution employed, there will always be unresolved small-scale eddies. Whereas the statistics of these eddies may be determined by the state of the large-scale flow, there exists an associated expected error to those statistics owing to the fact that the individual eddies are at different stages within their life cycles and thus possess considerably different amounts of energy. Between scales differing by several orders of magnitude, the number of small-scale eddies present within a single grid box is sufficient to allow cancellation yielding an average eddy kinetic energy very close to the expected value. Thus, there is little direct effect of the small-scale errors on the larger scales. However, when considering more nearly adjacent scales, the number of unresolved eddies within a grid box is relatively few and the average kinetic energy within the grid box can vary significantly from the expected value, allowing errors to efficiently transfer upscale. Under regimes dominated by two-dimensional turbulence, the time scale of the unresolved eddies is on par with the large-scale eddies and so a halving of the initial error results in a doubling of the predictability limit. Under regimes dominated by three-dimensional turbulence (e.g., convection), however, the turnover time of eddies decreases as the scale decreases, and so improving initial uncertainty becomes

an exercise of diminishing returns in terms of increasing predictability (Lorenz 1969; Tennekes 1978).

One shortcoming in the argument of Lorenz (1969) is that the while he employed the two-dimensional vorticity equation, the main conclusions are based on the assumption of a -5/3 energy spectrum (Leith 1971). Lorenz (1969) explains that the -5/3 spectrum was chosen because it "seems to place a reasonable amount of energy in the cumulus scales." However, earlier work in turbulence theory (e.g., Kraichnan 1967), suggests that two-dimensional homogeneous isoptropic turbulence should have a -3 energy spectrum; the -5/3 spectrum is expected for three-dimensional homogeneous isotropic turbulence. That is, when system kinetic energy is plotted as a function of wavenumber, there will exist a range for which the decline in energy with scale follows a p^{-k} power law, where $k = 3$ for two-dimensional turbulence and $k = 5/3$ for three-dimensional turbulence. Furthermore, early analyses of observations (e.g., Wiin-Nielsen 1967) detected a -3 power law for large-scale atmospheric flow. More recent observations (Nastrom and Gage 1985) replicate the -3 spectrum for large scales (> 800 km) transitioning to a -5/3 spectrum for smaller scales (< 500 km). Note that the upper end of the range for the -5/3 spectrum is too large to include isotropic three-dimensional turbulence. A conclusive explanation for the shallower slope in the mesoscales has not yet been provided (see discussions in Lilly 1983 and Skamarock 2004).

A closer look at the nature of turbulence in two and three dimensions reveals the importance of choosing an appropriate spectral slope. In barotropic, non-divergent (i.e., two-dimensional) flow, absolute vorticity is conserved along trajectories. In the absence of any vorticity sources or sinks, vorticity is simply advected by the flow itself.

However, random motions will induce shearing, which tends to redistribute the vorticity into filaments of increasing length and decreasing thickness. Hence, as the flow evolves, vorticity variations will move toward smaller and smaller scales, a property described as an "enstrophy cascade" (Leith 1968; Tennekes 1978; enstrophy is equal to half of the mean-square vorticity).

Whereas, in two dimensions, shearing acts to elongate and narrow vorticity filaments such that vorticity is conserved, as described above, the stretching induced by the differential vertical motion resulting from convergence and divergence acts to amplify vorticity in three dimensions. When motions are not constrained to be nondivergent, vorticity amplification occurs at all scales, but it tends to focus kinetic energy into scales smaller than the embedded flow (Tennekes 1978). Consequently, analogous to the enstrophy cascade present in two-dimensional turbulence, three-dimensional turbulence is characterized by an energy cascade toward smaller scales.

Applying scaling arguments to the above two scenarios results in a predicted -3 energy spectrum for two-dimensional turbulence and a predicted -5/3 spectrum for three-dimensional turbulence (Tennekes 1978). Consistent with the preceding discussion, scaling arguments show that for two-dimensional turbulence the intensity of the vorticity is independent of scale, while, for three-dimensional turbulence, the magnitude of the vorticity is inversely proportional to the scale of motion, such that vorticity increases as the spatial scale decreases. Eddy turnover times (i.e., the time in which an air parcel with an eddy makes a complete circuit) can be estimated by the reciprocal of the vorticity. Therefore, eddy turnover times are independent of scale for two-dimensional turbulence but decrease with the scale of the flow in three-dimensional turbulence.

Using the -3 energy spectrum that is more physically consistent with two-dimensional turbulence results in considerably longer predictability time estimates (Leith 1971; a different turbulence approximation is used as well, but the results are much more sensitive to the energy spectrum imposed than to the dynamical system, see Rotunno and Snyder 2008). These results lead Leith to conclude, in agreement with Thompson (1957), that atmospheric predictability could be continually lengthened by successive increases in the resolution of the observing network. However, repeating the experiment with a -5/3 energy spectrum (Leith and Kraichnan 1972), one finds predictability to be a function of the local eddy turnover time, consistent with Lorenz (1969). The loss of predictability is, though, somewhat slower when a dynamical system that is physically consistent with a -5/3 spectrum is used: ~10 times the local eddy time scale for three-dimensional turbulence compared to ~6 times the local eddy time scale for two-dimensional turbulence. Under the assumptions used by Leith and Kraichnan (1972), the implication for mesoscale forecasts is not as dire as the Lorenz (1969) estimates, but the basic scenario remains: the connection between predictability, eddy turnover times, and the scale of the flow means that there is an inherent limit to atmospheric predictability that cannot be extended by scientific or technological improvements.

Interestingly, when focusing on the spread of error energy between scales, the main difference between the two regimes is that, since there is more energy present in the small scales for the -5/3 scenario, large-scale errors can interact with the small scales resulting in a substantial downscale error-energy transfer (Rotunno and Snyder 2008). Consequently, while error energy is maximized in the slowly growing large scales for the -3 scenario, it is maximized in the smallest unsaturated scale for the -5/3 case (Rotunno

and Snyder 2008). This difference can be seen in the error growth curves (Fig. 1). After a brief period of initial adjustment error energy grows at a constant rate across the -3 spectrum (Fig. 1b). For the -5/3 case, the curve is actually composed of a series of increasingly long linear segments of increasingly shallow slope (Fig. 1a). Each segment represents a given length scale; after the error in each scale saturates (i.e., the error energy is equal to the total energy at that scale) error growth continues at the slightly slower rate characteristic of the next smallest unsaturated scale. This process is even more apparent when the number of scales of motion is reduced to two (Fig. 2). Initially, error growth is dominated by the "fast" component associated with the smaller scale until it saturates by $t = 2$ (non-dimensional time units) at which point the "slow" component associated with the larger scale takes over as if the smaller scales did not exist. Error growth continues until the slow component approaches saturation around $t = 10$ (see also Lorenz 2006; Zhang et al. 2007).

A novel critique of this standard view of atmospheric predictability claims the Lorenz (1969) model to be insufficiently nonlinear possessing too few degrees of freedom (Robert and Rosier 2001). Under a more realistic system, the downscale enstrophy cascade and upscale energy cascade present in turbulent flow lead to a form of self-organization of coherent structures that are significantly more predictable. An analogy is drawn to suggest that small-scale chaotic motions no more affect the predictability of the large-scale flow than molecular thermal agitation affects the general properties of the gas those molecules comprise. However, there are several reasons to suspect the relevance of this claim. One, it seems to rely on a stark separation between the coherent large scales and chaotic small scales, but Lorenz (1969) notes that direct

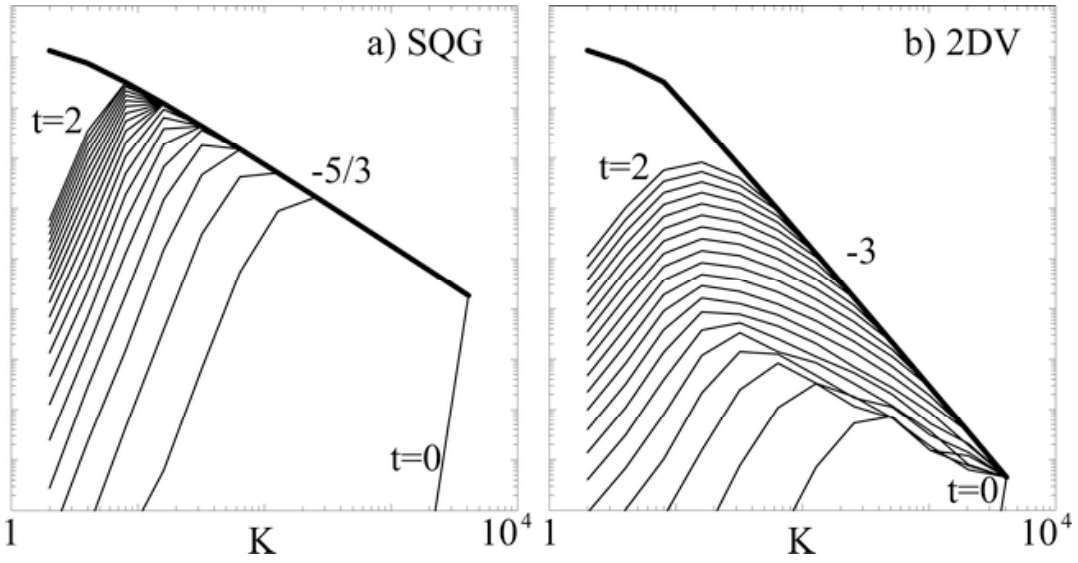

Fig. 1. Error energy per unit wavenumber for $t = 0, 2$ in steps of 0.1 for (a) surface quasi-geostrophic (SQG) turbulence and (b) two-dimensional vorticity (2DV) turbulence. The heavy solid line indicates the base-state kinetic energy spectra per unit wavenumber, which has a -5/3 slope for SQG and a -3 slope for 2DV. (From Rotunno and Snyder 2008.)

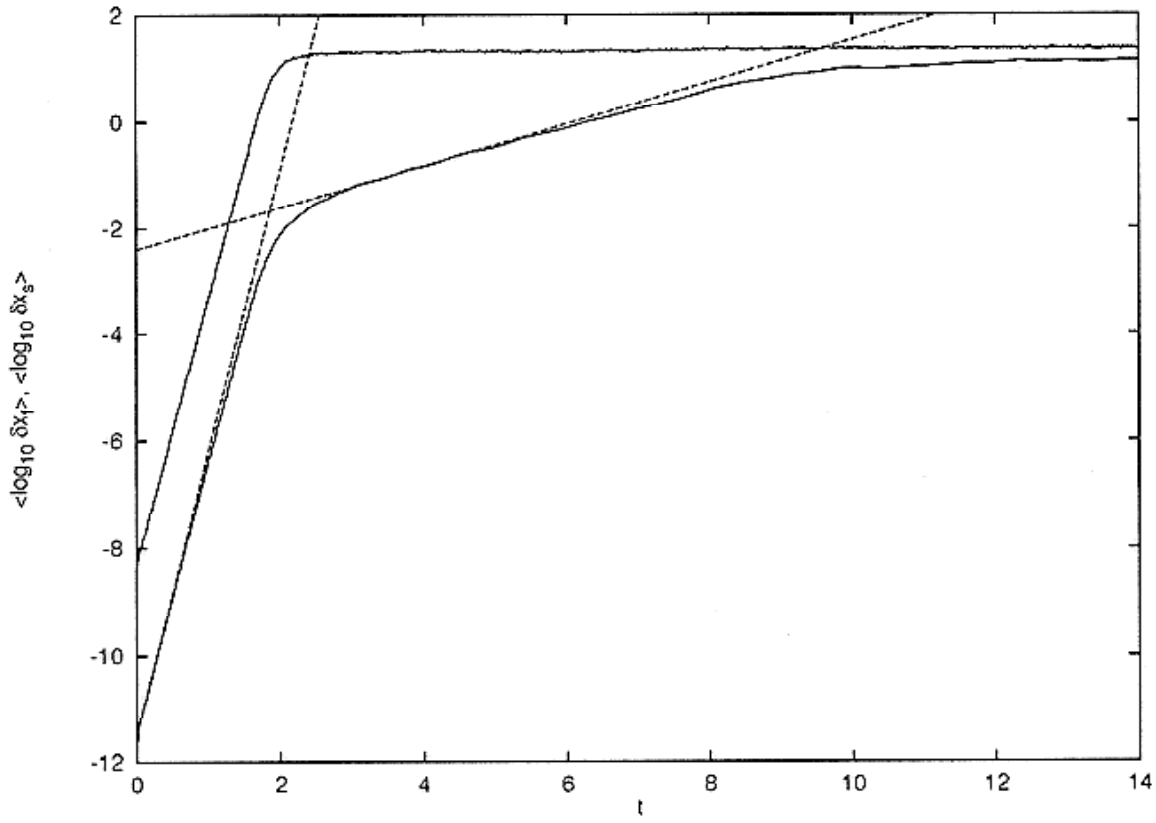

Fig. 2. Typical error growth for the fast component δx_f (upper curve) and for the slow component δx_s in coupled Lorenz models with $\delta x_f(0) = 10^{-8}$ and $\delta x_s(0) = 10^{-12}$, averaged over 500 samples. The dashed lines show the exponential growths with exponents $\lambda^{(f)}$ and $\lambda^{(s)}$. (From Boffetta et al. 1998.)

spreading of error energy from small scales to large scales is almost non-existent; the theory is based on flows possessing many scales of motion such that upscale error energy spreading occurs between nearly adjacent scales. Two, even within their constraints, there exist "separatrix" decision points in the flow between two distinct coherent structure solutions seemingly analogous to the decision points between the two lobes of the Lorenz (1963) attractor. Finally, Robert and Rosier (2001) acknowledge the oversimplification that their system admits only a single coherent structure whereas the atmosphere is composed of a finite set of such structures between which the nonlinear interactions can be chaotic. The nonlinear interactions can create intermediate scales, leading back to the first point above.

Still, the relevance of using turbulence-based models for predictability studies at the mesoscales and below raises an important question. The key dynamical phenomena at these scales appear to be gravity waves and convection (see Lilly 1983). Gravity waves are not treated at all by turbulence-based predictability models, while moist convection is highly intermittent—recall that the models assume the turbulent flow to be continuous in space and time—and involves the important, but untreated, process of latent heat release. Accounting for intermittency in turbulence models appears to steepen the spectral slope, at least initially (Basdevant et al. 1981; Lilly 1983), and extend the predictability limits (Crisanti et al. 1993). This steeper spectral slope indicates less small-scale energy, thereby reducing the downscale spread of error energy. However, the spectral slope predicted by Lilly (1983) for intermittent turbulent flow, $k = -2$, still indicates a bounded predictability horizon. Hence, whereas studies involving global atmospheric models (e.g., Lorenz 1982; Simmons et al. 1995) appear to support the

findings of Lorenz (1969) for large scales, the predictability of the mesoscale and smaller scales remains an open question (see also the discussion in Straus and Shukla 2005). Studies of mesoscale predictability awaited the development of mesoscale models.

b. Mesoscale models

The advent of numerical model simulation of the mesoscale brought with it claims of enhanced predictability at these scales (e.g., see the review by Anthes 1986) due to fixed forcing by terrain and other surface features. The conclusions by Anthes (1986) are based on only a few cases comparing the solution from a control run with that from a run with randomly perturbed initial conditions in which error growth (measured as the root-mean-square difference between the 500-hPa heights from two runs) is negligible over three days of integration. The conclusions were soon identified as being more emblematic of the experimental design than of the atmosphere. Limited computational capacity meant that focusing on the mesoscales required a shift from global to limited area models that, typically, still had relatively large grid spacing, on the order of 100 km. Grid spacing this large fails to resolve small-scale instabilities that may be responsible for fast early error growth (Berri and Paegle 1990). Indeed, Anthes (1986) includes a caution that the results would likely not apply to smaller scales where convective instability becomes more prominent. Three additional constraining influences affecting the error growth in Anthes (1986) have been recognized (Errico and Baumhefner 1987): strong numerical dissipation quashes small perturbations before they can grow; some of the initial perturbation projects onto gravity waves, which are dispersed through geostrophic adjustment and subsequently propagate out of the domain or are dissipated; the limited

domain combined with the use of identical lateral boundary conditions meant that as the numerical integration progresses differences are swept out of the downwind edge of the domain while identical conditions are introduced from the upstream edge. Along with error sweeping, a further effect of (one-way) lateral boundaries is that large-scale growth is prohibited (Vukicevic and Errico 1990). The large-scale solution is constrained to match the imposed lateral boundary conditions and scales larger than that determined by the model domain size are not present, though some large-scale error growth can be achieved by increasing the size of the limited area domain (Vukicevic and Errico 1990; Vukicevic 1991). Still, when using measures such as root-mean-squared-error, the two effects of the use of lateral boundaries are manifested as error growth curves that, after a period of initial growth, quickly level off to a value substantially less than expected (Laprise et al. 2000; Vannitsem 2003; Nutter et al. 2004).

Another shortcoming of many early studies of mesoscale predictability is a lack of moist physics (e.g., Farrell 1990; Vukicevic 1991; Ehrendorfer and Errico 1995). Ehrendorfer and Errico (1995) note that the lack of error growth may result from the fact that amplification of some perturbations may require nonlinear processes, particularly moist physics. Including moist processes increases the rate of forecast error growth by about an order of magnitude and reduces the scale of those errors, such that even parameterized convection is capable of producing new instabilities (Ehrendorfer et al. 1999).

The danger of misinterpreting model artifacts as properties of the atmosphere highlights the model dependence of any predictability study. Not only does the presence of model error place a severe restriction on the forecast improvement that can be obtained

by improving the initial analysis (Tribbia and Baumhefner 1988), but also it can prejudice one's view on the ultimately predictability inherent in the system (i.e., the atmosphere). Predictability cannot be removed from the context of the model framework being used (Schneider and Griffies 1999; Smith 2006). As a result, theoretical predictability limits can change as our level of understanding increases and models are improved (Smith 2006). This is true whether one is comparing model output to observations or employing the perfect model scenario (i.e., assuming that the model perfectly captures the full spectrum of physical processes at work in the atmosphere and thus would give an exactly correct forecast if exactly correct initial conditions could be supplied).

Additional insight into the predictability of mesoscale model forecasts can be gained by examining the error growth at different scales and the interactions between scales. Recall that in turbulence models, the error energy spread from the large scales to the small scales is of the same order of magnitude as that which occurs between adjacent small scales (Lorenz 1969; Rotunno and Snyder 2008). Typically, the small scales are found to be subject mainly to small-scale forcing with little impact on or from the larger-scales (Zhang and Fritsch 1986; van Tuyl and Errico 1989; Zhang et al. 2002). For experiments with low horizontal resolution, topography is the primary source of the small-scale forcing (e.g., van Tuyl and Errico 1989). The dominance of topography as a small-scale forcing mechanism leads to the conclusion that mesoscale flow is likely to be reasonable predictable, while acknowledging that predictability may be lower in regions where internal dynamics are predominant (Anthes 1986; van Tuyl and Errico 1989; Vukicevic and Errico 1990). When the grid spacing is reduced to about 30 km or less, moist convection becomes an important source of forecast error (e.g., Zhang and Fritsch

1986; Zhang et al. 2002). Consequently, for these higher resolution simulations, an overall finding of weak error growth is a result of higher-energy large scales, where error growth is restricted, dominating the lower-energy, but rapidly divergent, small scales (Laprise et al. 2000; Zhang et al. 2002). The important role of moist convection in small-scale error growth can be illustrated by removing the effects of convective latent heat release; the resulting pair of simulations exhibit substantially reduced small-scale error growth (Zhang et al. 2002).

The perceived degree of predictability is affected greatly by how error growth is measured. Fourier analysis is a common tool for scale decomposition. However, isolated features project onto a wide range of scales defined by Fourier functions, and so even very intense isolated maxima (such as convection) have less effect on the power spectra than broader-scale forcings such as topography (van Tuyl and Errico 1989). In addition, the specific atmospheric field examined can color the results. Fields such as 500-hPa heights or sea level pressure are indicative of what is occurring on the larger scales while precipitation may be strongly tied to small-scale errors. For example, the initial perturbations in Zhang et al. (2002) have a substantial impact on the subsequent evolution of the precipitation fields, but sea-level-pressure fields are affected only slightly.

c. Storm-scale models

Continued computational advances recently have enabled the use of storm-scale models (i.e., 2-3 km grid spacing) to examine further the role of convection in small-scale predictability and whether some enhanced predictability exists at these scales. In a follow up to Zhang et al. (2002), Zhang et al. (2003) apply a more systematic approach to

a poorly forecast snowstorm associated with a low-pressure system off the East Coast of the United States. In place of the initial conditions derived from separate operational analyses as in the earlier work, a height-independent sinusoidal temperature perturbation is added to the control run to initialize a twinned solution (i.e., the two runs are alike in all ways except the initial conditions). Errors grow rapidly, particularly over the first few hours, and the scale of the maximum error energy steadily increases with time, suggesting an upscale error cascade as predicted by Lorenz (1969). Also in line with the scenario presented by Lorenz (1969), error growth rates increase as the amplitude of the initial error is decreased. This behavior is most pronounced over the first 3-5 h, but continues throughout the 36-h simulation. For the largest amplitudes, error grows slowly, if at all (see also, Hohenegger and Schär 2007b). Recall that the error energy saturates at a level equal to the total energy at that scale; therefore, as the initial error magnitudes approach the total energy, little remaining error growth is possible. Large initial error magnitudes on the small scales could be another cause of the weak error growth present in the early mesoscale predictability studies.

Growth of the initial errors increases as the grid spacing is reduced from 10 km to 3.3 km (Zhang et al. 2003). The error doubling time for the higher-resolution runs is less than 1 h, indicating that reducing the initial errors by half would only extend forecast skill by less than 1 h. The rapid error growth begins at the scale of the perturbation but then quickly grows upscale, reaching the mesoscale within 2 h. Model error covariances suggest that errors can spread to scales of 1000 km or more (Zhang 2005). In studying summertime convection over a region covering southeast France, Switzerland, and northern Italy, Walser et al. (2004) similarly found that small-scale errors tied to

convection quickly contaminate larger scales. Echoing the early mesoscale studies, Walser and Schär (2004) note that one case of moderately strong convection located over the Alps shows a higher degree of predictability. Moreover, even within that case, the precipitation over the mountains is found to be more predictable than that which occurs over the nearby plains. In general, though, mesoscale perturbations are able to remain in the domain and grow over extended time periods, impacting the overall forecast uncertainty after about a day (Walser et al. 2004), somewhat later than that seen in Zhang et al. (2003).

Predictability at small scales appears to be tied closely to convection over or near mountains (Walser et al. 2004; Walser and Schär 2004), associated with a coastal low-pressure system (Zhang et al. 2003), within an idealized simulation of a baroclinic wave (Tan et al 2004), or in a weakly-forced heavy precipitation event over south Texas (Zhang et al. 2006). Zhang et al. (2003) note that the regions of strong error growth are coincident with convectively unstable regions. The error growth between simulations is tied to times and locations where small changes in the model fields can push the solution across a threshold (e.g., trigger convection) creating a nonlinear jump in the solution trajectory away from its twinned realization. (For example, imagine two convectively unstable regions: one exists in an area with a large cap, the other in an area with little inhibition. Small perturbations in the former scenario will likely have no impact, while they can have a significant impact on the latter. This leads to very different error growth in the two regions.) If the perturbations are too small to initiate convection, then perturbation growth must wait for convection to occur in the control run.

At lower resolutions ($\Delta x > 10$ km), early nonlinear growth is due to the "on-off" switches in the convective parameterizations, while at higher resolution ($\Delta x = 1$ to 4 km) the on-off switches reside in the microphysical schemes (Zhang et al. 2006). Despite still involving on-off switches in the microphysical schemes, non-convective precipitation does not correlate with regions of significant error growth (Tan et al. 2004). Typically, the stronger motions associated with convection produce larger error amplitudes than the weaker flows associated with non-convective precipitation, regardless of the fidelity of the model physics. Convection can introduce and amplify errors in both temperature fields (i.e., through latent heat release) and momentum fields (i.e., through shear production) (Zhang et al. 2007). While the latent heat release from non-convective precipitation can be substantial, the rate of latent heating is generally much lower and shear production is minimal. Thus, rapid error growth is not tied to precipitation in general, but specifically to regions of convection.

However, the presence of convective instability is not sufficient to guarantee rapid error growth. Along with the need for the perturbations to push fields beyond certain thresholds, it has been suggested (Hohenegger et al. 2006) that for perturbations to grow significantly, they must remain in areas of instability for an extended period; perturbations that rapidly move through unstable regions experience only transient growth. So, upstream propagation of gravity waves, by keeping the perturbations in convectively unstable regions longer, should be associated with lower predictability (Hohenegger et al. 2006). Indeed, gravity waves have been implicated as playing a major role in small-scale error growth. Hohenegger and Schär (2007b) show that some portion of the initial perturbations project onto both fast, small-amplitude waves (see also Errico

and Baumhefner 1987), which then disperse the initial error rapidly through the domain. These small-amplitude perturbations can experience substantial amplification through forced ascent over topography or, if the trigger thresholds are exceeded, through convective activity. (The faster, smaller-amplitude sound waves tend to simply propagate the perturbations out of the domain, while the slower, somewhat larger-amplitude gravity waves are more likely to play a role in subsequent error growth.) The larger-scale environment determines the saturation level, according to the amount of energy present at each scale in the environment, and the preferred regions of error growth, such as in the warm sector of a developing low-pressure system.

Zhang et al. (2007) expand on these error growth steps by offering a three-stage conceptual model explaining how initial small-scale error moves upscale to produce synoptic-scale errors within 24 h:

1.) During the first few hours (< 6 h) of the simulation convection is the dominant source of error growth. Individual convective cells quickly become displaced between a pair of control and perturbed runs, such that the magnitude of the errors reaches parity with the amplitude of the fields themselves, implying error saturation. Gravity waves then disperse the error energy away from the area of convection in the model.

2.) Over the next 6-12 h the unbalanced small-scale errors are transformed into errors in the larger-scale balanced motions, possibly through geostrophic adjustment (in response to the latent heating).

3.) Finally, the errors in the balanced fields can grow according to large-scale baroclinic instability.

It should be noted that the processes involved in earlier stages do not cease during the later stages. Convection can continue throughout the period along with the subsequent unbalanced response. In fact, while the balanced component of the difference fields increases with integration time, the unbalanced component remains nearly as strong (Zhang et al. 2007). This may explain why geostrophic adjustment is believed to play an important role in the upscale error growth by Zhang et al. (2007) whereas Errico and Baumhefner (1987) identify geostrophic adjustment as a one of the mechanisms limiting error growth in mesoscale models. In the latter simulations, geostrophic adjustment sweeps the initial, random perturbations out of the domain quickly, while in Zhang et al. (2007) the atmosphere is adjusting continually to the sustained latent heat release and shear production resulting from convective activity.

Interestingly, the three-stage model proposed by Zhang et al. (2007) presents a parallel between the large and small scales. At small scales, gravity waves transport the initial random perturbations to locations of convective instability where those errors can amplify. At larger scales, the inertia gravity wave response to the convectively generated latent heating transports errors to locations of baroclinic instability where those errors can amplify. Zhang et al. (2006, p.165) state it as follows: "It is possible that convective instability determines error growth at smaller scales, while large-scale baroclinic…instability dictates upscale energy transfer and determines error growth at larger scales." One result of the dominant roles of convective and baroclinic instability is that there should be forecast scenarios in which extended predictability is expected. Small-scale error growth should progress slowly in convectively stable environments and similarly for large-scale error in environments of low-baroclinicity (see Zhang et al.

2006). Indeed, casual observation of the weather bears this out, e.g., the highly persistent conditions often associated with high-pressure systems. However, many weather events of high societal impact occur outside of these more predictable scenarios.

One common feature of the Hohenegger and Schär (2007b) and Zhang et al. (2007) conceptual models of error growth is that to a large extent the mesoscales are ignored. Involved is an interplay between the storm-scale error source and synoptic scales which are both affected by the smaller-scale errors and help determine the location and timing of the convection. A similar interplay has been proposed between the meso- and planetary-scales (Stensrud 1996; Stensrud and Anderson 2001), whereby the latent heat release from persistent mesoscale convective activity (i.e., repeated MCSs in the same region over the course of several days) becomes a source region for Rossby-wave trains. These MCS-induced large-scale circulations can then create conditions favorable for continued persistent mesoscale convective activity.

d. Observational studies of MCS predictability

While MCSs have long garnered the attention of the meteorological community (e.g., Tepper 1950; Gerhardt 1963; Sanders and Paine 1975; Maddox et al. 1986; McAnelly et al. 1997; Parker and Johnson 2000), very little attention has been focused on the predictability of these systems. The challenge of correctly forecasting MCSs includes having a firm grasp of the large-scale conditions, accurately predicting the initial convective activity, capturing boundary layer disturbances caused by this initial convection, and forecasting the evolution and interaction of these disturbances in order to obtain the correct secondary activity (Carbone et al. 1990). Carbone et al. state that

quantitative prediction of convective precipitation presents a "formidable" challenge on even shorter (6-12 h) time scales. A more optimistic perspective is that squall lines that are forced by synoptic-scale features may be more predictable as a result of "inheriting" the predictability of the larger-scale forcing (Weygandt et al. 2002). However, many MCSs are not associated with strong, large-scale forcing (Porter et al. 1955; Stensrud and Fritsch 1993); in fact, many of the features of midlatitude squall lines can be reproduced in models lacking any large-scale forcing (Skamarock et al. 1994).

A difficulty in trying to apply the predictability limits suggested by Lorenz (1969) to MCSs involves determining the relevant scale. The convective portion of an MCS may be on the order of 10 km in the cross-line direction, but may extend for many hundreds of kilometers in the along-line direction. One perspective is that the smaller, cross-line dimension is more representative of the individual convective cells within the system, while the larger, along-line dimension is more representative of the system itself. Using the estimates from Lorenz (1969) one would then expect a predictability limit in the range of 10-15 h for the MCS, while the individual convective elements would be predictable for less then an hour. Studies of long time series of (nearly) national radar data (Germann et al. 2006; Vasic et al. 2007) set the limit at the somewhat lower values of 5-10 h for the scale of the along-line dimension (~250-500 km). Germann et al. (2006) also found a corridor of higher predictability (of up to 10 h) to extend from northwest Kansas through Michigan, a favored region for warm-season MCS activity. Furthermore, Hovmöller plots of radar-derived rain-rates display a recurrence of coherent structures that suggest a potentially greater degree of predictability for MCS- or MCC-type convective systems (Carbone et al. 2002). However, a subsequent examination showed a

pair of numerical weather prediction models to be incapable of capturing these coherent episodes, though the relatively coarse resolution (Δx = 10-22 km) and convective parameterizations appear to be contributors to the failure (Davis et al. 2003; see also Bukovsky et al. 2006). Also, it is not clear to what extent persistent longitudinally-averaged radar features would translate into useful local forecasts of convective activity.

e. Measuring predictability

The concept of forecast utility introduces a primary question that must be considered for any predictability study, namely, what is meant by predictability and how is it measured. Expanding on the procedure developed by Thompson (1957), Lorenz (1969) focused on the error kinetic energy and, thus, looked indirectly at the predictability of the flow field. Synoptic-scale studies of predictability (e.g., Lorenz 1982; Dalcher and Kalnay 1987; Molteni and Palmer 1993) focus almost exclusively on 500-mb height root-mean-squared error (rmse). With the development of mesoscale models the target switched to sensible weather elements, particularly precipitation fields (e.g., Du et al. 1997; Stensrud et al. 1999). This new focus entails some fundamental difficulties, however. Wind and height fields have the advantageous property of being continuous and relatively smooth, especially on larger scales, and thus are well suited to measures such as rmse. Examination of precipitation forecasts introduced the difficulty of evaluating highly intermittent fields, a problem that becomes even more pronounced as model resolution increases. In one sense, this involves a shift from predictability of fields to the predictability of phenomena, but this also can be viewed as merely one step removed from the underlying error dynamics. For example, one may be interested in

how initial errors may affect the development and evolution of an MCS, but these phenomenological differences are nothing more than the response to the accumulated errors in the temperature, wind, pressure, and moisture fields as a result of the eddy transfer of the initial errors. In other words, if the phenomenon of interest is predictable for a given period of time, then the constituent flows are likely predictable as well. However, the intermittency of convection can cause problems for standard predictability measures such as rmse, a fact noted as far back as 1972 in Leith and Kraichnan who observed that such measures are unable to distinguish between trivial displacements and more important changes associated with dynamic instabilities.

Lorenz (1969) defined the range of predictability as "the time interval within which the errors in prediction do not exceed some prechosen magnitude," but he gave that "prechosen magnitude" the rather broad range of somewhere between the observational error and the difference between any two randomly chosen states. Once more, the intermittent nature of convection causes problems. One or both of the pair of mesoscale states chosen completely at random is likely to contain only weak convection or even no convection whatsoever. Thus, the average difference between a pair of randomly chosen states is likely to be smaller than the differences encountered in a predictability experiment. Additionally, what are two randomly chosen states when one is working with idealized simulations? Islam et al. (1993) propose the process standard deviation, defined as the variability of the field throughout the full space-time domain, of the control forecast as a reference. However, for a solitary, translating feature, such as a squall line, the process standard deviation likely yields nothing more than the average strength of the system.

Focusing on the mesoscale, Warner and Keyser (1983) propose that the predictability limit should be viewed as "the time required for a specified error in the mesoscale structure of one or more variables to cause the prediction of a specific quantity to be sufficiently in error so that it has essentially zero utility." By emphasizing utility, Warner and Keyser imply that predictability is user dependent as much as an inherent property of the dynamical system. This definition suggests a phenomenological approach to predictability, that is, in approaching the problem of the predictability of a certain phenomenon, one is implicitly asking how long a forecast of this phenomenon can provide useful information to the user community. Under this approach, one can step away from the direct field comparisons and focus instead on various properties of the storms, including measures of storm intensity and size. This approach also alleviates the strong sensitivity to phase errors possessed by point-to-point measures, particularly rmse.

Even if the inherent properties of the atmosphere limit the ability to forecast the exact location and features of individual storms, correct prediction of certain features of the system can still provide much in the way of useful information to forecasters. For example, imagine a particular forecast that misplaces the MCS by 50 km too far west and does not extend the line far enough to the south, but it shows that the convective downdrafts are strong enough to reach through the stable nocturnal boundary layer and bring severe winds to the surface. Assuming the model has a history of correctly identifying such scenarios, a forecaster can easily make the spatial adjustments and extend a severe weather watch. As touched on in Chapter 1, SW04 followed this approach by using as their predictability measure an MCS forecast success rate, the percentage of ensemble members that produce an MCS within 100 km on either side of

the location of the control simulation. The same measure is adopted herein, along with the ensemble variability of the MCS size and intensity, with the latter identified by the maximum updraft strength and maximum surface (lowest model level) wind.

f. Summary of prior research

After nearly five decades of research on atmospheric predictability, a consensus on what sort of uncertainty can be expected on the mesoscale remains elusive. Early turbulence-based studies, which have gained support from more recent studies on synoptic-scale predictability, suggest that skillful mesoscale forecasts are limited to lead times of only a few hours. On the other hand, it has been proposed that fixed forcing from topography or surface inhomogeneities should increase the predictability of mesoscale flows near those features or that some mesoscale phenomena might inherit the predictability of the larger-scale environment in which they reside. Modeling studies lend some credence to the idea of enhanced predictability near fixed surface features, but the idea of inherit certainty from the larger scales seems to be pure conjecture. In addition, distinct patterns within observed precipitation as estimated by radar are interpreted to indicate that greater predictability is possible. Nevertheless, high-resolution experiments focused on small-scale predictability are more in agreement with the early pessimistic estimates, although it must be noted that these results are based on only a few cases. Hence, perhaps ironically, a great deal of uncertainty remains as to what sort of predictability is expected for mesoscale phenomena.

The purpose of this study is to begin the task of determining the predictability for MCSs, an important, intermittent mesoscale phenomenon. The predictability

experiments described in the preceding sections primarily have been either general, statistical-based studies or case studies using small ensembles with as few as two members. In contrast, herein is described an experiment examining a single phenomenon, an MCS, using large, storm-scale ensembles: 100 members for 2D runs and 50 members for 3D runs. Because the MCS is simulated in an idealized setting (detailed in the following section) with a single initial environment, the conclusions reached certainly are not the final word on MCS predictability. They are, however, a starting point that, as will be explained later, likely provides an upper bound for more realistic scenarios.

Chapter 3: The numerical model and perturbation methodology

a. Description of the numerical model

All simulations were performed with the National Severe Storms Laboratory Collaborative Model for Mesoscale Atmospheric Simulation (NCOMMAS; Wicker and Wilhelmson 1995). The model was developed to study supercell dynamics and tornado genesis, but has been employed successfully to study a wide range of tasks, including: a tornado outbreak associated with a landfalling hurricane (Romine and Wilhelmson 2002), convective initiation (Houston and Niyogi 2007), dryline morphology (Peckham and Wicker 2000), the evolution of convective cells within an incipient squall line (Jewett and Wilhelmson 2006), and even severe-weather forecasting (Wicker et al. 1997). Recently it has been used to simulate MCS evolution by Coniglio et al. (2006), who provide an in depth description of the model as it has been updated in the intervening years (see their Appendix).

The model uses a simple height coordinate in the vertical with pressure represented in the perturbation Exner function, $\pi' = \left(\frac{p'}{p_0} \right)^{R/C_p}$, where p_0 = 1000 mb, R is the gas constant (287 m^2 s^{-2} K^{-1}), and C_p is the specific heat at constant pressure (1004 m^2 s^{-2} K^{-1}). The time integration employs a third-order Runge-Kutta scheme while spatial derivatives are approximated by fifth-order finite differences. For the final iteration of the Runge-Kutta scheme, flux computations use a weighted essentially nonoscillatory scheme as a numerical filter. Subgrid-scale mixing is modeled using a turbulent kinetic energy based scheme similar to that of Deardorff (1980). An open boundary is used for the lateral boundary conditions (the *y*-coordinate boundaries in the 3D runs are periodic)

and a Rayleigh damping layer is used near the lid. Microphysical processes are modeled by the three-class ice parameterization of Gilmore et al. (2004), a variant of the well-known Lin et al. (1983) scheme, with particle density and intercept values for the mixing ratio distributions of rain, snow, and graupel of (1000, 8 x 10^6), (100, 3 x 10^6), and (400, 4 x 10^8), respectively. The latter set of parameters is such that small graupel is favored over large hail (Gilmore et al. 2004), as is appropriate for MCS simulations—severe hail reports associated with mesoscale systems, when they occur, are most common in the earliest stages of development or with isolated cells near the leading convective line (Houze et al. 1990).

For modeling convective events, Bryan et al. (2003) propose that grid spacing on the order of 100 m should be used to reproduce small-scale turbulent eddies. They also find, however, that resolving the bulk properties of a convective cloud may be possible with 1 km grid spacing. Similarly, Weisman et al. (1997) found that 4 km may be sufficient to represent the system-scale properties of mid-latitude squall-line-type convection. To achieve a balance between ensemble size and the fidelity of the simulated storms, the model grid spacing for the 2D and 3D experiment is Δx = 1 km and Δz = 500 m (the vertical grid spacing is 250 m in the lowest 1250 m and then stretches to 700 m near the top of the model domain, giving an average grid spacing of about 500 m), with a 10 s time step, covering a domain 800 km in length and 20 km in height. For the 3D runs the length of the domain is reduced to 700 km and the width is set to 200 km. The grid spacing and domain size are consonant with other recent studies of MCS evolution (e.g., Coniglio et al. 2006) and mesoscale predictability (e.g., Hohenegger and Schär 2007a). Nevertheless, the sensitivity of the results to the horizontal resolution is explored by re-

running a subset of the simulations with horizontal grid spacing of Δx = 500 m and 250 m. In agreement with Bryan et al. (2003), the qualitative differences one would expect from higher resolution (e.g., narrower, stronger updrafts) can be seen in individual simulations, but the statistics of the sets of simulations are not substantially altered and the conclusions drawn are not affected.

Convection is initialized in the 2D simulations using an identical 3 K warm bubble located at x = 100 km and z = 1500 m, with a horizontal radius of 10 km and a vertical radius of 1.5 km. For the 3D simulations, the warm bubble is replaced with a thermal line of the same magnitude extending the entire width of the domain. The line is located at x = 50 km and includes random perturbations with a maximum amplitude of 0.1 K along the line to encourage 3D structures. The strength of the bubbles is sufficient to guarantee convective initiation, although convective initiation is at least as sensitive to perturbations in the environment of a potential storm as is storm maintenance (e.g., Crook 1996). Since the focus of the present study is limited to the latter, *we explore at the predictability of MCS forecasts* ***assuming*** *convective initiation occurs*. As such, the results presented herein form an upper bound on MCS predictability overall.

The simulations are each run out to produce an 8 h forecast. The control sounding onto which the perturbations are added is a relatively moist sounding with approximately 2600 J kg^{-1} of surface-based convective available potential energy (CAPE) (Fig.3). The sounding has little convective inhibition (7 J kg^{-1}), so deep lifting likely is not necessary to initiate new cells in the simulations. The relative humidity has been capped at 90% in the boundary layer, and 85% above the boundary layer, to avoid spurious wave activity in saturated unstable layers. The wind profile for the control run increases linearly from 0.0

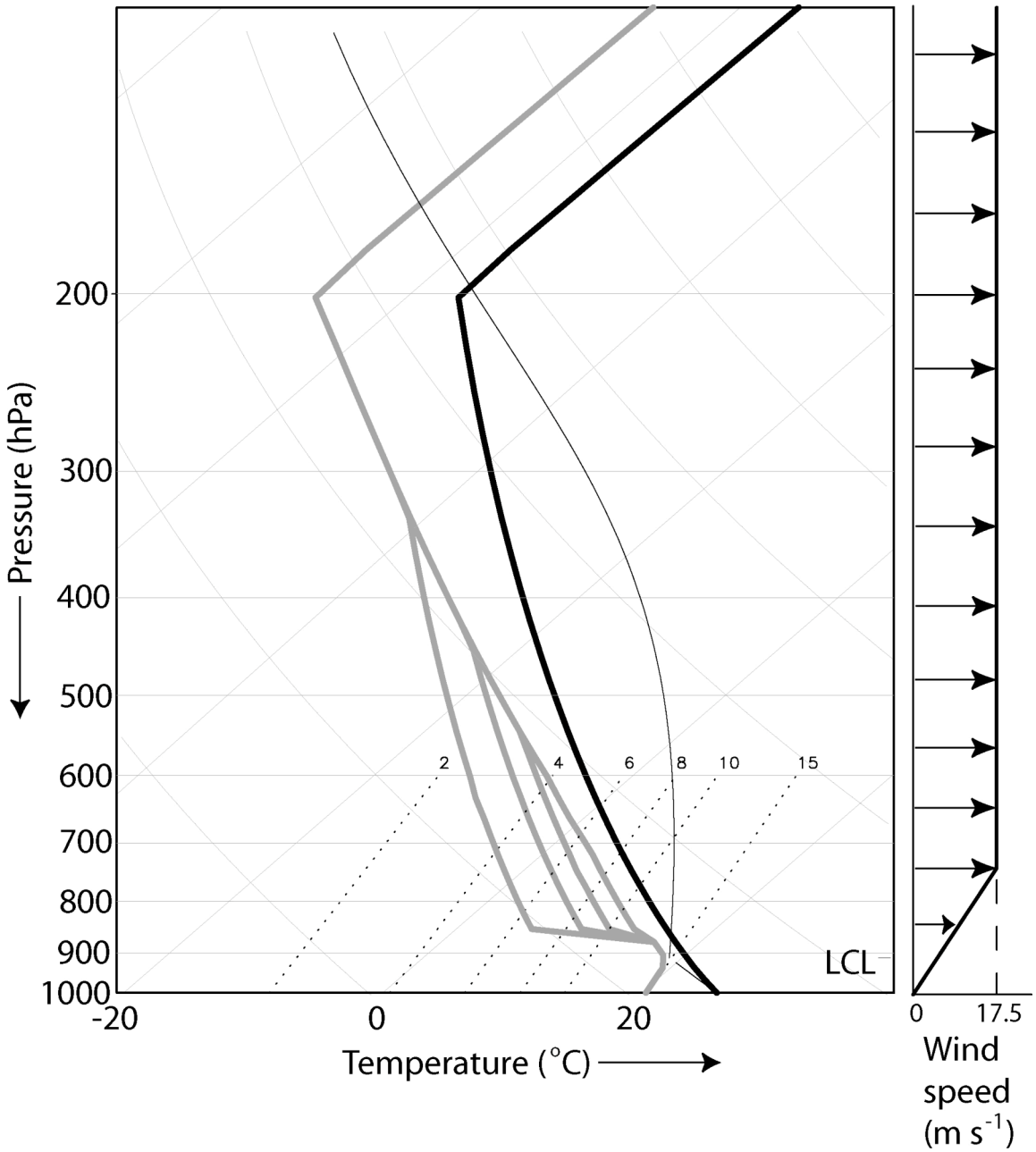

Fig.3. Control sounding for the ensembles: thick black for temperature, thin black for parcel ascent and gray for moisture profile. For the rightmost moisture profile, max(RH) = 85%. The additional moisture profiles represent control soundings for which the maximum relative humidity is set to 75, 65, and 50%.

m s^{-1} to 17.5 m s^{-1} at 2.5 km above the surface and no shear above that height (Fig. 3); this is the same wind profile as used by Rotunno et al. (1988).

b. Perturbation methodology

The most substantial difference between the SW04 simulations and the 2D results presented herein is that the former use only warm rain microphysics while the latter include a three-class ice microphysical parameterization. The addition of ice microphysics renders the simulations nearly independent of the perturbations used in SW04, likely due to the stronger downdrafts ice microphysics produces compared with warm rain only schemes (Srivastava 1987; Johnson et al. 1993). Based in part on the framework of Rotunno et al. (1988), SW04 include perturbations to CAPE (achieved by varying the tropopause temperature and the exponential used to derive the relative humidity) and wind speed for three different shear layer depths. This perturbation set appears to be insufficient for the present generation of numerical models with ice microphysics and thus a more complete description of uncertainty is required. The details of the perturbations are described further below. Instead of simply adjusting the exponential of the relative humidity profile, the profile is subjected to random perturbations in the same manner as the wind speed profile. This improved perturbation methodology produces mid-level dry layers similar to those often found in atmosphere and which can have a large impact on thunderstorm outflow (Gilmore and Wicker 1998).

Studies of the sensitivity of convection to initial perturbations (especially perturbations in the moisture field) often involve simulations of supercells and not MCSs, nevertheless, one would expect the results to generalize to some degree at least. The maximum updraft strength is shown to be particularly sensitive both to the surface

moisture (Crook 1996) and the moisture levels within the storm (Park and Droegemeier 2000). Moisture perturbations as small as 1% can produce substantial differences in accumulated rainfall within a few hours by altering the storm evolution and structure (Park 1999). On the other hand, Park and Droegemeier (2000) find the response to moisture perturbations to be not as strong in the precipitation fields as for the updraft strength. Strong midlevel dryness results in stronger initial modeled updrafts, but the storms then weaken as the stronger outflow pushes the low-level inflow away from midlevel mesocyclone (Gilmore and Wicker 1998). Stronger vertical wind shear can mitigate the effect of the midlevel dryness as the increased mixing leads to weaker outflow (Gilmore and Wicker 1998). Ducrocq et al. (2002) note a case in which an erroneously dry model atmosphere is unable to replicate an observed convective line. An inserted density current generates lifting along its leading edge, but the strength of the lift is not sufficient to initiate moist convection. Though not, strictly speaking, predictability or sensitivity studies, discriminant analyses connecting various environmental fields to MCS structure and evolution suggest a sensitivity to deep-layer shear and CAPE, among other variables (Coniglio et al. 2007; Cohen et al. 2007). Thus, one may expect the CAPE and relative humidity perturbations to affect the likelihood of producing a successful MCS forecast, while the contribution of the wind speed perturbations is somewhat more uncertain.

Perturbation sizes used in this study are based on forecast errors from the 12-km North American Model (NAM, formerly the Eta; Black 1994) from the National Centers for Environmental Prediction during May and June of 2006. The 20-km Rapid Update Cycle (RUC; Benjamin et al. 2004) model analyses are used as truth. The data from both

models are obtained from the NOAA National Operational Model Archive and Distribution System (NOMADS). The analyses from the NAM are bilinearly interpolated onto the RUC grid. The forecast errors are calculated within the region between 105° and 70° W and 25° and 49° N—the continental United States east of the Rocky Mountains. Finally, only errors from forecasts valid at 00Z are considered to focus on those errors one would expect to find around the peak hours of convective development. The spread of the 24-h NAM forecast errors thus provides the initial uncertainty for the developing MCS in the day 2 forecast period.

For wind speed, the standard deviation of the 24-h forecast errors (computed in space and time) increases slightly from 1000 hPa up to about 800 hPa, remains nearly constant up to near 400 hPa, and then increases to a peak value near 250 hPa (Fig. 4a). The 24-h relative humidity forecast error standard deviation increases from 1000 hPa to 800 hPa (by about 8%) and decreases above 400 hPa (Fig. 5a). In between, there is a (roughly 1%) peak just below 500 hPa with nearly equal values above and below this peak. For both variables, the error values used represent the nearly constant mid-level values (Figs. 4b, 5b). The distribution of the 24-h forecast errors for wind speed and relative humidity at 600 hPa are slightly peaked and right-skewed (Fig. 6a,b). However, these deviations from Gaussianity are sufficiently small that the standard deviation of the data is a reasonable measure of the spread of the forecast errors.

The increase in forecast uncertainty between 12 h and 24 h is minimal (Table 1). For example, the standard deviation of the wind speed errors only increases from 2.9 m s^{-1} for a 12-h forecast to roughly 3.1 m s^{-1} for a 24-h forecast. The forecast uncertainty for relative humidity increases meagerly from 20% at 12 h to 21% at 24 h. For the final error

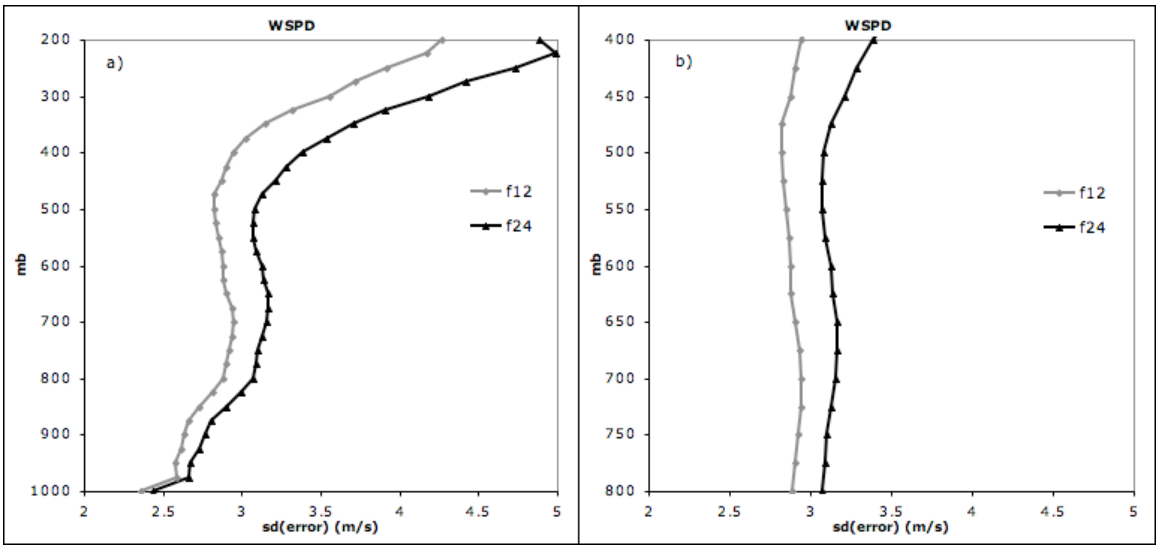

Fig. 4. Profile of the standard deviation of 24-h forecast errors of wind speed from the NAM compared against RUC analyses over the central United States for the period of May-June 2006 for (a) the full column from 1000 to 200 mb and (b) the mid-levels between 800 and 400 mb.

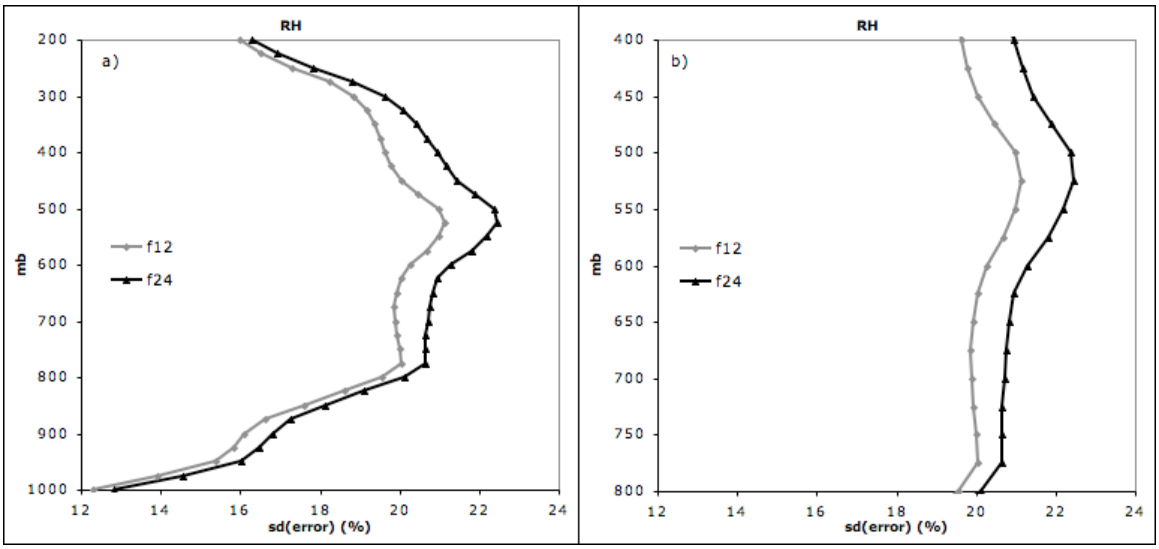

Fig. 5. Same as for Fig. 4, except for relative humidity.

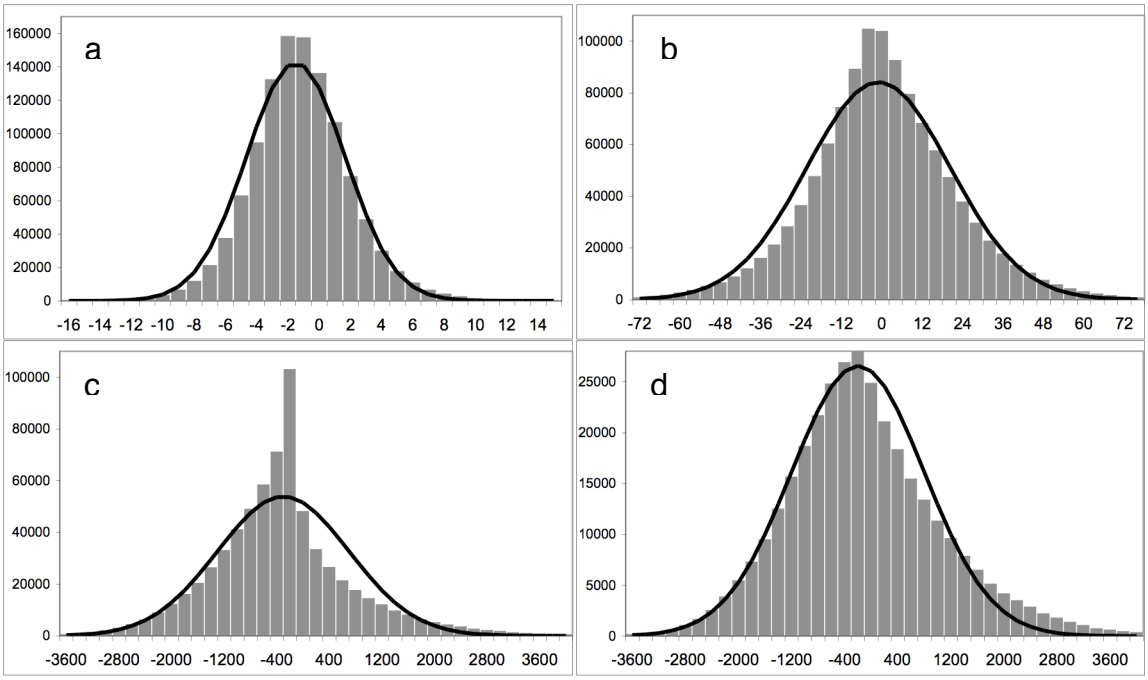

Fig. 6. Histograms of 24-h NAM forecast errors against RUC analyses at 600 hPa over May and June of 2006 for (a) wind speed, (b) relative humidity, (c) CAPE for all locations with positive CAPE, and (d) CAPE for all location with CAPE > 500 J kg^{-1}.

Table 1. The standard deviation of forecast errors valid at 0000 UTC from the NAM for 12- and 24-h lead times. Values represent averages throughout the mid-levels (800-400 hPa for relative humidity and 800-500 hPa for wind speed) where little dependence on pressure is found. The 1000 J kg^{-1} value is the 24-h forecast error standard deviation when only locations with CAPE > 500 J kg^{-1} are considered. Also, included are RUC analysis error estimates from Thompson et al. (2003; *) and Benjamin et al. (2004; **).

Forecast hour	CAPE (J kg^{-1})	Relative Humidity (%)	Wind Speed (m s^{-1})
00 h	300-500*	9**	1-3*,**
12 h	750	20	2.9
24 h	800/1000	21	3.1

field, two separate 24-h CAPE error standard deviations are calculated. To avoid having the CAPE errors dominated by zero CAPE values in stable regions, only locations with non-zero CAPE are evaluated, giving an error standard deviation of ~800 J kg^{-1}. Further restricting the data by requiring CAPE values to exceed 500 J kg^{-1}—since MCS development does not often occur in nearly neutral environments—increases the standard deviation of the forecast error to over 1000 J kg^{-1}. When the all locations with non-zero CAPE are considered the peakedness (or leptokurtosis) and skewness of the distribution of forecast errors is more pronounced (Fig. 6c); there is a strong tendency toward small, negative errors together with a relatively large number of big, positive errors. Removing the small CAPE locations restores to the distribution to one much closer to Gaussian (Fig 6d). Perturbations consistent with these error magnitudes given storm initiation at a forecast lead time of 24-h are used (Table 2). An additional set of perturbations is constructed using half the size of the 24-h errors instead of those consistent with the 12-h errors due to the relatively small differences between the 12- and 24-h forecast error magnitudes. Halving the 24-h forecast errors reduces the perturbations to a level roughly equal to typical analysis error (Table 1; Thompson et al. 2003; Benjamin et al. 2004). It should be noted that the errors in the RUC jump markedly between the analysis and the 1-h forecast, with little further change through 12 h (Benjamin et al. 2004). Thus, the reduced perturbations do not represent gradual forecast improvement (e.g., tomorrow's 24-h forecast being as good as today's 12-h forecasts), but a radical advance in numerical modeling.

The temperature profile for the control run is determined by the temperature at the surface and at the tropopause as well as an exponential defining the shape of the profile.

Table 2. The perturbations comprising the ensemble configurations. The ensembles are referred to by the leading digit of the perturbation sizes, i.e., rCRW, where C stand for CAPE, R stand for relative humidity, and W stands for wind speed. Hence, r823 denotes the ensemble based on current 24-h forecast errors.

Ensemble name	CAPE ($J\ kg^{-1}$)	Relative Humidity (%)	Wind Speed ($m\ s^{-1}$)
r1023	1000	20	3.1
r1021	1000	20	1.6
r1013	1000	10	3.1
r1011	1000	10	1.6
r823	800	20	3.1
r821	800	20	1.6
r813	800	10	3.1
r811	800	10	1.6
r523	500	20	3.1
r521	500	20	1.6
r513	500	10	3.1
r511	500	10	1.6

Therefore, the desired variance in CAPE is achieved by randomly perturbing the tropopause temperature (e.g., a 10 K change in tropopause temperature yields a 1000 J kg^{-1} change in CAPE). For the wind and relative humidity perturbation profiles, random perturbations are drawn from a Gaussian distribution with zero mean and a standard deviation matching the errors described in Table 1 and assigned to the vertical model levels every 2500 m starting at the surface. A cubic spline is used to obtain values for intermediate levels. This method ensures that the wind and relative humidity profiles are vertically coherent, so that, for example, realistic dry layers can be introduced into the initial soundings. While both positive and negative perturbations are allowed for the moisture profile, the negative perturbations dominate; the control sounding is already moist and positive perturbation sizes must be limited to prevent saturation. No changes are made to the moisture profile within the boundary layer so that the CAPE is affected by the temperature perturbations alone. Figure 7 shows sample perturbation profiles for temperature and wind speed, along with the relative humidity profiles. For each perturbation field, a single set of 100 perturbations is used and the different perturbation magnitudes are achieved by applying a multiplicative factor. Therefore, any differences between the different ensembles result from the perturbation magnitudes and not from a different sample of perturbations (e.g., more negative and fewer positive perturbations).

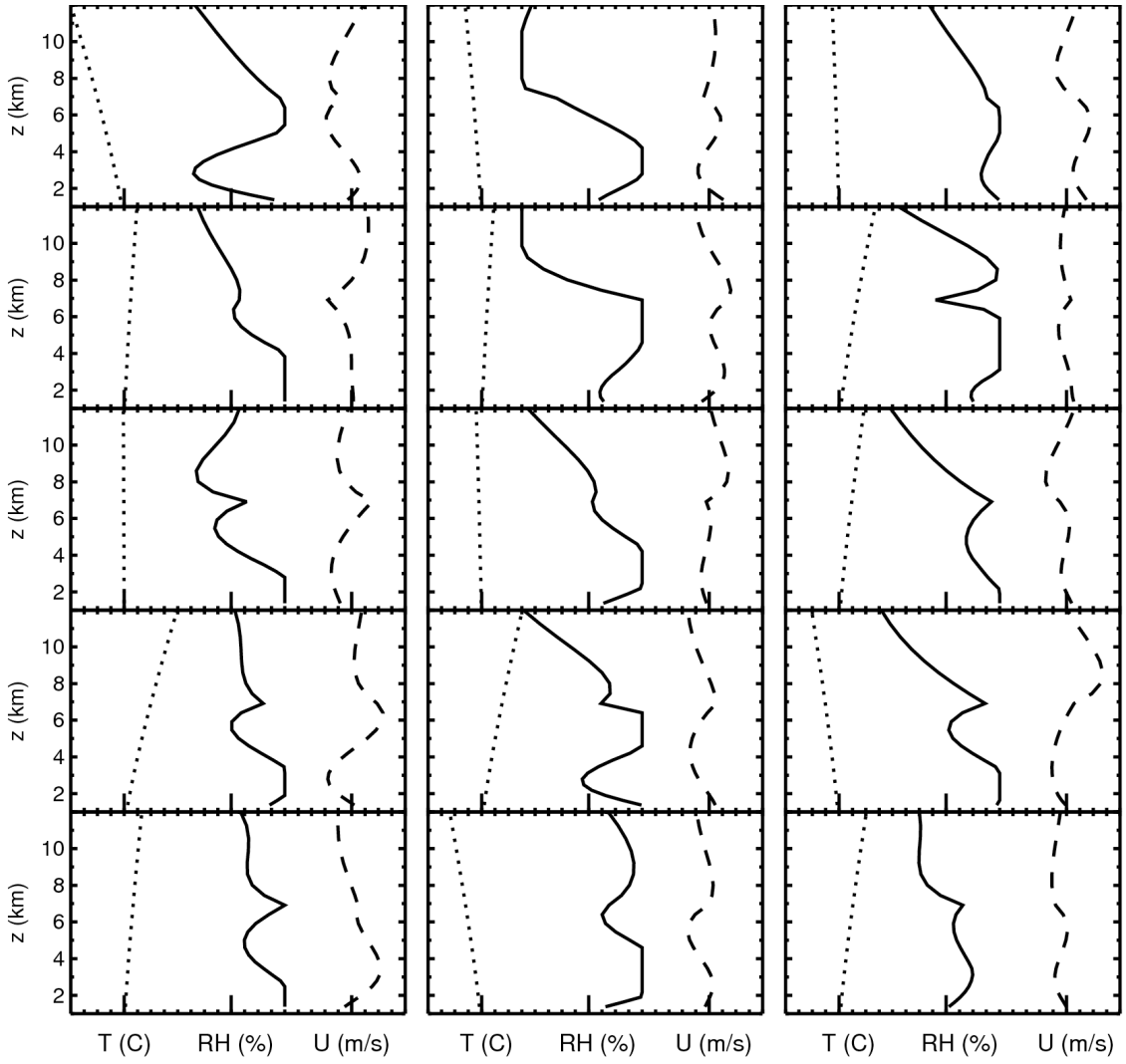

Fig. 7. Profiles of 15 perturbation structures consistent with the 24-h forecast errors, showing perturbation temperature (dotted), relative humidity (solid), and perturbation wind speed (dashed). Tick marks are every 1.75 °C, 10%, and 1 m s^{-1}. The large tick marks represent, from left to right, 0 °C, 50%, and 0 m s^{-1}. Relative humidity profiles are shown only above the 1-km deep boundary layer.

Chapter 4: Results from the 2D experiment

The base-state sounding described in the previous section defines the homogeneous environment for the control run (Fig. 3). The evolution of the MCS in this control run (Fig. 8) shows that the system steadily grows and moves across the domain at about 18 m s^{-1} (or 33 km per 30 minutes, slightly more than one tick mark each frame). The system at first leans upshear and appears to develop a leading stratiform region. However, it transitions to a system with trailing stratiform precipitation after 4 h.

Isolines of precipitation mixing ratios, at 8-h run time for 20 members of the ensemble indicate a variety of outcomes (Fig. 9), where the precipitation mixing ratio is defined as the sum of rain, snow and graupel mixing ratios. These perturbations are based on the 24-h forecast error estimates (r823, i.e., the perturbations have a standard deviation of 800 J kg^{-1}, 20%, and 3.1 m s^{-1} for CAPE, relative humidity, and wind speed, respectively; see Table 2). Substantial variability exists among the members ranging from no storm to small weak cells to a large, robust single cell to multicellular to a leading convective line with a trailing stratiform region extending over 150 km. The location of the leading edge of convection varies by up to 160 km (i.e., system propagation speed varies by 5-6 m s^{-1}).

a. Success rate

The measure of predictability proposed by SW04 is the MCS success mode rate, where an individual realization is said to have successfully simulated an MCS if there exists a continuous region of precipitation mixing ratio $\geq$ 0.2 g kg^{-1} between 3 and 5 km above the surface that is at least 20 km in length located within 100 km on either side of

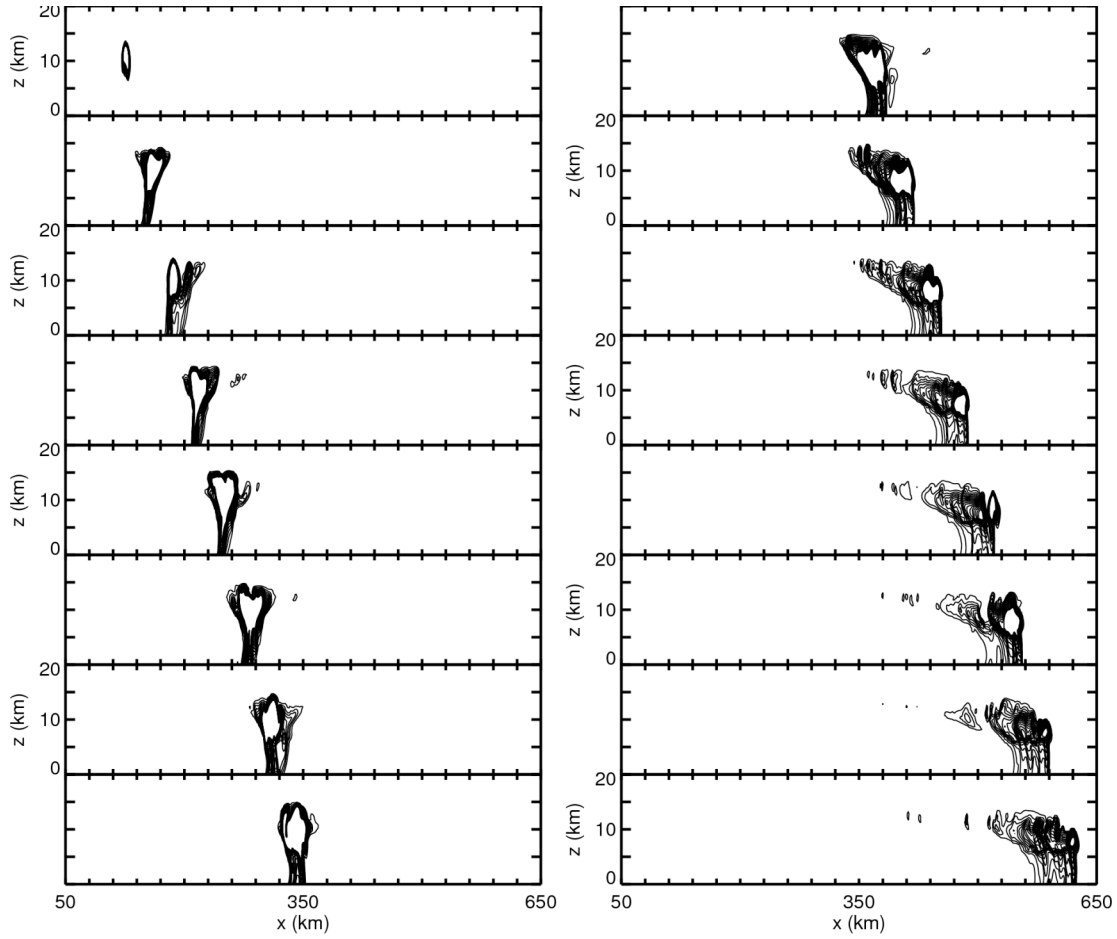

Fig. 8. Snapshots of the evolution of the control run every 30 minutes starting at 30 minutes into the simulation. Isolines depict precipitation greater than 0.2 g kg^{-1}. The region of the domain plotted extends from 50 km to 650 km. Tick marks are every 30 km along the horizontal axis and 5 km along the vertical axis.

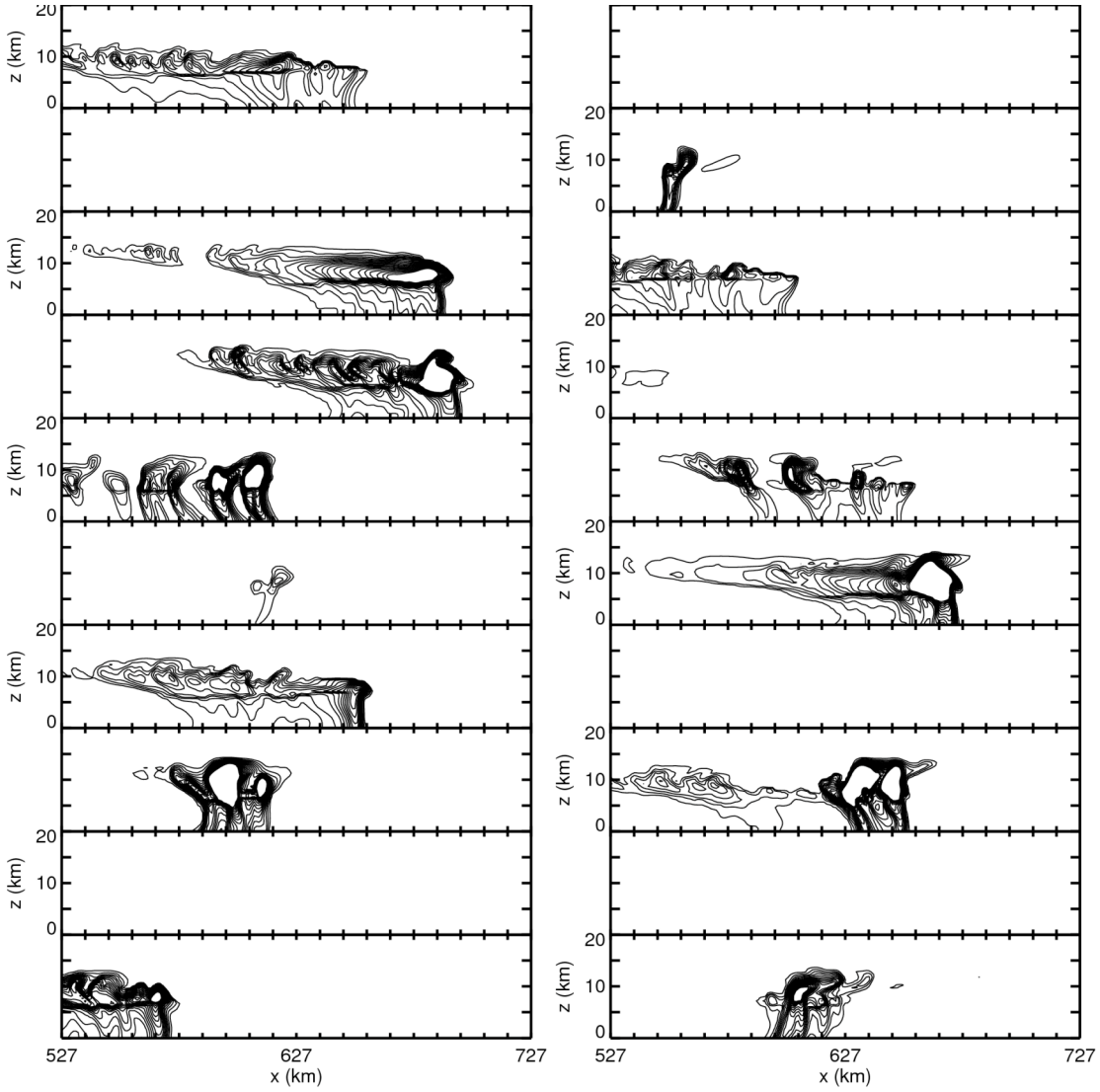

Fig. 9. Isolines of precipitation greater than 0.2 g kg^{-1} for 20 members of the r823 ensemble at 8 h (see text for the ensemble naming convention). The grid extends to 100 km on either side of the location of the control run at this time. Tick marks are located every 10 km along the horizontal axis and 5 km along the vertical axis.

the storm in the control run and the location of the maximum updraft defines the location of the storm in the control run. As seen in Fig. 9, this condition can result from a succession of convective cells or from the prototypical leading edge of convection following by a trailing stratiform region. The success mode definition is intended as a generous measure of forecast success: the model need only produce a contiguous region of precipitation resembling a minimum-size MCS within a broad window around the control run.

The MCS success mode rate possesses a modest sensitivity to the length threshold (Fig. 10). For the 20 km threshold, the success rate curve reaches an asymptotic value of around 70% about four hours into the simulation. The convective leading edge of observed MCSs is typically 10 to 50 km across (e.g., Smith et al. 2008). Increasing the minimum length threshold to 60 km decreases the asymptotic value to around 55% and increases the time required to reach this value to seven hours. Despite this reduction in the final success mode rate, each ensemble configuration appears to be affected similarly (not shown), and so the size criterion of 20 km is adopted herein. As mentioned above, the system movement varies by 5-6 m s^{-1}. Thus, using a narrower window for identifying successful forecasts results in a curve identical to that for the 200-km window until storms begin moving out either side of the window and the success rate declines (e.g., the success rate at 8 h drops from about 70% to just under 60% for a 100-km window; not shown). Since each ensemble consists of 100 members, the success mode rate is simply the numbers of members that meet the MCS criteria.

Time series of the MCS success mode rate for 12 ensemble configurations out to 8 h, plotted at 15-min intervals, is shown in Fig. 11. Over half of the runs produce an MCS

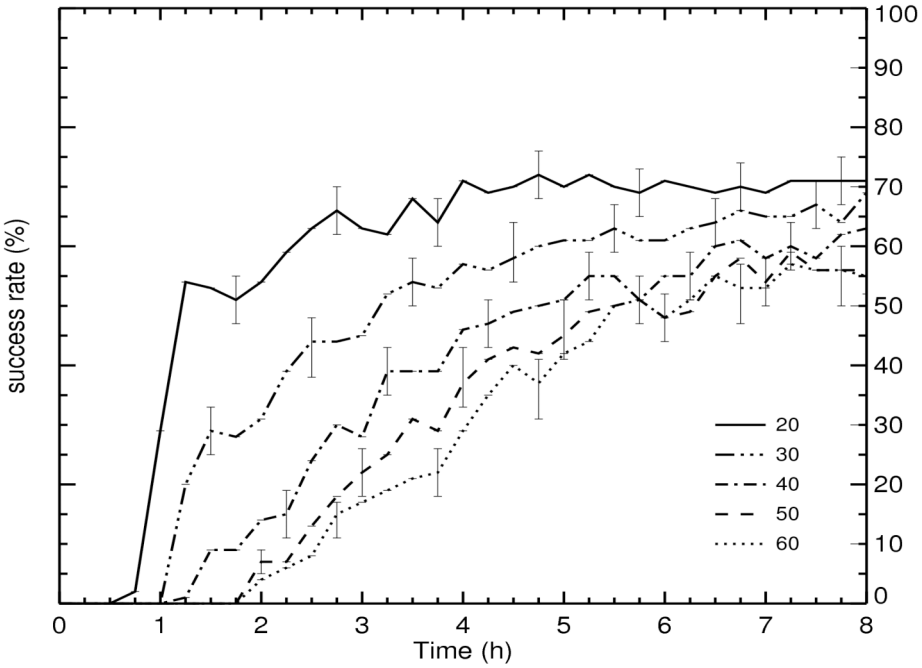

Fig. 10. MCS success mode (%)—the percentage of ensemble members meeting the MCS definition—for the r823 ensemble as a function of model run time for minimum areal extent thresholds ranging from 20 to 60 km. Error bars mark the central 50% credible interval, i.e., from the 25th to 75th percentiles.

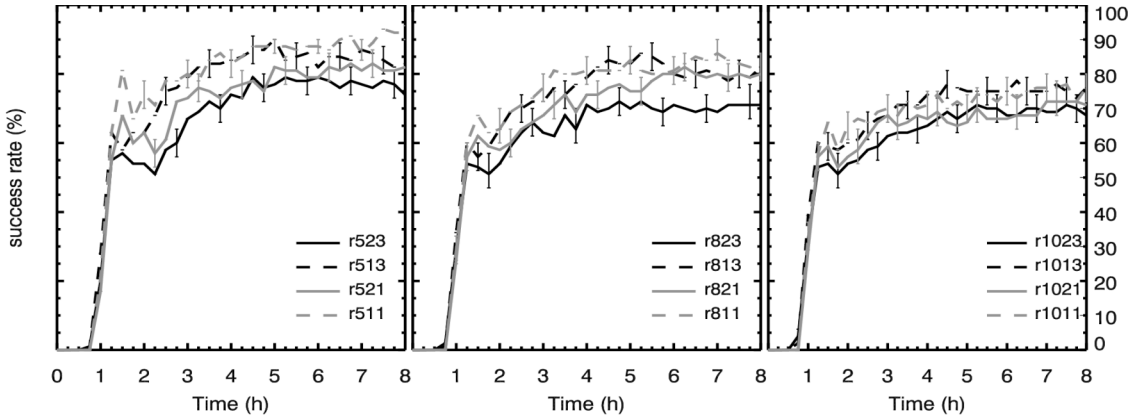

Fig. 11. As in Fig. 10, but for different ensemble configurations using the 20 km threshold. The ensemble configurations plotted are identified in each panel. See Table 2 for naming convention. Each panel shows the success rate for a given CAPE perturbation magnitude for different wind speed and relative humidity perturbation sizes. Dashed lines represent ensembles for which the relative humidity perturbations are reduced by half, while the gray lines represent the ensembles for which the wind speed perturbations are reduced by half.

within the first hour, with more of the runs meeting the MCS criteria as time increases until asymptoting to a constant success rate between 70 and 90%. Each panel shows the success rate for a given CAPE perturbation magnitude for different wind speed and relative humidity perturbation sizes. For the 24-h wind speed and relative humidity errors, reducing the CAPE uncertainty yields only marginal improvement: the ensembles using the 1000 and 800 J kg^{-1} CAPE errors (r1023 vs. r823) are nearly identical with about 70% of the members eventually producing an MCS, while the r523 ensemble approaches 80% before falling back to near 75%. However, if both the wind speed and relative humidity perturbations are reduced by half, the MCS success mode rate rises from 75 to 80 to 90% for r1011, r811, and r511, respectively. This could be viewed as an added sensitivity to CAPE uncertainty for the smaller wind speed and relative humidity perturbations or alternatively as an added sensitivity to wind speed and relative humidity uncertainty as the CAPE perturbations are reduced.

The success rate is consistently more sensitive to a reduction in the relative humidity perturbations than to a reduction in the wind speed perturbations as the success rate curves tend to group according to the size of the relative humidity perturbation (dashed curves vs. solid curves), though for the 1000 J kg^{-1} CAPE errors the two groups are visually distinct but statistically indistinguishable. The significant exception to this is the r821 ensemble, which closely tracks with the r823 ensemble for the first 2-3 h, but whereas the r823 curve then levels off at around 70% success rate, the r821 curve continues upward to 80%, matching the success rate of the r813 and r811 ensembles. There is also some hint of a separation during the final 2 h between the r521 and r523 curves and the r511 and r513 curves, with the smaller wind speed perturbations leading to

a higher success rate, but the improvement does not equal that gained by reducing the relative humidity perturbations.

In summary, an ensemble that has perturbation magnitudes consistent with estimates of current 24-h operational forecast errors of environmental conditions successfully produces an MCS in about 70% of the runs. Reducing those 24-h forecast errors by half increases the success mode rate to 90%. This means that achieving 90% success in forecasts of MCS development on a day 2 forecast, even leaving aside the problem of convective initiation, requires that errors in model forecasts of environmental conditions be reduced to the level of current analysis errors.

b. MCS size

The size of the MCSs (taken as the maximum extent of the continuous region with precipitation greater than 0.2 g kg^{-1} at any level between 3 and 5 km above the surface) is not greatly affected by the perturbations magnitudes (Fig. 12). The MCS growth slows initially for the reduced relative humidity and wind speed perturbations, especially for the 500 J kg^{-1} ensembles, but by 8 h those differences have vanished. For the 24-h forecast wind and humidity errors, the variability in the MCS size for the different CAPE ensembles (r523 vs. r823 vs. r1023) is nearly indistinguishable—the standard deviation rises to 40-45 km by 8 h. Even when both the wind and humidity perturbations are reduced, adjusting the size of the CAPE perturbations results in only about a 10% change in the uncertainty of the forecasted MCS size, and that change is mostly a result of the uncertainty increasing for the 1000 J kg^{-1} ensemble. To the extent that the wind speed

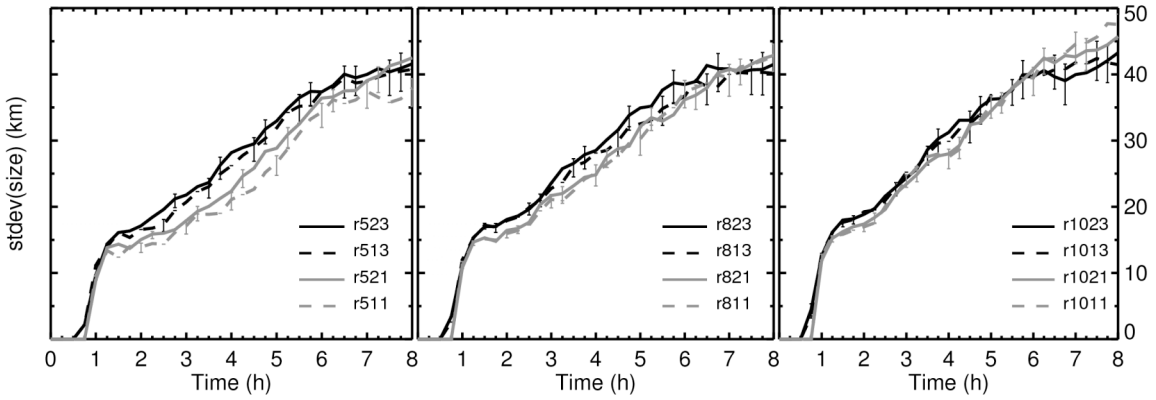

Fig. 12. As in Fig. 11, but for standard deviation of MCS size (km).

and relative humidity perturbations have an impact, the wind speed uncertainty is the more important of the two (compare black curves with gray curves).

c. Maximum updraft strength

The effect of perturbation magnitude on maximum updraft strength primarily occurs early in the simulation during the time when updrafts are the strongest, between 1.5 and 3 h into the simulation. The mean updraft strength during this stage is around 15 m s^{-1} with the strongest updrafts exceeding 45 m s^{-1}; the updrafts steadily weaken through the rest of the simulation. It should be noted that the ensemble distributions of maximum updraft strength are not Gaussian, but rather possess some skewness and are platykurtotic (i.e., they have a flat peak; Fig. 13). However, the distributions are close enough to Gaussian to allow use of the standard deviation as a measure of the ensemble spread. Reducing all perturbations by half decreases the uncertainty in the maximum updraft strength by 20% (Fig. 14). This improvement appears to come mostly from a reduction in the CAPE forecast errors, with a reduction to 500 J kg^{-1} improving the uncertainty by between 1.5 and 3 m s^{-1} (r1023 vs. r523). Interestingly, reducing the wind and relative humidity errors along with the CAPE reduction gives very little additional improvement to the updraft intensity uncertainty (e.g., r511 vs. r523), but halving the wind and relative humidity errors without also reducing the CAPE errors actually yields slightly greater uncertainty (e.g., r1011 vs. r1023). Since the distribution of maximum updraft strength is skewed—by definition, there cannot be an updraft with negative vertical velocity—changes to the standard deviation likely are due to increases in the maximum updraft. The impact of larger wind speed and relative humidity perturbations (e.g., greater dry air

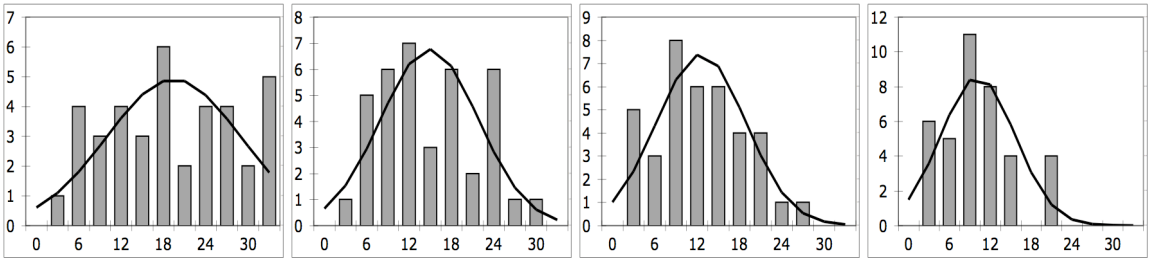

Fig. 13. Histograms for the maximum updraft from the r823 ensemble at t = 2, 4, 6, and 8 h.

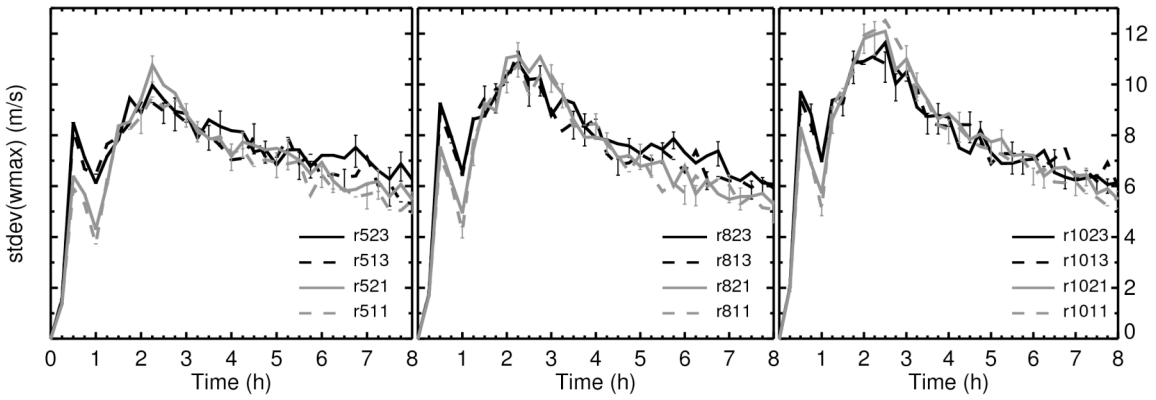

Fig. 14. As in Fig. 11, but for standard deviation of maximum updraft strength (m s^{-1}). perturbation ensembles show some sensitivity (~1-2 m s^{-1}) to the wind speed and relative humidity perturbations during the final 2-3 h of the simulations.

intrusion into the updraft) perhaps acts to constrain the potential impact of the large CAPE perturbations. While the 1000 J kg^{-1} ensembles are more sensitive to changes in the other perturbation magnitudes at the early peak updraft stage, the smaller CAPE perturbation ensembles show some sensitivity (~1-2 m s^{-1}) to the wind speed and relative humidity perturbations during the final 2-3 h of the simulations.

d. Maximum surface wind

Whereas the mean updraft strength steadily weakens after an early peak, the mean surface wind (actually lowest model level) climbs rapidly over the first 1.5 h and then holds nearly constant, at around 35 m s^{-1} for the last 6 h (not shown), with very little difference among the different ensembles. The maximum surface winds peak at over 60 m s^{-1} for the r1023 ensemble, but are mostly in the 45-50 m s^{-1} range. Histograms of the maximum surface wind (not shown) indicate that the standard deviation is again a suitable measure of ensemble spread. The separation among the different CAPE perturbations is more pronounced for this field (Fig. 15) than for the features previously examined, particularly for the smaller humidity and wind perturbations where the variability of the 500 J kg^{-1} ensemble is reduced substantially while the larger CAPE-perturbation ensembles are basically unaffected. Interestingly, the ensembles are more sensitive to a reduction in the relative humidity perturbations than for the wind speed (compare dashed vs. solid lines with black vs. gray lines). This suggests that the strength of the winds to be transported down to the surface is secondary to the presence of a dry layer capable of supporting downdrafts necessary for this downward transport. However, adding the wind-perturbation reduction to a humidity-perturbation reduction further

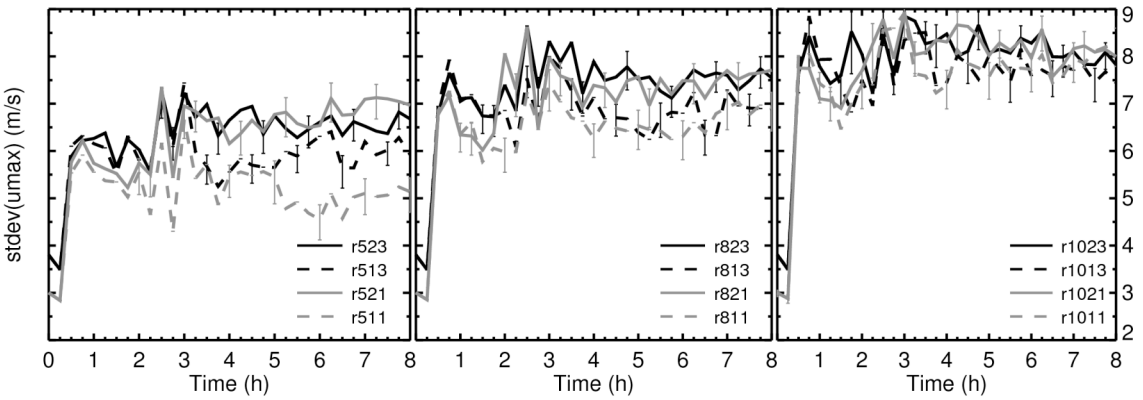

Fig. 15. As in Fig. 11, but for standard deviation of maximum surface wind (m s^{-1}).

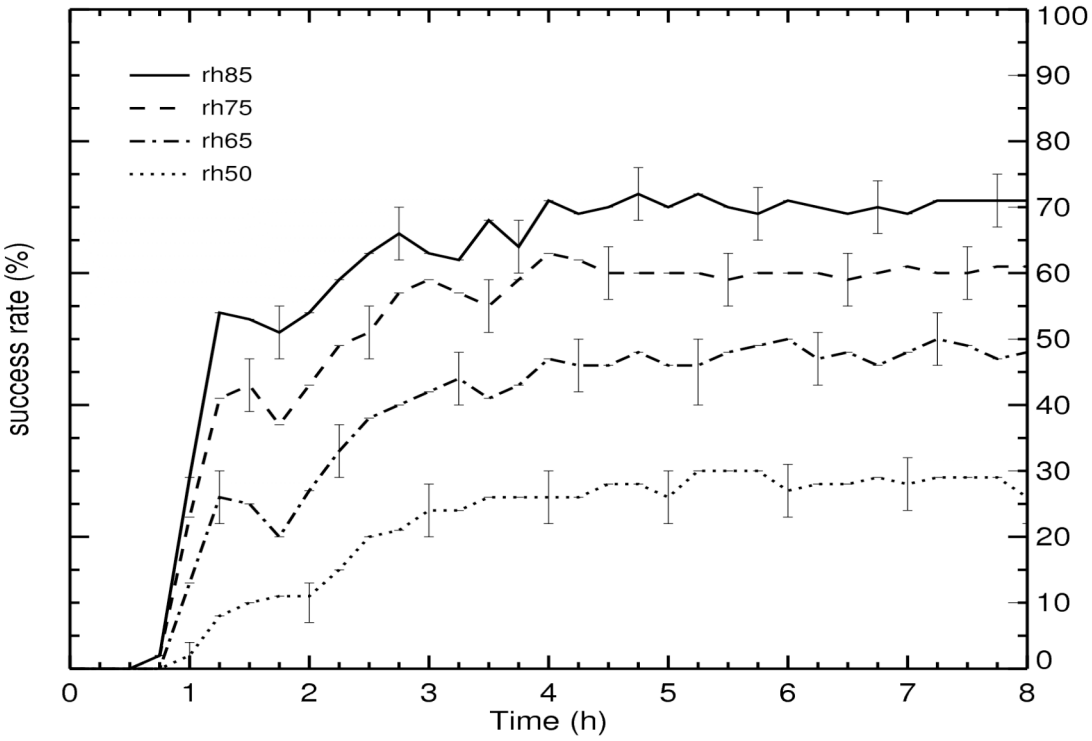

Fig. 16. As in Fig. 10, but the r823 ensemble for control soundings with different values of maximum relative humidity above the boundary layer, as shown in Fig. 3.

decreases the uncertainty for the 500 J kg^{-1} ensemble, such that the uncertainty for the r511 ensemble is 30-40% lower than for the r1023 ensemble.

e. Effect of deeper shear layer

When the sheer layer depth is increased from 2.5 km to 7.5 km, the success mode behavior is less dependent upon the CAPE perturbation magnitude; each ensemble group achieves about an 80% success mode rate. Deepening the sheer layer depth also reducesthe uncertainty of the size of the MCS by 10-15 km (a reduction of about one third). The effect of the deeper shear on the maximum updraft strength and surface winds is more modest. Specifically, the mean values are reduced by 1-3 m s^{-1} while the uncertainty is reduced by about same amount early in the simulations and is only marginally increased toward the end of the runs.

f. Effect of drier control sounding

As mentioned in Section 3a, the control sounding is fairly moist throughout the column (Fig. 3). One reason such a moist profile is used is to limit the number of ensemble members that fail to initiate convection using a warm bubble initialization. However, to examine the impact of the relative humidity of the control sounding, the r823 ensemble is rerun using successively drier control soundings in which the maximum relative humidity above the boundary layer is reduced from 85% to 75, 65, and 50% (Fig. 3). The success mode rate for these four ensembles shows that the cloud-layer relative humidity in the control sounding has a large impact (Fig. 16). Each 10% reduction in the relative humidity results in a reduction of the success mode rate of 10-20 %, dropping below half of all runs as the cloud-layer relative humidity is reduced below 65% and

down to a less than 30% success mode rate for the 50% maximum relative humidity sounding. A large portion of the additional MCS failures are from members in which storms are unable to initiate. In fact, for the driest profile, even the control run (not shown) cannot maintain a storm past two hours into the simulation (for this set, the location of the original control is used to define the success rate). Furthermore, for the dry control soundings, the use of small CAPE perturbations actually reduces the success mode rate (not shown), indicating that the higher CAPE values are needed to offset the effects of a drier cloud layer. MCS development can only occur with perturbations that lead to a substantial moistening of the column or substantial increases in instability, as stronger updrafts are needed to avoid entraining too much dry air, choking off the storm. The drier control soundings have little effect on the uncertainty of the maximum updraft strength or surface wind, except for the driest sounding, for which the variability declines by 3-4 m s^{-1} and 1-2 m s^{-1}, respectively (not shown).

g. Summary of 2D results

With current 24-h forecast errors, the ensemble suggests that a day 2 MCS forecast should have a success rate around 70% (using a generous criterion for identification as an MCS), or slightly higher if there is a deep shear layer present. This success rate drops below 50% as the cloud-layer relative humidity is reduced below 65%. Improvements to the ensembles resulting from reductions in forecast errors are not uniform. For example, halving the wind speed errors increases the success mode rate for both the 500 and 800 J kg^{-1} ensembles, but it only has a minor effect on the 1000 J kg^{-1} ensemble. Reducing the errors by half for each of the fields examined here (CAPE, wind

speed, and relative humidity) raises the MCS success rate to 90%. That is, achieving 90% confidence in forecasts of MCS development requires the tremendous advance of bringing 24-h forecast errors down to the level current analysis errors. This radical improvement in forecasting would yield a 10-20% reduction in the uncertainty for maximum updraft strength and a 25-35% reduction in the uncertainty for maximum surface wind.

The uncertainty in the success rate, maximum updraft strength, and maximum surface wind are all more sensitive to changes in the relative humidity perturbations than changes in the wind speed perturbations, while the opposite is true for the MCS size. The impact of the relative humidity and wind speed perturbations is greater when the CAPE perturbations are smaller for the success rate and maximum surface wind, but is greater when the CAPE perturbations are larger for the MCS size. For the maximum updraft strength the impact of the wind speed and relative humidity perturbations is greater for the large CAPE perturbations early in the runs, while the opposite is true toward the end of the simulations. The CAPE perturbations themselves have the biggest impact for the maximum surface wind, but have no impact on uncertainty in MCS size and only moderate impact on the success rate. Thus, it does not seem to be the case that improvements in the forecast of a particular field (e.g., CAPE vs. wind speed) increase confidence in MCS forecasts, generally.

Chapter 5: Results from the 3D experiment

As stated earlier, the configuration of the 3D simulations is identical to that for the 2D runs with the exceptions that the domain is extended 200 km in the *y* direction (with periodic boundaries), the domain is shortened in the *x* direction to 700 km (from 800 km) to reduce computational and storage requirements, and the warm bubble is replaced by a thermal line extending the full width of the domain with small random perturbations added to facilitate 3D structure. Because of the computational demands of an additional spatial dimension, the size of the ensembles is reduced from 100 to 50 members. Furthermore, only two CAPE perturbation sets have been retained, with the 800 J kg^{-1} standard deviation perturbations representing the 24-h forecast errors. MCS development in the 3D runs is somewhat more robust than in 2D, resulting in a substantially higher MCS success rate for the r823 ensemble. Thus to allow for a greater influence of the positive perturbations in the moisture field, the base-state RH profile is reduced from 85% to 75%. Use of the drier control sounding matches the success rate of the r823 ensemble in 3D to that from the 2D experiment, facilitating comparisons between the 2D and 3D runs of the impact of changes in the initial uncertainty.

The resulting control run (averaged in the along-line direction; Fig. 17) is very similar to the 2D control run (Fig. 8), with two notable exceptions: the 3D storm moves somewhat slower (~17 m s^{-1} compared with ~20 m s^{-1}), and the trailing stratiform region is almost non-existent in 3D. A shift in the perspective to a horizontal slice (depicting a 200 x 200 km window moving with the storm; Fig. 18), taken at z = 4 km, reveals a more complex scenario despite the horizontally homogeneous initial state. The maximum

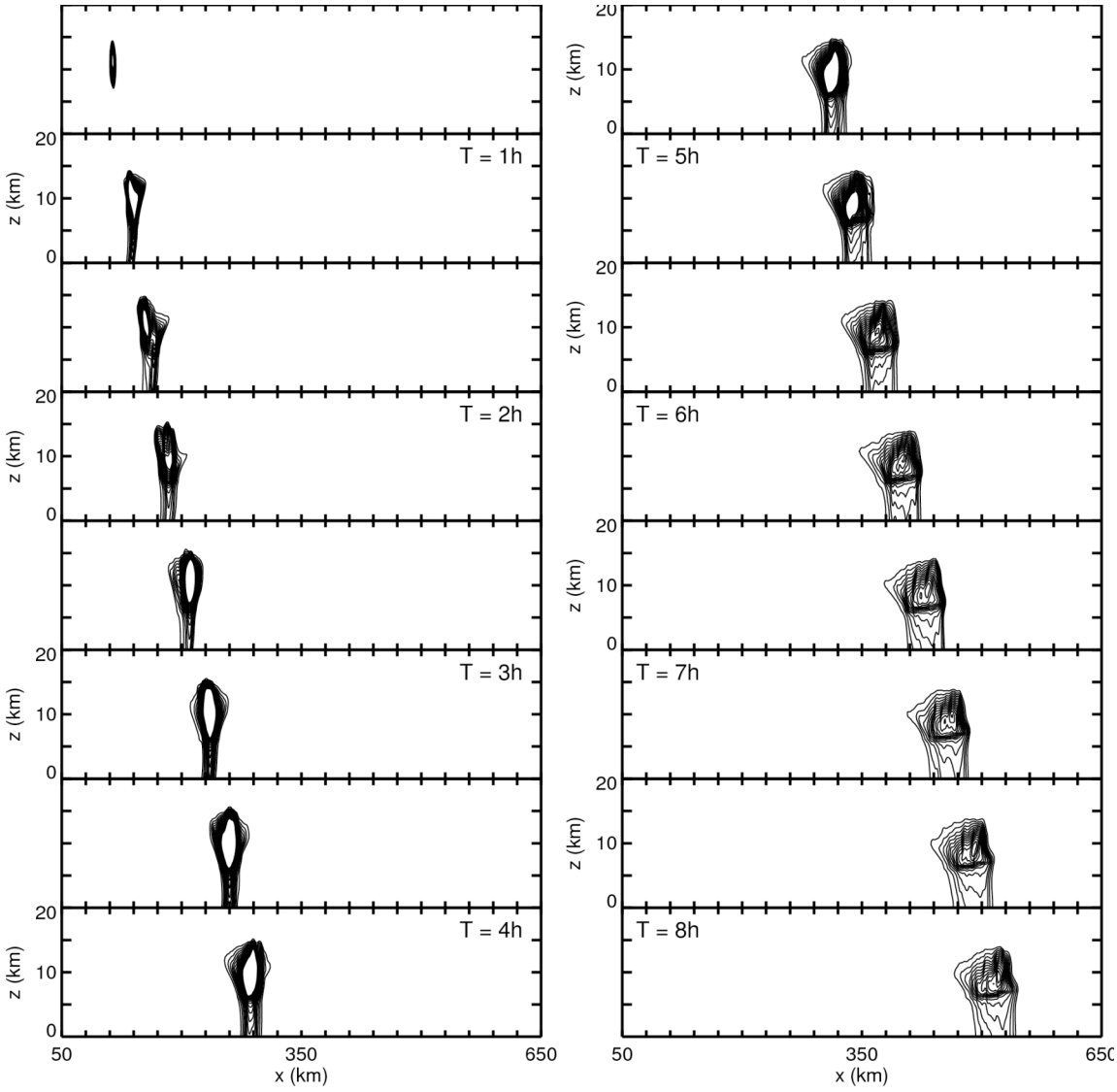

Fig. 17. As in Fig. 8, but for 3D control, averaged in the y-direction.

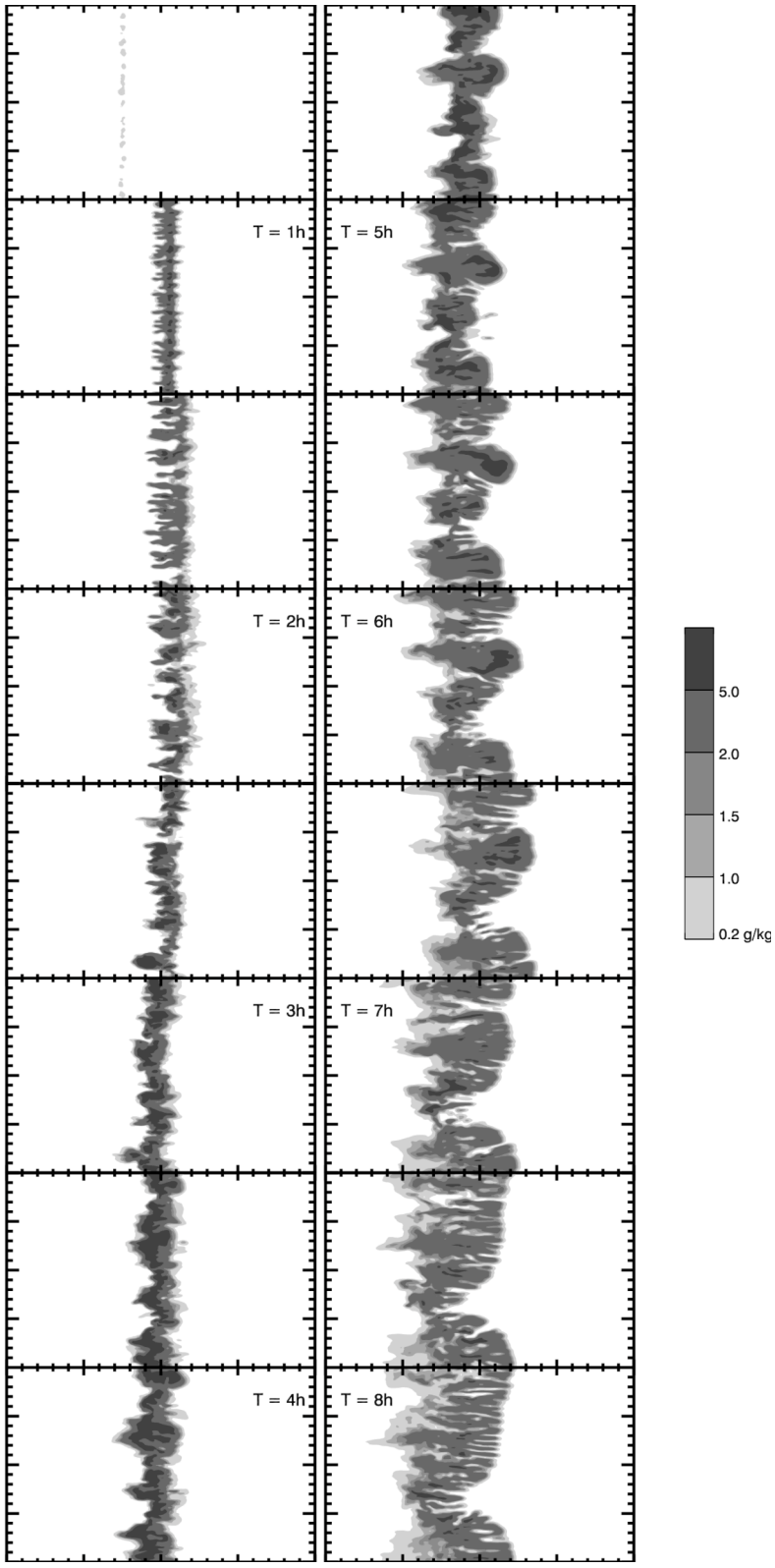

Fig. 18. As in Fig. 17, but for precipitation mixing ratio greater than [0.2, 1, 1.5, 2, 5] g kg^{-1} at z = 4 km in a 200 x 200 km window moving with the MCS.

width of the system 8 h into the simulation is nearly double that indicated in the line-averaged cross-section (Fig. 17). The fibrous nature of the MCS, particularly apparent in the final frame largely is an artifact of the grid spacing and the chosen low-level shear. Reducing the grid spacing leads to smaller and weaker updrafts, with correspondingly smaller gaps between cells, while increasing the strength of the low-level shear produces a more filled-in line (G. Bryan, personal communication).

As with the 2D ensembles, introducing perturbations equal to the 24-h forecast error, the r823 ensemble, yields a wide variety of solutions (Fig. 19), ranging from strong, narrow lines (e.g., member 01—number indicated in the lower right corner of each panel) to more broad, diffuse systems (e.g., member 16) to more isolated cells (e.g., member 42) to solutions completely unable to maintain convection (e.g., member 07). Some of the solutions (e.g., member 13) have the appearance of a bowing MCS, while others (e.g. member 4) are distinctly linear. Average system movement ranges from ~13 m s^{-1} to nearly 24 m s^{-1}. Reducing the uncertainty still allows for a range of structures (Fig. 20), but the members are noticeably more similar for the r511 ensemble, with much smaller differences in the speed of the systems, as well. Additionally, all members of the r511 ensemble are able to maintain some convective activity.

a. Success rate

Recall that for the 2D experiment, a forecast is considered a success if the region of precipitation mixing ratio exceeding 0.2 g kg^{-1} extends for at least 20 km in length anywhere in the z = 3-5 km layer. The along-line variability in the MCS structure in the 3D runs necessitates an additional constraint for labeling a forecast as a success, namely

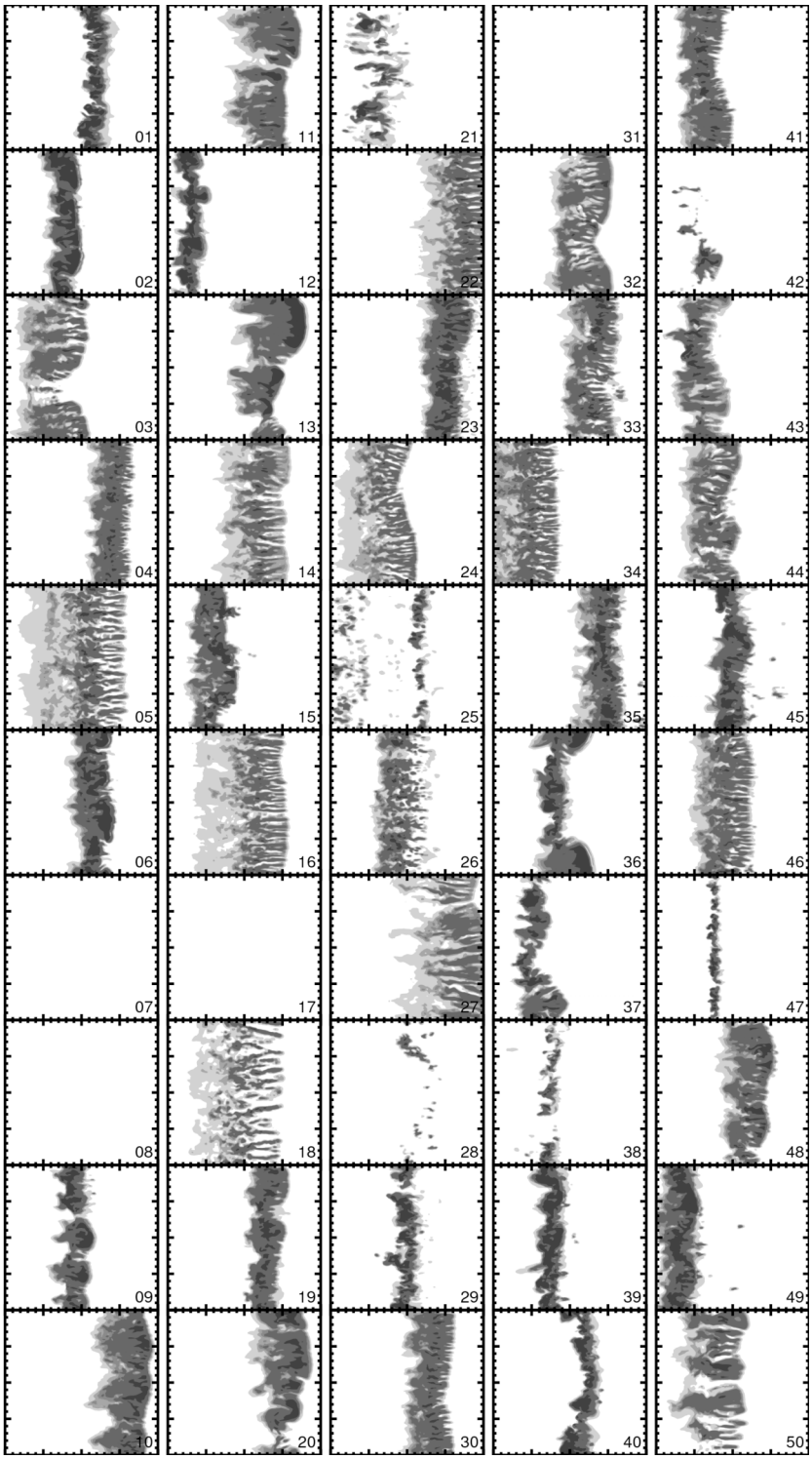

Fig. 19. Isolines of precipitation mixing ratio greater than [0.2, 1, 1.5, 2, 5] g kg^{-1} (shading from light to dark) for the 50 members of the 3D r823 ensemble at z = 4 km and t = 6 h. The grid extends to 100 km on either side of the location of the control run at this time, resulting in a 200 x 200 km window. The number in the lower left corner notes the ensemble member.

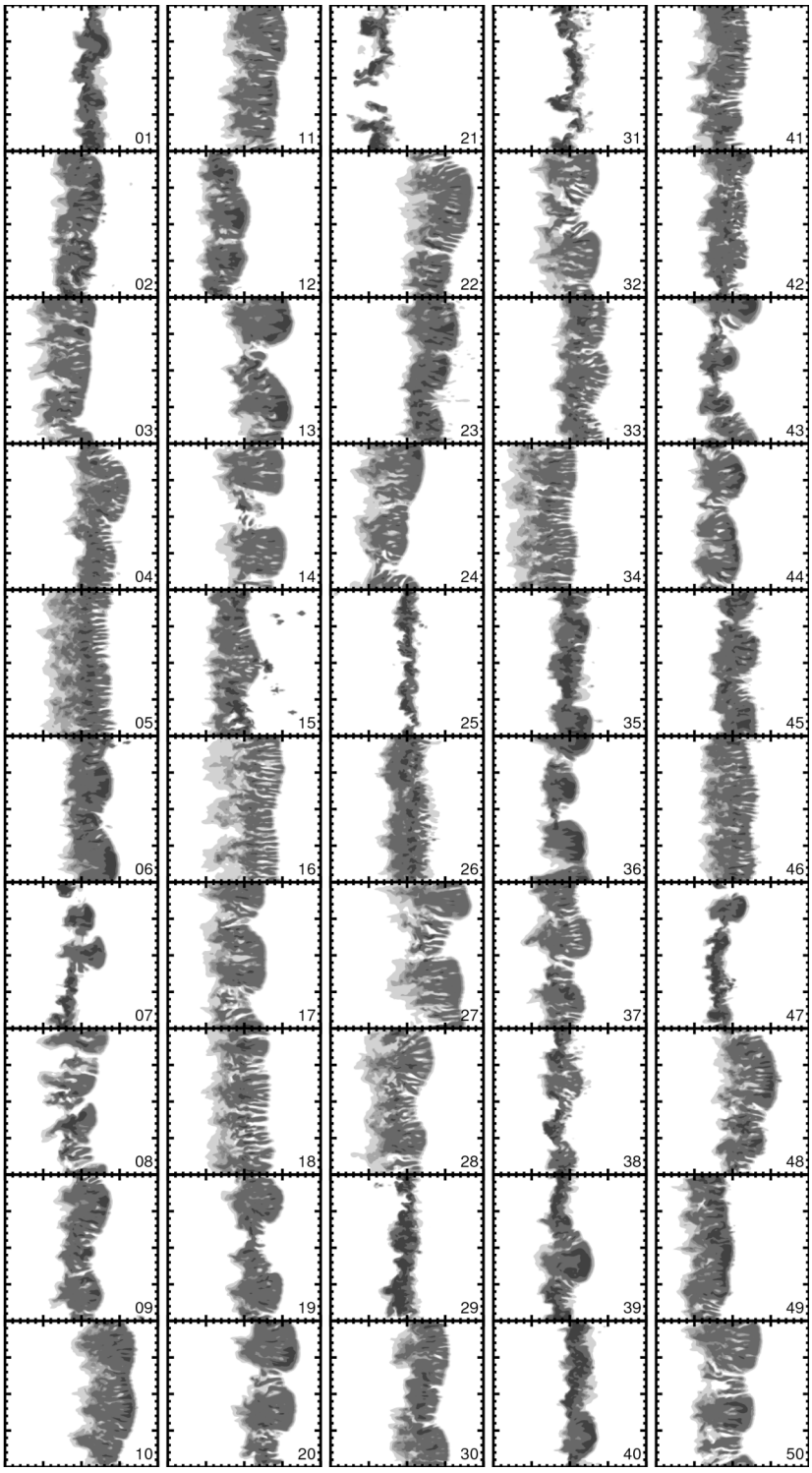

Fig. 20. As in Fig. 19, but for the r511 ensemble.

the 20-km width requirement must be met over at least 90% of the line and no gap in the line can exceed 10 km. Whereas the number of MCSs in the 2D runs rises drastically in the period t = [45, 75] minutes, immediate development occurs in only a third to half as many members in the 3D simulations, with subsequent successful MCS development requiring an additional 2-3 h (Fig. 19, cf. Fig. 11). As indicated above, the similar success rates between the 2D and 3D r823 ensembles (~70%) over the final 3-4 h of the simulations are due in part to the use of the drier control sounding in the 3D runs. The asymptotic success rate achieved by reducing the initial uncertainty for all three fields (r511) is 80-85%, or about 5-10 points lower than its 2D counterpart, which suggests that reducing the forecast uncertainty may be more difficult in 3D.

One interesting feature of the 3D success rate curves is the lack of organization according to the initial perturbations. For the 2D runs, the success rate steadily improves as the initial CAPE errors are reduced and the curves tend to group according to the degree of initial relative humidity uncertainty (Fig. 11). For the 3D runs, there is no such grouping nor is there any consistent dependence on the initial CAPE uncertainty (Fig. 21). Instead, all ensembles except r511 bunch together with many counterintuitive relationships. For example, reducing only the initial CAPE uncertainty from the 24-h forecast error level to the analysis levels (r523 vs. r823) actually decreases the success rate. Similarly, reducing the initial wind speed uncertainty also decreases the success rate marginally, by about 5%, between r523 and r521 and by about 10% between r823 and r821 during the period t = [4, 7] hours. No discernable sensitivity to the relative humidity uncertainty exists.

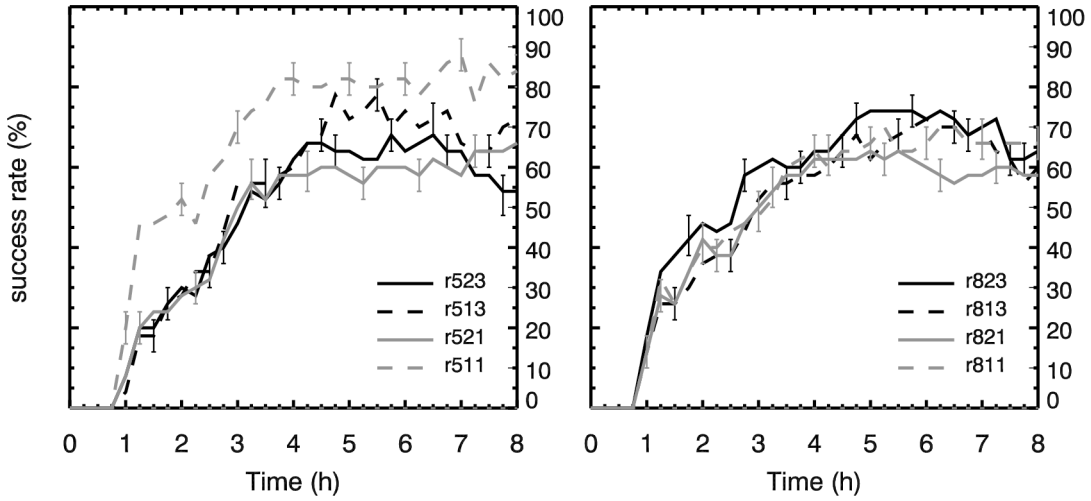

Fig. 21. As in Fig. 11, but for 3D simulations.

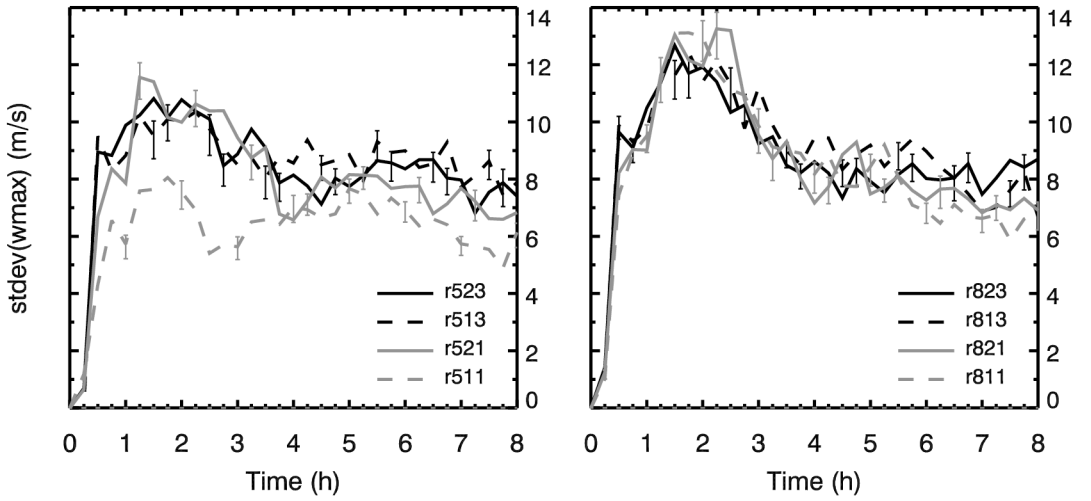

Fig. 22. As in Fig. 21, but for the standard deviation of maximum updraft strength (m s^{-1}).

b. Maximum updraft strength

The overall uncertainty in maximum updraft strength is comparable between the two- and three-dimensional configurations, although temporal differences exist (Fig. 22, c.f. Fig. 14). As with the 2D runs, histograms of 3D maximum updraft values, along with those of the fields described in subsequent sections, show that, while not purely Gaussian, the distributions are close enough to permit the use of the standard deviation as a measure of ensemble spread. The peak uncertainty occurs almost an hour earlier in the 3D simulations and the spread of the ensembles throughout the runs is 20-30% larger. As with the success rate curves, there is no concrete signal among the 800 J kg^{-1} ensembles. The separation between the two curves is minimal, but the spread of the r823 ensemble is consistently lower than that of r821 between t = 2 h and t = 6 h, again suggesting that reducing the initial wind speed uncertainty can lead to decreased confidence in the forecast. The peak of the 500 J kg^{-1} ensembles is somewhat lower than for the 800 J kg^{-1} ensembles, but by three to four hours into the simulations the uncertainty in the maximum updraft strength is the same for both ensemble groups. Most ensemble groups lack sensitivity to the initial perturbation magnitude; only reducing all fields (e.g., r511) yields substantial improvement in forecast confidence.

c. Maximum surface wind

The forecast uncertainty from the 3D runs for the maximum surface wind (actually lowest model level) undergoes a substantially different evolution than in the 2D runs (Fig. 23, c.f. Fig. 15). In 2D, the uncertainty increases rapidly over the first two hours and then changes very little for the remainder of the simulation period. In contrast,

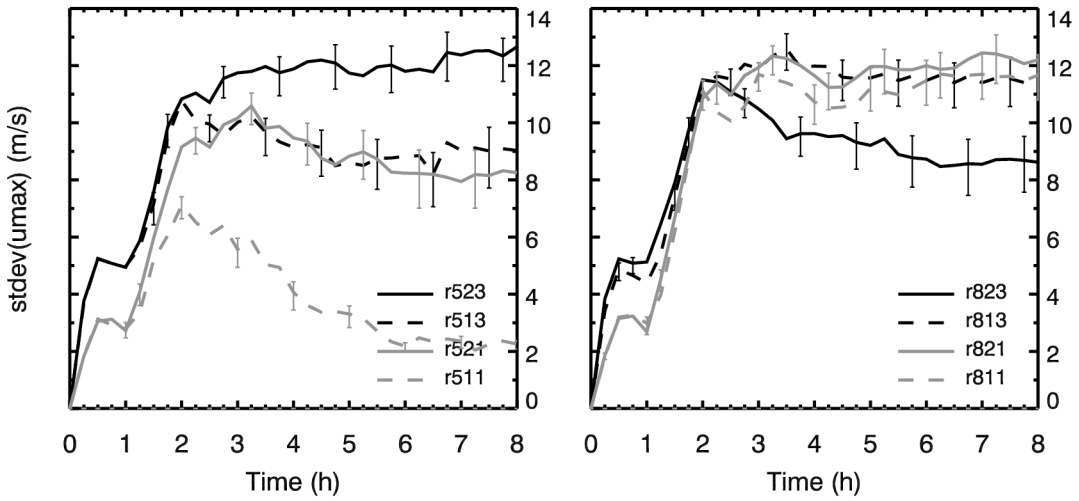

Fig. 23. As in Fig. 21, but for the standard deviation of maximum surface wind (m s^{-1}).

Table 3. Total rainfall integrated over time and space for the 3D runs. Units are in 10^5 mm.

Ensemble name	Mean	Standard Deviation
r823	8.68	3.50
r821	6.72	3.47
r813	7.51	3.46
r811	6.94	3.11
r523	7.14	3.59
r521	6.92	3.38
r513	7.74	2.96
r511	9.96	2.37

the 3D uncertainty peaks around t = 2 h (at levels 30-40% higher than in 2D) and then steadily falls; by the end of the simulations, the uncertainty has fallen by at least half. For the 3D ensembles, there is much greater variability in how quickly the systems develop leading to large uncertainty early in the runs. As the slower developing systems finally mature, the surface wind speeds become more similar. There are still few discernable differences between the 800 J kg^{-1} CAPE uncertainty ensembles, though, once again, there improving the initial uncertainty can lead to decreased forecast confidence (cf., r823 and r813). Similarly, the patterns that emerged among the 500 J kg^{-1} CAPE uncertainty groups in the success rate and maximum updraft plots are more pronounced for the maximum surface wind. While there is little difference between the r523, r513, and r521 ensembles, improving the initial uncertainty for all fields reduces the forecast uncertainty by about one-third, down to just a few meters per second.

d. Rainfall

As mentioned in the Chapter 1, one of the most important impacts of MCSs is hydrological; they are responsible for much of the warm-season rainfall over the eastern two-thirds of the United States (Fritsch et al. 1986), as well as floods and flash floods (Junkers et al. 1995; Schumacher and Johnson 2006). Two different rainfall measures are examined here: the maximum instantaneous rain rate and the 1-h accumulations, but first a look the rainfall totals integrated over the entire simulation.

The domain-averaged total rainfall among each ensemble group ranges from 6.72 x 10^5 mm for the r821 ensemble to 9.96 x 10^5 mm for the r511 ensemble (Table 3). Curiously, all the ensembles have similar means (6.7-7.5 x 10^5 mm), except the r823

(8.68 x 10^5 mm) and r511 (9.96 x 10^5 mm) ensembles—the ensemble representing current 24-h forecast errors and the one representing improvement in forecast errors for all three fields. Despite having the highest mean rainfall total, the r511 ensemble also has, by far, the least spread (25-40% lower). There is no simple relationship between the perturbations and either the mean or the standard deviation of the total rainfall, not even for relative humidity as one might expect for a rainfall forecast. In fact, as with the means, there is not a great deal of variability in the spread of the different ensembles with the exception of the r511 ensemble, as noted above, and the somewhat smaller spread of the r811 and r513 ensembles.

The maximum 1-h accumulations (Fig. 24), for the 800 J kg^{-1} simulations suggest a clustering according to the wind speed perturbations up to about t = 4 h, at which point the curves converge. There is even a hint that the curves then cluster according to the relative humidity perturbations over the last few hours. In contrast, for the groups with smaller initial CAPE uncertainty, the ensembles are similar up to about t = 4 h, then the spread of the r511 ensemble declines dramatically. Recall that for this ensemble, while not all members meet the MCS criterion, convection is present and robust in all members at t = 6h (Fig. 20). However, for some members (e.g., members 21 and 25, among others) the convection is not very extensive. Locations in these runs receive rain for a shorter period—translation speeds are similar for this ensemble—leading to lower accumulations for these members and greater spread for the ensemble. Two hours later (not shown), most of these members have continued to develop and now cover an area similar in scope to the other members, reducing the spread.

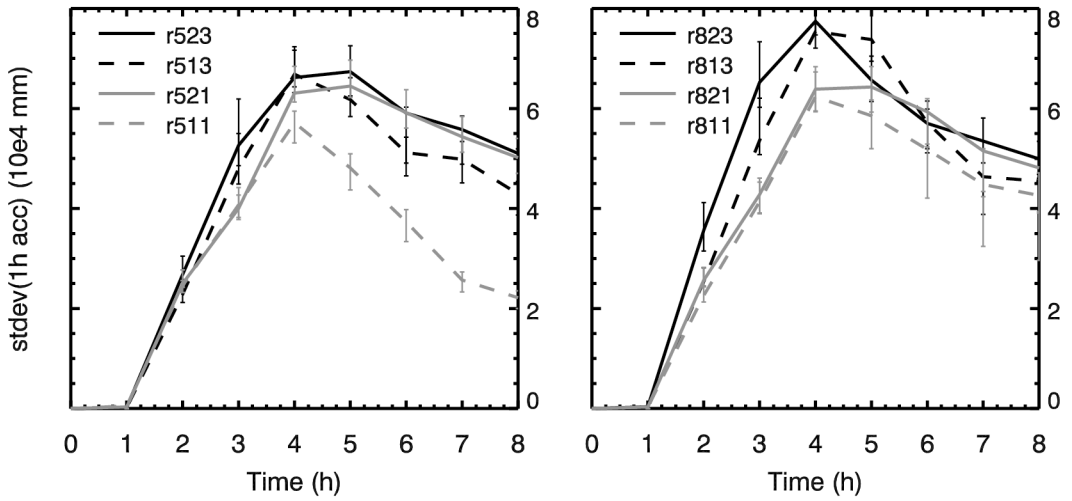

Fig. 24. As in Fig. 21, but for the standard deviation of maximum 1-h accumulated rainfall (mm).

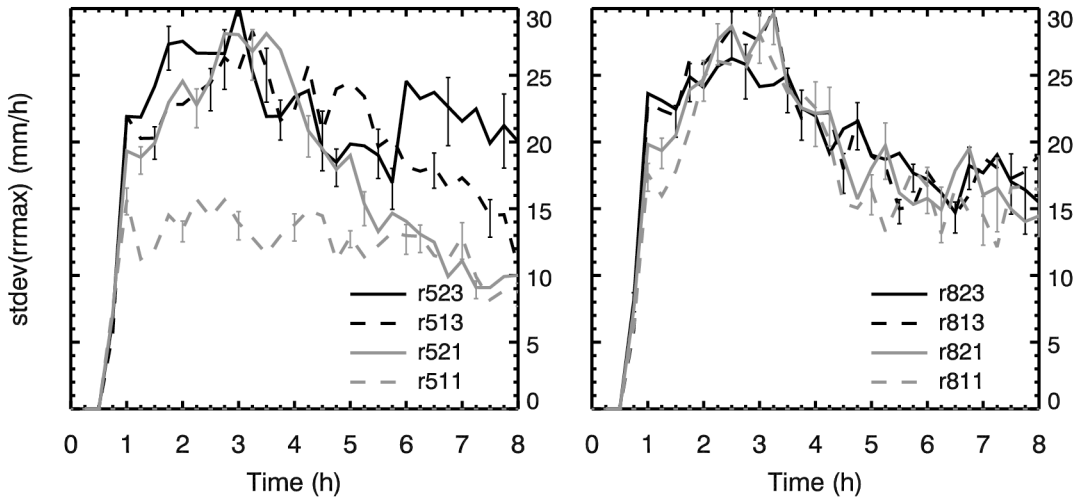

Fig. 25. As in Fig. 21, but for the standard deviation of maximum instantaneous rain rate (mm h^{-1}).

Continuing to even smaller temporal scales, the maximum instantaneous rain rate (Fig. 25) behaves much like the maximum updraft and maximum surface wind. The curves are indistinguishable among the 24-h CAPE uncertainty groups, along with the r523 ensemble; r521 follows these ensembles for the first 3-4 h and then declines by more than half over the rest of the run, finally matching the low spread of the r511 ensemble—after showing the largest spread of any ensemble near $t = 3$ h.

The correspondence ratio (Stensrud and Wandishin 2000) provides another perspective on the spread of the rain forecasts. The correspondence ratio (CR) is a numerical representation of the information contained in a Venn diagram, such that

$$\mathrm{CR} = \frac{I}{U},$$

where U is the area of the union of all specified field values and I is the intersection of these same specified field values. For example, U and I can be taken with respect to the grid points where precipitation exceeds a given threshold. Furthermore, I can be defined for a subset of the full data. For example, the intersection can be taken as the region over which at least half of the members of an ensemble exceed a given threshold; the correspondence ratio thus defined is identified as CR_f, where f is the fraction of members used to define the intersection. The CR is bounded by 0 and 1. Low values of the CR may be the result of either too few members exceeding the threshold value or phase errors in time or space.

The CR for 1-h accumulations exceeding 0.2 mm (basically a forecast of whether rain will occur) is shown in Fig. 26. The transient nature of the instantaneous rain rate is ill-suited for this measure. From top to bottom, the CR_f is given for $f = 10$, 20, 30, and 40

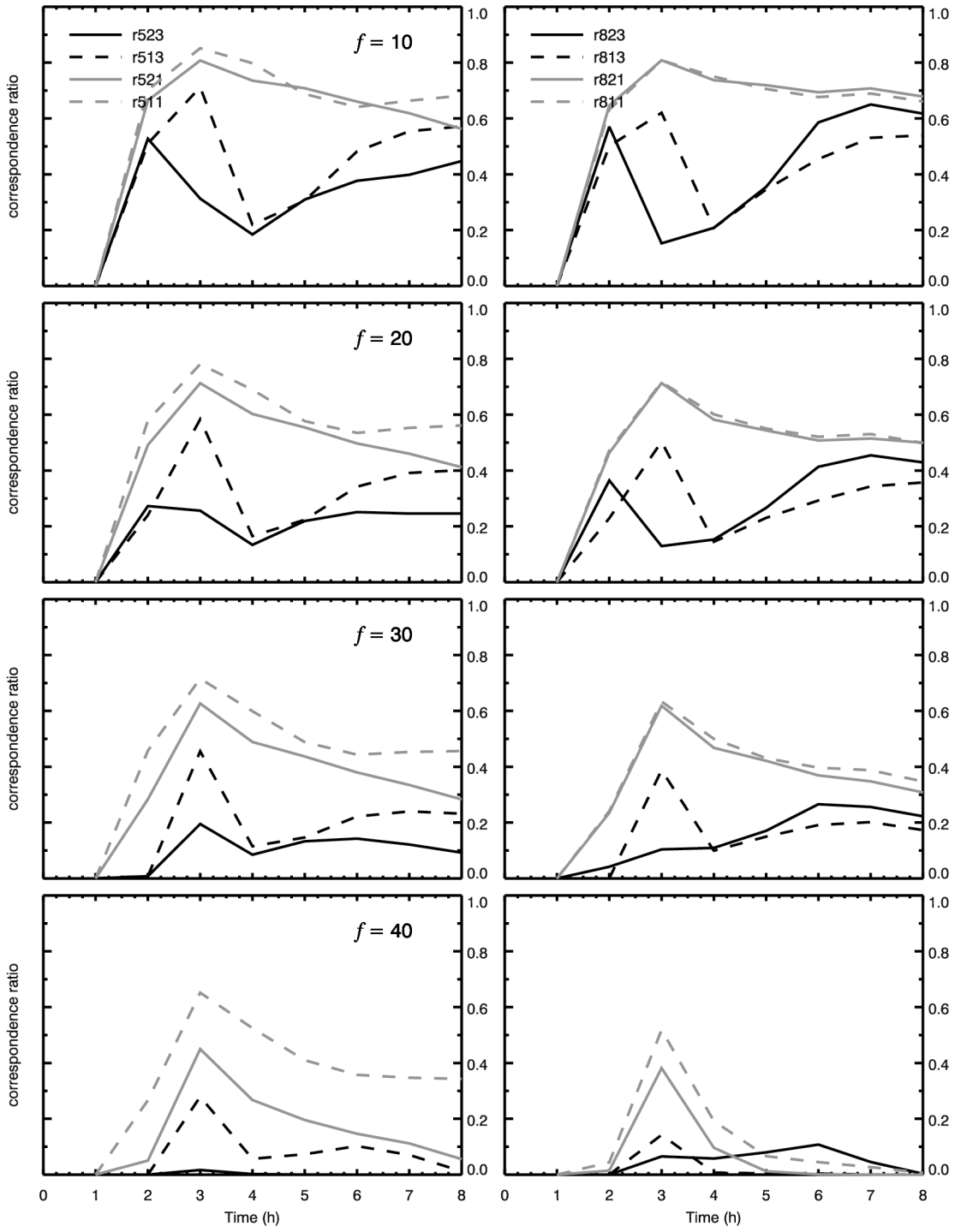

Fig. 26. Correspondence ratio, CR_f for 1-h accumulations greater than 0.2 mm for f = 10, 20, 30, and 40 members from top to bottom. Linestyles as in Fig. 21.

members. As expected, the CR declines substantially as *f* increases. The CR shows a strong dependence upon the magnitude of the initial wind speed uncertainty. Reducing this uncertainty results in an improvement of as much as 0.6 around $t = 4$ h for CR_{10}, with the curves converging as the runs progress. The greater certainty that is throughout thischapter for the r511 ensemble is not present for CR_{10}, but does manifest itself as more members are required for the intersection. The initial relative humidity and CAPE uncertainty has little impact on the scores. It is possible to create a map of the CR, or, more accurately, a map of the various intersections (Fig. 27). The light gray shading represents the union all members and the successively darker shading marks the area of intersection of at least 10, 20, 30, and 40 members. For reference, the contour indicates the location of the control run. The greater agreement within the r511 ensemble for the larger *f* values is easily seen, as is the growth of the area of union—both because of the growth in the MCSs and the increasing spread of the members. Since the union forms the denominator of the CR, this growth contributes to the gradual decline in scores after $t = 3$ h. Indeed, the size of the union appears to explain a great deal of the difference between the ensembles with the 24-h forecast wind speed errors and those with the reduced wind errors. The former group displays a tendency to somewhat larger systems and a greater range of translation speed (not shown) leading to substantially larger unions and thus smaller CRs, despite similar-sized areas of intersection for the lower the *f* thresholds.

Increasing the rainfall threshold to moderate levels, 2 mm (Fig. 28) and 5 mm (Fig. 29), results in only a slight decline in CR_f for $f = 10$ and 20 members, but a substantial reduction for the more restrictive $f = 30$ and 40 member thresholds. Furthermore, the separation of the curves based on the wind speed uncertainty becomes

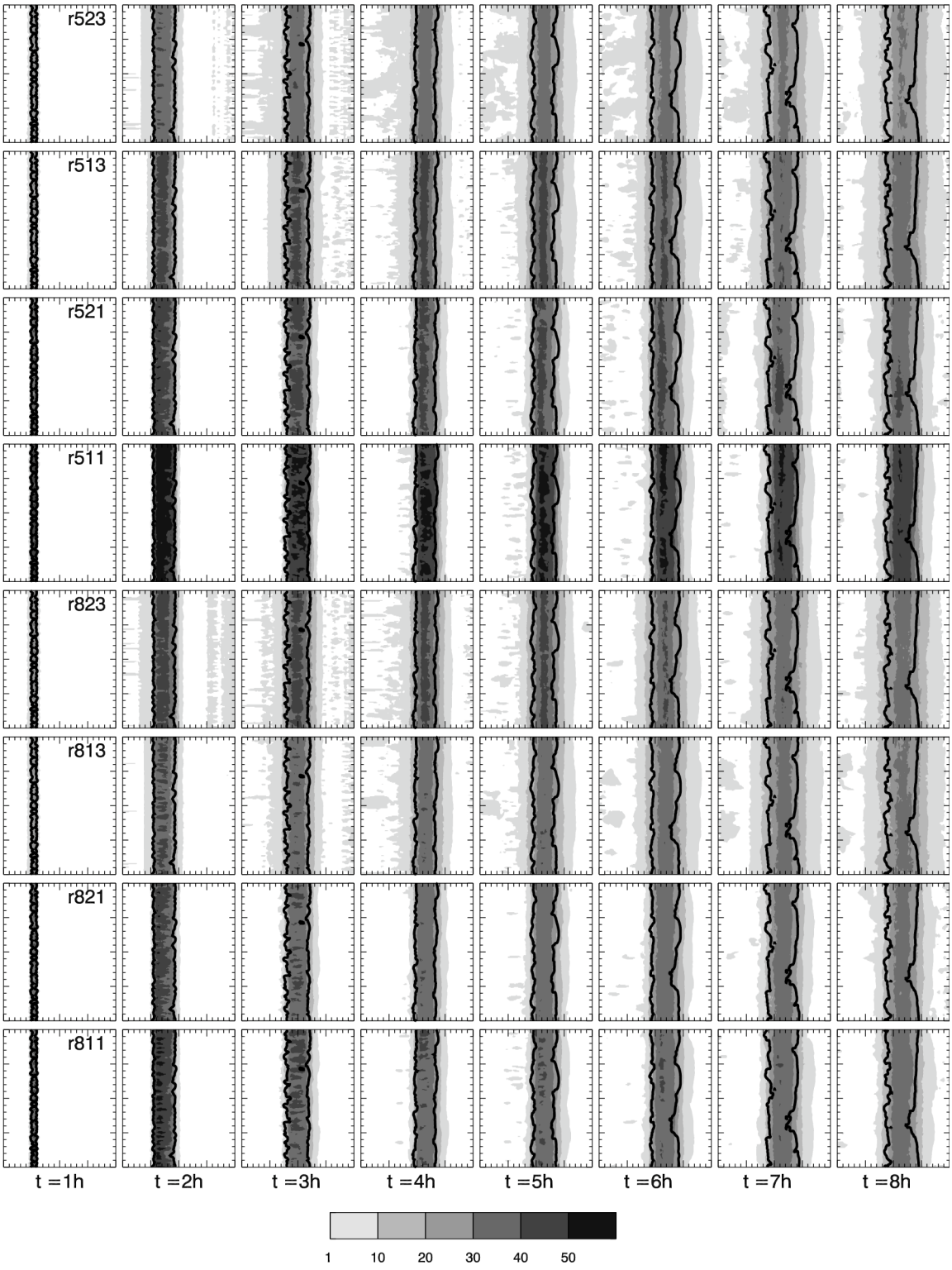

Fig. 27. Correspondence ratio maps for 1-h accumulations greater than 0.2 mm. Shading (from light to dark) denotes the intersection of at least 1 (the union), 10, 20, 30, 40, and 50 members. The dark contour marks the 0.2 mm contour from the control run.

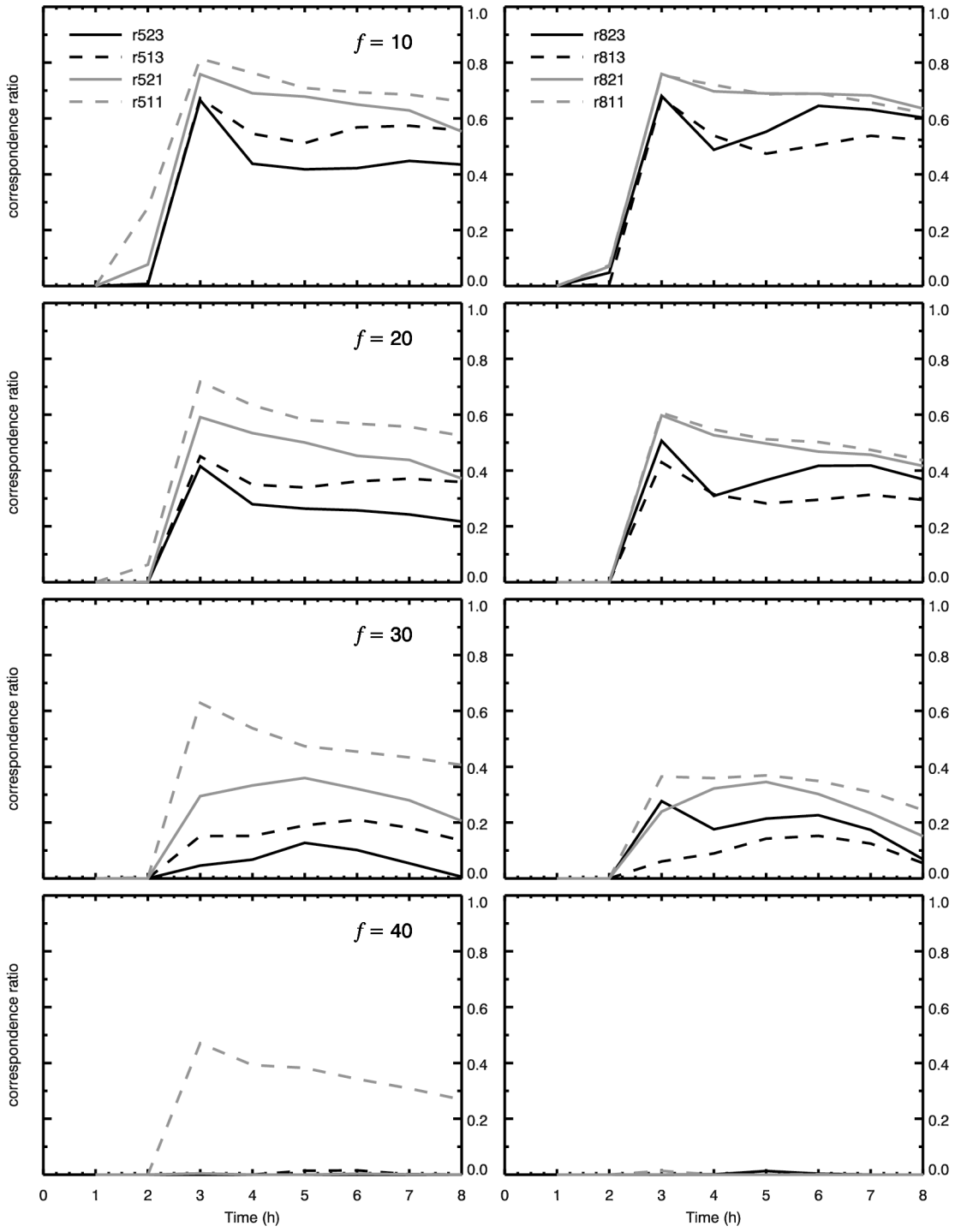

Fig. 28. As in Fig. 26, but for 1-h accumulated rainfall greater than 2 mm.

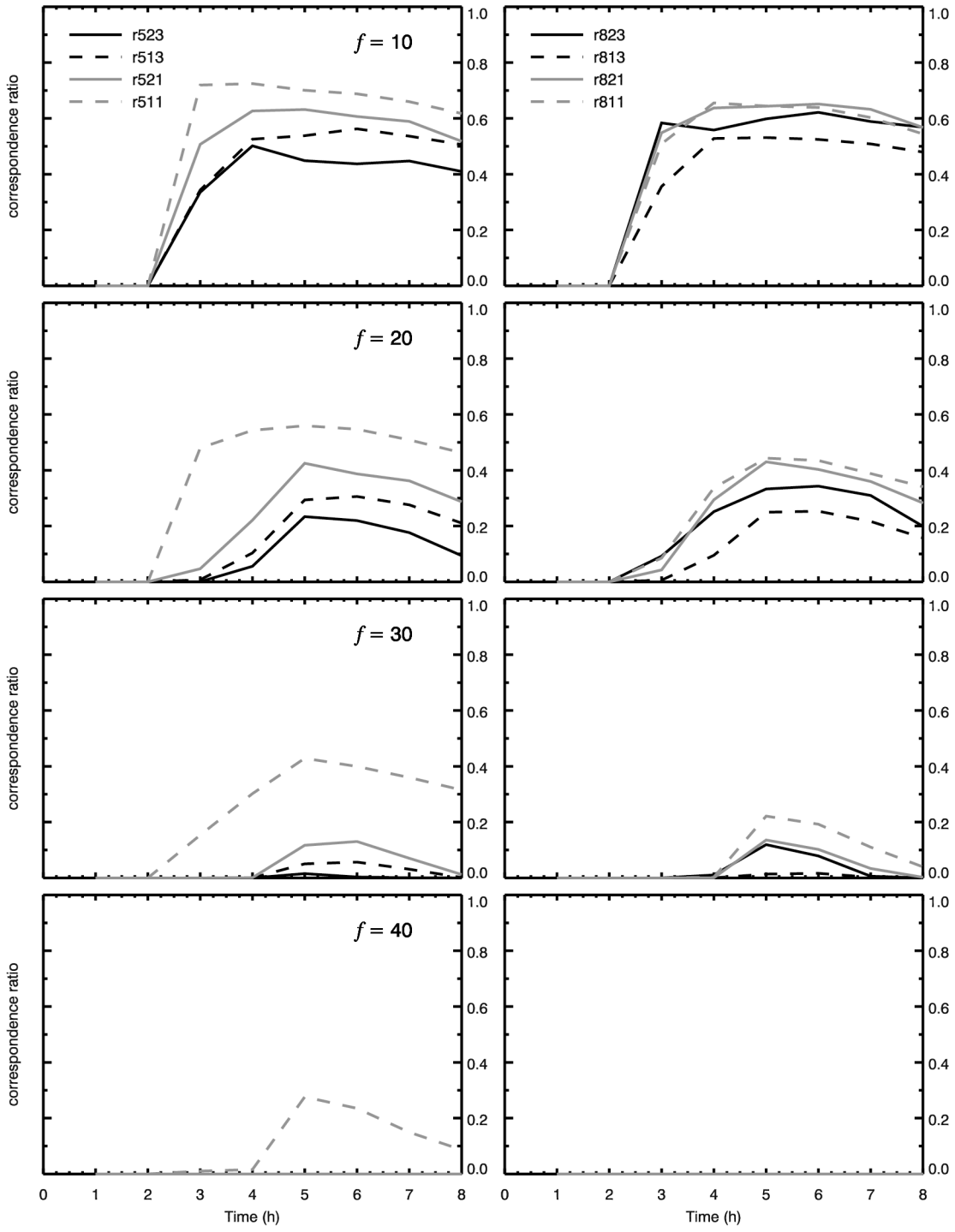

Fig. 29. As in Fig. 26, but for 1-h accumulated rainfall greater than 5 mm.

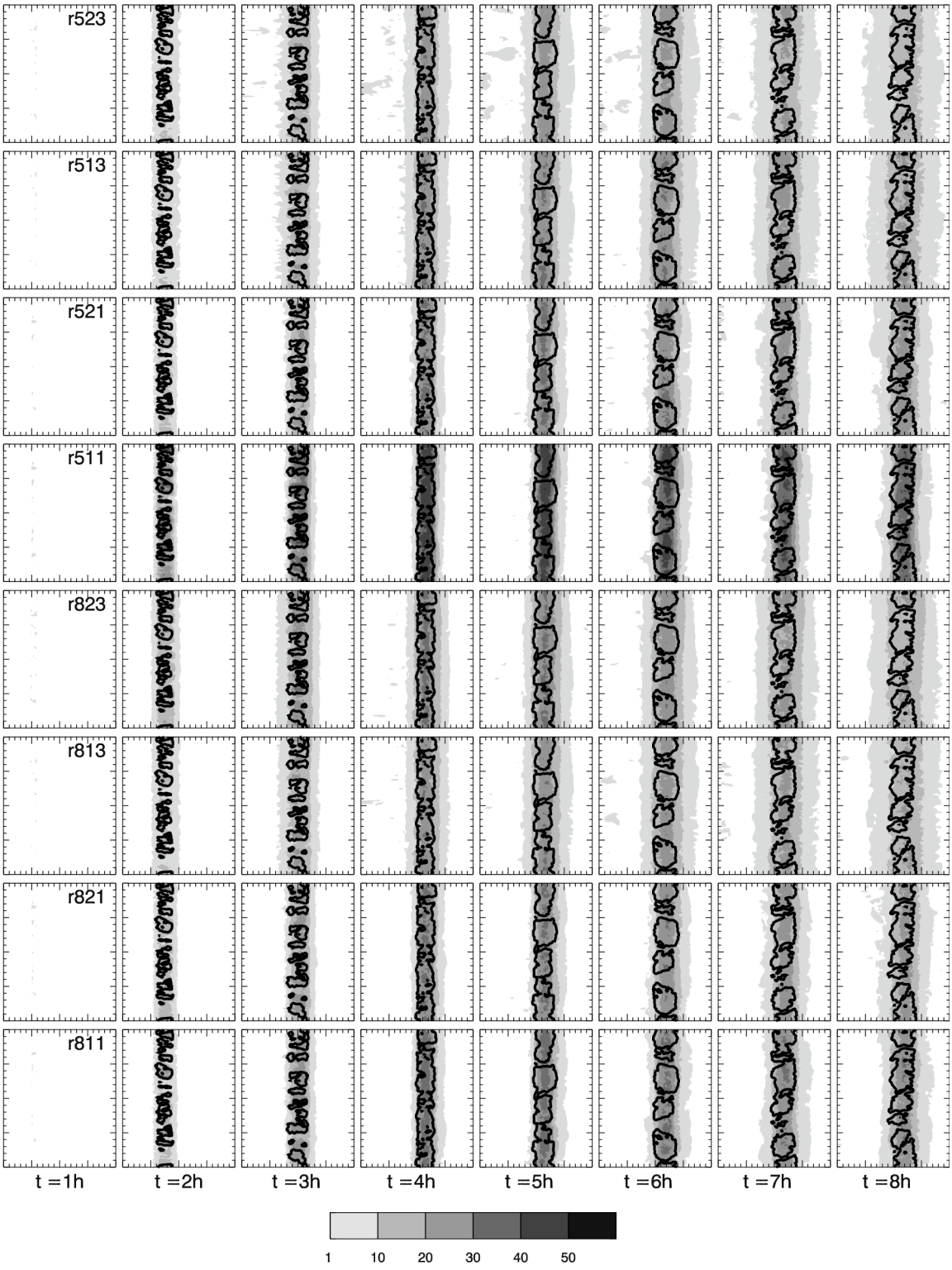

Fig. 30. As in Fig. 27, but for 1-h accumulated rainfall greater than 5 mm.

less pronounced. Note that the union of the 24-h uncertainty ensembles at the 5 mm threshold (Fig. 30) is closer in size to the reduced uncertainty ensembles. The large aerial coverage of the 0.2 mm unions is largely the result of a number of relatively weak, broadly diffused systems (e.g., member 5 or 16; Fig. 19) and the larger differences in translation speed of the systems for the ensembles based on 24-h wind speed forecast uncertainty (c.f., Figs. 19-20). The impact of the diffuse systems is reduced as the rainfall threshold is increased. Note, as well, that as the rainfall threshold is increased from 2 mm to 5 mm, the time at which the CR peaks shifts from t = 2-3 h to around t = 5 h. This later time also marks the peak in both the success rate curves (Fig. 21) and the curves marking the spread of the 1-h rain accumulations (Fig. 24); that is, the storms must reach maturity to produce agreement on the rain higher thresholds.

The peak areas for which at least 10 members of the ensembles agree on 1-h rainfall amounts exceeding 10 mm (Fig. 31) are still reasonably large—only a modest decline from the 5mm curves—but the CRs quickly drop away from that peak due to a combination of a smaller area of intersection and a larger union (not shown). Agreement for larger f thresholds is scant. Of course, one would not expect to find strong agreement on the timing and location of the stronger convective cores that are generally required to produce higher rainfall amounts. (The initial environment used here does not lead to the asymmetrical, training systems that can generate extremely large rainfall totals, as seen in Parker and Johnson 2000.)

A consistent but puzzling signal in the CR plots is that the r511 ensemble, which possesses a much stronger degree of self-agreement (i.e., high success rate, low spread) than all the other ensemble groups, tracks closely, for low rainfall or f thresholds, with the

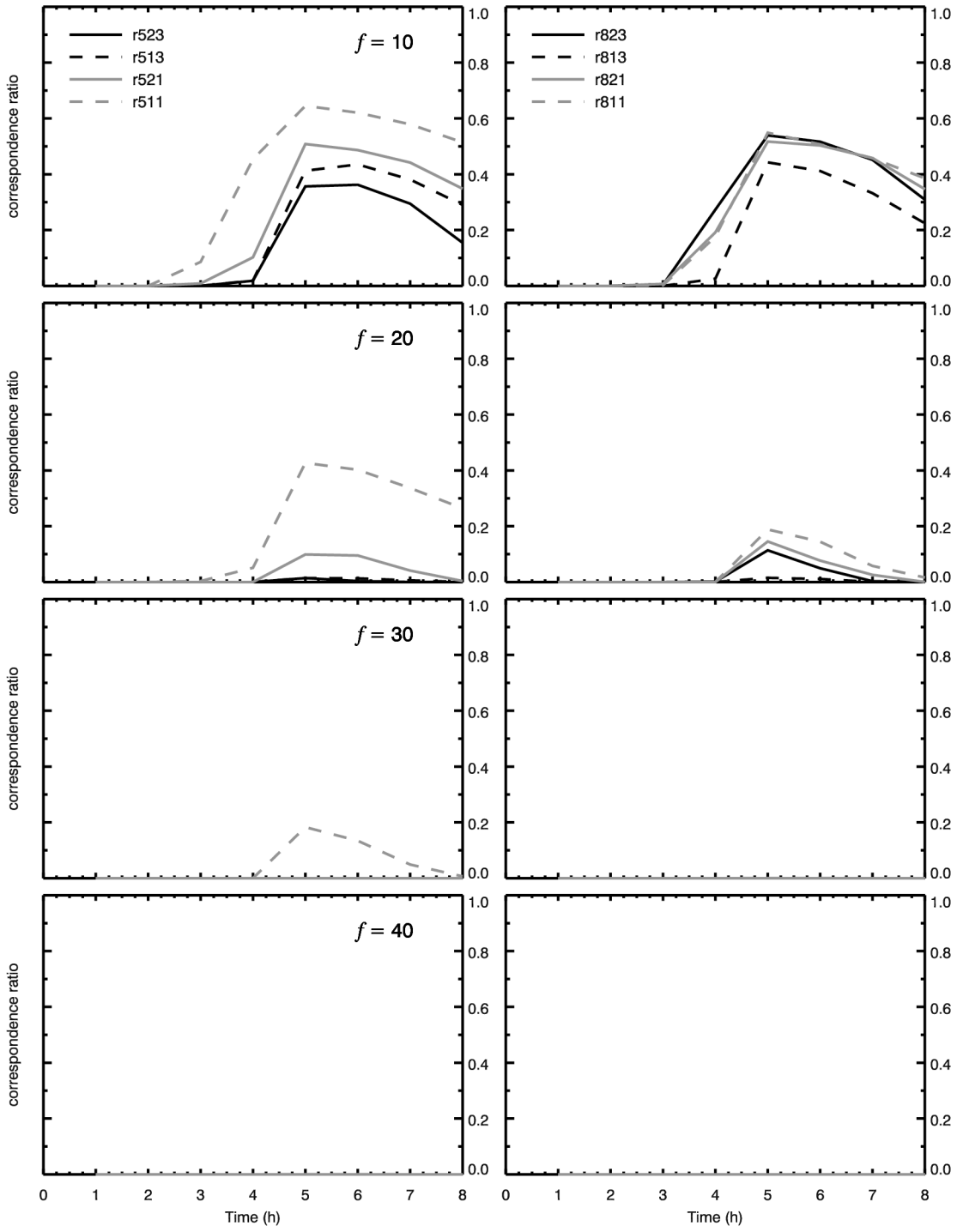

Fig. 31. As in Fig. 26, but for 1-h accumulated rainfall greater than 10 mm.

CR of the other ensembles with reduced initial wind speed uncertainty. As either the rainfall amount or the level of intersection, *f*, is increased, the r511 ensemble begins to separate itself from the other curves, but, for the most part, not to the same extent as seen for the other measures. It is not clear what different aspects of the simulations or the evaluation tools are responsible for the differing results.

Another signal unique to the CR plots, just mentioned above, is the dependence on the initial wind speed uncertainty. As discussed earlier, the larger initial wind speed uncertainty produces an anticipated greater range of system translation speeds, which creates a larger area for the union of all members, leading to smaller CR values. The success rate is not as sensitive to spatial displacements (the degree of sensitivity is a function of the size of the search window used) and so shows no sensitivity to the wind speed uncertainty. The other measures focus on extreme values and so are likewise insensitive to location errors. This highlights the importance of the choice of evaluation tool, as different tools will often key on different facets of forecast behavior. It is thus necessary to select an evaluation tool based on the forecast characteristics of interest. For example, the success rate should not be used (or should be used with a narrow search window) by someone sensitive to errors in forecast location.

e. Bowing segments

It was noted at the beginning of this chapter that among the variety of structures that develop in the 3D runs are several that resemble bow echoes, a feature frequently associated with damaging winds (Fujita 1978; Weisman 2001). To examine these features more closely, the evolution of three apparent bows from the r523 ensemble is

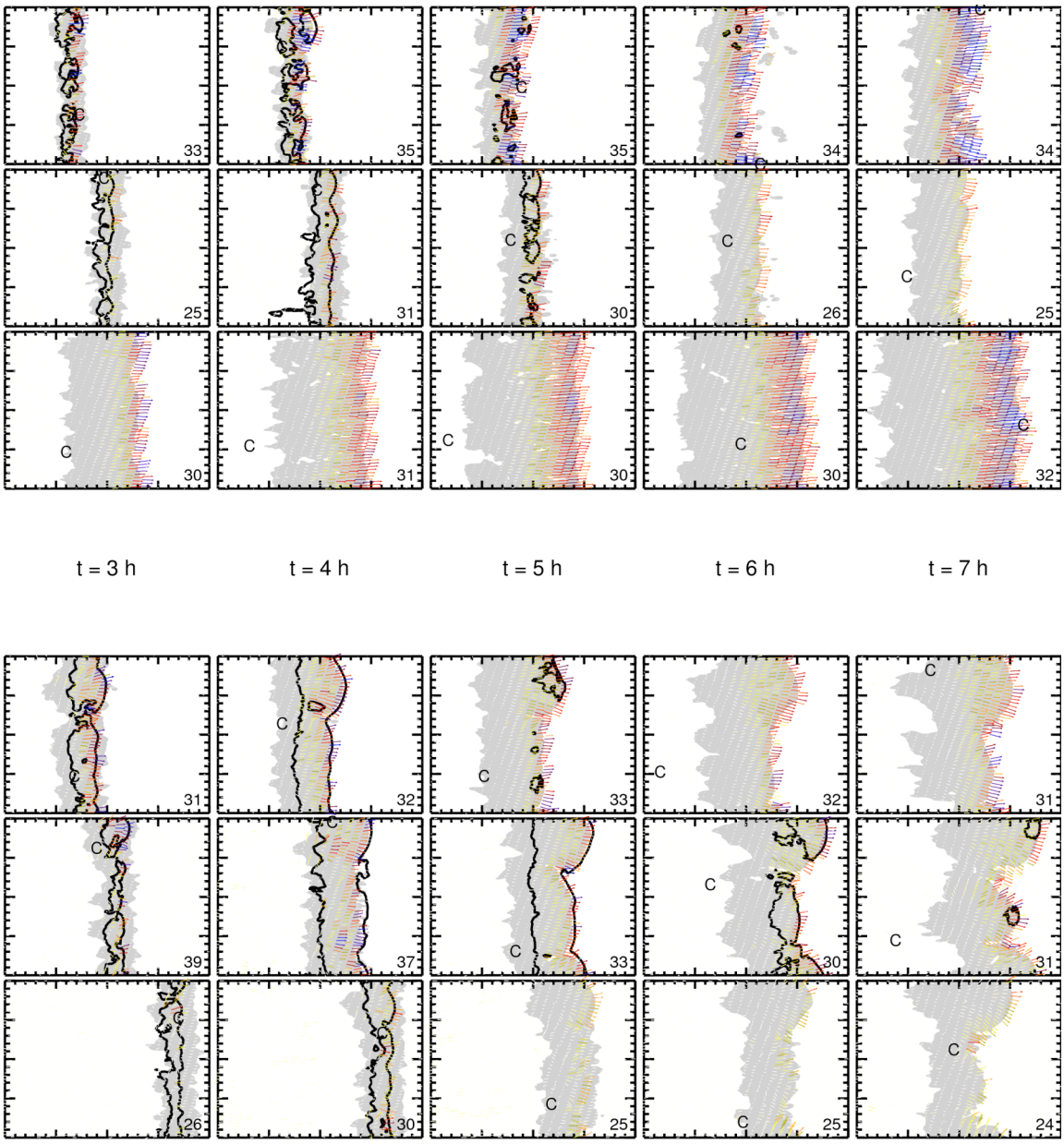

Fig. 32. Precipitation mixing ratio at z = 4 km greater than 0.2 g kg^{-1} (shading), together with the perturbation pressure greater than 3 mb (contour) and horizontal wind (arrows) at the lowest model level at t = 3-7 h for three linear members (top) and three bow-like members (bottom) of the r523 ensemble in a moving 200 x 200 km window. The 'C' in each panel marks the location of the minimum temperature perturbation. The number in the lower right corner of each plot is the maximum surface wind within the window.

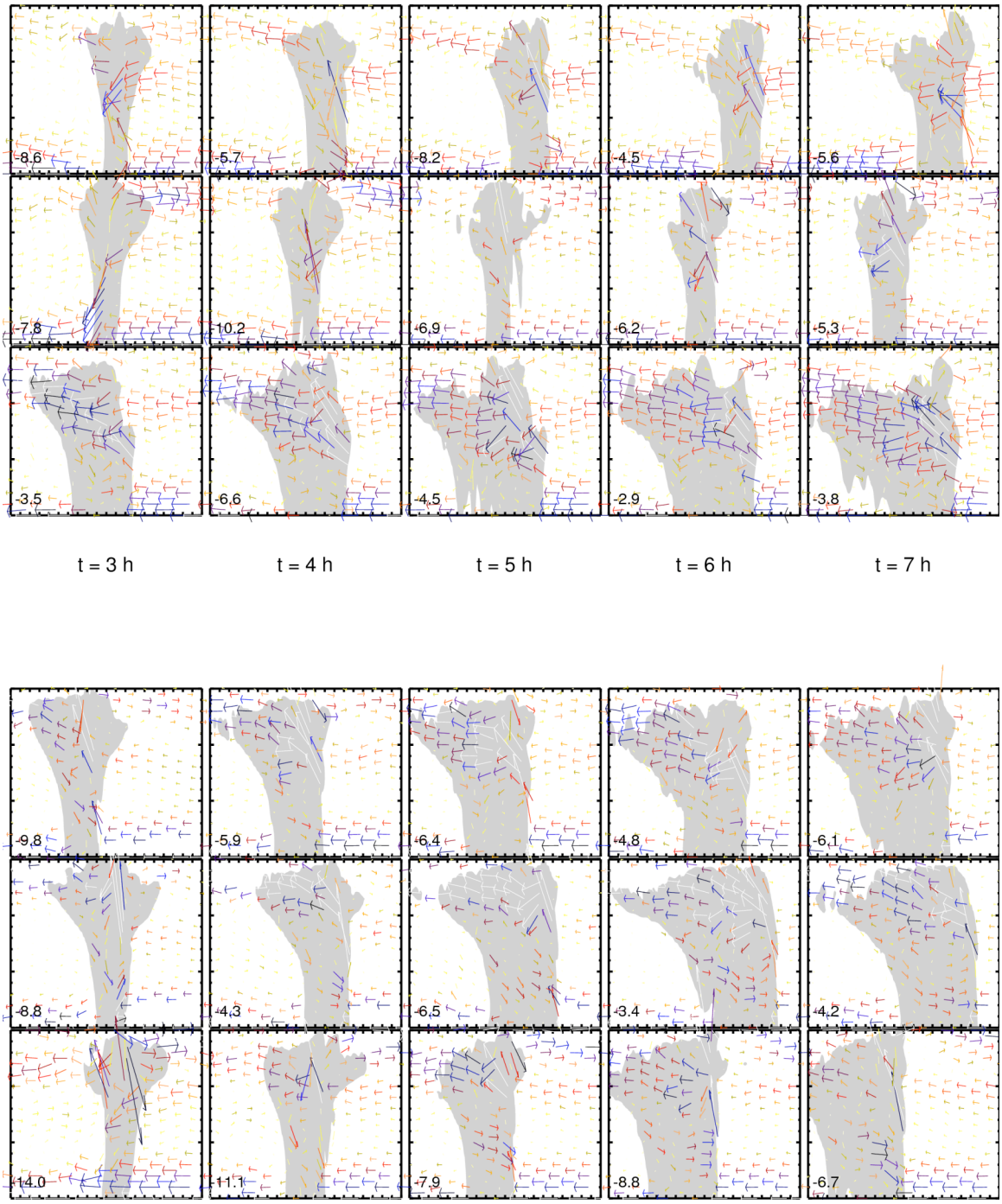

Fig. 33. As in Fig. 32, but for cross-sections (200 x 15 km). For the linear members (top), the cross-sections are taken through the middle of the domain (y = 100 km). For the bow-like members (bottom), the cross-sections are taken through the apex of the bow (y = 150, 175, and 175 km). Arrows show the (u, w) wind. The number in the lower left corner indicates the maximum downdraft in the window.

plotted along with three linear members (Figs. 32 and 33). The maximum surface wind (plotted in the lower right corner of each panel, Fig. 32) is similar for both groups; the bows (bottom) are no more likely to exceed the threshold for severe winds (~34 ms^{-1}) than are the lines (top). Similarly, there appears to be no connection between either the strength of the cold pool (not shown) or the location of the maximum temperature perturbation (marked by a 'C'). The perturbation pressure field (black contour = +3 mb) seems to be more extensive for the bowing members and is often coincident with the bowing segment, but given the presence of the +3 mb perturbation pressure field among some of the linear members, it would be difficult to devise a bow echo forecast rule from this field.

The cross-sections (Fig. 33) are taken through the middle of the domain for the linear members and through the apex of the bow for the bowing members. The movement of the system has been subtracted out to give systems-relative winds. The number in the lower left corner of each panel indicates the speed of the maximum downdraft. As with the maximum surface wind there is no signal from in the maximum downdrafts, neither group possesses a tendency toward stronger downdrafts. It does appear, however, that the descending rear-to-front flow, though not exceedingly strong, is more pronounced for the bowing members, while the front-to-rear flow is stronger in the linear members. That the members would develop different kinematics is not clear, however, from looking at the initial perturbations for these members (Fig. 34).

So, whereas there are suggestions that the bow-like solutions are dynamically different than the more linear solutions—stronger descending rear-to-front flow leading to stronger surface pressure perturbations—the two sets of solutions are indistinguishable

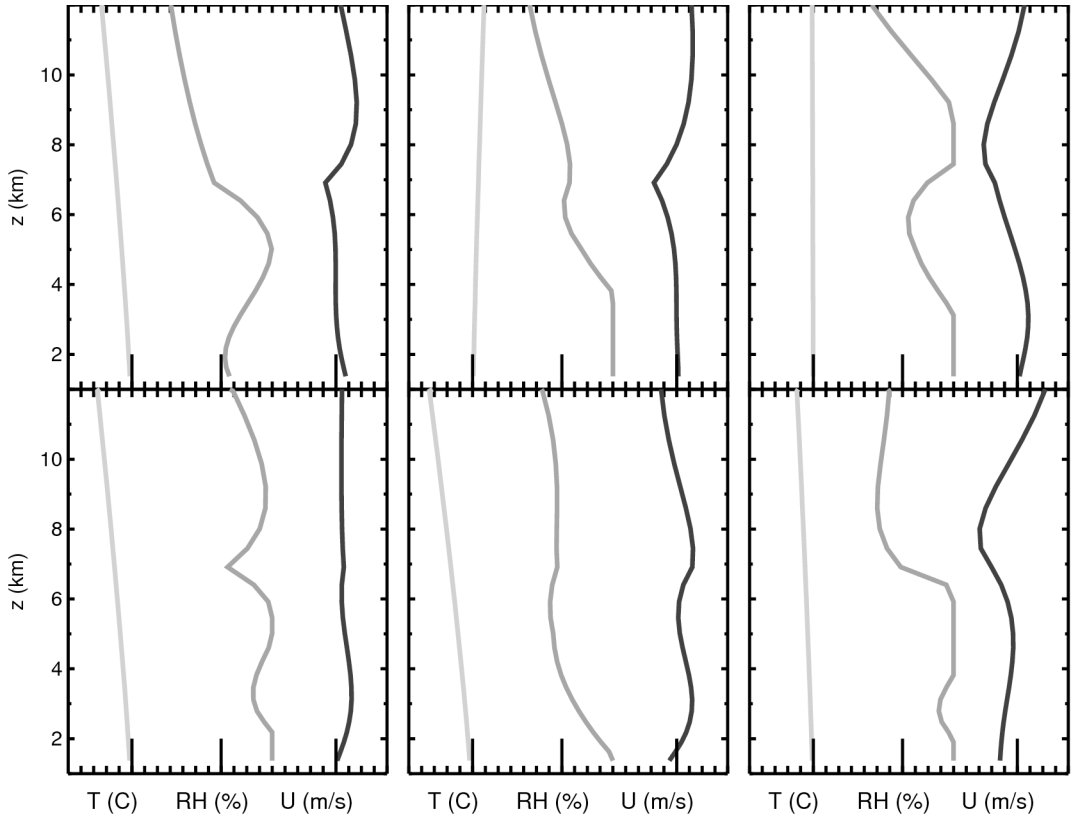

Fig. 34. As in Fig. 7, but for the three linear (top) and three bow-like (bottom) members.

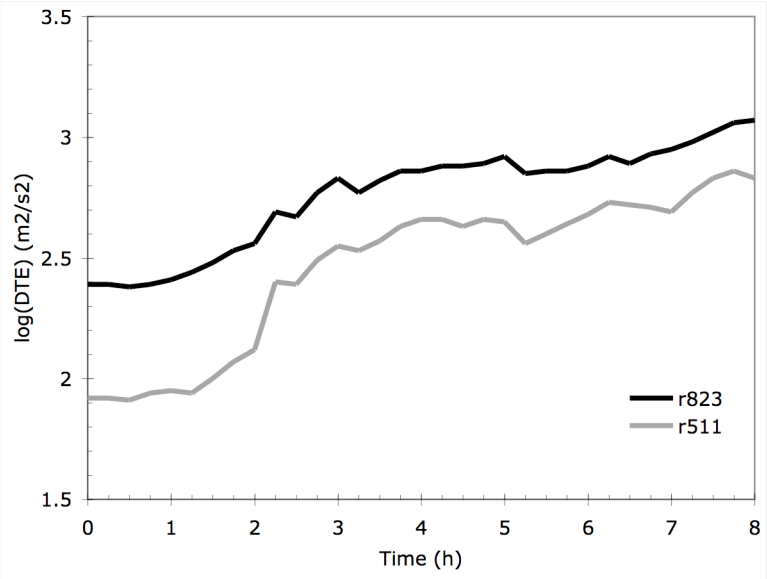

Fig. 35. Time series of the logarithm of the difference total energy (DTE) averaged over all members of the r823 (black) and r511 (gray) ensembles. See text for definition of DTE.

in terms of the strength of the downdrafts and surface winds. More importantly, from a forecast perspective, the two sets of solutions also are indistinguishable in terms of the perturbations from which they are initialized.

e. Difference total energy

In order to compare these results with those in Zhang et al. (2003), the difference total energy (DTE) is computed between each ensemble member and the control run, where

$$DTE = 1/2 \sum \left[(\delta u)^2 + (\delta v)^2 + \kappa (\delta T)^2 \right],$$

and where δu, δv, and δT are the difference wind components and difference temperature between two simulations, $\kappa = C_p/T_r$, T_r is a reference temperature of 287 K, and the summation is taken over the whole 3D domain. Error growth through integration is steady, but slow, although it possibly could be argued that the error growth for the r511_3D ensemble is slightly faster than for r823_3D (Fig. 35). While this result is somewhat surprising when compared to the MCS success rate results, it actually is to be expected. It is commonly seen (e.g., Zhang et al. 2003; Zhang et al. 2006; Hohenegger and Schär 2007) that error growth rates are strongly dependent on the amplitude of the initial perturbations; error growth slows as the initial perturbation amplitude increases. In fact, given the relatively large amplitude of the perturbations used here, it is perhaps surprising that there is any error growth whatsoever. The discrepancy between the DTE results and the MCS forecast success rate suggests that DTE may not be an appropriate predictability measure when a more phenomenological approach is taken, for the reasons noted by Leith and Kraichnan (1972). Namely, a squared difference measure, like DTE,

cannot distinguish between trivial and significant displacements of isolated features. Two simulations with storm locations separated by 50 km will have the same DTE as if the locations were separated by 500 km. However, the former would be counted as a forecast success by the criteria employed in this study, while the latter would not.

f. Summary of 3D results

Results indicate that the 2D experiments provide reasonable estimates of the 3D MCS predictability insofar as the success rate, along with the spread of the maximum updraft and maximum surface wind speed, are comparable between the two experiments. Looking more closely at the 3D results, however, uncovers some intriguing, and not easily explainable results. For example, in the 2D results there are definite sensitivities to the strength of the initial uncertainty in the three perturbed fields, such as a steady increase in agreement among ensemble members as the CAPE uncertainty is reduced. In contrast, no clear patterns emerge in the 3D runs; all ensembles tend to group together, with the notable exception of the r511 ensemble, for which the initial uncertainty is reduced for all fields.

The spread of maximum instantaneous rain rates behaves similarly to the other fields, except that the standard deviation of the r521 ensemble sharply declines after $t = 3$ h. The forecast uncertainty decreases for all ensembles over the last several hours, but the r521 drops by half, eventually reaching the same level as the r511 ensemble. Integrating the rainfall over 1 h increments begins to reveal some sensitivity to the initial uncertainty. For the 24-h forecast error CAPE uncertainty group, at least, the curves tend to separate according the strength of the wind speed uncertainty. This tendency is much

stronger when the correspondence ratio is used to measure forecast spread. Maps of the correspondence ratio reveal that the separation according to initial wind speed uncertainty is caused by greater variation in the translation speed of the runs based on the 24-forecast errors. The spread in the location of the MCSs increases the union of all members, while, for the lower rainfall amounts, the intersection of the members does not change substantially, with the result being lower CRs.

Several of the 3D runs obtain a structure resembling bow echoes. The bow-like solutions appear to have a somewhat more pronounced descending rear-to-front flow and consequent positive perturbation field, similar to a mesohigh. However, whereas bow echoes are often associated with damaging winds, the surface winds in the bow-like solutions are no stronger than the non-bowing solutions. Moreover, there is no clear signal in the initial perturbation profiles that could aid a forecaster in predicting the existence of bow echoes.

Chapter 6: Summary and Conclusions

Mesoscale convective systems are a prominent meteorological feature over the eastern two-thirds of the contiguous United States and are responsible for a majority of the warm-season rainfall over most of this region (Fritsch et al. 1986; Schumacher and Johnson 2006). Individual MCSs can carry the risk of flash flooding, damaging hail, and even tornadoes, while MCS activity over a monthly to seasonal timescale is associated with flooding and draughts.

Despite this prominence, little is known about the predictability of these systems. Historically, predictability studies have relied on simplified statistical models or low-resolution global models, where the focus is on fields such as 500-hPa heights. Extrapolations from these models suggest mesoscale predictability to be on the order of hours. Initial experiments with mesoscale models lead to claims of enhanced predictability on these scales, but they largely ignored sensible weather. Subsequent work called into question the enhanced predictability conclusion, noting several shortcomings in the experimental design. Recent, storm-scale experiments have concentrated on the predictability of precipitation and the role deep, moist convection plays in mesoscale predictability. Conclusions from these studies tend to favor the relatively short predictability limits indicated by the early experiments based on the statistical turbulence models, though fixed forcing mechanisms, such as orography, have been linked to extended predictability limits.

The research described in this document focuses on a single mesoscale phenomenon, the MCS, and explores the predictability of several features of these systems that are of interest to forecasters: the development and maintenance of the

system as a whole, maximum updraft strength, maximum surface wind, and precipitation. The study is undertaken from the perspective of a forecaster looking at a day 2 forecast from a currently operational mesoscale model. If the forecaster sees the model developing an MCS in the forecast area, what level of confidence can be placed in the occurrence of that event, or in the features of the modeled system? By using 24-h forecast errors to define the initial perturbation, the large, storm-scale ensemble provides an estimate of the predictability of an MCS in the day 2 time frame.

The use of 2D simulations allows for very large (i.e., 100-member) ensembles. The 2D simulations with current 24-h forecast errors suggest that a day 2 MCS forecast should have a success rate around 70% (judged by a generous criterion for identification as an MCS), or slightly higher if there is a deep shear layer present. This success rate drops below 50% as the basic state relative humidity at the cloud-layer is reduced below 65%. Increases in predictability resulting from reductions in forecast errors are not uniform. For example, halving the wind speed errors increases the success rate for both the 500 and 800 J kg^{-1} ensembles, but it only has a minor effect on the 1000 J kg^{-1} ensemble. Reducing the initial errors by half for each of the fields examined here (CAPE, wind speed, and relative humidity) raises the MCS success mode rate to 90%. That is, achieving 90% confidence in forecasts of MCS development requires the tremendous advance of bringing 24-h forecast errors down to the level current analysis errors. This radical improvement in forecasting would yield a 10-20% reduction in the uncertainty for maximum updraft strength and a 25-35% reduction in the uncertainty for maximum surface wind.

The success rate, along with the uncertainty in the maximum updraft strength and maximum surface wind, are all more sensitive to changes in the relative humidity perturbations than changes in the wind speed perturbations, while the opposite is true for the MCS size. The success rate and maximum surface wind are more sensitive to changes in the relative humidity and wind speed perturbations for the ensembles with smaller CAPE perturbations. In contrast, the MCS size is more sensitive to changes in these perturbation fields for the ensembles with larger CAPE perturbations. For the maximum updraft strength the impact of the wind speed and relative humidity perturbations is greater for the large CAPE perturbations early in the runs, while the opposite is true toward the end of the simulations. The CAPE perturbations themselves have the biggest impact for the maximum surface wind, but have no impact on uncertainty in MCS size and only moderate impact on the success rate. Thus, it does not seem that improvements in the forecast of a particular field (e.g., CAPE vs. wind speed) increase confidence in MCS forecasts, generally.

In many ways, the 2D experiments provide reasonable estimates of the 3D MCS predictability. The success rate, along with the spread of the maximum updraft and maximum surface wind speed, are comparable between the two experiments. (Recall that the base-state sounding was adjusted in the 3D experiment to match the 2D success rate for the r823 ensemble, but the improvement attained by reducing the initial uncertainty in 3D is comparable to that improvement seen in the 2D experiment.) Closer inspection of the 3D results uncovers some intriguing and not easily explained results. For example, in the 2D results there are definite sensitivities to the strength of the initial uncertainty in the three perturbed fields, such as a steady increase in agreement among ensemble members

as the CAPE uncertainty is reduced. In contrast, no such patterns emerge in the 3D runs; all ensembles tend to group together, with the notable exception of the r511 ensemble, for which the initial uncertainty is reduced for all fields.

The spread of maximum instantaneous rain rates behaves similarly to the other fields, except that the standard deviation of the r521 ensemble sharply declines after $t = 3$ h. The forecast uncertainty decreases for all ensembles over the last several hours, but the r521 ensemble drops by half, eventually reaching the same level as the r511 ensemble. Hourly rainfall begins to reveal some sensitivity to the initial uncertainty. For the 24-h forecast error initial CAPE uncertainty group, at least, the curves tend to separate according the strength of the wind speed uncertainty. This tendency is much stronger when the correspondence ratio is used to measure forecast spread. Maps of the correspondence ratio reveal that the separation according to initial wind speed uncertainty is caused by greater variation in the translation speed of the runs based on the 24-forecast errors. The spread in the location of the MCSs increases the union of all members, while, for the lower rainfall amounts, the intersection of the members does not change substantially, with the result being lower CRs.

Reducing the 24-h forecast errors to analysis error levels yields a somewhat smaller improvement in the 3D MCS success rate (80-85%) compared with the 2D results (90%), but substantial improvements in the uncertainty of maximum updraft strength (up to 33%) and maximum surface wind (up to nearly 75%). Similarly robust improvements are found for rainfall forecasts, as well. So, while the improvement from reducing the initial uncertainty in the forecast of the development of the MCS itself is not as great in 3D as in 2D, the improvement in forecasts of various characteristics of the system is

much greater in 3D. However, these substantial improvements are only achieved by reducing 24-h forecast errors to the level of current analysis errors—a formidable task.

Several of the 3D runs obtain a structure resembling bow echoes. The bow-like solutions appear to have a somewhat more pronounced descending rear-to-front flow and consequent positive surface perturbation pressure field, similar to a mesohigh. Whereas bow echoes are often associated with damaging winds, the surface winds in the bow-like solutions are no stronger than the non-bowing solutions. Moreover, there is no clear signal in the initial perturbation profiles that could aid a forecaster in predicting the existence of bow echoes.

It is important to emphasize the sensitivity of predictability estimates to the forecast evaluation tool used to derive those estimates. As noted as far back as Kraichnan and Leith (1972), tools designed for continuous fields are not suitable for highly intermittent phenomena. Thus, a measure such as the difference total energy shows very little error growth for the MCS modeled herein, suggesting extended predictability. In contrast, the success rate and the spread of the forecasts of maximum updraft, surface wind, and rain rates give a much different picture and indicate that even reducing the initial uncertainty by half, to levels on par with current analysis errors, is required to achieve substantial improvements in forecast confidence. In fact, for the 3D runs, no gain results from the reduction of initial errors by any single field, but rather all three fields examined here need to be improved to have an impact on the forecast uncertainty. Using the more restrictive correspondence ratio yields a somewhat different conclusion. Since it is sensitive to spatial dislocations, improving the initial wind speed errors increases the forecast confidence. Forecast users must be mindful of the forecast aspects

to which they are sensitive when examining predictability estimates. For example, a hydrologist concerned about flooding in a relatively small catchment would be more interested in the results from the CR than from the success rate.

There are a number of sources of uncertainty in these results, starting with the choice of 1-km grid spacing. Bryan et al. (2003) have shown that the internal structure of convection cannot be adequately represented with 1 km grid spacing. However, applying the argument of Lorenz (1969) suggests that using grid spacing less than 1 km should only reduce the predictability estimates by allowing for a more rapid upscale error cascade. Moreover, enhanced upscale error growth from the mesoscales in the presence of baroclinic instability, which is not considered here, would only serve to worsen predictability further (Tribbia and Baumhefner 2004).

Perhaps the biggest shortcoming is that convective initiation is nearly guaranteed in this study. Initiation is a major forecast concern and often not predicted well by either operational or research high-resolution models. Including the uncertainty of convective initiation would likely reduce the MCS success modes dramatically. Also, the results derive from a single, rather moist, control sounding. It is seen that reducing the mid-level moisture in the control sounding produces large decreases in the success rate. Similarly, a reduction in the boundary layer moisture of the control sounding from 15 g kg^{-1} to 13 g kg^{-1} results in a significant decline in the success rate (not shown). By using a homogeneous environment we have removed the problem of correctly forecasting the downstream environment into which the MCS is moving, which can play a decisive role in determining whether the system persists.

As stated previously, predictability is inextricably linked with the models from which the estimates are derived. Still, while the environment used in this study, e.g., unidirectional shear oriented perpendicular to the convective line, is expected to produce a squall line more two-dimensional in character, the resulting MCSs examined here display a large degree of along-line asymmetry. It would thus appear that fully three-dimensional turbulence is able to develop in the simulations. Adding additional complexity, such as removing the periodic boundaries in the *y* direction, does not seem likely to reduce forecast uncertainty.

Whereas the impact of each of these sources of uncertainty cannot be predicted, it would seem that, taken together, they suggest that the estimates obtained herein (e.g., 70% success rate for a day 2 forecast) are optimistic. That is, they represent a very optimistic estimate of an upper bound on the predictability of MCS forecasts.

Finally, let us return to the question presented earlier, how much confidence can a forecaster have in numerical model forecast of a MCS in the day 2 time frame? With a 70% success rate, fully 30% of MCS events would be missed. Added to that are false alarms, and while this research cannot assess false alarms directly, the 2D experiment does provide some indication. When the maximum cloud layer relative humidity is reduced to 50%, the control run is unable to produce an MCS, but roughly half of the ensemble members did maintain a system, suggesting a large false alarm rate. Assuming that a MCS is correctly forecast, there is still a question of severe potential. The National Weather Service severe criterion for surface winds is 26 m s^{-1} (58 mph), meaning that the maximum surface wind uncertainty is nearly one-third of the severe threshold. In other words, the modeled wind field must be well over the severe threshold to have strong

confidence that severe winds will occur—a model forecast as strong as 35 m s^{-1} (78 mph) would still have nearly a one-in-six likelihood of being a bust. In summary, high-resolution model forecasts are appealing for their ability to produce realistic-looking fields, but the results herein suggest that there is a substantial gap between the ability to produce a forecast that look like something that could happen and the ability to produce a forecast that accurately captures what will, in fact, happen.

Bibliography

Anthes, R.A., 1986: The general question of predictability. *Mesoscale Meteorology and Forecasting,* P.S. Ray, Ed., Amer. Meteor. Soc., 646-656.

Basdevant, C., B. Legras, R. Sadourny, and M. Beland, 1981: A study of barotropic model flows: Intermittency, waves and predictability. *J. Atmos. Sci.,* **38,** 2305-2326.

Benjamin, S.G., G.A. Grell, J.M. Brown, T.G. Smirnova, and R. Bleck, 2004: Mesoscale Weather Prediction with the RUC Hybrid Isentropic–Terrain-Following Coordinate Model. *Mon. Wea. Rev.,* **132,** 473–494.

____, D. Dévényi, S.S. Weygandt, K.J. Brundage, J.M. Brown, G.A. Grell, D. Kim, B.E. Schwartz, T.G. Smirnova, T.L. Smith, and G.S. Manikin, 2004: An Hourly Assimilation–Forecast Cycle: The RUC. *Mon. Wea. Rev.,* **132,** 495–518.

Berri, G.J., and J. Paegle, 1990: Sensitivity of local predictions to initial conditions. *J. Appl. Meteor.,* **29,** 256-267.

Black, T.L., 1994: The new NMC mesoscale Eta model: Description and forecast examples. *Mon. Wea. Rev.,* **9,** 265-278.

Boffetta, G., P. Giuliani, G. Paladin, and A. Vulpiani, 1998: An extension of the Lyapunov analysis for the predictability problem. *J. Atmos. Sci.,* **55,** 3409-3416.

Boucher, R.J., and R. Wexler, 1961: The motion and predictability of precipitation lines. *J. Meteor.,* **18,** 160-171.

Brooks, H.E., C.A. Doswell III, and L.J. Wicker, 1993: STORMTIPE: A forecasting experiment using a three-dimensional cloud model. *Wea. Forecasting,* **8,** 352-362.

Bryan, G.H., J.C. Wyngaard, and J.M. Fritsch, 2003: Resolution requirements for the simulation of deep moist convection. *Mon. Wea. Rev.,* **131,** 2394-2416.

Bukovsky, M.S., J.S. Kain, and M.E. Baldwin, 2006: Bowing convective systems in a popular operational model: Are they real? *Wea. Forecasting,* **21,** 307-324.

Carbone, R.E., J.W. Conway, N.A. Crook, and M.W. Moncrieff, 1990: The generation and propagation of a nocturnal squall line. Part I: Observations and implications for mesoscale predictability. *Mon. Wea. Rev.,* **118,** 26-49.

____, J.D. Tuttle, D.A. Ahijevych, and S.B. Trier, 2002: Inferences of predictability associated with warm season precipitation episodes. *J. Atmos. Sci.,* **59,** 2033-2455.

Clark, D.B., C.M. Taylor, A.J. Thorpe, R.J. Harding, and M.E. Nicholls, 2003: The influence of spatial variability of boundary-layer moisture on tropical continental squall lines. *Quart. J. Roy. Meteor. Soc.,* **129,** 1101-1121.

Cohen, A.E., M.C. Coniglio, S.F. Corfidi, and S.J. Corfidi, 2007: Discrimination of Mesoscale Convective System Environments Using Sounding Observations. *Wea. Forecasting,* **22,** 1045–1062.

Coniglio, M.C., D.J. Stensrud, and L.J. Wicker, 2006: Affects of upper-level shear on the structure and maintenance of strong quasi-linear mesoscale convective systems. *J. Atmos. Sci.,* **63,** 1231-1252.

____, H. E. Brooks, S. J. Weiss, and S. F. Corfidi, 2007: Forecasting the maintenance of quasi-linear mesoscale convective systems. *Wea. Forecasting,* **22,** 556–570.

Crisanti, A., M.H. Jensen, G. Paladin, and A. Vulipani, 1993: Intermittency and predictability in turbulence. *Phys. Rev. Lett.,* **70,** 166-169.

Crook, N.A., 1996: Sensitivity of moist convection forced b boundary layer processes to low-level thermodynamic fields. *Mon. Wea. Rev.,* **124,** 1767-1785.

Dalcher, A., and E.Kalnay, 1987: Error growth and predictability in operational ECMWF forecasts. *Tellus,* **39A,** 474-491.

Davis, C.A., K.W. Manning, R.E. Carbone, S.B. Trier, and J.B. Tuttle, 2003: Coherence of warm-season continental rainfall in numerical weather prediction models. *Mon. Wea. Rev.,* **131,** 2667-2679.

Deardorff J. W., 1980: Stratocumulus-capped mixed layer derived from a three-dimensional model. Bound.-Layer Meteor., 18, 495–527**.**

Derbyshire, S.H., I. Breau, P. Bechtold, J.-Y. Grandpeix, J.-M., Piriou, J.-L. Redelsperger, and P.M.M. Soares, 2004: Sensitivity of moist convection to environmental humidity. *Aurat. J. Roy. Meteor. Soc.,* **130,** 3055-3079.

Du, J., S.L. Mullen, and F. Sanders, 1997: Short-range ensemble forecasting of quantitative precipitation. *Mon. Wea. Rev.,* **125,** 2427-2459.

Ducrocq, V., D. Ricard, J.-P. Lafore, and F. Orain, 2002: Storm-scale numerical rainfall prediction for five precipitating events over France: On the importance of the initial humidity field. *Wea. Forecasting,* **17,** 1236-1256.

Elmore, K.L., D.J. Stensrud, and K.C. Crawford, 2002: Explicit cloud-scale models for operational forecasts: A note of caution. *Wea. Forecasting,* **17,** 873-884.

Ehrendorfer, M., and R.M. Errico, 1995: Mesoscale predictability and the spectrum of optimal perturbations. *J. Atmos. Sci.,* **52,** 3475-3500.

____, ____, and K.D. Raeder, 1999: Singular-vector perturbation growth in a primitive equation model with moist physics. *J. Atmos. Sci.,* **56,** 1627-1648.

Errico, R., and D. Baumhefner, 1987: Predictability experiments using a high-resolution limited-area model. *Mon. Wea. Rev.,* **115,** 488-504.

Farrell, B.F., 1990: Small error dynamics and the predictability of atmospheric flows. *J. Atmos. Sci.,* **47,** 2409-2416.

Fritsch J. M., and R. E. Carbone, 2004: Improving quantitative precipitation forecasts in the warm season. Bull. Amer. Meteor. Soc., 85, 955–965.

____, R. J. Kane, and C. R. Chelius, 1986: The contribution of mesoscale convective weather systems to the warm season precipitation in the United States. *J. Appl. Meteor.,* **25,** 1333–1345.

Fujita, T. T., 1978: Manual of downburst identification for project Nimrod. Satellite and Mesometeorology Research Paper 156, Dept. of Geophysical Sciences, University of Chicago, 104 pp. [NTIS PB-286048.]

Gerhardt, J., 1963: Mesoscale association of a low-level jet stream with a squall-line—cold-front situation. *J. Appl. Meteor.,* **2,** 49–55.

Germann, U., I. Zawadzki, and B. Turner, 2006: Predictability of precipitation from, continental radar images. Part IV: Limits to predictability. *J. Atmos. Sci.,* **63,** 2092-2108.

Gilmore, M.S., and L.J. Wicker, 1998: The influence of midtropospheric dryness on supercell morphology and evolution. *Mon. Wea. Rev.,* **126,** 943-958.

____, J.M. Straka, and E.K. Rasmussen, 2004: Precipitation uncertainty due to variations in precipitation particle parameters within a simple microphysics scheme. *Mon. Wea. Rev.,* **132,** 2610-2627.

Grabowski, W.W., 2004: An improved framework for superparameterization. *J. Atmos. Sci.,* **61,** 1940-1952.

____, X. Wu, M.W. Moncrieff, and W.D. Hall, 1998: Cloud-Resolving Modeling of Cloud Systems during Phase III of GATE. Part II: Effects of Resolution and the Third Spatial Dimension. *J. Atmos. Sci.,* **55,** 3264–3282.

Hohenegger, C. and C. Schär, 2007a: Atmospheric predictability at synoptic versus cloud-resolving scales. *Bull. Amer. Meteor. Soc.,* **88,** 1783-1793.

____, and ____, 2007b: Predictability and Error Growth Dynamics in Cloud-Resolving Models. *J. Atmos. Sci.,* **64,** 4467–4478.

____, D. Lüthi, and C. Schär, 2006: Predictability mysteries in cloud-resolving models. *Mon. Wea. Rev.,* **134,** 2095-2107.

Houston, A.L., and D. Niyogi, 2007: The Sensitivity of Convective Initiation to the Lapse Rate of the Active Cloud-Bearing Layer. *Mon. Wea. Rev.,* **135,** 3013–3032.

Houze, R.A., B.F. Smull, and P. Dodge, 1990: Mesoscale Organization of Springtime Rainstorms in Oklahoma. *Mon. Wea. Rev.,* **118,** 613–654.

Islam, S., R.L. Bras, and K.A. Emanuel, 1993: Predictability of mesoscale rainfall in the tropics. *J. Appl. Meteor.,* **32,** 297-310.

Jankov, I., and W.A. Gallus, Jr., 2004: MCS rainfall forecast accuracy as a function of large-scale forcing. *Wea. Forecasting,* **19,** 428-439.

Jewett, B.F., and R.B. Wilhelmson, 2006: The Role of Forcing in Cell Morphology and Evolution within Midlatitude Squall Lines. *Mon. Wea. Rev.,* **134,** 3714–3734.

Johnson, D.E., P.K. Chang, and J.M. Straka, 1993: Numerical simulations of the 2 August 1981 CCOPE supercell storm with and without ice microphysics. *J. Atmos. Sci.,* **32,** 745-759.

Junker N. W., R. S. Schneider, and R. A. Scofield, 1995: The meteorological conditions associated with the great Midwest flood of 1993. Preprints, 14th Conf. on Weather Analysis and Forecasting, Dallas, TX, Amer. Meteor. Soc., (J4)13–(J4)17.

Kessler, E., 1969: On the distribution and continuity of water substance on atmospheric circulation. *Meteor. Monogr.,* **10,** No. 32, 84 pp.

Khairoutdinov, M., and D.A. Randall, 2003: Cloud resolving modeling of the ARM 1997 IOP: Model formulation, results, uncertainties and sensitivities. *J. Atmos. Sci.,* **60,** 607-625.

Kraichnan, R.H., 1967: Inertial ranges in two-dimensional turbulence. *Phys. Fluids,* **10,** 1417-1423.

Laprise, R., M. Ravi Varma, B. Denis, D. Caya, and I. Zawadzki, 2000: Predictability of nested limited-area model. *Mon. Wea. Rev.,* **128,** 4149-4154.

Leith, C.E., 1968: Diffussion approximation for two-dimensional turbulence. *Phys. Fluids,* **11,** 671-673.

____, 1971: Atmospheric Predictability and Two-Dimensional Turbulence. *J. Atmos. Sci.,* **28,** 145–161.

____, R.H. Kraichnan, 1972: Predictability of turbulent flows. *J. Atmos. Sci.,* **29,** 1041-1058.

Lilly, D.K. 1983: Stratified Turbulence and the Mesoscale Variability of the Atmosphere. *J. Atmos. Sci.,* **40,** 749–761.

____, 1990: Numerical prediction of thunderstorms--has its time come? *Q. J. Roy. Meteor. Soc.,* **116,** 779-798.

Lin, Y.-L., R.D. Farley, and H.D. Orville, 1983: Bulk parameterization of the snow field in a cloud model. *J. Appl. Meteor.,* **22,** 1065-1092.

Lorenz, E.N., 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.,* **26,** 636-646.

____, 1969: The predictability of a flow which possesses many scales of motion. *Tellus,* **21,** 289-305.

____, 1982: Atmospheric predictability experiments with a large numerical model. *Tellus,* **34,** 505-513.

____, 2006: Predictability—a problem partly solved. *Predictability of Weather and Climate,* T. Palmer and R. Hagedorn, ed. Cambridge University Press, p. 40-58.

Maddox, R.A., K.W. Howard, D.L. Bartels, and D.M. Rodgers, 1986: Mesoscale convective complexes in midlatitudes. *Mesoscale Meteorology and Forecasting.* P. Ray, Ed., Amer. Meteor. Soc., 390-413.

McAnelly, R.L., J.E. Nachamkin, W.R. Cotton, and M.E. Nicholls, 1997: Upscale Evolution of MCSs: Doppler Radar Analysis and Analytical Investigation. *Mon. Wea. Rev.,* **125,** 1083–1110.

Moeng, C.-H., and Coauthors, 1996: Simulation of a stratocumulus-topped planetary boundary layer: Intercomparison among different numerical codes. *Bull. Amer. Meteor. Soc.,* **77,** 261-278.

Molteni, F., and T.N. Palmer, 1993: Predictability and finite-time instability of the northern winter circulation. *Quart. J. Roy. Meteor. Soc.,* **119,** 269-298.

Murphy, A.H., 1998: The Early History of Probability Forecasts: Some Extensions and Clarifications. *Wea. Forecasting,* **13,** 5–15.

Murzi, M., 2006: "Jules Henri Poincarè (1854-1912). *The Internet Encyclopedia of Philosophy*. http://www.iep.utm.edu/p/poincare.htm.

Nastrom, G.D., and K.S. Gage, 1985: A climatology of atmospheric wavenumber spectra of wind and temperature observed by commercial aircraft. *J. Atmos. Sci.,* **42,** 950-960.

Nutter, P., D. Stensrud, and M. Xue, 2004: Effects of Coarsely Resolved and Temporally Interpolated Lateral Boundary Conditions on the Dispersion of Limited-Area Ensemble Forecasts. *Mon. Wea. Rev.,* **132,** 2358–2377.

Park, S.K., 1999: Nonlinearity and predictability of convective rainfall associated with water vapor perturbations in a numerically simulated storm. *J. Geophys. Res.,* **104,** 31575-31587.

____, and K.K. Droegemeier, 2000: Sensitivity analysis of a 3D convective storm: Implications for variational data assimilation and forecast error. *Mon. Wea. Rev.,* **128,** 140-159.

Parker, M.D., and R.H. Johnson, 2000: Organizational Modes of Midlatitude Mesoscale Convective Systems. *Mon. Wea. Rev.,* **128,** 3413–3436.

Peckham S. E., and L. J. Wicker, 2000: The influence of topography and lower-tropospheric winds on dryline morphology. *Mon. Wea. Rev.,* **128,** 2165–2189.

Petch, J.C., 2004: The predictability of deep convection in cloud-resolving simulations over land. *Quart. J. Roy. Meteor. Soc.,* **130,** 3171-3187.

Poincaré, H., 1908: *Science et Mèthode.* Ernest Flammarion. [English Translation is Science and Method, Dover Publications, 288 pp., 1952.]

Porter, J.M., L.L. Means, J.E. Hovde, and W.B. Chappell, 1955: A synoptic study on the formation of squall lines in the north central United States. *Bull. Amer. Meteor. Soc.,* **36,** 390-396.

Randall, D., M. Khairoutdinov, A. Arakawa, and W. Grabowski, 2003: Breaking the cloud parameterization deadlock. *Bull. Amer. Meteor. Soc.,* **84,** 1547-1564.

Robert, R., and C. Rosier, 2001: Long range predictability of atmospheric flows. *Nonlinear Processes in Geophysics,* **8,** 55-67.

Robinson, G.D., 1967: Some current projects for global meteorological observation and experiment. *Q. J. Roy. Meteor. Soc.,* **93,** 409-418.

Romine G., and R. Wilhelmson, 2002: Numerical investigation of the role of mid-level dryness on tropical mini-supercell behavior. Preprints, *21st Conf. on Severe Local Storms,* San Antonio, TX, Amer. Meteor. Soc., 631–634.

Rotunno, R., 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.

____, J.B. Klemp, and M.L. Weisman, 1988: A theory for strong, long-lived squall lines. *J. Atmos. Sci.,* **45,** 463-485.

Ruelle, D., 1998: "Science et Mèthode by Henri Poincarè" *Nature,* **391,** 760.

Sanders, F., and R.J. Paine, 1975: The structure and thermodynamics of an intense mesoscale convective system in Oklahoma. *J. Atmos. Sci.,* **32,** 1563-1579.

Schneider, T., and S.M. Griffies, 1999: A conceptual framework for predictability studies. *J. Clim.,* **12,** 3133-3155.

Schumacher, R.S., and R.H. Johnson, 2006: Characteristics of U.S. extreme rain events during 1999-2003. *Wea. Forecasting,* **21,** 69-85.

Simmons, A.J., R. Mureau, and T. Petroliagis, 1995: Error growth and estimates of he predictability from the ECMWG forecasting system. *Quart. J. Roy. Met. Soc.,* **121,** 1739-1771.

Skamarock, W.C., 2004: Evaluating mesoscale NWP models using kinetic energy spectra. *Mon. Wea. Rev.,* **132.** 3019-3032.

____, M.L. Weisman, and J.B. Klemp, 1994: Three-Dimensional Evolution of Simulated Long-Lived Squall Lines. *J. Atmos. Sci.,* **51,** 2563–2584.

Smith, A.M., G.M. McFarquhar, R.M. Rauber, J.A. Grim, M.S. Timlin, B.F. Jewett, and D.P. Jorgensen, 2008: Microphysical and Thermodynamic Structure and Evolution of the Trailing Stratiform Regions of Mesoscale Convective Systems during BAMEX: Part I: Observations. *Mon. Wea. Rev.,* In Press.

Smith, L.A., 2006: Predictability past, predictability present. *Predictability of Weather and Climate,* T. Palmer and R. Hagedorn, ed. Cambridge University Press, p. 217-250.

Srivastava, R.C., 1987: A model of intense downdrafts driven by the melting and evaporation of precipitation. *J. Atmos. Sci.,* **44,** 2869-2889.

Stensrud, D.J., 1996: Effects of a persistent, midlatitude mesoscale region of convection on the large-scale environment during the warm season. *J. Atmos. Sci.,* **53,** 3503–3527.

____, and J.L. Anderson, 2001: Is midlatitude convection and active or a passive player in producing global circulation patterns? *J. Clim.,* **14,** 2222-2237.

____, and J.M. Fritsch, 1993: Mesoscale Convective Systems in Weakly Forced Large-Scale Environments. Part I: Observations. *Mon. Wea. Rev.,* **121,** 3326–3344.

____, and ____, 1994: Mesoscale convective systems in weakly forced large-scale environments. Part III: Numerical simulations and implications for operational forecasting. Mon. Wea. Rev., 122, 2084–2104.

____, and M.S. Wandishin, 2000: The correspondence ratio in forecast evaluation. *Wea. Forecasting,* **15,** 593–602.

____, and L.J. Wicker, 2004: On the predictability of mesoscale convective systems. Preprints, *AMOS 11th National Conf.,* Brisbane, Australia.

____, H.E. Brooks, J. Du, M.S. Tracton, and E. Rogers, 1999: Using ensembles for short-range forecasting. *Mon. Wea. Rev.,* **127,** 433-446.

Straus, D.M., and J. Shukla, 2005: The known, the unknown, and the unknowable in the predictability of weather. COLA Tech. Rep. 175, 20pp. [Available from the Center for Ocean-Land-Atmosphere Studies, 4041 Powder Mill Rd., Suite 302, Calverton, MD 20705.]

Tan, Z.-M., F. Zhang, R. Rotunno, and C. Snyder, 2004: Mesoscale predictability of moist baroclinic waves: Experiments with parameterized convection. *J. Atmos. Sci.,* **61,** 1794-1804.

Tennekes, H., 1978: Turbulent Flow In Two and Three Dimensions. *Bull. Amer. Meteor. Soc.,* **59,** 22–28.

Tepper, M., 1950: A proposed mechanism of squall lines: The pressure jump line. *J. Atmos. Sci.,* **7,** 21–29.

Thompson, R.L., R. Edwards, J.A. Hart, K.L. Elmore, and P. Markowski, 2003: Close Proximity Soundings within Supercell Environments Obtained from the Rapid Update Cycle. *Wea. Forecasting,* **18,** 1243–1261.

Thompson, P.D., 1957: Uncertainty of initial state as a factor in the predictability of large scale atmospheric flow patterns. *Tellus,* **9,** 275-295.

Tribbia, J.J., and D.P. Baumhefner, 1988: The reliability of improvements in deterministic short-range forecasts in the presence of initial state and model deficiencies. *Mon. Wea. Rev.,* **116,** 2276-2288.

____, and ____, 2004: Scale interactions and atmospheric predictability: an updated perspective. *Mon. Wea. Rev.*, **132**, 703-713.

Van Tuyl, A.H., and R.M. Errico, 1989: Scale interaction and predictability in a mesoscale model. *Mon. Wea. Rev.,* **117,** 495-517.

Vannitsem, S., 2003: Intrinsic error growth in a large-domain Eta regional model. *Mon. Wea. Rev.,* **131,** 2697-2704.

Vasic, S., C.A. Lin, I. Zawadzki, O. Bousquet, and D. Chaumont, 2007: Evaluation of precipitation from numerical weather prediction models and satellites using values retrieved from radars. *Mon. Wea. Rev.,* **135,** 3750-3766.

Vukicevic, T., 1991: Nonlinear and linear evolution of initial forecast errors. *Mon. Wea. Rev.,* **119,** 1602-1611.

____, and R.M. Errico, 1990: The influence of artificial and physical factors upon predictability estimates using a complex limited-area model. *Mon. Wea. Rev.,* **118,** 1460-1482.

Warner, T.T., and D. Keyser, 1983: Some practical insights into the relationship between initial state uncertainty and mesoscale predictability. *Proc. Conf. Predictability of Fluid Motions,* G. Holloway and B.J. West, Eds., La Jolla, California, American Institute of Physics, 271-286.

Walser, A., and C. Schär, 2004: Convection-resolving precipitation forecasting and its predictability in Alpine river cachements. *J. Hydrol.,* **288,** 57-73.

____, D. Lüthi, and C. Schär, 2004: Predictability of precipitation in a cloud-resolving model. *Mon. Wea. Rev.,* **132,** 560-577.

Weisman, M.L., 2001: Bow echoes: A tribute to T.T. Fujita. *Bull. Amer. Meteor. Soc.,* **82,** 97-116.

____, and R.J. Trapp, 2003: Low-level mesovortices within squall lines and bow echoes. Part I: Overview and dependence on environmental shear. *Mon. Wea. Rev.,* **131,** 2779–2803.

____, W.C. Skamarock, and J.B. Klemp, 1997: The resolution dependence of explicitly modeled convective systems. *Mon. Wea. Rev.,* **125,** 527-548.

Weiss, S.J., J.S. Kain, L.J. Wicker, R. Davies-Jones, D.R. Bright, J.J. Levit, G.W. Carbin, and M.E. Baldwin, 2005: Evaluating the skill of daily explicit predictions of mesocyclones in multiple high–resolution WRF model forecasts during the 2005 SPC/NSSL Spring Program. *Preprints, 11th Conf. on Mesoscale Processes,* Albuquerque, Amer. Meteor. Soc., 2M.2

Weygandt, S.S., A. Shapiro, and K.K. Droegemeier, 2002: Retrieval of model initial fields from single-Doppler observations of a supercell thunderstorm. Part II: Thermodynamic retrieval and numerical prediction. *Mon. Wea. Rev.,* **130,** 454-476.

Wicker, L.J., and R.B. Wilhelmson, 1995: Simulation and analysis of tornado development and decay within a three-dimensional supercell thunderstorm. *J. Atmos. Sci.,* **52,** 2675-2703.

____, M.P. Kay, and M.P. Foster, 1997: STORMTIPE-95: Results from a convective storm forecast experiment. *Wea. Forecasting,* **12,** 388-398.

Wiin-Nielsen, A., 1967: On the annual variation and spectral distribution of atmospheric energy. *Tellus,* **19,** 540-559.

Zeng, X., and R.A. Pielke, 1993: Error-growth dynamics and predictability of surface thermally induced atmospheric flow. *J. Atmos. Sci.,* **50,** 2817-2844.

Zhang, D.-L., and J.M. Fritsch, 1986: A case study of the sensitivity of numerical simulation of mesoscale convective systems to varying initial conditions. *Mon. Wea. Rev.,* **114,** 2418-2431.

Zhang, F., 2005: Dynamics and structure of mesoscale error covariance of a winter cyclone estimated through short-range ensemble forecasts. *Mon. Wea. Rev.,* **133,** 2876-2892.

____, A.M. Odins, and J.W. Nielsen-Gammon, 2006: Mesoscale predictability of an extreme warm-season precipitation event. *Wea. Forecasting,* **21,** 149-166.

____, C. Snyder, and R. Rotunno, 2002: Mesoscale predictability of the "surprise" snowstorm of 24-25 January 2000. *Mon. Wea. Rev.,* **130,** 1617-1632.

____, ____, and ____, 2003: Effects of moist convection on mesoscale predictability. *J. Atmos. Sci.,* **60,** 1173-1185.

____, N. Bei, R. Rotunno, C. Snyder, and C.C. Epifanio, 2007: Mesoscale predictability of moist baroclinic waves: Convection-permitting experiments and multistage error growth dynamics. *J. Atmos. Sci.,* **64,** 3579-3594.