ENSEMBLE KALMAN FILTER DATA ASSIMILATION IN REGIONS OF COMPLEX TERRAIN

by

Hailing Zhang

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Atmospheric Sciences

The University of Utah

May 2014

Copyright © Hailing Zhang 2014

All Rights Reserved

The University of Utah Graduate School

STATEMENT OF DISSERTATION APPROVAL

The dissertation of **Hailing Zhang**

has been approved by the following supervisory committee members:

Zhaoxia Pu	, Chair	**10/31/2013** Date Approved
Steven K. Krueger	, Member	**10/31/2013** Date Approved
Eric R. Pardyjak	, Member	**10/31/2013** Date Approved
Courtenay Strong	, Member	**10/31/2013** Date Approved
Joshua Hacker	, Member	**10/31/2013** Date Approved

and by **Kevin Perry** , Chair/Dean of

the Department/College/School of **Atmospheric Sciences**

and by David B. Kieda, Dean of The Graduate School.

ABSTRACT

Accurate weather forecasting in complex terrain is of great importance, yet it is a challenging problem due to a number of difficulties, including sparse observations, terrain misrepresentation in numerical models, and model errors related to terrain complexity. Owing to these limitations, few previous studies in data assimilation have emphasized regions of complex terrain. This dissertation presents the first comprehensive evaluation of data assimilation methods and forecast error characteristics for near-surface atmospheric variables in complex terrain. The mesoscale community Weather Research and Forecasting (WRF) model and an advanced ensemble Kalman filter (EnKF) data assimilation system are employed.

First, the capability of the advanced EnKF method in assimilating near-surface observations (2-m temperature and 10-m wind) is examined in an observing system simulation experiments framework and compared with the traditional three-dimensional variational data assimilation (3DVAR) method. Results indicated that the EnKF is able to effectively assimilate surface observations and improve the short-range weather forecasts, while the 3DVAR method has fundamental problems in assimilating surface observations.

Next, the performance of the WRF model in predicting near-surface atmospheric temperature and wind conditions under various terrain and weather regimes is examined. The WRF model is able to simulate these weather phenomena reasonably

well. Forecasts of near-surface variables in flat terrain generally agree well with observations. In complex terrain, forecasts not only suffer from the model's inability to reproduce accurate atmospheric conditions in the lower atmosphere but also struggle with representative issues due to mismatches between the model and the actual terrain. A statistical analysis during a 1-month period over the Dugway Proving Ground (DPG), Utah illustrates that forecast errors in near-surface variables depend strongly on the diurnal variation in surface conditions, especially when synoptic forcing is weak.

Finally, the impact of observations from the recent field experiments of the Mountain Terrain Atmospheric Modeling and Observations (MATERHORN) is examined with EnKF. Results illustrated that the quality of the EnKF/WRF analysis is generally high and the short-range forecast errors are comparable to those of the National Centers for Environmental Prediction (NCEP) North American Mesoscale Model (NAM) forecasts for both 10-m wind speed and direction.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I would like to express my sincere appreciation to my advisor Dr. Zhaoxia Pu for her continual support and guidance over the past years. I thank her for her hard work that made this research financially possible. I thank her for supporting me in attending many conferences and colloquia, for creating opportunities to communicate with inside and outside experts, for encouraging me in my research and classes, and for repeatedly revising and proofreading my manuscripts.

I would also like to thank my committee members, Drs. Steve Krueger, Court Strong, Eric Pardyjak, and Joshua Hacker, who gave me valuable input during this study. Some discussions with Prof. W. James Steenburgh and Mr. Jeff Massey were also very helpful.

I appreciate the MATERHORN science team members for their hard work. I have learned a lot from the team meetings and daily weather briefings during the field period. The field data has also been very helpful for my research. I also take this opportunity to thank all the faculty members of the Department of Atmospheric Sciences for their help and encouragement.

I am grateful to the NCAR WRF model development group and Data Assimilation Research Testbed (DART) team. Specifically, I would like to thank Jeff Anderson, Glen Romine, Nancy Collins, and Hui Liu for their help in using DART/WRF. In addition, I acknowledge the Meteorological Division of Dugway Proving Ground for providing

surface observations. Part of the surface observations used in this study were obtained from mesowest.utah.edu, managed by the University of Utah.

This research was supported by National Science Foundation Grant AGS-08339856 and AGS-1243027, and the Office of Naval Research Award No. N00014-11-1-0709. The computational support from the Center for High Performance Computing (CHPC) at the University of Utah is also gratefully acknowledged.

I would like to thank my office mates, Levi Thatcher, Zhan Li, Chao Lin, and Chris Pace, for their friendship and help. I also thank the many friends and meteorologists in the department who directly or indirectly lent a helping hand during these past years.

Finally, I would like to thank my husband, Xuebo, who was always there supporting me on this long journey toward a doctoral degree.

CHAPTER 1

INTRODUCTION

1.1 Motivation

About 25% of the Earth's land surface is covered by mountains, and 26% of the world's population lives in mountainous regions (Meybeck et al. 2001). Populations outside of the mountainous regions are greatly affected by mountain-related weather. Therefore, accurate weather forecasting in complex terrain is of great importance, yet it is a challenging problem due to a number of difficulties, including sparse observations, terrain misrepresentation in numerical models, and model errors related to terrain complexity. Previous studies of numerical models (Hart et al. 2005; Liu et al. 2008; Mass et al. 2002; Pu et al. 2013) have demonstrated the unavoidable errors of near-surface atmospheric simulation that are either related to or caused by complex terrain. Owing to these limitations, few previous studies in data assimilation have emphasized regions of complex terrain.

This dissertation conducts the first comprehensive evaluation of data assimilation methods and forecast error characteristics in regions of complex terrain with the mesoscale community Weather Research and Forecasting (WRF) model and an advanced ensemble Kalman filter (EnKF) data assimilation method.

1.2 Scope of This Dissertation

Accurate numerical weather forecasting is of great importance. Due to inadequate observations, our limited understanding of the physical processes of the atmosphere, and the chaotic nature of atmospheric flow, uncertainties always exist in modern numerical weather prediction (NWP). Recent developments in ensemble forecasting and ensemble-based data assimilation have proved that there are promising ways to beat the forecast uncertainties in NWP. Section 1.3 gives a literature review of fundamental problems and recent progress associated with ensemble forecasting and EnKF methods. The usefulness of these methods in improving high-impact weather forecasting is also discussed.

Surface observations represent the major source of available conventional observations. However, their use was limited in most operational NWP systems. In Chapter 2, the ability of EnKF to assimilate surface observations is examined with Observing System Simulation Experiments (OSSEs) and compared with the three-dimensional variational data assimilation (3DVAR) method in short-range forecasts. For simplicity, the study starts with single-observation experiments. Experiments are also conducted with the assimilation of multiple observations when a frontal system and a low-level jet system are present. The impact of surface observations on the forecast of these synoptic systems is examined in detail. The ability of the two methods to deal with terrain-mismatch and data-rejection is also investigated. The major discrepancies between the 3DVAR and EnKF and the sensitivity of each method to its configurations, particularly the background error correlation length scale and horizontal and vertical localization are examined. In addition, the relative capability of 3DVAR and EnKF in handling terrain misrepresentation in surface data assimilation is evaluated and

compared.

In Chapter 3, error characteristics of atmospheric near-surface variables are examined. Specifically, the performance of the WRF model in predicting near-surface atmospheric temperature and wind conditions under various terrain and weather regimes is investigated. Three individual events under strong synoptic forcing, namely, a frontal system, a low-level jet, and a persistent cold air pool, are first verified against observations in both flat and complex terrain. The model performance is examined with general verification of large-scale synoptic systems and statistical verification of surface conditions. The sensitivity of surface forecasts to the horizontal and vertical model resolution and the physical parameterization schemes is also explored. A statistical analysis is also performed for 120 forecasts during a 1-month period in fall 2011, to further investigate forecast error characteristics in complex terrain. The diurnal error characteristics under various synoptic forcing are compared.

The outcomes in Chapters 2 and 3 add to our knowledge about the forecast error characteristics of near-surface atmospheric variables from WRF simulations and the ability of EnKF data assimilation to improve surface analyses and forecasts. Inspired by these results, we conducted 3-hourly, continuous data assimilation cycles for a 1-month period at Dugway Proving ground (DPG) with EnKF, WRF, and observations collected during the fall 2012, MATERHORN field experiments to further investigate the analysis and forecast error characteristics in complex terrain, as discussed in Chapter 4. The impact of data assimilation on the model's prediction skill is also examined with selected cases.

Chapter 5 summarizes the results and highlights the main findings. A few concluding remarks are also made for future work.

1.3 Overview of Ensemble Forecasting and Ensemble-Based Data Assimilation in NWP

Numerical weather prediction (NWP) is an initial value problem: it forecasts the atmospheric state by integrating a numerical model with given initial conditions. Commonly, two fundamental factors account for an accurate numerical weather forecast: (1) the present state of the atmosphere must be characterized as accurately as possible and (2) the intrinsic laws, according to which the subsequent states develop out of the preceding ones, must be known (Bjerknes 1911).

Since the first successful NWP in the early 1950s by Charney et al. (1950), much progress has been made in enhancing the skill of NWP. These include efforts in improving initial conditions through advances in observing systems and the development of atmospheric data assimilation techniques. Many studies were also devoted to improving numerical modeling with advanced numerical methods, better representation of dynamics processes of the atmosphere, and improved physical parameterization schemes (Haltine 1971; Kalnay 2003; Thompson 1961). Today, NWP has become a major forecasting tool in many operational centers around the world.

However, due to inadequate observations, our limited understanding of the physical processes of the atmosphere, and the chaotic nature of the atmospheric flow, uncertainties always exist in both initial conditions and numerical models. Thus, reducing forecast errors caused by these uncertainties remains a large area of research and operational

implementation.

Recent developments have proved that ensemble forecasting and ensemble-based data assimilation are promising ways to beat the forecast uncertainties in NWP. This section gives a brief overview of the fundamental problems and recent progress associated with ensemble forecasting and ensemble-based data assimilation.

1.3.1 Forecast Uncertainties and Predictability

Predictability refers to the extent to which the future state of the atmosphere or a specific weather system may be predicted based on current ability of NWP. Corresponding to the aforementioned two fundamental factors that influence the numerical forecast, there are two kinds of predictabilities as addressed by Lorenz (1963): (1) attainable predictability, which is limited by the inaccuracy of measurement and (2) practical predictability, which is limited by our inability to express the precise equations of the atmosphere motion and physical processes in the numerical model. Errors in model equations rely much on the computational methods used to solve the equations and our current ability to understand the physical processes as well as the model resolution to resolve these physical processes in the numerical models. In this dissertation, we mainly address the practical predictability.

While attainable and practical predictabilities are associated with uncertainties in the initial conditions and imperfect models, what would the predictability be if the model (dynamical and physical processes) were perfect and the initial conditions were accurate? Lorenz (1963, 1965) asserted that the atmosphere, as a kind of unstable dynamical system, has a finite limit of predictability depending upon a particular flow. As is well known,

Lorenz (1965) found that a slight departure in initial conditions would evolve into totally different atmospheric states in the numerical forecasts regardless of how small the errors in the initial conditions were.

The chaotic nature of the atmosphere determines that the predictability of the model depends upon not only the realism of the model and the accuracy of initial conditions but also the system itself. Atmospheric motion, as a nonlinear dynamic system, is supposed to have finite limit predictability. The stochastic characteristics account for the extent to which the atmosphere could be predicted. The number of days we can forecast accurately in advance is dependent upon the evolution of the atmosphere.

The uncertain properties of the atmospheric system call for more suitable methods to represent the initial conditions and forecast the atmospheric states, instead of the traditional way that describes the initial values with the single analysis best state and integrates the single best guess forward.

1.3.2 The Stochastic Prediction and Ensemble Forecasting

In view of the uncertain properties of the atmospheric system, a theory of stochastic dynamic prediction was proposed by Epstein (1969). In a stochastic context, the initial and forecast states of the atmosphere are represented as probability distributions. That is, the probability density function (PDF) of the present model state should be estimated first according to all the prior information and available observations; then, a method for forecasting the evolution of this PDF forward in time is needed. Based on the stochastic dynamic prediction, it is possible to make the probabilistic forecasts in addition to a deterministic forecast using a single model with single initial conditions. Although early

experiments by Epstein were very different from the ensemble forecasting done today, the theory of stochastic dynamic predictions offers a stepping stone with which to develop ensemble forecasting.

The advance in parallel processing computers in the early 1990s and improved operational forecasting systems—improvements in both model physics and data assimilation—has led to operational stochastic dynamic prediction at the European Centre for Medium-Range Weather Forecasts (ECMWF), U. S. National Centers for Environmental Prediction (NCEP), and the Meteorological Service of Canada (MSC) in the early 1990s. These operational stochastic prediction systems are referred to as ensemble forecasting systems. Instead of using only one model with a single set of initial conditions, a group of forecasts with slightly different initial conditions are made in an ensemble forecast. The approach to ensemble prediction used at operational centers exhibits subtle differences when compared with the standard Monte Carlo method that was used in the stochastic dynamic prediction. In Monte Carlo, it is assumed that the initial probability density function (PDF) is known and that it is sampled randomly. In most of the methods used in current ensemble forecasting, the PDF is generally not sampled in a random way. There are different ways to generate the initial perturbations in the different operational ensemble systems, including the Breeding of Growing Modes (BGM; Toth and Kalnay 1993, 1997), the Singular vector (SV) method (Buizza and Palmer 1995; Molteni, et al. 1996), the Perturbed-observation approach (Houtekamer et al. 1996), and the Ensemble transform method (Bishop et al. 2001; Wei et al. 2008).

All of the methods discussed above only include perturbations in the initial conditions, assuming that the error growth due to model deficiencies is small compared to that due to

unstable growth of initial errors. However, in reality, uncertainties in model physical parameterizations cannot be ignored in many cases. Therefore, in addition to the aforementioned initial perturbation methods, ensemble forecast systems have also been designed to account for model errors and uncertainty. Current methods and progress include the multimodel ensemble (e.g., Kharin and Zwiers 2002; Krishnamurti et al. 2000), stochastic physical parameterizations (e.g., Buizza et al. 1999; Reynolds et al. 2008), nonlocal stochastic-dynamic parameterization schemes (Palmer 2001), kinetic energy backscatter (Shutts 2005), performing ensemble simulations with different time steps to study the impact of model truncation error (Teixeira et al. 2007), and using different parameterizations within the ensemble prediction system (Houtekamer et al. 1996). Krishnamurti et al. (2000) commented that the performance of multimodel ensemble forecasts shows superior forecast skill compared to all individual models used. Reynolds et al. (2008) illustrated that a stochastic convection scheme improves the ensemble performance in the tropics.

Since ensemble forecasting takes account of the uncertainties in NWP, it has major advantages over a single deterministic forecast (Tracton and Kalnay 1993). It improves the forecasting skill by reducing the nonlinear error growth and averaging out unpredictable components. It predicts the skill by relating it to the agreement among ensemble forecast members. If the ensemble forecasts are quite different from each other, it is clear that at least some of them are wrong, whereas if there is good agreement among the forecasts, there is more reason to be confident about the forecast. It provides an objective basis for forecasts in a probabilistic form. In a chaotic system such as the atmosphere, probabilistic information is recognized as the optimum format for weather

forecasts both from a scientific and a user perspective.

1.3.3 Bayes Theorem and Ensemble-based Data Assimilation

As mentioned in the previous section, uncertainties of the initial conditions are the major source of error in NWP. Thus, improved data assimilation techniques will be useful to beat the uncertainties in the initial conditions. We continue this subject with the stochastic dynamic prediction.

In a stochastic context, the initial and forecast states of the atmosphere are represented as probability distributions. Therefore, the probability density function of the present model state should be estimated first according to all the prior information and available observations and then a method for forecasting the evolution of this PDF forward in time is needed. Usually, getting the current PDF is referred to as Bayes data assimilation theory (Hamill 2006; Lorenc 1986).

In the application of data assimilation, Bayes' theorem can be expressed as

$$P(\mathbf{X}_t \mid \mathbf{Y}_t) = \frac{P(\mathbf{Y}_t \mid \mathbf{X}_t) P(\mathbf{X}_t \mid \mathbf{Y}_{t-1})}{denominator} \tag{1.1}$$

where $P(\mathbf{X}_t \mid \mathbf{Y}_t)$ denotes the probability density of the model state at time t, $\mathbf{X}$ and $\mathbf{Y}$ are the state variables, $P(\mathbf{Y}_t \mid \mathbf{X}_t)$ denotes the probability density of the observations at time t, and $P(\mathbf{X}_t \mid \mathbf{Y}_{t-1})$ is viewed as a kind of prior and represents the probability density of the prior ensemble forecast at time t. The denominator is a kind of normalization for guaranteeing that the total probability of all possible states is 1.

As shown above, equation (1.1) describes the way in which new observations are incorporated to modify the prior conditional probability density available from predictions based on earlier observations.

Taking an example, for Gaussian probability density, the prior is

$$\mathrm{P}\left(\mathbf{X}_t \mid \mathbf{Y}_{t-1}\right) = \mathrm{Normal}(\mu_p, \sigma_p^2) \tag{1.2}$$

where μ and σ are the mean and standard deviations, respectively. The subscript "p" denotes the "prior" state.

The observation PDF is given as

$$\mathrm{P}\left(\mathbf{Y}_t \mid \mathbf{X}_t\right) = \mathrm{Normal}(\mu_O, \sigma_O^2) \tag{1.3}$$

where the Gaussian probability density function given the mean μ and standard derivation error σ is

$$\mathrm{P}\left(\mu, \sigma^2\right) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{1.4}$$

Dividing the product (named P' temporarily) of P($\mathbf{Y}_t \mid \mathbf{X}_t$) times P($\mathbf{X}_t \mid \mathbf{Y}_{t-1}$) by a normalization denominator gives the posterior PDF. After the above processes, we get the posterior estimate,

$$\mathrm{P}\left(\mathbf{X}_t \mid \mathbf{Y}_t\right) = \mathrm{Normal}(\mu_U, \sigma_U^2) \tag{1.5}$$

Here,

$$\begin{aligned} \sigma_u^2 &= \left(\sigma_p^{-2} + \sigma_o^{-2}\right)^{-1} \\ \gamma_u &= \sigma_u^2\left[\sigma_p^{-2}\gamma_p + \sigma_o^{-2}\gamma_o\right] \end{aligned} \tag{1.6}$$

Although we can solve for the posterior PDF in the previous section, it is not easy to express the PDF of the observations and the prior information explicitly in real operational numerical implementation. Therefore, it is difficult to obtain the posterior PDF of initial conditions directly from Bayes theorem.

Fortunately, the implementation of the Monte Carlo method provided us an effective approach to simulate the desired PDF with a random sample and to some extent solve the

uncertainties of the initial conditions. However, the Monte Carlo method was conditionally effective only under an assumption that the number of sample members is sufficiently large in order to represent the PDF suitably. Consequently, the difficulty comes with the "large sample." For instance, for a common real model with 10^7 degree of freedom, a 10^7*10^7 dimension calculation for estimating the PDF will be involved. That is demanding, considering even the most recent computational advances.

Figure 1.1 shows schematically the forecast results under conditions when there are too few sample members doing the estimating. The mean forecast drifts away from the truth with time.

1.3.4 Ensemble Kalman Filter

Considering the limitations of traditional Bayes and Monte Carlo methods, a more practical technique is needed. With the most recent developments, ensemble Kalman filter data assimilation techniques, originated from the basic idea of the Monte Carlo theory and the well-known Kalman filter method, are successfully applied in many research and operational practices.

As a sequential data assimilation method, the implementation of the Kalman filter (Evensen 2003) includes two steps, which are named a forecast step and an analysis step. The model is integrated forward with time and used to update the model state by assimilating new observations when observations are available. The Kalman filter assumes that the prior conditional probability distribution is Gaussian and expresses it with its mean and covariance.

The analysis equation is

$$x^a = x^f + \mathbf{P}^f \mathbf{H}^T (\mathbf{H}\mathbf{P}^f \mathbf{H}^T + \mathbf{R})^{-1}[d - \mathbf{H}(x^f)] \tag{1.7}$$

where $\boldsymbol{x}^a$ is the vector of analysis variables, $\boldsymbol{x}^f$ is the vector of background fields (prior estimate), and **d** denotes observations. **H,** called the observational operator, connects the true state with observations within particular measurement errors:

$$\mathbf{d} = \mathbf{H}\mathbf{X}^t + \varepsilon \tag{1.8}$$

K is the so-called gain matrix:

$$\boldsymbol{K} = \mathbf{P}^f \mathbf{H}^T (\mathbf{H}\mathbf{P}^f \mathbf{H}^T + \mathbf{R})^{-1} \tag{1.9}$$

In theory, the error covariance of forecast estimation (background) is defined as

$$\mathbf{P}^f = \overline{(\boldsymbol{x}^f - \boldsymbol{x}^t)(\boldsymbol{x}^f - \boldsymbol{x}^t)^T} \tag{1.10}$$

However, we never know the true atmosphere state ($\boldsymbol{x}^t$). This makes the estimation of the background error covariance very difficult. In ensemble Kalman filter method, we assume the ensemble mean could be the best estimation of the true state. The ensemble Kalman filter employs a group of ensemble members to represent the covariance statistics of the analyzed state,

$$\mathbf{P}^f = \overline{(\boldsymbol{x}_k^f - \overline{\boldsymbol{x}^f})(\boldsymbol{x}_k^f - \overline{\boldsymbol{x}^f})^T} \tag{1.11}$$

Equation (1.11) indicates that a flow-dependent error covariance of forecast estimation can be obtained by using ensemble forecasting in practical implementation.

1.3.5 Implementation and Practical Issues on Ensemble Kalman Filter

Since the first attempt by Evensen (1994), ensemble Kalman filter methods have been developed rapidly and used widely in data assimilation applications. Based on whether to perturb the observations, there are two classes of basic approaches, referred to as perturbed observations and a square root filter, to implement the ensemble Kalman filter as aforementioned. The perturbed observation algorithm updates each ensemble member with a different set of observations perturbed with random noise. Because randomness is introduced in every assimilation cycle, the update is considered stochastic. The square root filter methods do not add stochastic noise to the observations and are called deterministic algorithms. Evensen (1994), Evensen and Van Leeuwen (1996), and Houtekamer and Mitchell (1998) originally implemented the ensemble Kalman filter with perturbed observations. Following this, Anderson (2001), Bishop et al. (2001), Baek et al. (2006), Corazza et al. (2007), Miyoshi and Yamane (2007), Hunt et al. (2007), Harlim and Hunt (2007), and Yang et al. (2009) contributed various square root filter algorithms including an ensemble adjustment Kalman filter (Anderson 2001), an ensemble transform Kalman filter (Bishop et al. 2001), a local ensemble Kalman filter (LEKF, Baek et al. 2006; Corazza et al. 2007) and a local ensemble transform Kalman filter (LETKF, Hunt et al. 2007; Miyoshi and Yamane 2007; Harlim and Hunt 2007; Yang et al. 2009). Whitaker and Hamill (2002) indicated that the perturbed observations approach might introduce another kind of sampling errors; thus, the square root algorithms methods are more accurate for a given ensemble size.

1.3.6 Comparison of Ensemble Kalman Filter with 4DVAR

Since the ensemble Kalman filter is becoming part of the operational choice, progress has been made to compare it with advanced data assimilation methods that are currently available. Specifically, a four-dimensional variational data assimilation (4DVAR) method has been widely adopted in operational centers around the world. Owing to its capability in assimilating asynchronous observations and high-resolution observations such as satellite radiance and radar reflectivity, 4DVAR method is indeed helpful for improving current numerical forecasting (Bauer et al. 2006; Köpken et al. 2004; Mahfouf et al. 2005). However, the requirement of the tangent linear and adjoint models made the 4DVAR method complicated in its implementation. Compared to 4DVAR, the major merit of the ensemble Kalman filter is its simplicity of implementation. It does not need to develop and maintain a tangent linear and adjoint model. It is model independent. One can easily switch to other models using ensemble methods (Lorenc 2003; Kalnay et al. 2007). In addition, the ensemble Kalman filter represents and forwards forecast covariance using the ensemble sample without much effort. The main disadvantage of the ensemble Kalman filter comes with the sampling problem. The low ensemble size brings up sampling errors in the estimation of the background error covariance. The inflation tuning is employed to adjust this sample error in the practice.

Fertig et al. (2007) studied the performance of 4DVAR and 4D-LETKF in assimilating the asynchronous observations using the Lorenz 96 model. Both schemes have comparable error when 4D-LETKF is cycled frequently and when 4DVAR is performed over a sufficiently long analysis time window. Yang et al. (2007) explored the relative advantages and disadvantages of the 4DVAR and LEKF using a quasi-

geostrophic model and asserted that the LEKF did better on both computational cost and accuracy when assimilating the same rawinsonde observations. Buehner et al. (2008) evaluated the operational performance of both methods in Environment Canada using the same model and observations and obtained equivalent forecast scores. Kalnay et al. (2007) offered a comprehensive comparison between 4DVAR and ensemble Kalman filter. Based on results obtained using operational models and both simulated and real observations, they concluded that currently the ensemble Kalman filter is becoming competitive with 4DVAR, and that the experience acquired with each of these methods can be used to improve the other.

In brief, due to its simple implementation and equivalent ability compared to 4DVAR, the ensemble Kalman filter is becoming an attractive operational choice in more centers. However, the current ability of the ensemble Kalman filter is not equal to 4DVAR in terms of assimilating satellite and radar observations. In order to utilize advantages from both methods, a hybrid approach, originally proposed by Hamill and Snyder (2000) has received significant attention. Lorenc (2003) asserted that hybrid approaches of variational methods and ensemble methods would be better than either single approach. Buehner et al. (2008) showed that a hybrid approach based on 4DVAR but using forecast covariance error estimation from the ensemble Kalman filter gave an improvement in 5-day forecasts in the southern hemisphere.

From the results of current studies, the hybrid method of the ensemble Kalman filter and 4DVAR has a promising future since it combines the advantages of both methods and eliminates the existing disadvantages.

1.3.7 Nonlinear Issues in Ensemble Kalman Filter

Previous studies have proven that ensemble Kalman filter is capable of dealing with data assimilation in nonlinear system (e.g., Pu and Hacker 2009). However, nonlinearity is still an important issue in the implementation of ensemble Kalman filter. Meanwhile, the equations introduced in section 1.3.4 are valid only when the error PDF is Gaussian. Unfortunately, in reality, even if the error PDF is Gaussian at the initial time, it would be non-Gaussian when the model is integrated forward due to the strongly nonlinear model. In the case, the error PDF cannot be represented by a Gaussian function. In addition, it is operationally impossible to assume a non-Gaussian error PDF, although it looks feasible based on the current ensemble.

There have been many studies devoted to dealing with the nonlinear and non-Gaussian problem, mainly focusing on the development and implementation of the ensemble Kalman filter. For instance, Van Leeuwen (2003) presented a true variance minimizing filter method. Its performance was tested by the Korteweg–DeVries equation and a quasi-geostrophic model. He addressed that the method works satisfactorily with a strongly nonlinear system. Hoteit et al. (2008) evaluated a new particle-type filter based on a Gaussian mixture representation of the state PDF using the Lorenz 96 model and discussed its application in real meteorological and oceanographic models. Yang and Kalnay (2010) applied the outer loop in LETKF to handle the nonlinear problem with the Lorenz 63 model. Results indicated that the LETKF with outer loop could use a longer assimilation window and improve the analysis accuracy during highly nonlinear time periods.

1.3.8 Applications of Ensemble Forecasting and Ensemble Kalman Filters to High-impact Weather Prediction

Owing to their advantages in beating the uncertainties and dealing with the nonlinearity, ensemble forecasting and ensemble-based data assimilation have received a lot of attention in the research and operational communities during the last decade. Specifically, they have been applied to high-impact weather forecasting. Many studies have documented results from these applications. The ensemble forecasting was used in short range ensemble forecasting (SREF; Du et al. 1997; Du and Traction 2001; Du et al. 2006; Mullen et al. 1999; Yuan et al. 2005), tropical cyclone forecasts (Aberson 2001; Zhang and Krishnamurti 1997), as well as the flooding warning (Mackey and Krishnamurti 2001), etc. The ensemble-based Kalman filtering techniques were also applied for the studying and numerical simulation of hurricanes (e.g., Zhang et al. 2009) and storm scale forecasts in the high resolution (e.g., Tong and Xue 2008; Xue et al. 2009).

Du et al. (1997) applied ensemble forecasting in quantitative precipitation forecasting (QPF). They found a remarkable reduction of root-mean-square error for QPF due to the ensemble application and asserted that the improvements from SREF techniques exceed the effect due to resolution doubling. After a short-range ensemble forecasting system was implemented in real time operational at NCEP in 2001 (Du and Traction 2001), Du et al. (2006) added another 6 members, which generated from a weather research and forecasting (WRF) model into the ensemble forecasting and obtained forecast improvements with increased ensemble spreads. Yuan et al. (2005) studied the QPFs and probabilistic QPFs (PQPFs) over the southwest United States, the area that is marked by

highly heterogeneous topography and diverse vegetation.

The hurricane track forecasting by Zhang and Krishnamurti (2001) showed that the ensemble forecasts are superior to the results from single-model control experiments and the track position errors are largely reduced by the ensemble prediction. Mackey and Krishnamurti (2001) combined ensemble forecasts with a high-resolution regional spectral model to postpredict the track, intensity, and flooding precipitation arising from Typhoon Winnie in August 1997. They evaluated the effectiveness of the ensemble forecasting and found that the ensemble mean track would be superior only if the forecast uncertainty is properly sampled.

Zhang et al. (2009) studied Hurricane Humberto (2007) using the ensemble Kalman filter method for assimilating Doppler radar radical velocity measurements. Results indicated that the ensemble Kalman filtering analysis improved the representation of the track and intensity of Humberto. Tong and Xue (2008) and Xue et al. (2009) used the ensemble Kalman filter method and radar reflectivity observations to correct errors in fundamental microphysical parameters that are of great importance to microphysics schemes. The results shows that the ensemble Kalman filter successfully corrected model errors in microphysical parameters.

1.3.9 Summary of Literature Review

NWP is an initial value problem: it forecasts the atmospheric state by integrating a numerical model with given initial conditions. Due to inadequate observations, our limited understanding in physical processes of atmosphere, and the chaotic nature of the atmosphere flow, uncertainties always exist in modern NWP. Enhancing the

predictability becomes a key issue in improving the skill of NWP.

In this paper, the ensemble forecasting and ensemble based Kalman filter methods, both derived from concepts of the stochastic prediction, are overviewed. It can be concluded as follows:

- Atmospheric motion, as an unstable system, has a finite predictability. NWP is strongly sensitive to the initial conditions. Uncertainties in the model physical parameterization also introduce errors into NWP. Due to strong nonlinearity and the chaotic nature of the atmospheric flow, unpredictable components exist in reality.
- Ensemble forecasting takes uncertainties into account in initial conditions and/or model physical parameterizations to help produce improved forecasts over a single deterministic forecast in NWP and also provide probabilistic forecasts.
- The Ensemble Kalman filter refines the Monte Carlo method and traditional Kalman filter. It uses ensemble forecasts to express the flow dependent error covariance of the forecast estimation. Ensemble Kalman filters present an effective way for data assimilation to improve model initial conditions, while at the same time also take uncertainties into account.

Owing to their advantages in beating the uncertainties and dealing with the nonlinearity in NWP, ensemble forecasting and ensemble-based data assimilation received a lot of attention in the research and operational communities during the last decade. Specifically, they have been applied to improve high-impact weather forecasting.

However, there are issues outstanding. As the ensemble forecasting requires large computational resources, many operational ensemble systems were implemented in

coarser resolutions compared with the high-resolution deterministic weather prediction models. Meanwhile, the small size of the ensemble could cause the underrepresentation problem when generating the background covariance for ensemble-based data assimilation. In addition, with perturbed initial conditions and various physical parameterizations, ensemble forecasts take into account both initial and model errors; however, there has not yet been a consensus regarding which one of these two methods is more efficient for accurate NWP in general. Moreover, the ensemble Kalman filter has many advantages over the current variational data assimilation systems. However, so far, the use of the ensemble Kalman filter in operational forecasts has been in a test phase. More studies are needed to make it a more powerful tool for assimilating real observations. In the meantime, a hybrid variational and ensemble Kalman filter method could be a promising technique in the near future.

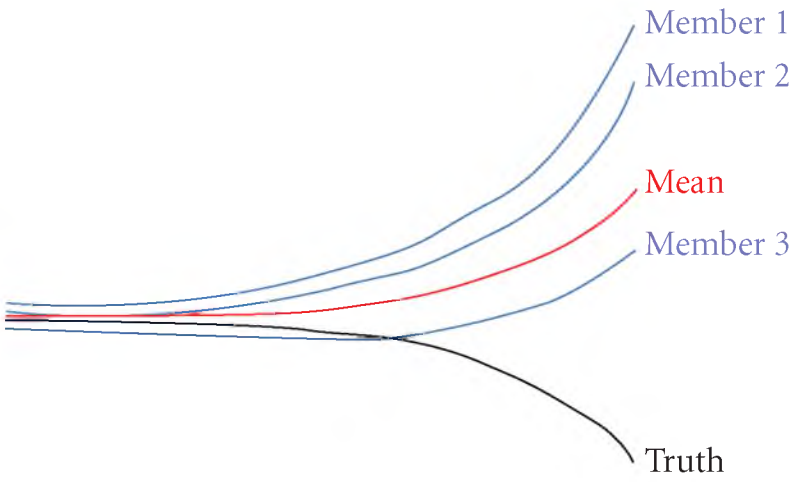


Figure 1.1 Ensemble forecasts with finite member.

CHAPTER 2

ENSEMBLE KALMAN FILTER ASSIMILATION OF SURFACE OBSERVATIONS IN COMPLEX TERRAIN: OSSES AND COMPARISON WITH 3DVAR

2.1 Background

It is critical to provide as accurately as possible near surface atmospheric conditions that are important for the analyses and short-term forecasts, especially for the forecasts in the atmospheric boundary layer (ABL). Although many studies (e.g., Fujita et al. 2007; McCaul and Cohen 2002; Stensrud et al. 2009) have addressed the importance of surface observations for weather forecasting, their use has proved difficult. Currently, only a limited number of surface observations are used in operational NWP. In the National Centers for Environmental Prediction/National Center for Atmospheric Research (NCEP/NCAR) 50-year reanalysis (Kalnay et al. 1996), only surface pressure observations were assimilated. In the North American Regional Reanalysis project (NARR, Mesinger et al. 2006), only surface pressure, wind at 10-meter height level, and relative humidity at the 2-meter level were assimilated into the model; temperature at the 2-m height level was found to be significantly detrimental to forecasts and was not assimilated into NARR. Similarly, no 2-m temperature observations were assimilated into the model for the European Centre for Medium-Range Weather Forecasts (ECMWF)

operational analysis and ECMWF 40-year reanalysis project (ERA 40; Simmons et al. 2004).

Similar problems were also found in the NCEP operational regional analysis system (Rogers et al. 2005). Experiments conducted in the summer of 2003 indicated that the assimilation of surface temperature observations over land in the North American Mesoscale (NAM, formerly Eta) model often degraded forecasts. It was determined that the Eta 3-dimensional variational data assimilation (3DVAR) system, using the step mountain coordinate, had difficulty in limiting the vertical influence of surface observations. The problem remains in the latest version of the operational Non-hydrostatic Mesoscale Model (NMM) core of the Weather Research and Forecasting (WRF) model. In order to meet the needs of weather analysis and forecasting as well as the verification of the National Digital Forecasting Database (NDFD), a 2DVAR scheme is used to assimilate surface observations into a real-time mesoscale analysis (RTMA) system (De Pondeca et al. 2011) at NCEP. This system is operated independently from the operational WRF model. Therefore, except for some selected Mesonet data that are assimilated into the Rapid Update Cycle (RUC; Benjamin et al. 2004) and Rapid Refresh (Benjamin et al. 2011), no other Mesonet data are assimilated into other components of the operational analysis system.

In addition, significant topography poses an extra challenge for surface data assimilation in NWP and data assimilation applications. Meanwhile, limited attention has been given to the use of the surface observations in data assimilation and parameter estimation (Dong et al. 2007; Fujita et al. 2007; Hacker et al. 2007; Hacker and Snyder 2005; Lee et al. 2005; Xie et al. 2005).

Specifically, considerable difficulty has been encountered when assimilating surface observations in complex terrain, as the accuracy of representation of the realistic terrain in NWP is usually limited by both the horizontal and vertical resolution of the forecast model. Because terrain data used in the model must be modified to conform to the model resolution, the variability of the terrain within each grid cell influences the overall resolved orographic and subgrid scale processes. This causes problems in 1) representing the orography in the dynamics and physics of the numerical model and 2) assimilating near-surface data. The latter is referred to as a "misrepresentation" problem and makes the process of quality control and the assimilation of these data complex and difficult.

These problems in near-surface data assimilation have also been well recognized for variational data assimilation methods (e.g., 3DVAR) in operational practice. During the last decade, ensemble Kalman filter (EnKF) techniques have gradually been advanced in the research community. Previous studies (e.g., Meng and Zhang 2008a,b) demonstrate that the EnKF outperforms 3DVAR. However, while all those studies integrate many types of observation in their data assimilation experiments, none emphasizes the assimilation of surface observations. In addition, progress has been made in examining the impact of surface data assimilation on short-range weather forecasts. For instance, Stensrud et al. (2009) assimilated surface observations into the WRF model with an EnKF and found that the analyses reproduced the cold pools beneath the precipitation system. Ancell et al. (2011) demonstrated that the WRF EnKF surface analyses and subsequent short-term forecasts are generally better than the forecasts from the NCEP Global Forecast System (GFS) and North America Model (NAM). In addition, Hacker and Snyder (2005) showed that assimilation of surface observations in a one-dimensional

column model resulted in error reductions throughout the atmospheric boundary layer (ABL). However, in these studies, little work has been reported evaluating whether the EnKF overcomes the existing problems for 3DVAR in surface data assimilation.

In order to initiate this investigation, a series of observing system simulation experiments (OSSEs) are performed with two data assimilation methods: a three-dimensional variational data assimilation (3DVAR), which is widely applied in current operational practice, and an ensemble Kalman filter (EnKF). The purpose of the study is to examine their respective problems and advantages in assimilating near-surface observations, not only to understand the fundamental problems in assimilating surface observations with the current 3DVAR method but also to evaluate the ability of the EnKF to deal with surface observations. Specifically, we perform OSSEs with idealized settings to investigate the problems associated with the assimilation of the surface wind at the 10-m height level (10-m wind hereafter) and temperature at the 2-m height level (2-m temperature hereafter) into the WRF model. With OSSEs in a short-range forecast, we can isolate various factors that affect the use of near-surface observations and thus enhance our understanding of the major factors that could limit our ability to assimilate surface observations.

Section 2.2 briefly describes the WRF model and its 3DVAR and EnKF data assimilation systems. Section 2.3 introduces the setup of the observing system simulation experiments (OSSEs). Sections 2.4 and 2.5 present the results from single observation and multiple observation experiments respectively. Results from both 3DVAR and the EnKF are compared. Section 2.6 examines the impact of the misrepresentation of terrain

in surface data assimilation. Advantages and disadvantages of 3DVAR and the EnKF are further discussed. Section 2.7 provides discussion and summaries of this Chapter.

2.2 Description of WRF Model and Data Assimilation Systems

2.2.1 WRF Model

This study employs an advanced research version 3.1 of the WRF model (ARW WRF). The ARW WRF is based on a Eulerian solver for the fully compressible nonhydrostatic equations, is cast in flux conservation form, and uses a mass (hydrostatic pressure) vertical coordinate. The solver uses a third-order Runge–Kutta time integration scheme coupled with a split-explicit second-order time integration scheme for the acoustic and gravity-wave modes. Fifth-order upwind-biased advection operators are used in the fully conservative flux divergence integration; second- to sixth-order schemes are run-time selectable. The ARW WRF carries multiple physical options for cumulus, microphysics, planetary boundary layer (PBL), and radiation physical processes. Details of the model are referred to Skamarock et al. (2008).

2.2.2 WRF 3DVAR

The WRF 3DVAR system was developed based on the NCAR/Penn State University Mesoscale Model Version 5 (MM5) 3DVAR system (Barker et al. 2004a,b). The 3DVAR provides an analysis $\boldsymbol{x}^a$ via the minimization of a prescribed cost function $J(x)$,

$$J(x) = J^b + J^o = \frac{1}{2}(x - x^b)^T B^{-1}(x - x^b) + \frac{1}{2}\sum_{i=1}^{n}(y - y^o)^T O_i^{-1}(y - y^o) \qquad (2.1)$$

In Eq. (2.1), the analysis $\boldsymbol{x}^a$ represents an a posteriori maximum likelihood (minimum variance) estimate of the true atmospheric state given two sources of information: the background (previous forecast) $\boldsymbol{x}^b$ and observations $\boldsymbol{y}^o$ (Lorenc 1986). The analysis fit to these data is weighted by the estimates of their errors: **B** and **O** are the background and observational error covariance matrices, respectively; $\boldsymbol{y}=\mathbf{H}(\boldsymbol{x})$, and **H** is a linear or nonlinear operator projecting the grid point state $\boldsymbol{x}$ to estimated observations.

The configuration of the WRF 3DVAR is based on a multivariate incremental formulation (Courtier et al. 1994). The preconditioned control variables are stream function, unbalanced potential velocity, unbalanced temperature, unbalanced surface pressure, and pseudo-relative humidity. Users have an option to generate the background error covariance term (**B**) to achieve consistency between the background error term and the model resolution.

In this study, the background error covariance is generated with the so-called NMC (National Meteorological Center, now known as NCEP) method (Barker et al. 2004a,b; Parrish and Derber 1992; Wu et al. 2002) to fit the specific region and season. Statistics of the differences between 24-h and 12-h forecasts valid at 0000UTC and 1200UTC for one month (June 2008) are paired (i.e., a total of 60 samples) to generate the background error covariance using the following steps: 1) Convert the WRF forecast variables to preconditioned control variables, then remove the mean for each variable and each model level; 2) Calculate "unbalanced" control variables and conduct regression analysis to determine multivariate correlations between perturbation fields; 3) Project the perturbations from model levels onto a climatologically averaged (in time, longitude and latitude) eigenvector using empirical orthogonal functions (EOFs) to obtain the vertical

component of the background error covariance; 4) Perform linear regression of the horizontal correlations to calculate a recursive filter length-scale that will be used to calculate the horizontal components of the background error in 3DVAR application.

The **B** term plays an important role in a 3DVAR system. It influences the analysis fit to the background and observations and also defines the domain of influence of observations. Horizontally the background error correlations are assumed to be a Gaussian function:

$$\mathbf{B}(r) = \mathbf{B}(0)\exp(-\frac{r^2}{s^2}) \tag{2.2}$$

where r is the distance between the model grid point and the observation location; s is the length-scale that determines how far the observation information can extend spatially. $\mathbf{B}(0)$ is the value of the background error covariance at the observation location and $\mathbf{B}(r)$ is the background error covariance at the model grid point at a distance r away from the observation location. The observational information is spread using recursive filters (Baker et al. 2004b; Wu et al. 2002), while the vertical relation is represented by applying the empirical orthogonal decomposition technique. Since the **B** in 3DVAR is generated by statistics, it is static throughout the data assimilation experiment.

2.2.3 WRF EnKF

Different from the 3DVAR method, the background error covariance term is estimated using an ensemble of forecasts in EnKF. As mentioned in Section 1.3, significant errors in NWP are often subject to initial condition and model errors. An ensemble is expected to represent the uncertainties by ensemble spreads. One can perturb either the initial condition or the model to generate an ensemble. In practice, different

methods can be applied to account for initial condition (e.g., Singular Vectors, Buizza and Palmer 1995; Fixed Covariance Perturbations, Torn et al. 2006) and model uncertainties (e. g, Stochastic Kinetic-Energy Backscatter Scheme; Berner et al. 2009).

The theory and implementation were introduced in Section 1.3. A description of the EnKF used in this Chapter is briefly given as follows. Readers are also referred to Evensen (2003) and Zhang and Pu (2010) for details. The ensemble mean is supposed to be the best estimate of the true state. The analysis is updated via the equation,

$$\boldsymbol{x}^a = \boldsymbol{x}^f + \mathbf{K}[\boldsymbol{y}^o - \mathbf{H}(x^f)] \tag{2.3}$$

with a Kalman gain matrix,

$$\mathbf{K} = \mathbf{P}^f\mathbf{H}^T(\mathbf{H}\mathbf{P}^f\mathbf{H}^T + \mathbf{R})^{-1} \tag{2.4}$$

where $\boldsymbol{y}^o$ is the observation vector, the superscripts f and a indicate, respectively, forecast (background) and analysis. $\mathbf{H}$ is a linearized observation operator, which relates the model state $\boldsymbol{x}$ and observation $\boldsymbol{y}^o$ by

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \varepsilon \tag{2.5}$$

and $\boldsymbol{\varepsilon}$ is a Gaussian variable with mean zero and covariance $\mathbf{R}$. $\mathbf{P}^f$ is the background error covariance and is estimated using an ensemble of k forecasts $x_k^f(t_i)$ (Evensen 1994),

$$\mathbf{P}^f \approx \frac{1}{K-1}\sum_{k=1}^{K}(\boldsymbol{x}_k^f - \overline{\boldsymbol{x}}^f)(\boldsymbol{x}_k^f - \overline{\boldsymbol{x}}^f)^T \tag{2.6}$$

where the overbars represent the ensemble average. As ensemble forecasts are used in generating the background error term, the background error covariance in the EnKF is flow-dependent.

Research on the EnKF started with Evensen (1994) in oceanography and Houtekamer and Mitchell (1998) in atmospheric science. Their methods can be classified as the EnKF

with perturbed observations. Another type of EnKF is a class of deterministic square root filters (Anderson 2001; Bishop et al. 2001; Whitaker and Hamill 2002), which consists of a single analysis based on the ensemble mean and in which the analysis perturbations are obtained from the square root of the Kalman filter analysis error covariance. Tippett et al. (2003) described serial implementations of the square root filters and argued the square root filter increases efficiency by avoiding the inversion of large matrices.

This study uses an ensemble adjustment Kalman filter (EAKF) data assimilation for WRF developed at NCAR (Anderson et al. 2009) with the Data Assimilation Research Testbed (DART/WRF).

Small ensemble size and model errors affect the performance of the EnKF. Thus, localization and covariance inflation are commonly used in many applications. Specifically, sampling errors are present due to the use of small ensemble size to reduce the computational cost. Spuriously large error covariance estimates between a state variable and a remote observation can be produced by using a small ensemble. Spatial localization is a practical strategy that eliminates the impact of observations beyond a cutoff distance (Anderson 2012; Hamill et al. 2001; Houtekamer and Mitchell 1998). It has been demonstrated that localization can mitigate the spurious correlations to some degree (Hacker et al. 2007; Hamill et al. 2001; Houtekamer and Mitchell 1998, 2001). A covariance inflation increases the prior ensemble estimates of the state variance and can reduce the impact of model error and avoid filter divergence (Anderson 2007; Buehner 2012; Miyoshi 2011). In particular, DART/WRF uses a hierarchical Bayesian approach (Anderson 2007) in which covariance inflation values are adaptively estimated and can vary temporally and spatially.

2.3 Observing System Simulation Experiments (OSSEs)

OSSEs (Atlas 1997; Lahoz et al. 2005) rather than real data assimilation experiments are performed in this study to simplify the surface data assimilation and also to examine the key factors that affect it. OSSEs are commonly carried out to check a data assimilation system before conducting a real data experiment. In an OSSE, synthetic "true" states and measurements are obtained by introducing random errors in the states. OSSEs make the direct comparison between model simulations and "true" states possible.

A time period of 0000 UTC to 1200 UTC 5 June 2008 was chosen arbitrarily for a case study. During this period, a cold front was passing over the western US (complex terrain), and its eastern extension was a stationary front. It was moving southeastward and its eastern part was approaching the Great Plains. Under the influence of the front, a low-level jet (LLJ) was evolving over the Great Plains. For OSSEs, the nature run (i.e., the "truth") was generated by integrating the ARW WRF model for a 12-h period, initialized from the NCEP NAM analysis at 0000 UTC 5 June 2008. Figure 2.1 shows the weather map of 0000 UTC 5 June 2008 from the nature run. The cold front and LLJ were clearly revealed at this time.

Model terrain is the same in the nature run and control run in most of the following numerical experiments except for a set of experiments discussed in Section 2.6 that examines the effect of the terrain misrepresentation. Model grid spacing (horizontal resolution) was set at 27-km, which includes 135 and 195 grid points in a south-north and west-east direction, respectively. The model's vertical structure consisted of 36 η levels in the terrain-following hydrostatic-pressure coordinate with the top of the model set on 50 hPa, where $\eta = (p_h - p_{ht})/(p_{hs} - p_{ht})$. While p_h is the hydrostatic component of the

pressure, p_{hs} and p_{ht} refer to pressure values along the surface and top boundaries, respectively. The η levels were placed closely together in the low-levels (below 500hPa) and are relatively coarsely spaced above.

In order to account for the initial and boundary errors, the control run was randomly chosen to begin with a 6-h forecast valid at 0000 UTC made during the first week of June 2008 (e.g., 1 June 2008, which eliminates errors in the diurnal variation but still produces enough differences between the "control" and the "truth"), initialized by the NCEP Global Forecast System (GFS) Final (FNL) analysis at 1 x 1 degree resolution. As shown in Figure 2.1b, the control run completely missed the cold front and LLJ. The differences between the initial conditions of the "nature" run and "control" run were sufficient to demonstrate the effect of the data assimilation.

Since errors can also exist in numerical models, different physical parameterizations were used for the nature run and control run to account for the model errors. Specifically, model physics options used in the nature run include the Yonsei University (YSU) PBL scheme (Hong and Pan 1996), the thermal diffusion land surface scheme, the Lin microphysical scheme (Chen and Sun 2002; Lin et al. 1983), the Kain–Fritsh cumulus parameterization scheme (Kain and Fritsch 1993), the longwave Rapid Radiative Transfer Model (RRTM, Mlawer et al. 1997) and the Dudhia shortwave parameterization model (Dudhia 1989). In the control run, the Mellor–Yamada–Janjic (MYJ) PBL scheme (Mellor and Yamada 1982), the Noah land surface model (Chen and Dudhia 2001), as well as the WRF single-moment 6-class microphysics (WSM6) scheme (Hong and Lim 2006) are used while other physics are the same as those used in the nature run.

For the experiments with both the EnKF and 3DVAR, simulated hourly surface observations were generated by interpolating the nature run to the surface station locations (as shown in Figure 2.2) with unbiased random errors, which are not larger than the statistics of the observational errors. Consistent with the statistics of a large sample of the data and Hacker and Snyder (2005), the observational variances were specified as 1.0 K^2 and 2.0 m^2 s^{-2} for 2-m temperature and 10-m u and v components, respectively. For the data assimilation experiments, 2-m temperature and/or 10-m winds are assimilated into the ARW WRF model.

All the EnKF experiments used 32 members. Ensemble initial perturbations were obtained by the fixed covariance perturbations (FCPs, Torn et al. 2006) 6 h prior to the data assimilation. In FCPs, ensemble perturbations were derived by drawing random perturbations from the 3DVAR system, which were scaled by 1.5 to ensure the ensemble spread is comparable to the root-mean-square (RMS) error of the ensemble mean (a similar scale was used in Torn and Hakim 2008, 2009). All members were then spun-up by running forward for 6 h until the observations were available at the beginning of the data assimilation. Since 3DVAR experiments commonly use a 6-h forecast as a first guess for the data assimilation experiment, we chose a short ensemble spin-up period in the EnKF to ensure a fair comparison. In order to examine the influence of observations and the impact of different background terms from 3DVAR and the EnKF on surface data assimilation, experiments were first conducted by assimilating observation(s) from a single observation station. The experiments with multiple observations then were carried out.

2.4 Results of Single Observation Experiments

One observation station over mountainous terrain (41.04°N, 112.98°W) and one station over flat terrain (38.0°N, 85.0°W) were randomly selected for single-observation experiments to represent stations in different terrain. All single station assimilations were conducted at 0000 UTC 5 June 2008. The observations were generated from nature run with random observational errors added, and representative errors caused by horizontal interpolation from the model grids to observation locations. Then the observations were assimilated into the data assimilation experiments.

2.4.1 3DVAR

Figure 2.3 shows the 3DVAR analysis increments of temperature and u and v wind components at the lowest model level over the mountainous area by assimilating temperature only, wind (u and v components) only, and both temperature and wind. Apparently, the analysis increments showed large-scale features and spread over the mountainous regions. Figure 2.4 illustrates the analysis increments in flat terrain at the lowest model level. Except for the differences in the phase of increment shapes due to the sign of the innovation vector at the observation point (**O-B**), the major features of the analysis increments were similar to those in complex terrain (Figure 2.3). Note that the analysis increments of the u (or v) wind component from assimilating temperature only were much smaller than those from assimilating the u and v components only. Therefore, Figures 2.3e, 2.3h, 2.4e, and 2.4h are similar to 2.3f, 2.3i, 2.4f, and 2.4i, respectively.

Figure 2.5 illustrates the correlation and cross-correlation functions for multivariate optimal interpolation analysis derived using the geostrophic increment assumption (after

Gustafsson 1981 and Kalnay 2003). It demonstrates the response of the analysis increment from one variable (e.g., a geopotential or thermodynamic variable such as temperature) to another variable (e.g., u and v wind components), or vice versa. Comparing Figures 2.3 and 2.4 with Figure 2.5, it is apparent that the shapes of the analysis increments from 3DVAR in both mountainous and flat terrain followed classical correlation and cross-correlation functions of variables prescribed by the geostrophic increment assumption. If we count the changes of contour shapes due to the linear combination of the variables (i.e., Figure 2.3e and 2.4e corresponding to the sum of Figure 2.5e and 2.5f and so on), the shapes of the analysis increments revealed by Figures 2.3 and 2.4 were equivalent to those structures in Figure 2.5, showing strong dependence of the 3DVAR analysis increments on the prescribed correlation functions in the background error covariance term.

In addition, as shown in Figure 2.3, the 3DVAR analysis increments from the observation station within the mountain valley area had been spread across the mountains. Since it is expected that the air temperature and wind conditions can be inhomogeneous over the mountain valley and cross-mountain areas, the cross-mountain analysis increments could be unrealistic.

In 3DVAR, the influence of a single observation on its surrounding area is determined by a horizontal correlation length-scale in the background error covariance term (Eq. 2.2). Figure 2.6 shows the analysis increments of temperature at the lowest model level from the assimilation of both 2-m temperature and 10-m wind with various background correlation length-scales. A simple sensitivity experiment indicates that a very small length-scale (25% of the default value that was specified in the NMC method

when generating the **B** term for this study) is needed in order to avoid unrealistic cross-mountain analysis increments. However, with such a small length-scale, the analysis increment and the influence of the single observation could be minimized. Therefore, it is not realistic to use a very small length-scale. Meanwhile, since a relatively large length-scale has to be used, the 3DVAR method typically can have problems in assimilating data in complex terrain. Therefore, the definition of length-scale is an influencing factor that could impose significant impact on the near-surface data assimilation with 3DVAR in the regions of complex terrain. A sensitivity study is necessary in order to determine optimal length-scale for 3DVAR.

2.4.2 EnKF

With the EnKF, the analysis increments resulted from observations at a single observational station in mountainous and flat terrain are shown in Figures 2.7 and 2.8 (corresponding to Figures 2.3 and 2.4 for 3DVAR), respectively. Compared with Figures 2.3 and 2.4, the analysis increments from the EnKF were much more localized. They also had less correspondence with the correlation functions shown in Figure 2.5 since the EnKF defines the background error covariance flow-dependently using ensemble forecasts, Figure 2.9 shows ensemble spreads of temperature and u and v wind components over the area surrounding the observational station. The shapes and structures (in terms of magnitudes and gradients of the increment contour lines) of the analysis increments of temperature and u and v wind components from the EnKF corresponded well to their ensemble spreads. Specifically, large analysis increments were

more associated with large ensemble spread (uncertainties) rather than the distance between the observational location and a specific grid point.

More importantly, the analysis increments of EnKF in complex terrain resulting from the assimilation of observations in the valley remained inside the valley, unlike those of 3DVAR. No cross-mountain analysis increment was found (Figure 2.7). In addition, similar to Figures 2.3e and 2.3f (2.3h and 2.3i), Figures 2.7e and 2.7f (2.7h and 2.7i) were very similar.

The spatial range of influence from an observation in the EnKF can be limited by the specification of the horizontal localization scale. The sensitivity test of analysis increments to the horizontal localization radius proved that analysis results are sensitive to the horizontal localization radius (Figure 2.10). However, in the EnKF, even with a large localization radius (Figure 2.10c), the analysis increments from the valley station spread widely through the surrounding area but still remain mostly inside the valley.

Overall, the major differences between the 3DVAR and the EnKF shown by the above experiments can be attributed to differences in their background error terms. Since 3DVAR uses a static background term, analysis increments from a single observation tend to be similar in various cases as they depend heavily on the prescribed correlation functions. The EnKF's flow-dependent background term enables the observations to have a greater influence on the areas where the ensemble spreads are larger. Figure 2.11a compares the structures of the estimated background error standard deviation of the streamfunction in 3DVAR (static) and the EnKF averaged over the entire data assimilation period. In 3DVAR, the background error variance is homogeneous in each statistical bin and has no correlation with terrain and the synoptic situation. In contrast,

error variances in the EnKF reflected the structure of the synoptic system. As shown in Figure 2.11b, large variances aligned well with the LLJ, indicating the flow-dependent nature of the EnKF background term.

2.5 Results of Multiple Observations Experiments

Surface observations have high spatial and temporal resolution and are among the most widespread observations of the lower atmosphere. The above results from assimilating observations from a single station show advantages of EnKF over 3DVAR in assimilating near-surface observations. This section further explores whether the EnKF outperforms 3DVAR in short-range weather forecasting if only surface observations are assimilated.

Observations from multiple stations (Figure 2.2) were assimilated to assess the impact on the short-range forecasts as well as to examine and compare the ability of both the 3DVAR and the EnKF to extend the information from single-level surface observations to the ABL. For both experiments of 3DVAR and EnKF, the data assimilation was performed for the period of 0000 UTC to 0600 UTC 5 June 2008, assimilating 2-m temperature and 10-m wind in a total of seven hourly data assimilation cycles. Then, 6-h forecasts followed the data assimilation.

Sensitivity experiments were first conducted to determine the optimal length-scale for the 3DVAR and vertical and horizontal localization scales for the EnKF. It was found that the default value of the horizontal and vertical length-scale, as defined by the NMC method, performed best for the 3DVAR method in terms of obtaining minimum root-mean-square errors (RMSEs) for both analyses and forecasts. Three sets of experiments

with various radii of maximum vertical localization (i.e., 1000 m, 3000 m, 5000m) for the EnKF were performed. Figure 2.12 shows the time-height RMSEs of wind speeds and temperatures with respect to the nature run over key LLJ and frontal regions. The 3000-m radius of maximum vertical localization produced the best analyses/forecasts. Similar sensitivity experiments were also conducted to determine the optimal horizontal localization scale. Among several tested options (e.g., 120-km, 240-km, 360-km), a half-radius of the maximum horizontal localization of 240 km was selected.

Based on the aforementioned sensitivity experiments, optimal results from the 3DVAR and the EnKF were further discussed to compare their relative performance.

Figures 2.1c and d show the wind direction and speed at 50 m above ground level (AGL) and temperature at 700 hPa at 0000 UTC 5 June 2008 after the first cycle of data assimilation with the 3DVAR and the EnKF. Compared with the corresponding fields in the nature run and control run (Figures 2.1a and b), the 3DVAR reproduced the LLJ over the Great Plains but missed the cold frontal system in complex terrain. The EnKF, however, reproduced most parts of the cold front. The LLJ was also well reproduced.

At the end of the data assimilation cycle, namely, at 0600 UTC 5 June 2008 (Figure 2.13), 3DVAR results showed only part of the front while sharing almost the full range of the LLJ but with weaker intensity in terms of wind speed. The EnKF captured the whole front and more accurate intensity in terms of wind speed magnitude and the area of coverage of the LLJ. Overall, the EnKF performed better for both the LLJ and cold front cases. Further evaluations were conducted for the LLJ and cold front systems as follows.

2.5.1 LLJ Over the Great Plains

We first evaluated the representation of the LLJ over the Great Plains during the analysis and forecast periods. Following Whiteman et al. (1997), the definition of the LLJ is when maximum wind speed reaches 12 ms^{-1} with a falloff value greater than 6 ms^{-1} from the wind speed maximum upward to the next wind speed minimum at or below the 3000 m above ground level (AGL). With the assimilation of surface observations, both the 3DVAR and the EnKF were able to reproduce the LLJ in the WRF model (e.g., Figures 2.1 and 2.13).

Figure 2.14 compares the wind profiles averaged over the key region of the LLJ from different experiments for 0000 UTC (beginning of the data assimilation cycle), 0600 UTC (end of the data assimilation cycle), and 1200 UTC (end of the forecast). It shows that both the 3DVAR and the EnKF resulted in an improved analysis and forecast of the LLJ compared with the control run. Specifically, at the beginning of the data assimilation (0000 UTC), the EnKF results were clearly better than 3DVAR in terms of representing the magnitude of the mean wind speed, especially for the low level (up to 1000 m). At the end of the data assimilation cycle (0600 UTC), both the 3DVAR and the EnKF resulted in mean wind speeds that were close to the nature run at heights below 500 m. However, the maximum mean wind speed height was near 500 m in 3DVAR, while it was higher than 500m in both the EnKF and the nature run. After 6-h forecast (1200 UTC), results in 3DVAR and the EnKF were still closer to the nature run than those in the control run. Compared with the 3DVAR, the EnKF analysis led to a better forecast in terms of the vertical structure of the mean wind speed and the height of the maximum wind speed over the LLJ area.

2.5.2 A Cold Front Over Complex Terrain

A cold front was passing over the Intermountain West (complex terrain) during the study period. Figure 2.15 shows the time-latitude cross-section of temperature averaged over the main frontal area (6^{o} longitude ranging from 114^{o} W to 108^{o} W) at 500 m AGL. In contrast to the nature run, the control run missed the major cold front. With both the 3DVAR and the EnKF, the front was reproduced, although the temperature over the frontal region in both the 3DVAR and the EnKF experiments was higher than that in the nature run. However, compared with the 3DVAR, the temperature in the EnKF experiment was closer to that in the nature run over the frontal area.

To further compare the 3DVAR and the EnKF experiments, Figure 2.16 illustrates the differences in the RMSEs (data assimilation experiments against the nature run) between the EnKF and the 3DVAR for both the temperature and wind fields (the negative numbers indicate that the analysis or forecast errors in the EnKF are smaller than those in 3DVAR). Here, the RMSEs were calculated over all observation stations in the key frontal region. The figure clearly reveals that the EnKF outperformed the 3DVAR at all levels in both the analysis and forecast periods.

2.6 Impact of Terrain

Terrain mismatch—namely, the discrepancy between model and realistic terrain heights—is common in numerical models in complex terrain due to the limitation of model resolution in resolving detailed terrain features. In order to test the impact of this type of terrain mismatch on surface data assimilation, we performed an additional set of experiments, in which terrain heights were perturbed by using coarser resolution terrain

data (2-degree versus 10-minute resolution) in the control and data assimilation experiments. A common way to deal with the terrain misrepresentation is to reject the data over the area where discrepancies between the model and realistic terrain are large. Therefore, another set of experiments was conducted in which we rejected the data when terrain differences between the experiments and nature run were greater than 25 m. Figure 2.17 shows the stations where the surface data were rejected in data assimilation. Since observations at many stations were rejected in the mountainous region, the length-scale in 3DVAR and the horizontal localization scale in EnKF needed to be adjusted to adapt to changes of the density of observations being assimilated. Sensitivity experiments were conducted with increased length scales in the 3DVAR and localization scales in the EnKF to determine the new optimal scales. By comparing the RMSEs from these sensitivity experiments, the length (localization) scale that produced the smallest RMSEs in 3DVAR (EnKF) was chosen as a new optimal scale to produce data assimilation results for further comparison.

Figure 2.18 illustrates results from the EnKF. Compared with the RMSEs of temperature and wind in the boundary layer over the key frontal region (complex terrain) in the original (default) experiment (as mentioned in Section 2.5, without perturbing the terrain heights and data rejection; Figures 2.18a and b), the EnKF analysis and subsequent forecasts with perturbed terrain heights achieved very similar accuracy, as the discrepancies in RMS errors between two experiments (Figures 2.18c and d) were almost negligible. This result indicates that the EnKF has a good ability to handle the terrain mismatch. However, when the data were rejected over the areas of terrain mismatch, the data assimilation and forecast results were degraded (Figures 2.18e and f). Since there are

fewer observation stations in complex terrain in the western US than in the eastern US, the degraded analysis and forecast here can be attributed to the lack of data over complex terrain. Outcomes from the 3DVAR experiments showed that the data assimilation results were degraded under mismatched terrain compared with the results before perturbing the terrain (results not shown), indicating that 3DVAR has less ability to handle the mismatched terrain than the EnKF does. In addition, under the data rejection case, 3DVAR results were even worse than those from the EnKF (Figure 2.19).

2.7 Summary and Discussion

Surface observations are known as the main source of conventional observations. However, their use in modern numerical weather prediction, especially in complex terrain, remains a challenge. In this section, a series of observing system simulation experiments (OSSEs) are performed with two popular data assimilation methods, a 3DVAR and an EnKF. The problems and advantages of both methods in assimilating near-surface observations are examined.

Even with a single case, this study provides many details that explain the differences between the 3DVAR and the EnKF in the context of surface data assimilation. Results from the assimilation of surface temperature and wind from a single observation station demonstrate that there are fundamental problems in assimilating surface observations in complex terrain with the 3DVAR. Specifically, the analysis increments from a valley station can unrealistically spread over the mountain areas even with a reasonable specification of the horizontal background length-scale. The EnKF can overcome some of these limitations through its flow-dependent background error term. Overall, major

discrepancies between the 3DVAR and EnKF in the single observation experiment can be attributed to their different ways of defining the background error terms. Since the 3DVAR uses a static background covariance, analysis increments from a single observation tend to be similar in various cases as they depend heavily on the prescribed correlation functions. Owing to its flow-dependent background error covariance term, the EnKF enables observations to have more influence on areas where the ensemble spreads (forecast uncertainties) are larger. In addition, due to sampling errors with limited ensemble size, data assimilation results from the EnKF are sensitive to the choice of the horizontal and vertical localization scales.

More comprehensive comparisons are conducted using a synoptic case with two severe weather systems: a front over complex terrain in the western US and a low-level jet over the Great Plains. It is found that both the 3DVAR and the EnKF are capable of extending information from surface level to the atmospheric boundary layer. Over flat terrain, the EnKF does better in terms of the analysis and forecast of the low-level jet system while the 3DVAR also simulates the low-level jet system generally well. Over complex terrain, the EnKF performs much better than the 3DVAR in general. Specifically, the EnKF has a better ability to handle surface data under terrain misrepresentation. Since a common way to deal with terrain misrepresentation is to reject data over the area where discrepancies between the model and the actual terrain are large, a data-rejection experiment is performed. However, since observations are sparse in complex terrain, data rejection results in degraded analyses and forecasts, suggesting that this may not be the best solution for dealing with errors due to model terrain representation.

It should be noted that the OSSEs performed in this section are in idealized settings in order to isolate various factors that affect the assimilation of near-surface observations. In addition, the terrain data used in OSSEs are all from model terrain, which is much smoother than the actual terrain. Therefore, while the results in this study can help us understand what factors limit our ability to assimilate surface observations, the real data assimilation and prediction problems are expected to be more complicated. The following sections further investigate the problems and challenges in assimilating surface observations in complex terrain with real observations and realistic terrain.

In addition, hybrid 3DVAR and EnKF data assimilation schemes (e.g., Hamill and Snyder 2000; Wang et al. 2008) have been developed in both research and operational communities. These hybrid data assimilation systems are expected to overcome some deficiencies of the pure 3DVAR method. Meanwhile, a four-dimensional variational data assimilation (4DVAR) method is also widely used for operational prediction. Comprehensive evaluation of these advanced methods in assimilating surface observations is desired and will be addressed in the future study.

Furthermore, model errors due to coarser resolution grids, terrain misrepresentation, and land-atmosphere thermal coupling could be issues in near-surface data assimilation. In order to improve our understanding of the problem, the following chapters will discuss forecast error characteristics and ensemble data assimilation in real data context.

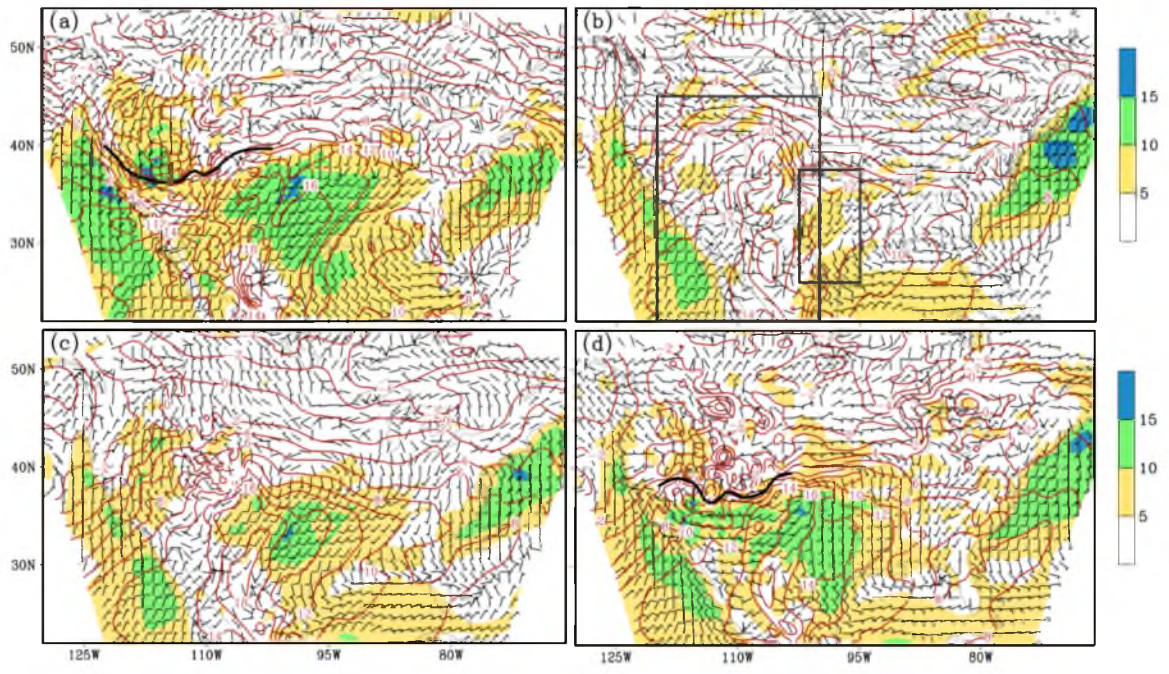


Figure 2.1 Synoptic maps at 0000 UTC 05 June 2008 for (a) nature run, (b) control run, (c) 3DVAR analysis, and (d) EnKF analysis. The wind barbs denote the wind fields at 50 m AGL. Shaded contour represents the wind speeds. The contours show the temperature on 700hPa (2°C interval). The thick black line denotes the cold front. The large box denotes the key front region that used for calculations in Figures 2.12, 2.16, 2.18, and 2.19. The small box denotes the key LLJ region that was used for the calculation in Figure 2.12.

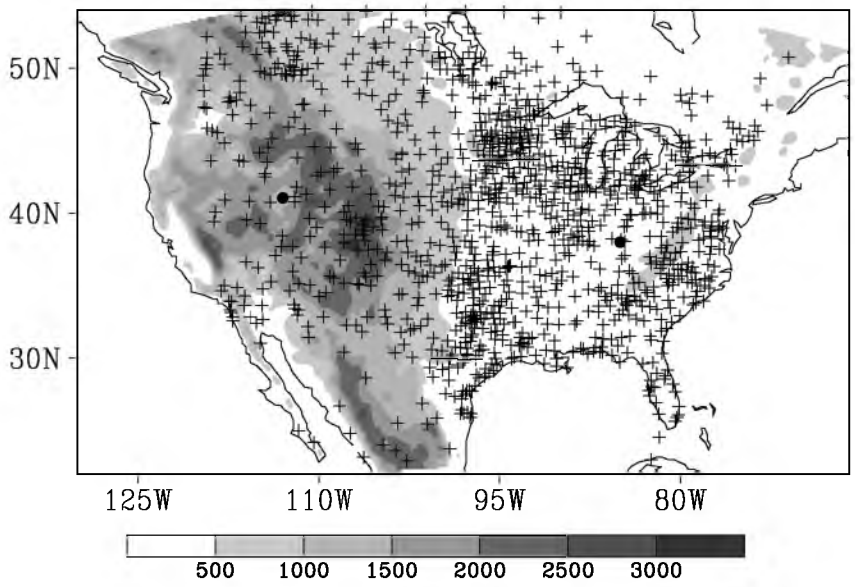


Figure 2.2 Distribution of surface observation stations (plus signs) and terrain heights (shaded contour; unit: m). The two solid dots denote the locations of the two stations in single station experiments.

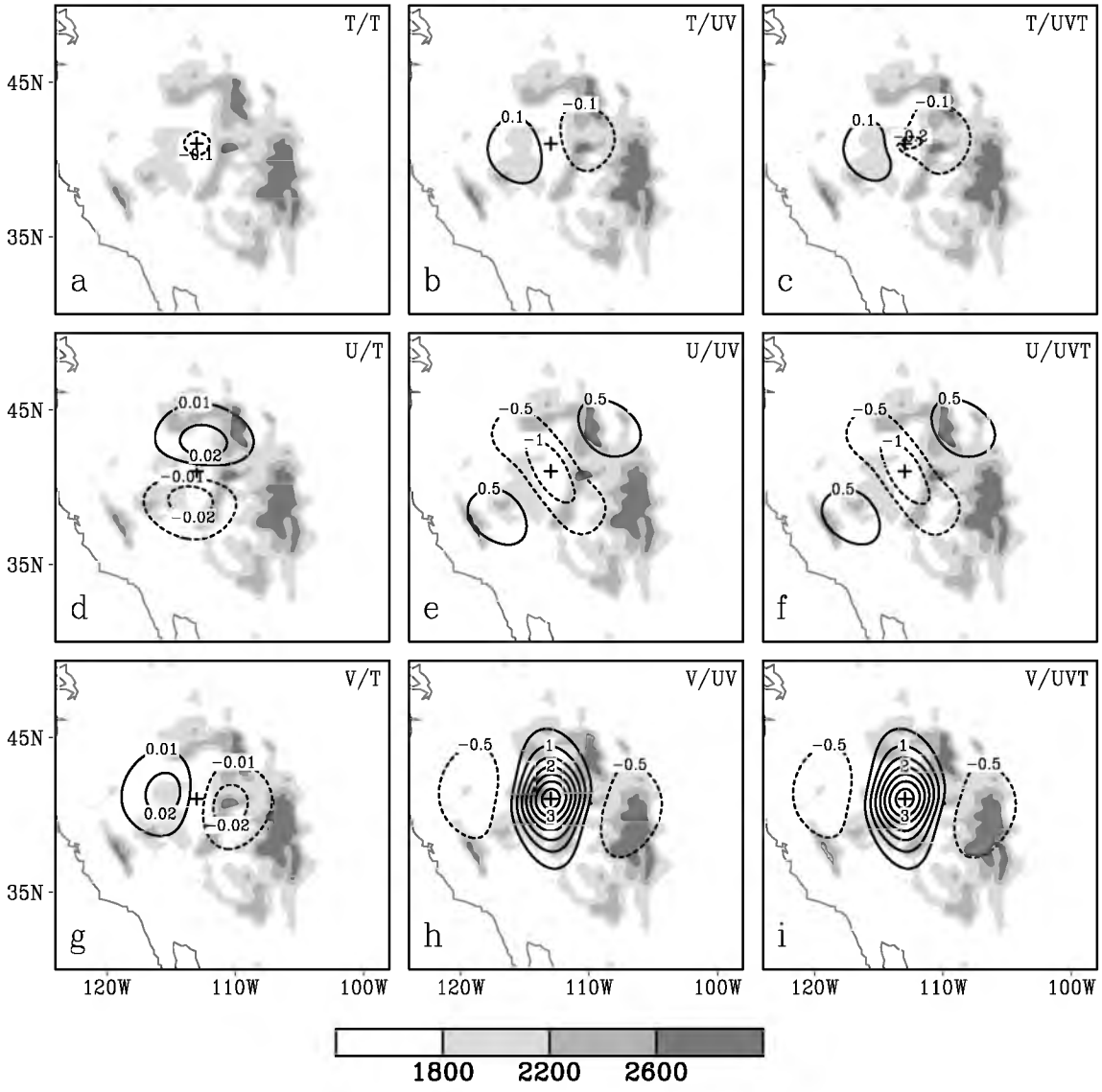


Figure 2.3 The 3DVAR analysis increments of temperature (K; top row), u-component (ms^{-1}; middle row), and v-component (m s^{-1}; bottom row) of wind at lowest model level with assimilation of 2-m temperature (left column), 10-m winds (middle column), and both 2-m temperature and 10-m winds (right column) from a single observation station over complex terrain. The shaded contours show the terrain heights (unit: m). "+" denotes the location of the observation station.

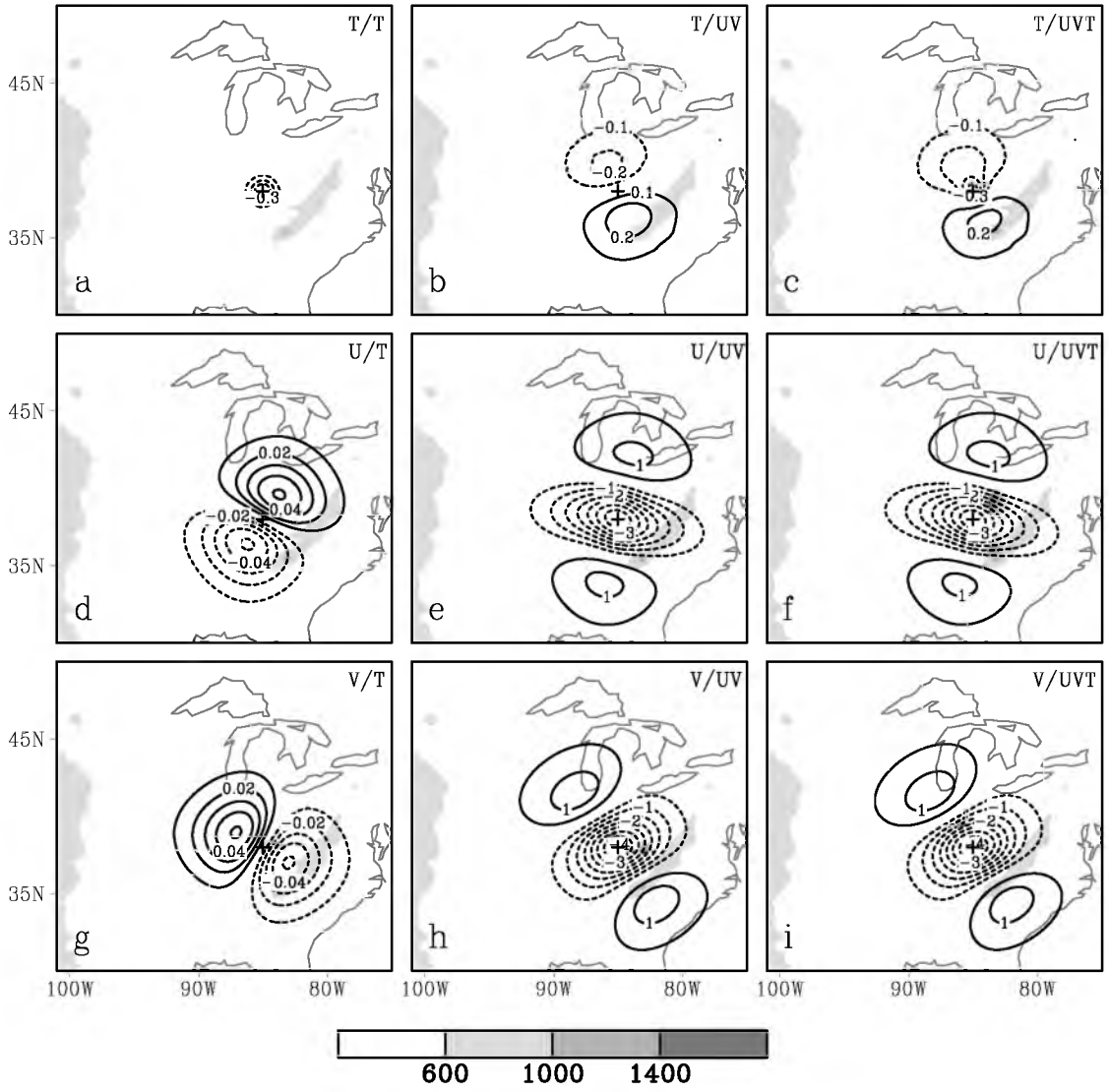


Figure 2.4 Same as Figure 2.3, except over flat terrain.

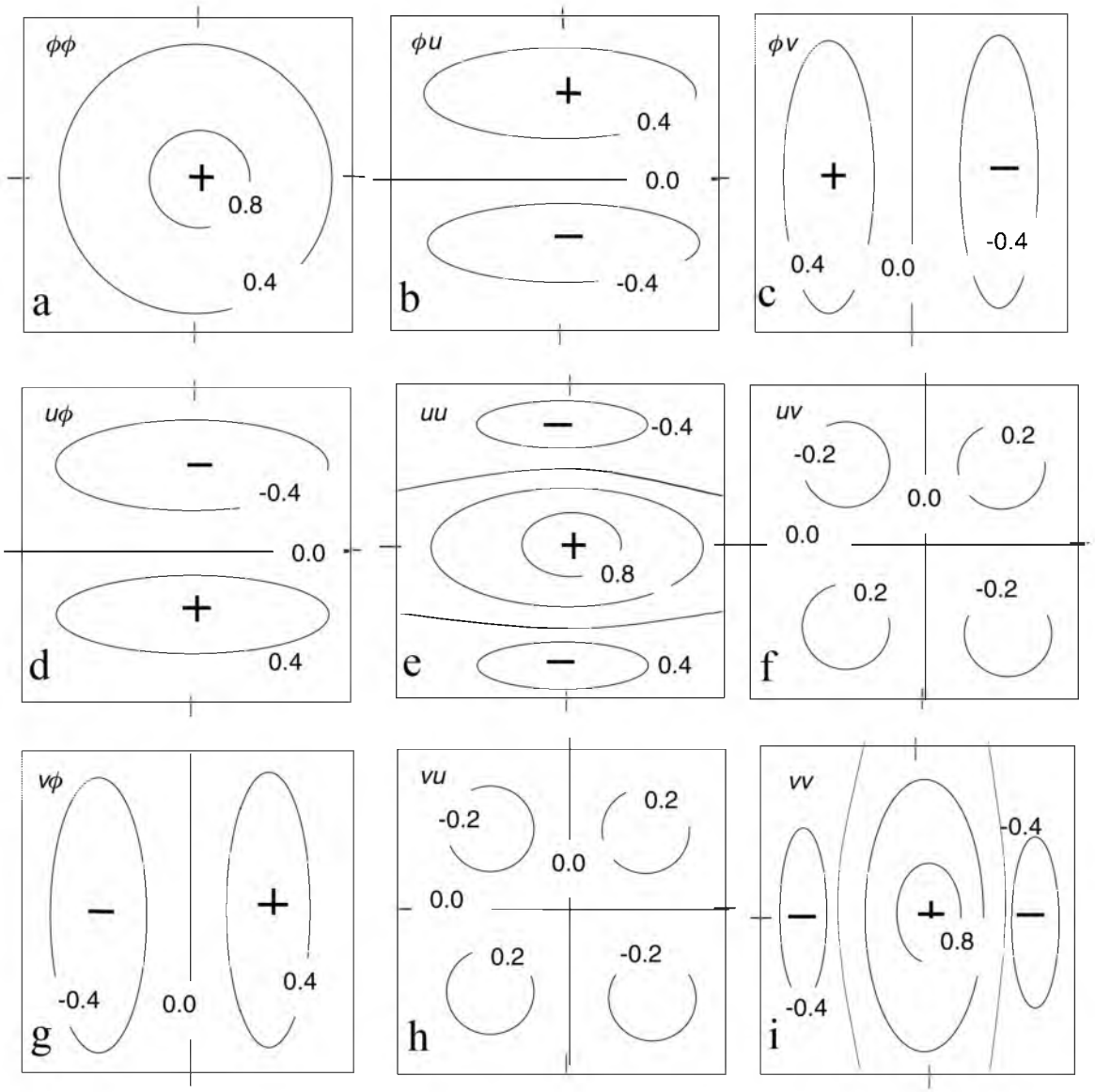


Figure 2.5 Schematic illustration of the correlation and cross-correlation functions for multivariate optimal interpolation analysis derived using the geostrophic increment assumption (courtesy of Gustafsson 1981 and Kalnay 2003). "ϕ" is thermodynamic variable related to the temperature. u and v denote the horizontal components of wind.

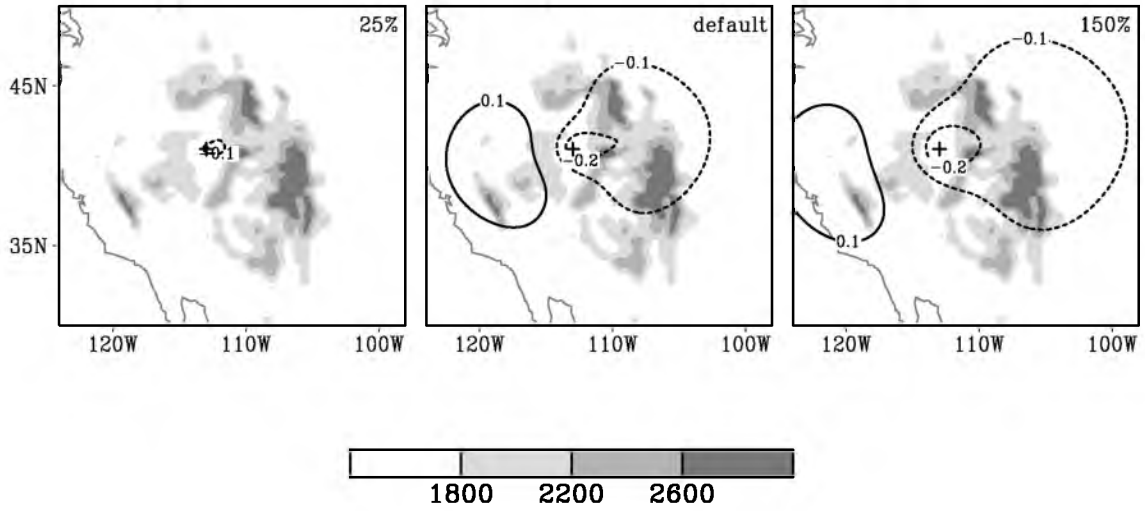


Figure 2.6 The 3DVAR analysis increments of temperature (K) with assimilation of both 2-m temperature and 10-m winds from a single observational station over complex terrain using different horizontal correlation length-scales: a default value (middle), 25% of the default value (left), and 150% of the default value (right). The shaded contour shows the terrain heights (unit: m).

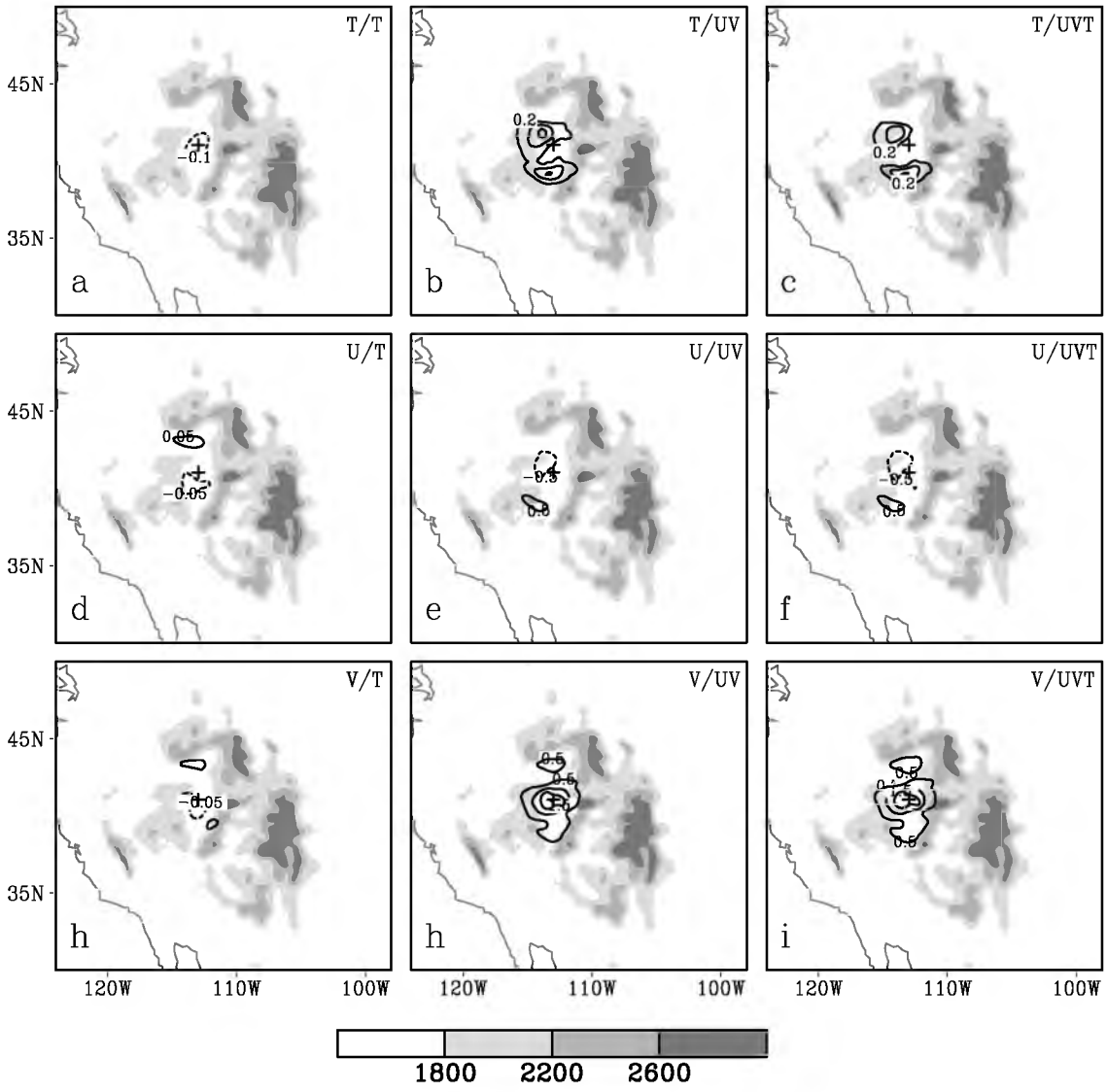


Figure 2.7 Same as Figure 2.3, except for the EnKF analysis increments. The half radius of the horizontal localization used in the experiments is 320km. No vertical localization is applied.

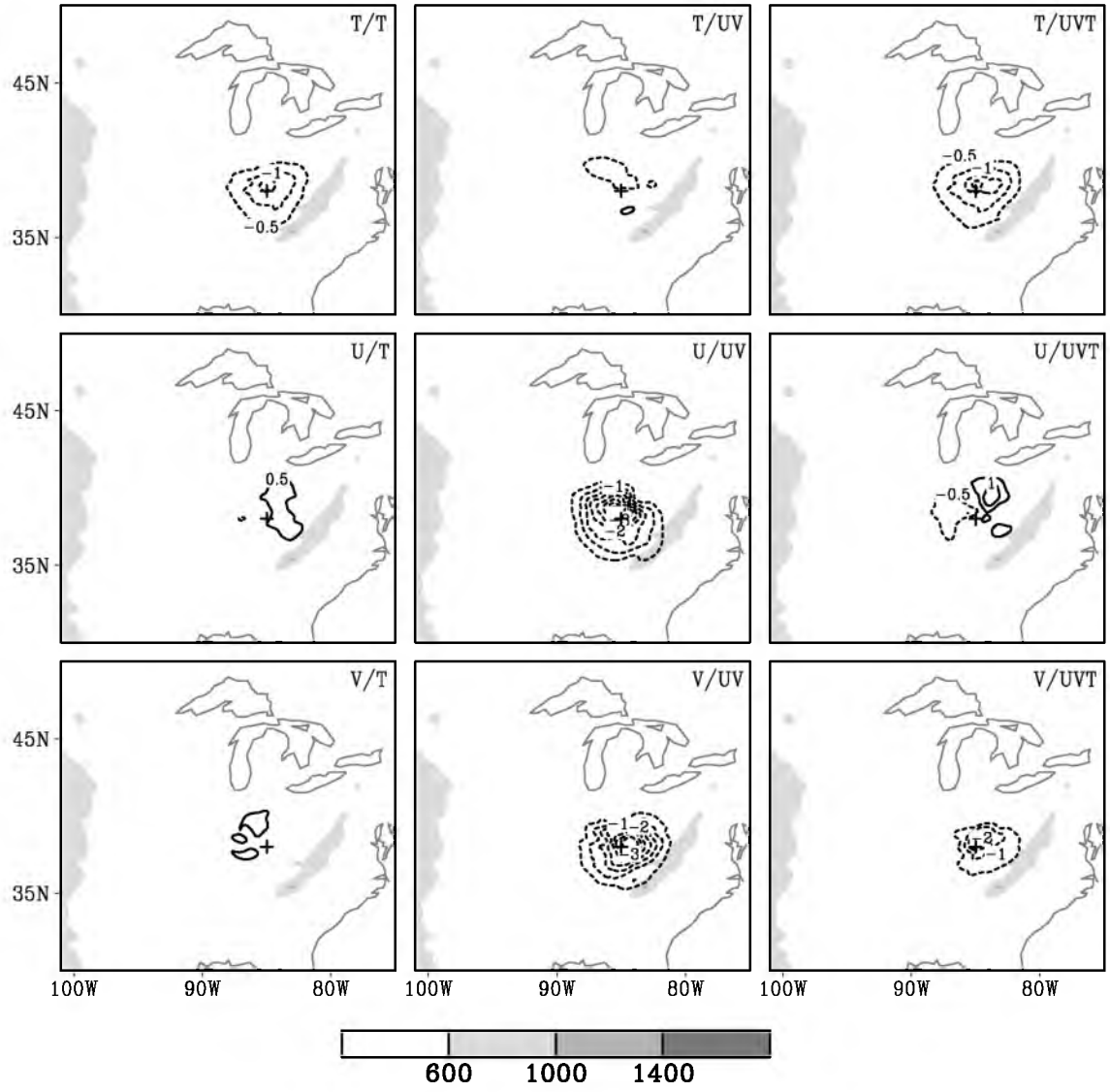


Figure 2.8 Same as Figure 2.7, except over flat terrain.

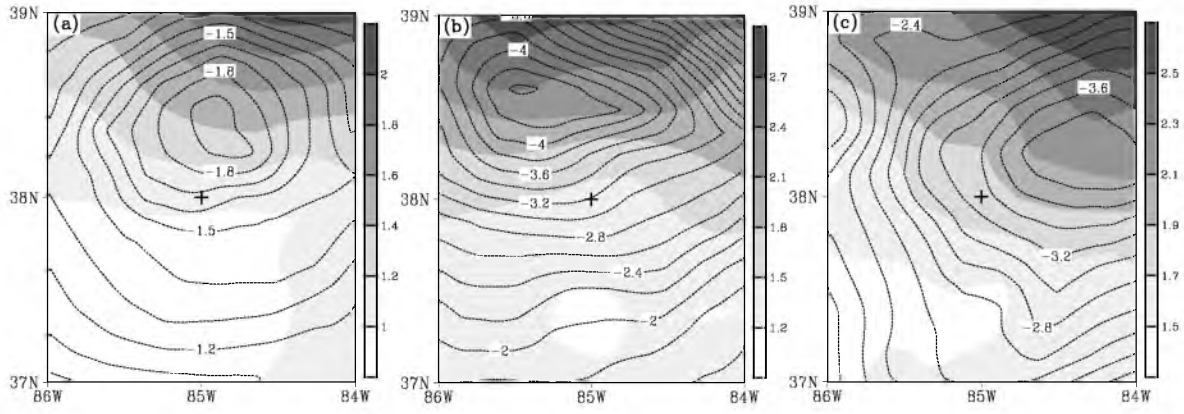


Figure 2.9 Ensemble spread (shaded) and analysis increments (contour) for (a) temperature (0.1 K interval), (b) u-component (0.2 m s^{-1} interval), and (c) v-component (0.2 m s^{-1} interval) with assimilation of both 2-m temperature and 10-m wind using EnKF. The "+' sign denotes the observational location.

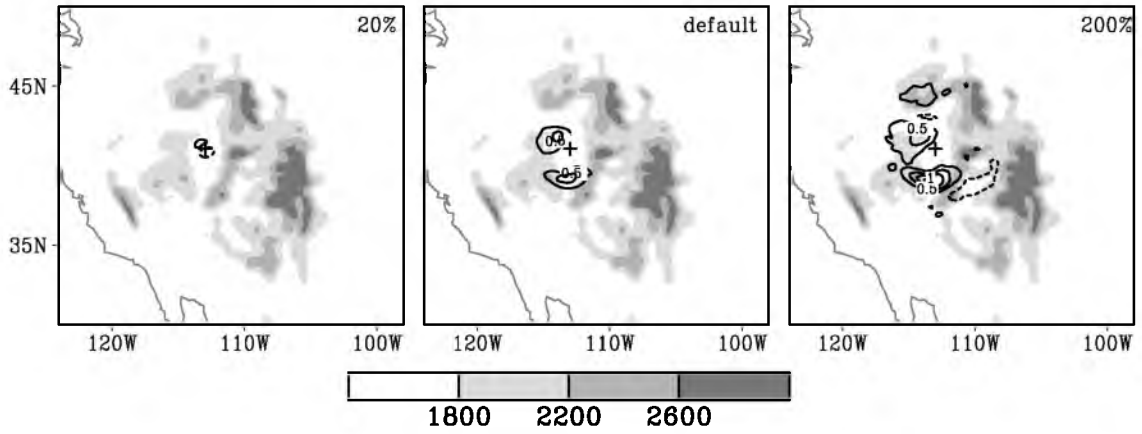


Figure 2.10 Same as Figure 2.6, except for EnKF with different horizontal localization scales. The default value of half radius of the horizontal localization is 320 km (middle). The smaller (left) and larger (right) scales are 20% and 200% of the default value, respectively.

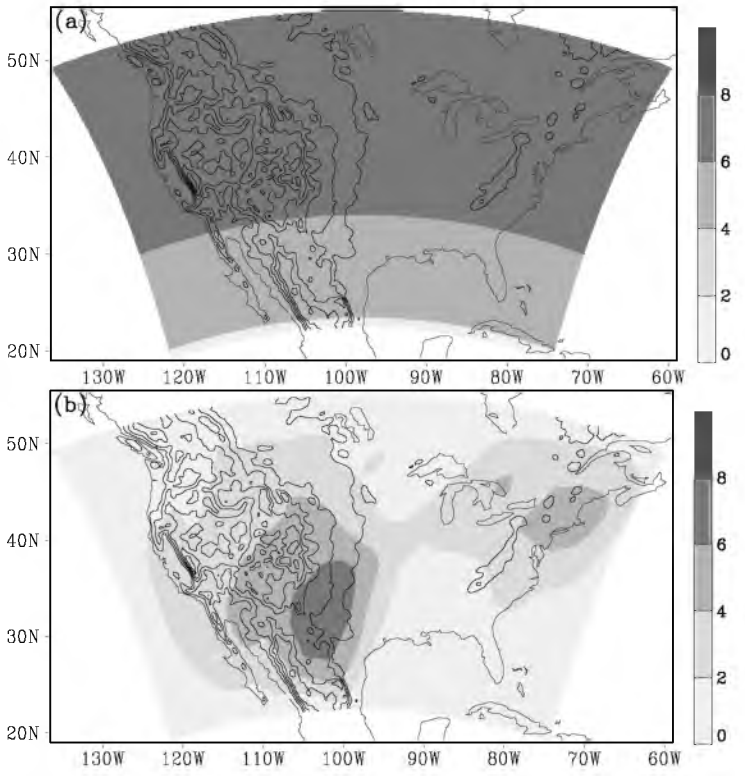


Figure 2.11 Estimated background error standard deviation of the streamfunction (shaded contour; unit: 10^5 m^2 s^{-1}) in 3DVAR (a; static in time) and EnKF (b; averaged over the data assimilation period [0000 UTC to 0600 UTC 5 June 2008]) near 800 m AGL. Contour lines denote the terrain heights (interval: 500 m)

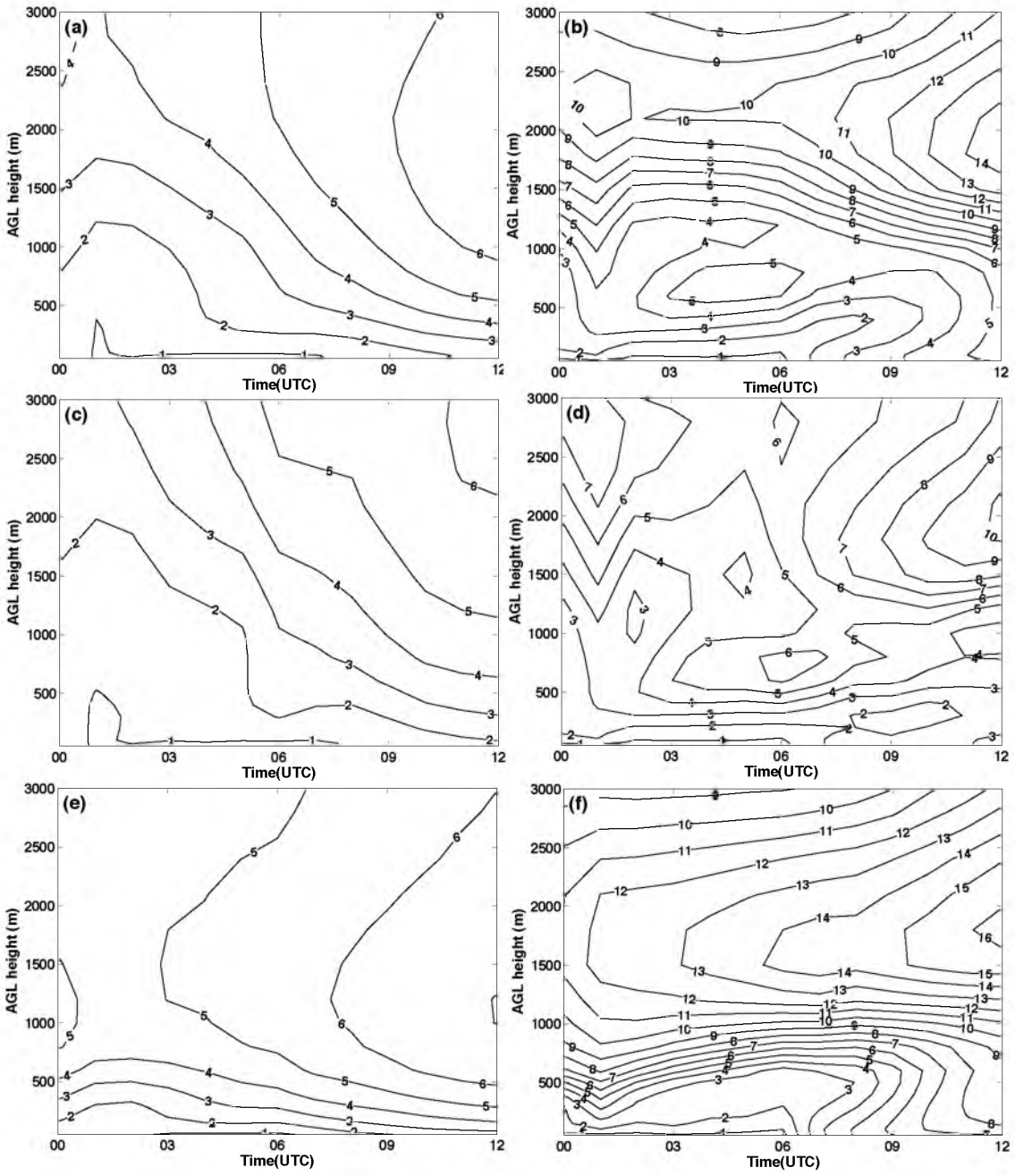


Figure 2.12 Time-height root-mean-square errors (against the nature run) of temperature (K; left column) averaged over a key front region and wind speed (m s^{-1}; right column) averaged over a key LLJ region for EnKF experiments with various maximum radii of vertical localization scales: 1000 m (a and b), 3000 m (c and d), and 5000 m (e and f). The key frontal region and the key LLJ region are marked in Figure 2.1b.

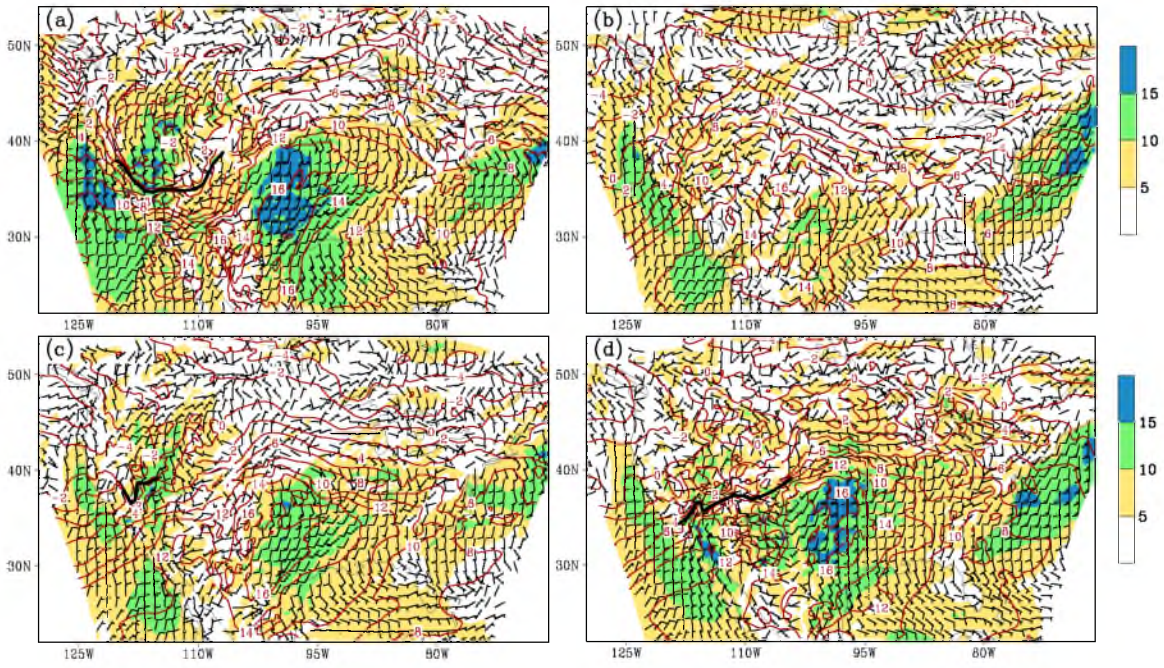


Figure 2.13 Same as Figure 2.1, except for 0600 UTC 05 June 2008.

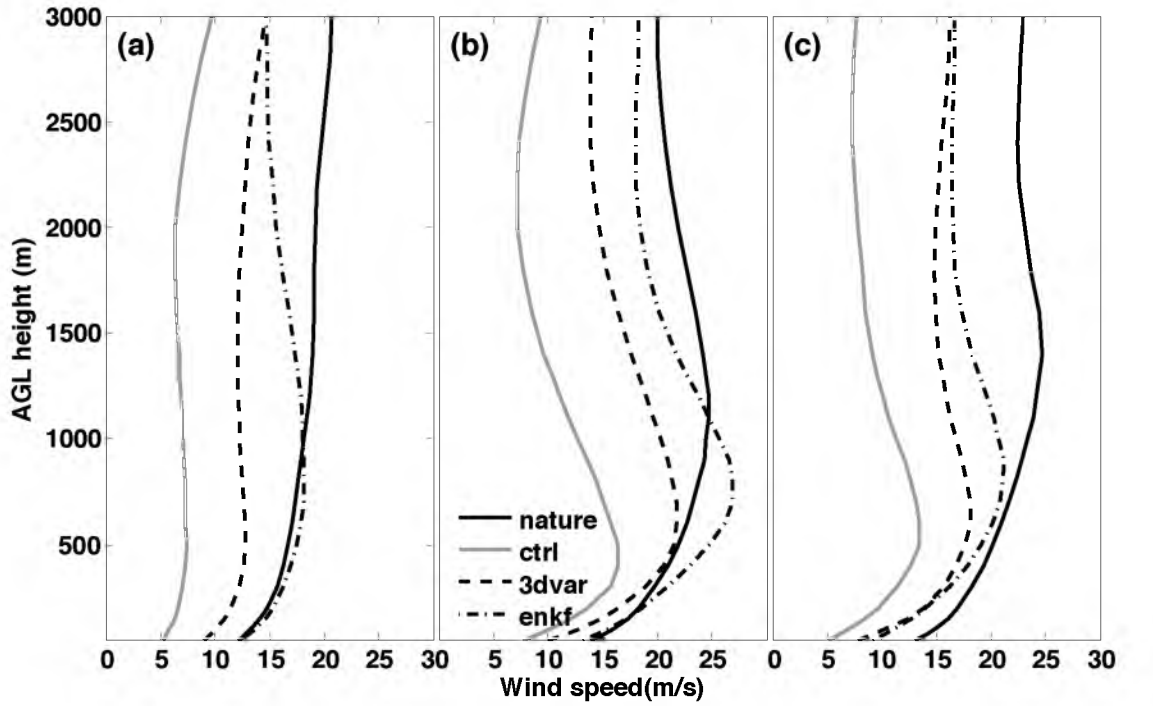


Figure 2.14 Vertical profiles of mean wind speed (m s^{-1}) over the key regions of LLJ. Over a box of (32° N-38° N; 103°W-97°W) after the first data assimilation cycle at 0000 UTC 5 June 2008 (a), (28° N-38° N; 103°W-95°W) at the end of data assimilation cycle at 0600 UTC 5 June 2008 (b), and (32° N-40° N; 105°W-95°W) after 6 h forecast at 1200 UTC 5 June 2008 (c).

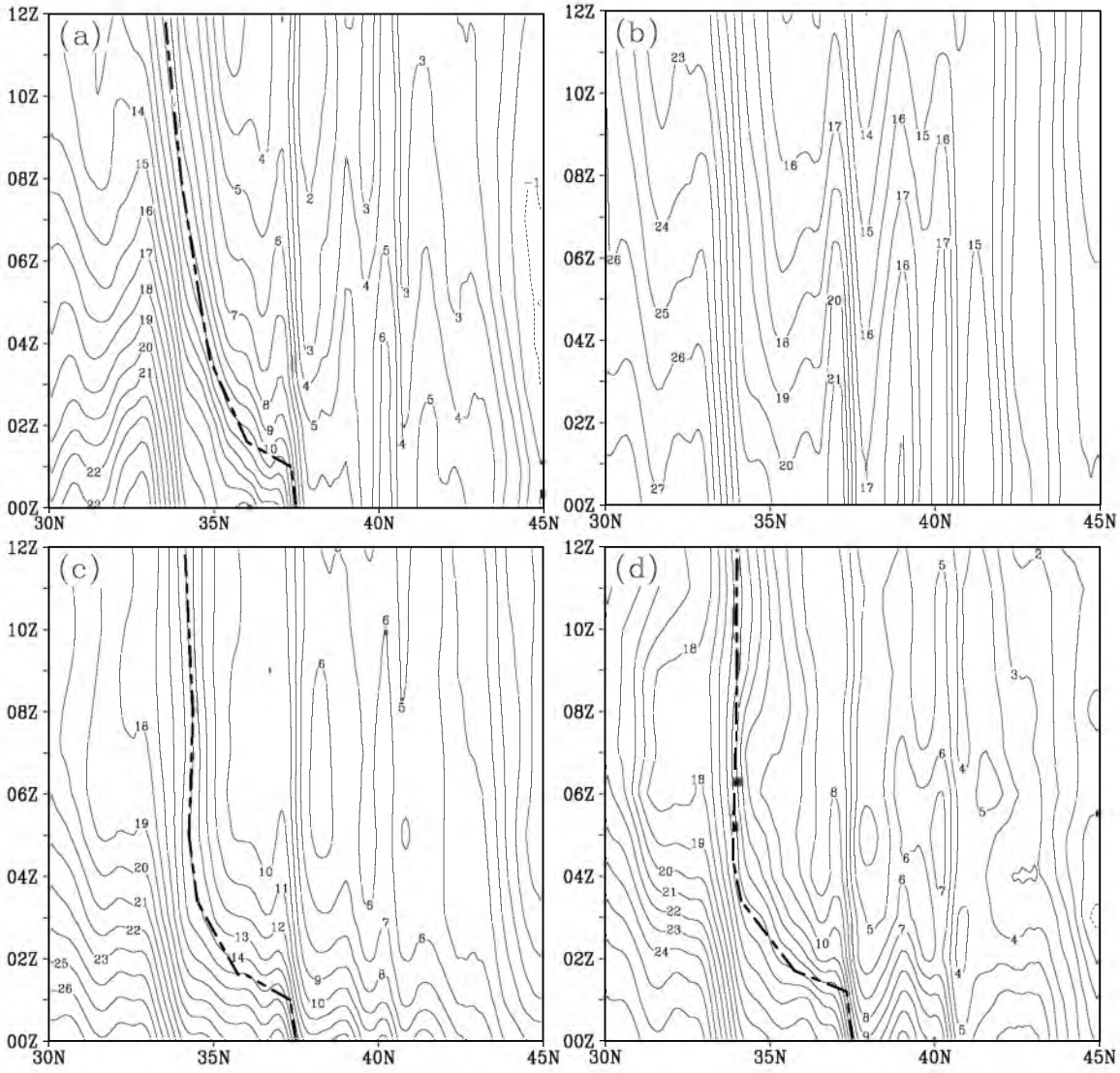


Figure 2.15 Time-latitude cross-section of temperature averaged over the main front zone in 6° longitude ranging from 114° W to 108° W at 500 m AGL for nature run (a), control run (b), 3DVAR analysis (c), and EnKF analysis (d). The dashed bold lines denote the cold front.

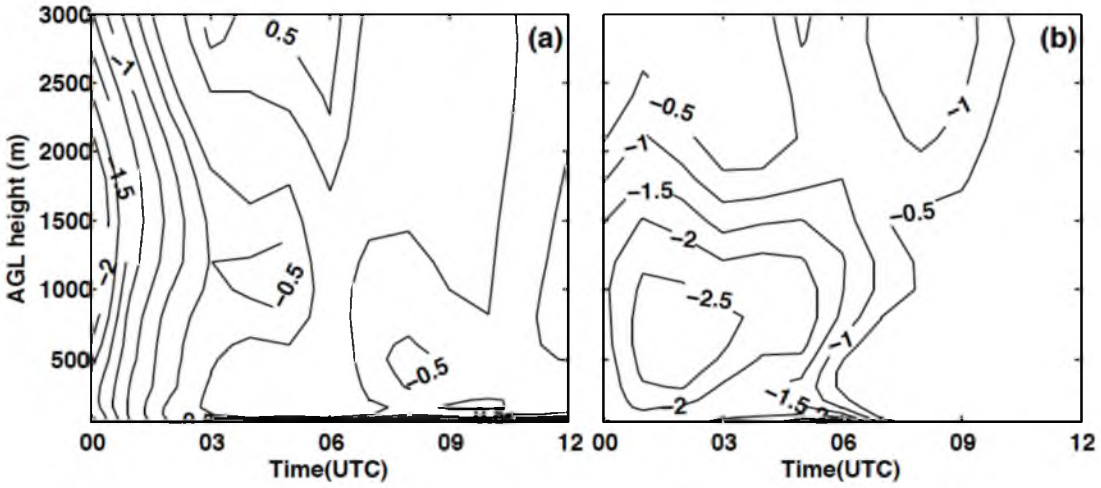


Figure 2.16 Time-height differences of the root-mean-square (RMS) errors (against the nature run) between the EnKF and the 3DVAR for (a) temperature (K) and (b) wind speed (ms^{-1}), calculated over all stations in the key front region as marked in Figure 1b. The negative numbers imply EnKF has smaller RMS errors, relative to the 3DVAR.

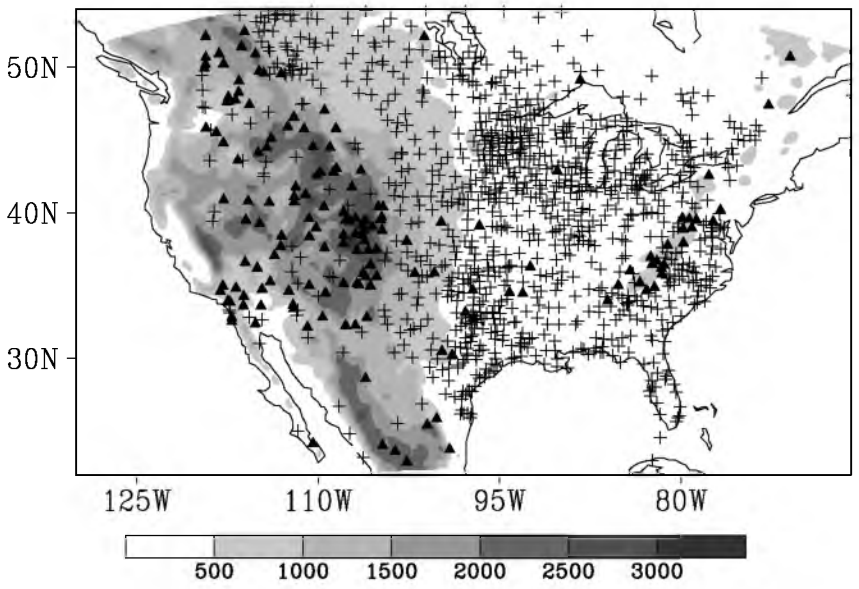


Figure 2.17 Same as Figure 2.2, except triangle signs denote rejected stations.

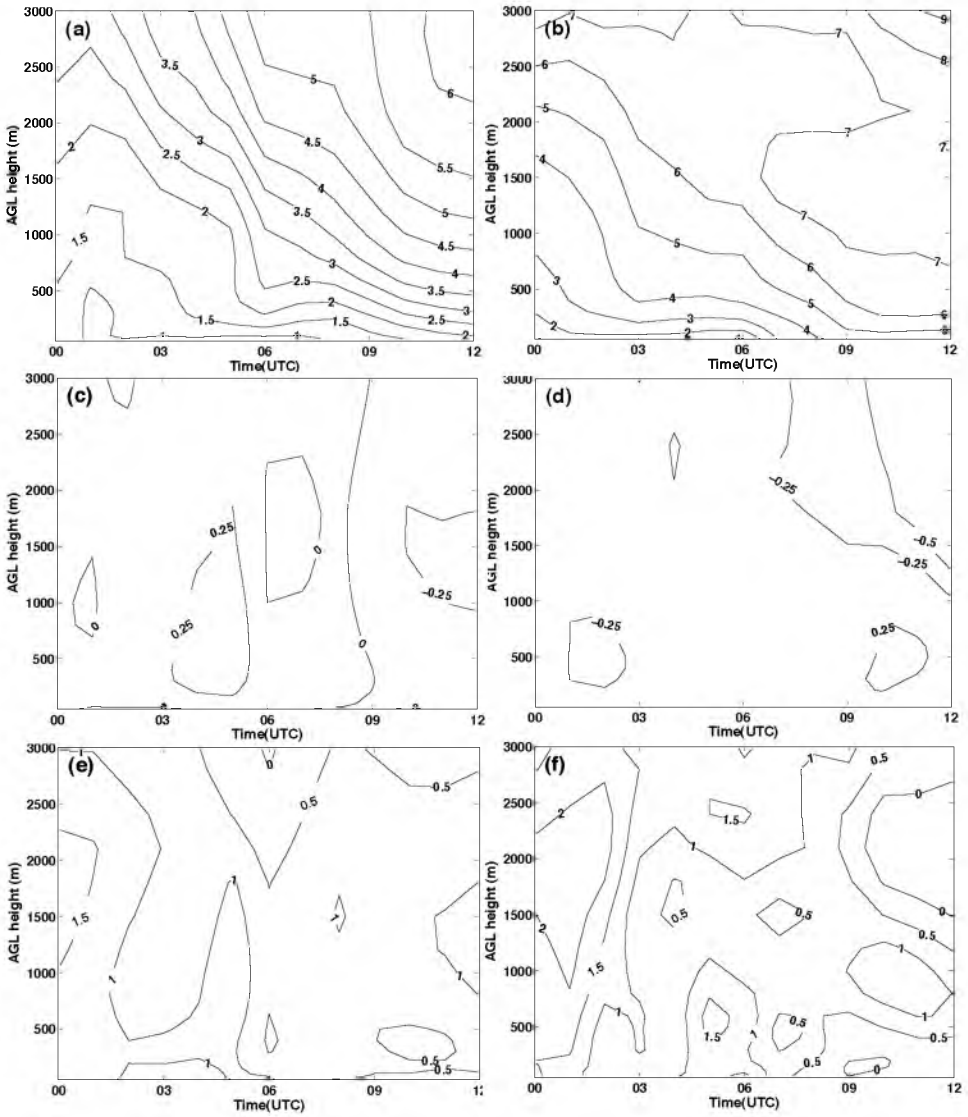


Figure 2.18 Time-height root-mean-square errors (RMSEs) (against the nature run) of temperature (K; left column) and wind speed (m s^{-1}; right column) of the EnKF analysis and subsequent forecast in the key front region as marked in Figure 1b for the *default* experiment without terrain perturbation and data rejection (a and b), RMSE differences between the experiment with terrain perturbation and default experiment (c and d; positive numbers denote the degraded analysis and forecast from the experiment with terrain perturbation), and RMSE differences between the experiment with data rejection and the default experiment (positive numbers mean the degraded analysis and forecast from the data rejection experiment).

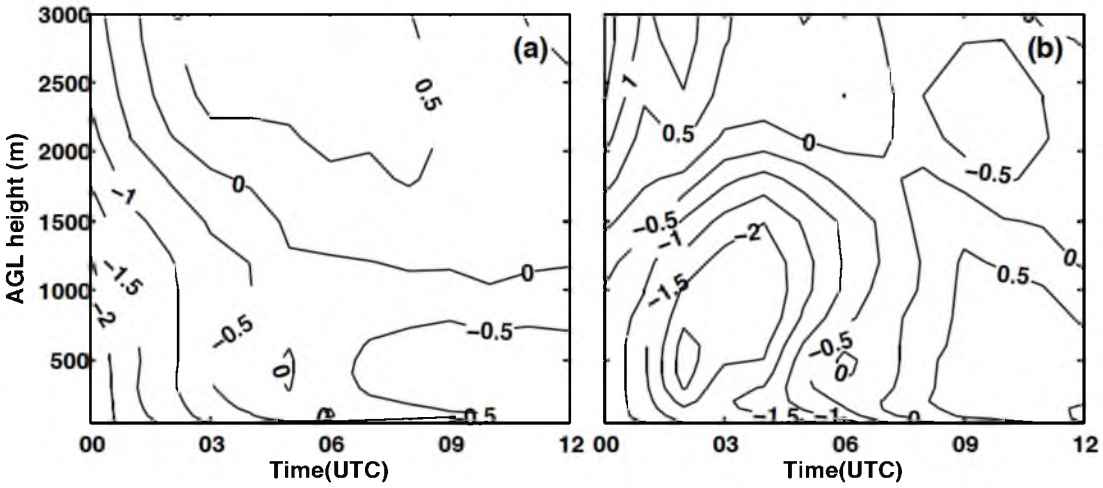


Figure 2.19 Same as Figure 2.16, except for the data rejection experiment.

CHAPTER 3

EXAMINATION OF ERRORS IN SURFACE TEMPERATURE AND WIND FROM WRF NUMERICAL SIMULATION IN COMPLEX TERRAIN

With an OSSE framework, Chapter 2 indicated that the EnKF data assimilation method is able to effectively assimilate surface observations and improve the short-range weather forecasts in complex terrain. Analyses and forecasts were verified in terms of the major influencing phenomena and severe weather systems. However, despite their important roles in interaction with atmospheric boundary layer and local weather systems, near-surface atmosphere variables, such as temperatures and winds, present a unique forecast difficulty. For instance, since the near-surface variables have their own diurnal variation, significant errors could still occur even when the synoptic and mesoscale forecasts are skillful (Zhang et al. 2013). In light of this fact, this Chapter presents a comprehensive verification of the WRF model's performance in forecasting near-surface variables. This is a necessary step for model validation and data assimilation due to our limited understanding of the uncertainties in near-surface atmospheric processes. The main purpose here is to demonstrate the error characteristics of the surface variables in complex terrain and also to examine the potential challenges in assimilating surface variables. Individual cases in flat terrain are also presented in this chapter as a

contrast.

3.1 Background

The near-surface atmosphere, namely, the bottom 10% of the atmospheric boundary layer, is unique due to its direct interaction with the earth's surface (Stull 1988). For instance, near-surface temperature is characterized by diurnal variation, with a maximum at local afternoon and a minimum just before sunrise. This is very different from the free atmosphere in which temperature shows little diurnal variation. Turbulence causes the wind field in the atmospheric boundary layer (ABL) and the near surface layer to behave differently from that in the free atmosphere because the ABL transports momentum, heat, and moisture between the earth's surface and the air above. Due to its unique features, accurate forecasts of near-surface atmospheric conditions are very important in many applications such as wind energy, agriculture, aviation, and fire weather forecasts. However, difficulties in forecasting near-surface variables such as temperature and wind have long been recognized and studied (Hanna and Yang 2001; Zhang and Zheng 2004).

To accurately simulate near-surface atmospheric conditions, several factors must be represented properly in numerical models. These include the land use, terrain representation, surface heat flux transport, and various characteristics of the lower atmosphere (Cheng and Steenburgh 2005; Hacker and Angevine 2012; Lee et al. 1989; Shafran et al. 2000; Wolyn and McKee 1989). Thus, the accurate simulation of near-surface atmospheric diurnal variation is one of the most important and difficult tasks in numerical weather prediction (NWP). Owing to our limited understanding of near-surface atmospheric processes and the uncertainties in model physics parameterizations, a

comprehensive verification of the NWP models' performance in forecasting near-surface variables becomes a necessary step for model improvement.

Hanna and Yang (2001) found that the uncertainties regarding wind speed and direction in the lower atmosphere are primarily due to the turbulent processes that are not appropriately represented in the numerical model, as well as errors in subgrid terrain and land use. They also argued that the models tend to underestimate the vertical temperature gradients in the lowest 100 m during the nighttime. Thus, the simulated boundary layer stability is not as strong as the observed. Considering the different capabilities of planetary boundary layer (PBL) parameterization schemes to reproduce atmospheric structures in the lowest few kilometers, Zhang and Zheng (2004) tested the performances of different PBL schemes in the MM5 model in simulating near-surface temperature and wind speed and direction. Their results revealed that the model could reproduce diurnal variations in surface temperature and wind direction. However, all the boundary layer schemes underestimated (overestimated) wind speeds during the daytime (nighttime). Their study was conducted over the central United States in summer, where little organized convection and topographical forcing was present.

The problem becomes more complicated in complex terrain. Liu et al. (2008b) conducted an inter-range comparison of the model analyses and forecasts of five Army test and evaluation command (ATEC) ranges over a 5-year period using a four-dimensional weather (4DWX, Liu et al. 2008a) system, which is a meteorological support infrastructure developed by NCAR. They concluded that forecast errors vary from range to range and season to season. They also found that larger errors are typically associated with complex terrain. Zhong and Fast (2003) compared three mesoscale numerical

models and evaluated the simulations in the Salt Lake Valley for cases influenced by both weak and strong synoptic scenarios. They found a cold bias in the valley extending from the surface to the top of the atmosphere. The simulated nocturnal inversion was much weaker than observed. There were significant errors in wind forecasts even under strong synoptic forcing. Hart et al. (2005) validated surface forecasts over mountainous terrain during wintertime. They employed the MM5 model at high resolution with three nested domains at 36-km, 12-km, and 4-km horizontal grid spacings. Simulation results did show improved wind and precipitation forecasts in the 4-km horizontal grid spacing domain, compared with those at the 12-km and 36-km domains. However, temperature forecasts did not benefit from the high-resolution simulations. It is noteworthy that although the model properly simulated the persistent nocturnal cold-air pools along with the better-resolved orography at higher resolution, it did not improve the 2-m temperature forecasts. Apparently, forecasting surface air conditions in complex terrain is a challenging problem.

Mass et al. (2002) presented an objective multiyear verification of the University of Washington real-time MM5 forecasts. In their study, triple one-way nested domains at 36-km, 12-km, and 4-km horizontal grid spacings were used and the forecasts of near-surface atmospheric conditions (i.e., 2-m temperature, 10-m wind direction and speed, precipitation, and sea level pressure) from all three domains were verified at observation locations in western Washington State. Their results suggested that the forecasts benefited significantly from decreasing the grid spacing from 36-km to 12-km. However, little improvement was found with further reduction in grid spacing from 12-km to 4-km. These results contrasted with the conclusions of some other studies. For instance, Rife

and Davis (2005) suggested that the gains in forecast accuracy from finer grid spacing are generally incremental. With a regional climate simulation in complex terrain, Leung and Qian (2003) found that a higher resolution simulation improves not only the spatial distribution and regional mean precipitation during summer but also snowpack during winter. However, they also commented that the accuracy of snow simulation is limited by factors such as deficiencies in the land surface model or biases in other model variables. The disagreement among these different studies further indicates the complexity of numerical prediction in complex terrain. Nevertheless, most of these previous studies emphasized the verification of synoptic cases at large and mesoscales. Little attention has been paid to the assessment of the forecast of the near-surface atmospheric conditions.

In this Chapter, we attempt to assess the accuracy of *the near-surface atmospheric conditions, specifically the 2-m temperature and 10-m wind speed and direction,* in numerical simulations produced by the WRF model. In particular, Version 3.3 of an advanced research version of the WRF model (ARW, Skamarock et al. 2008) was used for three typical weather events (i.e., a low-level jet, a cold front, and a wintertime persistent inversion) over the Southern Great Plains (SGP) and the Intermountain West of the US. Our purposes are not only to examine the ability of the ARW model to predict near-surface atmospheric conditions, but also to compare the predictability of near-surface conditions in flat and complex terrain. The sensitivity of numerical forecasts of near-surface atmospheric conditions to various PBL schemes and model resolutions is also investigated. In addition, forecasts during a 1-month period are evaluated to further reveal the characteristics of the forecast errors of near-surface variables under different synoptic forcings in complex terrain.

Section 3.2 briefly describes three synoptic events in two individual cases as well as the numerical simulations and verification methods. Sections 3.3 and 3.4 detail the simulation and verification results for the three synoptic events. Error characteristics of the near-surface variables are also evaluated. Section 3.5 examines the sensitivity of numerical simulations of near-surface atmospheric conditions to various PBL parameterization schemes and model vertical resolutions. Section 3.6 characterizes the errors in near-surface variables statistically with forecasts during a one-month period. Section 3.7 summarizes the results and discussion.

3.2 Description of Cases, Numerical Simulations, and Verification Methods

3.2.1 Cases

3.2.1.1 A Frontal System and a Low-Level Jet: 1 to 3 June 2008

There are two events of interest during 1 to 3 June 2008: a front evolved over the north-central US and a nocturnal low-level jet occurred over the Southern Great Plains (SGP). Surface maps (Figure 3.1) show that a cold front initially located north of North Dakota at 1200 UTC 1 June entered North Dakota at 1500 UTC 1 June. It arrived in South Dakota at 0100 UTC 2 June, then changed to a stationary front at 1200 UTC 2 June and evolved to a cold front again as it moved southward. A temperature gradient of 7°C along with wind direction changes were found between the two stations closest to the front on both sides, with a southwest wind on the south side and a northeast wind on the north side.

A low-level jet was dominant from the surface up to 1800 m above ground level (AGL)

over the entire SGP. It influenced near-surface conditions by interacting with the surface and the lower atmosphere. Radar wind profiles in Jayton, Texas, from the National Oceanic and Atmospheric Administration (NOAA) Profiler Network indicated two periods of evidently greater wind speed during 1 to 3 June (Figure 3.2a). One was between 0200 UTC and 1400 UTC 1 June, and the other was between 0200 UTC and 1400 UTC 2 June.

3.2.1.2 A Persistent Inversion: 1 to 3 December 2010

A persistent inversion began on 29 Nov 2010 and was maintained over the Salt Lake Valley, Utah, for 7 successive days. During this period, extremely strong surface cooling occurred during the night of 1 to 2 December, accompanied by a clear sky and a strong temperature inversion layer aloft. Very low temperatures were observed in the Salt Lake Valley and the adjacent mountains. A high-pressure system controlled the area and helped build and maintain the persistent cold air pools during this period through subsidence. The strong inversion layer extended from the surface up to 2000 m during this time.

3.2.2 Brief Description of Numerical Simulations

Numerical experiments were conducted to simulate the aforementioned cases using the ARW model with *one-way* nested domains. The initial and boundary conditions were derived from the National Centers for Environmental Prediction (NCEP)'s Northern American Mesoscale model (NAM) analysis by WRF preprocessing. A topography dataset at 30 arc-second (about 1000 m) resolution and an updated land use dataset with

27 land use categories (instead of 24 land use categories provided by WRF version 3.3 release) from the United States Geological Survey (USGS) were used in order to ensure more accurate surface conditions, especially for playa and desert regions in the western U.S. The Noah land surface model was used because it predicts the land states, such as surface temperature and soil moisture and temperature, in each layer with time. Table 3.1 lists the configurations of horizontal resolutions, model domains, and physical schemes used for each simulation in this study. Since PBL parameterization schemes contain key physical factors that strongly influence the model's prediction skill for near-surface atmospheric conditions, the PBL scheme used for each individual simulation (i.e., the control simulation) was chosen from the sensitivity studies (as described in Section 3.5).

3.2.3 Verification Methods

3.2.3.1 Synoptic Verification

Verification was first conducted to evaluate the accuracy of the numerical simulation of each synoptic event. Simulation results were compared with available observations and analyses.

3.2.3.2 Verification of Near-Surface Atmospheric Conditions

The major emphasis of this study is to characterize errors in the near-surface atmosphere. In order to quantify these errors, we used surface Mesonet observations (Horel et al. 2002) to verify the model's performance in terms of the near-surface variables, namely, 2-m temperature and 10-m wind speed and direction. According to Horel et al. (2002), quality control algorithms and data monitoring programs were

performed for all available data. The quality-controlled data were then made available hourly with quality flags. In this study, only those observations with a quality flag of "OK" (the highest quality) were used for verification.

Since there are case-by-case variations in near-surface atmospheric conditions due to various synoptic systems and terrain, verification of near-surface atmospheric conditions was performed for each synoptic event. Representativeness errors due to discrepancies between the model and actual terrain heights are commonly present in the forecasts of surface variables. Therefore, model terrain heights were compared against actual terrain heights for each case to examine the representative errors. Model performance was then checked for each case for each variable over time. In this study, we used variable mean, mean absolute error (MAE), and bias error (BE) of 2-m temperature and 10-m wind speed and direction against observations to characterize the errors in numerical simulations. We also calculated time-averaged mean absolute errors (TMAE or cumulative MAE) to average the MAEs over the whole simulation period. Because observational errors in wind direction are usually larger at lower wind speeds, only those observations with wind speeds greater than 1.5 ms^{-1} were used to verify wind direction. To verify the model simulation, simulation results at model grid points were interpolated to observation locations using a bilinear method. The statistical calculations are as follows:

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|F_i - O_i\right| \tag{3.1}$$

$$\mathrm{BE} = \frac{1}{n}\sum_{i=1}^{n}(F_i - O_i) \tag{3.2}$$

$$\mathrm{TMAE} = \frac{1}{n}\frac{1}{m}\sum_{t=1}^{m}\sum_{i=1}^{n}\left|F_{it} - O_{it}\right| \tag{3.3}$$

where i denotes the ith observation, t denotes the observation at time t, O_i represents the value of the observation at the ith location, F_i denotes the forecast value interpolated to that observation location, n is the total number of stations, and m represents the total number of times used to calculate TMAE.

3.3 Simulation and Verification: 1 to 3 June 2008

Triple-level, one-way nested domains (D01, D02, and D03 in Figure 3.3a) with 27-km, 9-km, and 3-km horizontal grid spacings (hereafter referred to as the 27-km domain, 9-km domain, and 3-km domain) were utilized in this simulation. There were 37 vertical levels from the surface up to 50 hPa. The innermost domain (i.e., the 3-km domain) focused on the two weather systems of interest over the north-central US and the SGP. The model was initialized at 0000 UTC 31 May, 24 hours ahead of the verification.

Physical parameterization options, as listed in Table 3.1, include the WSM 6-class microphysics scheme (Hong and Lim 2006), the Yonsei University (YSU) PBL scheme (Hong and Pan 1996), the Kain–Fritsch cumulus parameterization scheme (Kain and Fritsch 1993), the Noah land surface model (Chen and Dudhia 2001), the rapid radiative transfer model for longwave radiation (RRTM, Mlawer et al. 1997), and the Dudhia shortwave radiation scheme (Dudhia 1989). The cumulus scheme was used only in the 27-km and 9-km domains.

3.3.1 Synoptic Verification

3.3.1.1 The Frontal System

Sounding observations from the National Weather Service (NWS) were compared with the simulated temperature, geopotential height, and wind barbs in a weather map at the 850 hPa pressure level at 0000 UTC 2 June 2008 (Figure 3.4). The simulated geopotential heights almost overlap with the observations, indicating the front was well simulated. Over the frontal region, the observed and simulated temperature fields were almost identical. The larger temperature gradient over North Dakota and the wind barbs representing a realistic frontal system were reproduced by the model simulation (Figure 3.4b).

3.3.1.2 The Low-Level Jet

To verify the forecast of the low-level jet, a comparison was first made using the NCEP Northern American Regional Reanalysis (NARR, Mesinger et al. 2006) data products. Figure 3.5 compares wind speeds from the NARR data and the model simulations at 850 hPa at 0900 UTC 2 June 2008. The similarities between the wind fields in the NARR and the simulation, in terms of jet coverage and intensity, suggest that the model has simulated the low-level jet from a large-scale perspective.

Wind profile observations from the NOAA Profiler Network in Jayton, Texas, clearly reveal the nocturnal jet. A time series of vertical profiles with wind speeds and vectors from 0000 UTC 1 to 0000 UTC 3 June 2008 is displayed in Figure 3.2a. Compared with observed wind speeds and vectors, the simulated winds (Figure 3.2b) reasonably represent the structure of the low-level jets, although the simulations underestimate the

wind speed and intensity of the low-level jets.

Meanwhile, observed surface winds (10-m wind from Mesonet surface stations) were relatively calm during the night because the turbulence ceased after sunset. However, the simulated 10-m wind speeds during the night were much higher than observations, indicating the model has not captured the thermal decoupling between the surface and higher-level air after sunset. Figure 3.2c shows that the near-surface air cooled quickly after sunset due to radiative cooling. The air at 2-m is 3–5 ^{o}C colder than the air above since radiative cooling begins at the surface. The near-surface air temperature decreases immediately after sunset and then decouples with the air in the residual layer above that stays relatively warmer during the night. The simulated temperature (Figure 3.2d) in the boundary layer captures the observed temperature inversions on both nights in the higher level air. However, the simulated 2-m temperatures during both nights were much warmer than the observations. As a result, the model reproduced only a weak decoupling between the surface and the air above.

3.3.2 Verification of Near-Surface Atmospheric Conditions

3.3.2.1 The Frontal Case Over Flat Terrain

Figure 3.6a shows the area and Mesonet observation stations used for verifying the front. There are over 400 observations available hourly. Figure 3.6b compares the realistic and model terrain in these stations. The model terrain at all resolutions generally matches the actual terrain, while the 3-km domain (D03) makes the best match.

The simulated 2-m temperatures in the frontal area generally agree well with observations in all domains (Figure 3.7a), showing no systematic bias. However, large

MAEs occur in 2-m temperature at the end of the simulation when the stationary front changes to a cold front (Figure 3.7d). The MAEs of the 2-m temperature are smallest in the 3-km domain, specifically during the second day (Figure 3.8d).

The model captures the southwest-to-northeast wind direction change accompanying the frontal passage (Figure 3.7b). Particularly, wind directions in the simulation agree well with the observations from 1200 UTC 1 to 1200 UTC 2 June 2008. Relatively larger errors in wind direction occur near the end of the simulation period (Figure 3.7e) and are caused mainly by the rapid transitions from a cold front to a stationary front and then again to a cold front. These changes complicate the frontal event and pose extra difficulty for the numerical simulation. Overall, the errors in 10-m wind direction are similar in all three domains in the first 36 hours of forecasts. The 3-km domain performs better during the last 12 hours (Figure 3.8e), when rapid transitions take place. A diurnal feature of the errors in wind direction, characterized by larger errors during nighttime and smaller errors during daytime, can also be seen in Figure 3.7e.

The observations clearly depicted diurnal variations in wind speeds, with higher speeds during the daytime and lower speeds at night. The model well simulates the diurnal signals but generates larger errors at night (Figures 3.7c and 3.7f) during the nocturnal jets. In particular, the model does well in simulating wind speeds between 0000 UTC to 0200 UTC 1 June (corresponding to 1800 to 2000 31 May in Central Standard Time, CST) and 1500 UTC 1 June to 0100 UTC 2 June (corresponding to 0900 CST to 1900 CST 1 June); both of these periods are in the daytime. However, larger errors (Figure 3.7f), characterized by positive biases (Figure 3.7c), are found for nocturnal wind speeds. These positive bias errors in 10-m wind speed can be attributed to the incomplete

representation of the thermal decoupling between the higher layer air and the near-surface atmosphere in the simulation (similar to that shown in Figures 3.5b and 3.5d). In addition, the forecasts in the 3-km domain outperform the 9-km and 27-km domains during the daytime, but not at night (Figure 3.8f). This is mainly because the coarser resolution domains do not resolve the intensity of the low-level jets as well as the higher resolution domain does during the nighttime. The wind speeds of the low-level jets were weaker in the 9-km and 27-km domains than those in the 3-km domain. Meanwhile, the surface wind speeds were strongly influenced by the air above the surface layer (the low-level jet) because they were coupled as the model failed to represent the thermal decoupling. Therefore, the coarser resolution domains produced weaker low-level jets and weaker surface wind speeds, resulting in smaller positive biases of surface wind speeds. In other words, due to the model's inability to represent the thermal decoupling between the near-surface layer and boundary layer above (as mentioned in Section 3.3.1), the coarser resolution domains outperform the high-resolution domain during nighttime.

3.3.2.2 The Low-Level Jet Over Flat Terrain

Figure 3.6c shows the Mesonet observation stations used for verifying the low-level jets. Figure 3.6d compares the actual and model terrain heights over these stations. The model terrain matches the actual terrain very well at most stations. The terrain heights in the 3-km domain (the innermost domain), again, best match the actual terrain heights.

The 3-km domain resulted in the smallest errors in 2-m temperature during 1500 UTC 1 June to 1200 UTC 2 June (Figures 3.9a and 3.10d). However, it produced the largest errors during the first night and errors that were comparable to the other domains in the

early morning for both days. The errors peaked at 1200 UTC 1 June and 1200 UTC 2 June when the boundary layers were most stable. Southerly flow dominated during the simulation period. The errors in 10-m wind direction were relatively small due to the strong southerly forcing (Figure 3.9b). Owing to the influence of the low-level jet and the inaccurate representation of the nocturnal decoupling and radiative cooling as mentioned above, there were relatively larger errors in 10-m wind speed during the nighttime (Figure 3.9c).

Accurate simulation of the transition boundary layer and the typical stable boundary layer near the ground is still one of the challenges in numerical simulation. The method to parameterize the stable boundary layer has also been an active research area in recent studies (Brown and Wood 2003, Teixeira et al. 2008). A discussion of the best way to overcome forecast errors in a stable boundary layer is beyond the scope of this study. However, accurate forecasts of near-surface conditions depend on the model's ability to simulate the stable boundary layer.

3.4 Simulation and Verification: 1 to 3 December 2010

The simulation was initialized at 0000 UTC 31 November, and the results from 1 to 3 December 2010 were used to verify the persistent inversion. Three-level, one-way nested domains (D01, D02, and D03) at 12-km, 4-km, and 1.33-km horizontal grid spacings (hereafter referred to as the 12-km domain, 4-km domain, and 1.33-km domain) were used. The innermost domain (i.e., the 1.33-km domain) focused on the Salt Lake Valley and its surrounding mountains. The model also included 37 vertical levels from the surface up to 50 hPa.

Physical parameterization configurations for this case (see Table 3.1) were the same as for the 1 to 3 June 2008 case, except for the Mellor–Yamada–Janjic (MYJ) TKE PBL scheme (Mellor and Yamada 1982). The cumulus scheme was used only in the 12-km domain. The choice of the physics schemes in each case was based on a set of sensitivity studies. The schemes (experiments) that produced the smallest RMSEs were taken for analysis.

3.4.1 Synoptic Verification

Sounding observations, obtained every 12-h from the station at the Salt Lake City International Airport (KSLC), were available to examine the structure of the atmospheric boundary layer. Model simulated winds and temperatures were interpolated to the sounding locations for comparison. Figure 3.11 shows the evaluation of the temperature and wind fields for both soundings and simulations throughout the 2-day period. Observations showed warmer air (relative to near-surface air) above the surface; namely, an inversion layer was present throughout the entire period, although the inversion was more intense in the late stages. The top of the inversion varied from 700 to 1200 m AGL during this period. The simulation reproduced the persistent inversion for the entire period. The wind shears, which were present as southeasterly-to-southerly beneath the inversion layer and southwesterly above it at 1200 UTC 2 and 0000 UTC 3 December 2010, were well captured in the simulations, although the transition heights were slightly different from the observations. The simulated heights of the inversion layer were lower than those in the sounding observations. The simulated temperature gradient at the bottom of the boundary layer was not as strong as the observed. These results were

similar to those of Hanna and Yang (2001). Apparently, discrepancies between simulations and observations can be attributed mainly to errors in the simulation of the near-surface atmospheric conditions.

3.4.2 Verification of Near-Surface Atmospheric Conditions

Figure 3.12a shows the distribution of Mesonet observation stations used for verifying this case. In complex terrain, observations are distributed unevenly in the Salt Lake Valley and the surrounding mountains. Figure 3.12b compares the model and actual terrain heights for all three domains. The 1.33-km domain represents the actual terrain substantially better than the 12-km and 4-km domains. The 12-km domain misrepresents lower/higher terrain (less/greater than 2000 m in the valley/mountains) with higher/lower heights. Consequently, the coarser resolution domain does not resolve the deep valley and sharp mountains accordingly. The 4-km domain has an intermediate ability to represent the terrain compared with the 1.33-km and 12-km domains.

Because of the large differences in terrain height between the stations in the valley and those in the mountains, the stations were separated into two groups during the verification: stations inside the Salt Lake Valley and those in the surrounding mountains (valley stations and mountain stations hereafter). A station with higher (lower) than 2000 m terrain height was categorized as a mountain (valley) station. A similar separation was used in Hart et al. (2005).

3.4.2.1 Salt Lake Valley: Valley Stations

Consistent with synoptic verification results, large forecast errors were found for 2-m temperatures at the valley stations, especially during the night of 2 December 2010 (Figure 3.13a) when the valley was undergoing extreme nocturnal cooling induced by the persistent inversion. The 12-km domain produced the smallest MAEs during that night while the 4-km domain produced the largest MAEs (Figures 3.13a, 3.14a and 3.14b). This degradation accompanying the decrease of grid spacing from 12-km to 4-km can be attributed to the prediction skill of the intense inversion and the terrain representation in the ARW model. Specifically, the terrain heights in the 4-km domain are typically lower than those in the 12-km domain at the valley stations. During the inversions, the temperature usually increases with height inside the valley. However, the model cannot fully capture the intense cold air pools, resulting in a temperature profile that either decreases with height or does not increase with height with a proper lapse rate in the near-surface layer. Therefore, the 4-km domain produced warmer surface temperatures than the 12-km domain due to its deeper valley representation, causing even larger warm biases. A similar argument was presented by Hart et al. (2005) using the MM5 model.

It is interesting that the MAEs of the 1.33-km domain were intermediate between those of the 12-km and 4-km domains (Figures 3.13a, 3.14a and 3.14b). As seen in Figure 3.11, the 1.33-km domain captured the inversion in the lower atmosphere and properly represented the temperature lapse rate (i.e., increase with height). With a deeper topography and proper lapse rate representation, the 1.33-km domain produced better 2-m temperature forecasts than the 4-km domain. However, MAEs of the 2-m temperature were still larger in the 1.33-km domain than in the 12-km domain because the

temperature lapse rate in the lower atmosphere (namely, the inversion intensity) resolved by the 1.33-km domain was much weaker than that of the sounding observations (as shown in Figure 3.11).

The TMAEs in wind direction over all the stations were about 80 degrees (Figure 3.16b). Both the 4-km and 1.33-km domains produced degraded wind direction forecasts (Figures 3.13c, 3.14c, 3.14d, and 3.16b) because of the contradiction between the model's failure to fully capture the strong inversion and the better terrain representation in the higher resolution domains, as discussed above.

The MAEs of 10-m wind speed at the valley stations were significantly reduced in the 4-km domain (Figures 3.13e, 3.14e, 3.14f, and 3.16c). Much greater wind speeds were produced in the 12-km domain because it has a shallower valley floor (at a higher elevation than those of the 4-km and 1.33-km domains). There were larger MAEs in wind speed since the observed winds were relatively calm at the surface.

3.4.2.2 Salt Lake Valley: Mountain Stations

Meanwhile, decreasing the model's grid spacing from 12-km to 4-km and 1.33-km substantially improves the temperature forecasts over the mountain stations (Figures 3.13b, 3.15a, 3.15b, and 3.16a). These forecasts benefit from the sharper and better representation of mountain terrain at the higher resolution while the mountain stations are above the cold pools.

The wind direction forecasts are improved in the 4-km domain from the 12-km domain but are degraded in the 1.33-km domain from the 4-km domain (Figures 3.13d, 3.15c, 3.15d, and 3.16b). The mountain stations are more connected to the free

atmosphere and not influenced by the poor prediction skill of the WRF model for the intense inversions. Therefore the 4-km domain benefitted from the higher resolution and showed improved forecasts of wind directions over the 12-km domain. The reasons for the degradation in the 1.33-km domain are unknown, reflecting the challenges in producing accurate forecasts of near-surface atmospheric conditions. In reality, accurate forecasting of near-surface wind direction is extremely difficult as it depends not only on the terrain representation in the model, but also on the accurate prediction of wind speed and the thermal and dynamical forcings. In addition, low wind speeds in the surface layer pose extra difficulties in forecasting wind direction, although observations with speeds less than 1.5 ms^{-1} are already excluded in the comparison.

The simulation in the 12-km domain produced the smallest MAEs in wind speed. Slight degradations were found in the forecasts of the 1.33-km domain, and relatively larger degradations were found in the 4-km domain (Figures 3.13f, 3.15e, 3.15f, and 3.16c).

3.4.2.3 Dugway Proving Ground

The great complexity of the terrain in the Salt Lake Valley presents a significant challenge to forecast verification. To further examine error characteristics in near-surface variables in complex terrain, an additional verification was conducted against observations located in the Dugway Proving Ground (DPG) area. The DPG is located approximately 80 miles southwest of Salt Lake City, Utah. It is characterized by complex terrain and is surrounded on three sides by mountain ranges. Therefore the DPG area represents typical complex terrain.

Figure 3.12c shows the DPG area map and stations used for verification. Currently, there are a total of 31 automatic surface stations in the DPG area. However, six stations (denoted by circles) were not used for verification in this case because they did not begin providing observations until 2011. Figure 3.12d compares the actual and model terrain heights. Similar to the previous section, the 1.33-km domain has the best terrain representation. There is evidently a large error in both the 12-km and 4-km domains in representing the actual terrain at station DPG16, located on a mountaintop at an elevation of 2149 m. For instance, the terrain height at DPG16 in the 12-km domain is only 1450 m.

The effect of nocturnal cooling in the DPG area is shown in Figure 3.17. The temperature error was as high as 9^{o} C during the night. The 12-km and 4-km domains produced the smallest and the largest MAEs, respectively, while the 1.33-km domain produced intermediate MAEs. The 2-m temperature errors displayed similar features to those of the valley stations, as seen in Figures 3.13a and 3.16a, since they were under the same weather regime and synoptic environment. The MAEs of wind direction and speed also showed patterns similar to those of the valley stations (not shown). In order to eliminate the impact of the mismatched terrain at DPG16, statistical analyses were re-conducted by excluding this station. The results remained almost the same. This implies that the representative error caused by terrain mismatch is not the sole reason for the errors in simulated near-surface variables in complex terrain.

3.5 Sensitivity to PBL Schemes and Vertical Resolution

It has been recognized that PBL parameterization schemes have a substantial influence on the simulation of surface variables (Hu et al. 2010, Shin and Hong 2011). It

is also commonly believed that simulations could be improved with increased model vertical resolution, especially in the boundary layer, in which the model can better resolve small-scale processes. To examine the impact of PBL schemes and model vertical resolution on the simulation of near-surface atmospheric conditions, additional experiments were conducted and discussed in this section.

3.5.1 Sensitivity to Various PBL Schemes

PBL schemes parameterize unresolved turbulent vertical fluxes of heat, momentum, and moisture within the planetary boundary layer and throughout the atmosphere in numerical models. The ARW WRF model has multiple PBL scheme options, characterized by different closure methods, prognostic variables, and other aspects. The first set of sensitivity experiments used various PBL schemes while keeping other configurations the same, as specified in the control simulations (Table 3.1). These experiments were designed to evaluate the sensitivity of numerical simulation of near-surface temperature and wind fields to different PBL schemes.

Five PBL schemes in the ARW model were tested and compared—YSU, MYJ, quasi-normal scale elimination (QNSE, Sukoriansky et al. 2005), Mellor–Yamada– Nakanishi–Niino Level 2.5 (MYNN2, Nakanishi and Niino 2004), and Asymmetric Convective Model version 2 (ACM2, Pleim 2007). Among these schemes, the YSU and ACM2 are first-order, nonlocal schemes. They do not require additional prognostic equations to describe the effects of turbulence on mean variables. They are both based on the K-profile in determining the diffusivity in the boundary layer and consider nonlocal mixing by convective large eddies. The YSU scheme expresses nonlocal mixing by simply

adding a nonlocal gradient term which is a correction to the local gradient that incorporates the contribution of the large-scale eddies to the total flux in the turbulence diffusion equation (Hong and Pan 1996). The YSU scheme is modified in ARW version 3 (Hong et al. 2008) to enhance mixing in the stable boundary layer. The ACM2 scheme expresses the nonlocal flux explicitly. The MYJ, QNSE, and MYNN2 schemes are classified as 1.5-order turbulent kinetic energy (TKE) closure schemes. They calculate eddy diffusion coefficients by predicting TKE. They differ in how they define the diffusion coefficients. More details about these PBL schemes can be found on the ARW WRF user page at http://www.mmm.ucar.edu/wrf/users, as well as in Shin and Hong (2011) and Hu et al. (2010).

Figure 3.18 depicts the sensitivity of near-surface variable forecasts to model PBL schemes by comparing the time-averaged MAEs (TMAEs) of surface variables over the whole simulation period in each case. For temperature (Figure 3.18a), the TMAEs of the cold front and the low-level jet in summer 2008 were smaller than those of the inversion case in winter 2010. For the frontal and low-level jet cases, all schemes generated similar TMAEs. For the inversion case, however, the TMAEs of temperature were sensitive to the choice of PBL scheme. The MYJ and QNSE schemes produced smaller TMAEs than the others at both valley and mountain stations. All PBL schemes generated similar errors in wind direction for all cases (Figure 3.18b), except that the MYJ and QNSE schemes produced relatively larger TMAEs at the mountain stations. In terms of wind speed, the ACM2 and MYNN2 schemes performed best in the frontal and low-level jet events (Figure 3.18c). For the inversion, the YSU, MYNN2, and ACM2 schemes performed equally well at both valley and mountain stations. The MYJ and QNSE schemes, which

produced the best temperature forecasts for the inversion (as seen in Figure 3.18a), were the least accurate in wind speed forecasts.

Overall, all schemes performed similarly in the frontal and low-level jet cases in terms of temperature and wind direction forecasts. The MYJ and QNSE schemes produced relatively larger TMAEs in wind speed in the frontal and low-level jet cases. They improved the temperature forecasts at both valley and mountain stations for the winter of 2010 inversion but at the same time produced the least accurate wind speed and direction forecasts at the mountain stations.

Figure 3.19a further examines the sensitivity of 2-m temperature simulation to model PBL schemes as a function of forecast lead time in the Salt Lake Valley for the December of 2010 simulation. The five schemes split into two groups. The group with the MYJ and QNSE schemes performed much better than the group with the three other schemes, especially during the first day of the simulation. The MYJ and QNSE schemes, however, had larger MAEs in the 10-m wind speed forecast, as shown in Figure 3.19b.

The above results indicate that no single PBL scheme leads to overall improvement in the forecasts of near-surface wind and temperature fields for all cases and regions. The model is more sensitive to the choice of PBL scheme in the inversion case in complex terrain than in the frontal and low-level jet cases in flat terrain. In addition, one PBL scheme that results in better wind (temperature) forecasts does not reproduce improved temperature (wind) forecasts. Because the MYJ and YSU schemes have been frequently used in recent applications and also because of their respective performances in accurately forecasting wind and temperature fields, they were chosen for the control experiments for the low-level jet in flat terrain and the inversions in complex terrain, as

listed in Table 3.1.

3.5.2 Sensitivity to Model Vertical Resolution

The other set of experiments is used to examine the sensitivity of model simulations to vertical resolution. In these experiments, the vertical levels are increased to 70 (70L) instead of 37 (37L) as in the control simulations. Figure 3.20 shows sketches of vertical model levels in both configurations. The increase in the vertical levels is more obvious below 6 km AGL, especially the lowest 2 km AGL. As seen in Figure 3.11c, simulations with increased vertical resolution can better resolve the structure of the persistent inversion in the boundary layer and produce more reasonable inversion depth and intensity. However, the higher vertical resolution did not improve the forecasts of near-surface variables. Figure 3.21 shows that both experiments (with 37 and 70 vertical levels) perform almost identically in terms of the TMAEs for both 2-m temperature and 10-m wind speed and direction for almost all cases and over all regions. Overall, improvement in surface forecasts was not ensured by increased vertical resolution, although it can help to better resolve the structures of the mesoscale phenomena in the boundary layer.

3.6 Characteristics of Flow-Dependent Errors: Statistics Over a 1-Month Period in the Fall of 2011

The above results from the three typical weather patterns indicate the case-by-case variability of the errors in forecasts of near-surface variables. In order to further understand the general characteristics of the errors of the near-surface forecasts in complex terrain, verification was conducted for forecasts over a 1-month period in the

DPG area. We chose the DPG area as the focus for multiple case statistics for two reasons: 1) as mentioned above, the DPG area better represents common complex terrain features and 2) forecasts of near-surface variables behave similarly over the DPG area and the Salt Lake Valley (as seen in Figure 3.17 and described in Section 3.4.2), making it easier to examine the forecast errors associated with synoptic forcings.

A near real-time forecasting system was built using Version 3.3 of the ARW model. From 15 September to 14 October 2011, near real-time forecasts were performed four times daily at 0000 UTC, 0600 UTC, 1200 UTC, and 1800 UTC to produce a 48-h forecast each time. Four one-way nested domains with 30-km, 10-km, 3.33-km, and 1.1-km horizontal grid spacings were used (Figure 3.22). The innermost domain (1.11-km) focused on the DPG area. Initial and boundary conditions were derived from the analyses and forecasts produced by the North American Mesoscale (NAM) Forecast System at NCEP at 0000 UTC, 0600 UTC, 1200 UTC, and 1800 UTC. Over a 1-month period, a total of 120 forecasts were generated. As shown in Table 3.1, in addition to using physical schemes similar to those in the aforementioned case study, the Purdue Lin microphysics scheme (Chen and Sun 2002; Lin et al. 1983), and the YSU PBL scheme were used.

3.6.1 Overall Evaluation

MAEs and BEs were employed to characterize the forecast errors in near-surface variables. A statistical calculation was done for each of the four initialization times. For example, all forecasts initialized at 0000 UTC during the month were averaged over all stations to calculate the MAEs and BEs as a function of forecast lead time.

Figures 3.23a, b, and c show the MAEs calculated in the DPG area in the 10-km, 3.33-km, and 1.11-km domains for the forecasts initialized at 0000 UTC. Figures 3.23d, e, and f show the differences of MAEs between the 1.11-km domain and 10-km domain. Results confirm that simulations at high-resolution do not always outperform those at coarser resolution. However, a clear diurnal pattern was found in the errors of all variables produced by all model domains. Specifically, the temperature error peaked twice per day, around 3 am Mountain Standard Time (MST) and 3 pm MST (corresponding to 1000 UTC and 2200 UTC). There were also two error minima for temperature at around 7 am and 7 pm MST (corresponding to 1400 UTC and 0200 UTC). Wind speed and direction followed the same error trends, with a maximum in the early evening or before sunrise and a minimum in the afternoon.

Using results from the 1.11-km domain, the dependence of the surface forecasts on initialization time was examined. Figure 3.24 shows that the error trends are independent of initialization time and forecast lead time and follow the same diurnal variation. However, compared with the forecasts initialized during the daytime (0000 UTC and 1800 UTC), relatively large errors occured in the first 2 to 3 hours in 2-m temperature for the forecasts initialized at night (0600 UTC and 1200UTC). The large errors in the nighttime-initiated forecasts could be caused by the erroneous soil temperature initialization in the NAM analysis. Apparently, the large errors associated with initial conditions in the nighttime-initiated forecasts do not persist beyond a few hours. This may indicate that a better local-scale initialization using data assimilation can help reduce forecast errors within the first few hours, but its impact may subsequently vanish. Therefore, cycled data assimilation with a short time window may be required to mitigate

the problem.

Overall, compared with the previous study by Liu et al. (2008b), the statistical errors were moderate for the whole period, with maximum errors of 3°C in 2-m temperature and 2 m s^{-1} in 10-m wind speed. Figure 3.25 further demonstrates the diurnal patterns of the bias errors in 2-m temperature over the whole month. Positive (warm) biases were found at night and negative (cold) biases were present during the daytime. No systematic biases were found in wind direction and speed (figures not shown).

3.6.2 Strong vs. Weak Synoptic Forcing Cases

In Sections 3.3 and 3.4, it was concluded that the errors in 2-m temperature were significant under strong synoptic forcing. With 1-month forecasts that include both quiescent periods and strong synoptic forcing cases, we categorized the forecasts into strong and weak synoptic forcing cases by checking the weather maps at the surface and at the 700 hPa and 850 hPa pressure levels. A strong forcing case was identified when a cold front, a closed low, a trough, a low-pressure system, or large wind speeds (greater than 5 ms^{-1}) was present on the surface, 700 hPa, or 850 hPa weather maps. In contrast, a high-pressure system, a ridge, or low wind speeds (less than 5 ms^{-1}) at 700 hPa or 850 hPa was identified as a weak forcing case. Overall, three weak synoptic forcing cases (i.e., 0000 UTC 21 to 1800 UTC 23 September; 0000 UTC 27 to 1800 UTC 29 September; and 0000 UTC 13 to 1800 UTC 15 October) and three strong forcing cases (i.e., 0000 UTC 16 to 1800 UTC 18 September; 0000 UTC 3 to 1800 UTC 5 October; and 0000 UTC 5 to 1800 UTC 7 October) were identified. Figure 3.26 compares the errors between the weak and strong forcing cases. For each case, four forecasts with different initial

times were compared. It is apparent that diurnal patterns were present in the forecast errors for the weak forcing cases. The errors were independent of initialization time and forecast lead time (Figure 3.26a, c, e). The strong forcing cases, however, showed flow-dependent features (Figure 3.26b, d, f). The forecast errors did not follow a diurnal pattern, implying that the errors become more closely related to the influence of the weather systems. In addition, the magnitude of the errors was generally greater in the strong forcing than in the weak forcing cases.

3.7 Summary and Discussion

In this study, the performance of version 3.3 of the ARW WRF model in predicting near-surface atmospheric temperature and wind conditions under various terrain and weather regimes was evaluated. Three individual events under strong synoptic forcing, namely, a frontal system, a low-level jet, and a persistent cold air pool were first verified against observations over both flat and complex terrain. It was found that the ARW model is able to produce reasonable simulations of weather phenomena. Verification of near-surface conditions (i.e., 2-m temperature and 10-m wind) indicated the complexity in forecasting these surface variables. For the frontal case and low-level jet case over the central US, the model terrain matches the actual terrain and thus mitigates representative errors. The forecasts of surface variables generally agreed well with observations. However, errors still occurred depending on the model's ability in forecasting the structures in the lower atmospheric boundary layer. For the inversion case over the Salt Lake Valley, different error characteristics were found over the mountain and valley stations. Terrain mismatch and the ARW model's limited ability to simulate near-surface

atmospheric conditions make the forecasting errors even more complicated.

Overall, forecast errors in near-surface atmospheric variables showed flow-dependent features in all three individual cases when strong synoptic forcings were present. To better understand the characteristics of flow-dependent errors in complex terrain as they relate to near-surface forecasting, 1-month forecasts (from 15 September to 14 October 2011) were conducted in complex terrain in the western US. It was found that the forecast errors of surface variables depend to a large degree on the diurnal cycle of the surface variables themselves, especially when the synoptic forcing is weak. The forecast errors for 2-m temperature reached two daily maxima at 3 am and 3 pm local time, and two daily minima at 7 am and 7 pm local time. Errors in wind speed and direction followed the same trends, with a maximum at night and a minimum in the afternoon. Forecast errors followed the same trends regardless of the initialization time, showing that forecast errors are independent of the initialization time and forecast lead time. Further analyses revealed positive (warm) temperature biases at night and negative (cold) biases during the daytime. In contrast to the 2-m temperature, wind direction and speed had no systematic biases from a long-term perspective. Under strong synoptic forcing, diurnal patterns in forecast errors were absent, while flow-dependent errors were clearly shown.

Finally, it is apparent that simulations at finer resolutions did not outperform those at coarser resolutions in most cases. This was explained well in Hart et al. (2005), who found that the inability of the numerical model to depict near-surface structures (such as strong temperature inversions) results in worse forecasts, even with better terrain representation. Meanwhile, increasing the model's vertical resolution did not help

improve its forecast skill of near-surface variation, although it improved the forecasts of mesoscale weather phenomena. Numerical forecasts of near-surface atmospheric conditions are also sensitive to the PBL scheme in the ARW WRF model, but there is no single PBL scheme that performs better than the others. These factors illustrate the complexity and challenges involved in near-surface simulation over complex terrain. Future work should emphasize investigating the decoupling between near-surface variables and the atmospheric boundary layer and its impact on the model's prediction skill of near-surface variables. The sensitivity of numerical predictions of near-surface variables to terrain representation, land surface parameters, and model errors will also need to be examined. In addition, uncertainties in model initial conditions could contribute to forecast errors. Therefore, it is necessary to improve model initial conditions using advanced data assimilation methods in order to enhance prediction skill.

Table 3.1 Configurations of numerical simulations

Case	1-3 June 2008	1-3 Dec. 2010	Fall 2011 (Sep. 15 - Oct. 14)
Number of domains	3	3	4
Horizontal grid spacing (km)	27/9/3	12/4/1.33	30/10/3.33/1.11
Microphysics scheme	WSM 6	WSM 6	Purdue Lin
Planetary boundary layer scheme	YSU	MYJ	YSU
Cumulus scheme	Kain-Fritsch (*not applied for grid spacing less than 9km*)		
Land surface scheme	Noah		
Longwave radiation	Rapid Radiative Transfer Model		
Shortwave radiation	Dudhia		

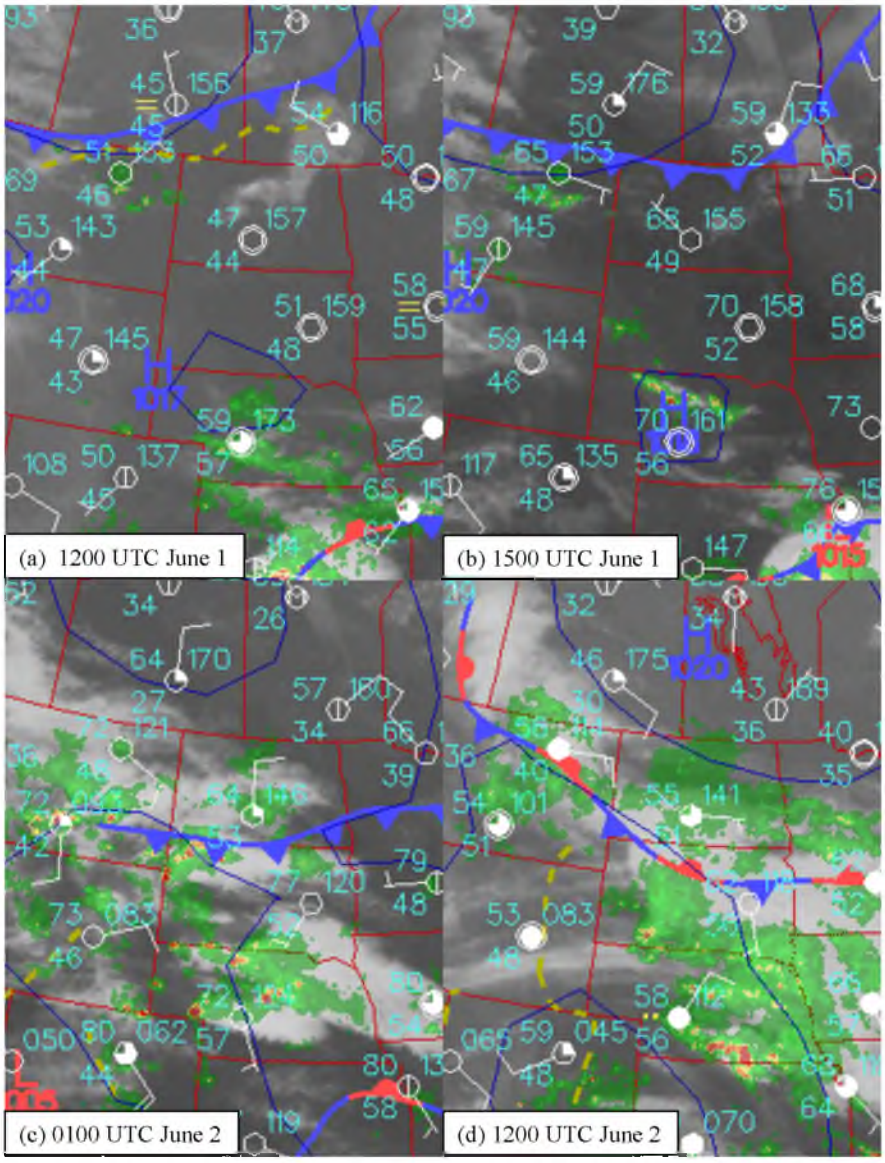


Figure 3.1 Surface weather maps showing the evolution of the front for June 1–3 front case. (Images are adapted from NCAR image archive at http://locust.mmm.ucar.edu/)

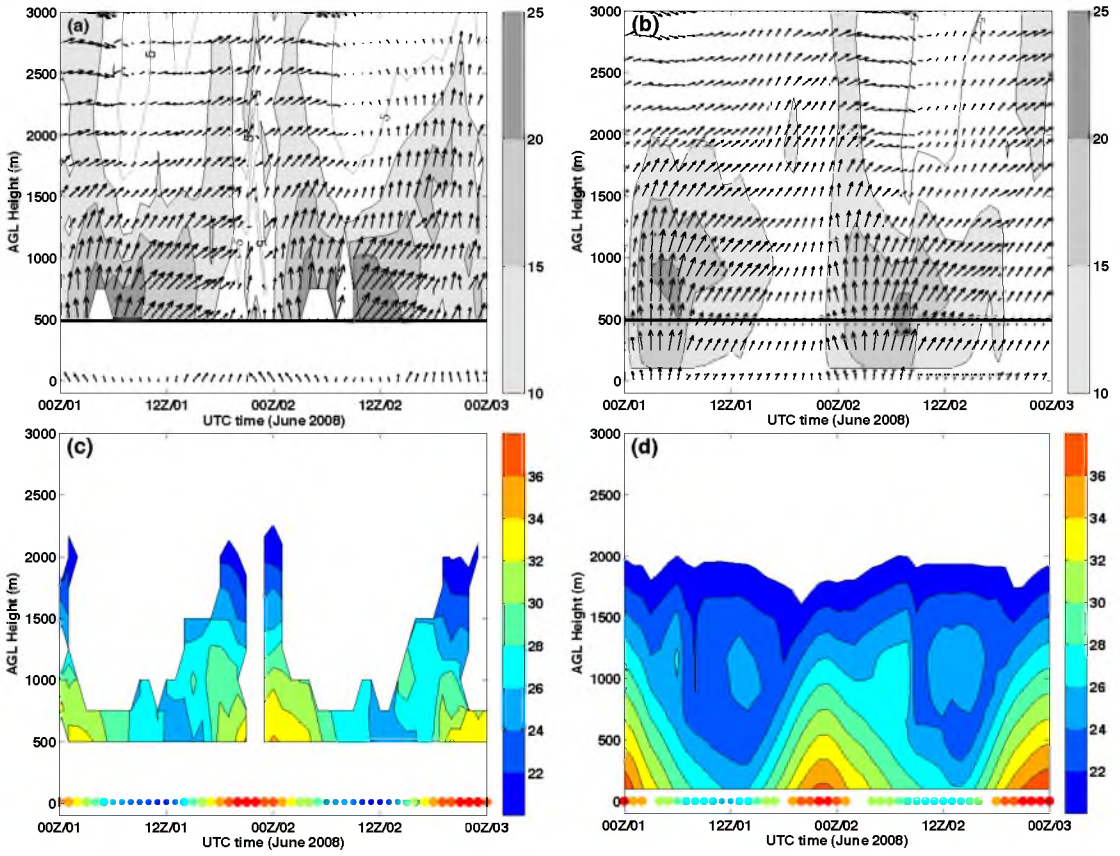


Figure 3.2 Time series of vertical profiles from 0000 UTC 1 to 0000 UTC 3 June 2008 at Jayton, Texas (33.01°N, 100.98°W; elevation: 707 m): (a) wind speeds (contour interval: 5 ms^{-1}) and vectors obtained from NOAA radar profiler and near-surface (10-m) winds from surface Mesonet observations, (b) wind speeds and vectors of ARW model simulation from the 3km domain, (c) temperatures obtained from NOAA radar profiler and near surface (2-m) temperature from surface Mesonet, and (d) temperatures of model simulation from the 3-km domain. In (a) and (b), wind speeds greater than 10 ms^{-1} are shaded. NOAA profiler data is not available below 490 m AGL.

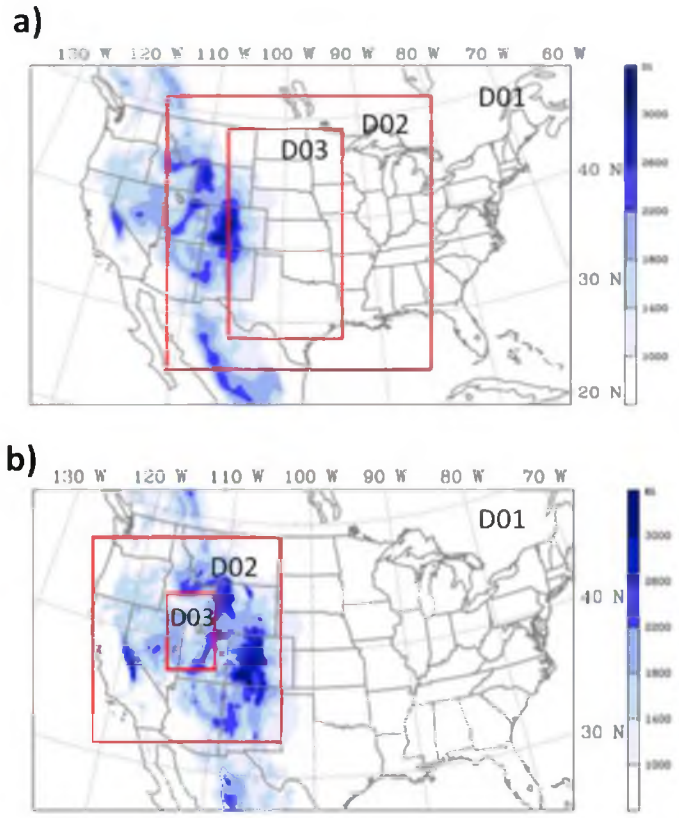


Figure 3.3 Locations of model domains for numerical simulations: (a) 0000 UTC 1 to 0000 UTC 3 June 2008 and (b) 0000 UTC 1 to 0000 UTC 3 December 2010. Shaded contours denote the terrain heights.

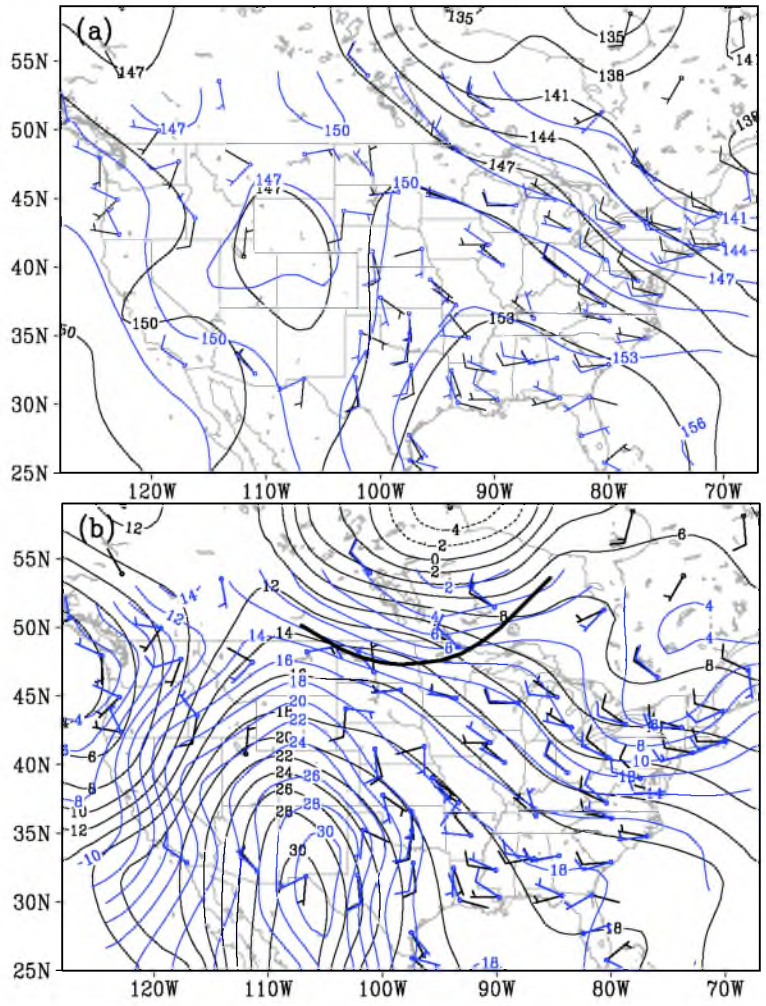


Figure 3.4 Weather maps at 850 hPa valid at 0000 UTC 2 June 2008. (a) Geopotential heights (contour interval: 30 m) and wind barbs and (b) Temperature (contour interval: 2 °C) with wind barbs. Black contour lines and wind barbs represent observations from the upper level observation network. Blue contour lines and wind barbs denote model simulations. The thick black curve in (b) marks the cold front.

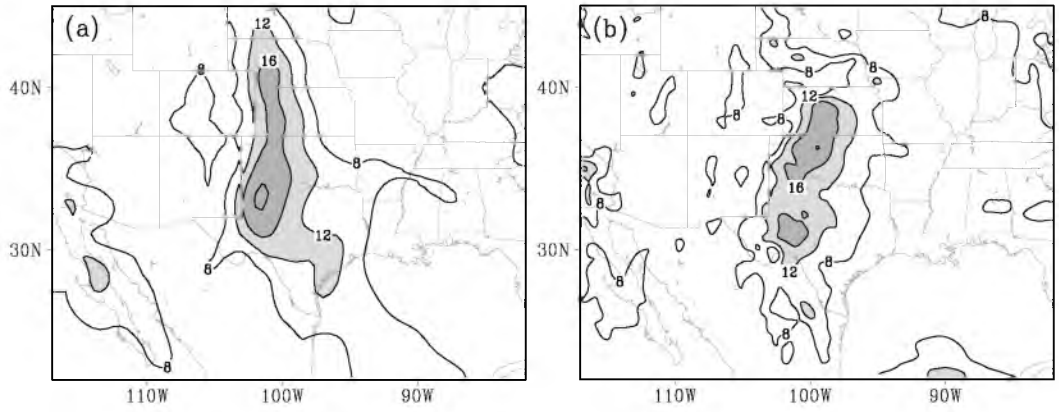


Figure 3.5 Horizontal wind speeds (contour interval: 4 ms^{-1}) valid at 0900 UTC 2 June 2008 at 850 hPa: (a) NARR reanalysis and (b) model simulation of the 3-km domain. Wind speeds greater than 12 ms^{-1} are shaded.

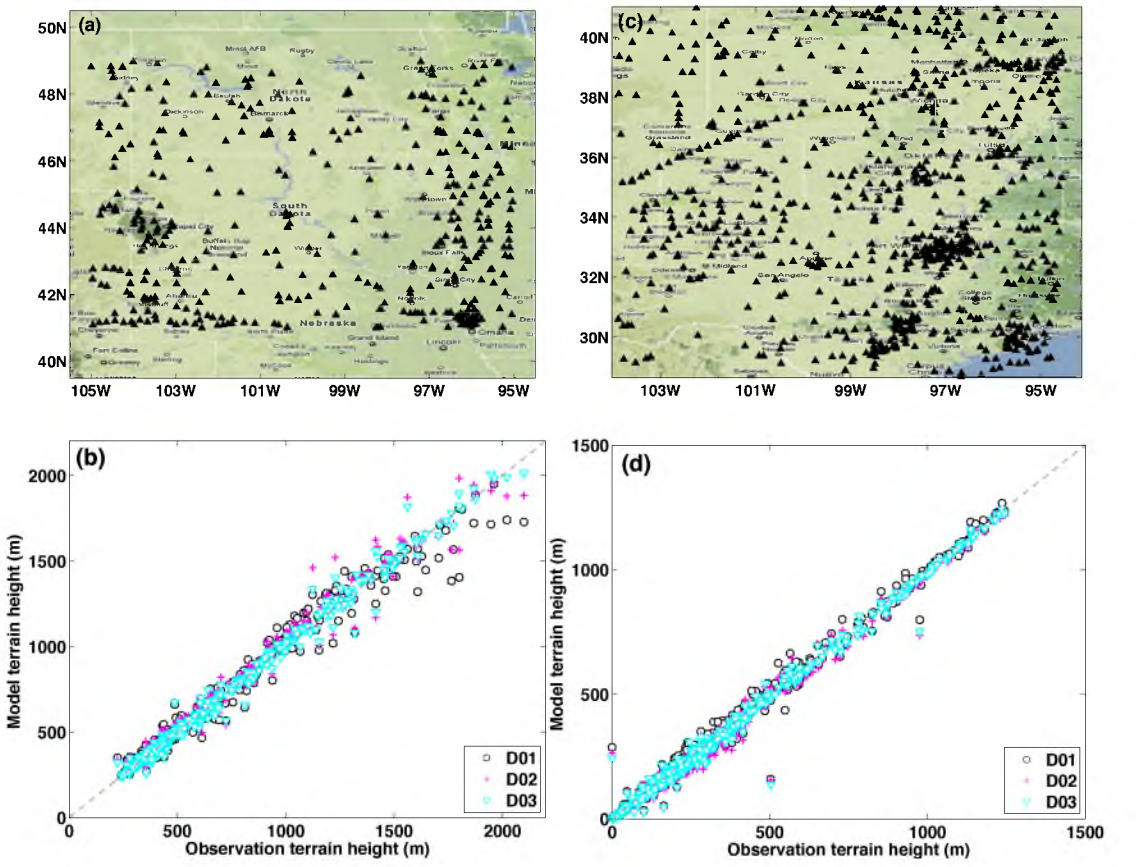


Figure 3.6 The area and Mesonet observation stations used for verification: (a) for the front case and (c) for the low-level jet case. (b) and (d) compare the actual and model terrain heights for stations in (a) and (c), respectively. D01, D02, and D03 represent the domains at 27-km, 9-km, and 3-km grid spacings, respectively. The straight dashed lines in (b) and (d) denote Y = X.

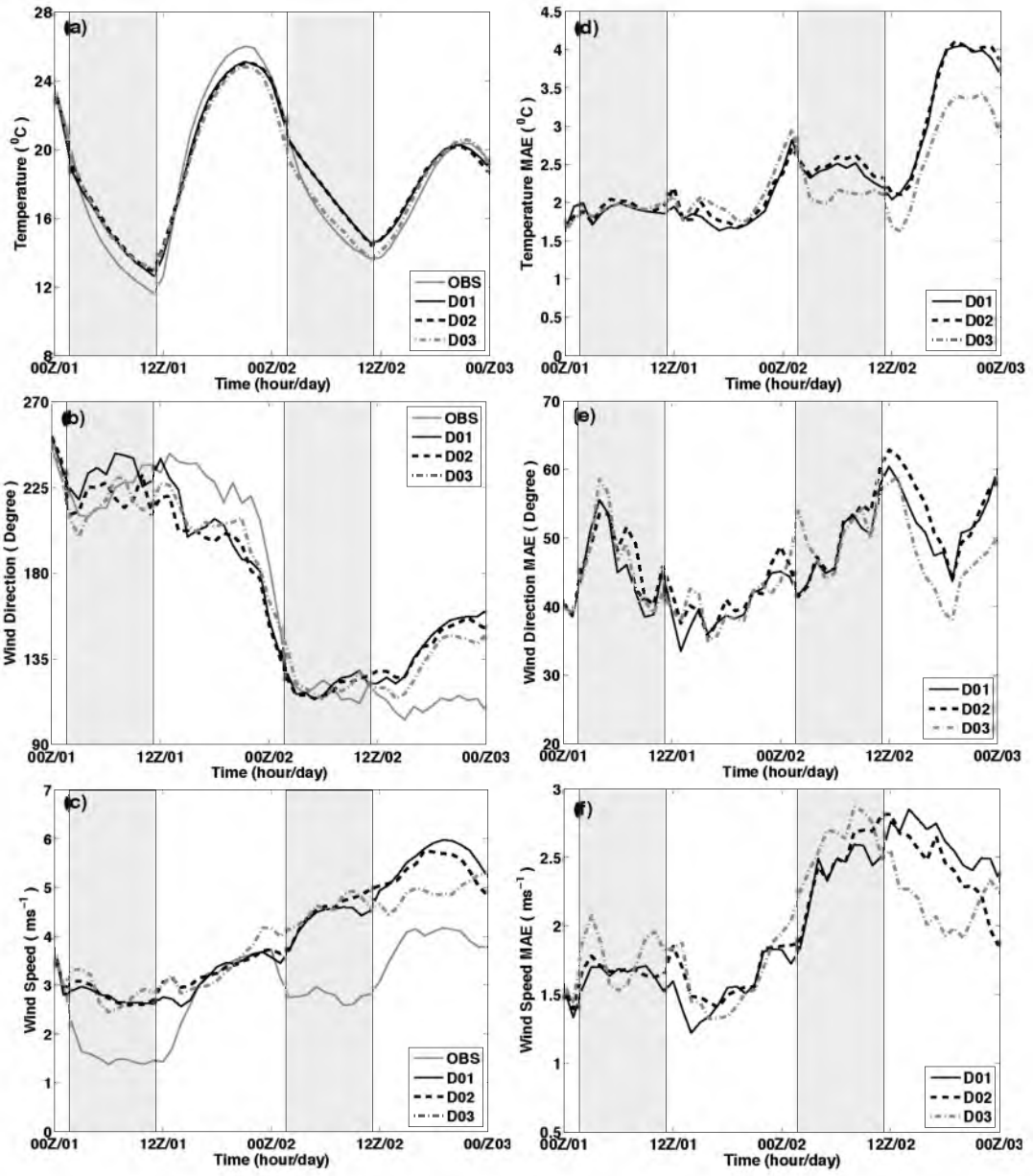


Figure 3.7 Comparison of (a) 2-m temperature, (b) 10-m wind direction, and (c) 10-m wind speed between observations and simulations averaged over the area of the frontal system (e.g., the domain covered in Fig. 31a) and mean absolute errors of simulated (d) 2-m temperature, (e) 10-m wind direction, and (f) 10-m wind speed. D01, D02, and D03 represent results from model domains at grid spacings of 27-km, 9-km, and 3-km, respectively. *The gray shaded areas represent the nighttime, hereafter.*

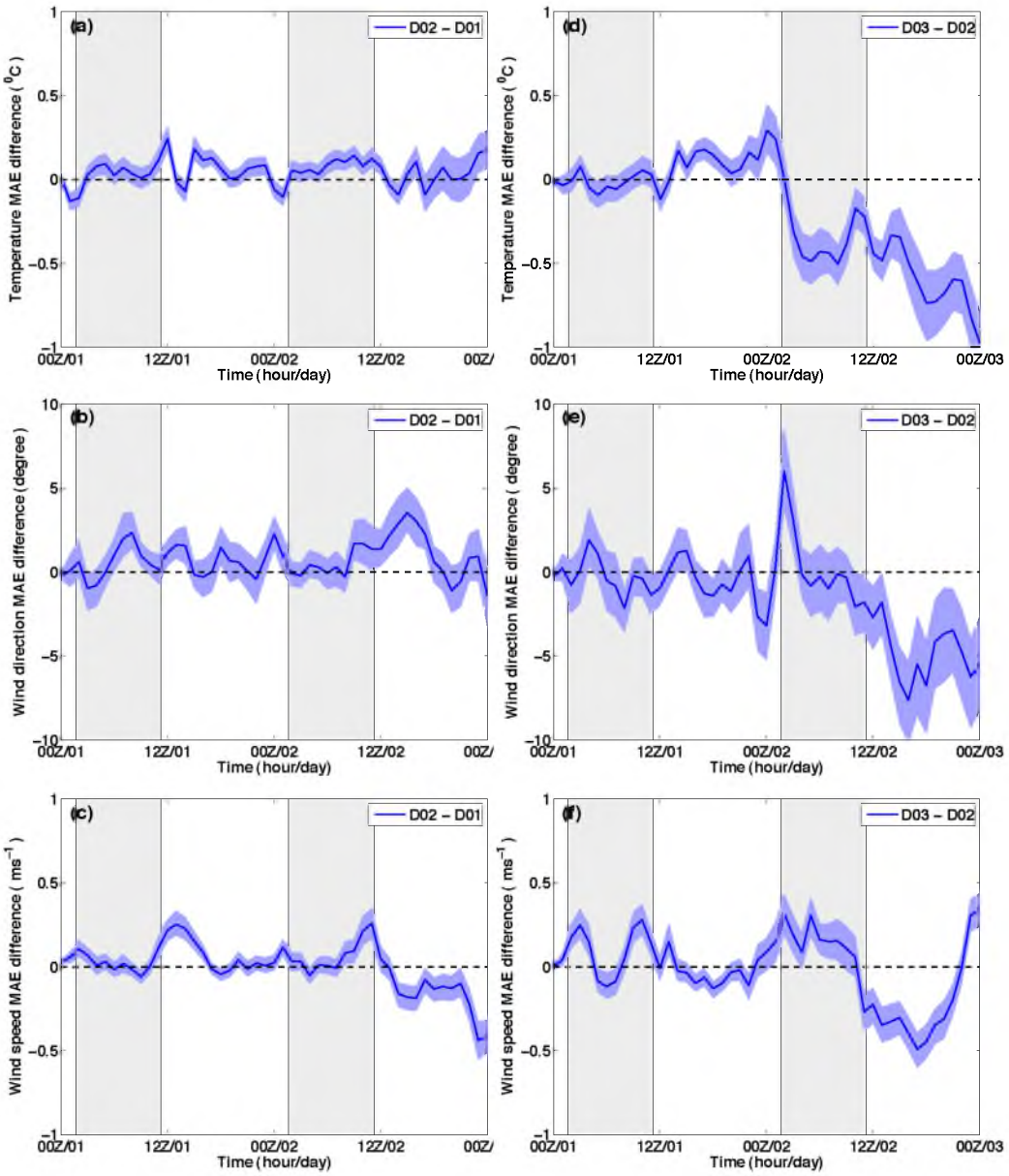


Figure 3.8 Differences of MAEs of (a) and (d) 2-m temperature, (b) and (e) 10-m wind direction, and (c) and (f) 10-m wind speed between D02 and D01 (D02 minus D01, a, b, and c) simulations and between D03 and D02 (D03 minus D02, d, e, and f) simulations averaged over the area of the frontal system. D01, D02, and D03 represent results from model domains at grid spacings of 27-km, 9-km, and 3-km, respectively. Blue shaded area represents 90% confidence intervals derived using the bias-corrected and accelerated (BCa) bootstrapping technique.

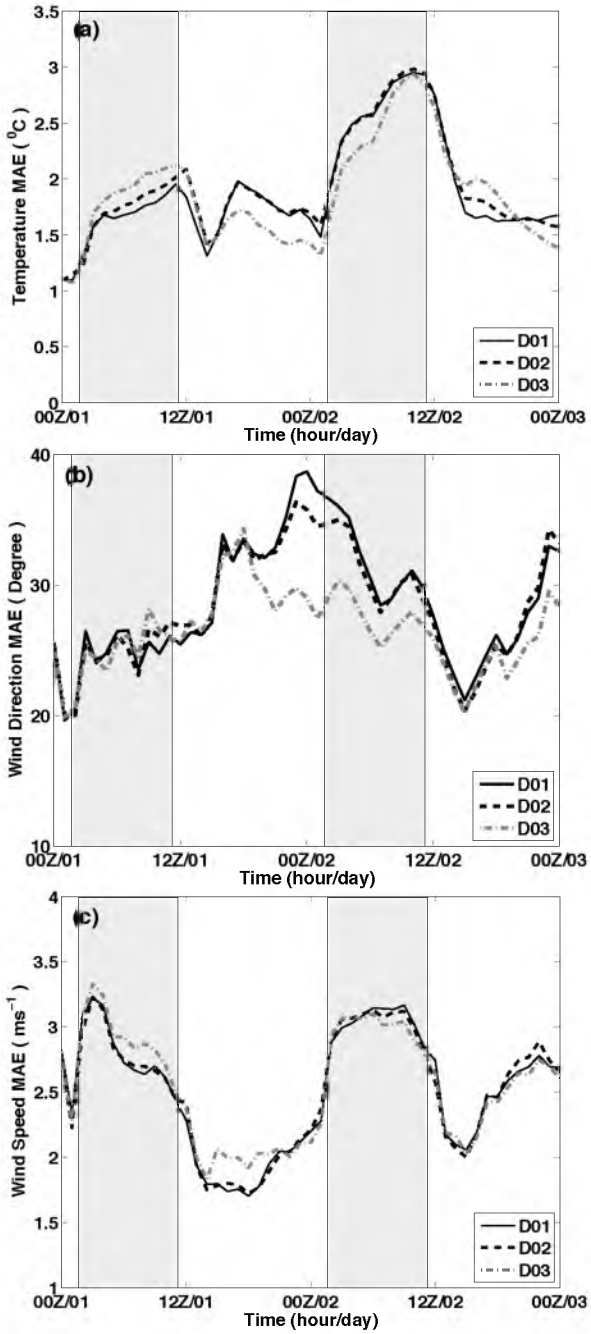


Figure 3.9 Mean absolute errors over low-level jet area of simulated (a) 2-m temperature, (b) 10-m wind direction, and (c) 10-m wind speed. D01, D02, and D03 represent results from model domains at grid spacings of 27-km, 9-km, and 3-km, respectively.

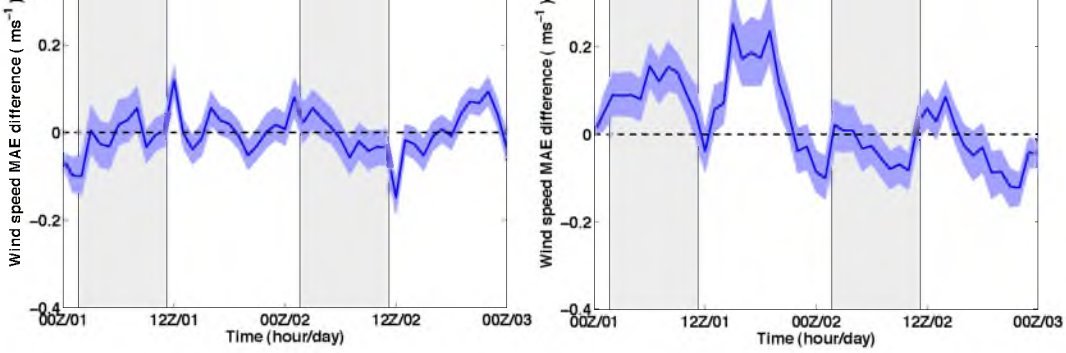
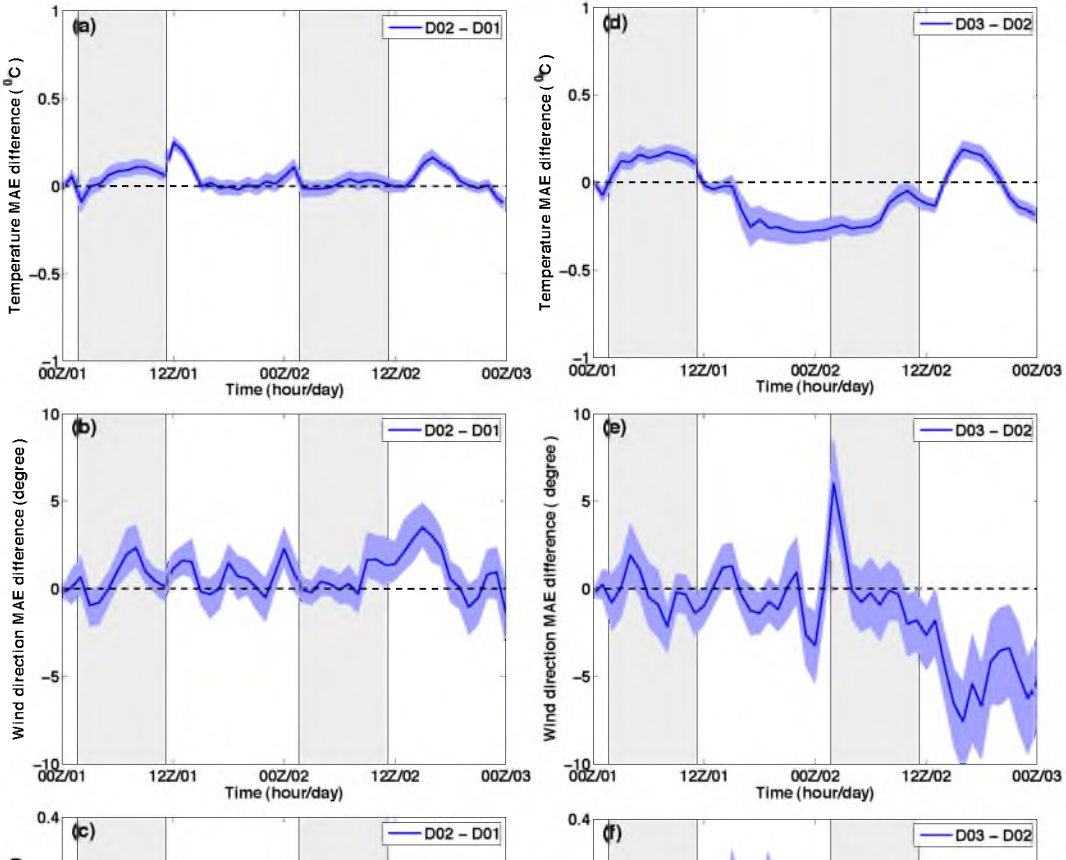


Figure 3.10 Same as Figure 3.8, but for the low-level jets.


(a)
D02 – D01
Temperature MAE difference (°C)
1
0.5
0
−0.5
−1
00Z/01
12Z/01
00Z/02
12Z/02
00Z/03
Time (hour/day)
(d)
D03 – D02
Temperature MAE difference (°C)
(b)
D02 – D01
Wind direction MAE difference (degree)
10
5
0
−5
−10
(e)
D03 – D02
Wind direction MAE difference (degree)
(c)
D02 – D01
0.4
(f)
D03 – D02

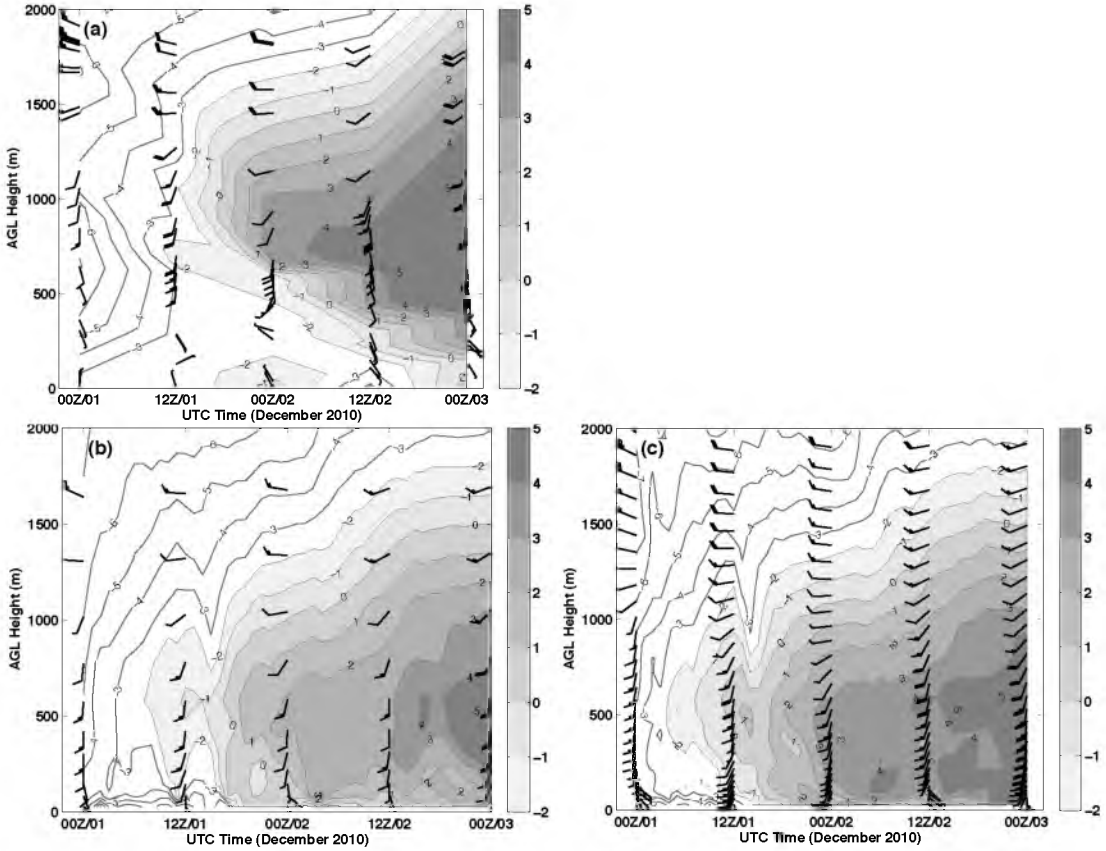


Figure 3.11 Time series of vertical profiles of temperature (contour interval: 1 °C) and wind at the Salt Lake City International Airport, Utah (KSLC, 40.77°N, 111.85°W, elevation: 1289 m): (a) sounding observations, (b) ARW simulation from the 1.33-km domain with 37 vertical levels, and (c) ARW simulation from the 1.33-km domain with 70 vertical levels. Shaded contours represent temperatures greater than 2 °C.

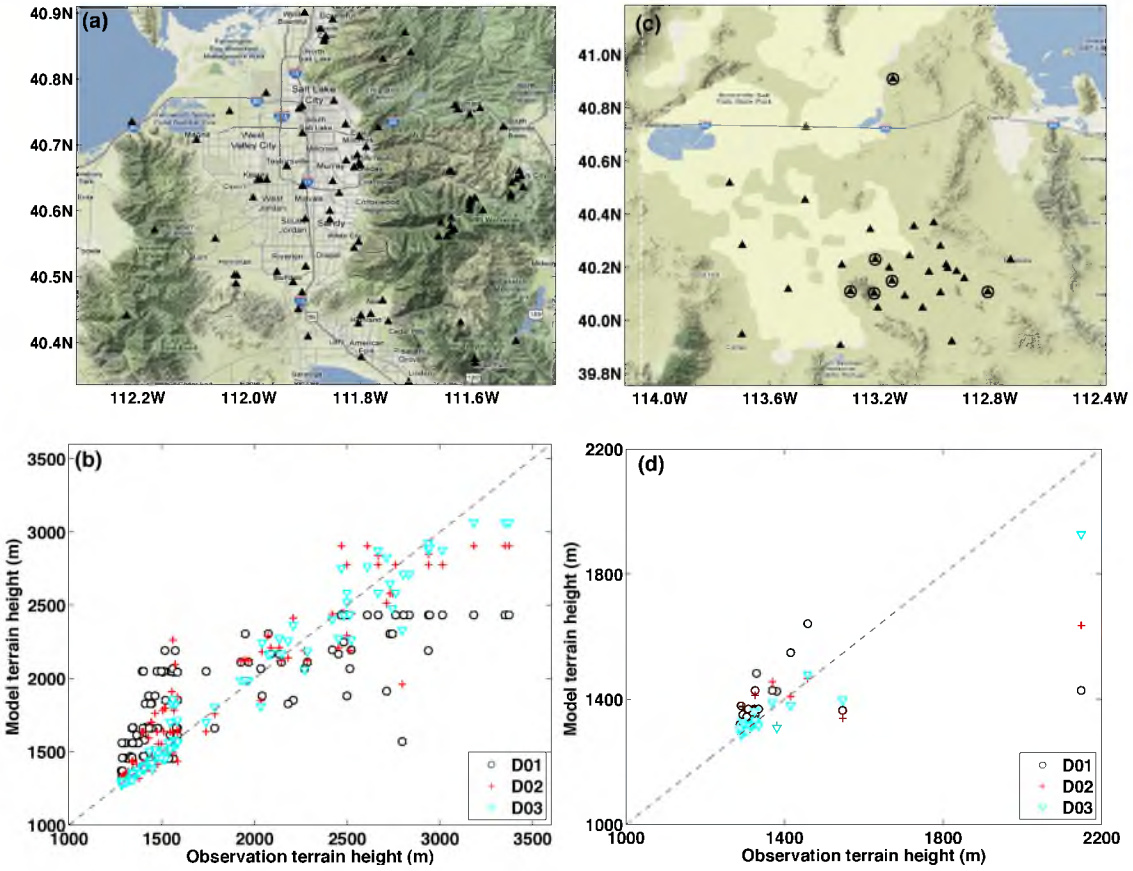


Figure 3.12 Same as Figure 3.6, except for the persistent inversion case over the Salt Lake Valley (a and b), and the Dugway Proving Ground (c and d). D01, D02, and D03 represent the domains at 12-km, 4-km, and 1.33-km horizontal grid spacings, respectively.

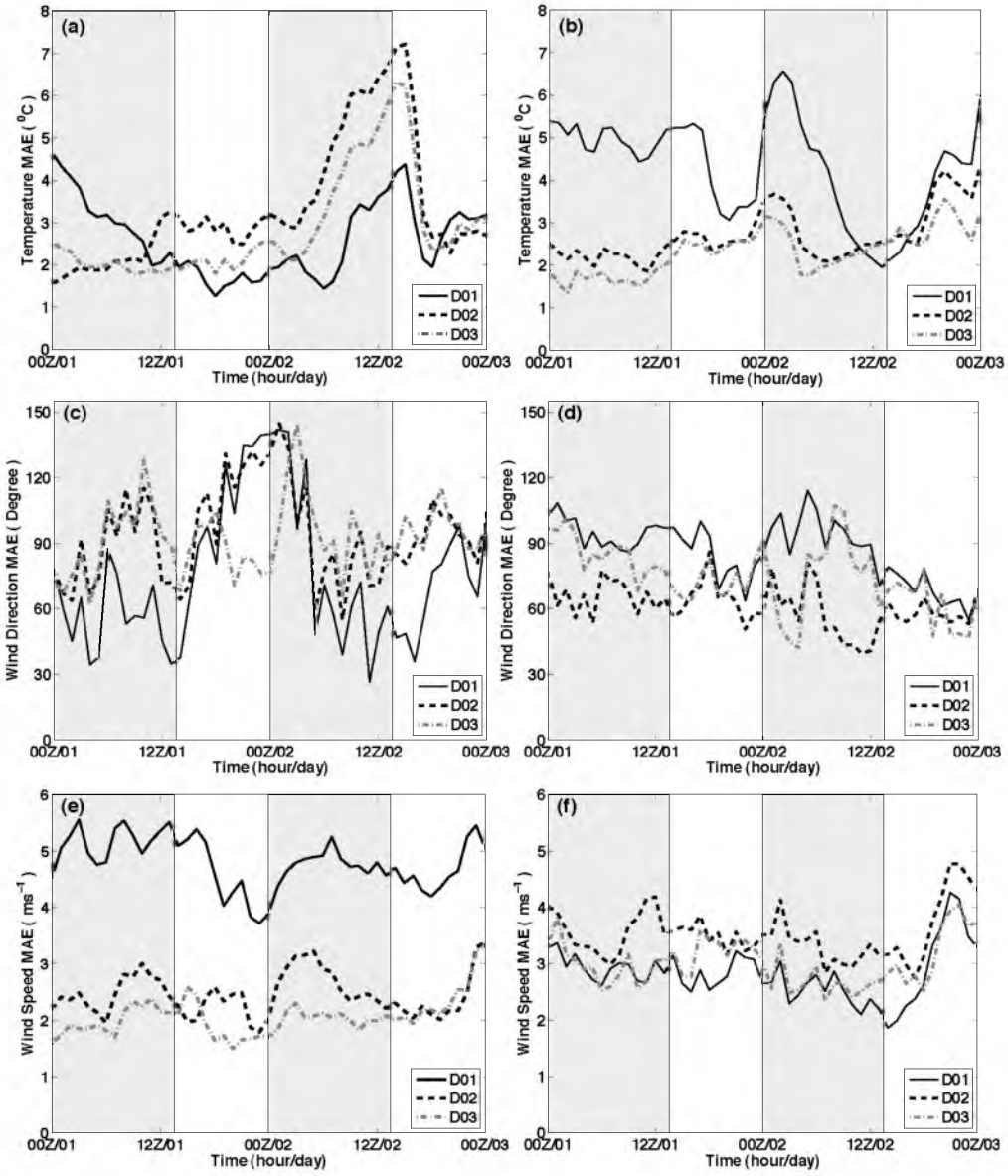


Figure 3.13 Mean absolute errors (MAEs) by station types for numerical simulations of 2-m temperature [(a) and (b)], 10-m wind direction [(c) and (d)], and speed [(e) and (f)]. Figures on the left column [(a), (c), and (e)] represent MAEs for valley stations, and these on the right column [(b), (d), and (f)] represent mountain stations. D01, D02, and D03 represent the domains at 12-km, 4-km, and 1.33-km horizontal grid spacings, respectively.

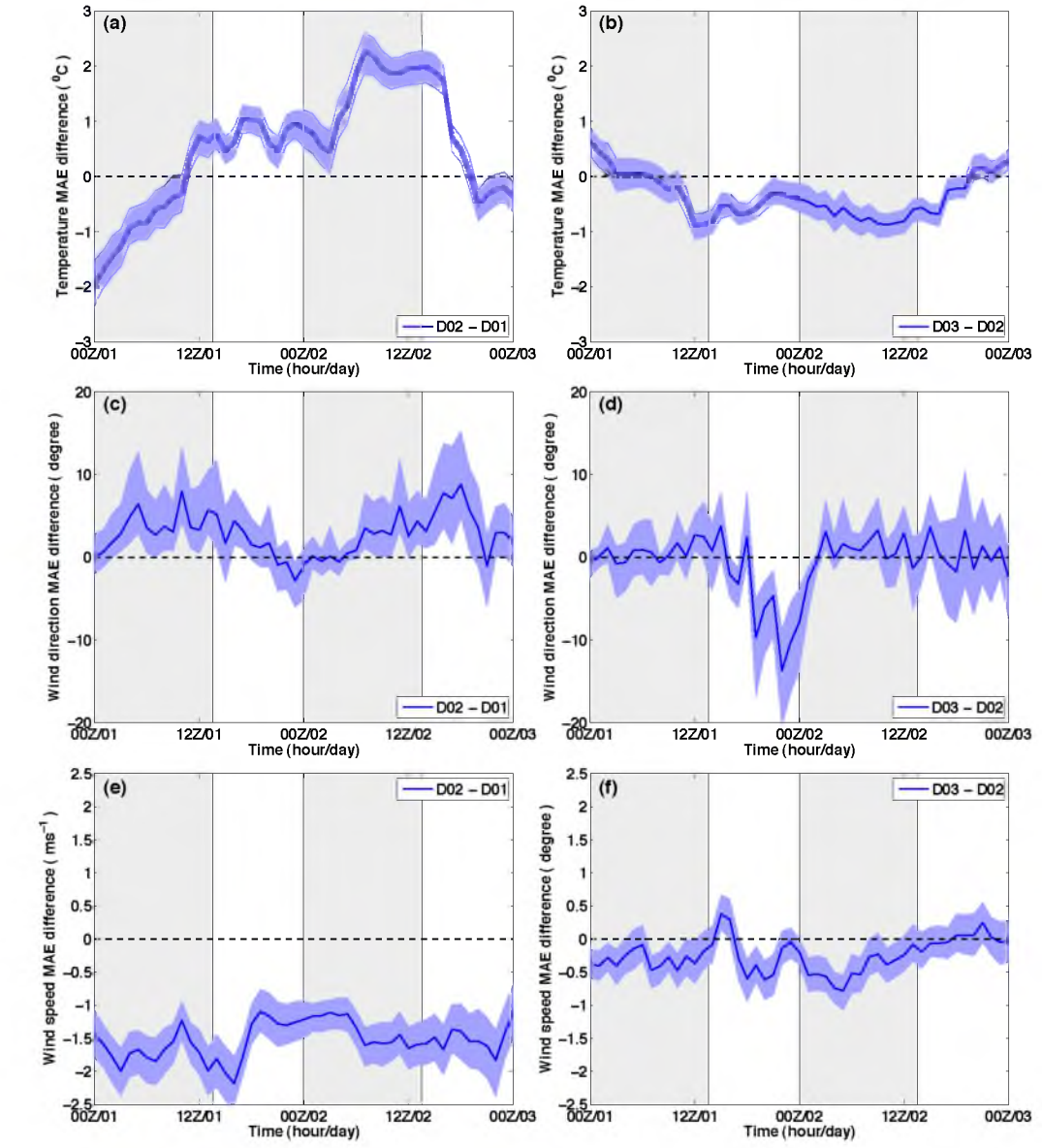


Figure 3.14 Differences of MAEs of (a) and (b) 2-m temperature, (c) and (d) 10-m wind direction, and (e) and (f) 10-m wind speed between D02 and D01 (D02 minus D01, a, c, and e) simulations and between D03 and D02 (D03 minus D02, b, d, and f) simulations averaged over valley stations. D01, D02, and D03 represent the domains at 12-km, 4-km, and 1.33-km horizontal grid spacings, respectively. Blue shaded area represents 90% confidence intervals derived using the BCa bootstrapping technique.

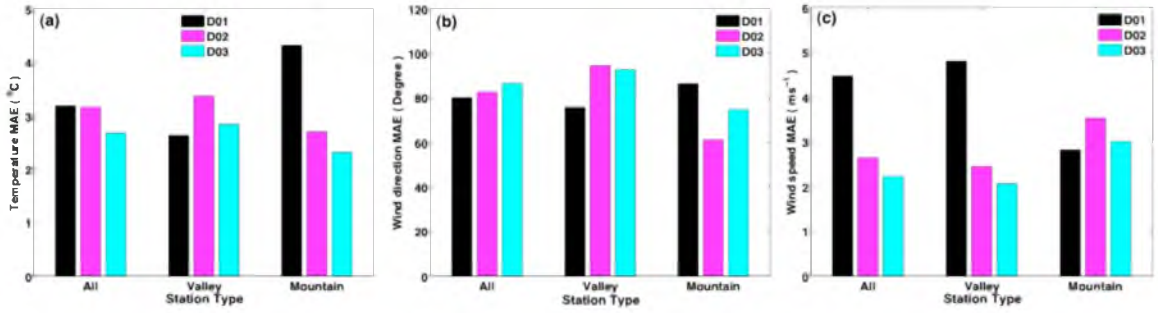


Figure 3.15 Cumulative MAEs of (a) 2-m temperature, (b) 10-m wind direction, and (c) 10-m wind speed by station types. D01, D02, and D03 represent the simulation from the domains at 12-km, 4-km, and 1.33-km horizontal grid spacings, respectively.

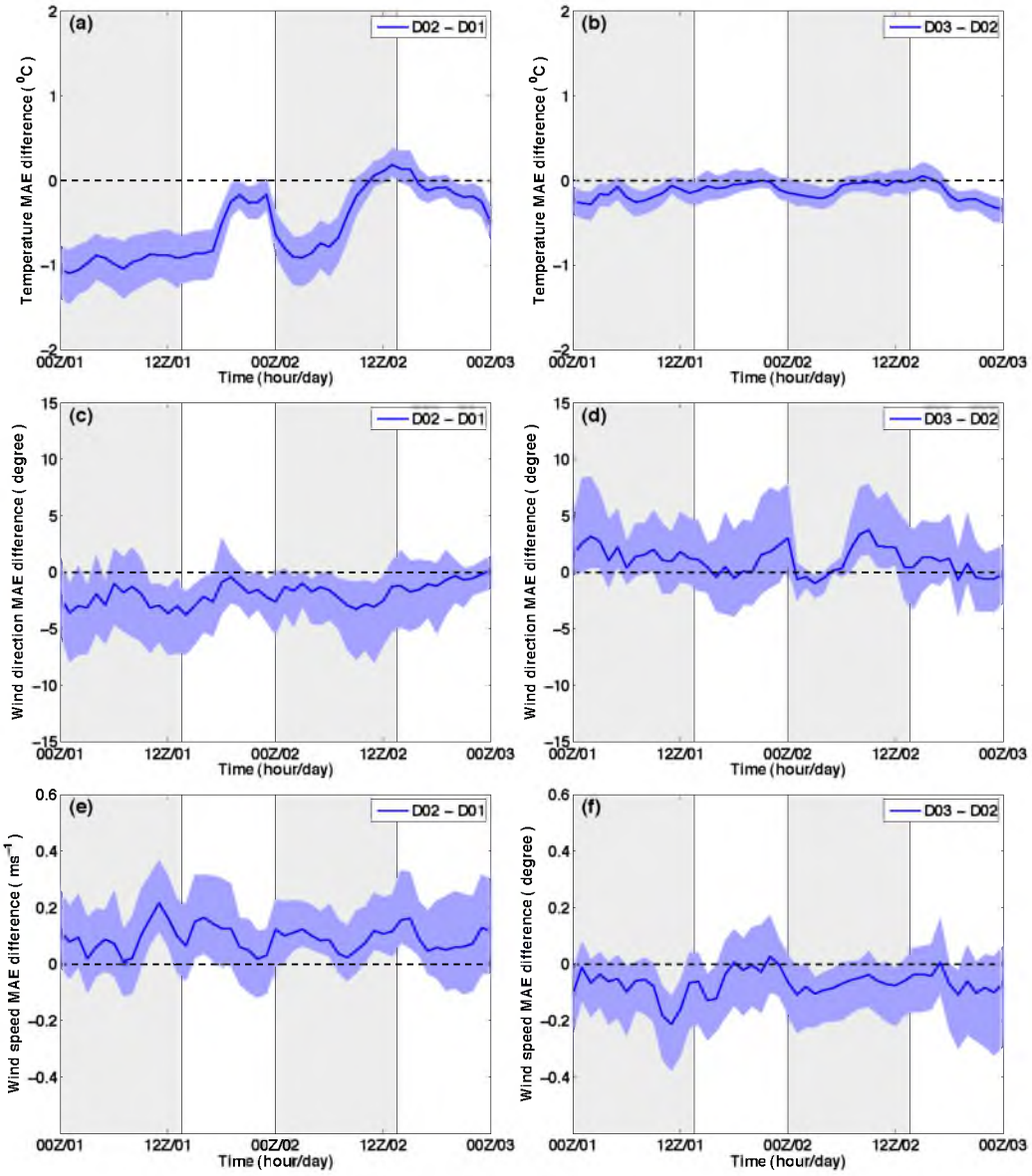


Figure 3.16 Same as Figure 3.14, but for mountain stations.

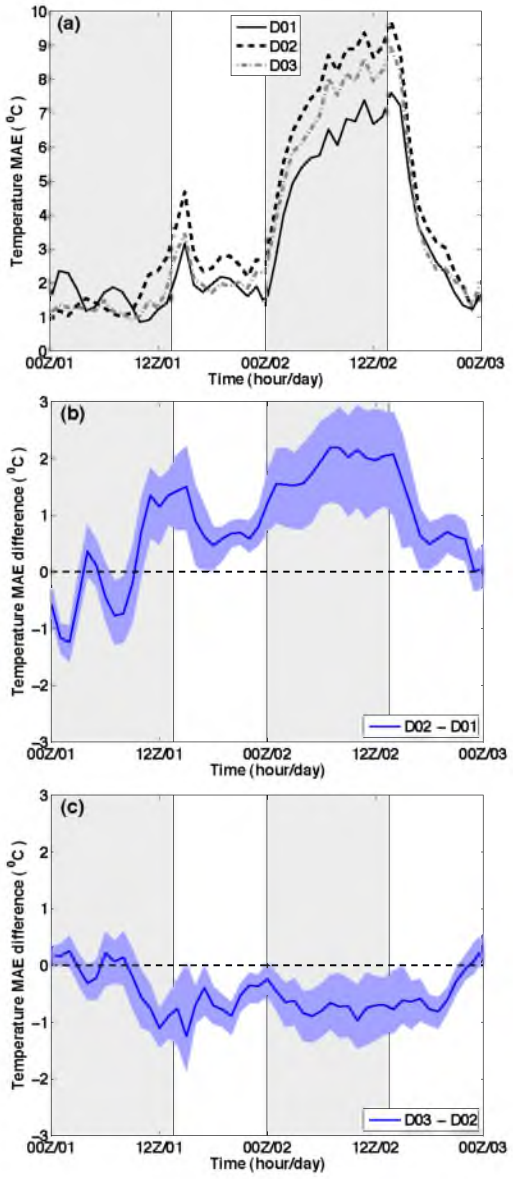


Figure 3.17 Comparison of Mean absolute errors of 2-m temperature among different domains: (a) Mean absolute errors of 2-m temperature at Dugway Proving Ground (DPG). Differences of MAEs of 2-m temperature between (b) D02 and D01 (D02 minus D01) simulations and between (c) D03 and D02 (D03 minus D02) simulations at Dugway Proving Ground. D01, D02, and D03 represent the simulation from the domains at 12-km, 4-km, and 1.33-km horizontal grid spacings, respectively. Blue shaded area represents 90% confidence intervals derived using the BCa bootstrapping technique.

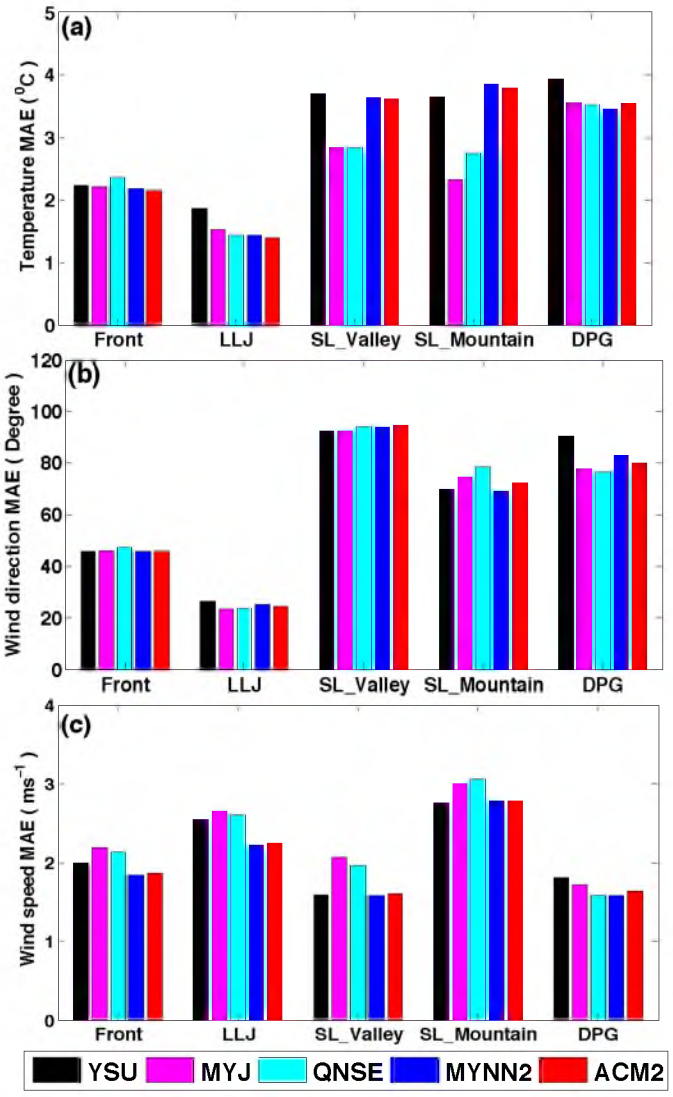


Figure 3.18 Cumulative MAEs of (a) 2-m temperature, (b) 10-m wind direction, and (c) 10-m wind speed for various cases and locations with various PBL schemes. All results are from the innermost domain.

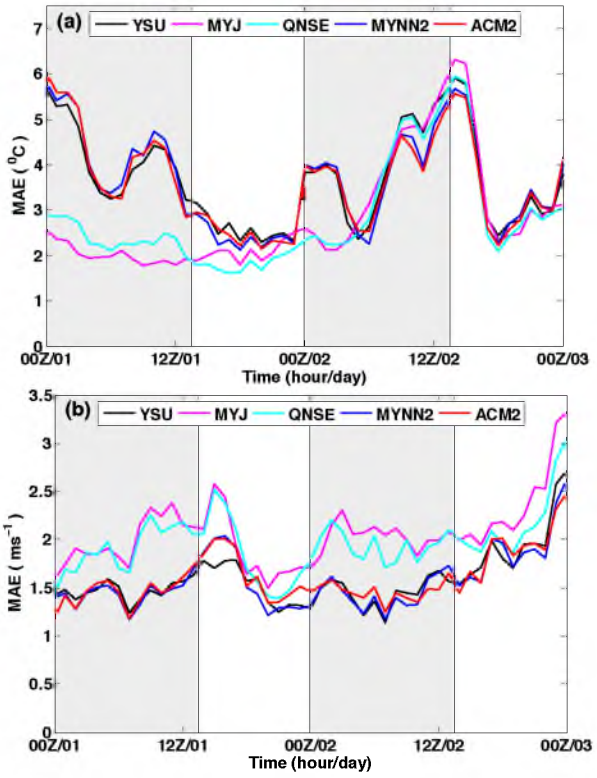


Figure 3.19 MAEs of (a) 2-m temperature, (b) 10-m wind speed with various PBL schemes from the 1.33-km domain for Salt Lake Valley stations.

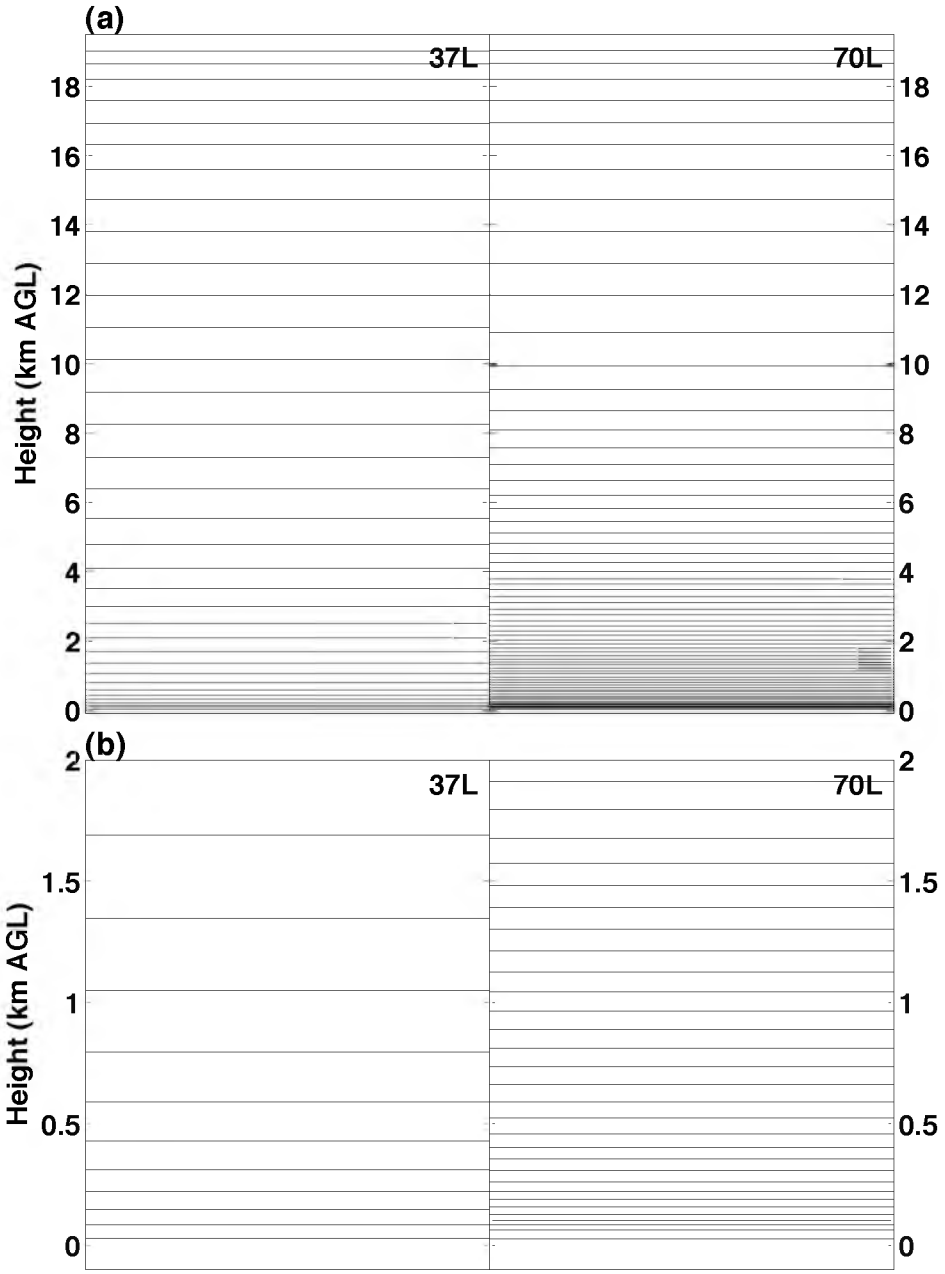


Figure 3.20 Sketches of vertical model vertical levels with AGL heights: (a) full levels and (b) levels in the bottom 2 km. The left panels are for 37 vertical levels and the right panels are for 70 vertical levels.

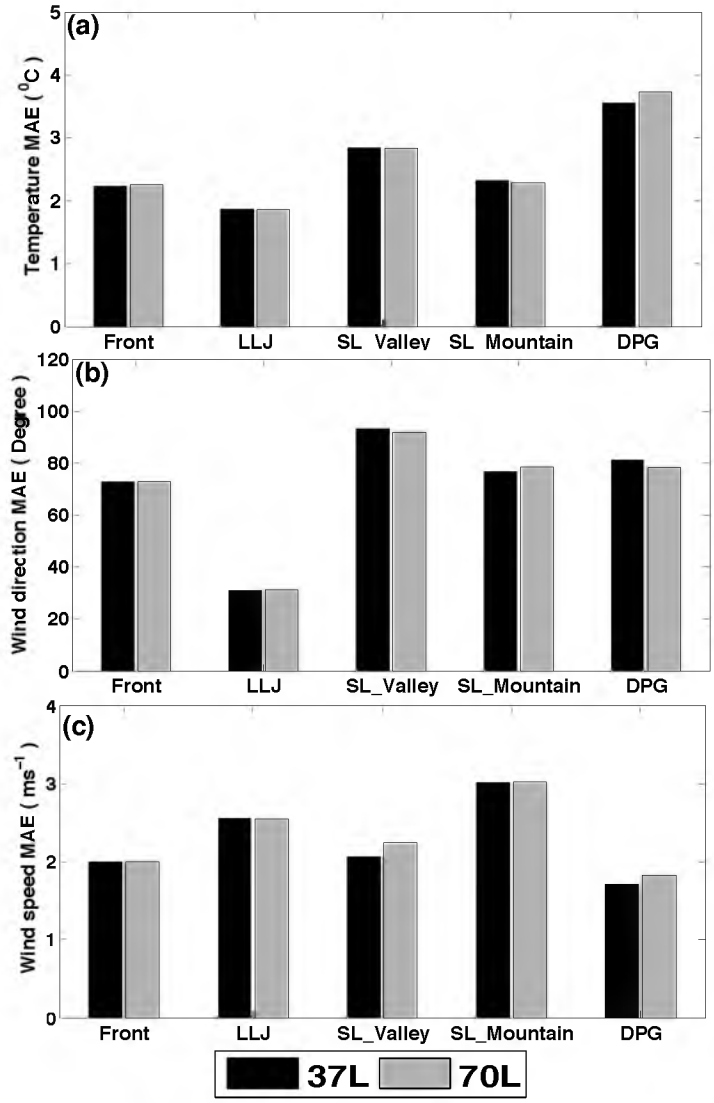


Figure 3.21 Same as Figure 3.18, except for different model vertical resolutions.

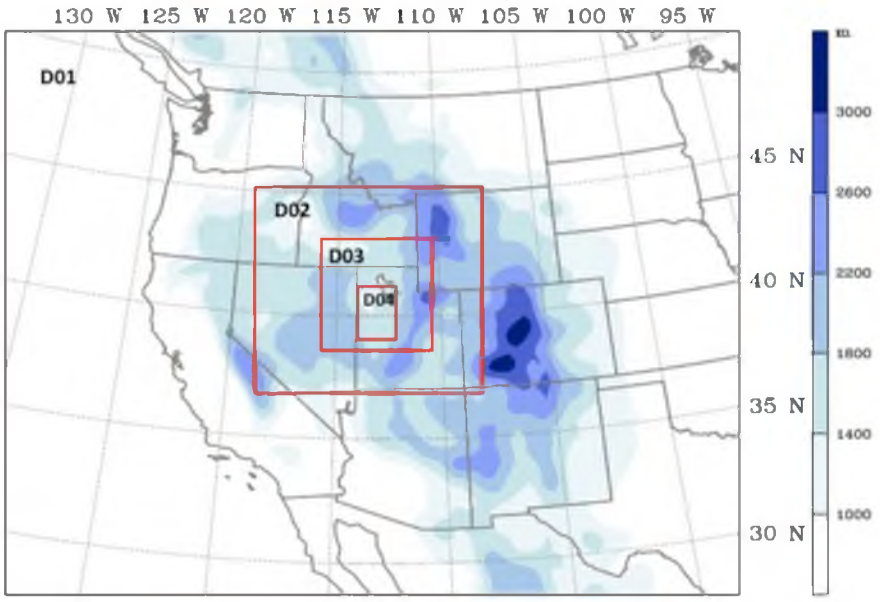


Figure 3.22 Locations of model domains for near real-time forecasting from 15 September to 14 October 2011.

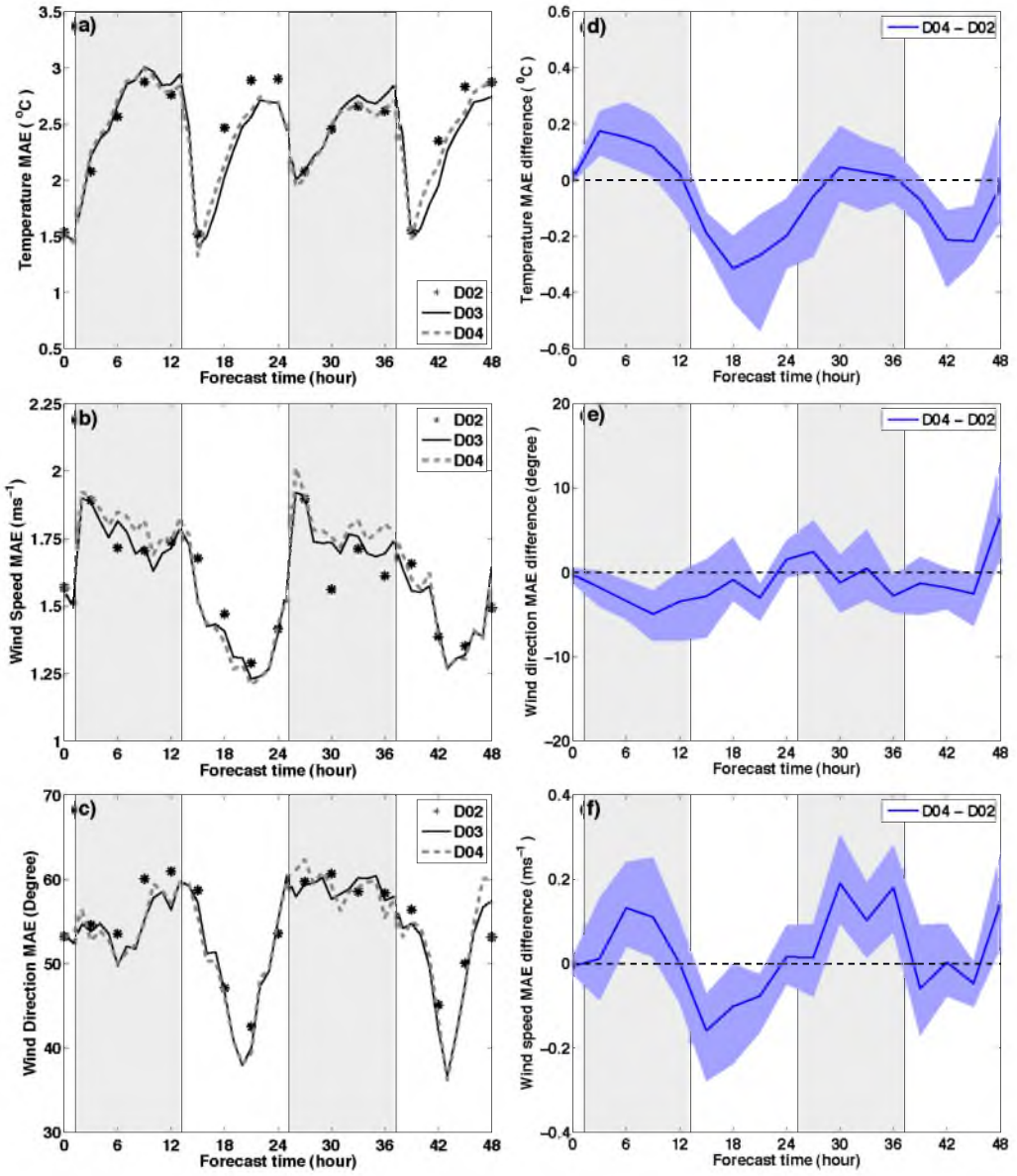


Figure 3.23 Mean absolute errors of simulated near-surface variables for different model domains (a, b, and c) and differences of MAEs of (d) 2-m temperature, (e) 10-m wind direction, and (f) 10-m wind speed between D04 and D02 (D04 minus D02) simulations. Results are from 0000 UTC forecasts and D02, D03, and D04 represent results from model domains at horizontal resolutions of 10-km, 3.33-km, and 1.11-km, respectively. The forecasts are output every 3 hours in the 10-km domain (D02), and every hour in the 3.33-km and 1.11-km domains. Blue shaded area represents 90% confidence intervals derived using the BCa bootstrapping technique.

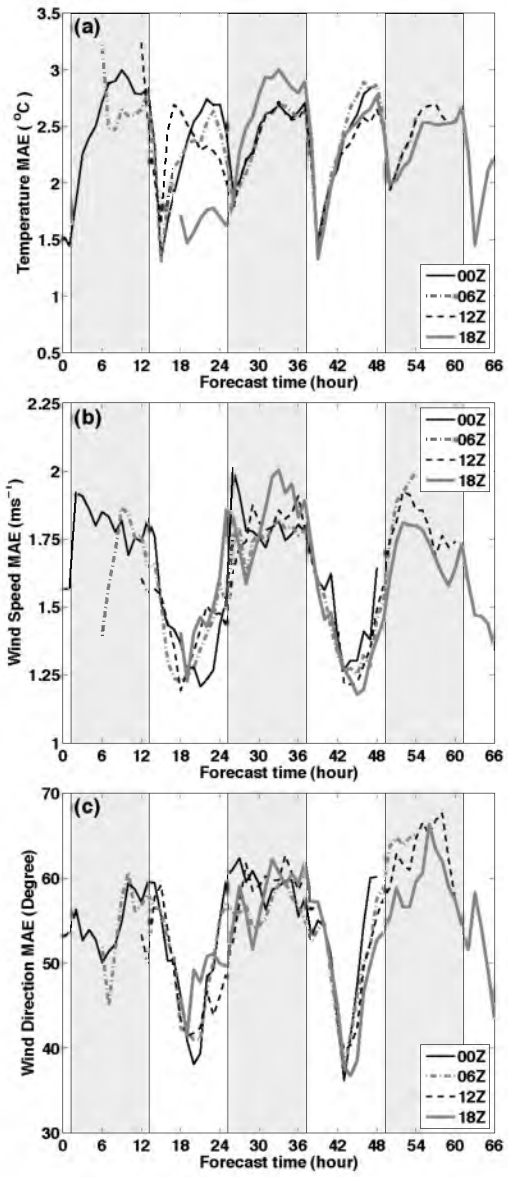


Figure 3.24 Mean absolute errors of simulated near-surface variables for various initialization times. Results are from the 1.11-km domain and various curves represent forecasts initialized at different times. The forecasting period for all forecasts is 48 h.

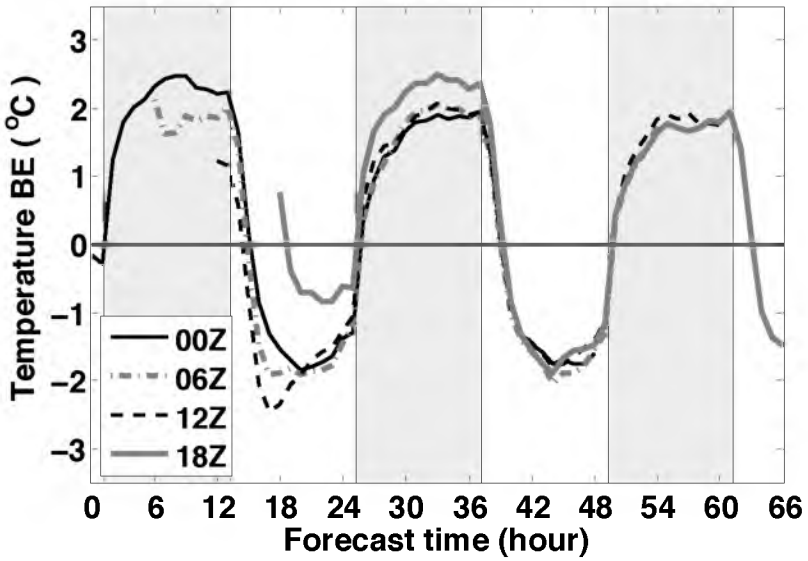


Figure 3.25 Bias error of simulated 2-m temperature from the 1.11-km domain with various initialization times. The forecasting period for all forecasts is 48 h.

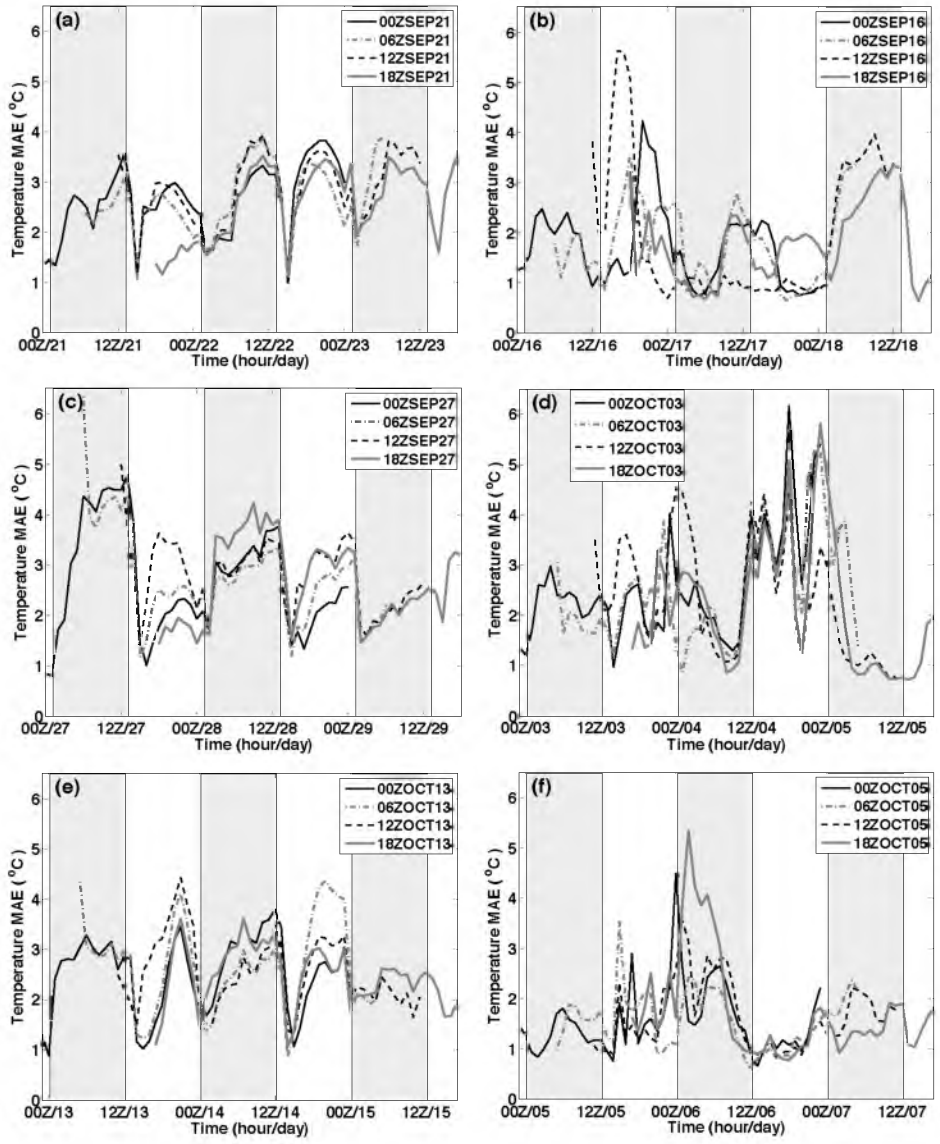


Figure 3.26 Mean absolute error of simulated 2-m temperature from the 1.11-km domain for various cases: (a) 0000 UTC 21 to 1800 UTC 23 September, (b) 0000 UTC 16 to 1800 UTC 18 September, (c) 0000 UTC 27 to 1800 UTC 29 September, (d) 0000 UTC 3 to 1800 UTC 5 October, (e) 0000 UTC 13 to 1800 UTC 15 October, and (f) 0000 UTC 5 to 1800 UTC 7 October. Four forecasts with different initial times are available for each case. The forecasting period for all forecasts is 48 h. Specifically, left panels (a, c, and e) represent weak synoptic forcing cases, and right panels (b, d, and f) denote strong synoptic forcing cases.

CHAPTER 4

ENSEMBLE KALMAN FILTER ASSIMILATION OF OBSERVATIONS DURING MATERHORN FIELD PROGRAM

4.1 Introduction

Weather forecasting in complex terrain is a challenge in NWP due to a number of difficulties, including sparse observations, terrain misrepresentation in numerical models, and model errors related to the complexity of surface conditions. Owing to these limitations, few previous studies have emphasized data assimilation in complex terrain. The most recent field experiments of the Mountain Terrain Atmospheric Modeling and Observations (MATERHORN, Fernando and Pardyjak 2013) Program collected comprehensive observations over mountainous regions, providing an opportunity to study the predictability of atmospheric conditions over complex terrain. Specifically, MATERHORN is designed to identify and study the limitations of current mesoscale models in predicting weather in mountainous terrain and to develop scientific tools to help realize improvements in prediction skill. During fall 2012 and spring 2013, comprehensive observations were collected of near-surface atmospheric conditions, profiling measurements from multiple platforms (e.g., tethersondes, lidar, radiosondes, etc.), soil states, and surface energy budgets over Dugway Proving Ground (DPG), Utah.

Inspired by the findings from Chapters 2 and 3, namely the capability of the EnKF in assimilating surface observations and the error characteristics associated with the near-surface variables in complex terrain, in this chapter, we attempt to investigate the ability of EnKF data assimilation in producing analyses and forecasts that represent atmospheric states in mountainous terrain and seek ways to mitigate the errors and eventually improve the model's prediction skill for the near-surface variables.

Sections 4.2 and 4.3 describe the observations and EnKF data assimilation configurations. Section 4.4 verifies the EnKF analyses and forecasts. Section 4.5 presents a case study for an intensive observation period (IOP) during MATERHORN fall 2012. Section 4.6 summarizes the results.

4.2 Observations

As mentioned in Chapter 3, DPG lies approximately 137 kilometers southwest of the Great Salt Lake, Utah. It is located in the Great Salt Lake Desert and characterized by complex terrain, featuring two major land surfaces of playa and sparsely vegetated desert. The latter one is characterized with sparse vegetation and little human influence. The land surface of the sparsely vegetated desert is known by low albedo, dry soil, and low evaporation, while the playa could be much moister (Malek 2003). Due to the distinguished properties of the two major land surfaces, stations were categorized into two groups—"Playa" stations over the playa land surface and "Sagebrush" stations over the sparsely vegetated desert land surface. Figure 4.1 shows the MATERHORN ground-based instrumentation system and their locations. Figure 4.2 shows the zoomed DPG map and surface stations. The "Playa" stations and "Sagebrush" stations are marked with

different colors. Considering the data availability, the DPG surface observations and radiosondes from the "Playa" and "Sagebrush" launching locations were assimilated into the EnKF data assimilation cycles.

The quality-controlled DPG surface observations come with quality flags. Only those observations with the highest quality flag are used for verification.

4.3 Data Assimilation Configuration

Following the methodology described in Chapters 2 and 3, DART/WRF ensemble data assimilation system and WRF Version 3.3 was employed for analysis update and forecast.

The model domain configuration was the same as that of the near real time experimental forecasts operated in fall 2011, shown in Figure 3.22. Four-level, one-way nested domains (D01, D02, D03, and D04) at 30-km, 10-km, 3.33-km and 1.11-km horizontal grid spacings were used. The innermost domain focused on the DPG area. The WRF model included 37 vertical levels; the model's top was set to 50 hPa. Physical parameterization configurations were also the same as those used in real-time experiments during fall 2011, referenced in Chapter 3.

Land-surface type and soil texture are attached at each model grid in the WRF Preprocess (WPS) based on the United States Geological Survey (USGS) land-use dataset (Skamarock et al. 2008). Parameters of the land-surface (e.g., albedo, roughness length, emissivity, and thermal inertia) and soil texture (e.g., porosity, field capacity, saturation soil suction, and hydrualic conductivity at saturation) are defined in lookup tables that were obtained from previous studies (e.g., Mahfouf et al. 1995; Peters-Lidard et al. 1998). In 2012, the land-use types in the WRF model were updated to 33 categories

(from previous 27-category), which include the type of Playa and other types. The soil texture was also increased to 19 classes from 16 classes, which includes the playa, white sand, and lava soil classes. These changes in land-surface and soil and their corresponding parameters in lookup tables make the land-surface and soil conditions represented more realistically in WRF model.

Based on the data availability, EnKF analysis was conducted for the fall 2012 MATERHORN field campaign, covering the period of September 21 to October 21, 2012. We consider the first several days a spin-up period as the first intensive observation period (IOP) started at 2000 UTC September 25.

A continuous, 3-hourly cycled EnKF data assimilation was performed for the whole period with 60 ensemble members. The initial perturbations were generated at 0000 UTC September 21, 2012 by adding ensemble perturbations to the deterministic initial condition using the fixed covariance perturbations (Torn et al. 2006). The data assimilation was conducted with the DART ensemble Kalman filter system, and the 3-h forecasts during the cycles were integrated with the WRF ARW model. The assimilated observations included the NCEP PREBUFR (NCEP Automated Data Processing Global Upper Air and Surface Weather Observations in PREBUFR format) conventional observations, the Mesonet surface observations and radiosoundes from MATERHORN field campaign. Figure 4.3 gives the schematic diagram of the cycled data assimilation procedure.

We used 60 ensemble members, many more than were used in previous studies, to relax the restriction of the localization that plays a more important role when the ensemble is smaller. Although with the larger the ensemble size, the less important the

localization is, we have to limit the ensemble size in order to keep the computational cost affordable. Considering the computational resources available, even using 60 ensemble members is very expensive! It took 45 continuous computing days to finish the 1-month cycles.

A covariance inflation that varies temporally and spatially was used in this data assimilation system (DART/WRF) to avoid filter divergence and to reduce the impact of model error (Anderson 2007).

4.4 Results and Verification

With the outcomes of a month-long analysis and forecast, we first evaluated their quality by comparing them with available observations and NCEP NAM forecast.

4.4.1 General Statistics

We calculated various statistical metrics to characterize the error of near surface variables. For instance, root mean square errors (RMSEs) and bias errors (BEs) were calculated against surface observations to evaluate the overall performance of the EnKF/WRF analyses and forecasts.

At each analysis time, RMSEs and BEs were calculated against surface observations over all DPG stations for both the analyses and forecasts. Figures 4.4–4.6 show the RMSEs of the analyses and 3-h forecasts (background field for data assimilation) and the differences of the MAEs between the analyses and forecasts with 90% confidence intervals during the entire period for 10-m wind speed and direction and 2-m temperature. Overall, the EnKF analyses fit observations better than the short-range forecasts,

indicating the improvements to atmospheric states from the data assimilation, especially for the wind fields. During the first few days (the spin-up period), the forecasts differed from the analyses with larger errors. Then, as more observations were assimilated, forecast errors tended to be stable and closer to analysis errors, indicating that the short-range (3-h) forecasts behave similarly to the analyses in a continuously cycled data assimilation period due to the influence from the observations.

4.4.2 Diurnal Variation

Figures 4.7–4.9 examine the diurnal error characteristics of the 10-m wind speed and direction and 2-m temperature during the entire month. The maximum RMSE of wind speed (Figure 4.7) was about 2 m/s for EnKF/WRF forecasts and smaller than 1.8 m/s for EnKF/WRF analyses. The biases of wind speed forecasts were positive between 0600 UTC and 1800 UTC and negative between 1800 UTC and 0600 UTC in the forecasts. For the analysis, all the times show negative biases except for the early morning time (1200 UTC and 1500 UTC).

The RMSEs of 10-m wind direction peaked at the two transitional times (Figure 4.8a), 0300 UTC (early evening) and 1500 UTC (early morning). The BEs were mostly positive and negligibly small and occured during nighttime when the atmospheric boundary layer was stable.

The RMSEs of 2-m temperature (Figure 4.9) show strong diurnal features and peak twice per day, one in the middle of the night and the other one in the afternoon. The peaks of RMSE correspond to a warm bias maximum in the middle of the night and a cold bias maximum in the afternoon. There are also two bias error minima per day at

nearly 0000 UTC and 1200 UTC, respectively. The minima could be caused by more observations being available at 0000 UTC and 1200 UTC in the data assimilation cycle. However, this is just a part of the reason since similar error characteristics were also shown in Chapter 3 in the cases without data assimilation. Overall, warm biases are present during the nighttime and cold biases are present during daytime. The diurnal errors are independent of forecast lead time and initialization time.

Since the near surface atmospheric conditions vary with underlying land surfaces, RMSEs and BEs are also compared between station types (Playa and Sagebrush). Figures 4.10–4.12 shows RMESs and BEs for different station types. While the statements made for all stations in Figures 4.7–4.9 are still valid, the error characteristics over Playa and Sagebrush stations are different in their magnitudes. The biases in the Playa stations are the main contributor to the diurnal pattern of the 10-m wind speed and direction, and the Sagebrush stations show lower wind speed and direction biases at all times. These results imply that the surface winds in Sagebrush stations are more predictable than those in the Playa stations. The Playa stations show overall negative wind speed biases and positive (clockwise) wind direction biases. In contrast, the Playa stations have smaller RMSEs and BEs of 2-m temperature than the Sagebrush stations.

Overall, the comparison of the diurnal variations of the errors in analyses and forecasts indicates that the data assimilation can reduce the errors and biases of those surface variables to some degrees. However, it cannot eliminate the diurnal errors and biases.

4.4.3 Comparison with NAM Forecasts

In order to further examine the error characteristics in the month-long EnKF/WRF analyses and forecasts, the EnKF/WRF 3-h forecasts were compared with NAM 3-h forecasts, shown in Figures 4.13–4.15. Since the initialization times for NAM forecasts are at 0000 UTC, 0600 UTC, 1200 UTC and 1800 UTC, the corresponding times for the NAM 3-h forecasts are at 0300 UTC, 0900 UTC, 1500 UTC and 1200 UTC. *Overall, the errors of the prior field inside of analysis cycle (3-h forecasts, hereafter) and the NAM 3-h forecasts are very comparable. Most of the RMSEs of the wind speed and direction forecasts from EnKF are smaller than those of the NAM forecasts.*

The forecasts of 2-m temperature from the EnKF/WRF are a little worse than the NAM forecasts in some cases. This result is not surprising because NAM is a well-tuned, operational system that treats the land surface and soil initialization in a better way. The soil moisture and temperature fields in the 4-layer (Noah) land-surface model (LSM) used by the operational NAM/NDAS (NAM Data Assimilation System) are continuously cycled without soil moisture nudging (e.g., to climatology). The fields used in the NAM are the sole product of model physics and internal NDAS surface forcing (e.g. precipitation and surface radiation). During the forecast portion of the NDAS, the model predicted precipitation, which would normally be used as the forcing to the Noah scheme, is replaced by the hourly merged Stage II/IV precipitation analyses (http://www.emc.ncep.noaa.gov/mmb/ylin/pcpanl/).

It should be pointed out that the boundary conditions used in the EnKF data assimilation cycles were NAM forecasts instead of NAM analyses. The boundary conditions, especially the lower boundary conditions, play a very important role in

surface forecasts. Although the NAM 3-h forecasts had smaller BEs and RMSEs in some cases for 2-m temperature, the systematic daytime cold biases and nighttime warm biases are also present.

Overall, while the surface temperature forecasts are strongly influenced by diurnal signal and substantially related to the lower boundary conditions and soil state, current EnKF/WRF analyses did not outperform the NAM in terms of near surface temperature forecasts. However, since the surface wind forecasts benefit from the high resolution WRF forecasts/analyses cycles, the EnKF/WRF outperforms the NAM in near surface wind forecasts.

4.5 Case Study

The third intensive observation period (IOP-3) during the MATERHORN fall 2012 field campaign stared at 0200 MDT (0800 UTC) 3 October and ended at 0200 MDT (0800 UTC) 4 October, lasting for a 24-h period. The IOP-3 was a daytime transitional case. A dry cold front entered into Utah in the local morning of 4 October. We chose this IOP for a case study because it well represented the multiscale weather patterns and their interactions. The ensemble forecast from the EnKF analysis was expected to help capture the multiscale transitional weather patterns along the frontal passage.

Figure 4.16 shows the NOAA surface weather maps of the frontal system evolution during the IOP-3. The satellite surface images indicate that the front entered into Utah about 2100 UTC 3 October, several hours earlier than the time documented by the MATERHORN team who claimed the front passed the Playa site around 0925 MDT (Mountain Daytime Time) on 4 October. Figure 4.17 shows the NOAA 700 hPa synoptic

analyses of temperature and geopotential heights. It shows that the front was located north of Utah on 1200 UTC 3 October, and entered into Utah early on 4 October. Note the shortwave trough right on the border of Utah and Idaho on 0000 UTC 4 October was the system that influenced Utah's weather in the following hours.

Figure 4.18 shows the simulated 700 hPa synoptic evolutions during the IOP-3 period from the ensemble mean and the corresponding WRF deterministic forecasts (initialized by NAM analysis). Both the ensemble forecast and the deterministic forecast predicted the frontal passage on 700 hPa, but the ensemble forecast outperformed the deterministic forecast with more accurate timing and location of the frontal passage. Specifically, compared with the analysis (Figure 4.18), the deterministic forecast predicts earlier entrance of the front. The ensemble forecasts also capture more detailed synoptic structures, such as the shortwave trough on the border of Utah and Idaho as well as the location of the main trough.

In addition, the ensemble forecasts also provide uncertainties of the weather system with its ensemble spread. Figure 4.19 illustrates the spaghetti plots of the 0 ^{o}C temperature isotherm and 312 dm geopotential height isoline on 700 hPa from all ensemble members. Figure 4.19a reveals how the analysis differed from the background and especially how the analysis was pulled away from the background field due to the assimilation of observations. Figure 4.19b shows an example that reduces analysis uncertainties by assimilating observations into the background field. Before the frontal system entered Utah, the ensemble members predicted the upstream geopotential height with large spread at 0000 UTC 4 October, indicating high uncertainties of this front

system. After the data assimilation, the ensemble spread of geopotential height decreased with reduced uncertianties.

The frontal passage is demonstrated in Figure 4.20 by the vertical north-south cross-section through the station DPG24 located in central playa. At 1800 UTC October 3 (Figure 4.20a), northwest flows prevail in the lower boundary layer in DPG under the influence of the upper level northwesterly flow. Wind speeds are between 4–8 m/s below 800-hPa and about 10 m/s on 700-hPa. Wind directions change to the northeast at about 0000 UTC 4 October (Figure 4.20c) below 750 hPa while northwesterly flows are still maintained above 750 hPa. This reveals that the front was passing the DPG region around this time. The 0300 UTC October 4 panel (Figure 4.20d) clearly shows that the strong northeasterly flows dominate the atmosphere below 780-hPa, indicating a shallow layer of cold air intruding in the area after the frontal passage. We also noted that the wind speeds in the layer between the upper level and the front layer are very small, around 2–4 m/s.

Figures 4.21 and 4.22 compare skew-T plots between the EnKF/WRF 3-h ensemble forecasts and radiosonde observations at the two radiosonde-launching stations SLTEST (Surface Layer Turbulence and Environmental Science Test, Klewicki 2008) and Sagebrush (as shown in Figure 4.1). Figures 4.23 and 4.24 show the comparison between the deterministic forecast (initialized by NAM) and the radiosonde observations. The ensemble forecast obviously fit the observations much better than the deterministic forecast does. For the rest of the diagnoses, we focus on the ensemble results.

At the SLTEST station, the temperature profiles from the ensemble mean agreed with the observations well in the first night (Figures 4.21a and b). The nocturnal inversions

were accurately captured. During the daytime (Figures 4.21c and d), the ensemble temperature profiles generally matched the observations, although the observations showed more detailed structure in the lower boundary layer. The boundary layer was well mixed during the local noontime as shown in both the observations and the ensemble mean in Figure 4.21e. The ensemble mean showed some disagreement in the lower boundary layer in the second night (Figure 4.21f) as the ensemble forecast of the strong nocturnal inversion is later than the observation. Similar results are found at the Sagebrush station (Figure 4.23). The ensemble forecast is in good agreement with the observations during the daytime (Figures 4.23a, b, and c) but disagree with the observations during the local nighttime (Figure 4.23d) of October 3. Large discrepancies are mainly found in the moisture field in the lower boundary layer.

4.6 Summary

This chapter investigates the ability of EnKF data assimilation to represent the atmospheric states over mountainous terrain by conducting cycled data assimilation during the MATERHORN fall 2012 field period. Three-hourly cycled EnKF data assimilation is performed with 60 ensemble members during a 1-month period. Month-long EnKF analyses and forecasts are obtained for general statistical analysis and case studies. As expected, the EnKF analyses improve the overall representation of near surface temperature and wind as the errors and biases are significantly reduced from the short-range forecasts.

The RMSEs of the 2-m temperature have two diurnal peaks with one in the middle of the night and the other one in the afternoon. The two RMSE peaks correspond to a warm

bias maximum in the middle of the night and a cold bias maximum in the afternoon. There are also two diurnal bias error minima at nearly 0000 UTC and 1200 UTC, respectively. These results are consistent with those obtained from the real-time run operated in fall 2011. Overall, warm biases are present during the nighttime and cold biases are present during the daytime.

The RMSEs of the 10-m wind speed are mostly smaller than 2 m/s except for the spin-up period. The RMSEs of the 10-m wind direction peak at the two transitional times, 0300 UTC (early evening) and 1500 UTC (early morning). The BEs are negligibly small and occur during the nighttime when the atmospheric boundary layer is stable.

Overall, the EnKF/WRF analysis reduced the magnitudes of the errors in near-surface variables but it cannot overcome the systematic errors in diurnal variations. The forecast errors in this study are strongly related to the soil states (e.g., soil temperature and moisture) and parameterizations in the land surface model. Future work should address the issues related to the land surface model and soil.

Compared with NAM, the 3-h forecasts from the EnKF/WRF analyses in this study have overall smaller RMSEs than the NAM 3-h forecasts for both the 10-m wind speed and direction. It is obvious that the surface wind speed and direction forecasts benefit from the high resolution WRF model. However, comparison of 2-m temperatures shows that NAM performs better in some cases. This is mainly because the near surface temperature forecasts are strongly influenced by the diurnal signal and surface and soil states.

In addition, a transitional IOP (IOP-3) was selected for a case study. The comparison between forecasts from the EnKF/WRF analyses and observations was conducted to

investigate the performance of the EnKF simulations. The ensemble forecasts from EnKF/WRF analyses outperform the deterministic forecasts initialized by NAM and successfully capture the large-scale synoptic weather patterns during the frontal passage. Simulated temperature profiles agree with the radiosonde observations, although disagreements still occur in the lower boundary layer.

More comprehensive diagnoses and comparison will be continued in future work.

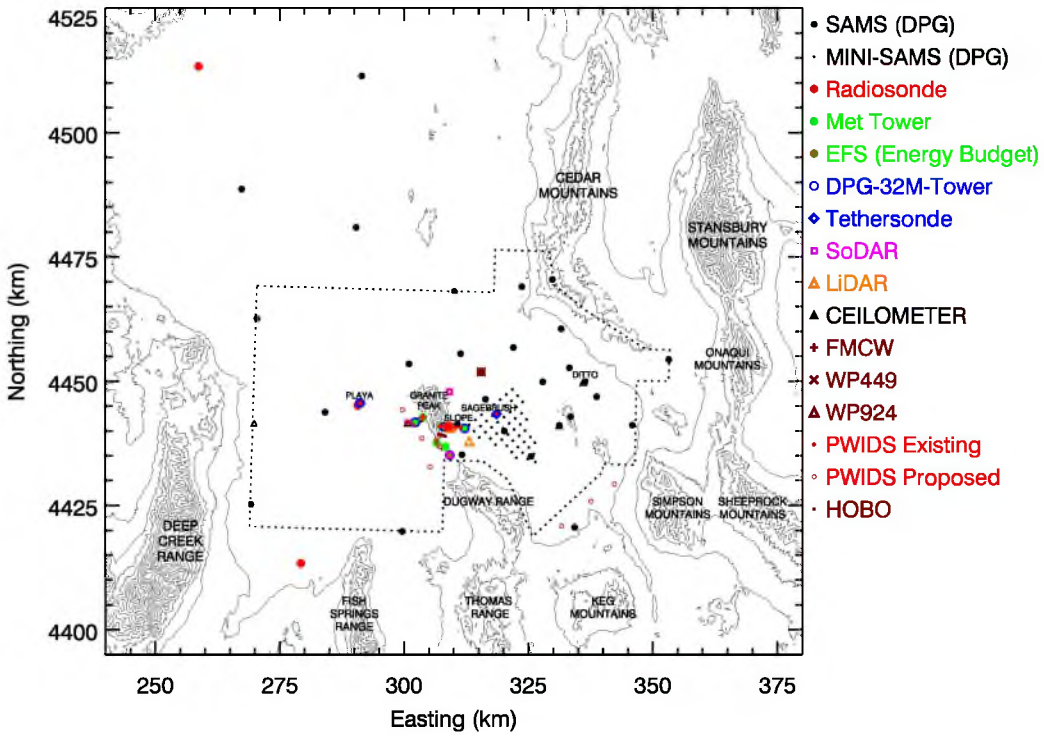


Figure 4.1 Ground-based instrumentations and their locations during the MATERHORN fall 2012 experiment. Red dots at the Playa and Sagebrush sites denote radiosonde launching locations. Black dots denote DPG surface stations. The dotted line indicates approximate boundaries of DPG. Contours are elevation with a 200 m interval (courtesy of MATERHORN field program plan by Pardyjak et al. 2012).

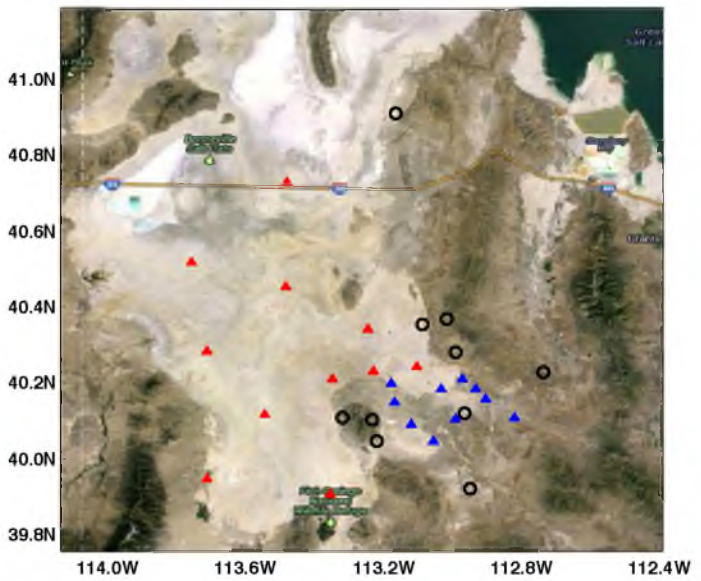


Figure 4.2 Map of Dugway Proving Ground region and surface stations. Red triangles represent "Playa" stations and blue triangles represent "Sagebrush" stations. Black circles denote those stations that do not count for either "Playa" or "Sagebrush."

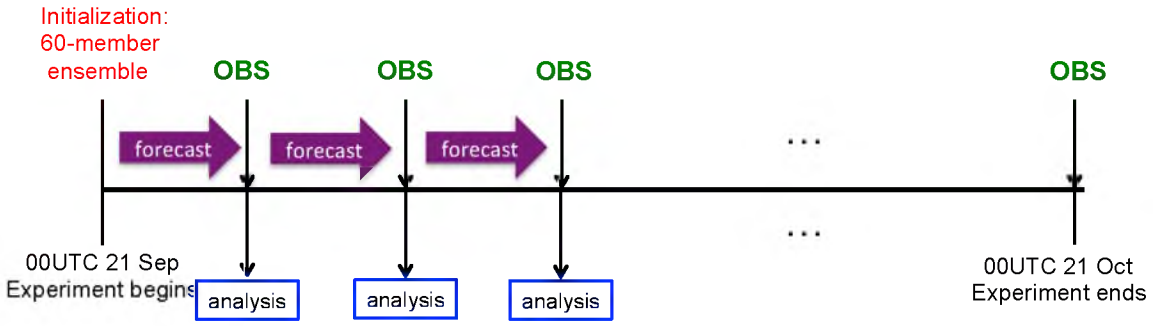


Figure 4.3 Schematic diagram of the analysis (data assimilation) and forecast cycles.

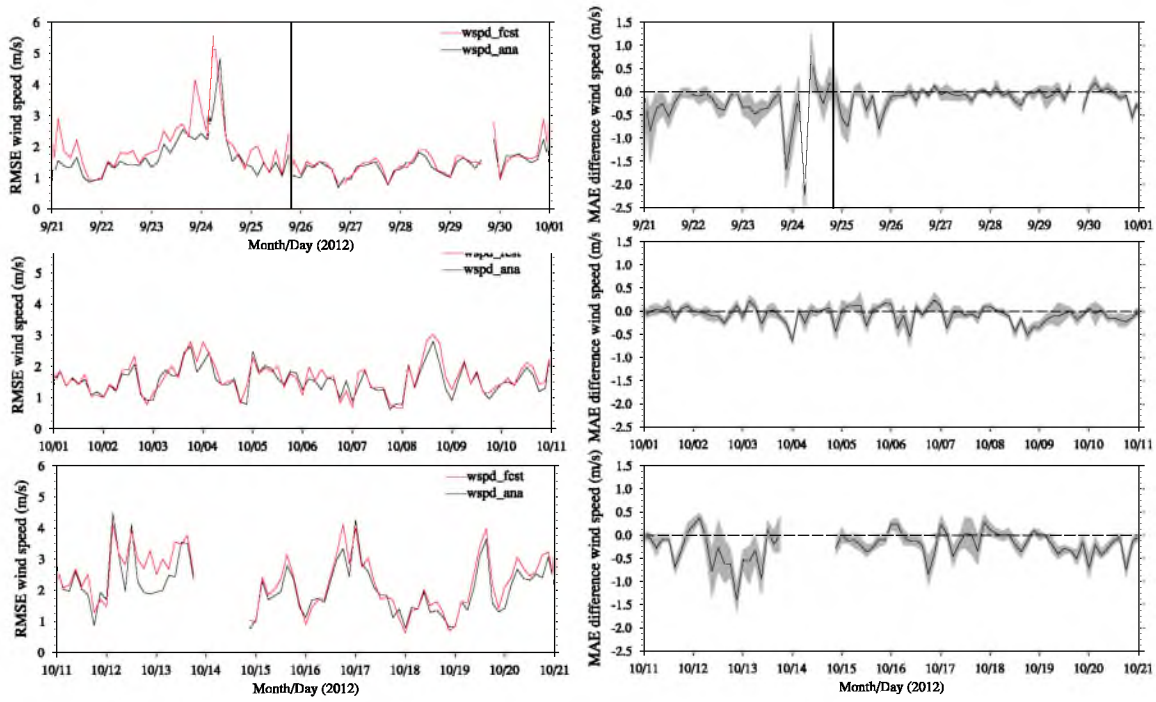


Figure 4.4 Time series of (a) RMSEs of 10-m wind speed against surface observations averaged over all stations in DPG for EnKF/WRF analyses and 3-h forecasts and (b) differences of MAEs between analyses and forecasts (analyses minus forecasts). Negative MAE differences mean that the analyses fit the observations better than the forecasts do. Shaded area represents 90% confidence intervals derived using the BCa bootstrapping technique. Statistics are from the innermost domain (1.11-km). The period prior to the black vertical line is the spin-up period.

Figure 4.5 Same as Figure 4.4, but for 10-m wind direction.

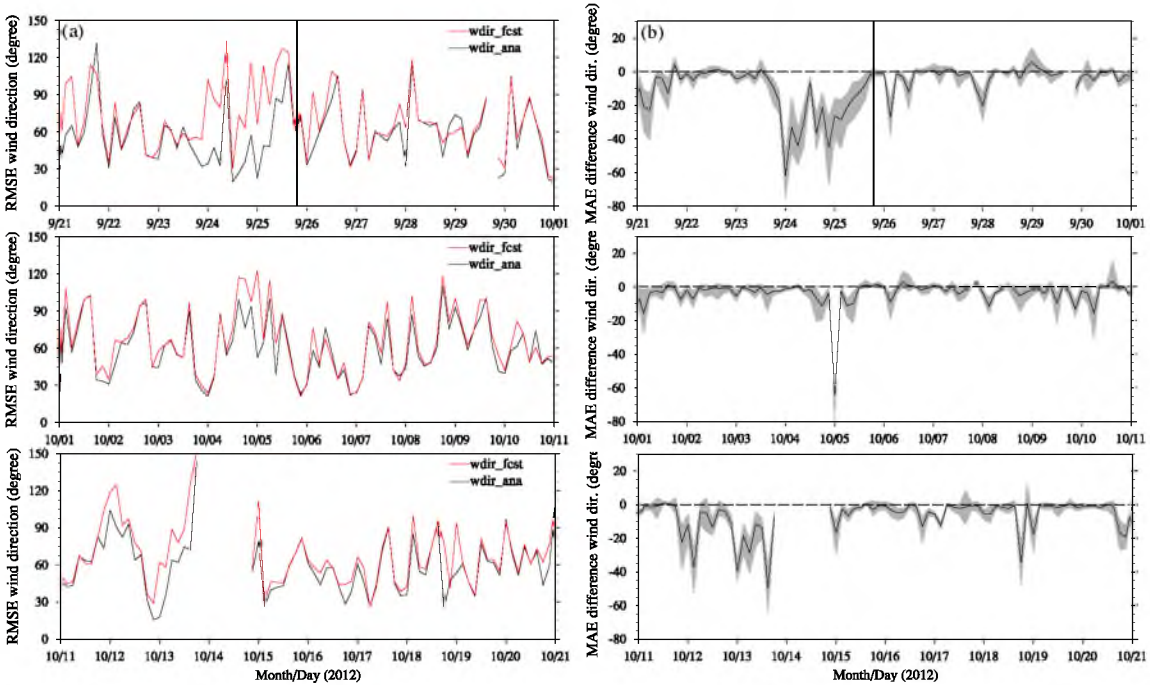
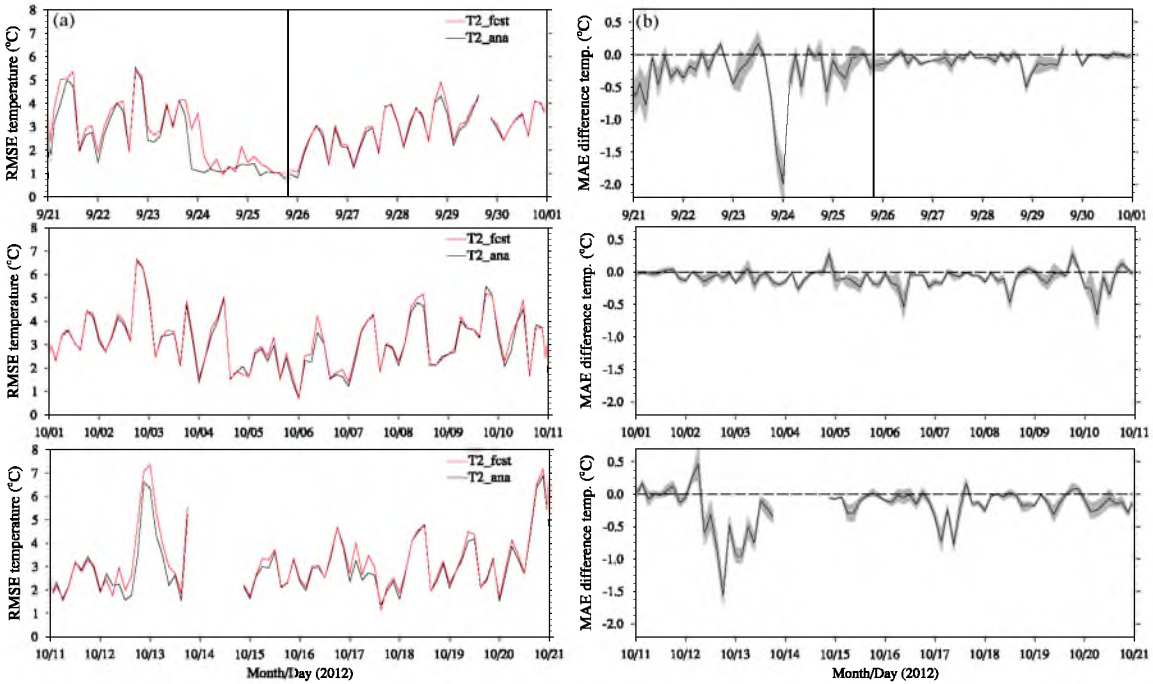

(a)
(b)
wdir_fcst
wdir_ana
RMSE wind direction (degree)
MAE difference wind dir. (degree)
Month/Day (2012)
Month/Day (2012)

Figure 4.6 Same as Figure 4.4, but for 2-m temperature.


(a)
(b)
T2_fcst
T2_ana
RMSE temperature (°C)
MAE difference temp. (°C)
Month/Day (2012)
9/21
9/22
9/23
9/24
9/25
9/26
9/27
9/28
9/29
9/30
10/01
10/02
10/03
10/04
10/05
10/06
10/07
10/08
10/09
10/10
10/11
10/12
10/13
10/14
10/15
10/16
10/17
10/18
10/19
10/20
10/21

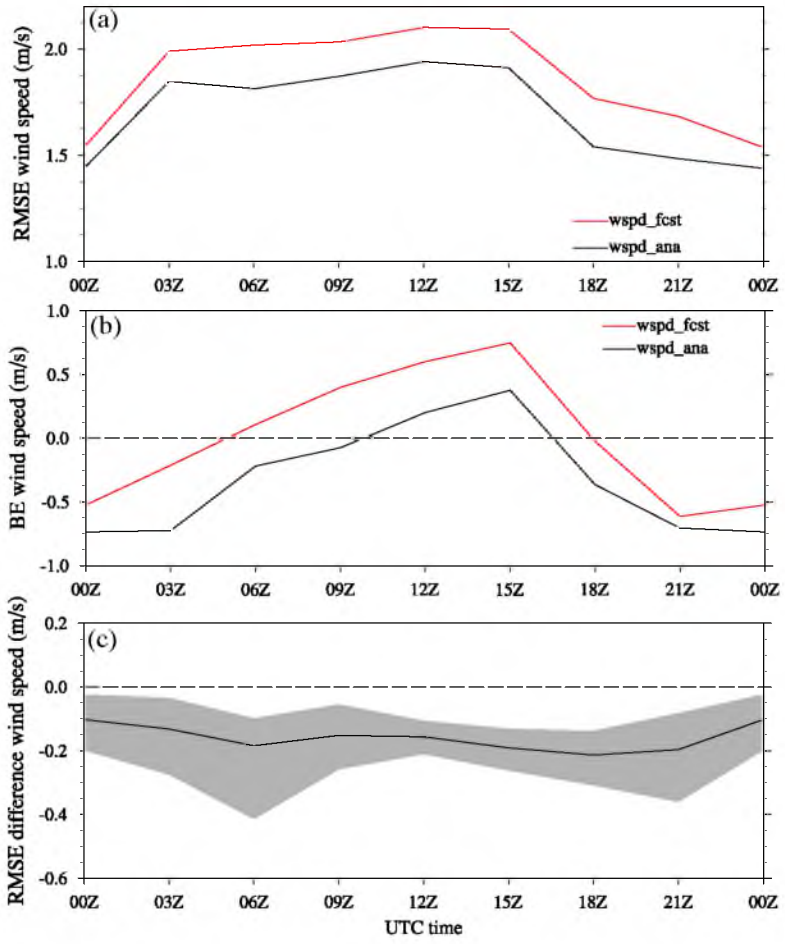


Figure 4.7 Diurnal variation of the (a) RMSEs and (b) BEs of 10-m wind speed against surface observations averaged over all DGP stations and the entire analysis period for EnKF/WRF analyses and 3-h forecasts, and (c) differences of RMSEs between analyses and forecasts (analyses minus forecasts). Negative RMSE differences mean that the analyses fit the observations better than the forecasts do. The shaded area represents 90% confidence intervals derived using the BCa bootstrapping technique.

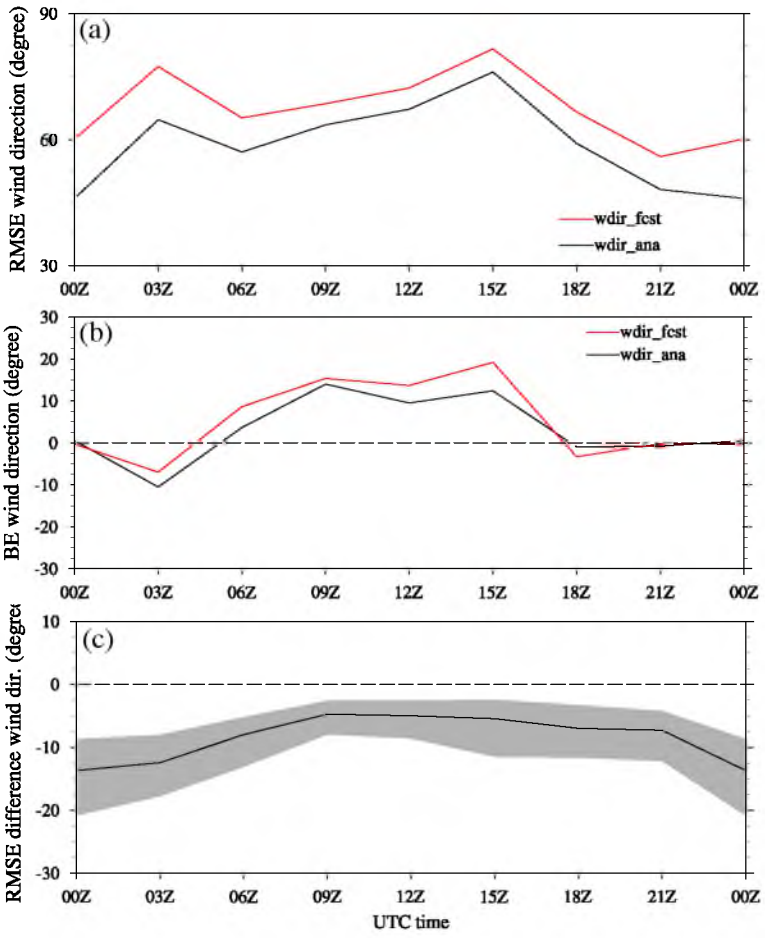


Figure 4.8 Same as Figure 4.7, but for 10-m wind direetion.

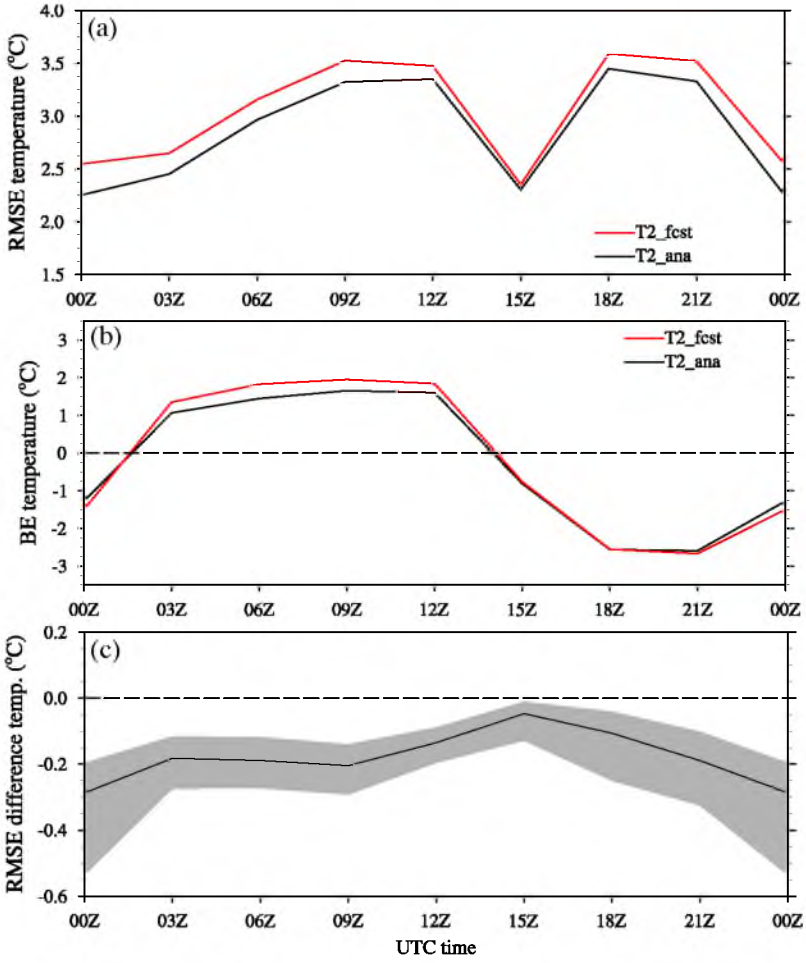


Figure 4. 9 Same as Figure 4.7, but for 2-m temperature.

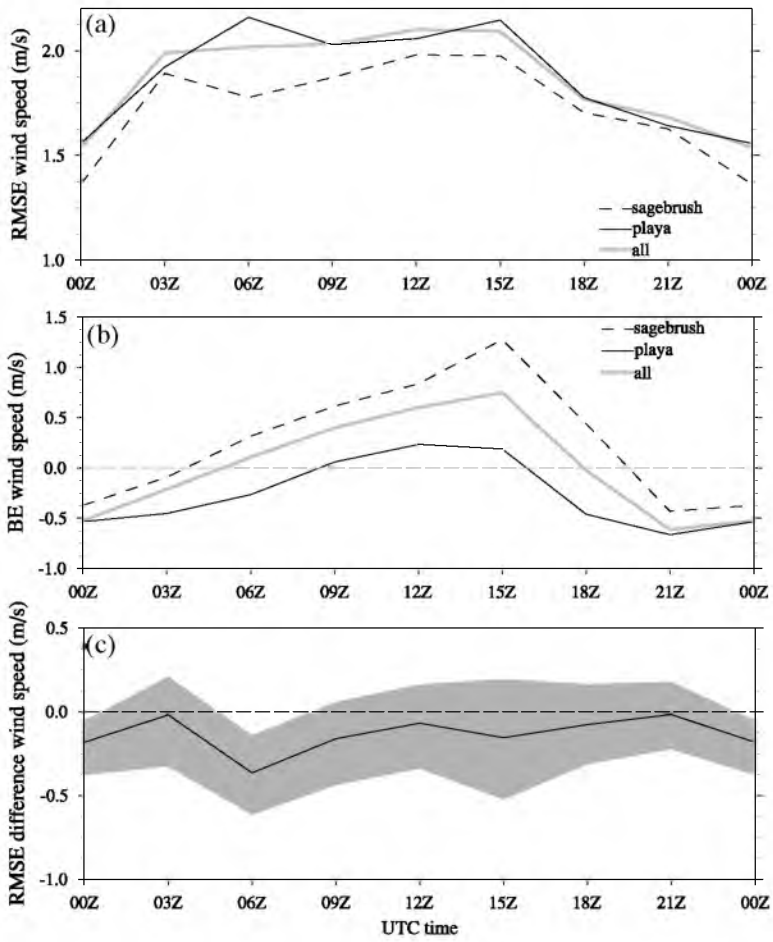
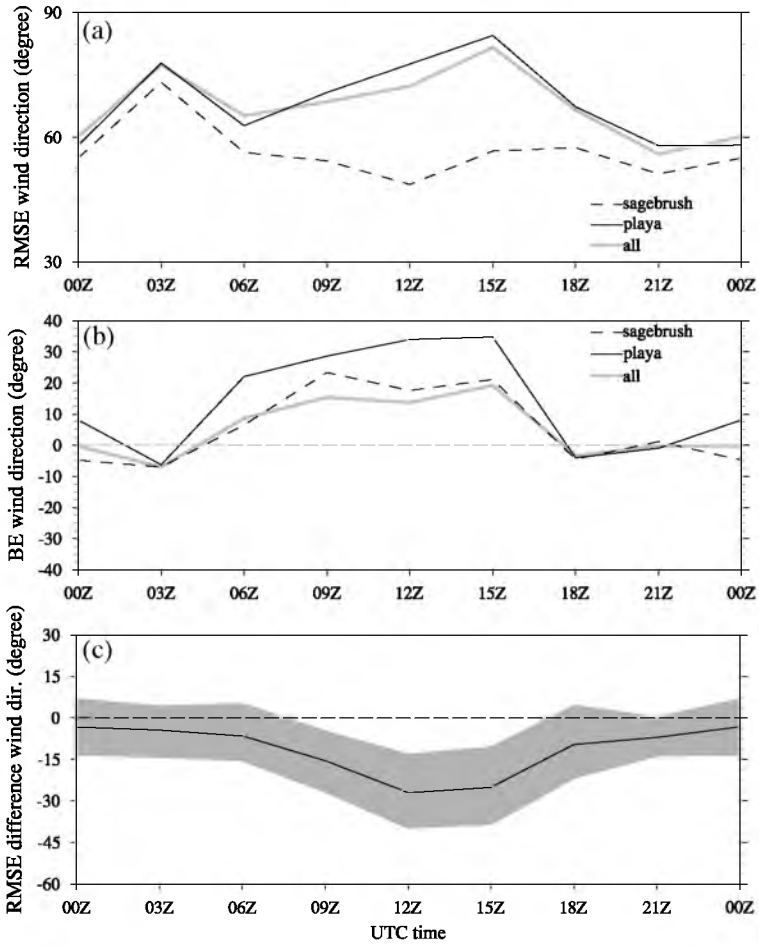


Figure 4.10 Comparison of diurnal variations of the RMSEs and BEs of 10-m wind speed forecasts against DGP surface observations averaged over "Sagebrush", "Playa" and all stations in DPG during whole analysis period. The shaded area represents 90% confidence intervals derived using the BCa bootstrapping technique.

Figure 4.11 Same as Figure 4.10, but for 10-m wind direction.


(a)
RMSE wind direction (degree)
90
60
30
sagebrush
playa
all
00Z
03Z
06Z
09Z
12Z
15Z
18Z
21Z
00Z
(b)
BE wind direction (degree)
40
30
20
10
0
-10
-20
-30
-40
sagebrush
playa
all
(c)
RMSE difference wind dir. (degree)
30
15
0
-15
-30
-45
-60
UTC time

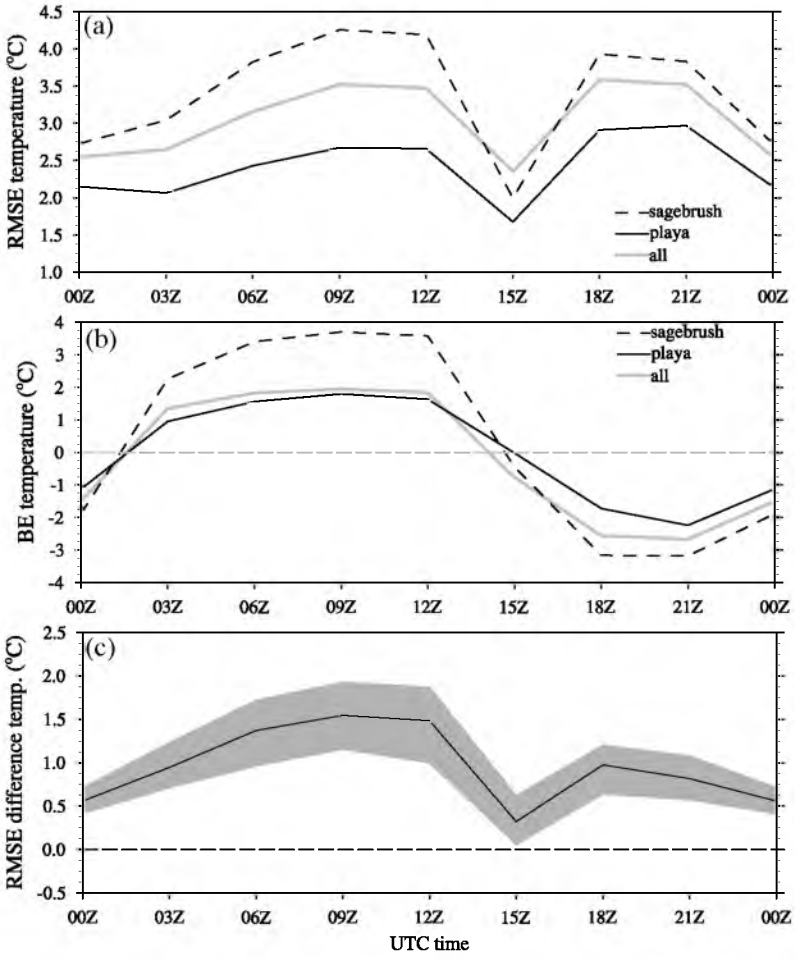


Figure 4.12 Same as Figure 4.10, but for 2-m temperature.

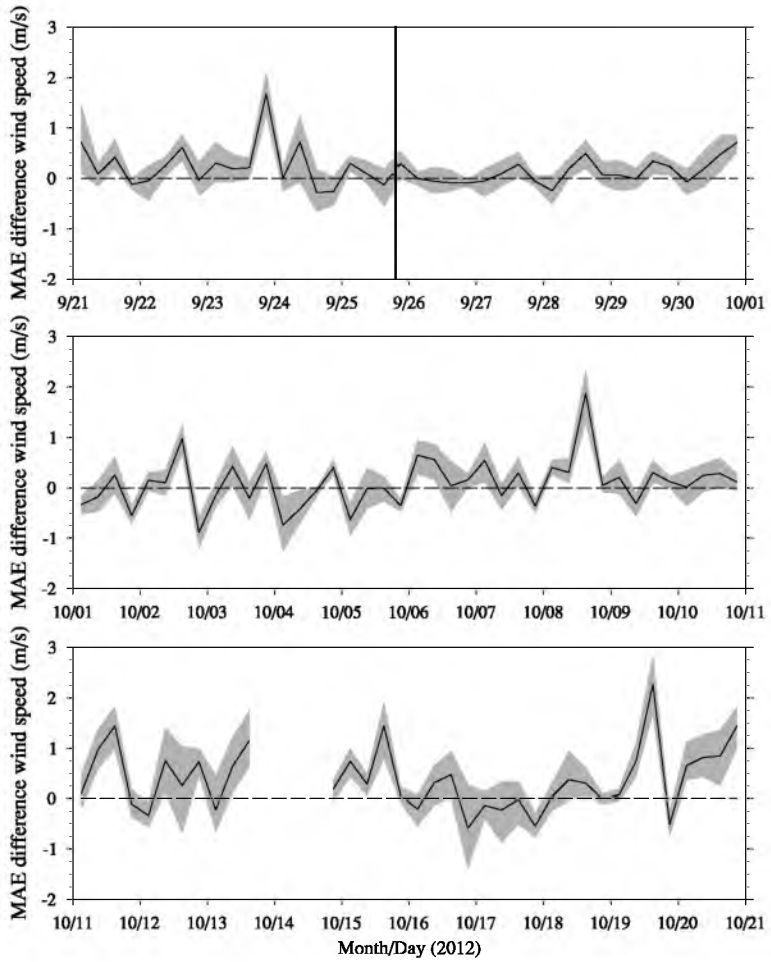


Figure 4.13 Difference of RMSEs of 10-m wind speed between EnKF/WRF 3-h forecasts initialized by EnKF/WRF analysis and NAM 3-h forecasts against surface observations averaged over all DGP stations. The period prior to the black vertical line is the spin-up period.

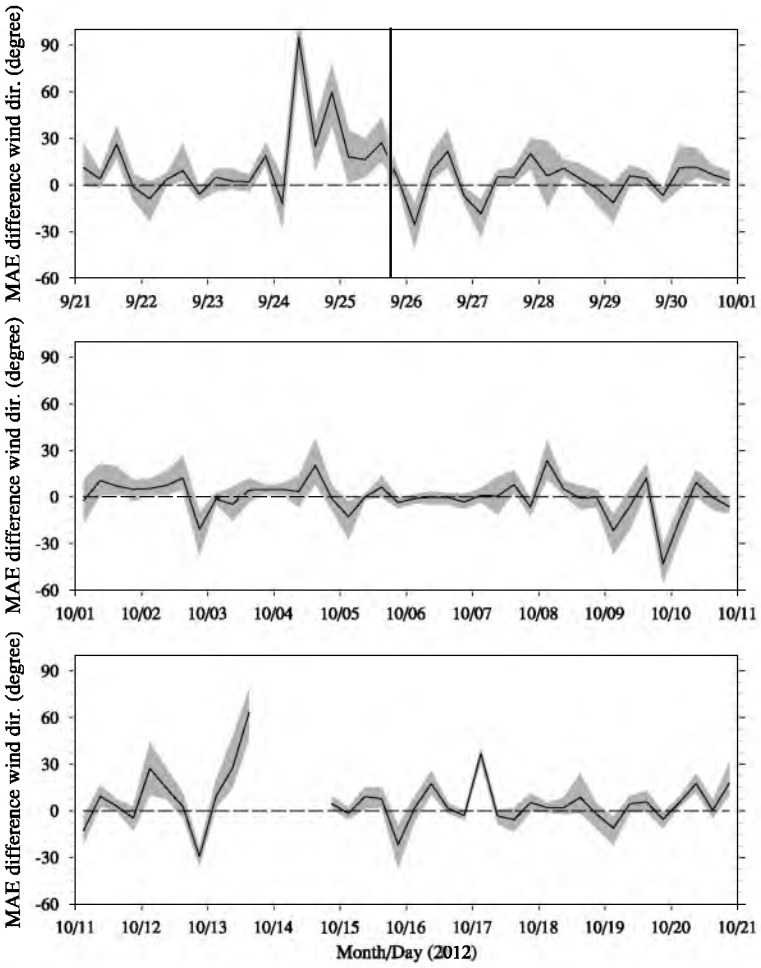


Figure 4.14 Same as Figure 4.13, but for 10-m wind direction.

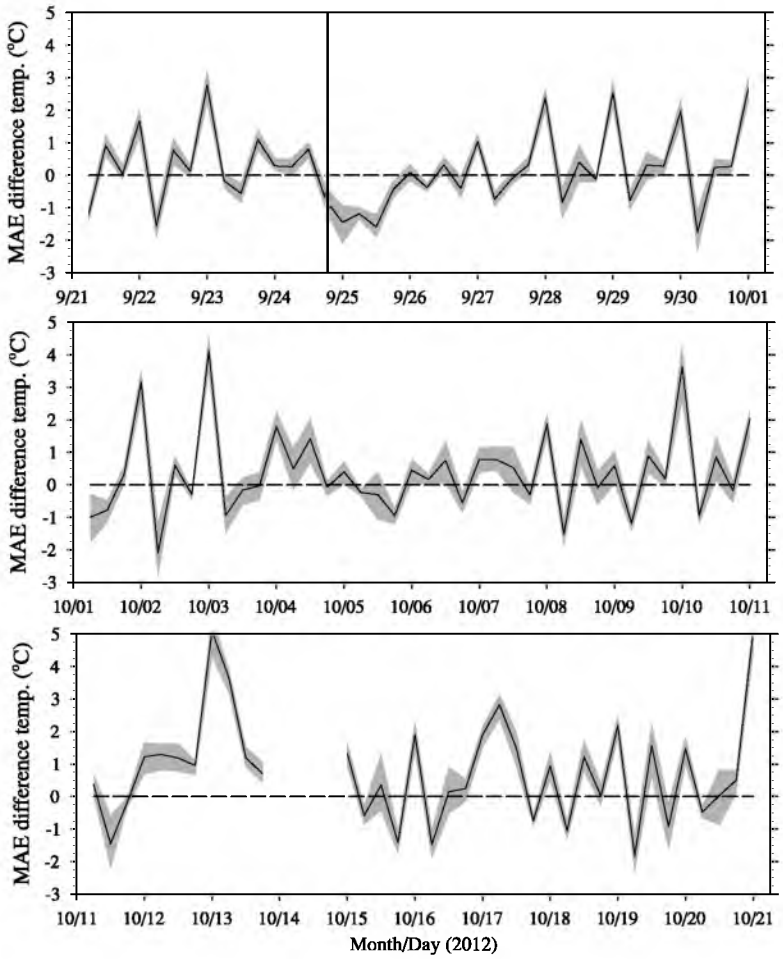


Figure 4.15 Same as Figure 4.13, but for 2-m temperature.

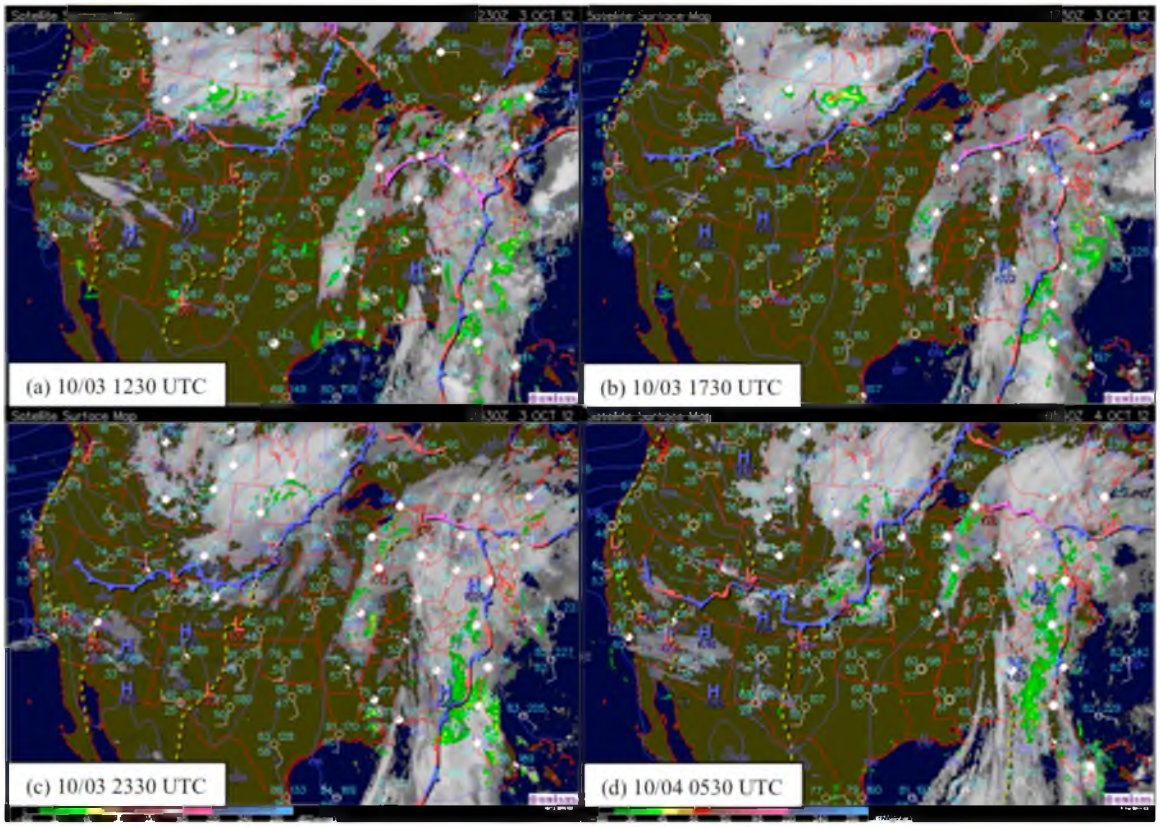


Figure 4.16 Surface weather maps (about every 6-h) during IOP-3 of fall 2012 MATERHORN. (Images adapted from NCAR image archive at http://locust.mmm.ucar.edu/)

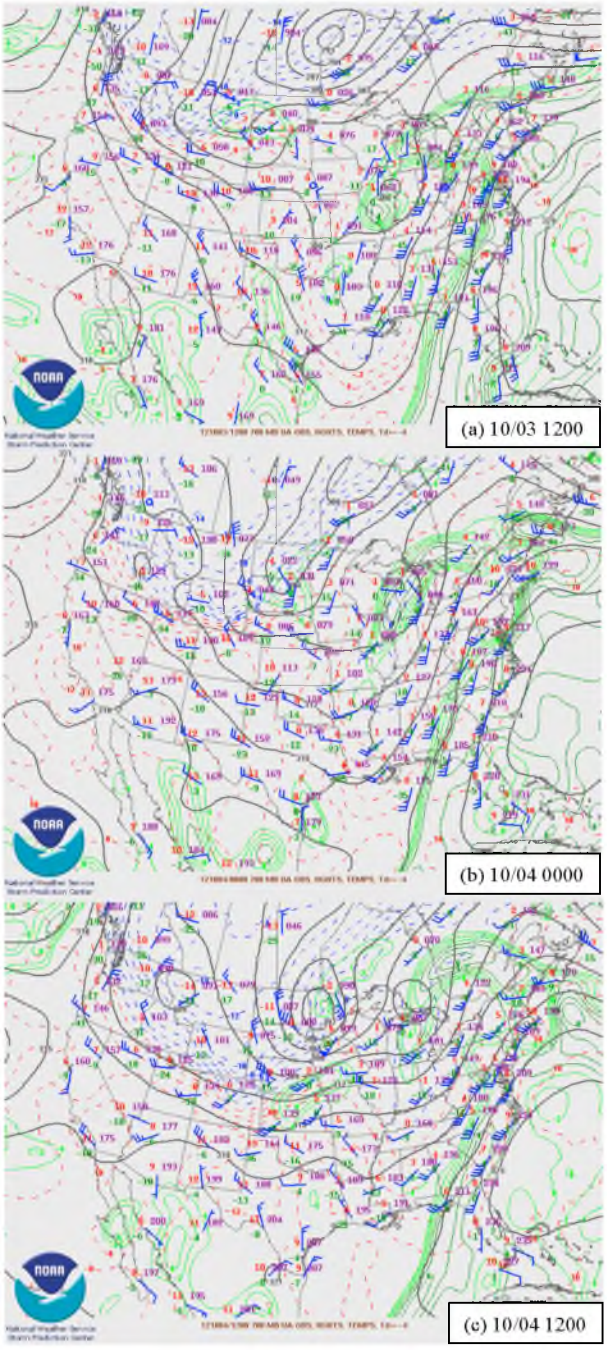


Figure 4.17 700 hPa synoptic analysis (every 12-h) during IOP-3 of fall 2012 MATERHORN. The solid black curves are the geopotential heights. The dashed lines are the temperatures. (Images adapted NCAR image archive at http://locust.mmm.ucar.edu)

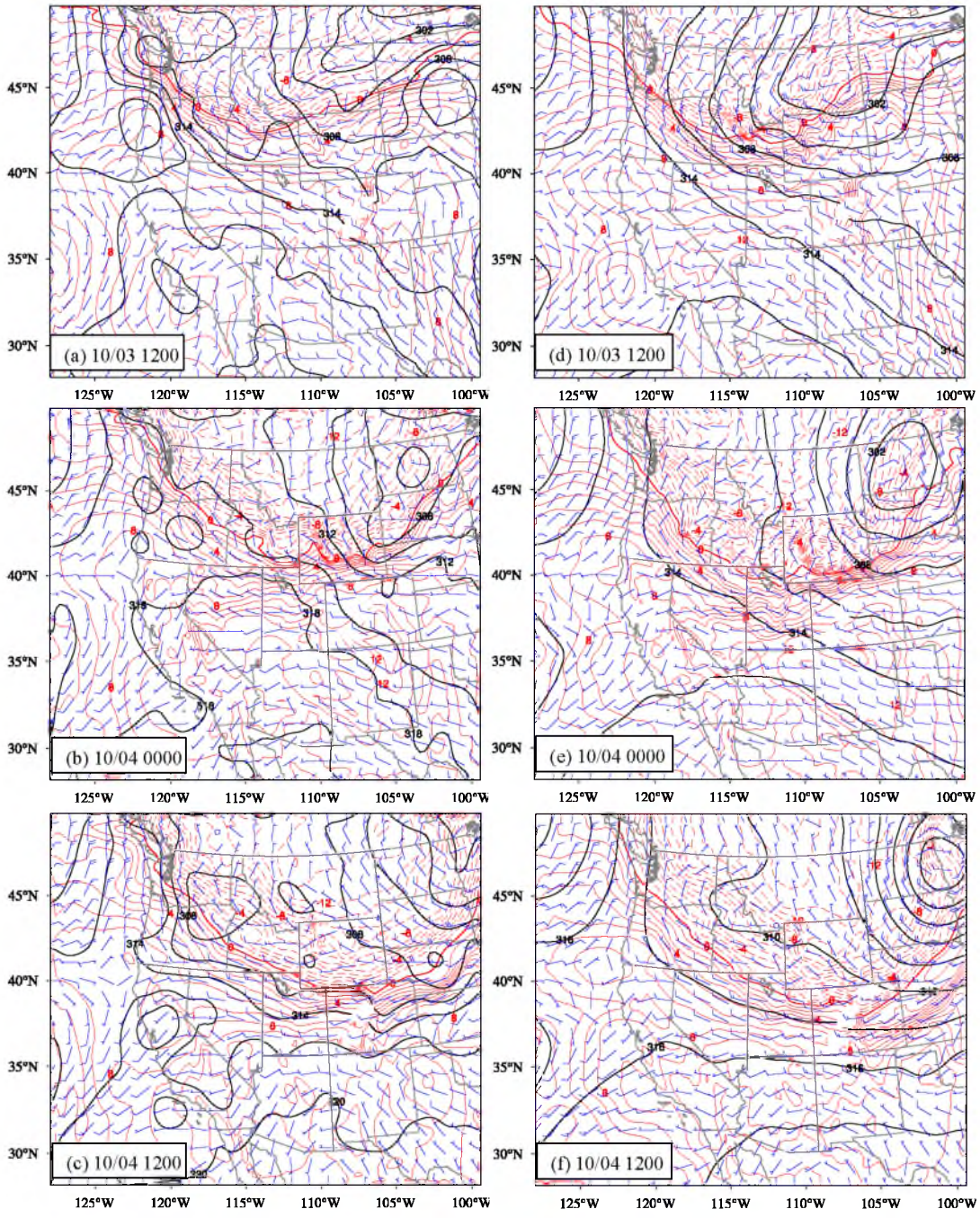


Figure 4.18 700 hPa synoptic maps, showing the evolution of a front system during IOP-3 of the fall 2012 MATERHORN. Black contours represent geopotential height (unit: dm). Red contours are temperature (Unit: ^{o}C). The top panels (a, b, and c) are the results of ensemble mean from the EnKf/WRF ensemble forecast; the bottom panels (d, e, and f) are the results of deterministic simulation initialized at 0000 UTC October 3 using NAM.

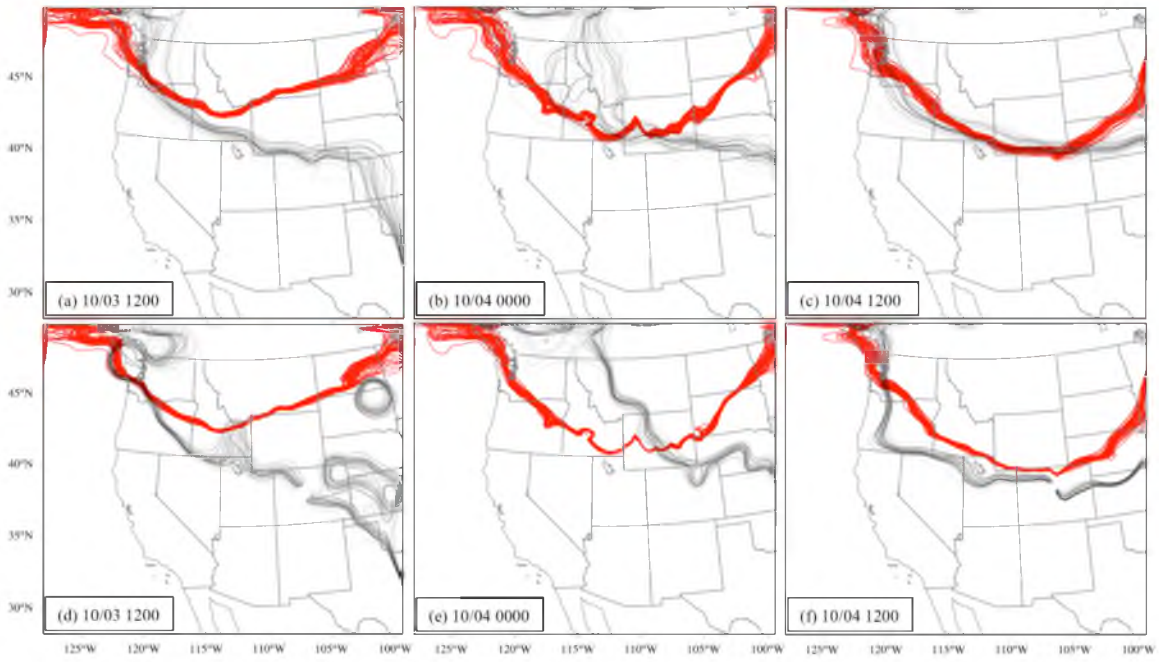


Figure 4.19 The EnKF/WRF ensemble forecast of the 0 °C temperature isotherm and 312 dm geopotential height isoliner on 700mb. a, b, and c are result of the ensemble forecast and d, e, and f are result of the ensemble analysis.

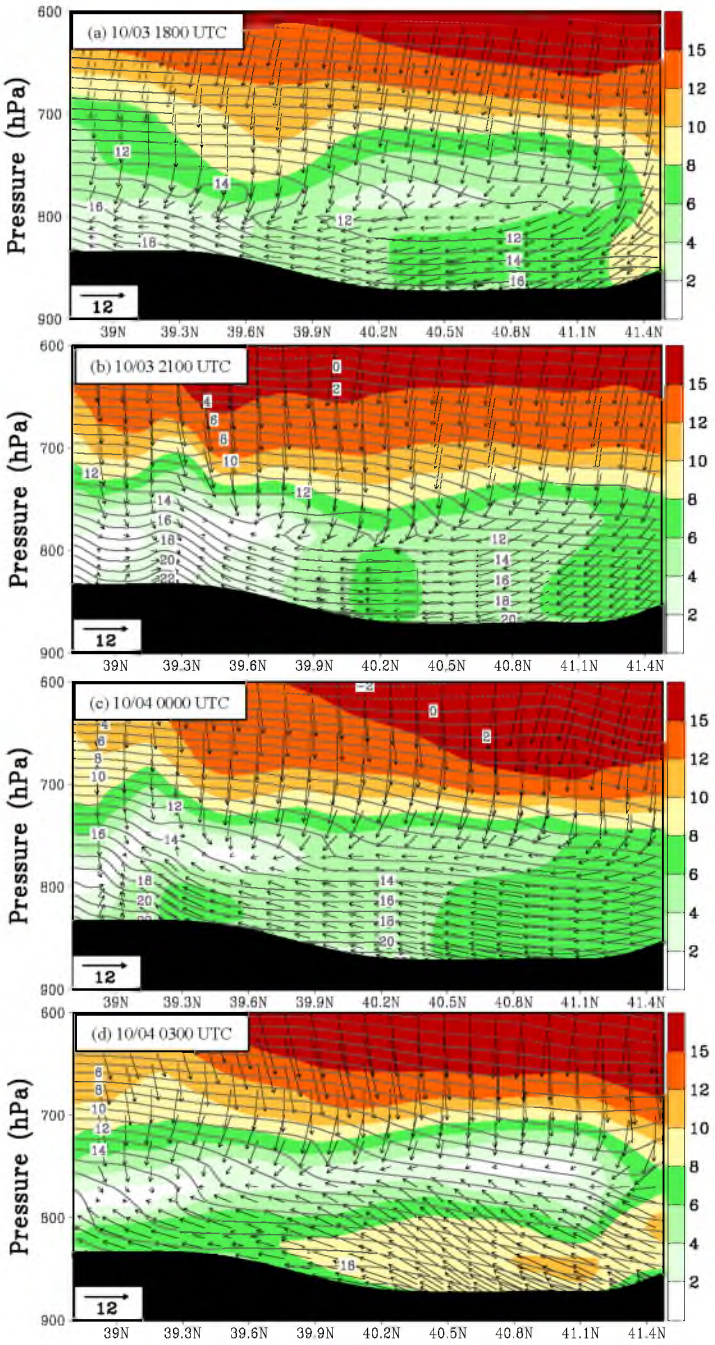


Figure 4.20 Vertical north-south cross-section of temperature (gray contours), wind speed (shaded coutours) and wind vectors through DPG24 station (in Playa) during IOP-3. In these north-south cross-sections, leftward arrows represent north winds and downward arrows represent west winds.

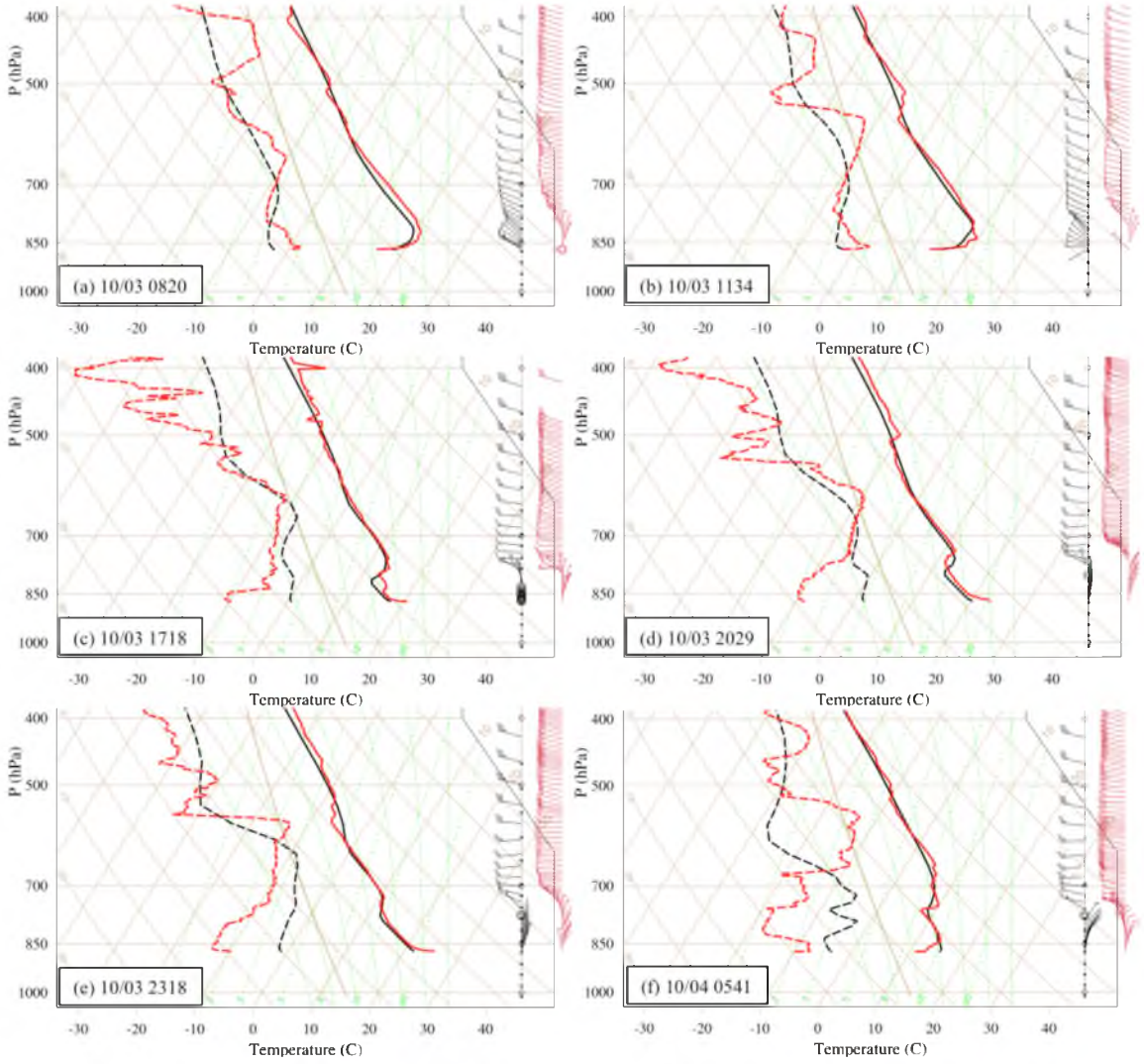


Figure 4.21 Comparison of skew-T between observations and the EnKF/WRF forecasts (ensemble mean) at the "SLTEST" (Playa) station.

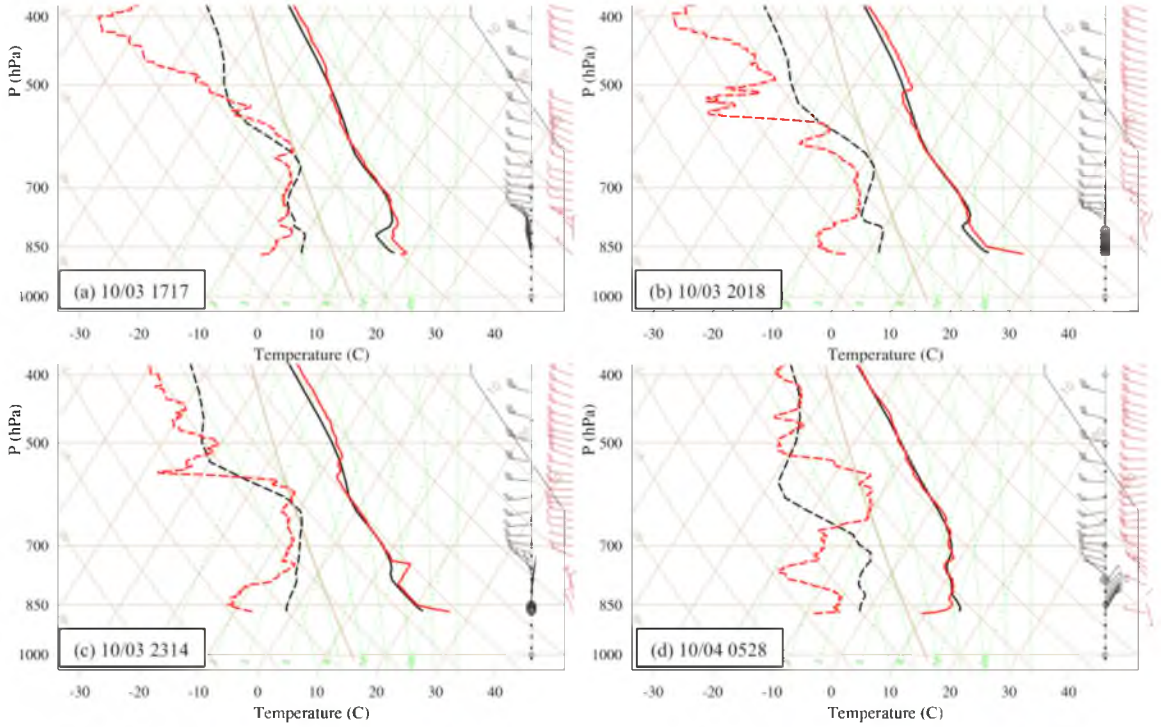


Figure 4.22 Same as Figure 4.21, but at the "Sagebrush" radiosonde launching station.

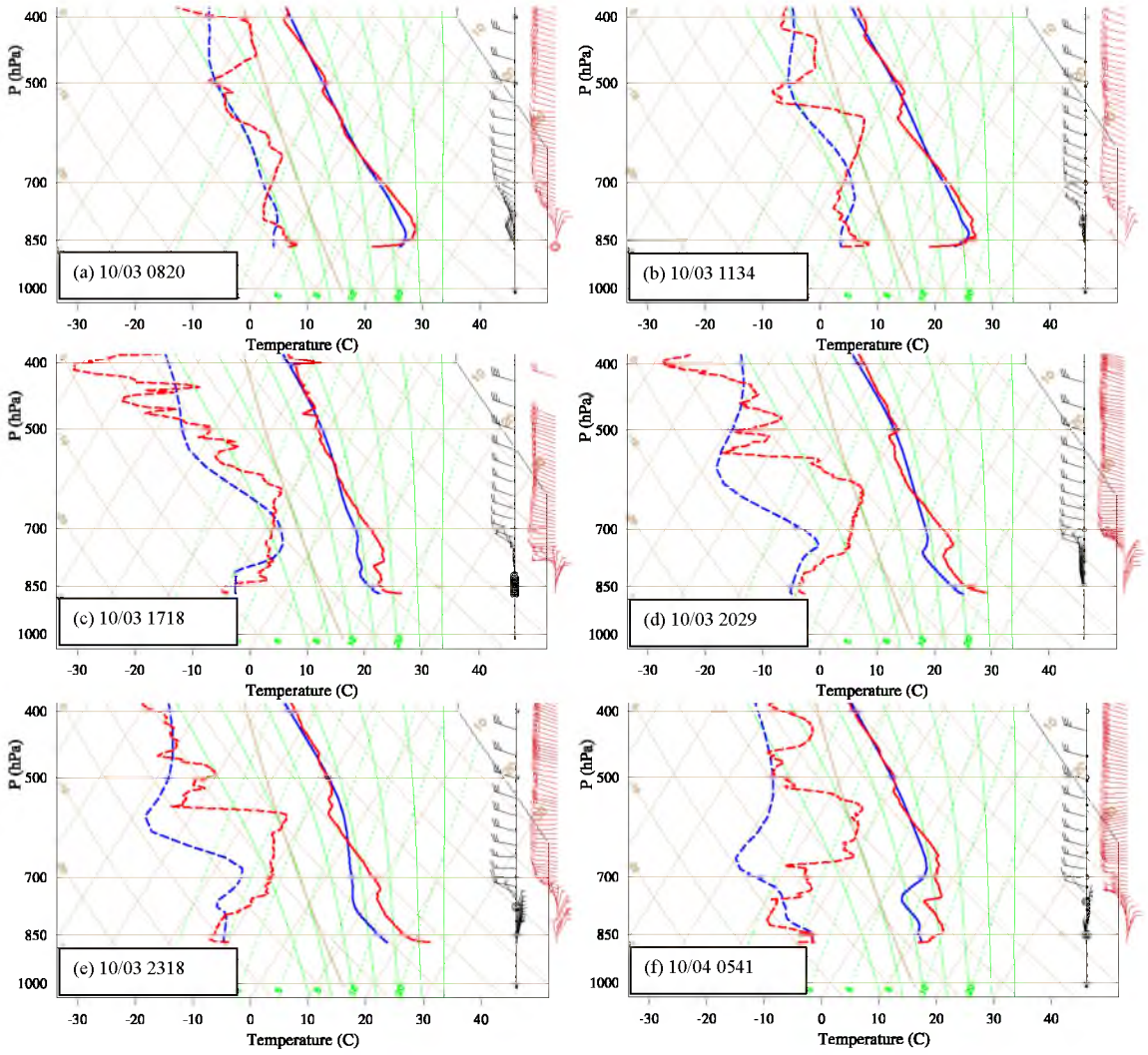


Figure 4.23 Comparison of skew-T between observations and the deterministic forecast initialized at 0000 UTC 3 October using NAM analysis at the "SLTEST" (Playa) radiosonde launching station.

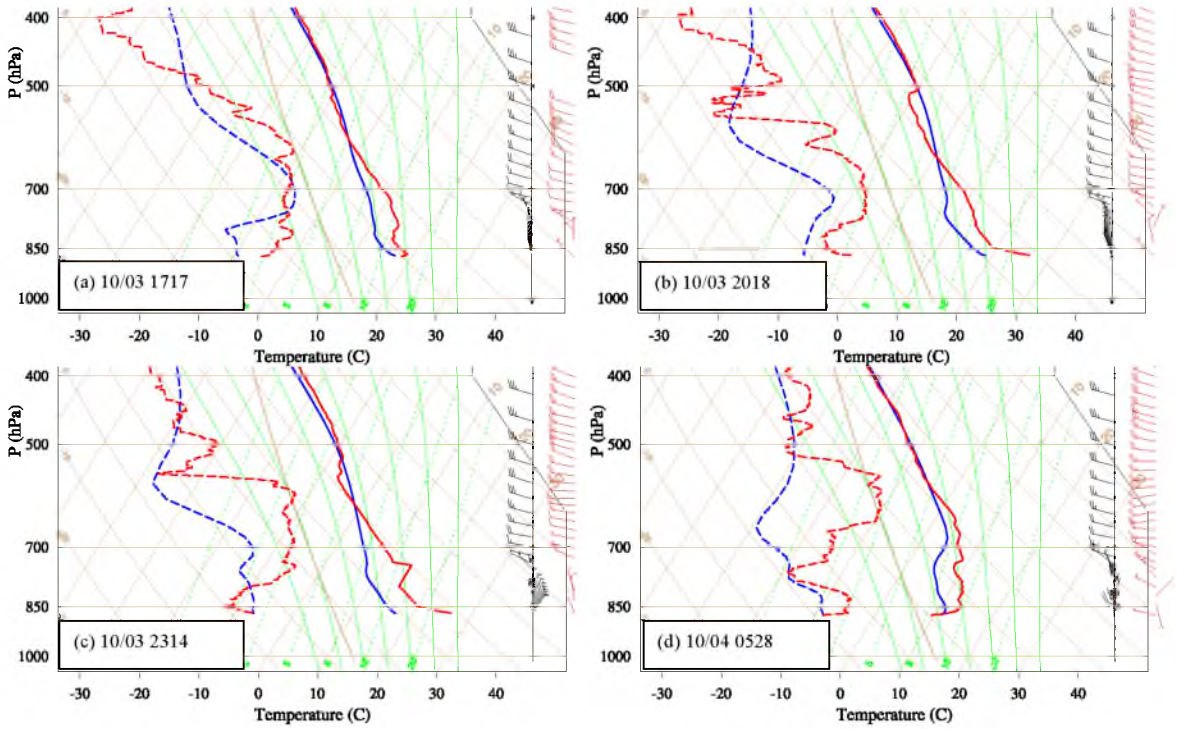


Figure 4.24 Same as Figure 4.23, but at the "Sagebrush" radiosonde launching station.

CHAPTER 5

SUMMARY AND DISCUSSION

Despite the importance of accurate, high-resolution atmospheric surface analyses and short-range forecasts in regions of complex terrain, they remain quite difficult to produce in current research and operational practice. In this dissertation, the ability of the ensemble Kalman filter assimilation of near-surface atmospheric observations is first evaluated in a framework of observing system simulation experiments. Then, forecast errors in near-surface temperature and wind from an advanced research version of the weather research and forecasting (WRF) model are characterized. Finally, data assimilation results with the recent MATERHORN field program are presented.

As the main source of conventional observations, surface data are helpful for weather prediction. However, their use in current NWP and data assimilation, especially in complex terrain, remains a challenge since the traditional 3DVAR method has a fundamental difficulty in assimilating surface observations. In an OSSE framework, we compare the ability of the EnKF and 3DVAR data assimilation methods to assimilate surface observations. Numerical experiments are conducted for both single observation experiments and multiple-observation cases.

The single-observation experiments simplify the problem and help us understand the fundamental differences between the 3DVAR and EnKF in how they deal with surface

observations. Results show that the 3DVAR method has fundamental problems in assimilating surface observations. Analysis increments from a valley station can spread unrealistically over mountainous areas even with a reasonable specification of the horizontal background length scale. On the other hand, the EnKF can overcome this limitation through its flow-dependent background error term. Even with a large localization scale, analysis increments from a valley station can still be retained in the valley. Overall, the major discrepancies between the 3DVAR and EnKF in the single-observation experiment can be attributed to the way the two methods define the background error terms in the data assimilation system. Since the 3DVAR uses a static background covariance, analysis increments from a single observation tend to be similar in various cases as they depend heavily on the prescribed correlation functions. Those prescribed correlation functions are based on large-scale dynamic balances and are suitable only for the large-scale free atmosphere. In contrast, the EnKF uses its ensemble to create the background error covariance terms. Owing to its flow-dependent background error covariance term, the EnKF enables observations to have more influence on areas where the ensemble spreads (forecast uncertainties) are larger.

The EnKF method is very sensitive to the choice of horizontal and vertical localization scales due to its small ensemble size that causes sampling error. Although the larger the ensemble size, the less important the localization, we have to compromise the ensemble size in order to keep the computational cost affordable. For the OSSEs, we conduct sensitivity experiments to select reasonable localization scales.

Comparison between the 3DVAR and EnKF method is also conducted using a synoptic case with two severe weather systems: a front over complex terrain in the

western US and a low-level jet over the Great Plains. Both methods are capable of extending information from the surface level to the atmospheric boundary layer. In flat terrain, the EnKF does better in terms of the analysis and forecast of the low-level jet while the 3DVAR also generally simulates the low-level jet well. In complex terrain, the EnKF performs much better than the 3DVAR does in general. To explore the performance of both methods with terrain misrepresentation, we add terrain perturbations to the terrain height of the nature run and carry out data-rejection experiments since a common way to deal with terrain misrepresentation is to reject data over the area where discrepancies between the model and the actual terrain are large. However, since observations are sparse in complex terrain, data rejection results in degraded analyses and forecasts, suggesting that this may not be the best solution for dealing with errors due to model terrain representation. But the EnKF still has a better ability to handle the data.

We used OSSEs to isolate various factors that affect the assimilation of near-surface observations. While the results of this study can help us understand what factors limit our ability to assimilate surface observations, real data assimilation and prediction problems are expected to be more complicated.

In order to better understand the forecast error characteristics of near-surface atmospheric variables, we investigate the performance of the ARW WRF model version 3.3 in predicting near-surface atmospheric temperature and wind conditions under various terrain and weather regimes, as described in Chapter 3. Three individual events under strong synoptic forcing, namely, a frontal system, a low-level jet, and a persistent cold air pool are first verified against observations in both flat and complex terrain. The examination begins with general verification of large-scale synoptic systems and is

followed by statistical verification of surface conditions. It was found that the ARW model was able to produce reasonable simulations of weather phenomena. However, verification of near-surface conditions (i.e., 2-m temperature and 10-m wind speed and direction) indicated the complexity in forecasting surface variables. For the front and low-level jet in the central US, the model terrain matched the actual terrain and thus mitigated representative errors. The forecasts of surface variables generally agreed well with observations. However, errors still occurred, depending on the model's ability to forecast the structures of the lower atmospheric boundary layer. For the inversion over the Salt Lake Valley, different error characteristics were found over the mountain and valley stations. Terrain mismatch and the ARW model's limited ability to simulate near-surface atmospheric conditions made the forecasting errors even more complicated.

Overall, forecast errors in near-surface atmospheric variables show flow-dependent features in all three individual cases when strong synoptic forcings are present. To better understand the characteristics of flow-dependent errors in complex terrain as they relate to near-surface forecasting, 1-month of forecasts (15 September to 14 October 2011) were conducted in complex terrain of the western US (i.e., DPG). It was found that the forecast errors of surface variables depend to a large degree on the diurnal cycle of the surface variables themselves, especially when the synoptic forcing is weak. The forecast errors for 2-m temperature reached two daily maxima at around 3 am and 3 pm local time, and two daily minima at around 7 am and 7 pm local time. Errors in wind speed and direction followed the same trends, with a maximum at night and a minimum in the afternoon. Forecast errors followed the same trends regardless of the initialization time, showing that forecast errors are independent of the initialization time and forecast lead

time. Further analyses revealed positive (warm) temperature biases at night and negative (cold) biases during the daytime. In contrast to 2-m temperature, wind direction and speed had no systematic biases from a long-term perspective. Under strong synoptic forcing, diurnal patterns in forecast errors were absent, while flow-dependent errors were clearly shown.

We also found that simulations at finer resolutions do not outperform those at coarser resolutions in most cases. This was explained well in Hart et al. (2005), who found that the inability of the numerical model to depict near-surface structures (such as strong temperature inversions) resulted in worse forecasts, even with better terrain representation. Meanwhile, increasing the model's vertical resolution does not help improve its prediction skill for near-surface variation, although it does improve the forecasts of mesoscale weather phenomena. Numerical forecasts of near-surface atmospheric conditions are also sensitive to the PBL scheme in the WRF model, but there is no single PBL scheme that performs always better than the others. These factors illustrate the complexity and challenges involved in near-surface simulation over complex terrain.

In order to develop scientific tools to help improve prediction skill, the most recent field experiments of the Mountain Terrain Atmospheric Modeling and Observations (MATERHORN) Program were designed to identify and study the limitations of current mesoscale models in predicting weather in mountainous terrain. Specifically, the MATERHORN field campaign collected comprehensive observations from multiple platforms (e.g., tethersondes, lidar, radiosondes, etc.) over mountainous regions, providing an opportunity to study the predictability of atmospheric conditions over

complex terrain.

Inspired by the findings from Chapters 2 and 3, namely, the capability of the EnKF in assimilating surface observations and the error characteristics associated with the near-surface variables in complex terrain, we investigated the ability of EnKF data assimilation to represent the near-surface and atmospheric states in mountainous terrain by conducting cycled data assimilation during the fall 2012 MATERHORN field period, as discussed in Chapter 4. A continuous, 3-hourly cycled EnKF data assimilation and forecast was performed for the whole period with 60 ensemble members. Conventional data, surface observations, and radiosondes collected in the field experiments were assimilated into the data assimilation system in 3-hourly cycles throughout the month. Month-long EnKF/WRF analyses and forecasts were obtained for general statistical analysis and case studies.

The analyses had overall smaller errors and biases than the forecasts for all surface variables: 2-m temperature and 10-m wind speed and direction. They benefits from the assimilation of observations in the continuous EnKF data assimilation cycles.

For 2-m temperature, the RMSEs peaked twice a day, one in the middle of the night and the other one in the afternoon. The two RMSE peaks corresponded to a warm bias maximum in the middle of the night and a cold bias maximum in the afternoon. There were also two bias error minima per day, one at nearly 0000 UTC and the other at 1200 UTC. It seems that the minima were caused by more observations being available at 0000 UTC and 1200 UTC in the data assimilation cycle, but it is not all of the reason. These results are consistent with those in Chapter 3, which showed the error peaks and lows without data assimilation. Overall, warm biases were present during the nighttime and

cold biases were present during the daytime. The diurnal errors were independent of forecast lead time and initialization time. RMSEs and BEs are also compared between NAM 3-h forecasts and the prior fields (EnKF/WRF 3-h forecasts) in the data assimilation cycle. The NAM forecasts also showed diurnal error and bias features, and the errors of the prior field inside the analysis cycle and the NAM 3-h forecasts are comparable. Although the NAM 3-h forecasts have overall smaller BEs and RMSEs, they still show systematic daytime cold biases and nighttime warm biases.

For 10-m wind speed, most of the RMSEs were less than 2 m/s except for the spin-up period. The biases were positive between 0600 UTC and 1800 UTC and negative between 1800 UTC and 0600 UTC in the forecasts.

For 10-m wind direction, RMSEs peak at the two transitional times, 0300 UTC (early evening) and 1500 UTC (early morning). The BEs were mostly positive and negligibly small and occurred during nighttime when the atmospheric boundary layer was stable.

The EnKF/WRF 3-h forecasts had overall smaller RMSEs than the NAM 3-h forecasts for both the 10-m wind speed and direction. This result was different from the comparison of 2-m temperature. That is because the surface temperature forecasts were influenced by a strong diurnal signal, while the surface wind speed and direction forecasts benefited from the high-resolution WRF model, which can better represent the terrain heights and the atmospheric structures in the lower boundary layer.

An intensive observation period (IOP-3) was selected as a case study. This was a transitional dry cold front that was passing over Utah during the IOP. The forecasts from the EnKF analysis successfully captured the large-scale synoptic weather patterns during the frontal passage. Simulated temperature profiles generally agreed with the radiosonde

observations although some disagreements did occur in the lower boundary layer.

In summary, this dissertation is the first comprehensive study of the use of ensemble Kalman filter data assimilation in regions of complex terrain. While many achievements have been made in this dissertation through advanced numerical models and data assimilation systems, a number of issues are pending for further studies to improve the prediction skill in complex terrain. For example, model errors from land-atmosphere thermal coupling are still an unsolved issue in near-surface data assimilation. In a recent study, with a single column model, Hacker and Angevine (2012) suggested that the parameterization errors that account for the inaccurate representation of thermal coupling between land and atmosphere are difficult to estimate. They also pointed out that soil and flux measurements (i. e., soil temperature and moisture and downwelling shortwave radiation) have large impacts on the forecast of near-surface variables. Future work should emphasize assimilating additional data types, dealing with model errors and investigating the thermal decoupling between near-surface variables and the atmospheric boundary layer and its impact on the model's prediction skill of near-surface variables.

REFERENCES

Aberson, S. D., 2001: The ensemble of tropical cyclone track forecasting models in the north Atlantic basin (1976–2000). *Bull. Amer. Meteor. Soc.*, **82**, 1895–1904.

Ancell, B. C., C. F. Mass, and G. J. Hakim, 2011: Evaluation of surface analyses and forecasts with a multiscale ensemble Kalman filter in regions of complex terrain. *Mon. Wea. Rev.*, **139**, 2008–2024.

Anderson, J. L., 2001: An ensemble adjustment Kalman filter for data assimilation. *Mon. Wea. Rev.*, **129**, 2884–2903.

Anderson, J. L., 2007: An adaptive covariance inflation error correction algorithm for ensemble filters. *Tellus*, **59**, 210–224.

Anderson, J. L., 2012: Localization and sampling error correction in ensemble Kalman filter data assimilation. *Mon. Wea. Rev.*, **140**, 2359–2371.

Anderson, J. L., T. Hoar, K. Raeder, H. Liu, N. Collins, N. Torn, and A. Avellano, 2009: The data assimilation research testbed: A community facility. *Bull. Amer. Meteor. Soc.*, **90**, 1283–1296.

Atlas, R. M., 1997: Observing system simulation experiments: Methodology, examples and limitations. *World Meteorological Organization CGC/WMO Workshop*, Geneva, Switzerland. Available at aosc.umd.edu/~seminar/data/y11spring/umd_aosc_110623_atlas.pdf

Baek, S.-J., B. R. Hunt, E. Kalnay, E. Ott, and I. Szunyogh, 2006: Local ensemble Kalman filtering in the presence of model bias. *Tellus*, **58**, 293–306.

Barker, D. M., M. Lee, Y.-R. Guo, W. Huang, Q.-N. Xiao, and R. Rizvi, 2004a: WRF variational data assimilation development at NCAR. *WRF/MM5 Users' Workshop*, Boulder, CO, USA. Available at http://www.mmm.ucar.edu/mm5/workshop/ws04/Session5/dale.pdf.

Barker, D. M., W. Huang, Y.-R. Guo, A. J. Bourgeois, and Q.-N. Xiao, 2004b: A three dimensional data assimilation system for use with MM5: Implementation and initial results. *Mon. Wea. Rev.*, **132,** 897–914.

Bauer, P., P. Lopez, A. Benedetti, D. Salmond, A. Benedetti, S. Saarinen, and M. Bonazzola, 2006: Implementation of 1D+4D-Var assimilation of precipitation-affected microwave radiances at ECMWF. II: 4D-Var. *Quart. J. Roy. Meteor. Soc.*, **132,** 2307–2332.

Benjamin, S. G., D. Devenyi, S. Weygandt, K. J. Brundage, J. Brown, J. Grell, D. Kim, B. Schwartz, T. G. Smirnova, and T. L. Smith, 2004: An hourly assimilation-forecast cycle: The RUC. *Mon. Wea. Rev.*, **132**, 495–518.

Benjamin, S. G. and Coauthors, 2011: Progress in NOAA hourly-updated models – Rapid Refresh, HRRR. *Alaska Weather Symposium*, Fairbanks, AK, USA.

Berner, J., G. J. Shutts, M. Leutbecher, and T. N. Palmer, 2009: A spectral stochastic kinetic energy backscatter scheme and its impact on flow-dependent predictability in the ECMWF ensemble prediction system. *J. Atmos. Sci.*, **66**, 603–626.

Bishop, C. H., B. J. Etherton, and S. J. Majumdar, 2001: Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. *Mon. Wea. Rev.*, **129**, 420–436.

Bjerknes, V., 1911: *Dynamic Meteorology and Hydrography: Kinematics.* Carnegie Institute of Washington, 175 pp.

Brown, A. R., and N. Wood, 2003: Properties and parameterization of the stable boundary layer over moderate topography. *J. Atmos. Sci.,* **60**, 2797–2808.

Buehner, M., 2012: Evaluation of a spatial/spectral covariance localization approach for atmospheric data assimilation. *Mon. Wea. Rev.*, **140**, 617–636.

Buehner, M., C. Charente, B. He, 2008: Intercomparison of 4-DVar and EnKF systems for operational deterministic NWP. Available at http://4dvarenkf.cima.fcen.uba.ar/Download/ Session 7/Intercomparison 4D-Var EnKF Buehner.pdf.

Buizza, R., and T. N. Palmer, 1995: The singular-vector structure of the atmospheric global circulation. *J. Atmos. Sci.*, **52**, 1434–1456.

Buizza, R., M. Miller, and T. N. Palmer, 1999: Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. *Quart. J. Roy. Meteor. Soc.,* **125**, 2887–2908.

Charney, J. G., R. Fjörtoft, and J. V. Neumann, 1950: Numerical integration of the barotropic vorticity equation. *Tellus*, **2**, 237–254.

Chen, F., and J. Dudhia, 2001: Coupling an advanced land surface–hydrology model with the Penn State–NCAR MM5 modeling system. Part I: Model implementation and

sensitivity. *Mon. Wea. Rev.*, **129**, 569–585.

Chen, S.-H., and W.-Y. Sun, 2002: A one-dimensional time dependence cloud model. *J. Meteor. Soc. Japan*, **80**, 99–118.

Cheng, W. Y., and W. J. Steenburgh, 2005: Evaluation of surface sensible weather forecasts by the WRF and the Eta models over the western United States. *Wea. Forecasting*, **20**, 812–821.

Corazza, M., E. Kalnay, and S. C. Yang, 2007: An implementation of the local ensemble Kalman filter in a quasi-geostrophic model and comparison with 3D-Var. *Nonlinear Processes in Geophysics*, **14**, 89–101.

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

De Pondeca, M. S. F. V., G. S. Manikin, G. DiMego, S. G. Benjamin, D. Parrish, R. Purser, W.-S. Wu, J. Horel, D. Myrick, Y. Lin, R. Aune, D. Keyser, B. Colman, G. Mann, and J. Vavra, 2011: The real-time mesoscale analysis at NOAA's National Centers for Environmental Prediction: Current status and development. *Wea. Forecasting*, **26,** 593–612.

Dong, J., M. Xue, and K. Droegemeier, 2007: The impact of high-resolution surface observation on convective storm analysis with ensemble Kalman filter. Preprints, *22nd Conf. on Weather Analysis and Forecasting/18th Conf. on Numerical Weather Prediction*, Park City, UT, Amer. Metero. Soc., P1.42.

Du, J., J. McQueen, G. DiMego, Z. Toth, D. Jovic, B. Zhou and H.-Y. Chuang, 2006: New dimension of NCEP short-range ensemble forecasting (SREF) system: Inclusion of WRF members. *WMO Expert Team Meeting on Ensemble Prediction System*, Exeter, UK. Available at http://www.emc.ncep.noaa.gov/mmb/SREF/WMO06_full.pdf

Du, J., and M. S. Tracton, 2001: Implementation of a real time short-range ensemble forecasting system at NCEP: An update. Proceedings, *9th conf. on Mesoscale processes*, Fort Lauderdale, FL, Amer. Metero. Soc., P4.9.

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

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

Epstein, E. S., 1969: Stochastic dynamic prediction. *Tellus*, **6**, 739–759.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99**, 10143–10162.

Evensen, G., and P. J. Van Leeuwen, 1996: Assimilation of Geosat altimeter data for the agulhas current using the ensemble Kalman filter with a quasi-geostrophic model. *Mon. Wea. Rev.*, **124**, 85–96.

Evensen, G., 2003: The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dynamics*, **53**, 343–367.

Fernando, H. J. S., and E. Pardyjak, 2013: Field studies delve into the intricacies of mountain weather. *Eos*, **94**, 36.

Fertig, E. J., J. Harlim, and B. R. Hunt, 2007: A comparative study of 4D-VAR and a 4D ensemble Kalman filter: Perfect model simulations with Lorenz-96. *Tellus*, **59**, 96–100.

Fujita, T., D. Stensrud, and D. Dowell, 2007: Surface data assimilation using an ensemble Kalman filter approach with initial condition and model physical uncertainties. *Mon. Wea. Rev.*, **135**, 1846–1868.

Gustafsson, B. 1981: The convergence rate for difference approximations to general mixed initial-boundary value problems. *SIAM J. Numer. Anal.* **18**, 179–190.

Hacker, J. P. and W. M. Angevine, 2012: Ensemble data assimilation to characterize land-atmosphere coupling errors in numerical weather prediction models. *Mon. Wea. Rev.*, **141**, 1804–1821.

Hacker, J. P., J. L. Anderson, and M. Pagowski, 2007: Improved vertical covariance estimates for ensemble-filter assimilation of near-surface observations, *Mon. Wea. Rev.*, **135**, 1021–1036.

Hacker, J. P., and C. Snyder, 2005: ensemble Kalman filter assimilation of fixed screen -height observations in a parameterized PBL. *Mon. Wea. Rev.*, **133**, 3260–3275.

Haltine, G. J., 1971: *Numerical Weather Prediction*. John Wiley and Sons, 326 pp.

Hamill, T. M., 2006: Ensemble based atmospheric data assimilation. *Predictability of Weather and Climate*, T. Palmer and R. Hagedorn, Eds., Cambridge Press, 124–156.

Hamill, T. M., J. Whitaker, and C. Snyder, 2001: Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Mon. Wea. Rev.*, **129,** 2776–2790.

Hamill, T. M., and C. Snyder, 2000: A hybrid ensemble Kalman filter–3D variational analysis scheme. *Mon. Wea. Rev.,* **128**, 2905–2919.

Hanna, S. R., and R. Yang, 2001: Evaluations of mesoscale models' simulations of near-surface winds, temperature gradients, and mixing depths. *J. Appl. Meteor.,* **40**, 1095–1104.

Harlim, J., and B. R. Hunt, 2007: Four-dimensional local ensemble transform Kalman filter: Numerical experiments with a global circulation model. *Tellus*, **59**, 731–748.

Hart, K. A., W. J. Steenburgh, and D. J. Onton, 2005: Model forecast improvements with decreased horizontal grid spacing over finescale intermountain orography during the 2002 Olympic Winter Games. *Wea. Forecasting*, **20**, 558–576.

Hong, S.-Y., and J.-O. Lim, 2006: The WRF single-moment 6-class microphysics scheme (WSM6). *J. Korean Meteor. Soc.*, **42**, 129–151.

Hong, S.-Y., and H.-L. Pan, 1996: Nonlocal boundary layer vertical diffusion in a medium-range forecast model. *Mon. Wea. Rev.,* **124**, 2322–2339.

Hong, S.-Y., S.-W. Kim, J. Dudhia, M.-S. Koo, and K.-H. Seol, 2008: Stable boundary layer mixing in a vertical diffusion package. *9th Annual WRF User's Workshop*, Boulder, CO. Available at http://www.mmm.ucar.edu/wrf/users/workshops/WS2008/presentations/3-3.pdf.

Horel. J., M. Splitt, L. Dunn, J. Pechmann, B. White, C. Ciliberti, S. Lazarus, J. Slemmer, and D. Zaff, 2002: Mesowest: Cooperative mesonets in the western United States. *Bull. Amer. Meteor. Soc.*, **83**, 211–226.

Hoteit, I., D.-T. Pham, G. Korres, and G. Triantafyllou, 2008: Particle Kalman filtering for data assimilation in meteorology and oceanography. Proceedings, *3rd International Conf. on Reanalysis*, Tokyo, Japan, WCRP, 6.

Houtekamer, P. L., L. Lefaivre, J. Derome, H. Ritchie, and H. L. Mitchell, 1996: A system simulation approach to ensemble prediction. *Mon. Wea. Rev.,* **124**, 1225–1242.

Houtekamer, P. L., and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.,* **126**, 796–811.

Houtekamer, P. L., and H. L. Mitchell, 2001: A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.,* **129,** 123–137.

Hu, X., J. Nielsen-Gammon, and F. Zhang, 2010: Evaluation of three planetary boundary layer schemes in the WRF model. *J. Appl. Meteor. Climatol.,* **49**, 1831–1844.

Hunt, B. R., E. J. Kostelich, and I. Szunyogh, 2007: Efficient data assimilation for

spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D*, **230**, 112–126.

Kain, J. S., and J. M. Fritsch, 1993: The role of the convective "trigger function" in numerical prediction of mesoscale convective systems. *Meteor. Atmos. Phys.*, **49**, 93–106.

Kalnay, E., H. Li, T. Miyoshi, S.-C. Yang, and J. Ballabrera-Poy, 2007: 4-D-Var or ensemble Kalman filter. *Tellus*, **59**, 758–773.

Kalnay, E., 2003: *Atmospheric Modeling, Data Assimilation and Predictability.* Cambridge University Press, 364 pp.

Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Wollen, Y. Zhu, A. Leetmaa, R. Reynolds, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K. C. Mo, C. Ropelewski, J. Wang, R. Jenne, and D. Joseph, 1996: The NCEP/NCAR 40-Year reanalysis project. *Bull. Amer. Meteor. Soc.*, **77**, 437–471.

Kharin V. V., and F. W. Zwiers, 2002: Climate predictions with multimodel ensembles: *J. Climate*, **15**, 793–799.

Klewicki, J. C., P. J. A. Priyadarshana, and M. M. Metzger, 2008: Statistical structure of the fluctuating wall pressure and its in-plane gradients at high Reynolds number. *J. Fluid Mech.* **609**, 195–220.

Köpken, C., G. Kelly, and J.-N. Thépaut, 2004: Assimilation of Meteosat radiance data within the 4D-Var system at ECMWF: Assimilation experiments and forecasts impact. *Quart. J. Roy. Meteor. Soc.*, **130**, 2277–2292.

Krishnamurti, T. N., C. M. Kishtawal, Z. Zhang, T. LaRow, D. Bachiochi, E. Williford, S. Gadgil, and S. Surendran, 2000: Multimodel ensemble forecasts for weather and seasonal climate. *J. Climate*, **13**, 4196–4216.

Lahoz, W. A., R. Brugge, D. R. Jackson, S. Migliorini, R. Swinbank, D. Lary, and A. Lee, 2005: Observing system simulation experiment to evaluate the scientific merit of wind and ozone measurements from the future SWIFT instrument. *Quart. J. Roy. Meteor. Soc.*, **131**, 503–523.

Lee, T. J., R. A. Pielke, R. C. Kessler, and J. Weaver, 1989: Influence of cold pools downstream of mountain barriers on downslope winds and flushing. *Mon. Wea. Rev.*, **117**, 2041–2058.

Lee, S.-J., D. F. Parrish, W.-S. Wu, M. De Pondeca, D. Keyser, and G. DiMego, 2005: Use of surface mesonet data in NCEP regional grid statistical-interpolation system. Preprints, *17th Conf. on Numerical Weather Prediction*, Washington, DC, Amer.

Meteor. Soc., 6A.7.

Leung, L. R., and Y. Qian, 2003: The sensitivity of precipitation and snowpack simulations to model resolution via nesting in regions of complex terrain. *J. Hydrometeor*, **4**, 1025–1043.

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

Liu, Y., and Coauthors, 2008a: The Operational mesogamma-scale analysis and forecast system of the U.S. Army test and evaluation Command. Part I: Overview of the modeling system, the forecast products, and how the products are used. *J. Appl. Meteor. Climatol.*, **47**, 1077–1092.

Liu, Y., and Coauthors, 2008b: The operational mesogamma-scale analysis and forecast system of the U.S. army test and evaluation command. Part II: Interrange comparison of the accuracy of model analyses and forecasts. *J. Appl. Meteor. Climatol.*, **47**, 1093–1104.

Lorenc, A., 2003: The potential of the ensemble Kalman filter for NWP—A comparison with 4D-Var. *Quart. J. Roy. Meteor. Soc.*, **129**, 3183–3203.

Lorenc, A., 1986: Analysis methods for numerical weather prediction, *Quart. J. Roy. Meteor. Soc*, **112**, 1177–1194.

Lorenz, E. N., 1963: The predictability of hydrodynamics flow, *Transactions of the New York Academy of Sciences Series II*, **25**, 409–432.

Lorenz, E. N., 1965: A study of the predictability of a 28-variable atmospheric model. *Tellus*, **12**, 321–333.

Mackey, B. P., and T. N. Krishnamurti, 2001: Ensemble forecast of a typhoon flood event. *Wea. Forecasting*, **16**, 399–415.

Mahfouf, J. F., A. O. Manzi, J. Noilhan, H. Giordani, and M. Deque, 1995: The land surface scheme ISBA within the Meteo-France climate model ARPEGE. Part I: Implementation and preliminary results. *J. Climate*, **8**, 2039–2057.

Mahfouf, J. F., P. Bauer, and V. Marécal, 2005: The assimilation of SSM/I and TMI rainfall rates in the ECMWF 4D-Var system. *Quart. J. Roy. Meteor. Soc.*, **131**, 437–458.

Malek, E., 2003: Microclimate of a desert playa: Evaluation of annual radiation, energy, and water budgets components. *J. Climatol.*, **23**, 333–345.

Mass, C. F., D. Ovens, K. Westrick, and B. A. Colle, 2002: Does increasing horizontal resolution produce more skillful forecasts? *Bull. Amer. Meteor. Soc.*, **83**, 407–430.

McCaul, E. W., and C. Cohen, 2002: The impact on simulated storm structure and intensity of variations in the mixed layer and moist layer depths. *Mon. Wea. Rev.*, **130**, 1722–1748.

Mellor, G. L., and T. Yamada, 1982: Development of a turbulence closure model for geophysical fluid problems. *Rev. Geophys. Space Phys.*, **20**, 851–875.

Meng, Z., and F. Zhang, 2008a: Tests of an ensemble Kalman filter for mesoscale and regional-scale data assimilation. Part III: Comparison with 3DVAR in a real-data case study. *Mon. Wea. Rev.*, **136,** 522–540.

Meng, Z., and F. Zhang, 2008b: Tests of an ensemble Kalman filter for mesoscale and regional-scale data assimilation. Part IV: Comparison with 3DVAR in a month-long experiment. *Mon. Wea. Rev.*, **136,** 3671–3682.

Mesinger, F., G. DiMego, E. Kalnay, K. Mitchell, P. Shafran, W. Ebisuzaki, D. Jović, J. Woollen, E. Rogers, E. Berbery, M. Ek, Y. Fan, R. Grumbine, W. Higgins, H. Li, Y. Lin, G. Manikin, D. Parrish, and W. Shi, 2006: North American Regional Reanalysis. *Bull. Amer. Meteor. Soc.*, **87**, 343–360.

Meybeck, M., P. Green, and C. Vo ro smarty, 2001: A new typology for mountains and other relief classes: an application to global continental water resources and population distribution. *Mountain Research and Development,* **21**, 34–45.

Miyoshi, T., 2011: The Gaussian approach to adaptive covariance inflation and its implementation with the local ensemble transform Kalman filter. *Mon. Wea. Rev.*, **139**, 1519–1535.

Miyoshi, T., and S. Yamane, 2007: Local ensemble transform Kalman filtering with an AGCM at a T159/L48 resolution. *Mon. Wea. Rev.*, **135**, 3841–3861.

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. *J. Geophys. Res.*, **102**, 16663–16682.

Molteni, F., R. Buizza, T. N. Palmer, and T. Petroliagis, 1996: The ECMWF Ensemble prediction system: Methodology and validation. *Quart. J. Roy. Meteor. Soc.*, **122**, 73–119.

Mullen, S. L., J. Du, and F. Sanders, 1999: The dependence of ensemble dispersion on analysis-forecast systems: Implications to short-range ensemble forecasting of precipitation. *Mon. Wea. Rev.*, **127**, 1674–1686.

Nakanishi, M., and H. Niino, 2004: An improved Mellor–Yamada level-3 model with condensation physics: Its design and verification. *Bound.-Layer Meteor.*, **112**, 1–31.

Palmer, T. N., 2001: A nonlinear dynamical perspective model error: A proposal for non-local stochastic-dynamic parametrization in weather and climate prediction models. *Quart. J. Roy. Meteor. Soc.*, **127**, 279–304.

Parrish, D. F. and J. C. Derber, 1992: The National Meteorological Center's spectral statistical interpolation analysis system. *Mon. Wea. Rev.*, **120**, 1747–1763.

Peters-Lidard, C. D., E. Blackburn, X. Liang, and E. F. Wood, 1998: The effect of soil thermal conductivity parameterization on surface energy fluxes and temperatures. *J. Atmos. Sci.*, **55**, 1209–1224.

Pleim, J. E., 2007: A combined local and nonlocal closure model for the atmospheric boundary layer. Part I: Model description and testing. *J. Appl. Meteor. Climatol.*, **46**, 1383–1395.

Pu, Z., H. Zhang, and J. A. Anderson, 2013: Ensemble Kalman filter assimilation of near-surface observations over complex terrain: Comparison with 3DVAR for short-range forecasts. *Tellus*, **65**, 19620.

Pu. Z., and J. Hacker, 2009: Ensemble-based Kalman filters in strongly nonlinear dynamics. *Adv. Atmos. Sci.*, **26**, 373–380.

Reynolds, C. A., J. Teixeira, and J. G. McLay, 2008: Impact of stochastic convection on the ensemble transform. *Mon. Wea. Rev.*, **136**, 4517–4526.

Rife, D. L., and C. A. Davis, 2005: Verification of temporal variations in mesoscale numerical wind forecasts. *Mon. Wea. Rev.*, **133**, 3368–3381.

Rogers, E., M. Ek, B. S. Ferrier, G. Gayno, Y. Lin, K. Mitchell, M. De Pondeca, M. Pyle, V. Wong, and W.-S. Wu, 2005: The NCEP North American Mesoscale Modeling System: Final Eta model/analysis changes and preliminary experiments using the WRF-NMM. Proceedings, *21st Conf. on Weather Analysis and Forecasting/17th conf. on Numerical Weather Prediction,* Washington, DC, Amer. Meteor. Soc., 4B.5.

Shafran, P. C., N. L. Seaman, and G. A. Gayno, 2000: Evaluation of numerical predictions of boundary layer structure during the Lake Michigan Ozone Study (LMOS). *J. Appl. Meteor.*, **39**, 412–426.

Shin, H. H., and S.-Y. Hong, 2011: Intercomparison of planetary boundary-layer parameterizations in the WRF model for a single day from CASES-99. *Bound.-Layer Meteor.*, **139**, 261– 281.

Shutts, G., 2005: A kinetic energy backscatter algorithm for use in ensemble prediction

systems. *Quart. J. Roy. Meteor. Soc.*, **131**, 3079–3102.

Simmons, A. J., P. Jones, V. Bechtold, A. Beljaars, P. Kallberg, S. Saarinen, S. Uppala, P. Viterbo, and N. Wedi, 2004: Comparison of trends and low-frequency variability in CRU, ERA-40, and NCEP/NCAR analyses of surface air temperature. *J. Geophys, Res.*, **109**, D24115.

Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, M. Barker, K. G. Duda, X.-Y. Huang, W. Wang, and J. G. Powers, 2008: A description of the advanced research WRF Version 3. *NCAR Tech.* Note, NCAR/TN-475+STR, 113 pp.

Stensrud, D., N. Yussouf, D. Dowell, and M. Coniglio, 2009: Assimilating surface data into a mesoscale model ensemble: Cold pool analyses from spring 2007. *Atmos. Res.*, **93**, 207–220.

Stull, R. B., 1988: *An Introduction to Boundary Layer Meteorology*. Springer, 666 pp.

Sukoriansky, S., B. Galperin, and V. Perov, 2005: Application of a new spectral theory of stably stratified turbulence to atmospheric boundary layers over sea ice. *Bound.-Layer Meteor.*, **117**, 231–257.

Teixeira, J., C. A. Reynolds, and K. Judd, 2007: Time step sensitivity of nonlinear atmospheric models: Numerical convergence, truncation error growth, and ensemble design. *J. Atmos. Sci.*, **64**, 175–189.

Teixeira, J., and Coauthors, 2008: Parameterization of the atmospheric boundary layer: A view from just above the inversion. *Bull. Amer. Meteor. Soc.*, **89**, 453–458.

Thompson, P., 1961: *Numerical Weather Analysis and Prediction*. Macmillan, 170 pp.

Tippett, M. K., J. L. Anderson, C. H. Bishop, T. M. Hamill, and J. S. Whitaker, 2003: Ensemble square root filters. *Mon. Wea. Rev.*, **131**, 1485–1490.

Tong, M., and M. Xue, 2008: Simultaneous estimation of micro-physical parameters and atmospheric state with simulated radar data and ensemble square root Kalman filter. Part II: Parameter estimation experiments. *Mon. Wea. Rev.*, **136**, 1649–1668.

Torn, R. D., G. J. Hakim, and C. Snyder, 2006: Boundary conditions for limited-area ensemble Kalman filters. *Mon. Wea. Rev.*, **134**, 2490–2502.

Torn, R. D., and G. J. Hakim, 2008: Performance characteristics of a pseudo-operational ensemble Kalman filter. *Mon. Wea. Rev.*, **136,** 3947–3963.

Torn, R. D., and G. J. Hakim, 2009: Ensemble data assimilation applied to RAINEX observations of Hurricane Katrina (2005). *Mon. Wea. Rev.*, **137**, 2817–2829.

Toth, Z., and E. Kalnay, 1993: Ensemble forecasting at NCEP: The generation of

perturbations. *Bull. Amer. Meteor. Soc.*, **74**, 2371–2330.

Toth, Z., and E. Kalnay, 1997: Ensemble forecasting at NCEP and the breeding method. *Mon. Wea. Rev.*, **125**, 3297–3319.

Tracton, M. S., and E. Kalnay, 1993: Operational ensemble prediction at the National Meteorological Center: Practical aspects. *Wea. Forecasting*, **8**, 379–398.

Van Leeuwen, P. J., 2003: A variance-minimizing filter for large-scale applications. *Mon. Wea. Rev.*, **131**, 2071–2084.

Wang, X., D. M. Barker, C. Snyder, and T. M. Hamill, 2008: A hybrid ETKF–3DVAR data assimilation scheme for the WRF model. Part II: Real observation experiments. *Mon. Wea. Rev.*, **136**, 5132–5147.

Wei, M., Z. Toth, R. Wobus, and Y. Zhu, 2008: Initial perturbations based on the ensemble transform (ET) technique in the NCEP global operational forecast system. *Tellus*, **60**, 62–79.

Whitaker, J. S. and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130**, 1913–1924.

Whiteman, C. D., X. Bian, and S. Zhong, 1997: Low-level jet climatology from enhanced rawinsonde observations at a site in the southern Great Plains. *J. Appl. Meteor.*, **36**, 1363–1376.

Wolyn, P. G., and T. B. McKee, 1989: Deep stable layers in the intermountain western United States. *Mon. Wea. Rev.*, **117**, 461–472.

Wu, W.-S., R. J. Purser, and D. F. Parrish, 2002: Three-dimensional variational analysis with spatially inhomogeneous covariances. *Mon. Wea. Rev.* **130**, 2905–2916.

Xie, Y., S. E. Koch, J. A. McGinley, S. Albers, and N. Wang, 2005: A sequential variational analysis approach for mesoscale data assimilation. Preprints, *21st Conf. on Weather Analysis and Forecasting/17th Conf. on Numerical Weather Prediction*, Washington, DC, Amer. Meteor. Soc., 15B.7.

Xue, M., M. Tong, and G. Zhang, 2009: Simultaneous state estimation and attenuation correction for thunderstorms with radar data using an ensemble Kalman filter: Tests with simulated data. *Quart. J. Roy. Meteor. Soc.*, **135**, 1409–1423.

Yang, S.-C., M. Corazza, A. Carrassi, E. Kalnay, and T. Miyoshi, 2007: Comparison of ensemble-based and variational-based data assimilation schemes in a quasi-geostrophic model. Proceedings, *10th Symposium on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface*. Available at http://ams.confex.com/ams/pdfpapers/101581.pdf.

Yang, S.-C., and E. Kalnay, 2010: Handling nonlinearity and non-Gaussianity in ensemble Kalman filter: Experiments with the three-variable Lorenz model. Available at http://www.atmos.umd.edu/~ekalnay/YangKalnay2010Rev.pdf.

Yang, S.-C., E. Kalnay, B. Hunt, and N. E. Bowler, 2009: Weight interpolation for efficient data assimilation with the local ensemble transfom Kalman filter. *Quart. J. Roy. Meteor. Soc.*, **135**, 251–262.

Yuan, H., S. L. Mullen, X. Gao, S. Sorooshian, J. Du, and H.-M. H. Juang, 2005: Verification of probabilistic quantitative precipitation forecasts over the southwest United States during winter 2002/03 by the RSM Ensemble system. *Mon. Wea. Rev.*, **133**, 279–294.

Zhang, D.-L., and W. Zheng, 2004: Diurnal cycles of surface winds and temperatures as simulated by five boundary-layer parameterizations. *J. Appl. Meteor.*, **43**, 157–169.

Zhang, H., and Z. Pu, 2010: Beating the uncertainties: Ensemble forecasting and ensemble-based data assimilation in modern numerical weather prediction. *Adv. Meteorol.*. Available at http://dx.doi.org/10.1155/2010/432160

Zhang, H., Z. Pu and X. Zhang, 2013: Examination of errors in near-surface temperature and wind from WRF numerical simulations in regions of complex terrain. *Wea. Forecasting*, **28**, 893–914.

Zhang, F., Y. Weng, J. A. Sippel, Z. Meng, and C. H. Bishop, 2009: Cloud-resolving hurricane initialization and prediction through assimilation of doppler radar observations with an ensemble Kalman filter. *Mon. Wea. Rev.*, **137**, 2105–2125.

Zhang, Z., and T. N. Krishnamurti, 1997: Ensemble forecasting of hurricane tracks. *Bull. Amer. Meteor. Soc.*, **78**, 2785–2795.

Zhong, S., and J. Fast, 2003: An evaluation of the MM5, RAMS, and Meso-Eta models at subkilometer resolution using VTMX field campaign data in the Salt Lake Valley. *Mon. Wea. Rev.*, **131**, 1301–1322.