Mississippi State University
Scholars Junction

Theses and Dissertations | Theses and Dissertations

5-6-2017

Sources of Ensemble Forecast Variation and their Effects on Severe Convective Weather Forecasts

Erin Amanda Thead

Follow this and additional works at: https://scholarsjunction.msstate.edu/td

Recommended Citation

Thead, Erin Amanda, "Sources of Ensemble Forecast Variation and their Effects on Severe Convective Weather Forecasts" (2017). *Theses and Dissertations*. 3998.
https://scholarsjunction.msstate.edu/td/3998

This Dissertation - Open Access is brought to you for free and open access by the Theses and Dissertations at Scholars Junction. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of Scholars Junction. For more information, please contact scholcomm@msstate.libanswers.com.

Sources of ensemble forecast variation and their effects on severe convective weather forecasts

By

Erin Amanda Thead

A Dissertation
Submitted to the Faculty of
Mississippi State University
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
in Earth and Atmospheric Sciences
in the Department of Geosciences

Mississippi State, Mississippi

May 2017

Copyright by

Erin Amanda Thead

2017

Sources of ensemble forecast variation and their effects on severe convective weather forecasts

By

Erin Amanda Thead

Approved:

_____
Andrew Mercer
(Major Professor)

_____
Jamie Dyer
(Committee Member)

_____
Michael Brown
(Committee Member)

_____
Song Zhang
(Committee Member)

_____
Renee M. Clary
(Graduate Coordinator)

_____
Rick Travis
Interim Dean
College of Arts & Sciences

Name: Erin Amanda Thead

Date of Degree: May 6, 2017

Institution: Mississippi State University

Major Field: Earth and Atmospheric Sciences

Major Professor: Dr. Andrew Mercer

Title of Study: Sources of ensemble forecast variation and their effects on severe convective weather forecasts

Pages in Study 170

Candidate for Degree of Doctor of Philosophy

The use of numerical weather prediction (NWP) has brought significant improvements to severe weather outbreak forecasting; however, determination of the primary mode of severe weather (in particular tornadic and nontornadic outbreaks) continues to be a challenge. Uncertainty in model runs contributes to forecasting difficulty; therefore it is beneficial to a forecaster to understand the sources and magnitude of uncertainty in a severe weather forecast. This research examines the impact of data assimilation, microphysics parameterizations, and planetary boundary layer (PBL) physics parameterizations on severe weather forecast accuracy and model variability, both at a mesoscale and synoptic-scale level. NWP model simulations of twenty United States tornadic and twenty nontornadic outbreaks are generated. In the first research phase, each case is modeled with three different modes of data assimilation and a control. In the second phase, each event is modeled with 15 combinations of physics parameterizations: five microphysics and three PBL, all of which were designed to perform well in convective weather situations. A learning machine technique known as a support vector machine (SVM) is used to predict outbreak mode for each run for both the

data assimilated model simulations and the different parameterization simulations. Parameters determined to be significant for outbreak discrimination are extracted from the model simulations and input to the SVM, which issues a diagnosis of outbreak type (tornadic or nontornadic) for each model run. In the third phase, standard synoptic parameters are extracted from the model simulations and a *k*-means cluster analysis is performed on tornadic and nontornadic outbreak data sets to generate synoptically distinct clusters representing atmospheric conditions found in each type of outbreak. Variations among the synoptic features in each cluster are examined across the varied physics parameterization and data assimilation runs. Phase I found that conventional and HIRS-4 radiance assimilation performs best of all examined assimilation variations by lowering false alarm ratios relative to other runs. Phase II found that the selection of PBL physics produces greater spread in the SVM classification ability. Phase III found that data assimilation generates greater model changes in the strength of synoptic-scale features than either microphysics or PBL physics parameterization.

Key words: Numerical weather prediction, severe weather, convective weather, model parameterization, microphysics, planetary boundary layer physics, data assimilation, support vector machines, mesoscale meteorology, synoptic meteorology

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

1.1 Overview of the Project

Severe weather outbreak forecasting has improved in recent years, such that the threat for a major severe weather outbreak can be anticipated multiple days in advance in ideal outbreak setups. However, forecasts for these events often carry significant levels of error more than a day in advance of the anticipated event. Further, the predominant mode of the outbreak is sometimes unclear to forecasters until hours prior to the event. In fact, on 24 August 2016, a localized tornado outbreak that produced an EF-3 tornado occurred in Indiana, despite forecasts not suggesting an outbreak hours prior to its onset (Frame 2016). In this unexpected tornado outbreak, as well as similar events, NWP models did not simulate the environmental conditions that meteorologists associate with tornado outbreaks. The forecasts of numerical weather prediction (NWP) models, while providing an essential aid to forecasters, contain inherent uncertainty.

Uncertainty in a model forecast can come from numerous sources. One source of uncertainty is the non-uniform, spatially discrete network of atmospheric observation sites, as well as the numerous temporal discontinuities of data reports and non-uniformity in observing platforms. In contrast, the atmosphere is a fluid that is spatially and temporally continuous in its changes; therefore, some missed observations are unavoidable with a discrete observation system. Uncertainty can also exist in NWP

model forecasts due to the fact that the mathematical equations programmed into these models are approximations of the physical equations that govern the atmosphere. In gridpoint models, finite differencing approaches for numerically solving these differential equations rely on methods such as Taylor series, which are infinite and must be truncated in real applications (Kalnay 2003). This truncation introduces error into the model. NWP models that do not solve the governing equations in discrete gridpoints avoid finite differencing issues by employing spectral methods (such as in the Global Forecast System – GFS), which utilize Fourier transforms for representing waves, which also introduces error into the model. In addition, many approximations of physical processes in the atmosphere, known as physics parameterizations, exist in computer models, with individual parameterizations either fine-tuned to perform particularly well with specific types of atmospheric phenomena or generalized for acceptable operational forecasts of most types of weather. Physics parameterizations make varying assumptions about the variables in the governing equations, yielding an additional source of error.

A common method of examining NWP forecast uncertainty is the use of ensemble modeling. Ensembles are produced with the purpose of modeling uncertainty in the location, timing, and strength of meteorological features and providing indications of the overall level of confidence in a forecast. These ensembles can be generated from mathematically modeled variations in initial conditions, different configurations of model physics parameterizations, or variations in the types (or sources) of data used as input for the model. The spread of the model output values indicates the overall level of uncertainty in a forecast.

This research aimed to quantify the uncertainty within numerical weather predictions for severe weather outbreak forecasts. The experiments examined model uncertainty in predicted atmospheric variables that are important in severe weather outbreak prediction, in both a synoptic-scale and a mesoscale forecast. The uncertainty was introduced to the model by varying the selection of physics parameterizations and the sources of data in the model input. Varying input data introduced added uncertainty from instrumental error in the model runs performed with additional data, while decreasing uncertainty from spatio-temporal discreteness in the observation network by increasing the density of data across the model domain. Varying physics parameterizations introduced uncertainty from the chosen approximations of physical processes that each parameterization scheme employed. This simulated uncertainty was quantified through examination of the model ensemble range of atmospheric variables important to severe weather, and a deterministic forecast of severe weather outbreak mode was generated to indicate the effects of this model uncertainty in an operational forecasting context. Overall, the objective of this research was to determine the effect of each source of model variation on NWP forecasts of outbreak mode in severe weather outbreaks.

1.2 Literature Review

1.2.1 Review of Ensemble Modeling in Severe Weather Forecasting

1.2.1.1 Data Assimilation

Several methods of data assimilation exist in NWP modeling. Three-dimensional variational data assimilation (3DVAR—Parrish and Derber 1992) is a form of assimilation that performs error analysis on data points in three spatial dimensions but at

a single time in the model run. In 3DVAR analysis, data points to be assimilated to a background field modify the background field in three dimensions, assuming that each data point to be ingested passes quality control checks. The assimilation uses error covariance information from the background field to modify the values in the immediate vicinity of the new data point. The assimilation process assumes that the new data would follow the same pattern as that in the background field, though the new data may be different in magnitude and location. Physical governing equations are therefore part of the 3DVAR procedure.

Four-dimensional variational data assimilation (4DVAR—Gauthier et al. 2007) is, as the name implies, a four-dimensional extension of the 3DVAR procedure, with time as the new dimension. Data points to be assimilated to a background field are compared to the expected temporal progression of the meteorological feature, as extrapolated by the assimilation tool using physical equations of the atmosphere. In 4DVAR analysis, data assimilation can correct for inaccurate speeds of weather feature development or movement in the model background fields in addition to inaccurate magnitudes.

Ensemble Kalman filtering (EnKF—Houtekamer and Mitchell 1998) is an assimilation technique that uses stochastic ensemble forecasts—multiple NWP model simulations of the same case with perturbations in the initial fields—to generate flow-dependent statistical information about the background fields. Two sets of ensembles are generated in EnKF. The error covariance information derived from one set of ensembles is used in assimilation of data to the other set; this technique is employed to avoid excessive internal referencing. Unlike 3DVAR and 4DVAR, EnKF requires the simulation of numerous model runs of a given case; however, because it produces flow-

dependent error fields (in contrast with 3DVAR, which is flow-independent), it has been found to reduce error in the assimilation procedure (Houtekamer and Mitchell 2007).

An older data assimilation approach is optimal interpolation (OI—Kalnay 2003). This approach employs least-squares methods to interpolate assimilated data. As in other forms of NWP data assimilation, data in OI are weighted according to error statistics. Background fields also have associated error fields. The OI equations produce assimilation fields that have minimized the amount of error in the final result.

In recent years, improvement of data assimilation techniques has been an important area of focus among NWP researchers. Yussouf et al. (2013) modeled the May 2003 tornado outbreak sequence with a selection of three microphysics schemes in combination with radar and conventional observation assimilation. The assimilation of radar and use of a double-moment physics scheme aided in the modeling of a prominent supercell thunderstorm that occurred on 8 May. Fierro et al. (2012) examined the assimilation of lightning strike observations in a modeling of the 24 May 2011 tornado outbreak over Oklahoma, finding that this data set improved the location of individual thunderstorms.

The assimilation of surface, rawinsonde, and aircraft observations in a 12-hour, 30-member initial condition and model physics ensemble forecast has been examined by Wheatley et al. (2012) for an analysis set consisting of multiple severe weather outbreaks. Their research compared control ensembles to ensembles with assimilated observations of altimeter setting, temperature, dewpoint, and horizontal winds. They calculated the ensemble-mean significant tornado parameter (STP) and the probability that the STP > 1 for model grid points closest to official tornado reports for 24 severe weather outbreaks.

Their research found that the differences in these calculations between the control ensembles and the data-assimilation ensembles were significant at the 95% level and that the assimilation ensemble produced values closer to observations.

Jones and Stensrud (2012) examined the impact of temperature and mixing ratio profiles computed by the Atmospheric Infrared Sounder (AIRS—Aumann et al. 2003) on forecasts of convective activity. They examined two severe weather events from 2009 and 2010, comparing high-resolution (3 km) ensemble forecasts with assimilated AIRS and conventional observations against the same ensemble forecasts with only assimilated conventional observations. Their results for the two cases indicated that AIRS temperature and mixing ratio observations produced improvement in the ensemble mean dewpoint forecast in levels of the atmosphere most associated with severe convective development.

As this sampling of studies has shown, data assimilation generally improves NWP forecast accuracy relative to control observations. Modern assimilation processes employ physically based approaches when adding the observation data to a background field, a strategy that should—and usually does—reduce error in comparison with simply adding observations to the background field.

1.2.1.2 Physics Parameterization

Physics parameterization has also been examined extensively in the context of severe weather forecasting. Numerous studies have examined the impact of various model physics parameterization schemes on the modeling of convection processes, cloud ice production, and mesoscale weather events. The effect of microphysics on cold pool formation is especially well-documented. Morrison and Milbrandt (2011) compared the

Morrison (Morrison et al. 2009) and Milbrandt-Yau (Milbrandt and Yau 2005) microphysics schemes in their modeling of idealized supercell thunderstorms, finding that the Morrison microphysics produced a stronger cold pool and higher levels of precipitation. Varying intra-cloud parameters such as ice levels, raindrop breakup, and drop size in each scheme produced broad differences in the modeling of the cold pool associated with convective storms. It should be noted that this study did not examine a real-world case, but rather, an idealized modeling of convection.

Cintineo et al. (2014) examined a set of five PBL and four microphysics schemes with regard to the modeling of cloud cover over the contiguous United States. They found that microphysics parameterization choice greatly influenced the modeling of upper-level cloud features. Of the schemes they examined, the Milbrandt-Yau and Morrison microphysics schemes produced much more upper-level cloud cover than the Thompson (Thompson et al. 2006) and WRF double-moment 6-class scheme (Lim and Hong 2010). The variations they found among different PBL schemes were small. Li et al. (2015) studied nested WRF simulations of marine cumulus clouds and cold pool development with regard to variations introduced by the Thompson and Morrison microphysics parameterization. They found that the Thompson scheme modeled less cloud cover and liquid precipitation than the Morrison scheme, resulting in a stronger cold pool with the Morrison scheme. This result reinforces the Morrison and Milbrandt (2011) result, which also found that the Morrison scheme generated strong cold pools.

Case studies have also been performed with varied parameterization schemes. As an example, the 3 May 1999 Oklahoma tornado outbreak—a frequently examined high-impact event—was modeled as a six-member physics parameterization ensemble in an

early work by Stensrud and Weiss (2002), who found that their ensemble scheme modeled the forcing mechanisms for this outbreak well. More recently, McMillen and Steenburgh (2015) examined the effects of microphysics on the modeling of a lake-effect snowstorm in Utah, finding that the Thompson microphysics scheme most comparably modeled snow in the event relative to real observations, and that the WRF double-moment 6-class microphysics scheme produced larger amounts of precipitation in the form of graupel than other tested microphysics.

In a study on a very different topic, Gibbs et al. (2011) examined the effect of three planetary boundary layer (PBL) parameterization schemes on the modeling of a dry convective boundary layer in two cases in Oklahoma in dryline and dry cold front cases. They found that the Mellor-Yamada-Janjić boundary layer scheme (MYJ—Janjić 1994) most closely approximated observed values of wind magnitude in the dryline PBL environment, whereas the Yonsei University (YSU—Hong et al. 2006) and refined Asymmetric Convection Model (ACM2—Pleim 2007) more closely modeled wind magnitudes for the dry cold front. YSU PBL most closely modeled heat flux values in both the dryline and the dry cold front cases, and ACM2 modeled flux values least accurately. The MYJ scheme produced the most accurate modeling of near-surface turbulence in the dryline case, but the least accurate in the dry cold front case. However, in their study, all three PBL schemes produced heat and moisture values that were reasonably similar to each other. The authors emphasized that in a dry PBL case, whether dry cold front or dryline, the choice of PBL scheme may not introduce very much uncertainty into a NWP forecast. Their study did not examine moist PBL cases, which are more commonly associated with convective instability.

Hu et al. (2010) also examined these three planetary boundary layer schemes—Mellor-Yamada-Janjić, Yonsei University, and Asymmetric Convection Model—for performance in modeling summertime boundary layer processes in Texas and eastward to Mississippi, a decidedly moist environment, in contrast with the model domain of the Gibbs et al. study. They found that, although all three examined PBL schemes introduced cold and moist biases in the 0-2 km atmospheric layer compared to an observation data set, the MYJ scheme produced the largest biases. Close examination of the modeled physical variables revealed that vertical mixing in the PBL was stronger in the YSU and ACM2 runs than the MYJ runs.

1.2.1.3 Initial Condition Perturbation

Initial condition perturbation has been extensively studied in the context of severe weather forecasting. Current operational weather models are run in ensemble mode, and—depending on the model—several dozen initial condition perturbation runs are produced. Initial conditions can be perturbed either by a randomized method of generation (Monte Carlo ensembles) or by a method that takes into account the flow pattern (Kalnay 2003). The flow-aware perturbations are preferred in operational forecasting for greater realism, and several schemes for generating variations exist. One approach, known as bred vectors, involves the differencing of a perturbed forecast from a non-perturbed forecast and the assimilation of this error into the next model time step. Another approach, which is used in the European Centre for Medium-Range Weather Forecasts (ECMWF) model, is known as singular vector perturbation (Leutbecher and Palmer 2007). This approach employs eigenvector solutions to a vector-norm equation as the basis for its perturbations. The perturbation method that is included in the Weather

Research and Forecasting (WRF) model is known as stochastic kinetic-energy backscatter (SKEBS, Berner et al. 2011). This method attempts to estimate the dissipation of kinetic energy at higher resolutions than the runtime configuration of the model, using the calculated amount of energy that is not dissipated in the next step in the model. A more detailed review of perturbation ensemble modeling methods is beyond the scope of this research, as this method of ensemble generation is not examined in the study.

However, approaches for quantifying uncertainty in perturbation ensembles still have application to this research, so a brief pair of examples is given. Clark et al. (2004) examined the growth of ensemble spread in a prolonged period of convective weather. A physics parameterization ensemble and a combination physics-perturbation ensemble at two different grid resolutions were compared to each other. The comparison values were the ensemble variance of standard meteorological fields (geopotential height, mean sea level pressure, temperature, dewpoint, and wind magnitude) at various pressure levels. For the combined ensembles and the parameterization-only ensembles, the finer grid resolution produced faster spread growth rates than the coarser grid. Additionally, lower atmospheric variables exhibited a greater proportion of ensemble spread from physics parameterization than from initial condition perturbation. This finding is relevant to this research because it suggests that physics parameterization, even at small grid spacing, introduces significant uncertainty into convective weather NWP forecasts.

Tapiador et al. (2012) modeled a severe hail storm in Spain with physics parameterization ensembles (without perturbations) and initial condition perturbation ensembles (without physics parameterization variations). They varied cloud

microphysics, cumulus physics, and land surface physics in the parameterization ensemble, and they used MYJ PBL physics for all ensemble members. Similar to Clark et al. (2004), they found that the physics parameterization ensemble produced statistically significantly greater ensemble spread than the perturbation ensemble in the modeling of simulated reflectivity and precipitation.

1.2.2 Review of Support Vector Machines

Verification of a numerical weather prediction model in the context of severe weather outbreak forecasting is a contentious issue. An operational forecast of a specific type of severe weather is made subjectively by a human forecaster with NWP forecasts, current observations, local climatology, and persistence as guidance, and the accuracy of such a forecast is often highly subjective as well. Therefore, other methods of forecasting severe weather and verifying severe weather outbreak predictions within a NWP forecast will be employed in this study. For the verification data set, the research will use the definitions of tornadic and nontornadic outbreaks employed in Shafer and Doswell (2010—hereafter SD10), specifically their N15 index.

Since human forecast decisions are often highly subjective, inconsistencies in the forecast outcomes from such a subjective determination suggest the need for an objective classification method for outbreak type. Support vector machines (SVMs - Burges 1998, Hearst et al. 1998) are a type of learning machine well suited to binary classification applications, such as tornadic/nontornadic outbreak types. SVMs employ an optimized decision hyperplane to demarcate classes in a binary fashion, with all points on one side corresponding to one outcome and all on the other side to the other outcome (Fig. 1.1a).

SVMs also utilize kernel methods to project nonlinearly separable data into a higher dimensional hyperspace where they are linearly separable, enhancing their utility for classification applications (Fig. 1.1b).

SVMs have been used in meteorological research before. SVMs have been used in aerosol modeling (Ackerman et al. 2004), classification of satellite radiance data into cloud types (Lee et al. 2003), and downslope windstorm forecasting in Colorado (Mercer et al. 2008). Most pertinently to this study, Mercer et al. (2009—hereafter M09) examined the use of SVMs to distinguish between tornadic and nontornadic severe weather outbreaks. They noted the inherent classification ability of SVMs in selecting them as the primary objective classification method. Following studies by Weisman and Klemp (1984), Droegemeier et al. (1993), McNulty (1995), Colquhoun and Riley (1996), Stensrud et al. (1997), Rasmussen and Blanchard (1998), Markowski (2002), and Davies (2004), M09 initially examined 15 meteorological parameters found to be important in severe convective weather prediction before using a permutation test (Efron and Tibshirani 1993) to determine the diagnostic variables that were most distinct between the two outbreak types, after which a SVM was trained on these significant parameters. The M09 study yielded low false alarm ratios (FAR; less than 0.3) while maintaining a high skill at discriminating outbreak type (greater than 0.7). However, they noted that the FAR values remain too high, suggesting additional work is needed in improving outbreak discrimination.

Other statistical methods exist for producing a probabilistic forecast of a binary outcome, such as logistic regression. In the context of outbreak mode prediction, logistic regression is linear with respect to the probability of a "hit" (here, a tornado outbreak).

The kernel function of an SVM, however, may be linear, polynomial, or Gaussian. Using this kernel function, the SVM iteratively tunes its decision hyperplane function to optimize separation between data points, learning from predictor-predictand sets of data. The learning and nonlinear approaches give the SVM added flexibility to adapt its decision hyperplane function for a given volume of data.

These examples use a set of data in which color indicates a distinction between two classes. The dashed diagonal line in the right panels of Figure 1.1a and Figure 1.1b is the decision hyperplane. The points closest to the decision hyperplane are the "support vectors" of the algorithm; it is the distance between them that must be maximized.

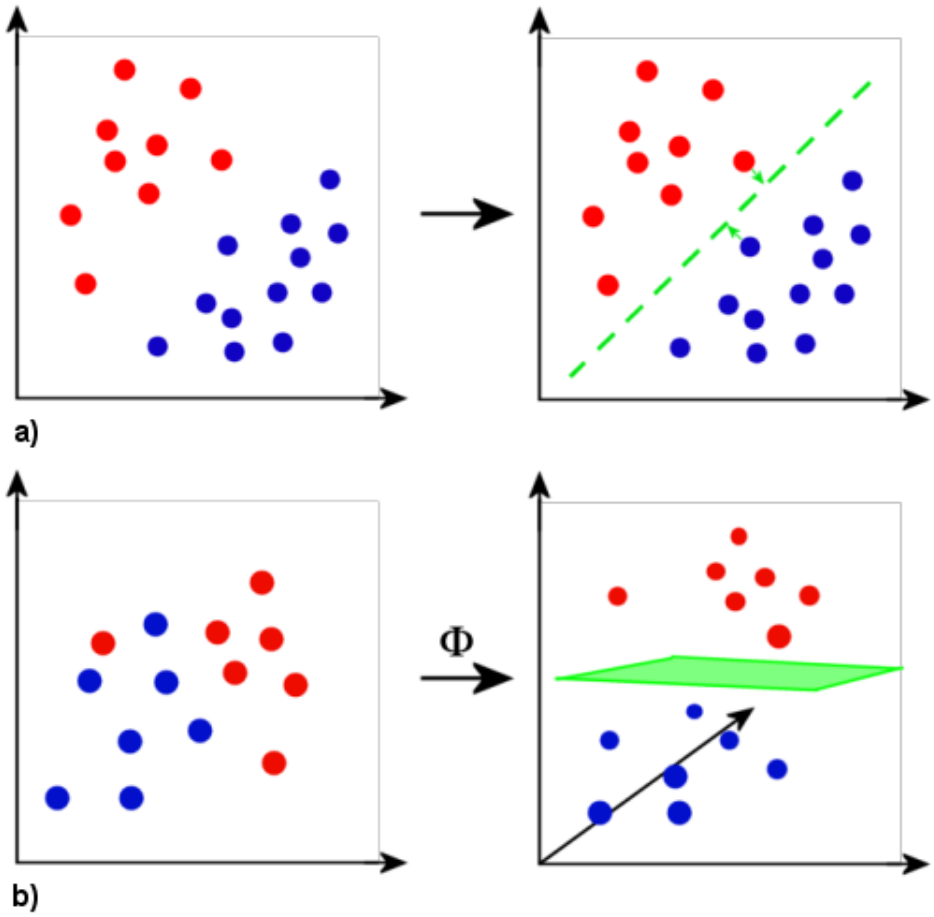

Figure 1.1 Support Vector Machine (a) and Higher-Dimensional Kernelization (b) with a Kernel Function Φ

1.3 Research Objectives

While nontornadic severe weather outbreaks can be highly dangerous and destructive, tornado outbreaks account for a far greater casualty and property damage toll. Operational forecasters try to predict outbreak mode as far in advance as possible, given costs to government and business associated with preparing for high-impact weather events and the psychological "crying wolf" problem when high false alarm rates exist. Minimizing the false alarm rate for all types of extreme weather forecasts is an ongoing

goal in operational meteorology, but the specific danger of tornado outbreaks makes it especially critical for this type of event.

This research proposes two overarching questions to be answered:

- Whether model physics parameterization or data assimilation produces the greater model spread in atmospheric parameters that are important for severe weather.
- Which model configuration within each set—a data assimilation ensemble and a physics parameterization ensemble—produces the most accurate outbreak forecast

These questions require a three-part research project to model and examine variation across a representative set of severe weather events. Phase I involves the assimilation of data from two sources—conventional observations and satellite infrared radiances—to determine the impact of this data assimilation on severe weather outbreak forecasts. Phase II involves the use of a microphysics and PBL physics parameterization ensemble to determine the effect of parameterization scheme choices on outbreak mode prediction. Phase III will examine the impact of both data assimilation and physics parameterization on synoptic-scale atmospheric fields. Within each phase of the project, additional research questions specific to that phase are raised.

In the data assimilation phase, the key questions to answer are whether the assimilated data used in the project produce an improvement above a control in the accuracy of SVM severe weather outbreak classifications, and if so, which type or types of assimilated data produce the most significant improvement to an outbreak mode forecast. It is hypothesized that the assimilation of conventional data in conjunction with high-resolution satellite data produces the most skilled outbreak forecast. The effects of

conventional data assimilation on model accuracy (with respect to an observation data set) are well documented, with researchers such as Wheatley et al. (2012) finding that assimilating these data brought a tornadic parameter closer to observed data. Jones and Stensrud (2012) found that the use of AIRS satellite sounding data improved the forecast of middle- to lower-atmospheric dewpoint temperature, an atmospheric parameter that is important in forecasting severe weather. These prior studies suggest that the assimilation of both conventional and satellite observations should produce improvements in an explicit probabilistic outbreak-mode forecast, relative to a control forecast and to forecasts made with the assimilation of only one such type of data.

In the model physics parameterization phase, the key question to answer is which combination of model physics provides the most accurate modeling of severe weather mode across the case set. Despite the volume of research concerning the effects of parameterization schemes on specific physical processes and individual cases, no study has formally identified the performance of a suite of physics parameterizations in severe weather outbreak mode forecasting; therefore, the goal of this phase of the research is to identify the individual impact of certain physics parameterization schemes on the SVM's ability to distinguish tornadic and nontornadic outbreaks within a NWP framework. Given that M09 found the most statistically significant parameters for distinguishing outbreak mode to be largely near-surface or lower-atmospheric, it is hypothesized that the choice of planetary boundary layer scheme produces a greater impact on outbreak mode forecasts than other examined types of model physics.

In the third and final phase of the study, the research question is what effect that data assimilation (as conducted in phase I) and microphysics/PBL parameterization (as

conducted in phase II) have on the modeling of synoptic-scale weather features in tornadic and nontornadic outbreaks. With regard to parameterization, it is hypothesized that microphysics parameterization will influence the modeling of upper-level cloud and wind features more strongly, whereas PBL physics will heavily influence lower-level thermodynamic and wind variables. With regard to data assimilation, it is hypothesized that conventional observations will most strongly influence thermodynamic variables and satellite radiances will influence cloud features and wind. The reasoning for this is that the conventional data set contains direct observations of atmospheric moisture, whereas a data assimilation system must calculate moisture values from satellite infrared radiance temperatures, and these calculations are highly sensitive to the quality of the satellite radiance data.

1.4 Data Description

All three phases of the research use model simulations produced by the Weather Research and Forecasting-Advanced Research WRF (WRF-ARW) model, version 3.4.1 (Skamarock et al. 2007). All ensembles of the WRF model are initialized with the North America Regional Reanalysis (NARR) data (Mesinger 2006). The NARR are provided on a 32 km Lambert-Conformal grid with 29 vertical levels and 3-hour temporal resolution. The NARR data are reanalyzed from recorded weather observations using the North American Mesoscale (NAM) model, formerly known as the Eta model, and assimilated with the Eta data assimilation system. Known sources of error with the initialization data include imprecision in surface wind stress (Ebisuzaki and Rutledge 2004) and diurnal inaccuracies in 2m temperature fields (Mesinger 2006). Due to the fact that this research study focuses on severe convective weather events that are driven in

part by boundary-layer thermodynamic and wind shear conditions (Kerr and Darkow 1996, Thompson et al. 2007), these errors in the NARR data are of significance and provide a significant opportunity for improvement through data assimilation, as well as potential modeling improvement through optimal choices of physics parameterization schemes, which in turn could improve outbreak classification by the WRF.

1.4.1 Event Selection

The research examines a set of 40 severe weather events in North America that took place between 2008 and 2011, as listed in Table 1.1. Doswell et al. (2006—hereafter D06) and Shafer and Doswell (2010—hereafter SD10) define a tornado outbreak as six or more tornado reports within a single synoptic-scale system over a 24 hour period (1200 UTC to 1200 UTC the following day). SD10 went further and ranked outbreaks based on their severity using a variety of outbreak intensity indices, based on characteristics of the outbreak that are listed in Table 1.1. They found that their N17 index was most representative of tornado outbreaks. As such, 20 tornado outbreaks ranked in the top 35 according to SD10 and 20 nontornadic outbreaks (those with five or fewer tornado reports that exhibited a large number of hail or wind reports) were selected for comparison in this study (Table 1.2). All storm reports were taken from the Storm Prediction Center's Storm Data database (NCDC 2010). A selection of 20 events was chosen from the top 35, rather than simply the top 20, because the satellite radiation sounder data used in this research were not available until 2008, and gaps exist in the satellite data set for certain days in which severe weather outbreaks occurred.

Table 1.1 SD10 Variables Used for Ranking Outbreaks

Total number of severe reports	Total number of violent tornadoes
Total number of tornadoes	Number of long-track tornadoes
Total number of hail reports	Number of killer tornadoes
Total number of wind reports	Destruction Potential Index (D06)
Total number of significant hail reports	Total path length
Total number of significant wind reports	Fatalities
Total number of significant tornadoes	Middle-50% parameter (D06)

Table 1.2 Severe Weather Outbreaks Simulated in the Study

Tornado Outbreaks	N17 Index	Report Count			Nontornadic Outbreaks	N17 Index	Report Count		
		Tor	Hail	Wind			Tor	Hail	Wind
7-Jan-08	2.325	47	171	119	15-Jun-08	0.193	3	267	274
5-Feb-08	8.889	85	145	313	8-Jul-08	-0.099	2	42	272
15-Mar-08	3.039	46	209	87	20-Jul-08	0.301	0	64	348
9-Apr-09	1.737	24	207	121	2-Aug-08	0.194	1	193	337
10-Apr-09	3.490	62	387	133	11-Feb-09	0.077	3	7	375
8-May-09	2.079	46	128	159	15-May-09	0.150	5	260	99
24-Apr-10	5.252	35	119	149	3-Jun-09	-0.180	1	117	68
10-May-10	4.034	69	126	60	18-Jun-09	0.140	5	173	312
5-Jun-10	2.268	46	3	55	24-Jul-09	0.200	3	166	116
17-Jun-10	4.056	73	136	87	9-Aug-09	0.177	2	78	241
4-Apr-11	2.261	46	116	1088	6-Apr-10	-0.078	0	188	60

Table 1.2 (Continued)

15-Apr-11	4.707	67	195	95	28-May-10	-0.250	0	89	120
16-Apr-11	5.096	55	51	130	15-Jun-10	-0.047	0	88	325
19-Apr-11	2.865	80	443	513	18-Jun-10	0.398	1	126	387
25-Apr-11	3.091	50	162	230	23-Mar-11	0.369	4	278	167
26-Apr-11	7.748	105	224	207	3-Apr-11	0.154	0	304	64
27-Apr-11	29.342	172	202	329	11-Apr-11	-0.226	1	5	112
22-May-11	4.879	47	487	279	20-Apr-11	-0.079	2	144	57
24-May-11	6.420	48	234	210	4-Jun-11	-0.099	0	98	122
25-May-11	3.242	90	467	496	9-Aug-11	0.446	1	129	132

Each phase of the research considered the same case set. These cases were simulated in the WRF model according to the requirements of each phase of the study, as detailed in the following sections of this document. Phases I and II involved subsetting a large regional model domain (described in sections following) into smaller, outbreak-centric domains, whereas phase III used the large regional domain for analysis. The smaller domains used in phases I and II are centered on the severe weather outbreaks. These outbreak centers are estimates of the geometrical centers of storm report mappings for each outbreak. The SD10 data set includes some dates in which more than one synoptically distinct outbreak occurred in North America; in these cases, SD10 treats each synoptic weather feature as a separate outbreak. When more than one synoptic weather feature produced severe weather reports on a given day, this research used only

reports associated with the weather feature of interest for estimation of an outbreak center.

1.4.2 Outbreak Valid Time

A temporal difference in the initiation of tornadic and nontornadic outbreaks was observed, with the tornado outbreaks generally beginning three to six hours earlier in the day than the nontornadic outbreaks. The nontornadic events also tended to last several hours longer than the tornadic events, while tornadic activity generally ended comparatively quickly. Seasonal differences in the frequency of tornadic and nontornadic events affected the timing of these events, with longer periods of heating in the summer contributing to more nighttime severe wind events. These average timing differences between outbreak modes indicated that using a single valid time for all outbreaks would not produce optimal forecasts from the SVMs; therefore, it was necessary to determine the outbreak peak time for each case.

This peak time was determined by calculating the mean time of all severe weather reports of the appropriate kind (tornado or thunderstorm wind) that were reported within the SVM input field (a 32 x 32 gridpoint grid, as described in the remaining sections) and rounding up or down to the nearest 3-hour period. Table 1.3 shows the calculated outbreak peak time for each of the 40 outbreaks.

Table 1.3 Outbreak Peak Times

Date	Peak (UTC)	Date	Peak (UTC)
7 Jan 2008	00	15 Jun 2008	21
5 Feb 2008	00	8 Jul 2008	21
15 Mar 2008	21	20 Jul 2008	00
9 Apr 2009	00	2 Aug 2008	03
10 Apr 2009	21	11 Feb 2009	00
8 May 2009	18	15 May 2009	21
24 Apr 2010	18	3 Jun 2009	00
10 May 2010	21	18 Jun 2009	21
5 Jun 2010	03	24 Jul 2009	00
17 Jun 2010	21	9 Aug 2009	21
4 Apr 2011	21	6 Apr 2010	21
15 Apr 2011	21	28 May 2010	00
16 Apr 2011	21	15 Jun 2010	00
19 Apr 2011	03	18 Jun 2010	00
25 Apr 2011	00	23 Mar 2011	00
26 Apr 2011	06	3 Apr 2011	00
27 Apr 2011	00	11 Apr 2011	00
22 May 2011	21	20 Apr 2011	00
24 May 2011	00	4 Jun 2011	21
25 May 2011	21	9 Aug 2011	00

The remaining sections of this dissertation will document specific details about the methods that are used in each of the three phases of this research, including background information on the tools, module-specific data choices, and configuration details for each phase. After the module-specific chapters, this document will contain a final chapter summarizing the findings of each phase and linking the three modules to the overall research objectives.

CHAPTER II

PHASE I—OUTBREAK PREDICTION WITH DATA ASSIMILATION

2.1 Methodology

2.1.1 Overview

The first phase of this project entails the generation of NWP ensembles with variation in the sources of data used by the NWP model. Numerous data observation platforms exist, including balloon soundings, terrestrial weather reporting stations, satellite observations, radar, ship reports, and aircraft reports. Assimilated data sources are of interest in this study because these observations are not all of uniform quality, and some are difficult to assimilate into a NWP model without degrading the quality of the forecast. This research does not examine all possible data observing platforms for assimilation. It focuses on conventional meteorological observations and satellite radiances, because these data types are readily available to meteorologists and are assimilated into operational NWP models. The fact that these observations include numerous important meteorological variables—temperature, moisture, wind, pressure—and are recorded at many levels of the atmosphere makes them desirable to assimilate.

2.1.2 Background Field

All data assimilation procedures require a background field onto which data may be assimilated. The WRF-ARW model was used to generate this field. The model configuration used a 12 km Lambert-conformal grid background field (Fig. 2.1)

encompassing much of North America east of the Rocky Mountains. This background domain was identical for all 40 severe weather cases. The time step of this configuration was 72 seconds, with output files generated for each hour. The model output spanned a 42-hour period beginning at 1800 UTC on the day before the outbreak (to allow for model spin-up) to 1200 UTC on the day following the outbreak. The geographical domain was chosen to encompass the spatial extent of all the outbreaks. The vertical dimension included 28 vertical levels with a model top of 100 hPa. Model physics used in generating this background field are listed in Table 2.1 below.

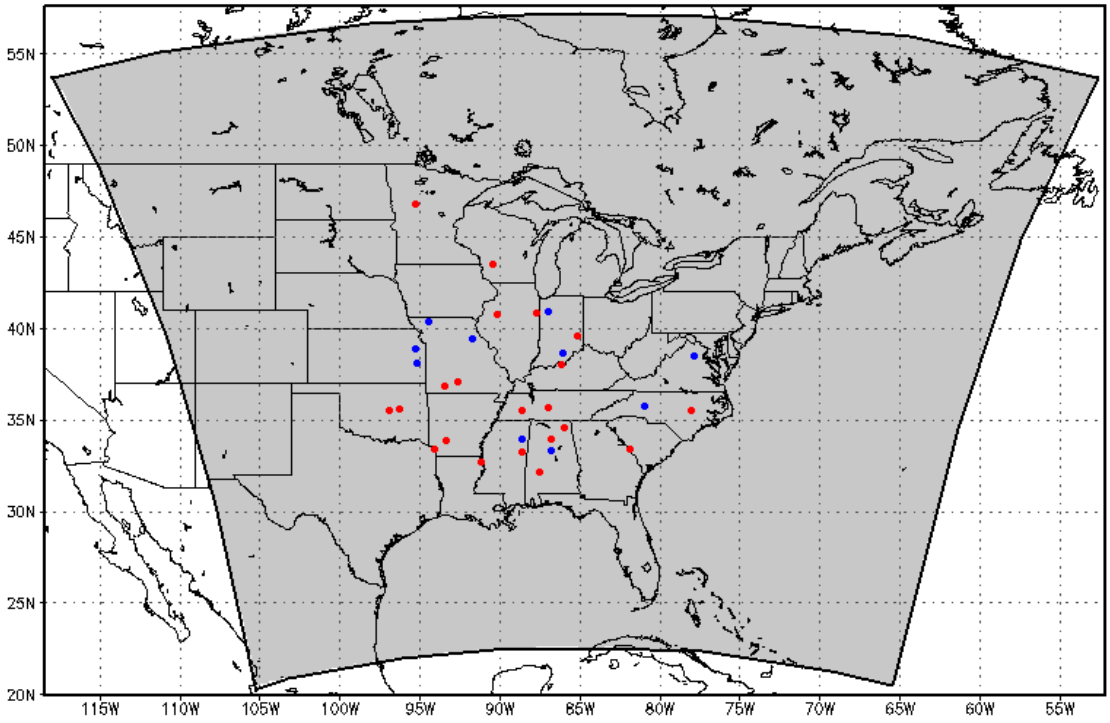

Figure 2.1 WRF Model Domain for Background Simulations with Tornadic (Red) and Nontornadic (Blue) Outbreak Centers Superimposed

Table 2.1 Phase I WRF Physics Parameterizations.

WRF physics option	Configuration
Cloud microphysics	Thompson et al. (2008)
Longwave radiation	Rapid Radiative Transfer Model (Mlawer et al. 1997)
Shortwave radiation	Dudhia (Dudhia 1989)
Surface layer	MM5-derived (Beljaars 1994)
Land surface	5-layer thermal diffusion (Dudhia 1996)
Urban surface	None
Planetary boundary layer	Yonsei University (Hong et al. 2006)
Cumulus physics	Kain-Fritsch (Kain 2004)

2.1.3 Data Assimilation

It is common to separate severe outbreak modes into tornadic and primarily nontornadic outbreak types for such forecast applications (Shafer et al. 2009, Mercer et al. 2009, Mercer et al. 2012, others). This phase of the study sought to identify the importance of data assimilation in outbreak type discrimination. This was accomplished through the assimilation of two different types of data: conventional observations from the NCEP ADP Global Upper Air and Surface observation database (NCEP et al. 2008) and High Resolution Infrared Radiation Sounder-4 (HIRS-4—NCEP et al. 2009) radiance data. The NCEP ADP Global Upper Air and Surface observations (hereafter conventional observations) data set contains temperature, moisture, and wind observations from a wide variety of sources and coverage, including aircraft, buoys, land-based recording stations, ships, rawinsondes, and satellite soundings. The High Resolution Infrared Radiation Sounder-4 (HIRS-4) is a sensor on the NOAA-18 and METOP-2 polar orbiting satellites, though observations from NOAA-19 were not available in the data set until 2010. It has 19 infrared channels, one visible channel, and a resolution of 10 km at nadir. Brightness temperatures from the HIRS-4 are utilized in the

data assimilation methods employed. Both the satellite and the conventional data were available daily in 6-hour bins beginning at 0000 UTC.

Once the background fields for each case were generated, assimilation of each individual type of observation (conventional and HIRS-4) as well the two in combination, was done using the Gridpoint Statistical Interpolation (GSI—Wu et al. 2002) software, version 3.1. The GSI was configured for three-dimensional variational data assimilation (3DVAR—Kleist et al. 2009). Although more sophisticated methods of data assimilation exist, namely EnKF and 4DVAR, the 3DVAR approach has been used effectively in operational NWP models. This method has the advantage of simplicity over EnKF in that it does not require the production of full stochastically perturbed ensembles. The choice of 3DVAR over 4DVAR implies that some temporal displacement of meteorological features in the assimilated fields is possible in the experimental model configuration if the WRF natively models their progression significantly inaccurately. It was determined that this possible model error would be accepted in the study due to the fact that the 4DVAR assimilation option was untested in the GSI 3.1 release.

A background error covariance field was produced for each case using the National Meteorological Center method (NMC—Parrish and Derber 1992). Background error covariance fields are used in data assimilation to determine the effect of assimilated observations on surrounding areas and the weighting of variables. To produce these background error covariance fields, 12 km WRF runs initialized at 0600 UTC and 1800 UTC the day of the outbreak were generated, using NARR data as initial input. Times used to generate the background error covariance fields were 1200 UTC, 1800 UTC, 0000 UTC the next day, 0600 UTC, and 1200 UTC. The WRF tool “gen_be”

(Descombes et al. 2014) was used to generate background error statistics for each case from these output fields. Once the background error covariance statistics were produced, conventional and satellite radiance observations were then assimilated to the 42-hour WRF output at each 6-hourly time interval in which they were available.

Channels 2-15 from the HIRS-4 unit were assimilated. Cycling was used to determine satellite bias coefficients, with GSI North American Mesoscale (NAM) regional default values used for the initial time (1800 UTC). These error fields are computed from the NCEP NAM model's forecast error (compared to observations). The satellite bias coefficients for each successive time step of the case used error calculations computed for the previous time step by the GSI. Conventional observations were assimilated without thinning. HIRS-4 observations were assimilated at a 4 km grid in GSI's data thinning process, a value determined after initial experimentation at coarser thinning resolutions, including 12 km, yielded low observation assimilation counts. For instance, assimilation at 12 km thinning resolution reduced the number of HIRS-4 observations by 67 percent in some outbreaks compared to assimilation at 4 km. Due to the varying view angle of the polar-orbiting satellite as it traversed its orbit, some observation bins, such as 12Z, already contained low counts relative to other bins (such as 18Z). Using a coarse thinning resolution yielded few to no observations being assimilated for these time blocks.

It is important to note that while fine-tuning of satellite radiance assimilation is performed for case study research to generate optimal data fields, the assimilation procedures used herein were not individually tuned for each of the 40 cases beyond the cycling procedure. This research examined cases of varying outbreak modes with the

aim of determining the benefits of radiance data assimilation—as a general procedure—toward predicting severe outbreak mode in an operational forecast setting. In common with other statistical tools, the learning machine used for making these predictions is designed to be trained on data having uniform characteristics (except for the predictor variables that are to be compared to each other); therefore, manually tuning data assimilation in a different manner for each case diminishes the ability to diagnose impacts of individual assimilation data types in outbreak prediction since it would be uncertain whether the individual fine-tuning or the distinct type of assimilated data caused any changes in classification performance. Additionally, an objective of this work was to implement a support vector machine classification scheme on the assimilated fields, and any inconsistencies among the assimilation procedures would create biases in the predictor variables used in the learning machine and potentially degrade the classification performance of the method.

The precise number of satellite observations assimilated varied by date and time due to variations in the satellite orbit in the spatial coverage of the sensor with respect to the background field domain. The data set used in this study included HIRS-4 observations from just two satellites (and only METOP-2 prior to 2010). Fig. 2.2 depicts a box plot showing the 40-case spread of HIRS-4 observations assimilated at each time step. Though precise totals sometimes varied slightly if observations were assimilated in conjunction with conventional observations, these variations were observed to be on the order of ± 1%. As shown in the figure, the largest numbers of HIRS-4 observations from this sensor were assimilated at 1800 UTC, indicating peak satellite observation coverage during or immediately prior to most of the outbreaks.

In the assimilation process, cost function for each observation bin is calculated by GSI. This cost function takes into account observation biases and error:

$$J = (x_a - x_b)^T B^{-1}(x_a - x_b) + (Hx_a - o)^T O^{-1}(Hx_a - o) + J_c$$

where x_a and x_b are analysis and background fields, B is a background error covariance matrix, H is an observation operator, o is an observation field, O is observation error covariance, and J_c is a placeholder for constraint terms such as dynamical and moisture constraints.

This function is iteratively updated with each phase of the data assimilation process. The assimilation procedure requires three iterations for each time step. In each step, the GSI defines a separate variable:

$$y = B^{-1}x \tag{2.2}$$

where x is the difference of the analysis and background fields. Equation (1) is reformulated in terms of equation (2) as follows:

$$J = y^T B^T y + (HBy - o)^T O^{-1}(HBy - o) + J_c \tag{2.3}$$

For both equation (1) and equation (3), the gradients of the background field and observation field terms of the cost function are calculated. The GSI then attempts to simultaneously minimize both gradient calculations.

One useful metric for this process is the difference of this cost function after the final iteration and after the initial one. Fig. 2.3 depicts a histogram of the per-observation final-initial cost function difference for HIRS-4 radiances. As the chart indicates, the final cost function was lower than the initial in every assimilation, indicating that the GSI software was operating as expected through the iterations.

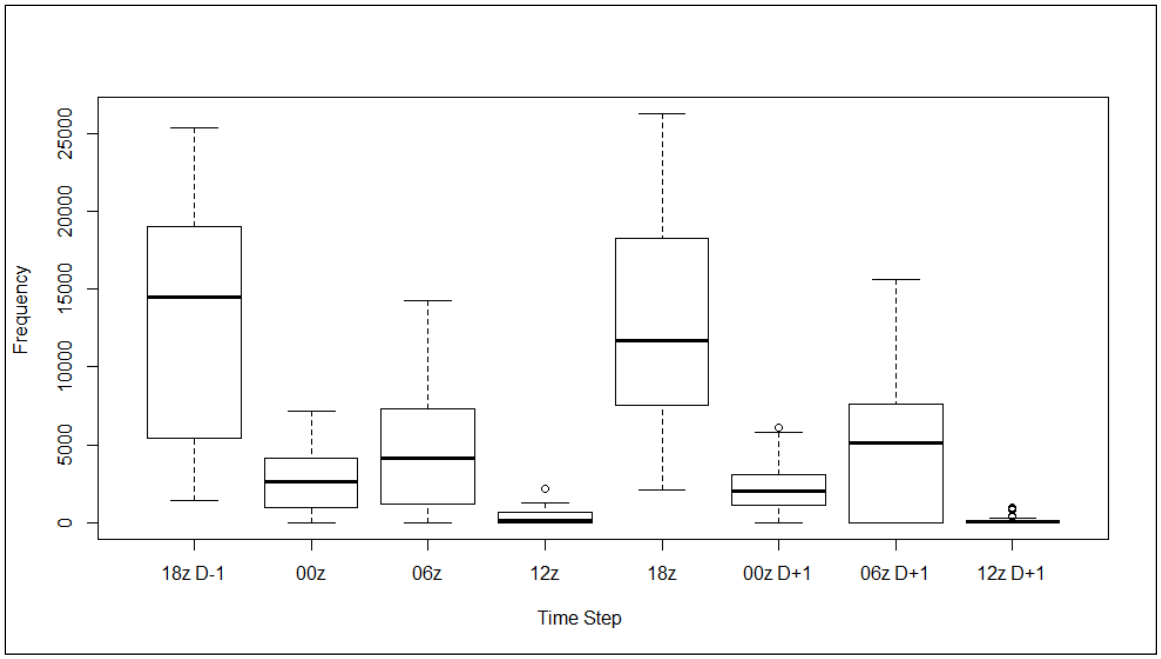

Figure 2.2 Spread of HIRS-4 Observation Counts Assimilated Among the 40 Cases for Each Time Step

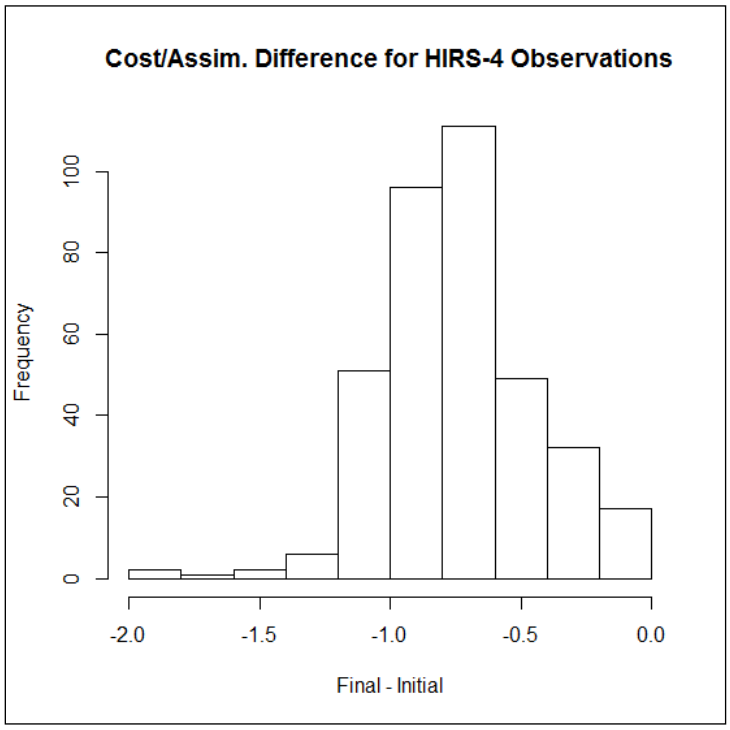

Figure 2.3 Distribution of Final Minus Initial Iteration GSI Cost Function for HIRS-4 Observations

2.1.4 High-Resolution Outbreak Domains

The output of the GSI data assimilation procedure yielded a 12 km domain identical in spatial extent to that of Fig. 2.1. The three possible updated background field combinations (HIRS-4, conventional, both) were used in the initialization of a new WRF simulation conducted at 4 km spatial resolution for the same 42 hour time period. Additionally, a control run in which no assimilation was done was formulated for each case, yielding four possible WRF simulations. Model physics used in these new WRF runs were identical to those used in the 12 km background simulations, except that no cumulus physics parameterization was used in the 4 km simulations (consistent among all four assimilation experiments). Model output was generated in 3-hour time intervals.

The domains of the 4 km WRF simulations were centered on the regions in which the tornado or other severe weather outbreaks occurred (i.e. the points in Fig. 2.1). The size of the domain varied according to the spatial extent of an outbreak. Fig. 2.4 shows an example of the domain of one case superimposed on the background field and associated official tornado reports.

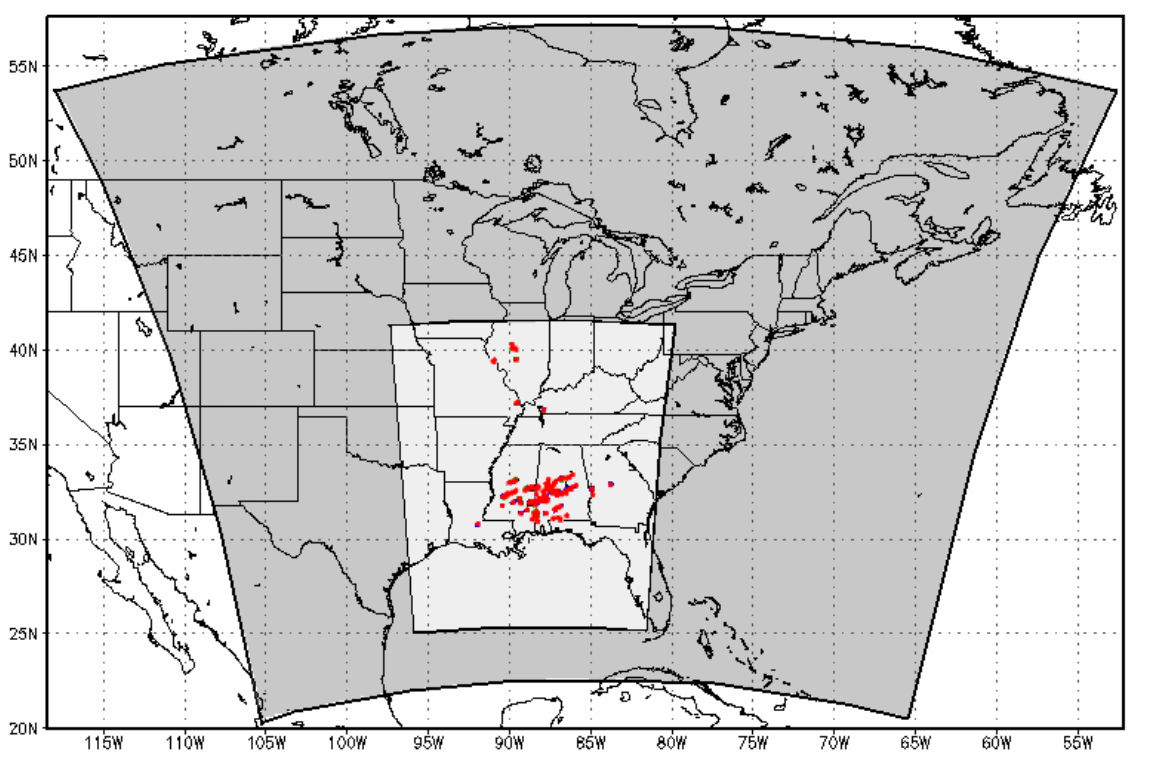

Figure 2.4 An Example of a Single Case Subdomain (15 April 2011)

Observed tornado reports are superimposed in red.

2.1.5 SVM Classification of Outbreak Mode

The output from the assimilated WRF simulations is not capable of discerning outbreak type directly. M09 dealt with this by using a support vector machine (SVM) to classify outbreaks as tornadic or nontornadic with modest success. This study utilized the same SVM approach on the output of the 4 km WRF simulations to establish which type of outbreak the WRF output is suggesting.

Output from the 4 km assimilated WRF fields at the outbreak valid time (section 1.4.2) was upscaled to 12 km by data thinning. The upscaling was deemed necessary due

to the generation of data noise by WRF at isolated locations at high horizontal resolution, which reduced SVM forecast accuracy at the "raw" 4 km resolution, and the computational time needed to run the SVM analysis on large matrices. For strictly research purposes, high computation time is not a concern, but the SVM was employed as a possible tool to aid operational forecasting, which is inherently time-sensitive. In addition, M09 found that SVMs trained on the same severe weather covariates for the same task—differentiation of outbreak mode—performed well on 18 km data.

A 32 x 32 gridpoint grid 12 km in resolution centered on the outbreak was retained for each of the 40 cases of the diagnostic predictor variables deemed important at 24 hours prior to the outbreak by M09. These variables were storm-relative helicity (SRH) at 0-1 km, SRH at 0-3 km, surface-based convective inhibition (CIN), bulk shear at 0-1 km, the product of 0-1 km bulk shear and surface-based convective available potential energy (CAPE), lifted condensation level, and energy-helicity index (EHI) at 0-1 km. Once these variables were retained, an S-mode varimax rotated principal component analysis (Richman 1986) was conducted, yielding between five and nine rotated principal component (RPC) scores (contingent on the input assimilation field) that were used as predictors in the SVM. RPC scores were used as SVM input because the PCA reduces the size of the input data matrix, and therefore the computation time for the SVM, without sacrificing information about the underlying data set.

The SVM method is based on establishing a decision hyperplane between two classes (here tornadic and nontornadic outbreaks) using discriminating variables (e.g. predictors). The method is unique from linear discriminant methods in that the predictors are first projected into a nonlinear hyperspace by a nonlinear map function φ. The dot

product of this map function for one predictor and the same map for another predictor yields a similarity matrix in nonlinear hyperspace, often referred to as a kernel matrix **K**. In addition to the projection into hyperspace, a quadratic programming optimization routine is used to find the optimal margin of separation between the two classes in the projected hyperspace, with greatest weight given to points lying on the margin of separation (known as support vectors). A penalty function can be assigned to give less weight to points far from the margin, affecting the results of the optimization routine. This penalty function, in addition to the selection of the kernel matrix formulation and its tunable parameters, yields a theoretically infinite number of possible configurations for the SVM. As such, cross-validation routines were used to identify the strongest combination of cost and kernel function for discriminating outbreak type.

In this study, the following kernel functions were considered:

Linear: (2.4)

Polynomial: (2.5)

Radial Basis (Gaussian): (2.6)

where **x** and **y** are vectors in original linear space, γ is a user-defined Gaussian spread parameter associated with the polynomial and radial basis kernels, and *d* is the degree of the polynomial (the user dictates this value). In this study, we considered all three kernel functions and multiple degree *d* values (2, 3, 4, 0.5, and 0.333), yielding a total of seven kernel functions considered. Additionally, cost function values ranging from 1 (no penalty to points far from the margin) to 10000 (severe penalty to points far from the margin) on a $\log_{10}$ scale, and γ values of 0.01, 0.05, and 0.1 were considered. This led to

a total of 95 kernel-cost-gamma combinations (the linear kernel is not a function of γ, so only five linear kernel experiments were done for the five cost values).

To identify the best kernel-cost-gamma combination of the 95 tested, a bootstrap 2-fold cross-validation routine was implemented (Efron and Tibshirani 1993), withholding 20 of the 40 events randomly for each bootstrap replicate. Each kernel-cost-gamma combination was tested with the same 1000 random samples of training and testing sets to ensure pairwise comparisons were possible.

Determination of forecast skill was made by calculating contingency statistics on the forecasts of the SVMs. Probability of detection (POD), false alarm ratio (FAR), bias, and Heidke skill score (HSS) were computed (Wilks 2011). The bootstrap confidence intervals of HSS associated with each kernel-cost-gamma combination were used to diagnose statistical significance among the experiments. By identifying those kernel-cost-gamma combinations whose lower HSS confidence limit was higher than the median HSS of other combinations, it is possible to diagnose combinations with statistically significantly better skill. Table 2.2 shows the best-performing SVMs for each of the model configurations.

Table 2.2 SVM with Highest Mean Heidke Skill Score

Ensemble Member	**SVM**
WRF only	Radial basis, cost=10, γ=0.1
HIRS-4	Radial basis, cost=10, γ=0.1
Conventional observations	Radial basis, cost=100, γ=0.01
HIRS-4 + Conventional	Radial basis, cost=100, γ=0.1

2.1.6 Euclidean Distance Calculations and Outbreak Composites

Following determination of the most skilled SVM for each assimilation run, calculations of Euclidean distance between nontornadic and tornadic outbreak values of the seven covariates were conducted. Graphical composites of the covariates were also generated for tornadic and nontornadic outbreaks for each type of data assimilation. These calculations allow for examination of covariates individually to determine the variables that were most strongly modified by data assimilation of a particular type, and the outbreak mode in which the modification occurred.

2.1.7 Observations Test

Finally, a test was conducted to determine whether the data assimilation procedure generated output fields that were closer to observed values than the control. This test was done to validate the assimilation procedure itself, as well as to indicate whether the SVM outbreak mode forecasts were reflecting real improvements in the modeling of data points. Root mean square error (RMSE) analysis was performed for the four model runs for the entire set of 40 cases. Fields examined were geopotential height at 500 mb; *u*-winds at 925, 850, 700, 500, 400, 300, 250, 200, 150, and 100 mb; *v*-wind at the same pressure levels; specific humidity at 700 and 500 mb; and temperature at 925, 850, 700, and 500 mb.

The conventional data set was used for observation points. Although this same data set was used in data assimilation for two of the runs, the assimilation procedure does not simply superimpose values onto the model grid of the background field, but performs physical calculations and weighting of data points based on error analysis and parameters given to the GSI. Additionally, the 4 km outbreak domains were used as the data to be

tested in the RMSE analysis, which meant that the WRF model itself had been run on the immediate output of the GSI. Therefore, RMSE analysis of the conventional and HIRS-4/conventional runs was not a self-comparison. In the RMSE procedure, standard synoptic fields at each 0000 UTC and 1200 UTC period in the model simulation were examined.

2.2 Results

The best output from the cross-validation step of the kernel-cost-gamma optimization was compared among all four phase I ensembles using contingency statistics (probability of detection—POD, false alarm ratio—FAR, bias, Heidke skill score—HSS; Wilks 2011) to establish improvements from data assimilation. Contingency statistics for the SVMs indicated that data assimilation improved outbreak discrimination ability above the control run. This was true for all three data assimilation runs examined. The HIRS-4/conventional member exhibited the highest skill at predicting severe outbreak mode, followed by the conventional run and the HIRS-4 run.

HIRS-4/conventional assimilation produced an HSS of 0.620 (Fig. 2.5c), with a 95% bootstrap confidence interval of [0.607, 0.633]. Conventional data assimilation produced an HSS of 0.603 with a confidence interval of [0.591, 0.615]. The HIRS-4 run produced an HSS of 0.581 and a confidence interval of [0.568, 0.594]. The control run had an HSS of 0.563 and a confidence interval of [0.550, 0.576]. The differences in skill between each member are all statistically significant at the 95% level.

The top performance of the HIRS-4/conventional run was brought about by lowered FAR (Fig. 2.5b) and improved POD (Fig. 2.5a) relative to the control. FAR especially was much lower for this run than any of the others. Bias statistics (Fig 2.5d)

indicate that this run exhibited a small under-prediction bias, which likely accounts for its low FAR; however, the POD of the HIRS-4/conventional run was not significantly lower than that of the conventional-only run. The conventional run demonstrated the lowest bias, with its mean bias value extremely close to 1.00 (unbiased). The conventional assimilation run had the highest POD, though it was within the margin of error of the second-best (for POD) HIRS-4/conventional run.

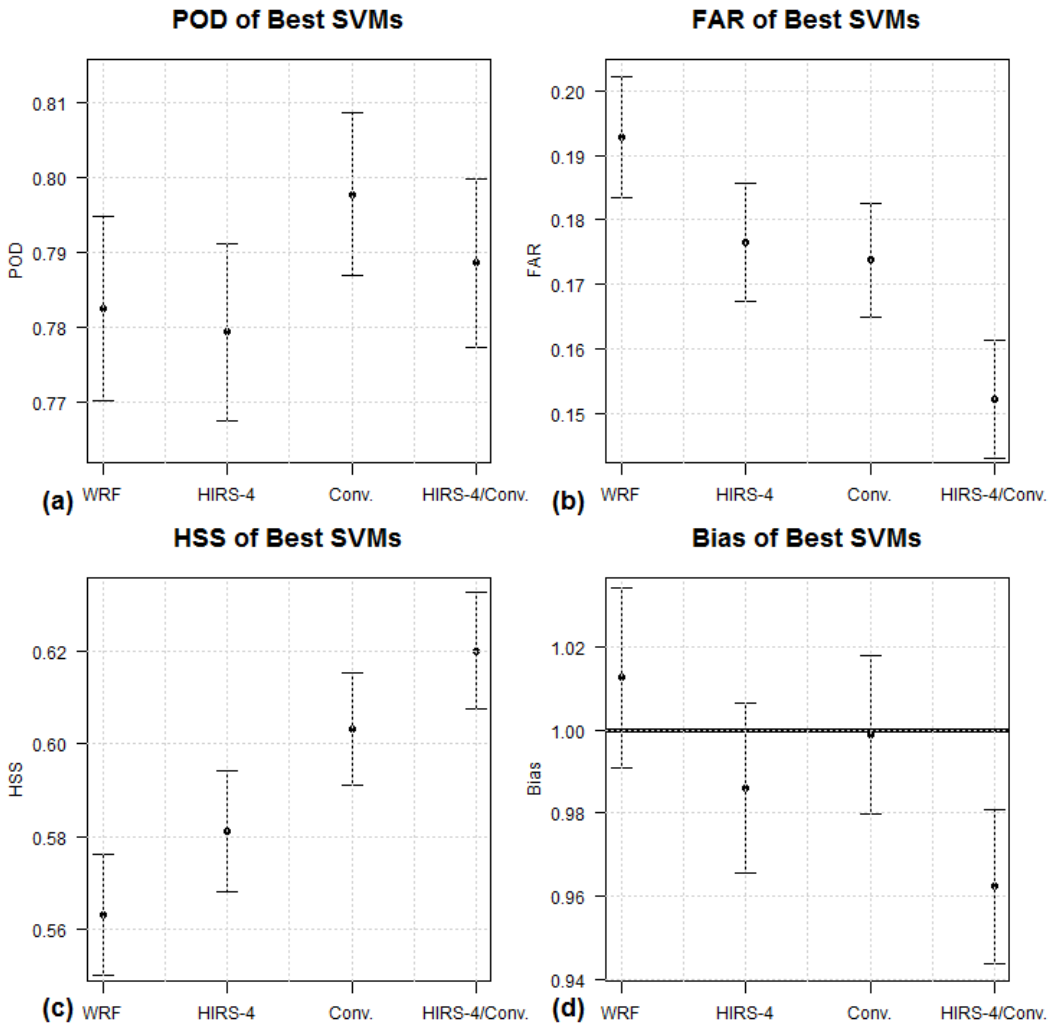

Figure 2.5 95% Confidence Intervals for POD (a), FAR (b), Heidke Skill Score (c), and Bias (d) of the Best-Performing SVMs for Phase I

The solid line in (d) represents an unbiased result.

2.2.1 Data Assimilation and Covariate Values

M09 identified the importance of thermodynamic, shear, and helicity parameters in outbreak discrimination, so it was important to identify the impact of the data

assimilation methods on the average magnitudes of these variables. To determine the impact of assimilation of satellite data and conventional meteorological observations on each type of outbreak, an average composite of all 20 covariate fields for each outbreak type was formulated for each of the four model ensemble members. The composites were generated across the 32 x 32 (12 km grid) outbreak-centered field that was input to the SVM. The contrasts in these composites help demonstrate the exact impacts of the assimilation on the simulations.

Tornadic and nontornadic composites were generated for this data set for 0-1 km bulk shear (Fig. 2.6, Fig. 2.10), 0-1 km bulk shear x CAPE (Fig. 2.7, Fig. 2.11), 0-1 km EHI (Fig. 2.8, Fig. 2.12), and 0-3 km SRH (Fig. 2.9, Fig. 2.13) for all four runs.

2.2.1.1 Tornadic Outbreak Composite

The 0-1 km bulk shear composites show the striking effect of assimilating data in simulations of tornado outbreaks (Fig. 2.6). While the HIRS-4 composite (Fig. 2.6c) depicts a slightly larger area of high bulk shear, the inclusion of conventional observations sharply increased composite values relative to the control (Fig. 2.6a) for tornado outbreaks at peak outbreak times. The area encompassed by high bulk shear values was only slightly larger in the conventional (Fig. 2.6b) and HIRS-4/conventional (Fig. 2.6d) composites, but the area of the highest values was notably increased and the maximum value raised.

The 0-1 km bulk shear x CAPE parameter for tornado outbreaks (Fig. 2.7) was also profoundly influenced by conventional data assimilation. As the figure indicates, the conventional (Fig. 2.7b) and HIRS-4/conventional (Fig. 2.7d) assimilation runs depict a region in the southern part of the composite domain that contains significantly higher

values of this parameter than the control (Fig. 2.7a) and HIRS-4 (Fig. 2.7c) runs. The gradient on the edges of this field of high CAPE x shear is also tighter than in the runs that do not contain conventional data.

The other combined thermodynamic-dynamic parameter, 0-1 km EHI (Fig. 2.8), exhibited a similar pattern from the conventional data assimilation. The maximum EHI values in the composite were markedly larger in the HIRS-4/conventional (Fig. 2.8d) and conventional (Fig. 2.8b) data assimilation runs. Additionally, the assimilation of conventional observations shifts the region of high EHI to the east. The composites also indicate that the assimilation of HIRS-4 observations decreases values of this field. The HIRS-4 composite (Fig. 2.8c) exhibits smaller areas of high EHI than the control composite (Fig. 2.8a), and the HIRS-4/conventional composite depicts smaller areas of high EHI and lower maximum EHI values than the conventional run.

The composites of 0-3 km SRH and 0-1 km SRH displayed the same pattern of effects among the ensemble; therefore, for the sake of examining a covariate over a larger vertical span, the 0-3 km SRH composites are discussed in detail (Fig. 2.9). This field exhibited an interesting pattern with data assimilation. Assimilation of any data resulted in an increase of the maximum values compared to the control, but these maxima were increased much more with the assimilation of conventional data, whether alone (Fig. 2.9b) or, especially, with HIRS-4 radiances (Fig. 2.9d). These maximum outbreak-mean values occurred in the northeastern region of the composite. However, assimilating conventional data also produced a decrease in the values in the entire western half of the composites, especially the far northwest corner. This had the effect of producing a sharper SRH gradient in the conventional and HIRS-4/conventional runs than in the

control and HIRS-4 runs. Finally, the assimilation of HIRS-4 data in combination with conventional data appeared to produce a very sharp boundary (Fig. 2.9d) on the western periphery of the area of highest 0-3 km SRH that was not present in the conventional-only composite. Since satellite radiance values can denote the location and intensity of clouds, it is very likely that this SRH boundary feature in Fig. 2.9d is enhanced by the assimilation of data that, among other effects, can aid the model in simulating storms.

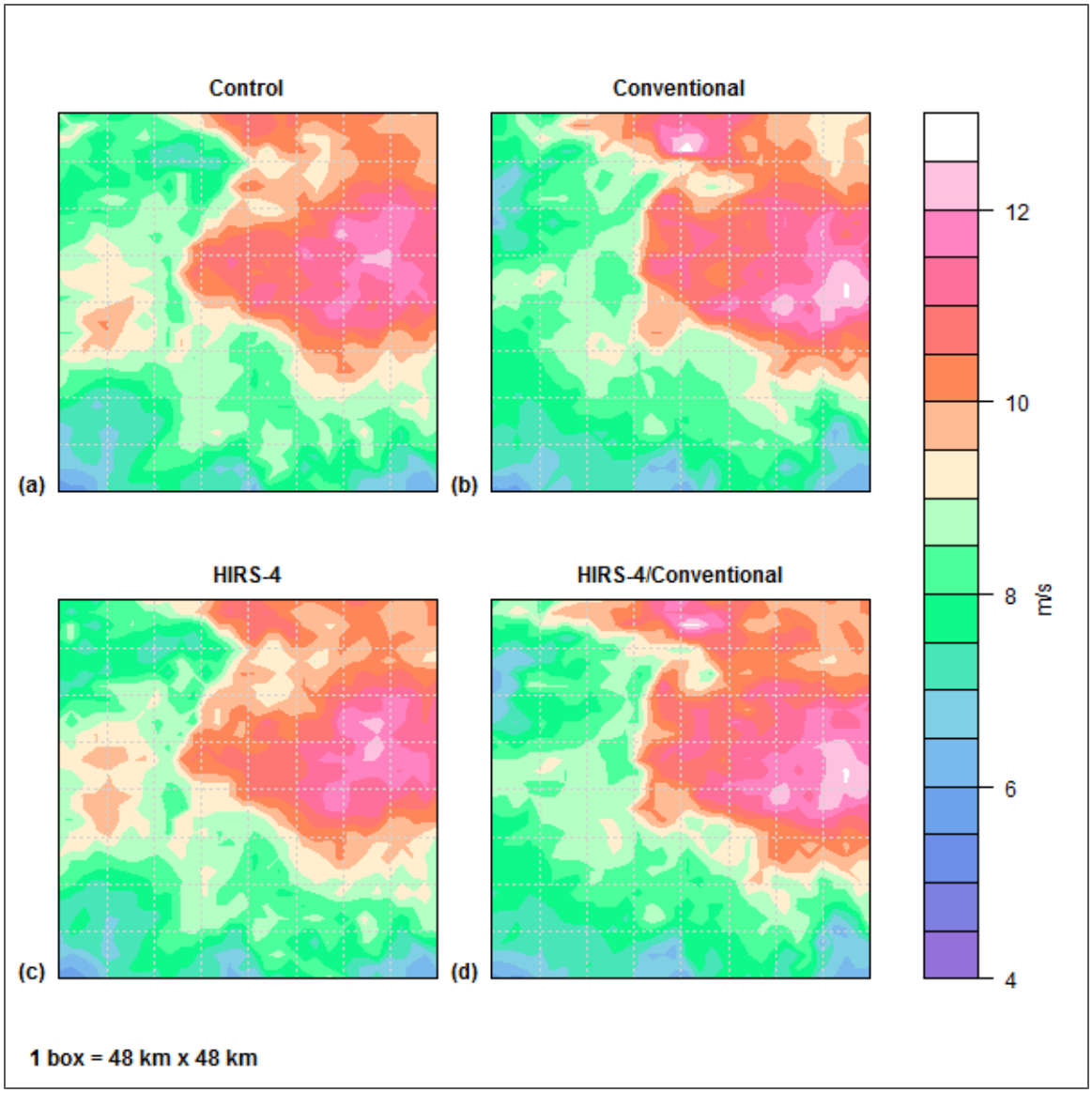

Figure 2.6 Phase I Tornadic Composites of 0-1 km Bulk Shear

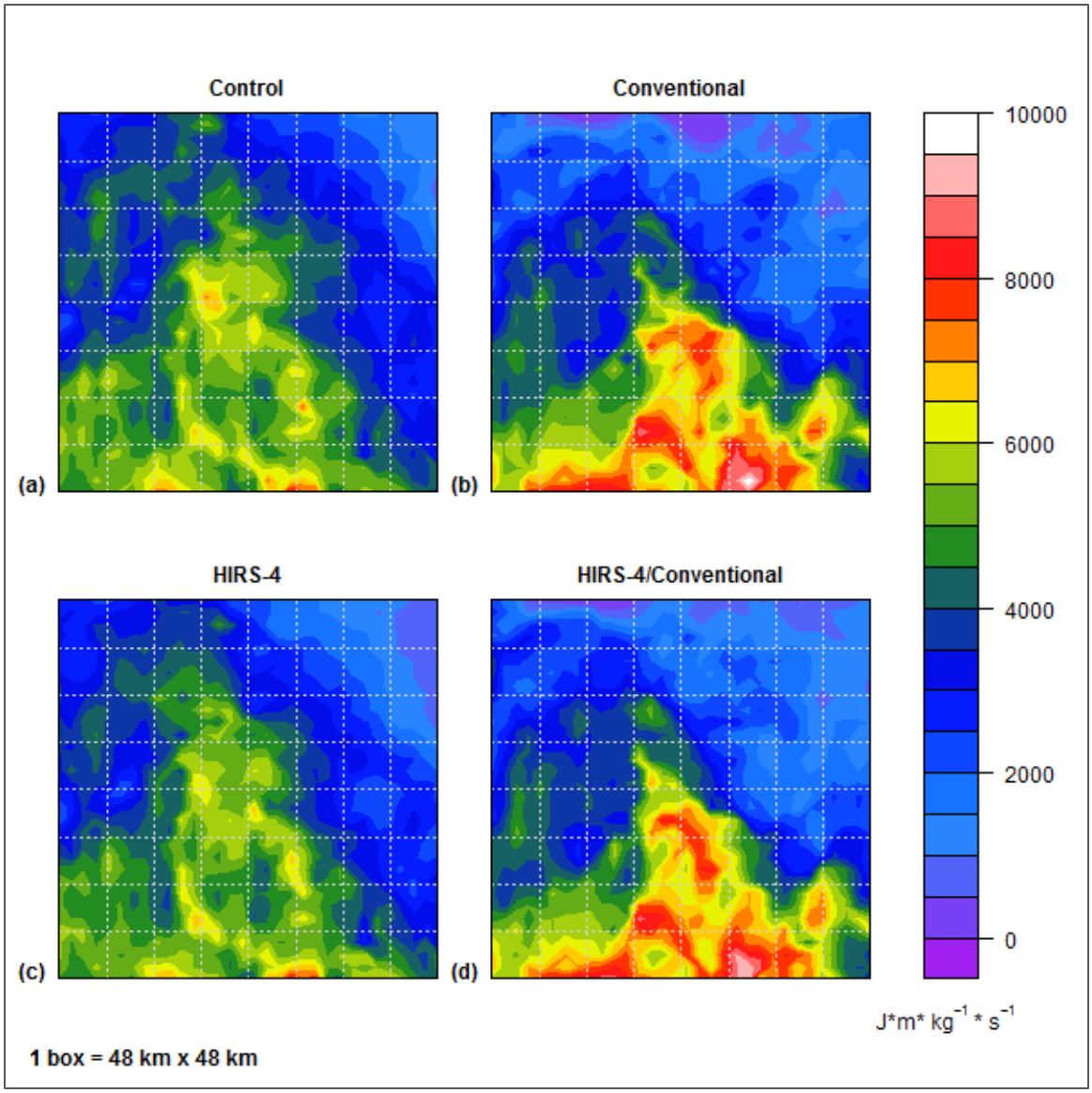

Figure 2.7 Phase I Tornadic Composites of 0-1 km Bulk Shear x CAPE

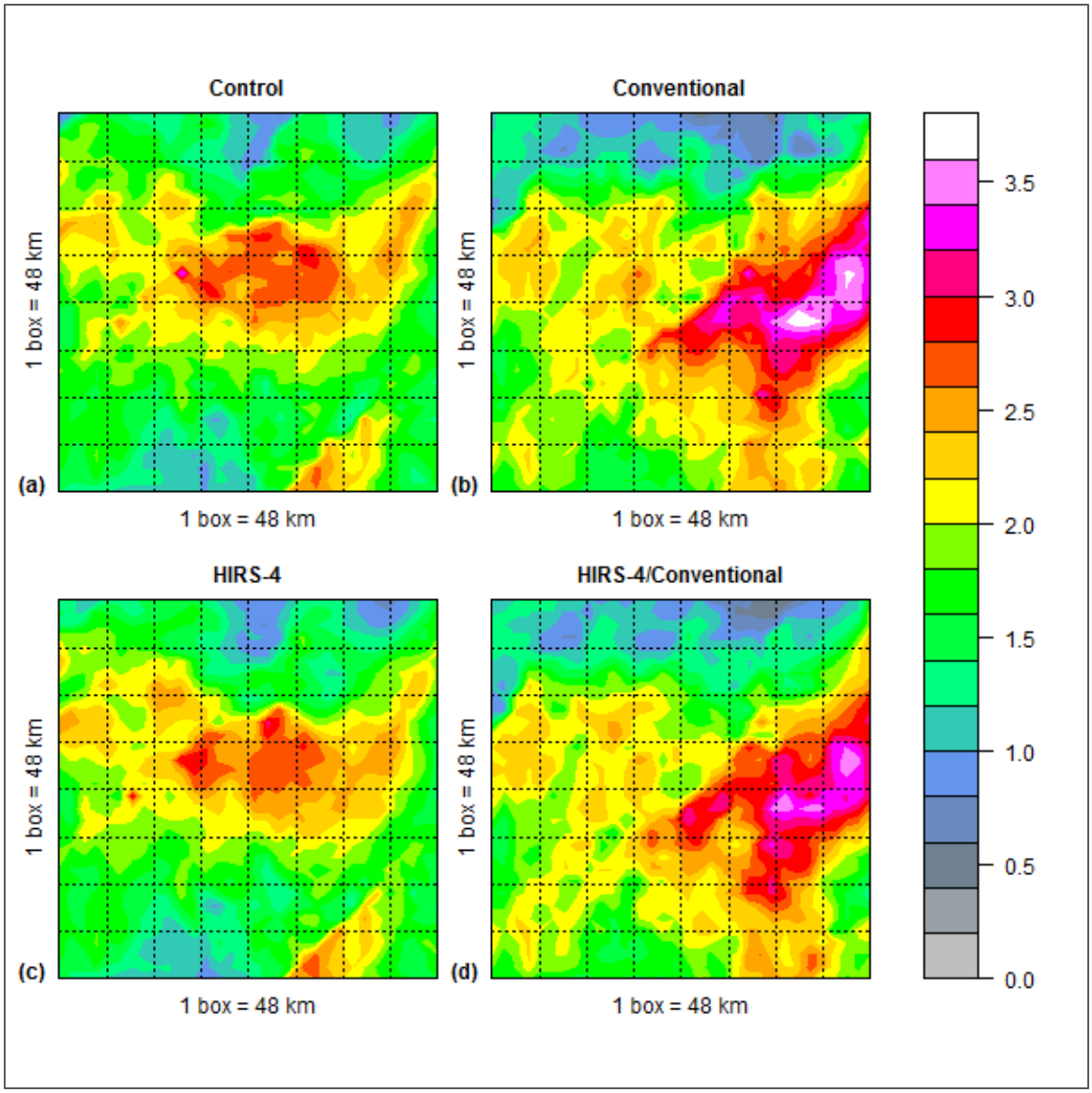

Figure 2.8 Phase I Tornadic Composites of 0-1 km EHI

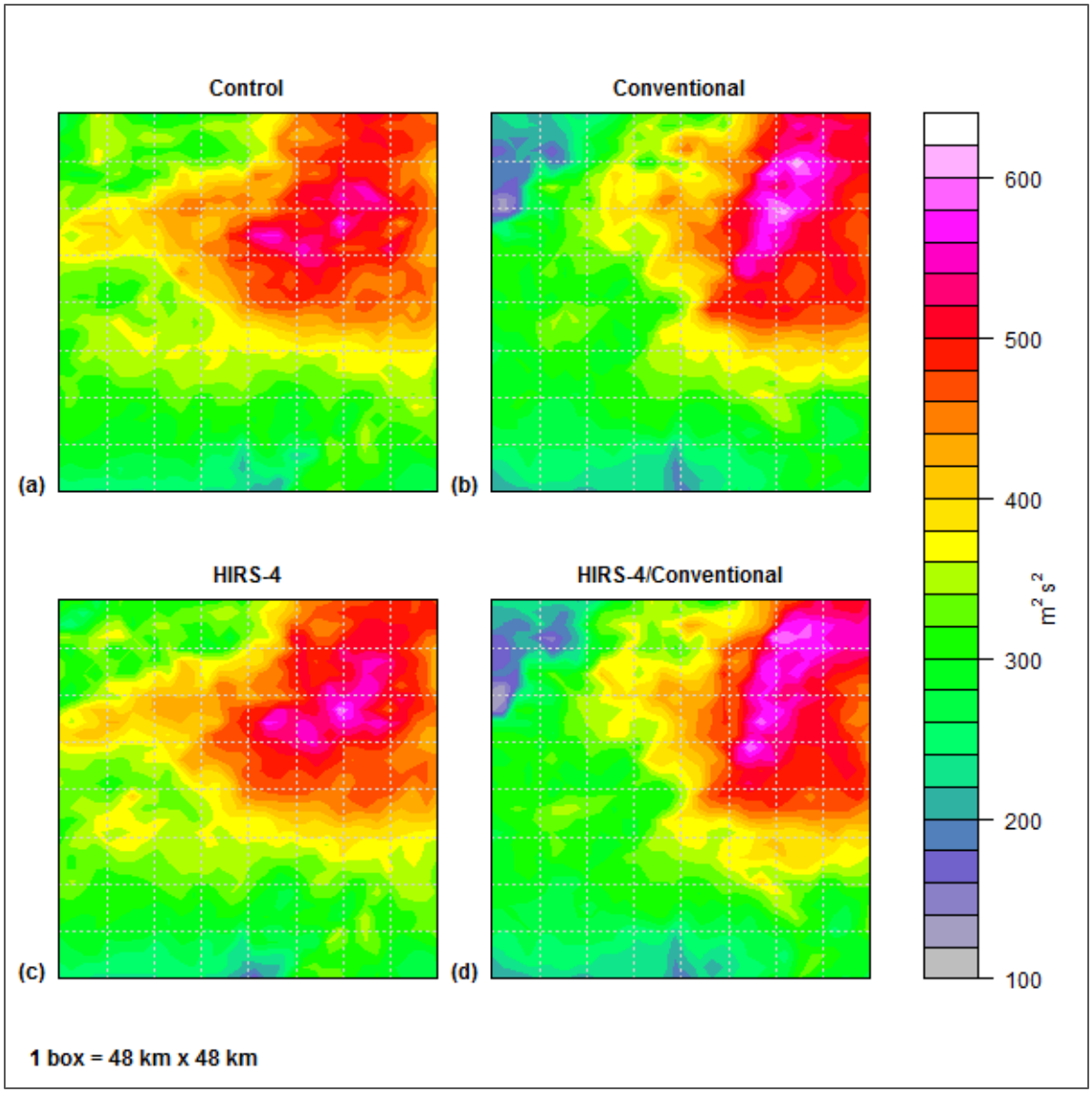

Figure 2.9 Phase I Tornadic Composites of 0-3 km SRH

2.2.1.2 Nontornadic Outbreak Composite

The 0-1 km bulk shear nontornadic composites (Fig. 2.10) exhibit a pattern similar in one way to that of the bulk shear tornado outbreak composites. Conventional (Fig. 2.10b) and HIRS-4/conventional (Fig. 2.10d) runs depict small regions of higher

bulk shear maxima than the maxima of the control (Fig. 2.10a) and HIRS-4 (Fig. 2.10c) runs. However, other than these small areas of high shear, there are not many clear differences among the nontornadic 0-1 km bulk shear composites. The composite-relative locations of high bulk shear in the nontornadic events are comparable to the locations of high bulk shear in the tornadic events, except that the small regions of maxima in the conventional and HIRS-4/conventional nontornadic composites are located northwest of the corresponding bulk shear maxima in the conventional and HIRS-4/conventional tornadic composites.

For CAPE x 0-1 km bulk shear (Fig. 2.11), a clear difference was observed between the conventional assimilation runs and the other two. The assimilation of conventional observations decreased this value for the nontornadic outbreaks, whereas such assimilation increased it for the tornado outbreaks. The HIRS-4/conventional nontornadic composite for this parameter (Fig. 2.11d) shows lower values in most of the field than the conventional composite (Fig. 2.11b). By increasing tornado outbreak values of this parameter and decreasing it for nontornadic outbreaks, conventional assimilation—alone or in combination with HIRS-4 observations—clearly helped to increase the difference between outbreak types for this combined thermodynamic-dynamic covariate.

The same overall effect was noted for 0-1 km EHI (Fig. 2.12) with respect to conventional observations. The conventional (Fig. 2.12b) and HIRS-4/conventional (Fig. 2.12d) runs exhibit significantly lower values across most of the field, and lower maxima, than the control (Fig. 2.12a) and HIRS-4 (Fig. 2.12c) runs. For this parameter, HIRS-4/conventional assimilation did introduce a small (approximately 12 km x 12 km)

maxima that is higher than the maxima of the conventional-only run, but to the northwest of this point, an area of lower EHI exists that is spatially larger than the corresponding low-EHI region in the conventional composite. As was the case with CAPE x 0-1 km bulk shear, conventional data assimilation increased this combined thermodynamic-dynamic parameter for tornado outbreaks while decreasing it for nontornadic outbreaks, and the effect was greater for nontornadic outbreaks with HIRS-4/conventional assimilation than conventional alone.

Magnitudes of 0-3 km SRH in the nontornadic composites were decreased by the assimilation of conventional data, alone or in conjunction with HIRS-4 radiances (Fig. 2.13). A slightly greater decrease, relative to the control, was observed in the HIRS-4/conventional composite (Fig. 2.13d) than in the conventional-only composite (Fig. 2.13b). Values were decreased in all regions of the composite, in contrast with the 0-3 km composites for tornado outbreaks, which exhibited increased values in the northeast sector compared to the tornadic control. The conventional and HIRS-4/conventional composites in particular resembled the conventional and HIRS-4/conventional tornadic composites in the spatial distribution of high and low values.

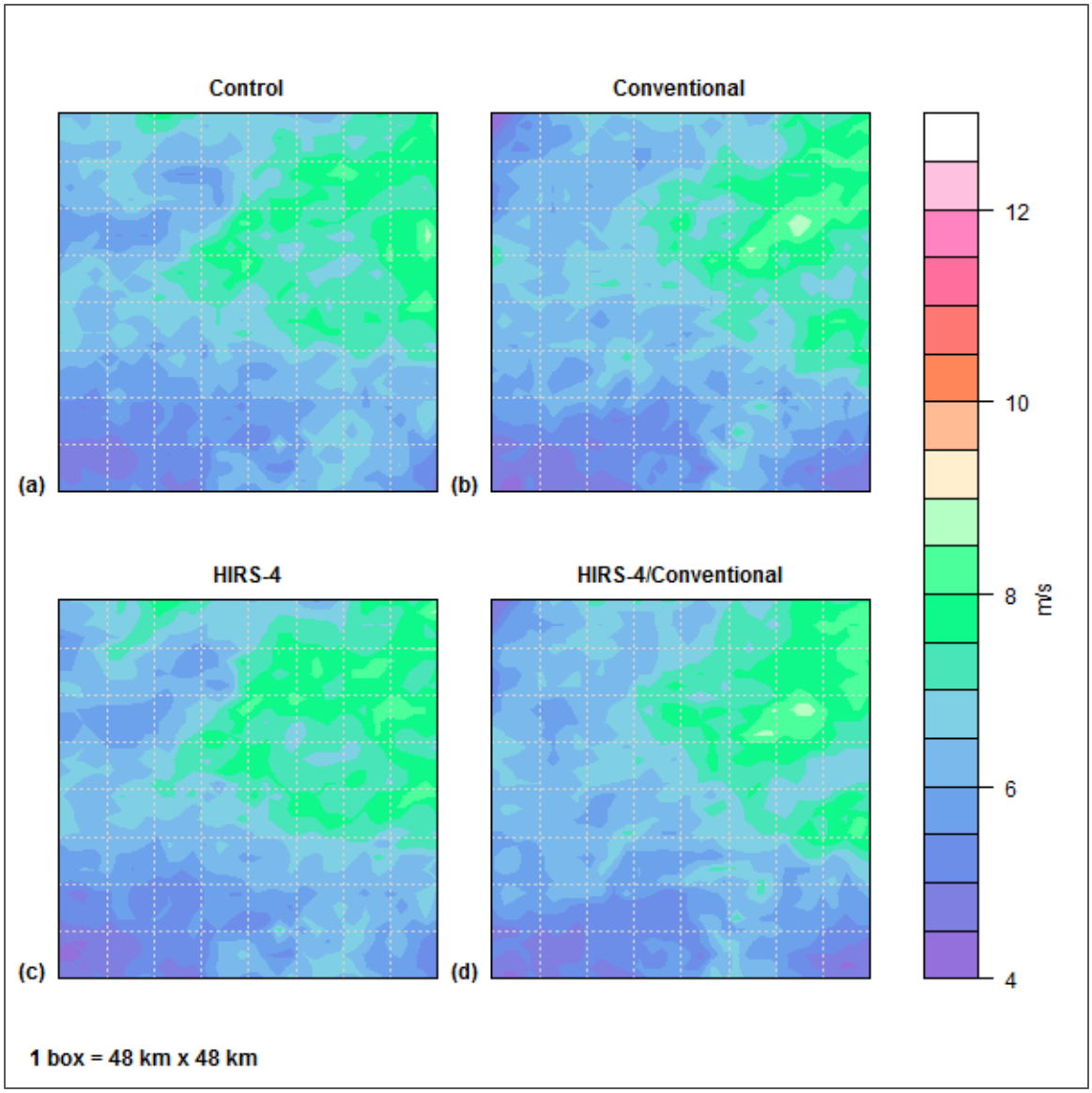

Figure 2.10 Phase I Nontornadic Composites of 0-1 km Bulk Shear

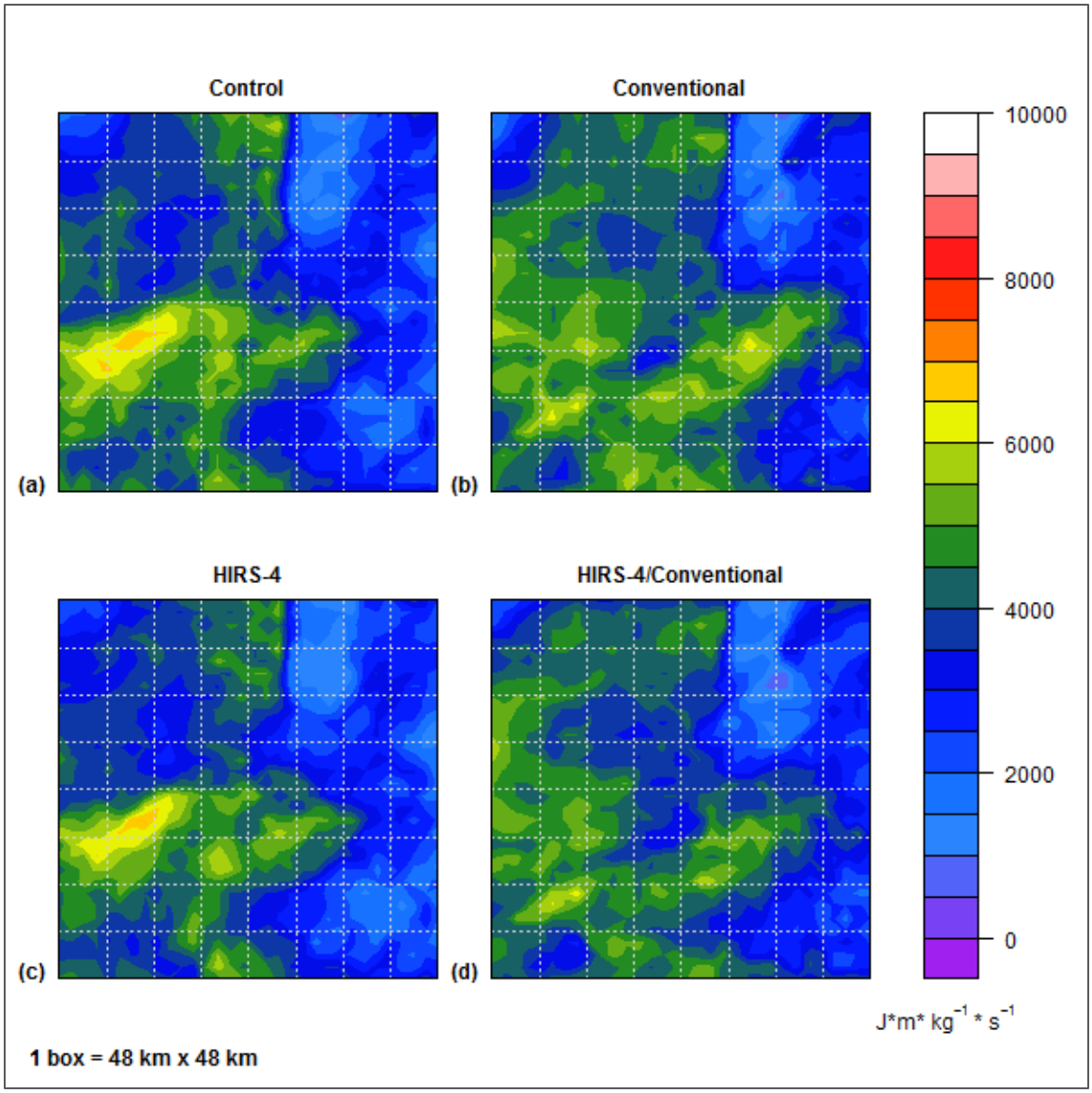

Figure 2.11 Phase I Nontornadic Composites of CAPE x 0-1 km Bulk Shear

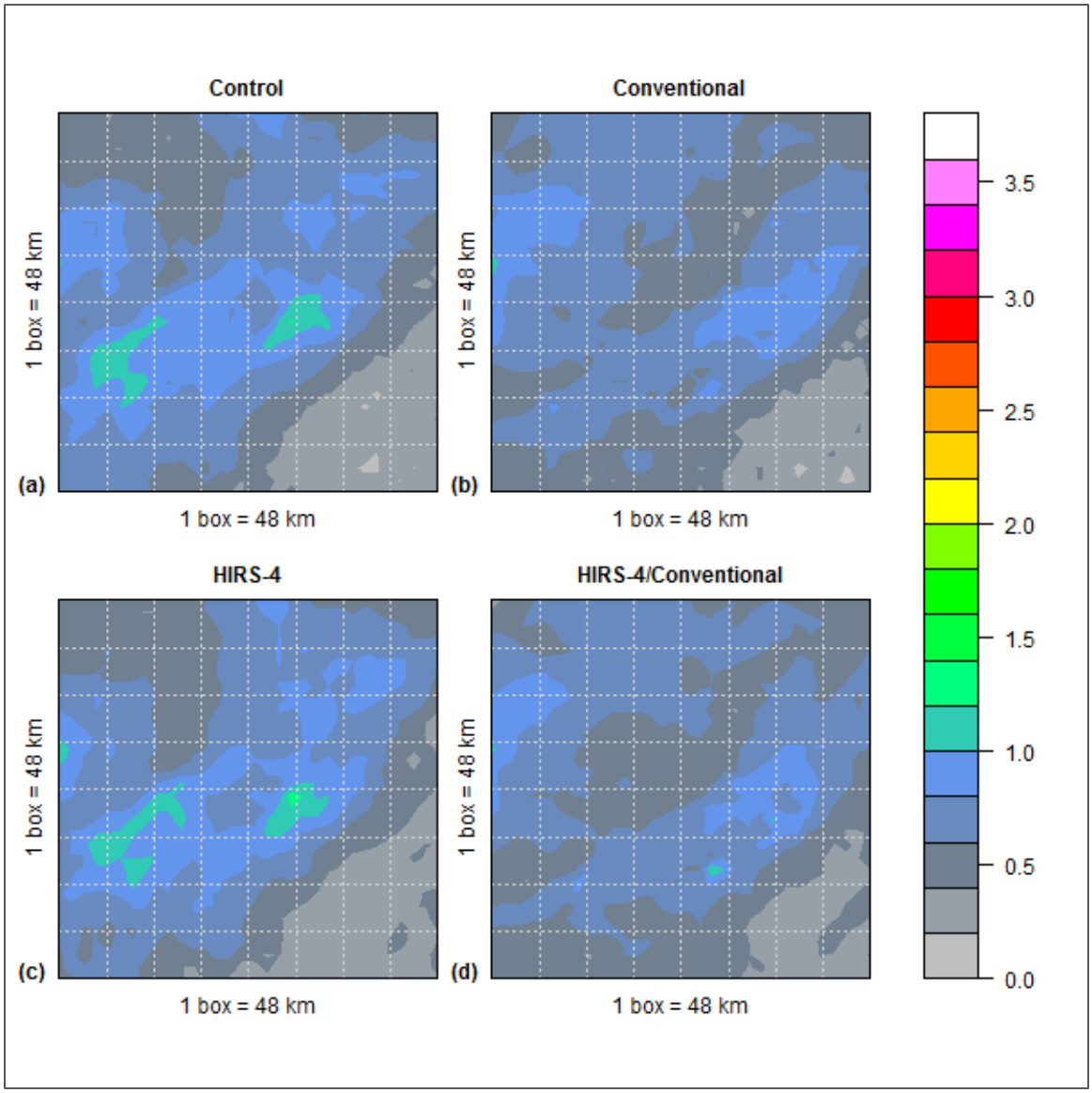

Figure 2.12 Phase I Nontornadic Composites of 0-1 km EHI

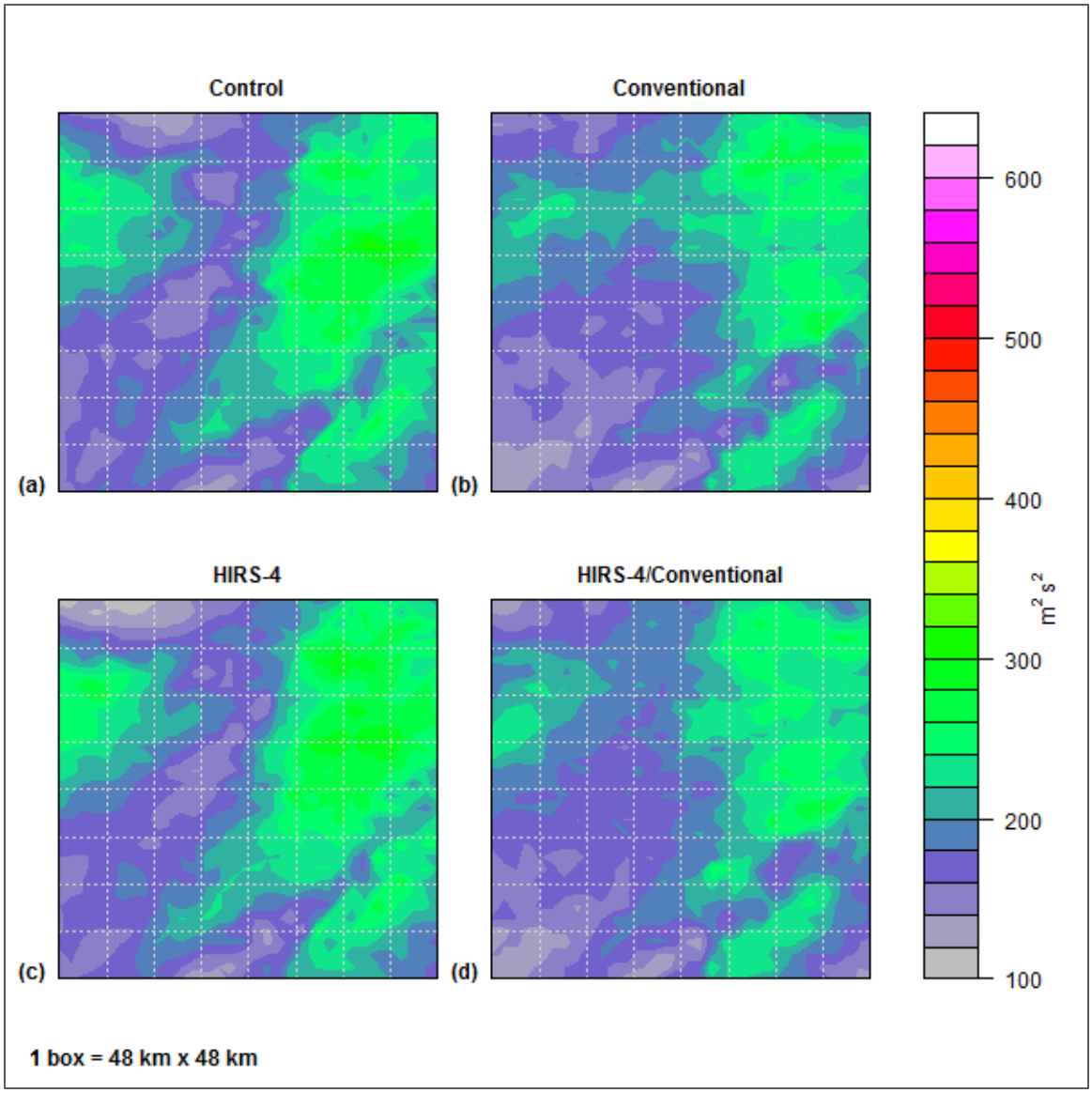

Figure 2.13 Phase I Nontornadic Composites of 0-3 km SRH

2.2.2 Euclidean Distance Calculations

Improved SVM discrimination capability should result when the statistical (Euclidean) distance between the predictor covariate fields is maximized, as this

maximizes the differences in the two classes. As such, the bootstrap Euclidean distance between the mean field for each outbreak type was formulated for each covariate.

Distance calculations are shown for all covariates and model runs in Table 2.3 with the largest distance in bold, and for select covariates in Fig. 2.14. The analysis showed that the HIRS-4/conventional run maximized the distance for 0-1 km bulk shear (Fig. 2.14a). This run and the conventional run maximized the distances for the CAPE x 0-1 km bulk shear product (Fig. 2.14b) and CIN (Fig. 2.14c). The conventional assimilation run maximized the distance for 0-1 km EHI (Fig. 2.14d). For LCL, the control generated the maximum distance between outbreak modes (Table 2.3).

The outbreak-type distances for SRH at 0-1 and 0-3 km layers were decreased by the assimilation of conventional observations. However, the HIRS-4/conventional run exhibited a slightly (and not statistically significantly) weaker decrease than the conventional data–only run (Table 2.3). Examination of 0-3 km SRH composites for tornadic and nontornadic outbreaks shows that assimilation of conventional data did decrease the magnitudes of SRH for nontornadic outbreaks while increasing the magnitudes of high SRH in tornado outbreaks (compared to the control). However, conventional data assimilation also sharply decreased the magnitudes of 0-3 km SRH in regions of the composite that did not exhibit the strongest levels of SRH. The result was that 0-3 km SRH fields for tornadic and nontornadic outbreaks exhibited similar geographical distributions of high and low values (relative to each type of outbreak) after conventional assimilation. This would decrease the outbreak-mode Euclidean distance for SRH, even though conventional data assimilation increased the highest values of this parameter for tornado outbreaks while not doing so for nontornadic outbreaks.

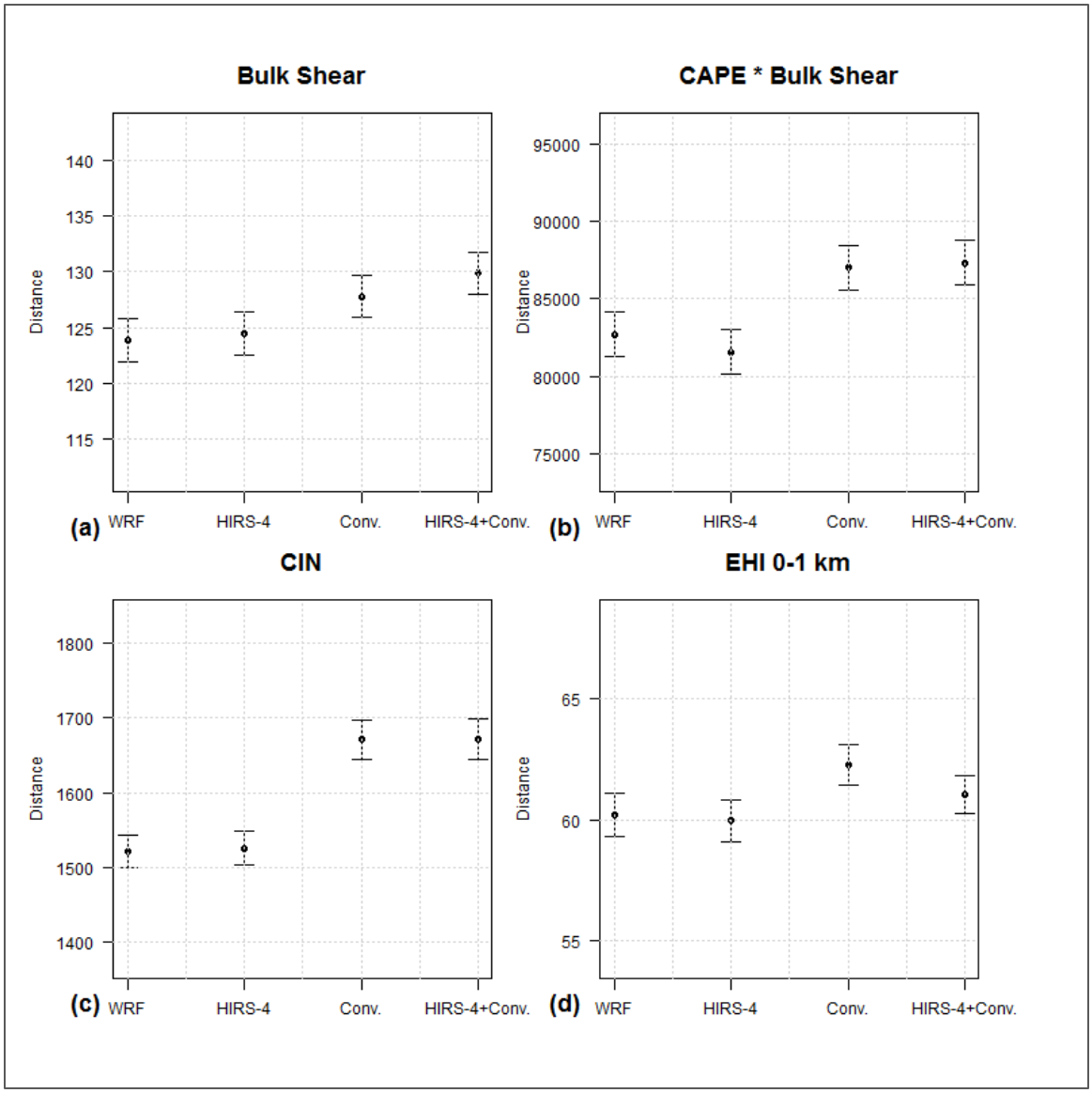

Figure 2.14 95% Confidence Intervals on Euclidean Distance Calculations Between Tornadic and Nontornadic Values of 0-1 km Bulk Shear (a), CAPE x 0-1 km Bulk Shear (b), CIN (c), and 0-1 km EHI (d) for Phase I

Table 2.3 Mean Bootstrapped Euclidean Distances Between Tornadic and Nontornadic Outbreaks for Each Covariate Type and Assimilation Mode

	0-1 km SRH	**0-3 km SRH**	**CIN**	**0-1 km Bulk Shear**	**CAPE x 0-1 km Bulk Shear**	**LCL**	**0-1 km EHI**
No assimilation	6968.69	7483.75	1521.56	123.88	82718.4	**12026.8**	60.19
HIRS-4	**7050.88**	**7504.32**	1525.71	124.43	81563.8	11583.3	59.97
Conventional	6917.80	7125.11	**1671.35**	127.79	87028.3	11784.3	**62.27**
HIRS-4/Conv.	6968.83	7158.99	1671.34	**129.90**	**87327.7**	11439.8	61.04

2.2.3 Comparison to Observed Values

A common concern in data assimilation is whether the assimilation procedure has brought the model analysis closer to an observation data set than a control run. This research did not have as its primary goal to examine this aspect of forecasting, except in the context of the determination of outbreak mode. However, to ensure that the data assimilation process employed in phase I was operating correctly and that the positive SVM results reflected a real modeling improvement, it was determined that such a comparison must be made.

As described in section 2, root mean square error (RMSE) analysis was performed on the 4 km data assimilation and control runs for several thousand data points. The difference in RMSE between each of the three data assimilation runs and the control run was also calculated. If data assimilation brought a variable closer to its observed value, the error for the data assimilation run for that point would be lower than that of the

control, and therefore the difference would be a negative value. Fig. 2.15 shows RMSE differences for all experimental runs.

The RMSE analysis indicated that for approximately 80% of examined data points, the assimilation of conventional observations did bring the model output closer to the observational data set than the no-assimilation control run, indicating that the assimilation procedure was operating as expected. The HIRS-4/conventional run exhibited similar error improvement relative to the control. The assimilation of HIRS-4 observations in the absence of conventional data did not improve the whole domain compared to an observational data set, as the figure indicates the HIRS-4-control RMSE differences to be symmetrically distributed. The assimilation of satellite radiance observations is a difficult research problem. Discernible improvements in modeling (in this context, "improvement" means that a model run is closer to an observation data set) generally require case-specific iterative fine-tuning of the assimilation parameters, which was not done in this study—as mentioned previously in section 2. Therefore, this finding is not unexpected.

Previous data assimilation studies (Wheatley et al. 2012, for example) have also examined the temporal progression of error throughout the model simulation. In the conventional and HIRS-4/conventional assimilation runs of this research, the mean error was significantly lower than the error of the control for each 12-hour period of the WRF simulation (Fig. 2.16). Model error increased over time in all runs, due to the accumulation of finite differencing approximations of the governing equations and approximations of specific conditions made in model physics parameterizations. These types of errors occur at both spatial and temporal scales, since the model grid and model

time step are discrete. However, compared to the control and HIRS-4 runs, the lines for conventional and HIRS-4/conventional model runs exhibited decreased slopes over time. This result indicates that the ingestion of additional conventional data at each time step reduced error in the simulations relative to the control run.

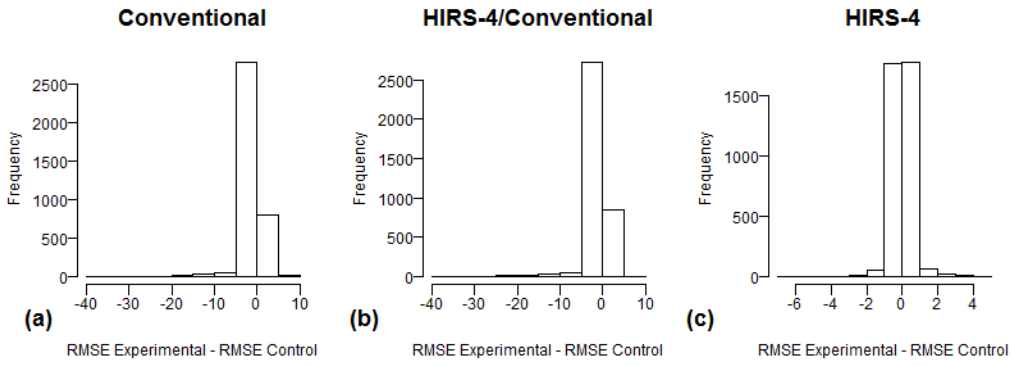

Figure 2.15 Experimental—Control RMSE Differences for Conventional (a), HIRS-4/Conventional (b), and HIRS-4 (c) Runs

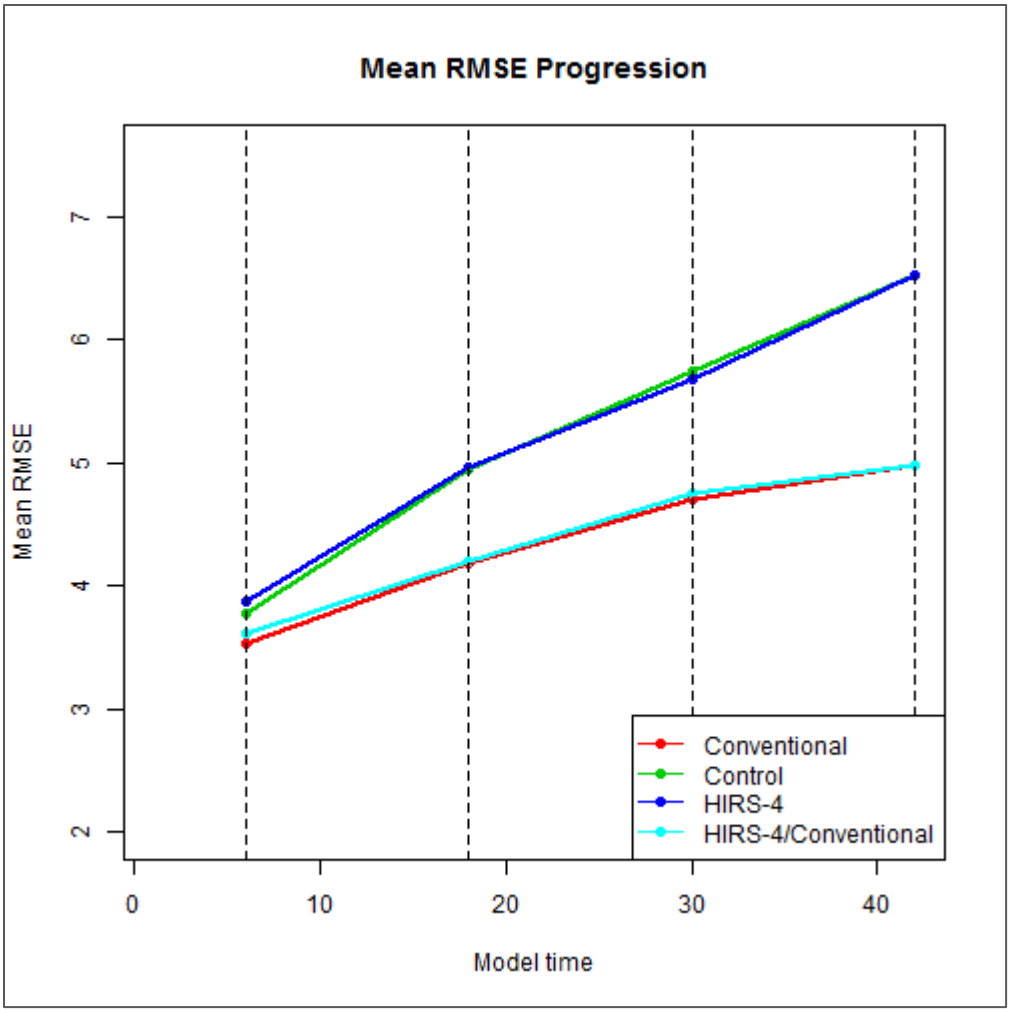

Figure 2.16 Mean RMSE for Data Assimilation Runs for Model *t*=6, 18, 30, and 42 Hours

2.3 Discussion

The most important findings of phase I are that the HIRS-4/conventional assimilation run generated the most accurate outbreak-type forecasts, and that the SVM for this model run produced its lowered FAR by increasing the difference between tornadic and nontornadic outbreak 0-1 km bulk shear and CAPE x 0-1 km bulk shear. Bootstrapped composite fields of tornadic and nontornadic outbreaks show that

assimilation of HIRS-4 and conventional observations together increased mean values of 0-1 km bulk shear in high-shear regions of the domain for both types of outbreaks. Assimilation of conventional data, alone or in combination with HIRS-4 radiances, produced a very strong effect on CAPE x 0-1 km bulk shear, increasing it for tornado outbreaks and decreasing it nontornadic outbreaks, an effect that further accentuated the differences between outbreak types. Although conventional data assimilation did not increase the Euclidean distance between outbreaks for SRH, it did have another notable effect. In the regions of high 0-3 km and 0-1 km SRH in tornado outbreaks, the magnitudes of this covariate increased even higher with data assimilation, whereas in other parts of the domain—and in the entire nontornadic outbreak domain—data assimilation decreased this covariate. Research has found that 0-1 km bulk shear is a better indicator of tornadic environments than SRH (Rasmussen and Blanchard 1998), but the consistent effects of data assimilation on both fields are useful to operational forecasting in differentiating between the types of outbreaks.

These results, in conjunction with results showing that conventional data assimilation brought modeled values closer to observations, indicate that data assimilation is correcting WRF-ARW model output that, on average, shows shear and helicity too low in tornado outbreaks in regions where it should be higher, and too high in other modeled situations. The NARR data set, which was used as the background field for the initial 12 km WRF run to which data were assimilated, is known to contain imprecision in surface wind stress (Ebisuzaki and Rutledge 2004) and diurnal inaccuracies in 2m temperature fields (Mesinger 2006). It is apparent that conventional data assimilation corrects some of these background field errors by providing additional

wind and temperature observations at the near-surface. The data assimilation procedure identifies the meteorological features present in the background field, the 12 km WRF run, and adjusts values based on the variables that it processes from the new data.

The HIRS-4 radiance set would provide an additional boost to the effect that conventional data assimilation provides by further indicating the locations of features such as clouds. Radiance brightness temperatures are especially good at identifying the location of deep convection. In a severe weather outbreak, towering cumuli and supercells would indicate the presence of local updrafts, areas of high shear, and spots of strong SRH. Although conventional data include some observations at high altitudes, the spatial coverage of satellite data is much larger and denser than the conventional data network at such elevations. In essence, assimilating both types of data provides observational checks to the model output at both near-surface and high-altitude regions of the atmosphere.

CHAPTER III

PHASE II—OUTBREAK PREDICTION WITH A PHYSICS PARAMETERIZATION ENSEMBLE

3.1 Methodology

The second phase of the study sought to identify the importance of model physics in outbreak type discrimination. This was accomplished through the variation of model microphysics and planetary boundary layer (PBL) physics in the WRF-ARW model simulations. Convective (cumulus) parameterization was not examined experimentally in this research. Although convective physics parameterization can certainly influence the modeling of environments associated with severe weather, it is typically employed only for coarser model resolutions to parameterize physical processes within convective clouds that occur on the mesoscale or microscale level. The parameterizations in a convective physics scheme do not scale well to model grids that are finer than approximately 10 km (Kain 2004). This research analyzes WRF output from 4 km outbreak-centered subdomains and a 12 km parent domain. A convective parameterization scheme was employed for the 12 km domain, but since the data analyzed experimentally were computed from the 4 km domains (with explicit modeling of convection), convective physics was not varied for the 12 km domains.

Five microphysics schemes were selected for examination: the Goddard single-moment, Morrison double-moment, Thompson double-moment, WRF double-moment 6-

class (WDM-6), and WRF single-moment 6-class (WSM-6) microphysics. Three PBL schemes were chosen: the Yonsei University (YSU), Mellor-Yamada-Janjić (MYJ), and refined Asymmetric Convection Model (ACM2) schemes. These microphysics and PBL physics schemes were chosen for their suitability for mesoscale convective weather modeling and frequent occurrence in previous research of the effects of microphysics on cloud process modeling. The PBL scheme selection was also influenced by the differing performance results found by other researchers in specific boundary layer case studies. It is important to note that this phase, however, does not seek to compare modeled values to an observation data set, but rather, to quantify the effect of parameterization scheme choice on a severe weather outbreak mode forecast. A brief overview of the microphysics and PBL physics schemes used in this research follows.

3.1.1 Microphysics Parameterization Schemes

The Goddard microphysics scheme (Tao et al. 1989) was designed to handle ice (as snow, cloud ice, and graupel) and vapor processes within convective clouds. It is specifically formulated to calculate condensation levels required to remove supersaturation, or evaporation levels required to remove areas of subsaturation. This scheme is based on the Goddard Cumulus Ensemble, a cloud-resolving model that has been used in many tropical and midlatitude convective case studies (Molthan et al. 2010, Lang et al. 2010, Lang et al. 2014, others). The WRF parameterization scheme is a single-moment scheme, meaning that it models only the mixing ratio of each category of water particle.

The Morrison microphysics scheme is a two-moment scheme developed to model the mixing ratios and number concentrations of five classes (droplets, ice, snow, rain, and

graupel) in convective clouds. It has been used extensively in severe convective weather modeling (e.g., Morrison and Milbrandt 2011, Hastings and Richardson 2016, others). The double-moment classes in Morrison microphysics are rain, cloud ice, snow, and graupel. The Thompson microphysics scheme (Thompson et al. 2008) was developed for high-resolution modeling of cloud ice, graupel, and snow, and it has also been used extensively in severe weather research (Tapiador et al. 2012, Wheatley et al. 2012, Clark et al. 2013, others). This scheme is double-moment, with number concentrations calculated for rain and cloud ice.

The WRF single-moment 6-class (Hong and Lim 2006) and WRF double-moment 6-class (Lim and Hong 2010) model water vapor, cloud droplets, ice, snow, rain, and graupel categories of cloud water. The WDM-6 scheme implements double-moment cloud droplets and rain. These schemes, or earlier versions of them, have been used in severe convective weather modeling as well (e.g., M09, Adams-Selin et al. 2013). It was desired to include the Milbrandt-Yau seven-class scheme (Milbrandt and Yau 2005)—which models hail, in addition to the other six classes—for examination as well, but the very high computational requirements of this scheme made this infeasible. This research thus included two single-moment and three double-moment microphysics schemes.

Studies exist comparing single- and double-moment microphysics schemes, though rarely in the context of severe thunderstorms or outbreaks. As one example, Molthan and Colle (2012) examined several microphysics schemes—including Morrison, Thompson, and WSM-6—in the simulation of a synoptic-scale snow event caused by a mid-latitude cyclone. They found that Thompson and Morrison microphysics represented water vapor and saturation best of the schemes, compared to observational

data. These schemes also modeled the amount of snow better than WSM-6, which modeled much of the snow as cloud ice instead.

In another study, Jankov et al. (2010) examined the effects of microphysics schemes on the modeling of an atmospheric river over the Pacific. They considered, among others, the WSM-6, Thompson, and Morrison schemes, finding that these three schemes modeled moisture in the atmospheric river and cloud infrared brightness temperatures comparably to each other. The Thompson and Morrison physics exhibited higher skill at forecasting brightness temperatures than WSM-6, but WSM-6 and Morrison modeled banding patterns in the clouds better than other schemes, including Thompson, which modeled a "smooth" cloud top. The Thompson scheme exhibited a slight warm bias compared to observational data. In the context of an atmospheric river, brightness temperatures can reflect the depth and density of clouds.

Adams-Selin et al. (2013) examined the simulation of a bow echo with several microphysics parameterization schemes, including WSM-6, WDM-6, and Morrison. They found that the Morrison scheme did not produce more accumulated precipitation than single-moment microphysics schemes, in contrast with previous studies involving this scheme. The reason for this result is that in the bow echo case, Morrison microphysics produced smaller graupel pellets, which led to higher rates of melting and evaporation. Adams-Selin et al. also found that WDM-6 microphysics produced an area of convection much larger than that produced by WSM-6, but that this did not result in a larger area of precipitation, due to increased evaporation from small graupel sizes. The WDM-6 scheme produced much more intense convection than Morrison microphysics, as well as a very strong cold pool.

The Jankov et al. (2012) and Molthan and Colle (2012) studies suggest that WSM-6 and Morrison microphysics may model severe thunderstorms especially intensely. These storms contain large amounts of cloud ice and often have overshooting tops indicative of powerful thermodynamically driven updrafts. Additionally, the Adams-Selin et al. (2013) study suggests that WDM-6 could model strong convection and cold pools well due to its handling of cloud graupel processes. However, it should be noted that the Adams-Selin et al. study employed MYJ PBL physics in all of its model runs, whereas this research examines two additional PBL schemes as well as MYJ.

3.1.2 PBL Physics Parameterization Schemes

The MYJ boundary layer scheme (Janjić 1994) is a local PBL scheme used in the operational Eta model. It employs a turbulent kinetic energy model to measure turbulence in the boundary layer. Janjić (1994) developed this scheme to address modeling of heavy spurious maritime precipitation in the Eta model that was found to be generated from excessive vertical turbulent heat and moisture flux and observed that spurious precipitation was reduced relative to other NWP forecasts. Later work with the MYJ scheme involving land events found that the local closure of this scheme resulted in its production of a cold, moist bias in summertime thermodynamically driven PBL development (Hu et al. 2010), and the underdevelopment of a dry convective boundary layer (Gibbs et al. 2011). The scheme does not perform vertical PBL mixing of nonlocal plumes or eddies. This characteristic is significant for severe convective weather modeling because updrafts, downdrafts, inflow jets, and rising air parcels all cross vertical layers of the PBL.

The revised Asymmetric Convection Model (ACM2) scheme (Pleim 2007) is a nonlocal PBL scheme that employs first-order eddy diffusion in unstable PBL environments and uses local turbulence closure in stable PBLs. ACM2 attempts to address a known limitation of local-closure PBL schemes such as MYJ, namely that these schemes assume that turbulent eddies must occur at a scale smaller than the vertical grid spacing. ACM2 was developed to perform mixing at both the local (one vertical layer) and nonlocal (across vertical layers) scale (Pleim 2007). In the Hu et al. (2010) study, ACM2 produced heat and moisture flux values in a humid summertime environment similar to those produced by YSU, a result that they attributed to the nonlocal mixing capability of this scheme. In the Gibbs et al. (2011) study, ACM2 modeled heat flux values of a dry PBL the least accurately of the three PBL schemes tested. The study also found, however, that ACM2 modeled potential temperature in a dry cold front case closely to observations.

The Yonsei University PBL scheme (YSU) is a nonlocal general-purpose PBL scheme. Hong et al. (2006) developed the scheme to resolve problems with vertical boundary layer mixing in the Medium-Range Forecast (MRF) PBL scheme (Hong and Pan 1996). The MRF scheme produced excessive mixing in PBL environments with high wind shear and too little mixing in primarily thermodynamically driven convective environments. Hong et al. (2006) performed a case study of 10 November 2002, a frontally driven tornado outbreak, to test the YSU against the MRF scheme, using the WRF model. They found that YSU produced a more intense frontal boundary and stronger convection for this tornado outbreak than MRF, and that the physical cause of this difference was higher prefrontal CAPE in the YSU simulation due to a shallower but

moister PBL. Prefrontal convection was also diminished with YSU, which more closely matched observations. As intended, the scheme minimized boundary layer mixing across a frontal boundary in high-shear environments associated with katabatic fronts. As described in section 1 of this document, YSU was also found by Hu et al. (2010) to increase vertical mixing in a non-frontal summertime PBL environment in Texas. Gibbs et al. (2011) further found that YSU produced the most accurate simulations of PBL heat flux in a dry boundary layer in Oklahoma, in both dry cold frontal and dryline conditions.

3.1.3 WRF Model Configuration

This phase, in common with all three phases of this study, employed the WRF model (version 3.4.1). The WRF runs for phase II were initialized at 1800 UTC the day preceding a given outbreak (to allow for model spin-up) to 1200 UTC the day after the outbreak (a 42 hour simulation). A two-way nested configuration was used, and phase II domains were as close as possible to those of phase I in dimensionality and geographical location. For phase II, the outer domain was a 12 km grid identical in spatial extent and geographical dimensions to the background fields of phase I (Fig. 2.1). The inner domains were 4 km grids that very closely matched the 4 km outbreak domains generated in phase I after data assimilation (Fig. 2.4). Some of these inner domains were not strictly identical in size or geographical location to their phase I counterparts, due to WRF model requirements of the positioning of nested domains relative to the parent domain. However, these differences were no larger than two grid points (8 km) in any direction.

The vertical dimension encompassed 28 vertical levels with a model top of 100 hPa. Model physics that were not varied in the simulations are shown in Table 3.1, while

the variants chosen for study are shown in Table 3.2. In the 12 km parent domains, Kain-Fritsch cumulus physics were used, but cumulus physics were disabled for the 4 km nests. The Kain-Fritsch cumulus scheme is commonly used in convective event modeling (e.g., Wheatley et al. 2012).

Table 3.1 Phase II WRF Model Physics Schemes Used for All Runs

WRF physics option	Configuration	Reference
Longwave radiation	Rapid Radiative Transfer Model	Mlawer et al. 1997
Shortwave radiation	Dudhia	Dudhia 1989
Surface layer	*For ACM2 and YSU PBL:* MM5-derived	Dudhia 1996
	For MYJ PBL: Monin-Obukhov/Eta similarity	Janjić 2002
Land surface	Noah land surface model	Tewari et al. 2004
Urban surface	None	
Cumulus physics	*For 12 km nest:* Kain-Fritsch	Kain 2004
	For 4 km nest: None	

Table 3.2 Phase II Cloud Microphysics and PBL Physics Variations

Physics Option	Reference
Goddard single-moment microphysics	Tao et al. 1989
Morrison double-moment microphysics	Morrison et al. 2009
Thompson double-moment microphysics	Thompson et al. 2008
WRF Double-Moment 6-class microphysics	Lim and Hong 2010
WRF Single-Moment 6-class microphysics	Hong and Lim 2006
Yonsei University PBL physics	Hong et al. 2006
Mellor-Yamada-Janjić PBL physics	Janjić 1994
Asymmetric Convection Model PBL physics	Pleim 2007

3.1.4 SVM Classification of Outbreak Mode

As in phase I, support vector machines were used to discriminate between outbreak modes. The data were prepared for the SVM in an identical manner to that of phase I. The same seven important severe weather covariates—0-1 km SRH, 0-3 km SRH, 0-1 km bulk shear, 0-1 km EHI, CIN, CAPE x 0-1 km bulk shear, and LCL height—were calculated on the 4 km domains for each of the outbreaks at outbreak valid time (section 1.4.2). These values were interpolated to 12 km resolution by data thinning, to minimize the effects of data noise generated by WRF at isolated locations at high horizontal resolutions. A 32 x 32 grid was extracted from this interpolated data set, and an S-mode rotated principal component analysis was conducted on each ensemble member (each microphysics and PBL combination, a total of 15) to generate PCA scores. These scores were then input to the SVMs.

In this phase, identically to phase I, the following kernel functions were considered:

Linear: $K(\mathbf{x},\mathbf{y}) = \mathbf{x}^{\mathrm{T}}\mathbf{y}$ (3.1)

Polynomial: $K(\mathbf{x},\mathbf{y}) = (\gamma\mathbf{x}^{\mathrm{T}}\mathbf{y})^{d}$ (3.2)

Radial Basis (Gaussian): $K(\mathbf{x},\mathbf{y}) = \exp(-\gamma|\mathbf{x}-\mathbf{y}|^{2})$ (3.3)

where **x** and **y** are vectors in original linear space, γ is a user-defined Gaussian spread parameter associated with the polynomial and radial basis kernels, and *d* is the degree of the polynomial (the user dictates this value). In this study, we considered all three kernel functions and multiple degree *d* values (2, 3, 4, 0.5, and 0.333), yielding a total of seven kernel functions considered. Additionally, cost function values ranging from 1 (no penalty to points far from the margin) to 10000 (severe penalty to points far from the margin) on a $\log_{10}$ scale, and γ values of 0.01, 0.05, and 0.1 were considered. This led to a total of 95 kernel-cost-gamma combinations (the linear kernel is not a function of γ, so only five linear kernel experiments were done for the five cost values).

To identify the best kernel-cost-gamma combination of the 95 tested, a bootstrap 2-fold cross-validation routine was implemented (Efron and Tibshirani 1993), withholding 20 of the 40 events randomly for each bootstrap replicate. Each kernel-cost-gamma combination was tested with the same 1000 random samples of training and testing sets to ensure pairwise comparisons were possible.

Determination of forecast skill was made by calculating contingency statistics on the forecasts of the SVMs. Probability of detection (POD), false alarm ratio (FAR), bias, and Heidke skill score (HSS) were computed (Wilks 2011). Bootstrap intervals (1000

iterations of the bootstrap procedure) were calculated for the contingency statistics, yielding results for the most skilled SVM. The highest-skilled SVM for each model run is shown in Table 3.3.

Table 3.3 SVM with Highest Median Heidke Skill Score for Parameterization Runs

WRF Run	SVM
Goddard/ACM2	Linear, cost=100
Goddard/MYJ	Linear, cost=10,000
Goddard/YSU	Linear, cost=1000
Morrison/ACM2	Radial basis, cost=1000, γ=0.01
Morrison/MYJ	Radial basis, cost=10,000, γ=0.01
Morrison/YSU	Linear, cost=10,000
Thompson/ACM2	Radial basis, cost=10,000, γ=0.01
Thompson/MYJ	Radial basis, cost=1000, γ=0.05
Thompson/YSU	Radial basis, cost=10,000, γ=0.05
WDM-6/ACM2	Linear, cost=10,000
WDM-6/MYJ	Radial basis, c=1000, γ=0.01
WDM-6/YSU	Linear, cost=1000
WSM-6/ACM2	Linear, cost=1000
WSM-6/MYJ	Linear, cost=10,000
WSM-6/YSU	Linear, cost=100

3.1.5 Euclidean Distance Calculations and Outbreak Composites

As in phase I, bootstrapped Euclidean distances between tornadic and nontornadic outbreaks were calculated for the mean fields of each covariate for each of the 15 parameterization ensemble members. Outbreak mean bootstrapped composites for the 32x32 grid were also generated for the 15 ensemble members for tornadic and nontornadic outbreak types, using the same methodology as in phase I.

Since one of the goals of the research was to quantify certain sources of model uncertainty in severe weather outbreak forecasting, it was desirable to determine the

effect of PBL physics parameterization and cloud microphysics parameterization individually on each covariate. However, microphysics or PBL physics could produce such a strong effect on covariate values that the influence of the other type of model physics could be difficult to discern, rendering statistical techniques such as *k*-means cluster analysis (Wilks 2011) unsuitable for determining the effect. Instead, fields of each covariate were averaged across all ensemble members that had been run with a given microphysics or PBL physics option. For example, an average of 0-1 km SRH was generated from the Goddard/ACM2, Goddard/MYJ, and Goddard/YSU runs for the outbreak domain at valid time for each of the 40 outbreaks, producing a "Goddard microphysics" mean of each outbreak. Outbreak averages were conducted in this manner for each of the microphysics options and PBL options for every covariate, for a total of 40 mean values per covariate per microphysics or PBL physics scheme.

3.2 Results

After the SVMs were run for the phase II physics parameterization ensemble, the best SVM output from the cross-validation step of the kernel-cost-gamma SVM optimization was compared among all fifteen WRF simulations using contingency statistics (probability of detection—POD, false alarm ratio—FAR, bias, Heidke skill score—HSS; Wilks 2011) to establish effects from physics parameterization. Contingency statistics for the SVMs (Fig. 3.1) indicated that the use of the WRF double-moment 6-class microphysics scheme (WDM-6) and YSU PBL physics produced the highest skill of any microphysics-PBL physics combination, generating an HSS of 0.658 (Fig. 3.1c). This score was significantly (at the 95% confidence level) greater than the HSS of any other physics combination examined. This run produced the lowest FAR

(Fig. 3.1b; tied with Morrison/YSU and WSM-6/ YSU) and highest POD (Fig. 3.1a; tied with Goddard/ YSU). The Goddard/ YSU, Morrison/Yonsei, and WSM-6/ YSU runs also produced the highest HSS after the WDM-6/Yonsei run.

Examination of the contingency scores of the same microphysics or PBL scheme across the model runs revealed other interesting patterns. PBL physics generated greater differences in skill than microphysics. YSU PBL physics had the highest skill of the three PBL physics schemes with every microphysics option examined, and its scores were highly significant for every microphysics option except Thompson. MYJ PBL physics generally performed poorest of the three, with the exception of the Morrison microphysics, for which the ACM2 PBL physics produced the lowest skill at modeling outbreaks.

PBL physics had a strong impact on FAR. The YSU PBL scheme produced the lowest FAR of the three examined PBL physics parameterizations for every microphysics option (Fig. 3.1b). Except for the Morrison/ACM2 run, ACM2 PBL physics produced lower FAR than MYJ PBL physics for every microphysics, though the difference was not significant for the Goddard runs. No discernible pattern of FAR was apparent with microphysics variation, except that the Goddard runs were more tightly clustered than any other set of microphysics runs.

POD (Fig. 3.1a) scores also exhibited patterns more strongly linked with PBL physics than with microphysics. YSU PBL physics yielded significantly highest POD for Goddard, Morrison, and WDM-6. The bootstrap means were also highest of the three PBL schemes for Thompson and WSM-6, though these statistics were not significantly higher. Most runs exhibited a clear over-prediction bias (Fig. 3.1d). The lowest-skilled

Morrison/ACM2 run was close to zero bias at its mean, and zero bias was within the confidence intervals of Thompson/MYJ and WSM-6/ YSU.

The sensitivity of microphysics to PBL physics variation, and PBL physics to microphysics variation, was also examined. The magnitude of microphysics sensitivity to PBL varied with each microphysics option. The Morrison and WDM-6 microphysics runs contained a greater range of HSS among their PBL options than did the Goddard, Thompson, and WSM-6 microphysics runs. Morrison microphysics proved to be most sensitive to PBL physics variation, with a difference of 0.133 between the means of the highest- and lowest-skilled Morrison runs. WDM-6 microphysics exhibited a range of 0.117 between the highest- and lowest-skilled runs. In comparison, Goddard microphysics runs had a range of 0.058, Thompson runs had a range of 0.069, and WSM-6 runs had a range of 0.046. This sharp difference between the highly sensitive microphysics and the other three is attributable to a poorly performing MYJ run for WDM-6, a poorly performing ACM2 run for Morrison, and skilled YSU runs for both.

The magnitude of PBL physics sensitivity to microphysics did not vary nearly as much. ACM2 PBL physics proved most sensitive to microphysics variations, with a range of 0.087 HSS between its highest-skilled (WDM-6) and lowest-skilled (Morrison) runs. MYJ PBL physics was the least sensitive to microphysics variation, showing a range of 0.0514 HSS between its highest-skilled (Goddard) and lowest (Thompson) runs (Fig. 3.1c). YSU PBL physics had a range of 0.0539 HSS between its highest-skilled (WDM-6) and lowest-skilled (Thompson) runs.

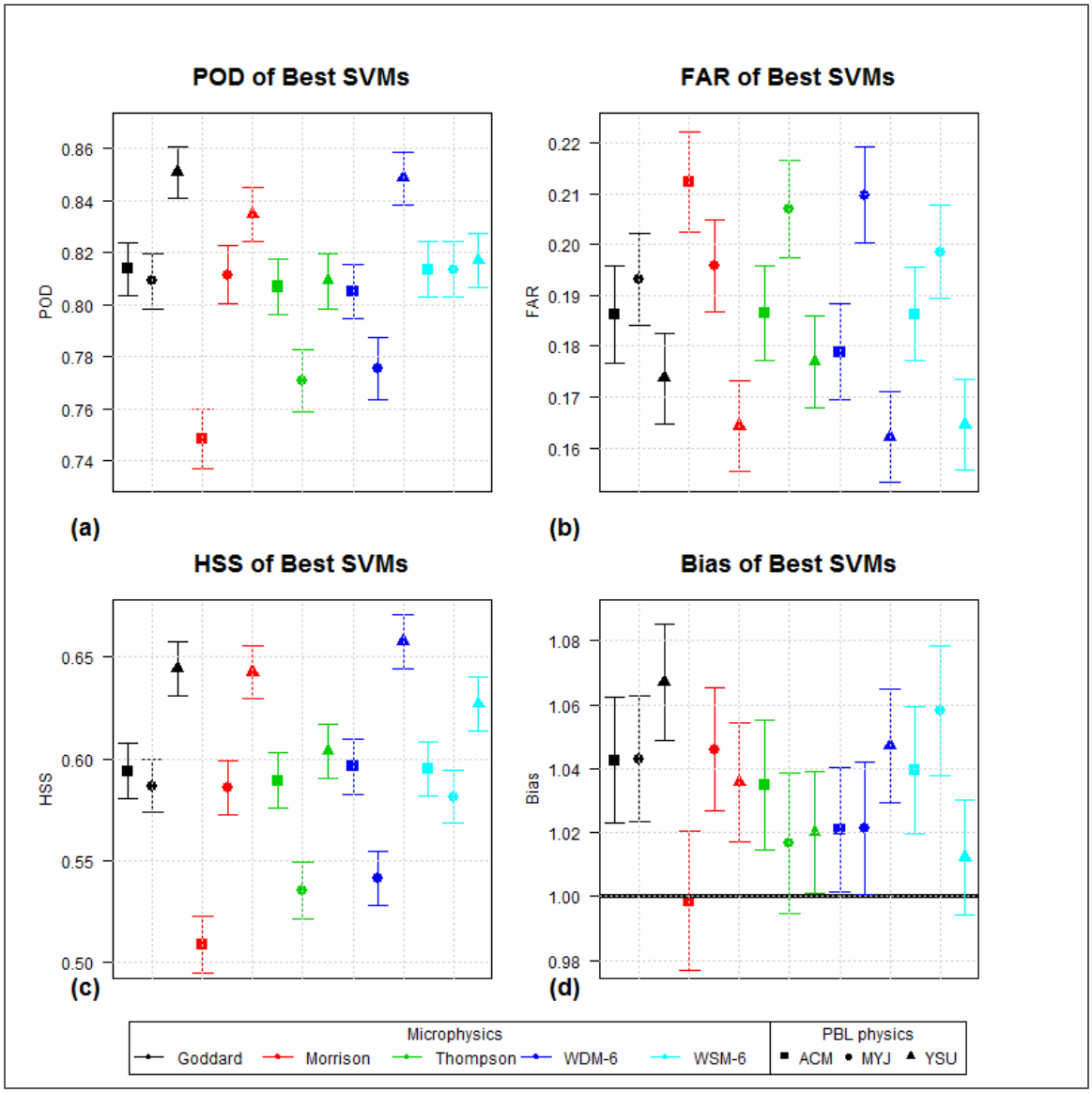

Figure 3.1 95% Confidence Intervals for POD (a), FAR (b), Heidke Skill Score (c), and Bias (d) of the Best-Performing SVMs for Phase II

3.2.1 Covariate Mean Analysis

As described previously, average values of each covariate were generated for each outbreak event for all ensemble members that had been run with a given microphysics or PBL physics scheme. The goal of this analysis was to determine which covariates, if any,

exhibited large variations in magnitude among microphysics or PBL physics runs. These mean values were examined in box and whisker plots, one diagram for each model physics scheme (five microphysics and three PBL), to compare data ranges and extreme values (Fig. 3.2). Tornadic and nontornadic outbreaks were analyzed separately. The analysis found that, for both tornadic and nontornadic outbreaks, greater variability in the mean fields existed among PBL runs for all covariates.

Figure 3.2 depicts box and whisker plots for 0-1 km bulk shear, 0-1 km SRH, and LCL height means for tornadic and nontornadic cases. As is evident in Fig. 3.2a and 3.2b, LCL height varied significantly for PBL physics runs and much less so for microphysics runs. Similar patterns appear in tornadic and nontornadic outbreaks for this covariate. ACM2 PBL physics generated higher LCL values than the other two PBL physics, and MYJ generated much lower LCL heights, especially in the nontornadic outbreaks (Fig. 3.2b). The outlier values for MYJ were also lower than the outliers for other runs. Variability did exist among microphysics for this covariate, with the Thompson physics scheme generating larger LCL height means in the third quartile (Q3)—and the fourth quartile (Q4) as well for nontornadic outbreaks—than the other microphysics options. This indicates that this scheme produced greater PBL mixing and vertical growth. More variability existed in LCL height means for the nontornadic outbreaks than the tornadic outbreaks for the microphysics runs, with most of this variability occurring in the Q3 and Q4.

WDM-6, the microphysics scheme used in the ensemble member that had the most skill at forecasting outbreak mode, had its quartiles distributed fairly symmetrically for this covariate for tornado outbreaks. Notably, its Q1 and Q4 had larger ranges than

those of the other microphysics. For the nontornadic outbreaks, WDM-6 had a smaller upper whisker than the other microphysics, indicating that the maximum LCL heights in Q4 were lower than those of the other microphysics runs.

Tornado outbreak 0-1 km SRH and 0-1 km bulk shear means exhibited some features in common on the box plots. A long upper whisker was apparent for both fields in all model physics, indicating few—but high—upper-end values in Q4 for these covariates. Thompson microphysics appeared to increase 0-1 SRH values for all quartiles, while increasing 0-1 km bulk shear values noticeably more for Q1 and Q2 (data points representing domain-averaged less intense events for this covariate) than Q3 or Q4. However, PBL physics generated more variability for both covariates for tornado outbreaks than microphysics. Close examination of Fig. 3.2c and Fig. 3.2e reveals that the PBL physics variability is entirely due to the stark difference between the plots for ACM2 and MYJ; YSU PBL physics generates box plots largely similar to those generated by the five microphysics options (with the exception of having higher values in Q4 for both covariates). ACM2 physics generates shear and helicity values that are lower than any other physics option in every quartile. MYJ physics has the highest Q1, Q2, and Q3 mean values of these covariates for tornado outbreaks, but as noted, its highest Q4 values are not as high as the YSU values or the Morrison and Thompson Q4 values.

Nontornadic outbreak 0-1 km SRH and 0-1 km bulk shear fields exhibited some patterns in common with tornado outbreak fields, while other patterns were distinct to this category of outbreak. Once again the MYJ physics modeled shear and helicity values higher than the other model physics, and ACM2 physics modeled 0-1 km SRH low overall. Interestingly, however, YSU physics modeled 0-1 km bulk shear Q4 upper

values significantly lower than any other model physics, including ACM2. Among the microphysics options, 0-1 km SRH means followed a similar pattern to 0-1 km SRH in tornado outbreaks for microphysics. For both kinds of outbreaks, Thompson physics produced a higher Q3 maximum. Goddard and WSM-6 microphysics produced lower Q3 maxima than the other microphysics.

An interesting pattern was apparent with YSU PBL physics in the shear and helicity covariate means. The Q4 maximum for this scheme for tornadic 0-1 km SRH and bulk shear is higher than the Q4 for any other physics scheme. In the nontornadic outbreaks, this scheme produced outlier values of both covariates that are higher than the outliers for any other physics scheme, but not Q4 values. For 0-1 km bulk shear, the YSU Q4 was, as noted, lower than that of any other physics scheme. It should be recalled that these box plots describe 40 data points apiece, each point representing the average of the entire outbreak domain at valid time, averaged again over three or five model runs (depending on whether the plot is for a microphysics or PBL physics scheme average). Apparently, YSU modeled a few nontornadic outbreaks with high shear and helicity values, but this occurred infrequently enough that these averages appear as outliers. This may have aided the SVM in discrimination.

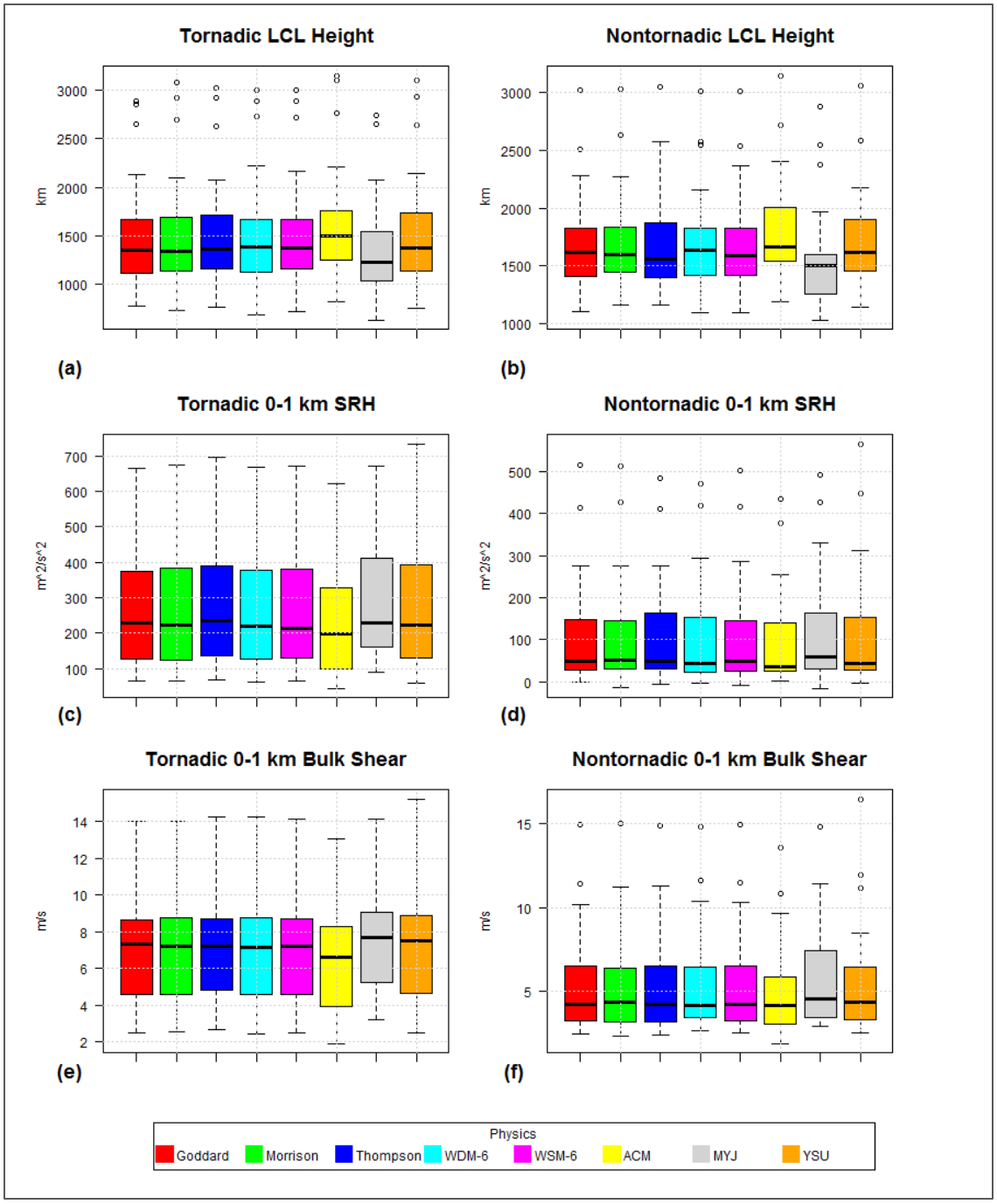

Figure 3.2 Box Plots of LCL Height, 0-1 km SRH, and 0-1 km Bulk Shear Means

3.2.2 Tornadic and Nontornadic Covariate Composites

To determine the specific physical impact of microphysics and PBL physics parameterization on each type of outbreak, an average composite of LCL, 0-1 km SRH, and 0-1 km bulk shear for each outbreak type was formulated for the WDM-6/ YSU, WSM-6/ YSU, Thompson/ACM2, Goddard/ACM2, Morrison/MYJ, and WDM-6/MYJ runs. These runs were selected to include the highest-skilled (WDM-6/MYJ) and lowest-skilled (Morrison/MYJ) variants and a range of all other microphysics and PBL options. These covariates were chosen in order to include the one (LCL) whose distance was maximized by the most-skilled WDM-6/ YSU run and two covariates that were dynamically driven to balance this thermodynamic covariate. The composites were generated across the 32 x 32 (12 km grid) outbreak-centered field that was input to the SVM. The contrasts in these composites help demonstrate the exact impacts of model physics parameterization on the simulations.

3.2.2.1 Tornadic Outbreak Composite

LCL height in tornado outbreaks exhibited a similar pattern among all parameterization runs, although differences did exist. As was expected, PBL physics generated greater differences in the spatial characteristics of covariate fields than did microphysics. As Fig. 3.3 shows, runs that were modeled with MYJ PBL physics (Fig. 3.3c and 3.3f) produced lower minima for this covariate, and larger areas thereof, than the other two PBL physics. However, MYJ runs also exhibited a more diffuse gradient between low and high LCL than runs modeled with YSU (Fig. 3.3a and 3.3d) and ACM2 (Fig. 3.3b and 3.3e) PBL physics.

Since the composites represent the bootstrapped means of the approximate outbreak centers at outbreak peak times, this gradient likely represents a temperature or moisture boundary. Two outbreaks featured a dryline passage, and four cases involved the passage of quasi-linear mesoscale convective systems (MCS) in the absence of a synoptic cold front in the region. In an MCS, storm outflow creates a cold pool in the wake of the storms, producing a mesoscale thermal-moisture boundary between the line of storms and the pool. In the remaining 14 tornado outbreaks, a synoptic cold front was present, but for approximately half of these cases, the frontal passage over the center of the outbreak domain occurred approximately three to six hours after outbreak peak time. A quasi-linear convective system (QLCS) or broken line of thunderstorm cells traversed the domain at outbreak valid time. These observations show that the feature that appears in the LCL height composites for tornado outbreaks is likely an averaged mesoscale cold front associated with cold pool growth from storm passage. This indicates that fronts were modeled more intensely with YSU and ACM2 PBL physics even at the mesoscale.

Of these two PBL physics, YSU generated lower LCL height minima than ACM2. Lower LCL heights are strongly associated with tornado development in conducive environments, since a lower LCL height indicates a less mixed, moister PBL, as well as a shorter distance over which vertical rotation between the surface and the cloud base must remain intact. This effect was more pronounced for the single-moment microphysics parameterization WSM-6 than for the double-moment WDM-6, with the WSM-6 runs having lower LCL heights than the WDM-6 runs in the prefrontal region. However, WSM-6 also generated lower LCL heights than WDM-6 west of the boundary in the LCL composites, which coincides with the findings of Adams-Selin et al. (2013) in

which WDM-6 generated a very strong cold pool in the wake of a bow echo. Cloud icing and graupel processes could modify the PBL in cases where strong prefrontal convection developed, but the WDM-6 does not use a second moment for frozen hydrometeors, so this factor alone cannot explain the difference between WDM-6 and WSM-6. Adams-Selin et al. (2013) also found, however, that WDM-6 modeled smaller graupel pellets than WSM-6 in their study, and this resulted in greater evaporative cooling. The WDM-6 does model rain and cloud droplets with a second moment, and this may have resulted in more evaporative cooling in general in this research due to the tendency of double-moment schemes to model smaller hydrometeors than single-moment schemes.

For 0-1 km SRH (Fig. 3.4) and 0-1 km bulk shear (Fig. 3.5), pronounced differences are once more apparent among the PBL physics parameterization runs. MYJ PBL physics generated the highest values of 0-1 km SRH (Fig. 3.4c, Fig. 3.4f) and 0-1 km bulk shear (Fig. 3.5c, Fig. 3.5f). However, the physics scheme mean analysis (section 3.2.1) found that MYJ tornadic runs, on average, increased SRH and bulk shear in the less strongly sheared outbreak domains, whereas YSU produced (in its casewide averages) higher values than MYJ in highly sheared outbreak domains. This result is not readily apparent in the composites in Fig. 3.4 and 3.5, but these composites use gridpoint bootstrap averages, whereas the analysis in section 3.2.1 averaged the covariate across the entire domain for each case.

Of the two MYJ runs examined in the composites, MYJ with Morrison microphysics displayed higher values of 0-1 km SRH and 0-1 km bulk shear than MYJ with WDM-6. YSU PBL physics generated the second-highest values of 0-1 km SRH (Fig. 3.4a, Fig. 3.4d) and 0-1 km bulk shear (Fig. 3.5a, Fig. 3.5d). Not much difference

was noted between the fields of WDM-6/YSU (Fig. 3.4a, Fig. 3.5a) and WSM-6/YSU (Fig. 3.4d, Fig. 3.5d) for these covariates, though WDM-6/ YSU did produce a slightly larger area of high values of each covariate than WSM-6/ YSU. For these two covariates, ACM2 PBL physics generated fields with the lowest values (0-1 km SRH—Fig. 3.4b, Fig. 3.4e; 0-1 km bulk shear—Fig. 3.5b, Fig. 3.5e).

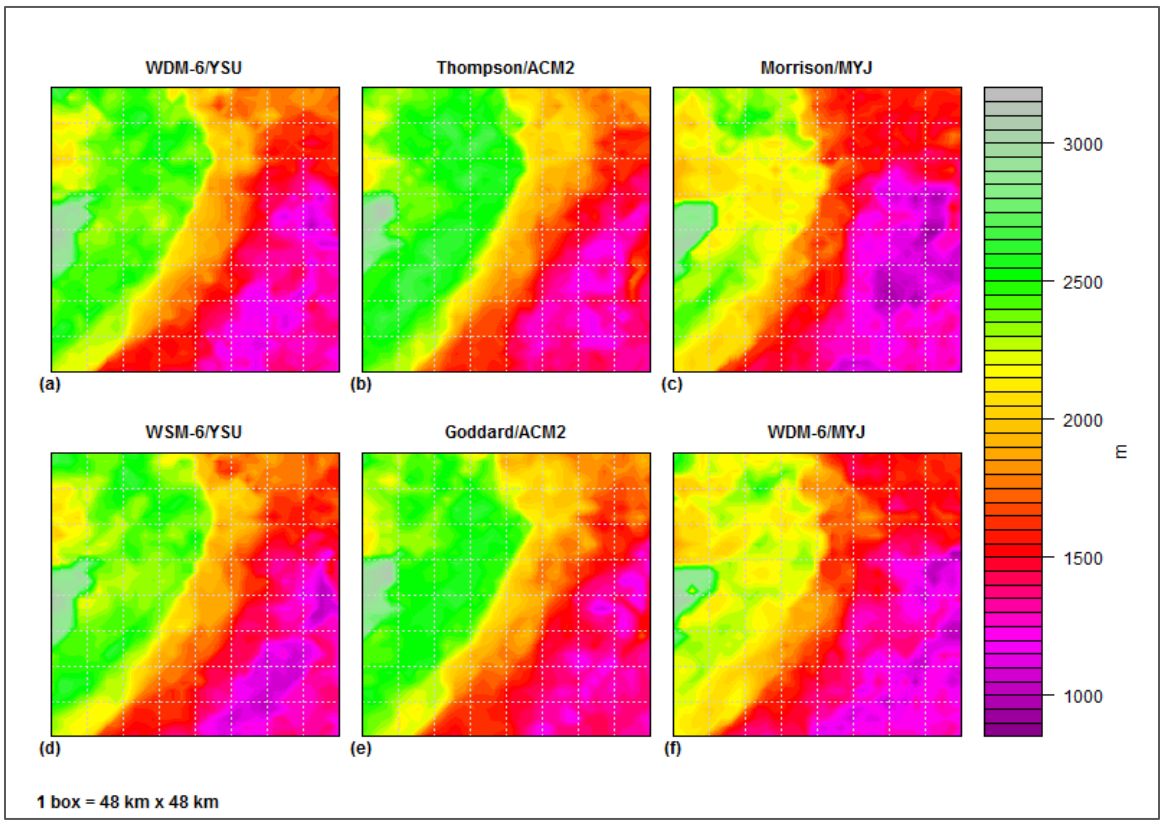

Figure 3.3 Phase II Tornadic Composites of LCL Height

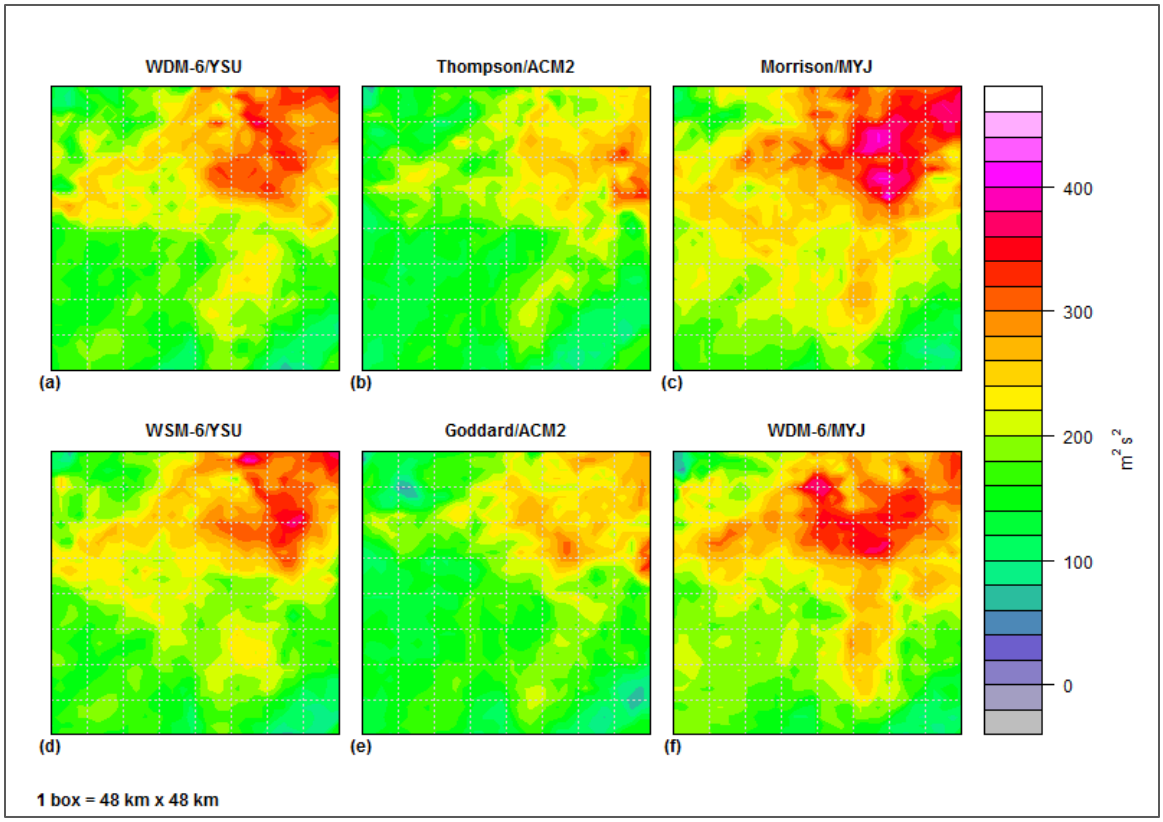

Figure 3.4 Phase II Tornadic Composites of 0-1 km SRH

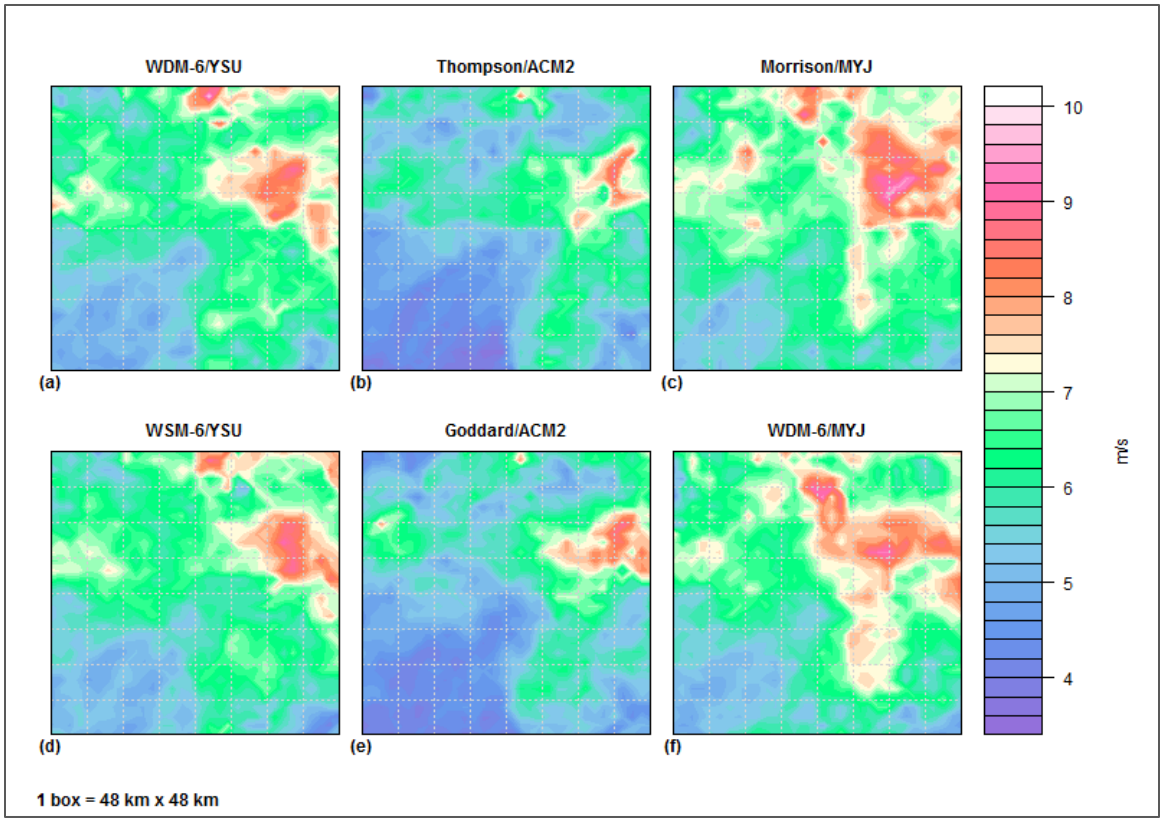

Figure 3.5 Phase II Tornadic Composites of 0-1 km Bulk Shear

3.2.2.2 Nontornadic Outbreak Composite

For LCL height, a PBL physics-based pattern was apparent in the nontornadic model runs. As was the case for the tornado outbreak composites of this covariate, MYJ PBL physics produced a larger region of low LCL heights than the other two PBL physics (Fig. 3.6). The overall spatial distribution and visual appearance of the covariate means was similar across all the ensemble members. However, the strong difference in the modeling of apparent frontal boundaries that was present in the tornadic LCL composites was not present in the nontornadic composites. Research conducted in phase III of this study implies that the likely reason for this was that many nontornadic

outbreaks were not associated with a katabatic cold front, deep trough, or mid-latitude cyclone, whereas all the tornado outbreaks were. Furthermore, examination of radar and surface observations for the nontornadic cases—as was conducted for the tornadic cases—found a difference in the orientation of organized storm systems between the two types of outbreaks. The tornado outbreaks usually had a line of storms oriented southwest-northeast and which traveled generally west to east. The nontornadic outbreaks contained storm clusters of varying shapes which more often had a greater north to south direction in their movements.

SRH from 0-1 km (Fig. 3.7) and 0-1 km bulk shear (Fig. 3.8) for the nontornadic outbreaks also displayed a strong PBL physics-based distinction. As was the case with the tornado outbreaks, ACM2 PBL physics decreased values of 0-1 km SRH (Fig. 3.7b, Fig. 3.7e) and 0-1 km bulk shear (Fig. 3.8b, Fig. 3.8e) relative to the other two PBL physics parameterizations. The role of microphysics in 0-1 km SRH and 0-1 km bulk shear nontornadic outbreak variations was very difficult to discern, with no consistent patterns appearing between single-moment and double-moment schemes.

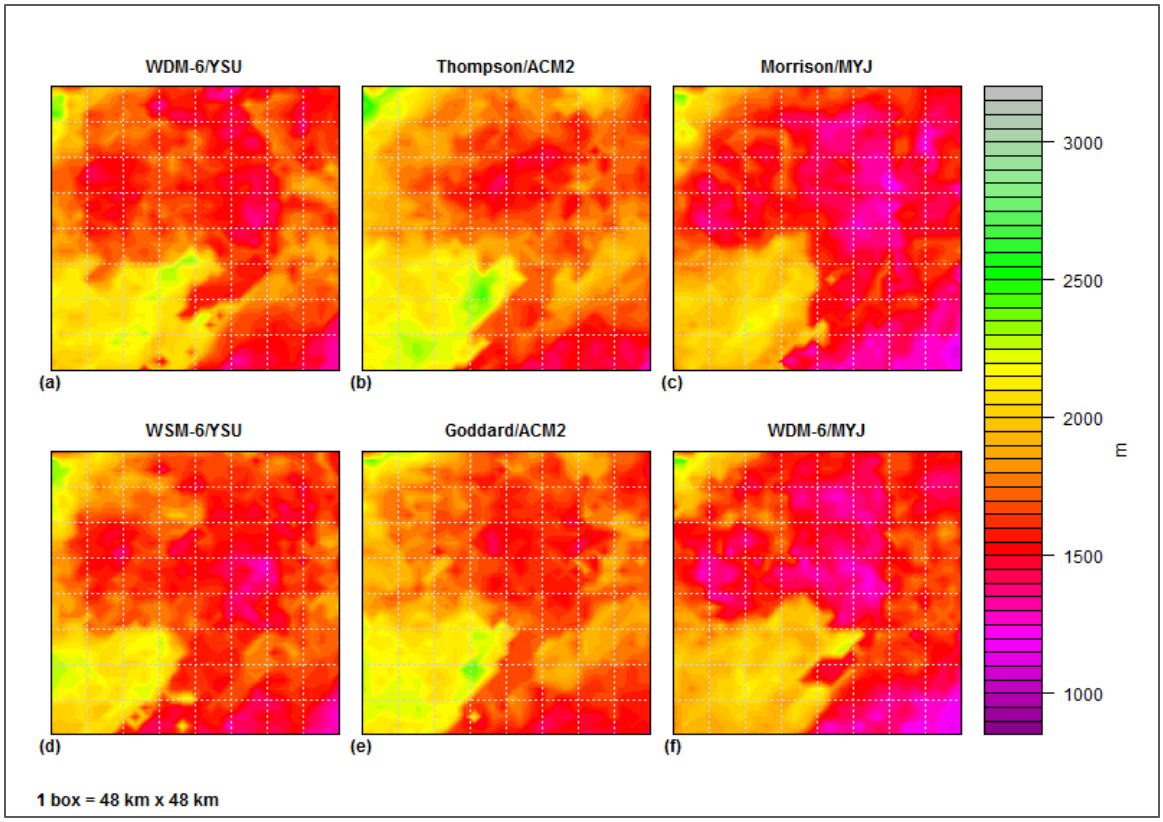

Figure 3.6 Phase II Nontornadic Composites of LCL Height

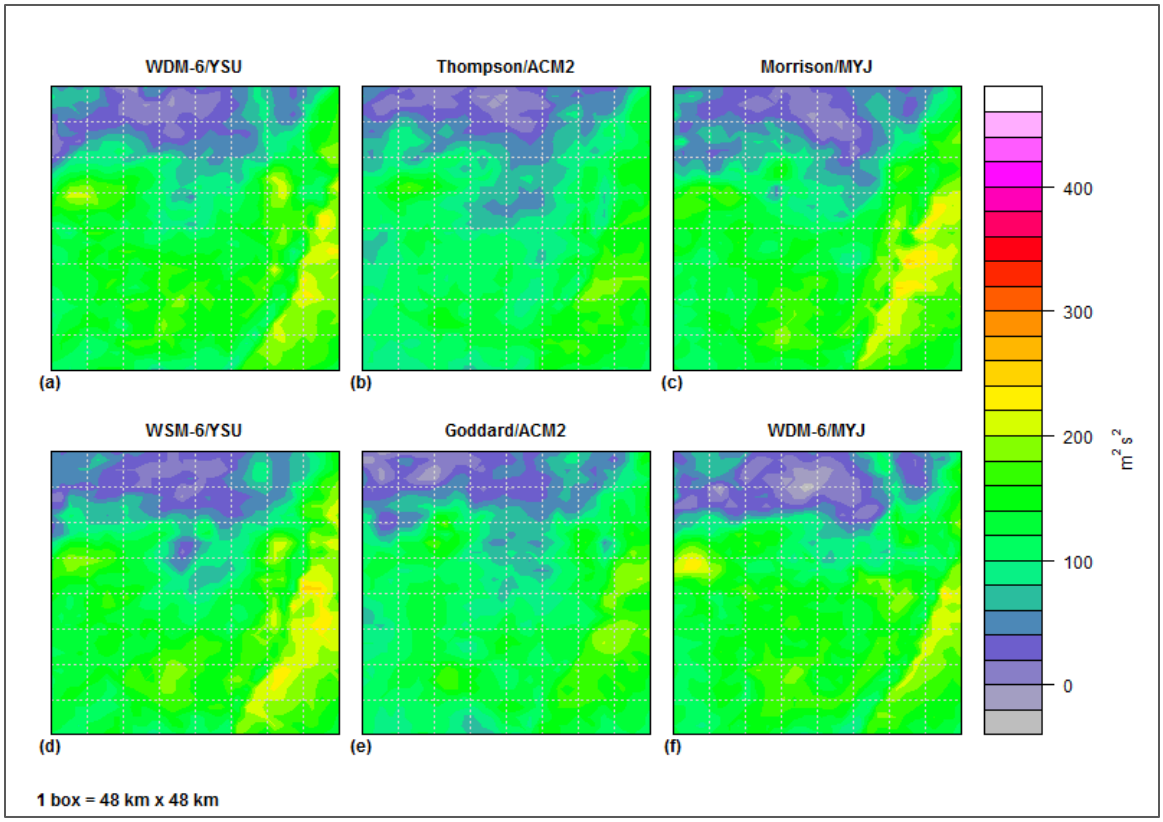

Figure 3.7 Phase II Nontornadic Composites of 0-1 km SRH

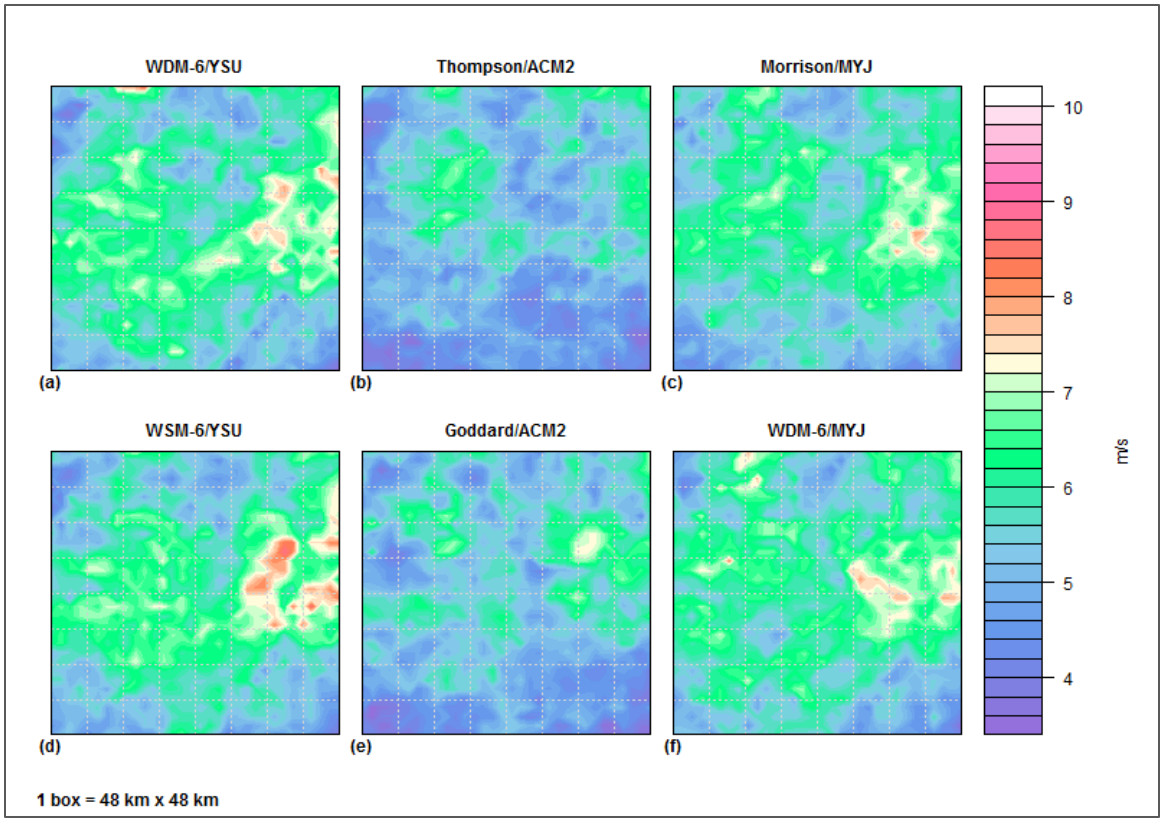

Figure 3.8 Phase II Nontornadic Composites of 0-1 km Bulk Shear

3.2.3 Euclidean Distance Calculations

As in phase I, it was important to identify the impact of microphysics and PBL physics parameterization choices on the average magnitudes of these variables. Improved SVM discrimination capability should result when the statistical (Euclidean) distance between the predictor covariate fields is maximized, as this emphasizes the differences in the two classes. The bootstrap Euclidean distance between the mean field for each outbreak type was formulated for each covariate.

Table 3.4 shows all covariates and ensemble runs for this phase, with the largest distance for each covariate in bold. Clear patterns emerged from this analysis regarding

PBL scheme effects on modeled environment. The ACM2 PBL physics scheme produced the minimum distances of the three PBL schemes for every covariate that included shear or helicity: 0-1 km SRH (Table 3.4), 0-3 km SRH (Fig. 3.9a), 0-1 km bulk shear (Fig. 3.9b), CAPE * 0-1 km bulk shear (Fig. 3.10a), and 0-1 km EHI (Fig. 3.10b). This effect occurred with every microphysics scheme examined, including the Thompson physics. It can be inferred from this analysis that ACM2 does not model differences in low-level shear between tornadic and nontornadic environments as strongly as the other two PBL schemes studied, at least in conjunction with the WRF configuration employed for this study. The MYJ PBL physics produced consistently greater outbreak mode Euclidean distances than the YSU PBL scheme for 0-1 km bulk shear and 0-1 km SRH. This effect was decreased for 0-3 km SRH, 0-1 km EHI, and the product of CAPE and 0-1 km bulk shear. The MYJ scheme produced lesser distances between tornadic and nontornadic LCL for every microphysics, though the difference between it and YSU was not significant for the WSM-6 microphysics. The implications of the LCL results for SVM classification are discussed later in the document.

The choice of microphysics had little effect on the distance calculations for 0-1 km bulk shear, 0-1 km SRH, and 0-3 km SRH. However, thermodynamic and partially thermodynamic covariates did exhibit sensitivity to microphysics. Among each PBL scheme, the Thompson microphysics scheme maximized distances for CAPE * 0-1 km bulk shear (Fig. 3.10a), CIN (Fig. 3.11a), and 0-1 km EHI (Fig. 3.10b). It is also notable that for CIN, Thompson microphysics maximized the distances for CIN very clearly. Thompson microphysics also produced larger outbreak mode distances for 0-1 km EHI than any other microphysics when paired with MYJ and YSU PBL physics. These

results indicate that Thompson microphysics models CAPE-related differences between tornadic and nontornadic environments the most distinctly of the microphysics studied.

Interestingly, the highest-skilled WDM-6/YSU run produced the maximum distance among the fifteen total runs for LCL only, although it performed middling to well for all the covariates. Its success at modeling distinctions between outbreak mode may be attributed to generally good modeling of all of these parameters rather than performing exceptionally well with a select number of them. Alternatively, the classification algorithm for the WDM-6/YSU SVM may have heavily weighted LCL as a discriminating factor. This possibility is explored in the following section.

Table 3.4 Mean Bootstrapped Euclidean Distances Between Tornadic and Nontornadic Outbreaks for Each Covariate and Parameterization Set

	0-1 km SRH	**0-3 km SRH**	**CIN**	**0-1 km Bulk Shear**	**CAPE x 0-1 km Bulk Shear**	**LCL**	**0-1 km EHI**
Goddard/ACM2	4644.60	6701.81	1009.50	60.75	86786.2	13400.6	48.00
Goddard/MYJ	6253.82	**8015.74**	990.25	75.24	116950.0	12519.5	62.52
Goddard/YSU	5591.81	7660.86	1084.83	69.11	112756.8	13178.6	63.18
Morrison/ACM2	4772.42	6854.94	1057.83	59.44	87143.0	13291.3	51.23
Morrison/MYJ	6035.54	7731.14	1113.72	73.41	119267.5	12788.3	65.20
Morrison/YSU	5506.08	7460.78	1137.78	70.15	118545.2	13445.3	65.07
Thompson/ACM2	4822.03	6909.63	1235.73	62.07	88771.3	13387.3	52.96
Thompson/MYJ	**6401.19**	7950.04	**1359.34**	**76.00**	121096.4	12760.7	**71.92**
Thompson/YSU	5660.07	7680.35	1292.67	70.68	**122446.7**	13249.8	68.82
WDM-6/ACM2	4576.09	6539.98	1104.75	59.21	84221.4	13970.9	48.09
WDM-6/MYJ	5902.21	7484.08	1158.60	72.16	116569.4	13203.2	63.46
WDM-6/YSU	5462.90	7400.42	1157.89	67.95	111690.3	**14153.8**	61.48
WSM-6/ACM2	4626.76	6538.27	1125.88	56.86	84576.7	13603.7	49.36
WSM-6/MYJ	6078.96	7763.67	1076.36	72.76	120675.3	13156.4	65.83
WSM-6/YSU	5504.99	7443.54	1185.10	68.01	112037.6	13314.7	62.70

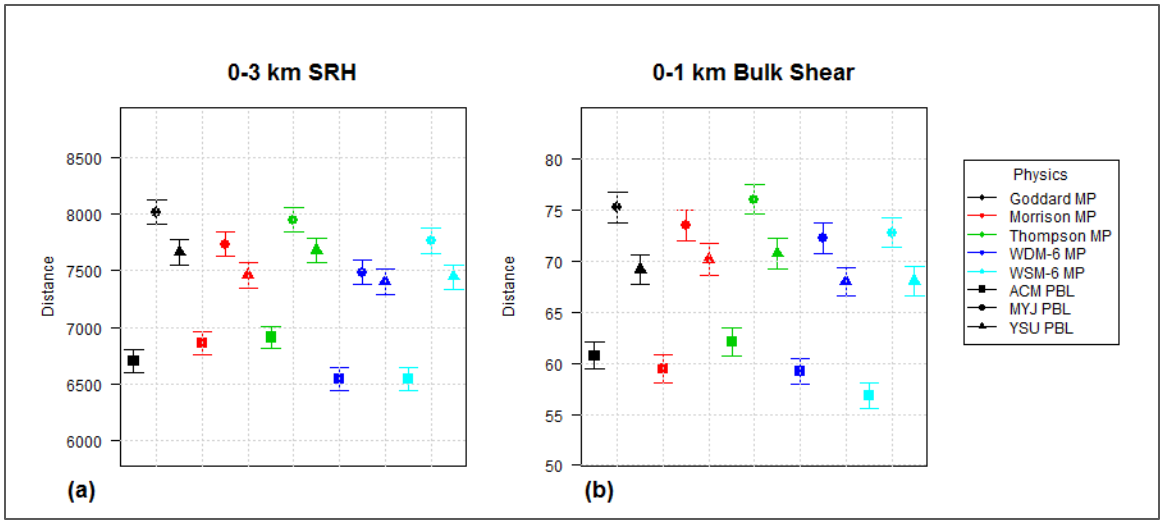

Figure 3.9 95% Confidence Intervals on Euclidean Distance Calculations Between Tornadic and Nontornadic Values of 0-3 km SRH (a) and 0-1 km Bulk Shear (b) for Phase II

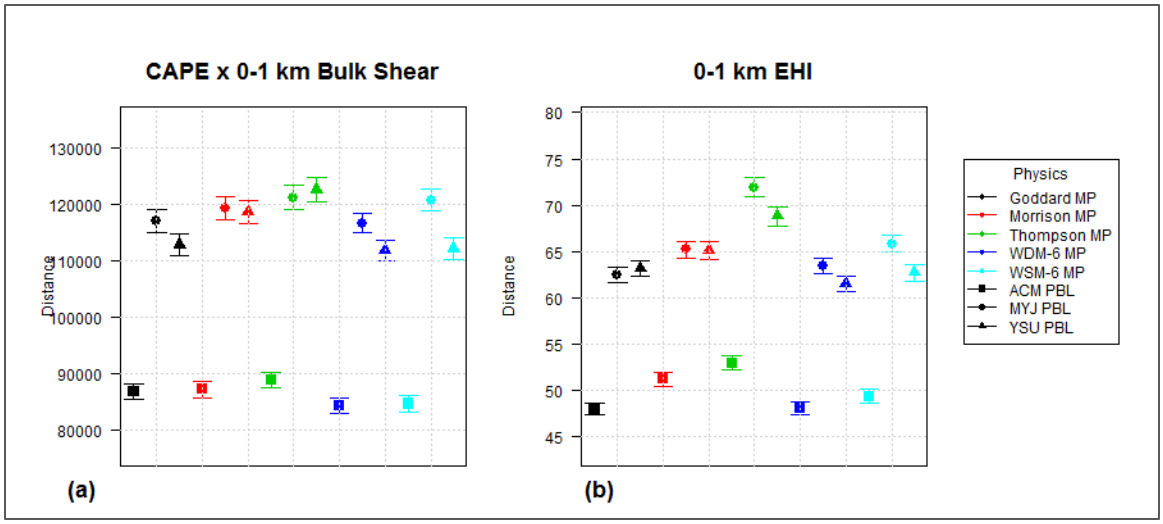

Figure 3.10 95% Confidence Intervals on Euclidean Distance Calculations Between Tornadic and Nontornadic Values of CAPE x 0-1 km Bulk Shear (a) and 0-1 km EHI (b) for Phase II

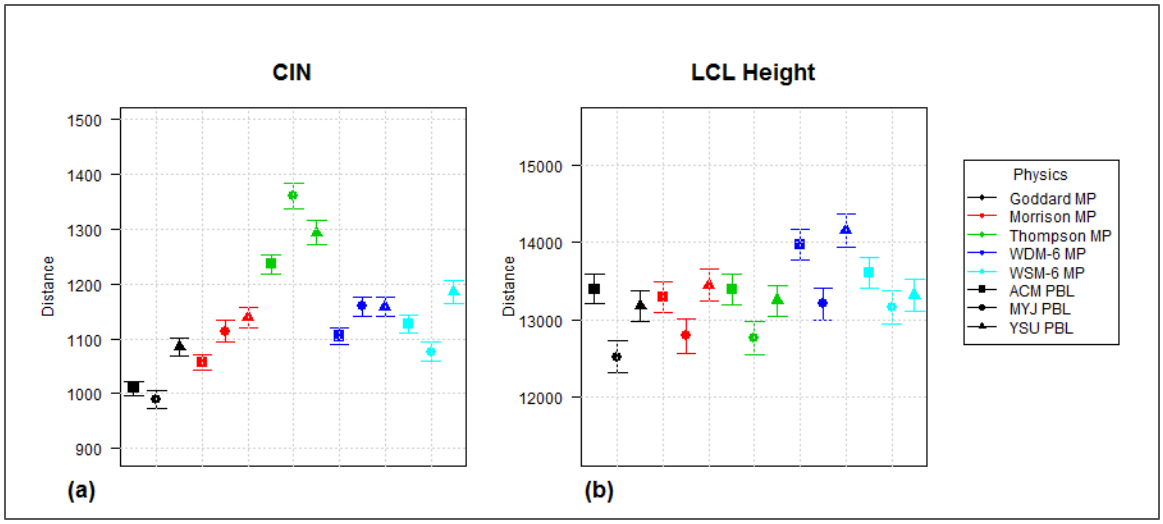

Figure 3.11 95% Confidence Intervals on Euclidean Distance Calculations Between Tornadic and Nontornadic Values of CIN (a) and LCL Height (b) for Phase II

3.3 Discussion

The predominant finding of phase II was that PBL physics parameterizations greatly influence the modeling of significant severe weather covariates as compared to microphysics parameterization. The YSU PBL scheme, which is commonly used in severe convective weather modeling studies, performed especially well in generating NWP forecasts with high skill at predicting severe weather outbreak mode. The revised Asymmetric Convection Model PBL scheme performed second-best of the examined PBL schemes, and the Mellor-Yamada-Janjić scheme generated NWP forecasts with the lowest level of skill at predicting outbreak mode.

The domain-mean covariate analysis of section 3.2.1 indicates that MYJ physics models 0-1 km SRH values higher than YSU physics in most nontornadic cases and the less intense tornado outbreaks. However, examination of individual cases indicates that

the overall domain mean of 0-1 km SRH is higher in all the YSU runs than the MYJ runs for 7 January 2008, 25 April 2011, 26 April 2011, and 27 April 2011. In addition, 5 February 2008, 9 April 2009, 15 April 2011, 16 April 2011, and 24 May 2011 had YSU and MYJ domain-mean 0-1 km SRH values that were similar, and in some of these cases YSU did produce higher 0-1 km SRH means with certain microphysics parameterization schemes. This selection of cases includes most of the strongest tornado outbreaks examined in this research. As the analysis of section 3.2.1 indicates, the MYJ scheme produced higher domain-mean 0-1 km SRH values for nontornadic outbreaks, but the YSU scheme produced higher domain-mean 0-1 km SRH in "outlier" cases. These cases were 11 February 2011 and 3 April 2011, a pair of spring-season outbreaks driven by powerful mid-latitude cyclones. These results indicate that MYJ may inflate 0-1 km SRH for less intense cases, while YSU increases it for events with more powerful dynamics. This result supports previous work (Hong et al. 2006) finding that YSU modeled fronts more intensely than a control PBL physics scheme in a simulation of a frontally driven tornado outbreak. A stronger front generally implies a deeper warm sector, a region of a low pressure system in which low-level winds have a southerly component but upper-level winds follow the jet stream with a west-to-east flow. Increased wind flow and wind shear in this sector lead to higher helicity values.

Similarities exist between the patterns of the outbreak bootstrap composites in this study and the findings of previous researchers. Notably, the MYJ PBL scheme generated apparent moist biases in the modeling of LCL height for both tornadic and nontornadic outbreaks. MYJ modeled LCL height in the nontornadic composite the lowest of all the PBL schemes. In the tornadic composite, the prefrontal side of the mean domain was

modeled with lower LCL heights with MYJ than with the other two schemes, and the frontal boundary itself was more diffuse. These results support the findings of Hu et al. (2010), who also found that this scheme generated moist biases in convective environments. The sharp gradient of the frontal boundary in the YSU tornadic LCL composites supports the findings of Hong et al. (2006), who also found that this scheme produced intense frontal boundaries in a single tornado outbreak case. It is noted that the Hong et al. case was an autumn tornado outbreak, but that the scheme's generation of a sharply defined frontal boundary has now been shown to appear in many spring tornado outbreaks as well. Considering that the frontal boundary was also sharply defined with the ACM2 scheme in the tornadic composites, a scheme that—like YSU—utilizes nonlocal closure, it is likely that nonlocal closure in the YSU and ACM2 accounts for the well-defined front. Unlike MYJ, these schemes can model mixing in the vertical dimension across PBL layers, rather than limiting mixing to each discrete layer.

The SRH and bulk shear results are more difficult to explain. One possible reason for those results is that in both tornadic and nontornadic composites, 0-1 km SRH and 0-1 km bulk shear were highest in the MYJ composites and lowest in the ACM2 composites. The MYJ scheme uses only local closure and the YSU scheme is nonlocal, but the ACM2 PBL scheme is a hybrid, utilizing both local and nonlocal mixing. It is possible that by employing mixing both within PBL layers and across them, the ACM2 scheme mixed excessively, minimizing the wind shear effects through friction. Tentative support for this hypothesis exists in the nontornadic LCL composites, which show the highest values—and thus the greatest amount of vertical mixing in a convective PBL—in the ACM2 runs. The sharp boundary of the tornadic composites makes it difficult to say

with certainty that this is also occurring in the tornado outbreaks, though the higher LCL heights (relative to both YSU and MYJ composites) in the prefrontal sectors of the ACM2 tornadic composites are suggestive.

The MYJ scheme may have produced higher shear and SRH values than the nonlocal YSU scheme because of its known cold bias in convective environments. A shallower PBL will have lower amounts of friction and turbulence of PBL origin, permitting dynamically driven winds to exert a stronger influence. Another possible source of higher shear and SRH values in MYJ model runs is the local closure of this scheme. The scheme performs mixing within individual vertical layers and does not mix across this dimension. Explicitly modeling vertical transport across layers may decrease horizontal wind magnitudes within some layers. Further research into this topic should be conducted.

The role of microphysics in LCL height simulation was less apparent than that of PBL physics, but a slight effect was observed between the WDM-6 and WSM-6 runs. The double-moment scheme produced higher LCL heights than the single-moment scheme across tornado outbreak domains, both in front of and behind the mesoscale cold front that appeared in most of the tornadic cases. Behind the boundary, this result can be readily explained by previous studies (such as Adams-Selin et al. 2013) that found that WDM-6 produced very strong cold pools behind thunderstorms. In the warm sector ahead of the boundary, the slightly higher LCL may be caused by greater evaporative cooling and drying due to smaller rain and cloud droplet sizes in WDM-6.

Examination of observed data for these events indicates that the boundary in the tornadic LCL composites was probably the result of the passage of a QLCS or a broken

line of discrete thunderstorms. Most of the cases did feature a synoptic cold front, but in about half of the cases with a synoptic front, it passed over the outbreak domain center up to 6 hours after the outbreak peak time. This boundary was not present in the nontornadic LCL composites to nearly the same degree, because many of these cases were not linearly oriented or did not travel west to east, and it is likely that the SVMs—especially the most-skilled WDM-6/YSU, which maximized tornadic-nontornadic Euclidean distance for the covariate—used this strong distinction between outbreak types heavily in their classification functions. The fact that the boundary in tornadic composites appears to be the result of the passage of the line of storms itself indicates that it is not a truly prognostic variable, but instead, that the SVM is identifying an artifact of tornadic outbreak occurrence after the outbreak has already happened for part of the model domain. However, this research did not examine SVM forecasts with NWP model output from before the outbreak valid time, so the possibility that the SVM might identify such a pattern as highly significant was "built in" to the experimental design. Within an NWP framework, this possibility should not be a concern for operational forecasters, who do not only use model output valid for times before the peak of a severe weather outbreak, but also the model prediction of the outbreak peak time. The strong distinction between tornadic and nontornadic LCL height fields may be an artifact of cold pool development from the main severe storm line, but if so, the SVM outbreak prediction results still indicate that the NWP model is simulating the type of outbreak well in the set of experimental cases.

CHAPTER IV

PHASE III—SYNOPTIC COMPOSITES OF DATA ASSIMILATION AND PHYSICS PARAMETERIZATION RUNS

4.1 Methodology

In this third and final phase of the study, the objective is to determine the effects that data assimilation (as conducted in phase I) and microphysics/PBL parameterization (as conducted in phase II) have on the modeling of synoptic-scale weather features in tornadic and nontornadic outbreaks. With regard to parameterization, it is hypothesized that microphysics parameterization will influence the modeling of upper-level cloud and wind features more strongly (as these features are heavily influenced by microphysics parameterization selection), whereas PBL physics will heavily influence lower-level thermodynamic and wind variables. With regard to data assimilation, it is hypothesized that conventional observations will most strongly influence thermodynamic variables given their direct measurements of atmospheric moisture, while satellite radiances will influence cloud features and upper-level winds since it is inferred from satellite infrared radiance temperatures, making those observations highly sensitive to data quality issues.

4.1.1 WRF model

Examining synoptic-scale fields requires a large enough domain to capture the extent of these features. The 4 km outbreak-centered domains developed for phases I and II were, for some smaller outbreaks, insufficiently large to effectively show the full

extent of synoptic features such as troughs and mid-latitude cyclones. For this reason, 12 km input fields were modeled. For all 40 cases, the WRF model was configured with a domain identical to that in Fig. 3.

The background fields generated by the GSI in phase I could not be used directly for synoptic analysis of the data assimilation runs because these fields were not processed by WRF at 12 km resolution after the data assimilation procedure, but instead were used as boundary conditions for the 4 km outbreak-centric domains. A new set of 12 km WRF model output was generated for analysis. Instead of NARR observations as initial fields for the WRF model, the final output fields from the GSI were used in the WRF model. Model physics for this set of runs were identical to those used in the 12 km background field stage of phase I and are shown in Table 4.1.

Table 4.1 Phase III Assimilation Ensemble WRF Physics Parameterizations

WRF physics option	Configuration
Cloud microphysics	Thompson et al. (2008)
Longwave radiation	Rapid Radiative Transfer Model (Mlawer et al. 1997)
Shortwave radiation	Dudhia (Dudhia 1989)
Surface layer	MM5-derived (Beljaars 1994)
Land surface	5-layer thermal diffusion (Dudhia 1996)
Urban surface	None
Planetary boundary layer	Yonsei University (Hong et al. 2006)
Cumulus physics	Kain-Fritsch (Kain 2004)

The 12 km parent domains generated in phase II were also unsuitable for synoptic analysis. These runs were produced with a two-way nesting configuration, which produced feedback from the smaller 4 km domains in the 12 km domains. No such feedback existed in the 12 km data assimilation runs generated for phase III. Therefore, a new set of 12 km parameterization ensemble runs was generated for phase III as well.

These runs used the NARR data as initial input. Model physics used in all cases for this set of runs are shown in Table 4.2, with microphysics and PBL physics variations in the parameterization ensemble shown in Table 4.3. There were four data assimilation runs and 15 physics parameterization runs, for a total of 19 variations for phase III.

Table 4.2 Phase III Parameterization Ensemble WRF Model Physics Schemes Used for All Runs

WRF physics option	Configuration	Reference
Longwave radiation	Rapid Radiative Transfer Model	Mlawer et al. 1997
Shortwave radiation	Dudhia	Dudhia 1989
Surface layer	*For ACM2 and* YSU *PBL:* MM5-derived	Dudhia 1996
	For MYJ PBL: Monin-Obukhov/Eta similarity	Janjić 2002
Land surface	Noah land surface model	Tewari et al. 2004
Urban surface	None	
Cumulus physics	*For 12 km nest:* Kain-Fritsch *For 4 km nest:* None	Kain 2004

Table 4.3 Phase III Cloud Microphysics and PBL Physics Variations

Physics Option	Reference
Goddard microphysics	Tao et al. 1989
Morrison double-moment microphysics	Morrison et al. 2009
Thompson microphysics	Thompson et al. 2008
WRF Double-Moment 6-class microphysics	Lim and Hong 2010
WRF Single-Moment 6-class microphysics	Hong and Lim 2006
Yonsei University PBL physics	Hong et al. 2006
Mellor-Yamada-Janjić PBL physics	Janjić 1994
Asymmetric Convection Model PBL physics	Pleim 2007

4.1.2 Synoptic Composites

Following the methodology of Mercer et al. (2012), synoptic composites of geopotential height, relative humidity, temperature, and *u* and *v* wind components were generated for all data assimilation and physics parameterization variations modeled in phases I and II. Pressure levels and corresponding approximate altitudes at which they were extracted are shown in Figure 4.1. Since this was a synoptic-scale analysis, and to minimize the effect of mesoscale noise in the analysis, these variables were linearly interpolated from NWP model output to a preset latitude-longitude grid of 72 km resolution.

A T-mode rotated principal component analysis (Richman 1986) was employed in generating analysis matrices. A T-mode RPCA was chosen instead of an S-mode RPCA because in an S-mode analysis, the correlations examined are between gridpoints, whereas in a T-mode, correlations between entire cases are examined. The size of the correlation matrices in an S-mode RPCA for this data set would have made the analysis

computationally infeasible even with a 72 km grid. For tornadic and nontornadic outbreaks (20 cases each), RPCA scores were generated separately for each of the 19 model variations. This RPCA entailed a correlation that was computed along the dimension of number of outbreaks, producing 20x20 correlation matrices for tornadic and nontornadic outbreaks for each run. The RPCA was conducted on 60 different variable-height fields for each case, with horizontal dimensions of 34 x 34 data points on the 72 km grid.

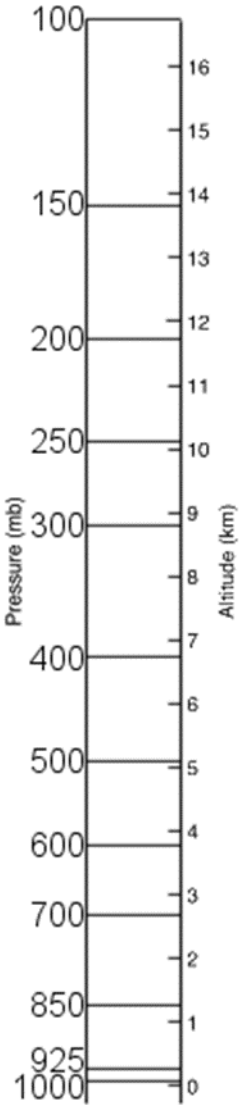

Figure 4.1 Pressure Levels and Altitudes of Examined Variables

4.1.2.1 K-means Cluster Analysis

Since a variety of atmospheric "setups" and thermodynamic or forcing mechanisms can contribute to the generation of severe weather, and such differences would be averaged out, it was possible that synoptic composites could be produced that did not resemble the actual synoptic pattern of many of the tornadic or nontornadic cases.

To prevent this problem from occurring, *k*-means clustering analyses were conducted on the nontornadic and tornadic outbreaks. This method of clustering data requires a prior expectation of the number of clusters to be generated. Since this number was unknown, a variety of clusters and PCA loadings were examined. The optimal number of loadings and clusters was determined by calculating the silhouette coefficient (Rousseeuw 1987). This number is calculated by:

$$Silhouette(\mathbf{a}) = \frac{Separation(\mathbf{a}) - Cohesion(\mathbf{a})}{\max[Separation(\mathbf{a}), Cohesion(\mathbf{a})]} \tag{4.1}$$

Separation in this equation is the distance between the nearest cases in each of the clusters. Cohesion represents the distance of each case within a cluster from the cluster's center point. Therefore, for the most distinct set of clusters, maximizing separation and minimizing cohesion is ideal. It was found that, for tornadic and nontornadic outbreaks, three PC loadings and four clusters optimized the silhouette coefficient, for a total of eight clusters (four for each type of outbreak). This result was found to be optimal for every model configuration examined across the data assimilation and the physics parameterization ensembles, although the individual membership of each cluster varied among some of the model runs. The tornadic clusters had average silhouette widths of 0.35 to 0.41, and the nontornadic clusters had average widths of 0.5 to 0.55. Synoptic variables as listed—geopotential height, relative humidity, temperature, and *u* and *v* wind components—were then extracted and averaged for each cluster to generate synoptic fields.

Further following the methodology of Mercer et al. (2012), manual examination of the resulting synoptic composites was conducted to determine variations between tornadic and nontornadic cases and the effect upon the analyzed synoptic fields by physics parameterization and assimilation of satellite radiances and conventional observations.

4.2 Results

Differences existed among the model ensemble in terms of which tornado outbreaks were assigned to each cluster. The nontornadic outbreaks were all assigned to the same clusters irrespective of data assimilation or model physics parameterization variations.

Table 4.4 shows cluster assignments for each outbreak (note that the numbers simply represent cluster membership, not priority or intensity). Since all nontornadic outbreaks sorted into the same clusters irrespective of data assimilation or model physics, and since all data assimilation runs of tornado outbreaks sorted into the same clusters, these events are represented by single columns in the table. Cluster mean fields were generated from the individual fields of each cluster's constituent outbreaks, and standard synoptic variables and levels were then visualized for each cluster.

Table 4.4 Synoptic Cluster Assignments for Outbreak and Model Configuration.

Tornado Outbreaks	Data Assim. Runs	Goddard/ACM	Goddard/MYJ	Goddard/YSU	Morrison/ACM	Morrison/MYJ	Morrison/YSU	Thompson/ACM	Thompson/MYJ	Thompson/YSU	WDM-6/ACM	WDM-6/MYJ	WDM-6/YSU	WSM-6/ACM	WSM-6/MYJ	WSM-6/YSU	Nontornadic Outbreaks	All Model Runs
7 January 2008	2	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	15 June 2008	4
5 February 2008	2	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	8 July 2008	3
15 March 2008	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	20 July 2008	3
9 April 2009	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2 August 2008	3
10 April 2009	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	11 February 2009	2
8 May 2009	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	15 May 2009	1
24 April 2010	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	3 June 2009	1
10 May 2010	1	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	18 June 2009	3
5 June 2010	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	24 July 2009	4
17 June 2010	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	9 August 2009	3
4 April 2011	2	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	6 April 2010	2
15 April 2011	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	28 May 2010	4
6 April 2011	4	4	4	4	4	4	4	4	4	4	1	4	4	1	4	4	15 June 2010	3
19 April 2011	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	18 June 2010	3
25 April 2011	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	23 March 2011	2
26 April 2011	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	3 April 2011	2
27 April 2011	4	4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	11 April 2011	1
22 May 2011	3	1	1	3	3	1	3	3	1	1	3	1	3	3	1	3	20 April 2011	1
24 May 2011	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	4 June 2011	3
25 May 2011	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	9 August 2011	3

4.2.1 Composites of Tornadic Outbreaks

The four tornado outbreak clusters exhibited several synoptic features in common with each other. The most obvious differences were of magnitude rather than type. All clusters showed a pronounced trough at 500 mb, with areas of high mid-level (700 mb) relative humidity in the base of the trough and temperature advection occurring at 850 mb due to a thermal gradient and warm sector. A southwest-to-northeast temperature gradient existed for all clusters at 850 mb, with the warmest values apparent in clusters 1 and 3. These warm isotherms were 5 K warmer than the warmest of clusters 2 and 4. However, in clusters 2 and 4, the minimum isotherms were 10 K cooler than the minima of clusters 1 and 3. Based on manual analysis of the maps and calculation of mean fields at 850 mb, thermal advection was weaker in clusters 2 (0.11 K/hr mean) and 3 (0.09 K/hr) due to the broader pressure gradient and flatter pressure areas south of the trough. Strong thermal advection was present in clusters 1 (0.13 K/hr) and 4 (0.18 K/hr).

Synoptic composites of temperature, humidity, pressure, and wind fields are shown for tornado outbreaks in Fig. 4.2 and Fig. 4.3. These composites represent the outbreaks as modeled with Thompson microphysics and YSU PBL physics for phase II. This set of microphysics and PBL parameterizations was also used for the phase I data assimilation runs, although the phase I runs employed a different land surface physics.

Cluster 1 depicted a trough of 5500 m at 500 mb and a tight pressure gradient to the southeast of the trough. A comma cloud shape was apparent in relative humidity fields at 700 mb (Fig. 4.2a). A jet streak was present at 300 mb, with a very tight wind gradient on the northwest side of the jet (Fig. 4.3a). Moderate wind shear was present at the mid-to-lower levels of the atmosphere, with winds exhibiting a strong westward

component at 500 mb and a southerly component at 925 mb. Wind speeds were generally 5 to 15 knots faster at 500 mb than at 925 mb ahead of the jet stream, indicating that this was primarily a directional shearing pattern. Beneath the jet stream itself, the difference was up to 60 knots.

Cluster 2 showed a strong negatively tilted trough as well (a height field of 5550 m), though isohypses south of the trough were broader and more attenuated. This cluster also exhibited a comma cloud shape at 700 mb in relative humidity values (Fig. 4.2b). Winds in this cluster were very strong, with a large region of 80 knot winds at 300 mb. High directional and speed shear existed at lower levels, with winds from the west or northwest at 500 mb and winds from the south or southwest at 925 mb (Fig. 4.3b). The speed difference between these two altitudes ranged from 5 to 25 knots in front of the jet.

Cluster 3 was similar in appearance to cluster 2 at 500 mb, but it was less intense, with the trough having a height of 5650 m. The comma shape present in the 700 mb relative humidity fields for clusters 1 and 2 was less pronounced in this cluster (Fig. 4.2c). Wind fields for this cluster illustrated a right rear quadrant jet stream flow pattern, with the highest magnitude winds at 300 mb in the northeast corner of the composite and a southwest-west directional shear pattern apparent from 925 mb to 500 fields (Fig. 4.3c). This composite was the most weakly forced of the tornado outbreak composites.

Cluster 4 depicted a deep, spatially broad trough with a height of 5450 m at the center and a tight height gradient. At 700 mb, an area of high relative humidity existed in the low, but this cluster lacked the north-south area of high humidity at this level that was present, to varying degrees, in the other three clusters (Fig. 4.2d). Wind fields at 300 mb

were characterized by a southwest-northeast jet with peak wind speeds up to 100 knots and high speed shear from 925 to 500 mb (Fig. 4.3d).

Overall, all four tornado outbreak clusters exhibited important synoptic similarities. These clusters all depicted a trough, a typical synoptic-scale feature associated with spring-season tornado outbreaks. Cluster 4 featured the most intense system in terms of height, followed by cluster 1 and cluster 2. Cluster 3 depicted the system with the weakest dynamic forcing. However, clusters 1 and 4 held areas of apparent stronger thermal advection than clusters 2 and 3, based on visual analysis of temperature and height fields. These two clusters' similarity in terms of strong thermal advection accounts for the sorting of the record outbreak 27 April 2011 into cluster 1 or cluster 4, depending on the model configuration.

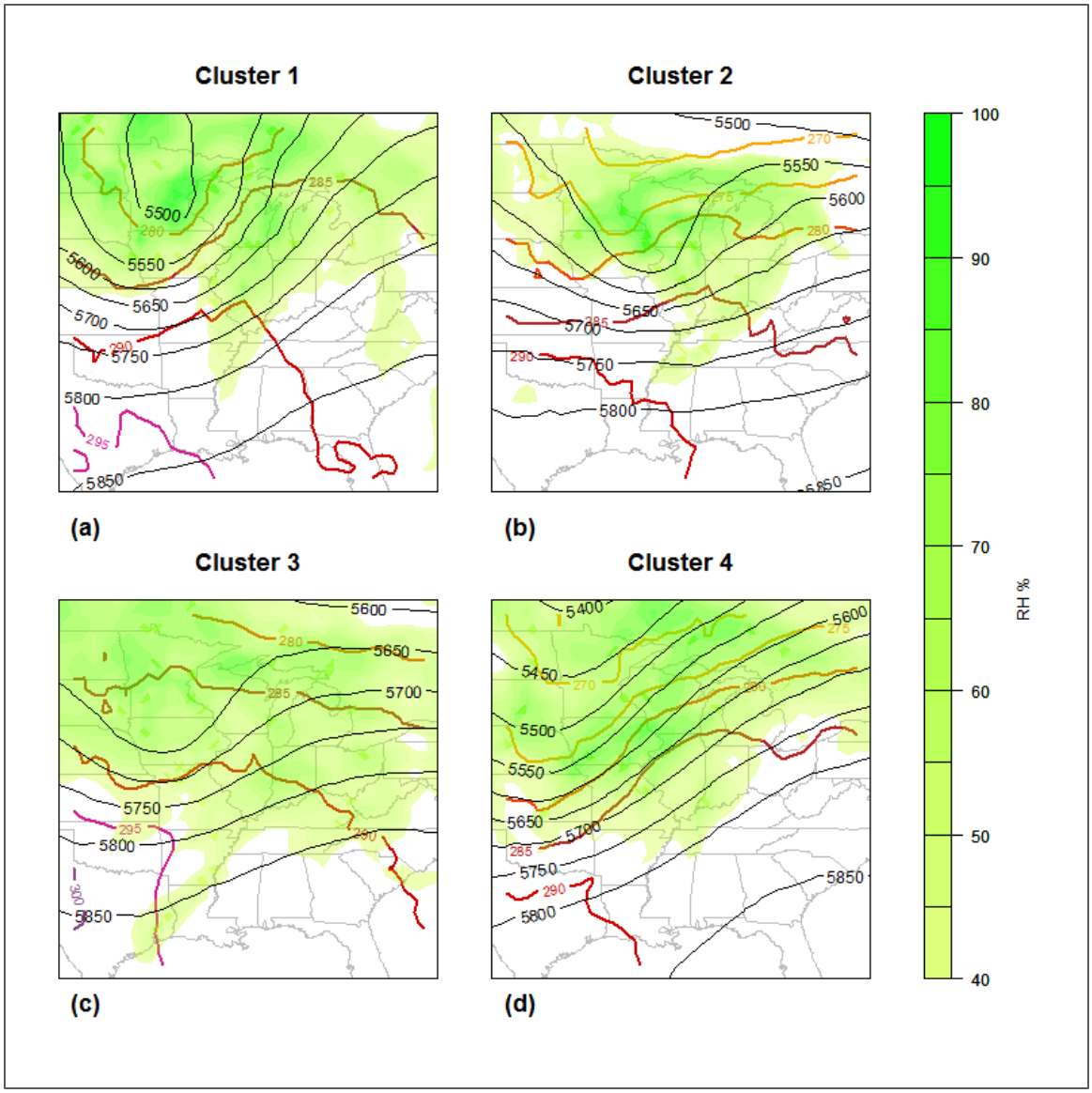

Figure 4.2 Composites of 500 mb Height, 700 mb RH, and 850 mb Temperature in Tornadic Clusters, Modeled with Thompson Microphysics and YSU PBL

Geographic underlays are provided for scale only and do not reflect the geographical location of synoptic features in all the constituent outbreaks.

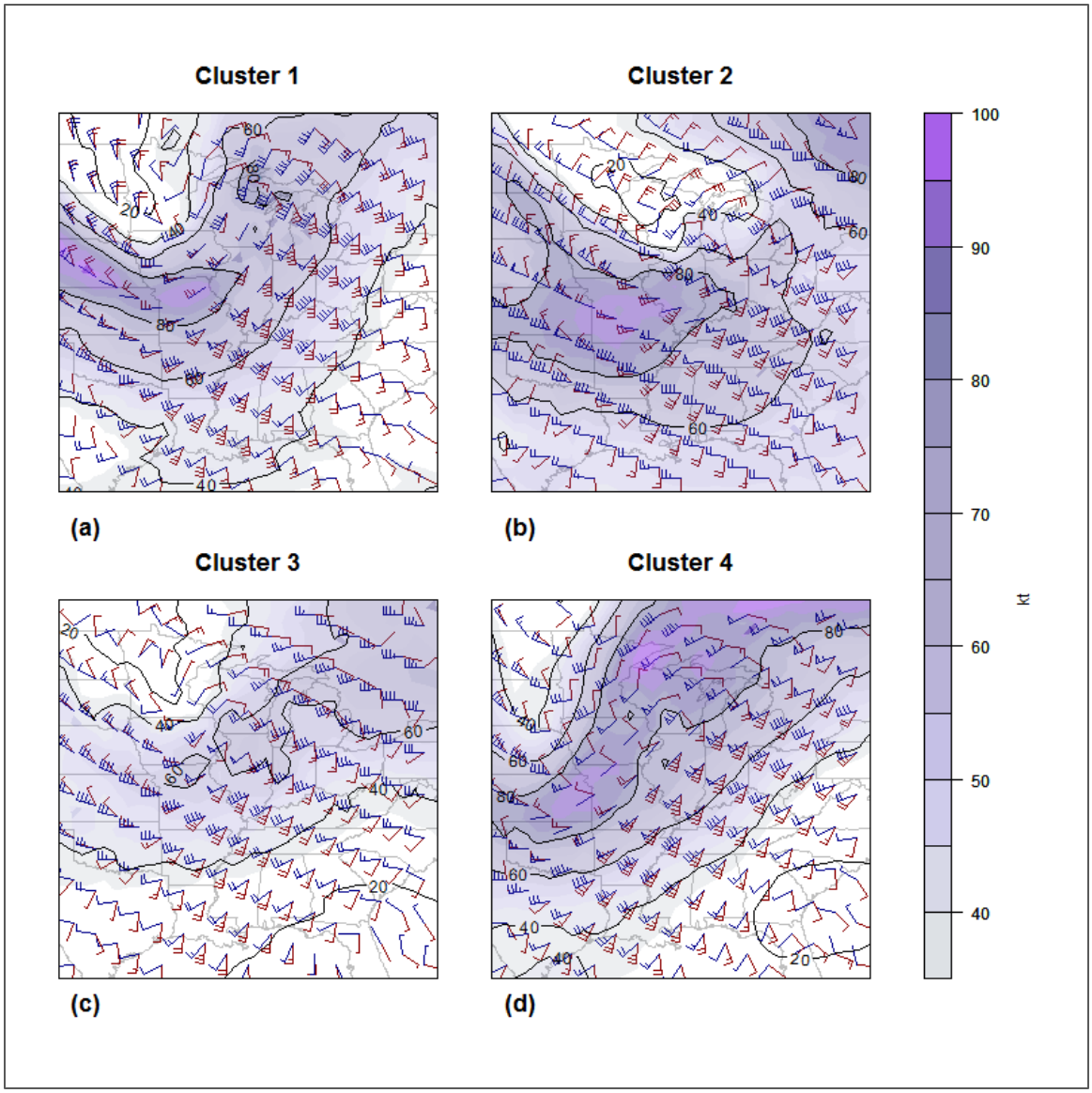

Figure 4.3 Composites of 300 mb Wind Magnitude, 500 mb Magnitude-Direction, and 925 mb Magnitude-Direction in Tornadic Clusters, Modeled with Thompson Microphysics and YSU PBL

500 mb wind barbs are blue; 925 mb wind barbs are red. Wind speeds are in knots. Geographic underlays are provided for scale only and do not reflect the geographical location of synoptic features in all the constituent outbreaks.

4.2.1.1 Assimilation Ensemble Modeling Variations

As Table 4.4 indicates, several tornado outbreaks were assigned to different clusters for the four "phase I" data assimilation runs (three runs with data assimilated and one run with no data assimilation but the land surface physics different from the "phase II" runs). The outbreaks of 7 January 2008, 5 February 2008, and 4 April 2011 were assigned to cluster 2 in these four runs and to cluster 4 in all the physics parameterization runs. 10 May 2010 was assigned to cluster 1 in the four "phase I" runs and cluster 3 in the "phase II" runs. 27 April 2011 was assigned to cluster 4 in these runs, but to cluster 1 in all the physics parameterization runs except one. 22 May 2011 was assigned to cluster 3 in these four runs, but its assignment varied among the physics parameterization runs.

The physics parameterization runs also exhibited differences among themselves in cluster assignment for 15 April 2011 (assigned to cluster 1 in WDM-6/ACM2 and WSM-6/ACM2, and to cluster 4 in all other runs), 27 April 2011 (assigned to cluster 4 in Goddard/ACM2 and cluster 1 in all other runs), and 22 May 2011 (assigned to cluster 1 or 3 approximately equally).

It is evident from the "disagreement" among the model runs that 22 May 2011 was not a clear-cut fit for either cluster 1 or 3, and as described previously, these two clusters exhibited similar forcing patterns, though cluster 3 showed weaker dynamics. Although this event is best remembered for the Joplin, MO EF-5 tornado, a swath of scattered tornadoes developed from extreme northeastern Oklahoma along a diagonal into northern Wisconsin. That the SD12 highest-ranked tornado outbreak, 27 April 2011, was assigned to different clusters was unexpected. An examination of the observed synoptic fields for this event reveals that a large, deep trough 5460 m in height was centered over

the upper Midwest, which more closely resembles the 500 mb fields of cluster 1, but the wind pattern at 300, 500, and 925 mb more closely resembled the intense jet stream pattern of cluster 4. During the height of the outbreak, the jet stream had 100 knot streaks at 300 mb. Cluster 4 depicted up to 100 knot winds in the jet stream.

Differences in synoptic modeling between the data assimilation runs and the physics parameterization runs may be attributable to the use of different land surface physics (5-layer thermal diffusion in the assimilation runs and Noah land surface in the parameterization runs), or to different cluster assignments in the assimilation runs of several tornado outbreaks. Therefore, variations among the synoptic fields are considered within each category of model runs separately.

Within the assimilation runs and no-assimilation control, the troughs in all clusters were not as deep in the two runs containing conventional observations as they were in the control run and the HIRS-4 run. Slight differences in orientation also occurred, with the troughs slightly more negatively tilted in the control and HIRS-4 runs. The exception was cluster 2 in the HIRS-4/conventional run, in which the trough was slightly deeper than in any other assimilation run. Differences in the 700 mb moisture field were very small, but a slight pattern was observed in clusters 2 and 3 with drier moisture fields in a small area of the southwestern part of the domain in the conventional and HIRS-4/conventional runs. This indicates weaker dynamically driven moisture flow toward the trough. The 850 mb temperature field in cluster 2 extended warm isotherms farther north in the conventional and HIRS-4/conventional runs, but this pattern was not observed for the other clusters. Overall the addition of conventional observations slightly weakened the magnitude of the troughs and associated temperature and moisture

advection, indicating that—since the data assimilation procedure was found in phase I to be functioning correctly—the WRF model configuration used in this phase, or the NARR data, natively modeled the troughs too deep. A possible cause is the Thompson microphysics scheme used in the data assimilation ensemble, which was found in phase II to increase values of certain thermodynamic fields compared to other microphysics schemes.

Far clearer differences existed among the wind fields for the assimilation runs. The conventional and HIRS-4/conventional runs demonstrated stronger 300 mb winds and larger areas of high 300 mb winds in all four clusters than did the runs not containing conventional observations. HIRS-4 radiances alone also generated a larger area of high 300 mb winds in clusters 1 and 2 than were apparent in the control run. The HIRS-4 run also produced higher winds at 500 mb in cluster 2 than any other assimilation run or the control. The conventional run exhibited greater directional shear from 925-500 mb in cluster 3 than any other assimilation run.

The wind results of this phase mirror the results of phase I, in which 12 km tornadic outbreak composites with conventional data assimilated (both the conventional-only and the HIRS-4/conventional runs) had notably higher values of 0-1 km bulk shear, 0-1 km EHI, and CAPE x 0-1 km bulk shear than the control and the HIRS-4 composites. The phase III results indicate that assimilation of conventional observations increases wind magnitudes above the near-surface layer of the atmosphere and at a synoptic scale. It is also important to recall that the observations test of phase I, which found that conventional and HIRS-4/conventional runs exhibited lower RMSE, was performed on standard atmospheric variables (geopotential height at 500 mb; *u*-winds at 925, 850, 700,

500, 400, 300, 250, 200, 150, and 100 mb; *v*-wind at the same pressure levels; specific humidity at 700 and 500 mb; and temperature at 925, 850, 700, and 500 mb). These variables include fields examined in the synoptic analysis of phase III, indicating that conventional data assimilation improves the accuracy of NWP output of these cases in addition to increasing the values of these variables.

4.2.1.2 Parameterization Ensemble Modeling Variations

The physics parameterization runs also generated slight differences in the modeling of tornado outbreak synoptic features, though these features were not apparent for all clusters and all microphysics options. Morrison and Thompson microphysics modeled the trough in cluster 1 at a lower height than Goddard microphysics, especially in conjunction with ACM2 PBL physics (Fig. 4.4, Fig. 4.5). WDM-6 (Fig. 4.6) and WSM-6 microphysics modeled this feature at a lower height still. However, the trough extended slightly farther south with Goddard microphysics. The 290 K isotherm in the 850 mb temperature field of cluster 3 extended farther to the north with the Morrison and Thompson microphysics than the other microphysics, especially in conjunction with MYJ PBL physics (Fig. 4.7).

Examination of the wind fields across the physics parameterization ensemble reveals that, with ACM2 PBL physics, the Goddard microphysics parameterization models the jet streak in cluster 1 over a smaller region than any other microphysics (Fig. 4.8). This result is expected given the less intense trough modeled by this set of parameterizations. The temperature difference in cluster 3 did not translate into any visible differences in wind at 300 mb among the microphysics runs with MYJ PBL physics, but the 925-500 mb crossover shear for Morrison and especially Thompson

microphysics is slightly greater than that of other microphysics-MYJ runs over the region where the 290 K isotherm extends in these runs. These schemes model rain and cloud ice with two moments, unlike the other microphysics schemes in this research (the WDM-6 models rain and cloud water droplets with two moments). The region of the 290 K isotherm in which the highest shear occurs is very slightly to the southeast of the highest 300 mb winds, which also coincides with an area of deep 700 mb moisture. The double-moment cloud ice resolution of Morrison and Thompson physics likely produced, in this case, a more intense area of strong convection. Phase II also found that Thompson microphysics increased 0-1 km bulk shear values more in the less intense tornado outbreaks. Cluster 3 was the most weakly forced of the tornado outbreak clusters and, in the Thompson/MYJ run, contained all but one of the tornado outbreaks with 0-1 km bulk shear domain averages below the 50^{th} percentile rank for the tornado outbreak data set. The evaporative effect that has resulted in weaker convection in past research with schemes that model double-moment cloud ice may be less pronounced in severe convective cases that involve weaker winds. Nonetheless, the shear magnitude difference was still small, approximately 5-10 knots.

Composites of selected physics parameterization fields are shown in figures 4.4, 4.5, 4.6, 4.7, and 4.8. In all composites, geographical maps are provided for scale only and do not reflect the geographical location of synoptic features in all the constituent outbreaks.

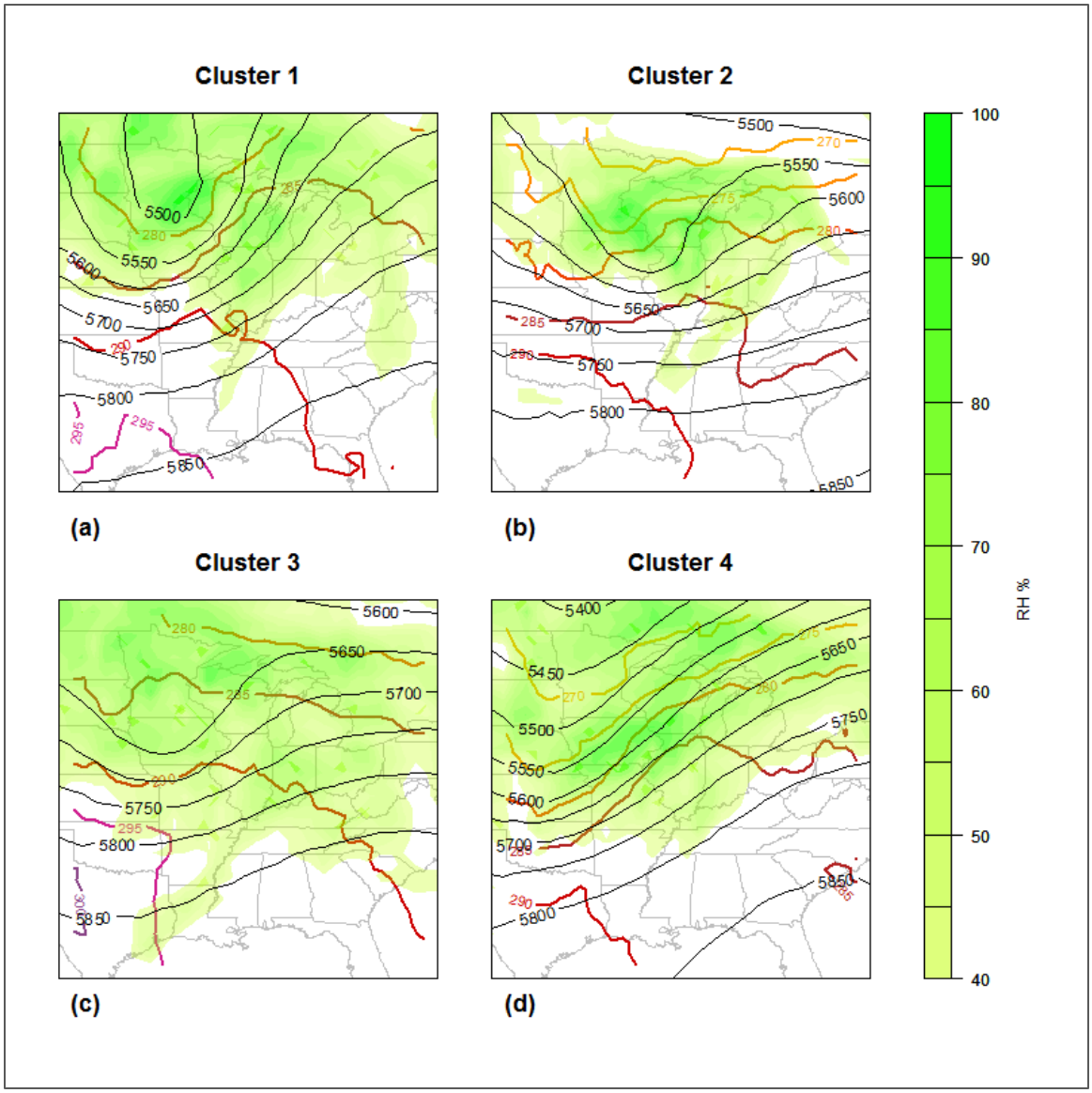

Figure 4.4 Composites of 500 mb Height, 700 mb RH, and 850 mb Temperature in Tornadic Clusters, Modeled with Morrison Microphysics and ACM2 PBL

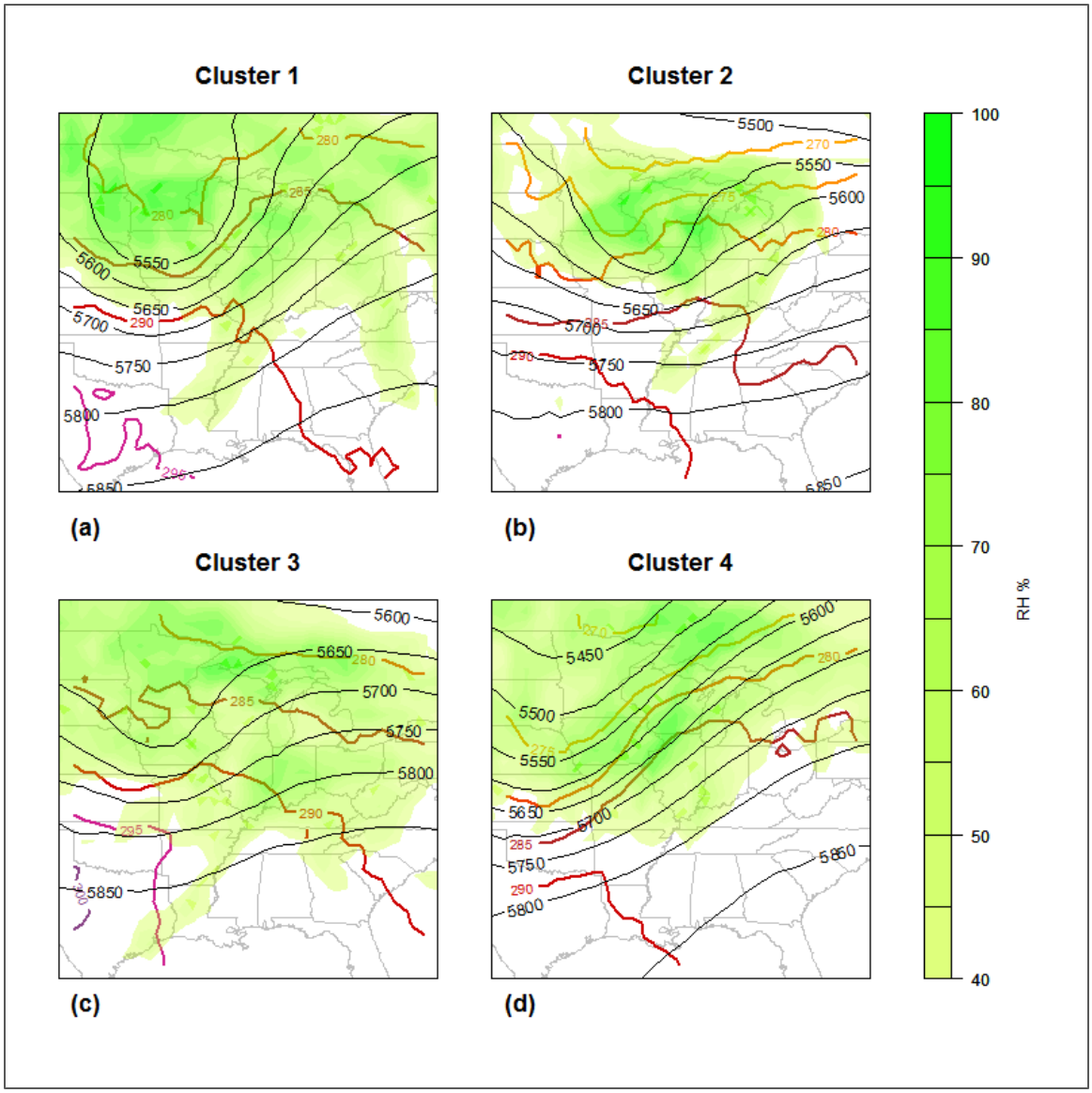

Figure 4.5 Composites of 500 mb Height, 700 mb RH, and 850 mb Temperature in Tornadic Clusters, Modeled with Goddard Microphysics and ACM2 PBL

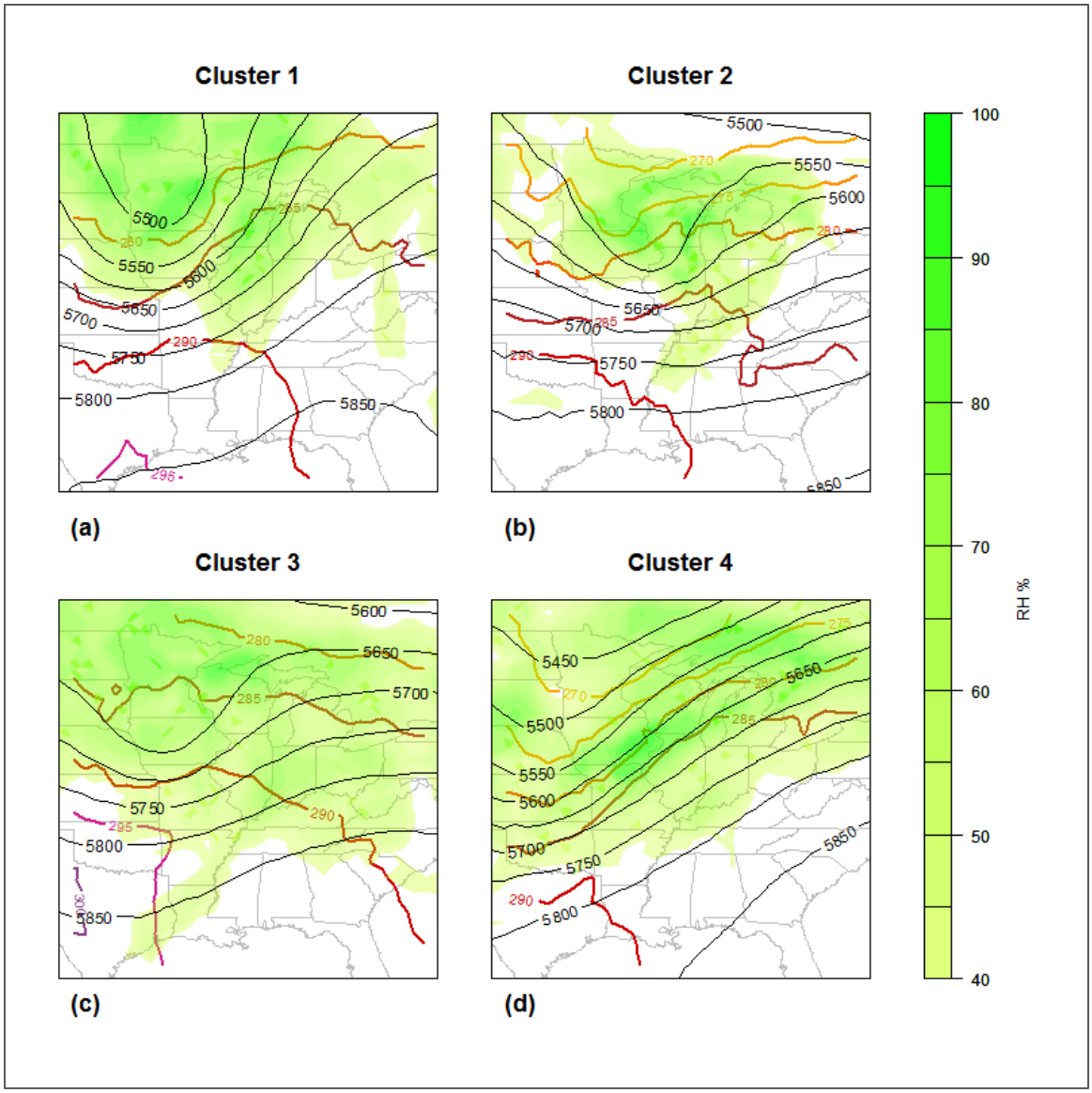

Figure 4.6 Composites of 500 mb Height, 700 mb RH, and 850 mb Temperature in Tornadic Clusters, Modeled with WDM-6 Microphysics and ACM2 PBL

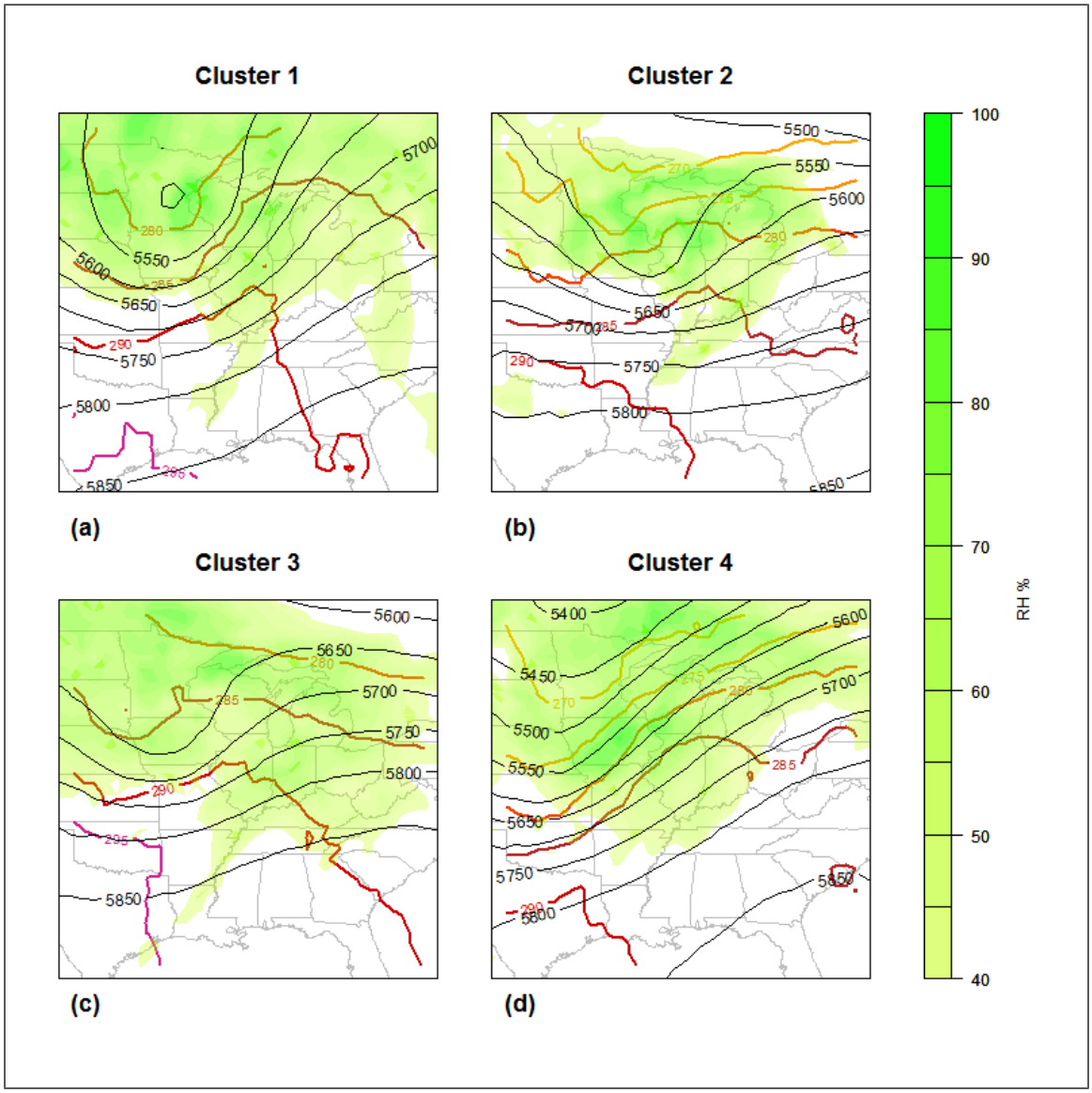

Figure 4.7 Composites of 500 mb Height, 700 mb RH, and 850 mb Temperature in Tornadic Clusters, Modeled with Morrison Microphysics and MYJ PBL

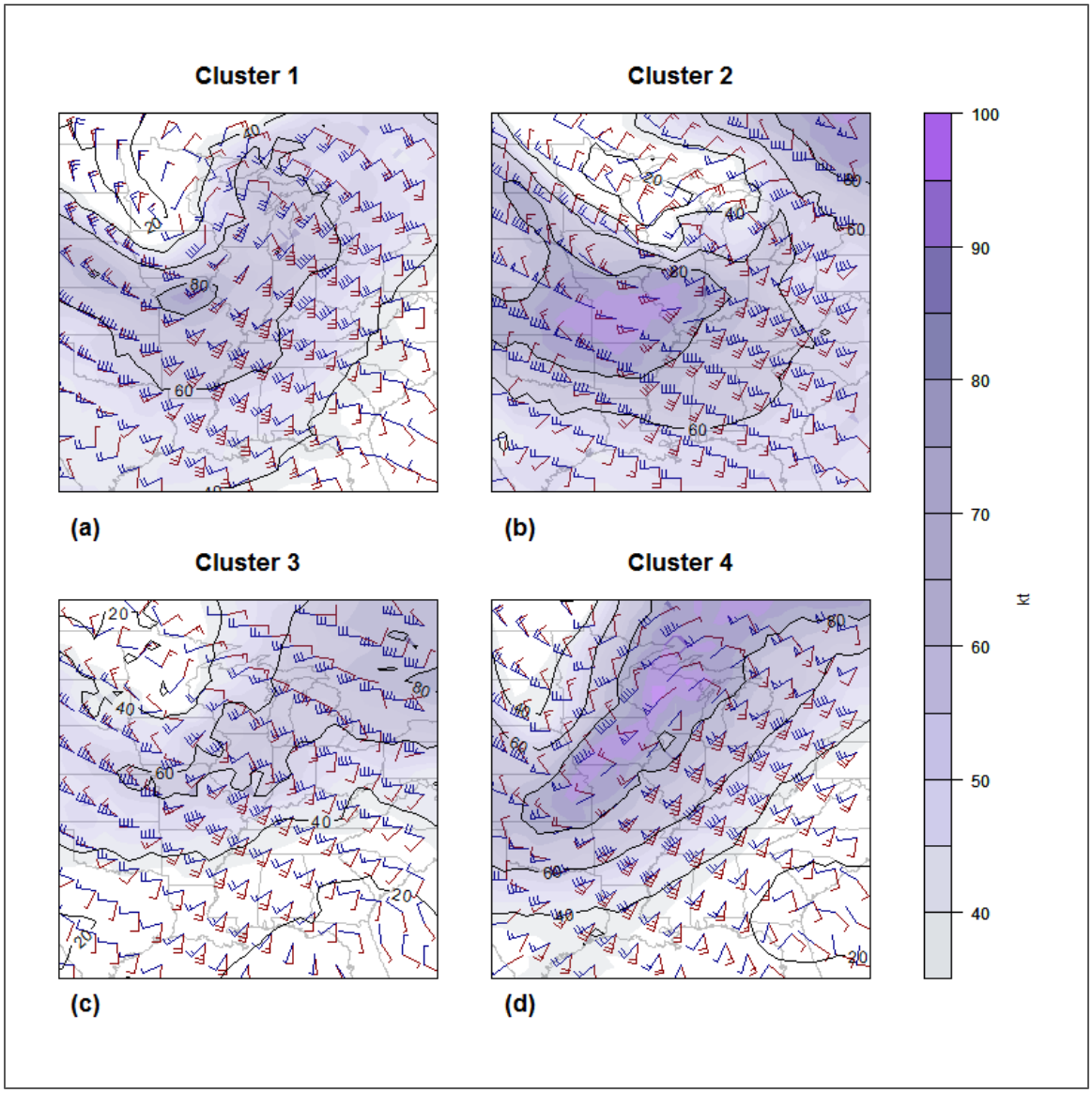

Figure 4.8 Composites of 300 mb Wind Magnitude, 500 mb Magnitude-Direction, and 925 mb Magnitude-Direction in Nontornadic Clusters, Modeled with Goddard Microphysics and ACM2 PBL

4.2.2 Composites of Nontornadic Outbreaks

In contrast with the tornadic outbreaks, the nontornadic outbreaks sorted into four very different clusters. Cluster 1 depicted a broad trough and tight geopotential height/temperature gradients across the north side of the domain. Cluster 2 was the most similar to the tornado outbreak clusters, showing a trough 5400 m in height at 500 mb and a very tight pressure gradient. Cluster 3 was quite distinct, showing a ridge of 5900 m and a temperature gradient on the north side of the ridge. Cluster 4 was characterized by a trough to the northeast and a northwesterly flow pattern to the west of the trough.

In cluster 1, isotherms at 850 mb were mostly parallel with isohypses at 500 mb, with the exception of a large region of high temperatures with a north-south bend south of the trough. A region of high relative humidity at 700 mb with a southwest-northeast orientation was also aligned with the height and temperature gradient (Fig. 4.9a). Wind patterns for this cluster (Fig. 4.10a) illustrated a strong jet streak in the northeast region of the cluster. A southwesterly flow pattern was evident in the southern region, indicating a right rear quadrant jet streak pattern.

Cluster 2, as mentioned, showed a classic intense mid-latitude cyclone pattern, with a very pronounced comma cloud shape of mid-level moisture and a deep warm sector (Fig. 4.9b). The wind fields for this cluster exhibited high magnitudes and a strong directional shearing pattern (Fig. 4.10b). 300 mb winds were very strong and a deep dip in the jet was apparent. The mid- and lower-level wind fields in this cluster were suggestive of a left front jet streak entry pattern. Altogether this cluster depicted a very strongly forced synoptic pattern.

Clusters 3 and 4 depicted environments not characterized by a classic trough, but examination of observational data for individual cases within clusters 3 and 4 indicated that the modeled flow patterns were representative of the constituent outbreaks. The environment of cluster 3 was characterized by southwesterly flow, veering to westerly, across the northern extent of a ridge and through a generally east-west thermal boundary (Fig. 4.9c, Fig. 4.10c). This environment is associated with derechos and quasi-linear convective systems (QLCS). Indeed, all cases comprising cluster 3 were summertime QLCS and derecho events. These events occurred in the months of June, July, and August. The springtime nontornadic events—11 February 2009, 6 April 2010, 23 March 2011, 3 April 2011, 11 April 2011, and 20 April 2011—sorted into clusters 1 and 2, the clusters that most closely resembled mid-latitude storm setups in the spring months. These findings indicate that a seasonally based sorting did occur to an extent. As the cluster analysis shows, nontornadic severe weather outbreaks are commonly produced by a variety of very different atmospheric setups, which occur at different times of the year, whereas most non-tropical tornado outbreaks are produced by a similar type of weather feature that is unusual in summer months.

Cluster 4 was composed of only 3 events, but these events all exhibited a trough and northwesterly flow pattern (Fig. 4.10d) suggestive of an upper-level low across a northwest-southeast 850 mb temperature gradient that defined the cluster (Fig. 4.9d). The upper-level flow pattern for this cluster was weak, though strong directional shear existed from 500 to 925 mb due to consistent—but weak—south-southwesterly flow toward the trough at this level.

The outbreaks that were primarily hail—15 May 2009, 3 June 2009, 6 April 2010, 23 March 2011, and 3 April 2011—sorted into clusters 1 and 2, the trough and mid-latitude cyclone clusters. The outbreaks that were mixed-mode—15 June 2008, 24 July 2008, and 9 August 2011—sorted into clusters 3 and 4. No clear pattern existed in terms of the clustering of nontornadic outbreaks that were primarily wind; these outbreaks appeared in all four clusters and during both spring and summer. Although hail-dominant and mixed-mode outbreaks occurred in both the spring and summer, a seasonal bias does appear in the type of synoptic weather patterns that produced these outbreaks. The hail-dominant events—with the exception of 3 June 2009—were springtime outbreaks, whereas the mixed-mode outbreaks occurred in the summer. This result is generally reflective of hail and severe wind climatology. Doswell et al. (2005), for instance, found that hail events in the United States have notably higher probabilities of occurrence during the April-June period than severe wind events, which have peak occurrences from June-August.

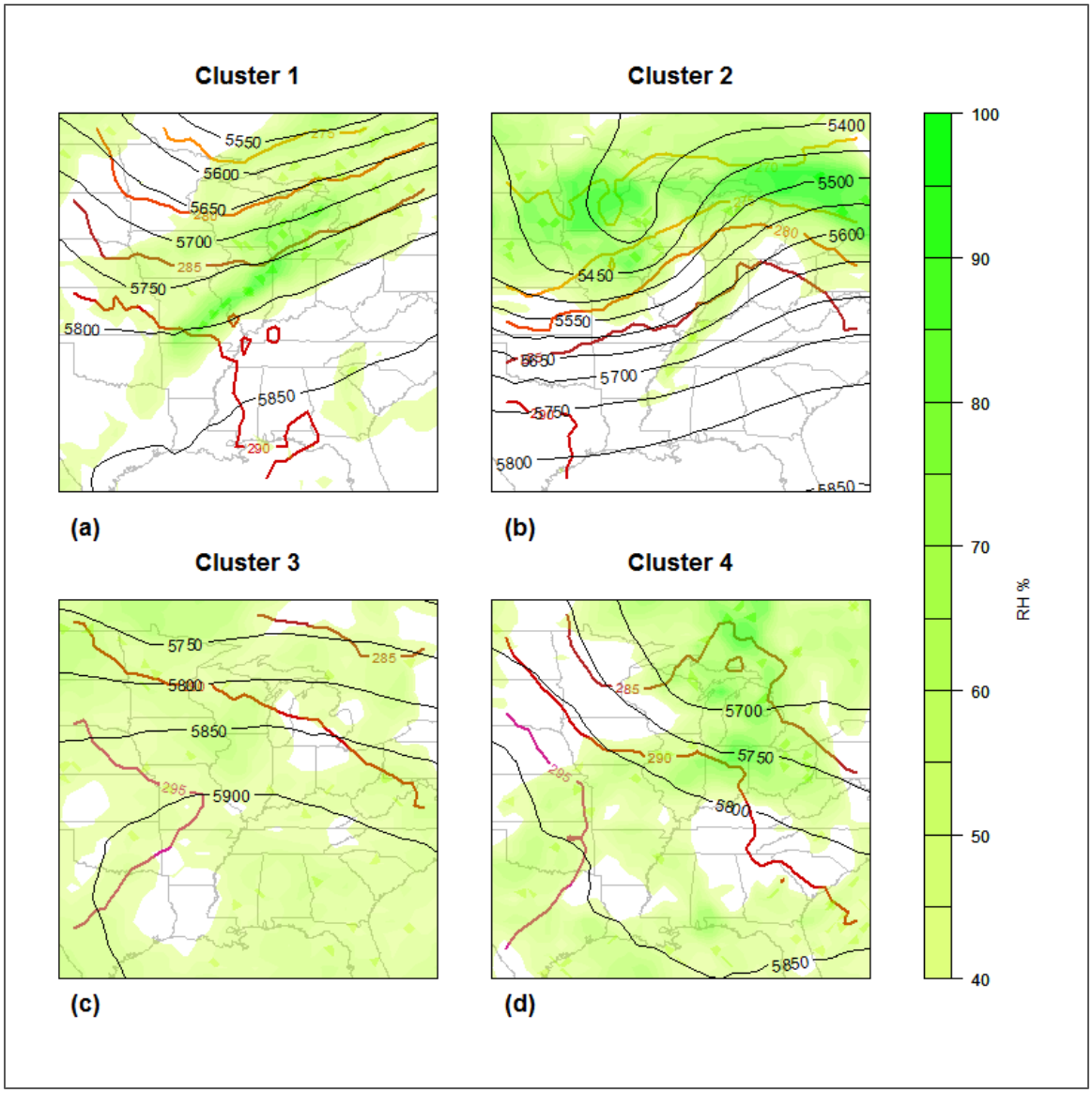

Figure 4.9 Composites of 500 mb Height, 700 mb RH, and 850 mb Temperature in Nontornadic Clusters, Modeled with Thompson Microphysics and YSU PBL Physics

Maps are provided for scale only and do not reflect the geographical location of synoptic features in all the constituent outbreaks.

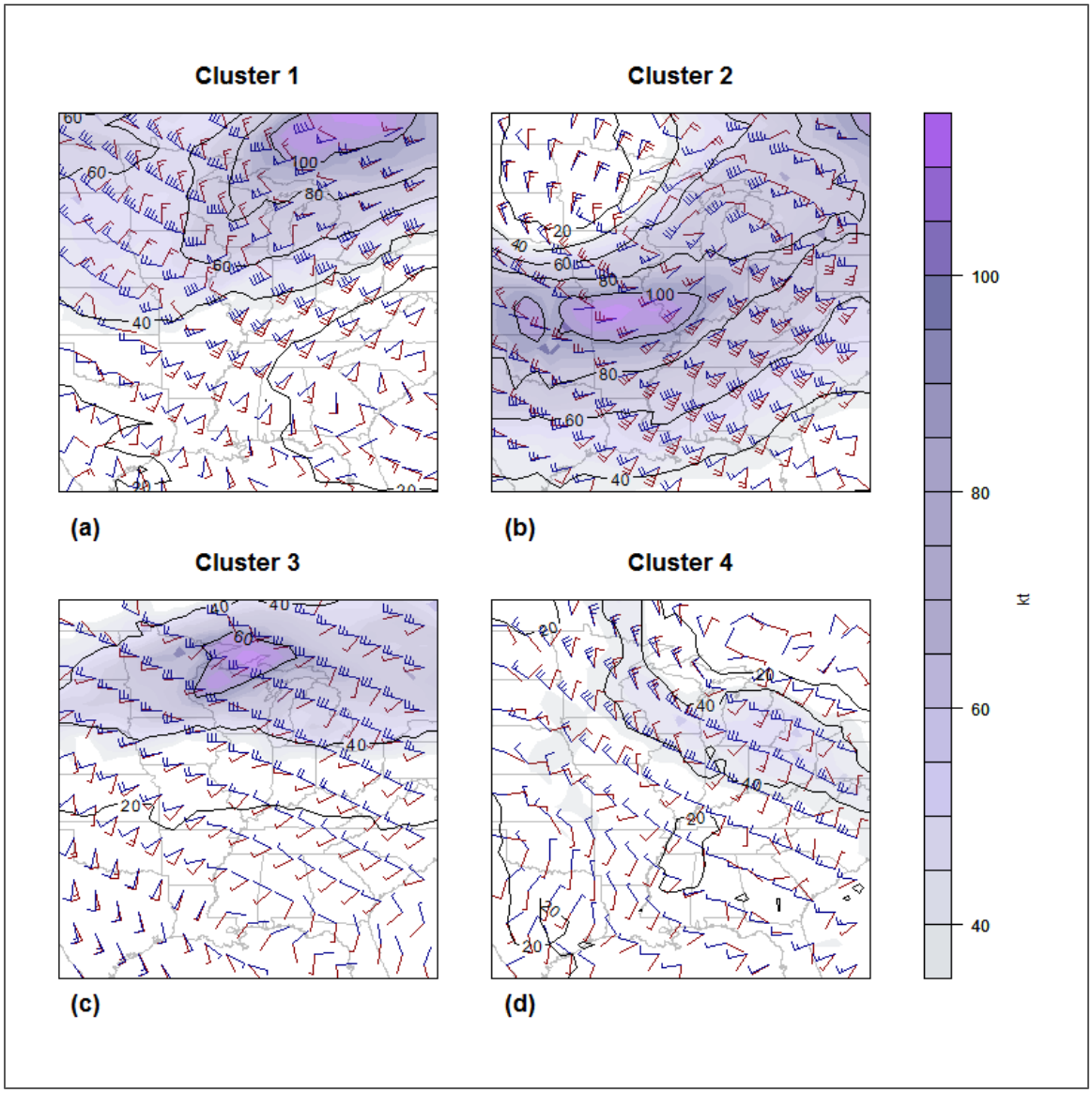

Figure 4.10 Composites of 300 mb Wind Magnitude, 500 mb Magnitude-Direction, and 925 mb Magnitude-Direction in Nontornadic Clusters, Modeled with Thompson Microphysics and YSU PBL

500 mb wind barbs are blue; 925 mb wind barbs are red. Wind speeds are in knots. Maps are provided for scale only and do not reflect the geographical location of synoptic features in all the constituent outbreaks.

4.2.2.1 Assimilation Ensemble Modeling Variations

Each nontornadic outbreak was assigned to the same cluster irrespective of model configuration. This prevented the large variations among some of the clusters that were apparent in the tornado outbreaks. It is suspected that the reason that the nontornadic outbreaks were always assigned to the same clusters was that the clusters themselves were markedly different in the types and locations of synoptic features present in each one, whereas the tornado outbreak clusters all indicated a typical spring-season trough of varying amplitude. Nonetheless, differences in model configuration did influence the modeling of these features for the nontornadic outbreaks as well.

Data assimilation generated variations among thermodynamic and wind field composites for all four nontornadic outbreak clusters, but the nature of the variation was different for certain clusters. Clusters 1 and 2 exhibited decreased strength of the trough in the conventional and HIRS-4/conventional runs as compared to the HIRS-4 and control runs. This decrease was apparent in the 500 mb height fields and the 850 mb temperature fields. Cluster 4, however, demonstrated a strengthened trough in the conventional assimilation runs as compared to the runs without this data assimilation. In cluster 3, which was characterized by the north side of a strong ridge, the northeastward extent of this ridge was lowered in the conventional and HIRS-4/conventional runs and the top of the ridge was slightly to the west. The decrease in trough strength for nontornadic clusters 1 and 2 mirrors the tornado outbreak composites for tornadic clusters 1, 3, and 4, in which the same effect occurred with conventional data assimilation. The Thompson microphysics scheme is suspected to be the cause of this in the tornado outbreaks, due to the increase in CAPE that it generated in phase II. As previously noted

in phase II, the Thompson scheme's warming effect was even more pronounced for the nontornadic outbreaks.

Wind fields showed expected patterns among the assimilation composites, given the thermodynamic fields. In clusters 1 and 2, the amplitude of the jet stream was weaker in the conventional and HIRS-4/conventional runs than in the HIRS-4 and control runs. In cluster 3, the northward extent of a jet streak was greater in the HIRS-4 and control runs. In cluster 4, weak upper-level winds characterized the composite field for all runs, but the amplitude of the wave in the upper-level jet is greater in the HIRS-4/conventional and conventional runs. Lower-level winds were less strongly influenced than upper-level winds in all clusters, but the conventional and HIRS-4/conventional runs appeared to exhibit small increases in 500 mb and 925 mb wind magnitudes, especially in the comparatively weakly forced clusters 3 and 4. Cluster 1 also exhibited a slight increase in wind magnitudes at these levels. Interestingly, the most strongly forced nontornadic composite, cluster 2, did not exhibit a readily discernible difference in lower-level wind magnitude among the assimilation runs.

The nontornadic composites reinforced the results of phase I for combination thermodynamic-dynamic covariates, which were that the conventional and HIRS-4/conventional model runs generated lower values of 0-1 km EHI and CAPE x 0-1 km bulk shear for nontornadic outbreaks than did the control and HIRS-4 runs. In addition, the slightly increased values of 0-1 km bulk shear apparent in the phase I nontornadic composites for conventional and HIRS-4/conventional runs may be due to the slight increase of lower-level shear found in phase III in the weakly forced nontornadic events, even though this increase was not apparent in the strongly forced cluster 2. However, the

strength of the trough in cluster 2 at 500 mb was decreased in the conventional and conventional/HIRS-4 runs. Again in reference to the observations test conducted in phase I, which found that conventional data assimilation brought atmospheric parameters closer to observed values, these results indicate that assimilation of these data aids the WRF model in simulating tornadic and nontornadic environments distinctly.

4.2.2.2 Parameterization Ensemble Modeling Variations

PBL physics parameterization produced significant differences for some nontornadic clusters and variables. In clusters 3 and 4, MYJ PBL physics (Fig. 4.11) produced a much smaller area of 295 K temperatures at 850 mb than the other two PBL physics. This effect was observed very strongly for all microphysics options for these clusters. The effect was also apparent in clusters 1 and 2, though it was not observed as strongly. These clusters were characterized by a broad trough and a mid-latitude cyclone pattern respectively, whereas clusters 3 and 4 were characterized by westerly flow over the top of a ridge and northwesterly flow around the western edge of an upper-level low. In cluster 2, the deepest region of the trough extended slightly farther west with MYJ PBL physics than ACM2 or YSU PBL physics. These results support the previously mentioned cold bias of the MYJ scheme (relative to other PBL schemes) in convective environments. The results of phase II implied such a bias for this scheme at the mesoscale and the lowest layers of the atmosphere, and the results for this phase confirm it on the synoptic scale.

In clusters 3 and 4, YSU University PBL physics (Fig. 4.9) produced a slightly larger area of 295 K temperatures than ACM2 PBL physics (Fig. 4.12). In cluster 2, YSU physics produced a larger area of 280 K and 275 K temperatures than ACM2

physics, but little difference was apparent between these two PBL schemes in cluster 1. 700 mb moisture fields in these clusters exhibit noticeable differences among the PBL physics schemes. In clusters 2, 3, and 4 especially, the RH fields are significantly moister in the ACM2 runs, while the MYJ runs are the driest.

In phase II of the research, the YSU PBL physics model runs generated values of CAPE x 0-1 km bulk shear that were higher than ACM2 physics runs and lower than the MYJ runs for nontornadic outbreaks. LCL heights among the YSU nontornadic runs were higher than those of the MYJ runs and lower than those of the ACM2 runs. The phase III results indicate that moisture levels, rather than temperatures, are likely the primary physical cause of the phase II thermodynamic covariate differences among the PBL runs. Dry air at 700 mb is usually associated with higher values of CAPE in convective environments.

The wind fields depict expected patterns among the PBL physics, given the patterns observed in pressure and thermodynamic fields. In clusters 1 and 3, the jet streak is smaller in size with MYJ physics (Fig. 4.14) than with the other two PBL physics. This feature could account for the more northerly extent of warm 850 mb temperature fields with YSU and ACM2 PBL physics. In cluster 3, the 925 mb winds immediately due south of the jet streak on the north side of the ridge were west-northwesterly with ACM2 and YSU PBL physics, whereas these winds were west-southwesterly with MYJ PBL physics.

Microphysics parameterization did not consistently influence the modeling of most features, comparably to the tornado outbreak clusters. However, in cluster 4, the 265 K isotherm at 850 mb extended farther northeast with Morrison and Thompson (Fig.

4.9, Fig. 4.11) microphysics than Goddard (Fig. 4.13), WDM-6, or WSM-6. This pattern was observed with all PBL physics parameterizations.

Composites of selected physics parameterization fields are shown in figures 4.11, 4.12, 4.13, and 4.14. In all composites, geographical maps are provided for scale only and do not reflect the geographical location of synoptic features in all the constituent outbreaks. In the thermodynamic maps, 700 mb RH is shown in green and 850 mb temperature is shown in colored isotherms. In the wind map, 500 mb wind barbs are blue; 925 mb wind barbs are red. Wind speeds are in knots.

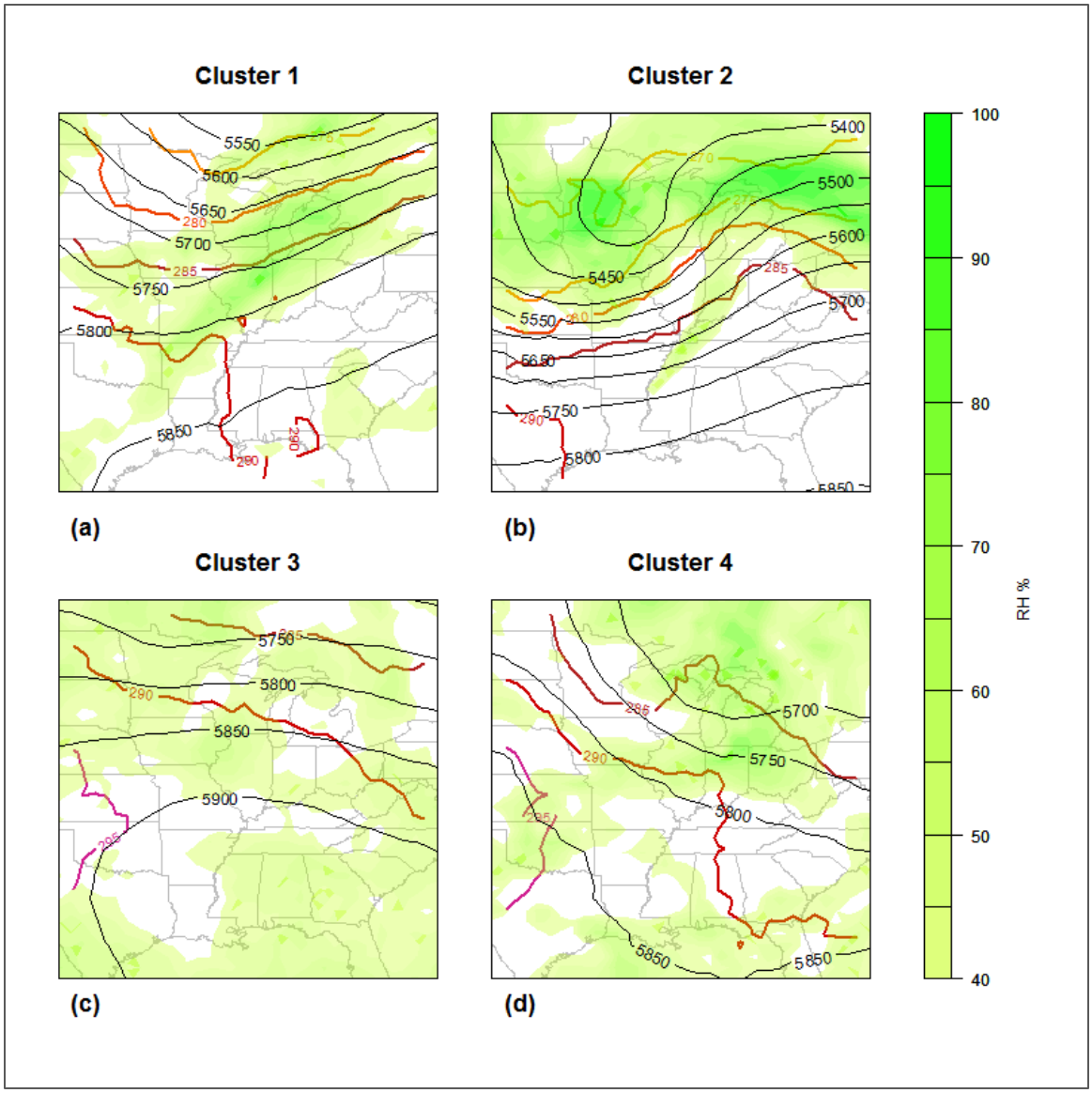

Figure 4.11 Composites of 500 mb Height, 700 mb RH, and 850 mb Temperature in Nontornadic Clusters, Modeled with Thompson Microphysics and MYJ PBL

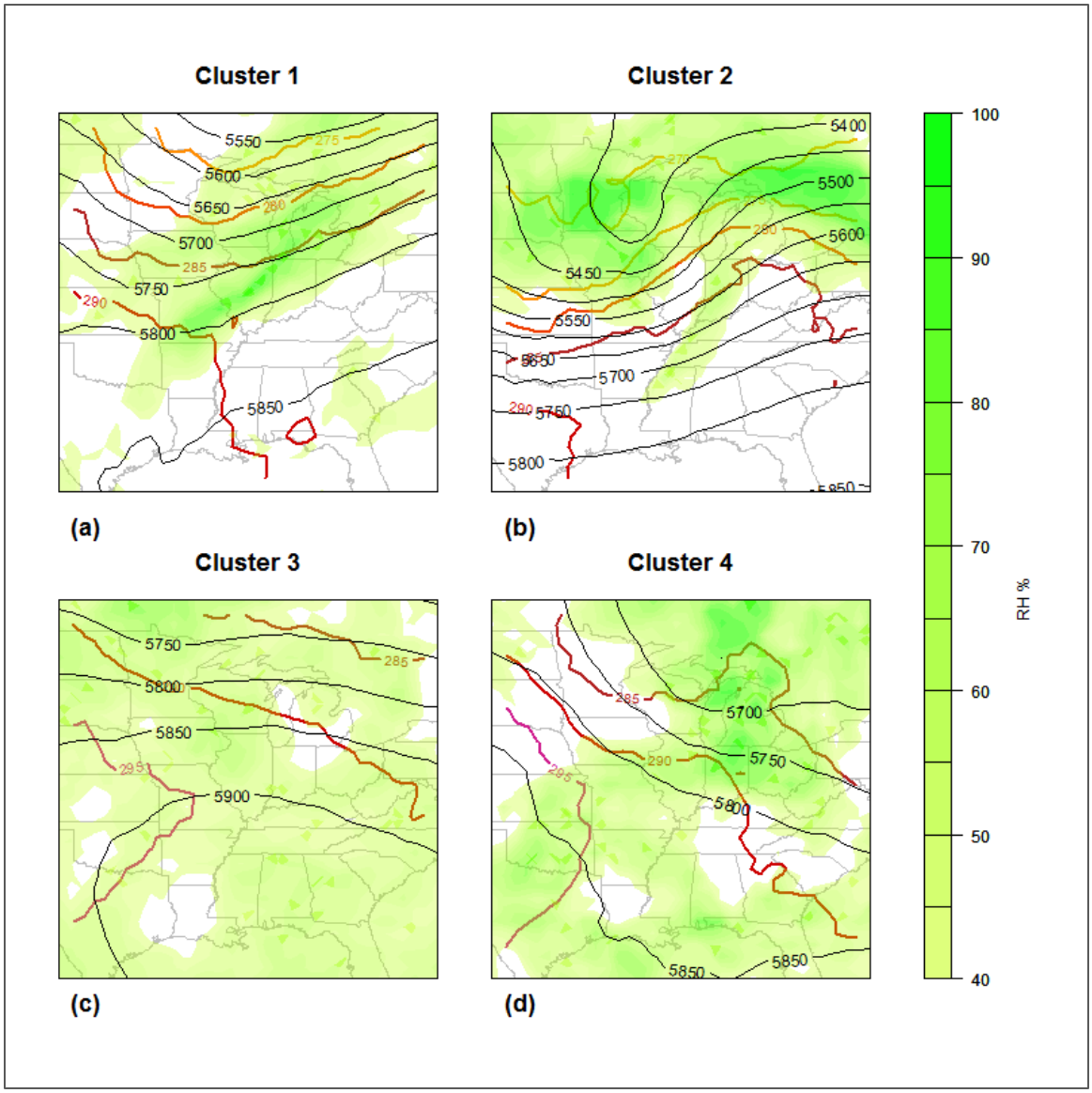

Figure 4.12 Composites of 500 mb Height, 700 mb RH, and 850 mb Temperature in Nontornadic Clusters, Modeled with Thompson Microphysics and ACM2 PBL

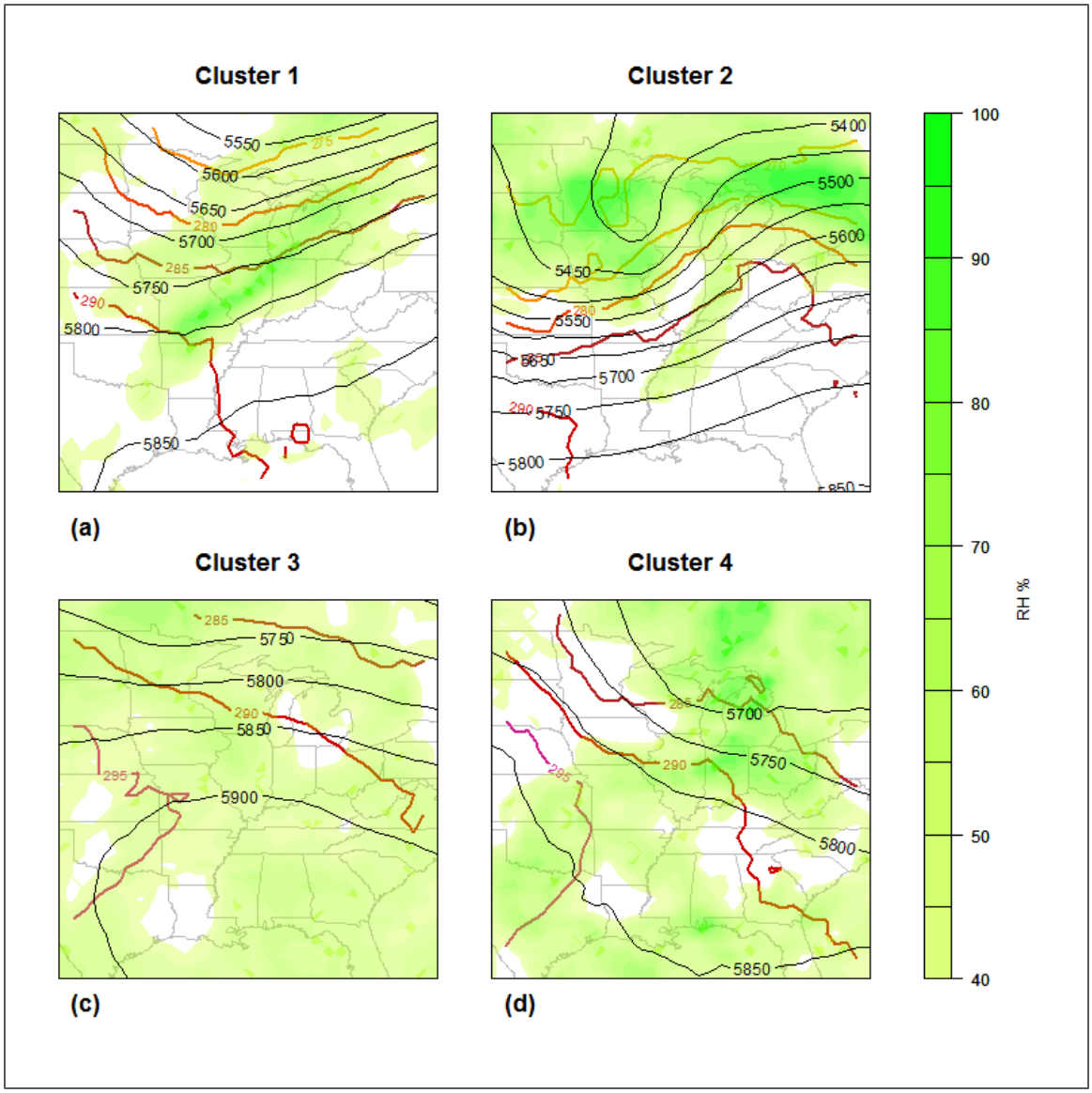

Figure 4.13 Composites of 500 mb Height, 700 mb RH, and 850 mb Temperature in Nontornadic Clusters, Modeled with Goddard Microphysics and ACM2 PBL

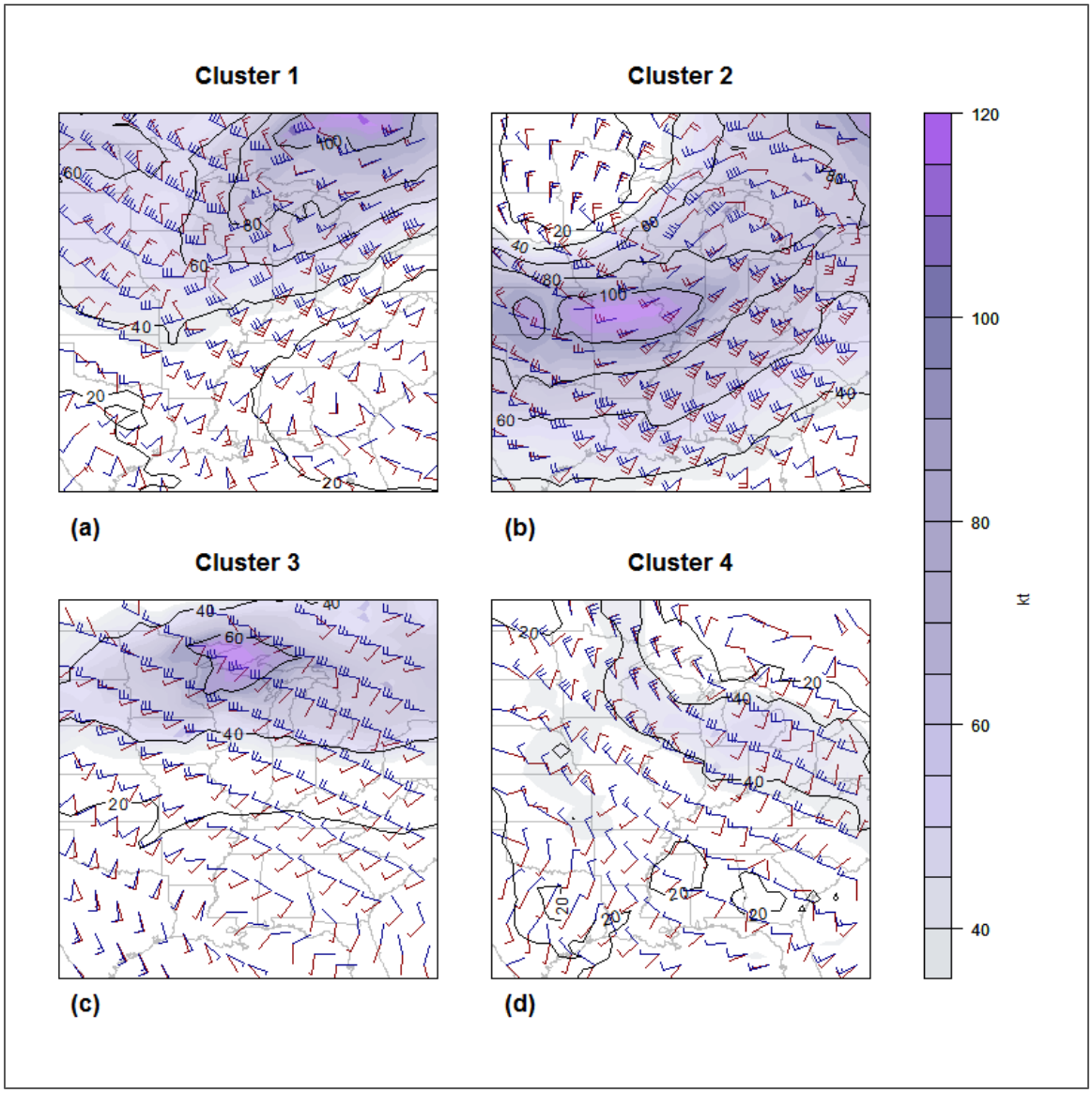

Figure 4.14 Composites of 300 mb Wind Magnitude, 500 mb Magnitude-Direction, and 925 mb Magnitude-Direction in Nontornadic Clusters, Modeled with WSM-6 Microphysics and MYJ PBL

4.3 Discussion

Data assimilation, especially of conventional observations, produced a consistent effect in tornadic and nontornadic outbreak environments that featured strong troughs: slightly weakening these troughs and decreasing the northern extent of warm areas at 850 mb. In the nontornadic cluster that featured a weak trough and northwesterly flow, this trough was strengthened by data assimilation. In the nontornadic cluster that featured a ridge, this meteorological feature was very slightly edged west by data assimilation. The effect of data assimilation on 700 mb moisture fields was largely to move the areas of high moisture slightly east, rather than any observable increase or decrease.

Despite weakening the strong ridges and the northern extent of the warm sectors, data assimilation increased the magnitudes of upper-level winds and lower-level wind shear in tornado outbreaks. In severe weather outbreaks, high-shear environments and intense jet streaks do not always correspond with extremely warm temperatures in the warm sector, especially in late winter and early spring. High shear can produce tornadic environments even when instability is comparatively low. Notably, this effect did not occur in the two "moderate-strong trough" clusters of nontornadic outbreaks. Although instability was decreased in these clusters with data assimilation, wind values were also decreased. These results suggest that the WRF model, in this configuration, simulated environments that were too unstable in trough environments, while at the same time not translating the strength of the trough to high enough wind shear values in tornado outbreaks. Further data assimilation research should be done to determine if this is a problem with the physics suite used in this phase (particularly the Thompson microphysics), the NARR data, or if the WRF model itself exhibits this behavior. Data

assimilation corrected the model background fields to more accurately represent the high-shear environments in tornado outbreaks.

Effects of physics parameterization on the synoptic clusters were less obvious than effects of data assimilation, especially consistent effects on the wind fields of the tornado outbreaks. The WSM-6 and WDM-6 microphysics resulted in a deeper trough in one of the tornadic clusters, but not the other clusters. Although this may seem to suggest that the issue of low wind magnitudes in tornado outbreaks is caused by the WRF or the NARR data, a direct comparison between the physics ensemble and the assimilation ensemble is not possible due to the use of different land surface physics.

Microphysics did not consistently influence the modeling of the nontornadic clusters either. However, PBL physics did affect these clusters. The MYJ PBL physics generated a cold bias in all nontornadic clusters at 850 mb, confirming similar results in phase II of the work as well as previous research showing such a bias with this scheme. YSU physics produced warmer temperatures than the other PBL physics in three of the four nontornadic clusters, as well as drier 700 mb moisture fields. This result is consistent with the vertical mixing mechanism of this scheme. The 700 mb layer of the atmosphere is usually above the PBL, especially in the eastern United States, where most of these outbreaks occurred. A PBL physics scheme such as ACM2, which uses both local and nonlocal mixing in the vertical, was shown in phase II to produce high LCL heights relative to the other model runs, likely due to excessive vertical mixing. This would result in the boundary layer itself being drier than with other PBL schemes, but its vertical extent would be higher and the vertical gradient of moist to dry air—and therefore convective instability—would be weaker with this very intense vertical mixing.

CHAPTER V

SUMMARY AND CONCLUSIONS

The purpose of this research was to determine the impact of data assimilation, microphysics, and PBL physics on the modeling of important features in tornadic and nontornadic severe weather outbreaks. The research had three phases:

1. A phase I to determine the effect of conventional observation and HIRS-4 satellite radiance assimilation on especially significant outbreak-discriminating atmospheric covariates and a support vector machine (SVM) outbreak type forecast;
2. A phase II to determine the effect of cloud microphysics and planetary boundary layer (PBL) physics parameterization on these same covariates and SVM forecasts;
3. A phase III to examine the effects of assimilation and physics parameterization on the modeling of synoptic weather features.

In phase I, conventional meteorological observations and HIRS-4 satellite radiance observations were assimilated by the Gridpoint Statistical Interpolation (GSI) software to a WRF domain for each of a set of 20 tornadic and 20 nontornadic outbreaks. These observations were assimilated individually and in combination with each other. The WRF model was then run again at a higher resolution with the raw assimilated fields as initial background fields, with model domains centered on each outbreak. A control

run without data assimilation performed was also created for comparisons. Seven covariates found by previous research to be important in distinguishing between tornadic and nontornadic outbreaks—0-1 km storm relative helicity (SRH), 0-3 km SRH, CIN, 0-1 km bulk shear, CAPE x 0-1 km bulk shear, LCL height, and 0-1 km Energy-Helicity Index (EHI)—were extracted from the high-resolution model output. SVMs were trained on a subset of RPCA scores derived from the raw covariate data, and tested on the remaining scores. Contingency statistics were computed for the forecasts of each SVM configuration to determine the highest skilled SVM for every model run (three data assimilation types and one control). These highest-performing SVMs were then compared against each other to determine which form of data assimilation produced the most accurate outbreak mode forecasts. Bootstrap mean Euclidean distances between tornadic and nontornadic covariate fields, and bootstrap mean composite fields of each covariate for tornadic and nontornadic outbreaks, were calculated as well to determine the impact of data assimilation on each of these covariates and outbreak types.

In phase II, the same 20 tornadic and 20 nontornadic outbreaks were modeled with cloud microphysics and PBL physics parameterizations varied. Five microphysics—Goddard, Morrison, Thompson, WRF Double Moment-6 class (WDM-6), and WRF Single Moment-6 class (WSM-6)—were used, and three PBL physics—Asymmetric Convection Model (ACM2), Mellor-Yamada-Janjić (MYJ), and Yonsei University (YSU)—were used, for a total of 15 model runs per case. The WRF model was run in a nested domain configuration with outer domains identical to those of phase I and inner nested domains at the same resolution and very similar spatial dimensions to those of phase I. The seven outbreak covariates were computed from the inner nest and

processed through an RPCA, with the same subset of resulting scores as in phase I used to train a group of SVMs. The SVMs were then tested on the remaining subset, as in phase I. Contingency scores were calculated for each SVM configuration, and the highest-skilled ones for each of the 15 model physics configurations were then compared to each other to determine the combination of microphysics and PBL physics that produced the highest skill at outbreak mode prediction. Euclidean distances between tornadic and nontornadic outbreaks for each covariate, and outbreak mean composites of each covariate, were computed as in phase I. For this phase, domain-mean analyses were performed on tornadic and nontornadic fields of each covariate for each microphysics and PBL physics scheme, with the goal of averaging out the effects from the other type of model physics and determining covariate value ranges among PBL physics and microphysics means.

In phase III, the same 20 tornadic and 20 nontornadic outbreaks were simulated in the WRF model for all model variations in the research: the 4 data assimilation runs and the 15 model physics parameterization runs. The model was run with a domain identical to the parent domain of phases I and II. Synoptic-scale fields were extracted from the model output at many atmospheric levels from 1000 mb to 100 mb, and *k*-means clustering analysis was performed on the tornadic and nontornadic outbreak fields separately to prevent the generation of composite mean fields that averaged two or more dissimilar synoptic setups together. The resulting cluster mean fields were then

examined visually for differences among the data assimilation and physics parameterization runs.

5.1 SVM Outbreak Mode Forecasting

Data assimilation and physics parameterization produced clear winners in SVM forecast skill. Data assimilation in particular produced a progression of forecast skill levels that was experimentally expected and which supports the idea that additional amounts and types of data, when ingested well by assimilation procedures, improve forecasting skill. With physics parameterization, a clear difference was noted between PBL physics and microphysics with regard to effects on forecast skill for outbreak mode classification. The seven covariates were mostly computed in the lower atmosphere, and PBL physics parameterization more significantly influenced the modeling of these covariates and thus the skill of the SVMs trained on the model data.

5.1.1 Data Assimilation

Phase I determined that, of the three data assimilation runs and the control, the run that produced the highest-skilled SVM at predicting outbreak mode was the HIRS-4/conventional run. This run contained the greatest amount of input data of all the runs for this phase. The second-highest-skilled SVM was the one for the conventional assimilation run, which contained the second-most observations. The third-best performer was the HIRS-4 run. The control run, which contained no assimilated data, performed the most poorly in terms of forecast skill. Each difference of skill between a given pair of runs was statistically significant.

Root mean square error comparison of each assimilation run to a set of observations demonstrated that the conventional data assimilation in particular markedly improved the WRF simulations. These observations consisted of geopotential height, specific humidity, temperature, *u*-wind, and *v*-wind measured at standard meteorological analysis levels for each parameter. In the conventional and HIRS-4/conventional runs, error was decreased for 80 percent of data points, examined over all levels of the atmosphere, relative to the control. These runs also exhibited lower model error over time than the control and HIRS-4 runs, demonstrating that assimilation of the conventional data was reducing cumulative model error relative to a control run. These results indicate that the improved skill of the conventional and HIRS-4/conventional runs in predicting severe weather outbreak mode can be attributed to better modeling of atmospheric features, particularly troughs and low-level winds, with the additional data.

Tornadic-nontornadic Euclidean distance calculations were maximized for 0-1 km bulk shear by the HIRS-4/conventional run. This run and the conventional run maximized the distances for the CAPE x 0-1 km bulk shear product and CIN. The conventional assimilation run maximized the distance for 0-1 km EHI. For LCL, the control generated the maximum distance between tornadic and nontornadic outbreak modes. The Euclidean distance for SRH at 0-3 km was decreased from the control distance by the assimilation of conventional observations, 4.79% for the conventional-only run and 4.33% for the HIRS-4/conventional run. The HIRS-4-only run produced the maximum 0-1 km SRH distance; the conventional-only run decreased 0-1 km SRH distance from the HIRS-4 distance by 1.88% and the HIRS-4/conventional run by 1.16%. Compared to the control distance for 0-1 km SRH, the conventional-only run

decreased the distance by only 0.7% and the HIRS-4/conventional run produced an almost identical distance.

The general pattern of the data assimilation effect on tornadic-nontornadic Euclidean distance was that conventional observations, either alone or in combination with HIRS-4 observations, increased outbreak distances for thermodynamic and dual thermodynamic-dynamic covariates. The assimilation of these observations produced the opposite effect on SRH, though not bulk shear. Examination of outbreak-mean composite fields for 0-3 km SRH, which demonstrated the same pattern as 0-1 km SRH, found that conventional assimilation increased the higher values of SRH in tornado outbreaks while decreasing the magnitudes of this variable elsewhere in the domain. In the nontornadic composite, values were decreased by conventional data assimilation throughout the domain. The result of conventional data assimilation, alone or with HIRS-4 radiances, was to produce a very strong similarity between tornadic and nontornadic outbreaks in the location of high and low SRH values, decreasing the statistical distance between the two outbreak types. Nonetheless, SRH has been found by previous researchers to be less predictive of tornadic environments and the formation of supercells than bulk shear (Rasmussen and Blanchard 1998), so operational forecasters should not use it in isolation to predict outbreaks.

The finding that conventional data assimilation produces more accurate model output, while at the same time decreasing the modeled difference between tornadic and nontornadic outbreak SRH, may be an undesirable outcome for operational forecasters. However, deterministic forecast tools such as the SVMs used in this research are often able to compute probabilities with high degrees of skill using several different covariates.

In addition, conventional data assimilation in tornado outbreaks did increase SRH values in domain regions where values were already high, a result that did not occur for nontornadic outbreaks. Finally, the decreased false alarm ratio and increased skill of the SVMs that used conventional and HIRS-4 model runs indicate that such tools may provide a useful form of guidance for those in operational forecasting.

5.1.2 Physics Parameterization

Phase II determined that the WRF double-moment 6-class microphysics scheme (WDM-6) and YSU PBL physics produced the highest outbreak-mode forecasting skill of any microphysics-PBL physics combination, with an HSS of 0.658. This result was significantly greater than the skill of any other run at the 95% confidence level. This parameterization run produced this result by generating the lowest FAR and highest POD. Other strong performers were Goddard/YSU, Morrison/YSU, and WSM-6/YSU.

The most dominant theme of phase II was that PBL physics choice appeared to have a much greater impact than microphysics on all aspects of forecasting observed in this research. This finding was affected by the fact that the seven highly significant covariates for discriminating between outbreak mode were all lower-atmospheric, measured either within or immediately above the PBL. Variables such as LCL height, shear, helicity, and combined parameters that utilize these fields are especially affected by mixing in the PBL. Although all PBL schemes produced acceptable levels of skill in outbreak-mode forecasts, the YSU PBL scheme, a commonly used parameterization for both general modeling and severe convective weather, generated high forecast discrimination skill with any microphysics choice examined. MYJ PBL physics performed at the lowest level of skill, with the revised Asymmetric Convection Model

PBL physics performing in the middle of the group of three PBL physics parameterizations.

Euclidean distance calculations found that the variables that exhibited more noticeable sensitivity to microphysics were thermodynamic covariates. Cloud microphysics parameterization influences the way that the model handles water and ice processes, leading directly to evaporation and condensation effects in the atmosphere, as well as effects on cloud thickness and solar heating of the surface. These processes strongly influence heating and cooling at all levels of the atmosphere. Thompson microphysics in particular strongly influenced the modeling of CIN, 0-1 km EHI, and CAPE * 0-1 km bulk shear in tornadic and nontornadic outbreaks. This finding coincides with previous studies showing a warm bias with the scheme.

PBL physics, however, again produced larger differences between tornadic and nontornadic outbreaks for the remaining covariates. ACM2 physics generated small distances between outbreak types for shear and helicity covariates; MYJ physics generated the largest distances for these covariates, though this notably did not increase the skill of MYJ runs to the level of most YSU runs. It is not surprising that boundary layer physics parameterization has a greater impact on wind patterns at the 0-1 and 0-3 km layers, than microphysics, but these findings empirically confirm it. A possible cause of the differences in wind shear covariate modeling among PBL schemes is the distinct way that each scheme handles vertical transport of air parcels through the boundary layer. The MYJ scheme uses only local closure and the YSU scheme is nonlocal, but the ACM2 PBL scheme utilizes both local and nonlocal mixing. It is possible that by employing mixing both within PBL layers and across them, the ACM2 scheme minimized the wind

shear effects through friction. Tentative support for this hypothesis exists in the nontornadic LCL composites, which show the highest values—and thus the greatest amount of vertical mixing in a convective PBL—in the ACM2 runs. Higher LCL heights (relative to both YSU and MYJ composites) in the prefrontal sectors of the ACM2 tornadic composites suggest that this pattern may be occurring in tornado outbreak cases as well as the nontornadic outbreaks. The MYJ scheme may have produced higher shear and SRH values than the nonlocal YSU scheme because of its known cold bias in convective environments. A shallower PBL will have lower amounts of friction and turbulence of PBL origin, permitting dynamically driven winds to exert a stronger influence. Another possible source of higher shear and SRH values in MYJ model runs is the local closure of this scheme. The scheme performs mixing within individual vertical layers and does not mix across this dimension. Explicitly modeling vertical transport across layers may decrease horizontal wind magnitudes within some layers. Further research into this topic should be conducted.

The covariate that appeared to have the greatest impact on SVM outbreak-mode forecasting skill was LCL height. The highest-skilled run, WDM-6/YSU, generated the maximum outbreak-mode distance for this covariate. The results of this study concerning PBL physics and LCL height are interesting, and offer an added opportunity to examine the implications of this research in relation to other studies on PBL physics described in the literature review. The finding that MYJ physics produces a diffuse, somewhat broken gradient of LCL heights for tornado outbreaks—different from the other two PBL physics, which produced a sharp gradient indicative of a cold front—as well as a shallow PBL in tornadic and nontornadic outbreaks supports the findings of Hu et al. (2010), who

found cold and moist biases in the modeling of the boundary layer with this parameterization. The YSU LCL patterns also support the work of Hong et al. (2006), who found that YSU PBL physics modeled a cold front in an autumn tornado outbreak strongly and in closer concurrence with observations than a control PBL scheme.

The role of microphysics in LCL height simulation was less apparent than that of PBL physics, but a comparison of the WDM-6 and WSM-6 runs shows that WDM-6 produced higher LCL heights than WSM-6 across tornado outbreak domains, both in front of and behind a boundary that appeared in most of the tornadic cases. WDM-6 has been shown in previous work (e.g., Adams-Selin et al.) to produce strong cold pools behind thunderstorms. In the prefrontal region, the slightly higher LCL in WDM-6 than WSM-6 may be caused by evaporative cooling and drying due to smaller rain and cloud droplet sizes in the double-moment scheme.

Examination of observed data for tornado outbreaks indicates that the boundary in the tornadic LCL composites was probably the result of thunderstorm passage. This boundary was not present in the nontornadic LCL composites to nearly the same degree, and it is likely that the SVMs—especially the WDM-6/YSU—used this distinction between outbreak types heavily in their classification functions. If so, this indicates that the SVM is identifying an artifact of tornadic outbreak occurrence after the outbreak has already happened for part of the model domain. However, this research did not examine SVM forecasts with NWP model output from before the outbreak valid time, so the possibility that the SVM might identify such a pattern as highly significant was "built in" to the experimental design. Within an NWP framework, this possibility should not be a concern for operational forecasters, who use model output valid for times before and

during the peak of a severe weather outbreak to issue a severe weather forecast. The SVM outbreak prediction results still indicate that the NWP model is simulating the type of outbreak well.

5.2 Synoptic Variable Modeling

The findings of Phase III were not as starkly clear as those of phases I and II, though some phase III results did reinforce the results of phases I and II. The synoptic features of the tornado outbreak composites were influenced in part by the distribution of outbreaks in each of four *k*-means clusters, a result that did not occur for the nontornadic outbreaks. Rather than model differences, it is strongly suspected that the reason for this result is that a meteorologically similar synoptic pattern occurred during all the tornado outbreaks, whereas four very distinct synoptic patterns, likely caused by seasonal biases for certain types of atmospheric forcing, produced the nontornadic events. However, some conclusions about the synoptic-scale modeling effects of data assimilation and physics parameterization can nevertheless be drawn.

5.2.1 Data Assimilation

In the tornado outbreaks, troughs were weakened slightly in the conventional and HIRS-4/conventional runs as compared to the HIRS-4 and control runs, except in cluster 2. In this cluster, which contained several high-impact events, warm isotherms at the 850 mb level extended farther north, indicating a stronger warm sector in the models. A possible explanation for the weaker troughs in the conventional and HIRS-4/conventional runs in the other three clusters is the use of the Thompson microphysics in the assimilation ensemble. This scheme is known to produce warm biases and in phase II

increased the values of CAPE-related covariates above other microphysics parameterizations.

Upper-level winds in the modeled jet streams were also more intense for the tornadic clusters with conventional and HIRS-4/conventional runs. The HIRS-4 run also generated stronger upper-level winds in two tornado outbreak clusters than the control, though this result did not generate a visibly stronger trough at lower levels. Conventional data are acquired from all levels of the atmosphere, so in the absence of other experimental tests, it is difficult to say if this increase in trough intensity in the conventional and HIRS-4/conventional runs is due to the assimilation of surface or of upper-level data. Satellite radiances, however, are interpreted by the assimilation procedure to indicate moisture fields, most typically clouds, so it is highly likely that the upper-level winds in HIRS-4 were due to the assimilation of data at upper levels of the atmosphere.

The wind results of this phase reinforce the results of phase I, in which tornadic outbreak composites with conventional data assimilated (both the conventional-only and the HIRS-4/conventional runs) had notably higher values of 0-1 km bulk shear, 0-1 km EHI, and CAPE x 0-1 km bulk shear than the control and the HIRS-4 composites. The phase III results indicate that assimilation of conventional observations increases wind magnitudes above the near-surface layer of the atmosphere and at a synoptic scale. The observations test of phase I, which found that conventional and HIRS-4/conventional runs exhibited lower RMSE, was performed on standard atmospheric variables (geopotential height at 500 mb; *u*-winds at 925, 850, 700, 500, 400, 300, 250, 200, 150, and 100 mb; *v*-wind at the same pressure levels; specific humidity at 700 and 500 mb;

and temperature at 925, 850, 700, and 500 mb). These variables include fields examined in the synoptic analysis of phase III, indicating that conventional data assimilation improves the accuracy of NWP output of these cases in addition to increasing the values of these variables.

In nontornadic outbreaks, the effect of data assimilation on a cluster depended on the type of synoptic setup modeled in that cluster. The two nontornadic clusters that were characterized by classic troughs exhibited decreases in the magnitude of the trough when conventional data were assimilated, either alone or in conjunction with HIRS-4 radiances. The small cluster characterized by a trough pattern suggestive of an upper-level low and weak northwesterly flow, however, had this feature strengthened by conventional and HIRS-4/conventional data assimilation. The "derecho cluster," which was characterized by a ridge, showed a westward shift of this ridge by the assimilation of conventional and conventional/HIRS-4 data, likely indicating a slight placement error by the model.

The nontornadic composites reinforced the results of phase I for combination thermodynamic-dynamic covariates, which were that the conventional and HIRS-4/conventional model runs generated lower values of 0-1 km EHI and CAPE x 0-1 km bulk shear for nontornadic outbreaks than did the control and HIRS-4 runs. In addition, the slightly increased values of 0-1 km bulk shear apparent in the phase I nontornadic composites for conventional and HIRS-4/conventional runs may be due to the slight increase of lower-level shear found in phase III in the weakly forced nontornadic events, even though this increase was not apparent in the strongly forced cluster 2, which most closely resembled a tornado outbreak cluster.

5.2.2 Physics Parameterization

The effects of physics parameterization on cluster modeling were less pronounced than the effects of data assimilation, especially in the tornado outbreaks, but effects were still noticeable within this ensemble. The trough of one tornadic cluster was modeled more intensely with WSM-6 and WDM-6 physics than other microphysics, but this result did not appear consistently in other tornadic clusters for these microphysics options. Wind fields did not exhibit any clear patterns across microphysics or PBL physics that appeared in all four clusters.

For the nontornadic outbreaks, differences among parameterization runs were more readily apparent, at least among PBL physics runs. Microphysics variation did not generate consistently and significantly different synoptic fields for nontornadic outbreak clusters. MYJ PBL physics produced a strong cold bias (compared to the other two PBL physics) in the modeling of clusters 3 and 4, which was observed with all microphysics. This result confirms existing literature that found a cold bias with PBL scheme, as well as the phase II results of this study. A less significant cold bias was observed with this PBL physics option for clusters 1 and 2. YSU PBL physics generated warmer 850 mb temperatures in clusters 3 and 4.

In clusters 3 and 4, YSU PBL physics produced a slightly larger area of 295 K temperatures than ACM2 PBL physics. In cluster 2, YSU physics produced a larger area of 280 K and 275 K temperatures than ACM2 physics, but little difference was apparent between these two PBL schemes in cluster 1. 700 mb moisture fields in these clusters exhibit noticeable differences among the PBL physics schemes. In clusters 2, 3, and 4,

the RH fields are significantly moister in the ACM2 runs, while the MYJ runs are the driest.

In phase II of the research, the YSU PBL physics model runs generated values of CAPE x 0-1 km bulk shear that were higher than ACM2 physics runs and lower than the MYJ runs for nontornadic outbreaks. LCL heights among the YSU nontornadic runs were higher than those of the MYJ runs and lower than those of the ACM2 runs. The phase III results indicate that moisture levels, rather than temperatures, are likely the primary physical cause of the phase II thermodynamic covariate differences among the PBL runs. Dry air at 700 mb is usually associated with higher values of CAPE in convective environments.

5.3 Ensemble Variance Analysis

The first question requires a final statistical analysis on the data to determine the variance within similar groups of model runs. Using the bootstrapped tornadic and nontornadic composite fields generated in phases I and II for the seven critical covariates, variances were calculated for the following groups of model runs:

- Conventional, HIRS-4/conventional, HIRS-4, and assimilation control for an assimilation ensemble variance;
- The three PBL physics variations used with each of the five microphysics schemes (for example, the variance for a covariate over the combined data set of WDM-6/ACM2, WDM-6/MYJ, and WDM-6/YSU), for a set of five measures of variability due to PBL physics;
- The five microphysics variations used with each of the three PBL physics schemes, for a set of three measures of variability due to microphysics.

This analysis was conducted for each of the seven covariates for tornadic and nontornadic outbreaks separately. The variance groups for the five microphysics schemes (variability due to PBL physics) and the three PBL schemes (variability due to microphysics) were averaged for each covariate for tornadic and nontornadic outbreaks. Table 5.1 shows the variances for each covariate for tornadic outbreaks and Table 5.2 shows the variances for each covariate for nontornadic outbreaks, with the ensemble-type averages for each covariate in bold.

As the tables show, tornado outbreak variances were largest for the assimilation ensemble for every covariate except LCL height and 0-1 km EHI, for which the PBL mean variance was the largest. PBL physics produced the second-largest variance for all of the other variables except for CIN, which had a variance that was slightly greater for the microphysics mean than for the PBL mean. The assimilation ensemble produced variances in the dynamic fields (bulk shear and SRH) that were especially large compared to either the microphysics or PBL physics mean.

In the nontornadic outbreaks, a different and in some ways opposite pattern was observed. The PBL ensemble variance mean was largest for 0-1 km SRH, 0-3 km SRH, CAPE x 0-1 km bulk shear, and 0-1 km EHI, followed by the microphysics variance mean, and the assimilation ensemble variance was smallest. For CIN, the microphysics ensemble variance mean was the largest. However, for bulk shear and LCL height, the assimilation ensemble variance was largest and the microphysics mean variance was smallest.

These results indicate that certain characteristics of outbreak modes influence the type of model uncertainty—observation-based or physics-based—that will be most

prominent in a given NWP forecast. Tornado outbreaks are associated with higher values of low-level shear than nontornadic outbreaks, and, as the phase I tornadic composites of 0-1 km bulk shear and 0-3 km SRH illustrated, data assimilation—particularly of conventional observations—greatly affected these values, sharply increasing their magnitudes relative to the control run in regions of strong shear or helicity. Conventional data are recorded from many sources across the vertical dimension of the atmosphere, which significantly increases spatial coverage. It is likely that the reason these dynamic covariates experienced such a stronger effect from data assimilation in tornado outbreaks is that the typical values of these variables are much higher in tornado outbreaks than nontornadic outbreaks. However, even in nontornadic outbreaks, data assimilation produced a larger ensemble variance for 0-1 km bulk shear than microphysics or PBL physics. This result illustrates the importance of having a spatially dense set of observations when forecasting low-level wind shear.

It is also intuitive that PBL physics would have a greater effect on lower-atmospheric parameters than cloud microphysics. The temperature, humidity, and height of the PBL have a strong influence on both thermodynamic and dynamic variables, due to the effect of mixing. As has been stated, the three PBL schemes examined in this research perform mixing in three distinct ways, which has been shown to lead to warm and dry or cool and wet biases in the PBL.

Microphysics produced the largest variance for only one covariate, nontornadic-outbreak CIN. It also produced variances slightly smaller than the largest ensemble-type variance for tornadic LCL height and nontornadic 0-1 km EHI. These are all variables that are at least in part thermodynamic. Cloud microphysics affects evaporation and

condensation processes in clouds in the model, resulting in variations in the heat and moisture content of the atmosphere through changes in latent heat release, surface heating, precipitation, and many other processes.

Table 5.1 Model Ensemble Variances for Severe Covariates (Tornadic)

	0-1 km SRH	**0-3 km SRH**	**CIN**	**0-1 km Bulk Shear**	**CAPE x 0-1 km Bulk Shear**	**LCL**	**0-1 km EHI**
Assimilation	**5656.373**	**8217.829**	**935.672**	**2.032**	**3132014**	**185125.7**	**0.295**
Goddard (PBL varied)	4157.259	7126.021	250.718	1.076	2877220	239392.0	0.462
Morrison	3384.210	6280.115	198.395	0.916	2990122	242495.5	0.502
Thompson	3155.090	6321.352	204.230	0.845	2733552	250423.9	0.527
WDM-6	3360.033	6269.920	222.872	0.975	2393577	231679.3	0.501
WSM-6	3500.944	6517.727	209.632	1.003	2803307	239282.2	0.514
PBL Mean	**3511.507**	**6503.027**	**217.170**	**0.963**	**2759556**	**240654.6**	**0.501**
ACM2 (microphysics varied)	2129.441	4617.677	200.972	0.770	840933	230841.1	0.247
MYJ	3802.787	7118.151	224.249	0.805	2774320	212710.4	0.600
YSU	2984.251	6508.103	234.121	0.685	1729447	246383.2	0.433
Microphysics Mean	**2972.160**	**6081.310**	**219.781**	**0.754**	**1781567**	**229978.2**	**0.427**

Table 5.2 Model Ensemble Variances for Severe Covariates (Nontornadic)

	0-1 km SRH	0-3 km SRH	CIN	0-1 km Bulk Shear	CAPE x 0-1 km Bulk Shear	LCL	0-1 km EHI
Assimilation	**2251.776**	**1343.953**	**390.745**	**0.726**	**1210707**	**71084.6**	**0.033**
Goddard (PBL varied)	2625.024	2391.550	403.887	0.578	1967885	59624.2	0.117
Morrison	2531.496	2019.281	441.764	0.550	1889523	65814.8	0.106
Thompson	2375.027	2180.465	585.954	0.589	2449881	62427.7	0.152
WDM-6	2448.140	2012.615	379.913	0.611	1721382	61330.0	0.114
WSM-6	2632.119	2078.058	379.111	0.640	2084309	63696.3	0.115
PBL Mean	**2522.361**	**2136.394**	**438.126**	**0.594**	**2022596**	**62578.6**	**0.121**
ACM2 (microphysics varied)	1744.174	1727.964	391.184	0.495	1223403	52917.4	0.105
MYJ	2808.247	2198.095	541.546	0.497	2535675	61181.2	0.143
YSU	2754.273	2333.139	477.086	0.549	1718236	53004.5	0.111
Microphysics Mean	**2435.565**	**2086.399**	**469.939**	**0.513**	**1825771**	**55701.0**	**0.120**

5.4 Synthesis of Results

The two main research questions that this project sought to answer were which model configuration in the data assimilation and physics parameterization ensembles produces the most accurate outbreak forecast, and whether model physics parameterization or data assimilation introduces produces greater model spread in atmospheric parameters that are important for severe weather. The concluding answers to these questions are discussed.

5.4.1 Forecast Skill

As a preface, it should be noted again that land surface physics in phase II were different from those of phase I. In phase I, 5-layer thermal diffusion physics were used, whereas in phase II, Noah land surface physics were used. This was due to the fact that certain experimental physics parameterizations required the use of Noah land surface physics. This difference in model configuration between phase I and II does not present a problem for individual phase results analysis since the model configurations were the same by phase of the work, but it does limit the ability of the data assimilation results to be directly compared to the physics parameterization results.

It was determined that, among data assimilation runs, the HIRS-4/conventional assimilation run generated the most skilled forecasts of severe weather outbreak mode, and among physics parameterization runs, the WDM-6/YSU run had the highest skill. The reasons for why these runs were the most skilled were complex. Analysis of covariate fields, outbreak-mode statistical distances, and observation error analysis indicated that the assimilation of HIRS-4 and conventional data brought model output

closer to an observed reality while, at the same time, differentiating between tornadic and nontornadic outbreak parameters more strongly than other model runs.

In the case of the physics parameterization runs, the WDM-6/YSU run performed with the highest skill, apparently by modeling differences reasonably strongly for all seven analyzed covariates, whereas other physics ensemble runs failed to do so for at least one variable. In addition, it seems likely that the SVM for this run heavily weighted LCL height, which exhibited a very distinct appearance between nontornadic and tornadic outbreak composites. The tornadic LCL fields in particular assumed the appearance of a sharp boundary, and examination of surface analyses showed that in each tornadic case, either a dryline, a synoptic cold front, or a line of severe thunderstorms capable of producing a strong cold pool passed over the outbreak domain at valid time. The WDM-6 scheme was shown in previous research to produce stronger convection and stronger cold pools than other schemes due to its handling of graupel processes, and the YSU was shown to produce more intense fronts in a tornado outbreak case.

5.4.2 Model Variability

The other research question required three phases of research to answer. In tornado outbreaks, data assimilation, especially of conventional observations, creates the greatest amount of variation in model output, followed by PBL physics and microphysics respectively. However, for the nontornadic outbreaks, PBL physics parameterization produces the greatest degree of model variation.

For data assimilation, this finding was apparent at both the mesoscale (phase I) and synoptic-scale (phase III) levels of atmospheric analysis. The variance analysis indicates that in tornado outbreaks, wind shear covariates are very strongly affected by

data assimilation, most notably in the form of increases of shear values in these events. Conventional data assimilation appears to improve the modeling of weather systems of all types, as evinced in phase I by the superior skill of these runs. In phase III, it was possible to determine the effect of data assimilation on large-scale features that produce severe weather outbreaks. Wind shear was increased in tornado outbreaks by the assimilation of this data set and decreased in nontornadic outbreaks. More unusual synoptic features associated with severe weather, such as a northwesterly-flow regime, were strengthened by the assimilation of conventional data, but since this flow pattern was not apparent in tornado outbreaks, the strengthening of this feature likely only served to reinforce the difference between tornadic and nontornadic outbreaks to the SVMs in phase I.

The consistent theme of physics parameterization analyses was that PBL physics had a strong impact on atmospheric covariates associated with severe weather, especially shear and helicity, whereas microphysics as a whole did not introduce as much variation to any covariate. It was surprising that the PBL physics result was less apparent in the phase III synoptic composite analysis. However, this effect did exist to a degree in the composites. The cold bias of the MYJ scheme in particular was confirmed to exist at both the mesoscale (phase II) and synoptic scale (phase III). YSU physics also generated slightly warmed 850 mb temperatures in nontornadic outbreak composites, but the most pronounced difference was in the moisture fields. The ACM2 PBL scheme produced the most moisture at 700 mb in the nontornadic composites and the MYJ the least, which likely accounts for the higher CAPE values apparent in the phase II MYJ clusters. The less pronounced (though still identifiable) synoptic-scale differences among model runs

indicate that the effects of PBL physics become most apparent at the mesoscale—and possibly microscale—levels of analysis rather than the synoptic. Given the highly localized, often chaotic, nature of boundary layer phenomena, this result is not surprising.

Modelers and operational forecasters can derive useful information from this research. It is increasingly common for local National Weather Service offices to have custom model configurations for their local areas, and there is a desire for information about the best ways of fine-tuning of such local and regional models when particular types of weather are expected. The results of this research indicate that, for modeling severe weather at the mesoscale in the eastern United States, the modeling choices that appear to matter the most for obtaining high forecast skill are the type and amount of high-quality data assimilated, and the choice of planetary boundary layer scheme.

REFERENCES

Ackerman, T. P., A. J. Braverman, D. J. Diner, T. L. Anderson, R. A. Kahn, J. V. Martonchik, J. E. Penner, P. J. Rasch, B. A. Wielicki, and B. Yu, 2004: Integrating and Interpreting Aerosol Observations and Models within the PARAGON Framework. *Bulletin of the American Meteorological Society,* 85, pp. 1523-1533.

Adams-Selin, R. D., S. C. van den Heever, and R. H. Johnson, 2013: Sensitivity of Bow-Echo Simulation to Microphysical Parameterizations. *Weather and Forecasting*, 28, pp. 1188-1209.

Aumann, H. H., M.T. Chahine, C. Gautier, M. Goldberg, E. Kalnay, L. McMillin, H. Revercomb, P.W. Rosenkranz, W. L. Smith, D. Staelin, L. Strow, and J. Susskind, 2003: AIRS/AMSU/HSB on the Aqua Mission: Design, Science Objectives, Data Products and Processing Systems. *IEEE Transactions on Geoscience and Remote Sensing*, 41, 253–264.

Beljaars, A. C. M., 1994: The Parameterization of Surface Fluxes in Large-Scale Models Under Free Convection. *Quarterly Journal of the Royal Meteorological Society*, Vol. 121, 255–270.

Berner, J., S.-Y. Ha, J. P. Hacker, A. Fournier, and C. Snyder, 2011: Model Uncertainty in a Mesoscale Ensemble Prediction System: Stochastic Versus Multiphysics Representations. *Monthly Weather Review,* 139, pp. 1972-1995.

Burges, C. J. C., 1998: A Tutorial On Support Vector Machines for Pattern Recognition. *Data Mining and Knowledge Discovery 2*, pp. 121-167.

Cintineo, R., J. A. Otkin, M. Xue, and F. Kong, 2014: Evaluating the Performance of Planetary Boundary Layer and Cloud Microphysical Parameterization Schemes in Convection-Permitting Ensemble Forecasts Using Synthetic GOES-13 Satellite Observations. *Monthly Weather Review,* 142, 163-182.

Clark, A. J., W. A. Gallus Jr., M. Xue, and F. Kong, 2010: Growth of Spread in Convection-Allowing and Convection-Parameterizing Ensembles. *Weather and Forecasting,* 25, 594-612.

Colquhoun, J. R., and P. A. Riley, 1996: Relationships Between Tornado Intensity and Various Wind and Thermodynamic Variables. *Weather and Forecasting*, 11, 360–371.

Davies, J. M., 2004: Estimations of CIN and LFC Associated with Tornadic and Nontornadic Supercells. *Weather and Forecasting*, 19, 714–726.

Descombes, G., T. Auligné, F. Vandenberghe, and D. M. Baker, 2014: Generalized Background Error Covariance Matrix Model (GEN_BE v2.0). *Geoscientific Model Development Discussions,* 7, 4291-4352.

Doswell, C. A., H. E. Brooks, and M. P. Kay, 2005: Climatological Estimates of Daily Local Nontornadic Severe Thunderstorm Probability for the United States. *Weather and Forecasting,* 20, 577-595.

Doswell, C. A., R. Edwards, R. L. Thompson, J. A. Hart, and K. C. Crosbie, 2006: A Simple and Flexible Method for Ranking Severe Weather Events. *Weather and Forecasting,* Vol. 21, 939-951.

Droegemeier, K. K., S. M. Lazarus, and R. Davies-Jones, 1993: The Influence of Helicity on Numerically Simulated Convective Storms. *Monthly Weather Review,* 121, 2005–2029.

Dudhia, J., 1989: Numerical Study of Convection Observed during Winter Monsoon Experiment Using a Mesoscale Two-Dimensional Model. *Journal of the Atmospheric Sciences,* 46, 3077-3107.

Dudhia, J., 1996: A Multi-Layer Soil Temperature Model for MM5. *Preprints,* The Sixth PSU/NCAR Mesoscale Model Users' Workshop, 22-24 July 1996, Boulder, CO, 49-50.

Ebisuzaki, W., and G. Rutledge (ed.), 2004: Data Documentation for NOAA Operational Model Archive and Distribution System (NOMADS) North America Regional Reanalysis (NARR) "Merge" Dataset. National Climatic Data Center.

Efron, B., and R. Tibshirani, 1993: *An Introduction to the Bootstrap*. Chapman and Hall, Boca Raton, Florida, 436 pp.

Fierro, A. O., E. R. Mansell, C. L. Ziegler, and D. R. MacGorman, 2012: Application of a Lightning Data Assimilation Technique in the WRF-ARW Model at Cloud-Resolving Scales for the Tornado Outbreak of 24 May 2011. *Monthly Weather Review,* 140, pp. 2609-2627.

Frame, J., 2016: This Is How the "Surprise" Indiana and Ohio Tornado Outbreak of August 24, 2016 Happened. U. S. Tornadoes. Retrieved 3 October 2016 from http://www.ustornadoes.com/2016/08/26/surprise-indiana-ohio-tornado-outbreak-august-24-2016-happened/

Gauthier, P., M. Tanguay, S. Laroche, S. Pellerin, and J. Morneau, 2007: Extension of 3DVAR to 4DVAR: Implementation of 4DVAR at the Meteorological Service of Canada. *Monthly Weather Review,* 135, 2339-2354.

Gibbs, J. A., E. Fedorovich, and A. M. J. van Eijk, 2011: Evaluating Weather Research and Forecasting (WRF) Model Predictions of Turbulent Flow Parameters in a Dry Convective Boundary Layer. *Journal of Applied Meteorology and Climatology,* 50, 2429-2444.

Hastings, R., and Y. Richardson, 2016: Long-Term Morphological Changes in Simulated Supercells Following Mergers with Nascent Supercells in Directionally Varying Shear. *Monthly Weather Review,* 144, 471-499.

Hearst, M. A., S. T. Dumais, E. Osuna, J. Platt, and B. Schölkopf, 2002: Support Vector Machines. *Intelligent Systems and Their Applications, IEEE,* Vol. 13, Issue 4, pp. 18-28.

Hong, S.-Y., and J.-O. J. Lim, 2006: The WRF Single-Moment 6-Class Microphysics Scheme (WSM6). *Journal of the Korean Meteorological Society*, Vol. 42, pp. 129–151.

Hong, S.-Y., Y. Noh, and J. Dudhia, 2006: A New Vertical Diffusion Package with an Explicit Treatment of Entrainment Processes. *Monthly Weather Review*, 134, 2318–2341.

Hong, S.-Y., and H.-L. Pan, 1996: Nonlocal Boundary Layer Vertical Diffusion in a Medium-Range Forecast Model. *Monthly Weather Review,* 124, 2322-2339.

Houtekamer, P. L., and H. L. Mitchell, 1998: Data Assimilation Using an Ensemble Kalman Filter Technique. *Monthly Weather Review,* 126, 796-811.

Hu, X.-M., J. W. Nielsen-Gammon, and F. Zhang, 2010: Evaluation of Three Planetary Boundary Layer Schemes in the WRF Model. *Journal of Applied Meteorology and Climatology,* 49, 1831-1844.

Janjić, Z. I., 1994: The Step-Mountain Eta Coordinate Model: Further Developments of the Convection, Viscous Sublayer, and Turbulence Closure Schemes. *Monthly Weather Review,* 122, 927-945.

Janjić, Z. I., 2002: Nonsingular Implementation of the Mellor-Yamada Level 2.5 Scheme in the NCEP Meso Model. *NCEP Office Note No. 437,* 61 pp.

Jones, T. A., D. J. Stensrud, 2012: Assimilating AIRS Temperature and Mixing Ratio Profiles Using an Ensemble Kalman Filter Approach for Convective-Scale Forecasts. *Weather and Forecasting*, Vol. 27, 541–564.

Kain, J. S., 2004: The Kain-Fritsch Convective Parameterization: An Update. *Journal of Applied Meteorlogy,* Vol. 43, 170-181.

Kalnay, E., 2003: *Atmospheric Modeling, Data Assimilation, and Predictability.* Cambridge. 341 pp.

Kerr, B. W., and G. L. Darkow, 1996: Storm-Relative Winds and Helicity in the Tornadic Thunderstorm Environment. *Weather and Forecasting,* Vol. 11, pp. 489-505.

Kleist, D.T., D. F. Parrish, J. C. Derber, R. Treadon, W.-S. Wu, and S. Lord, 2009: Introduction of the GSI Into the NCEP Global Data Assimilation System. *Weather and Forecasting*, 24, 1691-1705.

Lang, S. E., W.-K. Tao, X. Zeng, and Y. Li, 2010: Reducing the Biases in Simulated Radar Reflectivities from a Bulk Microphysics Scheme: Tropical Convective Systems. *Journal of the Atmospheric Sciences,* 68, pp. 2306-2320.

Lang, S. E., W.-K. Tao, J.-D. Chern, D. Wu, and X. Li, 2014: Benefits of a Fourth Ice Class in the Simulated Radar Reflectivities of Convective Systems Using a Bulk Microphysics Scheme. *Journal of the Atmospheric Sciences,* 71, pp. 3583-3612.

Lee, Y., G. Wahba, and S. A. Ackerman, 2004: Cloud Classification of Satellite Radiance Data by Multicategory Support Vector Machines. *Journal of Atmospheric and Oceanic Technology,* Vol. 21, pp. 159-169.

Leutbecher, M. and T. N. Palmer, 2008: Ensemble Forecasting. *Journal of Computational Physics,* Vol. 227, pp. 3515-3539.

Li, Z., P. Zuidema, P. Zhu, and H. Morrison, 2015: The Sensitivity of Simulated Shallow Cumulus Convection and Cold Pools to Microphysics. *Journal of the Atmospheric Sciences,* 72, 3340-3355.

Lim, K.-S. S., and S.-Y. Hong, 2010: Development of an Effective Double-Moment Cloud Microphysics Scheme with Prognostic Cloud Condensation Nuclei (CCN) for Weather and Climate Models. *Monthly Weather Review,* 138, 1587-1612.

Markowski, P. M., 2002: Mobile Mesonet Observations on 3 May 1999. *Weather and Forecasting*, 17, 430–444.

McMillen, J. D. and W. J. Steenburgh, 2015: Impact of Microphysics Parameterizations On Simulations of the 27 October 2010 Great Salt Lake-Effect Snowstorm. *Weather and Forecasting,* 30, 136-152.

McNulty, R. P., 1995: Severe and Convective Weather: A Central Region Forecasting Challenge. *Weather and Forecasting*, 10, 187–202.

Mercer, A. E., M. B. Richman, and H. B. Bluestein, 2008: Statistical Modeling of Downslope Windstorms in Boulder, Colorado. *Weather and Forecasting,* Vol. 23, pp. 1176-1194.

Mercer, A. E., C. M. Shafer, C. A. Doswell III, L. M. Leslie, and M. B. Richman, 2009: Objective Classification of Tornadic and Nontornadic Severe Weather Outbreaks. *Monthly Weather Review*, Vol. 137, pp. 4355-4368.

Mercer, A. E., C. M. Shafer, C. A. Doswell, L. M. Leslie, and M. B. Richman, 2012: Synoptic Composites of Tornadic and Nontornadic Outbreaks. *Monthly Weather Review*, 140, 2590-2608.

Mesinger, F., G. DiMego, E. Kalnay, K. Mitchell, P. C. Shafran, W. Ebisuzaki, D. Jović, J. Woollen, E. Rogers, E. H. Berbery, M. B. Ek, Y. Fan, R. Grumbine, W. Higgins, H. Li, Y. Lin, G. Manikin, D. Parrish, and W. Shi, 2006: North American Regional Reanalysis. *Bulletin of the American Meteorological Society*, Vol. 87, pp. 343-360.

Milbrandt, J. A., and M. K. Yau, 2005: A Multimoment Bulk Microphysics Parameterization. Part I: Analysis of the Role of the Spectral Shape Parameter. *Journal of the Atmospheric Sciences,* 62, 3051-3064.

Milbrandt, J. A., 2005: A Multimoment Bulk Microphysics Parameterization. Part II: A Proposed Three-Moment Closure and Scheme Description. *Journal of the Atmospheric Sciences.,* 62, 3065-3081.

Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997: Radiative Transfer for Inhomogeneous Atmospheres: RRTM, a Validated Correlated–k Model for the Longwave. *Journal of Geophysical Research*, Vol. 102, 16663–16682.

Molthan, A. L., and B. A. Colle, 2012: Comparisons of Single- and Double-Moment Microphysics Schemes in the Simulation of a Synoptic-Scale Snowfall Event. *Monthly Weather Review,* 140, pp. 2982-3002.

Molthan, A. L., W. A. Petersen, S. W. Nesbitt, and D. Hudak, 2010: Evaluating the Snow Crystal Size Distribution and Density Assumptions within a Single-Moment Microphysics Scheme. *Monthly Weather Review,* 138, pp. 4254-4267

Morrison, H., G. Thompson, and V. Tatarskii, 2009: Impact of Cloud Microphysics On the Development of Trailing Stratiform Precipitation in a Simulated Squall Line: Comparison of One- and Two-Moment Schemes. *Monthly Weather Review,* 137, 991-1007.

Morrison, H., and J. Milbrandt, 2011: Comparison of Two-Moment Bulk Microphysics Schemes in Idealized Supercell Thunderstorm Simulations. *Monthly Weather Review,* 139, 1103-1130.

National Centers for Environmental Prediction, National Weather Service, NOAA, U.S. Department of Commerce, 2008, updated daily: NCEP ADP Global Upper Air and Surface (PREPBUFR and NetCDF formats) Weather Observations, May 1997 - Continuing. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory. http://rda.ucar.edu/datasets/ds337.0.

National Centers for Environmental Prediction, National Weather Service, NOAA, U.S. Department of Commerce. 2009, updated daily: NCEP GDAS Satellite Radiance Data. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory. http://rda.ucar.edu/datasets/ds735.0.

National Climatic Data Center, National Environmental Satellite, Data, and Information Service, NOAA, U.S. Department of Commerce, 2010, updated monthly: NCDC Storm Data. http://www.ncdc.noaa.gov/IPS/sd/sd.html

Parrish, D. F. and J. C. Derber, 1992: The National Meteorological Center's Spectral Statistical-Interpolation Analysis System. *Monthly Weather Review,* 120, 1747–1763.

Pleim, J. E., 2007: A Combined Local and Nonlocal Closure Model for the Atmospheric Boundary Layer. Part I: Model Description and Testing. *Journal of Applied Meteorology and Climatology,* 46, 1383-1395.

Rasmussen, E. N., and D. O. Blanchard, 1998: A Baseline Climatology of Sounding-Derived Supercell and Tornado Forecast Parameters. *Weather and Forecasting*, 13, 1148–1164.

Richman, M., 1986: Rotation of Principal Components. *Journal of Climatology*, 6, 293-335.

Rousseeuw, P. J., 1987: Silhouettes: A Graphical Aid to the Interpretation and Validation for Cluster Analysis. *Journal of Computational Applied Mathematics,* 20, 53–65.

Shafer, C. M, A. E. Mercer, L. M. Leslie, M. B. Richman, and C. A. Doswell III, 2010: Evaluation of WRF Model Simulations of Tornadic and Nontornadic Outbreaks Occurring in the Spring and Fall. *Monthly Weather Review,* 138, 4098-4119.

Shafer, C. M, and C. A. Doswell III, 2010: A Multivariate Index for Ranking and Classifying Severe Weather Outbreaks. *E-Journal of Severe Storms Meteorology,* Vol. 5, No. 1

Shafer, C. M., and C. A. Doswell III, 2011: Using Kernel Density Estimation to Identify, Rank, and Classify Severe Weather Outbreak Events. *E-Journal Severe Storms Met.,* 6 (2), 1-28.

Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, W. Wang, and J. G. Powers, 2007: A description of the Advanced Research WRF Version 2. NCAR Tech. Note TN-4681STR, 88 pp.

Stensrud, D. J., J. V. Cortinas, and H. E. Brooks, 1997: Discriminating Between Tornadic and Nontornadic Thunderstorms Using Mesoscale Model Output. *Weather and Forecasting*, 12, 613–632.

Stensrud, D. J., and S. J. Weiss, 2002: Mesoscale Model Forecasts of the 3 May 1999 Tornado Outbreak. *Weather and Forecasting,* Vol. 17, pp. 526-543.

Tao, W.-K., J. Simpson, and M. McCumber, 1989: An Ice–Water Saturation Adjustment. *Monthly Weather Review*, Vol. 117, 231–235.

Tapiador, F. J., W.-K. Tao, J. J. Shi, C. F. Angelis, M. A. Martinez, C. Marcos, A. Rodriguez, and A. Hou, 2012: A Comparison of Perturbed Initial Conditions and Multiphysics Ensembles in a Severe Weather Episode in Spain. *Journal of Applied Meteorology and Climatology,* 51, 489-504.

Tewari, M., F. Chen, W. Wang, J. Dudhia, M. A. LeMone, K. Mitchell, M. Ek, G. Gayno, J. Wegiel, and R. H. Cuenca, 2004: Implementation and Verification of the Unified NOAH Land Surface Model in the WRF Model. *20th Conference on Weather Analysis and Forecasting/16th Conference on Numerical Weather Prediction*, pp. 11–15.

Thompson, Gregory, P. R. Field, W. D. Hall, and R. M. Rasmussen, 2006: A New Bulk Microphysical Parameterization for WRF (& MM5). National Center for Atmospheric Research.

Thompson, R. L., C. M. Mead, and R. Edwards, 2007: Effective Storm-Relative Helicity and Bulk Shear in Supercell Thunderstorm Environments. *Weather and Forecasting,* Vol. 22, pp. 102-115.

Thompson, G., P. R. Field, R. M. Rasmussen, and W. D. Hall, 2008: Explicit Forecasts of Winter Precipitation Using an Improved Bulk Microphysics Scheme. Part II: Implementation of a New Snow Parameterization. *Monthly Weather Review*, Vol. 136, 5095–5115.

Weisman, M. L., and J. B. Klemp, 1984: The Structure and Classification of Numerically Simulated Convective Storms in Directionally Varying Wind Shears. *Monthly Weather Review*, 112, 2479–2498.

Wheatley, D. M, D. J. Stensrud, D. C. Dowell, and N. Yussouf, 2012: Application of a WRF Mesoscale Data Assimilation System to Springtime Severe Weather Events 2007-09. *Monthly Weather Review,* 140, 1539-1557.

Wilks, D. S., 2011: *Statistical Methods in the Atmospheric Sciences*. 3rd edition. Academic Press, 627 pp.

Wu, W., R. J. Purser, and D. F. Parrish, 2002: Three-Dimensional Variational Analysis with Spatially Inhomogeneous Covariances. *Monthly Weather Review,* 130, 2905–2916.

Yussouf, N., E. R. Mansell, L. J. Wicker, D. M. Wheatley, and D. J. Stensrud, 2013: The Ensemble Kalman Filter Analyses and Forecasts of the 8 May 2003 Oklahoma City Tornadic Supercell Storm Using Single- and Double-Moment Microphysics Schemes. *Monthly Weather Review,* 141, pp. 3388-3412.