UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

SUPERCELL PREDICTABILITY:

EXPLORING ENSEMBLE SENSITIVITY TO

INITIAL CONDITION SPREAD

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

MASTER OF SCIENCE IN METEOROLOGY

By

MONTGOMERY LEE FLORA
Norman, Oklahoma
2017

SUPERCELL PREDICTABILITY:
EXPLORING ENSEMBLE SENSITIVITY TO
INITIAL CONDITION SPREAD

A THESIS APPROVED FOR THE
SCHOOL OF METEOROLOGY

BY

_____
Dr. Corey Potvin, Chair

_____
Dr. Louis Wicker

_____
Dr. Steven Cavallo

_____
Dr. Xuguang Wang

DEDICATION

to

God, my wife Kristina , and my mother Trisha

For

Supporting me as I pursued by dreams

Acknowledgements

As I finish my Master's degree, all that comes to mind is... Oklahoma! If I told my younger self I was moving to Oklahoma, he probably would laugh at me. Nevertheless, if ever I waiver in the understanding of my life's purpose, I think back to Proverbs 3:5-6. "Trust in the Lord with all your heart, and lean not on your own understanding. In all ways acknowledge Him and He will direct your path." The opportunities presented to me throughout my life are directly from God. It is His plan that has been laid out in front of me allowing me to succeed in ways I could never have imagined. One way He executes His will is by placing incredible people in our lives that help us to grow as individuals and encourage us to succeed, even in the face of adversity. With that being the case, multiple people in my life have made the journey that much easier, and I would like to acknowledge them here.

First of all, I would like to acknowledge my beautiful wife, Kristina, who has been my rock. She made the same journey to Oklahoma, leaving both our families behind to support my dream. It is has been tough on her, but her support is unwavering in these trying times.

Before my wife, my mother, Trisha, was my primary support and in many ways that continues to this day. Growing up my mom instilled in me a yearning for education through which a better life could be achieved. Even though the same opportunities were not afforded to me as other students, my mother never allowed our financial or social status to define us. We experienced incredibly difficult times, but we were able to make it through and are stronger now for it. "God promises a safe landing, not a calm passage. If God brings you to it, He will bring you through it."

I would also like to acknowledge my advisor, Corey Potvin. He has been my guide and mentor since my very first scientific project as an REU student and continues to support me as I pursue my Ph.D. He may be a young scientist, but he has the

demeanor and expertise of a veteran. His approaches to research are often very methodical, well-thought out, and speak highly of his character. Ultimately, these are the traits I aspire to acquire as a developing scientist.

Last, but certainly not least, I would like to acknowledge my friend, Shawn. He also made the same journey to Oklahoma This similarity led to a connection that has blossomed into a remarkable friendship (even between my wife and his now fiance!). Having healthy friendships in graduate school is must, and Shawn has been an amazing friend who pushes me to succeed. He also has had to listen to me relentlessly talk about my research. So his patience is commendable.

In all sincerity, Oklahoma has been a friendly environment and the friends I have made at the National Weather Center make it a wonderful place to be.

Table of Contents

List of Figures

Abstract

The sensitivity of full-physics ensemble forecasts of supercells to initial condition (IC) uncertainty is investigated. The motivation for the study largely stems from the NOAA Warn-on-Forecast (WoF) program, where the primary objective is to develop a storm-scale ensemble prediction system that will generate probabilistic guidance for severe weather forecasts and warnings. Motivation also stems from the need to develop a suitable full-physics/real data numerical weather prediction (NWP) framework for studying storm-scale predictability.

Three sets of initial/boundary conditions for our simulations were generated from the real-time NSSL Experimental WoF System for Ensembles (NEWS-e) during the 2016 NOAA Hazardous Weather Testbed Spring Forecasting Experiment. The 3-km NEWS-e analyses were interpolated to a 1-km grid, then integrated for 3 hours. Our primary goal was not to replicate observed supercell evolution, but rather to isolate the effect of IC uncertainty (i.e., perfect-model assumption).

The forecast sensitivity to IC uncertainty was assessed by successively reducing the initial ensemble perturbations to 50%, 25%, and 10% of their original magnitudes. Forecast spread was substantially reduced with decreasing IC uncertainty for all examined supercell features in all three cases. This was especially true for the mid-level mesocyclone and rainfall. Experiments were also run exploring the importance of uncertainty within vs. outside of the storm. In general, reducing intra-storm ensemble perturbations to zero greatly reduced forecast spread early on, while later in the simulations forecasts benefited more from eliminating uncertainty in the storm environment. Of all the supercell features, forecasts of the mid-level mesocyclone benefited the most when the perturbations inside the storm were set to zero. Substantial improvements also occurred in low-level mesocyclone forecasts, suggesting that significantly improving tornado forecast lead times does not necessarily require reducing uncertainty in the storm environment.

Chapter 1: Introduction

1.1   Mesoscale Predictability

"Chaos: When the present determines the future, but the approximate present does not approximately determine the future." - Edward Lorenz

Thompson (1957) is regarded as the first study to introduce the concept of atmospheric predictability. Although it had promising results in the early days of numerical weather prediction (NWP), Lorenz (1963) eventually proposed the existence of a predictability limit due to the non-periodicity of turbulence and atmospheric instabilities. By the end of the decade, Lorenz (1969) quantified the predictability for separate scales of motion using a highly idealized turbulence model (See Fig. 2 from Lorenz 1969, our Fig. 1.1).
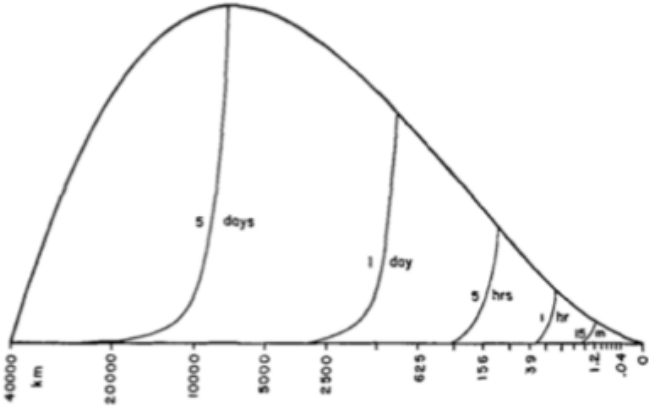
Figure 1.1: From Lorenz (1969), their Fig. 2. Basic energy spectrum (heavy curve), and error-energy spectra (thin curves) at 15 minutes, 1 hour, 5 hours, 1 day, and 5 days, as interpolated from (his) numerical solution in Experiment A. Thin curves coincide with heavy curve, to the right of their intersections with heavy curve. Horizontal coordinate is fourth root of wave length, labeled according to wave length. Resolution intervals are separated by vertical marks at base of diagram. Vertical coordinate is energy per unit logarithm of wave length, divided by fourth root of wave length. Areas are proportional to energy.

In the seminal work by Lorenz (1969), a primary mechanism for imposing finite limits of predictability was the rapid upscale cascade of small-scale, small-amplitude initial errors. The unaccountable disturbances either from the imprecise initial state or unresolved features by the model grew in magnitude to limit the synoptic- and planetary-scale predictability to roughly two weeks. This concept is famously known as the "butterfly effect" (or seagull effect if you ask Lorenz). In that illustration, the unaccountable disturbance from a single butterfly flapping its wings grows with time to impact weather on a global scale. However illustrative the concept is, it neglects the effects of viscosity, which arguably would dampen the disturbance of a single butterfly. In other words, it would take a plethora of unaccounted-for butterflies to disrupt larger-scale flows.

The original scope of Lorenz (1969) focused on larger-scale flows where the experiment limitations were acceptable and observations supported the theory. However, turbulence modeling was no longer applicable to the scope of predictability research shifted towards the mesoscale environment. For instance, continuous isotropic turbulence was no longer a valid assumption given the highly intermittent, organized weather systems present at mesoscales (Anthes et al. 1985). Although more sophisticated turbulence models than that in Lorenz (1969) were used to study predictability (e.g., Leith 1971; Leith and Kraichnan 1972; Metais and Lesieur 1986) it was not until Anthes et al. (1985) that a mesoscale numerical model was applied in a predictability study. In contradiction to previous studies, Anthes et al. (1985) found little to no forecast error growth over a 3-day period. The broader mesoscale predictability community disregards the results of Anthes et al. (1985) since perfect lateral boundary conditions were used (e.g., Vukicevic and Errico 1990). Nevertheless, it did demonstrate the ability of the mesoscale environment to inherit the extended predictability of the larger-scale flow. Inheriting predictability from the larger-scale has been recently supported by Ngan et al. (2009) who speculated that

the enhanced predictability could be related to wave motion. First, they noted that waves have can inhibit the loss of predictability (e.g., Basdevant et al. 1981). With that they argued as gravity waves dominate the domain and energy is transferred into the waves, then larger scales would become more predictable. Ultimately, this argument is tentative and would require further investigation.

Another issue with classic turbulence models diagnosing mesoscale predictability is the lack of moist processes. The effects of moisture are neglected for a simplified analysis, but such a simplification overlooks fundamental thermodynamic processes. In fact, in multiple mesoscale predictability studies, moist convection has been pinpointed as a primary source of limited mesoscale predictability (Ehrendorfer et al. 1999; Walser et al. 2004; Hohenegger et al. 2006; Zhang et al. 2002, 2003, 2006; Spyksma and Bartello 2008). Although Ehrendorfer et al. (1999) was the first to document rapid perturbation growth on the mesoscale Zhang et al. (2002) were the first to demonstrate the rapid perturbation growth at meso-β scales associated with moist processes. Evaluating meso-β-scale predictability in the presence of precipitation, Walser et al. (2004) concluded that convective activity could impact the predictability of scales > 100 km while also confirming that predictability decreased rapidly with progressively smaller scales similar to Lorenz (1969). While investigating the role of mesoscale processes in determining predictability within a cloud-resolving model, Hohenegger et al. (2006) found that diabatic forcing associated with moist dynamics was a prime source of rapid perturbation growth. They also found that significant loss of predictability only occurred over moist convectively unstable regions that were able to sustain energy propagation against the mean flow. In other words, there are regimes where the mean flow is sufficient in magnitude to "sweep" errors associated with gravity waves from convective cells out of the limited area domain. Zhang et al. (2007) generalized mesoscale predictability by using idealized, baroclinic waves to construct a multi-stage error-growth concep-

tual model. The first stage involved the rapid growth and saturation of small-scale, small-amplitude initial condition (IC) errors at convective scales. The second stage was then upscale error growth dictated by geostrophic adjustment and cold pool dynamics with saturation on the timescale of $O(2\pi f^{-1})$.

Multiple studies have demonstrated the rapid upscale error growth (with or without an association with moist dynamics), but recent studies are also finding a downscale error growth from larger scales (Durran and Gingrich 2014; Durran and Weyn 2016; Johnson et al. 2015). Using a modified version of Lorenzs spectral turbulence model, Durran and Gingrich (2014) showed that large-scale errors swamped the energy cascade by the unavoidable, unresolved small-scale errors, a result from Lorenz (1969) that, the authors noted, is often neglected. In an applied setting, Durran and Weyn (2016) found that after 3-4 h, idealized squall lines were more sensitive to small relative errors on scales of 100 km rather than larger relative errors on smaller scales. The large-scale, but unavoidable, errors rapidly transitioned downscale to the smallest resolved features and errors then propagated back upscale as if they had originated from the small scales.

To explain the connection between the upscale and downscale cascade of error/ perturbations to forecast error growth, Sun and Zhang (2016) explored the difference between intrinsic and practical predictability. The familiar type of predictability is intrinsic predictability, which is the longest possible forecast extent given nearly perfect atmosphere estimations and forecast models (e.g., Lorenz 1969; Melhauser and Zhang 2012). Practical predictability is defined as the predictability given current (or near-future) capabilities based on realistic forecast model and ICs errors (e.g., Lorenz 1982; Melhauser and Zhang 2012). Separating predictability into these two broad categories (Sun and Zhang 2016) found that practical and intrinsic predictability was tied to the scale and amplitude of the initial perturbations. Practical predictability was related to sufficiently large initial perturbations in both scale and

4

magnitude where the baroclinic growth was the primary mechanism. In that case, the forecast accuracy improved through the reduction of IC errors, especially those at larger scales, which supports Durran and Gingrich (2014) as well as Durran and Weyn (2016). On the other hand, intrinsic predictability was dictated by small-scale, small-amplitude errors (butterfly effect) associated with unresolved features in the model or the inability to specify the state of the atmosphere with infinite precision. This was in agreement with the original conclusion from Lorenz (1969).

1.2 Storm-Scale Predictability

1.2.1 General

"*The storm starts, when the drops start dropping. When the drops stop dropping then the storm starts stopping.*" - Dr. Seuss

Lewis Fry Richardson was the first to introduce the concept of NWP in his book *Weather Prediction by Numerical Process* published in 1922. In it, Richardson said,

> Perhaps some day in the dim future it will be possible to advance the computations faster than the weather advances and at a cost less than the saving to mankind due to the information gained. But that is a dream.

Fortunately, that dream has come to fruition for the synoptic and mesoscale environment. As for storm-scale NWP, the notion of numerically predicting individual convective cells and thunderstorms in real-time only began in the late 80s/early 90s (e.g., Lilly 1990). This idea then took nearly a decade and a half to manifest into the operational convective-allowing models employed today by National Centers for Environmental Prediction (NCEP). Thus, on the cusp of the century mark after the publication of *Weather Prediction by Numerical Process*, we finally have the capabilities to predict severe thunderstorms numerically.

Before a cloud-scale model is used to forecast explicit evolution and behavior of thunderstorms, the predictability of storms and the robustness (or stability) of the numerical solutions needs to be considered (Brooks et al. 1993). Based on Fig. 1.1, Lorenz (1969) set the initial estimate for storm-scale predictability at approximately $1\,\mathrm{h} \pm 30\,\mathrm{m}$ (i.e., dependent on the scale of the phenomena). Lilly (1990) re-evaluated the theoretical predictability limit using a similar approach to Lorenz (1969), but his estimate was 2 h. However, similar to organized, mesoscale systems, the predictability of storm-scale phenomena is not accurately being assessed by turbulence models (Droegemeier 1997). For example, 1) storms are highly intermittent (similar to organized, mesoscale systems) 2) storm interactions can cause forecast errors in one storm to result in forecast errors in other nearby or subsequent storms, 3) under certain regimes, errors are swept out of the domain whereas for Lorenz (1969), the cyclic boundary conditions recycled errors, and 4) sufficient helicity, which is relevant for supercells, can suppress the downscale transfer of energy (Lilly 1986), leading to enhanced predictability. A caveat of the last point is that naturally occurring supercell helicity may not exceed the 50-60% of maximum helicity to block the energy dissipation (Polifke and Shtilman 1989). As of yet, the impact of helicity on supercell predictability remains unresolved.

Storm Type Operational Research Model Including Predictability Evaluation (STORMTIPE) and STORMTIPE-95 (Brooks et al. 1993; Wicker et al. 1997) were the first studies to diagnose storm-scale predictability using a mesoscale numerical model in an applied framework. Both studies demonstrated the ability of numerical models to provide operational guidance, albeit in a highly idealized framework. Since STORMTIPE(-95) , there have been numerous studies on or related to storm-scale predictability and the predictability of severe thunderstorms (e.g., Park 1999; Elmore et al. 2002; Wandishin et al. 2008, 2010; Clark et al. 2010; Cintineo and Stensrud 2013; Durran and Weyn 2016; Melhauser and Zhang 2012; Wu et al. 2013;

Potvin and Wicker 2013; Potvin and Flora 2015; Zhang et al. 2015, 2016; Miglietta et al. 2016; Potvin et al. 2017).

Studying the predictability of convective rainfall using the ARPS model, Park (1999) found that the predictability limit of rainfall was beyond 140 min, with a primary contribution to total error coming from differences in location rather than amplitude/intensity errors. Elmore et al. (2002) investigated convective-storm lifetimes in an ensemble framework but had a limited focus on the prediction of convection regarding its bulk properties and not on the exact prediction of individual storm structures and evolution. Wandishin et al. (2008, 2010) explored the practical predictability of idealized mesoscale convective system (MCSs) in an ensemble framework. By reducing the IC uncertainty, the model maintained a MCS for a longer duration. They also found that minimizing forecast uncertainty in the individual storm features required reducing all perturbed variables. Consistent with the mesoscale predictability studies by Zhang et al. (2002, 2003) and Lorenz (1969), Clark et al. (2010) found a larger error growth with convection-permitting models ($\Delta x = 4$ km) versus their convection-parameterizing counterparts ($\Delta x = 20$ km). Melhauser and Zhang (2012) revealed diminishing returns when reducing the initial perturbation, meaning that the reduction of initial perturbations did not bring a corresponding linear decrease in forecast errors, which is contrary to results of the squall-line study of Wu et al. (2013), where a more linear decrease was present.

1.2.2 Supercell Predictability

Supercell thunderstorms can produce severe weather hazards such as large hail, $\geq$EF2 tornadoes, flash flooding and damaging straight-line winds. In fact, supercells produce a disproportionate amount of these life- and property-threatening severe weather hazards although they are the least common convective mode (Duda and

Gallus 2010). Even though supercells are regarded as prolific tornado and large hail producers, they are extreme precipitation producers (Hitchens and Brooks 2013). Given that supercells present such a threat, there has been some work related to the practical and intrinsic predictability of these storms. Predictability studies can also include sensitivity of forecasts to model configuration such as grid spacing (e.g., Potvin and Flora 2015) and other aspects of the ICs such as the resolution (e.g., Potvin et al. 2017), both of which will be discussed.

Cintineo and Stensrud (2013) was the first study to explore supercell predictability. To examine the upper limits of practical supercell predictability, they used Rapid Update Cycle (RUC) model forecast errors at 1-, 2-, and 3-h prior to convective initiation (CI), to perturb an initial sounding. Cintineo and Stensrud (2013) found storm location based on 40-dBZ was predictable out to 2 h, but mid-level mesocyclone location was only predictable out to 40 min, while 5-min heavy rainfall location and cold pool area were virtually unpredictable. As pointed out in Cintineo and Stensrud (2013), the use of idealized settings was a limitation of the work. Although modeling limitations were present, another limitation was the stringent definition of predictability, where times after the domain-max frequency for a given feature (e.g., mid-level mesocyclone, rainfall, etc.) was less than 60% were deemed unpredictable. Another possible limitation of the study was their stringent definition of predictability. For Cintineo and Stensrud (2013), unpredictability was determined when the domain-max frequency (i.e., similar to max probability of exceedance) for a given feature (e.g., mid-level mesocyclone) was less than 60%. The problem lies with the assumption that the numerical output no longer has value after the ensemble loses majority agreement. As a counter-example, ensemble agreement less 60 % on tornado potential at a 1-h lead time is valuable information since other sources of information (e.g., climatology or human forecasters) may not be as certain at that lead time.

Over the past few years, there has been a series of studies on supercell predictability guided by the Warn-on-Forecast initiative (See section 1.3) (Potvin and Wicker 2013; Potvin and Flora 2015; Potvin et al. 2017). In an idealized observing system simulation experiment (OSSE) framework, Potvin and Wicker (2013) found that low-level rotation probabilistic guidance was possible out to 30-60 min. Examining the sensitivity of individual supercell features evolution to grid spacing, Potvin and Flora (2015) concluded that 4-km grid spacing was too coarse to reliably simulate supercells whereas 1-km grid spacing sufficiently resolved most of the features. Finally, exploring the sensitivity of supercell evolution to IC resolution, Potvin et al. (2017) found the model was able to rapidly regenerate missing scales within 10 - 20 minutes. A corollary of this result is that the smaller-scale processes are heavily determined by the larger-scale flow.

Recently, two studies explored the practical and intrinsic predictability of the 20 May 2013 tornadic supercells (Zhang et al. 2015, 2016). In Zhang et al. (2015), using perturbed deterministic ICs to generate an ensemble (similar to Cintineo and Stensrud 2013), they explored the practical predictability of the event and concluded the diurnal planetary boundary layer (PBL) evolution, synoptic timing and low-level convergence influenced by southwest OK topography strongly modulated CI. Zhang et al. (2016) used much smaller magnitude perturbation than Zhang et al. (2015) and found the intrinsic predictability of the event to be 3-6 h. They also established a multistage conceptual supercell predictability model where small environment instabilities propagate and evolve via PBL turbulence, and then rapidly amplify in moist convective processes through a positive feedback mechanism with the updraft, phase change, and the cold pool. Another recent study by Miglietta et al. (2016) focused on the predictability of a supercell that formed over northeastern Italy where orographic effects can also play a significant role. In this particular case, the rainfall in a supercell was found to be relatively sensitive to different large-scale forcing and

initialization times than boundary layer parameterization.

A majority of past supercell predictability studies have been evaluated in an ensemble framework. However, their ensembles were not generated by an ensemble-based data assimilation (DA) method (e.g., Ensemble Kalman Filter). In such a framework, the IC perturbations become more flow-dependent rather than introduced randomly in the environment (e.g., Cintineo and Stensrud 2013; Zhang et al. 2015, 2016). Therefore, this study will be adopting an ensemble-based DA framework.

Snyder and Zhang (2003) and subsequent studies (Dowell et al. 2004; Tong and Xue 2005; Aksoy et al. 2009, 2010; Yussouf and Stensrud 2010; Dowell et al. 2011; Yussouf et al. 2013a) assimilated radar data using an ensemble Kalman filter (EnKF) approach, paving the way for storm-scale predictability experiments focused on observed thunderstorms and their associated hazards. In fact, the assimilation of satellite, radar, and in situ observations via an ensemble-based method (e.g., EnKF) coupled with adequate computing resources may yet prove helpful in providing severe weather warning guidance. Therefore, in addition to adapting an ensemble-based DA framework, this study will also include radar and satellite observations for more realistic ICs compared to past studies (Cintineo and Stensrud 2013; Zhang et al. 2015, 2016).

1.3   Warn-on-Forecast

"Whatever may be the progress of science, never will observers who are trustworthy, and careful of their reputation, venture to foretell the state of the weather"

- Francois Arago, 19th century French Mathematician

"But who wants to be foretold the weather? It is bad enough when it comes, without our having the misery of knowing about it before hand"

- Jerome K. Jerome, Three Men in a Boat

Successful forecasting of convective storms remains one of the most challenging problems of NWP today. Forecasting severe convective thunderstorms and their associated hazards (e.g., wind gusts, torrential rain, hail and sometimes tornadoes) is a crucial task since they present a serious threat to human lives and property. Since the 1950s, the United States National Weather Service (NWS) has issued hazardous weather warnings based on radar observations, spotters' reports or when an impending hazard is deemed imminent by NWS forecasters knowledge of the storm environment (e.g., Coleman et al. 2011; Brotzge and Donner 2013). Issuing warnings based on those factors is informally known as the warn-on-detection paradigm. Radar and satellite observations are heavily relied upon when nowcasting. Furthermore, establishing a nationwide Doppler weather radar network has increased tornado warning lead-time from 3 min in 1978 to 14 min in 2011 primarily by increasing the probability of detection (POD) (Stensrud et al. 2013). However, when Stensrud et al. (2013) further analyzed tornado warnings lead-times between 1986 and 2006, the estimate was closer to 17 min (see their Fig. 1, our Fig. 1.2). Unfortunately, tornado warning lead times have decreased recently with estimates at 8 min in 2015 (Samenow 2017). This may be attributed to NWS shift in philosophy towards reducing false alarms, which delays warnings issuance. Extending the current tornado warning lead-time using the current paradigm may be difficult. Although an increasing number of tools and storm-scale observations are becoming available in the "warn-on-detection" paradigm, warning lead times are primarily a function of human forecasters. This is not to say that human forecasters hinder warning lead-times. In fact, using conceptual models of severe weather has been fairly successful up to this point. However, the success of human forecasters (and the warn-on-detection paradigm) at longer lead times (e.g. beyond 30-60 min) remains highly in question.
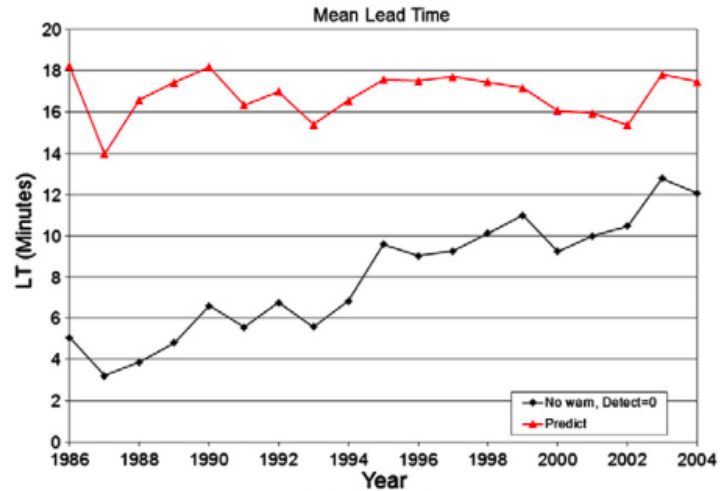
Figure 1.2: From Stensrud et al (2013), their Fig. 1. Annual, national mean lead time for tornado warnings issued by the US National Weather Service. Red line considers only those warnings issued prior to occurrence of a tornado in the warned area. Black line assigns a leadtime of 0 for any tornado that does not have a warning issued before the tornado occurs.

In recent years, observation platforms, numerical weather prediction (NWP) models, data assimilation algorithms, and computational resources have progressed considerably. It is becoming increasingly possible to incorporate satellite, radar, and in situ observations via an ensemble-based data assimilation method (e.g., EnKF) in real-time. This allows for the generation of more realistic IC for NWP models, which have proven to be helpful in providing severe weather warning guidance (Roebber et al. 2004; Stensrud et al. 2009, 2013). Thus, there hopes to be a transition from warn-on-detection to "warn-on-forecast" (WoF), where numerical guidance plays a more crucial role in the severe weather warning process by significantly extending warning lead times (Stensrud et al. 2009, 2013). Specifically, the vision is for a continuously updated, ensemble-based data assimilation and probabilistic forecast system that can provide quantitative information regarding the evolution of severe thunderstorm events on a 1-3 h time scale, potentially helping operational forecasters improve warning lead times. To be explicit, the system will not be limited to just

12

improving tornado lead times. Ideally, such a system can improve warning lead-times for all severe weather hazards (e.g., flash floods, severe winds, and hail).

Previous studies related to WoF have been with regard to the configuration of the ensemble prediction system, such as choice of parameterizations (Wheatley et al. 2014; Yussouf et al. 2013b), grid spacing (e.g., Potvin and Flora 2015), and storm-scale data assimilation methods (Gao et al. 2013; Yussouf et al. 2015; Potvin et al. 2017). Several studies have also examined the ability of a WoF-type system to assimilate observed tornadic supercells and provide 0-1 h probabilistic numerical forecasts of low-level vertical vorticity and/or updraft helicity, which are proxies for tornadoes in convective-allowing models (Dawson et al. 2012; Yussouf et al. 2013b,a, 2015, 2016; Potvin and Wicker 2013; Wheatley et al. 2015). Dawson et al. (2012) and Potvin and Wicker (2013) conducted experiments with horizontally homogeneous ICs and generated probabilistic forecasts of low-level rotation of supercell storms. Both studies concluded that short-range probabilistic forecasts of low-level rotation could be achieved with reasonable accuracy. Yussouf et al. (2013a, 2015) demonstrated the capability of the WoF-type system to provide relatively accurate estimates of intense low-level mesocyclone tracks that align well with the locations of radar-derived rotation tracks associated with the observed tornadic storm. In particular, Yussouf et al. (2013b) found that an improved representation of mesoscale heterogeneity in the near-storm environment produced more accurate EnKF analyses of tornadic supercell thunderstorms and improved probabilistic forecasts of low-level rotation. As for Yussouf et al. (2016), they concluded that both low-level rotation and rainfall probabilistic forecasting are possible through a WoF-type system. For instance, they found that rainfall probabilities agreed fairly well with observed rainfall and the greatest source of error after 6 h was due to phase errors, similar to Park (1999).

Recently, the WoF project has been implemented in real-time during the Spring

Forecast Experiments as the NSSL Experimental WoF System for ensembles (NEWS-e; Wheatley et al. 2015; Jones et al. 2016). NEWS-e is a 36 member ensemble generated using the Ensemble Adjustment Kalman Filter (EAKF) through the Data Assimilation Research Testbed (DART) software (Kalman 1960; Anderson 2001; Anderson et al. 2009). Radar, satellite, mesonet (when available), and other conventional observations are assimilated. The NEWS-e uses the WRF model on a 3-km grid one-way nested in a 15-km parent grid. Initial and boundary conditions for the parent and nested grids are downscaled from the 21 member 0000 UTC Global Ensemble Forecast System (GEFS) forecast cycle. Physics diversity (except for microphysics) is one technique used to increase the number of ensemble members from the GEFS to the NEWS-e system (i.e., some ensemble members are repeated, but with different physics). Another technique for generating additional ensembles includes using perturbations methods from DART. Wheatley et al. (2015) provides a thorough description of the NEWS-e and for additional details see their paper.

Using the NEWS-e during Spring 2013 and 2014, Wheatley et al. (2015) found the probabilistic ensemble forecasts were able to produce areas of intense low-level rotation approximately 30 min before the first observed tornado in cases of supercells and MCSs. Jones et al. (2016) assimilated satellite data (e.g., liquid and ice water path) in the NEWS-e, which was not included in Wheatley et al. (2015). Including the clear-sky satellite data, Jones et al. (2016) found that it reduced anomalous cloud cover and improved thermodynamics conditions leading to higher probabilistic forecasts of strong low-level rotation that corresponds well with observed tornado tracks. However, for the best results, Jones et al. (2016) argued that it might be necessary to combine clear-air reflectivity and satellite retrievals.

Chapter 2: Methods

2.1 Model Description

The numerical model used in all experiments was the Advanced Research Weather and Forecasting model (ARW-WRF; Skamarock et al. 2008). There were two domains with a 1-km grid one-way nested into a 3-km grid. The number of horizontal grid points in the inner domain varied from case to case, but the outer domain was 250×250 with a case-dependent position. There were 51 vertical levels with 11 in the lowest 2 km above ground level (AGL). A 1-km grid spacing has been used in past supercell predictability studies (e.g., Cintineo and Stensrud 2013; Zhang et al. 2015, 2016; Miglietta et al. 2016) and was a reasonable compromise between available computing resources and the ability of the model to represent essential physical processes relevant to supercell prediction such as the mid- and low-level mesocyclone. As further support, Potvin and Flora (2015) found a 1-km grid was sufficient for resolving the mid-level mesocyclone.

The initial and lateral boundary conditions were provided by the NEWS-e (Wheatley et al. 2015; Jones et al. 2016) experiments run during the 2016 Hazardous Weather Testbed Spring Forecasting Experiment. All simulations were integrated for 3 h with model output every 5 min. Physical parameterization schemes applied included Thompson microphysics (Thompson et al. 2008), MM5 similarity surface layer (Zhang and Anthes 1982), RUC Land Surface Model (Smirnova et al. 1997, 2000) , Mellor-Yamada-Nakanishi-Niino (MYNN) level-3.0 PBL (Nakanishi and Niino 2006), RRTM longwave radiation (Mlawer et al. 1997), and Dudhia shortwave radiation scheme (Dudhia 1989). The Thompson microphysics scheme calculates the radar reflectivity using cloud liquid water, rainwater, cloud ice, snow and graupel. Since mesoscale and storm scale predictability is greatly affected by moist convection (see section 1.1), a more sophisticated microphysics scheme may be preferable. However, Wang et al. (2012) found that on WoF timescales (e.g., 1-2 h), the differ-

ences in average perturbation evolution (i.e., both in amplitude and spatial pattern) were nominal when comparing simple (e.g., liquid only) vs. complex microphysics parameterizations. Thus, the 1.5 moment Thompson microphysics is sufficient for this study.

2.2   Cases

Given that the sensitivity of supercell evolution to IC spread may be case-dependent, multiple cases were used to evaluate the sensitivity of the results to different environments. The three dates and times were: 9 May, 17 May, and 25 May initialized at 2200, 0100, and 0000 UTC, respectively. Since 0100 and 0000 UTC were 8:00 and 7:00 pm CST, these two cases are referred to as 16 and 24 May. A brief description of each case is presented below. In addition, various bulk parameters of the initial storm inflow environment for all three cases is provided in Table 2.1

	May 9	May 16	May 24
CAPE (J Kg^{-1})	1800	1500-1800	1000-1500
CIN (J Kg^{-1})	-75	-200	-100
0-1 km SHR(ms^{-1})	10	11-14	11-14
0-6 km SHR(ms^{-1})	< 20	36	36-39
0-1km SRH(m^{2}s^{-2})	120	200-250	200
0-3 km SRH(m^{2}s^{-2})	250-280	450	> 500
SFC Td (^{o}F)	60-64	56-64	45-48

Table 2.1: For all three cases, the various bulk parameters of the storm inflow environment on the 3-km grid at t = 0.

2.2.1   9 May 2016

Multiple supercells formed in the early afternoon on 9 May 2016 over OK with two distinct regions in North Central and southeast OK (see Fig. 2.1). This study focused on the southern, isolated supercell of the northern storm cluster, which initiated in SW Noble County, OK around 2100-2200 UTC. The observed supercell produced one severe wind report and a single EF2 tornado. The storms were associated with a negatively tilted trough passing OK and TX with low-level convergence driven by the dryline. There was a low-level southwesterly jet advecting moisture northward from the Gulf of Mexico. Overall, the deep- and low-level shear was relatively low on the supercell spectrum, but there was sufficient moisture and static instability (see Table 2.1 and Fig. 2.1).

2.2.2   16 May 2016

A storm cluster was persistent over N TX and OK panhandle with initiation around 1900 UTC. They developed near the convergence of a weak stationary front and the dryline. An upper-level shortwave over KS provided upper-level support. The cluster was also being fed moisture by a strong low-level jet, similar to 9 May case. The storm of interest was an embedded supercell formed over the OK panhandle producing several severe wind and hail reports as well as multiple, weak tornadoes (e.g., EF0). Overall, there was sufficient moisture and unlike the 9 May case, the low-level shear was fairly strong (see Table 2.1 and Fig. 2.2).

2.2.3   24 May 2016

A storm cluster initiated over north central CO around 1900 UTC associated with a cold front and weak surface low (e.g., 1003 hPa) with a broad negatively tilted trough over the western US that extended over the Pacific Ocean (see Fig. 2.3). The

17
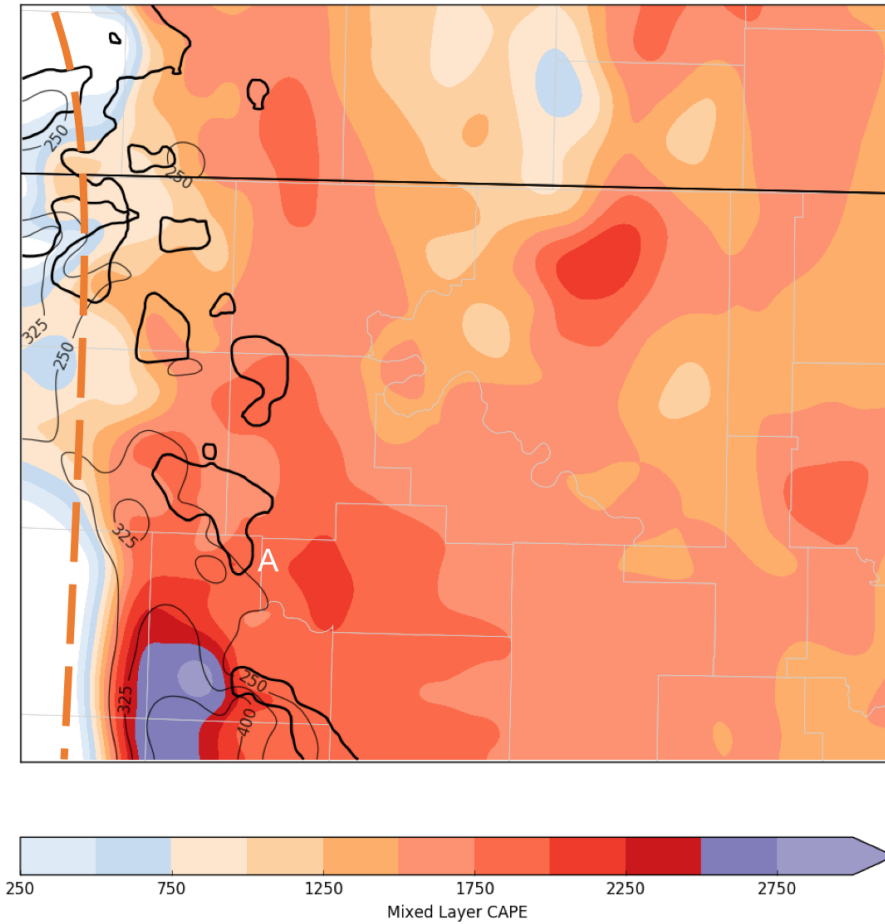
Figure 2.1: Mixed layer (lowest 75 hPa) CAPE (filled contour), 0-3 km storm-relative helicity ($m^2 s^{-2}$; thin lines), and reflectivity at 1.5 km AGL at 30 dBZ (thick black lines) are shown. The letter "A" denotes storm inflow region for the primary storm of interest. Approximate position of dryline shown with dashed brown curve.

storm of interest was a supercell on the southern edge of the cluster that produced multiple severe wind and hail reports and two tornadoes. The environment was noticeably drier than the other two cases, but had substantially stronger low- and deep-level wind shear (see Table 2.1).
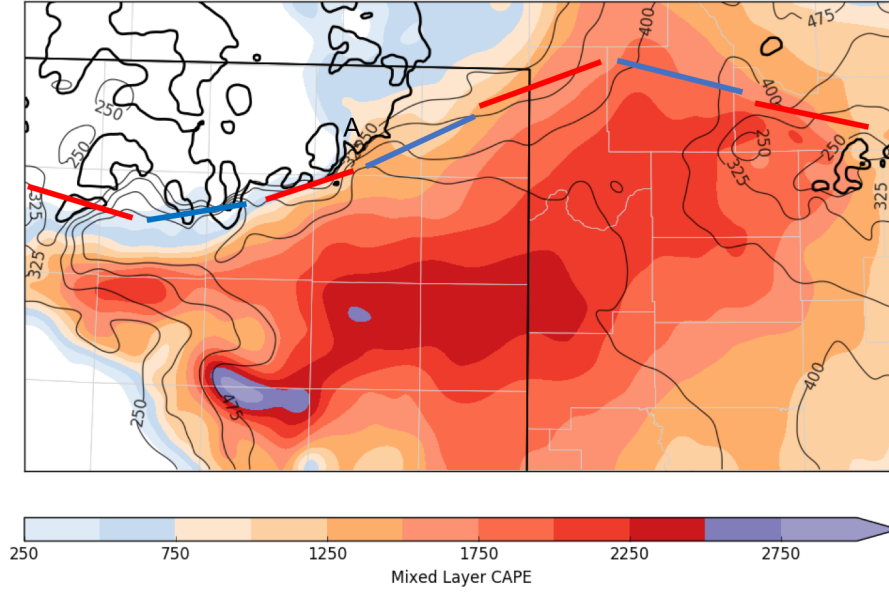
Figure 2.2: Same as Fig. 2.1. Approximate position of the stationary front shown with a red and blue dashed curve.

2.3 Experiment Set-up

2.3.1 Reducing Initial Ensemble Spread

Instead of introducing random perturbations at prescribed wavelengths or randomly in the environment as seen in past studies (Cintineo and Stensrud 2013; Zhang et al. 2015, 2016), flow-dependent perturbations were used. Using the original 3-km ensemble, a control member was chosen based on the smallest deviation from the ensemble mean for the initial state as determined by temperature, wind, and water vapor mixing ratio. Furthermore, even though explicitly comparing against observations was not a primary focus of this work, the ensemble mean in each case compared fairly well with the observed supercell (not shown). Therefore, another, but more subjective, criteria involved comparing potential control members against the early ensemble mean evolution (e.g., first 30 min). Past studies (e.g., Zhang et al. 2015) have initialized with the ensemble mean but in their case it was prior to CI. In this case, the ensemble mean itself was not used as the control since initializing with
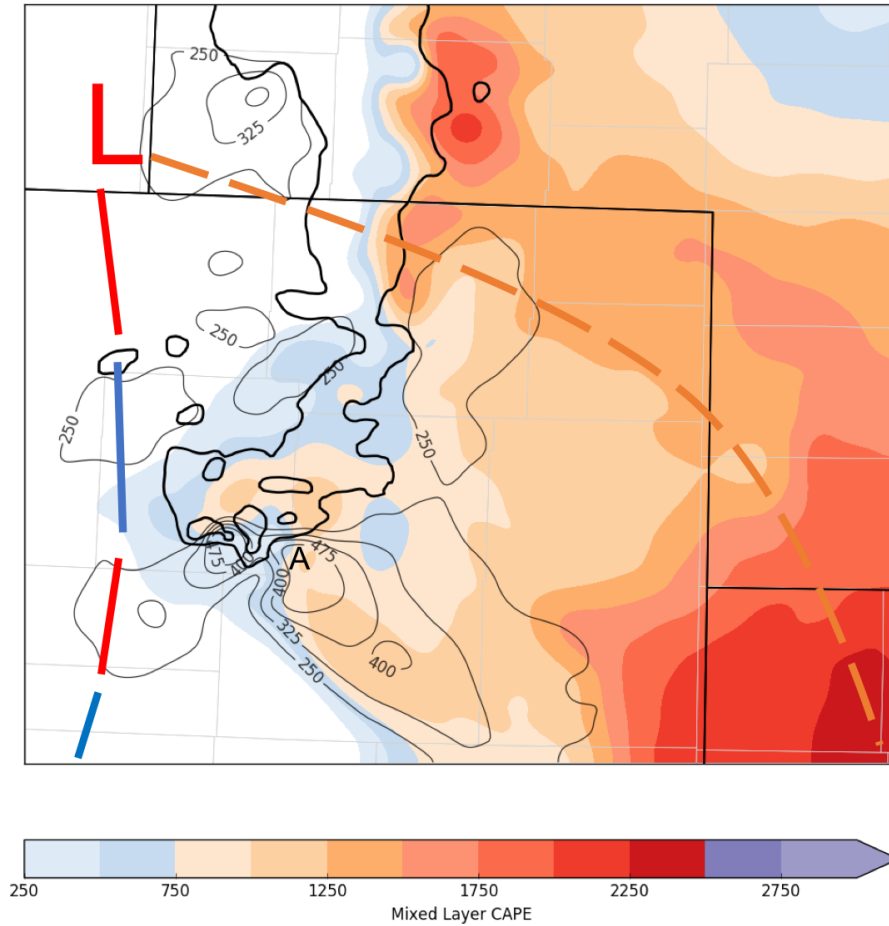
Figure 2.3: Same as Fig. 2.1. Approximate position of low pressure center, dryline, and stationary front are shown.

smoothed storms fields was deemed inappropriate. As a note, the naming convention when referring to the control member will be as such: May9_CNTL. As a reference, the reflectivity for all three control members at t=30,90, and 150 min are presented in Fig. 2.4. For 9 May (Fig. 2.4a-c), there was a right-moving supercell that was surrounded by additional, secondary storms early on and eventually became isolated as it decayed. By secondary storms, it refers to the additional storms in the domain near the primary storm of interest (e.g., may or may not be supercellular). For 16 May (Fig. 2.4d-f), there was a large supercell that was transitioning into an MCS. However, there is a clear hook signature, which is masked at times by convection

initiated off the gust front. Finally, for 24 May, there was a completely isolated right-moving supercell with a strong hook signature (Fig. 2.4g-i). There were secondary storms in this case, but they were generated from the primary storm of interest.
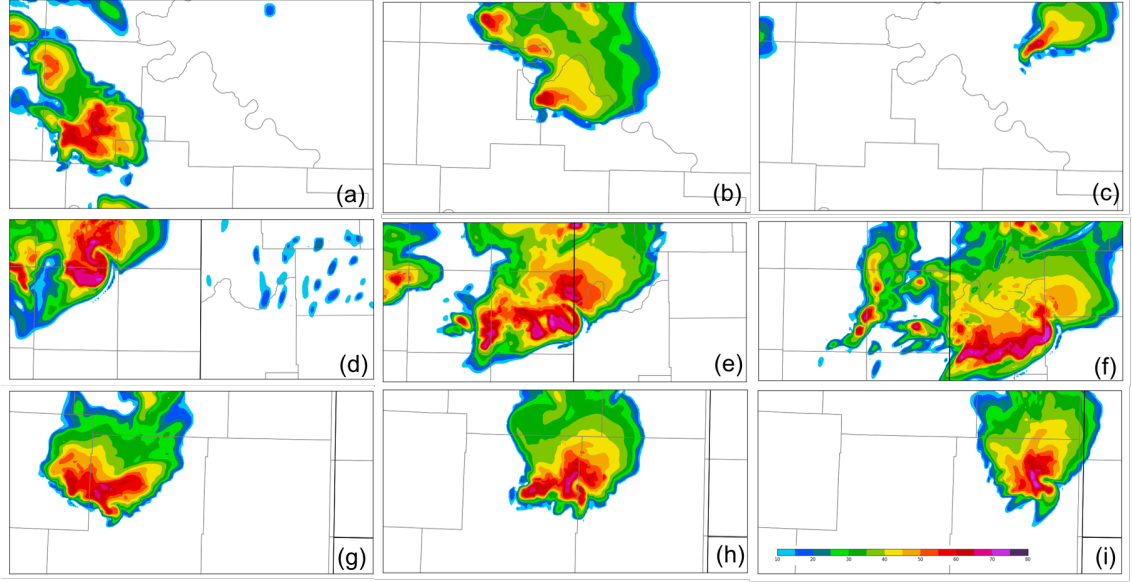
Figure 2.4: Reflectivity (dBZ) at z ≈ 1.5 km AGL. at t = 30 , 90, and 150 min. for 9 May(a-c), 16 May (d-f), and 24 May (g-i)

The following equation summarizes how the IC spread was reduced in the original ensemble:

$$X_{new} = X_c + [r * (X_{old} - X_c)]$$

X_{new} is the "new" ensemble with reduced IC spread, X_{old} is the original ensemble, X_c is the control member and r is the reducing factor. First, the control member X_c is subtracted from every single ensemble member in X_{old} for several variables on 3-km parent grid for the entire 3D domain. This difference between each ensemble member and the control are considered the set of perturbations. The perturbations were then reduced to 50%, 25%, and 10% of the original perturbation magnitudes (i.e., r = 0.5, 0.25, 0.1). The reduced perturbations were then added onto the control member generating a new ensemble with reduced IC spread. Following the

IC spread reduction, the 3-km ensembles were than downscaled to the 1-km grid. The perturbed variables include all three wind components (e.g., U, V, and W), perturbation potential temperature, water vapor mixing ratio, ice mixing ratio, rain water mixing ratio, graupel mixing ratio, cloud water mixing ratio, ice number concentration, rain number concentration, perturbation surface pressure of dry air, and perturbation geopotential height.

2.3.2   Initial Condition Spread Within vs. Outside The Storm

Even though storm-scale predictability studies have been performed in the past, a novel consideration in this study is the effect of IC spread inside and outside of the storm on ensemble forecasts of supercell evolution. Currently, the relationship between spread within the storm versus in the near-storm and mesoscale environment and the effects on forecast spread are unclear. Therefore, two experimental set-ups were used to explore reducing internal storm spread without reducing environment spread and vice versa. To reduce internal storm spread (i.e., having complete storm certainty), each ensemble member was initialized with the control member (i.e. no perturbations) for the primary storm of interest within a sub-domain centered on the storm. This means that the remaining 3-km domain remained unchanged (see Fig. 2.5 for the location of the prescribed domain for all three cases). Note that the perturbations were set to zero in the entire vertical column within the sub-domain to prevent vertical discontinuities from generating gravity waves and disrupting the solutions. However, this method does introduce horizontal discontinuities. Nevertheless, the horizontal discontinuities did not produce any noticeable impacts on the solutions. As for reducing the near-storm and mesoscale environment spread (i.e., having complete environment certainty) each ensemble member was initialized with
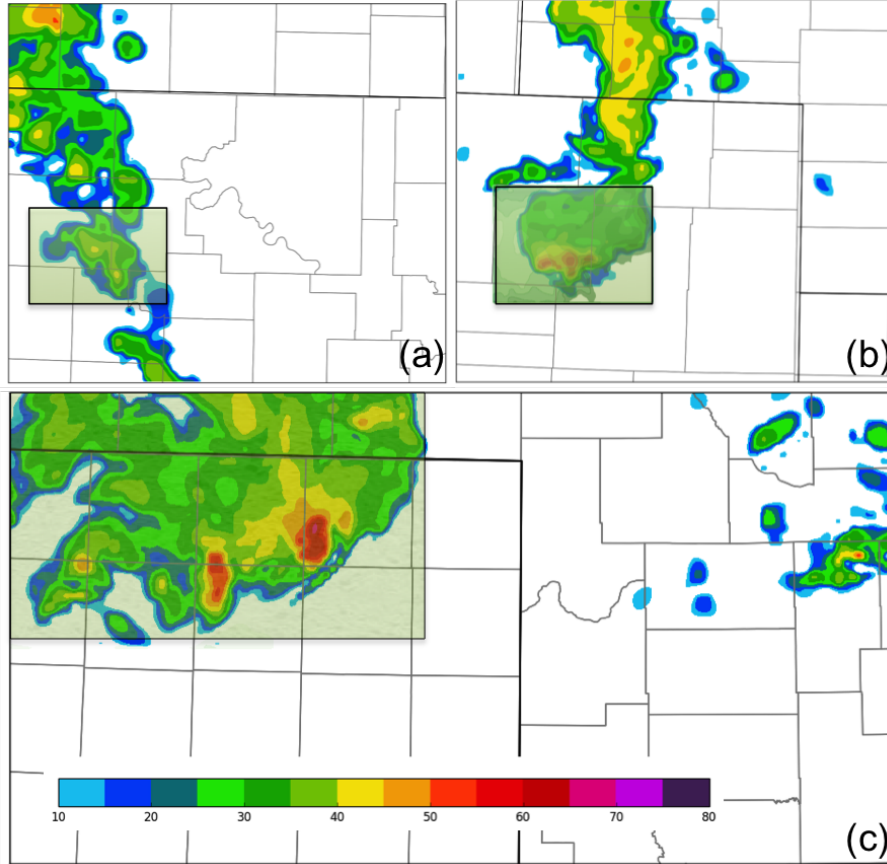
Figure 2.5: Sub-domain positions for (a) 9 May, (b) 24 May, (c) 16 May with reflectivity (dBZ) at 1.5 km AGL at t = 0.

the control member outside the subdomain with the storm of interest inside the subdomain remaining unchanged. Essentially, the two experiments test the effect of a single storm in multiple different environments vs. multiple storms in a single environment.

Even though this is a novel consideration, a study by Johnson and Wang (2016) explored the effects of IC spread at different scales (i.e., IC perturbations at larger and smaller scales, respectively). If one assumes that small-scale processes are primarily inside the storm and large-scale processes are outside the storm, then their approach is arguably similar to the current study. However, Johnson and Wang (2016) themselves do not explicitly suggest that the separation of scales is similar

to this separation of inside and outside the storm. Nevertheless, a comparison with their results are presented in section 3.2.4.

2.3.3 "Standard Range" as Measure of Dispersion

Ensemble spread is the tradition indicator of predictability in an ensemble framework (e.g., Lorenz 1984).The most widely used measure of spread is ensemble sample standard deviation (SD) defined as:

$$SD = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} (\phi_i - \overline{\phi})^2}$$

where N denotes the number of ensemble members and $\overline{\phi}$ is the ensemble mean of the variable ϕ (e.g., temperature, wind, water vapor mixing ratio, etc). This is reasonable for normal data distribution since the mean and SD together describes the entire distribution. We also know that 68% of data falls ± 1 SD about the mean with normal distributions. However, two implicit conditions must be met for SD to be the "best" measure of spread: the mean is a valid measure of central tendency and the data has no significant outliers (e.g., normal or nearly normal distribution). For example, in cases where the data distribution is highly non-normal (e.g., skewed strongly to the left or right), then the mean becomes less meaningful as the measure of central tendency, and therefore deviation from the mean becomes less meaningful. Furthermore, in cases of non-normal distributions, the portion of data within ± 1 SD becomes less certain. According to Chebyshev's Theorem from probability theory, at least 50% of the data will fall within ± 1 SD, independent of the distribution characteristics. Unfortunately, this is fairly imprecise. If the goal were to capture a

certain portion of the ensemble then interquartile range (IQR)

$$IQR = P_{75} - P_{25}$$

where P_{75} and P_{75} are the 75th and 25th percentile, respectively, would be more appealing. However, 50% does not capture a majority of the ensemble members. Therefore, a measure of spread that "combines" SD and IQR could be the difference of the approximate percentiles of a positive and negative SD in a normal distribution (i.e. 85th and 15th percentiles) or

$$SR = P_{85} - P_{15}$$

where such a measure could be called "standard" range (SR). SR is a more robust measure of spread than SD since no assumptions (e.g., implicit or otherwise) are made about the distribution characteristics and similar to IQR is immune to outliers. To note, if the data is normal and the exact percentiles were taken (i.e., 84.135th and 15.865th), then

$$SR = 2 \times SD$$

A digression about the robustness of SD, even with non-normal data is warranted. The time series of SD doubled versus SR of max intensity for the primary supercell features (e.g., mid-level UH, low-level vorticity, column-max updraft) from all three cases is shown in Fig. 2.6. In Fig. 2.6a-c, SR and SD are similar even though the distribution these statistics are generated from are fairly non-normal (not shown). However, in Fig. 2.6d-f, SR and SD are substantial different (e.g., difference of 15 m s^{-1} for updraft at t = 120 min, 450 m^2 s^{-2} for mid-level UH at t = 75 min, and 0.012 s^{-1} at t = 100 min). In these cases, the data is highly skewed and SD performed poorly by overestimating the spread. To further illustrate the

25
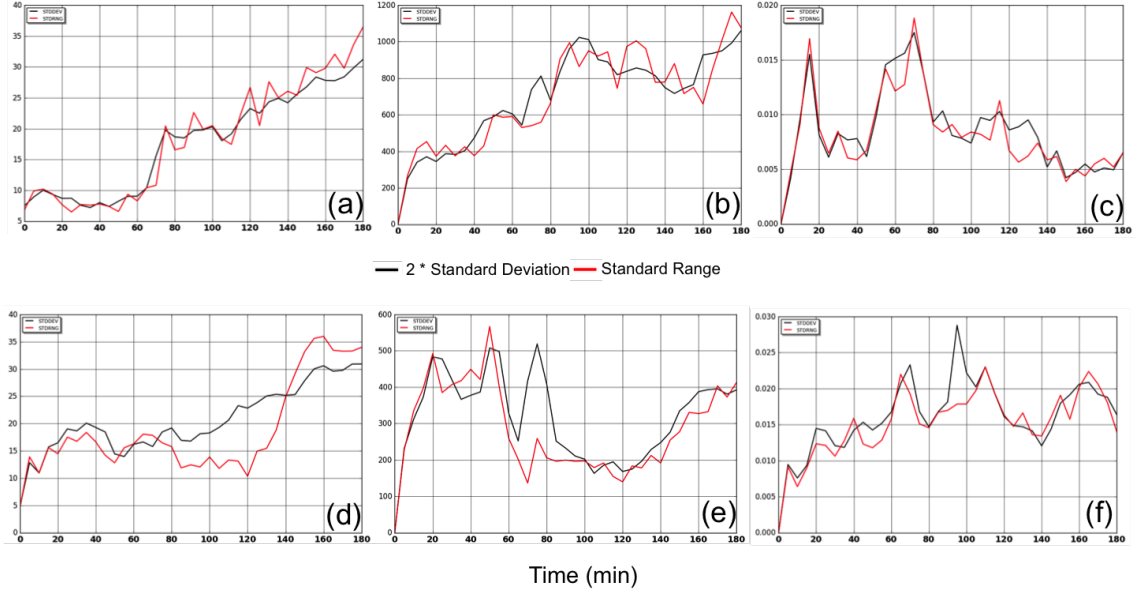
Figure 2.6: Time series of the standard deviation doubled median (black) versus standard range (red) of maximum intensity for (a) 24 May column-max updraft, (b) 24 May mid-level updraft helicity (UH), (c) 16 May low-level vorticity, (d) 9 May column-max updraft , (e) 16 May mid-level UH, and (f) 24 May low-level vorticity.

differences between SD and SR, 2D swaths of both parameters are shown in Fig. 2.7. In Fig. 2.7 we can see that SD can overestimate spread while in other cases it underestimates spread (e.g., top row vs. middle row, respectively). Moreover, even if SD captures the appropriate magnitude of spread, it can be displaced spatially from the SR towards outliers (Fig. 2.7, bottom row). Therefore, given that storm-scale distributions are characteristically non-normal, which causes SD to become unrepresentative of the true spread, the more robust statistic to use is SR.

Median vs. Mean

This same line of reasoning can be applied for employing the median rather the mean as the measure of central tendency. One characteristic of a non-normal distribution is the difference between the mean and median (i.e., they are equal in a normal distribution). The median, by definition, is at the center of a distribution while the mean may not be. Fig. 2.8 captures the evolution of the median and mean of max

26
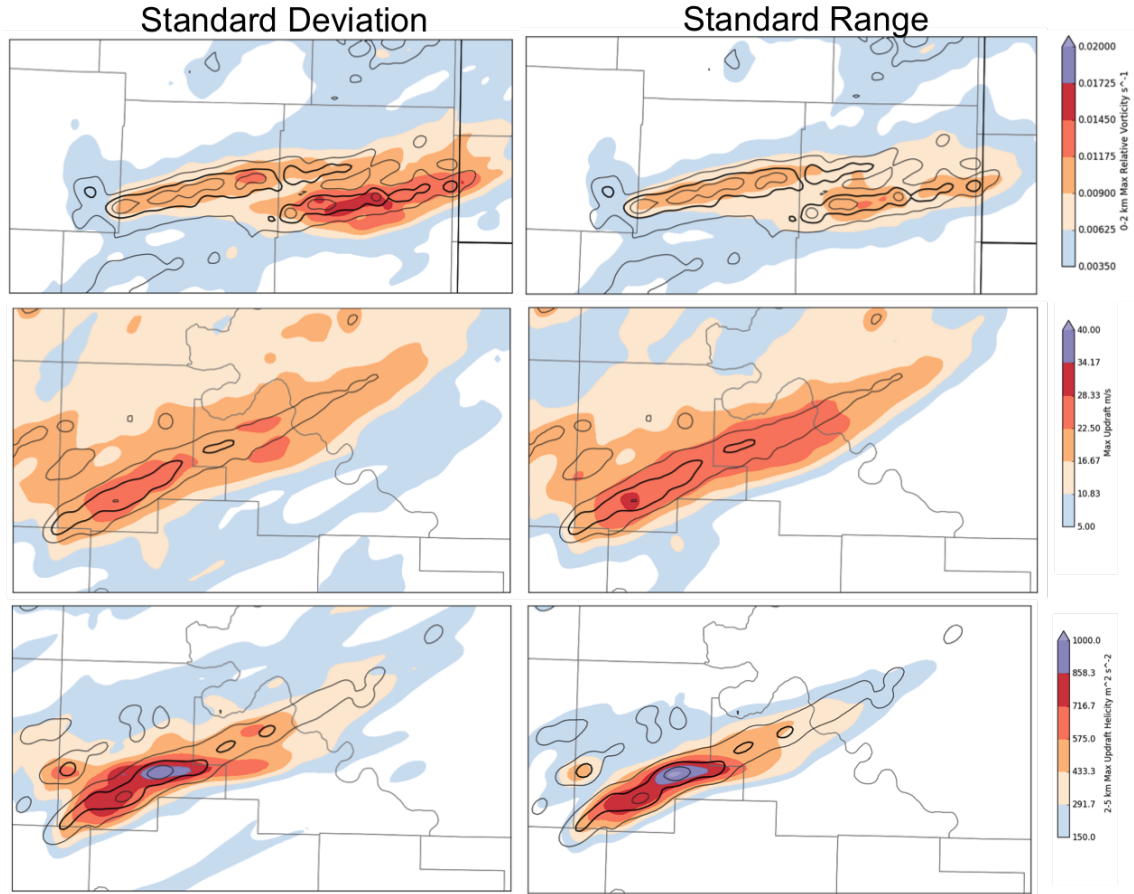
Figure 2.7: Swaths of standard deviation doubled (left column) and standard range (right column) with probability-matched median contours (see section 2.4.1). Top row: 0-2 km AGL relative vorticity (s^{-1}), middle row: column-max updraft ($m\ s^{-1}$), bottom row: 2-5 km AGL updraft helicity ($m^2\ s^{-2}$).

intensity for the primary supercell features for all three cases similar to the above comparison of SR and SD. The differences are smaller than was the case for SD vs. SR, but some systematic patterns exist. For example, in all three cases the median of max updraft was larger than the mean of max updraft (Fig. 2.8a,d) suggesting positive skewedness. As for low-level vorticity and mid-level UH, the mean was larger than the median, suggesting negative skewedness(Fig. 2.8b,c,e,f). Thus, the median is the robust measure of central tendency since the mean becomes skewed dependent on the supercell feature in question.
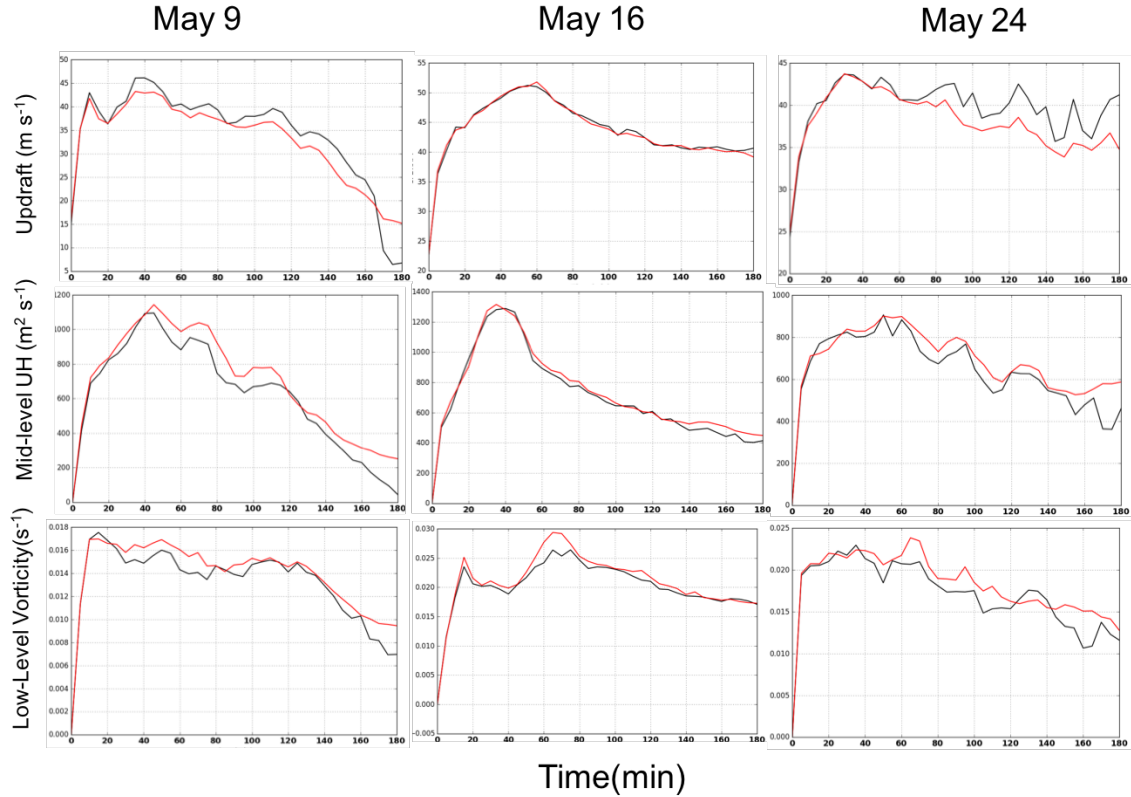
Figure 2.8: Median (black) versus mean(red) of max intensity for column-max updraft, max 0-2 km AGL vorticity, and 2-5 km AGL UH for all three cases.

2.4 Analysis Set-up

The common approaches to exploring predictability for large-scale atmospheric features (e.g., deviation in 500 hPa geopotential heights) are not appropriate for mesoscale/storm-scale (Warner et al. 1984). Due to the nature of the phenomena, diagnosing the predictability of individual storm features is more appropriate. Separating the features is also a more operational (e.g., practical) perspective (Warner et al. 1984). For example, the predictability of severe winds or extreme rainfall location would be quite useful for forecasters, especially in the WoF paradigm. Thus, the present study will take an operational approach to predictability based upon specific supercell features such as mid- and low-level mesocyclone, updraft and downdraft, rain, and severe surface winds. This approach is similar to that of previous studies

28

that investigated the predictability of mesoscale convective systems (MCSs) (Stensrud and Wicker 2004; Wandishin et al. 2008, 2010) as well as supercells (Cintineo and Stensrud 2013). Although such an approach has been taken in the past, the approach this study will take towards forecast spread, both in magnitude and spatially, warrants a thorough discussion on the figures and plots used for analysis.

2.4.1   2-Dimensional Spread Swaths

Evaluating the spatial characteristics of ensemble spread for various features is an appropriate first step. For spatial forecast spread, the SR of time-maximum values was used. It was found, however, that the ensemble median was altered in the process of IC spread reduction thereby introducing bias. Initially, the spatial swaths of SR were normalized by the ensemble median to account for this bias. However, the ensemble median decreased in magnitude as storm displacement increased. Thus, normalizing by ensemble median overinflated forecast spread values. A more appealing choice for normalization was probability-matched median (e.g., Ebert 2001). In Ebert (2001), they used probability-matched mean, but you can simply replace the mean with the median in their calculation to get the probability-matched median. Conceptually, probability matching blends the different spatial properties of the ensemble members to gain a more accurate spatial representation of the ensemble (for a greater description of probability matching see Ebert 2001). As a final note, the forecast spread normalization was only performed within subjectively-determined probability-matched median contours given that the probability-matched median goes to zero away from the location of the storms (i.e., primary or otherwise).

2.4.2 Probability of Exceedance

Spatial spread swaths are one tool for qualitatively diagnosing predictability, but additional information about the ensemble distribution is needed. Another useful metric is the probability of exceedance , which is the number of ensemble members that exceed a threshold (e.g. updraft $> 20m \ s^{-1}$) at a given horizontal grid point divided by the total number of ensemble members. In practice, instead of focusing on a single point, an ensemble member is included if it exceeds the threshold within some distance of any given point. This is to account for operationally insignificant storm displacement. For this study, a 3×3 neighborhood was used, which allows for spatial displacements less than 3 km from each horizontal grid point. Unfortunately, the choice of threshold for each supercell feature is rather subjective. For example, past studies have used remarkably low thresholds for mid-level UH (e.g., 50 $m^2 \ s^{-2}$; Cintineo and Stensrud 2013). Such low thresholds are more appropriate for coarser resolutions where the vertical vorticity magnitudes are smaller (e.g., Adlerman and Droegemeier 2002) and updrafts are under resolved (e.g., Potvin and Flora 2015). More recent studies have used a 180 $m^2 \ s^{-2}$ threshold (e.g., Zhang et al. 2015, 2016) based on the work by Naylor et al. (2012). However, 180 $m^2 \ s^{-2}$ was the lowest magnitude of mid-level UH that can be defined as a mesocyclone on a 1-km grid. The 180 $m^2 \ s^{-2}$ threshold was initially used in this study, but it was found to be a low threshold as well (not shown). To devise an more relevant threshold, a simply scale analysis argument was made. The expression for updraft helicity (UH) is given by (from section 3.1.1),

$$UH = \int_{z_b}^{z_t} w\zeta dz = w\zeta\Delta z$$

where w is vertical velocity, ζ is vertical vorticity, and z_b and z_t are bottom and top heights of some layer, respectively. On a 1-km grid, vertical vorticity on the order of 0.01 s^{-1} is reasonable (e.g., Adlerman and Droegemeier 2002) while vertical velocity

should be on the order of $10 \; m \; s^{-1}$. If we assume vertical velocity and vorticity are constant in the mid-level (e.g., 2-5 km AGL),

$$w\zeta\Delta z = (10 \; ms^{-1})(0.01 \; s^{-1})(3000 \; m) = 300 \; m^2 s^{-2}$$

then, $300 \; m^2 \; s^{-2}$ is a fairly appropriate threshold.

Thus, the probability of mid-level UH greater than $300 \; m^2/s^2$ within a 3×3 neighborhood corresponding to each horizontal grid point is used. As for the other supercell features similar, quasi-rigorous arguments could be made for the appropriate threshold, but they will not be presented here. As a complement to the probability of exceedance, time-maximum values were contoured at the threshold value for each given feature (i.e., colloquially referred to as "spaghetti" plots). Qualitative indications of location spread are also given by the spaghetti plots.

As will be shown, reducing the ensemble towards the control member reduced the ensemble bias. In light of that, there will be cases where the probabilities are rapidly increasing even when the normalized forecast spread is not decreasing. This is directly related to the IC spread collapsing to the control member, which did exceed the threshold. This does not necessarily prevent the results from being transferable to the real world since operational convecting-allowing ensembles are biased themselves. However, the probabilities should be interpreted with care since the change in the probabilities can arise from both spread and bias reduction rather than from spread reduction alone.

2.4.3   Phase and Amplitude Spread

The ensemble forecast spread is some function of two parts: phase and amplitude spread. Phase spread represents the degree to which storms become displaced in time and space while amplitude spread indicates the degree to which the magnitude

of a particular storm metric can vary. This division of forecast spread source is necessary as it allows for a thorough understanding of the total spread. Arguably, quantifying the storm displacement amongst the ensemble members quantifies phase (or location) spread. Therefore, the median distance between every combination of ensemble members storm location (indicated by max mid-level UH) was used.

To evaluate amplitude spread, we used time series of the SR of max intensity (e.g., spread in max updraft strength) normalized by the median max intensity. Similar to the 2D spread swaths, normalization was necessary due to the changes in bias. In essence, the amplitude spread represents the spread in intensity with little regard for location. The time series were time-averaged in a 20 min window centered on every 5 min output (i.e., ± 10 min) to smooth the curves and eliminate uncertainty about possible curve convergence (see section 3.1.1. for more details on curve (spread) convergence).

2.5 Important Considerations

Although real data were used, the primary goal was not to replicate case-specific observed supercell evolution, but rather to isolate the effect of realistic IC uncertainty by using a perfect-model assumption. Under a perfect model assumption, any ensemble dispersion is solely due to the IC spread. However, it is understood that the model physics can interact non-linearly with the environment making the results not only sensitive to the ICs, but also choice of parameterization.

Second, a key issue with storm-scale data assimilation is ensemble underdispersion, where one factor is insufficient IC spread. If the RUC errors from Cintineo and Stensrud (2013) are representative of storm-scale analysis errors, then the vertical profiles of root-mean-square difference (rmsd) for 9 May (which is representative of all three cases) for both wind horizontal wind components and relative humidity indicate that the ICs may be appropriately dispersed (cf. left and right columns

of Fig. 2.9). Temperature is an outlier with values roughly half that of the RUC
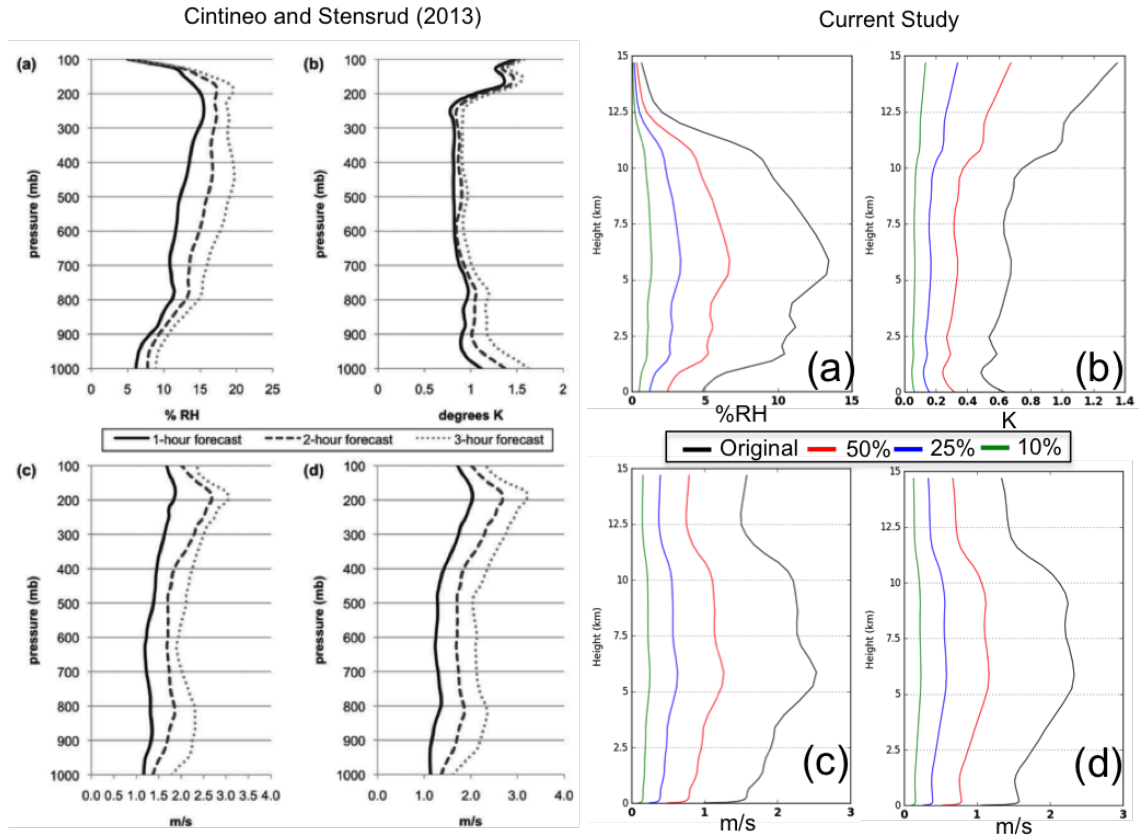
Figure 2.9: Left Column: Fig. 1 from Cintineo and Stensrud (2013). Root-mean-square errors (RMSE) of the (a) relative humidity (%), (b) temperature(K), and the (c) u and (d) y components of the wind (m s21) calculated at 0000 UTC from 1-, 2-,and 3-h 13-km RUC forecasts. Calculations only made in regions with SCP $\geq$ 1 over a 70-day period. Right Column: root mean square difference (RMSD) of (a) relative humidity (%), (b) temperature (K), (c) and (d) u- and v-component of the winds ($m\ s^{-1}$) from the ensemble mean. Calculations performed on the 3-km grid.

errors. However, given the inclusion of mesonet observations, radar, and satellite in the NEWS-e system, it is conceivable that the desired initial temperature spread has been reduced. Ultimately, this is not to suggest that the initial spread in our ensembles are "optimally" dispersed, but arguably capture current storm-scale analysis uncertainty. Even if this study is using underdispersed ensembles, it would still illustrate the effects of reducing current IC spread on forecast spread growth.

2.6 Summary of Goals

The time scale that storm-scale ensembles are reliable is essential knowledge for the WoF paradigm and therefore their robustness needs to be assessed. Given that supercells are a primary source of life-threatening hazards such as EF2+ tornadoes, > 1 in hail, and torrential rainfall, a thorough understanding of their predictability is warranted. This includes assessing the predictability, both intrinsic and practical, of primary supercell features, including the low-level mesocyclone and downdraft (which have been neglected in a majority of past supercell predictability studies) over multiple cases. Therefore, an initial goal of this study was to develop a suitable full-physics/real data NWP framework for studying storm-scale predictability. This is a crucial step as doing so facilitates the assessment of supercell predictability. To summarize, the steps taken to achieve a suitable framework are:

- Full-physics/real data rather than a homogeneous, idealized case

- Use of DA to generate the ensemble; including radar and satellite observation in the DA, which is critical for intra-storm regions

It is also important to consider non-traditional statistics and metrics because of the non-normality of storm-scale data. Although SD is the familiar measure of spread perhaps a more appropriate measure of spread for this application may be SR, which is immune to the non-normalities at storm-scales. Additionally, there is a need to move away from gridpoint-based metrics (e.g., RMSE) since storms are more like objects than continuous fields (highly discrete in time and space). Furthermore, given that supercell predictability may be sensitive to the environment or whether the supercell is tornadic versus non-tornadic, several cases need to be evaluated to ensure the robustness of the results. Exploring the nature of IC spread inside versus

outside the storm is a vital consideration for the WoF paradigm. For example, it will likely be a difficult or daunting task to substantially reduce IC spread in the entire domain. However, since the WoF paradigm is focused primarily on improving 1-3 h lead times results may show that reducing spread in only certain portions of the domain may prove beneficial (e.g., just within the storm).

To summarize, the primary goals of this work are to:

- Develop a suitable full-physics/real data NWP framework for studying storm-scale predictability (also includes non-traditional metrics).

- Diagnose the forecast spread growth of supercell features in an ensemble-framework (where the ensemble was generated from the EnKF) using real data/including physical parameterizations (as opposed to horizontally homogeneous data) on warn-on-forecast time and space scales (e.g., O(10 min, 1 km))

- Explore the importance of uncertainty within vs. outside of the storm.

- Examine a larger number of cases and supercell features than previous studies

2.7 Hypotheses

Given the goals presented, it is neccesary to present some of the expected (or at least hypotheized) results of this work. Some hypotheses are presented below:

- Intrinsic supercell predictability is on WoF timescales

- Intrinsic predictability is case-depedent

- Smaller scale supercell features should be less predictable than larger scale features

- Improving intra-storm certainty would provide greater benefits over storm environment certainty early on with the opposite at longer lead times.

A primary hypothesis for this work is that the intrinsic supercell predictability limit may be on WoF timescales (e.g., 1-3 h). This is motivated by theoretical turbulence models (e.g., Lorenz 1969; Lilly 1990) and more recent applied work (e.g., Zhang et al. 2016). As mentioned in section 2.2, it is hypothesized that the intrinsic predictability limit will be case-dependent. This is reasonable given that certain regimes are more or less predictable then others (e.g., Hohenegger et al. 2006; Boer 1994). Given that predictability was feature-dependent in Cintineo and Stensrud (2013), it is assumed that will be similar in this study. Furthermore, since predictability is scale-dependent (e.g., Lorenz 1969), it is hypothesized that smaller-scale features will be less predictable than larger scale features. Lastly, as mentioned in section 2.3.2, there has been no work on explicitly separating the effects of spread inside vs. outside the storm on supercell evolution in an ensemble framework. One hypothesis that seems clear is improving initial storm certainty should provide greater benefits early on compared to improved initial environment certainty with the opposite being true at longer lead times. However, there is uncertainty about the duration of the benefits and when they set in.

Chapter 3: Results

3.1 Original Experiments

3.1.1 Mid-Level Mesocyclone

The single distinguishing feature of a supercell, from other modes of convection, is the deep, persistent, quasi-steady rotating updraft known as the mid-level mesocyclone (MLM). A parameter frequently used for assessing the MLM is updraft helicity (UH), which is defined by

$$UH = \int_{z_b}^{z_t} w\zeta dz$$

where z_b and z_t are commonly 2 km and 5 km, respectively. In Fig. 3.1, the MLM evolution of all three control members is presented. Note that the 300 m^2 s^{-2} contour is overlaid, which will facilitate the interpretation of the probability of exceedance. To briefly summarize, the MLM in May9_CNTL rapidly intensified early on with mid-level (e.g. 2-5 km AGL) UH values near 1200 m^2 s^{-2} (see Fig. 3.1a). After the initial intensification, there was a brief period of cycling followed by a fairly steady-state evolution. The MLM in the May16_CNTL had the smallest mid-level UH values of the three cases (still fairly intense) but had a significantly larger spatial extent associated with a larger updraft driven by a surging gust front (Fig. 3.1b). Even though the storm was steadily growing upscale into an MCS, it did maintain sufficient mid-level UH to be classified as a supercell throughout the complete integration. Finally, The MLM in May24_CNTL was the strongest of the three cases with a brief intensification period around t= 60 min followed by a period of cycling where the MLM rapidly strengthened and decayed in the last 30-40 min (Fig. 3.1c).

Fig. 3.2a-d shows that reducing the IC spread greatly reduced the normalized mid-level UH forecast spread for the 9 May case. However, the probability-matched median (i.e. the black contours) were altered by the IC spread reduction (i.e., introducing bias).
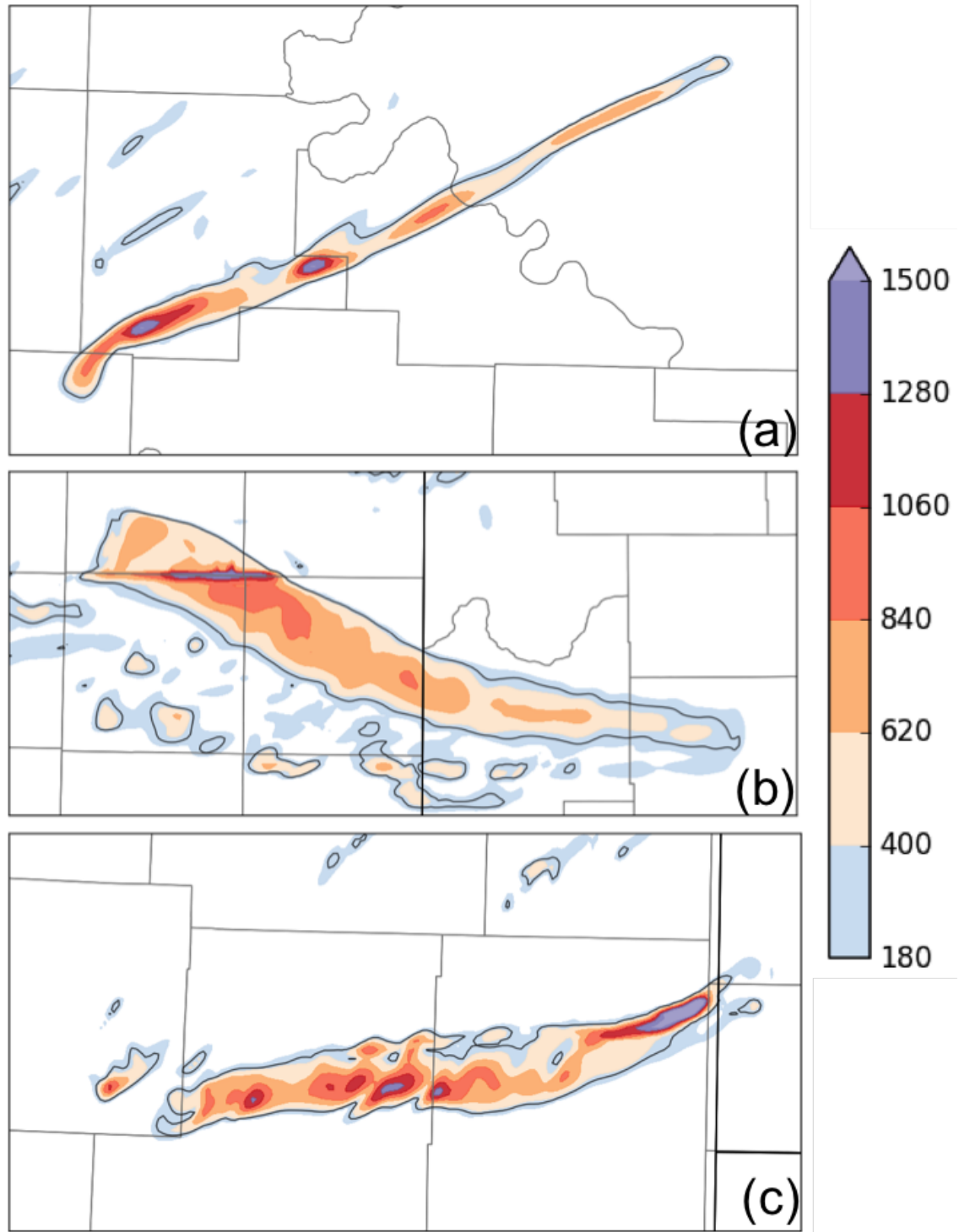
Figure 3.1: Control member evolution of 2-5 km AGL updraft helicity for (a) 9 May, (b) 16 May, and (c) 24 May.

However, the probability-matched median was increasing as the IC spread was reduced. The substantial change in the probability-matched median indicates that collapsing towards the control member is introducing bias as previously mentioned in section 2.4.2. Thus, normalizing the forecast spread was necessary to facilitate
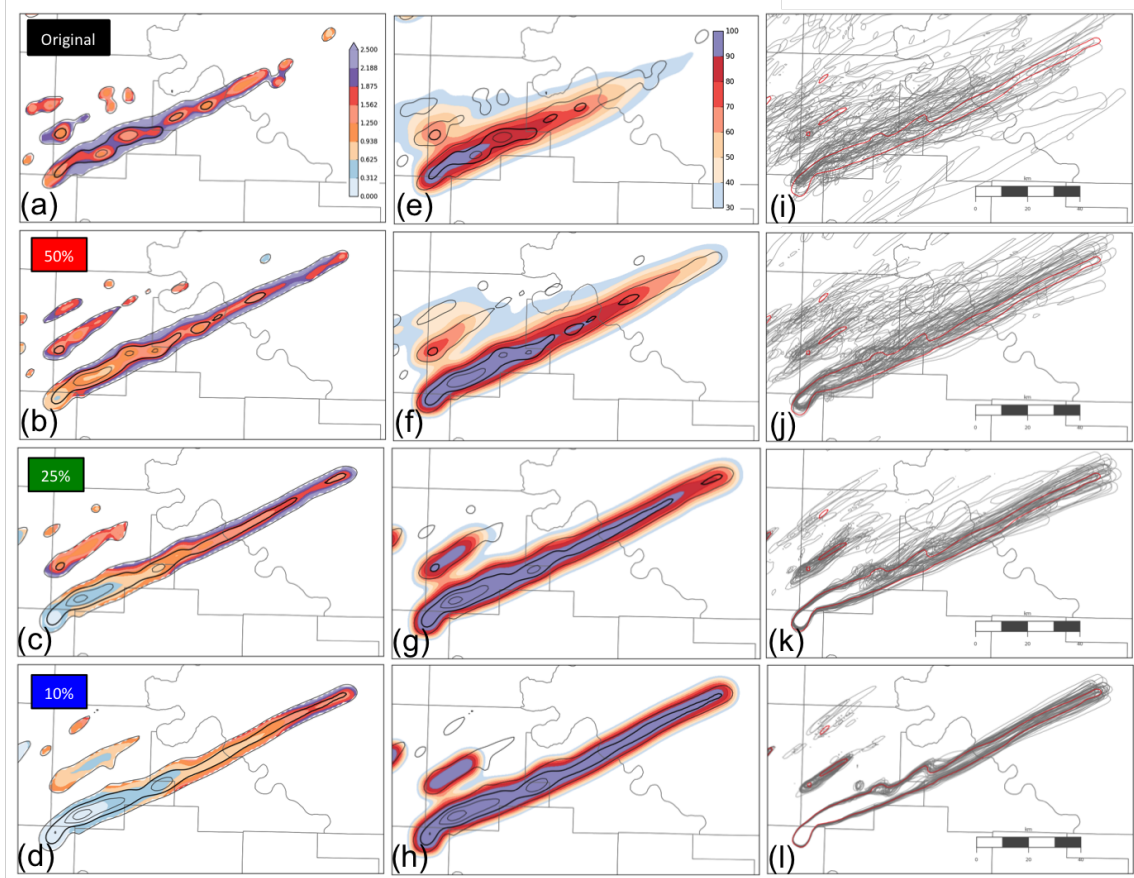
Figure 3.2: 9 May 2-5 km AGL updraft helicity. Left Column: (a-d) Normalized forecast spread swaths (filled contours) with probability-matched median contoured (black lines; thick line = 350 m^2s^{-2} with a contour interval of 250 m^2s^{-2}) with decreasing initial condition spread going down. Center column: (e-h) Probability of exceedance > 300 m^2s^{-2} (filled contour) within a 3 X 3 neighborhood corresponding to each horizontal grid point with probability-matched median contoured (black lines). Right column: (i - l) Spaghetti plot of 2-5 km AGL updraft helicity at 300 m^2s^{-2} for each ensemble member.

comparisons between the experiments (e.g., 25% vs. 10% experiment). Looking at Fig. 3.2c,d there are discernible differences between the normalized forecast spread with no indication of spread convergence. We define spread convergence in the

2D spread swaths as the point in time at which the normalized forecast spread in the 10% experiment is qualitatively indistinguishable from the 25% experiment. Arguably, the initial spread in the 10% experiments could be captured by a (distant) future observing network and modeling system. Even so, the initial 10% spread in all three cases is small enough such that any spread convergence may imply the intrinsic predictability limit has been met. Therefore, spread convergence in this study is interchangeable with the implied intrinsic predictability limit. As for the 16 and 24 May cases, reducing IC spread decreased the normalized mid-level UH forecast spread (Fig. 3.3a-d and Fig. 3.4a-d). However, unlike the 9 May case, there were distinct spread convergence points where the spread was qualitatively similar between the 25% and 10% experiments meaning an implied intrinsic predictability has likely been met. As will be shown shortly, these implied spread convergence points are collocated in time with normalized amplitude spread convergence.
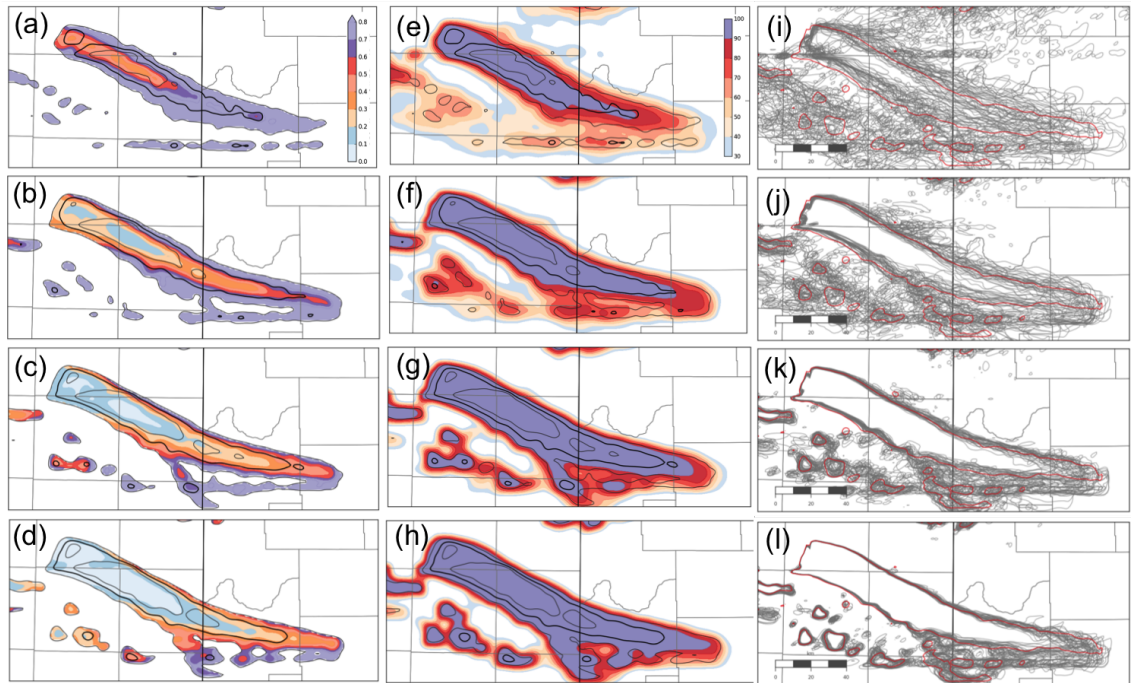
Figure 3.3: Same as Fig. 3.2 but 16 May 2-5 km AGL updraft helicity. Thick line = 450 m^2s^{-2} with a contour interval of 200 m^2s^{-2}.
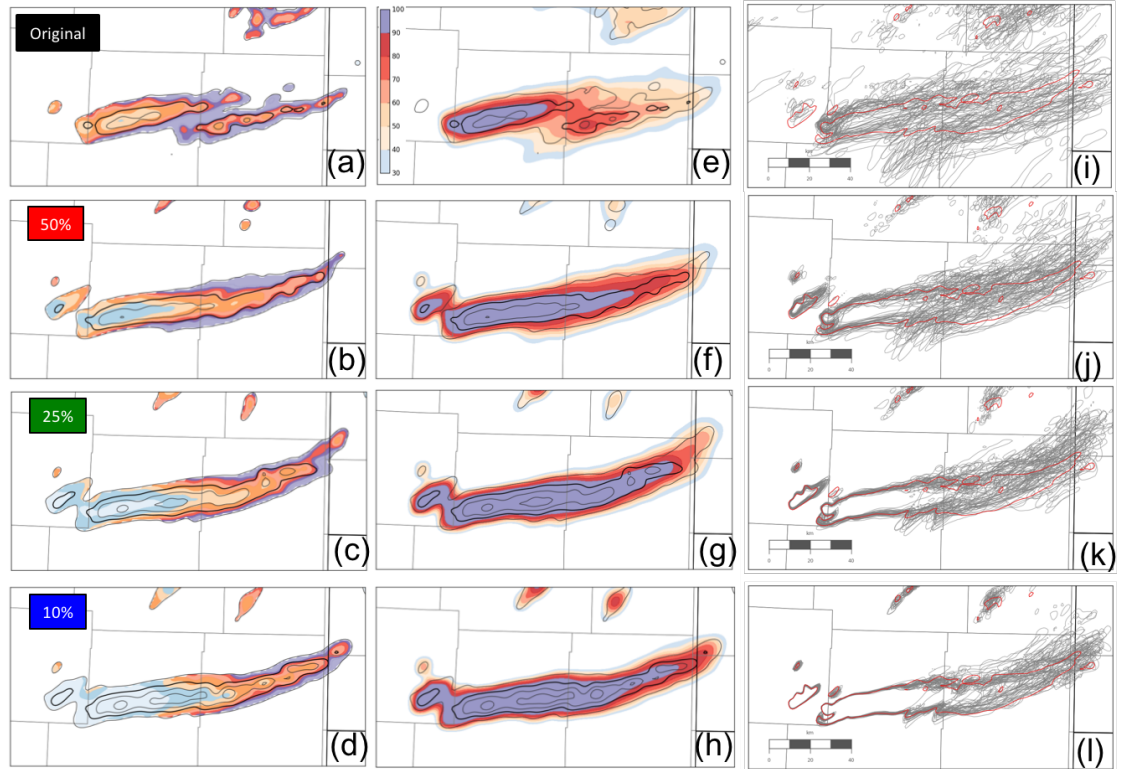
Figure 3.4: Same for Fig. 3.2 but for 24 May 2-5 km AGL updraft helicity.

Even though diagnosing an implied intrinsic predictability limit has merit, diagnosing current (or near future) supercell predictability has stronger operational implications. In all three cases, the original probability of exceedance's were >50 % throughout much of the integration (Fig. 3.2 - Fig.3.4e-h). Furthermore, the improvement in the probability of exceedance with a 50 % reduction in IC spread in all three cases signifies that the practical predictability limit can be significantly lengthened. In addition to providing information about practical predictability, the probability of exceedance swath also has operational utility. Probability of exceedance is more intuitive forecast tool than normalized forecast spread and is able to provide simple, but ample information about ensemble agreement on the path and evolution of any supercell feature.

In general, reducing IC spread substantially increased the probability of mid-

level UH > 300 m^2 s^{-2} for the 9 May case (Fig. 3.2e-g) . However, for the 16 and 24 May cases, there was virtually no increase in the probability of exceedance after the 50% reduction in IC spread (Fig. 3.3e-g and Fig. 3.4e-g). Even when the neighborhood size was set to zero, there were no discernible differences between the 10% and 25% experiments for both cases (not shown). The relationship between the spaghetti plots of 300 m^2 s^{-2} and probability of exceedance can shed some light on the cause of the probability convergence in the 16 and 24 May cases. Although it is not shown, a supermajority (i.e., > 60-80 %) of ensemble members exceeded the 300 m^2 s^{-2} threshold throughout the 3 h integration. It appropriate then to assume that the converging exceedance probabilities are primarily related to phase errors. For example, the 16 May ensemble exceeds the 300 $m^2 s^{-2}$ threshold, but due to little improvement in the MLM position spread (see Fig. 3.3i-l), the probabilities are converging in the 25 % and 10 % experiments. Converging probabilities as the IC spread is reduced can be an indication of location predictability. As we can see, location predictability is distinct from overall and amplitude predictability

Unlike the 16 May case, the spaghetti plots in both the 9 and 24 May cases show that the MLM position certainty improves with IC spread reduction (see Fig. 3.2i-l and Fig. 3.4i-l). The 24 May case is further distinct from the 16 May case since there is a clear spread convergence point, but no probability convergence. This distinction between overall and location predictability further supports the idea that they are truly separate. hus, it is warranted to separate the two sources of spread (i.e. location and amplitude) and address them individually to gain a better understanding of general supercell predictability.

Fig. 3.5 shows the quantification of location spread. The storm location spread is converging (i.e., an intrinsic storm location predictability limit <3 h) for the 16 May case (Fig. 3.5 b). Spread convergence in 16 May is expected as the spaghetti plots indicated little difference between the 25% and 10% experiments. Fig. 3.5a,c

verifies the qualitative analysis for the 9 and 24 May cases where there was no spread convergence. This implies that the intrinsic location predictability in these two cases is beyond 3 h. However, given the similarities between 25% and 10% experiments for the 9 and 24 May cases, the limit is not far beyond 3 h.
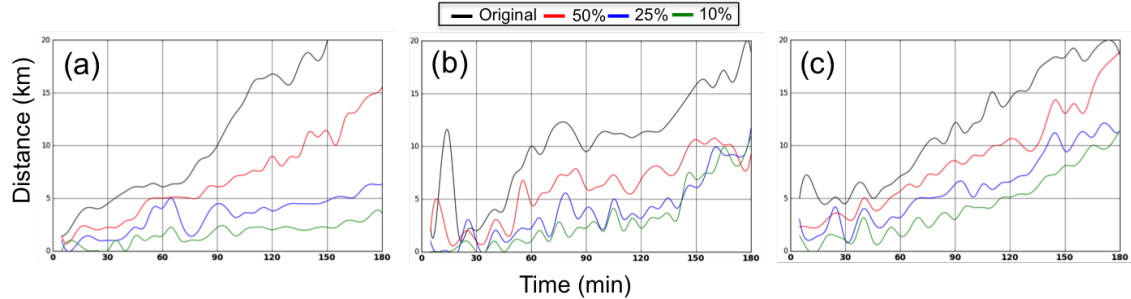
Figure 3.5: Time series of the median distance between every combination of ensemble members storm location (a) 9 May, (b) 16 May, and (c) 24 May. Max 2-5 km updraft helicity is used to identify storm location.

Although location spread convergence indicates the intrinsic location predictability, these results also shed light on practical location predictability. In all three cases, the 50% IC spread reduction lead to location spread reductions of as much as 5-10 km. Although the 50% reduction may be considered substantial, it still indicates that storm location can benefit from smaller IC spread reductions. The benefits of IC spread reduction on location spread are further supported by a closer inspection of Fig. 3.5. In all three cases, the location spread growth rate is weakly non-linear. This is distinct from the conceptual logarithmic spread growth that increases in rate at smaller IC spread. Thus, reducing IC spread benefits the location spread for a longer period than perhaps was initially thought.

The location spread is one component for understanding general supercell predictability and the complement is amplitude spread. Fig. 3.6a-c shows the quantification of amplitude spread as described in section 2.4.3. For the normalized amplitude spread time series, spread convergence is loosely defined here as either actual or implied curve convergence. It is also defined as when forecast spread loses

all information about its ICs. An example of this would be in Fig. 3.13, where the normalized forecast spread in the 10 % experiment is greater than the forecast spread in the 25% and 50% experiments. In that case, no curve convergence occurs, but the forecast spread in 10 % experiment has lost its connection with a more certain IC.

Looking at Fig. 3.6a, there was no normalized amplitude spread convergence for the 9 May case which agrees with the qualitative analysis of the 2D spread swaths. However, based on the similarities between 25% and 10% experiments, the intrinsic amplitude predictability limit is not far beyond 3h. As for the 16 and 24 May
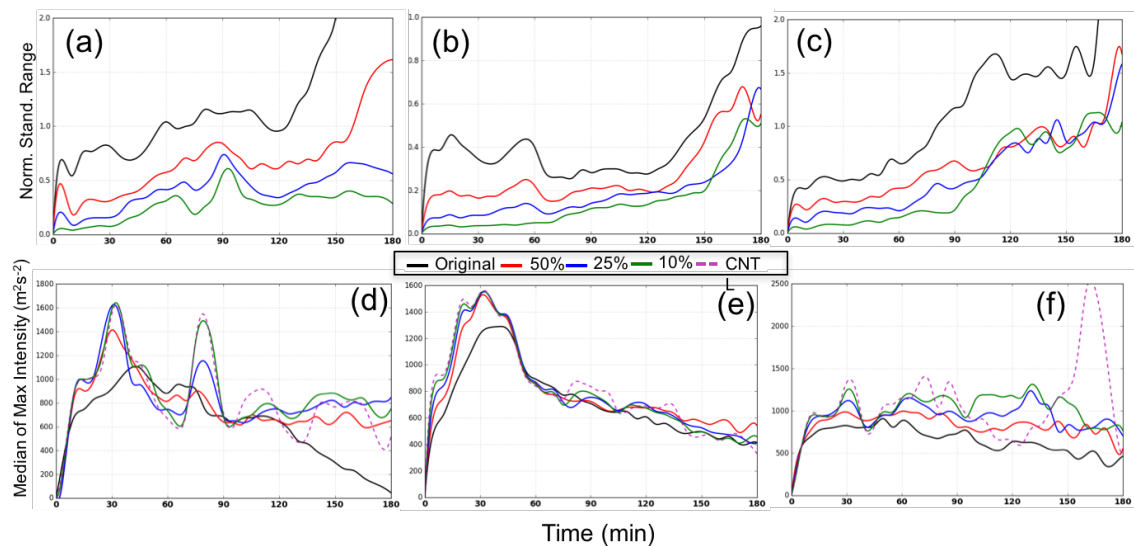
Figure 3.6: Time series of the normalized standard range of max 2-5 km AGL UH by the median of max 2-5 km AGL UH for (a) 9 May, (b) 16 May, and (c) 24 May. Time series of the median of max 2-5 km AGL UH along with the control member evolution (dashed magenta) for (d) 9 May, (e) 16 May, and (f) 24 May.

cases, there was normalized amplitude spread convergence points of 150 and 100 min, respectively (Fig. 3.6b-c). It is also possible to determine practical amplitude predictability from these plots. For instance, determining when the original 1-h amplitude spread occurs in the 50% experiment may be useful information to a forecaster. In the 9 May case, the magnitude of amplitude spread at 1-h is not

achieved in the 50% experiment until t = 160 min. One could interpret this as 100 min of additional practical predictability with a 50 % IC spread reduction. The 16 May case had a similar delay as 9 May where in the 24 May case the implied intrinsic predictability limit was met before the 50% experiment became equal to the original 1-h amplitude spread.

Another crucial component of predictability in an ensemble framework is that bias may be sensitive to the IC spread. To characterize this, the change in bias as the IC spread is decreased is used, where bias is the absolute difference between the control member and the ensemble median. Recall that the control member was chosen as the member closest to the ensemble center and therefore deviation of the ensemble median from the control member indicates bias. In Fig. 3.7, we can see in all three cases, the bias for mid-level UH increases as the IC spread *increases.* If the bias increases as the IC perturbations are increased, then the ensemble is
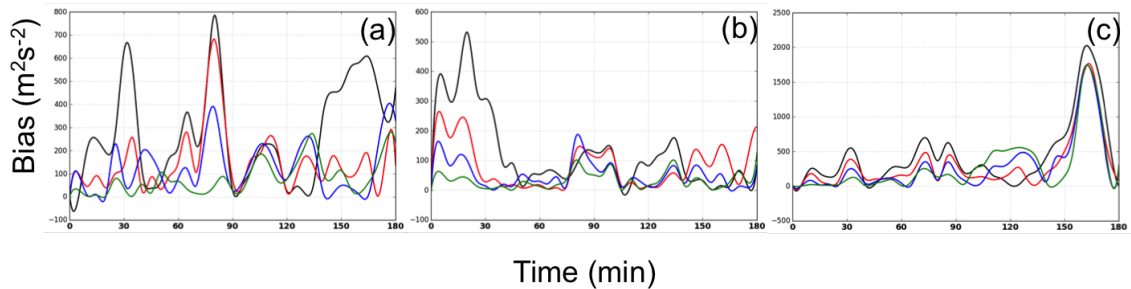
Figure 3.7: Time series of the absolute difference between the control member mid-level UH and median of max mid-level UH for (a) 9 May, (b) 16 May, and (c) 24 May.

becoming uncentered. The increase in "uncenteredness" can be an indication of less predictability. In particular, the 9 May case had the greatest difference between the original and 10% experiments, especially in the last hour of the simulation. This related to the fact that some ensembles did not sustain a supercell through the 3 hrs. Therefore, the implication is that 9 May is a "less" predictable case with current IC spread as the magnitude of the perturbations lead to a binary forecast in the

ensemble (i.e., having a storm and not having storm are likely).

Given that the implied location and amplitude predictability for the 16 and 24 May cases are quite different a discussion on the relationship between location and amplitude spread is warranted. Similar to the intrinsic predictability limit defined above, the intrinsic location (amplitude) predictability limit is the point at which the forecast spread in storm location (intensity) can longer be reduced by reducing IC spread. Arguably, the overall intrinsic predictability is some function of the intrinsic amplitude and location predictability. Even so, one could define the overall intrinsic predictability to be equal to the lesser (or greater) of the two limits. As we will see the location predictability can be greater than or equal to the amplitude predictability, but the implied predictability limit in the 2D spread swaths corresponds well with the amplitude predictability limit. Therefore, the overall intrinsic predictability limit will be determined by the amplitude predictability limit.

Given the implied location and amplitude predictability limits for the 16 May case, the overall mid-level UH intrinsic predictability limit is near 3 h. As for the 24 May case, the location predictability was beyond 3 h, but the amplitude spread convergence was well within 3 h. Evaluating the 2D mid-level UH spread for the 24 May case at individual times for the different experiments, it was found that spread convergence occurred approximately when the amplitude spread converged (not shown). Although the location spread did not converge, the overall intrinsic predictability limit was dictated by the amplitude spread convergence. This means storm intensity is more sensitive to IC perturbations than storm motion. This is consistent with past studies that found amplitude rather than phase spread contributes greater to the overall spread/error (Lorenz 1969; Park 1999; Yussouf et al. 2016). For a further discussion on amplitude and location predictability, see Ch. 4.

The relationship between location and amplitude spread for the 24 May case has strong operational implications. As discussed above, with a 50% reduction in

IC spread, there is a substantial reduction in both location and amplitude spread. These reductions mean that the practical predictability limit can be significantly lengthened. With significant IC reductions (> 50% reduction), forecast location spread in the 24 May case saw great benefits, but the forecast amplitude spread eventually converged. This relationship between location and amplitude spread suggests that in (distant) future, accurately forecasting storm position may be possible out to 3 h+, but the intensity will not necessarily be accurately forecasted. Ultimately, improving storm location forecasts will still provide critical guidance for operational forecasters. Note that the intrinsic predictability limit is not necessarily related to loss of forecast value and this will be discussed further in Ch. 4.

The very essence of intrinsic predictability is chaos, which is the sensitivity of error/spread growth to the IC uncertainty (i.e., regardless of the nature of growth). If the error/spread growth were insensitive to the IC uncertainty, then continual reductions in IC uncertainty would indefinitely reduce forecast uncertainty. However, when the forecast growth is extremely sensitive to the IC uncertainty, then there are diminishing returns on forecast uncertainty. In other words, a characteristic of extreme sensitivity to IC uncertainty is the rapid increase in forecast spread/error growth with smaller IC uncertainty. To demonstrate that forecast spread growth for supercells has such sensitivity to the IC uncertainty, presented in Fig. 3.8 is the forecast amplitude spread normalized by the initial amplitude spread for the 9 May case (which is representative of the other cases). Looking at Fig. 3.8, the 10% experiment has the largest forecast spread relative to the initial spread. The increase in relative forecast spread growth with IC uncertainty reduction illustrates that forecast uncertainty benefits are limited. Ultimately, an intrinsic predictability limit exists because the relative forecast spread increases at smaller IC uncertainty.
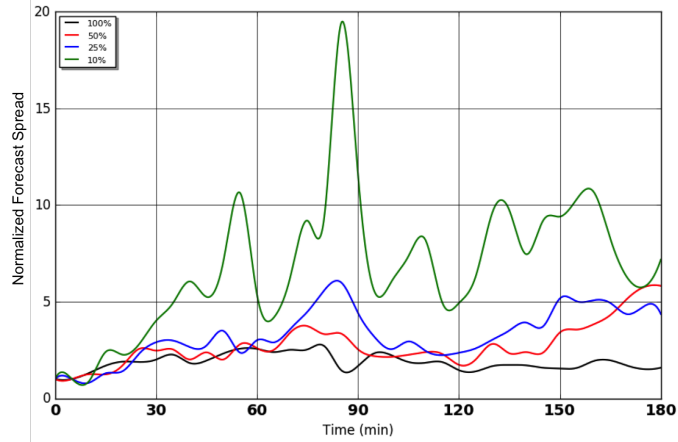
Figure 3.8: Time series of forecast spread of 2-5 km AGL updraft helicity normalized by IC spread for 9 May .

3.1.2 Updraft

The crucial component for any severe thunderstorm is the updraft. For this study, column-max updraft speed was used, but it will hereafter be referred to as the updraft. In all three cases, there were strong similarities between the evolution of the updraft and mid-level UH (cf. Fig. 3.1a-c and Fig. 3.9a-c). The similarities between the updraft and mid-level UH was not surprising given that the updraft is a primary component of the MLM.

The reduction in normalized updraft forecast spread for the 9 and 16 May cases were fairly substantial (Fig. 3.10(a-d) and Fig. 3.11(a-d)). Moreover, the reduction in forecast spread for these two cases appear to be greater than that for mid-level UH. The greater reduction in updraft forecast spread versus the mid-level UH indicates that the updraft is more sensitive to the IC spread. Given that UH is an integrated quantity this may explain why it is a more predictable(i.e., less sensitive to IC spread) quantity in these two cases. However, as we will see, this does not hold true for the 24 May updraft amplitude spread, which was more predictable than the mid-level UH amplitude spread. As for the location spread, the probability of exceedance

48
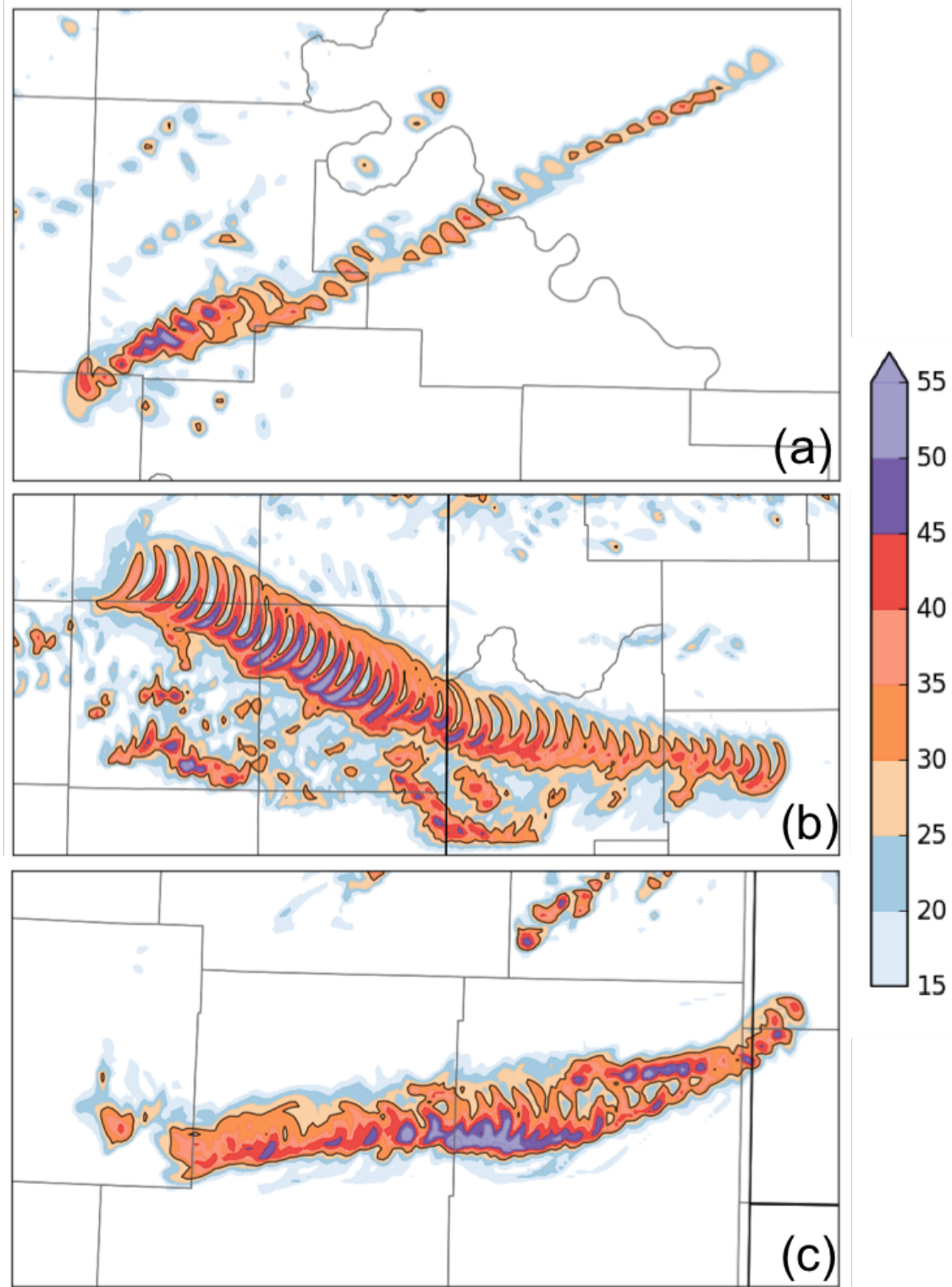
Figure 3.9: Control member evolution of column-max updraft for (a) 9 May, (b) 16 May, and (c) 24 May.

Figure 3.10: 9 May column-max updraft. Left Column: (a-d) Normalized forecast spread swaths (filled contours) with probability-matched median contoured (black lines; thick line $= 25 \ ms^{-1}$ with a contour interval of $10 \ ms^{-1}$) with decreasing initial condition spread going down. Center column: (e-h) Probability of exceedance $> 30 \ ms^{-1}$ (filled contour) within a 3 X 3 neighborhood corresponding to each horizontal grid point with probability-matched median contoured (black lines). Right column: (i - l) Spaghetti plot of column-max updraft at $30 \ ms^{-1}$ for each ensemble member.

in all three cases was qualitatively similar to the mid-level UH (Fig. 3.10e-g and Fig. 3.11e-g). This was not surprising given that the updraft location spread ought to be similar (if not the same) as the mid-level UH location spread. However, the spaghetti plots indicate that the updraft location within the ensemble is less than that of the mid-level UH. This is mainly related to the larger size of the $300 \ m^2 \ s^{-2}$ region compared to the $30 \ ms^{-1}$ region.

The 9 May updraft amplitude spread had no convergence similar to mid-level
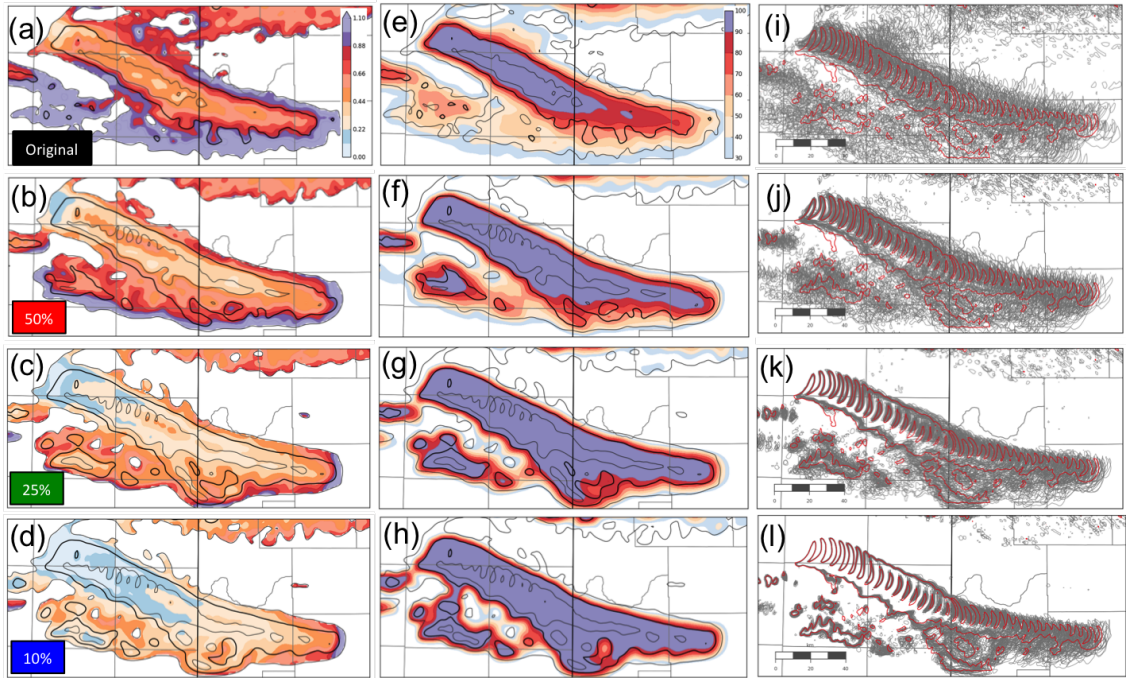
Figure 3.11: Same as Fig. 3.10 but for 16 May column-max updraft.
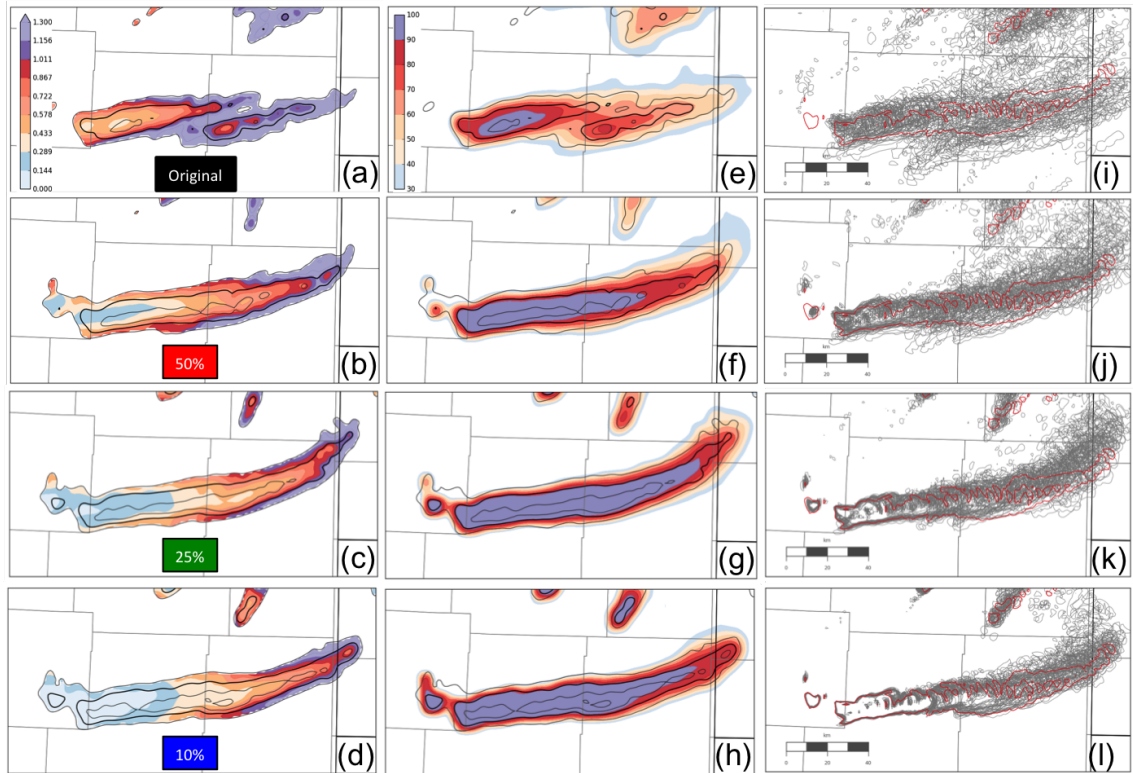
Figure 3.12: Same as Fig. 3.10 but for 24 May column-max updraft.

UH (Fig. 3.13a). The updraft amplitude spread convergence for the 16 May case can only be described by a range of 95 - 145 min (Fig. 3.13b). As for the 24 May case, Fig. 3.13c shows that the updraft amplitude has no spread convergence, but the 25% and 10% experiments are incredibly similar. In all three cases, the original 1-h amplitude forecast spread was delayed by approximately 100 min, more or less with a 50% reduction (similar to mid-level UH). For sake of comparison, Cintineo
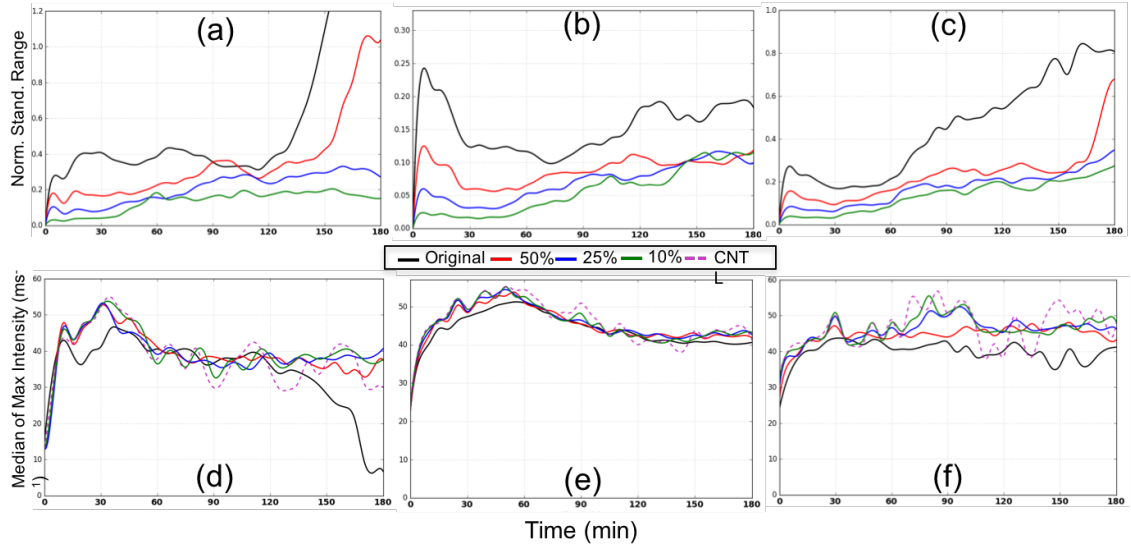
Figure 3.13: Same as Fig. 3.6 but for column-max updraft.

and Stensrud (2013) had a 12 ms^{-1} standard deviation (i.e., SR $\approx$ 24 ms^{-1}) in updraft speeds with 1-h errors (i.e., 1-h prior to CI). For the 9 and 16 May cases, the amplitude spread was less 24 ms^{-1} for the entire 3 h integration. As for 24 May, the updraft amplitude spread was only less than 24 ms^{-1} until the spread convergence point where the spread grew rapidly in magnitude. The smaller spread in this study could be related to difference in initialization (e.g., developing thunderstorm vs. prior to CI). The difference in updraft spread could also be related to the difference in average peak updraft speed between our study and theirs. One conclusion of Cintineo and Stensrud (2013) was that predictability decreases as the updraft speeds increase meaning that larger updraft speeds should have greater spread.

3.1.3 Low-Level Mesocyclone

As noted in section 1.3, a primary role for numerical guidance in the severe weather warning process is improving tornado forecasts. Operational model resolutions are far from resolving (or even remotely resolving) tornado-like vortices, so low-level vorticity and UH have been used as model proxies. At any rate, the low-level mesocyclone (LLM) has been neglected in most past supercell predictability studies (e.g., Cintineo and Stensrud 2013; Zhang et al. 2015, 2016; Miglietta et al. 2016).

Both May9_CNTL and May24_CNTL had two distinct periods of rapid LLM intensification (Fig.3.14a and Fig.3.14c). The second intensification in May24_CNTL produced large low-level vorticity (e.g., 0.05 s^{-1}) in the last 30-45 min (Fig. 3.14a and Fig. 3.14c). For 16 May, there were three LLM intensification periods. However, the low-level vorticity field eventually became muddled by vorticity generated near the gust front towards the latter half of the simulation (Fig. 3.14b). However, there were coherent pairs of positive and negative low-level vorticity to indicate low-level rotation.

Qualitatively, spatial distributions of normalized forecast spread of the low-level (i.e., 0-2 km AGL) vorticity were similar to mid-level UH and updraft for all three cases (Fig. 3.15 - 3.17a-d). Decreasing the IC spread substantially increased the probability of exceedance for the 9 May low-level vorticity. Even the 50% reduction in IC spread was able to substantially improve the handling of the early LLM intensification with a modest handling of the later LLM intensification. However, a 90% reduction in IC spread was needed to substantially improve the ensembles handling of the second development of the LLM. In the 16 May case, forecast spread along the strongest portion of the LLM (i.e., the northern part) was greatly reduced with no indication of spread convergence, unlike the previous features (Fig. 3.16a-d). The 16 May 50% experiment was able to improve the forecast by reducing the
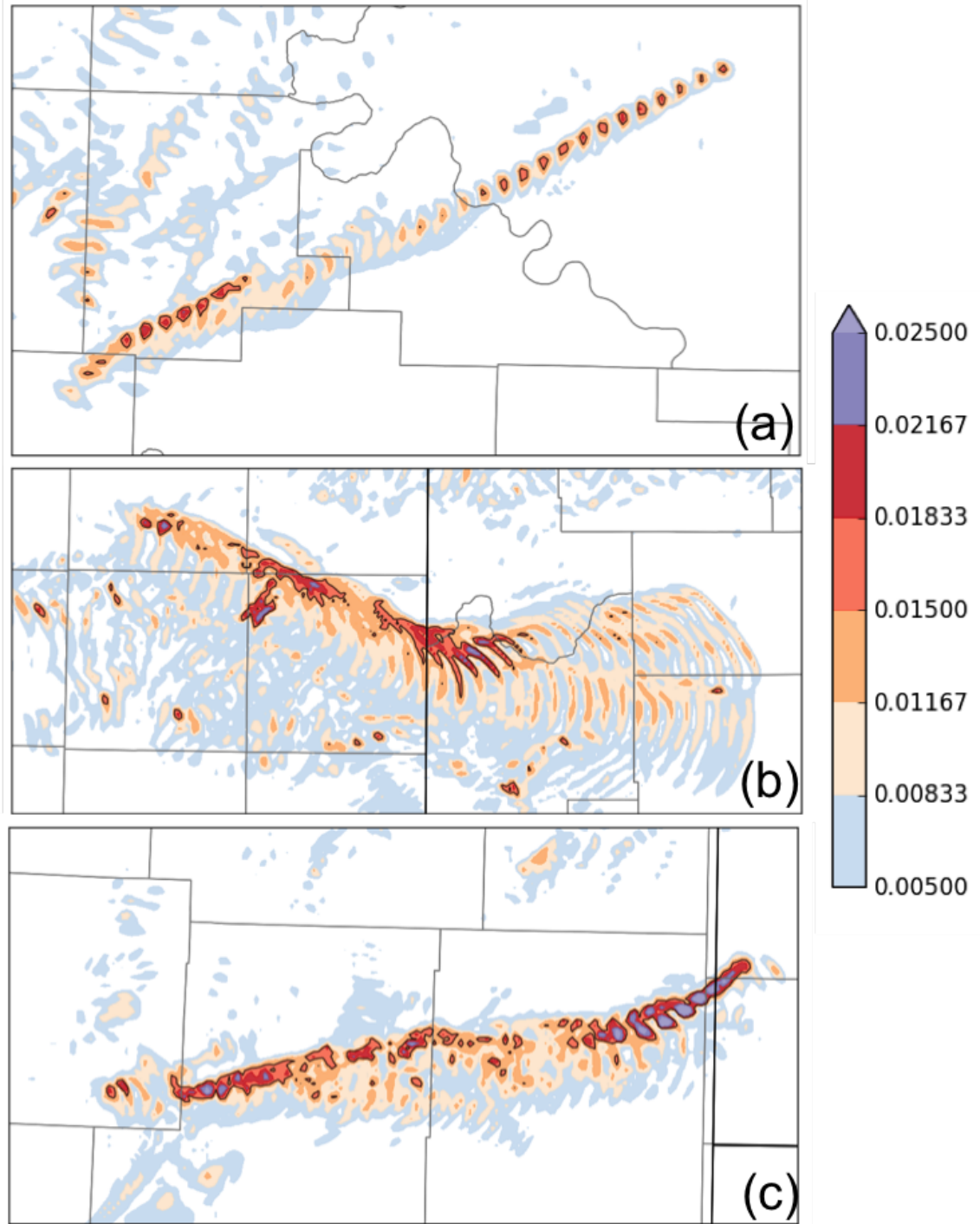
Figure 3.14: Control member evolution of the max 0-2 km AGL relative vertical vorticity for (a) 9 May, (b) 16 May, and (c) 24 May.

normalized spread and LLM position uncertainty. Probabilities were also increasing while the forecast the spread converged for the 24 May case as IC spread was

decreased, similar to the updraft and mid-level UH (Fig. 3.17e-g).

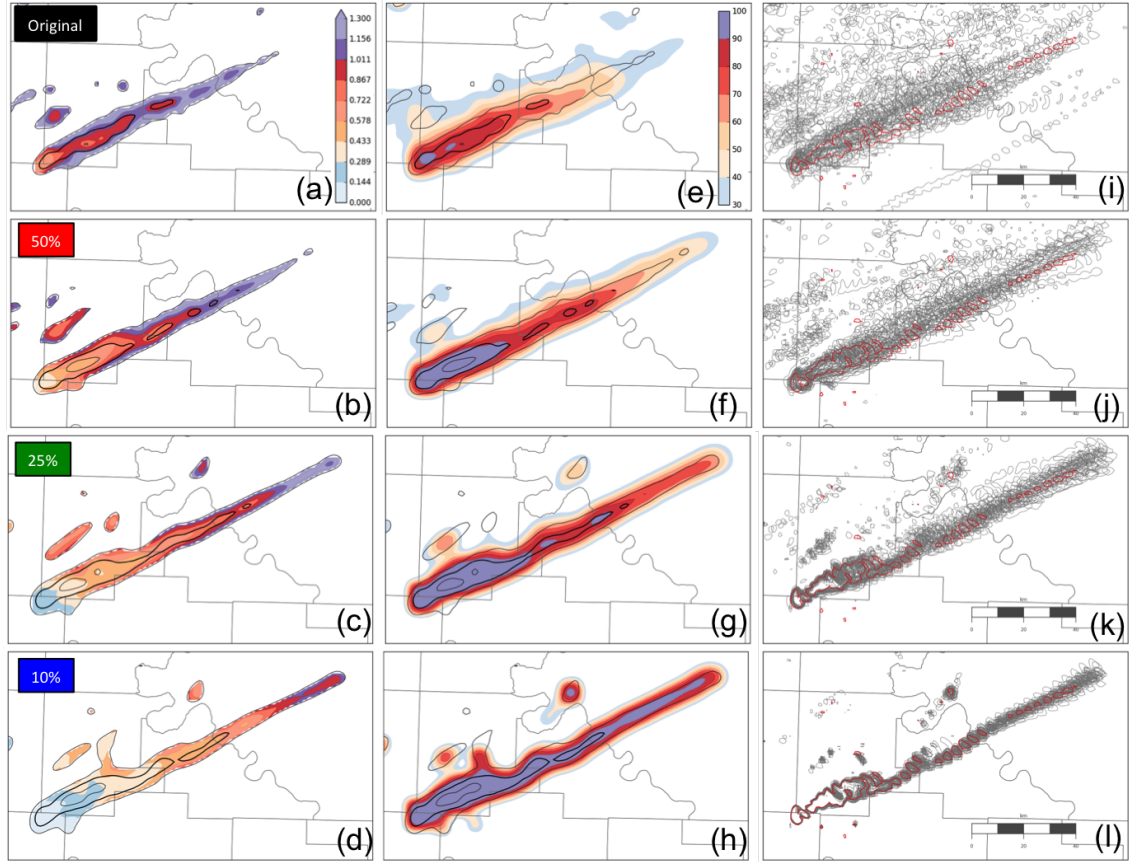Figure 3.15: 9 May max 0-2 km AGL relative vorticity. Left Column: (a-d) Normalized forecast spread swaths (filled contours) with probability-matched median contoured (black lines; thick line = 0.085 s^{-1} with a contour interval of 0.003 s^{-1}) with decreasing initial condition spread going down. Center column: (e-h) Probability of exceedance > 0.015 s^{-1} (filled contour) within a 3 X 3 neighborhood corresponding to each horizontal grid point with probability-matched median contoured (black lines). Right column: (i - l) Spaghetti plot of max 0-2 km AGL relative vorticity at 0.015 s^{-1} for each ensemble member.

As for the amplitude spread, Fig. 3.18b indicates there was no amplitude spread convergence for the 16 May case. However, the magnitude of the low-level vorticity associated with LLM in the 16 May case may be affected by baroclinic-generated vorticity associated with the gust front. As for the 24 May case (Fig. 3.18c), the amplitude spread convergence point was approximately t = 80 min, which is less than the previously examined features. This result from 24 May indicates that the
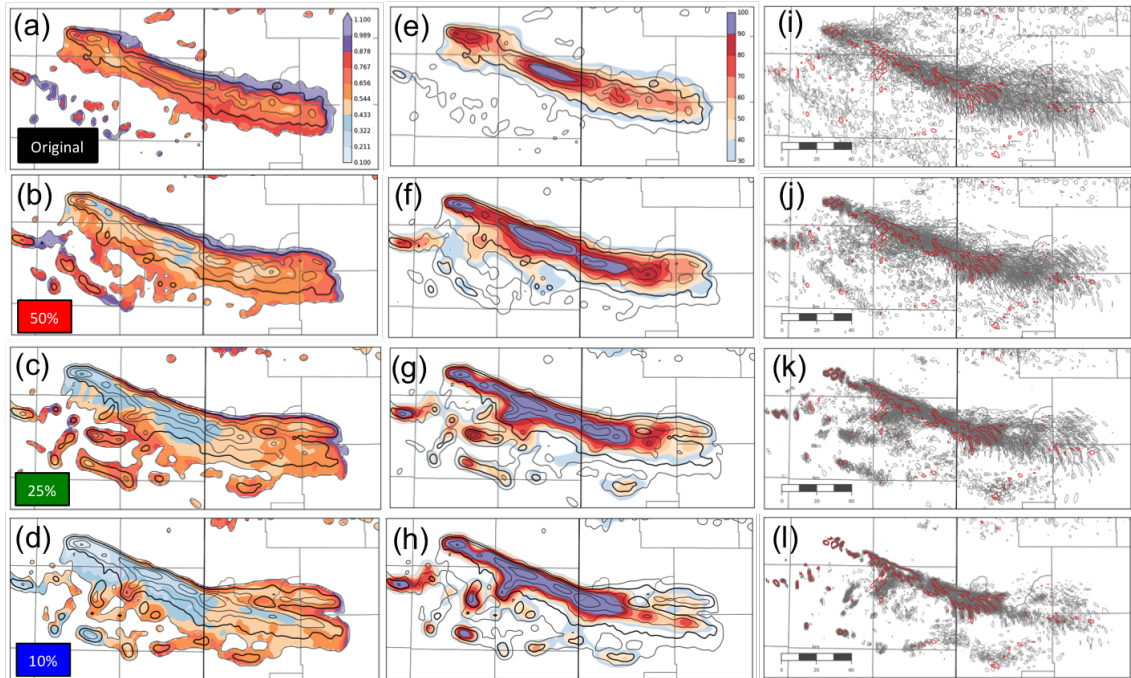
Figure 3.16: Same as in Fig. 3.15 but for 16 May max 0-2 km AGL relative vorticity.

predictability of the LLM, especially compared to the other features, is strongly case-dependent.

Improving ensemble forecasts of LLM forecasts is a primary goal (amongst many others) for WoF (Stensrud et al. 2009, 2013). Reducing the IC spread for 24 May reduced the LLM location spread beyond 3 h (see Fig. 3.18i-l), but the amplitude spread has a clear convergence point within 3 h. However, in the two other cases, there were substantial improvements in both location and amplitude spread by reducing IC spread. As for the practical predictability of the LLM, the results suggest that 1 h tornado forecasts will greatly benefit from IC spread reduction. However, the distinction between location and amplitude spread complicates how (distant) future tornado prediction (e.g., in WoF or otherwise) will be handled after 2 h. The LLM position can be substantially improved, but the intensity above a certain threshold can have considerable uncertainty. Ultimately, reducing IC spread will substantially improve ensemble forecasts of the LLM.
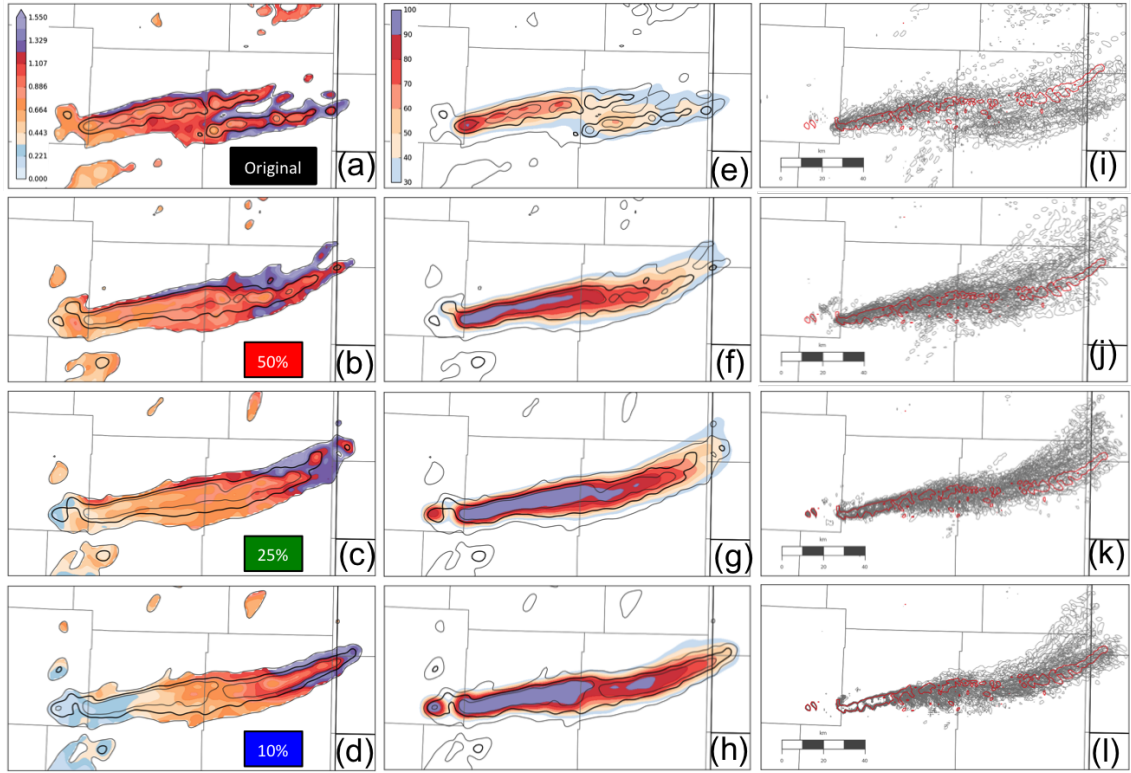
Figure 3.17: Same as in Fig. 3.15 but for 24 May max 0-2 km AGL relative vorticity.
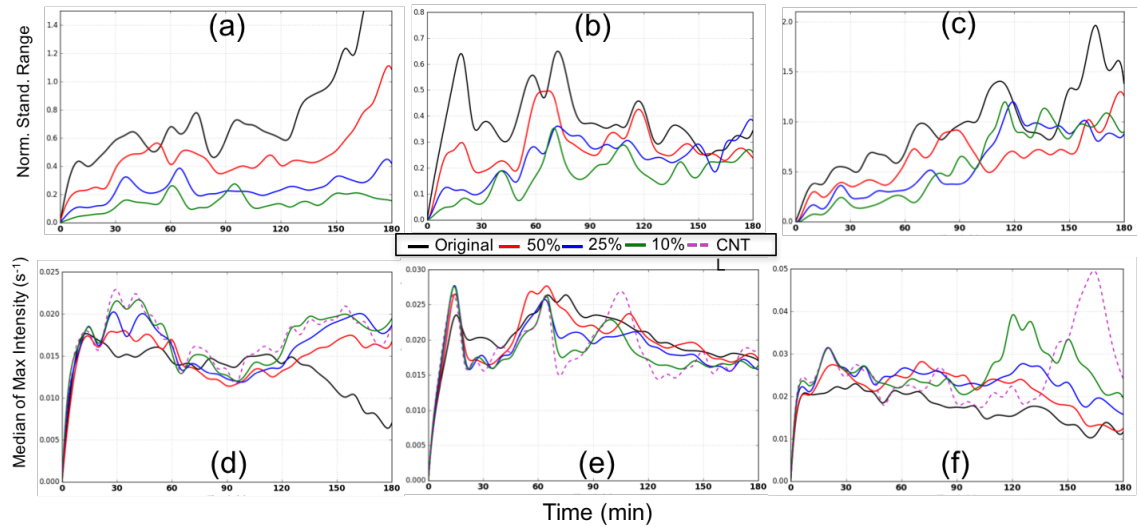
Figure 3.18: Same as in Fig. 3.6, but for max 0-2 km relative vorticity.

3.1.4 Downdraft

The downdraft is another feature that has been neglected in past supercell pre-
dictability studies. Neglection of the downdraft is surprising given how critical it is

for cold pool development and tornadogenesis. This neglect is understandable given that the downdraft is composed of two components, the forward- and rear-flank downdraft (FFD and RFD, respectively), which can difficult to separate and study individually. In this study, we will be using column-max downdraft and for the sake of simplicity will be making no distinction between the forward- and rear-flank downdraft (FFD and RFD, respectively).

The 9 May downdraft was fairly strong as the storm initially intensified. However, beyond that initial period, the downdraft was relatively weak (6-8 ms^{-1}) compared to the other two cases (Fig. 3.19a). 16 May had the strongest downdraft with the greatest spatial extent. This spatial extent of the downdraft was associated with a surging gust front, which lead to transition of the supercell into an MCS. The 24 May downdraft had a similar evolution as the updraft (cf. Fig. 3.19c and Fig. 3.9c).

The forecast spread reduction for the 9 May downdraft was quite dramatic and substantially greater than the previously examined features (Fig. 3.20a-d). Probability of exceedance substantially increased with a 50% IC spread reduction with no discernible increase as IC spread was further reduced. This is primarily related to the substantial location spread reduction as seen in Fig. 3.20i-l. As for the 16 May case, forecast spread reductions were qualitatively less than in the previously examined features (see Fig. 3.21a-d). In Fig. 3.21e-f, the probability of exceedance slightly increased with a 50% IC spread reduction with no discernible increase as the IC spread reduced further (similar to 9 May).

As for the 24 May case, downdraft forecast spread reductions were similar to the updraft (see Fig. 3.22e-g). However, the methodprobability of exceedance increases with IC spread reduction were greater than the updraft. This continues to support the idea that for the 24 May case, location predictability was greater than amplitude predictability. Although location predictability was greater than amplitude predictability for the 24 May case, this was not a general conclusion. Fig.3.21i-l

58
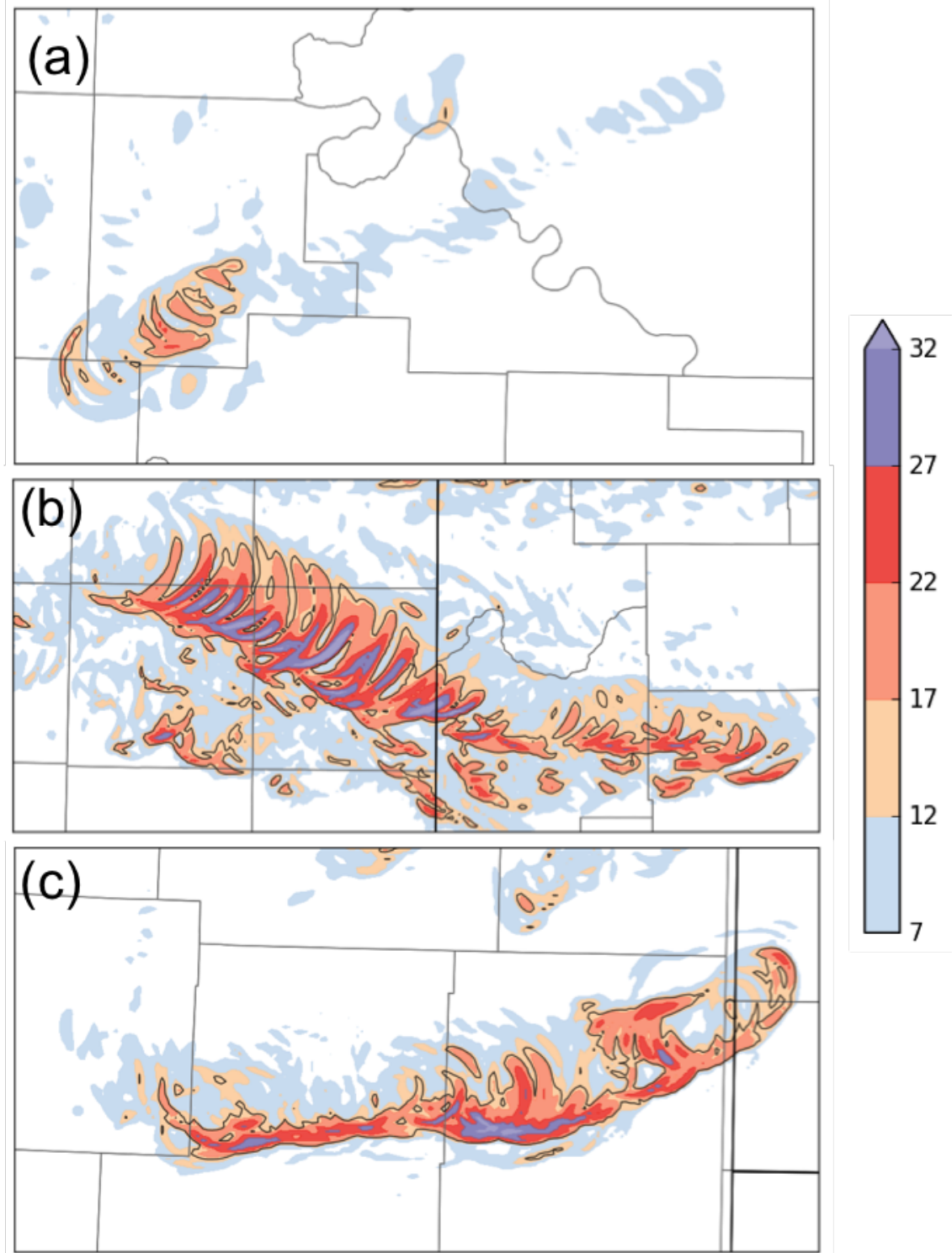
Figure 3.19: Control member evolution of column-max downdraft for (a) 9 May, (b) 16 May, and (c) 24 May.

shows that downdraft location spread has clearly converged for the 16 May case. However, the amplitude spread predictability limit for the 16 May case was between 90-120 min meaning the location predictability limit was lesser than or equal to the amplitude predictability limit. Thus, the greater of the two, location versus amplitude predictability, is case-dependent.
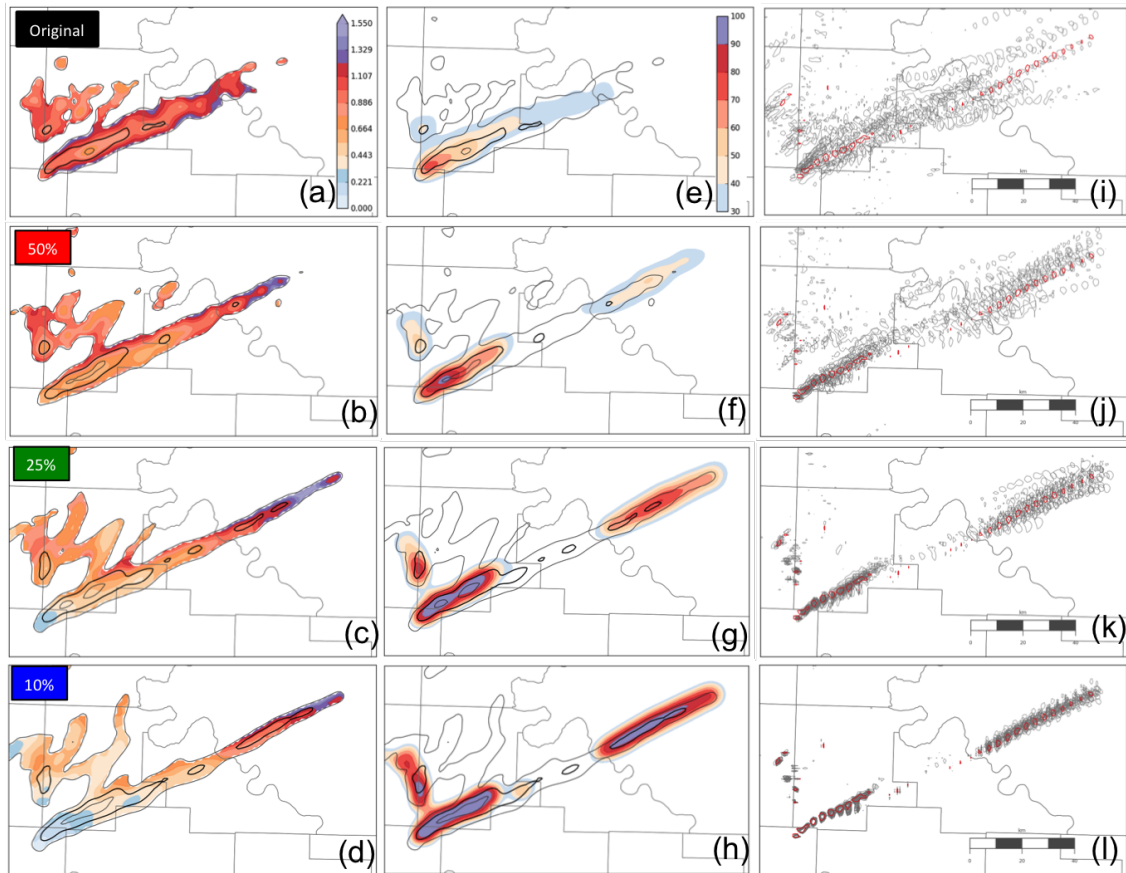

Figure 3.20: 9 May column-max downdraft. Left Column: (a-d) Normalized forecast spread swaths (filled contours) with probability-matched median contoured (black lines; thick line $= 12$ ms^{-1} with a contour interval of 4 ms^{-1}) with decreasing initial condition spread going down. Center column: (e-h) Probability of exceedance > 15 ms^{-1} (filled contour) within a 3 X 3 neighborhood corresponding to each horizontal grid point with probability-matched median contoured (black lines). Right column: (i - l) Spaghetti plot of column-max downdraft at 15 ms^{-1} for each ensemble member.
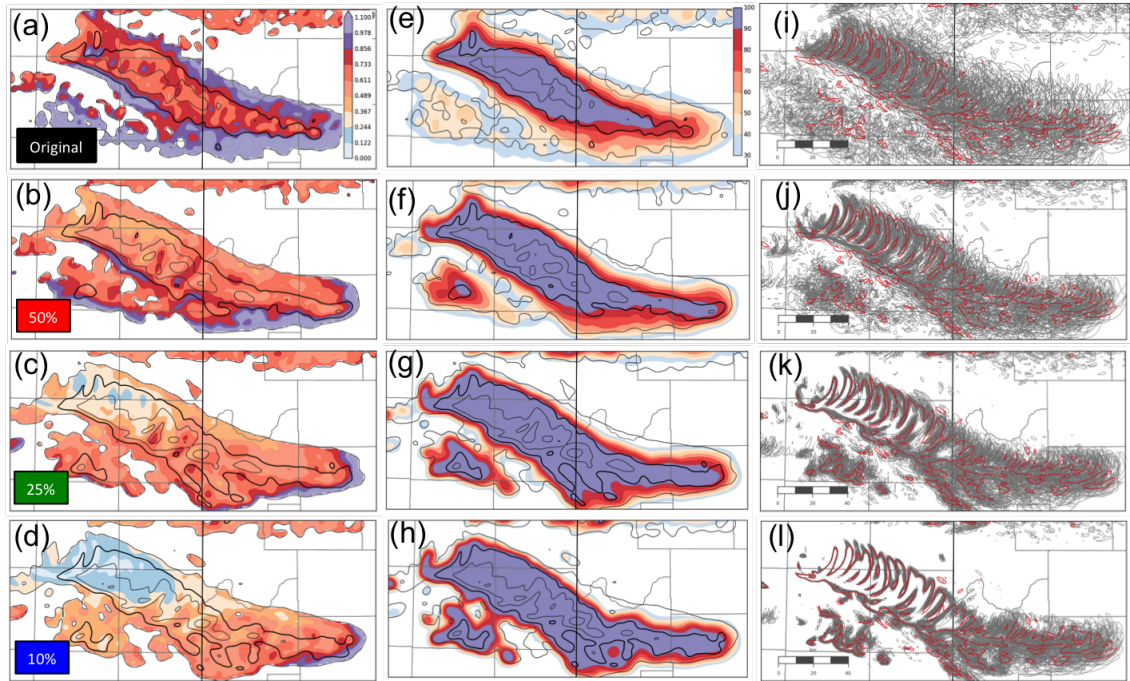
Figure 3.21: Same as Fig. 3.20, but for 16 May column-max downdraft. Thick lines $= 14 \ ms^{-1}$, contour interval $= 6 \ ms^{-1}$.

3.1.5 Rainfall

Intense rainfall is a general severe thunderstorm feature that impacts humans and property directly. Though supercells are responsible for most tornadoes, some can produce torrential rainfall and flash flood hazards. Given that the amount, location, and rate of accumulation rainfall from severe thunderstorms (such as supercells) are among the most important quantities in storm-scale prediction, several studies have focused exclusively on the predictability of it, both at the mesoscale and storm-scale (Park 1999; Walser et al. 2004; Zhang et al. 2006; Hohenegger et al. 2006; Done et al. 2012; Surcel et al. 2016). Understanding the predictability of the individual features is useful information, but from an operational-minded standpoint, having an understanding of the end products (such as rainfall) is more valuable information.

The 9 May supercell did not produce a substantial amount of rainfall as seen in Fig. 3.24. There is a small region of 3 h accumulate rainfall >1 in with 0.5 in
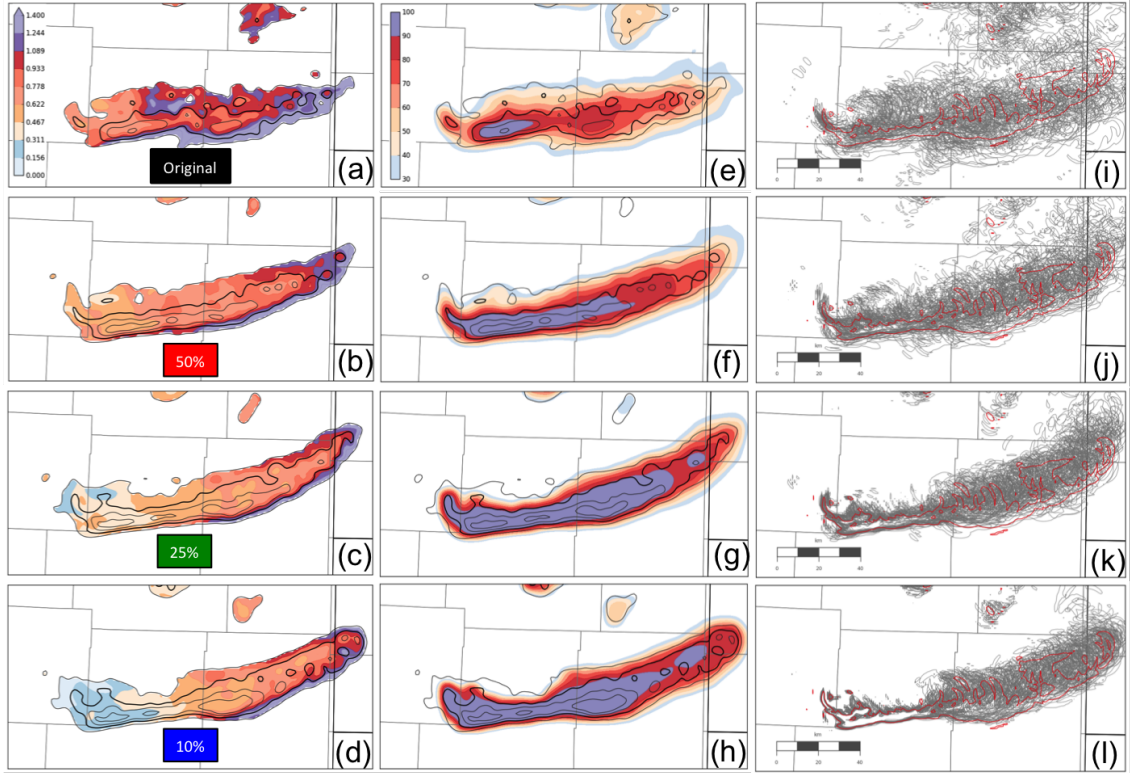
Figure 3.22: Same as Fig. 3.20, but for 16 May column-max downdraft. Thick lines = 12 ms^{-1}, contour interval = 4 ms^{-1}.
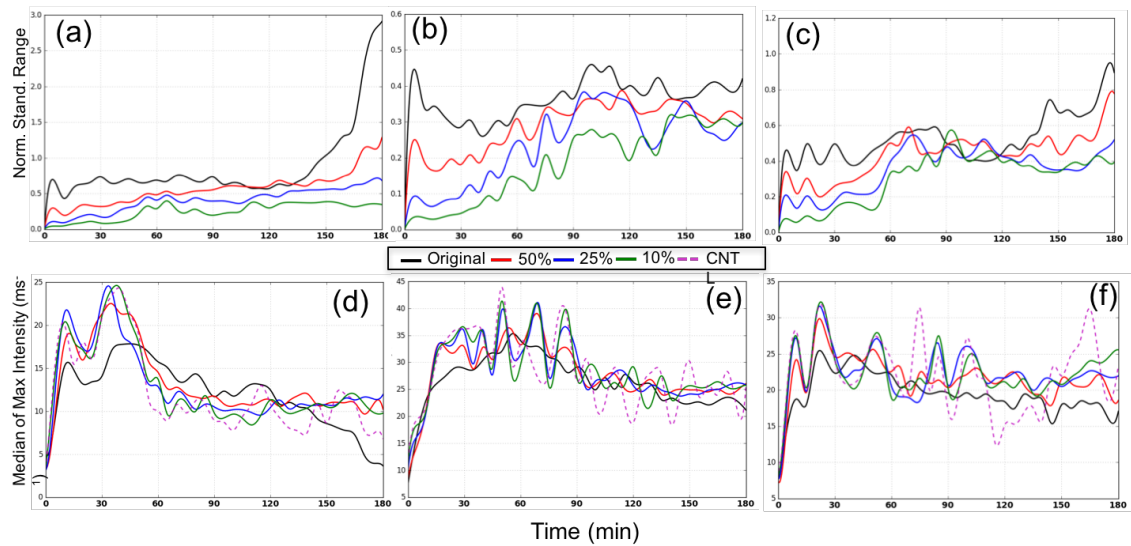
Figure 3.23: Same as Fig. 3.6, but for column-max downdraft.

elsewhere near the storm of interest. With this in mind, there were substantial reductions in normalized 3-h accumulated rainfall spread (Fig. 3.25). The majority

of the forecast spread reduction is related to the ensemble becoming surer about the lack of rain. Becoming more certain about the absence of a feature is incredibly important. In term of operational implications, reducing false alarm warning is just as important as increasing warnings when an event does happen. As for 5-min peak rainfall amplitude spread, there was no convergence point (Fig. 3.29). This means the timing of heaviest rainfall and the overall magnitude and location did not have an intrinsic predictability for the 9 May case.   16 May produced the most rainfall compared to the other two cases (Fig. 3.24). Given the isolation of the storms in the other two cases and the transition of the 16 May supercell into an MCS, this is not surprising. Reducing the IC spread substantially reduced the forecast spread in 3-h accumulated rainfall. Similar to the LLM, there was no indication of a spread convergence point. Analogous to the previous features, there was no distinguishable probability of exceedance differences between each experiment. However, unlike the other features, rainfall location was greatly improving based on the spaghetti plots. Given that this may be related to the low threshold (i.e., > 1 in) of 3-h accumulated rainfall, the probability of exceeding 2 in of 3-h accumulated rainfall is presented in Fig. 3.27. With the higher threshold, there are visible improvements even between the 25% and 10% experiments.

In all three cases, the amplitude predictability is relatively similar to that of Park (1999) who found convective rainfall predictability limits to be greater than 140 min (see 3.29). Even though the other features had a nearly similar saturation point for 16 May, there is arguably no convergence point for 5-min rainfall amplitude spread. Since rainfall had no convergence point for 16 May, it can be considered the most predictable quantity for this storm (this will be discussed further in Ch. 4). Overall, the results suggest that in two of the three cases, location was the greatest source of error consistent with past studies (Park 1999; Yussouf et al. 2016).
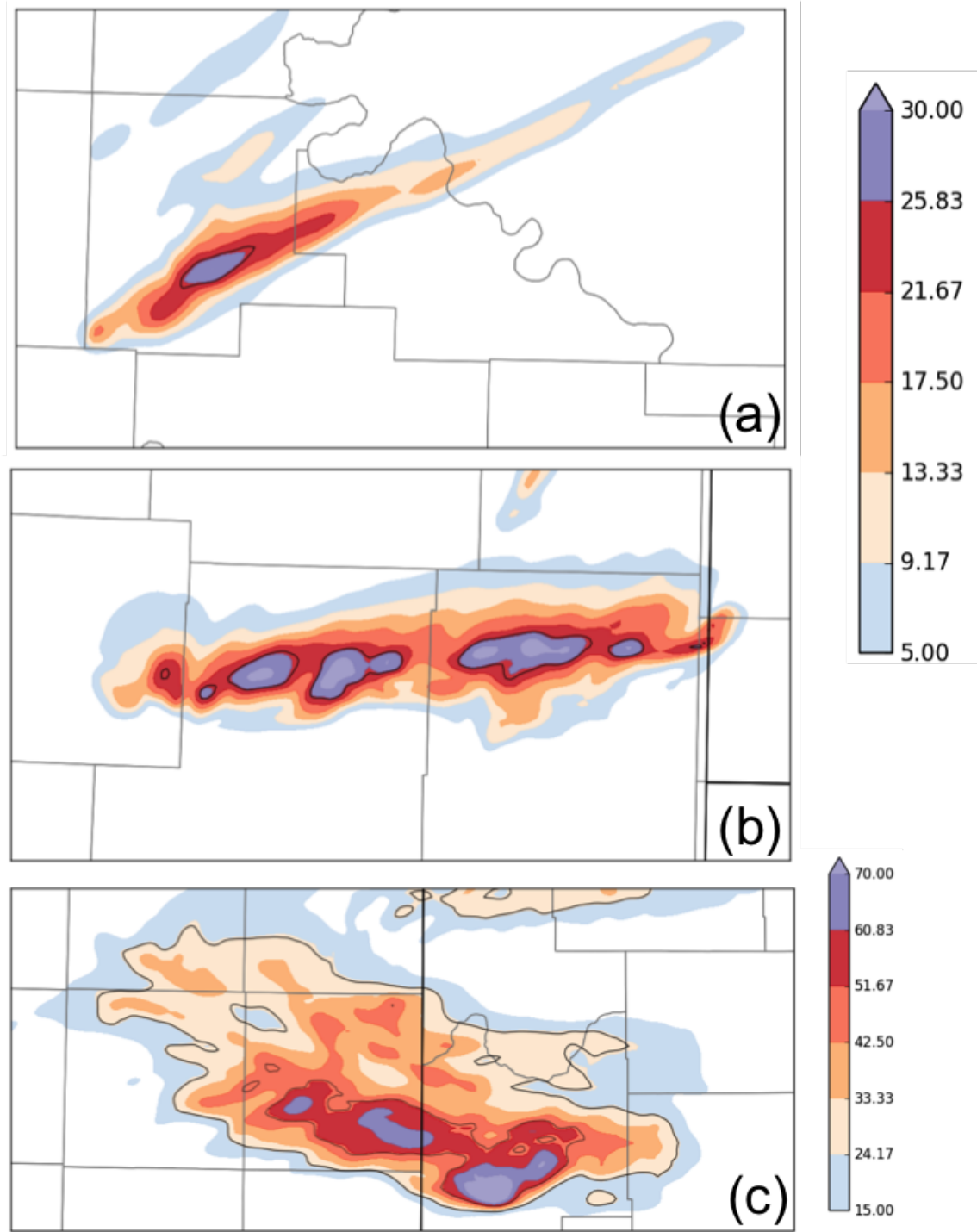
Figure 3.24: Control member 3-h accumulated rainfall for (a) 9 May, (b) 16 May, and (c) 24 May.
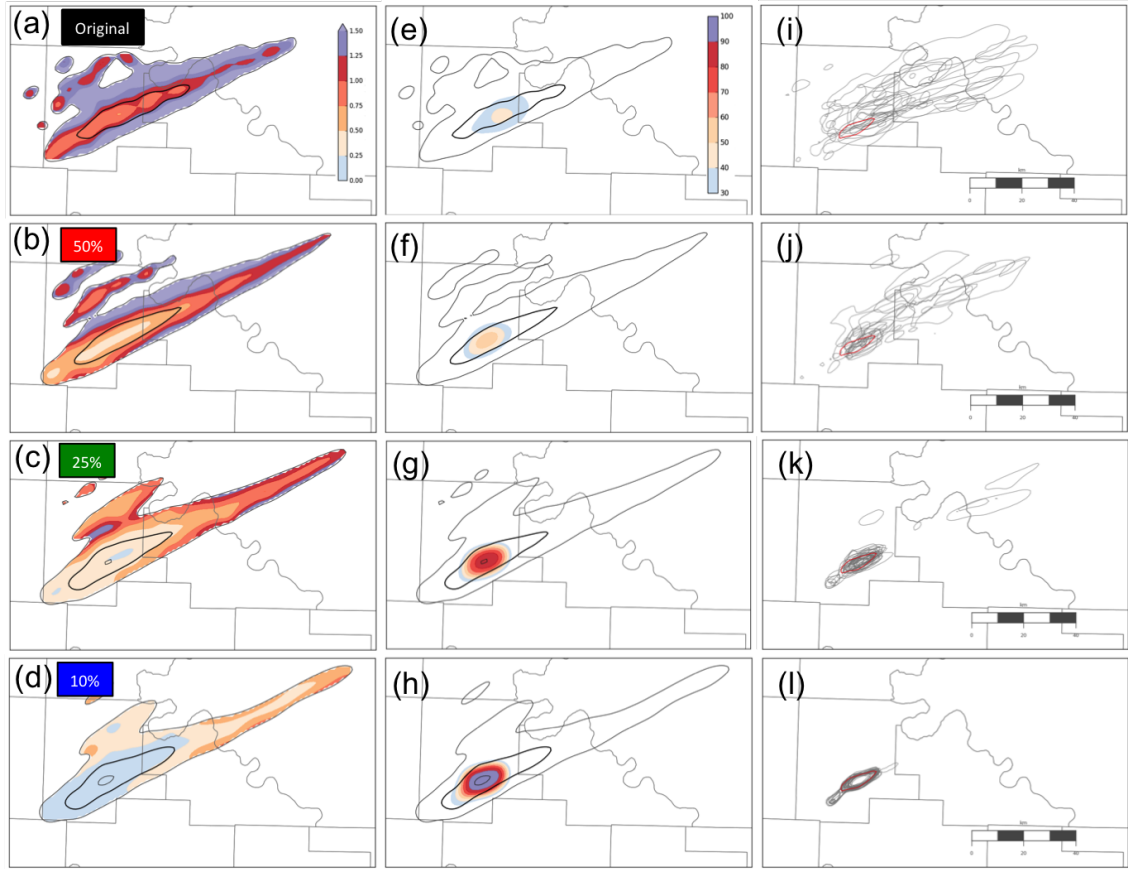
Figure 3.25: 9 May 3-h accumulated rainfall. Left Column: (a-d) Normalized forecast spread swaths (filled contours) with probability-matched median contoured (black lines; thick line $= 15\ mm$ with a contour interval of $10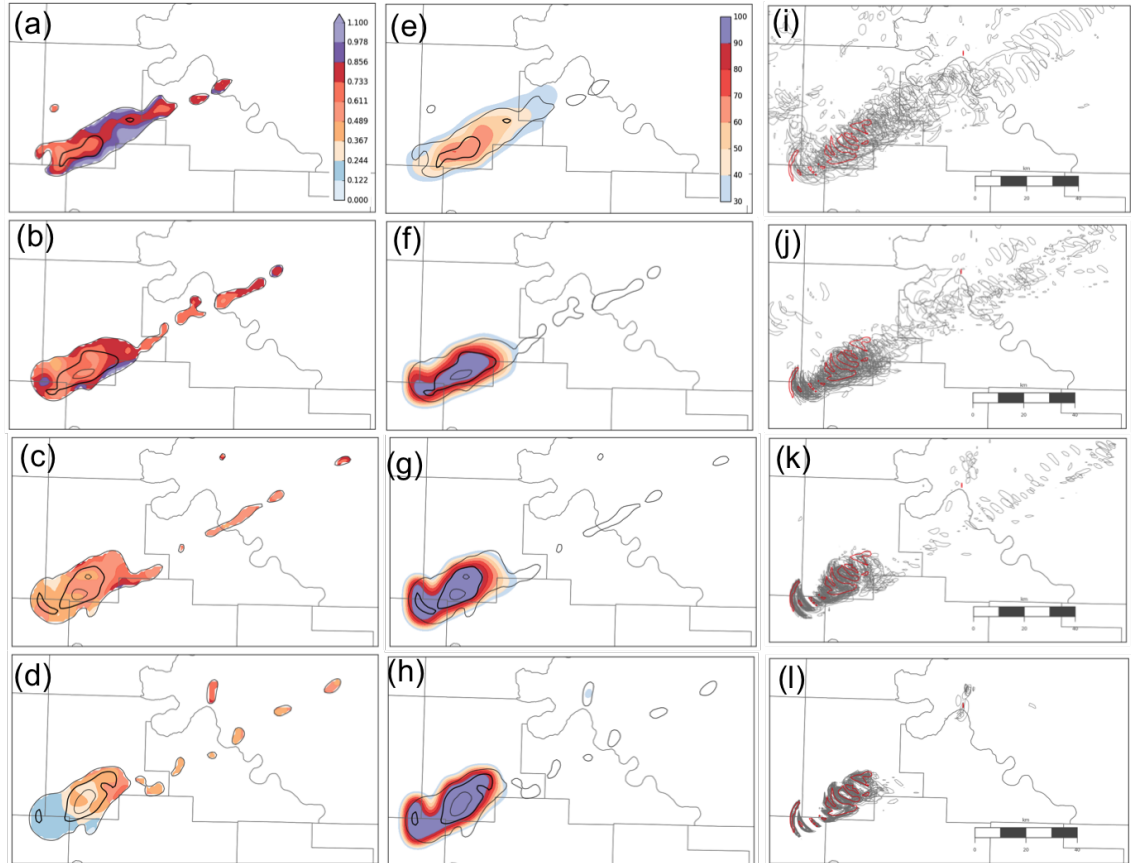\ mm$) with decreasing initial condition spread going down. Center column: (e-h) Probability of exceedance $> 25\ mm$ (filled contour) within a 3 X 3 neighborhood corresponding to each horizontal grid point with probability-matched median contoured (black lines). Right column: (i - l) Spaghetti plot of 3-h accumulated rainfall at $25\ mm$ for each ensemble member.

3.1.6 Severe Winds

Severe thunderstorm wind damage is more common than damage from tornadoes and accounts for half of all severe reports in the contiguous US (NSSL, 2017). Therefore, the ability to forecast severe winds is crucial for the WoF paradigm. However, the ability of the model to resolve severe wind speeds is often in question (sources). Recently, Potvin and Flora (2015) concluded that a 1-km grid spacing was able to
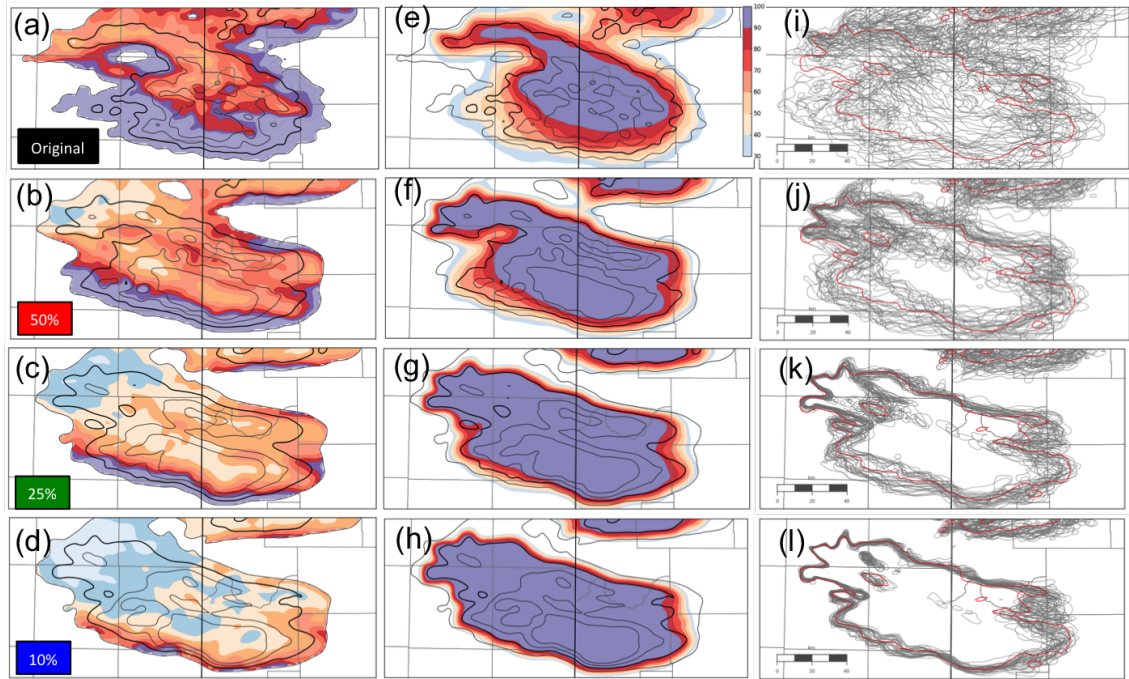
Figure 3.26: Same as Fig. 3.2, but for 16 May 3-h accumulated rainfall. Thick lines = 25 mm , contour interval = 10 mm.

resolve the wind speeds fairly well. Furthermore, max wind speeds $> 30 \ ms^{-1}$ in the 16 and 24 May case also supported a literal interpretation of the winds.

With that in mind, May9_CNTL did not produce severe winds. However, as previously discussed in the Rainfall section, reducing false alarm warnings is also crucial and other ensemble members did produce severe winds as seen in Fig. 3.30. The strongest winds were towards the beginning of the simulation due to the rapid intensification of the storm (see Fig. 3.30). With that in mind, the dramatic reduction in forecast spread as the IC spread is reduced as seen in Fig. 3.31 was expected.

As for the 16 May case, there were significant reductions in forecast spread with IC spread reduction (Fig. 3.32). However, there is an implied convergence point that occurs much sooner than for the other features. One of the processes that lead to the 16 May supercell growing upscale was a strong, surging gust front in association with a deep cold pool. In the initial development of the primary storm
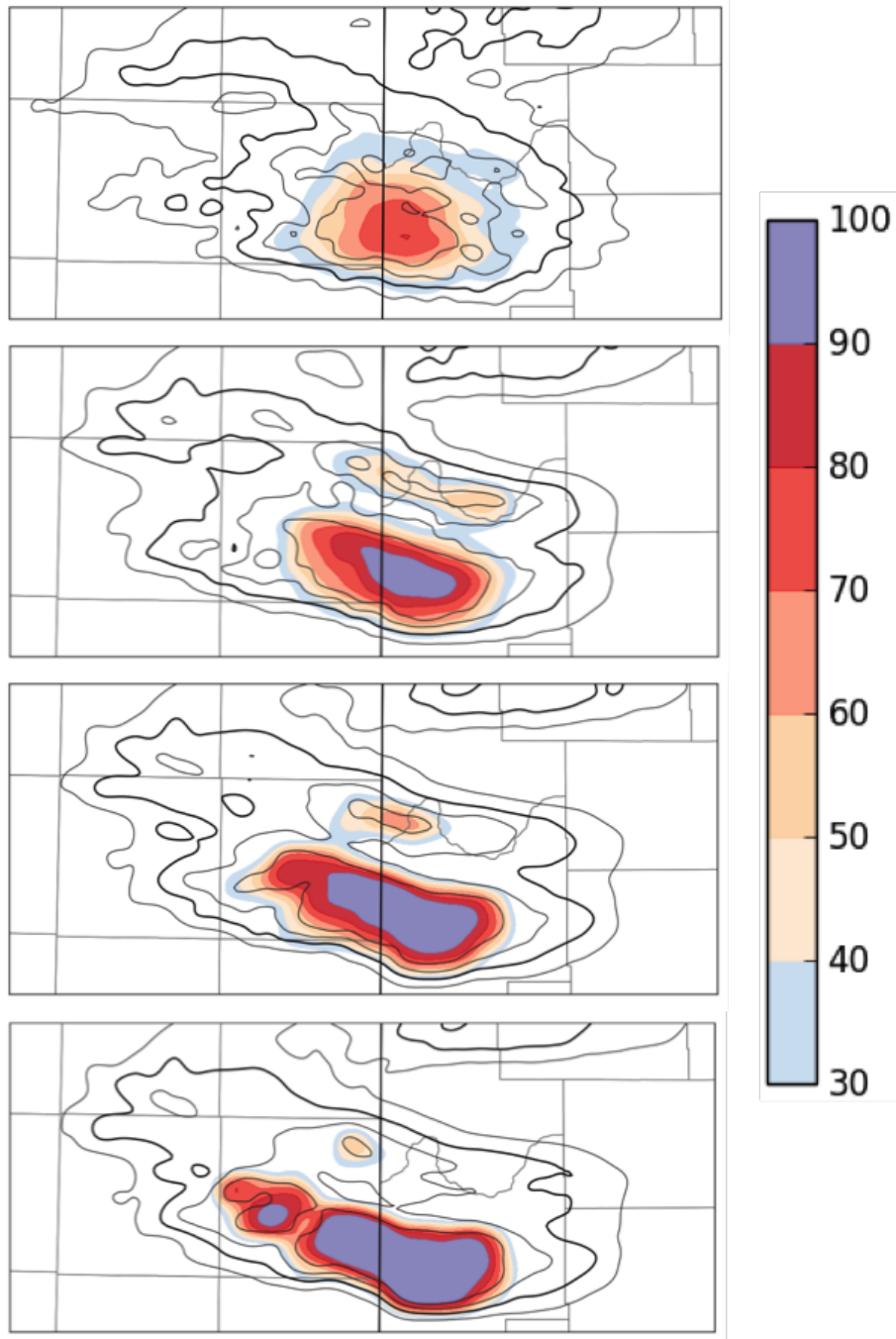
Figure 3.27: Probability of 3-h accumulated rainfall > 2 in for 16 May with decreasing IC spread going down.

of interest, there were relative strong winds. As the storm developed there appears to stronger winds on the southern flank of the storm, but some portion is due to secondary storms. Even so, the probability of exceedance completely fails to capture
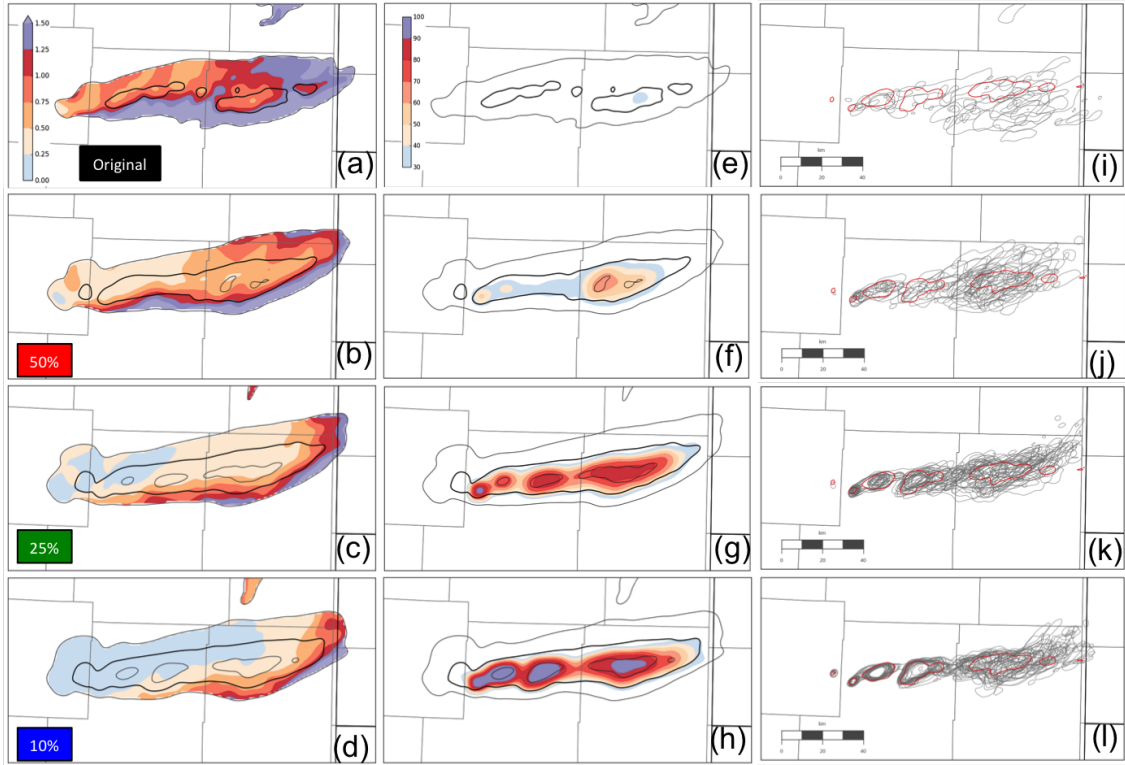
Figure 3.28: Same as Fig. 3.2, but for 24 May 3-h accumulated rainfall.
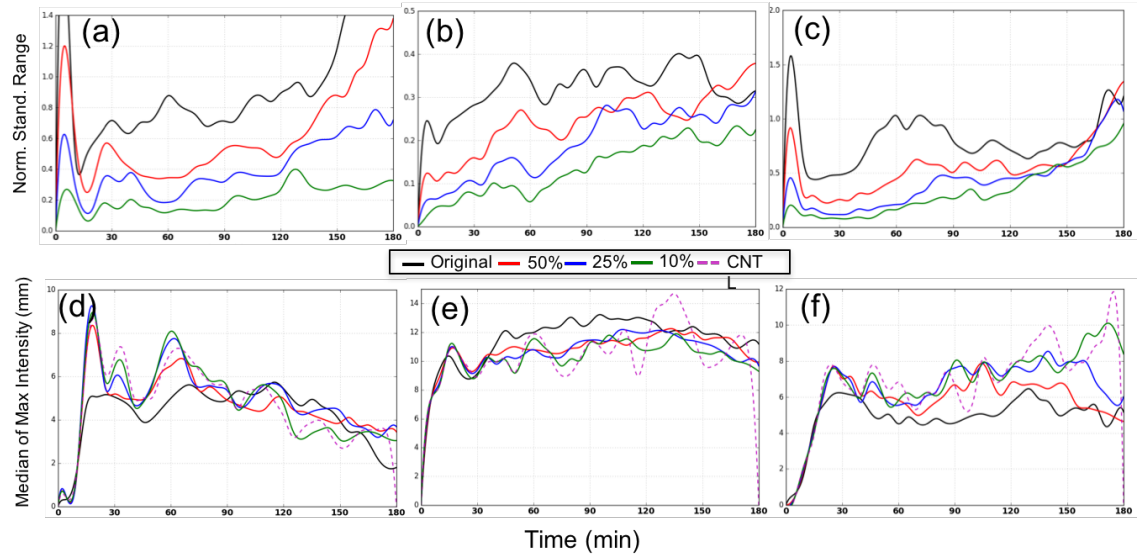
Figure 3.29: Same as Fig. 3.6, but for 5-min max rainfall.

these strong winds. This suggests that in certain cases severe winds may be the least predictability supercell feature.
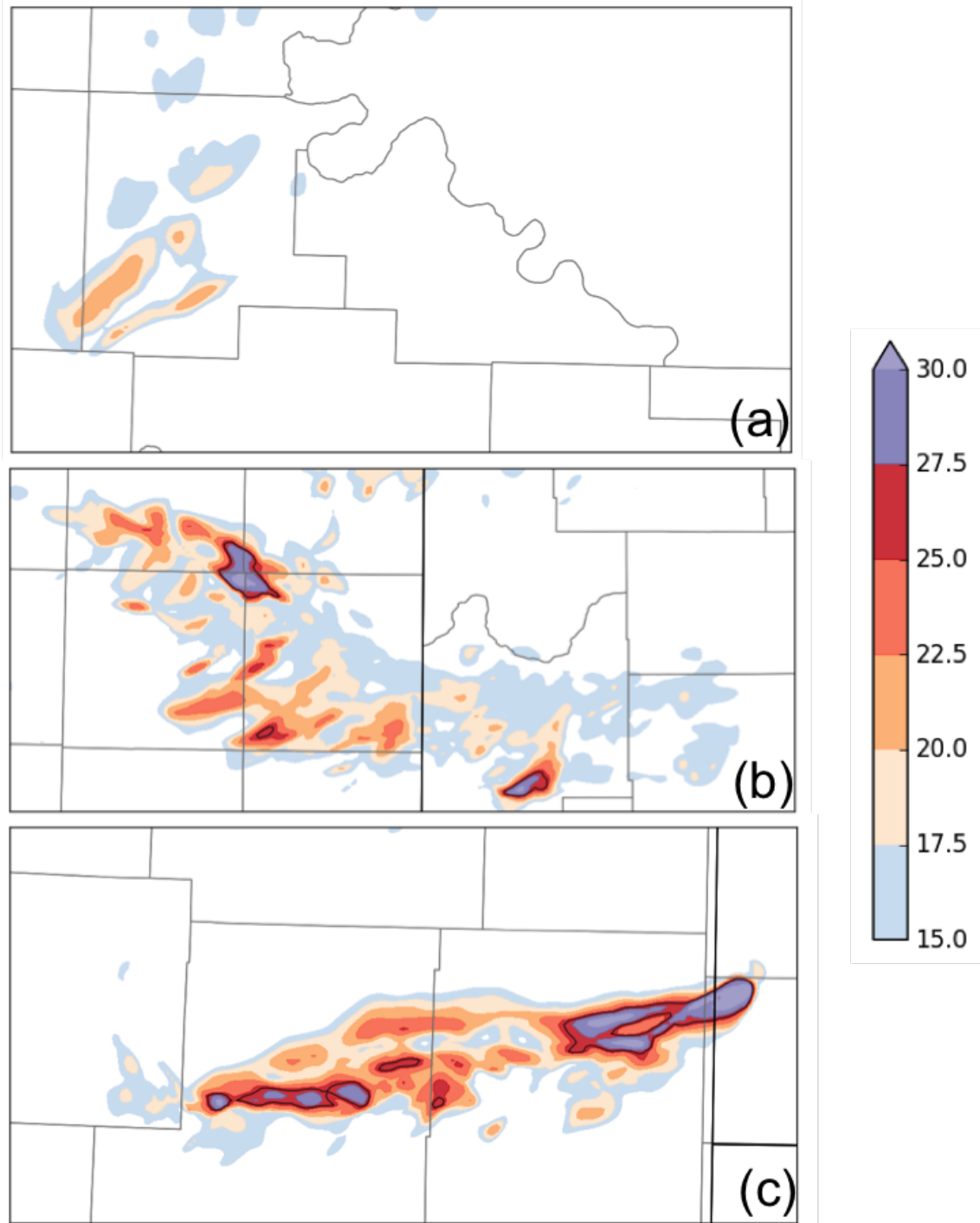
Figure 3.30: Control member evolution for max 10-m AGL winds for (a) 9 May, (b) 16 May, and (c) 24 May.

Similar to the 16 May case, the 24 May case also had a substantial reduction in forecast spread with IC spread reduction (see Fig. 3.33). There were also significant
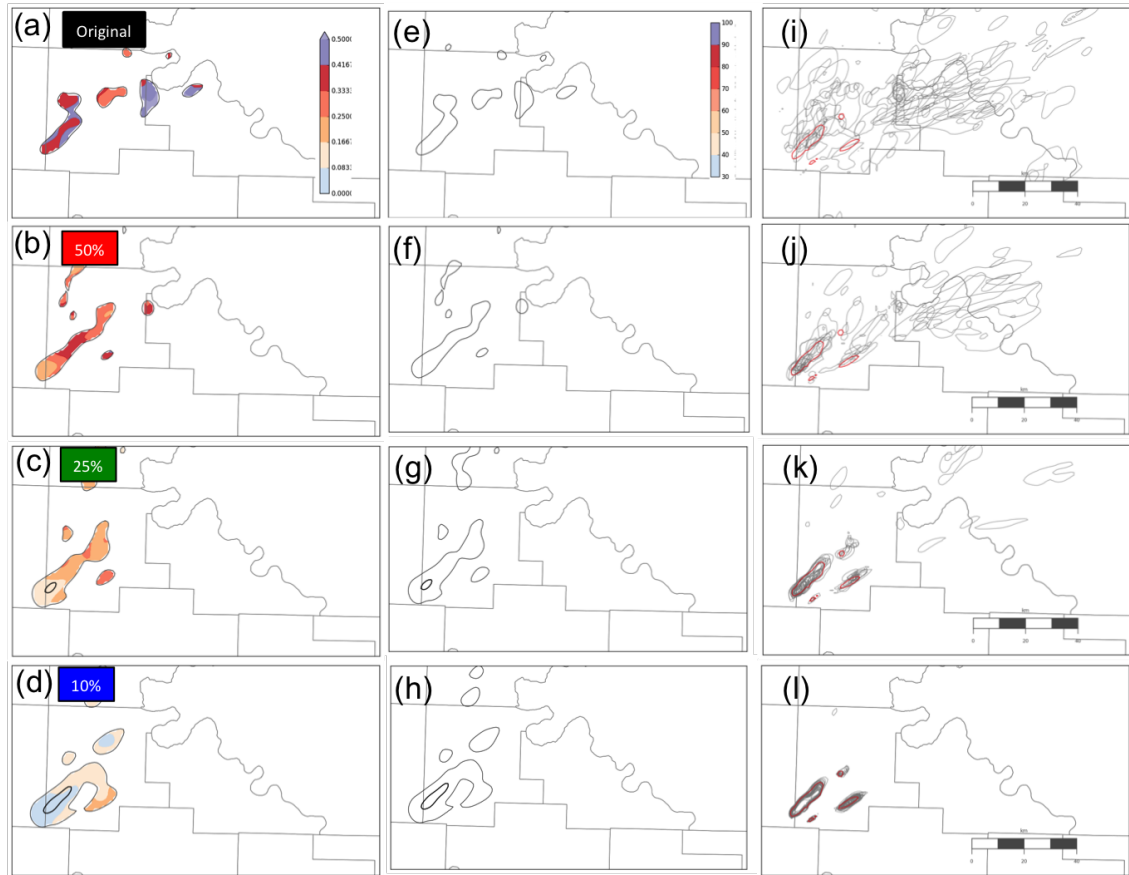
Figure 3.31: 9 May 10-m max wind speeds. Left Column: (a-d) Normalized forecast spread swaths (filled contours) with probability-matched median contoured (black lines; thick line = 19 ms^{-1} with a contour interval of 4 ms^{-1}) with decreasing initial condition spread going down. Center column: (e-h) Probability of exceedance > 26 ms^{-1} (filled contour) within a 3 X 3 neighborhood corresponding to each horizontal grid point with probability-matched median contoured (black lines). Right column: (i - l) Spaghetti plot of 10-m max wind speeds at 26 ms^{-1} for each ensemble member.

improvements in the probability of exceedance. Operationally, this that forecasts out to 3 h would benefit from IC spread reductions. Forecasters could issue severe wind warnings since the ensemble is confident that the winds are greater than 26 ms^{-1}. Unfortunately, intrinsic amplitude predictability would limit the degree of certainty in the magnitude of the severe winds.

As for amplitude spread (Fig. 3.34), the 16 May case had a relatively small normalized amplitude spread, but convergence still occurred around t = 60 min.
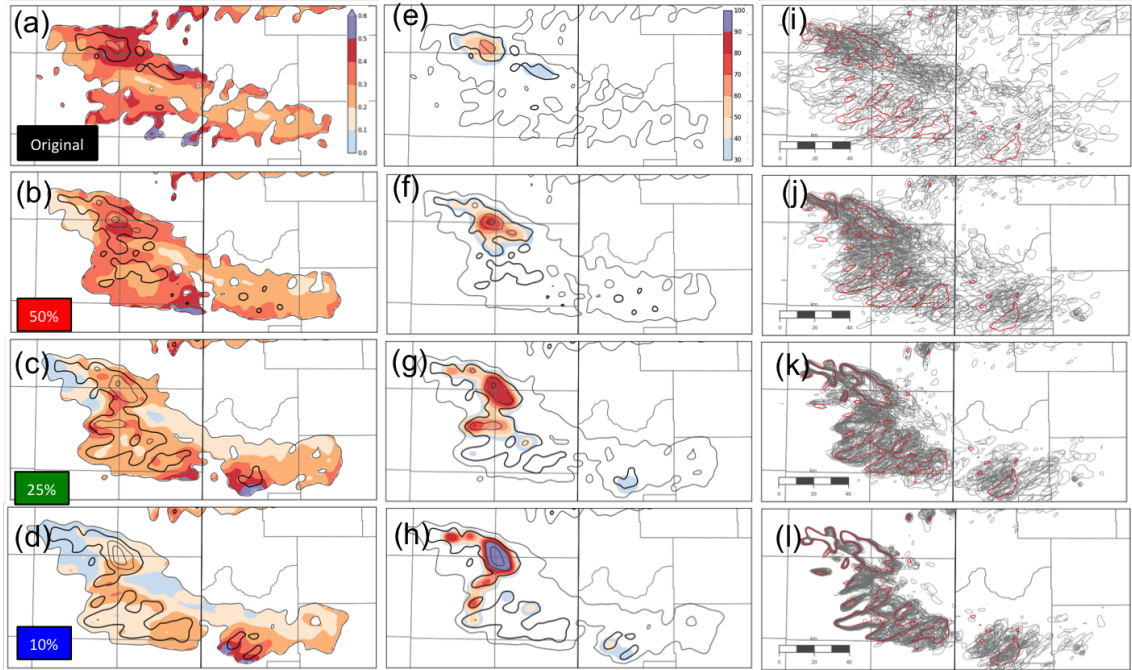
Figure 3.32: Same as Fig. 3.2, but for 16 May max 10-m AGL winds. Thick lines $= 19\ ms^{-1}$, contour interval $= 4\ ms^{-1}$.

This supports the idea aforementioned that severe winds, in this case, were the least predictability. As for 24 May, there were strong similarities to the other features with an amplitude predictability at t = 100 min.

3.2 Initial Condition Spread in the Storm vs. Environment Experiments

The relationship between forecast spread and IC spread within the storm versus in the storm environment is not properly understood. Therefore, we will be evaluating the effect of IC spread inside and outside of the storm on ensemble forecasts of supercell evolution. Therefore, two experiments set-ups were used to explore: reducing internal storm spread without reducing environment spread and vice versa. To reduce internal storm spread (i.e., having complete storm certainty), each ensemble member was initialized with the control member (i.e. no perturbations) for the primary storm of interest within a sub-domain centered on the storm. As for
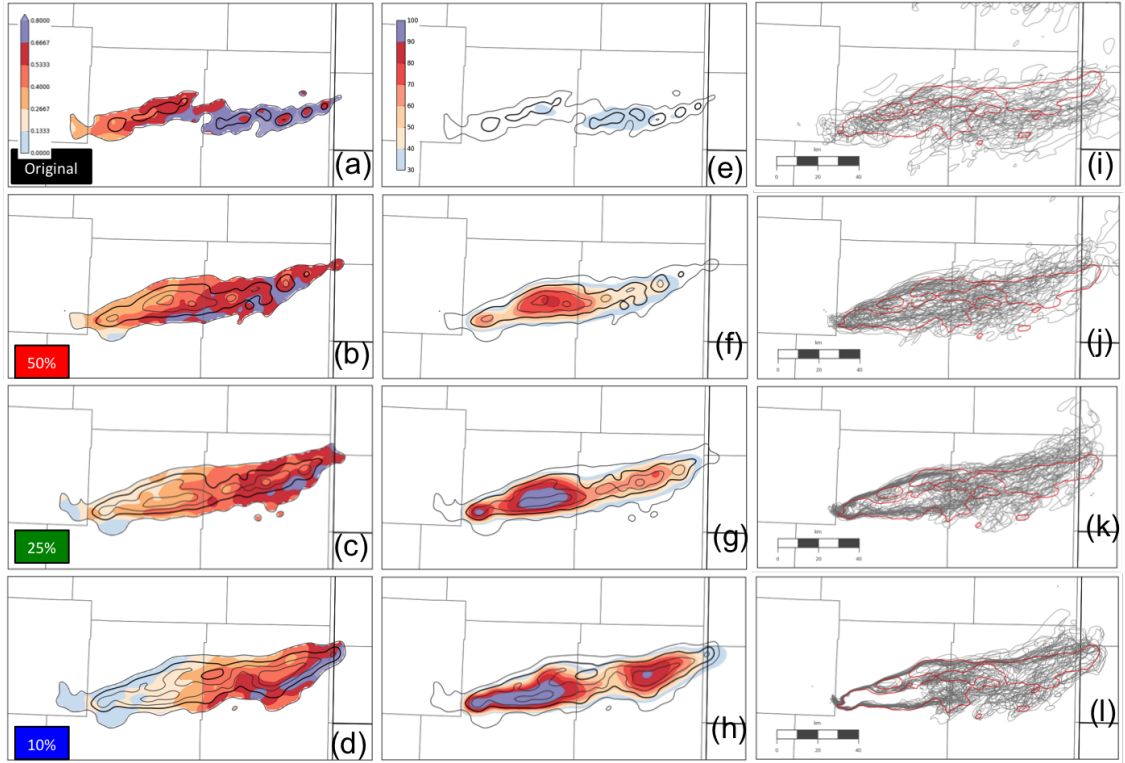
Figure 3.33: Same as Fig. 3.2, but for 24 May max 10-m AGL winds. Thick lines = 19 ms^{-1}, contour interval = 19 ms^{-1}.
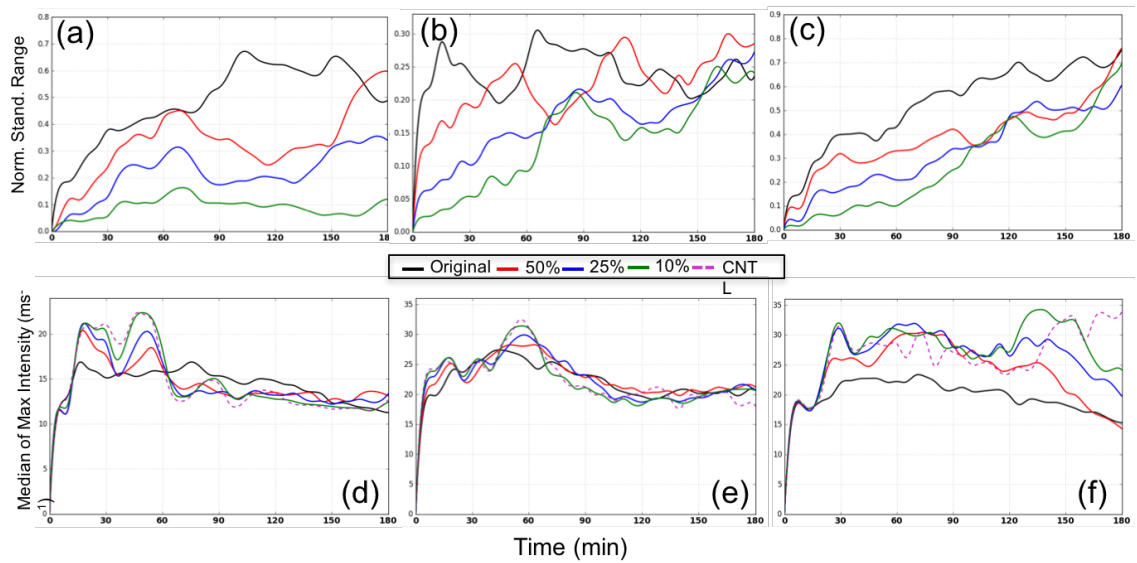
Figure 3.34: Same as Fig. 3.6, but for max 10-m AGL winds.

reducing the storm and mesoscale environment spread (i.e., having complete environment certainty) each ensemble member was initialized with the control member outside the subdomain with the storm of interest region remaining unchanged. The two experiments test the effect of a single storm in multiple different environments vs. multiple storms in a single environment.

As mentioned in section 2.3.2, there is novelty with these experiments, but an arguably similar line of work was done in Johnson and Wang (2016). In there study, they evaluated the effects of the scale of IC perturbations (e.g., from small to large scale) where the current practice is to introduce only large scale IC perturbations. They found that including small-scale IC perturbations had a noticeable impact on reflectivity forecasts at very short lead times. Assuming that small-scale processes are primarily inside the storm then their results indicate that storm evolution is sensitive to the intra-storm IC. As we will see, consistent with the results of Johnson and Wang (2016), reducing intra-storm uncertainty resulted in a corresponding decrease in forecast uncertainty in our experiments.

3.2.1 Mid-Level Mesocyclone

Eliminating perturbations inside the storm significantly reduced normalized mid-level UH spread in the 9 May case (Fig. 3.35c). However, rapid spread growth led to limited benefits (i.e., well within 3 h) as the spread equalled that of the original ensemble. Reducing perturbations outside the storm (i.e., in the environment), resulted in the forecast spread being qualitatively similar to the original ensemble throughout the integration (Fig. 3.35b).

As for the probability of mid-level UH $>$ 300 $m^2\ s^{-2}$, the greatest benefit was from reducing perturbations inside the storm (Fig. 3.35f). Looking at Fig. 3.35e, we can see that reducing environmental perturbations did increase exceedance proba-
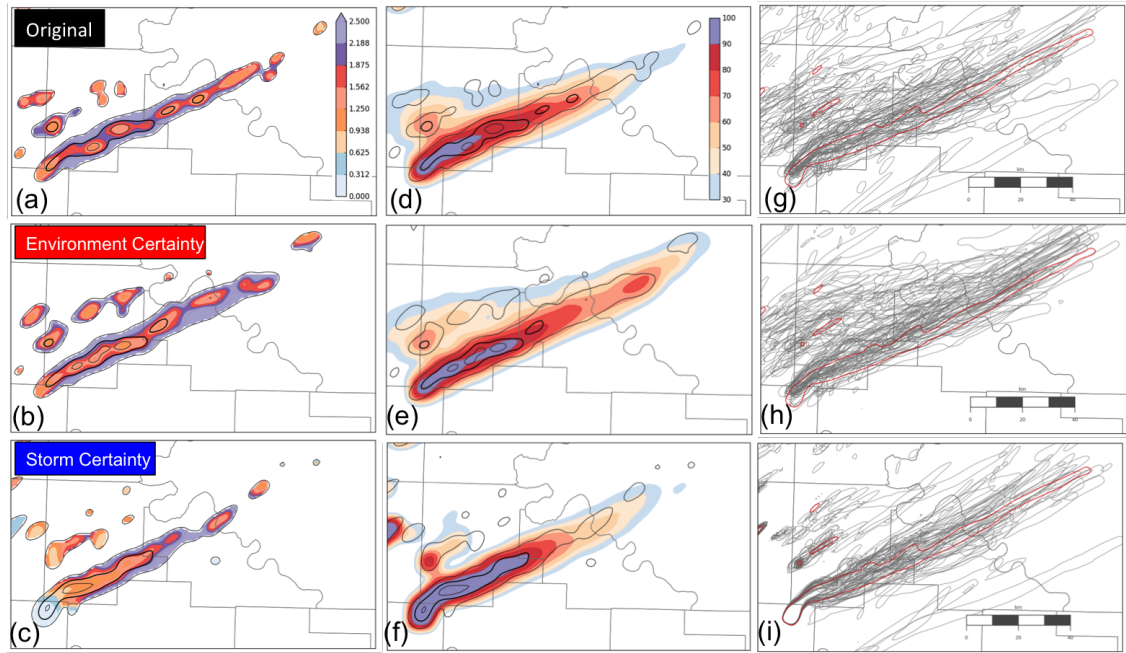
Figure 3.35: 9 May 2-5 km AGL updraft helicity. Left Column: (a-c) Normalized forecast spread swaths (filled contours) with probability-matched median contoured (black lines; thick line = 350 m^2s^{-2} with a contour interval of 250 m^2s^{-2}). Center column: (d-f) Probability of exceedance > 300 m^2s^{-2} (filled contour) within a 3 X 3 neighborhood corresponding to each horizontal grid point with probability-matched median contoured (black lines). Right column: (g - i) Spaghetti plot of 2-5 km AGL updraft helicity at 300 m^2s^{-2} for each ensemble member.
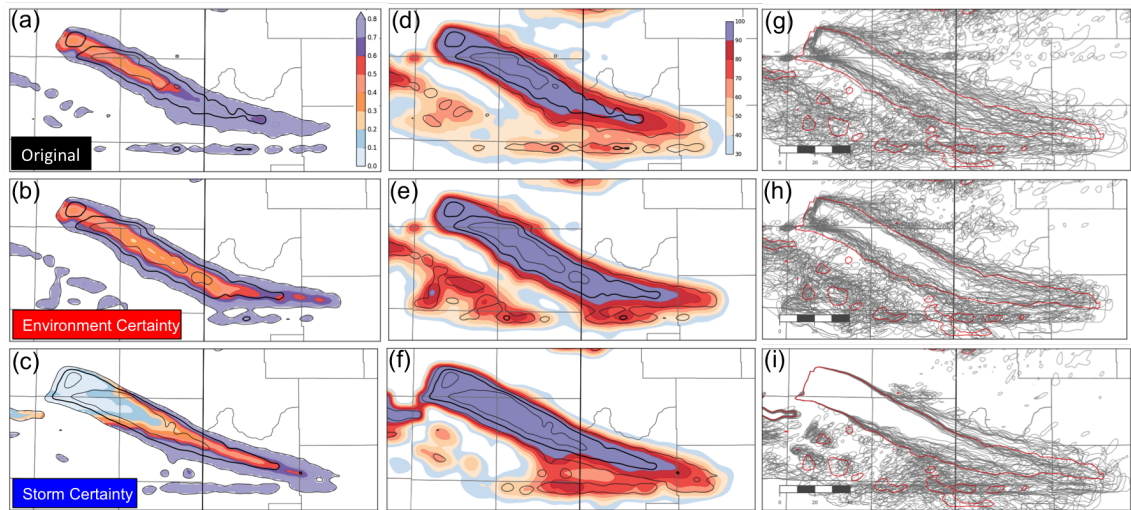
Figure 3.36: Same as in Fig. 3.35, but for 16 May 2-5 km AGL updraft helicity. Thick line= 450 $m^2 s^2$ with a contour interval of 200 $m^2 s^{-2}$.
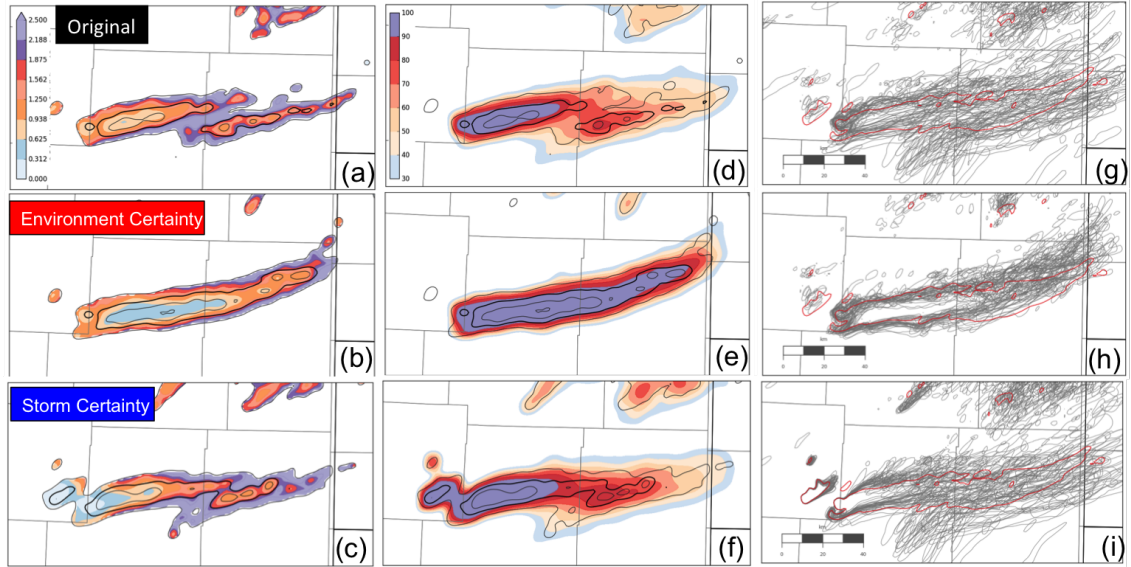
Figure 3.37: Same as in Fig. 3.35, but for 24 May 2-5 km AGL updraft helicity.

bilities but only near the end of the integration period. This increase was primarily related to a greater number of overlapping mid-level UH paths. However, some portion of the overlap may be due to paths of secondary storms. In all three cases, there were secondary storms that interacted (e.g., collisions) to some degree with the primary storms in some of their respective ensemble members.

Although the secondary storms were not a primary focus, neglecting their effects on the results is not possible. Reducing the perturbation inside the primary storm reduced additional storms in the 9 May (Fig. 3.35i) case while minimizing environmental perturbations did not. It was also found that improving certainty inside the storm eliminated several pop-up storms (i.e., non-supercellular) that were directly impacting the primary storm of interest (not shown). Given that the secondary storms were perfectly known (i.e., no IC uncertainty) in the environment certainty experiments, the results are surprising. These secondary storms were also evolving in the perfectly known environment, and although it is not shown, there spread was comparable to the primary storm of interest. For a further discussion on the secondary storms in all three cases see Ch. 4.

In the 16 May case, improving certainty inside the storm also significantly reduced normalized mid-level UH spread (Fig. 3.36c). Unlike the 9 May case, though, improving confidence in the environment ICs made a discernible reduction in forecast spread (Fig. 3.36b). This was true especially towards the latter half of the integration period. The increased IC certainty inside the 16 May storm also led to substantial increases in the probability of mid-level UH $>$ 300 m^2 s^{-2} (Fig. 3.36c). However, after an initial period (t 45 min), the probabilities were relatively indistinguishable between both experiments (i.e., storm vs. environment certainty).

As discussed in section 3.1.1, the probability of exceedance is somewhat sensitive to the location spread. Thus, the cause of the probabilities being indistinguishable is evident in spaghetti plots of 300 m^2 s^{-2} (Fig. 3.36h and i). The qualitative spread in MLM path in both experiments are becoming indistinguishable from each other. As a final note, the implied location spread convergence in these experiments is at a similar point as in the original experiments. This further supports the conclusion that the intrinsic location predictability limit for the 16 May case was well within 3 h.

The results of 9 and 16 May were dissimilar on the effects of environment certainty, but 24 May was an outlier compared to both. Unlike the 9 and 16 May cases, eliminating perturbations in the environment substantially reduced the normalized mid-level UH forecast spread (Fig. 3.37b). In fact, the reductions in forecast spread were comparable to the 50% and 25% experiments from the original experiments (cf. Fig. 3.4b-c). Some of this may be related to the dramatic change in ensemble bias. Looking at Fig. 3.37a and Fig. 3.37b, there is a huge difference in the probability-matched median contours. This change may also be related to lack of a secondary storm , which will be discussed shortly. As for reducing the IC uncertainty inside the storm, the benefits appear to be relatively short-lived (Fig. 3.37c). These differences between the two experiments for the 24 May case was also reflected in the

probability of exceedance (cf. Fig. 3.37e and f).

Reducing IC uncertainty in the environment did result in substantial spread reduction for the 24 May MLM. However, qualitatively, the spread eventually became similar to that in the original experiments. This further supports the conclusion that the diagnosed spread convergence (in the original experiments) indicates the approximate intrinsic predictability limit. Arguably, the average perturbation magnitudes with the environment known perfectly were much less the average perturbation magnitudes in the 10 % experiments. Therefore, the similarity in the 2D spread swaths between the 10 % experiment and the experiment with the environment known perfectly is a strong indication of the intrinsic predictability limit for the 24 May case.

Be that as it may, the cause of unpredictability in the 24 May case remains unclear. Unlike the 9 May case, reducing environment uncertainty eliminated the secondary, erroneous storm in the original 24 May ensemble. However, as mentioned above, spread convergence still occurred. This indicates that the secondary, erroneous storm did not have a substantial impact on 24 May's predictability. For potential causes of 24 Mays limited predictability see Ch. 4.

Recall that from the original experiments that in the 9 and 24 May case the implied location predictability limit was beyond 3 h while in the 16 May case it was well within 3 h. However, it is unclear how perturbations in the storm or the environment impact storm motion, and therefore location. In Fig. 3.38, we can see the quantification of the location spread for all three cases. In the 9 May case (Fig. 3.38a), storm certainty reduces the median distance between ensemble members greater over environment certainty (but < 5 km). Beyond an 1 h, both environment and storm certainty similarly decrease the median distance between ensemble members compared to the original ensemble. For the 16 May (Fig. 3.38b), storm certainty dramatically reduces the distance between ensemble members (slightly >

5 km) while environment certainty only slightly reduces the median distance after an 1h. Finally, environment certainty substantially reduces the median distance between ensemble members by as much as 10 km in the last 40 min of the simulation (Fig. 3.38c).
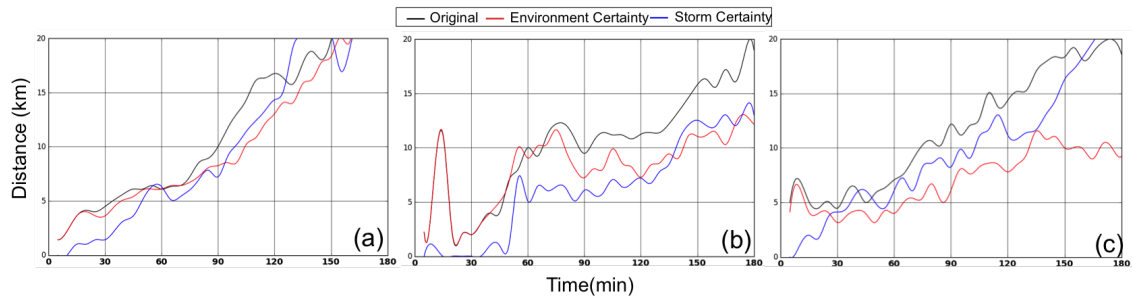
Figure 3.38: Time series of the median distance between every combination of ensemble members storm location for (a) 9 May, (b) 16 May, and (c) 24 May. Storm location was determined by max 2-5 km updraft helicity.

Based on the original experiments, we have some indication of the impacts on mid-level UH amplitude spread by reducing IC uncertainty in the entire domain. However, it is unclear what the benefits to amplitude spread will be by just eliminating perturbations inside or outside the storm. In all three cases, there were substantial reductions in amplitude spread by reducing IC uncertainty inside the storm. Looking at Fig. 3.39, the reduction for 9, 16, and 24 May lasted up to t = 100 min , beyond 180 min, and 110 min, respectively, compared to the original ensemble. As for reducing uncertainty outside the storm, the benefits to amplitude spread were delayed and case-dependent. Looking again at Fig. 3.39, the reduction for 9, 16, and 24 May were only after t = 60 min , 140 min, and 10 min, respectively, compared to the original ensemble. In addition, the benefits of the environment certainty remained until the end of the simulation. Keep in mind, the IC spread reduction in both experiments is (nearly) impossible. This may be especially true in the environment certainty experiments where virtually no perturbations exist in the entire 1-km domain. Therefore, the results are meant to more

illustrative of the effect that perturbations inside and outside of the storm have on the overall storm evolution. Therefore, a literal interpretation of the significant differences in amplitude spread (or overall spread in general) between the original and environment/storm certainty ensembles would be unwise.
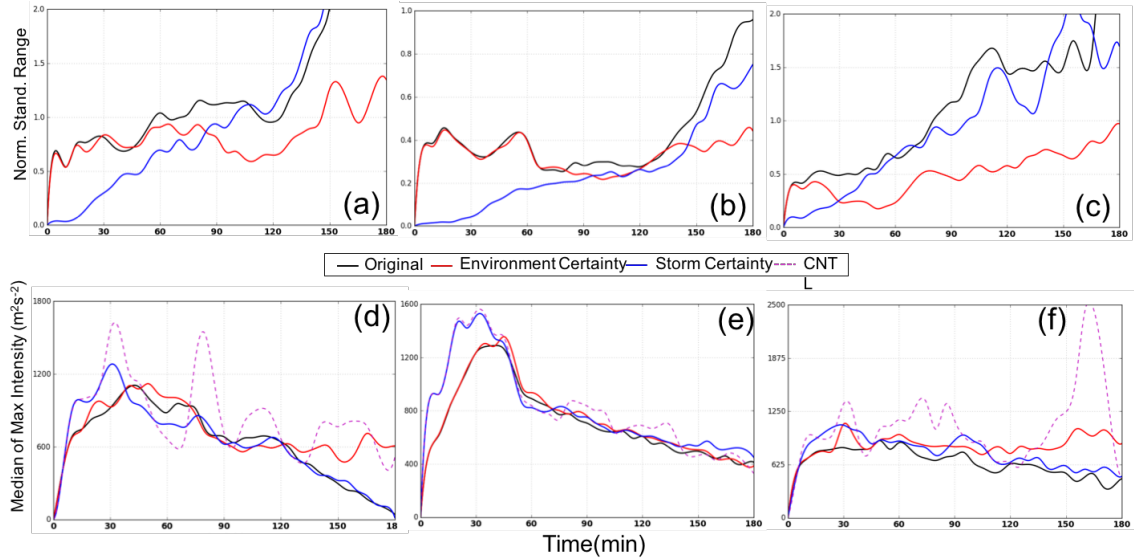
Figure 3.39: Top row: time series of the normalized standard range of max 2-5 km AGL UH by the median of max 2-5 km AGL UH for (a) 9 May, (b) 16 May, and (c) 24 May. Bottom row: time series of the median of max 2-5 km AGL UH along with the control member evolution (dashed magenta) for (a) 9 May, (b) 16 May, and (c) 24 May.

Given such substantial differences in the benefits of improving IC uncertainty inside and outside the storm for the 16 May case, more discussion is warranted. At this point, improvements in IC uncertainty inside the storm provides greater advantages than improving IC uncertainty outside the storm. This comes as a surprise given that the supercell in the 16 May case was organized on larger scales. One might assume that storms organized on larger scales would be more impacted by the broader, mesoscale environment than storm organized on smaller scales. However, this may not be true once the storm is already mature. At that point, for storms organized on larger scales, larger scales within the storm itself (mainly the

cold pool) primarily drive the storm evolution. Thus, a storm organized on larger scales could be greatly impacted by perturbations inside the storm as well as the environment. There was already strong agreement amongst the ensemble members in the original 16 May ensemble. This is evident in the probability of exceedance where there were 80-90% of ensemble members agreed on the MLM path throughout much of the 3-h period (see Fig. 3.36). Arguably, this ensemble agreement may be related to the 16 May storm already inheriting predictability from the larger-scale flow. This may also explain why reducing IC uncertainty inside storm has such a substantial impact since the environment was already having such a large impact on forecast spread.

3.2.2 Updraft

Reducing IC uncertainty inside the storm significantly reduced the normalized updraft spread in the 9 May case (similar to the mid-level UH; see Fig. 3.40c). However, analogous to the mid-level UH, there were diminishing returns as rapid spread growth limited the benefits. As for reducing IC uncertainty outside the storm, the normalized updraft spread reduction was fairly substantial (unlike mid-level UH; see Fig. 3.40b). Looking closer at Fig. 3.40b, we can see that the spread growth is not noticeably increasing, meaning there is little to no qualitative spread growth. To speculate, the environment in the 9 May case could be stunting the perturbation growth.

As for the 16 May case, reducing IC spread inside the storm significantly reduced spread along the northern flank (Fig. 3.41c). However, based on Fig. 3.9, we know that the primary updraft strength of the control member was along the southern side. As for environment certainty, Fig. 3.41i shows that there modest reductions in forecast spread and Fig. 3.41h shows that there was little reduction in updraft posi-

Figure 3.40: 9 May column-max updraft. Left Column: (a-c) Normalized forecast spread swaths (filled contours) with probability-matched median contoured (black lines; thick line $= 25\ ms^{-1}$ with a contour interval of $10\ ms^{-1}$). Center column: (d-f) Probability of exceedance $> 30\ ms^{-1}$ (filled contour) within a 3 X 3 neighborhood corresponding to each horizontal grid point with probability-matched median contoured (black lines). Right column: (g - i) Spaghetti plot of column-max updraft at $30\ ms^{-1}$ for each ensemble member.
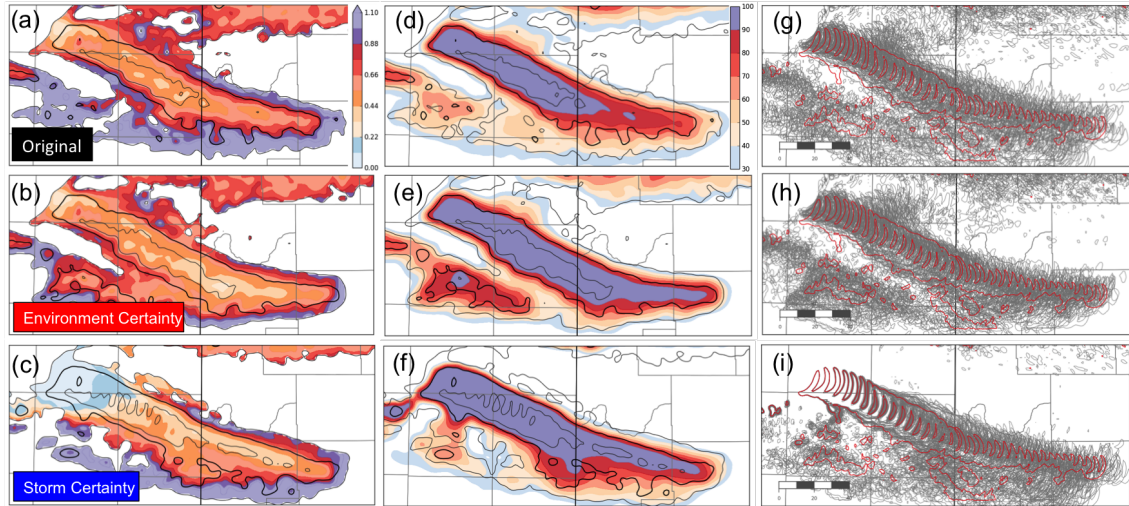
Figure 3.41: Same as in Fig. 3.40, but for 16 May column-max updraft.
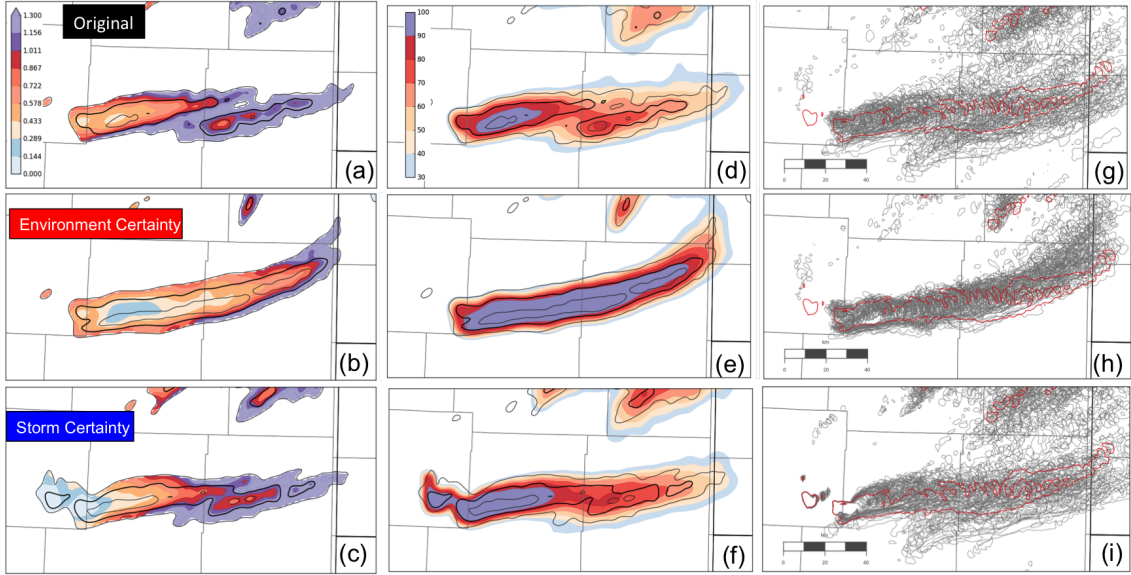
Figure 3.42: Same as in Fig. 3.40, but for 24 May column-max updraft.
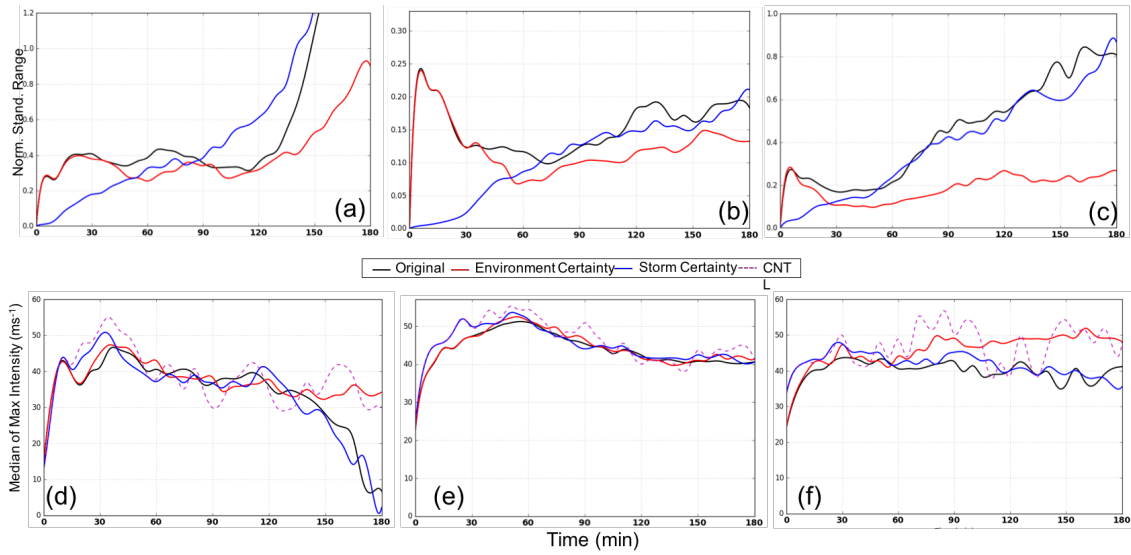
Figure 3.43: Same as in Fig. 3.39, but for column-max updraft.

tion spread. This indicates that improving uncertainty in the environment did not reduce location spread, but did reduce amplitude spread. This was a similar conclusion in the original experiments. Furthermore, after an initial period, there was a similar location spread for both experiments (i.e., IC uncertainty reduced inside and outside the storm). This further suggests that reductions in the overall spread

were primarily related to the amplitude spread reductions. As for the 24 May case, the updraft results were fairly similar to mid-level in that the IC spread reductions in the environment dramatically reduced forecast forecast spread (Fig. 3.42b). However, reducing IC spread inside the storm had even less benefit to forecast spread than the mid-level UH (Fig. 3.42c). In fact, the remaining supercell features for 24 May are related to the updraft or the mid-level UH. Given that, there will be limited discussion of 24 May expect where an important distinction was found. In two of the three cases, the reductions in updraft amplitude spread by reducing IC uncertainty inside the storm were dramatically less than that of the mid-level UH. Looking at Fig. 3.43, the reduction for 9, 16, and 24 May lasted up to t = 90 min , 60 min, and 40 min, respectively, compared to the original ensemble. These results indicate that improving the initial storm estimate has the capability of improving 1-h forecast of peak updraft.

However, this requires no perturbations in the initial storm state. Therefore, more realistic reductions in IC spread would lead to less benefits to the updraft on these lead times. As for reducing uncertainty outside the storm, the benefits to amplitude spread were far greater and less delayed than in the mid-level UH. Looking again at Fig. 3.43, the reduction for 9, 16, and 24 May were only after t = 40 min , 40 min, and 25 min, respectively, compared to the original ensemble. This means that 1 h+ forecasts of peak updraft speed could potentially benefit from reductions in IC spread outside the storm. It is also suggests that the updraft is more sensitive to the environment than the mid-level UH. We know that peak updraft speeds are mostly determined by the environment (e.g., CAPE), so this may not be a surprising result. However, all the storm features are strongly determined by the environment. Thus, it remains unclear as to why the updraft is more sensitive to environment than the mid-level UH.

3.2.3   Low-Level Mesocyclone

In Fig. 3.44c, we can see that improving the IC uncertainty inside the 9 May storm substantially reduced the low-level vorticity spread compared to the original. Fig. 3.44b shows there were also noticeable spread reductions with environment certainty as well (e.g., similar to the updraft, but unlike mid-level UH). In fact, improving IC uncertainty in the environment led to ensemble having a better grasp on the second LLM intensification. However, based on Fig. 3.44h, there is large LLM location spread. As mentioned above, some of those tracks may be related to secondary storms. In Fig. 3.44i we can see that only a few members have a LLM signature near the end of the simulation. Even though improving IC spread inside the storm reduced the number of LLM signatures, the remaining LLM paths have a greater collapse towards the control member. There are also appears to be no influence by secondary storms. This suggests that LLM forecasts would benefit greatly from improved initial storm states. Furthermore, improving storm certainty substantially improved the initial, rapid intensification of the LLM. This has important implications for LLM forecasting as well as WoF. If the goal is to improve tornado lead times to 1 h, then improving internal storm state would be valuable. However, if the aim is to improve lead times to beyond 2 h, then a substantial reduction in IC spread in the entire is needed (as shown in section 3.1.3).

The 16 and 24 May cases also support improving the initial storm estimate as a means to improve 1-h LLM forecasts (Fig. 3.45c and Fig. 3.45c). As for the 16 May case, storm certainty significantly reduced spread throughout the entire simulation with substantial improvements in the probability of exceedance and location spread as well. This is likely related to capturing the three intensification periods in the early portions of the simulation. When reducing the IC spread outside the storm, the benefits to forecast spread were delayed until near the end of the simulation
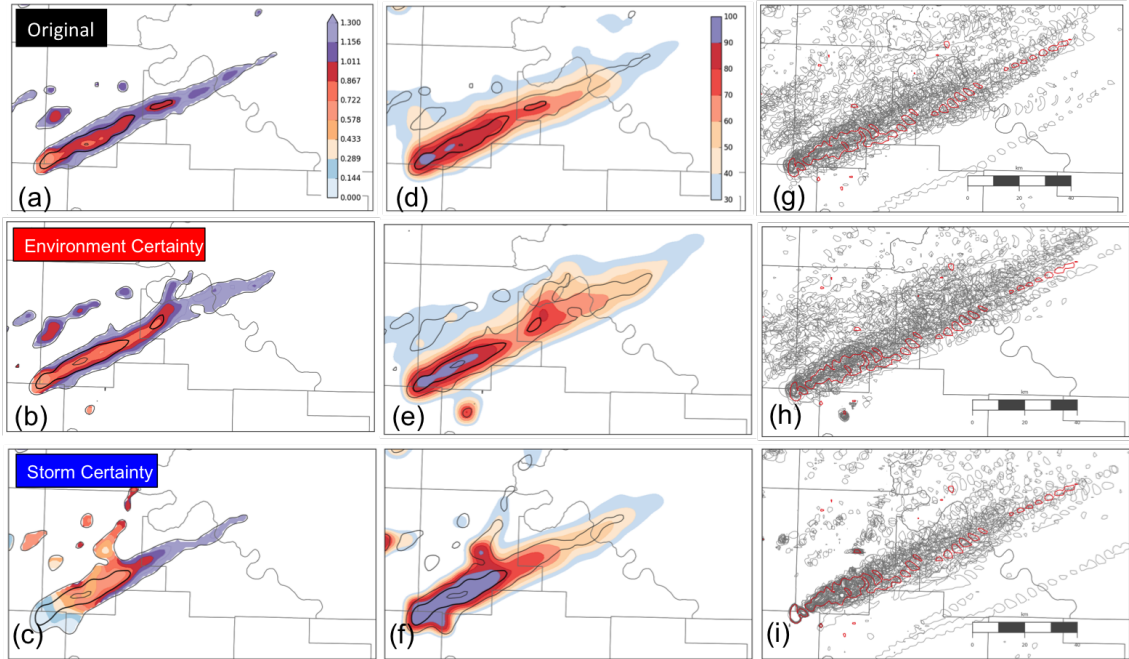
Figure 3.44: 9 May max 0-2 km relative vorticity. Left Column: (a-c) Normalized forecast spread swaths (filled contours) with probability-matched median contoured (black lines; thick line = 0.085 s^{-1} with a contour interval of 0.003 s^{-1}). Center column: (d-f) Probability of exceedance > 0.015 s^{-1} (filled contour) within a 3 X 3 neighborhood corresponding to each horizontal grid point with probability-matched median contoured (black lines). Right column: (g - i) Spaghetti plot of max 0-2 km relative vorticity at 0.015 s^{-1} for each ensemble member.
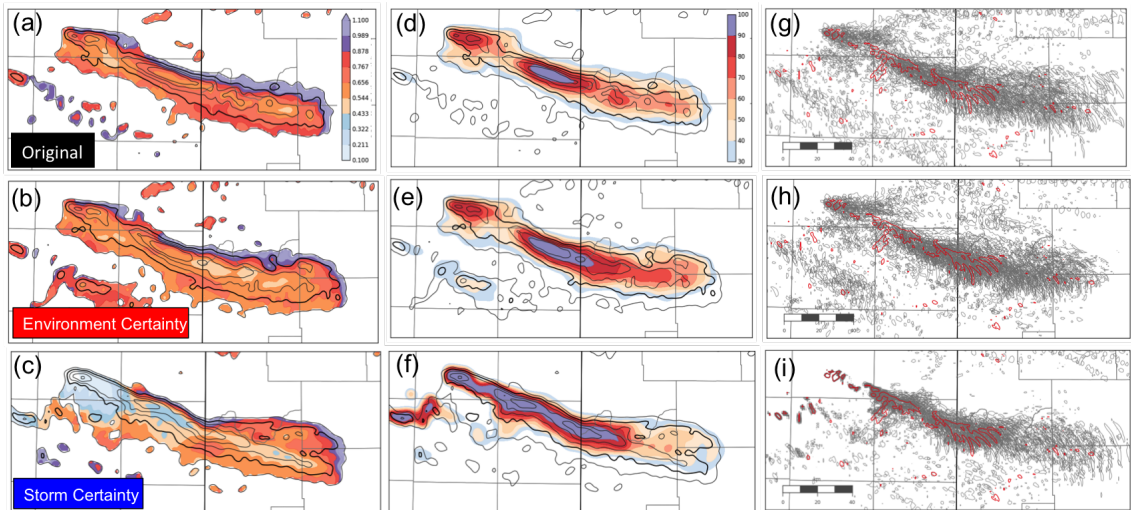
Figure 3.45: Same as in Fig. 3.44, but for 16 May max low-level vorticity. Thick lines = 0.009 s^{-1} with a contour interval = 0.002 s^{-1}.
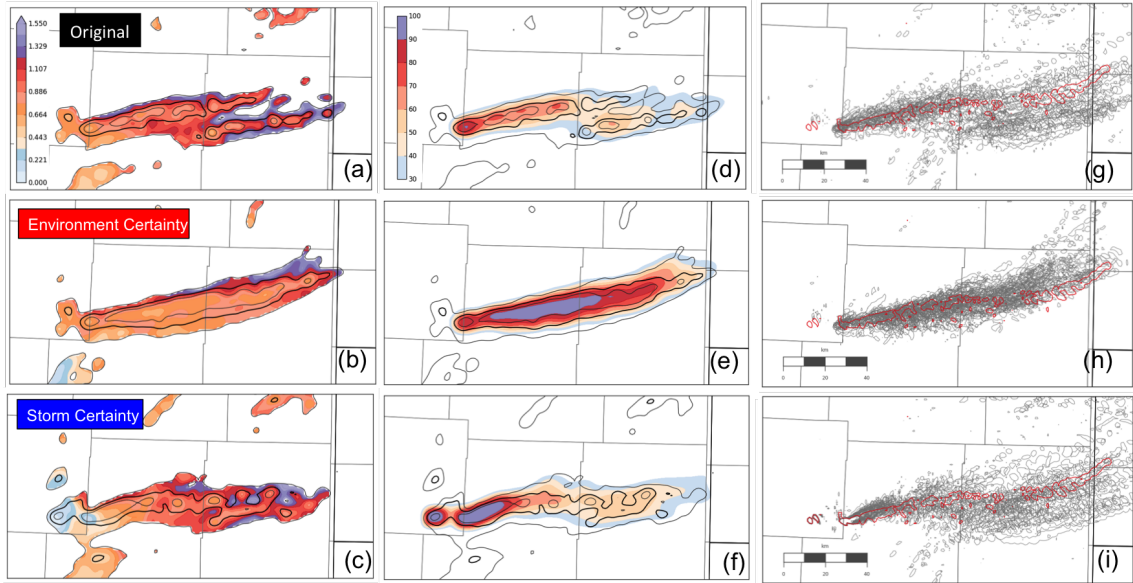
Figure 3.46: Same as in Fig. 3.44, but for 24 May max low-level vorticity.
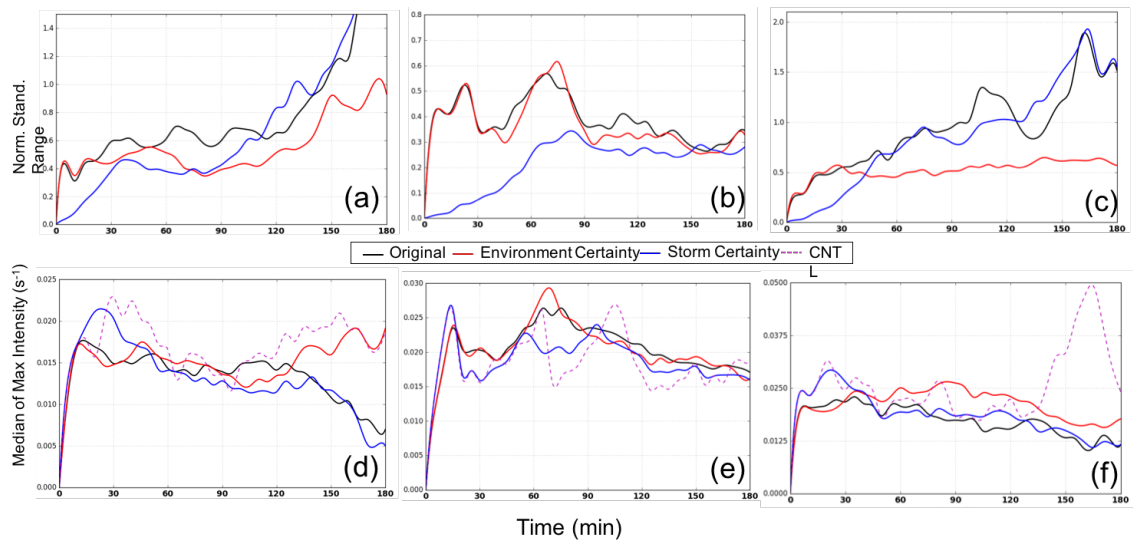
Figure 3.47: Same as in Fig. 3.39, but for max 0-2 km AGL relative vorticity.

(Fig. 3.45b). Regarding the 24 May case, storm certainty benefits were short-lived, but did provide benefits for approximately 1 h (Fig. 3.46c).

Looking at Fig. 3.47 we can see that the qualitative analysis of forecast spread with storm certainty corroborates with the amplitude spread. The reduction for 9, 16, and 24 May lasted up to t = 110 min , 150 min, and 50 min, respectively,

compared to the original ensemble. In the 9 and 24 May cases, we can also see that with environment certainty start at roughly t = 30 min and remain throughout the simulation. Ultimately, 1-h LLM forecasts would benefit significantly for improved initial storm estimates. However, improving the lead times beyond that requires a substantial reduction in IC spread both inside and outside the storm.

3.2.4 Downdraft, Rainfall, and Severe winds

Downdraft

The 9 May case did have a strong downdraft in the early portions of the simulation but was relatively weak throughout the remainder of the simulation. Due to this evolution, storm certainty was clearly more favorable than environment certainty (cf Fig. 3.48b and Fig. 3.48c). As for the 16 May case, Fig. 3.49b shows that reducing uncertainty in the environment had little to no impact on the downdraft forecast spread. However, there were slight improvements in downdraft position near the end of the simulation (see Fig. 3.49h) . Improving storm certainty reduced the forecast spread, but not in the most active regions (e.g., similar to updraft; Fig. 3.49c). As for the 24 May case, Fig. 3.50b shows that the forecast spread in column-max downdraft did not reduce as greatly as the other supercell features (but did have substantial benefits). It is unclear at this point why this was the case. Regarding amplitude spread, the benefits of the improved initial storm state only last t = 45 min in the 16 May case (see Fig. 3.51). After this time, the amplitude spread from both experiments was similar to the original. This means that improving uncertainty inside and outside the storm is necessary to reduce downdraft amplitude and location spread on at longer lead times (i.e., > 1 h).

Figure 3.48: 9 May column-max downdraft. Left Column: (a-c) Normalized forecast spread swaths (filled contours) with probability-matched median contoured (black lines; thick line $= 12 \ ms^{-1}$ with a contour interval of $4 \ ms^{-1}$). Center column: (d-f) Probability of exceedance $> 15 \ ms^{-1}$ (filled contour) within a 3 X 3 neighborhood corresponding to each horizontal grid point with probability-matched median contoured (black lines). Right column: (g - i) Spaghetti plot of column-max downdraft at $15 \ ms^{-1}$ for each ensemble member.

Figure 3.49: Same as in Fig. 3.48, but for 16 May column-max downdraft.
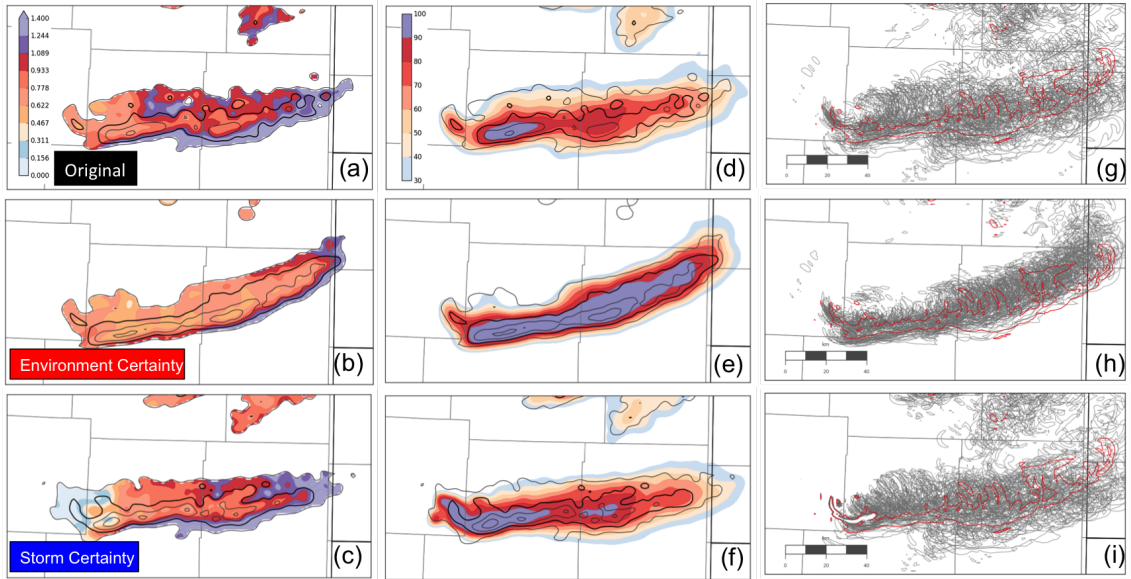
Figure 3.50: Same as in Fig. 3.48, but for 24 May column-max downdraft.
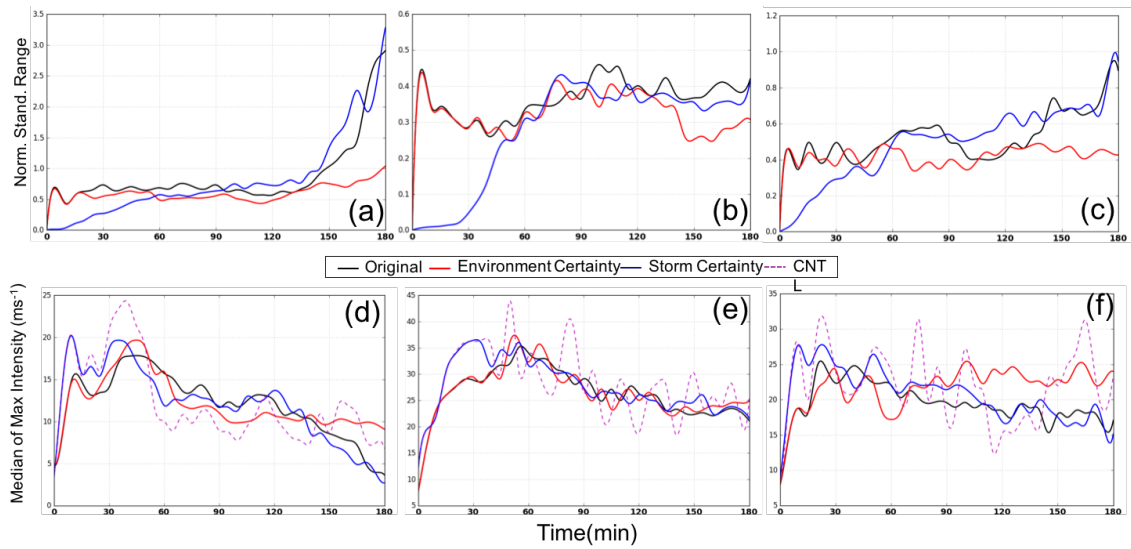
Figure 3.51: Same as in Fig. 3.39, but for column-max dowdraft.

Rainfall

As previously mentioned in section 3.1.5, the 9 May case did not produce a substantial amount of rainfall throughout its lifetime. The period of most intense rainfall was closer to the beginning of the simulation. Thus, improving uncertainty inside the initial storm provided greater benefits than improving the initial environment

(Fig. 3.52c). This is similar to the results for the downdraft. Reducing IC spread in the environment led to substantial spread reductions for the 16 May case. This was especially true in the regions of heaviest rainfall (see Fig. 3.55).
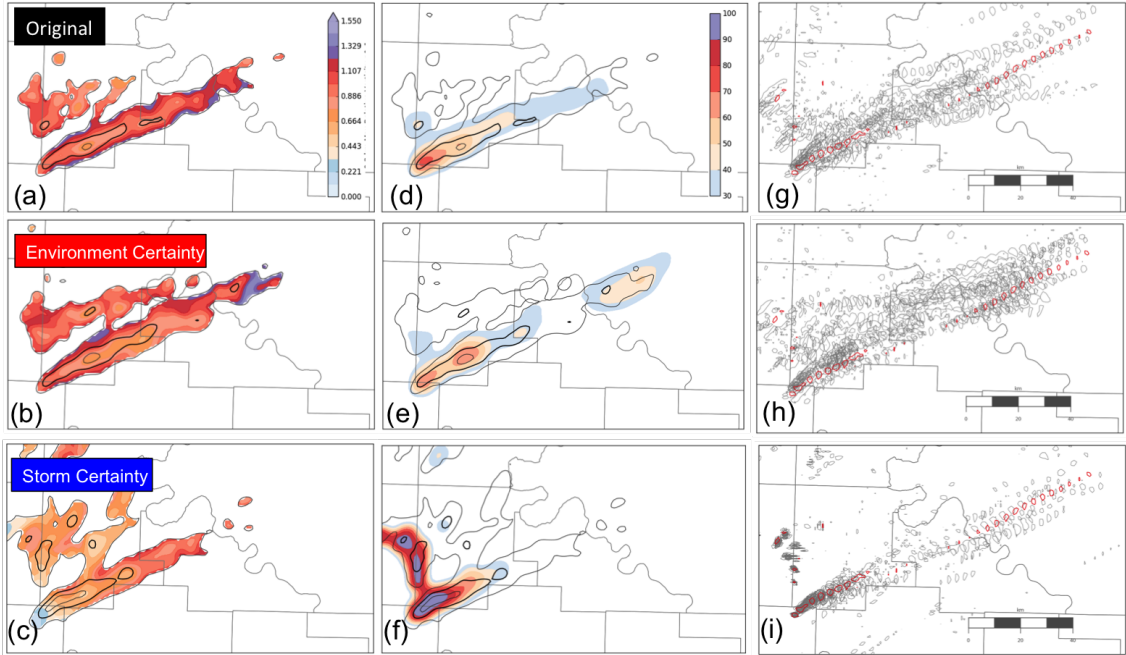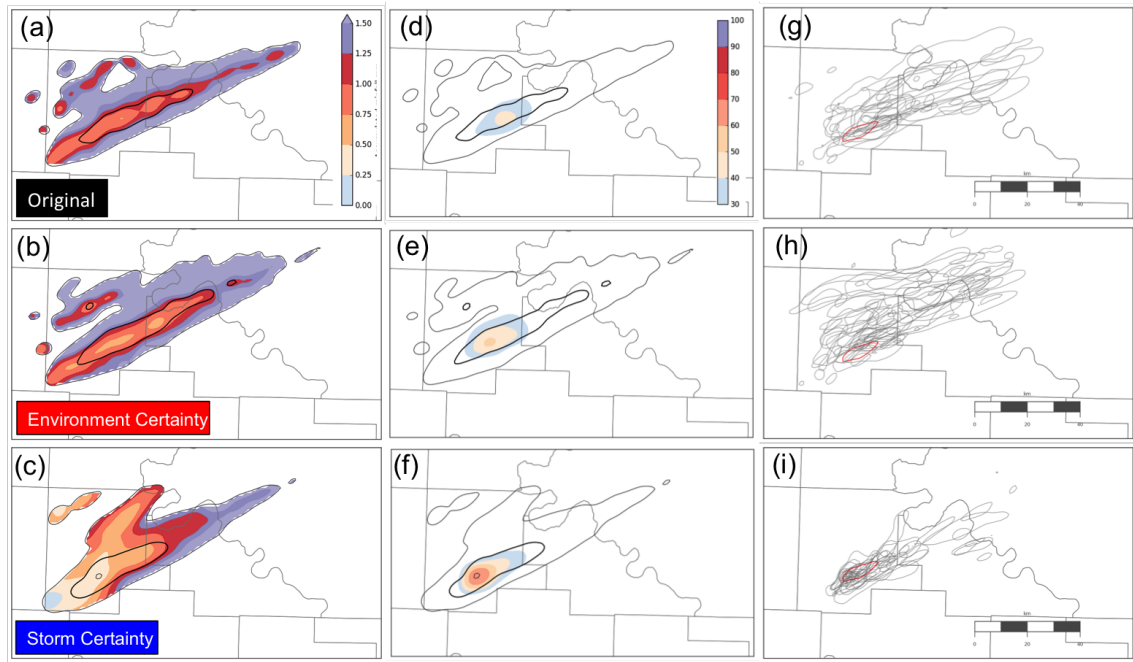

Figure 3.52: 9 May 3-h accumulated rainfall. Left Column: (a-c) Normalized forecast spread swaths (filled contours) with probability-matched median contoured (black lines; thick line = 15 mm with a contour interval of 10 mm). Center column: (d-f) Probability of exceedance > 25 mm (filled contour) within a 3 X 3 neighborhood corresponding to each horizontal grid point with probability-matched median contoured (black lines). Right column: (g - i) Spaghetti plot of 3-h accumulated rainfall at 25 mm for each ensemble member.

The improvement in the spread appears to be related to the substantial improvement in rainfall location. In Fig. 3.53b, we can see there are considerable reductions in location spread even compared to the storm certainty experiment. This is unlike the other features where environment certainty for 16 May only made substantially impacts on the amplitude spread. Improvements in the IC spread inside the storm substantially reduced forecast spread (Fig. 3.53c). However, the greatest reductions were not in the regions of heaviest rainfall (similar to the downdraft and updraft). As for the 24 May case, environment certainty lead to substantial improvement in
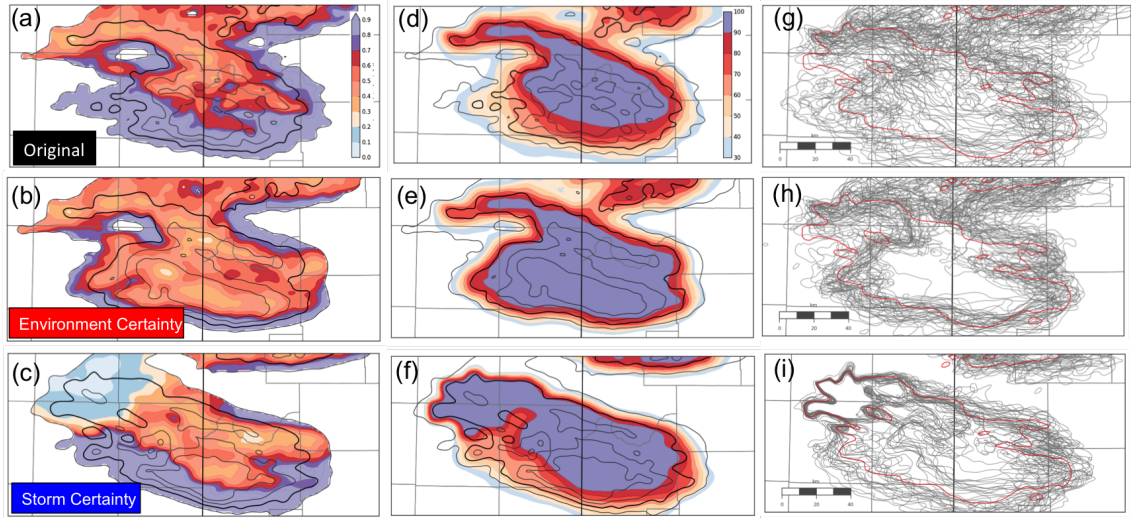
Figure 3.53: Same as in Fig. 3.52, but for 16 May 3-h accumulated rainfall. Thick line = 25 mm , contour interval = 10 mm.
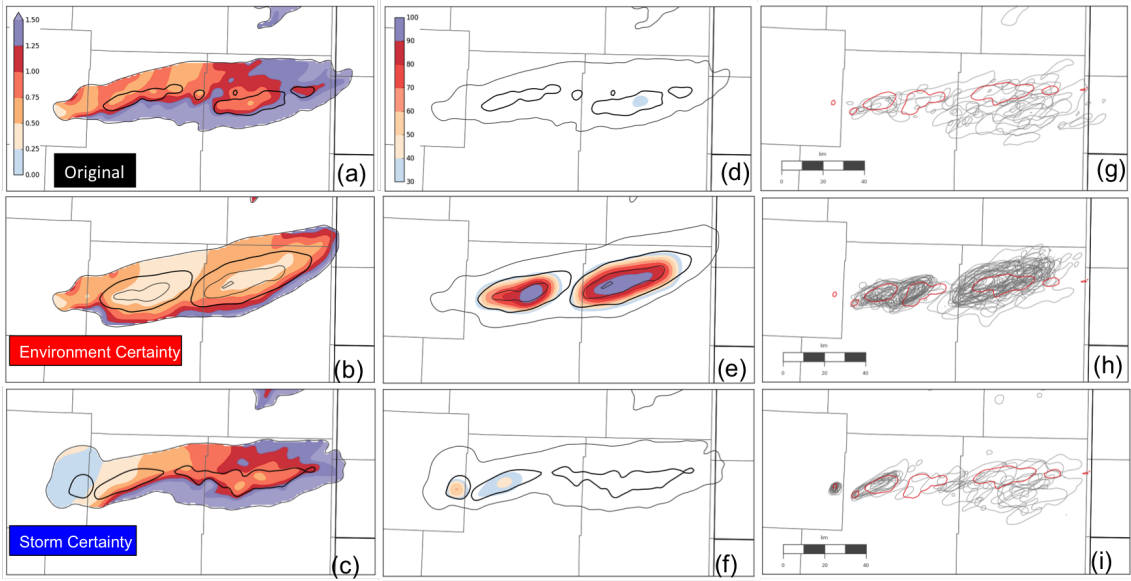
Figure 3.54: Same as in Fig. 3.52, but for 24 May 3-h accumulated rainfall.

probability of exceedance (i.e., > 1 in) and position even more so than the other features (Fig. 3.54e). In fact, the probabilities and position were comparable to the 10 % experiment for the original experiments (cf. Fig. 3.28d). However, given that the forecast spread was still converging, recall that the improvement in the probability is related to the ensemble bias reduction.
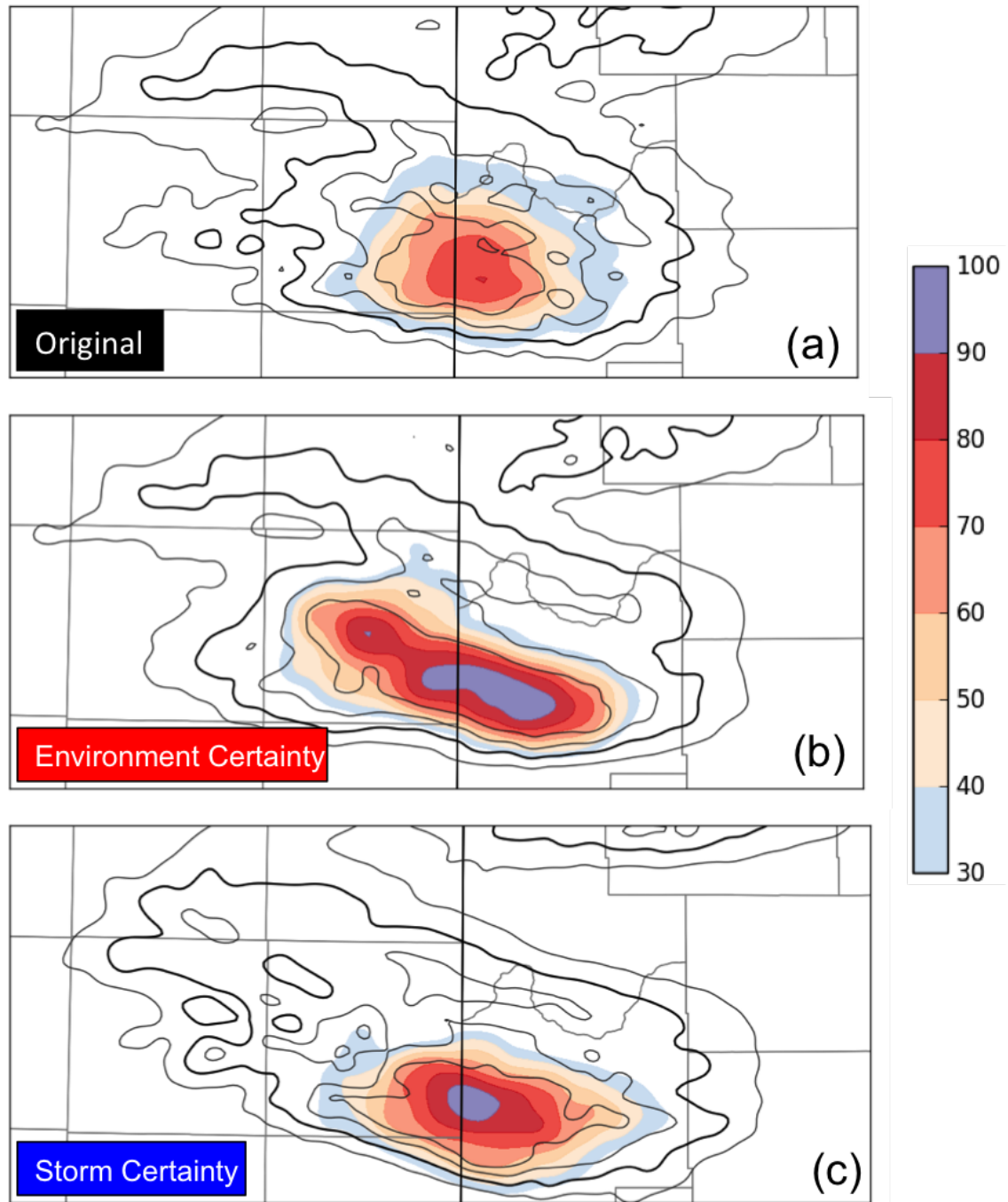
Figure 3.55: Probability of 3-h accumulated rainfall > 2 in

Reflectivity forecasts on verysmall scales and very short lead times are more skillful with the multiscale perturbations as a direct result ofthe smaller-scale IC perturbation energy
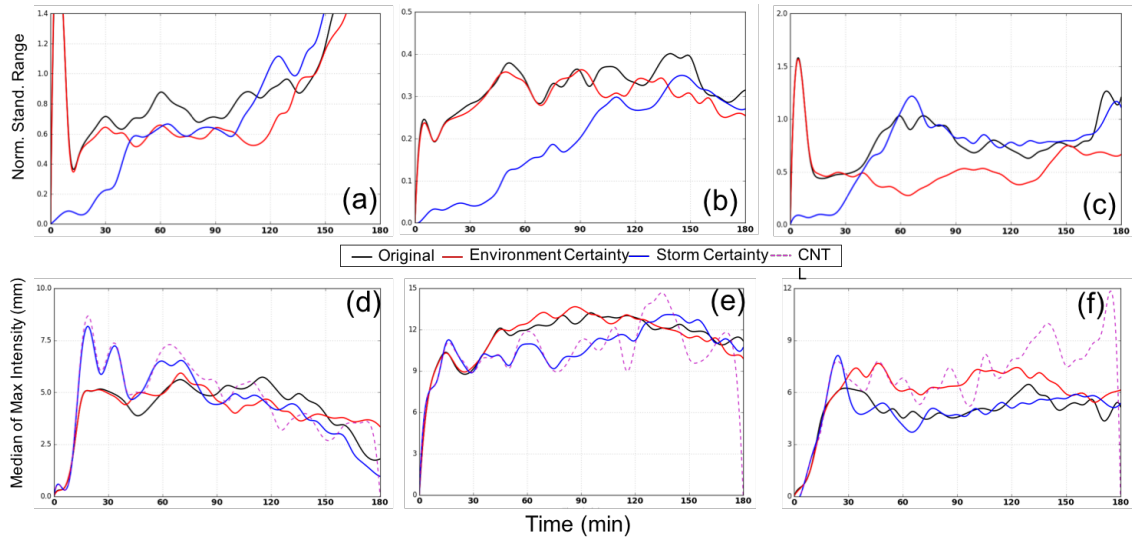
Figure 3.56: Same as in Fig. 3.39, but for 5-min max rainfall.

Severe winds

Finally, as mentioned in section 3.1.6, May9_CNTL did not produce severe winds. However, as previously discussed, reducing false alarm warnings is also crucial and other ensemble members did produce severe winds as seen in Fig. 3.57i. As we can in Fig. 3.57a-c, the strongest winds were towards the beginning of the simulation due to the rapid intensification of the storm. With that in mind, it is clear why the storm certainty is greatly improving forecast spread over environment certainty. In addition, storm certainty is also reducing the number of severe winds in the other ensemble members at later lead times (i.e., > 1 h).

As for the 16 May case, reducing IC uncertainty inside the storm benefited the strong winds (e.g., > 30 ms^{-1}) early in the simulation, but this is expected (Fig. 3.58c). The unexpected result was that both experiments had less 30% probability of exceedance throughout the remainder of the simulation (i.e., after the first hour or so). However, this may not be as surprising given that surface winds were the least predictable feature for 16 May. As for the 24 May case, Fig. 3.59c shows that the benefits of storm certainty appear to last for an even shorter period.
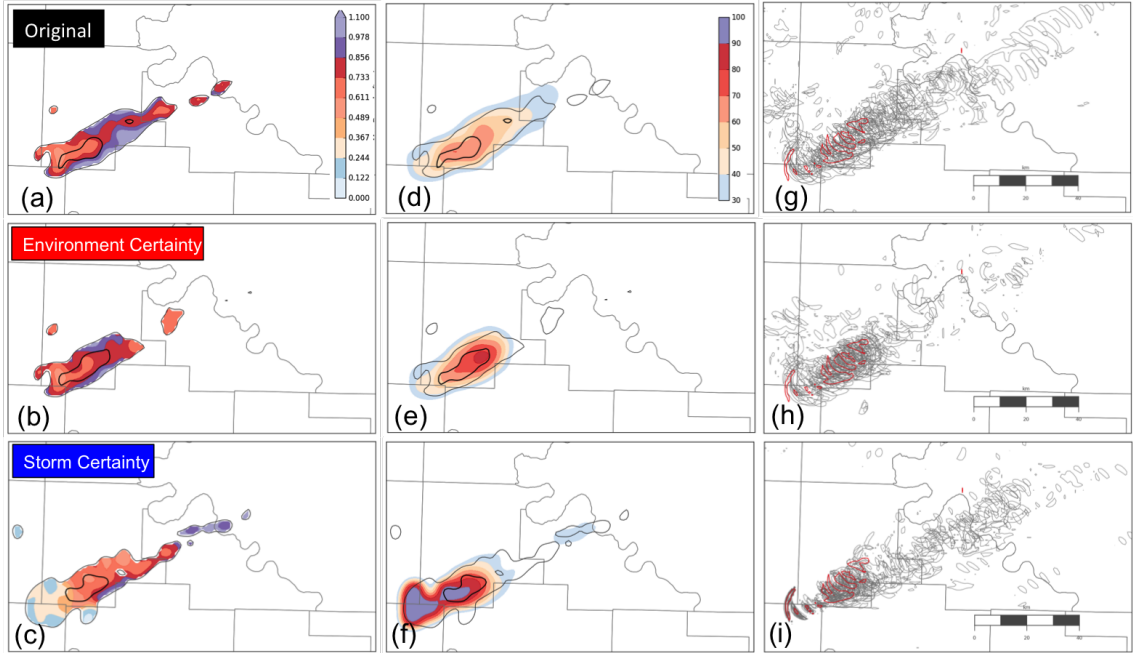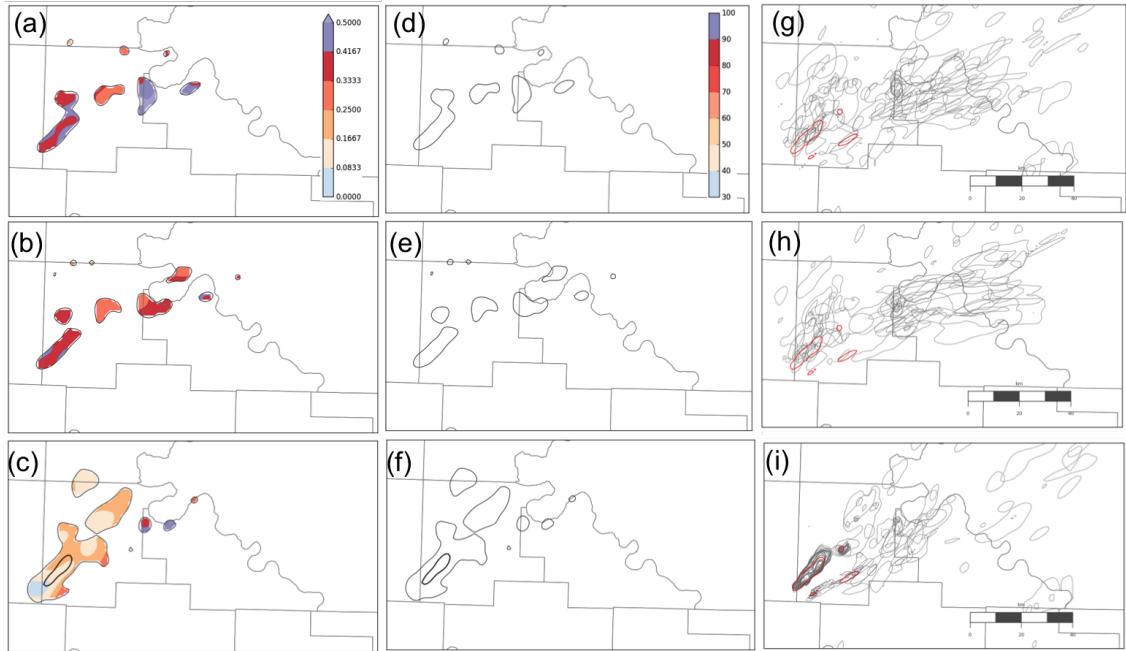
Figure 3.57: 9 May 10-m max wind speed. Left Column: (a-c) Normalized forecast spread swaths (filled contours) with probability-matched median contoured (black lines; thick line = 19 ms^{-1} with a contour interval of 4 ms^{-1}). Center column: (d-f) Probability of exceedance > 26 ms^{-1} (filled contour) within a 3 X 3 neighborhood corresponding to each horizontal grid point with probability-matched median contoured (black lines). Right column: (g - i) Spaghetti plot of 10-m max wind speed at 25 ms^{-1} for each ensemble member.
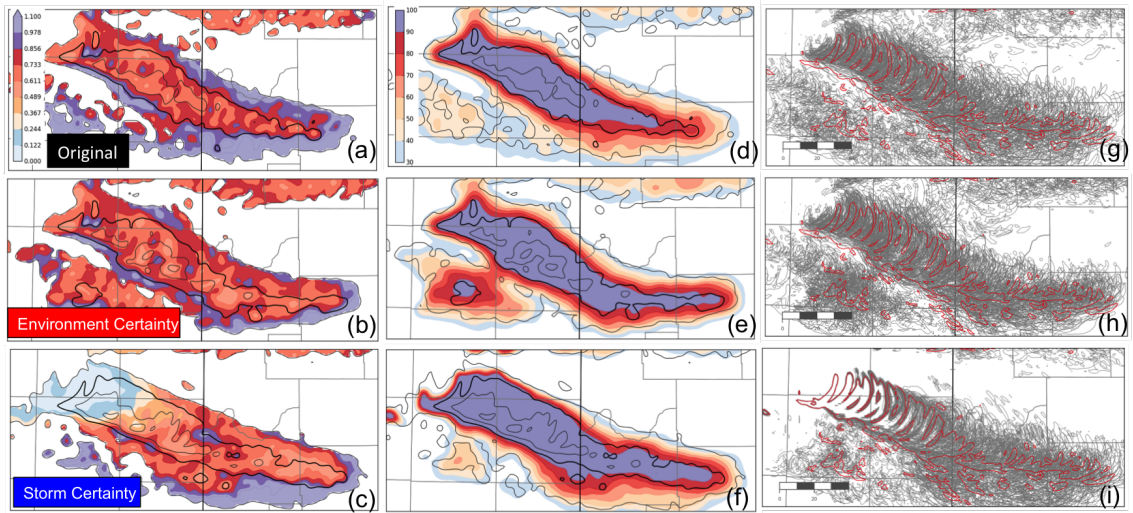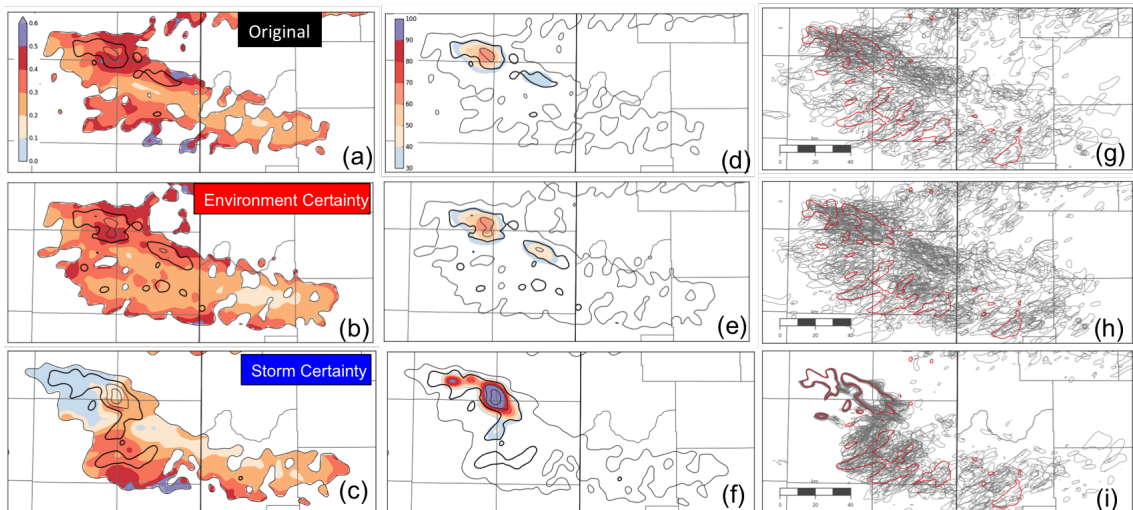

Figure 3.58: Same as in Fig. 3.35, but for 16 May max 10-m wind.
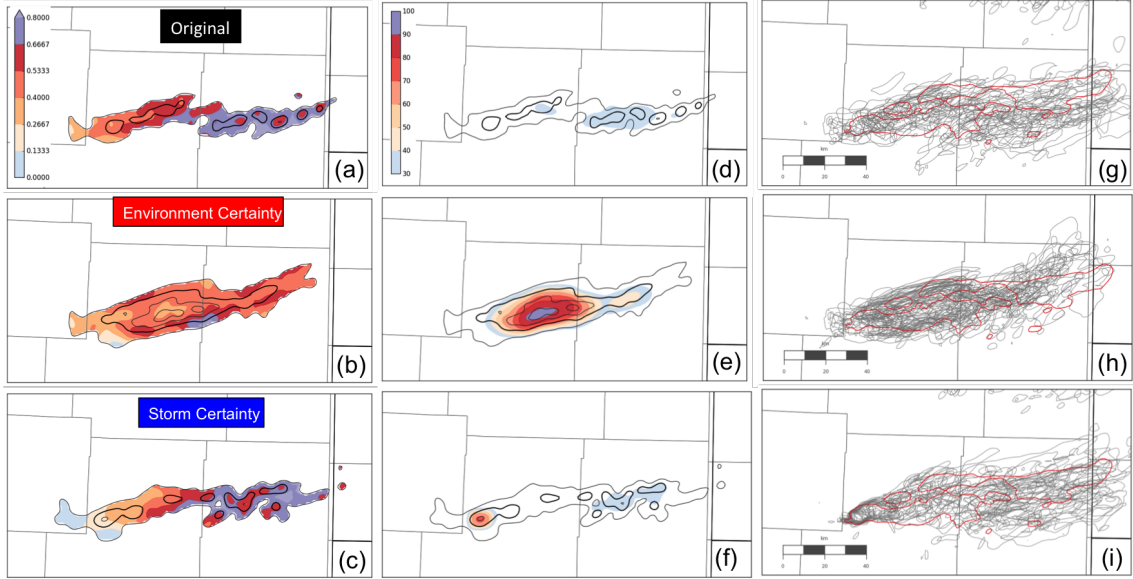
Figure 3.59: Same as in Fig. 3.35, but for 24 May max 10-m wind.
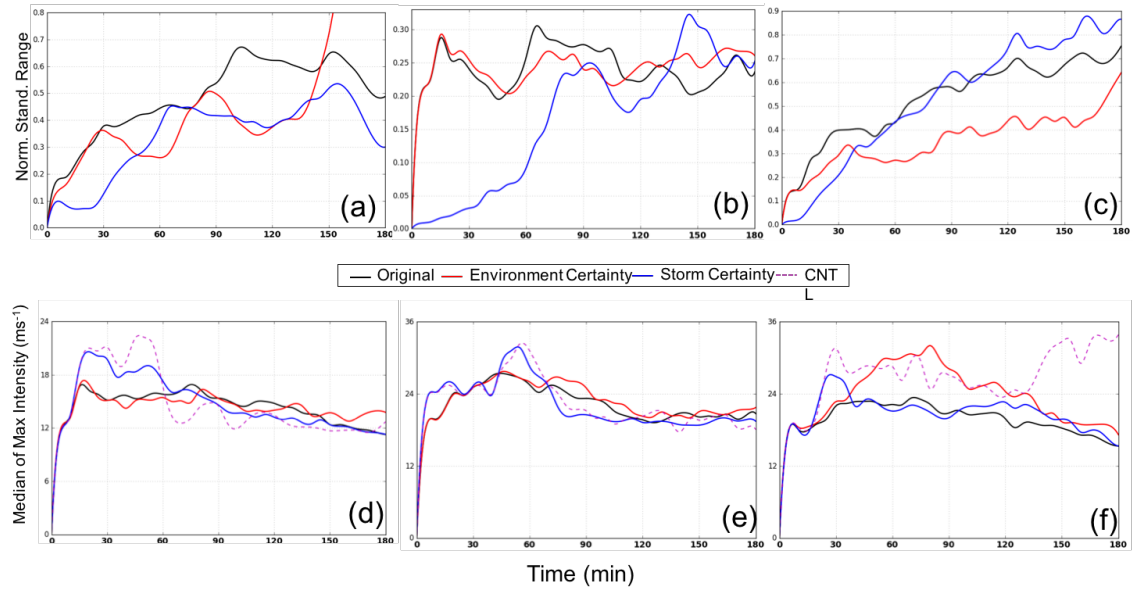
Figure 3.60: Same as in Fig. 3.39, but for max 10-m wind.

Improving the environment certainty greatly reduced the forecast spread (similar to the other features expect downdraft). There were also substantial increases in the probability of exceedance (analogous to the 10% experiment). As for amplitude spread, storm certainty lead to reductions for 16 and 24 May that lasted up to t =

130 min and 40 min, respectively, compared to the original ensemble (Fig. 3.60). As for environment certainty, there was no reduction in amplitude spread for 16 May. The difference between the amplitude spread and the 2D spread swaths for storm certainty in the 16 May case indicates that location predictability was the primary issue for predicting severe winds.

Chapter 4: Discussion

A primary reason for diagnosing atmospheric predictability is the implicit relationship it has with forecast skill. In this context, forecast skill is defined as the ability to provide an accurate forecast of the weather and the time extent to which that is possible. One way of measuring forecast skill in an ensemble framework is evaluating when the ensemble spread converges with the climatological spread. By climatological spread, we are not referring to nature's climatology, but the internal variability of a model, or the model's climatology. When the two spreads are equivalent, it is said that forecast skill has been lost. For example, if the ensemble spread of mid-level UH is approximately equal to the model's variability in mid-level UH then the ensemble is no more skillful than the climatology. Other baselines (e.g., persistence) can be used to define forecast skill, but climatological spread fits nicely into the context of intrinsic predictability in an ensemble framework. In the case of storm-scale prediction, even when forecast spread can no longer be improved beyond a certain point (i.e., as the IC spread is reduced), this does not represent a loss of forecast value. The difference between forecast skill and forecast value is an important distinction to make. For example, you may have two forecasts that are equally skillful at some lead-time, but one may be more valuable based on the forecast information it provides. To give an example, climatology (which could be considered a forecast) and numerical output past the intrinsic predictability limit are considered equally skillful. However, unlike at the synoptic-scale (even mesoscale), the storm-scale climatology is unknown (i.e., no benchmark or baseline to compare to). This means that in lieu of a storm-scale climatology, storm-scale ensemble forecast provide valuable information, before and after the intrinsic predictability limit.

We are far from having a storm-scale climatology that includes each convective mode (e.g., MCSs vs. supercells), the various severe storm environments (e.g., shear, CAPE, moisture), and the range of storm interactions that frequently occur.

In the absence of a suitable storm-scale climatology, human forecaster uncertainty may exceed the climatological uncertainty even at forecast lead times beyond the intrinsic predictability limit. Thus, ensemble forecasts will be valuable beyond the intrinsic predictability limit if they constrain the range of outcomes more than a human forecaster would in the absence of that guidance. As discussed in section 1.3, human forecasters are limited to the "warn-on-detection" paradigm where spatial and intensity extrapolation of storms are used. Therefore, valuable forecast information can be obtained by using storm-scale ensembles even beyond the intrinsic predictability limit. This may become especially true as storm-scale ensembles become more reliable.

Up to this point, it is not clear how these results translate to real-world settings. Although assuming a perfect model simplifies the analysis, it also limits using the results at face value. For example, when making the perfect-model assumption, we neglect the impact of model errors on forecast errors. Making the assumptions does not actually eliminate the model errors and their presence will bias the ensemble forecasts. To account for this model error and unbias the ensemble, physics diversity and/or stochastic physics schemes are often used. However, through the use of these methods, the initial ensemble spread will generally increase. In other words, the practical predictability diagnosed in this study is an upper-bound estimate of the limit that would occur for real ensemble systems. Fortunately, the relative practical predictability (i.e., relative to the IC spread in the three cases) of location vs. amplitude, one feature vs. another, etc. in this study should still be qualitatively valid.

The implied intrinsic predictability limit for the 24 May case was unusually short compared to the other two cases and merits further discussion. In section 3.2.1, it was determined that the intrinsic predictability limit was not greatly impacted (if at all) by the secondary, erroneous storm present in the original ensemble. As discussed

thoroughly in section 1.1, a primary cause of unpredictability at these scales is moist convective processes where the energy from phase changes has a large impact on perturbation growth (Spyksma and Bartello 2007).

In general, the environment in the 24 May case was noticeably drier than the other two cases (e.g., $T_d = 48$ °C). However, upon evaluating the environment in the 24 May control member, there was found a pocket of very moist air (e.g., 11 g/kg) south of the storm inflow at 1.5 km AGL. The region of moist air was associated with some weak convergence near the dryline. Eventually the moist air slowly made its way into the storm inflow, which was associated with the MLM intensifying. Arguably, the intensification was due to the condensation of the extra moisture that led to additional buoyant forcing in the updraft. The intensification of the MLM appears to be connected with the intensification of the downdraft. Recall that in Fig. 3.19, there was a notable strengthening of the column-max downdraft near the point of the spread convergence. This is important because rear-flank downdraft (RFD) intensification has been linked with cyclic mesocyclogenesis (Adlerman et al. 1999). Thus, the additional moisture in the updraft is arguably related to the disruption of the MLM, and it's eventual cycling. Though speculative, if the ensemble did not capture this intensification and cycling, then it likely explains why the intrinsic predictability limit was not beyond 3 h, similar to the other two cases.

Two other causes of unpredictability may be related to environmental vertical wind shear and updraft intensity. The 9 May case had the least vertical wind shear but was the most predictable whereas 24 May with the most vertical wind shear was the least predictable. 16 May had similar shear, but weaker than 24 May, but the organization on larger scales may have inhibited some of the perturbation growth. It is known that dry air entrainment into the updraft increases with vertical wind shear, which can potentially narrow the updraft and perhaps lessen its predictability. However, based on the spaghetti plots in Fig. 3.12i-l and Fig. 3.10i-l, the updrafts in

the 9 May case were more narrow than in the 24 May case. Therefore, entrainment was likely not playing a huge role in the predictability of the 24 May supercell. The second potential cause could be related to the updraft strength. Moreover, 24 May did have the strongest updraft of the two cases, and Cintineo and Stensrud (2013) found that predictability decreased as the updraft intensity increased. Arguably, the updraft intensity may have impacted the predictability of the storm, but this is rather speculative. The exact cause of unpredictability is uncertain, and further evaluation is needed to diagnose the ultimate cause of unpredictability in the 24 May case.

Another peculiarity with the 24 May case was the significant distinction between the amplitude and location predictability. For this case, the amplitude predictability limit for all the supercell features was well within 3 h (e.g., < 2 h). However, the location predictability limit was beyond 3 h (but not far beyond). This suggests that the storm motion was far less sensitive to the IC spread than storm intensity. This is surprising given that supercell propagation is dictated by storm intensity (e.g., rightward propagation off the mean flow). However, the storm motion of the ensemble was generally linear (i.e. moved along a quasi-straight line with no rightward propagation). Recall that the 24 May supercell was associated with a weak cold front and though it is speculative, the storm motion may be directly related to the direction of the boundary motion. In other words, given that each ensemble would be subject to the similar large-scale boundary, the location predictability is inheriting the predictability from the larger-scale flow.

The distinction between amplitude and intensity predictability is well documented in the tropical cyclone (TC) community. For instance, the prediction of TC intensity is lacking behind the prediction of TC position (Emanuel and Zhang 2016). Similar to a supercell, the TC path may be dictated by the larger-scale flow, and thus inheriting the greater predictability from this steering flow.

For two of the three cases, the implied predictability limits for all features were reasonably similar. For the 9 May case, there was no spread convergence (amplitude and location) for any of the supercell features within 3 h. In the 16 May case, the upper limit on the amplitude spread convergence for the most of the features was 150 min with LLM and rainfall beyond 3 h. For the MLM and LLM to have similar predictability limits is surprising given that they form and evolve in different manners (e.g., tilting streamwise vorticity vs. tilting baroclinic-generated vorticity). This correlation of features implies that smaller-scale phenomena may be inheriting predictability from the more predictable, large-scale phenomena. This was supported recently by Potvin et al. (2017) who found that in the absence of smaller scales in the ICs, the large-scale phenomena were able to accurately reproduce them quite rapidly (e.g. 10 - 20 min). However, given that the features of 24 May appear to be un-correlated, it would take more cases to find if correlations amongst feature can be expected and 24 May is simply an outlier.

Although there was a correlation in spread convergence amongst a majority of the features, rainfall had the greatest response to the IC spread reduction. In all three cases, accumulated rainfall had the most substantial reduction in forecast spread as the IC spread was reduced. Rainfall in all three cases had the longest predictability limit. Furthermore, in the 16 May case, max 5-min rainfall had no spread convergence even though the remaining supercell features did. To properly interpret this result, a discussion on chaos and predictability as it is understood in laymen terms is necessary.

If forecast spread for a given feature is insensitive to IC spread reduction, especially at future times (i.e., beyond an intrinsic predictability limit), then it is considered chaotic (i.e., in the rigorous sense). On the contrary, if the forecast spread were fairly sensitive to the IC spread reductions (with no predictability limit), then it can still be considered chaotic (e.g., if it were integrated until an intrinsic predictability

limit were met). Therefore rainfall in the 16 May case could be considered chaotic with no intrinsic predictability limit within 3 h. However, given such a strong sensitivity to the IC spread, one may consider rainfall to be unpredictable.

For an illustration of chaos and predictability, see Fig. 4.1. In the scenario A, the forecast ensemble spread at the reduced IC spread eventually grows to equal that of the original ensemble (with an implied intrinsic predictability limit). For scenario B, the forecast spread with reduced IC spread has not converged with the forecast spread of the original ensemble for the general lead time denoted here. However, even in scenario B, a feature (e.g., rainfall) may be considered unpredictable if the forecast spread is considerably greater than the IC spread. Therefore, rainfall is considered a unpredictable feature given current IC spread, but fundamentally it is a predictable quantity on 3 h lead times when initializing with developing thunderstorms.

Although the secondary storms were not highlighted in this study, in all three cases the forecast spread of those storms was modestly reduced. This suggests that disorganized supercellular convection were slightly less predictable than the more organized supercells in the domain. Although this may be intuitive, it still demonstrates that less organized convection will benefit from a reduction in IC spread. It was also found that perturbations in the primary storm of interest could impact other storms in the domain. In the environment certainty experiments, storms other than the primary storm of interest were known perfectly, but the perturbations in the primary storm of interest eventually impacted the forecast spread of those storms. This demonstrates the conclusion of Hohenegger et al. (2007) that gravity waves generated from convection can propagate and disrupt the predictability of the whole domain. Potvin et al. (2017) also found that introducing errors at discrete points in the domain ultimately impacted the forecast errors throughout the domain. In the end, this result is in its very essence chaos. Unless the ICs are known with complete certainty (e.g., infinite precision), then errors (no matter the amplitude or scale)
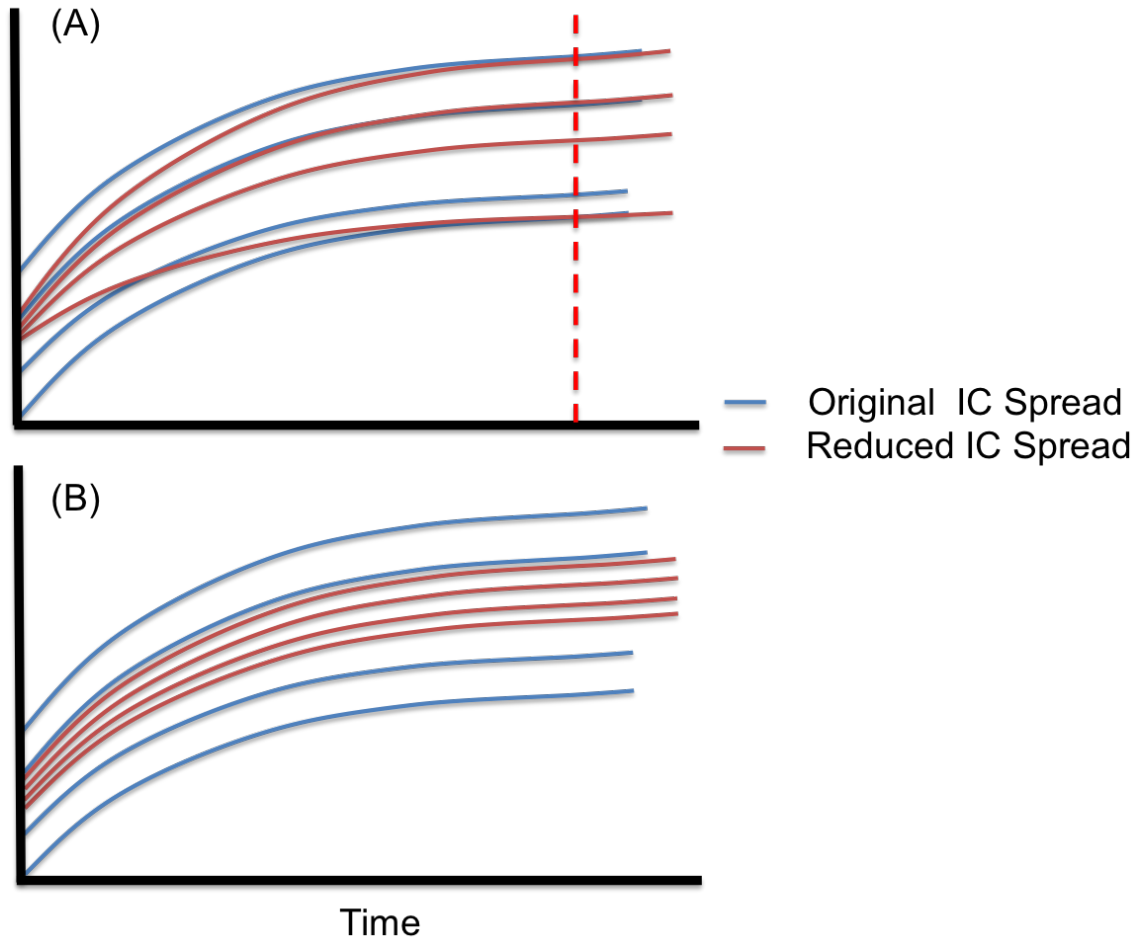
Figure 4.1: Illustration of Predictability and Chaos. Individual lines represent individual ensemble members. Dashed red indicates implied intrinsic predictability limit.

will grow rapidly and impact the entire domain.

Chapter 5: Conclusions and Future Work

5.1   Conclusions

The time scale that storm-scale ensembles are reliable is essential knowledge for the WoF paradigm and therefore their robustness needs to be assessed. Given that supercells are a primary source of life-threatening hazards such as EF2+ tornadoes, > 1 in hail, and torrential rainfall, a thorough understanding of their predictability is warranted. Furthermore, given that ensemble-based DA including radar and satellite observation is possible in real-time, developing a suitable full-physics/real data NWP framework for studying storm-scale predictability is a neccesary first step. To summarize, the methodology developed in this study includes:

- Full-physics /real data rather than a homogeneous, idealized settings.

- Use of DA to generate the ensemble; including radar and satellite observation in the DA, which is critical for intra-storm regions

- Non-traditional statistics, metrics, and analysis to evaluate model performance

In such a framework, we were able to assess the predictability, both intrinsic and practical, of most supercell features, including the low-level mesocyclone and downdraft (which have been neglected in past supercell predictability studies) over multiple cases. To summarize, the primary goals were to:

- Develop a suitable full-physics/real data NWP framework for studying storm-scale predictability (also includes non-traditional metrics).

- Diagnose the forecast spread growth of supercell features in an ensemble-framework (where the ensemble was generated from the EnKF) using real data/including physical parameterizations (as opposed to horizontally homogeneous data) on warn-on-forecast time and space scales (e.g., O(10 min, 1 km))

104

- Explore the importance of uncertainty within vs. outside of the storm.

- Examine a larger number of cases and supercell features than previous studies

Instead of introducing random perturbations at prescribed wavelengths or randomly in the domain, this study opted for reducing flow-dependent perturbations within a pre-existing ensemble towards a deterministic, control member. The largest reduction (i.e., 90%) in IC spread was arguably sufficient to evaluate intrinsic predictability, and the smallest reduction (i.e., 50%) was arguably an indication of practical predictability. Even the 0% reductions (i.e., the original ensemble) provides an indication of practical predictability (which is a very general concept).

In general, there were substantial benefits when reducing the IC spread for all supercell features in all three cases except for severe winds where the results were inconclusive. In particular, the MLM and rainfall had the greatest reduction in forecast spread in all three cases when IC spread was reduced. Recall from Ch. 4 that if the forecast spread of a feature is extremely sensitive to the IC spread then it could be considered unpredictable. This was particularly true for rainfall, which could be considered unpredictable given current IC spread. Nevertheless, given that rainfall forecast spread can improve with IC spread reductions, it is a fundamentally predictable quantity. As for the MLM, even though it would benefit greatly from IC spread reductions, it was shown to be fairly predictable even with 0% reductions.

As for the other features, their predictability was case-dependent. For 9 May, the predictability of every feature was beyond 3 h. As for 16 May, all features but the downdraft and severe wind arguably had a predictability limit greater than or equal to t = 150 min. Finally, for 24 May the MLM, LLM, and rainfall were more predictable than the updraft, downdraft, and severe wind with each having a quasi-distinct implied intrinsic predictability limits. Unfortunately, the caveat of these results is that it would take a tremendous reduction in model error and observation

uncertainty to reduce current IC spread by 75 - 90%. Nevertheless, there were large reductions in forecast spread for all supercell features in the three cases even with a 50% IC spread reduction (which still may be considered tremendous). For example, for the mid-level UH and updraft, a 50% IC spread reductions delayed 1-h amplitude forecast spread in the original ensemble by as much as 100 min. This is an important result for the supercell predictability that one could expect over the next decade or more.

A primary reason for diagnosing atmospheric predictability is the implicit relationship it has with forecast skill. Forecast skill can be defined as the ability to provide an accurate forecast of the weather and the time extent to which that is possible. Forecast skill in an ensemble framework is considered lost when the forecast spread is equivalent to the climatological spread. However, if the climatology is unknown, the ensemble still provides valuable information. Therefore, given that storm-scale climatology is unknown and human forecasters have limited success at longer lead times, storm-scale ensembles will continue to provide valuable forecast information.

Overall intrinsic predictability is some combination of location and amplitude predictability and therefore the two were separated and evaluated individually. Reducing IC spread greatly reduced location spread and due to the weakly non-linear growth, the benefits of reducing the IC spread lasted longer than expected. In fact, in two of the three cases, there was no spread convergence within the 3 h integration period. As for amplitude spread, the growth was fairly non-linear with spread convergence occurring in a majority of the supercell features in two of the three cases. This suggests that storm intensity was more sensitive than the mechanisms that control storm movement to the IC perturbations. Interestingly, this separation between location and amplitude predictability also occurs in TC forecasts. A possible explanation for cases where supercell location predictability was greater than am-

106

plitude predictability was also given in Ch. 4. In cases of weak storm propagation, the storm motion is primarily dictated by the larger-scale steering flow. Therefore, it may be possible for supercell location predictability to inherit the predictability of the larger-scale flow. Ultimately, since location predictability was greater than or equal amplitude predictability, the amplitude spread generally determined the overall intrinsic predictability.

Applying the mechanisms that lead to accelerated perturbation growth (e.g., moisture, updraft intensity, and vertical wind shear), the cause of 24 May's shorter predictability limit (compared to the other cases) was discussed. 24 May had a strong updraft and large vertical wind shear and was the least predictable, while 9 May on the other hand had the least shear and weakest updraft, and was the most predictable. 16 May had similar shear as 24 May, but it is speculated that the large-scale organization may have stunted perturbation growth. It was speculated that vertical wind shear was not greatly affecting the predictability. Ultimately, it was argued that if the MLM in multiple ensemble members did have cycling induced by extra moisture in the updraft then it was possible to have sufficient spread for convergence. The actual cause is unknown and warrants further investigation.

A novel consideration in this work was assessing the sensitivity of supercell evolution to uncertainty inside vs. outside the storm. Essentially, there were two experiments that evaluated the effect of a single storm in multiple different environments vs. multiple storms in a single environment. Of all the supercell features, mid-level UH in all three cases benefited the most when the perturbations inside the storm were set to zero. For example, the reductions in amplitude spread last for up to 100 min compared to the original ensemble. As for the other features, improving the initial storm uncertainty provided benefits over original ensemble uncertainty from 40 min to beyond 180 min while the benefits over improving initial environment uncertainty lasted from 25 min to beyond 180 min. Based on the results of

reducing uncertainty in the environment in all three cases, it is unclear exactly how environment perturbations impact supercell predictability. For example, in the 24 May case, the reductions in forecast spread with environment certainty for certain features were comparable to 25% and 10% experiments. However, for the other two cases, the benefits of environment certainty occurred after 60-90 min and were much less substantial. Finally, the vast improvements in the LLM forecast spread in all three cases by improving storm certainty suggests that significantly improving tornado forecast lead times does not necessarily require improving environment certainty.

5.2   Future Work

Two important supercell features that were not discussed here are large hail and the cold pool. Hail is another feature like severe winds and rainfall that impacts humans and property directly. However, predicting maximum hail size remains an elusive goal of modern NWP. As for the cold pool, it is vital for storm lifetime and development, and tornadogenesis. The gust front generated by the cold pool can cut-off the updraft leading to accelerated storm demise and tornadogenesis is sensitive to the cold pool temperature deficit.

Although we were unable to examine all the primary supercell features, we did evaluate multiple cases where past studies have only focused on a single case (Cintineo and Stensrud 2013; Zhang et al. 2015, 2016; Miglietta et al. 2016). However, as mentioned above, the predictability in each case was different and therefore more cases are needed to ensure the robustness of these conclusions. Another important consideration is that the observed supercells in all the cases were tornadic. It would be important to evaluate how predictability changes in cases where the observed supercell was non-tornadic.

Additional future experiments that are worth exploring is the use of a data as-

similation OSSE framework to assess the impacts of better observing particular variables/layers (e.g., mid-level moisture or PBL temperature). In a data assimilation framework, improving observations of certain variables in certain regions can lead to uncertainty reductions in other portions of the atmospheric state, which is not captured in the current methods. Similar to the line of reasoning for reducing perturbation inside and outside of the storm, exploring the perturbation reduction for a given variable within a particular layer may dramatically improve forecast spread. The use of a full-physics/real-data framework allows for the results/conclusions to be more applicable in an operational settings than in an idealized setting. However, it may be appropriate to perform idealized simulations of supercell convection where the causes of unpredictability can be isolated and rigorously evaluated. An additional motivation for performing idealized experiment is the muddling effects of the secondary storms in the domain, which affected all three cases.

Investigating the sensitivity of the results to different initialization times is also justified. For example, this could mean initializing prior to CI (similar to Cintineo and Stensrud 2013) as well as different times post CI (similar to Zhang et al 2016). Accurate prediction of CI remains an elusive goal of modern storm-scale NWP, but determining how supercell predictability changes when mesoscale perturbations grow prior to CI is crucial information for forecasts issued on the day of the event. As for the second framework, assessing the impacts of data assimilation over time on supercell predictability would be useful information for forecasts who may rely on continuous updates for storm-scale ensemble model output.

Finally, the perfect model assumption was made for this work, but the results may be sensitive to the physical parameterizations. The clearest example is turbulence which plays a vital role in the predictability of a fluid. Therefore, evaluating the sensitivity of the results to choice of PBL scheme is warranted.

References

Adlerman, E. J., and K. K. Droegemeier, 2002: The sensitivity of numerically simulated cyclic mesocyclogenesis to variations in model physical and computational parameters. *Monthly Weather Review*, **130 (11)**, 2671–2691, doi:10.1175/1520-0493(2002)130¡2671:TSONSC¿2.0.CO;2.

Adlerman, E. J., K. K. Droegemeier, and R. Davies-Jones, 1999: A numerical simulation of cyclic mesocyclogenesis. *Journal of the Atmospheric Sciences*, **56 (13)**, 2045–2069.

Aksoy, A., D. C. Dowell, and C. Snyder, 2009: A multicase comparative assessment of the ensemble kalman filter for assimilation of radar observations. part i: Storm-scale analyses. *Monthly Weather Review*, **137 (6)**, 1805–1824, doi: 10.1175/2008MWR2691.1.

Aksoy, A., D. C. Dowell, and C. Snyder, 2010: A multicase comparative assessment of the ensemble kalman filter for assimilation of radar observations. part ii: Short-range ensemble forecasts. *Mon. Wea. Rev.*, **138 (4)**, 1273–1292, doi: 10.1175/2009mwr3086.1.

Anderson, J., T. Hoar, K. Raeder, H. Liu, N. Collins, R. Torn, and A. Avellano, 2009: The data assimilation research testbed: A community facility. *Bulletin of the American Meteorological Society*, **90 (9)**, 1283–1296, doi: 10.1175/2009BAMS2618.1.

Anderson, J. L., 2001: An ensemble adjustment kalman filter for data assimilation. *Monthly Weather Review*, **129 (12)**, 2884–2903, doi:10.1175/1520-0493(2001)129¡2884:AEAKFF¿2.0.CO;2.

Anthes, R. A., Y.-H. Kuo, D. P. Baumhefner, R. M. Errico, and T. W. Bettge, 1985: *Advances in Geophysics*, Vol. 28. 159–202 pp., doi:10.1016/s0065-2687(08)60188-0.

Basdevant, C., B. Legras, R. Sadourny, and M. Bland, 1981: A study of barotropic model flows: Intermittency, waves and predictability. *J. Atmos. Sci.*, **38 (11)**, 2305–2326, doi:10.1175/1520-0469(1981)038¡2305:asobmf¿2.0.co;2.

Boer, G. J., 1994: Predictability regimes in atmospheric flow. *Monthly Weather Review*, **122 (10)**, 2285–2295, doi:10.1175/1520-0493(1994)122¡2285:priaf¿2.0.co;2.

Brooks, H. E., C. A. D. III, and L. J. Wicker, 1993: Stormtipe: A forecasting experiment using a three-dimensional cloud model. *Wea. Forecasting*, **8 (3)**, 352–362, doi:10.1175/1520-0434(1993)008¡0352:safeua¿2.0.co;2.

Brotzge, J., and W. Donner, 2013: The tornado warning process: Areview of current research, challenges, and opportunities. *BAMS.*, **94**, 1715–1733.

Cintineo, R. M., and D. J. Stensrud, 2013: On the predictability of supercell thunderstorm evolution. *J. Atmos. Sci.*, **70 (7)**, 1993–2011, doi:10.1175/jas-d-12-0166.1.

Clark, A. J., W. A. G. Jr., M. Xue, and F. Kong, 2010: Growth of spread in convection-allowing and convection-parameterizing ensembles. *Wea. Forecasting*, **25 (2)**, 594–612, doi:10.1175/2009waf2222318.1.

Coleman, T. A., J. S. K. R. Knupp, J. B. Elliott, and B. E.Peters, 2011: The history (and future) of tornado warningdissemination in the united states. *BAMS*, **92**, 567–582.

Dawson, D. T., L. J. Wicker, E. R. Mansell, and R. L. Tanamachi, 2012: Impact of the environmental low-level wind profile on ensemble forecasts of the 4 may 2007 greensburg, kansas, tornadic storm and associated mesocyclones. *Mon. Wea. Rev.*, **140 (2)**, 696–716, doi:10.1175/mwr-d-11-00008.1.

Done, J. M., G. C. Craig, S. L. Gray, and P. A. Clark, 2012: Casetocase variability of predictability of deep convection in a mesoscale model. *Quarterly Journal of the Royal Meteorological Society*, **138 (664)**, 638–648.

Dowell, D. C., L. J. Wicker, and C. Snyder, 2011: Ensemble kalman filter assimilation of radar observations of the 8 may 2003 oklahoma city supercell: Influences of reflectivity observations on storm-scale analyses. *Mon. Wea. Rev.*, **139 (1)**, 272–294.

Dowell, D. C., F. Zhang, L. J. Wicker, C. Snyder, and A. N. Crook, 2004: Wind and temperature retrievals in the 17 may 1981 arcadia, oklahoma, supercell: Ensemble kalman filter experiments. **132**, 1982–2005.

Droegemeier, K., 1997: The numerical prediction of thunderstorms: Chal- lenges, potential benefits and results from real-time operational tests. *WMO Bull.*, **46**, 324–336.

Duda, J. D., and W. A. Gallus, 2010: Spring and summer midwestern severe weather reports in supercells compared to other morphologies. *Weather and Forecasting*, **25 (1)**, 190–206, doi:10.1175/2009waf2222338.1, URL `http://dx.doi.org/10.1175/2009waf2222338.1`.

Dudhia, J., 1989: Numerical study of convection observed during the winter monsoon experiment using a mesoscale two-dimensional model. *J. Atmos. Sci.*, **46 (20)**, 3077–3107.

Durran, D. R., and M. Gingrich, 2014: Atmospheric predictability: Why butterflies are not of practical importance. *J. Atmos. Sci.*, **71 (7)**, 2476–2488, doi: 10.1175/jas-d-14-0007.1.

Durran, D. R., and J. A. Weyn, 2016: Thunderstorms do not get butterflies. doi: 10.1175/BAMS-D-15-00070.2.

Ebert, E. E., 2001: Ability of a poor man's ensemble to predict the probability and distribution of precipitation. *Monthly Weather Review*, **129 (10)**, 2461–2480, doi: 10.1175/1520-0493(2001)129¡2461:AOAPMS¿2.0.CO;2.

Ehrendorfer, M., R. M. Errico, and K. D. Raeder, 1999: Singular-vector perturbation growth in a primitive equation model with moist physics. *J. Atmos. Sci.*, **56 (11)**, 1627–1648, doi:10.1175/1520-0469(1999)056¡1627:svpgia¿2.0.co;2.

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 (4)**, 873–884, doi:10.1175/1520-0434(2002)017¡0873:ecsmfo¿2.0.co;2.

Emanuel, K., and F. Zhang, 2016: On the predictability and error sources of tropical cyclone intensity forecasts. *Journal of the Atmospheric Sciences*, **73 (9)**, 3739–3747, doi:10.1175/JAS-D-16-0100.1.

Gao, J., and Coauthors, 2013: A real-time weather-adaptive 3dvar analysis system for severe weather detections and warnings. *Weather and Forecasting*, **28 (3)**, 727–745.

Hitchens, N. M., and H. E. Brooks, 2013: Preliminary investigation of the contribution of supercell thunderstorms to the climatology of heavy and extreme precipitation in the united states. *Atmospheric Research*, **123**, 206 – 210.

Hohenegger, C., D. Lthi, and C. Schr, 2006: Predictability mysteries in cloud-resolving models. *Mon. Wea. Rev.*, **134 (8)**, 2095–2107, doi:10.1175/mwr3176.1.

Johnson, A., and X. Wang, 2016: A study of multi-scale initial condition perturbation methods for convection-permitting ensemble forecasts. *Mon. Wea. Rev.*, **144 (7)**, 2579–2604, doi:10.1175/mwr-d-16-0056.1.

Johnson, A., X. Wang, J. R. Carley, L. J. Wicker, and C. Karstens, 2015: A comparison of multiscale gsi-based enkf and 3dvar data assimilation using radar and conventional observations for midlatitude convective-scale precipitation forecasts. *Mon. Wea. Rev.*, **143 (8)**, 3087–3108, doi:10.1175/mwr-d-14-00345.1.

Jones, T. A., K. Knopfmeier, D. Wheatley, G. Creager, P. Minnis, and R. Palikonda, 2016: Storm-scale data assimilation and ensemble forecasting with the nssl experimental warn-on-forecast system. part ii: Combined radar and satellite data experiments. *Wea. Forecasting*, **31 (1)**, 297–327, doi:10.1175/waf-d-15-0107.1.

Kalman, R., 1960: A new approach to linear filtering and prediction problems. *ASME. J. Basic Eng.*, **82**, 35–45, doi:10.1111/j.2153-3490.1957.tb01885.x.

Leith, C. E., 1971: Atmospheric predictability and two-dimensional turbulence. *J. Atmos. Sci.*, **28 (2)**, 145–161, doi:10.1175/1520-0469(1971)028¡0145:apatdt¿2.0.co;2.

Leith, C. E., and R. H. Kraichnan, 1972: Predictability of turbulent flows. *J. Atmos. Sci.*, **29 (6)**, 1041–1058, doi:10.1175/1520-0469(1972)029¡1041:potf¿2.0.co;2.

Lilly, D. K., 1986: The structure, energetics and propagation of rotating convective storms. part ii: Helicity and storm stabilization. *J. Atmos. Sci.*, **43 (2)**, 126–140, doi:10.1175/1520-0469(1986)043¡0126:tseapo¿2.0.co;2.

Lilly, D. K., 1990: Numerical prediction of thunderstormshas its time come? *Quarterly Journal of the Royal Meteorological Society*, **116 (494)**, 779–798, doi:10.1002/qj.49711649402.

Lorenz, E., 1984: Some aspects of atmospheric predictability. URL `https://scholar.google.com/scholar?cluster=8630550962656270158`.

Lorenz, E. N., 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.*, **20**, 130–141, doi:10.1175/1520-0469(1963)020¡0130:dnf¿2.0.co;2.

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

Lorenz, E. N., 1982: Atmospheric predictability experiments with a large numerical model. *Tellus*, **34 (6)**, 505–513, doi:10.1111/j.2153-3490.1982.tb01839.x.

Melhauser, C., and F. Zhang, 2012: Practical and intrinsic predictability of severe and convective weather at the mesoscales. *J. Atmos. Sci.*, **69 (11)**, 3350–3371, doi:10.1175/jas-d-11-0315.1.

Metais, O., and M. Lesieur, 1986: Statistical predictability of decaying turbulence. *J. Atmos. Sci.*, **43 (9)**, 857–870, doi:10.1175/1520-0469(1986)043¡0857:spodt¿2.0.co;2.

Miglietta, M. M., A. Manzato, and R. Rotunno, 2016: Characteristics and predictability of a supercell during hymex sop1. *Quarterly Journal of the Royal Meteorological Society*, **142 (700)**, 2839–2853, doi:10.1002/qj.2872.

Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A.Clough, 1997: Radiative transfer for inhomogeneous atmo-spheres: Rrtm, a validated correlated-k model for thelongwave. *J. Geophys. Res.*, **102**, 16 663–16 682.

Nakanishi, M., and H. Niino, 2006: An improved mellor–yamada level-3 model: Its numerical stability and application to a regional prediction of advection fog. *Boundary-Layer Meteorology*, **119 (2)**, 397–407.

Naylor, J., M. S. Gilmore, R. L. Thompson, R. Edwards, and R. B. Wilhelmson, 2012: Comparison of objective supercell identification techniques using an idealized cloud model. *Monthly Weather Review*, **140 (7)**, 2090–2102.

Ngan, K., P. Bartello, and D. N. Straub, 2009: Predictability of rotating stratified turbulence. *J. Atmos. Sci.*, **66 (5)**, 1384–1400, doi:10.1175/2008jas2799.1.

Park, S. K., 1999: Nonlinearity and predictability of convective rainfall associated with water vapor perturbations in a numerically simulated storm. *Journal of Geophysical Research: Atmospheres (19842012)*, **104 (D24)**, 31 575–31 587, doi:10.1029/1999jd900446.

Polifke, W., and L. Shtilman, 1989: The dynamics of helical decaying turbulence. *Phys. Fluids*, **A (1)**, 2025–2033.

Potvin, C. K., and M. L. Flora, 2015: Sensitivity of idealized supercell simulations to horizontal grid spacing: Implications for warn-on-forecast. *Mon. Wea. Rev.*, **143 (8)**, 2998–3024, doi:10.1175/mwr-d-14-00416.1.

Potvin, C. K., E. M. Murillo, M. L. Flora, and D. M. Wheatley, 2017: Sensitivity of supercell simulations to initial condition resolution. *J. Atmos. Sci.*, doi:10.1175/jas-d-16-0098.1.

Potvin, C. K., and L. J. Wicker, 2013: Assessing ensemble forecasts of low-level supercell rotation within an osse framework. *Wea. Forecasting*, **28 (4)**, 940–960, doi:10.1175/WAF-D-12-00122.1.

Roebber, P., D.M.Schultz, and a. D. B.A.Colle, 2004: Toward improved prediction: High-resolution and ensemble modeling systems in operations. *Wea. Forecasting*, **19**, 936–949.

Samenow, J., 2017: Americans are getting less advance notice for tornadoes, as researchers struggle to understand why. URL https://www.washingtonpost.com/news/capital-weather-gang/wp/2017/04/20/americans-are-getting-less-advance-notice-for-tornadoes-as-researchers-struggle-to-understand-why/?utm$_t$erm$ = .3e4ab648e7dd$.

Smirnova, T. G., S. G. Benjamin, J. M. Brown, B. Schwartz, and D. Kim, 2000: Validation of long-term precipitation and evolved soil moisture and temperature fields in maps preprint. *15th Conf. on Hydrology, Long Beach, CA, Amer. Meteor. Soc.*

Smirnova, T. G., J. M. Brown, and S. G. Benjamin, 1997: Evolution of soil moisture and temperature in the maps/ruc assimilation. preprint. *13th Conf. on Hydrology, Long Beach, CA, Amer. Meteor. Soc.*,, 172175.

Snyder, C., and F. Zhang, 2003: Assimilation of simulated doppler radar observations with an ensemble kalman filter*. *Mon. Wea. Rev.*, **131 (8)**, 1663–1677, doi:10.1175//2555.1.

Spyksma, K., and P. Bartello, 2008: Predictability in wet and dry convective turbulence. *J. Atmos. Sci.*, **65 (1)**, 220–234, doi:10.1175/2007jas2307.1.

Stensrud, D. J., and L. J. Wicker, 2004: On the predictability of mesoscale convective systems. preprints. *AMOS 11th National Conf.*, Brisbane, Australia, AMOS, 62–67.

Stensrud, D. J., and Coauthors, 2009: Convective-scale warn-on-forecast system. *Bulletin of the American Meteorological Society*, **90 (10)**, 1487–1499, doi: 10.1175/2009bams2795.1.

Stensrud, D. J., and Coauthors, 2013: Progress and challenges with warn-on-forecast. doi:10.1016/j.atmosres.2012.04.004.

Sun, Y. Q., and F. Zhang, 2016: Intrinsic versus practical limits of atmospheric predictability and the significance of the butterfly effect. *J. Atmos. Sci.*, **73 (3)**, 1419–1438, doi:10.1175/jas-d-15-0142.1.

Surcel, M., I. Zawadzki, and M. K. Yau, 2016: The case-to-case variability of the predictability of precipitation by a storm-scale ensemble forecasting system. *Mon. Wea. Rev.*, **144**.

Thompson, G., P. R. Field, R. M. Rasmussen, and W. D. Hall, 2008: Explicit forecasts of winter precipitation using an improvedbulk microphysics scheme. part ii: Implementation of a newsnow parameterization. *Mon. Wea. Rev.*, **136**, 5095–5115.

Thompson, P. D., 1957: Uncertainty of initial state as a factor in the predictability of large scale atmospheric flow patterns. *Tellus*, **9**, 275–295, doi:10.1111/j.2153-3490.1957.tb01885.x.

Tong, M., and M. Xue, 2005: Ensemble kalman filter assimilation of doppler radar data with a compressible nonhydrostatic model: Oss experiments. *Mon. Wea. Rev.*, **133 (7)**, 1789–1807, doi:10.1175/mwr2898.1.

Vukicevic, T., and R. M. Errico, 1990: The influence of artificial and physical factors upon predictability estimates using a complex limited-area model. *Monthly Weather Review*, **118 (7)**, 1460–1482, doi:10.1175/1520-0493(1990)118¡1460:TIOAAP¿2.0.CO;2.

Walser, A., D. Lthi, and C. Schr, 2004: Predictability of precipitation in a cloud-resolving model. *Mon. Wea. Rev.*, **132 (2)**, 560–577, doi:10.1175/1520-0493(2004)132¡0560:popiac¿2.0.co;2.

Wandishin, M. S., D. J. Stensrud, S. L. Mullen, and L. J. Wicker, 2008: On the predictability of mesoscale convective systems: Two-dimensional simulations. *Wea. Forecasting*, **23 (5)**, 773–785, doi:10.1175/2008waf2007057.1.

Wandishin, M. S., D. J. Stensrud, S. L. Mullen, and L. J. Wicker, 2010: On the predictability of mesoscale convective systems: Three-dimensional simulations. *Mon. Wea. Rev.*, **138 (3)**, 863–885, doi:10.1175/2009mwr2961.1.

Wang, H., T. Aulign, and H. Morrison, 2012: Impact of microphysics scheme complexity on the propagation of initial perturbations. *Mon. Wea. Rev.*, **140 (7)**, 2287–2296, doi:10.1175/mwr-d-12-00005.1.

Warner, T., D. Keyser, and L. Uccellini, 1984: Some practical insights into the relationship between initial state uncertainty and mesoscale predictability. doi:10.1063/1.34276.

Wheatley, D. M., K. H. Knopfmeier, T. A. Jones, and G. J. Creager, 2015: Storm-scale data assimilation and ensemble forecasting with the nssl experimental warn-on-forecast system. part i: Radar data experiments. *Wea. Forecasting*, **30 (6)**, 1795–1817, doi:10.1175/waf-d-15-0043.1.

Wheatley, D. M., N. Yussouf, and D. J. Stensrud, 2014: Ensemble kalman filter analyses and forecasts of a severe mesoscale convective system using different choices of microphysics schemes. *Mon. Wea. Rev.*, **142 (9)**, 3243–3263, doi:10.1175/mwr-d-13-00260.1.

Wicker, L. J., M. P. Kay, and M. P. Foster, 1997: Stormtipe-95: Results from a convective storm forecast experiment. *Wea. Forecasting*, **12 (3)**, 388–398, doi:10.1175/1520-0434(1997)012¡0388:srfacs¿2.0.co;2.

Wu, D., Z. Meng, and D. Yan, 2013: The predictability of a squall line in south china on 23 april 2007. *Advances in Atmospheric Sciences*, **30 (2)**, 485–502, doi:10.1007/s00376-012-2076-x.

Yussouf, N., D. C. Dowell, L. J. Wicker, K. H. Knopfmeier, and D. M. Wheatley, 2015: Storm-scale data assimilation and ensemble forecasts for the 27 april 2011 severe weather outbreak in alabama. *Monthly Weather Review*, **143 (8)**, 3044–3066, doi:10.1175/MWR-D-14-00268.1.

Yussouf, N., J. Gao, D. J. Stensrud, and G. Ge, 2013a: The impact of mesoscale environmental uncertainty on the prediction of a tornadic supercell storm using ensemble data assimilation approach. *Advances in Meteorology*, **2013**, 1–15, doi:10.1155/2013/731647.

Yussouf, N., J. S. Kain, and A. J. Clark, 2016: Short-term probabilistic forecasts of the 31 may 2013 oklahoma tornado and flash flood event using a continuous-update-cycle storm-scale ensemble system. *Wea. Forecasting*, **31 (3)**, 957–983, doi:10.1175/waf-d-15-0160.1.

Yussouf, N., E. R. Mansell, L. J. Wicker, D. M. Wheatley, and D. J. Stensrud, 2013b: The ensemble kalman filter analyses and forecasts of the 8 may 2003 oklahoma city tornadic supercell storm using single- and double-moment microphysics schemes. *Mon. Wea. Rev.*, doi:10.1175/mwr-d-12-00237.1.

Yussouf, N., and D. J. Stensrud, 2010: Impact of phased-array radar observations over a short assimilation period: Observing system simulation experiments using an ensemble kalman filter. *Monthly Weather Review*, **138 (2)**, 517–538, doi: 10.1175/2009MWR2925.1.

Zhang, D., and R. A. Anthes, 1982: A high-resolution model of theplanetary boundary layersensitivity tests and comparisonswith sesame-79 data. *J. Appl. Meteor.*, **21**, 1594–1609.

Zhang, F., 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 (10)**, 3579–3594, doi: 10.1175/jas4028.1.

Zhang, F., A. M. Odins, and J. W. Nielsen-Gammon, 2006: Mesoscale predictability of an extreme warm-season precipitation event. *Wea. Forecasting*, **21 (2)**, 149–166, doi:10.1175/waf909.1.

Zhang, F., C. Snyder, and R. Rotunno, 2002: Mesoscale predictability of the surprise snowstorm of 2425 january 2000. *Mon. Wea. Rev.*, **130 (6)**, 1617–1632, doi: 10.1175/1520-0493(2002)130¡1617:mpotss¿2.0.co;2.

Zhang, F., C. Snyder, and R. Rotunno, 2003: Effects of moist convection on mesoscale predictability. *J. Atmos. Sci.*, **60 (9)**, 1173–1185, doi:10.1175/1520-0469(2003)060¡1173:eomcom¿2.0.co;2.

Zhang, Y., F. Zhang, D. J. Stensrud, and Z. Meng, 2015: Practical predictability of the 20 may 2013 tornadic thunderstorm event in oklahoma: Sensitivity to synoptic timing and topographical influence. *Mon. Wea. Rev.*, **143 (8)**, 2973–2997, doi: 10.1175/mwr-d-14-00394.1.

Zhang, Y., F. Zhang, D. J. Stensrud, and Z. Meng, 2016: Intrinsic predictability of the 20 may 2013 tornadic thunderstorm event in oklahoma at storm scales. *Mon. Wea. Rev.*, **144 (4)**, 1273–1298, doi:10.1175/mwr-d-15-0105.1.